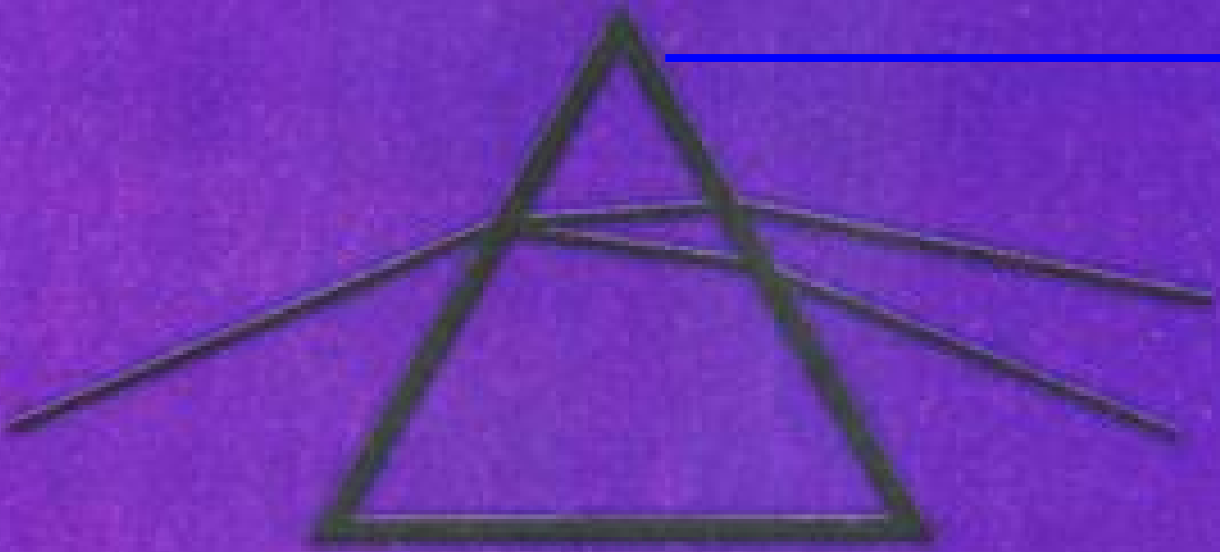
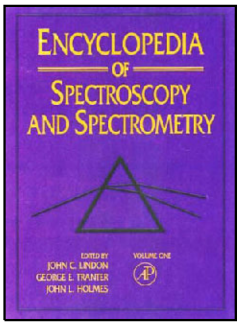
# ENCYCLOPEDIA OF SPECTROSCOPY AND SPECTROMETRY



EDITED BY
JOHN C. LINDON
GEORGE E. TRANTER
JOHN L. HOLMES

VOLUME ONE

# Encyclopedia of Spectroscopy and Spectrometry, Three-Volume Set
(Volume 1-3)




**George Tranter** (Editor)
*GlaxoWellcome Medicines Research, Beckenham, UK*
**John Holmes** (Editor)
*University of Ottawa, Ontario, Canada*
**John Lindon** (Editor-in-Chief)
*Imperial College of Science, Technology, and Medicine, London,* U.K.

## Features & Benefits

- Contains nearly 300 articles covering the whole area of spectroscopy and related areas
- Provides the essential facts and background for each topic
- Includes reading lists in each article to help the reader access more detailed information
- Allows easy access to required information through alphabetical listings, extensive cross-referencing to related articles, and a detailed subject index in each volume
- Includes numerous figures and tables that illustrate the text
- Contains color plate sections in each volume
- Includes initial access to the online version with extensive hypertext linking and advanced search tools Ongoing access is maintained for a minimum annual fee

## Description

The **Encyclopedia of Spectroscopy and Spectrometry** provides authoritative and comprehensive coverage of the whole topic of spectroscopy, from theory to applications. Short articles, each covering one aspect of spectroscopy, provide the professional spectroscopist working in academia or industry with the essential facts and background on areas of spectroscopy peripheral to their own. A list of further reading at the end of each article directs the reader to the level of detail required for professional purposes. Articles are arranged alphabetically, each having been named to facilitate logical access by the reader. Each article is flagged as to which area of spectroscopy it covers ("Mass Spectroscopy," "Magnetic Resonance," etc.) and whether it covers theory, methods and instrumentation, or applications. Users can refer to an alphabetical article listing, or to a listing arranged according to subject area to locate articles. Further reading lists at the end of each article allow easy access to the primary literature. Extensive cross-referencing, a complete subject index, numerous figures, and color plates are included in each volume.
Initial access to the online version offering extensive hypertext linking and advanced search tools is available to buyers of the print edition. Ongoing access is maintained for a minimum annual fee.

**Readership:**
Professional spectroscopists working in academia or industry (e.g., in pharmaceutical, chemical, and engineering industries), relevant libraries and individual research groups. CD-ROM version especially should be useful to teachers/ undergraduates as a source of course material and to postgraduates for practical tips and ideas for new ways of tackling problems.

**ISBN:** 0-12-226680-3
Book/Hardback

**Measurements:** 200 X 273 mm
**Pages:** 2581

**Imprint:** Academic Press

**Publication Date:** 21 January 2000

**Price:** £695.00

# Preface

This encyclopedia provides, we believe, a comprehensive and up-to-date explanation of the most important spectroscopic and related techniques together with their applications.

The *Encyclopedia of Spectroscopy and Spectrometry* is a cumbersome title but is necessary to avoid misleading readers who would comment that a simplified title such as the "Encyclopedia of Spectroscopy" was a misnomer because it included articles on subjects other than spectroscopy. Early in the planning stage, the editors realized that the boundaries of spectroscopy are blurred. Even the expanded title is not strictly accurate because we have also deliberately included other articles which broaden the content by being concerned with techniques which provide localized information and images. Consequently, we have tried to take a wider ranging view on what to include by thinking about the topics that a professional spectroscopist would conveniently expect to find in such a work as this. For example, many professionals use spectroscopic techniques, such as nuclear magnetic resonance, in conjunction with chromatographic separations and also make use of mass spectrometry as a key method for molecular structure determination. Thus, to have an encyclopedia of spectroscopy without mass spectrometry would leave a large gap. Therefore, mass spectrometry has been included. Likewise, the thought of excluding magnetic resonance imaging (MRI) seemed decidedly odd. The technique has much overlap with magnetic resonance spectroscopy, it uses very similar equipment and the experimental techniques and theory have much in common. Indeed, today, there are a number of experiments which produce multidimensional data sets of which one dimension might be spectroscopic and the others are image planes. Again the subject has been included.

This led to the general principle that we should include a number of so-called spatially-resolved methods. Some of these, like MRI, are very closely allied to spectroscopy but others such as diffraction experiments or scanning probe microscopy are less so, but have features in common and are frequently used in close conjunction with spectroscopy. The more peripheral subjects have, by design, not been treated in the same level of detail as the core topics. We have tried to provide an overview of as many as possible techniques and applications which are allied to spectroscopy and spectrometry or are used in association with them. We have endeavoured to ensure that the core subjects have been treated in substantial depth. No doubt there are omissions and if the reader feels we got it wrong, the editors take the blame.

The encyclopedia is organized conventionally in alphabetic order of the articles but we recognize that many readers would like to see articles grouped by spectroscopic area. We have achieved this by providing separate contents lists, one listing the articles in an intuitive alphabetical form, and the other grouping the articles within specialities such as mass spectrometry, atomic spectroscopy, magnetic resonance, etc. In addition each article is flagged as either a "Theory", "Methods and Instrumentation" or "Applications" article. However, inevitably, there will be some overlap of all of these categories in some articles. In order to emphasize the substantial overlap which exists among the spectroscopic and spectrometric approaches, a list has been included at the end of each article suggesting other articles in this encyclopedia which are related and which may provide relevant information for the reader. Each article also comes with a "Further Reading" section which provides a source of books and major reviews on the topic of the article and in some cases also provides details of seminal research papers. There are a number of colour plates in each volume as we consider that the use of colour can add greatly to the information content in many cases, for example for imaging studies. We have also included extensive Appendices of tables of useful reference data and a contact list of manufacturers of relevant equipment.

We have attracted a wide range of authors for these articles and many are world recognized authorities in their fields. Some of the subjects covered are relatively static, and their articles provide a distillation of the established knowledge, whilst others are very fast moving areas and for these we have aimed at presenting up-to-date summaries. In addition, we have included a number of entries which are retrospective in nature, being historical reviews of particular types of spectroscopy. As with any work of this magnitude some of the articles which we desired and commissioned to include did not make it for various reasons. A selection of these will appear in a separate section in the on-line version of the encyclopedia, which will be available to all purchasers of the print version and will have extensive hypertext links and advanced search tools. In this print version there are 281 articles contributed by more than 500 authors from 24 countries. We have persuaded authors from Australia, Belgium, Canada, Denmark, Finland, France, Germany, Hungary, India,

Israel, Italy, Japan, Mexico, New Zealand, Norway, Peru, Russia, South Africa, Spain, Sweden, Switzerland, The Netherlands, the UK and the USA to contribute.

The encyclopedia is aimed at a professional scientific readership, for both spectroscopists and non-spectroscopists. We intend that the articles provide authoritative information for experts within a field, enable spectroscopists working in one particular field to understand the scope and limitations of other spectroscopic areas and allow scientists who may not primarily be spectroscopists to grasp what the various techniques comprise in considering whether they would be applicable in their own research. In other words we tried to provide something for everone, but hope that in doing so, we have not made it too simple for the expert or too obscure for the non-specialist. We leave the reader to judge.

John Lindon
John Holmes
George Tranter

# Editor-in-Chief

John C. Lindon, *Biological Chemistry,*
*Division of Biomedical Sciences, Imperial College of Science,*
*Technology and Medicine, Sir Alexander Fleming Building,*
*South Kensington, London SW7 2AZ UK*

# Editors

George E. Tranter, *Glaxo Wellcome Medicines Research,*
*Physical Sciences Research Unit, Gunnells Wood Road,*
*Stevenage, Hertfordshire SG1 2NY, UK*

John L. Holmes, *University of Ottawa,*
*Department of Chemistry, PO Box 450,*
*Stn 4, Ottawa, Canada KIN 6N5*

# Editorial Advisory Board


Laurence D. Barron, *Department of Chemistry,*
*University of Glasgow, Glasgow G12 8QQ, UK*

Andries P. Bruins, *University Centre for Pharmacy,*
*State University, A Deusinglaan 1,*
*Groningen 9713 AV, Netherlands*

C.L. Chakrabarti, *Chemistry Department,*
*Carlton University, Ottawa,*
*Ontario K1S 5B6, Canada*

J. Corset, *Centre National de la Recherche Scientifique,*
*Laboratoire de Spectrochimie Infrarouge et Raman,*
*2 Rue Henri-Dunant, 94320 Thiais, France*

David J. Craik, *Centre for Drug Design & Development,*
*University of Queensland, Brisbane 4072,*
*Queensland, Australia*

James W. Emsley, *Department of Chemistry,*
*University of Southampton, Highfield,*
*Southampton SO17 1BJ UK*

A.S. Gilbert, *19 West Oak, Beckenham,*
*Kent BR3 5EZ, UK*

P.J. Hendra, *Department of Chemistry,*
*University of Southampton, Highfield,*
*Southampton SO9 5NH, UK*

James A. Holcombe, *Department of Chemistry,*
*University of Texas, Austin, Texas 7871-1167, USA*

Harry Kroto, *Department of Chemistry,*
*University of Sussex, Falmer,*
*East Sussex BN1 9QJ, UK*

Reiko Kuroda, *Department of Life Sciences,*
*Graduate School of Arts and Science,*
*The University of Tokyo, Komaba,*
*Tokyo 153, Japan*

N.M.M. Nibbering, *Institute of Mass Spectrometry,*
*University of Amsterdam, Nieuwe Achtergracht 129,*
*1018 WS Amsterdam, The Netherlands*

Ian C.P. Smith, *National Research Council of Canada,*
*Institute of Biodiagnostics, Winnipeg,*
*Manitoba MB R3B 1Y6, Canada*

S.J.B. Tendler, *Department of Pharmaceutical Sciences, University of Nottingham, University Park, Notttingham NG7 2RD, UK*

Georges H. Wagnière, *Physikalisch-Chemisches Institut, der Universität Zärich, Winterhurerstrasse 190 CH-8057 Zärich, Switzerland*

D.J. Watkin, *Chemical Crystallography Laboratory, University of Oxford, 9 Parks Road, Oxford OX1 3PD, UK*

# Acknowledgements

Without a whole host of dedicated people, this encyclopedia would never have come to completion. In these few words I, on behalf of my co-editors, can hope to mention the contributions of only some of those hard working individuals.

Without the active co-operation of the hundreds of scientists who acted as authors for the articles, this encyclopedia would not have been born. We are very grateful to them for endeavouring to write material suitable for an encyclopedia rather than a research paper, which has produced such high-quality entries. We know that all of the people who contributed articles are very busy scientists, many being leaders in their fields, and we thank them. We, as editors, have been ably supported by the members of the Editorial Advisory Board. They made many valuable suggestions for content and authorship in the early planning stages and provided a strong first line of scientific review after the completed articles were received. This encyclopedia covers such a wide range of scientific topics and types of technology that the very varied expertise of the Editorial Advisory Board was particularly necessary.

Next, this work would not have been possible without the vision of Carey Chapman at Academic Press who approached me about 4 years ago with the excellent idea for such an encyclopedia. Four years later, am I still so sure of the usefulness of the encyclopedia? Of course I am, despite the hard work and I am further bolstered by the thought that I might not ever have to see another e-mail from Academic Press. For their work during the commissioning stage and for handling the receipt of manuscripts and dealing with all the authorship problems, we are truly indebted to Lorraine Parry, Colin McNeil and Laura O'Neill who never failed to be considerate, courteous and helpful even under the strongest pressure. I suspect that they are now probably quite expert in spectroscopy. In addition we need to thank Sutapas Bhattacharya who oversaw the project through the production stages and we acknowledge the hard work put in by the copy-editors, the picture researcher and all the other production staff coping with very tight deadlines.

Finally, on a personal note, I should like to acknowledge the close co-operation I have received from my co-editors George Tranter and John Holmes. I think that we made a good team, even if I say it myself.

John Lindon
Imperial College of Science, Technology and Medicine
London
22 April 1999

| **Article Titles** | Authors, Pages |
|---|---|

**A**


| | |
|---|---|
| **Art Works Studied Using IR and Raman Spectroscopy** | Howell G M Edwards, Pages 2-17 |
| **Atmospheric Pressure Ionization in Mass Spectrometry** | W. M. A. Niessen, Pages 18-24 |
| **Atomic Absorption, Methods and Instrumentation** | Steve J Hill and Andy S Fisher, Pages 24-32 |
| **Atomic Absorption, Theory** | Albert Kh Gilmutdinov, Pages 33-42 |
| **Atomic Emission, Methods and Instrumentation** | Sandra L Bonchin, Grace K Zoorob and Joseph A Caruso, Pages 42-50 |
| **Atomic Fluorescence, Methods and Instrumentation** | Steve J Hill and Andy S Fisher, Pages 50-55 |
| **Atomic Spectroscopy, Historical Perspective** | C L Chakrabarti, Pages 56-58 |
| **ATR and Reflectance IR Spectroscopy, Applications** | U P Fringeli, Pages 58-75 |


**B**


| | |
|---|---|
| **Biochemical Applications of Fluorescence Spectroscopy** | Jason B Shear, Pages 77-84 |
| **Biochemical Applications of Mass Spectrometry** | Victor E Vandell and Patrick A Limbach, Pages 84-87 |
| **Biochemical Applications of Raman Spectroscopy** | Peter Hildebrandt and Sophie Lecomte, Pages 88-97 |
| **Biofluids Studied By NMR** | John C Lindon and Jeremy K Nicholson, Pages 98-116 |
| **Biomacromolecular Applications of Circular Dichroism and ORD** | Norma J Greenfield, Pages 117-130 |
| **Biomacromolecular Applications of UV-Visible Absorption Spectroscopy** | Alison Rodger and Karen Sanders, Pages 130-139 |
| **Biomedical Applications of Atomic Spectroscopy** | Andrew Taylor, Pages 139-147 |

## C


| | |
|---|---|
| **$^{13}$C NMR, Methods** | Cecil Dybowski, Alicia Glatfelter and H N Cheng, Pages 149-158 |
| **$^{13}$C NMR, Parameter Survey** | R Duncan Farrant and John C Lindon, Pages 159-165 |
| **Calibration and Reference Systems (Regulatory Authorities)** | C Burgess, Pages 166-171 |
| **Carbohydrates Studied By NMR** | Charles T Weller, Pages 172-180 |
| **Cells Studied By NMR** | Fátima Cruz and Sebastián Cerdán, Pages 180-189 |
| **Chemical Applications of EPR** | Christopher C Rowlands and Damien M Murphy, Pages 190-198 |
| **Chemical Exchange Effects in NMR** | Alex D Bain, Pages 198-207 |
| **Chemical Ionization in Mass Spectrometry** | Alex G Harrison, Pages 207-215 |
| **Chemical Reactions Studied By Electronic Spectroscopy** | Salman R Salman, Pages 216-222 |
| **Chemical Shift and Relaxation Reagents in NMR** | Silvio Aime, Mauro Botta, Mauro Fasano and Enzo Terreno, Pages 223-231 |
| **Chemical Structure Information from Mass Spectrometry** | Kurt Varmuza, Pages 232-243 |
| **Chiroptical Spectroscopy, Emission Theory** | James P Riehl, Pages 243-249 |
| **Chiroptical Spectroscopy, General Theory** | Hans-Georg Kuball, Tatiana Höfer and Stefan Kiesewalter, Pages 250-266 |
| **Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems** | Hans-Georg Kuball and Tatiana Höfer, Pages 267-281 |
| **Chromatography-IR, Applications** | George Jalsovszky, Pages 282-287 |
| **Chromatography-IR, Methods and Instrumentation** | Robert L White, Pages 288-293 |
| **Chromatography-MS, Methods** | W W A Niessen, Pages 293-300 |
| **Chromatography-NMR, Applications** | J P Shockcor, Pages 301-310 |
| **CIDNP Applications** | Tatyana V Leshina, Alexander I Kruppa and Marc B Taraban, Pages 311-318 |

| | |
|---|---|
| **Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism** | Christine L Maupin and James P Riehl, Pages 319-326 |
| **Cluster Ions Measured Using Mass Spectrometry** | O Echt and T D Märk, Pages 327-336 |
| **Colorimetry, Theory** | Alison Gilchrist and Jim Nobbs, Pages 337-343 |
| **Computational Methods and Chemometrics in Near-IR Spectroscopy** | Paul Geladi and Eigil Dåbakk, Pages 343-349 |
| **Contrast Mechanisms in MRI** | I R Young, Pages 349-358 |
| **Cosmochemical Applications Using Mass Spectrometry** | J R De Laeter, Pages 359-367 |

**D**

| | |
|---|---|
| **Diffusion Studied Using NMR Spectroscopy** | Peter Stilbs, Pages 369-375 |
| **Drug Metabolism Studied Using NMR Spectroscopy** | Myriam Malet-Martino and Robert Martino, Pages 375-388 |
| **Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy** | Volker Buss and Lutz Eggers, Pages 388-396 |

**E**

| | |
|---|---|
| **Electromagnetic Radiation** | David L Andrews, Pages 397-401 |
| **Electronic Components, Applications of Atomic Spectroscopy** | John C Lindon, Pages 401-402 |
| **Ellipsometry** | G E Jellison, Jr, Pages 402-411 |
| **Enantiomeric Purity Studied Using NMR** | Thomas J Wenzel, Pages 411-421 |
| **Environmental and Agricultural Applications of Atomic Spectroscopy** | Michael Thompson and Michael H Ramsey, Pages 422-429 |
| **Environmental Applications of Electronic Spectroscopy** | John W Farley, William C Brumley and DeLyle Eastwood, Pages 430-437 |
| **EPR Imaging** | L H Sutcliffe, Pages 437-445 |

| | |
|---|---|
| **EPR Spectroscopy, Theory** | Christopher C Rowlands and Damien M Murphy, Pages 445-456 |
| **EPR, Methods** | Richard Cammack, Pages 457-469 |
| **Exciton Coupling** | Nina Berova, Nobuyuki Harada and Koji Nakanishi, Pages 470-488 |

**F**

| | |
|---|---|
| **$^{19}$F NMR, Applications, Solution State** | Claudio Pettinari and Giovanni Rafaiani, Pages 489-498 |
| **Far-IR Spectroscopy, Applications** | James R Durig, Pages 498-504 |
| **Fast Atom Bombardment Ionization in Mass Spectrometry** | Magda Claeys and Jan Claereboudt, Pages 505-512 |
| **Fibre Optic Probes in Optical Spectroscopy, Clinical Applications** | Urs Utzinger and Rebecca R Richards-Kortum, Pages 512-528 |
| **Fibres and Films Studied Using X-Ray Diffraction** | Watson Fuller and Arumugam Mahendrasingam, Pages 529-539 |
| **Field Ionization Kinetics in Mass Spectrometry** | Nico M M Nibbering, Pages 539-548 |
| **Flame and Temperature Measurement Using Vibrational Spectroscopy** | Kevin L McNesby, Pages 548-559 |
| **Fluorescence and Emission Spectroscopy, Theory** | James A Holcombe, Pages 560-565 |
| **Fluorescence Microscopy, Applications** | Fred Rost, Pages 565-570 |
| **Fluorescence Polarization and Anisotropy** | G E Tranter, Pages 571-573 |
| **Fluorescent Molecular Probes** | F Braut-Boucher and M Aubery, Pages 573-582 |
| **Food and Dairy Products, Applications of Atomic Spectroscopy** | N J Miller-Ihli and S A Baker, Pages 583-592 |
| **Food Science, Applications of Mass Spectrometry** | John P G Wilkins, Pages 592-593 |
| **Food Science, Applications of NMR Spectroscopy** | Brian Hills, Pages 593-601 |
| **Forensic Science, Applications of Atomic Spectroscopy** | John C Lindon, Pages 602-603 |

**Forensic Science, Applications of IR Spectroscopy** Núria Ferrer, Pages 603-615

**Forensic Science, Applications of Mass Spectrometry** Rodger L Foltz, Dennis J Crouch and David M Andrenyak, Pages 615-621

**Forestry and Wood Products, Applications of Atomic Spectroscopy** Cathy Hayes, Pages 621-631

**Fourier Transformation and Sampling Theory** Raúl Curbelo, Pages 632-636

**Fragmentation in Mass Spectrometry** Hans-Friedrich Grützmacher, Pages 637-648

**FT-Raman Spectroscopy, Applications** R H Brody, E A Carter, H. G. M. Edwards and A M Pollard, Pages 649-657


**G**

**Gas Phase Applications of NMR Spectroscopy** Nancy S True, Pages 660-667

**Geology and Mineralogy, Applications of Atomic Spectroscopy** John C Lindon, Page 668

**Glow Discharge Mass Spectrometry, Methods** Annemie Bogaerts, Pages 669-676


**H**

**Halogen NMR Spectroscopy (Excluding $^{19}$F)** Frank G Riddell, Pages 677-684

**Heteronuclear NMR Applications (As, Sb, Bi)** Claudio Pettinari, Fabio Marchetti and Giovanni Rafaiani, Pages 685-690

**Heteronuclear NMR Applications (B, Al, Ga, In, Tl)** Janusz Lewiski, Pages 691-703

**Heteronuclear NMR Applications (Ge, Sn, Pb)** Claudio Pettinari, Pages 704-717

**Heteronuclear NMR Applications (La–Hg)** Trevor G Appleton, Pages 718-722

**Heteronuclear NMR Applications (O, S, Se and Te)** Ioannis P Gerothanassis, Pages 722-729

**Heteronuclear NMR Applications (Sc–Zn)** Dieter Rehder, Pages 731-740

**Heteronuclear NMR Applications (Y–Cd)** Erkki Kolehmainen, Pages 740-750

**High Energy Ion Beam Analysis** Geoff W Grime, Pages 750-760

**High Pressure Studies Using NMR Spectroscopy** Jiri Jonas, Pages 760-771

**High Resolution Electron Energy Loss Spectroscopy, Applications** Horst Conrad and Martin E Kordesch, Pages 772-783

**High Resolution IR Spectroscopy (Gas Phase) Instrumentation** Jyrki K Kauppinen and Jari O Partanen, Pages 784-794

**High Resolution IR Spectroscopy (Gas Phase), Applications** E Canè and A Trombetti, Pages 794-801

**High Resolution Solid State NMR, $^{13}C$** Etsuko Katoh and Isao Ando, Pages 802-813

**High Resolution Solid State NMR, $^{1}H$, $^{19}F$** Anne S Ulrich, Pages 813-825

**Hole Burning Spectroscopy, Methods** Josef Friedrich, Pages 826-836

**Hydrogen Bonding and Other Physicochemical Interactions Studied By IR and Raman Spectroscopy** A S Gilbert, Pages 837-843

**Hyphenated Techniques, Applications of in Mass Spectrometry** W M A Niessen, Pages 843-849

**I**

**In Vivo NMR, Applications, $^{31}P$** Ruth M Dixon and Peter Styles, Pages 851-857

**In Vivo NMR, Applications, Other Nuclei** Jimmy D Bell, E Louise Thomas and K Kumar Changani, Pages 857-865

**In Vivo NMR, Methods** John C Lindon, Pages 866-868

**Induced Circular Dichroism** Kymberley Vickery and Bengt Nordén, Pages 869-874

**Inductively Coupled Plasma Mass Spectrometry, Methods** Diane Beauchemin, Pages 875-880

**Industrial Applications of IR and Raman Spectroscopy** A S Gilbert and R W Lancaster, Pages 881-893

**Inelastic Neutron Scattering, Applications** Stewart F Parker, Pages 894-905

| | |
|---|---|
| **Inelastic Neutron Scattering, Instrumentation** | Stewart F Parker, Pages 905-915 |
| **Inorganic Chemistry, Applications of Mass Spectrometry** | Lev N Sidorov, Pages 915-923 |
| **Inorganic Compounds and Minerals Studied Using X-ray Diffraction** | Gilberto Artioli, Pages 924-933 |
| **Inorganic Condensed Matter, Applications of Luminescence Spectroscopy** | Keith Holliday, Pages 933-943 |
| **Interstellar Molecules, Spectroscopy of** | A G G M Tielens, Pages 943-953 |
| **Ion Collision Theory** | Anil K Shukla and Jean H Futrell, Pages 954-963 |
| **Ion Dissociation Kinetics, Mass Spectrometry** | Bernard Leyh, Pages 963-971 |
| **Ion Energetics in Mass Spectrometry** | John Holmes, Pages 971-976 |
| **Ion Imaging Using Mass Spectrometry** | Albert J R Heck, Pages 976-983 |
| **Ion Molecule Reactions in Mass Spectrometry** | Diethard K Böhme, Pages 984-990 |
| **Ion Structures in Mass Spectrometry** | Peter C Burgers and Johan K Terlouw, Pages 990-1000 |
| **Ion Trap Mass Spectrometers** | Raymond E March, Pages 1000-1009 |
| **Ionization Theory** | C Lifshitz and T D Märk, Pages 1010-1021 |
| **IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds** | Claudio Pettinari and Carlo Santini, Pages 1021-1034 |
| **IR Spectral Group Frequencies of Organic Compounds** | A S Gilbert, Pages 1035-1048 |
| **IR Spectrometers** | R A Spragg, Pages 1048-1057 |
| **IR Spectroscopy Sample Preparation Methods** | R A Spragg, Pages 1058-1066 |
| **IR Spectroscopy, Theory** | Derek Steele, Pages 1066-1071 |
| **Isotope Ratio Studies Using Mass Spectrometry** | Michael E Wieser and Willi A Brand, Pages 1072-1086 |
| **Isotopic Labelling in Mass Spectrometry** | Thomas Hellman Morton, Pages 1086-1096 |

## L


| | |
|---|---|
| **Labelling Studies in Biochemistry Using NMR** | Timothy R Fennell and Susan C J Sumner, Pages 1097-1105 |
| **Laboratory Information Management Systems (LIMS)** | David R McLaughlin and Antony J Williams, Pages 1105-1113 |
| **Laser Applications in Electronic Spectroscopy** | Wolfgang Demtröder, Pages 1113-1123 |
| **Laser Induced Optoacoustic Spectroscopy** | Thomas Gensch, Cristiano Viappiani and Silvia E Braslavsky, Pages 1124-1132 |
| **Laser Magnetic Resonance** | A I Chichinin, Pages 1133-1140 |
| **Laser Spectroscopy Theory** | Luc Van Vaeck and Freddy Adams, Pages 1141-1152 |
| **Laser Spectroscopy Theory** | David L Andrews, Pages 1153-1158 |
| **Light Sources and Optics** | R Magnusson, Pages 1158-1168 |
| **Linear Dichroism, Applications** | Erik W Thulstrup, Jacek Waluk and Jens Spanget-Larsen, Pages 1169-1175 |
| **Linear Dichroism, Instrumentation** | Erik W Thulstrup, Jens Spanget-Larsen and Jacek Waluk, Pages 1176-1178 |
| **Liquid Crystals and Liquid Crystal Solutions Studied By NMR** | Lucia Calucci and Carlo Alberto Veracini, Pages 1179-1186 |
| **Luminescence Theory** | Mohammad A Omary and Howard H Patterson, Pages 1186-1207 |


## M


| | |
|---|---|
| **Macromolecule–Ligand Interactions Studied By NMR** | J Feeney, Pages 1209-1216 |
| **Magnetic Circular Dichroism, Theory** | Laura A Andersson, Pages 1217-1224 |
| **Magnetic Field Gradients in High-Resolution NMR** | Ralph E Hurd, Pages 1224-1232 |
| **Magnetic Resonance, Historical Perspective** | J W Emsley and J Feeney, Pages 1232-1240 |
| **Mass Spectrometry, Historical Perspective** | Allan Maccoll†, Pages 1241-1248 |

| | |
|---|---|
| **Materials Science Applications of X-ray Diffraction** | Åke Kvick, Pages 1248-1257 |
| **Matrix Isolation Studies By IR and Raman Spectroscopies** | Lester Andrews, Pages 1257-1261 |
| **Medical Applications of Mass Spectrometry** | Orval A Mamer, Pages 1262-1271 |
| **Medical Science Applications of IR** | Michael Jackson and Henry H Mantsch, Pages 1271-1281 |
| **Membranes Studied By NMR Spectroscopy** | A Watts and S J Opella, Pages 1281-1291 |
| **Metastable Ions** | John L Holmes, Pages 1291-1297 |
| **Microwave and Radiowave Spectroscopy, Applications** | G Wlodarczak, Pages 1297-1307 |
| **Microwave Spectrometers** | Marlin D Harmony, Pages 1308-1314 |
| **Mössbauer Spectrometers** | Guennadi N Belozerski, Pages 1315-1323 |
| **Mössbauer Spectroscopy, Applications** | Guennadi N Belozerski, Pages 1324-1334 |
| **Mössbauer Spectroscopy, Theory** | Guennadi N Belozerski, Pages 1335-1343 |
| **MRI Applications, Biological** | David G Reid, Paul D Hockings and Paul G M Mullins, Pages 1344-1354 |
| **MRI Applications, Clinical** | Martin O Leach, Pages 1354-1364 |
| **MRI Applications, Clinical Flow Studies** | Y Berthezène, Pages 1365-1372 |
| **MRI Instrumentation** | Paul D Hockings, John F Hare and David G Reid, Pages 1372-1380 |
| **MRI of Oil/Water in Rocks** | Geneviève Guillot, Pages 1380-1387 |
| **MRI Theory** | Ian R Young, Pages 1388-1396 |
| **MRI Using Stray Fields** | Edward W Randall, Pages 1396-1403 |
| **MS-MS and $MS^n$** | W. M. A. Niessen, Pages 1404-1410 |
| **Multiphoton Excitation in Mass Spectrometry** | Ulrich Boesl, Pages 1411-1424 |

**Multiphoton Spectroscopy, Applications** Michael N R Ashfold and Colin M Western, Pages 1424-1433

**Multivariate Statistical Methods** R L Somorjai, Pages 1433-1439

**Muon Spin Resonance Spectroscopy, Applications** Ivan D Reid and Emil Roduner, Pages 1439-1450

**N**

**Near-IR Spectrometers** R Anthony Shaw and Henry H Mantsch, Pages 1451-1461

**Negative Ion Mass Spectrometry, Methods** Suresh Dua and John H Bowie, Pages 1461-1469

**Neutralization–Reionization in Mass Spectrometry** Chrys Wesdemiotis, Pages 1469-1479

**Neutron Diffraction, Instrumentation** A C Hannon, Pages 1479-1492

**Neutron Diffraction, Theory** Alex C Hannon, Pages 1493-1503

**Nitrogen NMR** G A Webb, Pages 1504-1514

**NMR Data Processing** Gareth A Morris, Pages 1514-1521

**NMR in Anisotropic Systems, Theory** J W Emsley, Pages 1521-1527

**NMR Microscopy** Paul T Callaghan, Pages 1528-1537

**NMR of Solids** Jacek Klinowski, Pages 1537-1544

**NMR Principles** P J Hore, Pages 1545-1553

**NMR Pulse Sequences** William F Reynolds, Pages 1554-1567

**NMR Relaxation Rates** Ronald Y Dong, Pages 1568-1575

**NMR Spectrometers** John C Lindon, Pages 1576-1583

**NMR Spectroscopy of Alkali Metal Nuclei in Solution** Frank G Riddell, Pages 1584-1593

| | |
|---|---|
| **Nonlinear Optical Properties** | Georges H Wagnière and Stanisaw Wozniak, Pages 1594-1608 |
| **Nonlinear Raman Spectroscopy, Applications** | W Kiefer, Pages 1609-1623 |
| **Nonlinear Raman Spectroscopy, Instruments** | Peter C Chen, Pages 1624-1631 |
| **Nonlinear Raman Spectroscopy, Theory** | J Santos Gómez, Pages 1631-1642 |
| **Nuclear Overhauser Effect** | Anil Kumar and R Christy Rani Grace, Pages 1643-1653 |
| **Nuclear Quadrupole Resonance, Applications** | Oleg Kh Poleshchuk and Jolanta N Latosiska, Pages 1653-1662 |
| **Nuclear Quadrupole Resonance, Instrumentation** | Taras N Rudakov, Pages 1663-1671 |
| **Nuclear Quadrupole Resonance, Theory** | Janez Seliger, Pages 1672-1680 |
| **Nucleic Acids and Nucleotides Studied Using Mass Spectrometry** | Tracey A Simmons, Kari B Green-Church and Patrick A Limbach, Pages 1681-1688 |
| **Nucleic Acids Studied Using NMR** | John C Lindon, Pages 1688-1689 |

**O**

| | |
|---|---|
| **Optical Frequency Conversion** | Christos Flytzanis, Pages 1691-1701 |
| **Optical Spectroscopy, Linear Polarization Theory** | Josef Michl, Pages 1701-1712 |
| **ORD and Polarimetry Instruments** | Harry G Brittain, Pages 1712-1718 |
| **Organic Chemistry Applications of Fluorescence Spectroscopy** | Stephen G Schulman, Qiao Qing Di and John Juchum, Pages 1718-1725 |
| **Organometallics Studied Using Mass Spectrometry** | Dmitri V Zagorevskii, Pages 1726-1733 |

**P**


| | |
|---|---|
| **$^{31}$P NMR** | David G Gorenstein and Bruce A Luxon, Pages 1735-1744 |
| **Parameters in NMR Spectroscopy, Theory of** | G A Webb, Pages 1745-1753 |
| **Peptides and Proteins Studied Using Mass Spectrometry** | Michael A Baldwin, Pages 1753-1763 |
| **Perfused Organs Studied Using NMR Spectroscopy** | John C Docherty, Pages 1763-1770 |
| **PET, Methods and Instrumentation** | T J Spinks, Pages 1771-1782 |
| **PET, Theory** | T J Spinks, Pages 1782-1791 |
| **Pharmaceutical Applications of Atomic Spectroscopy** | Nancy S Lewen and Martha M Schenkenberger, Pages 1791-1800 |
| **Photoacoustic Spectroscopy, Applications** | Markus W Sigrist, Pages 1800-1809 |
| **Photoacoustic Spectroscopy, Methods and Instrumentation** | Markus W Sigrist, Pages 1810-1814 |
| **Photoacoustic Spectroscopy, Theory** | András Miklós, Stefan Schäfer and Peter Hess, Pages 1815-1822 |
| **Photoelectron Spectrometers** | László Szepes and György Tarczay, Pages 1822-1830 |
| **Photoelectron Spectroscopy** | John Holmes, Page 1831 |
| **Photoelectron–Photoion Coincidence Methods in Mass Spectrometry (PEPICO)** | Tomas Baer, Pages 1831-1839 |
| **Photoionization and Photodissociation Methods in Mass Spectrometry** | John C Traeger, Pages 1840-1847 |
| **Plasma Desorption Ionization in Mass Spectrometry** | Ronald D Macfarlane, Pages 1848-1857 |
| **Polymer Applications of IR and Raman Spectroscopy** | C M Snively and J L Koenig, Pages 1858-1864 |
| **Powder X-Ray Diffraction, Applications** | Daniel Louër, Pages 1865-1875 |
| **Product Operator Formalism in NMR** | Timothy J Norwood, Pages 1875-1884 |
| **Proteins Studied Using NMR Spectroscopy** | Paul N Sanderson, Pages 1885-1893 |

**Proton Affinities** Edward P L Hunter and Sharon G Lias, Pages 1893-1901

**Proton Microprobe (Method and Background)** Geoff W Grime, Pages 1901-1905

**Pyrolysis Mass Spectrometry, Methods** Jacek P Dworzanski and Henk L C Meuzelaar, Pages 1906-1919

**Q**

**Quadrupoles, Use of in Mass Spectrometry** P H Dawson and D J Douglas, Pages 1921-1930

**Quantitative Analysis** T Frost, Pages 1931-1936

**R**

**Radiofrequency Field Gradients in NMR, Theory** Daniel Canet, Pages 1937-1944

**Raman and Infrared Microspectroscopy** Pina Colarusso, Linda H Kidder, Ira W Levin and E Neil Lewis, Pages 1945-1954

**Raman Optical Activity, Applications** Günter Georg Hoffmann, Pages 1955-1965

**Raman Optical Activity, Spectrometers** Werner Hug, Pages 1966-1976

**Raman Optical Activity, Theory** Laurence A Nafie, Pages 1976-1985

**Raman Spectrometers** Bernhard Schrader, Pages 1986-1992

**Rayleigh Scattering and Raman Spectroscopy, Theory** David L Andrews, Pages 1993-2000

**Relaxometers** Ralf-Oliver Seitter and Rainer Kimmich, Pages 2000-2008

**Rigid Solids Studied Using MRI** David G Cory, Pages 2009-2017

**Rotational Spectroscopy, Theory** Iain R McNab, Pages 2017-2028

S


| | |
|---|---|
| **Scanning Probe Microscopes** | J G Kushmerick and P S Weiss, Pages 2043-2051 |
| **Scanning Probe Microscopy, Applications** | C J Roberts, M C Davies, S J B Tendler and P M Williams, Pages 2051-2059 |
| **Scanning Probe Microscopy, Theory** | A J Fisher, Pages 2060-2066 |
| **Scattering and Particle Sizing, Applications** | F Ross Hallett, Pages 2067-2074 |
| **Scattering Theory** | Michael Kotlarchyk, Pages 2074-2084 |
| **Sector Mass Spectrometers** | R Bateman, Pages 2085-2092 |
| **$^{29}$Si NMR** | Heinrich C Marsmann, Pages 2031-2042 |
| **SIFT Applications in Mass Spectrometry** | David Smith and Patrik panl, Pages 2092-2105 |
| **Small Molecule Applications of X-Ray Diffraction** | Andrei S Batsanov, Pages 2106-2115 |
| **Solid State NMR, Methods** | J W Zwanziger and H W Spiess, Pages 2128-2136 |
| **Solid-State NMR Using Quadrupolar Nuclei** | Alejandro C Olivieri, Pages 2116-2127 |
| **Solid-State NMR, Rotational Resonance** | David L Bryce and Roderick E Wasylishen, Pages 2136-2144 |
| **Solvent Suppression Methods in NMR Spectroscopy** | Maili Liu and Xi-an Mao, Pages 2145-2152 |
| **Sonically Induced NMR Methods** | John Homer, Pages 2152-2159 |
| **SPECT, Methods and Instrumentation** | John C Lindon, Pages 2159-2161 |
| **Spectroelectrochemistry, Applications** | R J Mortimer, Pages 2161-2174 |
| **Spectroelectrochemistry, Methods and Instrumentation** | Roger J Mortimer, Pages 2174-2181 |
| **Spectroscopy of Ions** | John P Maier, Pages 2182-2189 |
| **Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy** | Carmen M Arroyo, Pages 2189-2198 |

| | |
|---|---|
| **Stars, Spectroscopy of** | A G G M Tielens, Pages 2199-2204 |
| **Statistical Theory of Mass Spectra** | J C Lorquet, Pages 2204-2211 |
| **Stereochemistry Studied Using Mass Spectrometry** | Asher Mandelbaum, Pages 2211-2223 |
| **Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules** | G E Hawkes, Pages 2224-2233 |
| **Structural Chemistry Using NMR Spectroscopy, Organic Molecules** | Cynthia K McClure, Pages 2234-2245 |
| **Structural Chemistry Using NMR Spectroscopy, Peptides** | Martin Huenges and Horst Kessler, Pages 2246-2260 |
| **Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals** | Alexandros Makriyannis and Spiro Pavlopoulos, Pages 2261-2271 |
| **Structure Refinement (Solid State Diffraction)** | Dieter Schwarzenbach and Howard D Flack, Pages 2271-2278 |
| **Surface Induced Dissociation in Mass Spectrometry** | S A Miller and S L Bernasek, Pages 2279-2294 |
| **Surface Plasmon Resonance, Applications** | Zdzislaw Salamon and Gordon Tollin, Pages 2294-2302 |
| **Surface Plasmon Resonance, Instrumentation** | R P H Kooyman, Pages 2302-2310 |
| **Surface Plasmon Resonance, Theory** | Zdzislaw Salamon and Gordon Tollin, Pages 2311-2319 |
| **Surface Studies By IR Spectroscopy** | Norman Sheppard, Pages 2320-2328 |
| **Surface-Enhanced Raman Scattering (SERS), Applications** | W E Smith and C Rodger, Pages 2329-2334 |
| **Symmetry in Spectroscopy, Effects of** | S F A Kettle, Pages 2335-2339 |

**T**

| | |
|---|---|
| **Tensor Representations** | Peter Herzig and Rainer Dirl, Pages 2342-2353 |
| **Thermospray Ionization in Mass Spectrometry** | W M A Niessen, Pages 2353-2360 |
| **Time of Flight Mass Spectrometers** | K G Standing and W Ens, Pages 2360-2365 |

**Tritium NMR, Applications** John R Jones, Pages 2366-2369

**Two-Dimensional NMR, Methods** Peter L Rinaldi, Pages 2370-2381

**U**

**UV-Visible Absorption and Fluorescence Spectrometers** G E Tranter, Pages 2383-2389

**V**

**Vibrational CD Spectrometers** Laurence A Nafie, Pages 2391-2402

**Vibrational CD, Applications** Günter Georg Hoffmann, Pages 2403-2414

**Vibrational CD, Theory** Philip J Stephens, Pages 2415-2421

**Vibrational, Rotational and Raman Spectroscopy, Historical Perspective** A S Gilbert, Pages 2422-2432

**X**

**X-ray Absorption Spectrometers** Grant Bunker, Pages 2447-2453

**X-ray Emission Spectroscopy, Applications** George N Dolenko, Oleg Kh Poleshchuk and Jolanta N Latoiska, Pages 2455-2462

**X-ray Emission Spectroscopy, Methods** George N Dolenko, Oleg Kh Poleshchuk and Jolanta N Latoiska, Pages 2463-2467

**X-ray Fluorescence Spectrometers** Utz Kramar, Pages 2467-2477

**X-ray Fluorescence Spectroscopy, Applications** Christina Streli, P Wobrauschek and P Kregsamer, Pages 2478-2487

**X-ray Spectroscopy, Theory** Prasad A Naik, Pages 2487-2498

**Xenon NMR Spectroscopy** Jukka Jokisaari, Pages 2435-2446

**Z**


**Zeeman and Stark Methods in Spectroscopy, Applications** — Ichita Endo and Masataka Linuma, Pages 2501-2504

**Zeeman and Stark Methods in Spectroscopy, Instrumentation** — Ichita Endo and Masataka Linuma, Pages 2505-2509

**Zero Kinetic Energy Photoelectron Spectroscopy, Applications** — K Müller-Dethlefs and Mark Ford, Pages 2509-2519

**Zero Kinetic Energy Photoelectron Spectroscopy, Theory** — K Müller-Dethlefs and Mark Ford, Pages 2519-2526


**APPENDICES**


**Appendix 1. Periodic Table of Elements** — Page 2528

**Appendix 2. Tables of SI and Related Units** — Pages 2529-2530

**Appendix 3. Wavelength Scale** — Page 2531

**Appendix 4. Colour, Wave Length, Frequency, Wave Number and Energy of Light** — Page 2532

**Appendix 5. Magnetic Susceptibilities at 25°C** — Page 2532

**Appendix 6. Electronic Configuration of Elements** — Pages 2533-2534

**Appendix 7. Properties of some Important Solvents** — Pages 2535-2536

**Appendix 8. Important Acronyms in Organic Chemistry** — Pages 2537-2538

**Appendix 9. Equilibrium Constants at 25°c/concentration Units for Solutions** — Page 2539

**Appendix 10. Acronyms and Abbreviations in Quantum Chemistry and Related Fields** — Page 2540

**Appendix 11. Standard Potentials in Aqueous Solutions** — Pages 2541-2544

**Appendix 12. Typical UV Absorptions of Unconjugated Chromophores** — Page 2545

**Appendix 13. Typical UV Absorption Maxima of Substituted Benzenes** — Page 2546

| | |
|---|---|
| **Appendix 14. Typical UV Absorption Maxima of Aromatic and Heteroaromatic Compounds** | Page 2546 |
| **Appendix 15. Common Isotopes for Mössbauer Spectroscopy** | Page 2547 |
| **Appendix 16. NMR Frequency Table** | Pages 2548-2551 |
| **Appendix 17. $^{19}F$ and $^{31}P$ NMR Chemical Shifts** | Page 2552 |
| **Appendix 18. Chemical Shift Ranges and Standards for Selected Nuclei** | Pages 2552-2553 |
| **Appendix 19. Abbreviations and Acronyms used in Magnetic Resonance** | Pages 2553-2556 |
| **Appendix 20. Symbols Used in Magnetic Resonance** | Pages 2556-2557 |
| **Appendix 21. EPR/ENDOR Frequency Table** | Pages 2557-2560 |
| **Appendix 22. Some Useful Conversion Factors in EPR** | Page 2560 |
| **Appendix 23. Mass Spectrometry: Atomic Weights.** | Pages 2561-2563 |
| **Appendix 24. Conversion Table of Transmittance and Absorbanceunits** | Page 2564 |
| **Appendix 25. Conversion Table of Energy and Wavelength Units** | Pages 2565-2566 |
| **Appendix 26. Optical Components used in FT-IR-Spectroscopy** | Page 2567 |
| **Appendix 27. Infrared and Raman Tables** | Pages 2568-2571 |
| **Appendix 28. Selected Force Constants and Bond Orders (According To Siebert) of Organic and Inorganic Compounds** | Pages 2572-2573 |
| **Appendix 29. Fundamental Physical Constants** | Page 2574 |
| **Appendix 30. List Of Suppliers** | Pages 2575-2581 |

# Subject Classification

## Atomic Spectroscopy

### *Historical Overview*


| | | |
|---|---|---|
| **Atomic Spectroscopy, Historical Perspective** | *C L Chakrabarti* | Pages 56-58 |


### *Theory*


| | | |
|---|---|---|
| **Atomic Absorption, Theory** | *Albert Kh Gilmutdinov* | Pages 33-42 |
| **Fluorescence and Emission Spectroscopy, Theory** | *James A Holcombe* | Pages 560-565 |


### *Methods and Instrumentation*


| | | |
|---|---|---|
| **Atomic Absorption, Methods and Instrumentation** | *Steve J Hill and Andy S Fisher* | Pages 24-32 |
| **Atomic Emission, Methods and Instrumentation** | *Sandra L Bonchin, Grace K Zoorob and Joseph A Caruso* | Pages 42-50 |
| **Atomic Fluorescence, Methods and Instrumentation** | *Steve J Hill and Andy S Fisher* | Pages 50-55 |


### *Applications*


| | | |
|---|---|---|
| **Biomedical Applications of Atomic Spectroscopy** | *Andrew Taylor* | Pages 139-147 |
| **Electronic Components, Applications of Atomic Spectroscopy** | *John C Lindon* | Pages 401-402 |
| **Environmental and Agricultural Applications of Atomic Spectroscopy** | *Michael Thompson and Michael H Ramsey* | Pages 422-429 |

| | | |
|---|---|---|
| **Food and Dairy Products, Applications of Atomic Spectroscopy** | *N J Miller-Ihli and S A Baker* | Pages 583-592 |
| **Forensic Science, Applications of Atomic Spectroscopy** | *John C Lindon* | Pages 602-603 |
| **Forestry and Wood Products, Applications of Atomic Spectroscopy** | *Cathy Hayes* | Pages 621-631 |
| **Geology and Mineralogy, Applications of Atomic Spectroscopy** | *John C Lindon* | Page 668 |
| **Pharmaceutical Applications of Atomic Spectroscopy** | *Nancy S Lewen and Martha M Schenkenberger* | Pages 1791-1800 |


## Electronic Spectroscopy

### *Theory*


| | | |
|---|---|---|
| **Chiroptical Spectroscopy, Emission Theory** | *James P Riehl* | Pages 243-249 |
| **Chiroptical Spectroscopy, General Theory** | *Hans-Georg Kuball, Tatiana Höfer and Stefan Kiesewalter* | Pages 250-266 |
| **Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems** | *Hans-Georg Kuball and Tatiana Höfer* | Pages 267-281 |
| **Colorimetry, Theory** | *Alison Gilchrist and Jim Nobbs* | Pages 337-343 |
| **Fluorescence Polarization and Anisotropy** | *G E Tranter* | Pages 571-573 |
| **Laser Spectroscopy Theory** | *David L Andrews* | Pages 1153-1158 |
| **Luminescence Theory** | *Mohammad A Omary and Howard H Patterson* | Pages 1186-1207 |

| | | |
|---|---|---|
| **Magnetic Circular Dichroism, Theory** | *Laura A Andersson* | Pages 1217-1224 |
| **Nonlinear Optical Properties** | *Georges H Wagnière and Stanisaw Wozniak* | Pages 1594-1608 |
| **Optical Spectroscopy, Linear Polarization Theory** | *Josef Michl* | Pages 1701-1712 |
| **Photoacoustic Spectroscopy, Theory** | *András Miklós, Stefan Schäfer and Peter Hess* | Pages 1815-1822 |
| **Scattering Theory** | *Michael Kotlarchyk* | Pages 2074-2084 |

***Theory and Applications***

| | | |
|---|---|---|
| **Exciton Coupling** | *Nina Berova, Nobuyuki Harada and Koji Nakanishi* | Pages 470-488 |

***Methods and Instrumentation***

| | | |
|---|---|---|
| **Fluorescent Molecular Probes** | *F Braut-Boucher and M Aubery* | Pages 573-582 |
| **Linear Dichroism, Instrumentation** | *Erik W Thulstrup, Jens Spanget-Larsen and Jacek Waluk* | Pages 1176-1178 |
| **Optical Frequency Conversion** | *Christos Flytzanis* | Pages 1691-1701 |
| **ORD and Polarimetry Instruments** | *Harry G Brittain* | Pages 1712-1718 |
| **Photoacoustic Spectroscopy, Methods and Instrumentation** | *Markus W Sigrist* | Pages 1810-1814 |
| **Spectroelectrochemistry, Methods and Instrumentation** | *Roger J Mortimer* | Pages 2174-2181 |
| **UV-Visible Absorption and Fluorescence Spectrometers** | *G E Tranter* | Pages 2383-2389 |
| **Zeeman and Stark Methods in Spectroscopy, Instrumentation** | *Ichita Endo and Masataka Linuma* | Pages 2505-2509 |

*Applications*


| | | |
|---|---|---|
| **Biochemical Applications of Fluorescence Spectroscopy** | *Jason B Shear* | Pages 77-84 |
| **Biomacromolecular Applications of Circular Dichroism and ORD** | *Norma J Greenfield* | Pages 117-130 |
| **Biomacromolecular Applications of UV-Visible Absorption Spectroscopy** | *Alison Rodger and Karen Sanders* | Pages 130-139 |
| **Chemical Reactions Studied By Electronic Spectroscopy** | *Salman R Salman* | Pages 216-222 |
| **Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism** | *Christine L Maupin and James P Riehl* | Pages 319-326 |
| **Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy** | *Volker Buss and Lutz Eggers* | Pages 388-396 |
| **Ellipsometry** | *G E Jellison, Jr* | Pages 402-411 |
| **Environmental Applications of Electronic Spectroscopy** | *John W Farley, William C Brumley and DeLyle Eastwood* | Pages 430-437 |
| **Fibre Optic Probes in Optical Spectroscopy, Clinical Applications** | *Urs Utzinger and Rebecca R Richards-Kortum* | Pages 512-528 |
| **Fluorescence Microscopy, Applications** | *Fred Rost* | Pages 565-570 |
| **Induced Circular Dichroism** | *Kymberley Vickery and Bengt Nordén* | Pages 869-874 |
| **Inorganic Condensed Matter, Applications of Luminescence Spectroscopy** | *Keith Holliday* | Pages 933-943 |
| **Interstellar Molecules, Spectroscopy of** | *A G G M Tielens* | Pages 943-953 |

| | | |
|---|---|---|
| **Laser Applications in Electronic Spectroscopy** | *Wolfgang Demtröder* | Pages 1113-1123 |
| **Laser Induced Optoacoustic Spectroscopy** | *Thomas Gensch, Cristiano Viappiani and Silvia E Braslavsky* | Pages 1124-1132 |
| **Linear Dichroism, Applications** | *Erik W Thulstrup, Jacek Waluk and Jens Spanget-Larsen* | Pages 1169-1175 |
| **Multiphoton Spectroscopy, Applications** | *Michael N R Ashfold and Colin M Western* | Pages 1424-1433 |
| **Organic Chemistry Applications of Fluorescence Spectroscopy** | *Stephen G Schulman, Qiao Qing Di and John Juchum* | Pages 1718-1725 |
| **Photoacoustic Spectroscopy, Applications** | *Markus W Sigrist* | Pages 1800-1809 |
| **Scattering and Particle Sizing, Applications** | *F Ross Hallett* | Pages 2067-2074 |
| **Spectroelectrochemistry, Applications** | *R J Mortimer* | Pages 2161-2174 |
| **Stars, Spectroscopy of** | *A G G M Tielens* | Pages 2199-2204 |
| **Zeeman and Stark Methods in Spectroscopy, Applications** | *Ichita Endo and Masataka Linuma* | Pages 2501-2504 |



## Fundamentals of Spectroscopy

### *Theory*



| | | |
|---|---|---|
| **Electromagnetic Radiation** | *David L Andrews* | Pages 397-401 |
| **Fourier Transformation and Sampling Theory** | *Raúl Curbelo* | Pages 632-636 |
| **Symmetry in Spectroscopy, Effects of** | *S F A Kettle* | Pages 2335-2339 |

| | | |
|---|---|---|
| **Tensor Representations** | *Peter Herzig and Rainer Dirl* | Pages 2342-2353 |

***Methods and Instrumentation***

| | | |
|---|---|---|
| **Calibration and Reference Systems (Regulatory Authorities)** | *C Burgess* | Pages 166-171 |
| **Laboratory Information Management Systems (LIMS)** | *David R McLaughlin and Antony J Williams* | Pages 1105-1113 |
| **Light Sources and Optics** | *R Magnusson* | Pages 1158-1168 |
| **Multivariate Statistical Methods** | *R L Somorjai* | Pages 1433-1439 |
| **Quantitative Analysis** | *T Frost* | Pages 1931-1936 |

**High Energy Spectroscopy**

***Theory***

| | | |
|---|---|---|
| **Mössbauer Spectroscopy, Theory** | *Guennadi N Belozerski* | Pages 1335-1343 |
| **Neutron Diffraction, Theory** | *Alex C Hannon* | Pages 1493-1503 |
| **Photoelectron Spectroscopy** | *John Holmes* | Page 1831 |
| **X-ray Spectroscopy, Theory** | *Prasad A Naik* | Pages 2487-2498 |
| **Zero Kinetic Energy Photoelectron Spectroscopy, Theory** | *K Müller-Dethlefs and Mark Ford* | Pages 2519-2526 |

***Methods and Instrumentation***

| | | |
|---|---|---|
| **High Energy Ion Beam Analysis** | *Geoff W Grime* | Pages 750-760 |

| | | |
|---|---|---|
| **Hole Burning Spectroscopy, Methods** | *Josef Friedrich* | Pages 826-836 |
| **Inelastic Neutron Scattering, Instrumentation** | *Stewart F Parker* | Pages 905-915 |
| **Mössbauer Spectrometers** | *Guennadi N Belozerski* | Pages 1315-1323 |
| **Neutron Diffraction, Instrumentation** | *A C Hannon* | Pages 1479-1492 |
| **Photoelectron Spectrometers** | *László Szepes and György Tarczay* | Pages 1822-1830 |
| **Proton Microprobe (Method and Background)** | *Geoff W Grime* | Pages 1901-1905 |
| **Structure Refinement (Solid State Diffraction)** | *Dieter Schwarzenbach and Howard D Flack* | Pages 2271-2278 |
| **X-ray Absorption Spectrometers** | *Grant Bunker* | Pages 2447-2453 |
| **X-ray Emission Spectroscopy, Methods** | *George N Dolenko, Oleg Kh Poleshchuk and Jolanta N Latoiska* | Pages 2463-2467 |
| **X-ray Fluorescence Spectrometers** | *Utz Kramar* | Pages 2467-2477 |

***Applications***

| | | |
|---|---|---|
| **Fibres and Films Studied Using X-Ray Diffraction** | *Watson Fuller and Arumugam Mahendrasingam* | Pages 529-539 |
| **Inelastic Neutron Scattering, Applications** | *Stewart F Parker* | Pages 894-905 |
| **Inorganic Compounds and Minerals Studied Using X-ray Diffraction** | *Gilberto Artioli* | Pages 924-933 |
| **Materials Science Applications of X-ray Diffraction** | *Åke Kvick* | Pages 1248-1257 |
| **Mössbauer Spectroscopy, Applications** | *Guennadi N Belozerski* | Pages 1324-1334 |

| | | |
|---|---|---|
| **Powder X-Ray Diffraction, Applications** | *Daniel Louër* | Pages 1865-1875 |
| **Small Molecule Applications of X-Ray Diffraction** | *Andrei S Batsanov* | Pages 2106-2115 |
| **X-ray Emission Spectroscopy, Applications** | *George N Dolenko, Oleg Kh Poleshchuk and Jolanta N Latoiska* | Pages 2455-2462 |
| **X-ray Fluorescence Spectroscopy, Applications** | *Christina Streli, P Wobrauschek and P Kregsamer* | Pages 2478-2487 |
| **Zero Kinetic Energy Photoelectron Spectroscopy, Applications** | *K Müller-Dethlefs and Mark Ford* | Pages 2509-2519 |


## Magnetic Resonance

### *Historical Overview*


| | | |
|---|---|---|
| **Magnetic Resonance, Historical Perspective** | *J W Emsley and J Feeney* | Pages 1232-1240 |


### *Theory*


| | | |
|---|---|---|
| **Chemical Exchange Effects in NMR** | *Alex D Bain* | Pages 198-207 |
| **Contrast Mechanisms in MRI** | *I R Young* | Pages 349-358 |
| **EPR Spectroscopy, Theory** | *Christopher C Rowlands and Damien M Murphy* | Pages 445-456 |
| **Magnetic Field Gradients in High-Resolution NMR** | *Ralph E Hurd* | Pages 1224-1232 |
| **MRI Theory** | *Ian R Young* | Pages 1388-1396 |
| **NMR in Anisotropic Systems, Theory** | *J W Emsley* | Pages 1521-1527 |

| | | |
|---|---|---|
| **NMR Principles** | *P J Hore* | Pages 1545-1553 |
| **NMR Pulse Sequences** | *William F Reynolds* | Pages 1554-1567 |
| **NMR Relaxation Rates** | *Ronald Y Dong* | Pages 1568-1575 |
| **Nuclear Overhauser Effect** | *Anil Kumar and R Christy Rani Grace* | Pages 1643-1653 |
| **Nuclear Quadrupole Resonance, Theory** | *Janez Seliger* | Pages 1672-1680 |
| **Parameters in NMR Spectroscopy, Theory of** | *G A Webb* | Pages 1745-1753 |
| **Product Operator Formalism in NMR** | *Timothy J Norwood* | Pages 1875-1884 |
| **Radiofrequency Field Gradients in NMR, Theory** | *Daniel Canet* | Pages 1937-1944 |


***Methods and Instrumentation***


| | | |
|---|---|---|
| **$^{13}$C NMR, Methods** | *Cecil Dybowski, Alicia Glatfelter and H N Cheng* | Pages 149-158 |
| **EPR, Methods** | *Richard Cammack* | Pages 457-469 |
| **In Vivo NMR, Methods** | *John C Lindon* | Pages 866-868 |
| **Laser Magnetic Resonance** | *A I Chichinin* | Pages 1133-1140 |
| **MRI Instrumentation** | *Paul D Hockings, John F Hare and David G Reid* | Pages 1372-1380 |
| **NMR Data Processing** | *Gareth A Morris* | Pages 1514-1521 |
| **NMR Microscopy** | *Paul T Callaghan* | Pages 1528-1537 |
| **NMR Spectrometers** | *John C Lindon* | Pages 1576-1583 |
| **Nuclear Quadrupole Resonance, Instrumentation** | *Taras N Rudakov* | Pages 1663-1671 |

| | | |
|---|---|---|
| **Relaxometers** | *Ralf-Oliver Seitter and Rainer Kimmich* | Pages 2000-2008 |
| **Solid State NMR, Methods** | *J W Zwanziger and H W Spiess* | Pages 2128-2136 |
| **Solvent Suppression Methods in NMR Spectroscopy** | *Maili Liu and Xi-an Mao* | Pages 2145-2152 |
| **Sonically Induced NMR Methods** | *John Homer* | Pages 2152-2159 |
| **Two-Dimensional NMR, Methods** | *Peter L Rinaldi* | Pages 2370-2381 |

### *Applications*

| | | |
|---|---|---|
| **Biofluids Studied By NMR** | *John C Lindon and Jeremy K Nicholson* | Pages 98-116 |
| **$^{13}$C NMR, Parameter Survey** | *R Duncan Farrant and John C Lindon* | Pages 159-165 |
| **Carbohydrates Studied By NMR** | *Charles T Weller* | Pages 172-180 |
| **Cells Studied By NMR** | *Fátima Cruz and Sebastián Cerdán* | Pages 180-189 |
| **Chemical Applications of EPR** | *Christopher C Rowlands and Damien M Murphy* | Pages 190-198 |
| **Chemical Shift and Relaxation Reagents in NMR** | *Silvio Aime, Mauro Botta, Mauro Fasano and Enzo Terreno* | Pages 223-231 |
| **Chromatography-NMR, Applications** | *J P Shockcor* | Pages 301-310 |
| **CIDNP Applications** | *Tatyana V Leshina, Alexander I Kruppa and Marc B Taraban* | Pages 311-318 |
| **Diffusion Studied Using NMR Spectroscopy** | *Peter Stilbs* | Pages 369-375 |
| **Drug Metabolism Studied Using NMR Spectroscopy** | *Myriam Malet-Martino and Robert Martino* | Pages 375-388 |
| **Enantiomeric Purity Studied Using NMR** | *Thomas J Wenzel* | Pages 411-421 |

| | | |
|---|---|---|
| **EPR Imaging** | *L H Sutcliffe* | Pages 437-445 |
| **$^{19}$F NMR, Applications, Solution State** | *Claudio Pettinari and Giovanni Rafaiani* | Pages 489-498 |
| **Food Science, Applications of NMR Spectroscopy** | *Brian Hills* | Pages 593-601 |
| **Gas Phase Applications of NMR Spectroscopy** | *Nancy S True* | Pages 660-667 |
| **Halogen NMR Spectroscopy (Excluding $^{19}$F)** | *Frank G Riddell* | Pages 677-684 |
| **Heteronuclear NMR Applications (As, Sb, Bi)** | *Claudio Pettinari, Fabio Marchetti and Giovanni Rafaiani* | Pages 685-690 |
| **Heteronuclear NMR Applications (B, Al, Ga, In, Tl)** | *Janusz Lewiski* | Pages 691-703 |
| **Heteronuclear NMR Applications (Ge, Sn, Pb)** | *Claudio Pettinari* | Pages 704-717 |
| **Heteronuclear NMR Applications (La–Hg)** | *Trevor G Appleton* | Pages 718-722 |
| **Heteronuclear NMR Applications (O, S, Se and Te)** | *Ioannis P Gerothanassis* | Pages 722-729 |
| **Heteronuclear NMR Applications (Sc–Zn)** | *Dieter Rehder* | Pages 731-740 |
| **Heteronuclear NMR Applications (Y–Cd)** | *Erkki Kolehmainen* | Pages 740-750 |
| **High Pressure Studies Using NMR Spectroscopy** | *Jiri Jonas* | Pages 760-771 |
| **High Resolution Solid State NMR, $^{13}$C** | *Etsuko Katoh and Isao Ando* | Pages 802-813 |
| **High Resolution Solid State NMR, $^{1}$H, $^{19}$F** | *Anne S Ulrich* | Pages 813-825 |
| **In Vivo NMR, Applications, $^{31}$P** | *Ruth M Dixon and Peter Styles* | Pages 851-857 |
| **In Vivo NMR, Applications, Other Nuclei** | *Jimmy D Bell, E Louise Thomas and K Kumar Changani* | Pages 857-865 |
| **Labelling Studies in Biochemistry Using NMR** | *Timothy R Fennell and Susan C J Sumner* | Pages 1097-1105 |

| | | |
|---|---|---|
| **Liquid Crystals and Liquid Crystal Solutions Studied By NMR** | *Lucia Calucci and Carlo Alberto Veracini* | Pages 1179-1186 |
| **Macromolecule–Ligand Interactions Studied By NMR** | *J Feeney* | Pages 1209-1216 |
| **Membranes Studied By NMR Spectroscopy** | *A Watts and S J Opella* | Pages 1281-1291 |
| **MRI Applications, Biological** | *David G Reid, Paul D Hockings and Paul G M Mullins* | Pages 1344-1354 |
| **MRI Applications, Clinical** | *Martin O Leach* | Pages 1354-1364 |
| **MRI Applications, Clinical Flow Studies** | *Y Berthezène* | Pages 1365-1372 |
| **MRI of Oil/Water in Rocks** | *Geneviève Guillot* | Pages 1380-1387 |
| **MRI Using Stray Fields** | *Edward W Randall* | Pages 1396-1403 |
| **Muon Spin Resonance Spectroscopy, Applications** | *Ivan D Reid and Emil Roduner* | Pages 1439-1450 |
| **Nitrogen NMR** | *G A Webb* | Pages 1504-1514 |
| **NMR of Solids** | *Jacek Klinowski* | Pages 1537-1544 |
| **NMR Spectroscopy of Alkali Metal Nuclei in Solution** | *Frank G Riddell* | Pages 1584-1593 |
| **Nuclear Quadrupole Resonance, Applications** | *Oleg Kh Poleshchuk and Jolanta N Latosiska* | Pages 1653-1662 |
| **Nucleic Acids Studied Using NMR** | *John C Lindon* | Pages 1688-1689 |
| **$^{31}$P NMR** | *David G Gorenstein and Bruce A Luxon* | Pages 1735-1744 |
| **Perfused Organs Studied Using NMR Spectroscopy** | *John C Docherty* | Pages 1763-1770 |
| **Proteins Studied Using NMR Spectroscopy** | *Paul N Sanderson* | Pages 1885-1893 |

| | | |
|---|---|---|
| **Rigid Solids Studied Using MRI** | *David G Cory* | Pages 2009-2017 |
| **$^{29}$Si NMR** | *Heinrich C Marsmann* | Pages 2031-2042 |
| **Solid-State NMR Using Quadrupolar Nuclei** | *Alejandro C Olivieri* | Pages 2116-2127 |
| **Solid-State NMR, Rotational Resonance** | *David L Bryce and Roderick E Wasylishen* | Pages 2136-2144 |
| **Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy** | *Carmen M Arroyo* | Pages 2189-2198 |
| **Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules** | *G E Hawkes* | Pages 2224-2233 |
| **Structural Chemistry Using NMR Spectroscopy, Organic Molecules** | *Cynthia K McClure* | Pages 2234-2245 |
| **Structural Chemistry Using NMR Spectroscopy, Peptides** | *Martin Huenges and Horst Kessler* | Pages 2246-2260 |
| **Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals** | *Alexandros Makriyannis and Spiro Pavlopoulos* | Pages 2261-2271 |
| **Tritium NMR, Applications** | *John R Jones* | Pages 2366-2369 |
| **Xenon NMR Spectroscopy** | *Jukka Jokisaari* | Pages 2435-2446 |


## Mass Spectrometry

### *Historical Overview*


| | | |
|---|---|---|
| **Mass Spectrometry, Historical Perspective** | *Allan Maccoll†* | Pages 1241-1248 |


### *Theory*


| | | |
|---|---|---|
| **Fragmentation in Mass Spectrometry** | *Hans-Friedrich Grützmacher* | Pages 637-648 |

| | | |
|---|---|---|
| **Ion Collision Theory** | *Anil K Shukla and Jean H Futrell* | Pages 954-963 |
| **Ion Dissociation Kinetics, Mass Spectrometry** | *Bernard Leyh* | Pages 963-971 |
| **Ion Energetics in Mass Spectrometry** | *John Holmes* | Pages 971-976 |
| **Ion Structures in Mass Spectrometry** | *Peter C Burgers and Johan K Terlouw* | Pages 990-1000 |
| **Ionization Theory** | *C Lifshitz and T D Märk* | Pages 1010-1021 |
| **Metastable Ions** | *John L Holmes* | Pages 1291-1297 |
| **Proton Affinities** | *Edward P L Hunter and Sharon G Lias* | Pages 1893-1901 |
| **Statistical Theory of Mass Spectra** | *J C Lorquet* | Pages 2204-2211 |

***Methods and Instrumentation***

| | | |
|---|---|---|
| **Atmospheric Pressure Ionization in Mass Spectrometry** | *W. M. A. Niessen* | Pages 18-24 |
| **Chemical Ionization in Mass Spectrometry** | *Alex G Harrison* | Pages 207-215 |
| **Chemical Structure Information from Mass Spectrometry** | *Kurt Varmuza* | Pages 232-243 |
| **Chromatography-MS, Methods** | *W W A Niessen* | Pages 293-300 |
| **Fast Atom Bombardment Ionization in Mass Spectrometry** | *Magda Claeys and Jan Claereboudt* | Pages 505-512 |
| **Field Ionization Kinetics in Mass Spectrometry** | *Nico M M Nibbering* | Pages 539-548 |
| **Glow Discharge Mass Spectrometry, Methods** | *Annemie Bogaerts* | Pages 669-676 |

| | | |
|---|---|---|
| **Inductively Coupled Plasma Mass Spectrometry, Methods** | *Diane Beauchemin* | Pages 875-880 |
| **Ion Molecule Reactions in Mass Spectrometry** | *Diethard K Böhme* | Pages 984-990 |
| **Ion Trap Mass Spectrometers** | *Raymond E March* | Pages 1000-1009 |
| **Laser Spectroscopy Theory** | *Luc Van Vaeck and Freddy Adams* | Pages 1141-1152 |
| **MS-MS and MS$^n$** | *W. M. A. Niessen* | Pages 1404-1410 |
| **Multiphoton Excitation in Mass Spectrometry** | *Ulrich Boesl* | Pages 1411-1424 |
| **Negative Ion Mass Spectrometry, Methods** | *Suresh Dua and John H Bowie* | Pages 1461-1469 |
| **Neutralization–Reionization in Mass Spectrometry** | *Chrys Wesdemiotis* | Pages 1469-1479 |
| **Photoelectron–Photoion Coincidence Methods in Mass Spectrometry (PEPICO)** | *Tomas Baer* | Pages 1831-1839 |
| **Photoionization and Photodissociation Methods in Mass Spectrometry** | *John C Traeger* | Pages 1840-1847 |
| **Plasma Desorption Ionization in Mass Spectrometry** | *Ronald D Macfarlane* | Pages 1848-1857 |
| **Pyrolysis Mass Spectrometry, Methods** | *Jacek P Dworzanski and Henk L C Meuzelaar* | Pages 1906-1919 |
| **Quadrupoles, Use of in Mass Spectrometry** | *P H Dawson and D J Douglas* | Pages 1921-1930 |
| **Sector Mass Spectrometers** | *R Bateman* | Pages 2085-2092 |
| **Spectroscopy of Ions** | *John P Maier* | Pages 2182-2189 |
| **Surface Induced Dissociation in Mass Spectrometry** | *S A Miller and S L Bernasek* | Pages 2279-2294 |

| | | |
|---|---|---|
| **Thermospray Ionization in Mass Spectrometry** | *W M A Niessen* | Pages 2353-2360 |
| **Time of Flight Mass Spectrometers** | *K G Standing and W Ens* | Pages 2360-2365 |

*Applications*

| | | |
|---|---|---|
| **Biochemical Applications of Mass Spectrometry** | *Victor E Vandell and Patrick A Limbach* | Pages 84-87 |
| **Cluster Ions Measured Using Mass Spectrometry** | *O Echt and T D Märk* | Pages 327-336 |
| **Cosmochemical Applications Using Mass Spectrometry** | *J R De Laeter* | Pages 359-367 |
| **Food Science, Applications of Mass Spectrometry** | *John P G Wilkins* | Pages 592-593 |
| **Forensic Science, Applications of Mass Spectrometry** | *Rodger L Foltz, Dennis J Crouch and David M Andrenyak* | Pages 615-621 |
| **Hyphenated Techniques, Applications of in Mass Spectrometry** | *W M A Niessen* | Pages 843-849 |
| **Inorganic Chemistry, Applications of Mass Spectrometry** | *Lev N Sidorov* | Pages 915-923 |
| **Ion Imaging Using Mass Spectrometry** | *Albert J R Heck* | Pages 976-983 |
| **Isotope Ratio Studies Using Mass Spectrometry** | *Michael E Wieser and Willi A Brand* | Pages 1072-1086 |
| **Isotopic Labelling in Mass Spectrometry** | *Thomas Hellman Morton* | Pages 1086-1096 |
| **Medical Applications of Mass Spectrometry** | *Orval A Mamer* | Pages 1262-1271 |
| **Nucleic Acids and Nucleotides Studied Using Mass Spectrometry** | *Tracey A Simmons, Kari B Green-Church and Patrick A Limbach* | Pages 1681-1688 |

| | | |
|---|---|---|
| **Organometallics Studied Using Mass Spectrometry** | *Dmitri V Zagorevskii* | Pages 1726-1733 |
| **Peptides and Proteins Studied Using Mass Spectrometry** | *Michael A Baldwin* | Pages 1753-1763 |
| **SIFT Applications in Mass Spectrometry** | *David Smith and Patrik panl* | Pages 2092-2105 |
| **Stereochemistry Studied Using Mass Spectrometry** | *Asher Mandelbaum* | Pages 2211-2223 |


## Spatially Resolved Spectroscopic Analysis

### *Theory*

| | | |
|---|---|---|
| **Neutron Diffraction, Theory** | *Alex C Hannon* | Pages 1493-1503 |
| **PET, Theory** | *T J Spinks* | Pages 1782-1791 |
| **Scanning Probe Microscopy, Theory** | *A J Fisher* | Pages 2060-2066 |
| **Surface Plasmon Resonance, Theory** | *Zdzislaw Salamon and Gordon Tollin* | Pages 2311-2319 |


### *Methods and Instrumentation*

| | | |
|---|---|---|
| **Neutron Diffraction, Instrumentation** | *A C Hannon* | Pages 1479-1492 |
| **PET, Methods and Instrumentation** | *T J Spinks* | Pages 1771-1782 |
| **Scanning Probe Microscopes** | *J G Kushmerick and P S Weiss* | Pages 2043-2051 |
| **SPECT, Methods and Instrumentation** | *John C Lindon* | Pages 2159-2161 |
| **Structure Refinement (Solid State Diffraction)** | *Dieter Schwarzenbach and Howard D Flack* | Pages 2271-2278 |
| **Surface Plasmon Resonance, Instrumentation** | *R P H Kooyman* | Pages 2302-2310 |

### *Applications*


| | | |
|---|---|---|
| **Fibres and Films Studied Using X-Ray Diffraction** | *Watson Fuller and Arumugam Mahendrasingam* | Pages 529-539 |
| **Inelastic Neutron Scattering, Applications** | *Stewart F Parker* | Pages 894-905 |
| **Inorganic Compounds and Minerals Studied Using X-ray Diffraction** | *Gilberto Artioli* | Pages 924-933 |
| **Materials Science Applications of X-ray Diffraction** | *Åke Kvick* | Pages 1248-1257 |
| **Mössbauer Spectroscopy, Applications** | *Guennadi N Belozerski* | Pages 1324-1334 |
| **Scanning Probe Microscopy, Applications** | *C J Roberts, M C Davies, S J B Tendler and P M Williams* | Pages 2051-2059 |
| **Surface Plasmon Resonance, Applications** | *Zdzislaw Salamon and Gordon Tollin* | Pages 2294-2302 |


## Vibrational, Rotational and Raman Spectroscopies

### *Historical Overview*


| | | |
|---|---|---|
| **Vibrational, Rotational and Raman Spectroscopy, Historical Perspective** | *A S Gilbert* | Pages 2422-2432 |


### *Theory*


| | | |
|---|---|---|
| **IR Spectroscopy, Theory** | *Derek Steele* | Pages 1066-1071 |
| **Nonlinear Raman Spectroscopy, Theory** | *J Santos Gómez* | Pages 1631-1642 |
| **Photoacoustic Spectroscopy, Theory** | *András Miklós, Stefan Schäfer and Peter Hess* | Pages 1815-1822 |

| | | |
|---|---|---|
| **Raman Optical Activity, Theory** | *Laurence A Nafie* | Pages 1976-1985 |
| **Rayleigh Scattering and Raman Spectroscopy, Theory** | *David L Andrews* | Pages 1993-2000 |
| **Rotational Spectroscopy, Theory** | *Iain R McNab* | Pages 2017-2028 |
| **Vibrational CD, Theory** | *Philip J Stephens* | Pages 2415-2421 |


***Methods and Instrumentation***


| | | |
|---|---|---|
| **Chromatography-IR, Methods and Instrumentation** | *Robert L White* | Pages 288-293 |
| **Computational Methods and Chemometrics in Near-IR Spectroscopy** | *Paul Geladi and Eigil Dåbakk* | Pages 343-349 |
| **High Resolution IR Spectroscopy (Gas Phase) Instrumentation** | *Jyrki K Kauppinen and Jari O Partanen* | Pages 784-794 |
| **IR Spectrometers** | *R A Spragg* | Pages 1048-1057 |
| **IR Spectroscopy Sample Preparation Methods** | *R A Spragg* | Pages 1058-1066 |
| **Microwave Spectrometers** | *Marlin D Harmony* | Pages 1308-1314 |
| **Near-IR Spectrometers** | *R Anthony Shaw and Henry H Mantsch* | Pages 1451-1461 |
| **Nonlinear Raman Spectroscopy, Instruments** | *Peter C Chen* | Pages 1624-1631 |
| **Raman and Infrared Microspectroscopy** | *Pina Colarusso, Linda H Kidder, Ira W Levin and E Neil Lewis* | Pages 1945-1954 |
| **Raman Optical Activity, Spectrometers** | *Werner Hug* | Pages 1966-1976 |
| **Raman Spectrometers** | *Bernhard Schrader* | Pages 1986-1992 |

| | | |
|---|---|---|
| **Vibrational CD Spectrometers** | *Laurence A Nafie* | Pages 2391-2402 |

### *Applications*

| | | |
|---|---|---|
| **Art Works Studied Using IR and Raman Spectroscopy** | *Howell G M Edwards* | Pages 2-17 |
| **ATR and Reflectance IR Spectroscopy, Applications** | *U P Fringeli* | Pages 58-75 |
| **Biochemical Applications of Raman Spectroscopy** | *Peter Hildebrandt and Sophie Lecomte* | Pages 88-97 |
| **Chromatography-IR, Applications** | *George Jalsovszky* | Pages 282-287 |
| **Far-IR Spectroscopy, Applications** | *James R Durig* | Pages 498-504 |
| **Flame and Temperature Measurement Using Vibrational Spectroscopy** | *Kevin L McNesby* | Pages 548-559 |
| **Forensic Science, Applications of IR Spectroscopy** | *Núria Ferrer* | Pages 603-615 |
| **FT-Raman Spectroscopy, Applications** | *R H Brody, E A Carter, H. G. M. Edwards and A M Pollard* | Pages 649-657 |
| **High Resolution Electron Energy Loss Spectroscopy, Applications** | *Horst Conrad and Martin E Kordesch* | Pages 772-783 |
| **High Resolution IR Spectroscopy (Gas Phase) Instrumentation** | *Jyrki K Kauppinen and Jari O Partanen* | Pages 784-794 |
| **Hydrogen Bonding and Other Physicochemical Interactions Studied By IR and Raman Spectroscopy** | *A S Gilbert* | Pages 837-843 |
| **Industrial Applications of IR and Raman Spectroscopy** | *A S Gilbert and R W Lancaster* | Pages 881-893 |

| | | |
|---|---|---|
| **IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds** | *Claudio Pettinari and Carlo Santini* | Pages 1021-1034 |
| **IR Spectral Group Frequencies of Organic Compounds** | *A S Gilbert* | Pages 1035-1048 |
| **Matrix Isolation Studies By IR and Raman Spectroscopies** | *Lester Andrews* | Pages 1257-1261 |
| **Medical Science Applications of IR** | *Michael Jackson and Henry H Mantsch* | Pages 1271-1281 |
| **Microwave and Radiowave Spectroscopy, Applications** | *G Wlodarczak* | Pages 1297-1307 |
| **Nonlinear Raman Spectroscopy, Applications** | *W Kiefer* | Pages 1609-1623 |
| **Photoacoustic Spectroscopy, Applications** | *Markus W Sigrist* | Pages 1800-1809 |
| **Polymer Applications of IR and Raman Spectroscopy** | *C M Snively and J L Koenig* | Pages 1858-1864 |
| **Raman Optical Activity, Applications** | *Günter Georg Hoffmann* | Pages 1955-1965 |
| **Surface Studies By IR Spectroscopy** | *Norman Sheppard* | Pages 2320-2328 |
| **Surface-Enhanced Raman Scattering (SERS), Applications** | *W E Smith and C Rodger* | Pages 2329-2334 |
| **Vibrational CD, Applications** | *Günter Georg Hoffmann* | Pages 2403-2414 |

# Plate 1

**Plate 1** Roman die, ca. AD 300, from archaeological excavations at Frocester Villa, Gloucester, UK. Raman spectroscopy has suggested the origin of the die as sperm whale ivory. *See Art Works Studied using IR and Raman Spectroscopy*.

# Plate 2

**Plate 2** Selection of ornamental jewellery consisting of three bangles assumed to be ivory but which were shown spectroscopically to be composed of modern resins, and a genuine ivory necklace.*See Art Works Studied using IR and Raman Spectroscopy*.

# Plate 3

**Plate 3** 'Ivory' cat, which was identified spectroscopically as a modern limitation composed of poly(methyl methacrylate) and polystyrene resins with added calcite to give the texture and density of ivory. Reproduced with permission from Edwards HGM and Farwell DW, Ivory and simulated ivory artefacts: Fourier-transform Raman diagnostic study, *Spectrochimica Acta, Part A,* **51**: 2073–2081 © 1995, Elsevier Science B. V. *See Art Works Studied using IR and Raman Spectroscopy*.

# Plate 4

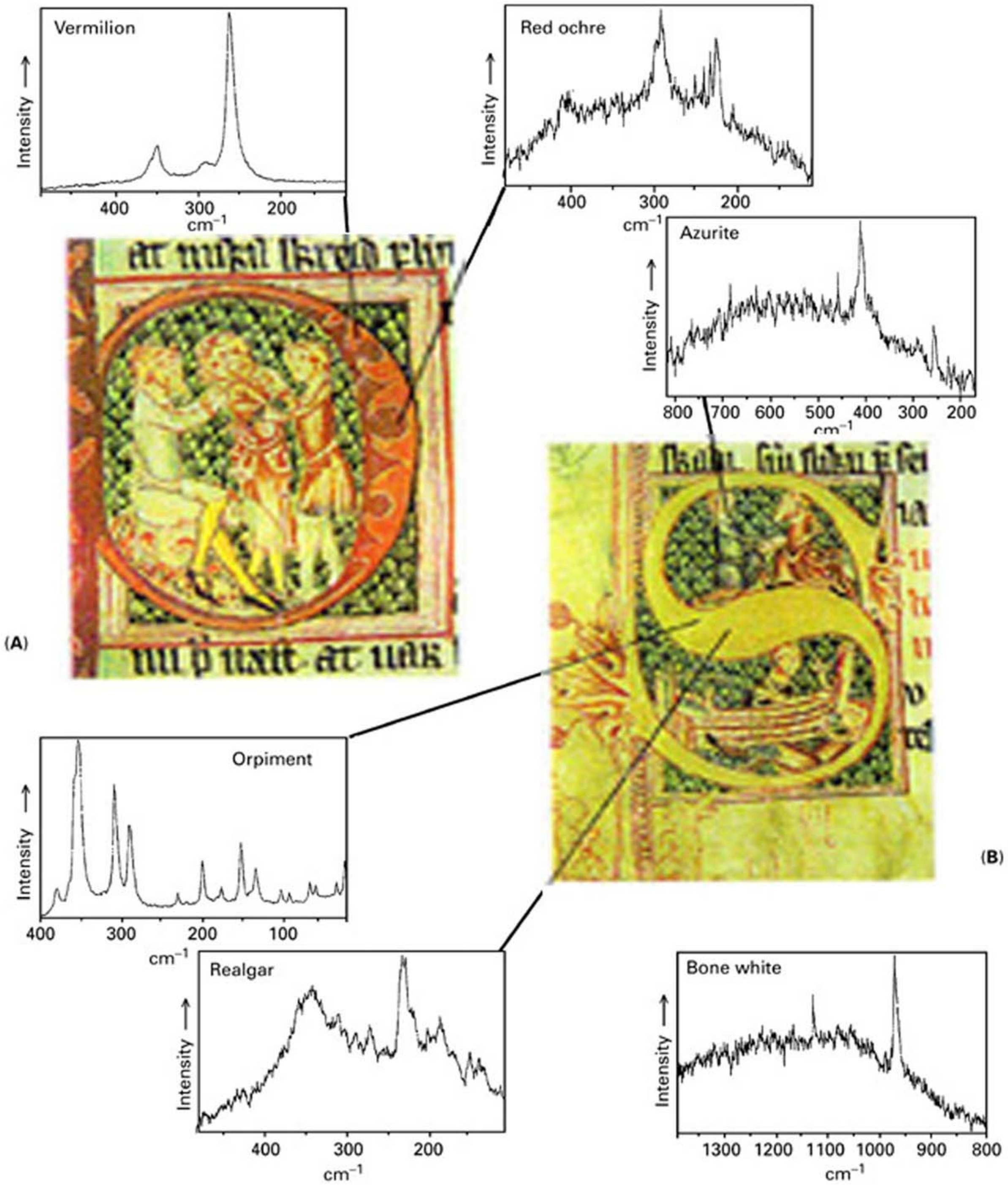



**Plate 4** The historiated initials (a) 'P' and (b) 'S' from the Icelandic Jók with the Raman spectra of pigments contained therein. The combination of vermilion and red ochre in the 'P' should be noted. The spectrum of bone white has been obtained from the letter 'H'(not shown). Reproduced with permission from Bent SP, Clark RH, Daniels MAM, Proter CA and Withnay R. Identification by Raman microscopy and visible reflectance spectroscopy of pigments on an Icelandic manuscript, (1995) *Studies in Conservation* **40**: 31–40. *See Art Works Studied using IR and Raman Spectroscopy.*

# Plate 5a

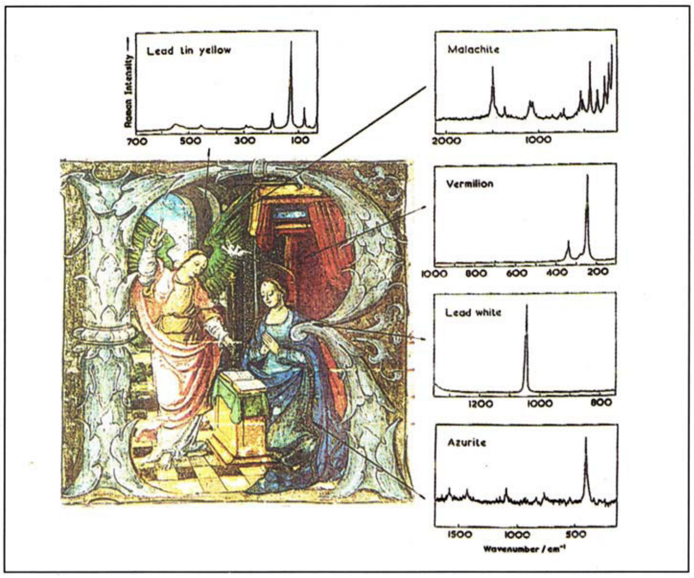



**Plate 5a** Elaborately historiated initial 'R' in sixteenth-century German choir book, with Raman microscopy spectra of selected pigmented regions. *See Art Works Studied using IR and Raman Spectroscopy*.

# Plate 5b

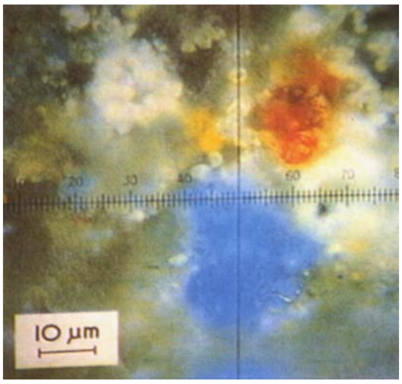



**Plate 5b** Portion of top of column of historiated initial ‘R’ shown in 5a the individual pigment-grains can be clearly seen under the x100 magnification and can be separately identified using Raman microscopy. 5a and b, reproduced with permission from Clark RJH (1995) Raman microscopy: application to the identification of pigments on medieval manuscripts. Chemical Society Reviews 187-196. © The Royal Society of Chemistry. *See Art Works Studied using IR and Raman Spectroscopy*.

# Plate 6

**Plate 6** Holy Sepulchre Chapel, Winchester Cathedral. Wall painting of ca. 1175–85 on the east well depicting the Deposition, Entombment, Maries at the Sepulchre and the Harrowing of Hell. Reproduced with permission from Edwards HGM, Brooke C and Tait JKF, An FT-Raman spectroscopic study of pigments of medieval English wall paintings, *Journal of Raman Spectroscopy* **28**: 95–98, © 1997 John Wiley and Sons Ltd. *See Art Works Studied using IR and Raman Spectroscopy*.

# Plate 7

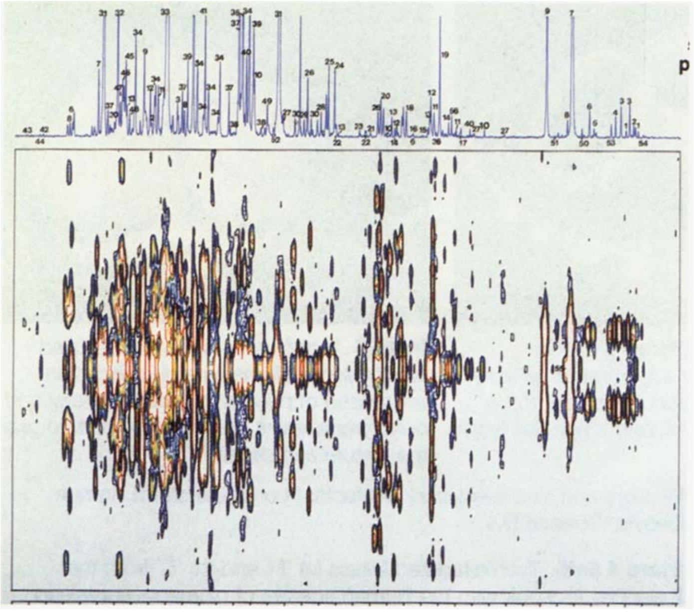
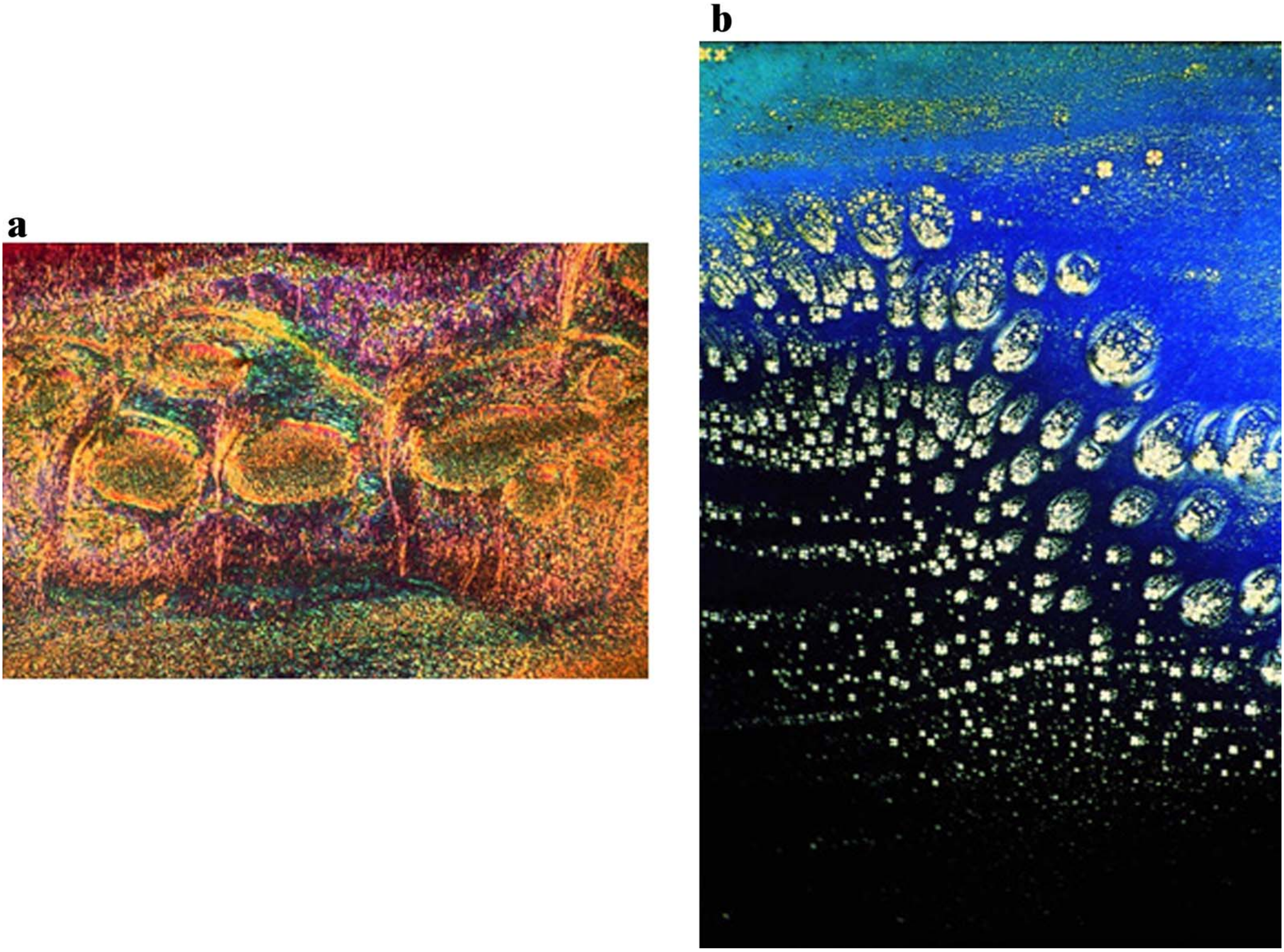


**Plate 7** A two-dimensional $^1$H-NMR spectrum of human urine, demonstrating the complexity of such mixtures. This was measured using the J-resolved pulse sequence and results in contour plot of spectrum intensity as a function of two frequency axes. The horizontal axis represents the usual $^1$H NMR chemical shift range and the vertical axis cover the $^1$H$^1$H spin coupling range. Each $^1$H NMR spin-coupled multiplet is rotated such that the overlap between closely spaced signals is minimised, thus aiding interpretation. *See Biofluids Studied by NMR.* *Reproduced with permission from John Lindon.*

# Plate 8




**Plates 8 a and b** Polarised light micrograph of liquid crystals. *See Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Liquid Crystals and Liquid Crystal Solutions Studied by NMR*. Reproduced with permission from Science Photo Library.

# Plate 9

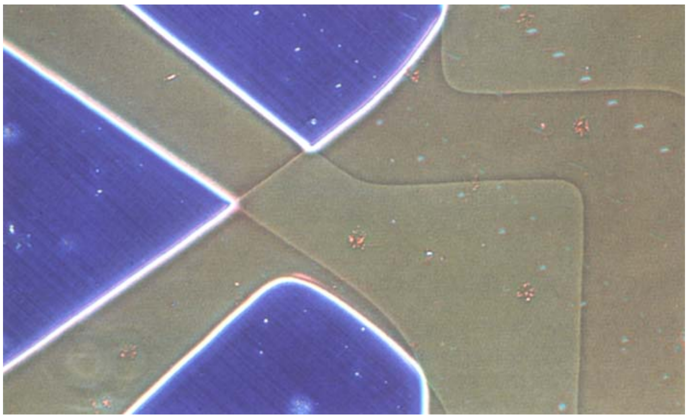


**Plate 9** Light micrograph of a liquid crystal display (LCD) of the type used to represent numerical figures. Although a liquid crystal (LC) can flow like a fluid its molecular arrangement exhibits some order. A LCD has a film of LC sandwiched between crossed polarizers & set on top of a mirror. In the 'off' state light is able to traverse both polarizers to reach the mirror because it gets rotated through 90 degrees by the LC. In the 'on' state an electric field applied across the LC alters its molecular alignment & hence its polarizing properties; light cannot traverse both polarizers to reach the mirror & the display appears black. *See Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Liquid Crystals and Liquid Crystal Solutions Studied by NMR.* Reproduced with permission from Science Photo Library.

# Plate 10

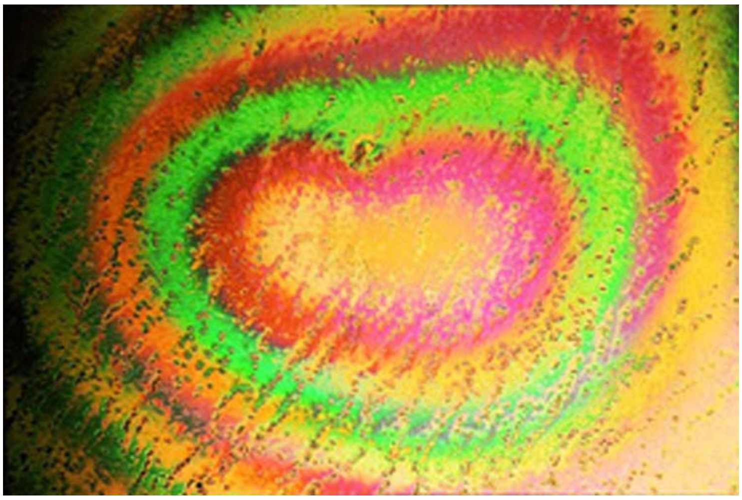


**Plate 10** Polarised light micrograph of nematic liquid crystals.*See Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Liquid Crystals and Liquid Crystal Solutions Studied by NMR*. Reproduced with permission from Science Photo Library.

# Plate 11

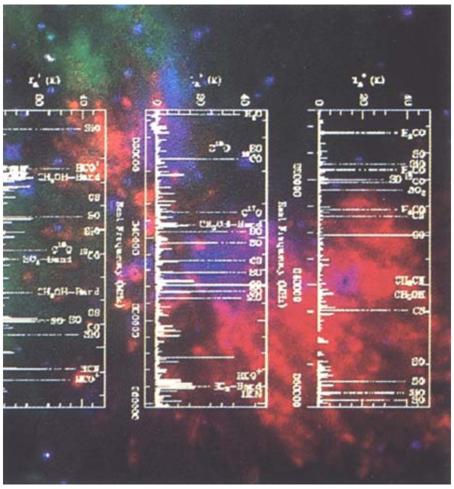


**Plate 11** The molecular composition of massive star-forming region in the galaxy. The false color background shows the heart of a massive star-forming region in the constellation of Orion obtained by the Hubble Space Telescope. The red colour is the emission of the molecular hydrogen excited by shocks driven by the powerful wind of a newly formed star in the center of the image. The bluish color is scattered light from this young star reflected in surrounding dust in its stellar nursery. The white spectrum shows the results of ground-based line survey of its region and reveals the presence of molecules such as methanol, formaldehyde, carbon monoxide, sulphur monoxide, and sulphur dioxide. *See Cosmochemical Applications Using Mass Spectrometry; Interstellar Molecules, Spectroscopy of.* Reproduced with permision from A.G.G.M. Tielens.

# Plate 12

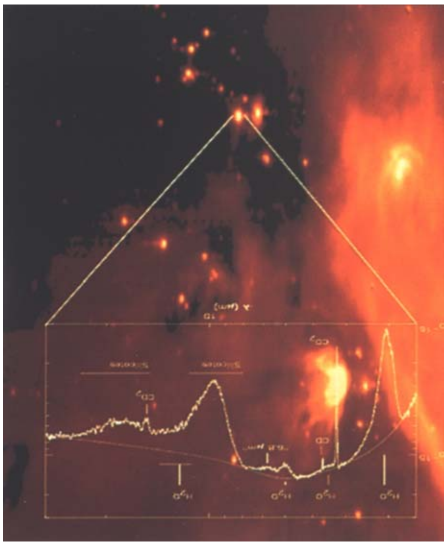


**Plate 12** The composition of ices in low-mass star forming regions. This is a false color image of an infrared image obtained by the European ISO satelite (Infrared Space Observatory). The colors indicate emission by warm dust and gas heated by nearby stars in the star-forming molecular cloud Rho Ophiucus. Two bright objects just below the center are newly formed solar-type stars. The spectrum displayed on top was taken by a spectrograph on board ISO towards the indicated source. This spectrum reveals the presence of various simple molecules such as water, carbon monoxide in ice-form in these regions. The composition of these ices is very similar to those of comets in the solar system and eventually, these ices are expected to coagulate together forming comets in the budding planetary system around this newly formed star. *See Cosmochemical Applications Using Mass Spectrometry; Interstellar Molecules, Spectroscopy of*. Reproduced with permission from A.G.G.M. Tielens.

# Plate 13

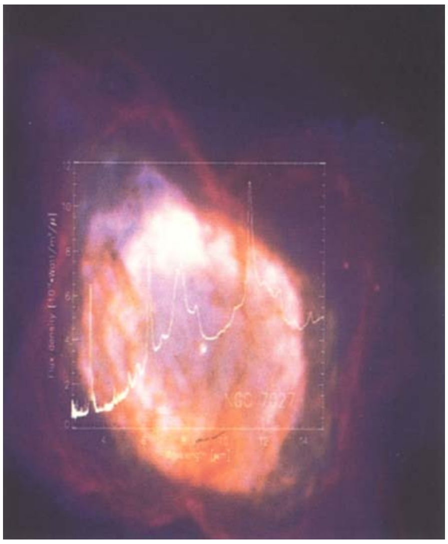


**Plate 13** Large molecules in the ejecta from a dying star. When solar-type stars grow old they shed much of their mass in the form of a wind much of that in the form of complex molecules. Eventually, the star becomes very hot and sets this ejected material aglow, forming so-called planetary nebulae. The material ejected during this stage will mix with other material present in space and eventually new stars and planetary systems like our own will form these stellar ashes. The false color image shows the emission of molecular hydrogen (red) and atomic hydrogen (white) in the planetary nebula NGC 7027 obtained by the ISO satelite (Infrared Space Observatory). This spectrum reveals the presence of large complex molecules called polycyclic aromatic hydrocarbons. These molecules are formed in a process akin to that of sooting flames in terrestrial envirnoments such as car engines. *See Cosmochemical Applications Using Mass Spectrometry; Interstellar Molecules, Spectroscopy of*. Reproduced with permission from A.G.G.M. Tielens.

# Plate 14

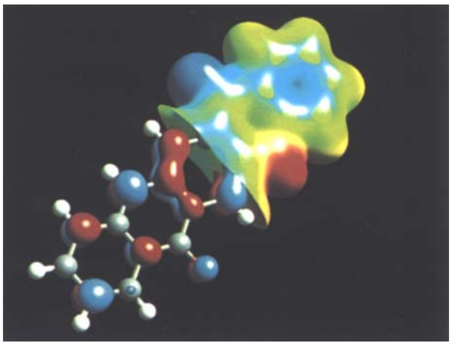


**Plate 14** A molecule of the pigment quinacridone. The coloured surfaces represent key molecular properties predicated using computational chemistry. *See Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy*. Reproduced with permission from Molecular Simulations (www.msi.com).

# Plate 15

**Plate 15** Sir Frederick William Herschel (1738–1822) German-British astronomer, and disocerer of Uranus, 1781. Reproduced with permission from Mary Evans Picture Library.

# Plate 16

**Plate 16** Sir Chandrasehara Venkata Raman. Indian (1888–1970) physcicist, Nobel Laureate in physics 1930 and discoverer of the Raman effect, a scattering of light by a sample which results in a small fraction of the light being of a different frequency. This phenomenon has applications in analytical chemistry and molecular structure. *See Electromagnetic Radiation; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective*. Reproduced with permission from Mary Evans Picture Library.

# Plate 17

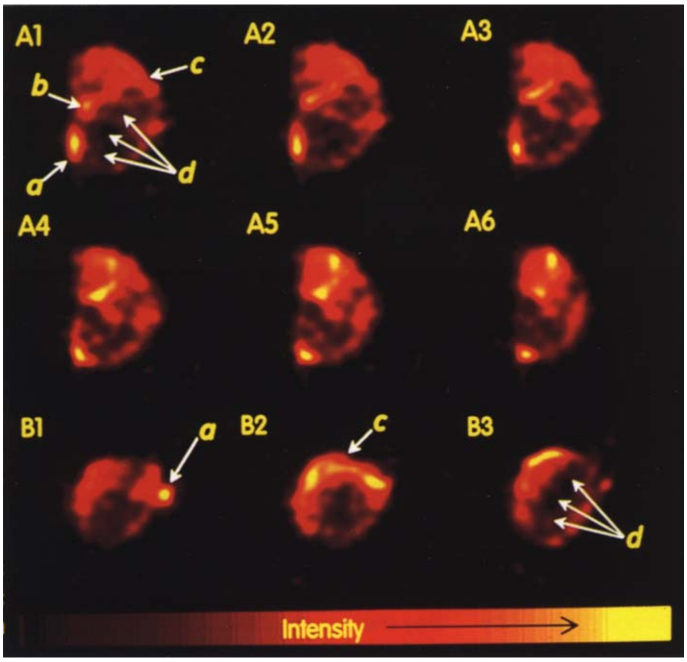



**Plate 17** Imaging of a rat kidney perfused with 0.5mM triayl-methyl (TAM) radical. A few representative slices (2424 $mm^2$), obtained from the 3D spatial image are shown. A1-A6 are vertucak slices (0.75 mm apart), B1-B3 are traverse slices (0.75 mm apart). The images show the structure of cannula (a), renal artery (b), cortex (c), and calysis (d). *See EPR Imaging*. Reproduced with permission from Peraiannan Kuppusamy/EPR Laboratories Johns Hopkins University School of Medicine.

Plate 18

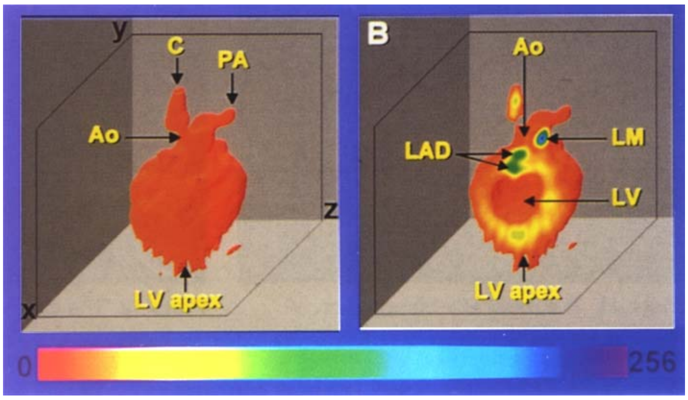



**Plate 59** Three-dimensional images (25  25  25 $mm^3$) of an ischaemic rat heart infused with a glucose char suspension. (A) Full view of the heart. (B) A longitudinal cut out showing the internal structure of the heart. C is the cannula; Ao the aorta; PA the pulmonary artery; LM the left main coronary artery; LAD the left anterior descending artery; and LV the left ventricular cavity. *See EPR Imaging*. Reproduced with permission from Periannan Kuppusamy/ERP Laboratories Johns Hopkins University School of Medicine.

# Plate 19

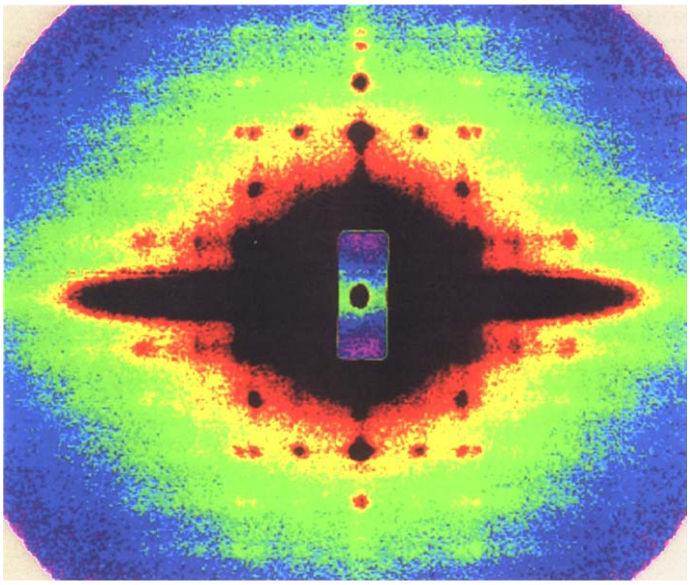


**Plate 19** Small angle X-ray fibre diffraction pattern recorded from Plaice fish muscle. *See Fibres and Films Studied Using X-Ray Diffraction*. Reproduced with permission from W. Fuller and A. Mahendrasingam.

# Plate 20

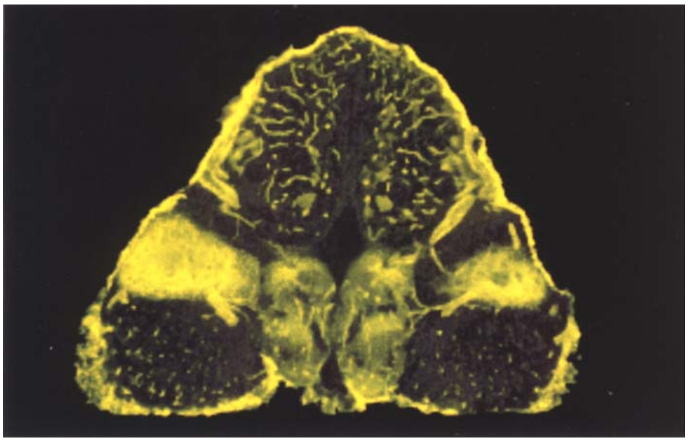


**Plate 20** An example of a fluorescent molecular probe, Endogenous alkaline phosphatase activity in the zebrafish brain was localized with an endogenous phosphatase detection kit. Enzymatic cleavage of the phosphatase substrate yields a bright yellow-green fluorescent precipitate at the site of enzyme activity. *See Fluorescent Molecular Probes*. Reproduced with permission from Greg Cox, Molecular Probes, Inc.

# Plate 21

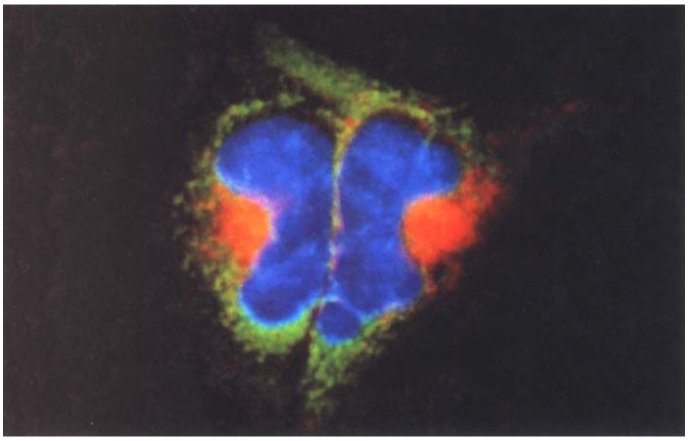


**Plate 21** A fluorescent molecular probe. Fixed and permeabilized osteosarcoma cells were simultaneously stained with the fluorescent lectins, Alexa 488 concanavalin A (Con A) and Alexa 594 wheat germ agglutinin (WGA). Con A selectively binds a-mannopyranosyl and a-glucopyranosyl residues, whereas WGA selectively binds sialic acid counterstained with blue-fluorescent Hoechst 33342 nucleic acid stain. *See Fluorescent Molecular Probes, Inc.*

# Plate 22

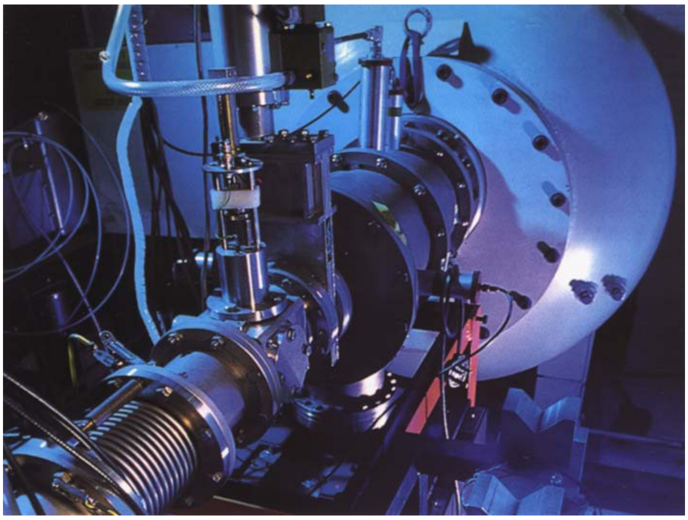


**Plate 22** Radiocarbon dating. View of a linear accelerator used as part of an accelerator mass spectrometer (AMS). This device is capable of counting the relatively few carbon-14 atoms in a radioactive sample. The proportion of carbon-14 to carbon-12 atoms in the sample may be used to determine the radiocarbon age of an organic object. This is then adjusted by various corrections to give the true age. *See Isotope Ratio Studies Using Mass Spectrometry*. Reproduced with permission from Science Photo Library.

# Plate 23

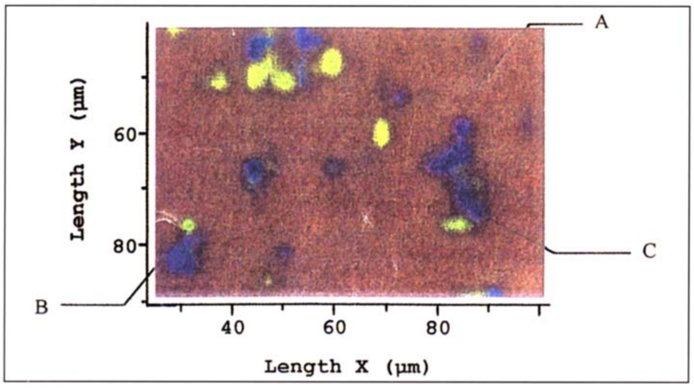



**Plate 23** Raman microscope image of a pharmaceutical powder showing the relative importance of three compound, A, B and C. *See Industrial Application of IR and Raman Spectroscopy*. Courtesy of A.S. Gilbert and R.W. Lancaster.

# Plate 24

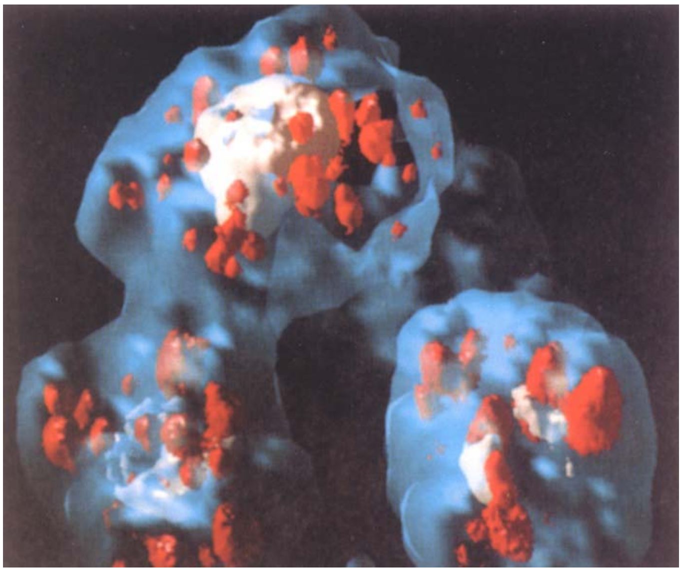


**Plate 24** 3-D reconstruction of a plant cell undergoing meiotic cell division, created from two-photon cross-sectioned images. *See Laser Spectroscopy Theory; Light sources and Optics*. Reproduced with permission from Spectra-Physics © W. Zipfel and C. Conley/Cornell University.

# Plate 25

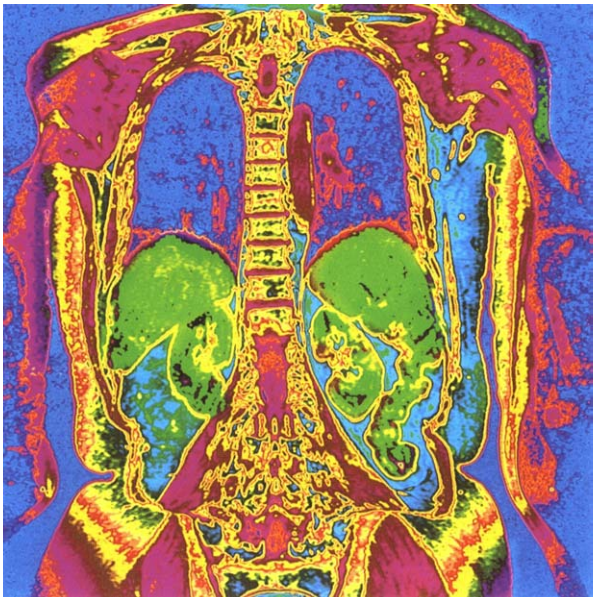


**Plate 25** Thorax and abdomen. Coloured Magnetic Resonance Imaging (MRI) scan of the thorax and abdomen of a woman aged 58 years, seen in posterior view. The arms are seen on either side, with the waist and hips at lower frame. The thorax contains the blue lung fields (upper frame) with bones of the thoracic spines at upper centre. In the abdomen are lobes of the liver (green, below the lungs) and a pair of kidneys (green, below the liver lobes). Bands of muscle can be seen around the lumbar spine (lower centre) and the shoulder (at top, magenta). MRI uses probes of radiowave energy in the presence of a magnetic field to create slices. *See MRI Applications, Clinical; MRI Instrumentation; MRI Theory; Contrast mechanisms in MRI*. Reproduced with permission from Science Photo Library.

# Plate 26

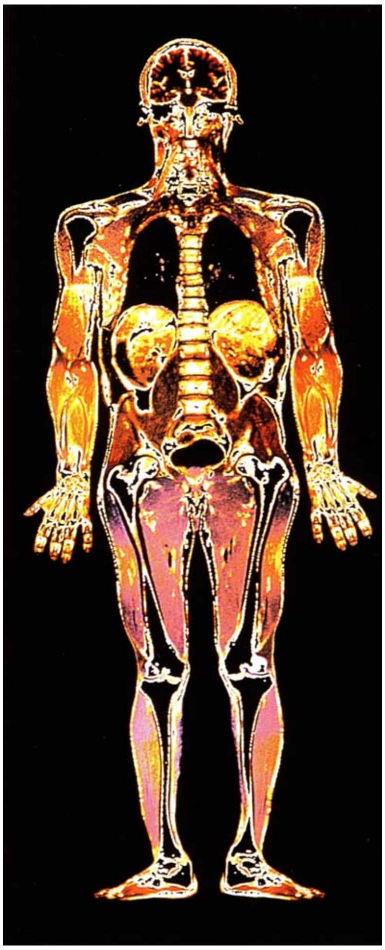


**Plate 26** False-colour magnetic resonance image (MRI) of a whole human body, a woman, taken in coronal (frontal) section. Various parts of the body are prominent: bone appears in black, wiht the hip and shoulder joints clearly defined. The lungs are also black. Vertebrae forming the spinal column are also obvious (in yellow and white). Part of the spinal cord is visible (in gold) beneath the base of the brain, the hemispherical structure of which is revealed in the section. This whole body image is the product of a number of MRI scans made along the length of the body. The complete image was rendered from data relating to the various sections stored on the scanner's computer. *See MRI Applications, Clinical; MRI Instrumentation; MRI Theory; Contrast mechanisms in MRI.* Reproduced with permission from Science Photo Library.

# Plate 27

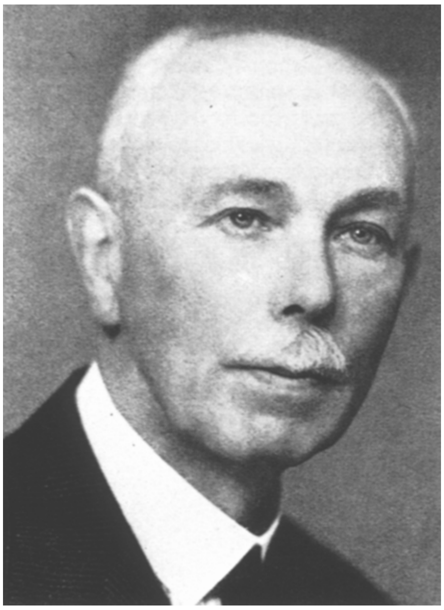


**Plate 27** Portrait of Francis Aston (1877–1945) British physicist and Nobel Laureate. After WW1, Aston helped Thompson in his studies of the deflection of ions in magnetic fields. He went on to improve Thompson's apparatus, designing it so as to make all atoms of a given mass fall on the same part of a photographic plate. Working with neon, he found that two lines were isotopes. He repeated this with chlorine with similar results. The device, called the mass spectrometer, showed that most stable elements had isotopes. His work earned him the 1922 Nobel Prize for Chemistry, and introduced a powerful new analytical tool to science. *See Mass Spectrometry, Historical Perspective*. Reproduced with permision from Science Photo Library.

# Plate 28

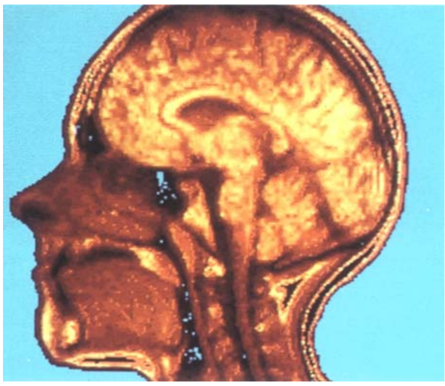


**Plate 28** Magnetic Resonance Image of Human female brain. *See MRI Applications, Clinical.* Reproduced with permission from © Medipics/Dan McCoy/Rainbow.

# Plate 29

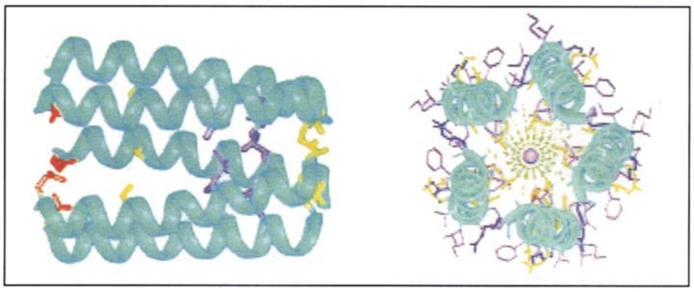


**Plate 29** Structure of membrane proteins. Membrane peptides often self assemble into controlled oligomeric forms to make molecular selective channels whose structure can be very difficult to resolve at the molecular level by most methods, although solid state NMR methods can make a contribution to their functional and structural description. Here the pentameric funnel-like bundle of the M2-peptide from the nicotinic acetly choline receptor has been resolved from $^{15}N$ NMR studies of oriented M2 peptides in lipid bilayers. The funnel has a wide mouth at the N-terminal, intracellular side of the pore. The pore lining residues has also been modelled and distances between residues in the channel estimated. The a-carbon backbone is in cyan, acidic residues in red and basic residues in blue, polar residues in yellow and lipophilic residues in purple. (Figure adapted from Opella et al., (1999) Nature St. Biology, 6: 374–379). *See Membranes Studied by NMR Spectroscopy.*

# Plate 30

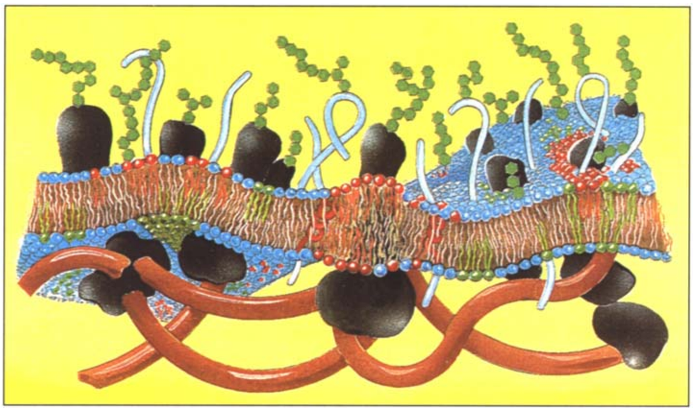


**Plate 30** Biological membranes which enclose all living cells, and are also present within the cells of higher life forms, are very complex and heterogeneous in their chemical make-up. Between 20 and 40% of components encoded by the genome in any one living cell eventually end up in the membranes of the cell. Any one typical membrane may be composed of over 100 different bilayer lipids and sterols, several hundreds of different proteins, some of which have polysaccharides attached to them, and the whole assembly may be supported by a protein scaffold underlying the membrane on the cytoplasmic side. The whole molecular complex is in dynamic equilibrium, with lipids and sterols rotating fast (about $10^9$ times per second) around their long axis, and diffusing laterally (covering about $10^8$ $cm^2$ in one second), whilst at the same time maintaining a regular and relatively ordered structure, on average. Such dynamic information has come from a range of spectroscopic methods, including NMR, in which it is usual to study just one, or a limited number of such components in a simplified model system, and try and understand them and their interactions with a limited number of other components. (Figure by Ove Broo Sorensen of the Technical University of Copenhagen, Denmark). *See Membranes Studied by NMR Spectroscopy*.

# Plate 31

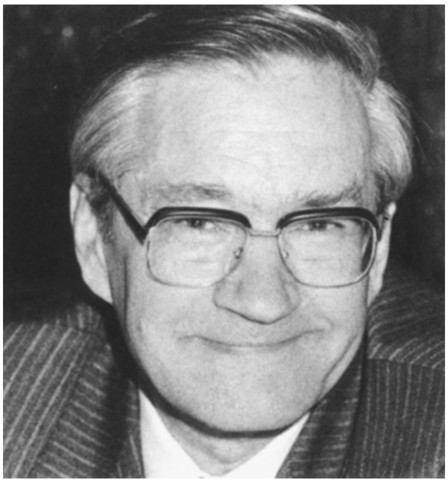


**Plate 31** Richard Ernst, Nobel Laureate in Chemistry 1991. Notes for this contribution to NMR spectroscopy and MRI. *See Magnetic Resonance, Historical Perspective*. Reproduced with permission from The Nobel Foundation.

# Plate 32

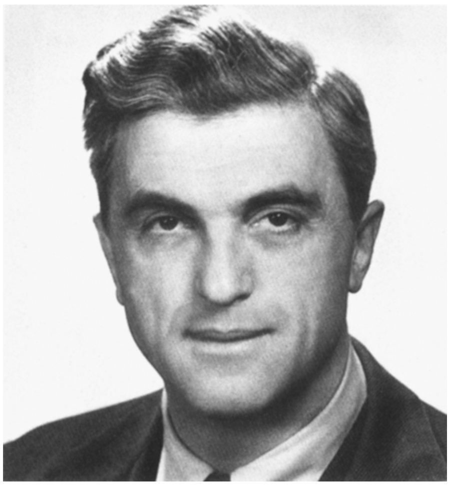


**Plate 32** Felix Bloch who was awarded the Nobel prize for physics in 1952 jointly with Edward Purcell. They led two independent research groups which, in 1945, first detected the nuclear magnetic resonance phenomenon in bulk matter. *See Magnetic Resonance, Historical Perspective*. Reproduced with permission from The Nobel Foundation.

# Plate 33

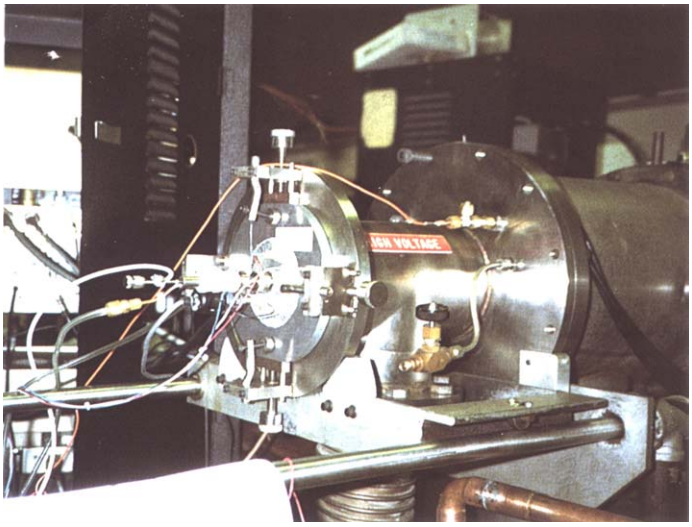


**Plate 33** The first fully functioning electrospray mass spectrometer built at Yale by Masamichi Yamashita in 1983. Reproduced with permission from John B. Fen.

# Plate 34

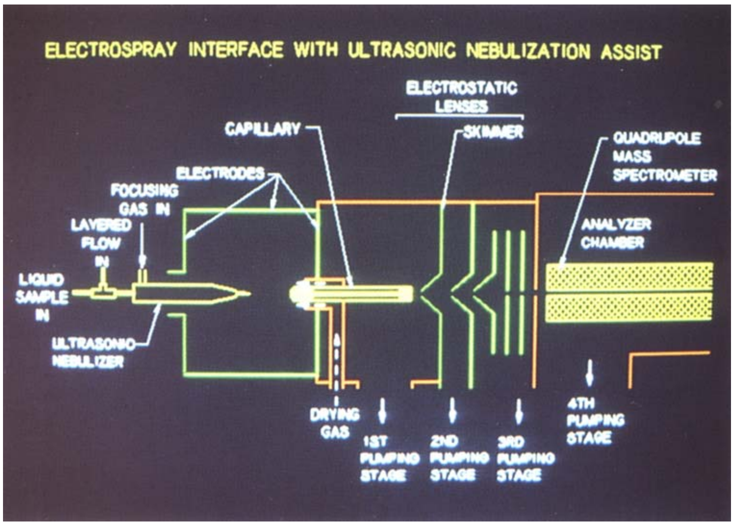



**Plate 34** A schematic representation of a typical electrospray mass spectrometer in which the mass analyser is a quadrupole mass filter. Reproduced with permission from John B. Fen.

# Plate 35

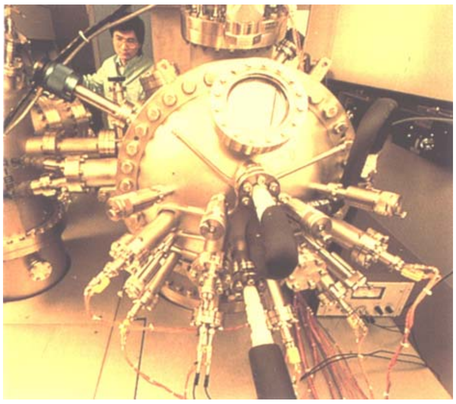


**Plate 35** A Bell Lab ZAB high field mass spectrometer. Reproduced with permission from © Medipics/Dan McCoy/Rainbow.

# Plate 36

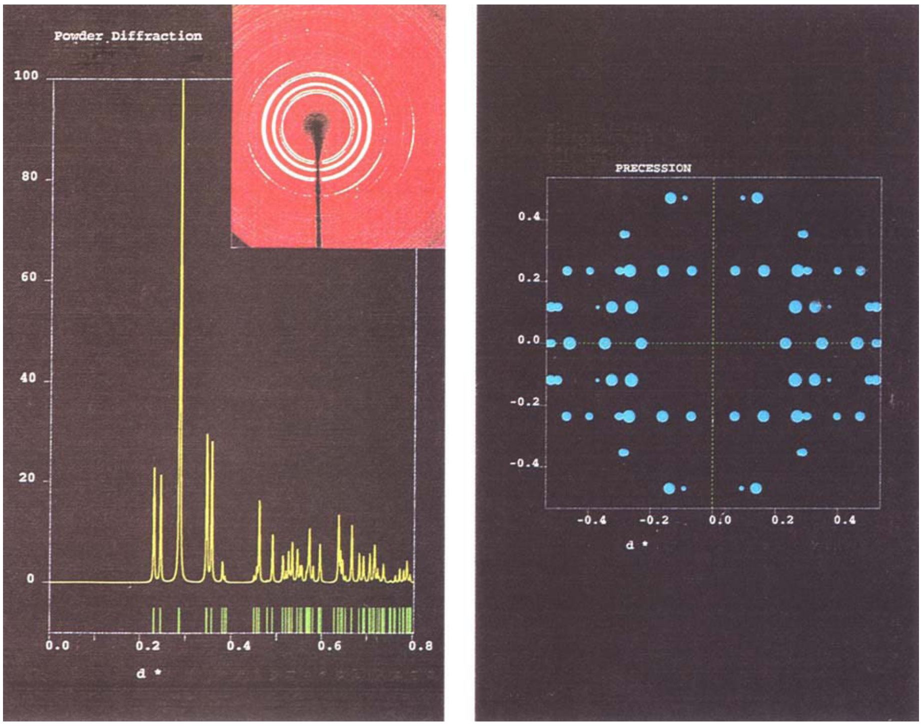



**Plate 36** The left-hand picture illustrates a powder diffraction pattern of $S_2N_2$ at a wavelength of 0.325 Å. The insert is the actualpattern which can be integrated and displayed as a function of *d*[ $d^* = 2 \sin q/l$]. The right-hand picture shows the result if one single $S_2N_2$ crystal is rotated in the X-ray beam. The size of the spots illustrates the difference in diffracted intensities from the separated Bragg reflections. *See Material Science Applications of X-ray Diffraction*. Courtesy of Svensson and Kvick, 1998.

# Plate 37

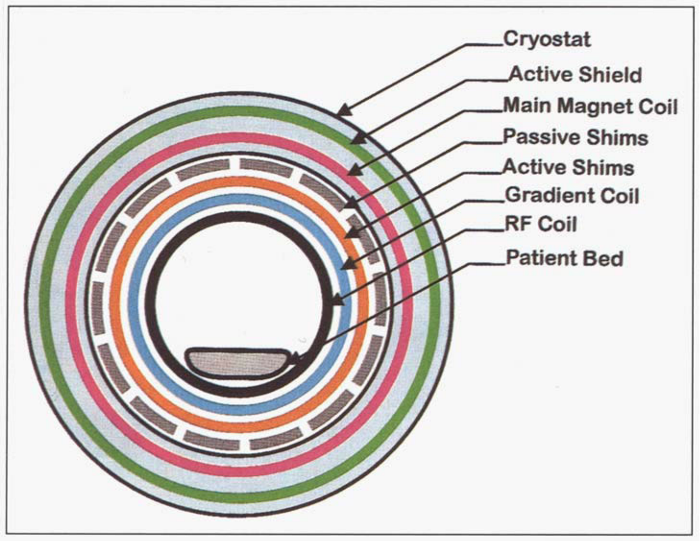



**Plate 37** Schematic cross-section through a typical superconducting clinical MR scanner. Within the cryostat (light blue) are the superconducting coils of the primary magnet (red) and active shield (green). In the bore of the magnet there are passive shim rods (grey), active shim coils (orange), gradient set (blue), whole body RF coil (black) and patient bed. The tractable diameter is generally half the magnet bore diameter. *See MRI Instrumentation*. Courtesy of P.D. Hockings, J.F. Hare and D.G. Reid.

# Plate 38

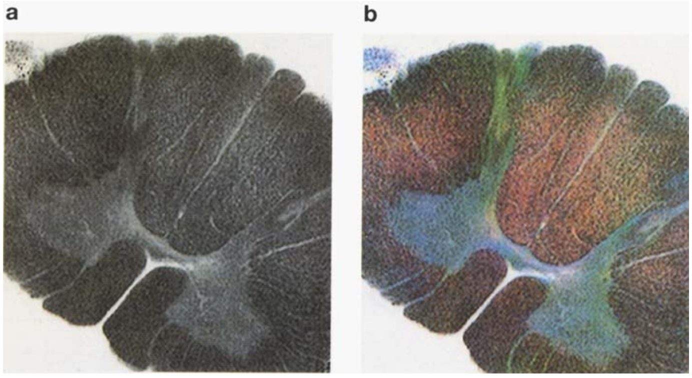



**Plate 38 a and b** Black and white & colourscale diffusion weighted spin echo magnetic resonance images of postmortem human cervical spinal cord using a micro-imaging probe at 600 MHz observation frequency. *See NMR Microscopy*. Reproduced with permission from Doty Scientific Inc.

# Plate 39

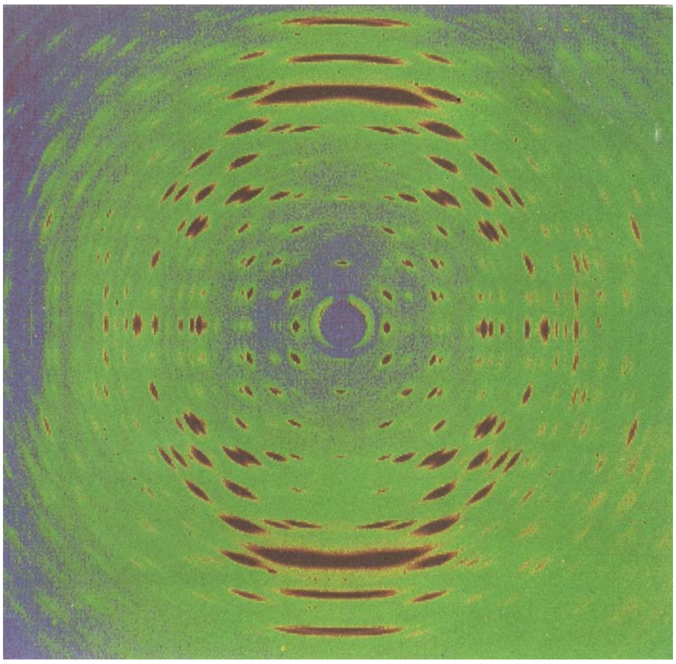


**Plate 39** Small angle X-ray fibre diffraction pattern recorded form of the DNA double-helix. *See Nucleic Acids and Nucleotides Studied Using Mass Spectroscopy; Nucleic Acids Studied by NMR*. Courtesy of A. Mahendrasigam and W. Fuller.

# Plate 40

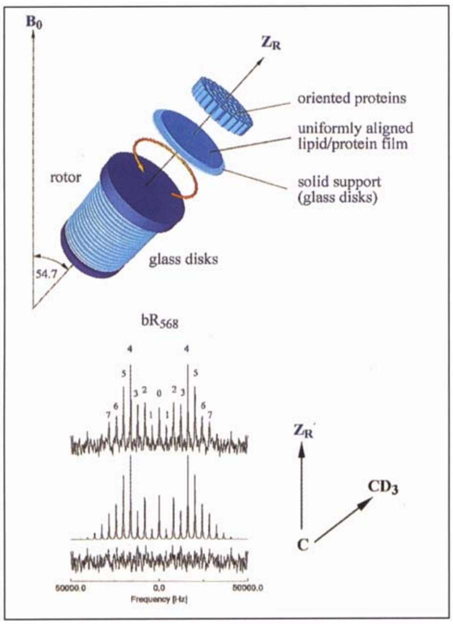



**Plate 40** The anisotrophy of chemical shift, dipolar and quadrupolar coupling usually broadens NMR spectra of very large, slowly tumbling complexes making a detailed structural analysis a difficult task. If however the sample is set spinning at an angle of 54.7° (MASS = magic angle sample spinning) with respect to the applied magnetic field, all anisotropies are scaled down and eventually collapse at sufficiently high speeds resulting in isotropic spectra. In a special version of MASS called MAOSS, oriented membranes are set spinning at moderate speeds (1–4 kHz) at the magic angle, and now the narrow spinning side bands, which arise from not-averaged anisotropic interactions, give information about the orientation of the group being observed (here a $CD^3$ group in deuterium NMR of a prosthetic group in a large protein embedded in membrane) at much higher sensitivity and resolution than could be obtained from a conventional spectrum of a statically positioned sample. Similar experimental condition for other nuclei ($^{13}C$, $^{15}N$), permit both the magnitude and direction of the dipolar couplings to be extracted, to give high resolution structural details within a large membrane complex, (courtesy of Dr C. Glaubitz; adapted from Glaubitz and Watts, (1998) J. Mag. Res. 130; 305–316). *See NMR in Anisotropic Systems, Theory; Solid State NMR, Methods.*

# Plate 41a

a

**Plate 41a** A superconducting NMR magnet operating at 18.8T for $^{1}$H NMR observation at 800 MHz demonstrating the size of these state-of-the-art magnets. Reproduced with permission from Bruker Instruments Inc., Billerica, MA, USA.

# Plate 41b

b

**Plate 41b** A modern high-resolution NMR spectrometer. A superconducting magnet is shown at the rear, in this case providing a field of 18.8T corresponding to a $^{1}H$ observation frequency of 800 MHz. Behind the operator is the single console containing the RF and other electronics and the temperature-control unit. The whole instrument is computer controlled by the workstation shown at the right. Reproduced with permission from Bruker Instruments Inc., Billerica, MA, USA. *See NMR Spectrometers*.

# Plate 42

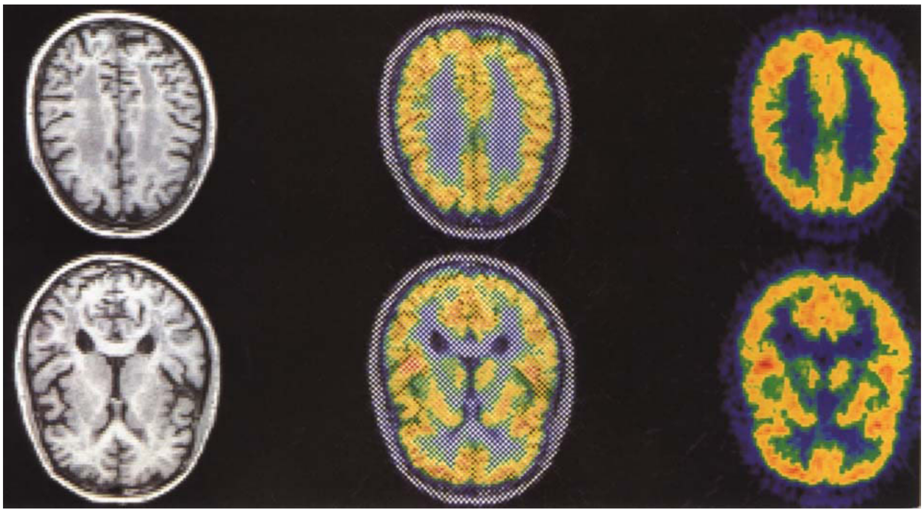


**Plate 42** PET functional images of glucose metabolic rate (MRGl) (right), MRI images (magnetic resonance images, left) showing anatomical detail, and coregistered (overlaid) PET/MRI (centre). *See PET, Methods and Instrumentation.*

# Plate 43

**Plate 43** X-ray photoelectron spectrometer. This is an analytical instrument, used mainly in metallurgy. It consists of an electron source, a sample stage and X-ray detectors. A steam of electrons is accelerated toward and focussed on the sample. When an electron strikes an atom in the target, it may cause the ejection of an electron in the atomic shell. If this is replaced from an electron in a higher-energy shell, a photon of a specific wavelength is emitted, normally in the X-ray region. This may be detected and analysed, giving an indication of the identity and quantity of given elements in the sample. *See Photoelectron Spectrometers*. Reproduced with permission from Science Photo Library.

# Plate 44

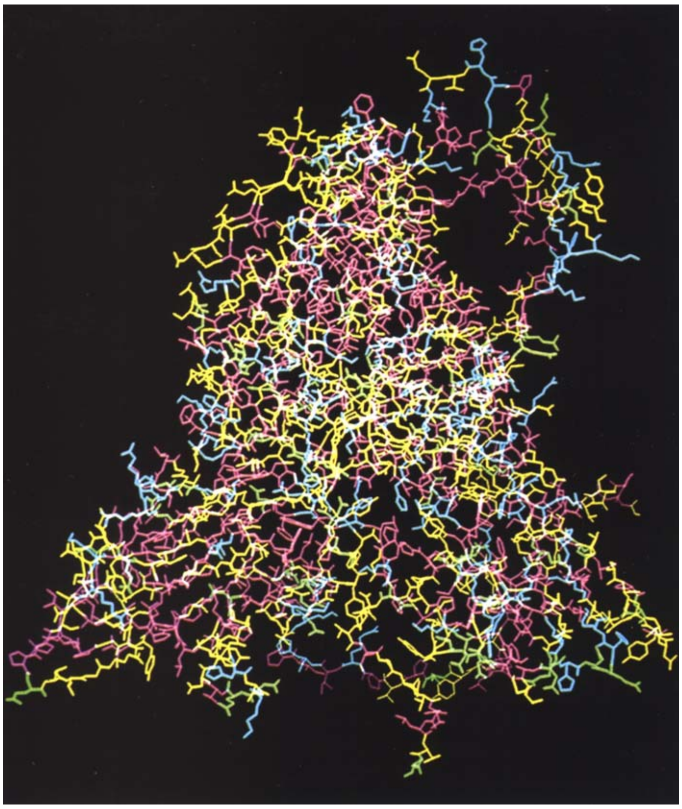


**Plate 44** Common cold virus protein: computer graphics representation of protein comprising 1 of the 60 faces of the icosahe-dral capsid (casing) of rhinovirus-14, a virus causing the common cold. This image is of the atomic backbone of the protein. Rhinovirus belongs to the picornavirus (small RNA viruses) group, which also includes the enteroviruses and the agent of foot and mouth disease in cattle. They are all relatively small (24–35 nanometres diameter), non-enveloped, with their genetic information held in the form of RNA. Rhinovirus infects the nose and throat, and is spread in droplets from talking, coughing and sneezing. *See Proteins Studied Using NMR Spectroscopy*. Reproduced with permission from Science Photo Library.

# Plate 45

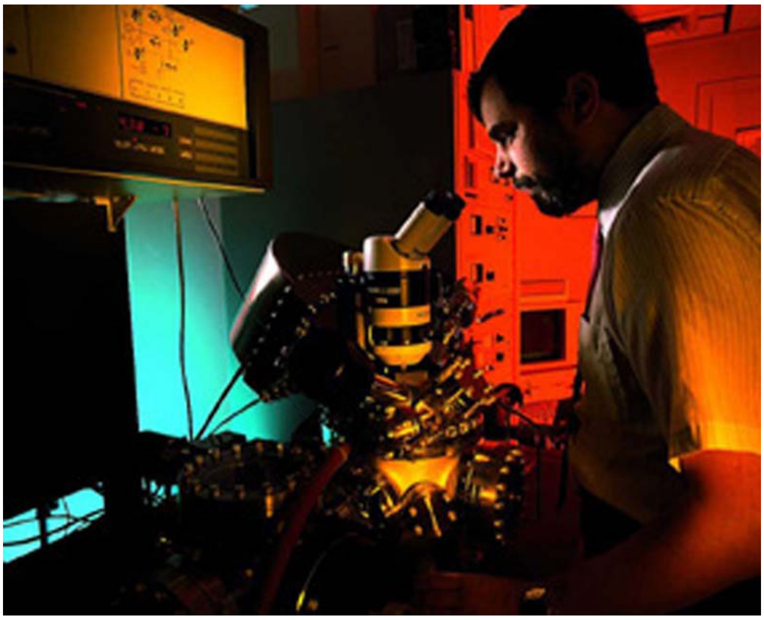


**Plate 45** A researcher at the BOC Group Technical Centre, Murray Hill, New Jersey, USA, using an advanced Electron Spectroscopy For Chemical Analysis (ESCA) unit. ESCA provides qualitative and quantitative analysis of the chemistry of elements present in the outermost layers of solid materials. The unit is used in the development of thin-film coatings, medical sensors, molecular sieves and catalysts. *See Photoelectron Spectrometers; Photoelectron Spectroscopy*. Reproduced with permission from Science Photo Library.

# Plate 46

**Plate 46** Photoacoustic Multi-gas Monitor. *See Photoacoustic Spectroscopy, Methods and Instrumentation*. Reproduced with permission from INNOVA AirTech Instruments.

# Plate 47

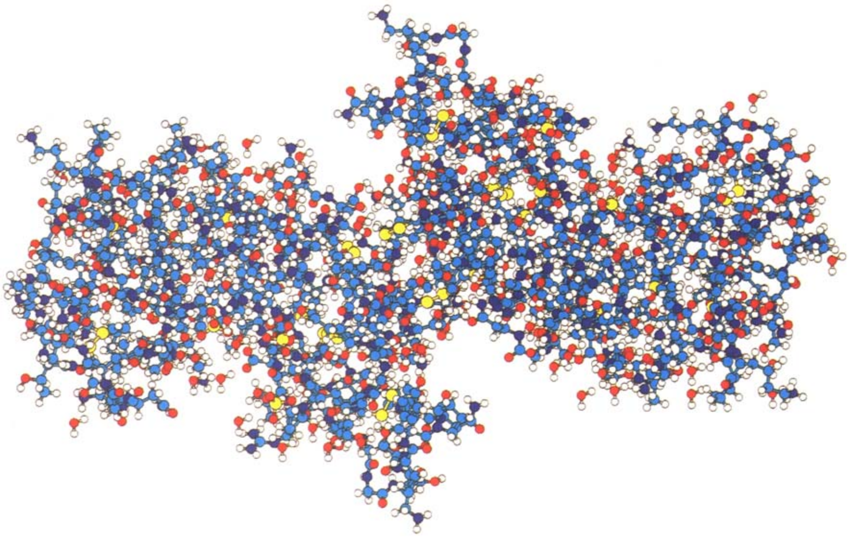


**Plate 47** The molecular structure of Bovine Seminal Ribonuclease, as determined using X-ray diffraction. The disulphide bridges are shown in yellow. The structure was drawn form the protein database entry drawn. *See Proteinss Studied Using NMR Spectroscopy*. Supplied by Dr G.G. Hoffmann.

# Plate 48

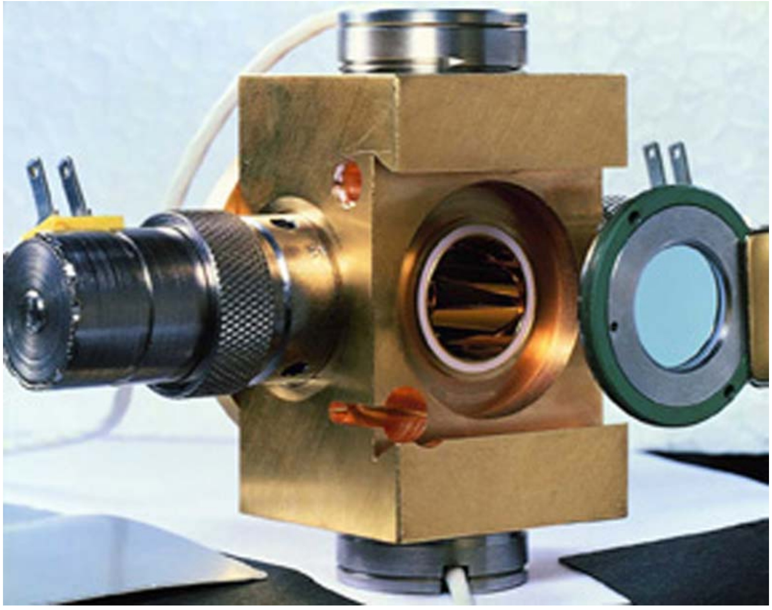


**Plate 48** A photoacoustic spectroscopy (PAS) measuring cell, shown with the germanium window half removed. The gas analysis is based on the same principle as conventional infra-red (IR) monitoring except that here the amount of IR light absorbed is measured directly by determining the amount of sound energy emitted upon the absorption. *See Photoacoustic Spectroscopy, Methods and Instrumentation*. Reproduced with permission from INNOVA AirTech Instruments.

# Plate 49

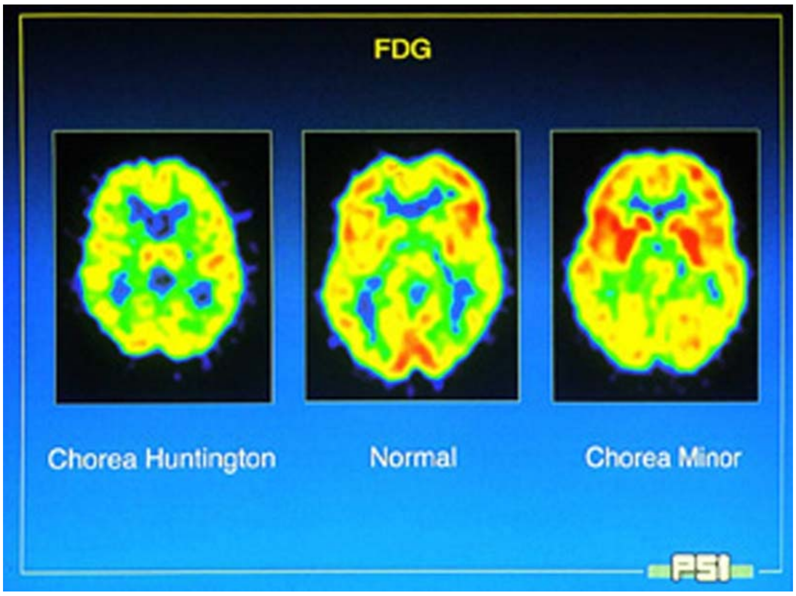



**Plate 49** Transaxial planes through the brains of three subjects representing glucose utilisation obtained by application of the radiotracer fluoro-deoxy-glucose (FDG) PET tomography. The top of the images are the front of the brain and the person is looked upon from below. The values of the images are colour coded according to a linear scale red being the highest value. The middle panel shows the distribution in a healthy person. The left panel is from a person with Huntington's disease illustrating a loss of tracer uptake at the centre as a consequence of local nervous tissue loss. The right panel is from a person who had similar features of disordered movement (cholera), but based on different brain pathology. The image in this person shows however a clear increase in tracer uptake in the central regions (called: the striatum) of the brain. *See PET, Methods and Instrumentation*. Reproduced with permission from Prof. Leenders.

# Plate 50

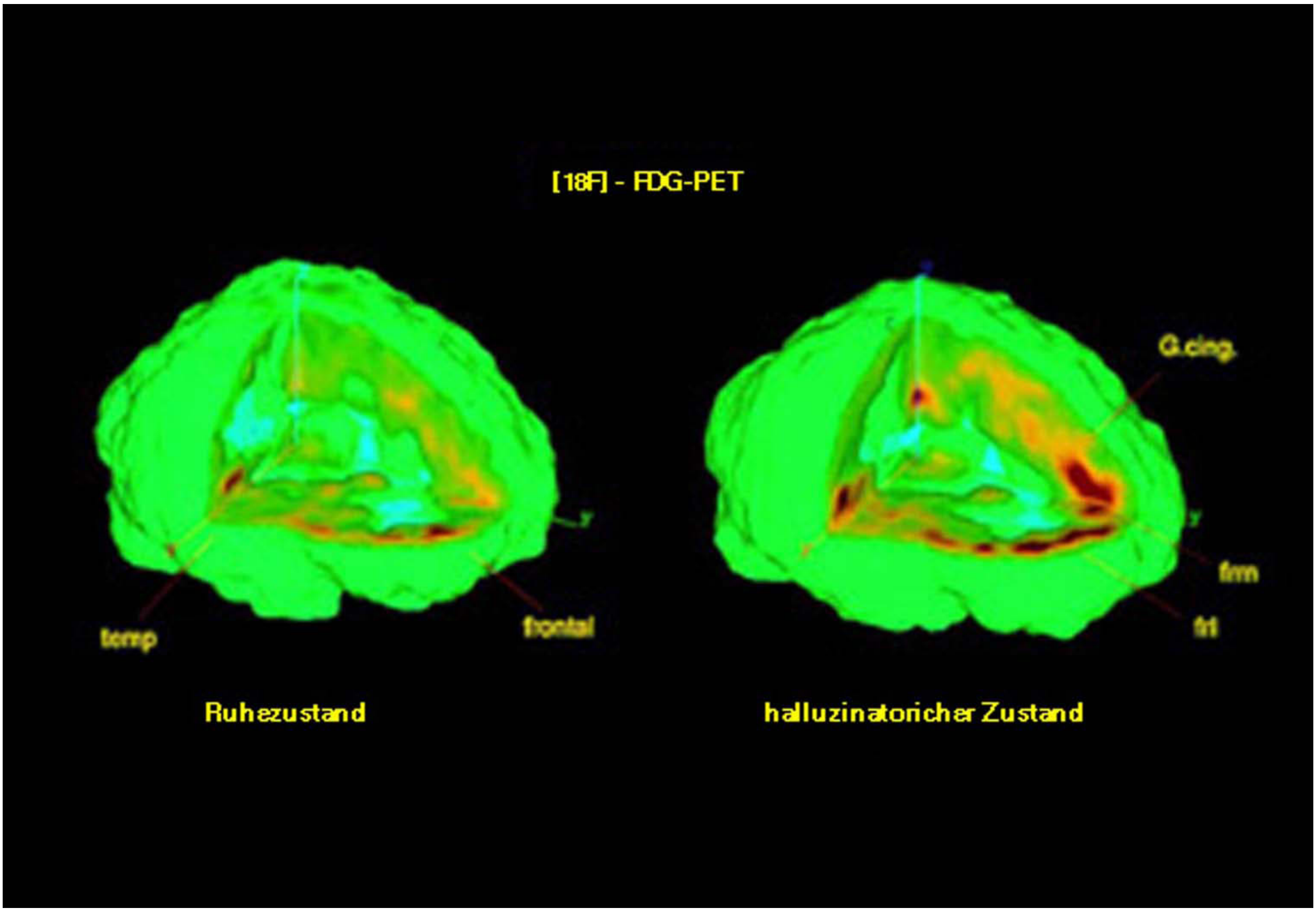



**Plate 50** 3-D computerized reconstruction of FDG uptake measured by PET in a healthy volunteer. Part of the right upper and frontal brain is "cut out". The left panel shows the tracer uptake in a rest condition. The right panel shows the uptake in the same person but now during intake of a hallucinogen. It can be seen that the frontal cortex region in the brain has a higher (more red) glucose uptake in the activated condition. *See PET, Methods and Instrumentation*. Reproduced with permission from Prof. Leenders.

# Plate 51

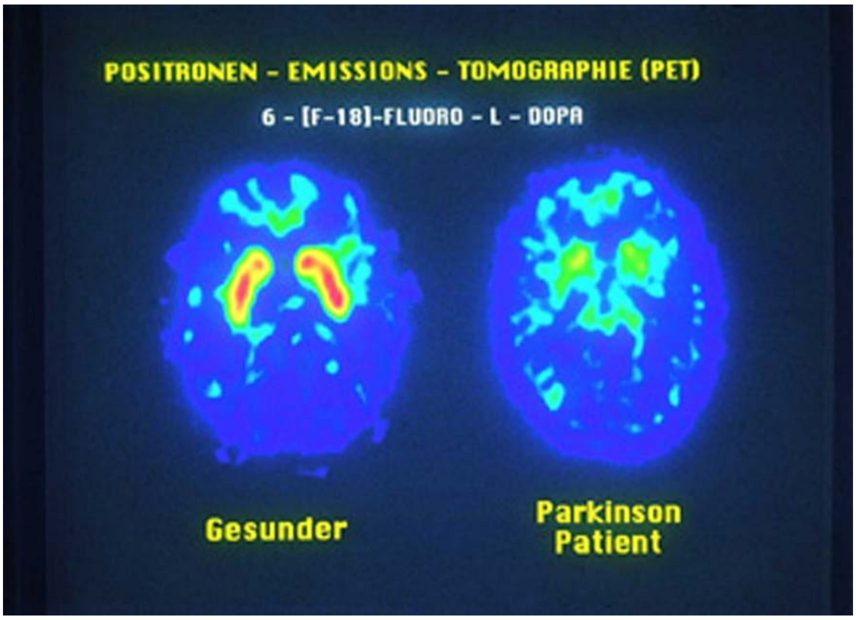



**Plate 51** Transaxial planes through the human brain after administration of the of the radiotracer fluoro-dopa indicating dopa decarboxylase capacity as measured by PET. The left panel shows a healthy volunteer whilst right panel shows a patient with Parkinson's disease. It can be seen that the tracer uptake measured with PET in the patient is markedly reduced within the regions of the basal ganglia. *See PET, Methods and Instrumentation*. Reproduced with permission from Prof. Leenders.

# Plate 52

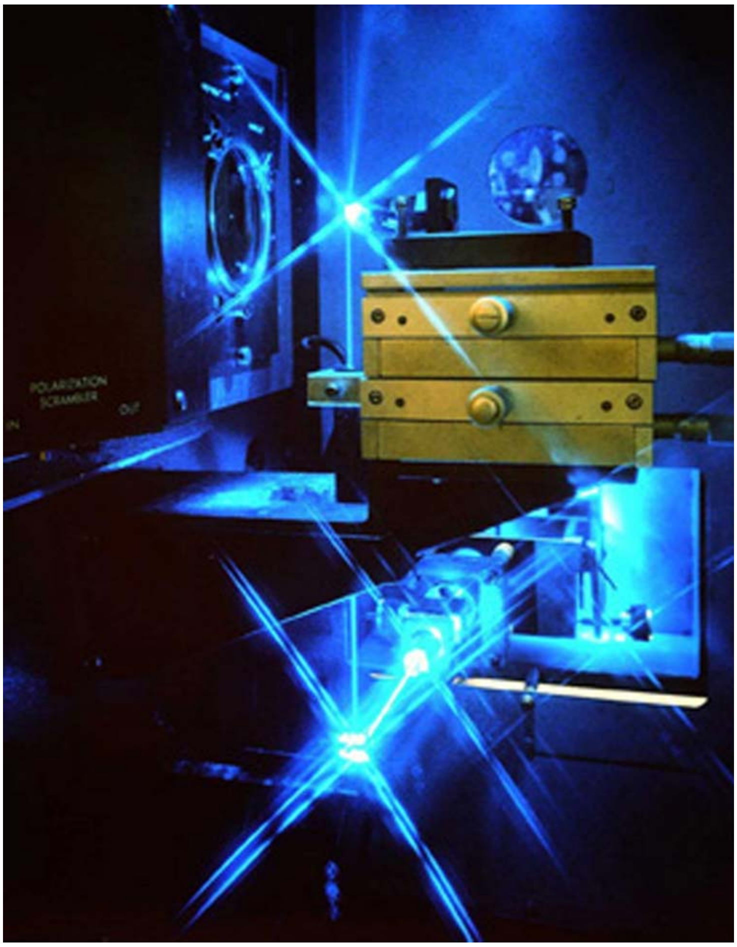



**Plate 52** Laser optical bench system used as an excitation source for a Raman spectrometer. Raman spectroscopy provides essentially the same sort of information about molecular structure and dimensions as do infrared and microwave spectroscopy. The laser is an ideal Raman source: it provides a narrow, highly mono-chromatic bearn of radiation which may be focused accurately into small smaple. *See Raman Spectrometers; FT-Raman Spectroscopy, Applications; Industrial Applications of IR and Raman Spectroscopy; Polymer Application of IR and Raman Spectroscopy; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds*. Reproduced with permission from Science Photo Library.

# Plate 53

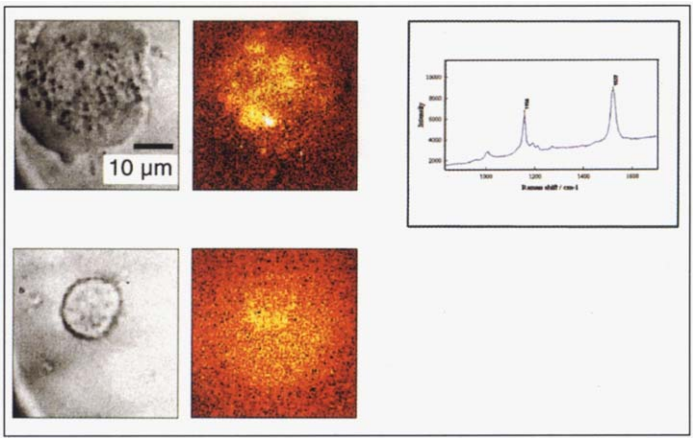



**Plate 53** Direct Raman imaging, using the **b**-carotene line at 1552cm $^{-1}$, of live corpus luteum cells. This provides a rapid insight into the distribution of different molecules within the cell (old cell below, young above). *See Raman and IR Microspectroscopy*. Reproduced with permission from Renishaw pic/Prof. D. N. Batchelder *et al*, Dept. Physics and Astronomy, University of Leeds.

# Plate 54

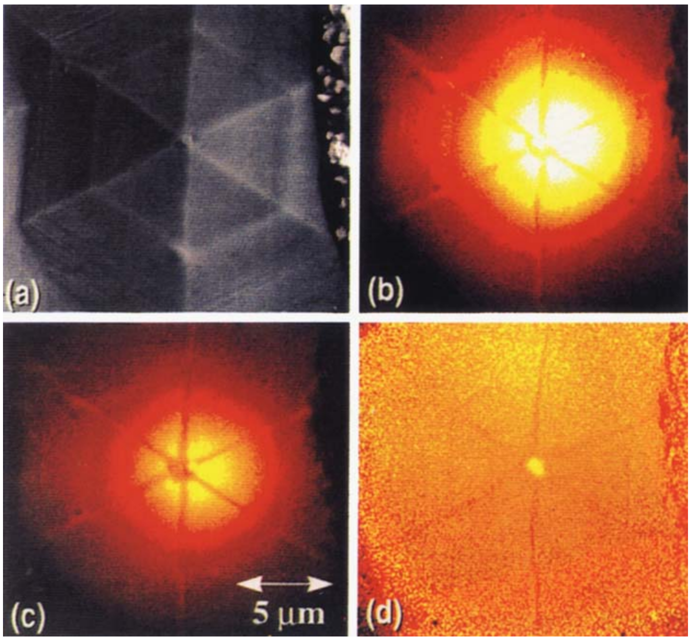



**Plate 54** Gallium Nitride Impurity Doping Images. *See Raman and IR Microspectroscopy*. Reproduced with permission from Renishaw pic.

# Plate 55

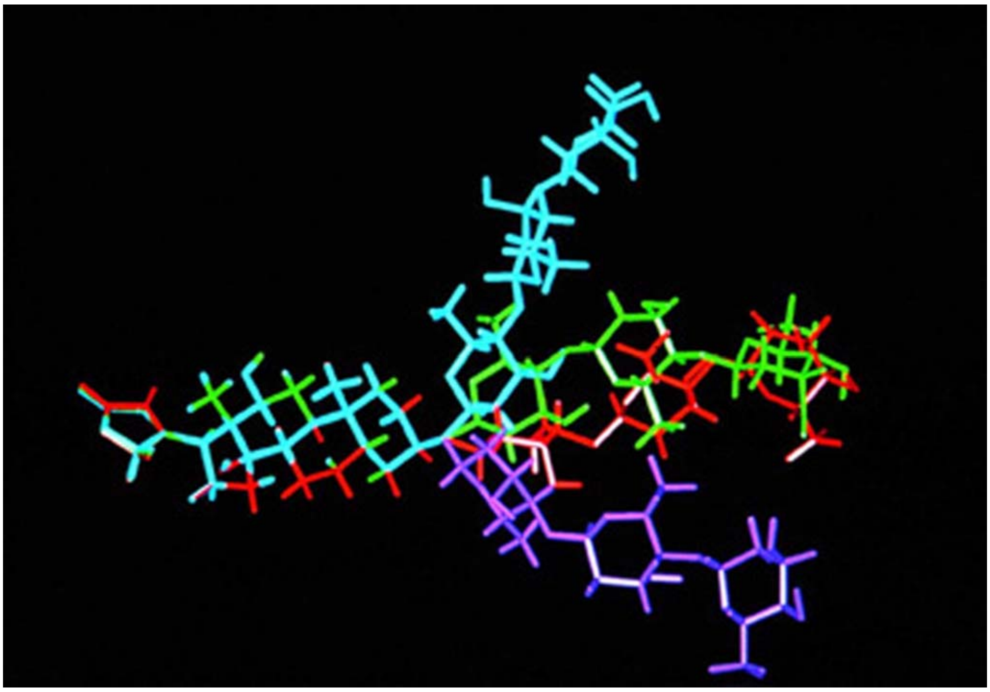


**Plate 55** A computer generated molecular model of the cardiac drug digoxin showing the molecular structure as derived by different techniques. The green structure is that determined in the solid state using X-ray diffraction. The red structure is calculated using a molecular dynamics approach. The cyan and magenta structures are the two possible structures determined in solution using NMR spectroscopy. *See Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals*. Reproduced with permission from John Lindon.

# Plate 56

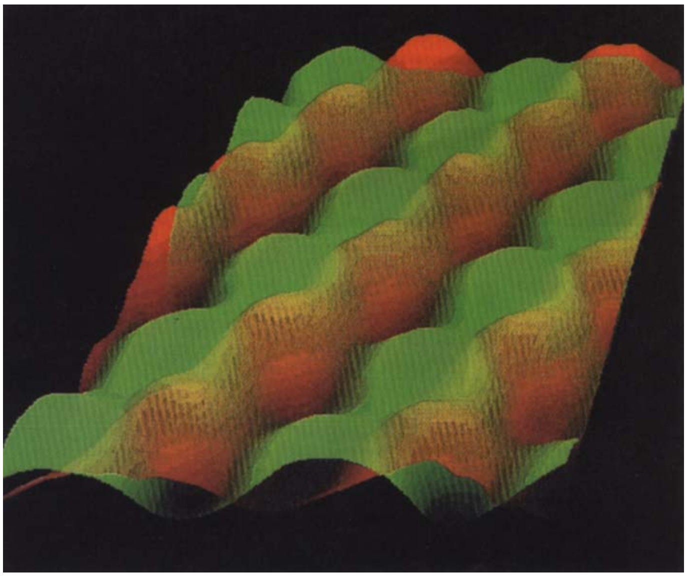


**Plate 56** A composite of two images of the GaAs(110) surface. The orange features obtained at positive sample bias are the Ga atoms, while green features obtained at a negative sample bias are the As atoms. Feenstra RM, unpublished results. *See Scanning Probe Microscopes*.

# Plate 57

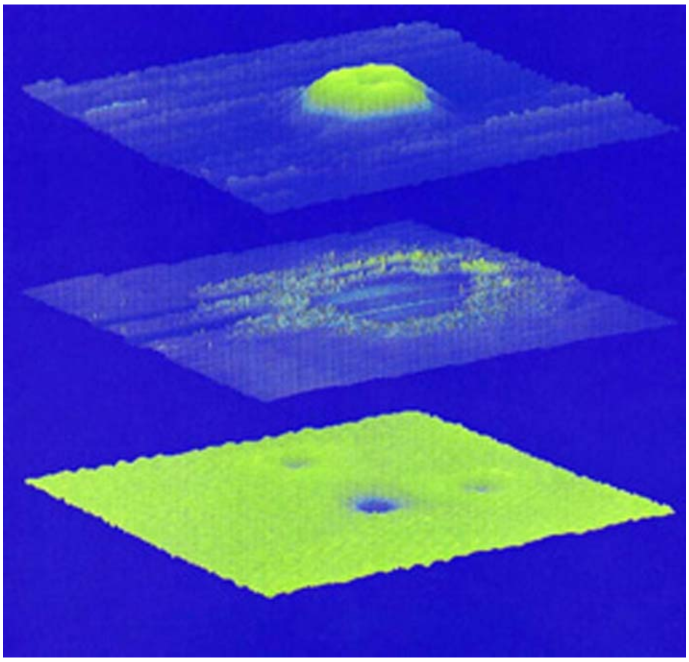


**Plate 57** Three STM images of a $Ni_3$ cluster adsorbed on a $MoS_2$ basal plane at 4 K. All three images show a 60Å area and are plotted as three-dimensional representations with the same aspect ratio and with the same angle of view. The images were acquired with sample biases of: +2 V (upper), +1.4 V (middle), and −2 V (lower). Reproduced with the permission of the American Chemical Society from Kushmerick JG and Weiss PS (1998) *Journal of Physical Chemistry* **B102**: 10094–10097. *See Scanning Probe Microscopes.*

# Plate 58

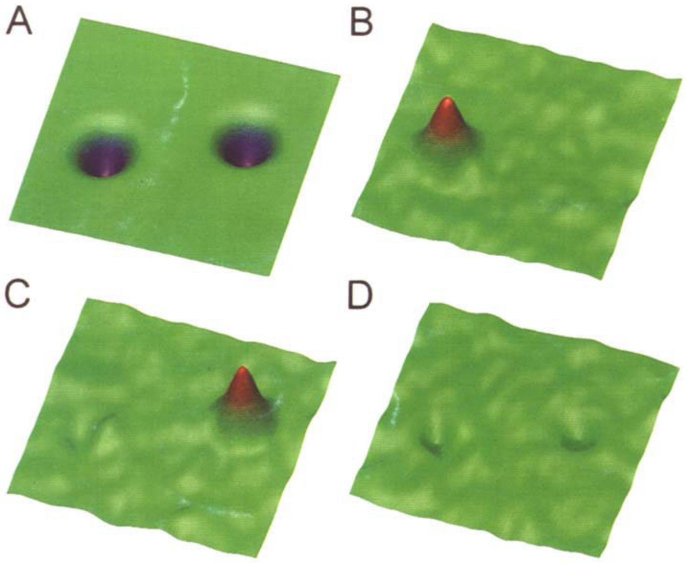



**Plate 58** Vibrational spectroscopic imaging of $C_2H_2$ and $C_2D_2$. (A) Constant current STM image of a $C_2H_2$ molecule (left) and $C_2D_2$ molecule (right). The d2/d $V^2$ images of the same area recorded with a bias voltage of (B) 358 mV, (C) 266 mV, and (D) 311 mV, with a 10 mV modulation. All images are 48 Å with 1 nA DC tunnelling current. Reproduced with the permission of the American Association for the Advancement of Science from Stipe BC *et al*. (1998) *Science* **280**: 1732–1735. *See Scanning Probe Microscopes*.

# Plate 59

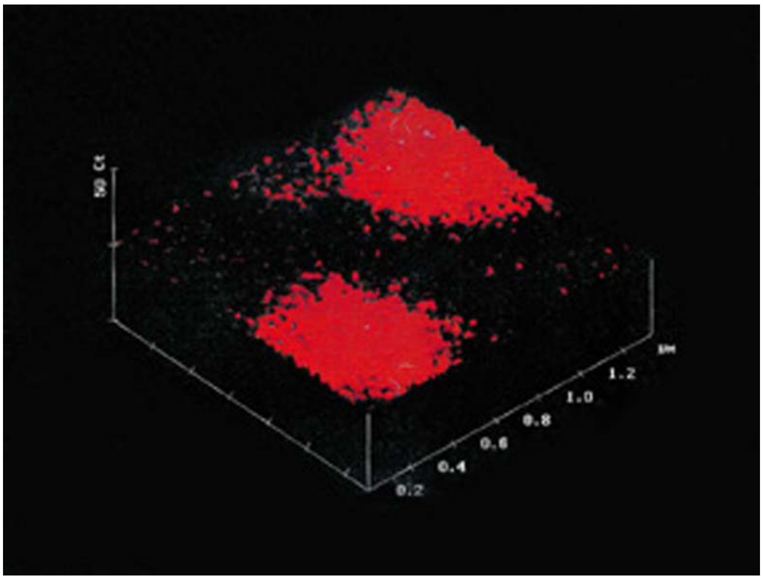



**Plate 59** Fluorescence emission near-field scanning optical microscope image (1.3 mm × 1.3 mm) of a photosynthetic membrane fragment. Reproduced with permission of the American Chemical Society from Dunn RC et al. (1994) Journal of Physical Chemistry 98: 30943098. *See Scanning Probe Microscopes.*

# Plate 60

**Plate 60** Sir Isaac Newton. He is shown using a prism to decompose light. *See Electromagnetic Radiation*. Reproduced with permission from Mary Evans Picture Library.

# A

## Antimony NMR, Applications

*See* **Heteronuclear NMR Applications (As, Sb, Bi).**

## Arsenic NMR, Applications

*See* **Heteronuclear NMR Applications (As, Sb, Bi).**

## Art Works Studied Using IR and Raman Spectroscopy

**Howell GM Edwards**, University of Bradford, UK

Copyright © 1999 Academic Press

**VIBRATIONAL ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

### Introduction

The scientific study of works of art and the materials that have been used in their creation has received impetus with the development of nondestructive, microsampling analytical techniques and an increasing awareness on the part of art historians, museum conservators and scientists of the importance of characterization for the attribution of the historical period and genuiness of an artefact. Useful information about ancient technologies and methods used in the construction of works of art is now forthcoming; in particular, spectroscopists are realizing the challenge that is being provided by the analytical characterization of ancient materials.

In this article, the applications of vibrational spectroscopic techniques to art are addressed and examples are taken to illustrate the following:

- identification of pigments, nondestructively, in preserved illuminated manuscripts, paintings and watercolours;
- characterization of genuine and fake artefacts, e.g. ivories;
- rock art and frescoes – effects of environmental and climatic degradation on exposed artwork;
- provision of information about archaeological and historical trade routes from the spectroscopic analysis of dyes, pigments and resins;
- identification of biomaterials in carved artwork, e.g. horn, hoof and tortoiseshell;
- information of critical importance to art restorers and museum conservation scientists – ill-restored items and the preservation of deteriorating material.

### Comparison of infrared and Raman spectroscopies

Although both Fourier transform-infrared (FT-IR) and Raman spectroscopies have been applied to the study of art and museum artefacts, it is possible to identify several indicators that will dictate the preferential use of either technique. Infrared studies have been reported from a much earlier date than Raman studies; hence there now exists much more comprehensive infrared database for pigments and natural or synthetic materials. Specific points also need to be considered relating to the sampling and composition of the specimens being studied. The presence of highly fluorescent coatings or impurities in ancient specimens dictated the use of infrared spectroscopy when the alternative was visible excitation for Raman spectroscopic studies; fluorescence often swamped the lower-intensity Raman signals from such samples.

However, the smaller scattering intensity of water and hydroxyl groups in the Raman effect, and their strong absorption in the infrared, generally favours the use of the Raman technique for the characterization of ancient, hydrated biomaterials such as linens and cottons. In the area of paintings and manuscripts, Raman spectroscopy has had extensive application in recent years because of the accessibility of the low-wavenumber regions of the vibrational spectrum (< 500 $cm^{-1}$), which are vitally important for the characterization of inorganic and mineral pigments. This region is observed only with difficulty in the infrared, and then with the use of special instruments with far-infrared capability.

The shape and reflectivity of surfaces are also important for specimens that cannot be subjected to mechanical or chemical pretreatment – the taking of small samples by drilling, scraping or excision is often prohibited. Curved surfaces are notoriously poor for infrared examination and specimens are better analysed after being subjected to a flattening or crushing process. The use of fibreoptic probes for 'remote' analysis is now advocated for many infrared and Raman applications, and art objects are no exception. However, the interpretation of the spectral data from remote-probe scanning experiments is dependent on accurate subtraction of probe background; this is particularly relevant for features below 1000 $cm^{-1}$, e.g. artists' pigments. However, the obvious advantage of having the capability of taking the Raman or infrared spectrometer into a museum environment for *in situ* examination of an artefact or painting has dictated the construction of several portable units that might well provide a new dimension to these applications.

Finally, the advent of FT-Raman spectroscopy with near-infrared excitation from a $Nd^{3+}$/YAG laser at 1064 nm has given a new dimension to art applications, especially for the study of naturally fluorescent biomaterials such as horn, hoof, tortoiseshell and ivory. Improvements in the generation of Raman spectra and their detection using CCD-visible Raman spectroscopy are now also providing valuable new ways of analysing often difficult museum specimens. The role of Raman microscopy (and also infrared microscopy) in the characterization of micro-sized specimens will be illustrated in this article. Generally, it is now possible to achieve good-quality Raman microscopy spectra from sample areas as small as 1–2 μm using visible excitation (~ 0.5 mm has been claimed in some instances) and about 5–15 μm using infrared and Raman FT techniques. The latest advances in confocal microscopy using visible excitation and CCD detection now means that it is possible to depth-profile suitable specimens with a resolution of about 1 μm or better; this is important for the analysis of coatings applied by artists to paintings.

## Applications of IR and Raman spectroscopies

### Plastics

Since the first commercial plastic, parkesine, in 1862, a huge range of functional articles have been produced in these media; many advantages were appreciated, including the ability to incorporate pigments and to mould large objects. Museums now have exhibitions devoted to plastic articles, but in recent years the problems in their conservation and storage and the control of degradation in damaged plastic exhibits has become acute. In particular, the acid ester plastics such as cellulose nitrate and cellulose acetate are prone to degradation, which has often resulted in major destruction of the articles concerned. Examples include early motion picture films, Victorian substitute 'tortoiseshell' articles, Bakelite and children's toys, such as dolls. FT-IR and Raman spectroscopy have been instrumental in discovering the methods of degradation and the source of possible 'triggers'; suggestions have thus been made for the proper storage and conservation of plastic articles for future generations, such as dolls from the 1940s, space suits from the 1960s and 'art nouveau' objects from the 1920s.

### Glass

Degradation processes in medieval stained glass windows have been studied using FT-IR spectroscopy and the pigments applied to early specimens have been characterized using the FT-Raman technique. Similar methods have been used to study early enamels and cloisonné specimens.

### Faience

Faience was produced in Egypt over 5000 years ago and may truly be considered as the first 'high-technology ceramic'. The faience body, formed from sand with additives such as lime and natron (a hydrated sodium carbonate found naturally as a mineral in dried lake beds), was shaped and modelled, and a glaze was applied and fired.

Sixteen pieces of Egyptian faience from Tell-el-Amarna have been analysed using Raman microscopy and the red and yellow pigments identified as red ochre and lead antimonate yellow, respectively.

A major problem in the characterization of the blue, white and green pigments resulted from a swamping of the pigment Raman signals by the silica in the intact glazes. Results for the red and yellow specimens were obtained from fractured samples.

### Biomaterials

The preservation and restoration of objects made from biomaterials is particularly challenging as their degradation products are complex and diverse. Examples of problems facing museum conservators include art objects made from ivory, horn, natural resins and textiles; these objects have often become fragile and restoration is of the utmost significance and importance. Often, earlier restorative procedures, which may have been incompletely documented, are no longer satisfactory and the results of chemical deterioration of applied restoratives under the influence of solar radiation and humidity changes in storage are sadly too often plain to see.

There are several good examples in the recent literature of the successful application of vibrational infrared and Raman spectroscopy to the characterization of art objects composed of biomaterials. Ivory, a generic name for the exoskeletal dental growths of certain mammalian species, has been appreciated as an art medium for thousands of years; it is soft enough to be carved and polished, yet hard enough to resist superficial weathering damage. However, the identification and attribution of archaeological ivories to species of elephant, sperm whale, narwhal, hog and hippopotamus, for example, is often fraught with difficulty, particularly where the ivory object is small and may only be a part of a larger specimen. The observation of Schreger lines and surface morphology is extremely difficult in such cases. Very recently, Raman spectroscopy has been used to provide a suggested protocol for ivory identification and characterization and has been used successfully to determine the animal origin (sperm whale) of a Roman die excavated at Frocester Roman villa (third century AD) in the UK (**Figure 1**).

There have also been several examples of the use of Raman spectroscopy to identify genuine and fake ivory articles. **Figures 2** and **3** shows some bangles, an Egyptian necklace and a carved cat, all of which were assumed to be genuine ivory dating from the seventeenth century or later; in some cases, the articles were found to be modern imitations. The case of the carved cat is very significant, as the Raman spectra show the presence of calcite that has been added to a polymer composite of poly(methyl methacrylate) and polystyrene to simulate the density of true elephant ivory. This specimen could not therefore be three hundred years old as these polymers have only been known in the last fifty years! The FT-Raman spectra of true ivories from different mammal sources are shown in **Figure 4** and of the fake specimens in **Figure 5**; the spectroscopic differences are clearly discernible and provide a means of identification between fake and real specimens.

'Scrimshaw' is a special name for carved whalebone and teeth of the sperm whale; many objects in museum collections date from the eighteenth-century production of carved decorative artwork in this material by whaling sailors for their families ashore. Genuine scrimshaw is now extremely valuable and items have been created to deceive the unwary.

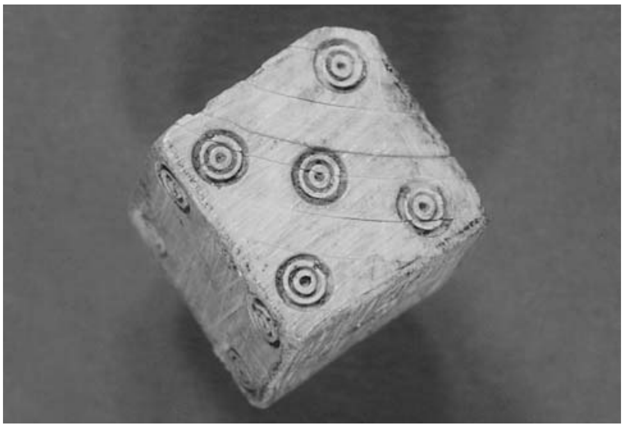


**Figure 1** Roman die, ca. AD 300, from archaeological excavations at Frocester Villa, Gloucester, UK. Raman spectroscopy has suggested the origin of the die as sperm whale ivory. (See Colour Plate 1).

**Figure 2** Selection of ornamental jewellery consisting of three bangles assumed to be ivory but which were shown spectroscopically to be composed of modern resins, and a genuine ivory necklace. (See Colour Plate 2).

**Figure 3** 'Ivory' cat, which was identified spectroscopically as a modern limitation composed of poly(methyl methacrylate) and polystyrene resins with added calcite to give the texture and density of ivory. Reproduced with permission from Edwards HGM and Farwell DW, Ivory and simulated ivory artefacts: Fourier-transform Raman diagnostic study, *Spectrochimica Acta, Part A*, **51**: 2073–2081 © 1995, Elsevier Science B. V. (See Colour Plate 3).

Spectroscopy has played a role in the characterization of genuine scrimshaw. **Figure 6** shows a stackplot of Raman spectra of carved solid and hollow sperm whale-tooth scrimshaw specimens, a staybusk, and a spill vase/quill pen holder; from these spectra the whale teeth and staybusk are confirmed to be genuine, but the spill vase/quill pen holder is identified as a modern fake made from polymer resin.

Similarly, infrared and Raman spectroscopic studies of ancient textiles are being undertaken to derive information about the possible processes of their degradation under various burial environments, e.g. Egyptian mummy wrappings, silk battle banners, Roman woollen clothing and artistic linens from funerary depositions. A classic, topical and ongoing controversy covers the spectroscopic studies associated with the Shroud of Turin and the nature of the pigmented areas on the ancient linen.

Ancient technologies and cultures often used naturally occurring biomaterials as decorative pigments and as functional repairing agents on pottery and glass. Recent studies have centred on the provision of a Raman spectroscopic database for native waxes and resins, which has been used for the nondestructive evaluation of archaeological items, e.g. 'Dragon's blood', and it is possible to identify different sources of this material from the spectra. FT-IR spectroscopy, too, has been successful in the characterization of conifer resins used in ancient amphorae and has been particularly useful for attribution of the Baltic geographical origins of amber jewellery through its succinic acid content and for the detection of modern fakes made from phenol–formaldehyde resins that are often passed off as examples of ancient ethnic jewellery.

The identification of dammar and mastic resins that have been used as spirit-soluble varnishes on

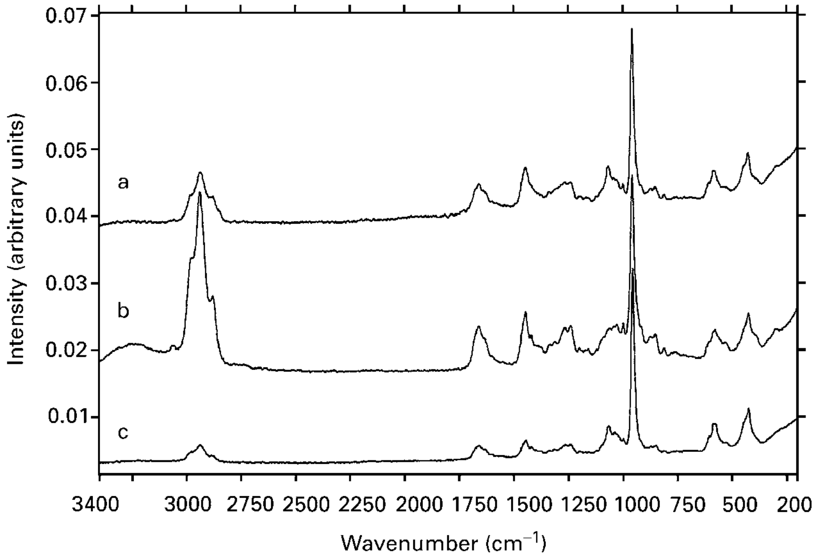



**Figure 4** FT-Raman spectra of true ivory; 1064 nm excitation, 500 spectral scans accumulated, 4 cm$^{-1}$ spectral resolution: (a) sperm whale ivory, (b) elephant ivory, (c) walrus ivory.

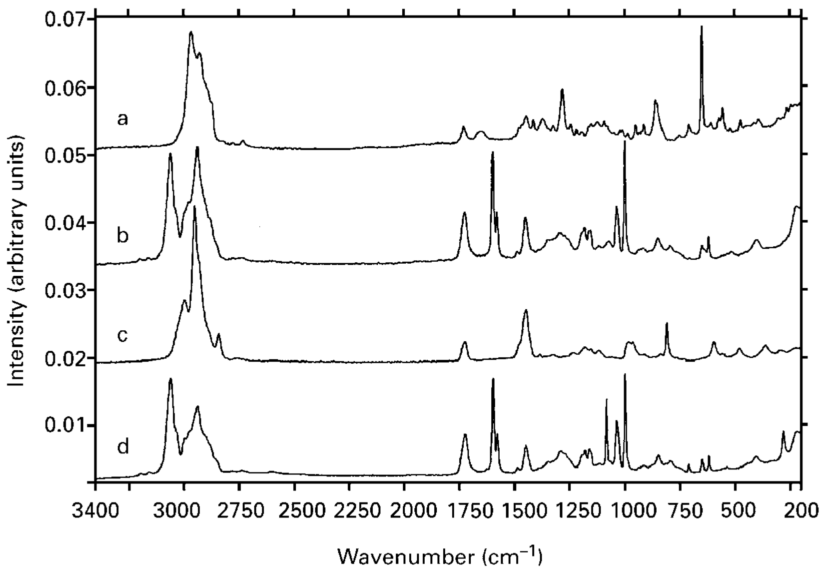



**Figure 5** FT-Raman spectra of fake ivory specimens; conditions as for **Figure 4**: (a) carved Victorian bangle, (b) large bangle, (c) small bangle, (d) cat. The absence of the characteristic proteinaceous features in true ivory near 1650 and 1450 cm$^{-1}$ and the strong phosphate mode near 960 cm$^{-1}$ should be noted. Also, the presence of the aromatic ring bands at 3060, 1600 and 1000 cm$^{-1}$ in (b) and (d) indicate a polystyrene resin content, while the carbonyl stretching band at 1725 cm$^{-1}$ in all fake specimens indicates the presence of poly(methyl methacrylate). In the cat specimen, the band at 1086 cm$^{-1}$ uniquely identifies a calcite additive in the specimens of imitation ivory studied. Reproduced with permission from Edwards HGM and Farwell DW, Ivory and simulated ivory artefacts: Fourier-transform Raman diagnostic study, *Spectrochimica Acta, Part A*, **51**: 2073–2081 © 1995, Elsevier Science B.V.

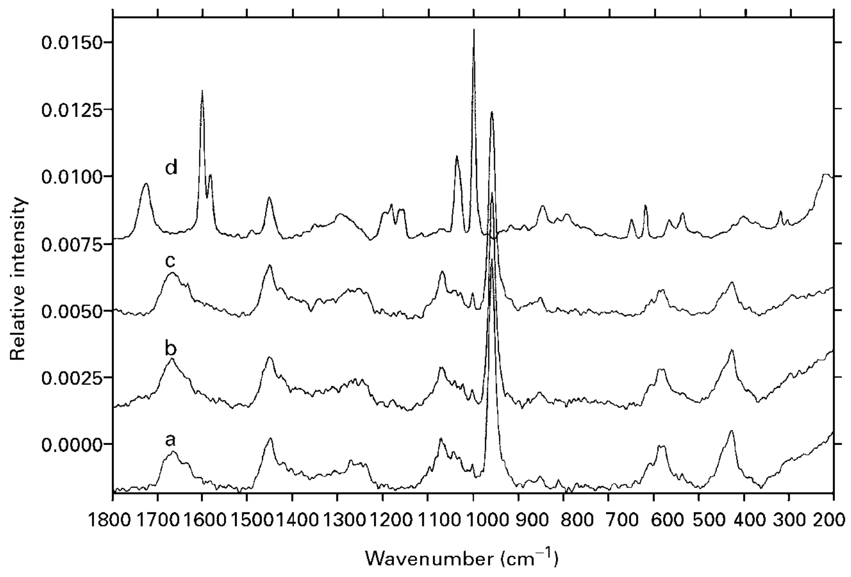



**Figure 6** FT-Raman stack-plot spectra of scrimshaw specimens: (a) hollow sperm whale tooth, (b) solid sperm whale tooth, (c) whalebone staybusk and (d) spill vase/quill pen holder. Minor spectroscopic differences confirm the whalebone origin of the staybusk. The modern resin composition of the spill vase/quill pen holder is also unambiguously identified from the aromatic ring stretching bands at 3060 cm$^{-1}$ and 1600 cm$^{-1}$. Reproduced with permission from Edwards HGM, Farwell DW, Sedder T and Tait JKF, Scrimshaw: real or fake? An FT-Raman diagnostic study, *Journal of Raman Spectroscopy*, **26**: 623–628 © 1995, John Wiley and Sons Ltd.

paintings and on early photographic prints, and the conservation of the latter, in particular, is very important because of the embrittlement of the substrate due to exposure to light and humidity changes. Raman spectroscopy has been used to identify wax coatings on early photographs from the American Civil War (ca. 1865) that are showing evidence of deterioration in museum collections.

### Paintings

The use of infrared and Raman spectroscopies for the characterization of paint pigments depends on two critical factors: first, the ability to record good-quality spectra nondestructively from small paint flakes or chips, and, second, the provision of a database of mineral, natural and synthetic pigments on which the basis of a comparison and attribution can be made. Some examples of the success of vibrational spectroscopic methods will now be given to illustrate the potential of these techniques for pigment analysis, and the conservation of degraded media.

Several databases now exist for the comparison of unknown pigments with contemporary specimens. A most useful basis for the attribution of specimens is provided by synthetic pigments, for which the dates of first manufacture or use are well established. For example, titanium(IV) oxide ($TiO_2$) is the most important white pigment in use today; its two major naturally occurring forms are rutile and anatase, which have been in use as paint media since 1923 and 1947, respectively. Their identification by vibrational spectroscopy is unambiguous; hence, if either is found on a disputed Renaissance painting, it could indicate a forgery or at least a recently restored work. It is interesting that the distinction between these two forms of $TiO_2$ is not achievable with X-ray fluorescence techniques normally used for pigment analysis in art.

A rather different example is provided by the work of Guys, who painted social life in France between 1843 and 1860. Seventy samples taken from 43 of Guys' works revealed a heavy dependence on newly synthesized and experimental pigments at that time, including Prussian blue, Cobalt blue and French ultramarine, sometimes in admixture and unique to an artist of his time. Restoration, therefore, needs careful attention to unusual combinations of experimental pigments. A similar study emerges from the FT-IR study of Morisot's work, *Bois de Boulogne*; this also showed the presence of novel combinations of pigments, but ones that were more stable than others that were in use at that time. Analysis of paint flakes from the *Virgin and Child*, a suspected fifteenth-century work, revealed the presence of expected organic binders and varnishes but also suggested that a very heavy reworking had taken place more recently, and that the work could be a forgery. On the 'Jónsbók' Icelandic manuscript, six pigments were identified by Raman microscopy, only bone white being indigenous to Iceland; the others, vermilion, orpiment, realgar, red ochre and azurite would all have been imported.

The most important naturally occurring and synthetic inorganic pigments used over periods of time and which feature in paintings from medieval to contemporary ages are given in **Tables 1** to **7** along with the year of manufacture or documented use; the tables are constructed according to colour or type, viz. blue, black, brown/orange, green, red, white and yellow. These tables form the basis of dating of paintings and manuscripts by vibrational spectroscopy. The stability of inorganic mineral pigments relative to 'fugative' organic dyes was realized in mediaeval times.

The principal dyes used by medieval dyers were indigo from woad for blue, alizarin and purpurin from madder for red, and luteolin from weld or crocetin from saffron for yellow; some had been used long before the Middle Ages and weld was known in the Stone Age. Organic pigments have also been used on manuscripts, notably saffron, weld, indigo, woad, Tyrian purple, madder and carmine.

Many natural products were extracted from lichens in the past and used to dye textiles. Notable among these were orchil for purple and crottle for brown. Other dye plants can yield greens, browns and blacks (in the last case, for example, marble gall from oak *Quercus* trees with added iron sulfate).

Organic pigments are prone to both fluorescence and photochemical degradation. Moreover, they often scatter only very weakly, perhaps owing to the fact that they may have been made into Colourfast lakes with a mordant such as alum; hence they are difficult to identify uniquely on a manuscript owing to the lack of concentration of pigment at the sampling point. The better known organic dyes and pigments are listed in **Table 8**.

The Raman and infrared spectra of organic pigments and dyes of relevance to artwork have now been recorded, including those of modern dyes such as methyl blue (a synthetic triarylmethane dye), methyl violet and perylene reds.

Perhaps the greatest advances in recent years in the field of nondestructive pigment identification in art works have come from historiated manuscripts that have been studied using Raman microscopy. *In situ* analyses of the brightly coloured initials and borders of the fourteenth-century Icelandic 'Jónsbók', Chinese manuscripts, illuminated Latin texts and early medieval Bibles have demonstrated the power of the technique. From samples that often represented only a 20 μm fragment that had fallen into the bindings of early manuscripts, FT-IR and Raman microscopy have made considerable advances in the knowledge of colour hue technology in medieval times; for example, the different blues in an historiated initial could be attributed to a finer particle size

**Table 1** Blue inorganic pigments

| *Pigment* | *Chemical name* | *Formula* | *Date / Source*[a] |
|---|---|---|---|
| Azurite | Basic copper(II) carbonate | $2CuCO_3{\cdot}Cu(OH)_2$ | Mineral |
| Cerulean blue | Cobalt(II) stannate | $CoO{\cdot}nSnO_2$ | 1821 |
| Cobalt blue | Cobalt(II)-doped alumina glass | $CoO{\cdot}Al_2O_3$ | 1775 |
| Egyptian blue | Calcium copper(II) silicate | $CaCuSi_4O_{10}$ | 3rd millennium BC/Mineral |
| Lazurite (from lapis lazuli) | Sulfur radical anions in a sodium aluminosilicate matrix | $Na_8[Al_6Si_6O_{24}]S_n$ | Mineral/1828 |
| Manganese blue | Barium manganate(VII) sulfate | $Ba(MnO_4)_2{+}BaSO_4$ | 1907 |
| Phthalocyanine blue (Winsor blue) | Copper(II) phthalocyanine | $Cu(C_{32}H_{16}N_6)$ | 1936 |
| Posnjakite | Basic copper(II) sulfate | $CuSO_4{\cdot}3Cu(OH)_2{\cdot}H_2O$ | Mineral |
| Prussian blue | Iron(III) hexacyanoferrate(II) | $Fe_4[Fe(CN)_6]_3{\cdot}14\text{–}16H_2O$ | 1704 |
| Smalt | Cobalt(II) silicate | $CoO{\cdot}nSiO_2(+K_2O+Al_2O_3)$ | ca. 1500 |
| Verdigris | Basic copper(II) | $Cu(O_2CCH_3)_2{\cdot}2Cu(OH)_2$ | Mineral |

[a] The pigment is either specified to be a mineral and/or the date of its manufacture is listed.

**Table 2** Black inorganic pigments

| *Pigment* | *Chemical name* | *Formula* | *Date / Source* |
|---|---|---|---|
| Ivory black[a] | Calcium phosphate + carbon | $Ca_3(PO_4)_2 + C + MgSo_4$ | 4th century BC ? |
| Lamp black | Amorphous carbon | C | ~ 3000 BC |
| Magnetite | Iron(II,III) oxide | $Fe_3O_4$ | Mineral |
| Mineral black | Aluminium silicate + carbon (30%) | $Al_2O_3{\cdot}nSiO_2 + C$ | Mineral |
| Vine black | Carbon | C | Roman |

[a] Bone black is similar to ivory black.

**Table 3** Brown/orange inorganic pigments

| *Pigment* | *Chemical name* | *Formula* | *Date / Source* |
|---|---|---|---|
| Cadmium orange | Cadmium selenosulfide | Cd(S,Se) or CdS (>5 μm) | Late 19th century |
| Ochre (goethite) | Iron(III) oxide hydrate | $Fe_2O_3{\cdot}H_2O$ + clay, etc. | Mineral |
| Sienna (burnt) | Iron(III) oxide | $Fe_2O_3$ + clay, etc. | Antiquity? |

[a] Bone black is similar to ivory black.

**Table 4** Green inorganic pigments

| *Pigment* | *Chemical name* | *Formula* | *Date / Source* |
|---|---|---|---|
| Atacamite | Basic copper(II) chloride | $CuCl_2{\cdot}3Cu(OH)_2$ | Mineral |
| Chromium oxide | Chromium(III) oxide | $Cr_2O_3$ | Early 19th century |
| Cobalt green | Cobalt(II) zincate | $CoO{\cdot}nZnO$ | 1780 |
| Emerald green | Copper(II) arsenoacetate | $Cu(C_2H_3O_2)_2{\cdot}3Cu(AsO_2)_2$ | 1814 |
| Green earth – a mix of celadonite and glauconite | Hydrous aluminosilicate of magnesium, iron and potassium | Variations on $K[(Al^{III}{\cdot}Fe^{III})(Fe^{II}{\cdot}Mg^{II})]$ $(AlSi_3{\cdot}Si_4)O_{10}(OH)_2$ | Mineral |
| Malachite | Basic copper(II) carbonate | $CuCO_3{\cdot}Cu(OH)$ | Mineral |
| Permanent green deep | Hydrated chromium(III) oxide + barium sulfate | $Cr_2O_3{\cdot}2H_2O{+}BaSO_4$ | Latter half of 19th century |
| Phthalocyanine green | Copper(II) chlorophthalocyanine | $Cu(C_{32}H_{15}ClN_8)$ | 1938 |
| Pseudo-malachite | Basic copper(II) phosphate | $Cu_3(PO)_4)_2{\cdot}2Cu(OH)_2$ | Mineral |
| Verdigris (basic) | Basic copper(II) acetate | $Cu(C_2(C_2H_3O_2)_2{\cdot}2Cu(OH)_2$ | Mineral and synthetic (BC) |
| Viridian | Hydrated chromium(III) oxide | $Cr_2O_3{\cdot}2H_2O$ | 1838 (?1850) |

**Table 5** Red inorganic pigments

| *Pigment* | *Chemical name* | *Formula* | *Date / Source* |
|---|---|---|---|
| Cadmium red | Cadmium selenide | CdSe | ca. 1910 |
| Chrome red | Basic lead(II) chromate | $PbCrO_4{\cdot}Pb(OH)_2$ | Early 19th century |
| Litharge | Lead(II) oxide | PbO | Antiquity |
| Realgar | Arsenic(Ii) sulfide | $As_2S_2$ | Mineral |
| Red lead (minimum) | Lead(II,IV) oxide | $Pb_3O_4$ | Antiquity |
| Red ochre | Iron(III) oxide + clay + silica | $Fe_2O_3{\cdot}H_2O$ + clay + silica | Mineral |
| Vermilion (cinnabar)[a] | Mercury(II) sulfide | HgS | Mineral and synthetic (13th century) |

[a] Limited lightfastness (→ black form).

**Table 6** White inorganic pigments

| *Pigment* | *Chemical name* | *Formula* | *Date / Source* |
|---|---|---|---|
| Anatase | Titanium(IV) oxide | $TiO_2$ | 1923 |
| Barytes | Barium sulfate | $BaSO_4$ | Mineral |
| Bone white | Calcium phosphate | $Ca_3(PO_4)_2$ | Antiquity |
| Chalk (whiting) | Calcium carbonate | $CaCO_3$ | Mineral |
| Gypsum | Calcium sulfate | $CaSO_4{\cdot}2H_2O$ | Mineral |
| Kaolin | Layer aluminosilicate | $Al_2(OH)_4Si_2O_5$ | Mineral |
| Lead white | Lead(II) carbonate (basic) | $2PbCO_3{\cdot}Pb(OH)_2$ | Mineral and synthetic (500–1500 BC) |
| Lithopone | Zinc sulfide and barium sulfate | $ZnS + BaSO_4$ | 1874 |
| Rutile | Titanium(I) oxide | $TiO_2$ | 1947 |
| Zinc white | Zinc oxide | $ZnO$ | 1834 |

**Table 7** Yellow inorganic pigments

| *Pigment* | *Chemical name* | *Formula* | *Date / Source* |
|---|---|---|---|
| Barium yellow | Barium chromate | $BaCrO_4$ | Early 19th century |
| Cadmium yellow | Cadmium sulfide | $CdS$ | Mineral (greenockite) + synthetic ca. 1845 |
| Chrome yellow | Lead(II) chromate | $PbCrO_4$ or $PbCrO_4{\cdot}2PbSO_4$ | 1809 |
| Cobalt yellow (aureolin) | Potassium cobaltinitrite | $K_3[Co(NO_2)_6]$ | 1861 |
| Lead antimonate yellow | Lead(II) antimonate | $Pb_2Sb_2O_7$ or $Pb_3(SbO_4)_2$ | Antiquity |
| Lead tin yellow | Lead(II) stannate | [1] $Pb_2SnO_4$<br>[1] $PbSn_{0.76}Si_{0.24}O_3$ | Antiquity ?<br>Antiquity ? |
| Massicot | Lead(II) oxide | $PbO$ | Antiquity |
| Ochre | Geothite + clay + silica | $Fe_2O_3{\cdot}H_2O$ + clay + silica | Mineral (and synthetic) |
| Orpiment | Arsenic(II) sulfide | $As_2S_3$ | Mineral |
| Strontium yellow | Strontium chromate | $SrCrO_4$ | Early 19th century |
| Zinc yellow | Zinc chromate | $ZnCrO_4$ | Early 19th century |

**Table 8** Organic pigments and dyes

| *Colour* | *Pigment* | *Formula / Composition* | *Origin (Date)* |
|---|---|---|---|
| Blue | Indigo | Indigotin $C_{16}H_{10}N_2O_2$ | Plant leaf (BC), synthetic (1878) |
| Black | Bitumen | Mixture of hydrocarbons | (BC) |
| Brown | Sepia | Melanin | Ink of cuttlefish (ca. 1880) |
| | Van Dyck brown | Humic acids | Lignite containing manganese (16th century?) |
| | | Allomelanins | |
| Green | Sap green | Organic dye | Buckthorn berry (14th century?), coal-tar dye |
| Purple | Tyrian purple | 6,6′-Dibromoindigotin, $C_{16}H_8Br_2N_2O_2$ | Marine mollusc (1400 BC), synthetic (1903) |
| Red | Carmine | Carminic acid, $C_{22}H_{20}O_{13}$ | Scale insect, cochineal (Aztec) |
| | | Kermesic acid, $C_{16}H_{10}O_8$ | Scale insect, kermes (antiquity) |
| | Madder | {Alizarin, $C_{14}H_8O_4$<br>{Purpurin, $C_{14}H_8O_5$ | Madder root (3000 BC)<br>Synthetic alizarin (1868) |
| | Permanent red | Various azo dyes | Synthetic (after 1856) |
| Yellow | Gamboge | $\alpha$- and $\beta$-Gambogic acids<br>$C_{38}H_{44}O_8$ and $C_{29}H_{36}O_6$ | Gum-resin (before 1640) |
| | Hansa yellow | Various azo dyes | Synthetic (1900) |
| | Indian yellow | Magnesium salt of euxanthic acid<br>$MgC_{19}H_{16}O_{11}{\cdot}5H_2O$ | Cow urine (15th century) |
| | Quercitron | Quercitrin $C_{21}H_{20}O_{11}$ | Inner bark of *Quercus* oak |
| | Saffron | Crocetin $C_{20}H_{24}O_4$ | Crocus flower stigma (antiquity) |
| | Weld | Luteolin $C_{15}H_{10}O_6$ | Plant foliage (Stone Age) |

Tables 1–8 are reproduced with some modifications from Clark RJH, Raman microscopy: application to the identification of pigments on medieval manuscripts, *Chemical Reviews*, 187–196. 1995 © The Royal Society of Chemistry.

and not to a dilution with other materials or colours. Useful information has also been provided about additives and binding media.

On some paintings and manuscripts, an inappropriate blend of adjacent colours or mixtures of pigments and binders was achieved; Raman microscopy has identified two examples of these in cadmium sulfide and copper arsenoacetate (which yields black copper sulfide) and egg tempera with lead white (which yields black lead sulfide). The term 'inappropriate' here refers to the instability of pigments and pigment mixtures resulting in a chemical reaction over periods of time and through aerial or substratal influences.

Where the pigment is coloured, the choice of exciting line for Raman spectroscopy is extremely important because absorption of the scattered light by the sample may affect the spectrum. In such cases, the exciting line is chosen to fall outside the contour of the electronic absorption bands of the pigment. Vermilion (HgS) and red ochre ($Fe_2O_3$) give poor-quality Raman spectra using green excitation but give strong Raman spectra when excited with red radiation. An interesting example of the effects due to absorption of laser radiation is provided by the Raman spectra of red lead ($Pb_3O_4$) obtained using 514.5 and 632.8 nm excitation (**Figure 7**). In each case, well-defined Raman spectra are obtained, but only with 632.8 nm excitation is the spectrum that of the genuine material; that obtained with 514.5 nm excitation matches that of massicot (PbO). These observations may be explained because red lead can be converted into massicot by heat. Red lead absorbs 514.5 nm radiation strongly, leading to localized heating that results in conversion of the irradiated particle or particles into massicot. Since 632.8 nm radiation is not absorbed by the red lead, there is minimal local heating with this exciting line and thus no decomposition.

For coloured pigments, the absorption of exciting radiation can be used to advantage to produce enhanced Raman scattering through the resonance Raman effect. This is particularly useful for weakly scattering species, and selection of excitation within the contour of an electronic absorption band of a chromophore can produce a spectral intensity several orders of magnitude larger than that obtained in conventional Raman spectra. A classic example of this is provided by the deep blue mineral lapis lazuli, $Na_8[Al_6Si_6O_{24}]S_n$, where the species $S_2^-$ and $S_3^-$ are responsible for the colour. The species $S_3^-$, although present to an extent of $< 1\%$ in the pigment, gives such an intense Raman spectrum with green-red excitation that no bands due to the host lattice are observed. Other techniques fail to discriminate the presence of $S_3^-$ in the aluminosilicate lattice.

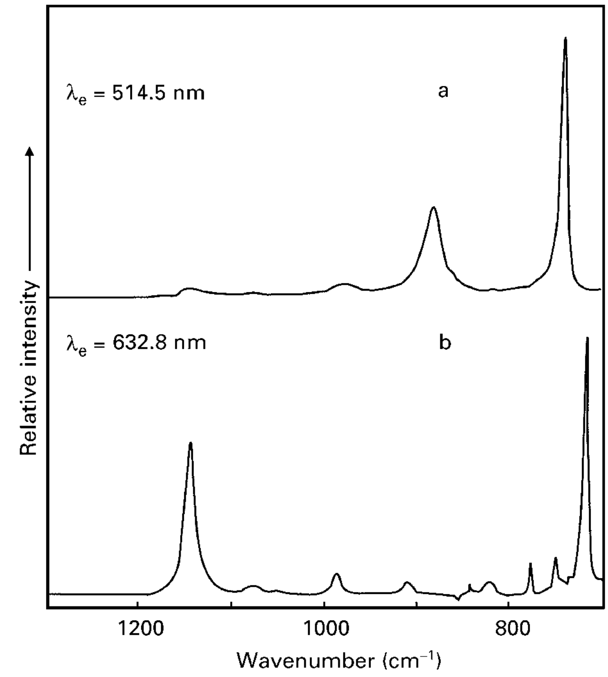



**Figure 7** Illustration of the effect of wavelength of laser excitation on the Raman spectra of a pigment, red lead ($Pb_3O_4$), excited with (a) 514.5nm and (b) 632.8nm radiation. The genuine spectrum is (b); the spectrum excited by green radiation in (a) corresponds to massicot (PbO), converted from $Pb_3O_4$ by localized heating in the laser beam. Reproduced with permission from Bert SP, Clark RJH and Withnall R, Non-destructive pigment analysis of artefacts by Raman microscopy, *Endeavour, New Series*, **16**: 66–73 © 1992, Elsevier Science.

The Raman spectrum is sensitive to both composition and crystal form, as is demonstrated by titanium(IV) dioxide, the most important white pigment in use today. White pigments normally present considerable problems in pigment identification, comprising sulfates, carbonates and phosphates, which are easily distinguished in Raman spectroscopy (**Figure 8**). The symmetric vibration of the anion gives rise to an intense band at ~1000 cm$^{-1}$ for sulfates, ~1050 cm$^{-1}$ for carbonates and ~960 cm$^{-1}$ for phosphates. The exact wavenumbers of these bands are also sensitive to the cation (1050 cm$^{-1}$ for $PbCO_3$ and 1085 cm$^{-1}$ for $CaCO_3$). This factor becomes extremely important in database construction since artistic vocabulary generally describes the colour rather than a precise mineral origin, e.g. cadmium yellow, although strictly CdS has also been designated for organic substitutes of similar hue. Old recipes for obtaining pigment colours are often vague and employ unidentifiable materials; the same chemical compound or mixture can even have different names according to geographical locality or historical period. For example, three contemporary samples designated Naples Yellow, assumed to be

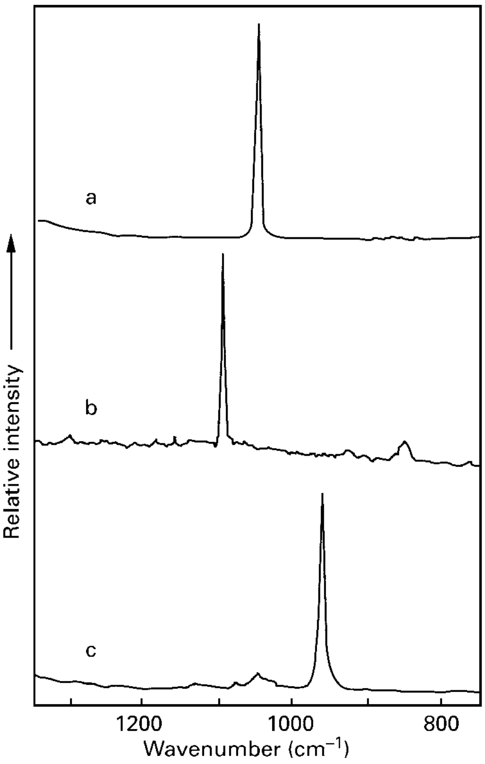



**Figure 8** Raman spectra of white pigments; differentiation between (a) lead white ($PbCO_3$) (b) chalk ($CaCO_3$) and (c) bone white ($Ca_3(PO_4)_2$). Reproduced with permission from Best SP, Clark RJH and Withnall R. Nondestructive pigment analysis of artefacts by Raman microscopy, *Endeavour, New Series* **16**: 66–73 © 1992, Elsevier Science.

$Pb_2\ (SbO_4)_2$ – a lead antimonate – have been shown by Raman spectroscopy to be a lead antimony oxide, $Pb_2Sb_2O_6$, another oxide, $Pb_2Sb_2O_3$, and the third sample mainly $Pb_2(CrO_4)_2$, but also containing $PbCO_3$. None of the samples tested actually proved to be Naples Yellow of the assumed formula!

## Medieval manuscripts

Most pigments in medieval manuscript illuminations are inorganic. They are coated with a small quantity of binding material and deposited on parchment that has previously been rubbed with pumice. The small size of the illuminations limits the number of spectroscopic samplings.

The blue colour in a set of six French manuscripts from the twelfth century was studied in different kinds of initial – historiated (with a descriptive scene), decorated, or simply coloured – which showed variable hues of pure blue, grey and sometimes violaceous. Raman microspectra obtained from all of the samples show common features: they are dominated by a strong band at 548 $cm^{-1}$, with weaker bands at 260, 585 and 1098 $cm^{-1}$. This group of four bands is characteristic of ultramarine blue and reveals the presence of only this pigment. In the blue-grey samples, microscopic observation showed tiny black particles that were identified by their Raman spectra as graphite.

**Table 9** The Raman bands of inorganic pigments in the Lucka Bible

| *Pigment* | *Raman bands ($cm^{-1}$)* |
|---|---|
| Azurite | 248w, 404vw, 770m, 838vw, 1098m, 1424w, 1578w |
| Lapis lazuli | 259w, 549vw, 807w, 1096m, 1355vw, 1641w |
| White lead | 1054s |
| Malachite | 225w, 274w, 355w, 437m, 516vw, 540w, 724w, 757vw, 1064w, 110w, 1372w, 1498m |
| Orpiment | 136m, 154m, 179w, 203m, 293s, 311s, 355s, 384m, 587vw |
| Realgar | 124vw, 143m, 166w, 172w, 183s, 193s, 214w, 222s, 329w, 345m, 355s, 370w, 376w |
| Red lead | 121vs, 152m, 223w, 232w, 313w, 391w, 477w, 549s |
| Vermilion | 254s, 281w, 344m |

Reproduced with permission from Best SP, Clark RJH, Daniels MAM and Withnall R, A bible laid open, *Chemistry in Britain*, **2**: 118–122 © 1993, The Royal Society of Chemistry.

Ultramarine has been identified in a manuscript commentary on *Ezekiel*, ca. AD 1000, in the abbey of St Germain, Auxerre, France. The pigment ultramarine is first mentioned in written sources in the thirteenth century. This discovery established for the first time that the pigment was in fact introduced into Europe almost two centuries earlier. The manuscript was to reveal more fascinating details. Part of the dedication image at the beginning of the manuscript is a figure representing the Abbot, Heldric, kneeling before St Germain, who died about 448 and is buried in the abbey. From the spectrum of the pigments, it was established that the Abbot's blue garments were painted with indigo (woad) while only the patron saint's more glorious garment was covered with the rare and expensive lapis lazuli, thus emphasizing his importance.

The letter 'I' in the Lucka Bible at the beginning of the Book of Genesis displays seven scenes representing the seven days of Creation. The varied palette (**Table 9**) has been well characterized using Raman microscopy (**Figure 9**).

The Skard copy of the Jónsbók Icelandic Law Code dates from ca. 1360 and is one of the most outstanding examples of an Icelandic medieval manuscript. The pigments have been analysed by Raman microscopy; three historiated initials and about 15 illuminated initials with associated background paintings and embellishments have been examined. Six pigments were identified unambiguously by

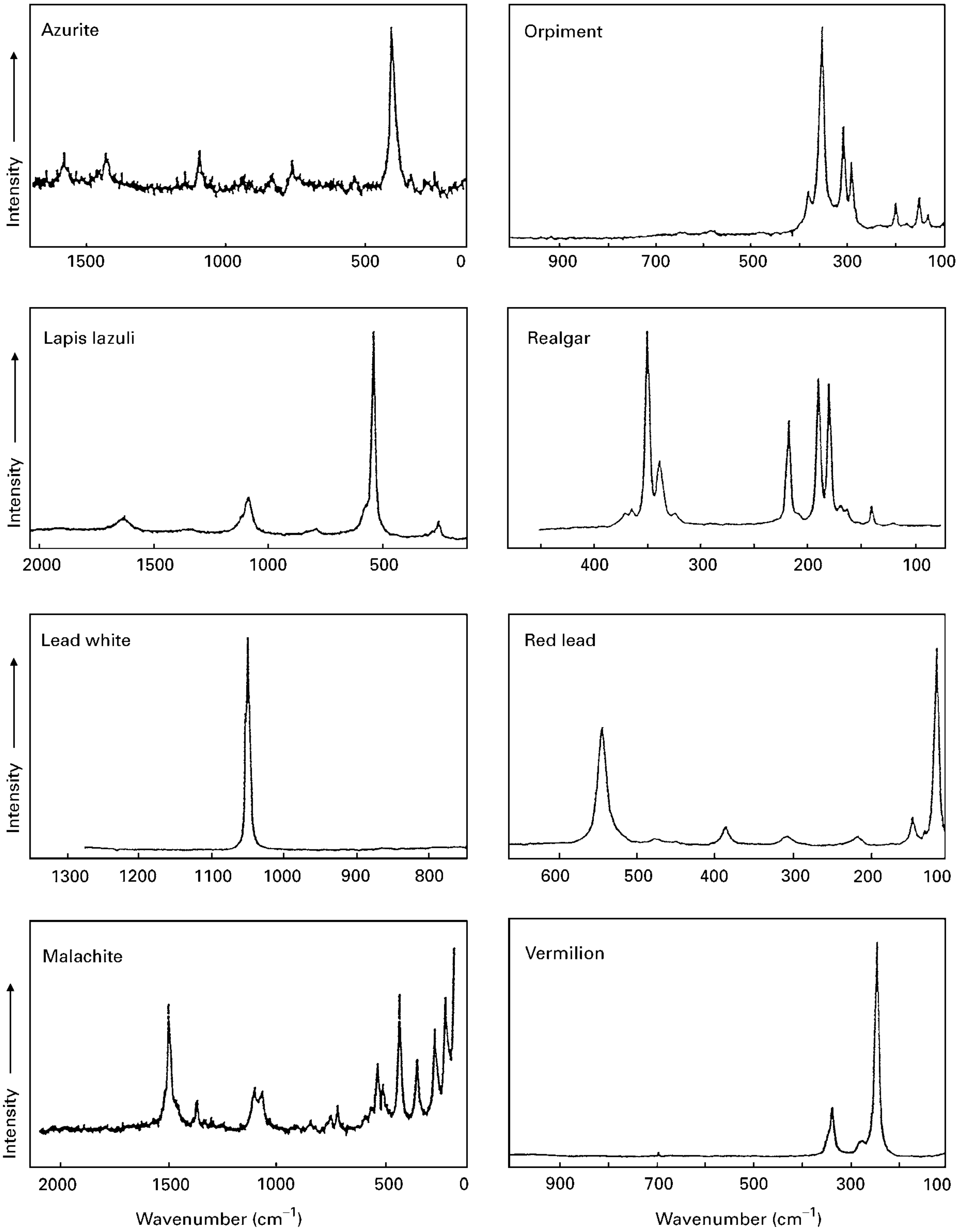



**Figure 9** Raman spectra of the historiated letter 'I' from the Book of Genesis, Lucka Bible, from which the data and identification of the mineral pigments in **Table 9** have been derived. Reproduced with permission of Clark RJH, Raman microscopy: application to the identification of pigments on medieval manuscripts, *Chemical Reviews*, 187–196 © 1995, The Royal Society of Chemistry.

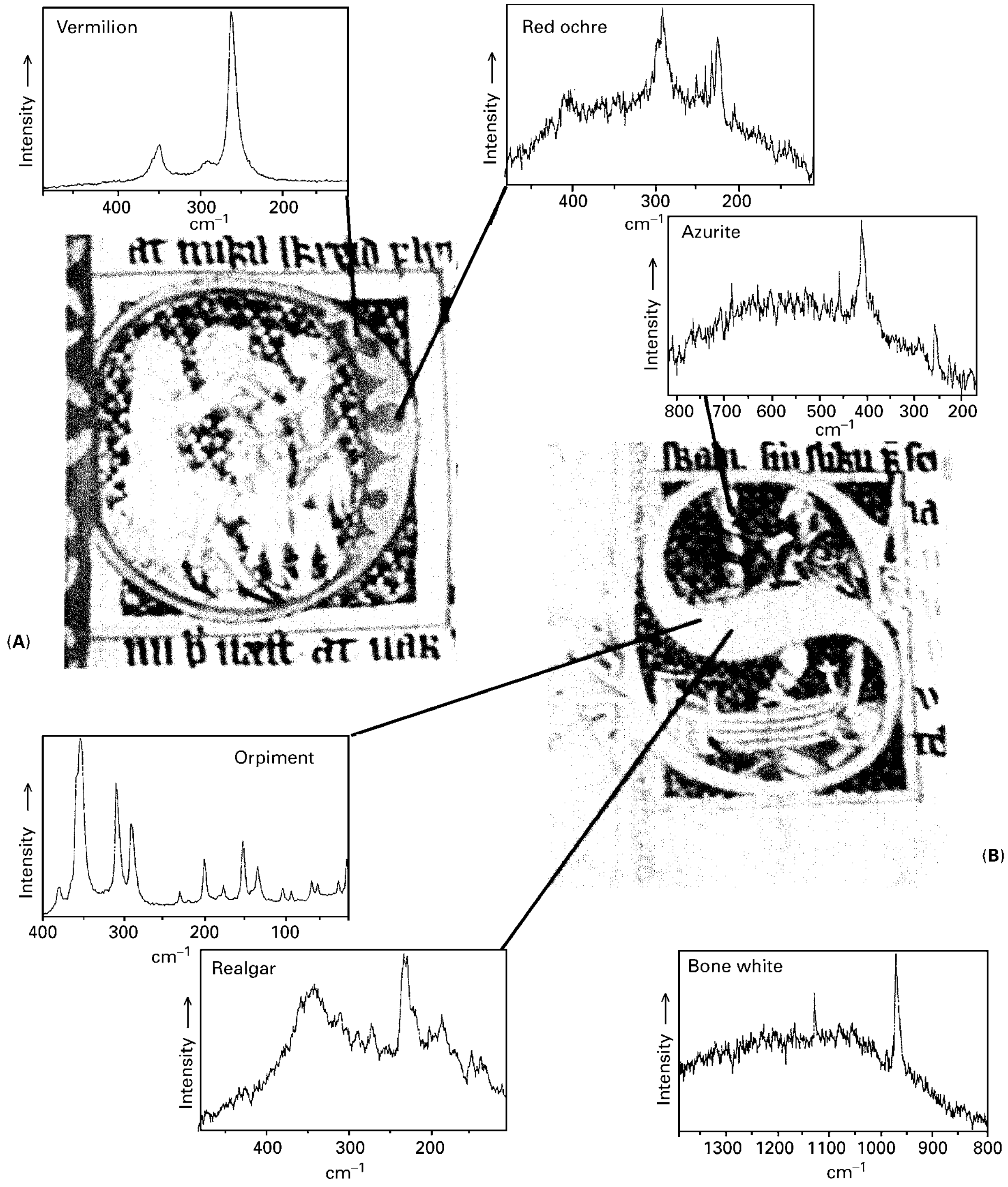



**Figure 10** The historiated initials (A) ‘P’ and (B) ‘S’ from the Icelandic Jónsbók with the Raman spectra of pigments contained therein. The combination of vermilion and red ochre in the ‘P’ should be noted. The spectrum of bone white has been obtained from the letter ‘H’ (not shown). Reproduced with permission from Best SP, Clark RJH, Daniels MAM, Porter CA and Withnall R (1995). Identification by Raman microscopy and visible reflectance spectroscopy of pigments on an Icelandic manuscript. *Studies in Conservation* **40**: 31–40. (See Colour Plate 4.)

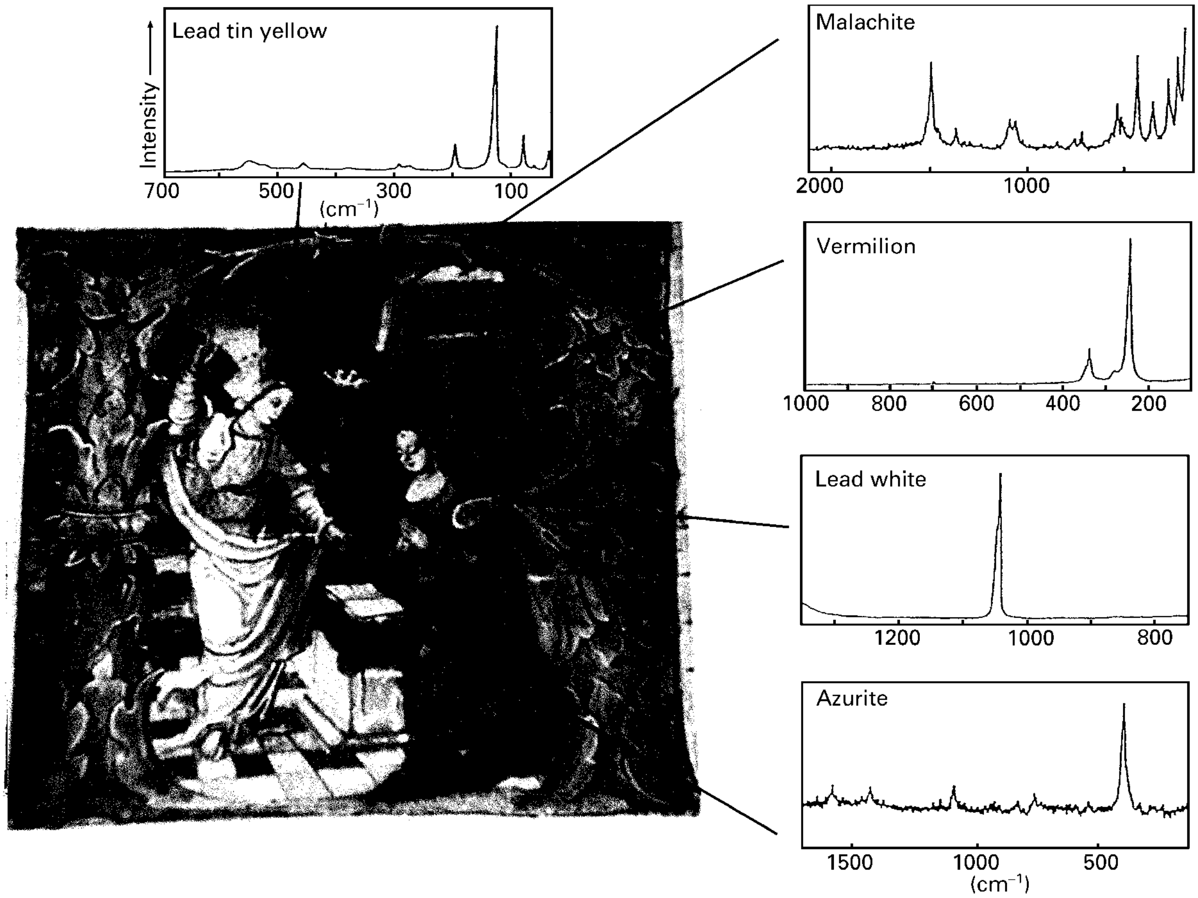



**Figure 11** Elaborately historiated initial 'R' in sixteenth-century German choir book, with Raman microscopy spectra of selected pigmented regions. Reproduced with permission of Clark RJH (1995) Raman microscopy: application to the identification of pigments on medieval manuscripts. *Chemical Society Reviews* 187–196 © The Royal Society of Chemistry. (See Colour Plate 5a).

Raman microscopy: vermilion, orpiment, realgar, red ochre, azurite and bone white. Neither red lead nor lead white, pigments commonly used in Northern Europe, was identified on the manuscript: this raises questions as to the availability of these pigments in Iceland, to which they are not native, and of the effectiveness of the trading routes. Examples of two historiated initials and the Raman spectra of the pigments used are shown in **Figure 10**.

An elaborately historiated initial 'R' on a sixteenth-century German choir book has been studied by Raman microscopy and no fewer than eight pigments have been identified, namely, azurite, lead tin yellow, malachite, vermilion, white lead, red lead, carbon and massicot (**Figure 11**).

Raman microscopy demonstrated that in this case the two shades of blue on the garments of the right-hand figure, although appearing to be due to different pigments, arise from the same pigment, azurite, the illuminator having used less azurite and proportionately more binder to produce the lighter shade without choosing to add any white pigment, such as lead white, to gain the same effect. The azurite used in the deeper blue robe arises from coarse grains of pigment (~30 μm diameter) whereas the lighter blue used in the undergarment arises from fine grains (~3 μm diameter). The effect of the particle size on the depth of colour of a powder is a common feature of powdered materials, whose colour is determined by diffuse reflectance. As the particle size is reduced, the average depth to which radiation penetrates before being scattered is also reduced and hence the depth of colour is reduced.

The angel's wing and podium are painted in malachite, which is a basic copper carbonate similar to azurite. Both azurite and malachite have a similar provenance in that they are both associated with secondary copper ore deposits, but their spectra are very different.

It is of particular interest that the dark grey colour of the pillar top is not obtained via a single pigment but by colour subtraction of a mixture of at least

seven pigments: white lead, carbon, azurite, as well as small amounts of vermilion, red lead, massicot and lead tin yellow type I. The mixture is evident at ×100 magnification by Raman microscopy, whereby each individual pigment grain may be identified (**Figure 12**).

## Wall paintings

The application of FT-Raman spectroscopy to the study of red pigments from English early medieval wall paintings in the Chapter House of Sherborne Abbey and the Holy Sepulchre Chapel in Winchester Cathedral (**Figure 13**) has been reported. The complexity of the spectra can be observed in **Figure 14**, where the pigment and substratal features are clearly differentiated. The operations of the two monastic houses were very different, as the Sherborne Abbey pigment consisted of pure vermilion, whereas that of Winchester Abbey was a 3:1 mixture of red ochre and vermilion. In both cases, spectral features due to sandstone and marble could be seen, but no identifying feature could be ascribed to a plaster substrate.

The biodeterioration of the Renaissance frescoes in the Palazzo Farnese, Caprarola, Italy, arising from aggressive colonization by the lichen *Dirina massiliensis* forma *sorediata* has been studied using Raman microscopy. Over 80% of the masterpieces painted by Zuccari in 1560 have now been destroyed. The role of the lichen metabolic products and the ability

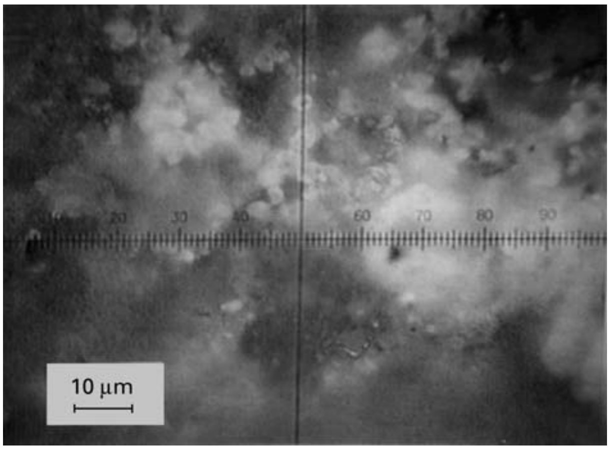



**Figure 12** Portion of top of column of historiated initial 'R' shown in **Figure 11**; the individual pigment grains can be clearly seen under the ×100 magnification and can be separately identified using Raman microscopy. Reproduced with permission of Clark RJH (1995) Raman microscopy: application to the identification of pigments on medieval manuscripts. *Chemical Society Reviews* 187–196 © The Royal Society of Chemistry. (See Colour Plate 5b).

**Figure 13** Holy Sepulchre Chapel, Winchester Cathedral. Wall painting of ca. 1175–85 on the east well depicting the *Deposition, Entombment, Maries at the Sepulchre and the Harrowing of Hell*. Reproduced with permission from Edwards HGM, Brooke C and Tait JKF, An FT-Raman spectroscopic study of pigments or medieval English wall paintings, *Journal of Raman Spectroscopy* **28**: 95–98, © 1997 John Wiley and Sons Ltd. (See Colour Plate 6).

of primitive organisms to survive hostile environments generated by large concentrations of heavy metals such as mercury, lead and antimony are now being understood as a result of these studies, which will assist conservators in future projects.

The red pigment on the Carlovingian frescoes in the crypt of the Abbey of St Germain at Auxerre shows up to 13 layers of paint identified by Raman microscopy, and vermilion and iron(II) oxide have both been found.

The North American palaeo-Indian shelters at Seminole Canyon, at the confluence of the Rio Grande and Devils Rivers, have provided the Raman spectra of the oldest cave paintings yet studied. From a limited palette, these 3500-year-old pictographs have yielded red, white and black colours identified

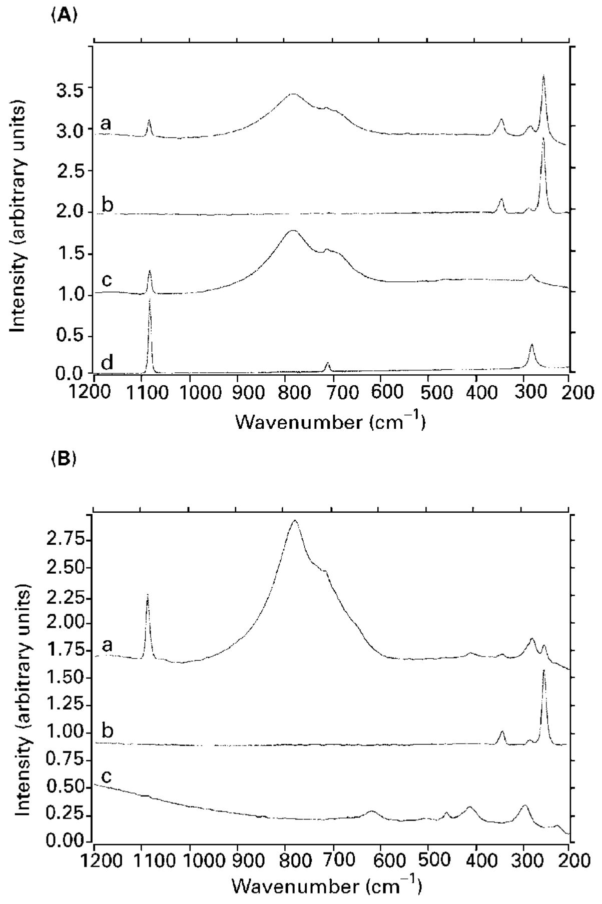



**Figure 14** FT-Raman spectra of twelfth-century wall paintings in (A) Sherborne Abbey ((a) red pigment, (b) vermilion, (c) unpainted stone, (d) calcite) and (B) Winchester Cathedral ((a) red pigment, (b) vermilion, (c) red ochre). Reproduced with permission from Edwards HGM, Brooke C and Tait JKF, An FT-Raman spectroscopic study of pigments on medieval English wall paintings, *Journal of Raman Spectroscopy* **28**: 95–98, © 1997 John Wiley and Sons Ltd.

as red ochre, calcium oxalate monohydrate (whewellite) and manganese(IV) oxide, respectively. In addition, the black pigment showed traces of an organic additive recently identified from DNA profiling as bison or deer bone marrow, presumably added in a symbolic ritual.

FT-IR spectroscopy, too, has had considerable success in the identification of pigments and associated materials in paintings and manuscripts. In particular, the use of infrared microspectroscopy has proved to be an essential prerequisite for the scientific examination of complex pigment mixtures. Generally, however, inorganic mineral pigments are not well characterized in the infrared because of their low band wavenumbers; hydroxyl groups in hydrated minerals cause problems in absorption and the substrates themselves, e.g. cellulose, linen or starch, often give rise to broad backgrounds. IR spectroscopy, therefore, is much more applicable to organic pigments, binders and mixtures.

One project dealing with medieval manuscripts has had as its objective the application of small-particle-analysis techniques to the study of pigments in medieval Armenian and Byzantine manuscripts. At the University of Chicago, 10 decorated manuscripts were sampled (**Table 10**). These manuscripts represent a broad chronological span ranging from the tenth century to the post-Byzantine era (sixteenth century or later).

Interpretation of IR spectra was complicated by the fact that the pigment samples often contained several components in addition to the pigment. These additives often served as binding media – thickening agents or extenders, for example. Because the amount of pigment is small relative to the amounts of these nonpigment components, the spectral bands of the pigment are often difficult to distinguish from interfering spectral bands of non-pigment components. Chromatographic separation prior to spectroscopic analysis was not feasible because of the insolubility and limited quantity of the samples. Spectral subtraction has been used in many cases as a means to 'separate' the components.

The infrared spectrum of a yellow pigment sample from manuscript (MS) 46 (Haskell gospels) shows several bands that can be attributed to egg yolk, a typical medium used in mediaeval times, as well as calcium carbonate and kaolin. The absorption bands at 3382, 3298 and 1600 $cm^{-1}$ suggested that a primary amine is present. Putrescine (1,4-diaminobutane), an amine found in decaying proteinaceous matter, has similar spectral features. Because parchment (being of animal skin origin) would be subject to putrefaction if not properly preserved, this amine could be a product of bacteriological degradation.

**Table 10** University of Chicago Special Collections manuscripts analysed by FT-IR spectroscopy

| *Manuscript number* | *Name* |
|---|---|
| 972 | Archaic Mark |
| 1054 | Elfleda Bond Goodspeed *Gospels* |
| 965 | Rockefeller-McCormick *New Testament* |
| 131 | Chrysanthus gospels |
| 232 | Greek gospels |
| 46 | Haskell gospels |
| 129 | Nicolaus gospels |
| 727 | Georgius gospels |
| 879 | Lectionary of Constantine the reader |
| 948 | Lectionary of St Menas the wonder worker |

Although the pigment was not identified in this case, the information that egg tempera is present as the binding medium is valuable; the binding medium in MS 965 is hide glue.

Pigment components have also been identified from their IR spectra. The infrared spectrum obtained from a purple pigment removed from MS 965 showed bands due to ultramarine blue and white lead. A red pigment (cochineal?) must also be present with the blue to produce the purple colour.

In MS 972 (Archaic Mark), the presence of Prussian blue from the infrared spectrum indicates possible restoration because Prussian blue, a synthetic dye, was not available until 500 years later.

Another example of the use of FT-IR spectroscopy in the authentication of paintings is provided by the analysis of the *Virgin and Child*, a fifteenth-century Italian painting on panel. There were already stylistic doubts and, although IR spectroscopy confirmed the presence of egg tempera as a binding medium, the blue pigment was not azurite or lapis lazuli. A prominent infrared absorption at 2091 $cm^{-1}$ identified the blue pigment on the Virgin's robe as Prussian blue, unknown before 1704. Hence, the painting has either been very heavily reworked or is possibly a nineteenth-century copy of an earlier work.

*See also:* **Colorimetry, Theory; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy; Fourier Transformation and Sampling Theory; Rayleigh Scattering and Raman Spectroscopy, Theory; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading

Adelantado JVG, Carbo MTD, Martinez VP and Reig BF (1996) FTIR spectroscopy and the analytical study of works of art for purposes of diagnosis and conservation. *Analytica Chimica Acta* **330**: 207–215.

Best SP, Clark RJH and Withnall R (1992) Non-destructive pigment analysis of artefacts by Raman microscopy. *Endeavour* **16**: 66–73.

Clark RJH (1995) Raman microscopy: application to the identification of pigments on mediaeval manuscripts. *Chemical Society Reviews* 187–196.

Coupry C and Brissaud D (1996) Applications in art, jewellery and forensic science. In Corset J and Turrel G. (eds) *Raman Microscopy: Developments and Applications*, pp 421–453. London: Academic Press.

Edwards HGM and Farwell DW (1995) FT-Raman spectroscopic study of ivory: a nondestructive diagnostic technique. *Spectrochimica Acta Part A* **51**: 2073–2079.

Hoepfner G, Newton T, Peters DC and Shearer JC (1983) FTIR in the service of art conservation. *Analytical Chemistry* **55**: 874A–880A.

Katon JE, Lang PI, Mathews TF, Nelson RS and Orna MV (1989) Applications of IR microspectroscopy to art historical questions about mediaeval manuscripts. *Advances in Chemistry Series* **220**: 265–288.

Learner T (1996) The use of FTIR in the conservation of 12th century paintings. *Spectroscopy Europe* 14–19.

Mills JS and White R (1994) *The Organic Chemistry of Museum Objects*, 2nd edn. London: Butterworth-Heinemann.

# Astronomy, Applications of Spectroscopy

*See* **Interstellar Molecules, Spectroscopy of; Stars, Spectroscopy of.**

# Atmospheric Pressure Ionization in Mass Spectrometry

**WMA Niessen**, hyphen MassSpec Consultancy, Leiden, The Netherlands

Copyright © 1999 Academic Press

MASS SPECTROMETRY
Methods & Instrumentation

## Introduction

Ion sources for atmospheric-pressure ionization (API) in mass spectrometry (MS) were first described in 1958 by Knewstubb and Sugden. The work of the research group of Horning and Carroll in the late 1960s and early 1970s resulted in a system commercially available from Franklin GNO Corp. These types of instruments were mainly used in the study of ion–molecule reactions in gases in the atmosphere. Furthermore, in 1974 this type of instrument was applied with an atmospheric-pressure corona discharge ion source in the coupling of liquid chromatography to MS (LC-MS). Despite the promising results, the technique did not attract much attention at that time.

Subsequently, several other API instruments were described and built. The most successful commercial instrument was the Sciex TAGA (trace atmospheric gas analyser), while API systems were also available from Extranuclear and Hitachi. In the late 1980s, the TAGA API-MS system was reengineered by Sciex for online LC-MS and was called the API-III. This system was applied by a growing group of scientists, especially within pharmaceutical industries. In the early 1980s, Yamashita and Fenn fundamentally investigated the potential of electrospray ionization in an API source. This research led to another approach to LC-MS using an API system. In 1988, Fenn and co-workers demonstrated that with this electrospray system it was possible to perform the MS analysis of high molecular-mass proteins via multiple charging. As a result of this observation, API-MS went through a period of tremendous growth, the end of which is certainly not yet here. In the mid 1980s, similar API technology was also applied in the coupling of inductively coupled plasmas (ICP) to MS.

At present, there are three major application areas of API-MS:

- air or gas analysis,
- online LC-MS, and
- ICP-MS.

In this paper, technical and instrumental aspects of API-MS are discussed and a number of typical applications are briefly reviewed.

## API ion source design

An API ion source generally consists of four parts:

1. the sample introduction device, the design of which is highly dependent on the type of application,
2. the actual ion source region, i.e. the region where the ions are generated,
3. an ion sampling aperture, where the ions are sampled from atmospheric pressure into the vacuum system of the MS, and
4. an ion transfer system, where the pressure difference between the atmospheric-pressure source and the high-vacuum mass analyser is bridged and the ions entering through the ion sampling aperture are preferentially transferred and focused into the mass analyser region.

The proper design of parts 3 and 4 is of utmost importance, as it determines to what extent the very high ionization efficiency that can be achieved in an API source actually becomes available to the analytical applications. The inevitable ion losses in the sampling and transfer regions should be kept to a minimum.

## Ion sampling aperture and ion transfer system

An API source system consists of a region where the ions are generated. This region can have a relatively large volume of several litres, although in practice only the ions that come close to the ion sampling aperture are efficiently sampled into the MS. Therefore, in many systems the ions are generated just in front of the ion sampling aperture. A general scheme of an API source is shown in **Figure 1**.

In the early API systems, the ion sampling aperture consisted of a 25–200 μm i.d. pinhole in a metal plate. The size of the pinhole is limited by the pumping capacity of the vacuum system of the MS.

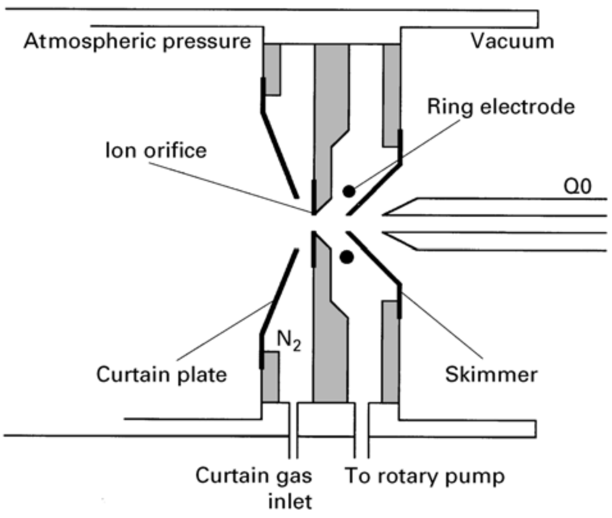



**Figure 1** Schematic diagram of a typical API source, as for instance used in LC-MS.

In the first prototypes, the pinhole inner diameter was only 25 μm, while in the Sciex TAGA a larger pinhole could be applied, because of the high pumping efficiency of the cryogenic vacuum-pumping system available in this instrument.

The sampling of ions through an orifice in a plate or in the tip of a cone is used in many commercial API systems. However, two alternative ion sampling devices based on the transport of ions through capillary tubes have been developed and commercialized. The first, designed by Fenn and co-workers and subsequently commercialized by Analytica of Branford, incorporates a 100–150 mm × 0.5 mm ID glass capillary. Both ends of the glass capillary are coated with metal, e.g. silver, gold or platinum, in order to define the potential at both tips. The use of an insulating glass capillary enables independent control over the potentials applied in the atmospheric-pressure ion source region and in the ion acceleration region inside the vacuum. The second alternative to the ion sampling orifice is a heated stainless steel capillary, initially proposed by the group of Chait and subsequently commercialized by Finnigan.

In the strong cooling that takes place upon the expansion of a gas into a vacuum chamber, as is taking place in an electrospray source, the ions act as condensation nuclei for water and solvent molecules. As a result, clusters of analyte ions and numerous solvent molecules are formed. This cluster formation can be counteracted or avoided in two ways: either by preventing the solvent ions from reaching the ion sampling orifice, or by increasing the temperature of the gas, vapour and ion mixture entering the vacuum. By flushing the area just in front of the sampling orifice with a stream of dry nitrogen, the solvent vapour can be swept away, while the ions are forced to pass through this nitrogen 'curtain' by the application of an electric field between the curtain plate and the orifice plate in **Figure 1**. An additional benefit of a curtain gas is the prevention of other neutral contaminants, such as particulate material, from entering the ion sampling orifice or capillary. This is important for the ruggedness required for day-to-day use of LC-MS, e.g. in the analysis of large series of biological samples. A curtain or countercurrent gas is applied in the sources built by Sciex, Analytica of Branford and Hewlett-Packard. Alternatively, the temperature of (part of) the ion source or of the ion sampling capillary can be increased to avoid the formation of clusters between analyte ions and solvent molecules. This obviates the need for a curtain gas, but also leaves the ion sampling device unprotected from contamination by particulate material.

Ions are sampled from the atmospheric-pressure region and together with nitrogen, which is applied as nebulizing gas as well as curtain gas, entering the vacuum system. The low-pressure side of the ion sampling pinhole or the low-pressure end of the ion sampling capillary acts as a nozzle, where the mixture of gas, solvent vapour and ions expands. The expanding jet is subsequently sampled by means of a skimmer into a second vacuum stage. In most systems, the region between the nozzle and the skimmer, which is pumped by a high-capacity rotary pump, contains an additional lens, e.g. a ring or tube electrode, to preferentially have the ions sampled by the skimmer. From a gas dynamics point of view, it might be important to position the skimmer just in front of the Mach disk of the expansion. However, because ions are preferentially sampled, the positioning of the skimmer relative to the Mach disk does not appear to be very important. In some systems, nozzle and skimmer are actually not precisely aligned in order to reduce the number of neutral molecules in the highly directed flow from the nozzle entering the region with higher vacuum. Optimum ion transfer is achieved by means of a tube lens between nozzle and skimmer.

At the lower pressure side of the skimmer a second expansion step takes place. In older systems, the ions entering through the skimmer were extracted, transferred and focused into the mass analyser region by means of a set of conventional flat lenses. Later, it was determined that a more efficient ion transfer and focusing can be achieved by means of an RF-only quadrupole, hexapole or octapole device (cf. **Figure 1**). These types of ion transfer and focusing devices are especially in use in systems for LC-MS. In other systems, as used for ICP-MS for instance, ion

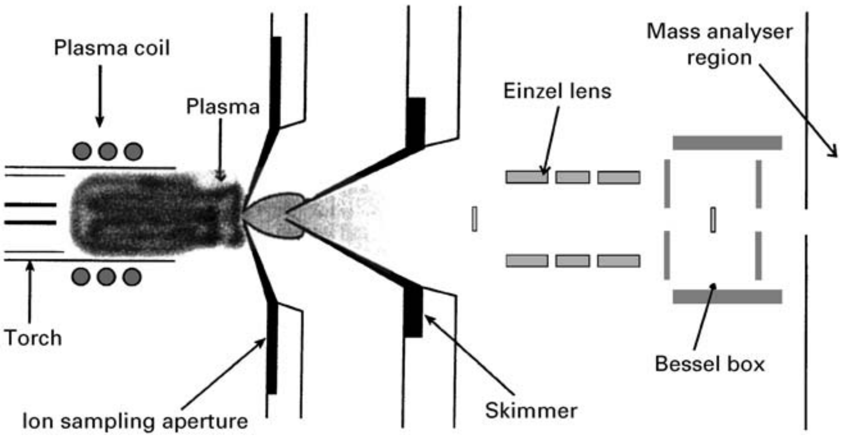



**Figure 2** Schematic diagram of a typical API source for ICP-MS.

focusing is achieved by means of a combination of Einzel lenses and a Bessel box (cf. **Figure 2**).

The source designs described so far are used in combination with a (triple) quadrupole instrument. Modifications of this design are required in the coupling to other types of mass analysers, e.g. the implementation of a gate or ion pulse electrode for use in combination with ion-trap and time-of-flight mass analysers.

## Sample introduction devices

A wide variety of sample introduction devices are available for gas analysis by API-MS. In the study of atmospheric processes, direct sampling of ionic species and clusters in various atmospheric layers may be performed. For these *in situ* studies, air is directly sampled into the system, eventually through specially designed sampling tubes. Alternatively, gases may be sampled and mass analysed after ionization by electrons from a corona discharge (similar to APCI).

For use in combination with LC-MS and other liquid-introduction techniques, two main types of sample introduction devices are available, i.e. one for electrospray ionization and one for APCI.

The devices for APCI consist of a heated nebulizer, which is a combination of a concentric pneumatic nebulizer and a heated quartz tube for droplet evaporation. Additional auxiliary gas is used to flush the evaporating droplets through the vaporizer into the actual ionization region. A schematic diagram of such a device is shown in **Figure 3**. Reactant ions for APCI are generated in a point-to-plane corona discharge in the atmospheric-pressure ion source. The discharge needle is operated at 3–5 kV and provides a corona emission current of a few µA.

For electrospray ionization, initially, 100 µm i.d. stainless-steel capillaries, i.e. hypodermic needles, were used for sample introduction. With such a device, the flow-rate is limited to ~10 µL $min^{-1}$, which is too high for many biochemical applications and too low for effective LC-MS coupling. Therefore, the initial system was modified.

For biochemical applications, the dimensions of the introduction capillary were decreased to a few µm i.d. capillaries, mostly a glass or fused silica. This approach is called nanoelectrospray as it allows the continuous introduction of between 5 and 100 nL$min^{-1}$ of a liquid into the API source. Nano-electrospray is especially useful for applications where the sample amount is limited. In this way, it is possible to perform a series of MS experiments with a minute amount of sample, e.g. 30 min of various MS experiments with only 1 µL of sample.

For LC-MS applications, the electrospray nebulization process was assisted by pneumatic nebulization, i.e. a nebulization gas at a high linear velocity is forced around the electrospray needle. This approach, introduced by Sciex as ionspray in 1986, was subsequently adopted by other instrument manufacturers and is now the most widely applied

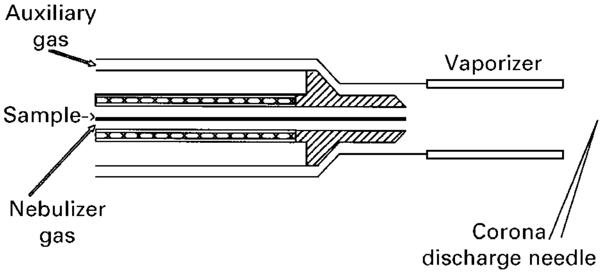



**Figure 3** Schematic diagram of a sample introduction device for APCI: a heated nebulizer.

approach to LC-MS via electrospray interfacing. A schematic diagram of a typical sample introduction device for LC-MS via pneumatically-assisted electrospray is shown in **Figure 4**.

Significant research has taken place in the past few years concerning the optimum geometry of the spray device in the atmospheric-pressure ion source. While initially the spray device was positioned axially relative to the ion sampling orifice, an approach which is still in use, other instrument manufacturers use a different setup, e.g. the spray device positioned orthogonally to the ion sampling orifice, or at an angle of 45°, eventually with a perpendicular heated gas probe to assist in droplet evaporation. The two main perspectives in this type of research are the ability to introduce higher liquid flow-rate while still avoiding the risk of clogging the ion sampling orifice, and the ability to introduce samples containing significant amounts of nonvolatile material, e.g. samples from biological fluids after minimum sample pretreatment, or LC mobile phases containing phosphate buffers. While the advantages of the alternative positioning of the spray device is well documented for biological fluids, the evidence for the possibilities of the use of phosphate buffers is less convincingly demonstrated.

For use in ICP-MS, ions are sampled directly from the inductively coupled plasma. Special precautions are required related to the high temperatures in the plasma. A schematic diagram of an ICP-MS system is shown in **Figure 2**. In most cases, the liquid sample, which can be introduced directly from a vial, of a liquid-phase separation technique such as LC is nebulized into a separate spray chamber. The aerosol, containing the smaller droplets, is transported by argon to the torch, where the ICP is generated and sustained. In this so-called ICP flame, the analytes are atomized and ionized. The ions are directly sampled from the flame by the ion sampling aperture for mass analysis.

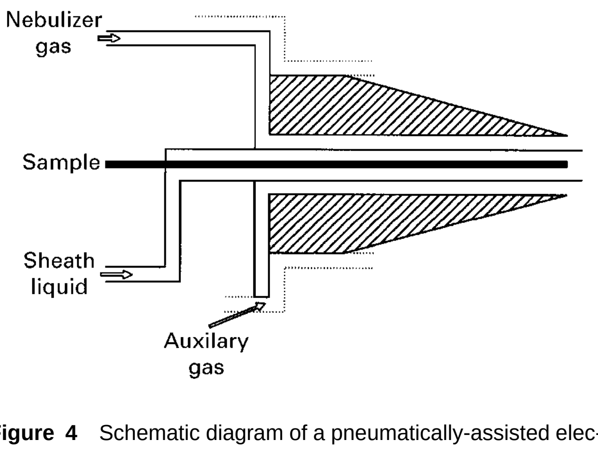



**Figure 4** Schematic diagram of a pneumatically-assisted electrospray sample introduction device.

## Ionization in API conditions

Ionization in an API source can take place in a variety of ways, depending on the type of applications. The various way are briefly reviewed here.

### Gas-phase ionization

In initial applications of API sources for gas analysis and monitoring of atmospheric ionic processes, either direct sampling of the ions present in flames or in the atmosphere was applied, or ions were generated by means of electrons from an electron-emitting $^{63}Ni$ foil or by means of a corona discharge. Currently, the corona discharge is the most widely applied approach for generating ions, in both gas analysis and APCI in LC-MS applications.

In an API source equipped with a $^{63}Ni$ foil, energetic electrons ionize the nitrogen present in a series of consecutive steps:

$$N_2 + e^- \rightarrow N_2^{\bullet+} + 2\,e^-$$

$$N_2^{\bullet+} + 2N_2 \rightarrow N_1^{\bullet+} + N_2$$

Similarly, other gases may be ionized. In the presence of traces of water in the ion source, these ionic species react with the water molecules:

$$N_4^{\bullet+} + H_2O \rightarrow H_2O^{\bullet+} + 2N_2$$

$$H_2O^{\bullet+} + H_2O \rightarrow H_3O^+ + HO^{\bullet+}$$

$$H_3O^+ nH_2O \rightarrow H_3O^+(H_2O)_n$$

The main ionic species generated, i.e. $N_4^{\bullet+}$ and $H_3O^+.(H_2O)_n$, may enter in ion–molecule reactions, leading to charge exchange of proton transfer reactions, respectively. This is valid for positive-ion acquisition, while negative ions may be generated either by electron-capture processes with electrons or via proton-transfer ion-molecule reactions starting from $OH^-.(H_2O)_n$ for instance.

The same reactant ions are formed in positive-ion mode in a point-to-plane corona discharge. Approximately 3–9 kV is applied at the corona discharge, creating a corona current of 1–5 μA, depending on the point-to-plane distance and the composition of the vapours in the ion source. Corona discharges are very complicated processes that involve avalanches of ionization reactions by ion–molecule and electron–molecule collisions together with quenching

processes, as discussed in detail by Meek and Craggs.

In the positive-ion mode, the ionization of analyte molecules follows the general rules of chemical ionization, i.e. proton transfer may take place when the proton affinity of the analyte molecule exceeds that of the reagent gas, leading to protonated molecules:

$$M + H_3O^+ \cdot (H_2O)_n \rightarrow [M + H]^+ + nH_2O$$

while, in addition, various adduct ions and cluster ions with solvent molecules may be generated. In the negative-ion mode, either electron-capture products $M^{-\bullet}$ or deprotonated molecules $[M–H]^-$ may be generated, although the electron-capture products may further react in ion–molecule reactions to form different ionic species.

The positive-ion mass spectra obtained from APCI and conventional medium-pressure CI might show differences due to the fact that the mass spectrum in medium-pressure CI is more determined by the reaction kinetics, while in APCI, due to the longer residences times of ions in the sources as well as the more frequent ion–molecule collisions, the mass spectrum more reflects the thermodynamic equilibrium conditions.

### Liquid-based ionization

The liquid-based ionization technique in API-MS is applicable in electrospray interfacing as well as with a heated nebulizer introduction under certain conditions, i.e. without switching the corona discharge electrode on, and for certain compounds. Here, the discussion is restricted to electrospray interfacing.

In electrospray ionization and interfacing, a high potential, typically 3 kV, is applied to a liquid emerging from a capillary. This causes the liquid to break into fine threads emerging from the so-called Taylor cone formed at the capillary tip. The threads subsequently disintegrate into small droplets. The electrohydrodynamic or Rayleigh disintegration results from autorepulsion of the electrostatically charged surface which exceeds the cohesive forces of surface tension (the Rayleigh limit). In this electrospray nebulization process uniform droplets in the 1-μm diameter range are formed. These charged droplets shrink by solvent evaporation, whereas electrohydrodynamic droplet disintegration again takes place as soon as the Coulomb repulsion of the surface charges exceeds the surface tension forces. The offspring droplets generated in the field-induced droplet disintegration process carry only 2% of the mass of the parent droplet and 15% of its charge. Analyte ions in these offspring droplets, present as preformed ions, desorb or evaporate into the gas phase and become available for mass analysis. The ion evaporation process, described first by Iribarne and Thomson and generally considered as the most important ionization mechanism in electrospray ionization, is a process which thermodynamically corresponds to bringing a solvated ion from the surface of a charged droplet to infinity. Below a certain size/charge ratio for a droplet, the Gibbs free energy of solvation of an ionic species present in the droplet will exceed the energy required to bring this species as a solvated ion from infinity to the surface of the droplet. Under these conditions, ion evaporation is possible. According to Iribarne and Thomson, ion evaporation will take place instead of electrohydrodynamic disintegration when the critical ion evaporation droplet radius exceeds the Rayleigh limit.

Electrospray ionization is especially effective for analytes that are present as preformed ions in solution, e.g. organic acids or bases that are present as deprotonated or protonated molecules, respectively, at suitable pH values, quaternary ammonium compounds, etc. For biomacromolecules like proteins, which exist in solution as a distribution of ionic species carrying different numbers of charge, electrospray ionization results in a spectrum containing an ion envelope of multiply charged ions (see **Figure 5**). To a first approximation, this mass spectrum reflects the charge-state distribution of the protein in solution. As the ion separation in a mass analyser is based on mass-to-charge ratio, the multiple charging allows the mass analysis of large molecules within the limited *m/z*-range of, for instance, a quadrupole mass spectrometer.

The observation of multiple charging of proteins in electrospray ionization attracted much attention: all major instrument manufacturers introduced an API system to be used in combination with an electrospray interface. At the same time, it was found that electrospray ionization is not only suitable for the mass analysis of large molecules, but is also a very efficient ionization technique of small polar or ionic molecules, e.g. drugs and their metabolites. In the past few years, hundreds of dedicated API-MS systems equipped with electrospray and APCI interfaces have found their way into many different laboratories, especially within pharmaceutical companies and biochemistry/biotechnology laboratories.

### Plasma-based ionization

A plasma is an ionized gas that is macroscopically neutral, i.e. an equal number of positive and negative particles are present. In ICP-MS, the gas used to

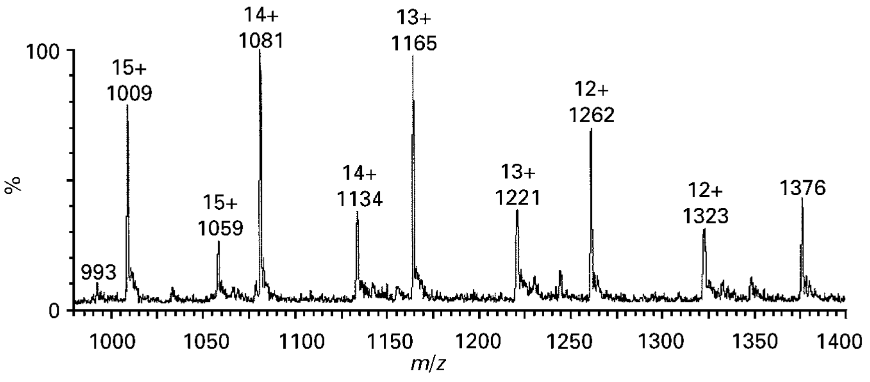



**Figure 5** Ion envelope of multiply charged ions, obtained in the electrospray ionization of haemoglobin. In the spectrum, the *m/z* and charge state of the ions are indicated. The spectrum consists of two series: one due to a Hb-α chain ($M_r$ 15126 Da) and one due to a Hb-$\beta^A$ chain ($M_r$ 15 867 Da).

generate the plasma is argon, which has a high ionization energy (15.76 eV). The hot argon plasma has the capability to atomize most samples, and excite and ionize most elements of the periodic table. The plasma is generated in a torch. A high-frequency generator, typically operated at 27 or 40 MHz and with a power of 1–2 kW, is used to produce a high-frequency field through an induction coil, which is positioned at the outside of the torch. In this way, the argon plasma is produced, i.e. the argon is ionized and the plasma is sustained. Liquid samples are nebulized and the aerosol is injected into the plasma for atomization and ionization. The ionized atoms generated in this way are sampled by the ion sampling aperture and mass analysed.

## Applications

### Gas analysis

There are a number of applications of API-MS in gas analysis, e.g. the study of environmental air contamination, human breath analysis, the study of (ionic) processes in the atmosphere, the detection of trace impurities in gases that are important in microelectronic manufacturing processes, or the identification of ions in flames. The power of this technique is perhaps best illustrated by the ability to achieve detection limits in the ppt range for oxygen, water, methane and carbon dioxide in high-purity gases like hydrogen, nitrogen, argon and helium. A proper characterization of the trace impurities in these gases is of utmost importance in semiconductor processing. Using a mobile API-MS system, such as the Sciex TAGA, real-time environmental monitoring of TNT, industrial emissions, PCBs and other environmental contaminants in the ground-level troposphere is possible.

### LC-MS

The use of API-MS for the online LC-MS coupling is the largest commercial field of application of API-MS. From the early 1990s, when the potential of API-MS for LC-MS, and especially its robustness and ‘user-friendliness’, were substantially demonstrated, an astonishingly large number of API-MS systems for LC-MS were purchased. Particularly in various stages of drug development within pharmaceutical industries, the use of API-based LC-MS is preferred over the more conventional LC-UV system. In general, better reliability, confirmation, selectivity and sensitivity are achieved in LC-MS systems. In this way, high-throughput quantitative bioanalysis is possible. In addition, API-based LC-MS systems using electrospray or atmospheric-pressure chemical ionization (APCI) interfacing and ionization are applied for the quantitative and qualitative analysis of a wide variety of polar compounds in numerous matrices in a number of other fields, e.g. pesticides, herbicides and their metabolites and degradation products in environmental samples, natural products in plant extracts and cell cultures, peptides, proteins and other biomacromolecules in biochemical and biotechnological applications.

### ICP-MS

ICP-MS is used in practically every discipline where inorganic analytical support is required. This includes environmental, geological, biological, medical, nuclear, metallurgical (semiconductor industry) and nutritional studies. An important advantage of ICP-MS in quantitative analysis, required in many fields of application, is the substantial gain in precision and accuracy that can be achieved by the use of isotope-dilution mass spectrometry, where stable isotopes or

long-lived radio-isotopes of the elements to be analysed are added as internal standards, e.g. the use of $^{111}Cd$ in the analysis of $^{114}Cd$. In addition, ICP-MS can be extremely useful in speciation. In this respect, there is a growing interest in the use of ICP-MS directly coupled to LC or capillary electrophoresis for speciation and analysis of elements of toxicological interest, like As, Se, Pb and Hg.

*See also:* **Chemical Ionization in Mass Spectrometry; Chromatography-MS, Methods; Cosmochemical Applications Using Mass Spectrometry; Inductively Coupled Plasma Mass Spectrometry, Methods; Inorganic Chemistry, Applications of Mass Spectrometry.**

## Further reading


Bruins AP, Covey TR and Henion JD (1987) Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry. *Analytical Chemistry* **59**: 2642–2646.

Carroll DI, Dzidic I, Horning EC and Stillwell RN (1981) Atmospheric pressure ionization mass spectrometry. *Applied Spectroscopic Review* **17**: 337–406.

Chowdhury SK, Katta V and Chait BT (1990) An electrospray-ionization mass spectrometer with new features. *Rapid Communications of Mass Spectrometry* **4**: 81–87.

Cole RB (ed) (1997) *Electrospray Ionization Mass Spectrometry*. Chichester, UK: Wiley.

Fenn JB, Mann M, Meng CK, Wong SF and Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**: 64–71.

Hieftje GM and Norman LA (1992) Plasma source mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **118/119**: 519–573.

Iribarne JV and Thomson BA (1976) On the evaporation of small ions from charged droplets. *Journal of Chemical Physics* **64**: 2287 and (1979) Field induced ion evaporation from liquid surfaces at atmospheric pressure. *Journal of Chemical Physics* **71**: 4451.

Knewstubb PF and Sugden TM (1958) Mass-spectrometric observation of ions in flames. *Nature* **181**: 474–475.

Knewstubb PF and Sugden TM (1958) Mass-spectrometric observation of ions in hydrocarbon flames. *Nature* **181**: 1261.

Meek JM and Craggs JD (1978) *Electrical Breakdown of Gases*. Chichester: Wiley.

Meng CK, Mann M and Fenn JB (1988) *Proceedings of the 36th ASMS Conference on Mass Spectrometry and Allied Topics*, June 5–10, San Francisco, CA, pp. 771–772.

Niessen WMA (1998) *Liquid Chromatography – Mass Specctrometry*, 2nd edn. New York: Marcel Dekker.

Niessen WMA (1998) Advances in instrumentation in liquid chromatography – mass spectrometry and related liquid-introduction techniques. *Journal of Chromatography A* **794**: 407–435.

Vela NP, Olson LK and Caruso JA, Elemental speciation with plasma mass spectrometry. *Analytical Chemistry* **65**: 585A–597A.

Yamashita M and Fenn JB (1984) Electrospray ion source. Another variation on the free-jet theme. *Journal of Physical Chemistry* **88**: 4451–4459.

Yamashita M and Fenn JB (1984) Negative ion production with the electrospray ion source. *Journal of Physical Chemistry* **88**: 4671–4675.


# Atomic Absorption, Methods and Instrumentation

**Steve J Hill** and **Andy S Fisher**, University of Plymouth, UK

Copyright © 1999 Academic Press

| ATOMIC SPECTROSCOPY |
| --- |
| Methods & Instrumentation |

## Introduction

Atomic absorption spectroscopy has become one of the most frequently used tools in analytical chemistry. This is because for the determination of most metals and metalloids the technique offers sufficient sensitivity for many applications and is relatively interference free. There are two basic atom cells (a means of turning the sample, usually a liquid, into free atoms) used in atomic absorption spectroscopy: (i) the flame and (ii) electrothermal heating of a sample cell. It is generally acknowledged that if sufficient analyte is present in the sample, then it should be determined using a flame technique because this has the added advantages of being rapid (assuming only a few elements need be determined) and, in comparison with alternative techniques, very simple to use. Electrothermal atomic absorption spectroscopy (ETAAS) requires more operator skill and is less rapid, but yields substantially superior limits of

detection when compared with flame atomic absorption spectroscopy (FAAS). This section describes some of the methods and instrumentation that have been developed for both flame and electrothermal techniques of atomic absorption spectroscopy.

## Instrumentation

The basic principle of both FAAS and ETAAS is that sample is introduced into the atom cell, where it is desolvated and then atomized. The analyte atoms so formed then quantitatively absorb light in a way that is proportional to the concentration of the atoms of the analyte in the cell. The light, which is at a specific wavelength, is then isolated from other wavelengths that may be emitted by the atom cell and then detected. Thus, much of the instrumentation used for electrothermal and flame absorption spectroscopy is identical. Both techniques require a similar light source, background correction system, line isolation device (monochromator or polychromator), detector (photomultiplier or charge coupled device) and readout system. Each of these components is discussed below, together with details of the individual atom cells (flame and the electrothermal atomizer) and sample introduction systems.

## Light source

The fundamental requirement of the light source is to provide a narrow line profile with little background. It should also have a stable and reproducible output with sufficient intensity to ensure that a high signal-to-noise ratio is obtained. Two basic types of light source are used for atomic absorption, of which the hollow-cathode lamp (HCL) is the more commonly used. This is a lamp in which the cathode is coated with the analyte metal of interest. Within the lamp, inert filler gas (neon or argon) is ionized by an electric current and these ions are then attracted by the cathode. The inert gas ions bombard the cathode and in so doing excite the metal ions coated on it. It is this excitation of the metal that produces the emission of radiation with wavelengths characteristic of the analyte. Hollow-cathode lamps are available for most metallic elements. A schematic diagram of a hollow-cathode lamp is shown in **Figure 1**.

Electrodeless discharge lamps are used less frequently than the hollow-cathode lamps except for analytes such as arsenic and selenium. These lamps may be excited using either microwave energy (although these tend to be less stable) or radiofrequency energy. The radiofrequency-excited lamps are less intense than the microwave-excited ones, but are still 5–100 times more intense than a standard hollow-cathode lamp. In the electrodeless discharge lamp, a bulb contains the element of interest (or one of its salts) in an argon atmosphere. The radiofrequency energy ionizes the argon and this in turn excites the analyte element, causing it to produce its characteristic spectrum. For analytes such as arsenic and selenium, these lamps give a better signal-to-noise ratio than hollow-cathode lamps and have a longer useful lifetime. A schematic diagram of an electrodeless discharge lamp is shown in **Figure 2**.

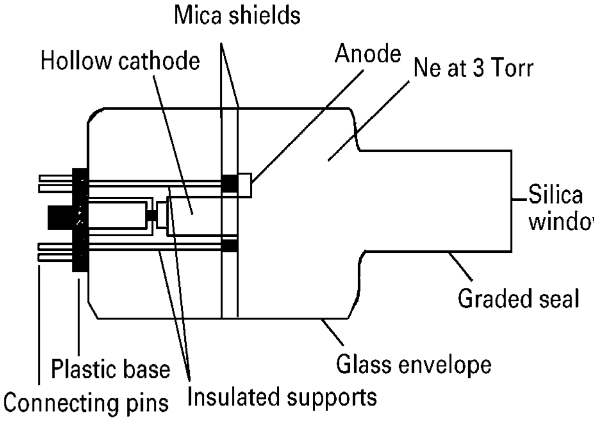



**Figure 1** Schematic representation of a hollow-cathode lamp. From Ebdon L (1998) *Introduction to Analytical Atomic Spectrometry*. Reproduced by permission of John Wiley & Sons Limited.

## Background correction systems

There are basically three types of automatic background correction system available for atomic absorption, although manual methods such as the use of nearby non-atomic absorbing lines to estimate background absorbance may also be used. The three main automatic methods are the deuterium or hydrogen lamp, the Zeeman effect, and Smith–Hieftje background correction. The deuterium lamp produces a continuum of radiation, some of which will be absorbed by molecular species within the atom cell. The amount of atomic absorption observed using the deuterium lamp is negligible and hence the atomic absorption signal is obtained by subtracting the absorbance from the continuum lamp from the total analyte absorbance from the hollow cathode lamp. The deuterium or hydrogen lamp is of most value at wavelengths in the UV region (< 350 nm). The Zeeman effect background correction system is more versatile. It relies on a strong magnetic field operating at approximately

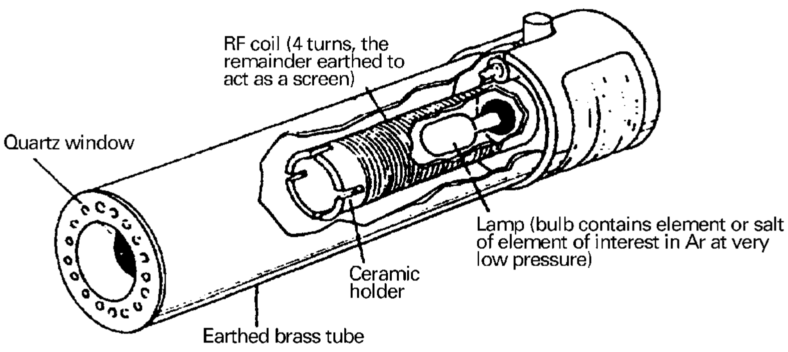



**Figure 2** Electrodeless discharge lamp. From Ebdon L (1998) *Introduction to Analytical Atomic Spectrometry.* Reproduced by permission of John Wiley & Sons Limited.

1 T and 50–60 Hz, placed either around the light source or, more commonly, around the atom cell to split the signal into a number of components. The $\pi$ component is at the normal analyte wavelength; the $\sigma$ components are typically 0.01 nm (depending on the strength of the magnetic field) either side of the $\pi$ component and hence lie outside the atomic absorption profile. Background can be corrected by subtracting the absorbance with the magnetic field 'on', from the signal with the magnet 'off'. It should be noted that this is a simplified description of the process, as different elements have different splitting patterns and the magnetic field may be applied longitudinally or transversely, which may also produce different splitting patterns.

When utilizing the Smith–Hieftje system, the hollow-cathode lamp is boosted periodically to a much higher current, causing the lamp to 'self-absorb'. In this state no atomic absorption occurs in the atom cell but the molecular absorption still remains. Background correction is achieved automatically by subtracting the signal obtained at high current from that obtained using the normal current. The process is particularly efficient at removing interferences such as that caused by phosphate or selenium determinations, although the high currents used may shorten the lifetime of the lamp. In modern instruments, modulation of the source is achieved electronically. This enables discrimination between absorption and emission signals. Previously, a rotating sector, often referred to as a 'chopper', placed between the source and the atom cell was used.

## Single-beam and double-beam instruments

The vast majority of instruments used for atomic absorption measurements have a single-beam configuration, using the optical layout shown in **Figure 3A**. The double-beam arrangement (**Figure 3B**) is far more complex and offers far fewer advantages for atomic absorption when compared with double-beam spectrometers in molecular spectroscopy. This is because the reference beam does not pass through the atom cell. Despite this, double-beam instruments can compensate for source drift and for warm-up and source noise.

## Line isolation devices

To ensure that only light of a wavelength specific to the analyte of interest is being measured, a line isolation device is required. Until recently, the line isolation device used for atomic absorption was a

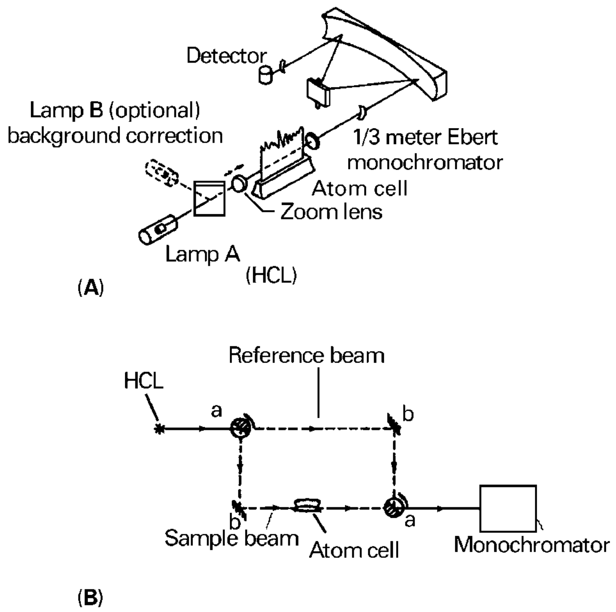



**Figure 3** Schematics of (A) single-beam and (B) double-beam spectrometers. From Ebdon L (1998) *Introduction to Analytical Atomic Spectrometry.* Reproduced by permission of John Wiley & Sons Limited.

monochromator. There are numerous types of monochromator, but modern instruments typically use one of three designs: Ebert, Czerny–Turner or Littrow configuration. These are shown schematically in **Figure 4**.

Light of all wavelengths enters the monochromator through an entrance slit and is then split into specific wavelengths using either a prism, or more commonly, a diffraction grating. By altering the position of this dispersing element, light of only the desired wavelength passes through the exit slit to the detector. Since the method of atomic absorption is so specific, very highly resolving monochromators are not required. Thus, the focal length of an atomic absorption monochromator is often 0.25 or 0.5 m compared with a minimum of 0.75 m required for conventional optical emission spectroscopy.

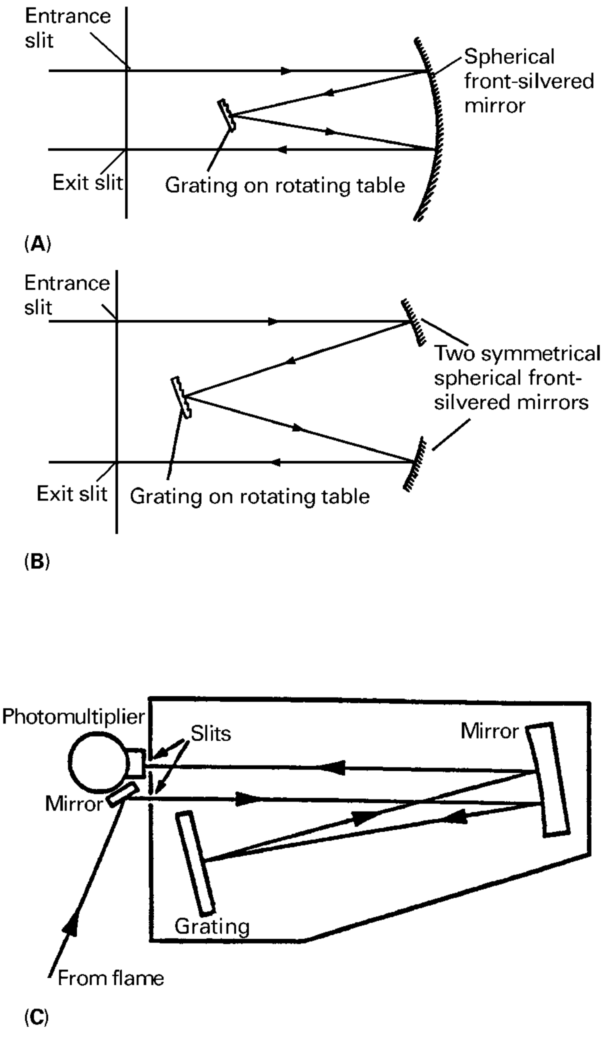



**Figure 4** Schematics of different monochromator types: (A) Ebert, (B) Czerny–Turner, (C) Littrow. From Ebdon L (1998) *Introduction to Analytical Atomic Spectrometry*. Reproduced by permission of John Wiley & Sons Limited.

A monochromator enables only one wavelength to be interrogated at any instant; this was something of a weakness of atomic absorption spectroscopy. New technology, however, has enabled the development of multielement spectrometers that use a bank of between 4 and 6 hollow-cathode lamps and an echelle-style polychromator employing orders of 100 or more and are thus capable of considerable dispersion. In addition, such instruments have no exit slit and so a huge number of wavelengths may be focused onto an array detector such as a charge-coupled device. A schematic diagram of such an instrument is shown in **Figure 5**. Research continues into producing a continuum light source that would enable 30–40 analytes to be determined simultaneously.

## Detection systems

Traditionally, detection of the light isolated by the monochromator has been accomplished using a photomultiplier tube (PMT). Several configurations exist, e.g. end-on and side-on, and the construction may be of different materials, to increase the efficiency at different wavelengths; but basically they all work in a similar way. Light enters the multiplier through a quartz window and impacts with a photocathode that is usually made from one of a number of alloys (e.g. Cs–Sb, Na–K–Sb–Cs or Ga–As), which then emits electrons. These electrons are then accelerated down a series of dynodes, each being at a more positive potential than the previous one. As the electrons impact with successive dynodes, further electrons are ejected and hence a cascade effect occurs. In this way a single photon may cause the ejection of $10^6$ electrons. The number of electrons is

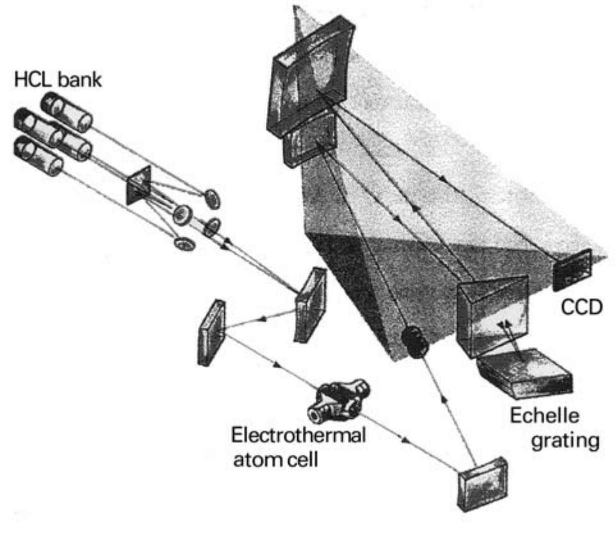



**Figure 5** Echelle-based polychromator.

then measured at the anode and the resulting current is proportional to the radiation reaching the PMT.

As described previously, a number of different materials may be used to coat the photocathode in the PMT, and each has a different response curve. Some, for example the Cs–Sb tube, have very little sensitivity above 600 nm, but others, for example Ga–As, may be used up to almost 900 nm. The response profiles of some commonly used photomultipliers are shown in **Figure 6**.

Conventional spectrometers use a detector that is capable of measuring only one signal. Multielement instruments use state-of-the-art diode arrays or charge-coupled devices that can measure numerous spatially separated signals and can therefore simultaneously determine the signals arising from a bank of hollow-cathode lamps. A charge-coupled device may be considered to be similar to an electronic photographic plate. The device consists of several hundred linear photodetector arrays on a silicon chip with dimensions of typically 13 × 18 mm. The line isolation device (often an echelle grating) separates the analytical wavelengths and these may then be detected on different regions of the array. A detailed description of how the array works is beyond the scope of this text.

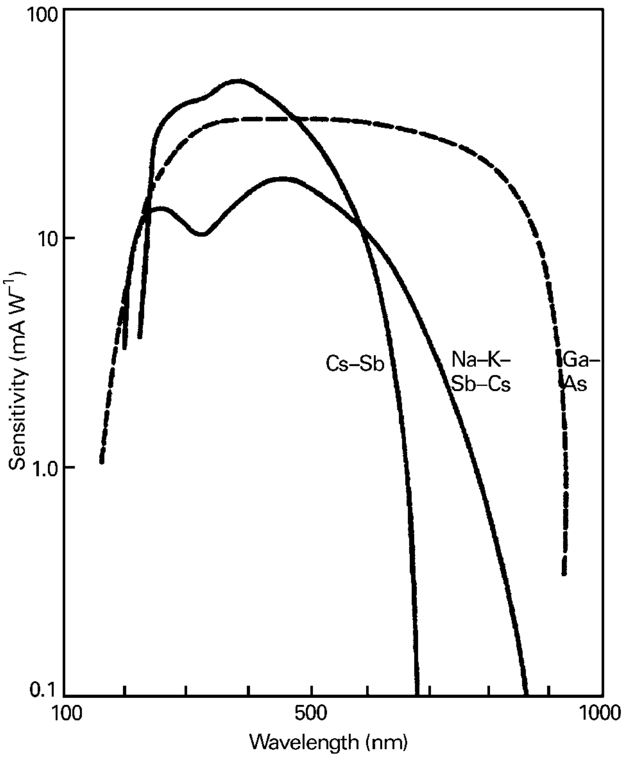



**Figure 6** Response curves for several commercial photomultiplier tubes. From Ebdon L (1998) *Introduction to Analytical Atomic Spectrometry*. Reproduced by permission of John Wiley & Sons Limited.

## Instrument control and output devices

Manual controls and dials have slowly disappeared, with the large majority of modern instruments being controlled by computer. In addition to controlling the instrumental parameters, the computer may be used to program autosamplers, save experimental parameters and data, calibrate with known standards, plot calibration curves, use scale expansion facilities (if necessary) and produce statistical data from the results. A continuous graphics mode enables the signal to be monitored over a period of time. This is especially useful when flow injection, chromatography or hydride-generation transients are to be detected and integrated. Older instrumentation relied on a chart recorder for the same purpose.

## Sample introduction and atom cells

For flame systems, the sample is introduced via a nebulizer and spray chamber assembly. Sample is drawn up the intake capillary by the Venturi effect. The liquid sample column comes into contact with the fast moving flame gases and is shattered into small droplets. In many systems it is further smashed into still smaller droplets by an impact bead. The flame gases then carry the aerosol (the nebular) through a spray chamber containing a series of baffles or flow spoilers. These act as a droplet size filter, passing the larger droplets to waste (85–90% of the sample) and allowing the finer droplets to be transported to the flame. For electrothermal atomizers, the sample is placed into the atom cell either using a hand-held micropipette or by an autosampler. Since most instruments are supplied with an autosampler and the precision associated with their use is superior to that of the micropipette, most laboratories make use of this method of sample introduction.

As described previously, there are two basic types of atom cell; flame and electrothermal tube, but both can have numerous modifications. The flame cell most commonly used is the 10 cm length air–acetylene burner. This provides a flame that is at approximately 2300°C (although the actual temperature will depend on the fuel/air ratio). The flame may be used in a fuel-lean mode, which produces a hot, oxidizing blue flame; in a fuel-rich mode, which produces a cooler, reducing yellow flame; or in a stoichiometric mode whose properties are between the two. Different analytes give different sensitivities depending on the flame type. Chromium, for instance, gives better sensitivity in a reducing flame, whereas for magnesium it is better to use a lean, oxidizing flame.

Different areas of the flame have different temperatures and different chemical properties. The

efficiency of atomization of analytes will therefore depend critically on the area within the flame. Hence, it is important that the light beam from the HCL passes through the region of the flame where atomization is optimal. Optimization of the burner height is therefore necessary to ensure that maximum sensitivity is achieved.

Although an air–acetylene flame is sufficient for the atomization of the majority of analytes, it is not sufficiently hot or reducing to atomize analytes that form refractory oxides (e.g. Al, Mo, Si and Ti). For analytes such as these a nitrous oxide–acetylene flame of length 5 cm is used. This flame is hotter (2900°C) and contains typically 2.5–3 times as much fuel as the air–acetylene flame. Other flame types, e.g. air–propane and air–hydrogen also exist, although they are less commonly used. The air–hydrogen flame (2200°C) is almost invisible and has the advantage of being transparent at lower wavelengths offering improved noise characteristics for elements such as lead or tin at around 220 nm. The inhomogeneity of flame chemistry and the need for burner height optimization hold true for all flame types.

One common modification to a flame cell is to place a quartz tube on the burner head so that the light beam passes through the length of the tube. This method is especially useful when using sample introduction by hydride generation or by gas chromatography. The tube is heated either by a flame or electrically using a wire winding. The analyte, which enters the tube via a hole or slit cut into the side, is atomized within the tube. The atoms then leave the tube but, because of their retention in the tube, spend longer in the light beam than in normal flame systems. An increase in sensitivity is therefore obtained. Another small modification of this system is that, instead of a quartz tube, a quartz T-piece may be used. However, the function is the same.

Electrothermal atom cells have changed radically since their inception in the late 1950s. The majority of electrothermal devices have been based on graphite tubes that are heated electrically (resistively) from either end. Modifications such as the West Rod Atomizer (a carbon filament) were also devised but were later abandoned. Tubes and filaments made from highly refractory metals such as tungsten and tantalum have also been made, but they tend to become brittle and distorted after extended use and have poor resistance to some acids. Their use continues, however, in some laboratories that need to determine carbide-forming elements. For example, silicon reacts with the graphite tube to form silicon carbide, which is both very refractory and very stable. The silicon is therefore not atomized and is lost analytically. Use of a metal vaporizer prevents this.

In electrothermal atomization the sample is introduced into the tube, which is then heated in a series of steps at increasing temperature. The sample is dried at a temperature just above the boiling point of the solvent (but not so hot as to cause frothing and spitting of the sample); ashed (charred at an intermediate temperature to remove as much of the concomitant matrix and potential interferences as possible without losing any analyte); and then atomized at a high temperature. During the atomization stage, the atoms leave the graphite surface and enter the light beam, where they absorb the incident radiation. The ashing and atomizing temperatures used will depend on the analyte of interest; for example, some analytes such as lead are relatively volatile and so cannot be ashed at temperatures above 450°C, otherwise volatile salts such as chlorides will be lost. Other analytes, for example, magnesium, are less volatile and can be ashed at temperatures close to 1000°C without analyte loss. Such elements require a much higher atomization temperature. The sensitivity of electrothermal atomization AAS is greater than for flame AAS because the atoms are formed within the confines of a tube and hence spend longer in the light beam. Also, since the sample is placed within the atom cell, 100% of it is available for analysis compared with the 10–15% available in flame systems. A comparison of characteristic concentrations (concentration that gives an absorbance of 0.0044) obtained for flame and electrothermal techniques for many analytes is shown in **Table 1**. It must be stressed that the figures for ETAAS will depend critically on the injection volume and so values tend to be given in absolute terms (i.e. a weight).

Overall, the Massmann design of electrothermal atomizer in which the tube is heated from either end is still the most common (**Figure 7B**), but more recently transversely heated tubes have been developed (**Figure 7A**). The longitudinally heated tubes have a temperature gradient along the tube, with the central portion being several hundred degrees hotter than the ends. This can lead to condensation of analyte at the cooler ends and subsequent re-atomization from the hot graphite surface. Several atomization peaks may therefore result. The transversely heated tubes do not have a temperature gradient and therefore do not suffer from this problem.

## Interferences

Flame techniques are regarded as being relatively free from interferences, but some distinct classes of interference do exist. These include a few spectral interferences (e.g. Eu at 324.753 nm on Cu at 324.754 nm), ionization interferences and chemical

**Table 1** Comparison of characteristic concentrations for flame and electrothermal AAS

| *Analyte* | *Flame AAS (µg L⁻¹)* | *Electrothermal AAS (pg)* |
|---|---|---|
| Ag | 30 | 5 |
| Al | 300 | 30 |
| As | 800[a],5[b] | 42 |
| Au | 100 | 18 |
| B | 8 500 | 600 |
| Ba | 200 | 15 |
| Be | 16 | 2.5 |
| Bi | 200 | 67 |
| Ca | 13 | 1 |
| Cd | 11 | 1 |
| Co | 50 | 17 |
| Cr | 50 | 7 |
| Cs | 40 | 10 |
| Cu | 40 | 17 |
| Fe | 45 | 12 |
| Ge | 1 300 | 25 |
| Hg | 2 200[a],0.1[b] | 220 |
| K | 10 | 2 |
| La | 48 000 | 7 400 |
| Mg | 3 | 0.4 |
| Mn | 20 | 6 |
| Mo | 280 | 12 |
| Ni | 50 | 20 |
| Pb | 100 | 30 |
| Rb | 30 | 10 |
| Sb | 300 | 60 |
| Se | 350[a],4[b] | 45 |
| Si | 1 500 | 120 |
| Sn | 400 | 100 |
| Sr | 40 | 4 |
| Te | 300 | 50 |
| Ti | 1 400 | 70 |
| Tl | 300 | 50 |
| U | 110 000 | 40 000 |
| V | 750 | 42 |
| Yb | 700 | 3 |
| Zn | 10 | 1 |

[a] Under normal flame conditions.
[b] With vapour generation.

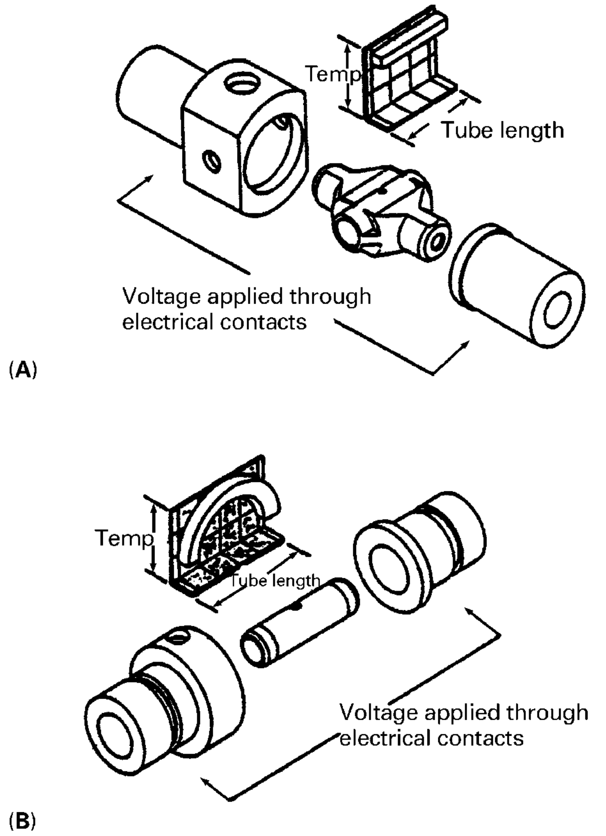



**Figure 7** Electrothermal tubes available commercially: (A) transversely heated graphite atomizer (THGA), (B) longitudinally heated Massmann atomizer. From Ebdon L (1998) *Introduction to Analytical Atomic Spectrometry*. Reproduced by permission of John Wiley & Sons Limited.

interferences. Ionization interference is a vapour-phase interference (in the past often termed cation enhancement) that occurs when the sample contains large amounts of an easily ionized element. The presence of large concentrations of easily ionized elements will lead to a large concentration of electrons in the flame. These electrons prevent the ionization of the analyte and hence lead to higher atomic absorption signals. If the easily ionized element is not present in the standards, the analyte may be partially ionized (and lost analytically) and hence serious overestimates of the true concentration will be obtained. This type of interference may be overcome by adding an excess of easily ionized element to all standards and samples.

Chemical interferences may exist in several different forms:

- Formation of less volatile compounds, e.g. when phosphate is present during the determination of calcium. Calcium phosphate is refractory and hence atomization will be retarded in comparison with calcium in the standards.
- Formation of more volatile compounds, e.g. chlorides.
- Occlusion into refractory compounds. Small amounts of analyte may become trapped in a refractory substance and hence not be atomized efficiently.
- Occlusion into volatile compounds. Some compounds sublime explosively and hence atomization may be enhanced.

Most of these interferences may be overcome by using a hotter flame (e.g. nitrous oxide–acetylene); by adding a chelating reagent, e.g. EDTA, to complex preferentially with the analyte; by adding a releasing agent, e.g. lanthanum, that will combine preferentially with phosphate; or by optimizing the flame conditions and viewing height.

Electrothermal AAS was renowned for being highly prone to interferences. However, modern methods and instrumentation have decreased this problem substantially. Interferences include memory effects, chemical interferences (loss of analyte as a volatile salt, carbide formation, condensation and recombination), background absorption (smoke), and physical interferences such as those resulting from placing the sample on a different part of the tube. The stabilized temperature platform furnace (STPF) concept has gone a long way to eliminating the problem. The concept is that a ‘matrix modifier’ is required for interference-free determinations. This is a material that either decreases the volatility of the analyte, enabling higher ash temperatures to be achieved and hence boiling away more interference (a common example is a mix of palladium and magnesium nitrates) or increases the volatility of the matrix (for example, the introduction of air ‘burns away’ many interferences leaving the analytes in the atomizer). The matrix modifier may be placed in the autosampler, and added to all samples and standards. Other requirements for the STPF concept include a rapid heating rate during atomization, integrated signals (rather than peak height), a powerful background correction system (e.g. Zeeman), fast electronics to measure the transient signal, and isothermal operation. Isothermal operation means that the analyte is vaporized from the graphite surface into hot gas. In normal electrothermal AAS, the analyte leaves the hot tube wall and enters the cooler gas phase. Under these circumstances it may recombine with some other species to form a compound and thus be lost analytically. To overcome this, platforms that are only loosely in contact with the tube walls have been developed. The analytes in this case are atomized not by the resistively heated graphite tube but by the surrounding gas. The transversely heated tubes described earlier have a platform as an integral part of the tube. A modification to this technique, called probe atomization, has also been developed. Here the sample is placed on a probe and is then dried and ashed in the normal way. The probe is then removed from the tube, which is heated to the atomization temperature. The probe is then re-introduced into the hot environment.

Most analytical techniques for use with furnace work now utilize the advantages of the STPF concept. Other interferences, e.g. formation of carbides, may be partially overcome by treating the atomizer with a carbide-forming element, for example by soaking in tantalum solution. The principle is that the tantalum occupies the active sites on the surface, thereby preventing the analyte from forming a carbide.

## Methods

Numerous methods have been described for use with AAS, but the majority require the sample to be in a liquid form. Some ETAAS systems allow the analysis of solids directly (e.g. by weighing small amounts onto sampling boats that may be slotted into specialized tubes), but usually, if solids are to be analysed, acid decomposition methods are required. An alternative is the analysis of slurries, using finely ground material (< 10 μm) dispersed in a solvent. Manual agitation of the slurry ensures homogeneity of the sample, enabling a representative aliquot to be introduced. Some autosamplers come equipped with an ultrasonic agitator that will perform the same task. This method of analysis is frequently used in ETAAS.

If samples containing high dissolved (or suspended) solids are to be analysed by FAAS, a flow injection technique (in which discrete aliquots of sample are introduced) may be used. This has also been referred to as ‘gulp sampling’. This prevents salting up of the nebulizer and blockage of the burner slot, which are obviously undesirable effects that have an adverse effect on sensitivity and signal stability. Flow injection techniques may also be used to preconcentrate analytes. If the sample flows through a column containing an ion-exchange resin, the analytes will be retained. Elution with a small volume of acid may yield very high preconcentration factors. This technique has been used for both FAAS and, more recently, for ETAAS, yielding limits of detection far superior to those achieved under normal conditions. A typical flow injection manifold suitable for this type of application is shown in **Figure 8**.

A more traditional method of preconcentration is solvent extraction. Flame AAS may be used as a detector for analytes in organic solvents (sensitivity

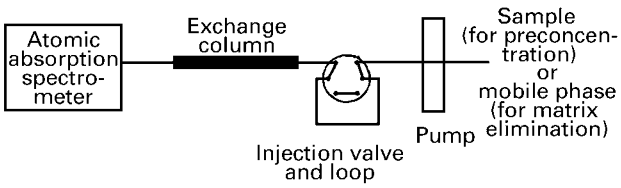



**Figure 8** Typical flow injection manifold for matrix analyte preconcentration or matrix elimination.

may even be improved because of improved nebulization efficiency), but occasional blocking of the burner slot by carbon may occur. This may be removed by gentle rubbing with a noncombustible material (such as a spatula). Organic solvents may also be introduced to ETAAS, but the 'dry'-stage temperature has to be modified to prevent sample loss by spitting.

Hydride generation is a common method for the detection of metalloids such as As, Bi, Ge, Pb, Sb, Se, Sn and Te, although other vapours, e.g. Hg or alkylated Cd, may also be determined. This technique improves the sensitivity of the analysis substantially. Since the sample is in the gas phase, the sample transport efficiency is close to 100%. The hydrides atomize readily in the flame, although this approach is usually used in conjunction with a quartz T-piece in the atom cell. Methods have been developed that trap the hydrides on the surface of a graphite tube for use with ETAAS. This leads to preconcentration and further improvements in detection limit.

Chromatography has been coupled to AAS to effect speciation analysis. Here different chemical forms of an analyte are separated (either by high-performance liquid chromatography or, if they are sufficiently volatile, by gas chromatography) prior to introduction to the atomic absorption instrument. Gas chromatography is usually coupled directly with a T-piece interface arrangement, but HPLC couplings introduce the sample through the conventional nebulizer/spray chamber assembly. A small postcolumn air bleed may be necessary to compensate for the differences in flow rate between the chromatograph (1–2 mL $min^{-1}$) and the uptake rate of the nebulizer (5–10 mL $min^{-1}$). Chromatography is not frequently coupled with ETAAS because the atom cell is not well suited to continuous monitoring.

The limited linear range offered by the AAS technique may be partially overcome by using alternative, less sensitive lines. This is a useful technique that avoids the requirement to dilute samples. Care must be taken, however, to ensure that viscosity effects do not cause a difference in nebulization efficiency between samples and standards. Alternatively, burner head rotation may be used. This shortens the path in the flame through which the light beam travels and hence decreases sensitivity. However, there is often an increase in noise as a result of using this approach.

## Conclusions

The sensitivity offered by the various techniques utilizing AAS for the determination of metals and metalloids is often sufficient for many applications. Direct aspiration of samples in solution into a flame atomic absorption instrument provides the analyst with rapid acquisition of data with good precision. The technique is remarkably free from interferences and those that do exist may usually be overcome by judicious choice of operating conditions. The instrumentation involved is relatively simple and, for flame work, analyses may easily be performed by non-expert operators. If increased sensitivity is required, the analyst has the option of preconcentrating the analyte using one of a number of methods prior to introduction to flame AAS, or alternatively may use electrothermal AAS. The latter technique, however, does require more experience if reliable results are to be obtained.

A potential disadvantage of using atomic absorption is that in the past instruments have been capable of determining only one analyte at a time. Modern instrumentation has improved this number to as many as six, although it must be stressed that compromise conditions (e.g. in ETAAS temperature programs) must be used. This may lead to a decrease in overall sensitivity compared with single-element determinations. However, the ongoing development of multielement AAS will certainly boost the use of this already popular and versatile technique.

*See also:* **Atomic Absorption, Theory; Atomic Fluorescence, Methods and Instrumentation; Atomic Spectroscopy, Historical Perspective; Light Sources and Optics.**

## Further reading

Dedina J and Tsalev DL (1995) *Hydride Generation Atomic Absorption Spectrometry*. Chichester: Wiley.

Ebdon L, Evans EH, Fisher A and Hill SJ (1998) *An Introduction to Analytical Atomic Spectrometry*. Chichester: Wiley.

Harnly JM (1996) Instrumentation for simultaneous multielement atomic absorption spectrometry with graphite furnace atomization. *Fresenius Journal of Analytical Chemistry* **355**: 501–509.

Haswell SJ (ed) (1991) *Atomic Absorption Spectrometry, Theory, Design and Applications*, Analytical Spectroscopy Library. Amsterdam: Elsevier.

Hill SJ, Dawson JB, Price WJ, Shuttler IL, Smith CMM and Tyson JF (1998) Atomic spectrometry update — advances in atomic absorption and fluorescence spectrometry and related techniques. *Journal of Analytical Atomic Spectrometry* **13**: 131R–170R.

Vandecasteele C and Block CB (1993) *Modern Methods for Trace Element Determination*. Chichester: Wiley.

# Atomic Absorption, Theory

**Albert Kh Gilmutdinov**, Kazan State University, Russia

Copyright © 1999 Academic Press

| ATOMIC SPECTROSCOPY |
|---|
| Theory |

Atomic absorption spectroscopy (AAS) is a technique for quantitative determination of metals and metalloids by conversion of a sample to atomic vapour and measurement of absorption at a wavelength specific to the element of interest. Owing to high sensitivity and selectivity, the technique is widely used for fundamental studies in physics and physical chemistry: measurements of oscillator strengths, diffusion coefficients of gas phase species, partial pressure of vapours, rate constants of homogeneous and heterogeneous reactions, etc. The widest application of AAS, however, is in analytical chemistry. Nowadays it is one of the most popular techniques for trace analysis of over 65 elements in practically all types of samples (environmental, biological, industrial, etc.).

The vast majority of substances to be analysed by AAS are in the condensed phase. At the same time, atomic absorption, like any other atomic spectrometry technique, can only detect free atoms that are in the gas phase. Thus, as analyte initially present in solution at a concentration $c$(liq) must first be transferred into the gas phase via the atomization process to produce $N$(g) free atoms in an atomizer volume. The analyte atoms are then detected by absorption of radiation from a primary source at a wavelength characteristic of the element resulting in an absorption signal, $A$. Thus, the general scheme of AAS can be presented as follows:

$$c(\mathrm{liq}) \xrightarrow{\text{atomization}} N(\mathrm{g}) \xrightarrow{\text{absorption}} A$$

The primary goal of the theory of AAS is to establish the relationship between the measured analytical signal, atomic absorbance, $A$, and the analyte concentration, $c$, in the sample. Theoretical description of the two stages involves entirely different sciences: theory of atomization is based on thermodynamics, kinetics and molecular physics, while description of absorbances is based on optics and spectroscopy. Below, the two stages will be considered separately.

## Atom production

In the first step of an AA determination the element of interest must be atomized. The ideal atomizer would provide complete atomization of the element irrespective of the sample matrix producing a well-defined and reproducible absorption layer of the analyte atoms. Excitation processes, however, should be minimal so that analyte and background emission noise is small. For achieving the best detection limits, the analyte vapour should not be highly diluted by the atomizer gas. Various types of electric discharges, laser radiation, electron bombardment and inductively coupled plasmas were tested as atomizers. Although not ideal, high-temperature flames and electrothermal atomizers have gained acceptance as the most common technique of atom production in AAS.

### Flame atomization

The longest practiced technique for converting a sample into atoms is the spraying of a sample solution into a combustion flame. The most popular flames used in AAS are the air–$C_2H_2$ flame providing a maximum temperature of about 2500 K and the $N_2O$–$C_2H_2$ flame with a maximum temperature exceeding 3000 K. Atomization occurs because of the high enthalpy and temperature of the flame, and through chemical effects. Owing the generation of cyanogen radicals, which are known to be an efficient scavenger for oxygen, the nitrous oxide-acetylene flame provides not only a hotter but also a more reducing environment for atom production: the lack of oxygen moves equilibria such as $MeO \rightleftharpoons Me + O$ to the right. Therefore a $N_2O$–$C_2H_2$ flame provides greater atomization efficiencies and thus better detection limits for refractory elements, such as Al, Ta Ti, Zr, Si, V and the rare earths.

Atom production in flames is an extremely complex process consisting of many stages and involving numerous sides processes. A quantitative theory of analyte atomization in flames is absent. Qualitatively, the process can be described as follows: the solution droplets sprayed into the flame are first dried, the resulting solid microparticles become molten and vaporize or thermally decompose to produce gaseous molecules that are finally dissociated into free atoms. The atoms may further be excited and ionized and form new compounds. These processes are dependent upon the temperature and the reducing power of the flame and occur within a few milliseconds – the time required by the sample to pass through the

flame. The processes taking place in flames are shown in more detail in **Figure 1**.

### Electrothermal atomization

In the majority of cases this occurs in a small graphite tube where the sample to be analysed is introduced via a small aperture in the upper tube wall. Metals such as tantalum or tungsten are also used to construct electrothermal atomizers as well as non-tubular geometry for the atomizer (graphite cups, rods, filaments, etc.). After a 5–50 μL droplet of the sample solution has been deposited into the graphite furnace, it is heated resistively through a series of preprogrammed temperature stages to provide: (i) drying of the droplet to evaporate the solvent (~ 370 K for about 30 s); (ii) thermal pretreatment (pyrolysis) to remove volatiles without loss of analyte (600–1500 K for 45 s) (this allows selective volatilization of matrix components that are purged from the atomizer by a flow of inert gas); (iii) analyte atomization (1500–3000 K achieved rapidly at a heating rate of about 2000 K $s^{-1}$). A major difference of electrothermal atomization from atomization in flames is that some matrix components are removed at the pyrolysis step and the atomization takes place within a confined geometry in an inert gas atmosphere. During atomization, the purge gas flow is shut off so that analyte atoms remain in the probing beam as long as possible.

The number of analyte atoms in the atomizer volume $N(t)$, generated during the atomization stage is described by the convolution:

$$N(t) = \int_0^t S(t')R(t,t')\,dt' \quad [1]$$

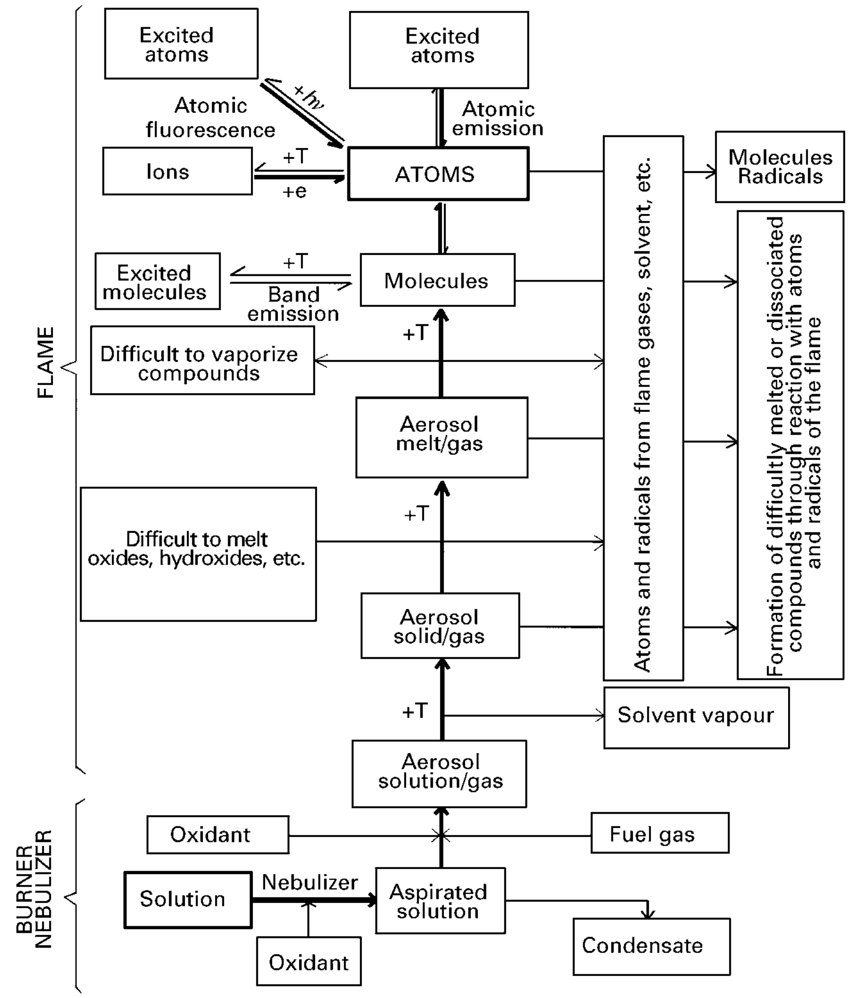



**Figure 1** Schematic representation of the processes taking place in a flame. The bold arrows show the pathways for analyte atom production and thin arrows show side processes. Reproduced with permission from Welz B (1985) *Atomic Absorption Spectrometry*, 2nd edn, p. 167. Weinheim: VCH.

Here *the supply function* $S(t')$, gives the rate (atoms $s^{-1}$) of the primary generation of analyte atoms from the deposited sample, whereas *the removal function*, $R(t,t')$ characterizes the subsequent transfer of the vaporized atoms in the atomizer volume. In the simplest case of vaporization of a monolayer of atoms from the atomizer surface the supply function is given as:

$$S(t) = N_0 \nu \exp\left\{ -\frac{E}{RT(t)} - \nu \int_0^t \exp\left[ -\frac{E}{RT(t')} \right] dt' \right\} \quad [2]$$

where $N_0$ is the total number of atoms in the deposited sample, $\nu$ and $E$ are the frequency factor and the activation energy of the vaporization process, respectively, $R$ is the gas constant and $T(t)$ is the time-dependent atomizer temperature. A typical example of the supply function in electrothermal AAS is given in **Figure 2**. The removal function $R(t,t')$ gives the probability of an analyte atom, vaporized at the time, $t'$, to be present in the atomizer volume at a later instant, $t > t'$. Early in time, the probability is equal to unity and decreases with increasing time because of loss processes. Generally, the loss processes include concentration and thermodiffusion, and convection. The knowledge of the removal function allows calculation of the mean residence time, $\tau$, of analyte in the atomizer volume. This parameter is defined as $\tau = \int_0^\infty R(t)\,dt$ and gives the life-time of evaporated atoms in the atomizer volume. When the analyte loss is governed only by concentration diffusion through the tube ends, the residence time can be estimated by the following relationship:

$$\tau = \frac{1}{8D}\left( L^2 - \frac{a^2}{3} \right) \quad [3]$$

where $L$ is the tube length, $a$ is the linear dimension of the area occupied by the sample on the atomizer surface, $D$ is the analyte diffusion coefficient. Typically, the analyte residence time in tube graphite atomizers is a few tenths of a second, which is about 100 times longer than in a flame. The long residence time, along with the high degree of analyte atomization caused by the reducing environment, leads to a 10- to 1000-fold increase in graphite furnace sensitivity compared with flame AAS.

The typical change in the total number of analyte atoms, $N(t)$, in the tube electrothermal atomizer obtained by convolution [1] of the supply function with the diffusion removal function is given in **Figure 2**.

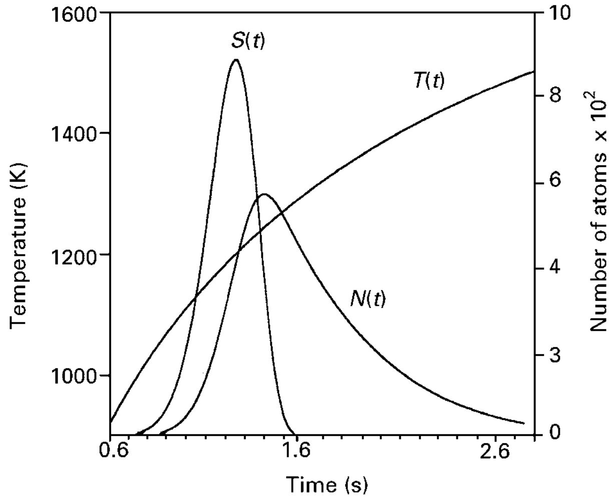



**Figure 2** Supply function $S(t)$ of silver atoms computed for the change in the atomizer temperature $T(t)$ ($E = 280$ kJ mol$^{-1}$, $\nu = 10^{13}$ s$^{-1}$); the change in the total number of silver atoms within the atomizer volume $N(t)$.

The actual atom production and dissipation processes in electrothermal atomizer are much more complex than that described above. In many cases the analyte is atomized as a result of a set of multistage processes taking place simultaneously. The basic processes that occur with analyte in a graphite furnace are shown schematically in **Figure 3** and listed in **Table 1**. The last column of the Table shows the elements for which the respective processes have been documented.

In addition to the above widely used atomization techniques, mention must also be made of 'the hydride generation technique'. The method is based on conversion of a hydride forming element (As, Bi, Ge, Sb, Se, Sn, Te) in an acidified sample to volatile hydride and transport of the released hydride to an atomizer (hydrogen diffuse flame, graphite furnace, heated quartz tube) where they are atomized to give free analyte atoms. Sodium borohydride is almost exclusively used as an agent for conversion of analyte to hydride:

$$NaBH_4 + 3H_2O + HCl \longrightarrow H_3BO_3 + NaCl + 8H \xrightarrow{El^{m+}} ElH_n + H_2\,(\text{excess})$$

where El, is the element of interest and $m$ may or may not equal $n$. The advantage of sample volatilization as a gaseous hydride lies in the analyte's preconcentration and separation from the sample matrix. This results in an enhanced sensitivity and in a significant

**Table 1** Selection of processes involving analyte in graphite tube atomizers

| *Process*[a] | *Ref. in* **Figure 3** | *Elements* |
|---|---|---|
| Me (s,l) → Me (g) | 1 | Ag, Au, Co, Cu, Ni, Pb, Pd |
| MeO (s,l) → Me (g) | 1 | Al, Cd, In, Pb |
| MeO (s,l) + C(s) → Me (g) | 1 | Mn, Mg |
| Me C (s,l) → Me (g) | 1 | Al, Ba, Ca, Cu, Mo, Sr, V |
| MeOH (s,l) → Me (g) | 1 | Rb |
| MeO (s,l) → MeO (g) | 1′ | Al, As, Ga, In, Tl |
| Me (s,l) → $Me_2$ (g) | 1′ | Se, Co |
| Me (ad) → Me (g) | 3 | As, Cu |
| Me (g) → Me (s,l) | 4 | Ag, Au, Cu, Mg, Mn, Pd |
| Me (g) + O (g) → MeO (g) | 5 | Al, Si, Sn |
| MeO (g) → Me (g) | 5′ | Al, As, Cd, Ga, In, Mn, Pb, Zn |
| MeO (g) + C(s) → Me (g) + CO | 6 | Al, Ga, In, Tl |
| Me (g) + C(s) → MeC (g) | 6′ | Al |

[a] s, l, g, ad stand for solid, liquid, gaseous and adsorbed species, respectively.

reduction of interferences during atomization. The main disadvantages of the method are interferences by substances in the solution that reduce the efficiency of hydride generation.

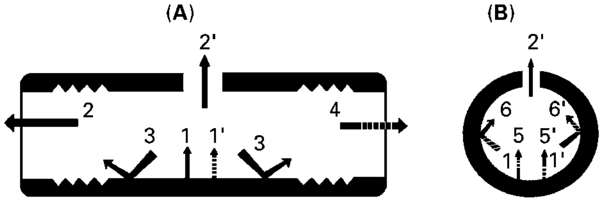



**Figure 3** Schematic representation of the basic physical and chemical processes taking place in a tube electrothermal atomizer. Solid arrows denote pathways of free analyte atoms, dotted arrows show the pathways of the analytes that are bound into molecules. *Primary generation* of the analyte vapour from the site of sample deposition as an atomic (1) or a molecular (1′) species. Irreversible *loss of analyte* from the furnace through its ends (2) and through the sample dosing hole (2′) by diffusion and convection. *Physical adsorption/desorption* at the graphite surface (3). Gas phase *condensation* (4) at the cooler parts of the atomizer. *Gas phase reactions* (5) that bind free analyte atoms into stable molecules or those (5′) that increase the free atom density. *Heterogeneous reactions* of analyte vapour with the atomizer walls: includes both production (6) and loss (6′) of free atoms at the furnace wall.

Relative detection limits achieved by the described atomization techniques are presented in **Table 2.**

**Table 2** Selected relative detection limits (μg $L^{-1}$) for different atomization techniques[a]

| *Element* | *Flame AA* | *GF AA* | *Hydride AA* | *Element* | *Flame AA* | *GF AA* | *Hydride AA* |
|---|---|---|---|---|---|---|---|
| Ag | 1.5 | 0.02 | | Mn | 1.5 | 0.035 | |
| Al | 45 | 0.1 | | Mo | 45 | 0.08 | |
| As | 150 | 0.2 | 0.03 | Na | 0.3 | 0.02 | |
| Au | 9 | 0.15 | | Ni | 6 | 0.3 | |
| B | 1000 | 20 | | P | 75 000 | 130 | |
| Ba | 15 | 0.35 | | Pb | 15 | 0.06 | |
| Be | 1.5 | 0.008 | | Pd | 30 | 0.8 | |
| Bi | 30 | 0.25 | 0.03 | Pt | 60 | 2 | |
| Ca | 1.5 | 0.01 | | Rb | 3 | 0.03 | |
| Cd | 0.8 | 0.008 | | Ru | 100 | 1 | |
| Co | 9 | 0.15 | | Sb | 45 | 0.15 | 0.045 |
| Cr | 3 | 0.03 | | Se | 100 | 0.3 | 0.03 |
| Cu | 1.5 | 0.1 | | Sn | 150 | 0.2 | 0.03 |
| Fe | 5 | 0.1 | | Sr | 3 | 0.025 | |
| Ge | 300 | 10 | 1 | Te | 30 | 0.4 | 0.03 |
| Hg | 300 | 0.6 | 0.009 | Ti | 75 | 0.35 | |
| Ir | 900 | 3 | | Tl | 15 | 0.15 | |
| K | 3 | 0.008 | | V | 60 | 0.1 | |
| Li | 0.8 | 0.06 | | Zn | 1.5 | 0.1 | |
| Mg | 0.15 | 0.004 | | | | | |

[a] Detection limits are based on 98% confidence level (3 standard deviations). The values for the graphite furnace technique are referred to sample aliquots of 50 μL.

Adapted with permission from *The Guide to Techniques and Applications of Atomic Spectroscopy* (1997) p 5. Norwalk, CT, USA: Perkin-Elmer.

## Atomic absorption

AAS is based on the interaction of the probing radiation beam from a primary source with gas phase analyte atoms. Two different types of primary source are used in AAS: line sources emitting narrow spectral lines of the element to be analysed and a continuum source (normally high-pressure xenon arc lamp) that produces radiation from below 180 nm to over 800 nm. In principle, the continuum source is best suited for AAS. It has a unique capability for multi-element determinations, and provides extended analytical range and inherent background correction. However, the lack of intensity found below 280 nm still limits the use of the source in research laboratories. In conventional AAS, the line radiation of the element of interest is generated in a hollow cathode lamp (HCL) or an electrodeless discharge lamp (EDL). The spectral profile, $J(\lambda)$, of the analytical line emitted by the source interacts with the absorption profile, $k(\lambda)$, of the analyte atoms in the atomizer (**Figure 4**). It is seen that radiation absorption is strongly dependent on the shape and relative position of the two spectral profiles. In order to describe quantitatively absorption, both emission and absorption profiles must first be well defined.

### Emission and absorption line profiles

Three major broadening mechanisms determine emission and absorption profiles in AAS: the Doppler effect, interatomic collisions and hyperfine splitting. Other types of broadening (natural broadening, Stark broadening, etc.) are negligible. The Doppler broadening originates from the thermal agitation of emitting and absorbing atoms that results in a Gaussian-shaped spectral line (curve G in **Figure 5**). The full width, $\Delta\nu_D$, at one-half of the maximum intensity (FWHM) of the Doppler-broadened atomic line is given by:

$$\Delta\nu_{\mathrm{D}} = \frac{7.16}{\lambda}\sqrt{\frac{T}{M}} \qquad [4]$$

where $\Delta\nu_D$ is in $cm^{-1}$, $\lambda$ in nm, $T$ in K and $M$ in g $mol^{-1}$. The width is inversely proportional to the wavelength of the transition, to the square root of the absolute temperature, $T$, and to the reciprocal of the square root of the molecular weight of the analyte, $M$. The absorption line widths predicted from this equation range from 1 pm to 10 pm for flame and electrothermal atomizers. The Doppler widths of the lines emitted by HCLs and EDLs are 2–3 times lower than those for absorption profiles in atomizers.

The second major broadening mechanism in AAS is collision or pressure broadening which originates from deactivation of the excited state due to interatomic collisions. There are two types of collisions. Collisions between two analyte atoms lead to resonance broadening which is negligible in analytical

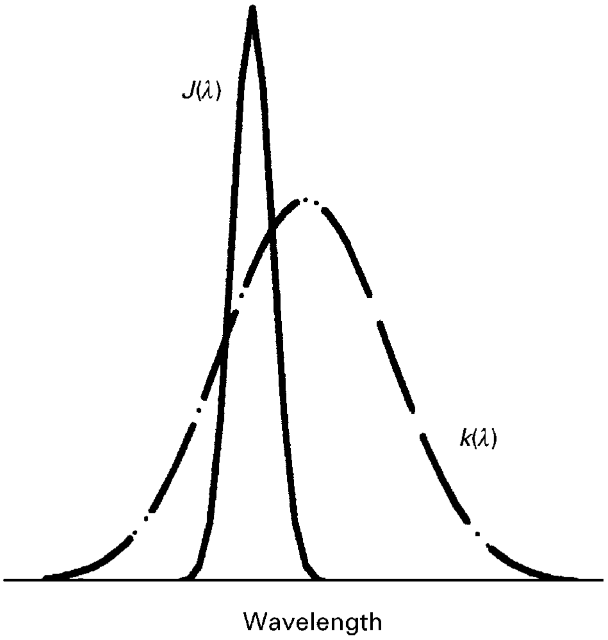



**Figure 4** Schematic diagram of the emission $J(\lambda)$ and absorption $k(\lambda)$ spectral profiles typical in AAS.

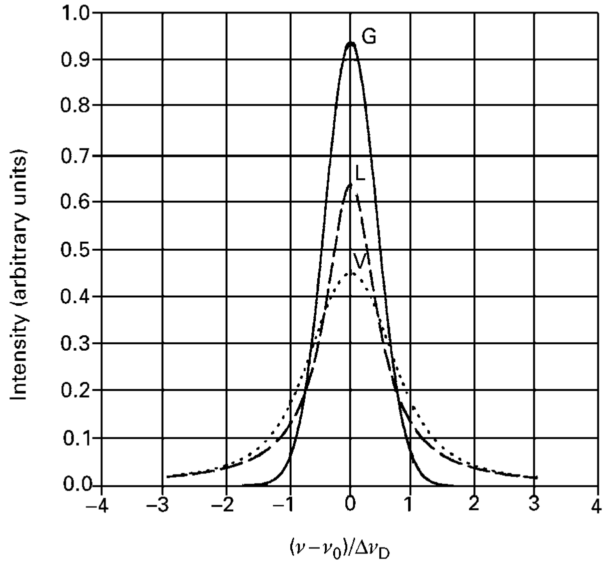



**Figure 5** Normalized theoretical Gaussian (G), Lorentzian (L) and Voigt (V) line profiles versus dimensionless frequency. The FWHMs and areas of the Gaussian and Lorentzian peaks are equal; the Voigt profile is the convolution of the (G) and (L) profiles.

AAS because the analyte concentration is small. However, Lorentz broadening, which involves collisions between analyte atoms and foreign species (Ar atoms in electrothermal atomizers) called perturbers, are significant. Compared to a Doppler spectral profile, the Lorentzian profile is broader, has a lower peak height and is described by a Lorentzian function f(curve L in **Figure 5**) with the FWHM given by:

$$\Delta\nu_L = N_p \sigma \sqrt{\frac{8RT}{\pi}\left(\frac{1}{M_A}+\frac{1}{M_p}\right)} \quad [5]$$

Here $N_p$ is the number density of perturbers ($cm^{-3}$), $\sigma$ is the collisional cross-section ($cm^2$), $M_A$ and $M_p$ is the molecular weight ($g\ mol^{-1}$) of the analyte and the perturbers, respectively. The collision width predicted by Equation [5] for absorption profiles is about the same order of magnitude as the Doppler line width (1–10 pm). The $\Delta\nu_L$ values for lines emitted by the HCLs and EDLs are much lower than those for the absorption profiles in atomizers because of low pressure in the primary sources. Therefore this part of the broadening is normally neglected in the sources. However, at high optical densities of an absorbing layer when the wings of the emission line play an important role, the Lorentzian component in the broadening of the emission line is also becoming important. In addition to broadening the lines, collisions also cause a small shift in the line centre (towards the longer wavelengths in an Ar environment) and an asymmetry in the far line wings. The shift is proportional to the collision width and theory predicts a value of 2.76 for the width-to-shift ratio (it is shown schematically in **Figure 4**).

The total spectral profile of an atomic line is given by the convolution of the Gaussian profile (due to Doppler broadening) and the Lorentzian profile (due to pressure broadening). The result is the Voigt function (curve V in **Figure 5**), $H(\omega,a)$:

$$H(\omega,a) = \frac{a}{\pi}\int_{-\infty}^{\infty} \frac{e^{-y^2}dy}{a^2+(\omega-y)^2} \quad [6]$$

where $a = \sqrt{\ln 2}\,(\Delta\nu_L/\Delta\nu_D)$ is the damping constant determining the shape of the Voigt profile: as the damping constant increases, the profile broadens and its peak height decreases. In the extreme cases of $a = 0$ and $a = \infty$, the Voigt function goes to the purely Gaussian function and purely Lorentzian function, respectively. Normally, a dimensionless frequency, $\omega$, normalized by the Doppler width $\Delta\nu_D$ of the absorption profile is used for computations: $\omega = (\nu/\Delta\nu_D)2\sqrt{\ln 2}$. Within about 1% accuracy, the FWHM of the Voigt profile can be estimated from the following empirical equation:

$$\Delta\nu_V = \frac{\Delta\nu_L}{2} + \sqrt{\left[\left(\frac{\Delta\nu_L}{2}\right)^2 + (\Delta\nu_D)^2\right]} \quad [7]$$

Finally, the third broadening mechanism is hyperfine splitting which is caused by isotope shifts and nuclear spin splitting. Generally a spectral line consists of $n$ hyperfine components with relative intensities, $b_i$, ($\Sigma_{i=1}^{n} b_i = 1$) that are located at a distance $\Delta\omega_i$ from the component with the minimum frequency. Each component is Doppler and collision broadened and described by the Voigt function [6]. Thus in the general case the emission $J(\omega)$ and absorption $k(\omega)$ profiles are presented as follows:

$$J(\omega) = \sum_{i=1}^{n} b_i H_i\left(\frac{\omega-\Delta\omega_i}{\alpha}; a_e\right),$$
$$k(\omega) = \sum_{j=1}^{n} b_j H_j(\omega - \Delta\omega_j + \Delta\omega_S; a) \quad [8]$$

where $\alpha$ is the ratio of the Doppler width of the emission line to the Doppler width of the absorption line. The equation takes into account that the emission profile, $J(\omega)$, is approximately $1/\alpha$ times narrower than the absorption profile, $k(\omega)$, in the scale of dimensionless frequency $\omega$ and that the absorption profile, is pressure-shifted relative to the emission profile to a value $\Delta\omega_S$. For many analytical lines used in AAS the hyperfine splitting is greater than Doppler and pressure broadening and determines the total FWHM.

In graphite furnace AAS, the damping constant, $a$, varies from 0.17 (Be resonance line, 234.9 nm) to 3.27 (Cs, 852.1 nm) for absorption lines and its value for emission lines, $a_e$, in the primary sources varies from 0.01 to 0.05. A relatively small value of the $a$-parameter means that the emission profile is primarily determined by the Doppler broadening.

**Figure 6** shows the emission and absorption spectral profiles of the lead 283.3 nm resonance line computed for the conditions of graphite furnace AAS using Equation [8]. The line consists of five hyperfine components that are clearly seen in the emission profile. Because of higher temperature and much higher pressure, the absorption profile is wider and shifted towards the longer wavelengths.

## Calibration curve shapes

It has been demonstrated that in AAS the atomization efficiency is close to 100% for the majority of elements in aqueous solutions. Assuming complete atomization, the shape of calibration curves in AAS is determined by the dependence of detected atomic absorbance on the number of absorbing atoms. The relationship $A = f(N)$ is called the *concentration curve* and its description is one of the primary goals of theoretical AAS.

The phenomenon of radiation absorption had already been studied quantitatively early in the 18th century, mainly on liquids and crystals. The investigations resulted in the Beer–Lambert law relating the radiant flux, $\Phi_{tr}$, transmitted through an absorbing layer to the optical properties of the layer. In modern formulation, the law applied to a spatially uniform absorbing layer of length $l$ and a uniform parallel probing beam is written as follows:

$$\Phi_{tr}(\lambda) = \Phi_0(\lambda)\exp[-\mu(\lambda)l] \qquad [9]$$

where $\Phi_0(\lambda)$ is the incident radiant flux incident at a wavelength $\lambda$, $\mu(\lambda)$ is a wavelength-dependent attenuation coefficient which is the sum of absorption, $k(\lambda)$, and scattering, $\varepsilon(\lambda)$, coefficients: $\mu(\lambda) = k(\lambda) + \varepsilon(\lambda)$. Curve $J_{tr}(\lambda)$ in **Figure 6** shows the spectral profile of the lead 283.3 nm resonance line that is transmitted through a uniform layer of $3 \times 10^{12}$ cm$^{-2}$ lead atoms. It was assumed that the spatially uniform incident radiation with the spectral profile $J(\lambda)$ is absorbed by the absorption profile $k(\lambda)$, without scattering, following Equation [9]. Generally, a non-uniform radiation with intensity distribution $J(x,y)$ over the beam cross section is used to probe a non-uniform layer of absorbing species with the number density $n(x,y,z)$. With non-uniform probing beams, a radiant flux passing through a surface $S$ within a spectral bandwidth $\{-\lambda^*, \lambda^*\}$ is expressed in terms of intensities as follows:

$$\Phi = \iint\limits_{S} \int_{-\lambda^*}^{\lambda^*} J(x,y)J(\lambda)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\lambda$$

where $J(\lambda)$ is a function describing the spectral composition of the radiation.

Absorbance, $A$, which is the analytical signal in AAS, is defined as the logarithm of the ratio of the incident radiant flux, $\Phi_0$, to the radiant flux, $\Phi_{tr}$, that is transmitted through the absorbing layer of the analyte atoms. Using the Beer–Lambert law [9] for expressing the transmitted radiant flux, the general relationship for the absorbance recorded by an AA spectrometer can be expressed as follows:

$$A = \lg\frac{\Phi_0}{\Phi_{tr}} = \lg\left\{\frac{\int_0^h\int_0^b\int_{-\lambda^*}^{\lambda^*} J(x,y)J(\lambda)\,\mathrm{d}\lambda\,\mathrm{d}x\,\mathrm{d}y}{\int_0^h\int_0^b\int_{-\lambda^*}^{\lambda^*} J(x,y)\,J(\lambda)\exp\left(-\int_0^l k(\lambda;x,y,z)\,\mathrm{d}z\right)\mathrm{d}\lambda\,\mathrm{d}x\,\mathrm{d}y}\right\} \qquad [10]$$

Here, $b$ and $h$ are the width and the height of the monochromator entrance slit, $l$ is the length of the absorbing layer; the limits of $-\lambda^*$ to $\lambda^*$ are the spectral bandwidth isolated by the monochromator. It was assumed that radiation attenuation occurs only because of atomic absorption, i.e. $\varepsilon(\lambda) = 0$. The absorption coefficient, $k(\lambda;x,y,z)$, depends on the number density of absorbing atoms in the ground state, $n_0$ (cm$^{-3}$), and the spectral profile, $k(\lambda)$, of the absorption line, and is presented as:

$$k(\lambda;x,y,z) = \frac{2\sqrt{\pi}\ln 2e^2}{mc}\cdot\frac{f}{\Delta\nu_D}\cdot n_o(x,y,z)\,k(\lambda) = c(\lambda,T)\,n_o(x,y,z) \qquad [11]$$

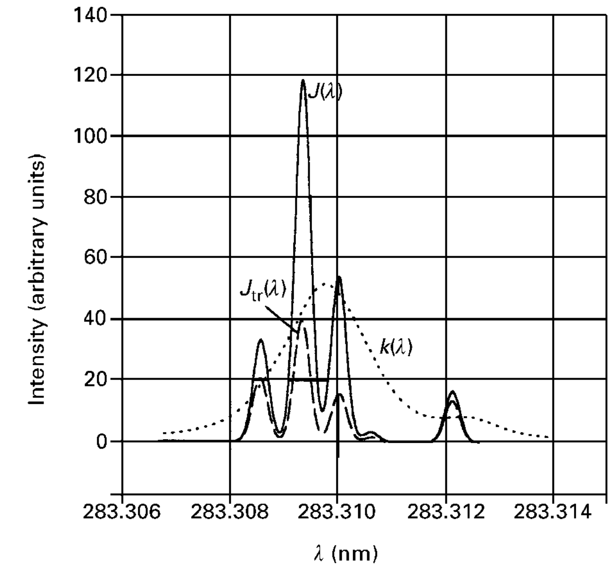



**Figure 6** Emission $J(\lambda)$ and absorption $k(\lambda)$ spectral profiles of the 283.3 nm lead resonance line computed for the conditions of graphite furnace AAS ($a_e = 0.01$, $a = 1.24$; $T = 2100$ K). $J_{tr}(\lambda)$ is the emission profile that is transmitted through a uniform layer of $3 \times 10^{12}$ cm$^{-2}$ lead atoms.

Here $e$ and $m$ are the charge and mass of the electron, respectively, $c$ is the velocity of light, $f$ is the oscillator strength of the spectral transition, $\Delta v_D$ is the Doppler width of the absorption line. The coefficient $c(\lambda, T)$ denotes the combination of values that depend only on the spectral features of the analysis lines.

The general relationship [10] is simplified significantly if the following assumptions are made: (i) radial distributions of the probing beam and the analyte atoms are uniform, i.e. $J(x,y) = \text{const}$, $n(x,y,z) = n(z)$; (ii) the atomizer is spatially isothermal, i.e. $T(x,y,z) = \text{const}$ and (iii) the emission line is a single and infinitely narrow line at $\lambda_0$ so that it is described by Dirac's delta-function: $J(\lambda) = J_0\,\delta(\lambda - \lambda_0)$. Substitution of these simplifications into Equation [10] gives:

$$A = (\lg e)\int_0^l k(z;\lambda_0)\,\mathrm{d}z = (\lg e)c(\lambda_0, T)N \qquad [12]$$

where $c(\lambda_0,T)$ is a constant for given transition and for a given temperature, $N(\text{cm}^{-2}) = \int_0^l n(z)\mathrm{d}z$ is the number of absorbing atoms per unit cross-sectional area along the radiation beam; in the case of uniform analyte distributions, this value is proportional to the total number of analyte atoms in the atomizer. Thus, if the above assumptions are valid, the recorded absorbance is proportional to the total number of absorbing atoms in the atomizer and the concentration curve $A = f(N)$ is a straight line. This is a basis for AAS to be an analytical technique, because the recorded signal, $A$, is directly proportional to the unknown number, $N$, of analyte atoms. Such a concentration curve computed for the case of a lead resonance line is presented by curve 1 in **Figure 7**.

The actual situation, however, differs substantially from that presented by simple Equation [12]. It follows from general relationship [10] that absorbance is dependent on three groups of factors: (1) the spectral characteristics of the analysis line (dependence of $J$ and $k$ on $\lambda$); (2) the cross-sectional distribution of intensity in the incident beam, $J(x,y)$; and (3) spatial distribution of the analyte and temperature within the atomizer (dependence of $k$ on $x,y,z$). Thus detected absorbance, $A$, depends not only on the number, $N$, of absorbing atoms but also on the spectral features of the analysis lines used to probe the absorbing layer and the spatial distribution of analyte atoms and radiation intensity in the probing beam:

$$A = f(N; \text{Spectral, Spatial}) \qquad [13]$$

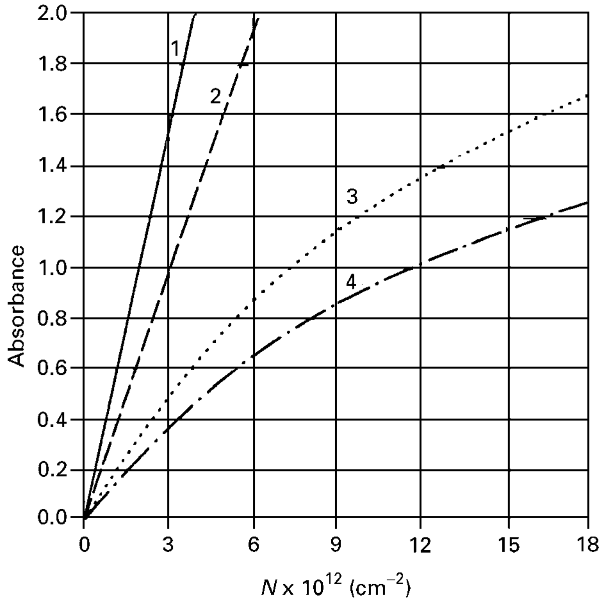



**Figure 7** The evolution of the 283.3 nm lead concentration curve as the spectral and spatial features are successively taken into account. Emission line is: single and infinitely narrow (1), single and broadened (2); all the spectral features of the line are taken into account including hyperfine splitting and the pressure shift (3); (4) as (3) plus absorbing layer is assumed to be non-uniform in the radial cross section with the analyte number density at the atomizer bottom 4 times greater than that in the upper part of the absorbing layer.

The effect of spectral features of the analysis line and spatial nonuniformities of the absorbing layer is illustrated in **Figure 7**. Curve 2 is the lead concentration curve when the resonance line is assumed to be a single but broadened one; accounting for all the spectral features of the emission and absorption profiles, including hyperfine structure and pressure shift, results in curve 3. It can be seen that broadening of the analysis line lead to a bending of the concentration curve and decreasing of its slope, the major effect being the hyperfine broadening. Additional accounting for the fact that the analyte is distributed nonuniformly in the atomizer cross section leads to further curvature and a decrease in the slope of concentration curve (curve 4). **Figure 7** illustrates clearly the importance of the *spectral* and *spatial* features in relationship [10]. Recent research shows, however, that these undesirable dependencies can be avoided if transmitted intensities are measured with spectral, spatial and temporal resolutions. The *multidimensional* absorbance detected in such a way is dependent only on the number of absorbing atoms irrespective of their spatial distribution and spectral features of the primary source.

### Interferences

An interference effect is a change of the analytical signal by the sample matrix as compared to the reference standard, typically an acidified aqueous solution. There are two types of interference in spectrochemical analysis, *spectral* and *nonspectral*. Spectral interference originates from incomplete isolation of the radiation emitted or absorbed by the analyte from other radiation detected by the spectrometer. Any reasons other than specific atomic absorption that cause radiation attenuation in equation [9] result in spectral interferences. They arise from: absorption of radiation by overlapping molecular bands or atomic lines of concomitants; scattering of source radiation by nonvolatilized microparticles (smoke, salt particles, condensed microdrops); foreign line absorption if the corresponding radiation happens to be emitted by the radiation source, in addition to the analysis line, within the spectral bandwidth of the monochromator. The first two sources of spectral interferences are most common and described as background attenuation or background absorption. In general, they are more pronounced at shorter wavelengths. Most background attenuation can be distinguished from the analyte absorption because the element only absorbs in the very narrow spectral region while the background attenuation is less specific and extends over a considerably broader wavelength band. Instrumental methods of correction for background attenuation effects include compensation using a continuum source and the Zeeman effect.

For interferences other than spectral, the analyte itself is directly affected. The nonspectral interferences are best classified according to the stage at which the particular interference occurs, i.e. solute-volatilization and vapour-phase interferences. A nonspectral interference is found when the analyte exhibits a different sensitivity in the presence of sample concomitants as compared to the analyte in a reference solution. The difference in the signal may be due to: analyte loss during the thermal pretreatment stage in the electrothermal atomizer: analyte reaction with concomitants in the condensed phase to form compounds that are atomized to a lesser extent, analyte ionization or change the degree of ionization caused by concomitants.

Reactions of the analyte with the tube material (graphite) or the purge gas are not normally considered to be interferences because they influence the analyte in the sample and in the standard to the same degree. A nonspectral interference is in many instances best detected by the use of the analyte addition technique. An interference exists if the slope of the additions curve is different from that of the calibration curve. For most systems, nonspectral interferences in electrothermal AAS can be effectively eliminated by adding a proper matrix modifier ($Pd\text{-}Mg(NO_3)_2$ is the most common modifier) that delays analyte volatilization until the atomizer temperature is sufficiently high and steady for efficient atomization, and buffers the gas phase composition.

## List of symbols

$a$ = damping constant; $a_e$ = damping constant for emission lines; $a$ = the linear dimension of the area occupied by the sample on the atomizer surface; $A$ = absorbance; $b$ = width of monochromator entrance slit; $c$(liq) = concentration of analyte in solution; $c$ = velocity of light; $D$ = analyte diffusion coefficient; $e$ = charge of electron; $E$ = activation energy; $f$ = oscillator strength; $h$ = height of monochromator entrance slit; $J(\lambda)$ = spectral profile; $k(\lambda)$ = absorption profile; $l$ = length of absorbing layer; $L$ = tube length; $m$ = mass of electron; $M$ = molecular weight of the analyte; $M_A$ = molecular weight of analyte; $M_p$ = molecular weight of perturbers; $n_o$ = number density of absorbing atoms in the ground state; $N$ = number of analyte atoms; $N$(g) = number of free atoms in an atomizer volume; $N_0$ = total number of atoms in the deposited sample; $N_P$ = number density of perturbers; $N(t)$ = number of analyte atoms in the atomizer volume; $R$ = gas constant; $R(t,t')$ = removal function; $S(t')$ = supply function; $t$ = time analyte atom present in atomizer volume; $t'$ = time of vaporization; $T$ = absolute temperature; $T(t)$ = time-dependent atomizer temperature; $\alpha$ = ratio of the Doppler width of the emission to the absorption line; $\Delta\nu_D$ = width of Doppler-broadened atomic line; $\Delta\nu_L$ = width of Lorentzian-broadened atomic line; $\varepsilon(\lambda)$ = scattering coefficient; $\lambda$ = wavelength; $\lambda$ = wavelength; $\mu(\lambda)$ = wavelength-dependent attenuation coefficient; $\nu$ = frequency factor; $\sigma$ = collisional cross section; $\tau$ = mean residence time; $\Phi$ = transmitted radiant flux; $\Phi_0$ = incident radiant flux; $\omega$ = frequency (dimensionless).

*See also:* **Atomic Absorption, Methods and Instrumentation; Atomic Fluorescence, Methods and Instrumentation; Fluorescence and Emission Spectroscopy, Theory.**

## Further reading

Alkemade CThJ, Hollander T, Snelleman W and Zeegers PJTh (1982) *Metal Vapours in Flames*. Oxford: Pergamon Press.

Chang SB and Chakrabarti CL (1985) Factors affecting atomization in graphite furnace atomic absorption spectrometry. *Progress in Analytical Atomic Spectroscopy* 8: 83–191.
Ebdon L (1982) *An Introduction to Atomic Absorption Spectroscopy*. London: Heyden.
Haswell SJ (Ed.) (1991) *Atomic Absorption Spectrometry*. Amsterdam: Elsevier.
Holcombe JA and Rayson GD (1983) Analyte distribution and reactions within a graphite furnace atomizer. *Progress in Analytical Atomic Spectroscopy* 6: 225–251.
L'vov BV (1970) *Atomic Absorption Spectrochemical Analysis*. London: Adam Hilger.
Mitchell ACG and Zemansky MW (1961) *Resonance Radiation and Excited Atoms*. Cambridge: Cambridge University Press.
Slavin W (1991) *Graphite Furnace AAS – A Source Book*, 2nd edn. Norwalk, CT: Perkin-Elmer Corporation.
Styris DL and Redfield DA (1993) Perspectives on mechanisms of electrothermal atomization. *Spectrochimica Acta Reviews* 15: 71–123.
Welz B (1985) *Atomic Absorption Spectrometry*, 2nd edn. Weinheim: VCH.
Winefordner JD (ed) (1976) *Spectroscopic Methods for Elements*. New York: Wiley-Interscience.

# Atomic Emission, Methods and Instrumentation

**Sandra L Bonchin**, Los Alamos National Laboratory, NM, USA
**Grace K Zoorob**, Biosouth Research Laboratories, Inc., Harahan, LA, USA
**Joseph A Caruso**, University of Cincinnati, OH, USA

Copyright © 1999 Academic Press

**ATOMIC SPECTROSCOPY**
**Methods & Instrumentation**

Atomic emission spectroscopy is one of the most useful and commonly used techniques for analyses of metals and nonmetals providing rapid, sensitive results for analytes in a wide variety of sample matrices. Elements in a sample are excited during their residence in an analytical plasma, and the light emitted from these excited atoms and ions is then collected, separated and detected to produce an emission spectrum. The instrumental components which comprise an atomic emission system include (1) an excitation source, (2) a spectrometer, (3) a detector, and (4) some form of signal and data processing. The methods discussed will include (1) sample introduction, (2) line selection, and (3) spectral interferences and correction techniques.

## Atomic emission sources

The atomic emission source provides for sample vaporization, dissociation, and excitation. The ideal excitation source will allow the excitation of all lines of interest for the elements in the sample, and do this reproducibly over enough time to encompass full elemental excitation. Excitation sources include but are not limited to (1) inductively coupled plasma (ICP), (2) direct current plasmas (DCP), (3) microwave induced plasmas (MIP), and (4) capacitively coupled microwave plasmas (CMP). Glow discharges are utilized for direct solids analyses, but will not be discussed here. An analytical plasma is a high-energy, slightly ionized gas (about 0.01 to 0.1% ionized).

### Inductively coupled plasmas

The most commonly used ion source for plasma spectrometry, the ICP, is produced by flowing an inert gas, typically argon, through a water-cooled induction coil which has a high-frequency field (typically 27 MHz) running through it (**Figure 1**). The alternating current in the coil has associated with it a changing magnetic field, which induces a changing electric field. The flowing gas is seeded with electrons by means of a Tesla coil. These electrons undergo acceleration by the electric field, and gain the energy necessary to excite and ionize the gaseous atoms by collision. This produces the plasma, self-sustaining as long as the RF and gas flows continue.

Sample particles entering the plasma undergo desolvation, dissociation, atomization, and excitation. The ICP has sufficiently long residence times and high enough temperatures so that the sample solvent is completely vaporized, and the analyte reduced to free atoms, which undergo excitation. This excitation results in the emission of light at specific frequencies for elements in the sample, which is

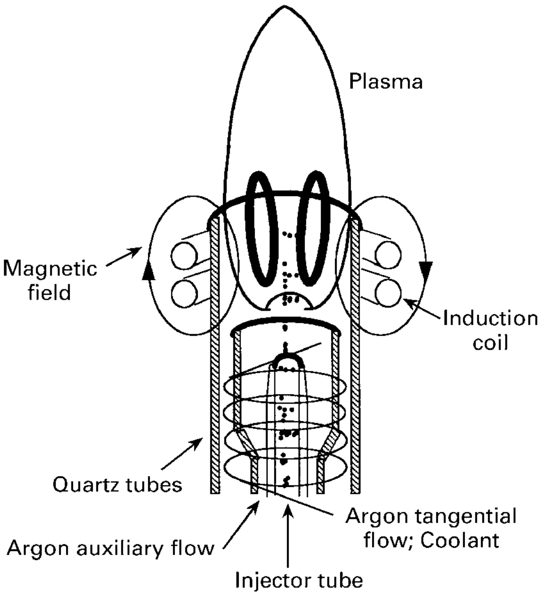



**Figure 1** Inductively coupled plasma and torch schematic.

proportional to their concentration. An emission spectrometer separates the frequencies of light into discrete wavelengths, which are unique to given elements in the sample.

### Direct current plasmas

In a direct current plasma, a dc current passing between two electrodes heats the plasma gas, again typically argon, and produces a discharge. The most common version is the 3-electrode system (**Figure 2**). This system has argon flowing around two graphite anodes and a tungsten cathode to produce the plasma. The sample is introduced between the anodes. Vaporization, atomization, ionization, and excitation occur. This technique is more tolerant of samples containing a high proportion of solids than the ICP method. However, it is less efficient due to lower plasma temperatures, and the electrodes need to be replaced frequently.

### Microwave induced plasmas

A microwave induced plasma (MIP) is an electrodeless discharge generated in a glass or quartz capillary discharge tube, often in a resonant cavity. These tubes generally have an inner diameter of the order of a few millimetres, and the plasma gas is an inert gas, such as helium or argon. The resonant cavity, which is hollow and of the order of a few centimetres diameter, allows coupling of the microwave power into the plasma gas flowing through the capillary discharge tube. The microwave power supply operates at a frequency of 2.45 GHz. Microwave plasmas can be produced at atmospheric pressure, if the design of the cavity allows. Helium plasmas typically require reduced pressure unless a 'Beenakker cavity' is employed (**Figure 3**). MIPs use lower power levels (hundreds of watts) than those required by the DCP and ICP. However, due to the decreased size of the MIP, the power densities produced are comparable. Argon has an ionization potential of 15.75 eV, which is not sufficient in the argon ICP to produce the excitation ionization energy needed to ionize some elements efficiently, such as the nonmetals. Helium plasmas are more efficient for these problem elements, as the ionization potential of helium is 24.6 eV. The MIP is not very adept at handling liquid samples, due to the low powers employed. However, at higher powers above 500 watts similar performance to the ICP may be obtained. The general lack of availability of these sources as part of commercially available instrumentation, has limited their application and use.

### Capacitively coupled microwave plasmas

A capacitively coupled microwave plasma is formed using a magnetron to produce microwave energy at 2.45 GHz. This is brought to a hollow coaxial electrode via coaxial waveguides. The microwave power is capacitively coupled into the plasma gas, usually argon or nitrogen, via the electrode. This is an atmospheric pressure source with a small plasma volume. Power levels can range from 10 to 1000

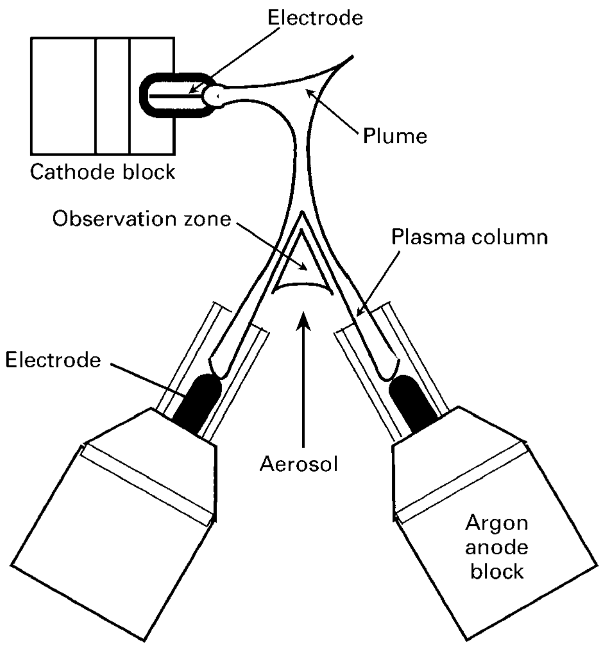



**Figure 2** Direct current plasma schematic.

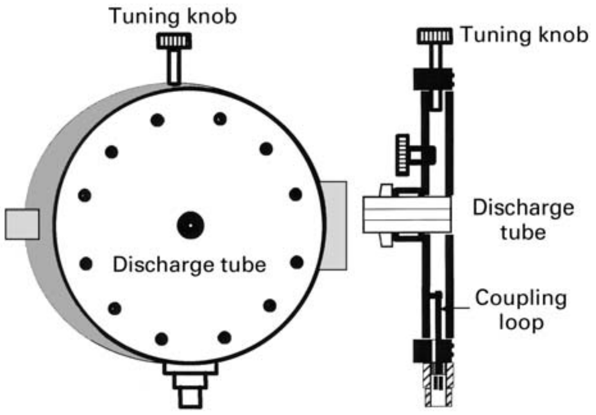



**Figure 3** Microwave induced plasma schematic illustrating the 'Beenakker Cavity' design.

watts, with frequencies from 200 kHz to 30 MHz. Unlike other MIPs, the CMP is able to handle greater amounts of solvent. However, the CMP has been found to be easily contaminated due to the microwave energy being conducted through a coaxial waveguide to the electrode, which must be replaced regularly.

## Spectrometers

The atomic emission source will produce unstable, excited atoms or ions, which spontaneously return to a lower energy state. The emission spectrum is produced when a photon of energy is generated during this transition. The basic assumption is that the emitted energy is proportional to the concentration of atoms or ions in the sample. The measurement of this energy is performed using the optics of the spectrometer to isolate the characteristic elemental emission wavelengths, and to separate this radiation from the plasma background. The spectrometer will need a high resolving power to be able to separate the lines of interest from adjacent spectral lines (at least 0.1 nm). The spectrometer consists of (1) a dispersive element, such as a grating, and (2) an image transfer assembly, which contains the entrance and exit slits, and mirrors or lenses.

The grating provides dispersion of the wavelength range of interest over a given angular range. Some commonly used grating spectrometers include (1) the Paschen–Runge spectrometer, which is used in both sequential and simultaneous instruments and has the advantage of extensive wavelength coverage, (2) the echelle grating spectrometer, with excellent dispersion and resolution with a small footprint, and (3) the Ebert and Czerny–Turner spectrometers, which are similar except that the latter has two mirrors to the Ebert's one.

The slit allows a narrow line of light to be isolated. The number of lines will represent the various wavelengths emitted by the plasma, with each line corresponding to the image produced by the spectrometer slit, and each wavelength corresponding to a specific element. There are two slits used: (1) the primary slit, which is where light enters the spectrometer, and (2) the secondary slit, where light exits, producing a line isolated from the rest of the spectrum. The imaging system consists either of lenses or of concave mirrors.

One of the most useful aspects of atomic emission spectrometry is its capability for multi-element analysis. This can be achieved using either the sequential monochromator, where elements in a sample are quickly read one at a time, or the polychromator, which allows the simultaneous measurement of the elements of interest.

### Monochromators

Multi-element determinations using a monochromator must be sequential, as the monochromator can only observe one line at a time due to it only having one secondary slit. The slit can be set to scan the wavelengths, which is slower than in a simultaneous instrument, but allows for the selection of a wider range of wavelengths. A scanning monochromator user a movable grating to find known individual spectral lines at fixed positions, typically a Czerny–Turner configuration. Although this allows the observance of only one individual wavelength at a time, it also allows all wavelengths in the range of the spectrometer to be observed.

### Polychromators

Polychromators have a permanently fixed secondary slit for certain individual wavelengths (individual elements). Each slit will have its own detector. This allows for truly simultaneous analyses of selected elements, resulting in much quicker analyses than the monochromator for elements that have a slit installed. However, the selection of possible analysis lines is limited. A simplified diagram of the beam path in a polychromator is shown in **Figure 4**. A polychromator can focus the emission lines on the circumference of the Rowland circle by using concave gratings; typically the Paschen–Runge configuration is used. In summary, the polychromator has a throughput advantage while the sequential scanning monochromator has flexibility as its advantage.

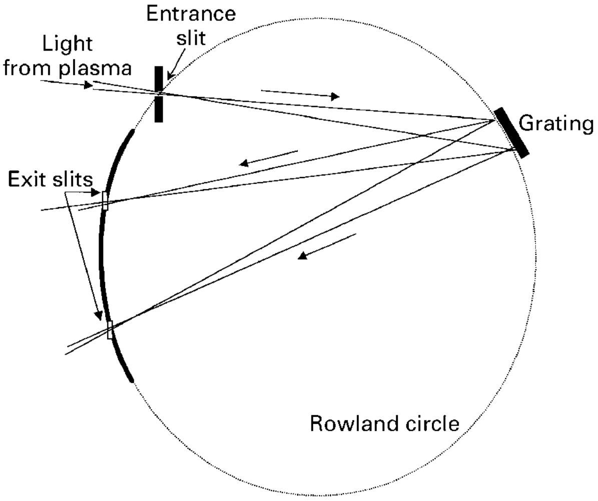



**Figure 4** Beam path through a polychromator.

## Detectors

### Photomultiplier tubes

A photomultiplier tube (PMT) consists of a photosensitive cathode, several dynodes and a collection anode. The dynodes are responsible for the increase in signal by electron multiplication. PMTs see the elemental line intensity per unit time proportionally with current and have wide dynamic ranges. A PMT, housed in a suitable mechanical movement, can scan the range of wavelengths in a spectrum sequentially, which involves longer analysis times. Alternatively, there can be a series of PMTs, each collecting the signal at discrete wavelengths at its assigned exit slit. Unfortunately, it always takes one PMT for each wavelength of light observed, whether or not the PMT is fixed or roving. Modern instruments have taken advantage of the advent of multichannel solid-state detectors to provide more flexibility in multi-element analyses.

### Charge-coupled and charge-injection devices

The charge-coupled and charge-injection devices are solid-state sensors with integrated silicon circuits. They are similar in that they both collect and store charge generated by the light from the emissions in metal-oxide semiconductor (MOS) capacitors. The amount of charge generated in a charge transfer detector is measured either by moving the charge from the detector element where it is collected to a charge-sensing amplifier (CCD), or by moving it within the detector element and measuring the voltage change induced by this movement (CID). The CCD is susceptible to 'blooming' in the presence of too much light, since there can be an overflow of charge from a full pixel to an adjacent one. CCDs are best used in very sensitive, low light level applications. The CID does not suffer from 'blooming' due to its method of nondestructively measuring the photon-generated charge. The CID allows the monitoring of any wavelength between 165 and 800 nm, whereas the range of the CCD is between 170 and 780 nm.

## Methods

### Sample introduction

Sample introduction into the plasma is a critical part of the analytical process in atomic emission spectroscopy (AES). Since the ICP is the most commonly used source, the sample introduction schemes described below will focus more on it than the other sources mentioned previously. Sample is carried into the plasma at the head of a torch by an inert gas, typically argon, flowing in the centre tube at 0.3–1.5 L min$^{-1}$. The sample may be an aerosol, a thermally or spark generated vapour, or a fine powder. Other approaches may also be taken to facilitate the way the analyte reaches the plasma. These procedures include hydride generation and electrothermal vaporization.

**Torches** One of the torch configurations more commonly used with the ICP is shown in **Figure 1**. The plasma torch consists of three concentric quartz tubes through which streams of argon flow. The nebulizer gas, which carries the analyte into the plasma, flows in the centre tube. The auxiliary gas flows around the centre tube and adjusts the position of the plasma relative to the torch. The coolant gas streams tangentially through the outer tube, serving to cool the inside walls and centre of the torch, and stabilize the plasma.

**Nebulizers** The great majority of analyses in ICP AES are carried out on liquid samples. The most convenient method for liquids to be introduced into the gas stream is as an aerosol from a nebulizer. The aerosol may be formed by the action of a high-speed jet across the tip of a small orifice or by the use of an ultrasonic transducer. A spray chamber is usually placed after the nebulizer to remove some of the larger droplets produced, and thereby improves the stability of the spectral emission.

The most commonly used nebulizer designs are pneumatic and ultrasonic, although other types (electrostatic, jet impact, and mono-dispersive generators) have been described. The selection of the appropriate nebulizer depends upon the characteristics of the

sample: mainly density, viscosity, organic content, total dissolved solids, and total sample volume. Additionally, the performance of a particular nebulizer can be described using several attributes such as droplet size distribution, efficiency, stability, response time, tendency to clog, and memory effects.

*Concentric nebulizers* Solution sample introduction has been associated with pneumatic nebulizers (concentric and cross-flow) almost universally for routine analysis, due to their simplicity and low cost. However, they provide low analyte transport efficiency (< 5%) to the plasma, and may be prone to clogging.

The most widely used ICP nebulizer is the one-piece Meinhard concentric nebulizer (**Figure 5A**). It is a general-purpose nebulizer with low tolerance for total dissolved solids, used for applications requiring low nebulizer gas flows. It operates in the free running mode whereby the solutions are drawn up by the pressure drop generated as the nebulizer gas passes through the orifice. The viscosity of the solution and the vertical distance through which the liquid is lifted affect the rate of liquid transfer. Although the concentric nebulizer is easy to use, the transport efficiency is low, owing to the wide range of droplet sizes produced. In addition, nebulizer blockage may occur due to the presence of suspended solids becoming lodged in the narrow, central sample uptake capillary. Filtering or centrifuging the sample may minimize the risk of blockage. If a sample with a high dissolved-solids content dries in the nebulizer, it may also cause nebulizer blockage or interfere with the operation of the nebulizer by decreasing the signal. Frequent cleaning of the nebulizer solves this problem.

*Cross-flow nebulizers* The cross-flow nebulizer shows much of the same general behaviour of the concentric nebulizers. The cross-flow nebulizer is less prone to blockage and salting effects, although they may still occur as the sample solution passes through a capillary. The cross-flow nebulizer operates when a horizontal jet of gas passes across the top of a vertical tube (**Figure 5B**). The reduced pressure that is generated draws the liquid up the tube, where, at the top, it is disrupted into a cloud of fine droplets.

*Babington type nebulizer* The Babington nebulizer is designed to allow a film of liquid containing sample to flow over the surface of a sphere (**Figure 5C**). A gas, which is forced through an aperture beneath the film, produces the aerosol. This design features the liquid sample flowing freely over a small aperture, rather than passing through a fine capillary, and is therefore more tolerant of high dissolved-solids. This kind of nebulizer can be used to introduce slurries into the system since the delivery of the sample is not constrained by a capillary. However, the Babington-type nebulizer shows extensive memory effects since the solution is allowed to wet the entire face of the sphere. A modification of the design is also in use, whereby the liquid sample passes through a V-groove and a gas is introduced from a small hole in the bottom of the groove. The main advantage of the V-groove nebulizer is its resistance to blockage. However, the design produces aerosol less efficiently, and it is of coarser size distribution than that produced by other concentric nebulizers.

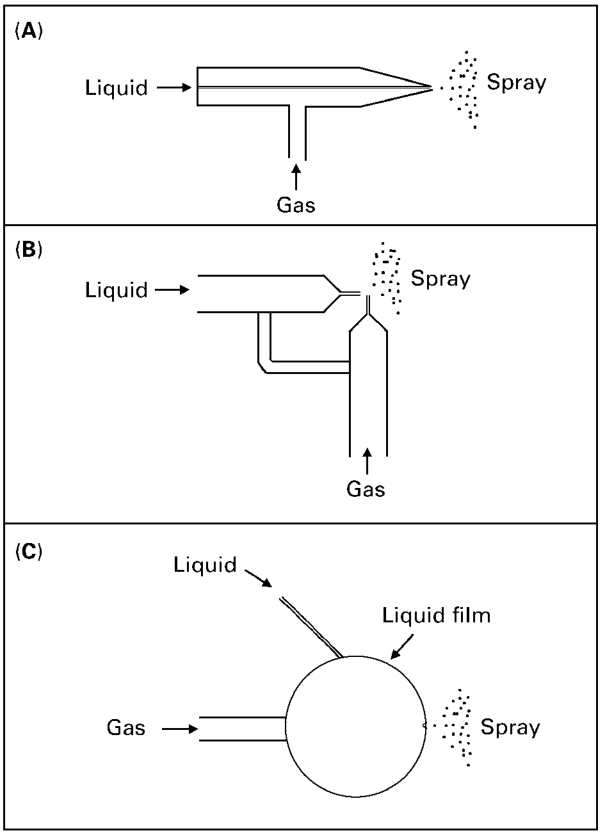



**Figure 5** Schematic of the (A) concentric nebulizer, (B) cross-flow nebulizer and (C) Babington nebulizer.

*Frit-type nebulizer* The concentric and cross-flow nebulizers are inefficient at the 1 mL $min^{-1}$ flow rate at which they usually operate, producing only about 1% of droplets of the correct size to pass into the plasma. An alternative design is the frit nebulizer, which produces droplets with a mean size of 1 μm. The glass frit nebulizer is depicted schematically in **Figure 6**. The sample flows over the fritted glass disc as the nebulizing gas is passed through the frit. This

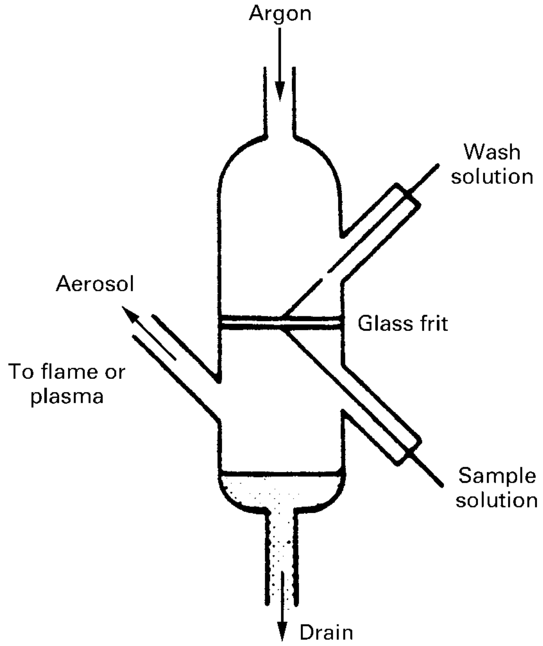



**Figure 6** Glass frit nebulizer schematic. Reprint permission from Ingle JD and Crouch SR (1988) *Spectrochemical Analysis*. p. 193, Figure 7-5. Englewood Cliffs, NJ: Prentice Hall.

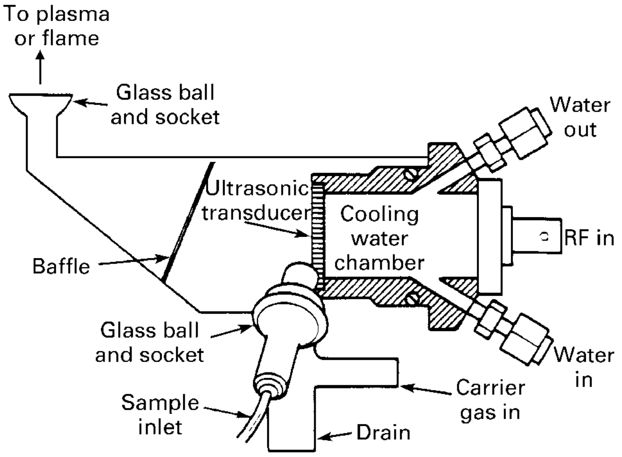



**Figure 7** Ultrasonic nebulizer schematic. Reprint permission from Ingle JD and Crouch SR (1988) *Spectrochemical Analysis*. p. 194, Figure 7-6. Englewood Cliffs, NJ: Prentice Hall.

design provides an excellent fine aerosol, but at the expense of the fritted disc clogging over time. The transport efficiency is very high compared to the pneumatic-type nebulizers but the operation is at flows of 100 μL $min^{-1}$ or less.

*Ultrasonic nebulizer* The ultrasonic nebulizer also produces fine aerosol and has a high sample-transport efficiency (**Figure** 7). The aerosol is generated when the solution is fed to the surface of a piezoelectric transducer which operates at a frequency between 0.2 and 10 MHz. Vibrations of the transducer cause the solution to break into small droplets which are then transported to the plasma. The ultrasonic nebulizer offers highly efficient aerosol production independent of the carrier-gas flow rate. With this nebulizer, more analyte can be transported into the plasma at a lower nebulizer gas flow rate than that seen with pneumatic nebulizers. This increases the sensitivity and improves detection limits since samples have a longer residence time in the plasma. However, the high efficiency also allows more water to enter the plasma, producing a cooling effect, which may decrease analyte ionization. Hence these nebulizers require desolvation to be realistically utilized. Apart from the analytical shortcomings of the ultrasonic nebulizer, the high cost of the commercial systems may be prohibitive.

*Electrothermal vaporization* Electrothermal vaporization can introduce liquid, solid, and gaseous (by trapping) samples to the plasma with high analyte transport efficiency (20–80%) and improved sensitivity over pneumatic nebulization. A small amount (5–100 μL) of sample is deposited into an electrically conductive vaporization cell (**Figure** 8). Initially, a low current is applied to the cell, which causes resistive heating to occur and dries the sample. A high current is then passed for a short time (typically 5 s) to completely vaporize the sample. An optional 'ashing' stage may be used to remove some of the matrix prior to the analyte vaporization stage. A stream of argon gas is passed through the unit and carries the sample vapour to the plasma. A variable current supply is required for controlling cell heating, and the sequence of heating steps is carried out using an electronic control system. Electrothermal vaporization allows sample matrix components, including the solvent, to be separated from the analytes of interest through judicious selection of temperature programming steps. This may reduce oxide formation and the number of spectroscopic interferences. Additionally, samples with high salt content and limited volume can be analysed with electrothermal vaporization.

*Hydride generation* The problems associated with liquid sample introduction, and the inefficiencies in analyte transport can be overcome by presenting the sample in a gaseous form to the plasma. Introducing the analyte in a gaseous form eliminates the use of a nebulizer and a spray chamber. It also provides nearly 100% analyte transport efficiency, which subsequently improves the detection limits. Additional

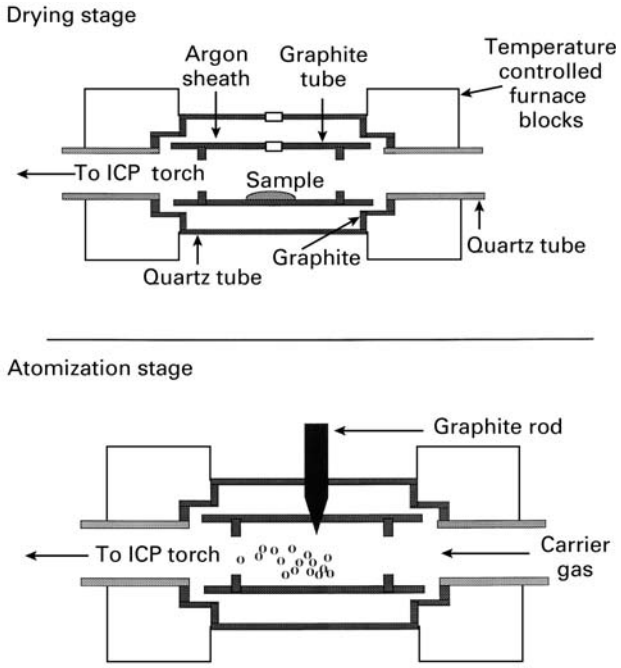



**Figure 8** Illustration of electrothermal vaporization graphite furnace.

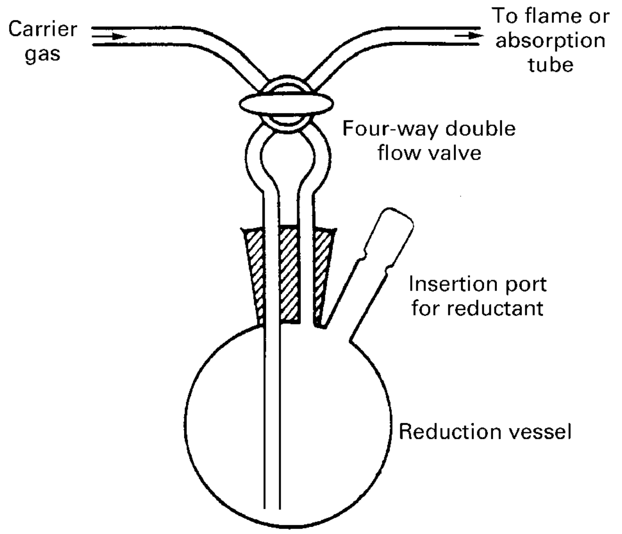



**Figure 9** Hydride generation apparatus. Reprint permission from Ingle JD and Crouch SR (1988) *Spectrochemical Analysis.* p. 280, Figure 10-6. Englewood Cliffs, NJ: Prentice Hall.

benefits include no blockage problems, the possibility of matrix separation and analyte preconcentration, and the absence of water, which reduces the levels of many polyatomic ion species. The hydrides of the elements arsenic, bismuth, germanium, lead, antimony, selenium, and tin are easily formed in acidic solutions and are gaseous at room temperature. The most frequently used reaction is:

$$NaBH_4 + 3H_2O + HCl \rightarrow H_3BO_3 + NaCl + 8H \rightarrow EH_n + H_2$$

where E is the hydride forming element of interest. **Figure 9** shows an apparatus for hydride generation. It allows the mixing of the sample and the reducing agent, followed by the transport of the volatile analyte species (hydride) by the carrier gas to the plasma.

The gaseous sample introduction line is connected directly to the central tube of the plasma torch, eliminating the need for the conventional nebulizer and spray chamber. Additional equipment to handle gas mixing and dilution at controlled flow rates may be required.

*Direct solids introduction* Direct insertion of solid samples into the ICP can be used for metal powders, salts, and geological samples. The sample is placed on a wire loop or cup made from graphite, molybdenum, tantalum, or tungsten. The probe is then moved along the axis of the torch and closer to the plasma where drying takes place. At the completion of this stage, the device is propelled rapidly into the core of the plasma and measurements of the analytes can be taken.

### Line selection

The choice of which line to use for a given sample type is a difficult one, as most elements have many lines available for analysis. Not only should the intensity of the line be considered, but also whether the line is free of spectral interferences from both the plasma and other sample constituents. Additionally, the nature of the emission, atomic or ionic, should be considered. The emission could originate from an excited neutral atom, which is termed an atomic transition, or an excited ionized form of the element, which is called an ionic transition. While analysis of elements undergoing transitions of the ionic type tend to experience fewer consequences due to changes in the plasma operating conditions, there are elements that produce no ionic lines, such as aluminium and boron. There are very good reference books which list tables of spectral lines and are included in the Further reading section. Often modern instrumentation has sufficiently 'intelligent' software to assist the operator with line selection.

### Spectral interferences and correction techniques

A spectrometer with good resolution will help greatly in the separation of adjacent lines from the

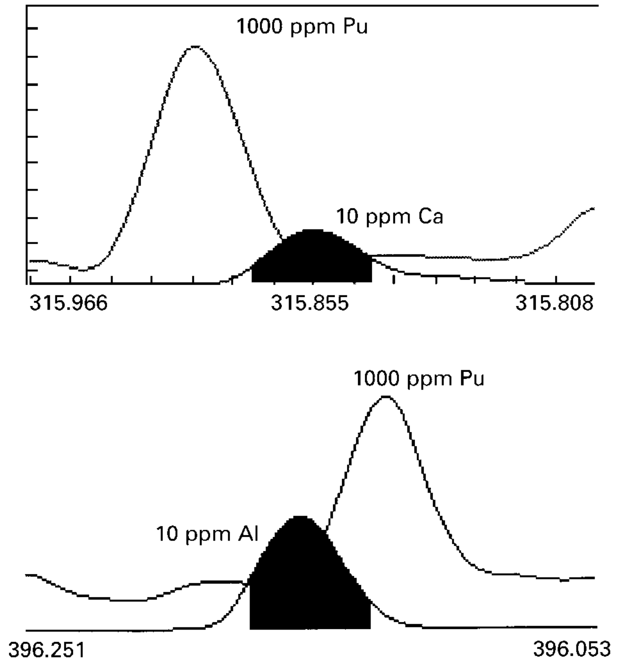



**Figure 10** Interelement effects of plutonium on calcium (top) and aluminium (bottom).

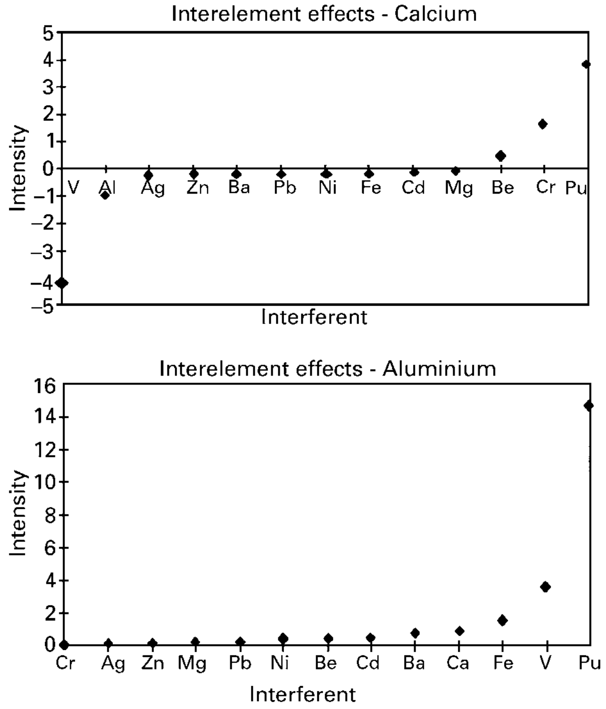



**Figure 11** Interelement equivalent concentration correction factors for several interfering elements on calcium (top) and aluminium (bottom) as calculated by a program written in-house by Gerth DJ.

spectral line of interest. The effect of partial overlap can be minimized in some cases with a high-resolution spectrometer, and in other cases, can be overcome using correction techniques.

When the interference is from the plasma emission background, there are background correction options available with most commercial instrumentation. The region adjacent to the line of interest can be monitored and subtracted from the overall intensity of the line. If direct spectral overlap is present, and there are no alternative suitable lines, the interelement equivalent concentration (IEC) correction technique can be employed. This is the intensity observed at an analyte wavelength in the presence of 1000 mg $L^{-1}$ of an interfering species. It is expressed mathematically as:

$$\mathrm{IEC} = (I_{\mathrm{init}}/I_{\mathrm{a}}) \times C_{\mathrm{a}}$$

where the correction is in milligrams of analyte per litre of solution, and $I_{\mathrm{init}}$ is the intensity read at the analyte wavelength in the presence of 1000 mg $L^{-1}$ of the interfering species, $I_{\mathrm{a}}$ is the intensity the instrument will produce for a certain analyte concentration at the analyte wavelength, and $C_{\mathrm{a}}$ is the concentration in mg $L^{-1}$ used to give $I_{\mathrm{a}}$. This correction can be applied to all the analytes in a method and some instrument manufacturers' software will automatically perform the necessary corrections. Examples of the effect of 1000 ppm plutonium as the interfering element on calcium and aluminium are shown in **Figure 10**. This shows the subarrays obtained for separate solutions of plutonium, calcium, and aluminium. Graphical representations of interfering elements on calcium and aluminium are shown in **Figure 11**. These graphs are produced by a program written at Los Alamos National Laboratory by DJ Gerth, and illustrate which interfering elements have the strongest effect on the analytes. Those elements with the strongest deviation from the average are those needing the interelement correction.

## List of symbols

$C_{\mathrm{a}}$ = analyte concentration used to give $I_{\mathrm{a}}$; $I_{\mathrm{a}}$ = intensity at certain analyte concentration and wavelength; $I_{\mathrm{init}}$ = intensity in presence of 1000 mg $L^{-1}$ of the interfering species.

*See also:* **Electronic Components, Applications of Atomic Spectroscopy; Fluorescence and Emission Spectroscopy, Theory; Inductively Coupled Plasma Mass Spectrometry, Methods.**

## Further reading

Boumans PWJM (1987) *Inductively Coupled Plasma Emission Spectroscopy – Part 1, Chemical Analysis Series*, Vol. 90. Chichester: Wiley.

Edelson MC, DeKalb EL, Winge RK and Fassel VA (1986) *Atlas of Atomic Spectral Lines of Plutonium Emitted by an Inductively Coupled Plasma*. Ames Laboratory, US Department of Energy.

Ingle JD and Crouch SR (1988) *Spectrochemical Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Montaser A and Golightly DW (eds) (1992) *Inductively Coupled Plasmas in Analytical Atomic Spectrometry* Weinheim: VCH.

Slickers K (1993) *Automatic Atomic Emission Spectroscopy*. Brühlsche Universitätsdruckerei.

Sneddon J (ed) (1990) *Sample Introduction in Atomic Spectroscopy, Analytical Spectroscopy Library, #4*. Amsterdam: Elsevier.

Sweedler JV, Ratzlaff KL and Denton MB (eds) (1994) *Charge-Transfer Devices in Spectroscopy*. Weinheim: VCH.

Thompson M and Walsh JN (1989) *Handbook of Inductively Coupled Plasma Spectrometry*. Glasgow: Blackie.

Varma A (1991) *Handbook of Inductively Coupled Plasma Atomic Emission Spectroscopy*. Boca Raton, FL: CRC Press.

Winge RK, Fassel VA, Peterson VJ and Floyd MA (1985) *Inductively Coupled Plasma-Atomic Emission Spectroscopy, An Atlas of Spectral Information*. Amsterdam: Elsevier.

# Atomic Fluorescence, Methods and Instrumentation

**Steve J Hill** and **Andy S Fisher**, University of Plymouth, UK

Copyright © 1999 Academic Press

**ATOMIC SPECTROSCOPY**
**Methods & Instrumentation**

## Introduction

Atomic fluorescence spectroscopy (AFS) has been used for elemental analysis for several decades. It has better sensitivity than many atomic absorption techniques and offers a substantially longer linear range. However, despite these advantages, it has not gained the widespread usage of atomic absorption or emission techniques. In recent years, the use of AFS has been boosted by the production of specialist equipment that is capable of determining individual analytes at very low concentrations (at the ng $L^{-1}$ level). The analytes have tended to be introduced in a gaseous form and hence sample transport efficiency to the atom cell is very high. This article describes the instrumentation and methods available for atomic fluorescence spectroscopy, although it should be emphasized that much of the instrumentation associated with this technique is often very similar to that used for atomic absorption spectroscopy (AAS). A schematic diagram of the different parts of an AFS instrumentation is shown in **Figure 1**. It can be seen that the light source, atom cell, line isolation device, detector and readout system used for AFS are very similar to those used in AAS, although there may be subtle differences and the components may be less sophisticated, as described below.

## Light sources

As can be seen from **Figure 1**, light sources for AFS are placed at right angles to the detector. This is to ensure that incidental radiation from the source is not detected. Since a more intense source will lead to greater sensitivity, in AFS a standard hollow-cathode lamp (HCL) is often insufficient for the majority of applications. A range of alternatives has therefore been developed. An electrodeless discharge lamp (EDL) is suitable because it has an intensity 200–2000 times greater than that of a hollow-cathode lamp. Boosted-discharge hollow-cathode lamps (BDHCL)

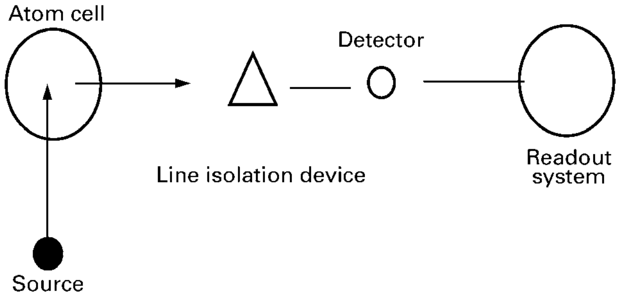



**Figure 1** A schematic representation of an AFS instrument.

are similar to standard hollow-cathode lamps but can be operated at greatly increased current. The light output is consequently several times more intense. A similar effect may be obtained by pulsing a standard hollow-cathode lamp to 100–1000 mA, although this will reduce the lifetime of the lamp. Continuum sources such as 150–500 W xenon lamps have also been used, but there are several problems associated with their use. The most obvious drawback is the relatively poor intensity offered at each individual wavelength by the black body irradiator. The problem is particularly bad in the ultraviolet region of the spectrum. Scatter has also been reported as being problematic.

Vapour discharge lamps have been available for some elements (e.g. Cd, Ga, Hg, In, K, Na and Tl) since the inception of commercial instrumentation, although such devices are limited to the volatile elements and are therefore not universal.

Lasers have the greatest output intensity and would therefore be expected to yield the highest sensitivity. Unfortunately, they are expensive and have, in the past, been produced to give light of very specific wavelengths, which may not correspond to the frequency required to excite the analyte atoms. The development of the tunable dye laser has helped to overcome this problem. Frequency doubling (i.e., halving the wavelength of the light output) has enabled wavelengths well into the ultraviolet region to be obtained. A nitrogen laser pump can be used to obtain a wavelength of 220 nm while a Nd:YAG (neodymium:yttrium aluminium garnet) laser allows 180 nm to be reached. An added advantage of high-intensity lasers is that they may cause saturation fluorescence (population inversion), which nullifies the effects of quenching and self-absorption. Another source that has been used to initiate fluorescence is an inductively coupled plasma. This may be used either as a continuum source (if the actual fireball is used), or as a line source (if the tail-flame is used). The use of an ICP as a light source to initiate fluorescence is not yet a routine technique.

As with AAS, modulation of the light source enables an optimal signal-to-noise ratio to be obtained. The frequency of modulation depends on the light source, but may be up to 10 kHz for an EDL.

## Atom cell

Several atom cells have been used for AFS. The majority of applications have used a circular flame atom cell (sometimes with a mirror placed to collect as much light as possible), but more recently inductively coupled plasmas and electrothermal atomizers have also been used. There is a risk of self-absorption at high analyte concentrations, i.e. light fluoresced by some atoms will be absorbed by others in different parts of the flame. A circular flame has a smaller path length compared with a slot burner, hence its popularity. A variety of fuels have been used for the flames. Air–acetylene, nitrous oxide–acetylene and argon–hydrogen diffusion flames have all been used successfully. Quenching (the nonradiative loss of energy) increases with increasing flame temperature and so the cooler air–acetylene and hydrogen flames are often used. Conversely, the hot nitrous oxide–acetylene flame aids the atomization of refractory compounds, The quenching crosssection of the colliding particles also effects the amount of quenching. The relative quenching cross-section of the different flame gases increases in the order argon (negligible) < hydrogen (low) < oxygen (high). It can therefore be seen that for easily atomized analytes an argon–hydrogen diffusion flame offers at the best properties for atomic fluorescence.

Inductively coupled plasmas atomize refractory analytes very efficiently because of their high temperature (8000 K) and therefore have a very high fluorescent yield. Care must be taken, however, to ensure that ionization of the analyte species does not occur. An added advantage is that the plasma is made up of argon and is therefore optically thin (i.e. it has a low quenching cross-section). The disadvantages of using a plasma as the atom cell are that it is more expensive and the instrumentation is more complex. When a plasma is used as an atom cell, the light source may be placed in one of two orientations. It may pass light through the side of the plasma although more recently it has been used axially, i.e. the source passes light through the entire length of the plasma. Since the second configuration has a longer path length, more analyte atoms may be excited and hence improved detection limits are obtained. The disadvantage with this configuration is that the light source must be protected in some way from the extreme temperatures of the plasma. This may be achieved by using either a water-cooled condenser containing a quartz window or by using a protective gas flow to protect the lens from the plasma tail flame. If mirrors or some other collection device are placed around the plasma, then all fluoresced light may be collected and directed to the detector. It has been noted that the optimum viewing height for fluorescence measurements using an ICP as the atom cell occurs substantially higher above the load coil than for emission measurements. Observation above 3 cm is routine and making measurements from the tail flame is common to ensure that a signal above the background and analyte emission noise is obtained.

Similarly, a decrease in plasma power has also been found to be beneficial.

Electrothermal atomizers for AFS share many of the advantages associated with their use for AAS. The argon inert gas that prevents oxidation of the graphite tube ensures that minimal quenching occurs, although a number of other interferences may occur. By and large, the interferences are very similar to those experienced in traditional electrothermal AAS, such as carbide formation (for some analytes), scatter by particulate matter, and losses during thermal pretreatment. A more comprehensive overview of interferences may be found in the sections on AAS. The use of matrix modifiers to assist in the separation of the analyte from the matrix is still often necessary.

Where a laser is used as light source, a technique termed laser-induced fluorescence electrothermal atomization (LIF-ETA) is possible. A schematic diagram of a typical LIF-ETA system is shown in **Figure 2A**. The light source passes through a pierced mirror and into the atomizer. Light fluoresced from the analytes within the atomizer then reflect off the mirror and onto the detector. An alternative configuration is shown in **Figure 2B**, in which the laser irradiates the sample through the injection port. The two configurations are known as longitudinal and transverse geometries, respectively. The technique is capable of extremely low detection limits (below 1 pg $g^{-1}$) and has therefore found use in the nuclear, biomedical, electronics and semiconductor industries.

A heated atom cell in not required in the case of mercury determinations. Mercury produces an atomic vapour at room temperature and therefore does not require a flame to cause atomization. Instead, the mercury vapour may simply be transported to an atom cell at room temperature. The radiation from the light source excites the mercury atoms in the normal way.

## Line isolation devices

A conventional monochromator (either Ebert, Czerny–Turner, Littrow or Echelle) may be used (see AAS sections for details), but some of the more basic instrumentation uses interference filters. These are optical filters that remove large bands of radiation in a nondispersive way. A dispersion element such as a prism or a grating is therefore not required. Only a relatively narrow band of radiation is allowed to pass to the detector. The disadvantage with such devices is that they are not particularly efficient and hence much of the fluoresced light is lost. An alternative development is the multi-reflectance filter. This is shown diagramatically in **Figure 3**, and has the

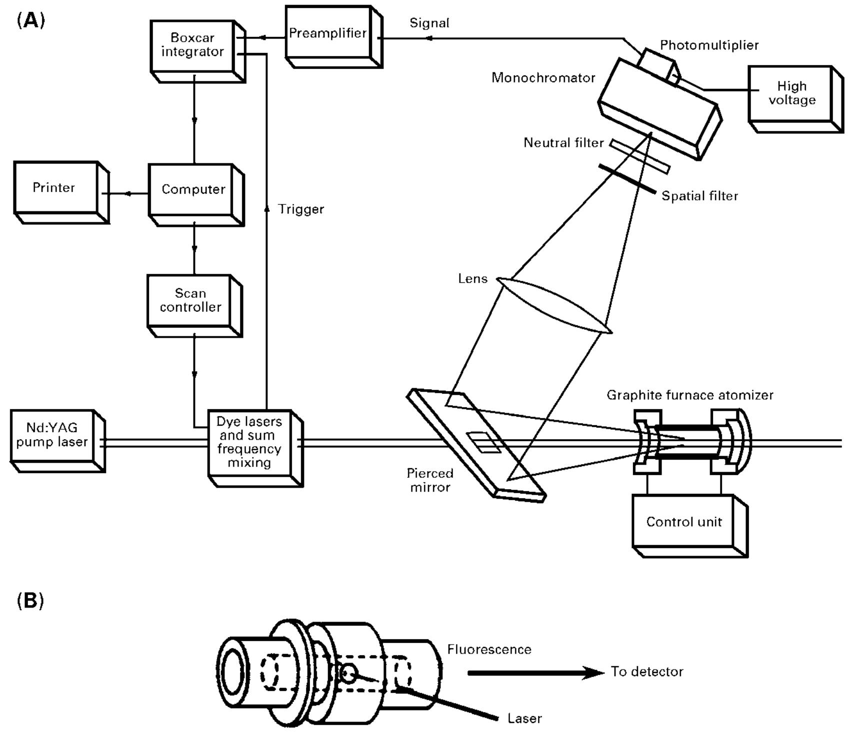



**Figure 2** (A) Schematic diagram of longitudinal laser-induced fluorescence with electrothermal atomization. (B) Transverse laser-induced fluorescence. From Ebdon L (1998) *Introduction to Analytical Atomic Spectrometry*. Reproduced by permission of John Wiley & Sons Limited.

advantages of being 80% efficient at transmitting the wavelengths of interest while virtually eliminating background noise. Such a device therefore has a superior performance to other wavelength filters.

Interference and multi-reflectance filters may be used when the analyte of interest has been separated from the matrix and all the concomitant elements. In such a case the analyte atoms are the only elements that will fluoresce when light from a monochromatic light source is used to excite them. Powerful dispersion gratings are therefore not required because, theoretically, light specific to only one element is produced. A schematic diagram of a very basic instrument that uses a multi-reflectance filter is shown in **Figure 4**. If other analytes are present in the atom cell, then either a monochromator is required to separate the other wavelengths (produced by emission), or, the source must be modulated so that distinction between emission and fluorescence signals may be made.

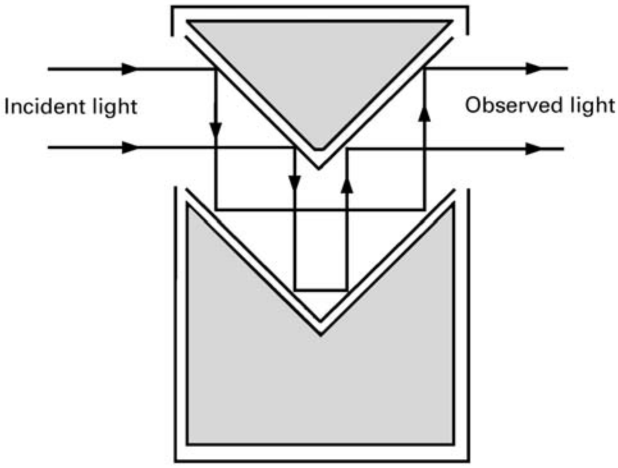



**Figure 3** A multi-reflectance filter.

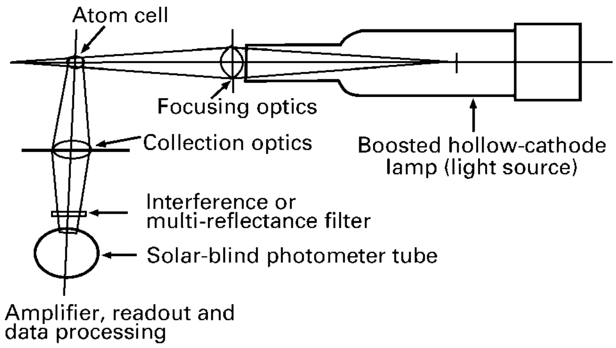



**Figure 4** Schematic of a very basic nondispersive AF spectrometer.

## Detectors

The light isolated by the line isolation device (monochromator, interference filter or multi-reflectance filter) is normally detected using a photomultiplier tube (PMT). A description of how these devices work may be found in the sections on AAS. It should be noted that some basic instruments (e.g. as shown in **Figure 4**), often use a solar-blind photomultiplier tube. This is a PMT that detects radiation in the UV region but does not detect light in the visible region of the spectrum.

## Readout systems and data handling

The readout and data-handling systems used for AFS are similar to those found for AAS instrumentation. The signal from the photomultiplier tube passes through an amplifier and then on to the readout device. This may be a digital display, a chart recorder, an integrator, or a computer with software dedicated to the instrument. The last alternative is the most commonly used in modern instrumentation. The software is likely to be able to control calibration, plot graphs, perform quality assurance tests and calculate means and standard deviations of the results. One advantage of using a computer to control the analysis is that the use of an autosampler becomes a possibility. This enables unattended operation of the instrumentation.

## Applications

As described previously, although techniques to utilize AFS were developed several decades ago, until recently they were not widely used. However, a new generation of simple AFS instruments have been developed to specifically detect the vapour-forming elements, such as those that form hydrides (As and Se) and mercury, which forms an atomic vapour. All of these analytes have a primary line below 260 nm and, since the analytes may be readily separated from the bulk matrix and concomitant elements, dispersion is not necessary. Instead these basic instruments originally used a simple interference filter, although these have now been superseded by the more efficient multi-reflectance filter.

The theory behind hydride generation is beyond the scope of this text, but in brief, if a sample is mixed with a reducing agent, such as sodium tetrahydroborate, then some elements, e.g. As, Bi, Ge, Pb, Se, Sb, Sn and Te, will form hydrides. It must be noted that the efficiency of hydride formation depends on the oxidation state of the analyte. Arsenic(III)

forms a hydride more efficiently than As(V), and Se(IV) forms a hydride whereas Se(VI) does not. It is therefore necessary to ensure that all the analyte atoms are in the same oxidation state; this is normally achieved using one of a variety of reducing agents.

Once formed, the hydride leaves the reaction cell and is flushed towards the AFS instrument using an inert gas (usually argon). During this process, hydrogen is formed as a breakdown product of tetrahydroborate, and can be swept along with the hydride in the argon to the atom cell, where it can be used as the fuel for the hydrogen–argon diffusion flame. More hydrogen may be added as necessary from an external source. Since a hydrogen–argon diffusion flame is used, the amount of water entering the flame must be minimal (otherwise quenching and flame instability result). The hydride and water vapour are therefore separated, often by passing through a membrane drier tube. This is a tube containing a semi-permeable hydroscopic membrane surrounded by a counterflow of dry gas. Water vapour in the hydride passes through the membrane and is removed by the countergas flow. The dry hydride continues through the tube to be flushed toward the AFS instrument by the flow of argon.

Mercury may be determined using very similar instrumentation, but it may also be introduced in a number of other ways. Since mercury is volatile, solid samples may be placed on a small platform and then heated. The sample is thus vaporized and the vapours are passed through a trap (sand coated with gold). The mercury is trapped onto the gold while the smoke and matrix constituents pass through and are hence separated from the analyte. Heating of the gold trap to 900°C releases the mercury, which may then be swept in a flow of argon to the AFS instrument. This technique has the advantage that the mercury may be preconcentrated on the trap, since several repeat samples may be vaporized and the mercury collected before it is subsequently removed from the trap. The use of the gold trap preconcentrator is equally applicable to the analysis of liquid samples. When a gold trap preconcentrator is used, the limit of detection for mercury is substantially below 1 ng $L^{-1}$, which is superior to that obtained using atomic absorption, emission or inductively coupled plasma–mass spectrometric (ICP-MS) techniques.

The instrumentation described above may also be used for speciation studies. In such cases a separation technique such as high-performance liquid chromatography (HPLC) or gas chromatography (GC) is coupled with the AFS instrument. The chromatography is used to separate different analyte species and some on-line chemistry is subsequently required to convert the species into a chemical form suitable for vapour generation. Although the chemistry behind these methods of speciation is beyond the scope of this text, it should be emphasized that the method used should be powerful enough to convert even very stable species, such as arsenobetaine or selenomethionine into atomic arsenic and selenium. The additional use of photolysis may therefore be necessary. Similarly, pyrolysis may be necessary to convert different mercury species into atomic mercury since this is the only species that will fluoresce.

As described previously, vapour introduction approaches are by far the most common application of atomic fluorescence. Despite this, mention of other methods should be made. If a conventional nebulizer and spray chamber assembly (see AAS section) is used, it is possible to introduce liquid samples directly to the atom cell. In circumstances such as these, it is necessary to use more robust air–acetylene or nitrous oxide–acetylene flame, or perhaps an ICP. The use of an ICP as an atom cell for AFS measurements has led to the development of a number of different techniques, e.g. ASIA, an acronym for atomiser, source, inductively coupled plasmas in AFS. This technique uses a high-powered ICP as a source and a low-powered ICP for the atom cell. It has been found that ICP-AFS yields linear calibrations over 4–6 orders of magnitude and is more sensitive than ICP-AES.

Several techniques have been developed that use an electrothermal atomizer as an atom cell. Laser-induced fluorescence (LIF) has been discussed previously, but other techniques such as laser-excited atomic fluorescence spectroscopy (LEAFS) also exist, although not used routinely.

## Interferences

Atomic fluorescence has the advantage of being less prone to spectral interferences than either AES or AAS. Molecular fluorescence is less of a problem than molecular absorption is in AAS. Scatter from the light source and quenching from the gaseous species in the atom cell are often the major sources of interference. For many applications where the analyte is separated from the matrix (e.g. vapour generation) chemical interferences may exist; for example, the presence of high concentrations of some transition metals may interfere in the hydride formation process. This will inevitably lead to errors in the measurement unless preventative steps are taken.

## Detection limits

A discussion of the detection limits for the different approaches utilizing atomic fluorescence is difficult

because they depend so heavily on the type of source used. Similarly, samples introduced in the vapour phase tend to yield substantially improved limits of detection (LODs) when compared with those introduced as a liquid. In general, if an EDL is used as a source and a circular flame as an atom cell, LODs at the $\mu g\ L^{-1}$ level are obtainable. Improvement will be made if a laser is used as a source. If the sample is introduced as a vapour, LODs at the $ng\ L^{-1}$ level are obtained. The use of an ICP as an atom cell leads to LODs at the $\mu g\ L^{-1}$ level if a normal HCL is used as the source, improving by approximately an order of magnitude if a boosted discharge HCL is used. If an electrothermal atomizer is used in conjunction with a laser, the best LODs of all are obtained.

## Conclusions

Atomic fluorescence spectroscopy is a popular technique for those analytes that readily form vapours, and specialized instrumentation is now available for individual elements. Such instruments are simple to operate, easily automated, and offer good sensitivity and freedom from interferences. The use of other flame-based fluorescence techniques has waned considerably over the years, but research continues into the use of electrothermal atomizers and inductively coupled plasmas as atom cells, reflecting the superior limits of detection that are potentially available. The availability of tuneable lasers may also encourage the resurgence of AFS as a routine analytical tool.

*See also:* **Atomic Absorption, Methods and Instrumentation; Light Sources and Optics; UV-Visible Absorption and Fluorescence Spectrometers.**

## Further reading

Greenfield S (1994) Inductively coupled plasmas in atomic fluorescence spectrometry. A review. *Journal of Analytical Atomic Spectrometry* **9**: 565–592.

Greenfield S (1995) Atomic fluorescence spectrometry — progress and future prospects. *Trends in Analytical Chemistry* **14**: 435–442.

Hill SJ, Dawson JB, Price WJ, Shuttler IL, Smith CMM and Tyson JF (1998) Advances in atomic absorption and fluorescence spectrometry and related techniques. *Journal of Analytical Atomic Spectrometry* **13**: 131R–170R.

Kirkbright GF and Sargent M (1974) *Atomic Absorption and Fluorescence Spectroscopy*. London: Academic Press.

# Atomic Force Microscopes

*See* **Scanning Probe Microscopes.**

# Atomic Force Microscopy, Theory

*See* **Scanning Probe Microscopy, Theory.**

# Atomic Spectroscopy, Historical Perspective

**CL Chakrabarti**, Carleton University, Ottawa, Ontario, Canada

Copyright © 1999 Academic Press

**ATOMIC SPECTROSCOPY**
**Historical Overview**

Professor B. V. L'vov, the inventor of the very powerful analytical technique graphite furnace atomic absorption spectroscopy, has rightly pointed out in the Introduction of his definitive book entitled *Atomic Absorption Spectrochemical Analysis* that, 'The discovery of atomic absorption and the history of research into it are integral parts of the entire history of spectroscopy and spectrochemical analysis'. Indeed, the early history of atomic spectroscopy, as far as spectrochemical analysis was concerned, consisted of the development of emission spectrochemical analysis, which was usually dependent on Fraunhofer lines (which are atomic absorption lines) for wavelength calibration.

Following Newton's study of the spectrum of the sun in 1666, there was a period of almost 136 years entirely concerned with emission spectroscopy. Only in 1802 did Wollaston report the presence of dark bands in the continuum emission spectrum of the sun and, after a more detailed study by Fraunhofer (1814), Brewster (1820) was able to ascribe them to absorption of radiation within the sun's atmosphere. Another 40 years passed before Kirchhoff and Bunsen showed that one of these dark bands in the emission spectrum of the sun corresponded exactly to the yellow emission band obtained when sodium vapour is heated in a flame. This led Kirchhoff to deduce that the Fraunhofer lines in the solar spectrum are absorption lines of elements whose flame emission spectra would contain lines at exactly the same position in the spectrum. Their work enabled Kirchhoff to develop the fundamental relationship between emission and absorption spectra: any species that can be excited to emit radiation at a particular wavelength will also absorb radiation at that wavelength. Thus Kirchhoff not only laid the foundations of atomic absorption methods of chemical analysis but also gave a striking example of their power. Indeed, it is difficult to imagine a more convincing and dramatic demonstration. Bunsen and Kirchhoff are thus rightfully considered to be founders of spectrochemical analysis. However, it is surprising that almost a century after the work of Bunsen and Kirchhoff the potentialities of atomic absorption measurements remained unexplored and unsuspected. Why?

As Alan Walsh has presented in his perceptive analysis of the reasons, it seems likely that one reason for neglecting atomic absorption methods was that Bunsen and Kirchhoff's work was restricted to visual observation of the spectra. In such visual methods the sensitivity of emission methods was probably better than that of absorption. Not only was the photographic recording of the spectra more tedious, but also the theory seemed to indicate they would only prove useful for quantitative analysis if observed under very high resolution. But possibly a more fundamental reason for neglecting atomic absorption is related to Kirchhoff's law (1859), which states that the ratio of the emissive power $E$ and absorptive power $A$ of a body depends only on the temperature of the body and not on its nature. Otherwise radiative equilibrium could not exist within a cavity containing substances of different kinds. The law is usually expressed as

$$E/A = K(\lambda T)$$

where $K(\lambda T)$ is a function of wavelength and temperature.

This law is perfectly correct but unfortunately much and possibly most of the subsequent teaching concerning it has been misleading. In most textbooks and lessons on this subject the enunciation of the law is followed by a statement to the effect that 'this means good radiators are good absorbers, poor radiators are poor absorbers'. This statement is patently absurd without any reference to temperature or wavelength.

This erroneous conclusion has been made, presumably, because Kirchhoff's constant, $K$, as he so clearly pointed out, is a function of wavelength and temperature. He was unable to find an analytical expression for it and it was not deduced until 1900 by Kirchhoff's successor at the University of Berlin, Max Planck, as Planck's distribution function. Many spectroscopists were misled by the widely held but misleading assumptions regarding the implications of Kirchhoff's law. What is even more surprising is that numerous spectroscopists wrote papers on atomic

emission methods and referred to the problems caused by the effects of self-absorption and self-reversal, and yet failed to make the small-step connection between atomic emission and atomic absorption! It was left to Alan Walsh, whose research experience in atomic emission spectroscopy and molecular absorption spectroscopy virtually compelled him to see the obvious connection. Walsh expressed this experience in his inimitable way in the following words: 'It appears to be true that "having an idea" is not necessarily the result of some mental leap: it is often the result of merely being able, for one sublime moment, to avoid being stupid!'.

Walsh in 1953 and Alkemade and Milatz in 1955 independently published papers indicating the substantial advantage of atomic absorption methods over emission methods for quantitative spectrochemical analysis. Alkemade and Milatz considered only the selectivity in atomic absorption methods, but Walsh discussed general problems of development of absolute methods of analysis. During his early experiments with flame atomizer–burners, Walsh encountered the problems which arise when measuring atomic absorption using a continuum source, which might have been responsible for the neglect of atomic absorption methods. The use of a sharp-line source such as a hollow-cathode lamp not only solved this problem, it also provided the atomic absorption method with one of its important advantages, that is, the ease and certainty with which one can isolate the analysis line. The concept of 'putting the resolution in the source' also permitted the use of a simple monochromator since the function of the monochromator, in such a case, is only to isolate the analysis line from the neighbouring lines and the background.

Graphite furnace atomic absorption spectroscopy (GFAAS) has been widely used as a spectrochemical trace-analytical technique during the last 30 years. L'vov pioneered most of the theoretical and experimental developments in GFAAS and provided a masterly treatment of the possibilities of GFAAS in a paper entitled 'Electrothermal atomization—the way towards absolute methods of atomic absorption analysis'. RE Sturgeon, who has been responsible for much of the development of GFAAS, presented a critical analysis of what it does and what it cannot do. AKh Gilmutdinov, who did much of the theoretical development of GFAAS in the 1990s, also elucidated the complex processes and reactions that occur in an electrothermal atomizer by digital imaging of atomization processes using a charge-coupled device camera in my research laboratories.

Atomic spectroscopy in the three variations that are most commonly used in spectrochemical analysis, atomic absorption, atomic emission and atomic fluorescence, are all mature techniques, with their particular areas of strengths and weaknesses now well recognized. Many a battle has been fought and won to establish the superiority of one or the other technique over the rival techniques. It is sometimes claimed that Inductively Coupled Plasma mass spectrometry (ICP-MS), which has the winning combination of high-temperature ionization of elements in a plasma with the detection of the ions by mass spectrometry, is the ultimate trace-analytical technique that will triumph over the rival techniques. However, such claims are based on the limited applications of these techniques by practitioners whose allegiance to their own techniques gives more credit to their loyalty than to their scientific objectivity, as elaborated in the following paragraph. Some have also predicted the imminent demise of graphite furnace atomic absorption spectroscopy as a trace-analytical technique. Such a prediction, based as it is on inadequate comprehension of the enormous complexities and extreme diversities of real-life situations requiring a variety of trace-analytical techniques which possesses some special capabilities not possessed by many analytical techniques, is destined to be false, as is shown below.

In my research laboratories in recent years, my 10 PhD students and three adjunct research professors have been doing research on the chemical speciation of potentially toxic elements in the aquatic, the atmospheric and the terrestrial environment. Because freshwaters and soils are systems that are usually far removed from equilibrium, our research is mostly directed to the development of kinetic schemes of chemical speciation which require routine use of inductively coupled plasma mass spectrometry and graphite furnace atomic absorption spectroscopy to measure the kinetics of metal uptake from aqueous environmental samples. As such, we routinely make hundreds of determinations of trace and ultratrace elements every week. For these determinations, we have in our research laboratories the latest models of two ICP-MS and several GFAAS systems. We have made objective studies of the performance characteristics of ICP-MS and GFAAS for kinetic studies of real-life, aqueous, environmental samples over a period covering 6 years from 1993 to 1998, and have come to the following inescapable conclusions. The analytical sensitivities of both ICP-MS and GFAAS for toxic metals are comparable. For kinetic runs on aqueous, environmental samples, ICP-MS has the great advantage of providing continuous sampling with numerous data points having adequate time resolution, whereas GFAAS provides discrete sampling with fewer data points having inadequate time resolution. However, in the most important and decisive

criterion of sample compatibility with the technique, ICP-MS fails completely on account of its inability to handle any solutions containing significant amounts of solute materials and/or corrosive chemicals such as hydrofluoric acid and strong inorganic acids and bases, whereas GFAAS can readily handle such samples. Since the aqueous, environmental samples for kinetic runs are always pretreated with solutes in the form of acid–base buffers, and since such solutions cannot be diluted in order to make them acceptable to ICP-MS, it is not possible to use ICP-MS for such kinetic studies without fouling the interior of the ICP-MS equipment with encrustations of the solute materials, which do serious damage. The alternative then is GFAAS, in spite of all its limitations. Because of the easy compatibility of GFAAS with the difficult sample type described above, GFAAS will continue to be used, as it is used now, until another new analytical technique which has all the advantages of ICP-MS without its fatal deficiencies mentioned above, is invented, developed and tested using real-life samples, i.e. for a long time.

*See also:* **Atomic Absorption, Theory; Atomic Emission, Methods and Instrumentation; Atomic Fluorescence, Methods and Instrumentation; Inductively Coupled Plasma Mass Spectrometry, Methods.**

## Further reading

Chakrabarti CL, Gilmutdinov AKh and Hutton JC (1993). Digital imaging of atomization processes in electrothermal atomizer for atomic absorption spectrometry. *Analytical Chemistry* **65**: 716–723.

L'vov BV (1970) *Atomic Absorption Spectrochemical Analysis*. London: Adam Hilger.

L'vov BV (1978) Electrothermal atomization—the way toward absolute methods of atomic absorption analysis. *Spectrochimica Acta, Part B* **33**: 153–193.

Sturgeon R (1986) Graphite furnace atomic absorption spectrometry: fact and fiction. *Fresenius' Zeitschrift für Analytische Chemie* **324**: 807–818.

Walsh A (1980) Atomic absorption spectroscopy—some personal reflections and speculations. *Spectrochimica Actya, Part B* **35**: 639–642.

# ATR and Reflectance IR Spectroscopy, Applications

**UP Fringeli**, University of Vienna, Austria

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

In Memory of N Jim Harrick.

## Introduction

Chemical reactions that occur at gas–solid and liquid–solid interfaces are of central importance to a variety of research and technological areas, including biomembranes, drug design, drug–membrane interaction, biosensors, chemical sensors, heterogeneous catalysis, thin film growth, semiconductor processing, corrosion and lubrication. Many methods are used for interface studies, ranging from most simple ones like the octanol–water two-phase system for mimicking the partition of a drug between a biomembrane and the surrounding water, to most specialized and expensive techniques such as low-energy electron diffraction (LEED), Auger electron spectroscopy (AES), X-ray photoelectron spectroscopy (XPS/ESCA) and ion scattering spectroscopy (ISS).

Among this palette of techniques, optical reflection spectroscopy in the mid- and near-IR range occupies an important complementary position. The basic equipment consists of a commercial IR spectrometer and a suitable reflection accessory that usually fits into the sample compartment of the spectrometer. Many reflection techniques permit *in situ* applications, and if applied in the mid IR, result in quantitative and structural information on a molecular level. Moreover, IR reflection spectroscopy features a very high performance-to-price ratio.

There is a wide range of different spectroscopic reflection techniques. First one should distinguish between internal (total) and external reflection. Attenuated total reflection (ATR) belongs to the first group. It makes use of the evanescent wave existing at the interface of the IR waveguide and the sample. Commercial ATR attachments differ mainly in shape and mounting of the internal reflection element (IRE) in the light path. Most IREs enable multiple internal reflections, a prerequisite for monolayer and

sub-monolayer spectroscopy, and a referred to as MIRE.

A variety of external reflection techniques are in use. In specular reflection (SR) the radiation reflected from the front surface of a bulk sample is collected. SR is often measured at or near normal incidence. Reflected spectral energy depends on the absorption behaviour of the sample. In regions of strong absorption the reflected energy is enhanced with respect to non-absorbing spectral regions, moreover, the reflection spectrum is usually very different from a corresponding absorption (AB) spectrum obtained by a transmission experiment. AB spectra may, however, be calculated from SR spectra by means of the Kramers–Kronig transformation (KKT). Corresponding software for SR data processing is supplied with most commercial IR instruments. While specular reflectance is measured at or near normal incidence, IR reflection–absorption spectroscopy (IRRAS) works from about 10° to grazing incidence. In this case the sample is placed on a reflecting substrate, usually a metal. The portion of reflected light from the sample surface is generally small compared with the energy reflected off the metal surface. Therefore, IRRAS data and transmission (T) data are analogous. From Fresnel's equations (see below), it follows that parallel ($\parallel$) and perpendicular ($\perp$) polarized electromagnetic waves undergo different phase shifts upon reflection. This phase shift is 180° for $\perp$-polarized light at non-absorbing interfaces. As a consequence, incoming and reflected beams cancel at the interface (node). On the other hand, at large angles of incidence $\parallel$-polarized incident light results in an enhanced electric field component perpendicular to the reflecting interface (*z*-axis). For thin samples, i.e. sample thickness ($d$) much smaller than the quarter wavelength ($\lambda/4$) of the reflected light, RA spectra report only partial information on orientation. It should be noted, however, that for a complete orientation analysis spectra obtained with light polarized in the plane of the interface (*x*,*y*-plane) is also necessary. ATR fulfils this requirement, in contrast to RA.

Diffuse reflectance (DR) is successfully applied to obtain IR spectra of rough (scattering) or dull surfaces, i.e. of media intractable by other reflection techniques. The interpretation of DR spectra, however, is sometimes handicapped by the fact that they may be a mixture of AB and SR spectra. DR spectroscopy is a sensitive tool, especially when used with an IR Fourier transform (FT) spectrometer (DRIFT).

Elucidation of structure–activity relationships is the aim of many applications of reflection spectroscopy to thin layers at interfaces. In this context, polarization measurements are of considerable importance, since molecules at interfaces exhibit often induced ordering.

Low signal intensities are common in different kinds of interface spectroscopy, especially when the sample consists of a monolayer or even submonolayer as usual in heterogeneous catalysis and substrate–biomembrane interaction. Although modern FTIR spectrometers exhibit very high stability, signal-to-noise (S/N) ratio enhancement by data accumulation is limited by environmental and instrumental instabilities. The fact that most commercial FT instruments are operated in a single-beam mode is disadvantageous in this respect, because the longer an experiment lasts, the greater is the time lag between sample and reference data, which facilitates the intrusion of instabilities. Several optional extras are available in order to reduce the time lag between acquisition of sample and reference spectra.

One possibility is the conversion of the single-beam instrument into a pseudo-double-beam instrument by means of a shuttle which moves alternately the sample and reference into the IR beam. Such attachments were first developed for transmission experiments, and were later adapted for ATR measurements. In the latter case the sample and the reference are placed on top of one another on the same trapezoidal MIRE. A parallel beam of half the height of the MIRE is directed alternately through the upper and lower half of the MIRE by computer-controlled vertical displacement of the ATR cuvette. This method is referred to as single-beam sample reference (SBSR) technique and is described in more detail below.

Polarization modulation (PM) in combination with IRRAS is a further possibility to enhance instrumental stability and background compensation when working at grazing incidence to a thin sample on a metal substrate. PM at about 50 kHz is achieved by means of a photoelastic modulator (PEM). Since under these experimental conditions the sample will only absorb light in the $\parallel$-polarized half-wave, the $\perp$-polarized half-wave of the signal is representative of the background, i.e. of the reference. Subtraction is performed by lock-in technique within each PM cycle, i.e. 50 000 times per second. As a consequence, environmental and instrumental contributions are largely compensated.

Finally, it should be noted that a more general application of modulation spectroscopy can be used to obtain selective information on an excitable sample. Modulated excitation (ME) spectroscopy can always be applied with samples allowing periodic stimulation via a periodic variation of any external thermodynamic parameter, e.g. temperature, pressure, concentration, electric field, light flux. ME causes a

periodic variation of the absorbance at those wavelengths that are typical for the molecules involved in the stimulated process. Phase-sensitive detection (PSD) by digital lock-in technique adapted for FTIR instruments permits spectral registration of the modulated, i.e. affected, part of the sample. A typical feature of ME with PSD is the comparison of sample and reference within each period of stimulation. Within this time interval environmental and instrumental parameters are usually stable so that a very good baseline is achieved. Moreover, if one or more relaxation times $\tau_i$ of the kinetic response of the stimulated sample fulfil the condition $0.1 < \omega\tau_i < 10$, where $\omega$ denotes the angular frequency of stimulation, significant phase lags $\phi_i$ between stimulation and sample responses will occur which are related to the reaction scheme and the rate constants of the stimulated process.

## Theory of reflectance spectroscopy

For a comprehensive description of the theory of reflectance the reader is referred to the Further reading section. In this article, theory will only be presented when necessary for a general understanding.

### Fresnel's equations

The theory of reflection and transmission of an electromagnetic wave by a plane boundary was first derived by Fresnel. The geometry of specular reflection and transmission is depicted in **Figure 1**. The incident (i) plane wave consists of the parallel (‖) and perpendicular polarized (⊥) electric field components $E_{i\|}$ and $E_{i\perp}$, respectively. The corresponding components of the reflected ($r$) and refracted (transmitted $t$) field components are denoted by $E_{r\|}$, $E_{r\perp}$, $E_{t\|}$, and $E_{t\perp}$. Fresnel's equations relate the reflected and transmitted components to the corresponding incident components.

For a nonabsorbing medium, i.e. the absorption indices $\kappa_1$ and $\kappa_2$ equal to zero, one obtains for the ratio $r$ between reflected and incident electric field

$$r_\| = \frac{E_{r\|}}{E_{i\|}} = \frac{n_2\cos\theta_i - n_1\cos\theta_t}{n_2\cos\theta_i + n_1\cos\theta_t}$$
$$r_\perp = \frac{E_{r\perp}}{E_{i\perp}} = \frac{n_1\cos\theta_i - n_2\cos\theta_t}{n_1\cos\theta_i + n_2\cos\theta_t} \qquad [1]$$

where ‖ and ⊥ denote parallel and perpendicular polarization, according to **Figure 1**. Equation [1] may be modified by introducing Snell's law of refraction:

$$n_1\sin\theta_i = n_2\sin\theta_t \qquad [2]$$

resulting in

$$r_\| = \frac{E_{r\|}}{E_{i\|}} = \frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)}$$
$$r_\perp = \frac{E_{r\perp}}{E_{i\perp}} = -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)} \qquad [3]$$

As concluded from Equation [3], perpendicular polarized incident light undergoes a phase shift of 180° upon reflection, i.e. there is a node at the reflecting interface resulting in zero electric field strength at this point. On the other hand, parallel polarized components remain in-phase. However, this conclusion holds no longer in the case of absorbing media.

The corresponding equations for the ratio $t$ between transmitted and incident electric fields are

$$t_\| = \frac{E_{t\|}}{E_{i\|}} = \frac{2n_1\cos\theta_i}{n_2\cos\theta_i + n_1\cos\theta_t}$$
$$t_\perp = \frac{E_{t\perp}}{E_{i\perp}} = \frac{2n_1\cos\theta_i}{n_1\cos\theta_i + n_2\cos\theta_t} \qquad [4]$$

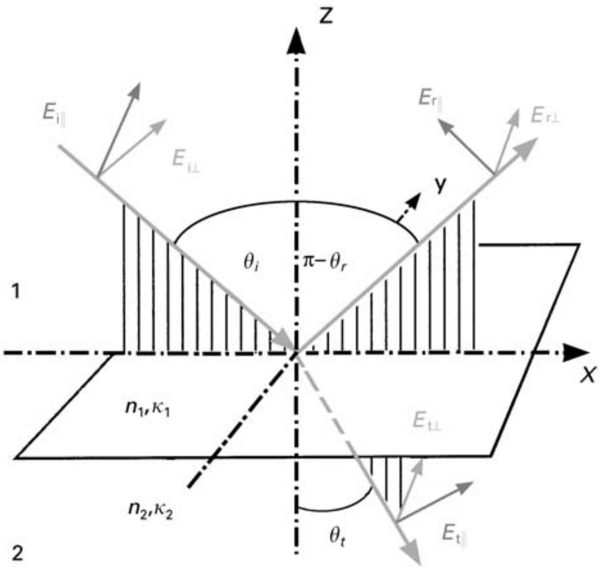



**Figure 1** Specular reflection and transmission. The angles of incidence ($i$), reflection ($r$) and refraction ($t$) are denoted by $\theta_i$, $\theta_r$ and $\theta_t$, respectively. The corresponding electric field components are denoted by $E$. They are split into orthogonal portions, one parallel to the plane of incidence ($x$,$z$-plane) and the other perpendicular to this plane (parallel to $y$-axis). Accordingly, electric fields are referred to as parallel (‖) and perpendicular (⊥) polarized, $n_1$, $n_2$, $\kappa_1$ and $\kappa_2$ denote the refractive and absorption indices in the two media.

In order to modify Equations [1], [3] and [4] for absorbing media one has to introduce the complex refractive index. For an incoming plane-wave it is given by

$$\hat{n} = n + i\kappa \qquad [5]$$

where $n$ is the refractive index and $\kappa$ is the absorption index. As a consequence, $r$ and $t$ become complex and the resulting phase shifts differ from 0° and 180° as mentioned above.

## Energy flux density

The flux density of an electromagnetic wave is described by the Poynting vector. For the case of the plane wave field one obtains for the time average in the direction of propagation

$$\bar{S} = \frac{1}{2}\sqrt{\varepsilon/\mu}E_0^2 \qquad [6]$$

where $\mu$ is the permeability, i.e. the product of the permeability of vacuum $\mu_0$ and the relative permeability $\mu_r$ which is unity for nonconductive materials, and $\varepsilon$ denotes the permittivity which is the product of the permittivity of vacuum $\varepsilon_0$ and the relative permittivity (dielectric constant) $\varepsilon_r$ which is complex for absorbing media, according to

$$\begin{aligned}\hat{\varepsilon}_r &= \hat{n}^2 = n^2 - \kappa^2 + i \cdot 2n\kappa \\ &= (n^2 + \kappa^2) \cdot \exp\left[i \cdot \operatorname{atan}\left(\frac{2n\kappa}{n^2 - \kappa^2}\right)\right] \\ &= |\hat{\varepsilon}_r| \cdot \exp(i \cdot \varphi)\end{aligned} \qquad [7]$$

Introducing the absolute value of Equation [7] into Equation [6] results in, for nonconducting media,

$$\bar{S} = \frac{1}{2}\sqrt{\varepsilon_0/\mu_0}\,\sqrt{n^2 + \kappa^2}\,E_0^2 \qquad [8]$$

Reflectance $\rho$ and transmittance $\tau$ are defined as the ratios of the corresponding energy fluxes $\bar{J}$. According to Equation [8] they are proportional to the square of the electric field, i.e.

$$\rho = \frac{\bar{J}_r}{\bar{J}_i} = rr^* \quad \text{and} \quad \tau = \frac{\bar{J}_r}{\bar{J}_i} = \frac{\sqrt{(n_2^2 + \kappa_2^2)}\cos\theta_t}{\sqrt{(n_1^2 + \kappa_1^2)}\cos\theta_i}tt^* \qquad [9]$$

The factor $\sqrt{(n_2^2 + \kappa_2^2)}/\sqrt{(n_1^2 + \kappa_1^2)}$, which reduces to $n_{21}$ for nonabsorbing media, results from the change of dielectrica (see Equation [8]), and $\cos\theta_t/\cos\theta_i$ takes account of the different cross-sections of the beam in media 1 and 2, respectively.

$rr^*$ and $tt^*$, become $r^2$ and $t^2$ for nonabsorbing media. In this case, Equations [1], [2] and [8] result in

$$\rho_\parallel = \left(\frac{n_{21}^2\cos\theta_i - \sqrt{n_{21}^2 - \sin^2\theta_i}}{n_{21}^2\cos\theta_i + \sqrt{n_{21}^2 - \sin^2\theta_i}}\right)^2$$
$$\rho_\perp = \left(\frac{\cos\theta_i - \sqrt{n_{21}^2 - \sin^2\theta_i}}{\cos\theta_i + \sqrt{n_{21}^2 - \sin^2\theta_i}}\right)^2 \qquad [10]$$

where $n_{21}$ denotes the ratio of refractive indices of media 2 and 1, respectively (see **Figure 1**). For normal incidence, i.e. $\theta_i = 0$, Equation [10] reduces to

$$\rho_\parallel = \rho_\perp = \left(\frac{n_{21} - 1}{n_{21} + 1}\right)^2 \qquad [11]$$

In order to obtain the reflectance of an absorbing medium one may introduce Equation [5] into Equation [10] or [11]. The result for normal incidence is

$$\rho_\parallel = \rho_\perp = \frac{(n_2 - n_1)^2 + (\kappa_2 - \kappa_1)^2}{(n_2 + n_1)^2 + (\kappa_2 + \kappa_1)^2} \qquad [12]$$

It should be noted that in many applications medium 1 is air or a nonabsorbing crystal, i.e. $\kappa_1 = 0$. It follows from Equation [12] that the reflectance increases with increasing absorption index of medium 2 ($\kappa_2$). In the limiting case of $\kappa_2 \to \infty$ one obtains $\rho \to 1$, i.e. a perfect mirror. Expressions for the more complicated case of oblique incidence to absorbing media have been derived (see Further reading).

## The Kramers–Kronig relations

For normal incidence ($\theta_i = 0$) and nonabsorbing medium 1 ($\kappa_1 = 0$, see **Figure 1**), one obtains from Equations [1] and [5] the following expression for the complex ratio $\hat{r}(\theta_i = 0)$ between reflected and

incident electric fields:

$$\hat{r}_{\parallel}(\theta_i=0)=\frac{E_{r\parallel}}{E_{i\parallel}}=\frac{n_2-n_1+i\kappa_2}{n_2+n_1+i\kappa_2}=\eta\exp(i\phi)$$
$$\hat{r}_{\perp}(\theta_i=0)=\frac{E_{r\perp}}{E_{i\perp}}=-\frac{n_2-n_1+i\kappa_2}{n_2+n_1+i\kappa_2}=\eta\exp[i(\phi+\pi)] \quad [13]$$

where $\eta$ and $\phi$ are the amplitude and phase of $\hat{r}(\theta_i = 0)$. They are functions of the wavenumber $\tilde{\nu}$ and related to each other by the Kramers–Kronig equations:

$$\ln\eta(\tilde{\nu})=\frac{2}{\pi}\int_0^{\infty}\frac{\nu\,\phi(\nu)}{\nu^2-\tilde{\nu}^2}\mathrm{d}\,\nu$$
$$\phi(\tilde{\nu})=\frac{2\,\tilde{\nu}}{\pi}\int_0^{\infty}\frac{\ln\eta(\nu)}{\nu^2-\tilde{\nu}^2}\mathrm{d}\,\nu \quad [14]$$

Experimentally, $\eta(\tilde{\nu})$ can be determined, since it is related to the reflectance at normal incidence by $\rho(\theta_i=0)=\eta(\tilde{\nu})\eta(\tilde{\nu})^*$, i.e. $\eta(\tilde{\nu})=\sqrt{\rho(\theta_i=0)}$; see Equation [12]. The Drude model may be used to extrapolate the measurement to $\tilde{\nu} = 0$ and $\tilde{\nu} = \infty$. From $\eta(\tilde{\nu})$ and $\phi(\tilde{\nu})$ the components of the refractive index can be calculated according to

$$n=\frac{1-\eta^2}{1-2\eta\cos\phi+\eta^2}$$
$$\kappa=\frac{2\eta\sin\phi}{1-2\eta\cos\phi+\eta^2} \quad [15]$$

For a detailed discussion, see the Further reading section.

### Internal reflection spectroscopy (ATR)

Internal reflection can only occur when the angle of the refracted beam $\theta_t$ is larger than the angle of incidence $\theta_i$. This means, according to Snell's law (Equation [2]), that the refractive index of medium 2 must be smaller than that of medium 1 ($n_2 < n_1$). This is contrary to the situation in **Figure 1** where $n_2 > n_1$ was assumed. The region of total reflection begins when $\theta_t$ reaches 90°, i.e. at the critical angle of incidence $\theta_c$. It follows from Equation [2] that

$$\sin\theta_c=\frac{n_2}{n_1}=n_{21} \quad [16]$$

It follows from Equations [2] and [16] for $\theta_i > \theta_c$ that $\sin\theta_t = \sin\theta_i/\sin\theta_c > 1$, resulting in a complex value for the corresponding cosine:

$$\cos\theta_t=\pm\, i n_{12}\sqrt{\sin^2\theta_i-n_{21}^2} \quad [17]$$

The ratio $r$ between internally reflected and incident electric field components is then obtained by introducing Equation [17] into Equation [1], resulting in

$$r_{\parallel}=\frac{E_{r\parallel}}{E_{i\parallel}}=\frac{n_{21}^2\cos\theta_i-i\sqrt{\sin^2\theta_i-n_{21}^2}}{n_{21}^2\cos\theta_i+i\sqrt{\sin^2\theta_i-n_{21}^2}}$$
$$r_{\perp}=\frac{E_{r\perp}}{E_{i\perp}}=\frac{\cos\theta_i-i\sqrt{\sin^2\theta_i-n_{21}^2}}{\cos\theta_i+i\sqrt{\sin^2\theta_i-n_{21}^2}} \quad [18]$$

The corresponding equations for medium 2 are obtained by introducing Equation [17] into Equation [4], resulting in

$$t_{\parallel}=\frac{E_{t\parallel}}{E_{i\parallel}}=\frac{2\cos\theta_i}{n_{21}(\cos\theta_i+i\sqrt{\sin^2\theta_i-n_{21}^2})}$$
$$t_{\perp}=\frac{E_{t\perp}}{E_{i\perp}}=\frac{2\cos\theta_i}{\cos\theta_i+in_{21}^2\sqrt{\sin^2\theta_i-n_{21}^2}} \quad [19]$$

Finally, it should be noted that incident electric fields undergo phase shifts in the ATR mode even if medium 2 is nonabsorbing. It follows from Equation [18] that

$$\tan\frac{\delta_{\parallel}}{2}=-\frac{\sqrt{\sin^2\theta_i-n_{21}^2}}{n_{21}^2\cos\theta_i}$$
$$\tan\frac{\delta_{\perp}}{2}=-\frac{\sqrt{\sin^2\theta_i-n_{21}^2}}{\cos\theta_i} \quad [20]$$

where $\delta_{\parallel}$ and $\delta\perp$ are the phase shifts per internal reflection (no absorption) of ∥-polarized and ⊥-polarized incident light. Since the phase shifts and amplitudes are polarization dependent, linearly polarized incident light is elliptically polarized after an internal reflection. This phenomenon, however, does not hinder polarization measurement in the ATR mode.

## Applications

### Diffuse reflectance

The geometry of a diffuse reflection experiment is shown in **Figure 2**. The incident beam (I) is collimated to the sample S by means of the ellipsoidal mirror $M_i$. Two reflection mechanisms must be considered, specular reflection, $R_s$, and diffuse reflection, $R_d$. The former occurs at the surface and is governed by the Fresnel equations (Equations [1], [3] and [10–12]). As a consequence of anomalous dispersion, specular reflected light exhibits S-shaped intensity changes at the wavelengths of sample absorption. In contrast, diffuse reflected light exhibits absorption bands at frequencies observed also with transmitted light, but with intensities deviating significantly from those measured in a transmission experiment. The intensity of the diffuse reflection spectrum may be described by the Bouguer–Lambert law (Eqn [21]), the analogous expression to the Lambert–Beer law in transmission spectroscopy.

$$R_d = I_0 \exp(-\alpha \bar{d}) \qquad [21]$$

where $\bar{d}$ is the mean penetrated layer thickness, i.e. the depth of light penetration into the surface layer which results in an intensity decrease by a factor of 1/e, and $\alpha$ denotes the napierian absorption coefficient.

Diffuse reflectance infrared Fourier transform spectroscopy (DRIFT) has become a frequently used technique to obtain IR spectra from materials intractable by transmission spectroscopy. A number of high-performance reflection accessories are available from different manufacturers (see below), allowing the detection of quantities down to the nanogram region. Nevertheless, DRIFT spectroscopy is confronted with two intrinsic problems: (i) the superposition of diffuse and specular reflected light (see **Figure 2**), which may lead to distorted line shapes, and (ii) the dependence of the mean penetration depth $\bar{d}$ on the absorption coefficient. $\bar{d}$ is found to be inversely proportional to the absorption coefficient $\alpha$, thus leading to a certain leveling of the band intensities.

The disturbance by specular reflection may be reduced considerably by technical means (trapping) on the reflection attachment. The resulting diffuse reflection spectrum then has to be corrected in order to correspond to the absorbance of a transmission spectrum. This mathematical procedure is generally performed according to the Kubelka–Munk theory. For a comprehensive and critical discussion of this theory the reader is referred to the Further reading section.

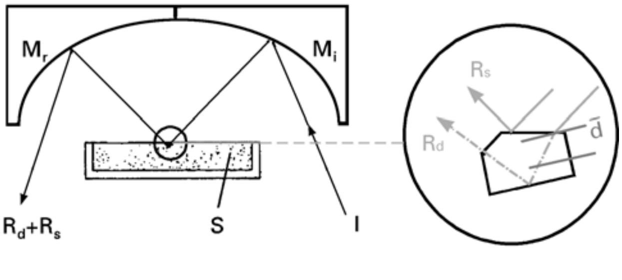



**Figure 2** Diffuse reflection experiment. $M_i$, $M_r$ = ellipsoidal mirrors for incident and reflected light; S = sample; I, $R_d$, $R_s$ = incident diffuse, and specular reflected beams, respectively. In the magnified circle a possible ray tracing through a surface particle is shown, demonstrating the formation of mixed diffuse and specular reflected light. $\bar{d}$ is the mean penetrated layer thickness according to the Bouguer–Lambert law (Equation [21]).

### Specular reflection spectroscopy (SRS)

In specular reflectance, only light reflected off the front surface is collected (see **Figure 2**). The reflected energy is generally small (<10%) for non-absorbing regions at normal or near-normal incidence. However, according to Equation [12], enhanced reflectance is observed in regions of sample absorption. As illustrated by **Figure 3**, radiation intensity is different to transmission intensity, since S-shaped bands result as a consequence of anomalous dispersion of the refractive index in the region of an absorption band. As a typical example, the specular reflectance spectrum of a black plastic is shown in **Figure 3**.

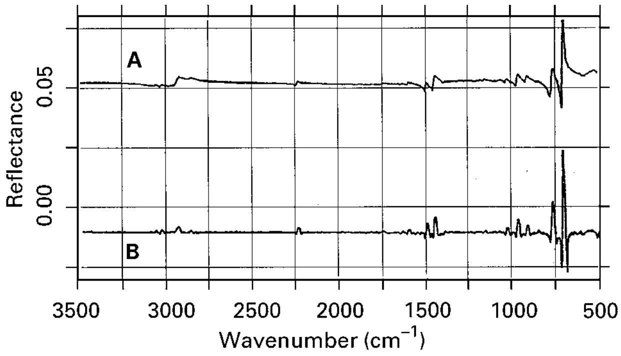



**Figure 3** (A) Specular reflectance (SR) spectrum of a black acrylonitrile–butadiene–styrene polymer film measured at near normal incidence, (B) Absorbance spectrum after data treatment of SR spectrum by a Kramers–Kronig transformation. Reproduced in part with permission of Elsevier Science from Zachman G (1995) *Journal of Molecular Structure* **348**: 453–456.

### Reflection absorption spectroscopy (RAS)

**RAS at near normal incidence** This is one of the most common and straightforward external reflection techniques. The IR beam is directed to the sample in the angular range 10–50°. The sample film must be on a reflective support. Under these conditions the RA spectrum is dominated by absorption since specular reflectance from the outer sample surface results in only 4–10% as shown by **Figure 4**. For this reason RA spectra resemble transmission spectra very closely. Accordingly, typical sample thicknesses are between 0.5 and 20 μm.

**RAS at grazing angle** Sensitivity at grazing angle is significantly enhanced with respect to near-normal reflectance and to transmission. The enhancement is explained by a polarization effect a the reflective surface. This effect is greatest for metal substrates and angles of incidence above 80°. As a general effect, ⊥-polarized incident light has a node at the interface (see above). The resulting reflectance will be very weak as long as the layer thickness is significantly smaller than $\lambda/4$. Therefore, ⊥-polarized reflectance may be used as reference spectrum. ||-Polarized incident light produces electric field components in the *x*- and *z*-directions. On metal surfaces, however, the *x*- and *y*-components vanish, but, the *z*-component is significantly enhanced owing to interference of incident and reflected beams. Compared with near-normal incidence, RAS magnification may be by more than one order of magnitude, thus permitting monolayer spectroscopy by a single reflection. Detailed information on RAS techniques applied to study carbon monoxide (CO) adsorbed on metal surfaces can be found in the Further reading section.

More recently, RAS at grazing angle has been applied successfully for *in situ* spectroscopy of lipid monolayers and proteins at the air–water interface of a Langmuir trough, using the water surface as reflector (see Further reading section).

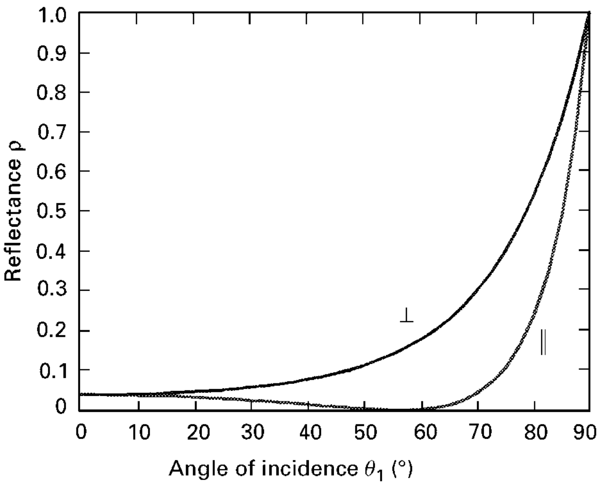



**Figure 4** Specular reflectance calculated according to Equation [10]. Refractive indices: $n_1$=1 and $n_2$=1.5, ||, ⊥ denote parallel and perpendicular polarized incident light. The Brewster angle, where $\rho_{||}$ = 0 for a nonabsorbing medium, is calculated as $\theta_i = \theta_B$ = 56.3°, according to tan $\theta_B = n_{21}$.

**Polarization modulation RAS** This technique makes use of the fact that ⊥-polarized incident light has a node at the reflecting interface resulting in zero absorbance at this point and nearly no absorbance of films significantly thinner than the quarter wavelength. Under these conditions, the ⊥-reflectance spectrum may be used as a reference spectrum for the ||-reflectance spectrum. Since monolayer and submonolayer quantities of organic molecules result in low-intensity spectra, such measurements are susceptible to instrumental and environmental instabilities. This problem may be overcome by polarization modulation (PM). For this purpose, a photoelastic modulator (PEM) is placed in the light path, leading to very fast periodic polarization changes of the incident light. The frequency range is 40–100 kHz, i.e. high enough to avoid interference's with the interferometer frequencies. Phase-sensitive demodulation of the PEM signal results in the interferogram of the difference between ||- and ⊥-RA spectra, which is then normalized by division by the stationary response featuring the sum of ||- and ⊥-mean reflectance according to

$$\frac{\Delta R}{R} = \frac{R_{||} - R_{\perp}}{R_{||} + R_{\perp}} \quad [22]$$

A description of an experimental setup and of the relevant equations for PM-IRRAS can be found in the Further reading section.

Instabilities are largely compensated because sample and reference spectra are measured and evaluated within one period, i.e. within 10–25 μs. If applied in the IR region, this technique is referred to as PM-IRRAS.

It should be noted, however, that significant baseline problems may occur owing to different transmittance of ||- and ⊥-polarized light by the spectrometer. This is demonstrated by the IRRAS data of a cadmium arachidate monolayer on a gold surface in **Figure 5**.

One should note that the usually intense absorption bands of $CH_2$ stretching, $\nu(CH_2)$ in the 2800–2950 $cm^{-1}$ region and of $CH_2$ bending near 1470 $cm^{-1}$ are very weak in the IRRAS, spectrum in

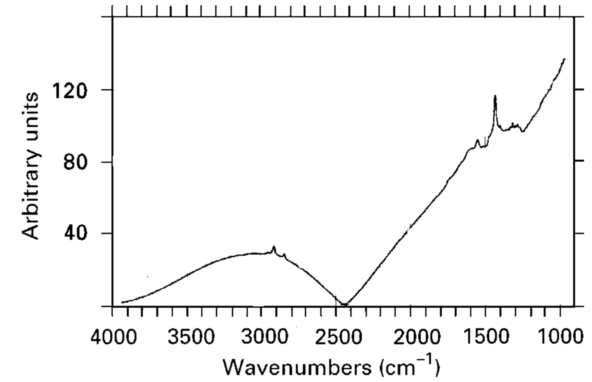



**Figure 5** PM-IRRAS spectrum of an arachidic acid monolayer on a gold surface. The spectrum was normalized according to Equation [22]. It should be noted that only molecular vibration can be detected by IRRAS when the corresponding transition moment exhibits a component normal to interface (*z*-direction). Reproduced in part with permission from Beccard B and Mapanowicz R (1995) *Nicolet Application Note AN-9542*. Madison, WI: Nicolet.

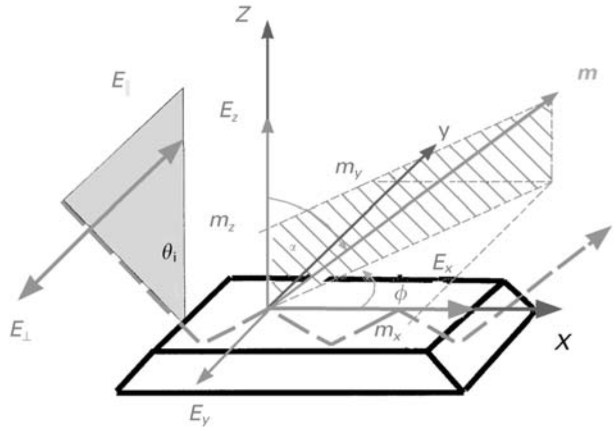



**Figure 6** ATR setup. Optical and structural features are related to the IRE fixed-coordinate system *x,y,z*. $E_{\|}$ and $E\perp$ denote the parallel and perpendicular polarized electric field components of the light incident to the IRE under the angle $\theta_i$. $E_{\|}$ results in the $E_x$, and $E_z$ components of the evanescent wave, while $E\perp$ results in the $E_y$ component. ***m*** denotes the unit vector in direction of the transition dipole moment vector of a given vibrational mode, and $m_x$, $m_y$, $m_z$ are the corresponding components in the IRE coordinate system. ***m*** goes off at an angle $\alpha$ with respect to the *z*-axis and the projection of ***m*** to the *xy*-plane goes off at an angle $\phi$ with respect to the *x*-axis. Reproduced with permission of the American Institute of Physics from Fringeli UP *et al.* (1998) *AIP Conference Proceedings* **430**: 729–747.

contrast to a corresponding DPPA monolayer ATR spectrum (see **Figure 9**). The reason is that IRRAS of thin layers on metallic surfaces offers only the electric field component normal to the interface, i.e. $E_z$, as already explained above. The intensities of $CH_2$ stretching and $CH_2$ bending in **Figure 5** are consistent with hydrocarbon chains of arachidic acid aligned along the normal to the interface. The lack of a $E_x$ and $E_y$ components, i.e. of an electric field component parallel to the interface, disables the determination of a tilt angle. In this respect, IRRAS is at a disadvantage with respect to ATR, the technique presented in the following section. Moreover, ATR allows significantly better baseline control, especially if special techniques such as SBSR (single-beam sample reference technique) or ME (modulated excitation) are applied.

### Attenuated total reflection (ATR) spectroscopy

A number of interesting conclusions may be drawn from Equations [18] and [19]. First, calculation of the reflectance according to Equation [9] results in $\rho_{\|} = \rho\perp = 1$, which means *total reflection*. However, if medium 2 is absorbing, one has to insert the complex refractive index (Equation [5]) into Equation [18], resulting in $\rho_{\|} \neq \rho\perp < 1$, which means *attenuated total reflection (ATR)*.

Second, in order to obtain more information on the nature of this process, one may calculate the propagation of a plane-wave in medium 2 under the conditions of total reflection. The result for a nonabsorbing medium 2 is that there is no transmittance normal to the interface, i.e. $\tau_z = 0$, however, there is an energy flux in the *x*,*y*-plane near the interface, i.e. $\tau_x, \tau_y \neq 0$. Hence, there is a electromagnetic wave beyond the interface, although the whole energy is totally reflected. This wave is referred to as *evanescent wave*. Straightforward calculation results in that the electric field strength of this wave decreases exponentially with distance *z* from the interface, according to

$$E_{x,y,z} = E_{0x,y,z} \exp\left[-\frac{z}{d_p}\right] \qquad [23]$$

The subscripts *x*, *y*, *z* stand for the electric field components of the evanescent wave in the corresponding directions. Subscript 0 denotes the value at $z = 0$ (interface in medium 2) and $d_p$ is the so-called *penetration depth*, which results in

$$d_p = \frac{\lambda/n_1}{2\pi\sqrt{\sin^2\theta_i - n_{21}^2}} \qquad [24]$$

where $\lambda$ denotes the wavelength in vacuum and $\lambda/n_1 = \lambda_1$ is the wavelength in medium 1, i.e. in the internal reflection element (IRE). A typical ATR setup is shown in **Figure 6**.

Hence the penetration depth is the distance in the $z$-direction within which the electric field is decreased by a factor of 1/e. This distance varies between a fraction of a micron and a few microns depending on the refractive indices of the IRE (e.g. $n_{Ge} = 4.0$, $n_{ZnSe} = 2.4$) and the medium 2($n_2 \approx 1.5$ for organic materials), as well as the angle of incidence. Harrick has given further practical details (see Further reading section). As a consequence of Equations [23] and [24], the ATR spectrum features information on materials within a distance of one or a few $d_p$ from the reflecting interface, resulting the highest sensitivity at the interface. Therefore, ATR spectroscopy is an optimum tool for *in situ* thin immobilized layer analysis. For reviews on membrane spectroscopy the reader is referred to the Further reading section.

**Quantitative analysis of ATR spectra**

*The concept of effective thickness* The concept of effective thickness was introduced by Harrick. The quantity $d_e$ indicates the thickness of a sample that would result in the same absorbance in a hypothetical transmission experiment, as obtained with the genuine ATR experiment. This concept enables the straightforward application of the Lambert–Beer's law on ATR data according to

$$T = 10^{-N\varepsilon c d_e} = 10^{-A} \qquad [25]$$

where $A = N\varepsilon c d_e$ denotes the absorbance resulting from $N$ internal reflections. For an isotropic layer extending from $z = z_i$ to $z = z_f$ one obtains

$$d_e^{\text{iso}} = \frac{1}{\cos\theta_i}\frac{n_2}{n_1}\frac{d_p}{2}E_{02}^{r^2} \times \left[\exp\left(-\frac{2z_i}{d_p}\right) - \exp\left(-\frac{2z_f}{d_p}\right)\right] \qquad [26]$$

According to Equation [26] $d_e$ turns out to be wavelength dependent via $d_p$. As a consequence, ATR spectra of bulk media generally show increasing intensity with increasing wavelength. However, if the thickness of the layer $d = z_f - z_i$ is small compared with $d_p$ then Equation [26] reduces to Equation [27], which is independent of the wavelength. $E_{02}^r$ denotes the relative electric field component at the reflecting interface of medium 2. For $z_i = 0$ one obtains

$$d_e^{\text{iso}} = \frac{1}{\cos\theta_i}\frac{n_2}{n_1} d\, E_{02}^{r^2} \qquad [27]$$

For a bulk medium extending from $z_i = 0$ to $z_f = \infty$ Equation [26] results in

$$d_e^{\text{iso}} = \frac{1}{\cos\theta_i}\frac{n_2}{n_1}\frac{d_p}{2}E_{02}^{r^2} \qquad [28]$$

A more detailed presentation including an approximate calculation of the effective thickness for intermediate layer thickness, i.e. $d \approx d_p$, and references for a rigorous application of the general formalism are given in the Further reading section.

*Relative electric field components* The relative electric field components $E_{02}^r$ are obtained from Equations [19] according to

$$\begin{aligned} &E_{02,\parallel}^r = \sqrt{(t_\parallel t_\parallel^*)};\ E_{02,\perp}^r = \sqrt{(t_\perp t_\perp^*)} \\ &E_{02,x}^r = E_{02,\parallel}^r \cdot \cos\theta_t;\ E_{02,z}^r = E_{02,\parallel}^r \cdot \sin\theta_t \\ &E_{02,y}^r = E_{02,\perp}^r \end{aligned} \qquad [29]$$

Explicit expressions of Equations [29] for thin films and bulk media can be found in references in the Further reading section.

*Validity of the effective thickness concept* Since the effective thickness concept permits the application of Lambert–Beer's law to ATR data, experimental validation may be performed easily by comparing spectra of the same sample measured by both, ATR and transmission (T). As long as the results do not differ significantly from each other the formalism described above is considered to be applicable. ATR and T measurements with aqueous solutions of $Na_2SO_4$ have shown that at a 1 M concentration Lambert–Beer's law is still fulfilled for the very intense $SO_4^{2-}$ stretching band at 1100 cm$^{-1}$. Even for the strong $H_2O$ bending [$\delta(H_2O)$] band of liquid water at 1640 cm$^{-1}$ the integral molar absorption coefficients determined by ATR with a germanium IRE at an angle of incidence of $\theta_i = 45°$ was found to be equal to T data within the experimental error. However, a few per cent deviations were found when peak values of the absorbance were used to determine the molar absorption coefficient. The latter indicates the onset of band distortion, a phenomenon well known in ATR spectroscopy under conditions of strong absorption. This finding is in accordance with calculations by Harrick using Fresnel's equations with complex refractive indices. For Ge in contact with liquid water and $\theta_i = 45°$ the analysis resulted in an upper limit of the absorption coefficient $\alpha_{max} \approx 1000$ cm$^{-1}$. The concept of effective thickness as described above may be considered to be

valid for $\alpha < \alpha_{max}$. For organic compounds this condition is generally fulfilled. In case of $\delta(H_2O)$ of liquid water, however, the absorption coefficient results in $\alpha = \varepsilon(1640\ \text{cm}^{-1})c = 1.82 \times 10^4\ \text{cm}^2\ \text{mol}^{-1} \times 5.56 \times 10^{-2}\ \text{mol cm}^{-3} = 1011.9\ \text{cm}^{-1}$, which indicates that the limit of validity of the approach is reached, in complete accordance with experimental data mentioned above. Furthermore, it turned out that the anomalous dispersion in the range of strong water absorption bands should be taken into account if sample absorption within this range is analysed quantitatively.

*Quantitative analysis taking sample absorption and thickness into account* Recently, the validity of Harrick's weak absorber approximations has been checked by comparison with the general thickness- and absorption-dependent model. It was found that the formalism depicted in the section above may be used for film thicknesses up to 20 nm. Especially if the film is in contact with a third bulk medium, e.g. water, the deviation between accurate and approximate calculation of relative electric field components according to Equation [29] was found to be below 3%, i.e. within the error of most experiments. A comprehensive description of ATR spectroscopy of polymers using the general formalism can be found in the Further reading section.

Since quantitative analysis of ATR data by the general formalism is very cumbersome, the use of more tractable approximations is recommended if possible. One possibility is the weak absorber approximations described above; another approach was derived for the study of electrochemical reactions by ATR spectroscopy.

Further more general information on quantitative methods and applications of ATR spectroscopy, see the Further reading section.

**Orientation measurements** Considering a transition dipole moment $\boldsymbol{M}$ associated with a vibrational mode of a given molecule and the electric field $\boldsymbol{E}$, responsible for vibrational excitation, the magnitude of light absorption depends on the mutual orientation of these two vectors according to Equation [30] which is the basis for orientation measurements. $M_x$, $M_y$ and $M_z$ denote the components of the transition dipole moment in the IRE fixed coordinate system shown in **Figure 6**.

$$\begin{aligned} A &= (\boldsymbol{E}\cdot\boldsymbol{M})^2 \\ &= |\boldsymbol{E}|^2 \cdot |\boldsymbol{M}|^2 \cdot \cos^2(\boldsymbol{E},\boldsymbol{M}) \\ &= (E_xM_x + E_yM_y + E_zM_z)^2 \end{aligned} \qquad [30]$$

It is usual to work with dimensionless relative intensities instead of absolute intensities in order to get rid of physical and molecular constants, e.g. the magnitude of the transition moment. Introducing the so-called dichroic ratio $R$, the absorbance ratio obtained from spectra measured with parallel and perpendicular polarized incident light, i.e.

$$\begin{aligned} R &= \frac{A_\parallel}{A_\perp} = \frac{d_{e,\parallel}}{d_{e,\perp}} \\ &= \frac{E_x^2\ \langle m_x^2\rangle + E_z^2\ \langle m_z^2\rangle + 2\ E_x\ E_z\ \langle m_x\ m_z\rangle}{E_y^2\ \langle m_y^2\rangle} \end{aligned} \qquad [31]$$

where $A$ and $d_e$ denote absorbance and effective thickness relative to $\parallel$- and $\perp$-polarized incident light, respectively. $E_x$, $E_y$ and $E_z$ denote the relative electric field components according Equation [29]. $\langle m_x^2\rangle$, $\langle m_y^2\rangle$, $\langle m_z^2\rangle$ and $\langle m_x\ m_z\rangle$ are ensemble mean values of the components of the unit vector in the direction of the transition dipole moment, see **Figure 6**. $\langle m_x\ m_z\rangle = 0$ for uniaxial orientation, e.g. isotropic distribution around all relevant orientation axes. On substitution of the unit vector components according to the geometry depicted in **Figure 6**, one obtains for the dichroic ratio

$$R^{\text{ATR}} = \frac{E_x^2}{E_y^2} + 2\ \frac{E_z^2}{E_y^2}\ \frac{\langle\cos^2\alpha\rangle}{1 - \langle\cos^2\alpha\rangle} \qquad [32]$$

where $\langle\cos^2\alpha\rangle$ denotes the mean square cosine of the angle between the transition moment and the normal to the interface. This quantity is accessible via two measurements, one with $\parallel$-polarization and the other with $\perp$-polarized incident light, resulting in $R$. The relative electric field components are available from Fresnel's Equations [19] with Equation [29]. For an isotropic sample, $\langle\cos^2\alpha\rangle_{\text{iso}} = 1/3$. Insertion of this value into Equation [32] result in

$$R_{\text{iso}}^{\text{ATR}} = \frac{E_x^2 + E_z^2}{E_y^2} \qquad [33]$$

It should be mentioned that an isotropic sample results in $R = 1$ in transmission but, according to Equation [33], not in ATR, where $R_{\text{iso}}^{\text{ATR}} \neq 1$. As an example, for total reflection with a bulk material 2 and $\theta_i = 45°$, one obtains $R_{\text{iso}}^{\text{ATR}} = 2.0$ irrespective of $n_1$ and $n_2$, except $n_1 > n_2$.

The segmental order parameter $S_{\text{seg}}$ is frequently used to characterize molecular ordering, e.g. To

describe the fluctuation of a functional group in a molecule via the polarized absorption bands of a typical group vibration. For uniaxial orientation the order parameter is defined according to

$$S_{\text{seg}} = \frac{3}{2}\langle\cos^2\alpha\rangle - \frac{1}{2} \qquad [34]$$

Thus, if the ATR geometry and the optical constants of the system are known, then $S_{\text{seg}}$ may be determined measuring the dichroic ratio $R$ of a given absorption band, followed by the calculation of the mean square cosine $\langle\cos^2\alpha\rangle$ and inserting this value into Equation [34]. A typical example is discussed later.

For more details and examples of application the reader is referred to the Further reading section.

**Determination of surface concentration** The effective thickness as indicated by Equations [26]–[28] holds for isotropic samples. Modification for oriented samples results

$$\begin{aligned} d_{ex} &= 3\,\langle m_x^2\rangle\, d_{ex}^{\text{iso}} = \tfrac{3}{2}(1-\langle\cos^2\alpha\rangle)\, d_{ex}^{\text{iso}} \\ d_{ey} &= 3\,\langle m_y^2\rangle\, d_{ey}^{\text{iso}} = \tfrac{3}{2}(1-\langle\cos^2\alpha\rangle)\, d_{ey}^{\text{iso}} \\ d_{ez} &= 3\langle m_z^2\rangle d_{ez}^{\text{iso}} = 3\,\langle\cos^2\alpha\rangle\, d_{ez}^{\text{iso}} \end{aligned} \qquad [35]$$

From Equation [25] and the relation $d_{e\|} = d_{ex} + d_{ez}$ and $d_{e\perp} = d_{ey}$ one obtains for the surface concentration $\Gamma$.

$$\begin{aligned} \Gamma = c\cdot d &= \frac{d\cdot A_\|}{\nu\cdot N\cdot d_{e\|}\cdot\varepsilon} = \frac{d\cdot\int A_\| d\tilde{\nu}}{\nu\cdot N\cdot d_{e\|}\cdot\int\varepsilon d\tilde{\nu}} \\ &= \frac{d\cdot A_\perp}{\nu\cdot N\cdot d_{e\perp}\cdot\varepsilon} = \frac{d\cdot\int A_\perp d\tilde{\nu}}{\nu\cdot N\cdot d_{e\perp}\cdot\int\varepsilon d\tilde{\nu}} \end{aligned} \qquad [36]$$

where $A_\|$ and $A_\perp$ denote the absorbance measured with parallel and perpendicular polarized incident light, respectively, $\varepsilon$ is the molar absorption coefficient, $\nu$ denotes the number of equal functional groups per molecule and $N$ is the number of active internal reflections. It should be noted that Equation [36] holds for peak absorbance and integrated absorbance, provided that the corresponding molar absorption coefficients are used.

## Special experimental ATR techniques

**Single-beam sample reference (SBSR) technique** Most FTIR spectrometers are working in the single-beam (SB) mode. As a consequence a single-channel reference spectrum has to be stored for later conversion of single-channel sample spectra into transmittance and absorbance spectra. This technique suffers inaccuracy owing to drifts resulting from the instrument, the sample or atmospheric absorption. In order to eliminate these unwanted effects to a great extent, a new ATR attachment has been constructed, converting a single-beam instrument into a pseudo-double-beam instrument. The principle features of this attachment are depicted in **Figure 7**. As usual, a convergent IR beam enters the sample compartment. However, the focal point is now displaced from the centre of the sample compartment by means of the planar mirrors M1 and M2 to the new position F. The off-axis

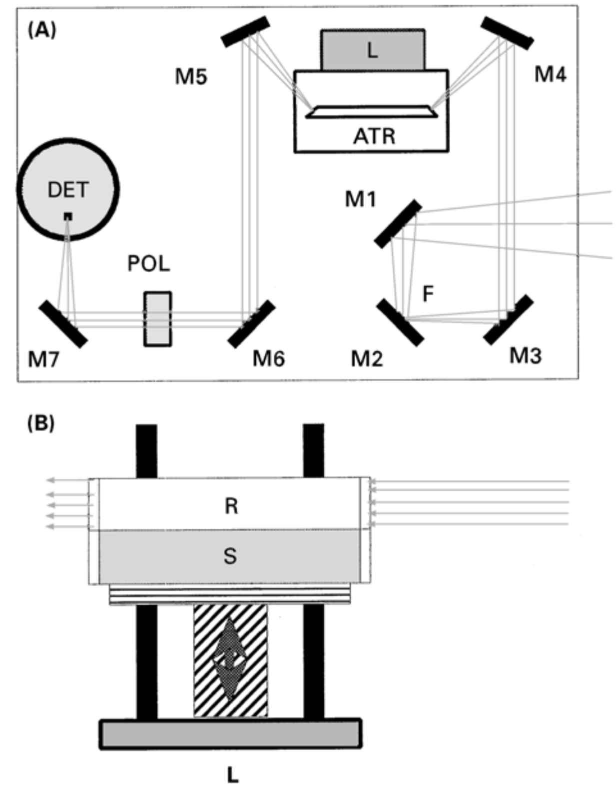



**Figure 7** Single-beam sample reference (SBSR) ATR attachment. (A) The focus in the sample compartment is displaced to the position F by the planar mirrors M1 and M2. The off-axis parabolic mirror M3 produces a parallel beam with a diameter of 1 cm, i.e. half of the height of the MIRE. The cylindrical mirror M4 focuses the light to the entrance face of the MIRE. M5, which has the same shape as M4, reconverts to parallel light directing it via the planar mirror M6 through the polarizer POL and it is then being focused to the detector DET by the off-axis parabolic mirror M7. (B) Alternating change from sample to reference is performed by computer-controlled lifting and lowering of the ATR cell body. Reproduced with permission of the American Institute of Physics from Fringeli UP *et al.* (1998) *AIP Conference Proceedings* **430**: 729–747.

parabolic mirror M3 performs a conversion of the divergent beam into a parallel beam with fourfold reduced cross-section. This beam is focused to the entrance face of a trapezoidal MIRE by a cylindrical mirror M4. Therefore, the ray propagation in the MIRE is still parallel to the direction of light propagation ($x$-axis), enabling subdivision of the two reflective faces ($x$,$y$-planes) of the MIRE alongside at half-height. One half of the MIRE is then used for the sample (S) and the other one for the reference (R). Both S and R were encapsulated by flow-through cuvettes, independently accessible by liquids or gases. This principle is referred to as the Single beam sample reference (SBSR) technique.

SBSR absorbance spectra are calculated from sample and reference single-channel spectra which have been measured with very short mutual time delay. A most favourable benefit of SBSR technique is that no waisted time for purging is required before starting a measurement after closing the sample compartment. Moreover, the whole sequence of single-channel spectra in the S and R channels is also available, allowing reconstruction of the history of each channel at any time by conventional data handling.

**Modulated excitation (ME) spectroscopy and 2D IR spectroscopy** Change of any external thermodynamic parameter generally exerts a specific influence on the state of a chemical system. The system response will be relaxation from the initial state (e.g. an equilibrium) to the final state (a new equilibrium state or a stationary state). In the case of a periodic change (modulation) of the parameter, the system response will also be periodic with angular frequency $\omega$, relaxing from the initial state to a stationary state. All absorption bands of the spectrum which result from stimulated molecules or parts of them will be labelled by the frequency $\omega$. As a consequence, it is possible to separate the modulated response of a system from the stationary response, resulting from parts of the system that were not affected by modulated excitation (ME) as well as from the background. Moreover, if the kinetics of the stimulated process is in the same time range as the period of external excitation, phase lags and damped amplitudes will result. Both depend characteristically on the stimulation frequency, and therefore one can derive relevant information on the reaction scheme and the kinetics of the stimulated process (see also caption to **Figure 8**).

A variety of ME experiments have been reported. (i) Temperature ME of poly-L-lysine was used to study induced periodic secondary structural changes as well as the sequence of transients. (ii) The classical ATR setup (see **Figure 6**) facilitates the application of electric fields to immobilized thin films, such as biomembrane assemblies or to bulk materials such as liquid crystals, since a Ge ATR plate, supporting the membrane, may be used as one electrode, and the back-wall of the cuvette as counter electrode. (iii) Hydration modulation was used to detect hydration sites of model membranes, and (iv) ME by UV radiation permitted kinetic studies of photoinduced chemical reactions. (v) ME by chemical substrates is a further versatile method to study chemically induced conformational changes of a sample immobilized to the MIRE. For that purpose, two computer-controlled pumps are used for periodic exchange the liquid (water) environment of the sample in a flow-through cell. An example demonstrating the sensitivity and high quality of background compensation of ME techniques is presented later. The principles of ME spectroscopy are depicted schematically in **Figure 8**.

**2D FTIR spectroscopy** Absorption bands in a set of modulation spectra that exhibit equal phase shifts with respect to the external stimulation are considered to be correlated. 2D correlation analysis is a statistical graphical means to visualize such a correlation in a 2D plot. Consequently, phase-resolved modulation spectra are data of a higher level and unambiguously allow a more direct and accurate evaluation. 2D plots look attractive, but, one should be aware that the information content is lower than that of the underlying modulation spectra, first because band overlapping may result in inadequate phase information, and second because 2D spectra are affected much more by baseline errors than the original modulation spectra. A comprehensive discussion can be found in the Further reading section.

### Sensitivity of ATR spectroscopy

**Sensitivity of stationary ATR measurements** Commercial multiple internal reflection elements MIRE permit up to 50 internal reflections. This is generally enough for thin-layer spectroscopy in the nanometre or even subnanometre region. As an example, **Figure 9** shows a dipalmitoylphosphatidic acid monolayer, i.e. a lipid monolayer of about 2 nm thickness, which has been transferred from the air–water interface to a germanium MIRE by means of the Langmuir–Blodgett technique.

The dominant bands in **Figure 9** result from the stretching vibrations of 28 $CH_2$ groups of the two saturated hydrocarbon chains of the DPPA molecule. Looking at three resolved weaker bands gives an impression of the absorbance to be expected from a monomolecular coverage by functional groups of medium or weak molar absorption. The first is the terminal methyl group of the hydrocarbon chains. The

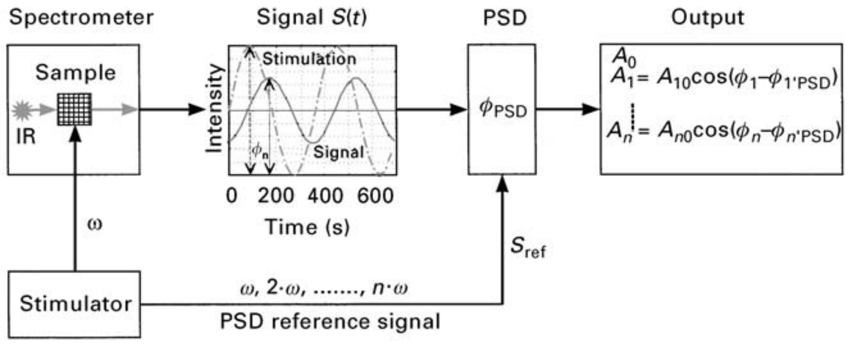



**Figure 8** Schematic setup for modulated excitation (ME) experiments. A periodic excitation is exerted on the sample with frequency $\omega$. The sample response $S(t)$, as sensed by IR radiation, contains the frequency $\omega$ and higher harmonics at wavelengths that are significant for those parts of the sample that have been affected by the stimulation. Selective detection of the periodic sample responses is performed by phase-sensitive detection (PSD), resulting in the DC output and $A_n$ of fundamental $\omega(n = 1)$ and their harmonics $n\omega$ ($n = 2, 3, \ldots$), as well as the phase shifts $\phi_n$ between the $n$th harmonic and the stimulation. This phase shift is indicative of the kinetics of the stimulated process and of the underlying chemical reaction scheme. Since the PSD output $A_n(n = 1, 2, \ldots, n$; frequency $n\omega)$ is proportional to $\cos(\phi_n - \phi_{n.\mathrm{PSD}})$, absorption bands featuring the same phase shift $\phi_n$ are considered to be correlated, i.e. to be representative of a population consisting of distinct molecules or molecular parts. $\phi_{n.\mathrm{PSD}}$ is the operator-controlled PSD phase setting. Because of the cosine dependence, different populations will have their absorbance maxima at different $\phi_{n.\mathrm{PSD}}$ settings, thus allowing selective detection. Moreover, since in the case that $0.1 < \omega\tau_i < 10$ ($\tau_i$ denotes the $i$th relaxation time of the system), $\phi_n$ becomes $\omega$ dependent, $\phi_n = \phi_n(\omega)$. The spectral information can then be spread in the $\phi_{n.\mathrm{PSD}} - \omega$ plane, resulting in a significant enhancement of resolution with respect to standard difference spectroscopy and time-resolved spectroscopy.

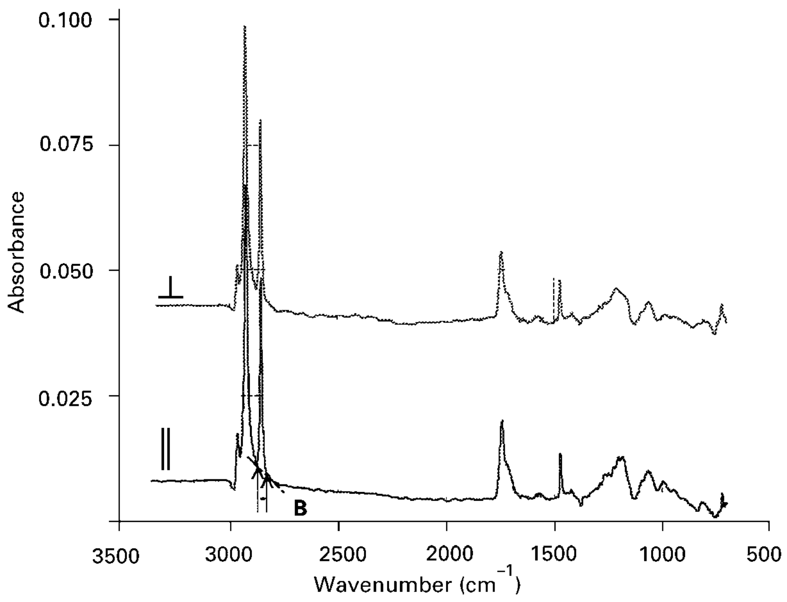



**Figure 9** Parallel (||) and perpendicular (⊥) polarized ATR absorbance spectra of a dipalmitoylphosphatidic acid (DPPA) monolayer transferred at 30 mN $m^{-1}$ from the aqueous subphase ($10^{-4}$ M $CaCl_2$) to a germanium multiple internal reflection element (MIRE). Spectra were obtained from the dry monolayer in contact with dry air. A surface concentration of $\Gamma = 3.93 \times 10^{-10}$ mol $cm^{-2}$ was calculated by means of Equation [36] using the dichroic ratio of the symmetric $CH_2$ stretching vibration at 2850 $cm^{-1}$ with respect to a linear baseline (B), resulting in $R^{\mathrm{ATR}}$ [$\nu^s(CH_2)$] = 0.923. Angle of incidence $\theta_i = 45°$; number of equal functional groups $\nu = 28$; number of active internal reflections $N = 39$.

antisymmetric stretching vibration, $\nu_{as}(CH_3)$ absorbs at ~2960 $cm^{-1}$. As concluded from **Figure 9**, this monolayer results in a peak absorbance of about 6 mAU. A weaker band is observed near 1420 $cm^{-1}$ and may be assigned to the bending vibration of the $\alpha$-methylene groups of the hydrocarbon chains, $\delta(\alpha\text{-}CH_2)$. Thus an approximate monolayer of $\alpha$-$CH_2$ groups results in an absorbance of only about 1 mAU. Third, a monolayer of phosphate head groups results in more intense absorption bands because of the larger transition dipolemoment of the polar group. The corresponding absorbances of $PO_3$ stretching

vibrations in the range 1000–1250 $cm^{-1}$ are within 5 and 10 mAU. It is concluded that conventional ATR measurements may allow significant access to bands of about 0.2–0.5 mAU, which corresponds to 20–50% of a monolayer of weak absorbers.

**Quantitative analysis of stationary ATR spectra** The DPPA monolayer spectra shown in **Figure 9** are now used to demonstrate the ease of application of the formalism for quantitative analysis of ATR spectra presented earlier.

*Dichroic ratio of symmetric $CH_2$ stretching* The dichroic Ratio according to Equation [32] was calculated from the integrated absorbances of the symmetric $CH_2$ stretching bands, $\nu_s(CH_2)$, using linear baselines as marked in **Figure 9** with lower and upper limits at 2828 and 2871 $cm^{-1}$, respectively. The corresponding integrals were found to be $\int A_{\|} d\tilde{\nu} = 0.381$ $cm^{-1}$, and $\int A_{\perp} d\tilde{\nu} = 0.413$ $cm^{-1}$ resulting in $R^{ATR} = 0.923$. This is the relevant experimental quantity.

*Mean orientation of hydrocarbon chains* Uniaxial orientation, i.e. isotropic distribution of DPPA around the $z$-axis, is assumed. The mean square cosine of the angle between the transition dipole moments of $\nu_s(CH_2)$ of the whole population of $CH_2$ groups of the molecule (28 groups in hydrocarbon chains, 1 in the glycerol part, slightly shifted in frequency) can be calculated from Equation [32], resulting in Equation [37].

$$\langle \cos^2 \alpha \rangle = \frac{(R^{\mathrm{ATR}} - E_x^2/E_y^2)(E_y^2/E_z^2)}{2 + (R^{\mathrm{ATR}} - E_x^2/E_y^2)(E_y^2/E_z^2)} \qquad [37]$$

The squares of relative electric field components at the interface ($z = 0$) in medium 2 as calculated from Equation [29] for $\theta_i = 45°$, $n_1 = 4.0$ (germanium), $n_2 = 1.5$ (DPPA monolayer) and $n_3 = 1.0$ (dry air) result in $E^2_{0x,2} = 1.991$, $E^2_{0y,2} = 2.133$, and $E^2_{0z,2} = 0.450$. It follows that $E^2_{0\|,2} = E^2_{0x,2} + E^2_{0z,2} = 2.441$ and $E^2_{0\perp,2} = E^2_{0y,2} = 2.133$. The dichroic ratio for an isotropic film under these conditions would result in, according to Equation [33], $R^{ATR}_{iso,2} = 1.144$. Explicit equations for relative electric components calculated by means of Harrick's weak absorber approximation can be found in the Further reading section.

Introducing the experimental value of $R^{ATR}$ and the calculated squares of relative electric field components into Equation [37], one obtains for the mean square cosine of the angle between the transition moment of $\nu_s(CH_2)$ and the $z$-axis $\langle \cos^2 \alpha \rangle = -0.025$. This value should not be negative because its minimum is zero; however, since it is small, we consider it to be within experimental and predominately systematic errors. Therefore, we set $\langle \cos^2 \alpha \rangle = 0$, resulting in $\alpha = 90°$. This result requires that all methylene groups of the hydrocarbon chains assume an all-*trans* conformation and, moreover, all hydrocarbon chains are aligned normal to the MIRE, i.e. parallel to the $z$-axis (tilt angle 0°). The exact wavenumber of the symmetric stretching vibration of the $CH_2$ group in glycerol is not known. However, overlapping with $\nu_s(CH_2)$ of the hydrocarbon chains is probable. Consequently, the bisectrice of the glycerol $CH_2$ group may also be concluded to be predominately parallel to the $x$,$y$-plane.

*Mean order parameter of $CH_2$ groups* The mean segmental order parameter resulting from Equation [34] is found to be $\bar{S}_{seg}[\nu_s(CH_2)] = -\frac{1}{2}$. This value is representative of a perfectly ordered molecular entity with isotropic arrangement of transition dipole moments around the $z$-axis and perfect parallel alignment to the interface ($x$,$y$-plane). It should be noted that for $\langle \cos^2 \alpha \rangle = 1$, i.e. transition moments perfectly aligned normal to the interface ($z$-axis), Equation [34] results in the upper limit $\bar{S}_{seg} = 1$. Lipids in natural biomembranes consist of a considerable amount of unsaturated hydrocarbon chains. Since double bonds cause unavoidably *gauche* defects in elongated hydrocarbon chains, which leads to a reduced chain ordering, $\bar{S}_{seg}[\nu_s(CH_2)]$ is increased, reaching zero for an isotropic chain arrangement, since $\langle \cos^2 \alpha \rangle = 1/3$ in this case.

It should be noted that the determination of order parameters of individual methylene groups in the hydrocarbon chains requires generally selective deuteration. In this respect, comprehensive deuterium NMR work should be mentioned (see Further reading section).

A more general case of sample geometry is that of a transition moment being inclined by an angle $\Theta$ with respect to the molecular axis $a$ and isotropically distributed around $a$. Furthermore, the molecular axis $a$ forms an angle $\gamma$ with respect to the tilt axis $t$, and is isotropically distributed around it, and finally, the axis $t$ forming a tilt angle $\delta$ with the $z$-axis and is isotropically distributed around it. In this case, the segmental order parameter, e.g. $S_{seg}[\nu_s(CH_2)]$, may be expressed as superposition of three uniaxial orientations according to

$$\begin{aligned} S_{\mathrm{seg}} &= \left(\frac{3}{2}\langle \cos^2 \delta \rangle - \frac{1}{2}\right) \cdot \left(\frac{3}{2}\langle \cos^2 \gamma \rangle - \frac{1}{2}\right) \\ &\quad \times \left(\frac{3}{2}\langle \cos^2 \Theta \rangle - \frac{1}{2}\right) \\ &= S_\delta \cdot S_\gamma \cdot S_\Theta \end{aligned} \qquad [38]$$

The angles $\delta$, $\gamma$ and $\Theta$ may be distinct or fluctuating (partly or all), describing a microcrystalline ultrastructure (MCU) and a liquid crystalline ultrastructure (LCU), respectively. $S\gamma$ is referred to as the molecular order parameter $S_{mol}$.

Applying Equation [38] to the DPPA monolayer under discussion, one obtains: $S_\delta = 1$, $S_\gamma = 1$, $S_\Theta = -\frac{1}{2}$, meaning no tilt ($\delta = 0$), molecular axis (hydrocarbon chain) normal to the interface ($\gamma = 0$), and transition dipole moment normal to the molecular axis ($\Theta = 90$).

*Surface concentration and area per molecule* The surface concentration may be calculated using Equation [36]. The following additional information is required: (i) the integrated molar absorption coefficient related to a linear baseline from 2828 to 2871 cm$^{-1}$ (see **Figure 9**) was $\int \varepsilon d\tilde{\nu} = 5.7 \times 10^5$ cm mol$^{-1}$, (ii) the real thickness of the layer was assumed to be $d = 2.5$ nm, (iii) the number of equal functional groups $\nu = 28$, and (iv) the effective thicknesses $d_e$ for parallel or perpendicular polarized incident light, which were calculated from Equations [27] and [35], resulted in $d_{e,\parallel} = 3.97$ nm and $d_{e,\perp} = 4.30$ nm. The mean surface concentration was found to be $3.93 \times 10^{-10}$ mol cm$^{-2}$, corresponding to a molecular cross-section of 0.427 nm$^2$ per molecule (42.3 Å$^2$ per molecule). This value leads to the conclusion that the two hydrocarbon chains of a DPPA molecule predominantly determine the area per molecule, since the cross-section of an elongated hydrocarbon chain is 20–21 Å$^2$.

*Conclusions* Quantitative analysis, including orientation measurements, has been shown to be straightforward when the formalism based on Harrick's weak absorber approximation is applied. For thin adsorbed layers, such as the DPPA monolayer under discussion, the results are fairly good. Application to bulk materials may introduce systematic errors as discussed above. If the weak absorber approximation is still to be applied, one should take care to work with an angle of incidence which is at least 15° larger than the critical angle, in order to avoid significant band distortions. In many cases it is possible to use quantitative data from transmission experiments to check the validity of the formalism applied to ATR data.

A general critical aspect concerning the baseline selection should be mentioned. A linear tangential baseline has been used for quantitative analysis of the symmetric $CH_2$ stretching vibration of DPPA (see **Figure 9**). Obviously the correct baseline is lower, i.e. the integrated absorbances used for analysis are systematically too small. The reason for this procedure is only to permit good reproducibility. While the determination of the dichroic ratio is indifferent with respect to the choice of the baseline, it is mandatory to use integrated or peak molar absorption coefficients which have been determined under the same conditions. Even then deviations in the range of several per cent may occur among different operators.

Finally, it should be noted that ATR spectroscopy allows very good background compensation, when adequate equipment is used.

**Sensitivity of modulated excitation (ME) ATR spectroscopy** An impression of the sensitivity of stationary measurements was given the last but one section. A limit of 0.2 mAU is suggested. This limit is beaten by one order or magnitude when the ME technique is applied. As mentioned above, the sample must fulfil the condition of a reversible stimulation by a periodically altered external thermodynamic parameter. Here, the excellent sensitivity and instrumental stability will be demonstrated, for example, with a chemical modulation experiment performed in liquid water, a very strong absorber in the 3400 and 1640 cm$^{-1}$ region.

In order to study the influence of immobilized charges on a lipid model membrane, an arachidic acid (ArAc) bilayer was prepared on a germanium MIRE by means of the Langmuir–Blodgett (LB) technique. The MIRE was transferred in the hydrated state from the LB trough into a flow-through cell and kept in permanent contact with an aqueous buffer solution. Since the carboxylic acid groups of the second monolayer were facing the aqueous phase, the degree of protonation could be controlled via the environmental pH. A periodic pH modulation between pH 3 and 10 induced a periodic protonation and deprotonation of the carboxylic acid group. It should be noted that the first ArAc LB layer was attached by head to the Ge MIRE. Obviously, this binding was so special that typical absorption bands of the carboxylic acid groups were not visible in the spectrum. Therefore, one may assume that the head group signals shown in **Figure 10** result predominately from the outer monolayer of ArAc. The stationary spectral intensity is comparable to that of the DPPA monolayer shown in **Figure 9**. Moreover, one should note that the experiments were performed in $H_2O$, where in the 1640 and 3400 cm$^{-1}$ regions there is very low spectral energy available, favouring perturbations by incomplete background compensation.

In this context, only the sensitivity and selectivity of ME techniques will be discussed. A comprehensive presentation and analysis of polarized pH modulation spectra will be given elsewhere.

*Sensitivity of ME-spectroscopy* Taking the pH-modulated spectrum shown in **Figure 10** as a typical example, one may estimate the sensitivity by comparing the most intense ME spectra ($\phi_{PSD} = 60°/\phi_{PSD} = 90°$, the maximum is expected at $\phi_{PSD} \approx 75°$, consequently, $\phi_{PSD} \approx 165°$ should result intensity zero) with the lowest intensity ME spectrum at $\phi_{PSD} = 0°$. In order to check the S/N ratio, the $\phi_{PSD} = 0°$ spectrum was expanded 25 times in the $CH_2$ wagging region and is plotted as a dashed inset in **Figure 10**. The ordinate scaling factor for the zoomed spectrum is $4.0 \times 10^{-5}$. Comparing it with the other ME spectra (scaling factor $1.0 \times 10^{-3}$) one can conclude that bands as weak as $1.0 \times 10^{-5}$ AU are still detectable.

*Selectivity of pH ME* The highest selectivity of ME spectroscopy is achieved if the stimulation frequency $\omega$ and the kinetics of the stimulated process are matched, i.e. if $0.1 < \omega\tau_i < 10$, where $\tau_i$ denotes the $i$th relaxation time of the system. $\tau_i$ is a function of the rate constants involved in the stimulated process. Under these conditions, significant $\omega$-dependent phase shifts are expected, resulting in $\phi_i = \arctan(-\omega\tau_i)$ for a linear system. Consequently, a molecular or conformational population represented by the relaxation time $\tau_i$ exhibits maximum absorbance in the ME spectrum at a PSD phase setting $\phi_i = \phi_{PSD}$, thus allowing selective detection and kinetic analysis by means of phase-resolved ME spectra. Moreover, $\omega$ acts as an additional experimental degree of freedom in this context, since information on selectivity and kinetics can be spread in the $\omega/\phi_{PSD}$ plane which is more selective than the unidirectional information resulting from conventional relaxation measurements.

In the actual case of pH modulation exerted on a monolayer of ArAc, there is no phase resolution observed, owing to the long modulation period of $\tau_m = 16$ min, i.e. no kinetic information is available. However, unambiguous discrimination between the protonated and deprotonated populations is possible. Only one characteristic example will be given here. The most prominent band from the protonated state is the C=O stretching vibration $\nu$(COOH) of the carboxylic acid group near 1700 cm$^{-1}$. All other bands in the ME spectrum that have the same phase belong to the protonated population, whereas the remaining bands featuring opposite sign are members of the deprotonated population. Consequently, if no

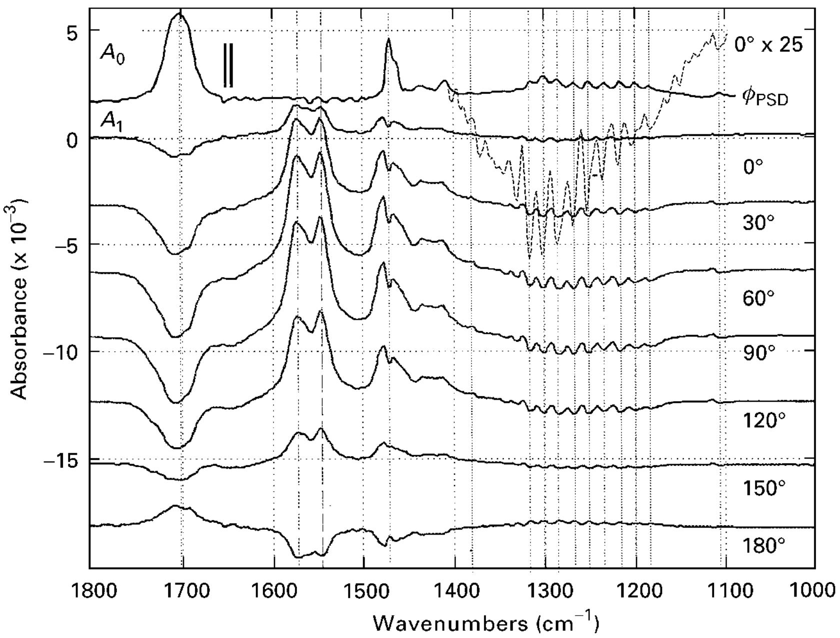



**Figure 10** pH-modulated excitation (ME) of an arachidic acid (ArAc) bilayer attached to a germanium multiple internal reflection element (MIRE). ME was performed by pumping alternatively two buffer solutions (100 mM NaCl, pH 3 and 100 mM NaCl, pH 10) through the ATR cuvette with a modulation period of $\tau = 16$ min. T = 10°C. Upper trace $A_0$; stationary spectrum of a protonated ArAc layer for comparison with modulation spectra. Traces $A_1$; modulation spectra at PSD phase settings $\phi_{PSD} = 0, 30,\ldots, 180°$. The 180° spectrum corresponds to the 0° spectrum with opposite sign, because the PSD output is proportional to cos $(\phi-\phi_{PSD})$, see also **Figure 8**. $\phi$ denotes the phase difference between a given band and the stimulation. Owing to the long period of $\tau_m = 16$ min, the observed bands in the modulation spectra exhibit only two resolved $\phi$ values, which are 180° apart, as a consequence of the fact that the chemical relaxation time of protonation/deprotonation of ArAc is much shorter than the stimulation period. In order to demonstrate the excellent S/N ratio, the ordinate of the weakest modulation spectrum has been expanded in the $CH_2$ wagging region by a factor of 25, i.e. the ordinate scaling factor for the dashed spectrum results in $4.0 \times 10^{-5}$ (see text).

phase resolution is achieved, ME spectra reduce to difference spectra, which, however, have a considerably better background and instability compensation than conventional difference spectra, since corresponding sample and reference spectra are measured and evaluated/accumulated within each period of stimulation.

Consider now the wagging region $\gamma_w(CH_2)$ of the spectra shown in **Figure 10**. The wagging motion $\gamma_w(CH_2)$ is described as in-phase displacement of both H atoms through the H–C–H plane of a methylene group, where the C atom remains predominately in place. In an all-*trans* hydrocarbon chain the transition dipole moment of $\gamma_w(CH_2)$ is expected to be parallel to the chain direction. Deviations may occur, however, from coupling with a polar end group. In the stationary absorbance spectrum $A_0$, one can observe nine weak bands between about 1180 and 1320 $cm^{-1}$. This sequence results from concerted wagging vibrations of all methylene groups in a hydrocarbon chain with an all-*trans* conformation. According to IR selection rules one has to expect $n/2$ IR-active vibrations for an even number $n$ of $CH_2$ groups in an all-*trans* conformation. ArAc has 18 $CH_2$ groups per chain, resulting in the above-mentioned sequence of nine bands in accordance with theory. Since these bands are found to be in phase with ν(COOH), one can conclude that deprotonation of COOH is paralleled by reversible disordering of the chain structure, most probably by introducing *gauche* defects.

Finally, it should be mentioned that $\gamma_w(CH_2)$ belongs to the group of weak absorption bands. One can conclude, therefore, that ME IR ATR spectroscopy allows significant quantitative studies on a molecular level with submonolayer quantities of weak absorbers.

## Manufacturers of reflection accessories

### Standard equipment for reflection spectroscopy

ASI Sense IR Technologies, 15 Great Pasture Road, Danbury, CT 06810, USA.

Bruker Optics, Wikingerstrasse 13, D-76189 Karlsruhe, Germany.

Graseby Specac Inc., 301 Commerce Drive, Fairfield, CT 06432, USA.

Harrick Scientific Corporation, 88 Broadway, Ossining, NY 10562, USA.

International Crystal Laboratories, 11 Erie Street, Garfield, NJ 07026, USA.

Spectra-Tech, Inc., Warrington WA3 7BH, UK.

### Special equipment for SBSR-ATR and ME-ATR spectroscopy

Optispec, Rigistrasse 5, CH-8173 Neerach, Switzerland.

## List of symbols

$A$ = absorbance (decadic); $d$ = sample thickness; $d_e$ = *effective thickness* (Harrick); $d_p$ = penetration depth; $\bar{d}$ = mean penetrated layer thickness; $\boldsymbol{E}$ = electric field; $E_x$ = electric field component in $x$-direction; $E_y$ = electric field component in $y$-direction; $E_z$ = electric field component in $z$-direction; $\bar{J}$ = energy flux (time average); $m_x$ = $x$-component of the unit vector in the direction of $\boldsymbol{M}$; $m_y$ = $y$-component of the unit vector in the direction of $\boldsymbol{M}$; $m_z$ = $z$-component of the unit vector in the direction of $\boldsymbol{M}$; $\boldsymbol{M}$ = transition dipole moment; $n$ = refractive index; $\hat{n}$ = complex refractive index; $N$ = number of active internal reflections; $r$ = ratio of reflected to incident field; $R^{ATR}$ = dichroic ratio related to ATR spectra; $R$ = dichroic ratio; $R$ = reflectance; $S_\gamma = S_{mol}$ = molecular order parameter; $S_\theta$ = tilt order parameter; $S_\theta$ = order parameter with respect to the molecular axis; $S_{seg}$ = segmental order parameter; $S$ = Poynting vector (time average); $t$ = ratio of transmitted to incident field; $T$ = transmittance; $z$ = distance from surface; ∥, ⊥ = parallel and perpendicular polarized light, respectively; * = conjugate complex; $\alpha$ = absorption coefficient; $\alpha$ = angle between transition dipole moment and $z$-axis; $\delta$ = phase shift; $\varepsilon$ = permittivity; $\varepsilon$ = molar absorption coefficient; $\varepsilon_0$ = permittivity of vacuum; $\varepsilon_r$ = relative permittivity (dielectric constant); $\eta$ = amplitude; $\theta_c$ = critical angle; $\theta_i$ = angle of incidence; $\theta_r$ = angle of reflection; $\theta_t$ = angle of refraction (transmission); $\kappa$ = absorption index; $\lambda$ = wavelength; $\mu$ = permeability; $\mu_0$ = permeability of vacuum; $\mu_r$ = relative permeability; $\nu$ = number of equal functional groups per molecule; $\tilde{\nu}$ = wavenumber; $\rho$ = reflectance; $\tau$ = transmittance; $\tau$ = relaxation time; $\Gamma$ = surface concentration; $\phi$ = phase lag; $\omega$ = angular frequency.

*See also:* **Electromagnetic Radiation; Industrial Applications of IR and Raman Spectroscopy; Polymer Applications of IR and Raman Spectroscopy; Raman and IR Microspectroscopy; Surface Studies by IR spectroscopy.**

## Further reading

Baurecht D and Fringeli UP (2000) Surface charge induced conformational changes in an arachidic acid Langmuir-Blodgett bilayer observed by pH-modulated excitation FTIR ATR spectroscopy. *Langmuir* (in preparation).

Beccard B and Hapanowicz R (1995) Polarization Modulation FT-IR spectroscopy. *Nicolet Application Note AN-9592*. Madison, WI: Nicolet.

Blaudez D, Turlet J-M, Dufourcq D, Bard D, Buffeteau T and Desbat B (1996) Investigation at the air-water interface using polarization modulation IR spectroscopy. *Journal of the Chemical Society, Faraday Transactions* **92**: 525–530.

Born M and Wolf E (1983) *Principles of Optics*, Chapter I. Oxford: Pergamon Press.

Fringeli UP (1992) *In situ* infrared attenuated total reflection membrane spectroscopy. In: Mirabella FM (ed.) *Internal Reflection Spectroscopy, Theory and Applications*, Chapter 10, pp 255–324. New York: Marcel Dekker.

Fringeli UP (1997) Simulatneous phase-sensitive digital detection process for time-resolved, quasi-simultaneously captured data arrays of a periodically stimulated system. *PCT International Patent Application*, WO97/08598.

Fringeli UP, Goette J, Reiter G, Siam M and Baurecht D (1998) Structural investigation of oriented membrane assemblies by FTIR-ATR spectroscopy. In: deHaseth JA (ed.) *Fourier Transform Spectroscopy*; 11th International Conference, AIP Conference Proceedings 430, pp. 729–747. Woodbury, New York: American Institute of Physics.

Galant J, Desbat B, Vaknin D and Salesse Ch (1998) Polarization-modulated infrared spectroscopy and X-ray reflectivity of photosystem II core complex at the gas-water interface. *Biophysical Journal* **75**: 2888–2899.

Greenler RG (1966) Infrared study of adsorbed molecules on metal surfaces by reflection techniques. *Journal of Chemical Physics* **44**: 310–315.

Hansen WH (1973) Internal reflection spectroscopy in electrochemistry. In: Delahay P and Tobias ChW (eds) *Advances in Electrochemistry and Electrochemical Engineering*, Vol 9, Muller RH (ed.) *Optical Techniques in Electrochemistry*, pp 1–60. New York: John Wiley & Sons.

Hapke B (1993) *Theory of Reflectance and Emittance Spectroscopy*. New York: Cambridge Univeristy Press.

Harrick NJ (1967) *Internal Reflection Spectroscopy*, New York: Interscience; 2nd edn (1979) Ossining, NY: Harrick, Scientific.

Hoffmann FM (1983) Infrared reflection-absorption spectroscopy of adsorbed molecules. *Surface Science Reports* **3**: 107–192.

Kortüm G (1969) *Reflectance Spectroscopy*. New York: Springer.

Mendelsohn R, Brauner JW and Gericke A (1995) External infrared reflection absorption spectroscopy of monolayer films at the air-water interface. *Annual Review of Physical Chemistry* **46**: 305–334.

Mirabella FM (ed) (1992) *Internal Reflection Spectroscopy, Theory and Application*. New York: Marcel Dekker.

Müller M, Buchet R and Fringeli UP (1996). 2D-FTIR ATR spectroscopy of thermo-induced periodic secondary structural changes of poly-(L)-lysine: A cross-correlation analysis of phase-resolved temperature modulation spectra. *Journal of Physical Chemistry* **100**: 10810–10825.

Picard F, Buffeteau T, Desbat B, Auger M and Pézolet M (1999) Quantitative orientation measurements in thin films by attenuated total reflection spectroscopy. *Biophysical Journal* **76**: 539–551.

Seelig J and Seelig A (1980) Lipid confromation in model membranes and biological membranes. *Quarterly Review of Biophysics* **13**: 19–61.

Tamm L and Tatulian S (1997) Infrared spectroscopy of proteins and peptides in lipid bilayers. *Quarterly Review of Biophysics* **30**: 365–429.

Urban MW (1996) *Attenuated Total Reflectance Spectroscopy of Polymers*. Washington, DC: American Chemical Society.

Wendlandt WWM and Hecht HG (1996) *Reflectance Spectroscopy*, New York: Interscience.

Wenzl P, Fringeli M, Goette J and Fringeli UP (1994) Supported phospholipid bilayers prepared by the "LB/Vesicle Method": A Fourier transform infrared attenuated total reflection spectroscopic study on structure and stability. *Langmuir* **10**: 4253–4264.

Wooten F (1972) *Optical Properties of Solids*. New York: Academic Press.

Zachman G (1995) A rapid and dependable identification system for black polymeric materials. *Journal of Molecular Structure* **348**: 453–456.

Zbinden R (1964) *Infrared Spectroscopy of High Polymers*, New York: Academic Press.

# B

# Biochemical Applications of Fluorescence Spectroscopy

**Jason B Shear**, University of Texas at Austin, TX, USA

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| Applications |

## Introduction

A remarkable range of fluorescence spectroscopy techniques has been developed in the last forty years to characterize fundamental properties of biological systems. The rapid progress over this period has led to strategies for monitoring events that transpire on the femtosecond time scale, for discerning features smaller than the Rayleigh resolution limit of light, and for detecting individual molecules. More routinely, fluorescence spectroscopy has provided invaluable insights into the structure, function, and concentrations of macromolecules, small molecules, lipids, and inorganic ions. Moreover, fluorescence analysis of living systems has begun to reveal directly how these biochemicals function and interact *in situ*. The initial section of this article presents an overview of the measurement formats used in fluorescence studies of biochemical systems. In the sections that follow, the application of diverse spectroscopic techniques to the study of biomolecular systems is examined, considering both the properties and uses of intrinsic biological fluorophores and also of fluorescent probes that have been designed to report on the presence or condition of biochemicals. In the Conclusion, several fluorescence approaches that may offer new capabilities in future biochemical studies are discussed. A fully comprehensive treatment of the diverse applications of fluorescence in biological studies is not possible in this article. For example, fluorescence studies of biological porphyrins – a vast field of spectroscopy – is mentioned in the most cursory fashion. Readers are encouraged to refer to other articles as listed later, for additional information, as well as to the sources in the Further reading list at the end of this article.

## Overview of measurement formats

### *In situ* measurements

In general, when information is sought regarding the spatial distribution of chemicals within cells or tissue, some variant of fluorescence microscopy is used – often in combination with transmission or Nomarski imaging. From an optical standpoint, fluorescence microscopy measurements either use a *wide-field* geometry, in which an entire field-of-view is illuminated simultaneously, or rely on raster scanning of a small resolution element to sequentially generate a fluorescence image. In the first approach, an arc lamp is commonly used as the excitation source, and large-format images can be acquired at video rate using an array detector. In scanning fluorescence microscopy, a laser typically provides the excitation light, which is focused critically to diffraction spot sizes as small as ~0.2 µm using a high numerical aperture (NA) microscope objective. The sequential nature of image formation in scanning microscopy makes this procedure relatively slow. Although large format two-dimensional (2D) images typically require half a second or more to acquire, small 2D images or line scans can be produced much more rapidly. A fundamental advantage of the two scanning techniques most commonly used – confocal and multiphoton microscopies – is an ability to acquire 3D images of various biological specimens by scanning the laser focal spot in a given plane, then shifting the fine focus of the microscope to repeat the process in a new plane. In confocal microscopy, out-of-plane fluorescence is prevented from reaching a single-element detector by placing a small aperture in a plane conjugate to the focal spot; multiphoton

microscopy achieves a similar degree of axial resolution through an inherent 3D localization of the excitation focal volume.

### Measurements with chemical separations

Though microscopy can be used to quantify simultaneously up to several chemical species by acquiring detailed spectroscopic information, when analysis of *many*, potentially unidentified, cellular components is desired, a separation procedure typically is used in combination with fluorescence detection. Commonly used techniques include high-pressure liquid chromatography (HPLC), capillary electrophoresis (CE) and related capillary chromatography techniques, and gel electrophoresis. In HPLC and CE, trace analysis frequently is performed by labelling nonfluorescent components either pre- or post-separation using fluorogenic reagents. Capillary techniques frequently offer several advantages over HPLC, including faster and more efficient separations, lower sample volume requirements, and alleviation of the need for high-pressure equipment. In DNA gel electrophoresis, fluorescent intercalating dyes can be used to stain oligonucleotide bands in gels after electrophoresis. For DNA sequencing in agarose gels, the Sanger method can be used to generate fluorescent fragments by labelling primers with fluorophores whose emission spectra encode a specific terminating dideoxyribo-nucleotide. Analysis of proteins using SDS-polyacrylamide slab gel electrophoresis or various capillary techniques can be accomplished by fluorescently labelling analytes before separation.

### Cuvette, flow, and 'chip' measurements

A large number of studies on isolated biochemical systems can be performed in cuvettes or deep-well slides. In particular, characterizations of macromolecular structure frequently are conducted using thermostatted cells with capabilities for stopped-flow exchange of various solution modifiers (e.g. ligands, denaturants) on a millisecond time scale. In many cases, instruments used for such studies also can perform nanosecond or microsecond time-resolved measurements of emission intensity or polarization in addition to standard steady-state detection after polarization.

In many instances, it is desirable to rapidly count or isolate a subpopulation of cells exhibiting a particular chemical characteristic. In flow cytometry, this property is linked to a corresponding difference in cellular fluorescence (e.g. by selectively labelling a relevant gene product), which provides a means for measuring relatively large numbers of cells in a flowing stream using automated instrumentation. Fluorescence-activated cell sorting (FACS) uses a flow cytometer in combination with a shunting system to isolate cells exhibiting characteristic fluorescence.

The presence of various oligonucleotides, peptides, and small molecules in solution can be analysed using solid state 'chip' based devices, in which receptor molecules (often antibodies or complementary oligonucleotides) immobilized on the chip surface bind the analyte(s) of interest. As a result of analyte binding, a characteristic fluorescence signal is generated from the receptor, analyte, or some other interacting species. In ELISA (enzyme-linked immunosorbent assay), the binding of analyte is ultimately linked to the enzymatic production of a fluorescent solution-phase molecule.

## Analyses based on intrinsic biological fluorescence

Ideally, all biological chemicals would carry highly specific spectroscopic signatures that define not only their identities but also every aspect of their physical states and environments, and they would yield optical signals intense enough to reveal the smallest changes relevant to a particular experiment. Biological analyses, of course, are not so straightforward, and generally provide very limited information in the absence of exogenous 'probe' chemicals. Nevertheless, there are several classes of intrinsically fluorescent biological molecules, and a host of elegant spectroscopic techniques are used to investigate the properties of these species.

The aromatic amino acids – tryptophan (Trp), tyrosine (Tyr) and phenylalanine (Phe) – have strong deep-UV absorption bands ($\lambda_{ex} < 230$ nm) corresponding to $S_0 \rightarrow S_2$ transitions, but commonly are excited to the $S_1$ state in fluorescence studies ($\lambda_{ex} \approx 260$–$280$ nm) to minimize photoreaction and enhance fluorescence quantum yields ($\Phi_f$). At these longer wavelengths, Trp has the largest molar extinction coefficient ($\varepsilon_{max} \approx 5600$) and quantum yield ($\Phi_f \approx 0.2$) of the three amino acids; for Phe, the values of $\varepsilon$ and $\Phi_f$ are so poor that this species is rarely useful in fluorescence studies. When subjected to UV irradiation, proteins with both Trp and Tyr (**Figure 1**) typically exhibit emission spectra whose shape is characteristic of Trp residues ($\lambda_{max} \approx 350$ nm) because of nonradiative energy transfer from Tyr to Trp.

Despite the fact that Trp and Tyr are not highly fluorescent, and undergo intersystem crossing and myriad photoreactions with large quantum yields, these species can be useful probes of protein structure. Emission from both Trp and Tyr is

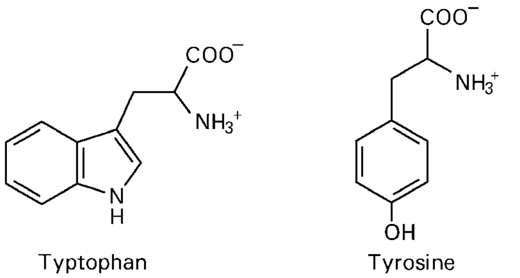



**Figure 1** Fluorescent amino acids.

quenched to different extents by a variety of intramolecular polar moieties (including carboxylic acids, amino groups, and imidazole groups), bound cofactors (e.g. $NAD^+$), prosthetic groups (e.g. heme) and small molecules in solution, making determination of protein structural changes possible through measurement of excited-state lifetimes or steady-state fluorescence intensities. Unlike Tyr, Trp fluorescence often is found to be higher in proteins than in free solution. In molecules containing multiple fluorescent residues, the overall modulation of fluorescence associated with protein conformational changes can be small due to an averaging of advantageous and deleterious quantum yield changes. For this reason, time-resolved measurements of multi-exponential excited state decay times sometimes can provide information unavailable from steady-state intensity measurements.

The Stokes shift for Trp also can depend significantly on environment. When Trp is sequestered in the hydrophobic core of a protein, little or no reorientation of solvent takes place after excitation of the chromophore, and emission is blue-shifted relative to a Trp residue on the exterior of a protein. Extreme examples include azurin, in which Trp appears to be completely isolated from $H_2O$ ($\lambda_{max} < 310$ nm), and denatured parvalbumin, in which the Trp emission spectrum is essentially identical to free Trp. For proteins that contain Tyr but not Trp, the shapes of the emission spectra closely match free Tyr ($\lambda_{max} \approx 305$ nm), regardless of local chemical environment.

A number of macromolecular diffusion and conformational properties can be studied using fluorescence anisotropy, fluorescence correlation spectroscopy (FCS), and fluorescence recovery after photobleaching (FRAP). These techniques most commonly are applied to proteins labelled with highly fluorescent probes, but can exploit intrinsic fluorescence in some instances. In fluorescence anisotropy studies, polarized light is used to selectively excite molecules whose transition dipole moments are aligned with the electric field vector. Steady-state measurements of fluorescence anisotropy can be accomplished using a continuous excitation source; provided that anisotropy decays monoexponentially after excitation, information can be obtained on properties such as rotational diffusion times in membranes and the existence of ligand–host interactions. Time-resolved measurements, in which anisotropy is measured as a function of time following pulsed excitation, can provide more detailed information on phenomena such as anisotropic rotational diffusion, and can be used to study complex processes such as electron transfer in light-harvesting chlorophyll complexes. FCS is a steady-state technique for measuring the concentration and diffusion coefficient of a fluorescent molecule based on the magnitude of signal fluctuations as molecules diffuse through a finite probe volume. Diffusion coefficients can also be determined with FRAP, a technique in which fluorescent molecules within the probe volume are photobleached with a high-intensity pulse of light. After the bleaching event, the kinetics of fluorescence recovery (representing the diffusion of 'fresh' molecules into the probe volume) are monitored with low-intensity irradiation. Depending on the circumstances, FCS or FRAP can be useful for measuring the local viscosity of cellular or isolated biochemical environments, for monitoring ligand/host interactions, and for determining changes in diffusion coefficients associated with macromolecular conformational changes.

Identification and measurement of UV fluorescent proteins in complex biological samples requires mixtures to be fractionated into individual components, often with CE or HPLC. CE with on-column UV fluorescence detection has been shown to be useful in measuring attomole quantities of protein in individual red blood cells. Peptide mapping can be accomplished using HPLC with post-column UV fluorescence detection by purifying individual proteins, then subjecting various fractions to proteolytic digests before separating the resulting peptide segments with chromatography.

Metabolic derivatives of Trp and Tyr also produce significant UV fluorescence. In some secretory cells, Trp is converted to the indolamine neurotransmitter, serotonin, which in turn can be processed into the pineal hormone, melatonin. Metabolic conversion of Tyr yields the catecholamines (dopamine, norepinephrine, and epinephrine). Although the cellular concentration of transmitters is typically low, their concentrations within secretory granules can approach 1 molar. The fluorescence properties of these neurotransmitters are analogous to the parent amino acids, with indolamines exhibiting stronger emission and at longer excitation and emission wavelengths than the catecholamines.

Intrinsic UV fluorescence has been used to perform *in situ* measurements of serotonin. Loss of serotonin from astrocytes can be monitored using wide-field UV fluorescence microscopy after various chemical treatments. In addition, three-photon scanning fluorescence microscopy has been shown to be useful for tracking secretion of serotonin from granules in cultured cells in response to cross-linking of cell surface IgE receptors. Fluorescence from indolamines and catecholamines separated with CE has provided a means to measure these species in the low- to mid-attomole range.

In some instances, protein fluorescence does not derive directly from the aromatic amino acids. Two such proteins native to the jellyfish genus *Aequorea* – *aequorin* and the *green fluorescent protein* (GFP) – have been particularly useful tools in biochemical studies. Aequorin emits light ($\lambda_{max} \approx 470$ nm) from the bound luminophore coelenterazine as the result of a $Ca^{2+}$-promoted oxidation reaction, and hence has been useful as an intracellular $Ca^{2+}$ probe both through microinjection and recombinant gene expression. The discovery and cloning of GFP has attracted much attention recently in part because of its utility as a reporter for various gene products when incorporated into fusion proteins. In GFP, the chromophoric unit has been identified as an imidazolone anion formed by cyclization and oxidation of tripeptide sequence (-Ser-Tyr-Gly-); fluorescence of this species is characterized by excitation maxima at ~400 and ~475 nm and peak emission at ~510 nm. The fluorescence from GFP is intense enough to detect individual molecules of this species immobilized in aqueous gels. Both aequorin and GFP can be genetically targetted to accumulate in specific organelles to measure localized properties, such as [$Ca^{2+}$] and protein diffusion coefficients in mitochondria.

A variety of protein cofactors have significant intrinsic fluorescence, including the reduced nicotinamide nucleotides (NADH and NADPH) and the oxidized flavins (**Figure 2**). The nicotinamides are excited in the near-UV and emit maximally at ~470 nm; flavins exhibit both near-UV and visible excitation maxima, and have a peak emission at ~515 nm. In both groups of cofactors, the adenine group partially quench fluorescence through collisions; in flavins, adenine also can quench fluorescence by forming an intramolecular complex with the fluorescent ring system. The significance of adenine quenching is underscored by the fact that flavin mononucleotide (FMN), a highly fluorescent cofactor that lacks adenine, has a fluorescence quantum yield tenfold greater than that of flavin adenine dinucleotide (FAD). Flavins that are covalently or non-

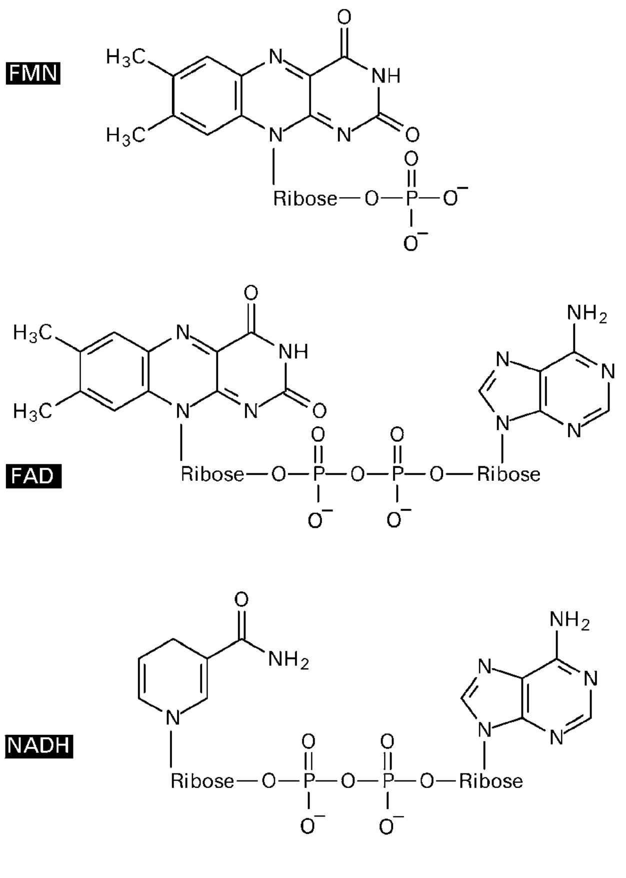



**Figure 2** Three fluorescent cofactors.

covalently bound to proteins display widely varying differences in fluorescence quantum yields, although flavoproteins often are relatively nonfluorescent. In contrast, NADH bound to proteins sometimes has substantially greater fluorescence than free NADH because of a decreased ability of adenine to quench the nicotinamide group. In dehydrogenases, for example, crystallographic studies indicate that nicotinamide cofactors are bound in an extended fashion, with the adenine and nicotinamide groups bound to different pairs of β-sheets. Moreover, conformational changes in lactate dehydrogenase induced by effector molecules can be monitored as changes in NADH fluorescence. As a rule, the redox cofactors – like the fluorescent amino acids – have spectroscopic properties that are sensitive to environment, and can be used to probe conformational changes in proteins.

NADH and oxidized flavin (flavoprotein) fluorescence can be used as an indicator of the redox states of cells challenged with a variety of chemical effectors, and has been used to correlate the oxidative status of cells *in situ* and *in vivo* to functions such as electrical activity in brain tissue. The fluorescence of NADH also has been used to characterize the *in vitro* activity of individual lactate dehydrogenase molecules, which reduce the nonfluorescent $NAD^+$

during oxidative production of pyruvate. This general strategy, in which single-enzyme molecules are characterized through the generation of a fluorescent product, also can be applied to reactions using engineered (nonbiological) fluorogenic substrates.

Most nucleic acids have poor fluorescence properties because of extremely short excited state lifetimes, and are most often studied using nonbiological intercalating or groove-binding dyes. One notable exception is the *Y base*, which has been used as an indicator of tRNA folding induced by divalent cation binding. In addition, some nucleotides exhibit moderate fluorescence under highly acidic conditions, and can be measured in attomole to femtomole amounts using capillary electrophoresis.

## Analyses using nonbiological probes

Although fluorimetric analysis of biological systems is experimentally simpler and less disruptive to inherent chemical properties when the use of synthetic fluorescent molecules can be avoided, in many instances the intrinsic properties of biological chemicals do not provide sufficient means for characterizing or quantifying desired properties. In the case of protein analysis, for example, many species lack Trp or Tyr residues altogether; in protein molecules that do contain these amino acids, the fluorescence characteristics may be inadequate to yield the needed information.

Numerous probes have been developed for investigating biological molecules *in vitro* and in living cells, with greatly differing strategies for uncovering desired information. No common photophysical or photochemical characteristics can be ascribed to this entire assortment of nonbiological fluorophores. Many probes have been engineered to have very large absorption cross sections and fluorescence quantum yields, and to be resistant to photoreaction. The probes based on laser dyes such as fluorescein ($\varepsilon_{max} > 50\ 000$; $\Phi_f > 0.9$) or one of the rhodamines are good examples of this class of molecules, and often are incorporated in a number of the calcium probes, fluorescent antibodies (see below) and reagents used to label proteins for conformational studies involving fluorescence anisotropy, FCS and FRAP. In addition, by conjugating two spectroscopically distinct probes to different sites on a molecule, Förster energy transfer between the fluorophores can be used to measure intramolecular distances. Despite the advantages of these highly fluorescent molecules, probes having *less* optimal fluorescence properties are deliberately selected for some applications – often as a compromise to achieve a greater *change* in optical characteristics under different chemical environments, or to avoid toxicity to living cells. For example, a dye whose fluorescence properties depend sensitively on the polarity and rigidity of its local environment may be more useful in many instances than a rhodamine for characterizing protein conformation. One such compound is NBD chloride, an amine-reactive reagent whose product has a markedly greater $\Phi_f$ value when occupying a hydrophobic site.

Reactive fluorogenic reagents (e.g. NBD chloride, fluorescamine and CBQCA) are essentially nonfluorescent until they react with analyte molecules, and are commonly used to tag biological species that share a particular reactive group (e.g. a thiol or amine). This low-specificity labelling approach is extremely useful for identification and measurement of many low-concentration species in a mixture when some means exist for fractionating the multiple components (e.g. HPLC). Because of the low level of reagent fluorescence, a large excess of these species typically can be present in the reaction mixture without interfering with analysis.

When characterizing samples containing many biomolecules, a probe that reacts with a particular chemical moiety sometimes does not provide adequate specificity, even when used with a highly efficient separation procedure. Analysis of the protein distribution in cellular specimens is perhaps the quintessential high-specificity requirement. Because of the value of such measurements in characterizing cellular composition, immunohistochemistry is one of the most ubiquitous biological applications of fluorescence. In routine determinations of gene products, cultured cellular samples or sectioned tissue are chemically fixed and subjected to sequential application of primary and secondary antibodies, followed by imaging with fluorescence microscopy. Primary antibodies can be raised against a tremendous diversity of cellular species, ranging from membrane proteins to small diffusable species such as neurotransmitters. Several species can be analysed in a single specimen by covalently attaching dyes with different excitation/emission spectra to different secondary antibodies. Most commonly, fixed and labelled samples are analysed using either wide-field or confocal laser scanning microscopy. When one seeks to analyse a solution sample for the presence of a known antigen, ELISA can be used. In one format, an immobilized antibody binds the antigen, removing it from solution. A second antibody which is conjugated to an enzyme capable of converting a nonfluorescent substrate into a fluorescent product then is applied, and binds to the immobilized antigen. After a washing step, the enzymatic reaction is initiated. In this way, large signal amplifications

provide extremely sensitive assays for analytes such as hormones and drug metabolites. Although immunological techniques can be useful for measuring analyte concentrations, uncertainties in cross-reactivity and matrix effects on the antibody-antigen conjugation efficiency often limit this approach to semiquantitative determinations.

Diffusable cytosolic species (e.g. second messengers) can be measured in living cells using highly specific fluorescent probes. The calcium-sensitive dyes, a broad class of molecules whose excitation or emission properties are dependent on chelation of $Ca^{2+}$, are ubiquitous in cellular biology studies. For such compounds, it is important that affinity for $Ca^{2+}$ is much greater than for other cationic species that may be present in the cytosol at much higher concentrations (e.g. $Mg^{2+}$). In general, cytosolic free $Ca^{2+}$ concentrations vary over the approximate range 100 to 1000 nM, although much higher levels are sometimes reached transiently. Common examples of $Ca^{2+}$-sensitive dyes include fluo-3 (100-fold intensity increase when bound to $Ca^{2+}$), fura-2 (excitation $\lambda_{max}$ changes when bound to $Ca^{2+}$), and indo-1 (emission $\lambda_{max}$ changes when bound to $Ca^{2+}$) (**Figure 3**). For these species, the $Ca^{2+}$ dissociation constant has been designed to approximately match free cytosolic levels, thus maximizing the modulation in fluorescence for a given change in [$Ca^{2+}$]. Like tryptophan, indo-1 has an indole-based chromophore, but in this case the $\pi$-system is further delocalized to electronically interact with carboxylate groups responsible for $Ca^{2+}$ chelation. In many instances, fura-2 and indo-1 may offer advantages over fluo-3, as ratiometric measurements of excitation or emission intensities at two different wavelengths can help avoid errors associated with uncertainties in intracellular dye concentrations. Calcium probes have played an invaluable role in identifying a range of cytosolic phenomena, such as [$Ca^{2+}$] spiking after hormone binding to cell surface receptors and $Ca^{2+}$ waves in oocytes following a fertilization event. Probes exist for other important cytosolic effectors, such as cyclic AMP (cAMP), which is monitored by the level of Förster energy transfer between fluorescein and rhodamine attached to a cAMP-binding enzyme.

Although it is possible to microinject dyes for probing cytosolic milieus through patch or intracellular pipettes, it is more common to incubate many cultured cells in a medium containing low molecular mass dyes in an esterified form (e.g. acetoxymethyl, or 'AM', ester). This parent species is hydrophobic, and thus can diffuse freely through cell membranes. Once the compound is localized to the cytosol, nonspecific esterases hydrolyse the hydrophobic group, generating a charged form of the probe that

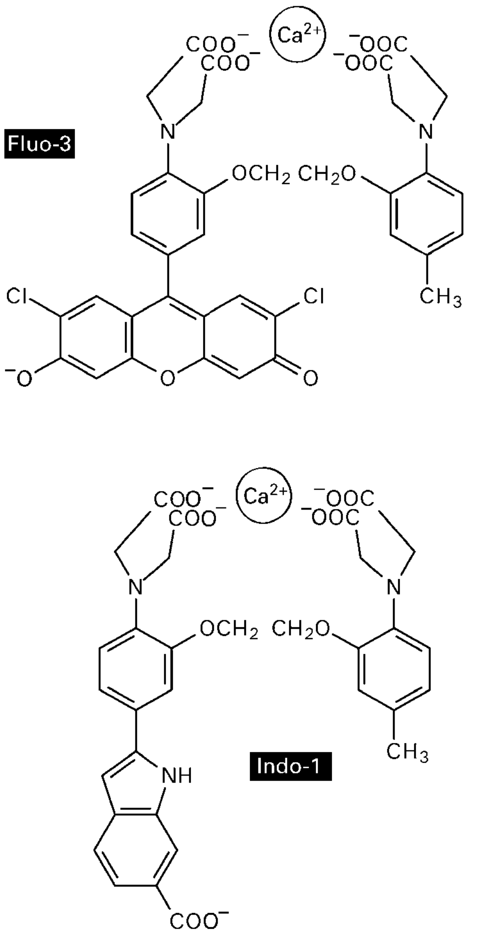



**Figure 3** Two common calcium probes.

can no longer diffuse through the membrane. Large numbers of cells also can be loaded rapidly using electroporation.

Structures in living cells often can be characterized using probes with specificity for *classes* of molecules rather than for particular target analytes. Hydrophobic or amphipathic dyes, for example, can be used to gauge cell membrane fluidity and diffusion, as well as exocytosis of secretory vesicles and subsequent membrane recycling. Examples of this class of dyes are 8-anilinonaphathalene sulfonate (ANS), which displays significant fluorescence only when bound to membranes (or protected hydrophobic regions of proteins), and FM-143, a compound used extensively to study regulated secretion from neurons. Some membrane dyes have been engineered (or serendipitously found) to alter their fluorescence properties based on the electric field across the phospholipid bilayer of cell or organelle membranes. Typically, such potentiometric dyes either respond with fast kinetics to electric field changes but display small modulations in fluorescence (e.g. the styrylpyridinium

probes), or undergo large fluorescence changes but only very slowly. Because large, rapid changes in fluorescence generally are not obtained with existing probes, it is difficult to perform voltage measurements on individual neurons and other excitable cells.

Other fluorogenic dyes, such as DAPI, ethidium bromide and the cyanine dyes, bind to DNA and/or RNA and undergo changes in fluorescence intensity of up to 1000-fold. Depending on the probe, these compounds can be used to image fixed cells or track mitotic events. Combinations of probes with different specificities for single-stranded, double-stranded and ribosomal RNA have been used to characterize the quantity and conformation of nucleic acids throughout the cell cycle.

## Conclusion

In the forty years since the intrinsic fluorescence of amino acids was first characterized, fluorescence spectroscopy has developed into one of the most versatile and powerful tools for investigating biochemical systems. In addition to the enormous range of applications that now exist, a variety of emerging analysis strategies provide evidence of a continued phase of rapid growth for fluorescence applications. Developments in instrumentation – laser sources, optics, and detectors – have made possible the characterization of individual enzyme molecules and highly fluorescent proteins such as GFP and β-phycoerythrin, a light-harvesting protein. Solid-state femtosecond sources recently have made it feasible to generate multiphoton-excited fluorescence of biological molecules in solution using relatively low integrated irradiation, and have provided reproducibility necessary for experiments involving living specimens.

Various technologies that rely on fluorescence for high sensitivity analysis have been made possible through advances in microfabrication and robotics. Examples include chip-based microarray sensors for detecting a multitude of ligands simultaneously, and chips containing arrays of microscopic electrophoresis channels for performing high-throughput, rapid separations of DNA fragments.

Other fluorescence techniques exploit the evanescent, or near-field, properties of light to excite fluorescence in highly restricted regions of space, thereby improving measurement sensitivity or the spatial resolution of fluorescence imaging. Near-field scanning optical microscopy (NSOM) can image structures smaller than the diffraction limit of light using fluorescence generated by an evanescent field that escapes from the aperture at the end of a drawn, metallized fibre. The resolution obtained with this approach is determined by the dimensions of the aperture, which is usually limited by losses in throughput as the fibre tip diameter is reduced. Thus far, only a few biological applications of this powerful technique have been reported.

New developments in the chemistry of fluorescent probes undoubtedly will open new applications for fluorescence in biochemistry. A growing number of companies and academic researchers are devising new biological probes, and in some cases are using nontraditional strategies – such as combinatorial chemistry – to generate hosts with high specificity and large binding constants for desired ligands. Genetic manipulation of cultured cells offers particularly exciting opportunities for *in situ* biochemical measurements. Technology now exists for using an enzyme as a reporter for neurotransmitter-activated transcription in individual mammalian cells, with sensitivities capable of measuring fewer than 50 gene product molecules. In an initial demonstration of this strategy, large amplifications of gene product expression were obtained by localizing an engineered substrate to the cytosol whose fluorescence properties change when it is degraded by the reporter enzyme. Through the judicious use of molecular and cellular biology, chemistry, and fluorescence spectroscopy, eventually it may be feasible to track the production and breakdown of individual molecules in living cells.

## List of symbols

$\varepsilon$ = molar extinction coefficient; $\lambda$ = wavelength; $\Phi$ = quantum yield.

*See also:* **Fluorescence Microscopy, Applications; Fluorescence Polarization and Anisotropy; Fluorescent Molecular Probes; Inorganic Condensed Matter, Applications of Luminescence Spectroscopy; Luminescence, Theory; Organic Chemistry Applications of Fluorescence Spectroscopy; UV-Visible Absorption and Fluorescence Spectrometers; X-ray Fluorescence Spectrometers; X-ray Fluorescence Spectroscopy, Applications.**

## Further reading

Cantor CR and Schimmel PR (1980) *Biophysical Chemistry*, Parts I–III. New York: Freeman.

Chalfie M, Tu Y, Euskirchen G, Ward WW and Prasher DC (1994) Green fluorescent protein as a marker for gene expression. *Science* **263**: 802–805.

Cobbold PH and Rink TJ (1987) Fluorescence and bioluminescence measurement of cytoplasmic free calcium. *Biochemical Journal* **248**: 313–328.
Craig DB, Arriaga E, Wong JCY, Lu H and Dovichi NJ (1998) Life and death of a single enzyme molecule. *Analytical Chemistry* **70**: 39A–43A.
Creed D (1984) The photophysics and photochemistry of the near-UV absorbing amino acids – I. Tryptophan and its simple derivatives. *Photochem. Photobiol.* **39**: 537–562.
Creed D (1984) The photophysics and photochemistry of the near-UV absorbing amino acids – II. Tyrosine and its simple derivatives. *Photochem. Photobiol.* **39**: 563–575.
Everse J, Anderson B and You K-S (1982) *The Pyridine Nucleotide Coenzymes*. New York: Academic Press.
Freifelder D (1982) *Physical Biochemistry*, 2nd edn. New York: W.H. Freeman.
Hoagland RP (1996) *Handbook of Fluorescent Probes and Research Chemicals*, 6th edn. Eugene, OR: Molecular Probes.
Lakowicz JR (ed) (1991) *Topics in Fluorescence Spectroscopy*, Vol 1, *Techniques*. New York: Plenum Press.
Lakowicz JR (ed) (1992) *Topics in Fluorescence Spectroscopy*, Vol 3, *Biochemical Applications*. New York: Plenum Press.
Lakowicz JR (1983) *Principles of Fluorescence Spectroscopy*. New York: Plenum Press.
Lillard SJ, Yeung ES, Lautamo RMA and Mao DT (1995) Separation of hemoglobin variants in single human erythrocytes by capillary electrophoresis with laser-induced native fluorescence detection. *J. Chromatogr. A* **718**: 397–404.
Permyakov EA (1993) *Luminescent Spectroscopy of Proteins*. Boca Raton, FA: CRC Press.
Rutter GA, Burnett P, Rizzuto R *et al* (1996) Subcellular imaging of intramitochondrial $Ca^{2+}$ with recombinant targeted aequorin: significance for the regulation of pyruvate dehydrogenase activity. *Proceedings of the National Academy of Science, USA* **93**: 5489–5494.
Stryer L (1995) *Biochemistry*, 4th edn. New York: W.H. Freeman.
Teale FWJ and Weber G (1957) Ultraviolet fluorescence of the aromatic amino acids. *Biochemical Journal* **65**: 476–482.
Tsien RY (1994) Fluorescence imaging creates a window on the cell. *Chemical Engineering News* **72**: 34–44.
Zlokarnik G, Negulescu PA, Knapp TE *et al* (1998) Quantitation of transcription and clonal selection of single living cells with beta-lactamase as reporter. *Science* **279**: 84–88.

# Biochemical Applications of Mass Spectrometry

**Victor E Vandell** and **Patrick A Limbach**,
Louisiana State University, Baton Rouge, LA, USA

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Applications |

Mass spectrometry is a powerful tool for the characterization of various biomolecules including proteins, nucleic acids and carbohydrates. The advantages of mass spectrometry are high sensitivity, high mass accuracy, and more importantly, structural information. Historically, biomolecules have proven difficult to characterize using mass spectrometry. Problems often arise with impure samples, low ion abundance for analysis due to inefficient ionization processes, and low mass accuracy for higher molecular weight compounds. Recent advances in the development of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) now permit the analysis of biomolecules with high sensitivity and good mass accuracy. These improvements now allow the use of mass spectrometry for the identification of unknown structures and are sui for applications focused on acquiring sequence information on the samples of interest.

This article provides a brief introduction to the various applications of mass spectrometry to biomolecule analysis. More detailed discussions of particular applications of mass spectrometry to such analyses can be found in other articles in this encyclopedia.

## Ionization and mass analysis

Mass spectrometric measurements of biomolecules involve the determination of molecular mass, or of the masses of various components of the original molecule, which can then be related to various structural properties, such as sequence. In either case, the crucial step lies in the conversion of liquid- or solid-phase solutions of the analytes into gaseous ions. Common ionization sources used in mass spectrometry for biomolecule analysis experiments are: fast-atom bombardment (FAB), ESI and MALDI. FAB was

historically the ionization method of choice, but has been largely replaced by ESI or MALDI at the present time.

ESI- and MALDI-based methods have particularly benefited the analysis of biomolecules because these techniques accomplish the otherwise experimentally difficult task of producing gas-phase ions from solution species that are both thermally labile and polar. **Table 1** summarizes the different instrument configurations used in the analysis of various classes of biomolecules.

MALDI has extremely high sensitivities with reports of detection limits at the sub-femtomolar level. A higher efficiency for protonated molecular ion production with MALDI has been observed relative to FAB. MALDI typically yields intact protonated molecular ions with minimal fragmentation and is therefore commonly referred to as a 'soft' ionization process. MALDI sources are generally coupled to time-of-flight (TOF) mass analysers. TOF mass analysers are characterized by high upper mass limits with reduced resolution at the higher masses. The production of high molecular weight ions and subsequent analysis of these ions makes the MALDI-TOF combination a powerful tool for biomolecule analysis.

ESI can be considered a complementary method to MALDI. As with MALDI, electrospray ionization of biomolecules yields protonated or cationized molecular ions with little or no fragmentation, and it is also referred to as a 'soft' ionization source. A particular advantage of ESI compared to MALDI is that the analyte is sampled from the solution phase. Under these conditions, ESI is readily coupled to high performance liquid chromatography (HPLC) or capillary electrophoresis (CE) separation systems. Such combinations permit online LC-MS or CE-MS experiments.

**Table 1** Summary of mass spectrometry techniques used in biomolecule analysis

| *Analyte* | *Ionization source* | *Mass analyser* |
|---|---|---|
| Proteins/peptides | FAB | Quadrupole, sector |
| | MALDI | TOF, FTICR |
| | ESI | Quadrupole, sector, FTICR, TOF |
| Oligonucleotides | FAB | Quadrupole, sector |
| | MALDI | TOF, FTICR |
| | ESI | Sector, FTICR, TOF |
| Oligosaccharides | FAB | Quadrupole, sector |
| | MALDI | TOF |
| | ESI | Quadrupole, sector, TOF |
| Lipids | FAB | Quadrupole |
| | MALDI | TOF |
| | ESI | Quadrupole, FTICR, TOF |

The overriding feature of an ESI-generated mass spectrum is the appearance of multiply charged ions. Multiple charging results from the loss or addition of multiple hydrogen ions or metal ions (e.g. potassium or sodium) to the biomolecule. The multiple charging effect is advantageous because it allows for the analysis of high molecular weight biomolecules using mass analysers with low $m/z$ limits. The disadvantage of multiple charging is that a spectrum has to be deconvoluted and thus spectral interpretation can be complicated, especially during the analysis of mixtures.

## Molecular weight determinations of biomolecules

Relative mass is an intrinsic molecular property which, when measured with high accuracy, becomes a unique and unusually effective parameter for characterization of synthetic or natural biomolecules. Mass spectrometry based methods can be broadly applied not only to unmodified synthetic biomolecules, but also to modified synthetic and natural biomolecules (e.g. glycosylated proteins). The level of mass accuracy one obtains during the measurement will depend on the capabilities of the mass analyser used. Quadrupole and TOF instruments yield lower mass accuracies than sector or Fourier transform ion cyclotron resonance (FTICR) instruments. High mass accuracy is not only necessary for qualitative analysis of biomolecules present in a sample, but is necessary to provide unambiguous peak identification in a mass spectrum.

The primary challenge to accurate molecular weight measurements of biomolecules is reduction or complete removal of salt adducts. In the majority of situations, the buffers used to prepare or isolate the analyte of interest contains Group 1 or Group 2 metal salts. These metal salts can potentially interfere with the accurate mass analysis of the analyte due to the gas-phase adduction of one to several metal cations to the analyte. Optimal results are obtained only after substantive (and in some cases, exhaustive) removal of these contaminants. Recent developments for sample purification involve the use of solution additives or online purification cartridges which reduce the presence of interfering salts while retaining the ability to characterize minimal amounts of analyte.

Measurement of molecular mass of biomolecules is now a suitable replacement for prior methods based on the use of gel electrophoresis. Molecular weight measurement for all classes of biomolecules is a relatively routine procedure, and the results obtained are typically of greater accuracy than those previously

obtained by gel electrophoresis. A common application of mass spectrometry and molecular weight measurement is for the identification of unknown proteins. The experimentally obtained molecular weight value can be searched against the available protein databases (e.g. Swiss–Prot or PIR) to potentially identify the protein. This application becomes particularly useful for the identification of unknown proteins when used in conjunction with enzymatically generated peptide fragments (see below).

## Determination of the primary sequence of biomolecules

One of the most popular and productive uses of mass spectrometry for biomolecule analysis is the sequence or structure determination through analysis of smaller constituents of the original molecule. Two different approaches for generating the smaller, sequence informative constituents from the original molecule are available. An indirect approach is to generate sequence-specific information by solution-based chemical or enzymatic reactions. Chemical or enzymatic digestion of an intact biomolecule results in a number of smaller constituents that are amenable to mass spectrometric analysis. The so-called direct approach involves fragmentation of the analyte in the gas phase. Fragmentation can be induced by the desorption or ionization process, or can be induced by collisions with neutral target molecules or by collisions with surfaces. The analysis of fragment ions initiated by gas-phase dissociation resulting from collisions with neutral molecules or surfaces are broadly referred to as tandem mass spectrometry (MS/MS) experiments. With both approaches, the sequence of the original molecule is obtained by interpreting the resulting fragments.

### Enzymatic approaches

Enzymatic digestion followed by mass spectrometric analysis of the resulting products is a popular and powerful approach to sequence determination. A number of enzymes are available which either are specific for a particular substituent of the biomolecule or are nonspecific but sequentially hydrolyse the analyte of interest. Trypsin is an enzyme commonly used to digest proteins into smaller peptides. Trypsin selectively cleaves proteins at the C-terminal side of lysine and arginine amino acid residues. Digestion of a protein using trypsin will generate a tryptic digest whose components are amenable to mass spectrometric analysis. In many cases, the masses of the tryptic fragments can be measured more accurately than the mass of the original molecule, thereby improving the identification of unknown proteins. As mentioned earlier, the masses of tryptic peptides, used in conjunction with molecular weight measurements of intact proteins, can be used to search known protein databases for efficient and accurate identification of unknown proteins.

Sequential digestion of biomolecules is an alternative approach to determining sequence information. The utility of any mass spectrometric sequencing method that relies on consecutive backbone cleavages depends on the formation of a mass ladder. The sequence information is obtained by determining the mass difference between successive peaks in the mass spectrum. For example, phosphodiesterases are enzymes which sequentially hydrolyse the phosphodiester linkage between oligonucleotides and nucleic acids. In the case of oligodeoxynucleotides, the expected mass differences between successive peaks will correspond to the loss of: dC = 289.5, dT = 304.26, dA = 313.27, and dG = 329.27 Da. Mass ladder methods have a distinct advantage for sequence determination, because it is the difference in two mass measurements that results in the desired information. A drawback to this approach is the limited size of the analyte that is amenable to sequential digestion.

### Tandem mass spectrometry approaches

The gas-phase approach to determining structural information about biomolecules is through tandem mass spectrometry (MS/MS). Tandem mass spectrometry involves isolation of the ion of interest (commonly referred to as the parent ion in MS/MS) and then dissociating this ion via collisions with neutrals or surfaces to produce fragment ions (commonly referred to as product ions in MS/MS) from which the primary sequence of the molecular ion of interest can be determined. Generally, the fragmentation process in MS/MS experiments generates product ions which contain sequence specific information from throughout the molecular ion of interest.

Tandem mass spectrometry is applicable to all types of biomolecules but typically requires specialized mass spectrometry instrumentation for implementation. The most common tandem mass spectrometer is a triple quadrupole instrument, wherein the first and third quadrupoles are used as mass analysers and the middle quadrupole is used as a collision chamber. MS/MS can be performed with double focusing sector instruments and in some cases with specialized TOF mass analysers. Quadrupole ion traps and FTICR mass spectrometers are ideally suited for MS/MS experiments, and due to the operational characteristics of these mass analysers

additional MS/MS experiments can be performed, allowing for $MS^n$ studies of biomolecules.

## Higher order gas-phase structure and non-covalent interactions studied using mass spectrometry

Similar to studies performed using NMR, hydrogen/deuterium (H/D) exchange experiments on biomolecules are feasible with mass spectrometry. Labile hydrogens can be exchanged in solution prior to mass spectrometric analysis, or alternatively H/D exchange experiments can be performed in the gas phase. The former studies have been used to localize sites of H/D exchange on biomolecules of interest and are used for mechanistic studies of biomolecule function. The latter studies typically involve ESI-MS and permit investigations into the gas-phase conformation of biomolecules in the absence of the solvent. The majority of H/D exchange experiments have focused on peptide and protein analysis, although the method is amenable to other classes of biomolecules.

A particular advantage of ESI-MS for biomolecule analysis is realized by generating the analyte ions from solution conditions that retain the secondary, tertiary and even quaternary structure of the biomolecules. Noncovalent binding of biomolecules has been observed in the ESI mass spectrum and, when operated under the appropriate conditions, the mass spectral data are a direct probe of the solution-phase biomolecule assembly. Protein assemblies, protein-nucleic acid complexes, duplex DNA and other noncovalently bound biomolecule assemblies have been studied using mass spectrometry.

*See also:* **Chemical Structure Information from Mass Spectrometry; Chromatography–MS, Methods; Fast Atom Bombardment Ionization in Mass Spectrometry; Fragmentation in Mass Spectrometry; Hyphenated Techniques, Applications of in Mass Spectrometry; Medical Applications of Mass Spectrometry; MS-MS and $MS^n$; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Peptides and Proteins Studied Using Mass Spectrometry; Quadrupoles, Use of in Mass Spectrometry; Sector Mass Spectrometers; Surface Induced Dissociation in Mass Spectrometry; Time of Flight Mass Spectrometers.**

## Further reading

Fenn JB, Mann M, Meng CK, Wong SF and Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**: 64–70.

Karas M, Bahr U and Hillenkamp F (1989) UV laser matrix desorption/ionization mass spectrometry of proteins in the 100,000 Dalton range. *International Journal of Mass Spectrometry and Ion Processes* **92**: 231–242.

Loo JA (1995) Bioanalytical mass spectrometry: many flavors to choose. *Bioconjugate Chemistry* **6**: 644–665.

McCloskey JA (ed) (1990) *Methods in Enzymology*, Vol. 193 San Diego: Academic Press, Inc.

Senko MW and McLafferty FW (1994) Mass spectrometry of macromolecules: Has its time now come? *Annual Review of Biophysics and Biomolecular Structure* **23**: 763–785.

Smith RD, Loo JA, Edmonds CG, Barinaga CJ and Udseth HR (1990) New developments in biochemical mass spectrometry: Electrospray ionization. *Analytical Chemistry* **62**: 882–899.

# Biochemical Applications of Raman Spectroscopy

**Peter Hildebrandt**, Max-Planck-Institut für Strahlenchemie, Mülheim, Germany
**Sophie Lecomte**, CNRS-Université Paris VI, Thiais, France

Copyright © 1999 Academic Press



Until the late 1960s, Raman spectroscopy was largely restricted to small inorganic and organic molecules. Concomitant to the substantial improvement of excitation sources and detection systems beginning in the early 1970s, biological molecules, which in most cases are weak Raman scatterers, became more and more accessible to this technique. Now, Raman spectroscopy has been developed to a versatile tool for studying biological systems on various levels of complexity, i.e. ranging from small amino acids to biopolymers and even to living cells.

Raman spectra – like IR spectra – include detailed information about the molecular structure and, hence, may provide a key for elucidating structure–function relationships. Unfortunately, the ability to extract structural data from the spectra is much less advanced than, for example, in NMR spectroscopy, since a sound vibrational analysis based on normal mode calculations is not straightforward. Empirical force fields only provide meaningful results in the case of small and/or symmetric molecules for which a large set of isotopomers are available. Alternatively, quantum chemical methods can be employed to calculate the force constants. Despite the recent progress in hard- and software development, this approach is as yet restricted to prosthetic groups and building blocks of biopolymers. Thus, in most cases the interpretation of the Raman spectra of biomolecules is based on empirical relationships derived from comparative studies of model compounds.

The large number of normal modes of biomolecules is also associated with an experimental difficulty inasmuch as individual Raman-active modes may be closely spaced so that the observed peaks include several unresolved bands. This is particularly true for those modes which originate from chemically identical building blocks of the biopolymers such as the amide vibrations of proteins. In these cases, however, the analysis of these bands provides valuable information about the structure of the ensemble of these building blocks, e.g. the secondary structure of proteins.

The selectivity of Raman spectroscopy can be substantially increased by the choice of an appropriate excitation wavelength. When the excitation energy approaches that of an electronic transition of a molecule, exclusively the Raman intensities of the modes originating from this chromophore are enhanced by several orders of magnitude (resonance Raman, RR). In this way, it is possible to probe the vibrational band pattern of this specific part of the macromolecule regardless of the size of the remaining optically transparent matrix.

## Specific experimental considerations

Under non-resonant conditions, quantum yields for the Raman effect are in the order of $10^{-9}$, implying that sample concentrations in the millimolar range are required for obtaining Raman spectra of satisfactory quality. Resonance enhancement substantially improves the sensitivity, but even RR intensities are too weak to obtain spectra of samples which exhibit a strong fluorescence or which contain strongly fluorescent impurities. In those cases, Fourier-transform near-infrared Raman spectroscopy may be a useful alternative approach.

Possible thermal denaturation or photochemical processes which may be induced by laser irradiation may impose serious constraints on Raman (RR) spectroscopic experiments of biomolecules. Such effects can be reduced by using appropriate sample arrangements such as flowing or rotating devices, or by lowering the temperature.

In any case, the photon flux through a volume element of a sample should be kept as small as possible to ensure the integrity of the biomolecule during the Raman experiment. These requirements are best fulfilled using CW-lasers as excitation sources. A second parameter which is essential for measuring the spectra of sensitive biomaterial is a short duration of the Raman experiment. Scanning monochromatic detection systems (photon counting devices) may require several hours to obtain Raman spectra of sufficient quality. Polychromatic detection systems can strongly reduce the total signal accumulation time and, in the case of CCD detectors, yield a spectral resolution and signal-to-noise ratio comparable to scanning systems.

## Structure of biomolecules

Raman spectroscopic studies which are directed to gain information about the structure of biomolecules are generally performed in the stationary mode. Such experiments have been carried out with all kinds of biopolymers and their building blocks. However, up to now, the level of spectral analysis and interpretation is such that reliable structural information can be extracted from the spectra only for proteins, nucleic acids and membranes.

### Proteins and amino acids

Non-resonant Raman spectra of proteins display bands originating from the protein backbone, i.e. the amide modes and the amino acid side chains, in particular, the aromatic amino acids tryptophan (Trp), tyrosine (Tyr) and phenylalanine (Phe). The strongest amide band in the non-resonant Raman spectra originates from the amide I mode which is essentially a pure C=O stretching of the peptide bond. Its frequency depends on the type of hydrogen bonding and dipole–dipole interactions associated with the peptide bond. As these interactions are different for the various elements of secondary structure of proteins, systematic Raman and IR spectroscopic studies of peptides and proteins with known three-dimensional structure allow the determination of amide I frequencies which are characteristic not only for the main secondary structure elements α-helix (1657 $cm^{-1}$) and β-sheet (1626 and 1635 $cm^{-1}$) but also for turns (1666 $cm^{-1}$) and β-sheet/turns (1679 $cm^{-1}$). For irregular or random coil peptide bonds the amide I is found at 1644 $cm^{-1}$. These findings constitute the basis for the secondary structure determination of proteins and peptides of unknown structure. In the most general way, the measured amide I envelope is fitted by a set of these bands with constant frequencies and half widths. The relative intensities, which are the only adjustable parameters, represent a measure for the relative contributions of the various secondary structure elements. This approach is also frequently used in IR spectroscopy which offers the advantage of a better signal-to-noise ratio but is restricted to measurements in $D_2O$ due to the strong absorption of water. With an increased set of reference proteins, the method has gained a considerable accuracy which can well compete with CD measurements.

Complementary information about the secondary structure can be obtained by using RR spectroscopy. Upon excitation in resonance with the $\pi \rightarrow \pi^*$ transitions of the peptide bonds at ~190 nm, the amide modes II, III and II′, which include the C–N stretching vibrations, are preferentially enhanced so that they appear with substantially higher intensity than the amide I mode (C=O stretching) which in turn gains intensity only via a higher lying electronic transition at ~160 nm. Like the amide I mode, the amide modes II, III and II′ are at different frequencies in the various structural elements. More important for secondary structure determination, however, are the RR intensities of these modes. The qualitatively different dipole–dipole interactions in α-helical and β-sheet structures lead to a respective decrease (hypochromic effect) and increase (hyperchromic effect) of the oscillator strength which in turn affects the Raman cross-sections of these amide modes. Thus, it was possible to derive (linear) relations between the RR intensities of these modes (determined relative to an internal standard) and the secondary structure. Due to the high energy of the deep UV laser pulses required as the excitation source, special attention has to be paid to avoiding photodegradation processes of the biomolecule.

Such effects are also a restriction on probing the aromatic amino acid side chains of proteins using excitation lines between 200 and 240 nm. Under these conditions, several ring vibrational modes are selectively enhanced via the $L_a$ (Phe, Tyr) and $B_{a,b}$ transitions (Trp). The frequencies and relative intensities of these modes can be used to study the local environment of these amino acids, i.e. to obtain information about this specific part of the tertiary structure. In particular, the modes of Tyr and Trp may provide more detailed information about the local protein structure. The *p*-hydroxy substituent of Tyr is generally involved in hydrogen bonding interactions with hydrogen bond acceptors or donors. These interactions are sensitively reflected by the modes $\nu_{7a}$, $\nu_{7a'}$ and $\nu_{9a}$ since their frequencies show linear relations with the solvent donor number. Monitoring these modes in proteins allows conclusions about the kind and the strength of hydrogen bonding interactions of Tyr. In the extreme case, when Tyr is deprotonated, the entire vibrational band pattern drastically changes and, due to the red-shift of the electronic transition, the resonance enhancement is altered. Also Trp can undergo hydrogen bonding interactions via the indole N–H group. In that case, a band at ~880 $cm^{-1}$ ($\nu_{10a}$, W17) is an appropriate marker in the RR spectra of proteins. The second important spectral marker of Trp, a Fermi doublet at ~1345 $cm^{-1}$, exhibits a moderate resonance enhancement and can only be identified in the spectra of relatively small proteins. This band is an indicator for the hydrophobicity of the local environment of the indole ring. Based on systematic studies of Trp and model compounds, an empirical

relationship has been derived between the torsional angle of the $C_2$–$C_3$–$C_\beta$–$C_\alpha$ entity and the frequency of the band at ~1550 cm$^{-1}$ (W3). This marker band, however, does not appear to be applicable for analysing the Trp structure in proteins inasmuch as it appears in a spectrally crowded region.

Although *a priori* the size of the protein is not a limitation for UV RR spectroscopic analysis of aromatic amino acids, increasing numbers of Phe, Tyr and Trp residues strongly complicate the interpretation of the spectra. Nevertheless, working with high quality spectra it is possible to probe subtle structural changes of a single amino acid even in relatively large proteins as demonstrated in the studies on haemoglobin by Spiro and co-workers. This protein consists of two (αβ) dimers, each of them including 15 Phe, 6 Tyr and 3 Trp residues. It had been assumed that the T → R transition of haemoglobin was associated with a change of hydrogen bonding interactions of one Tyr (out of 12) which in fact could be confirmed by UV RR (difference) spectra.

The application of Raman spectroscopy for probing amino acid side-chain structures in proteins is not restricted to excitation in the deep UV. Some ring vibrations of the aromatic amino acid side chains can also be identified in the non-resonance Raman spectra. These are, for instance, the Fermi doublet of Trp at ~1345 cm$^{-1}$ and a band pair of Tyr at 850 and 830 cm$^{-1}$, assigned to the fundamental mode $\nu_1$ and the overtone $2\nu_{16a}$, which together form a Fermi doublet as well. This doublet, which is only weakly resonance enhanced upon UV excitation, appears with considerable intensity under non-resonance conditions and is known to be a sensitive marker for hydrogen bonding interactions as derived from systematic studies on model compounds. In fact, the intensity distribution of this doublet can be used to monitor changes of hydrogen bonding interactions of Tyr in proteins.

Among the aliphatic amino acid side chains, cysteine (Cys) and methionine (Met) give rise to particular strong Raman bands which appear in less cogent spectral regions. The C–S stretching is found between 620 and 750 cm$^{-1}$ depending on the conformation of the side chain, e.g. *trans* or *gauche* conformation of Met. In addition, the C–S stretching vibrations of disulfide bridges, which constitute an important stabilizing factor for the tertiary structure of proteins, are observed at 630–650, 700 and 720 cm$^{-1}$ when the *trans* substituent X of the X–$C_\alpha$–$C_\beta$–S–S entity is a hydrogen, nitrogen and carbon atom, respectively. However, the most characteristic feature of disulfide bridges is the S–S stretching (500–525 cm$^{-1}$). There is a large body of experimental data for proteins and model compounds and evidence has been provided for a linear correlation between the frequency and the C–S–S–C dihedral angle.

### Chromoproteins

Chromoproteins are characterized by an electronic absorption band in the near-UV, visible or near-IR spectral range. These bands may arise from $\pi \rightarrow \pi^*$ transitions of prosthetic groups or from charge-transfer transitions of specifically bound transition metal ions. Thus, chromoproteins which may serve as electron transferring proteins, enzymes or photoreceptors, are particularly attractive systems to be studied by RR spectroscopy since an appropriate choice of the excitation wavelength readily leads to a selective enhancement of the Raman bands of the chromophoric site. Moreover, these chromophores generally constitute the active sites of these biomolecules so that RR spectroscopic studies are of utmost importance for elucidating structure–function relationships.

Photoreceptors may use light either for energy conversion (photosynthetic pigments, bacterial rhodopsins) or as a source of information to trigger a physiological response (rhodopsins, phytochromes). Typical chromophores are retinal Schiff bases which are covalently attached to the apoprotein [(bacterial) rhodopsins] and cyclic or linear tetrapyrroles (photosynthetic pigments, phytochromes). In the case of (bacterial) rhodopsins and phytochromes, light absorption initiates a photochemical reaction which is followed by a series of thermal relaxation processes. Thus, RR spectroscopic studies of these biomolecules encounter the difficulty that the exciting laser beam induces the reaction sequence. In order to avoid accumulation of various intermediate states, low temperature spectroscopy is frequently employed so that thermally activated reaction steps can be blocked. Alternatively, time-resolved RR spectroscopy may be an appropriate approach (e.g. bacteriorhodopsin) which, in addition, provides information about the chromophore–protein dynamics.

Primary interest of RR spectroscopic investigations of the visual pigment rhodopsin was directed to the analysis of the photochemical reaction. Single-beam experiments carried out at 77 K yielded a photostationary equilibrium of the parent state and the primary photoproduct. Using an additional laser beam of a wavelength in resonance with the electronic transition of the product, it was possible to shift the equilibrium largely to the parent state. Thus, subtracting the dual-beam- from the single-beam-excited spectrum yielded a pure spectrum of the parent state. The interpretation of these spectra was greatly facilitated by the large body of

experimental data of retinal model compounds including a variety of isotopomers as well as of the detailed investigations of bacteriorhodopsin (see below). Moreover, normal mode analyses based on empirical force field calculations support a reliable vibrational assignment allowing for the determination of the configurational and conformational state of the retinal chromophore. In this way, it could be shown that the primary photoprocess of the parent state in which the retinal is in the 11-*cis* configuration leads to a distorted all-*trans* configuration. Unusually strong intensities of the C–H out-of-plane vibrations of the retinal chain were taken as a measure of the torsion of the C–C single bonds of the chain. These intensities decrease in the intermediates formed during the subsequent thermal reactions indicating a relaxation of the chromophore geometry and the immediate protein environment. In the final step, the Schiff base, i.e. the covalent linkage to the apoprotein, was deprotonated as indicated by the shift of the C–N stretching vibration and its insensitivity towards H/D exchange. Furthermore, analysis of the RR spectra in terms of retinal–protein interactions contributed essentially to the understanding of colour regulation mechanism in vision.

In addition to its photolability, the plant photoreceptor phytochrome, whose chromophore is a linear tetrapyrrole, exhibits a relatively strong fluorescence which severely hinders RR spectroscopic studies under rigorous resonance conditions. Thus, Mathies and co-workers employed shifted-excitation Raman difference spectroscopy in which two spectra, measured with excitation lines separated by $10\ cm^{-1}$, were subtracted from each other. In this way, the fluorescence background was removed yielding only difference signals which were used to calculate the absolute RR spectrum. Similar spectra were obtained by an alternative approach using Fourier-transform Raman spectroscopy at low temperature (77 K). In that case, the excitation line in the near-infrared is shifted from the electronic transition of the chromophore to circumvent fluorescence (and unwanted photochemical processes) but yields a sufficient preresonance enhancement so that the spectra are dominated by the RR bands of the tetrapyrrole. In addition to the spectra of the parent states, those of the intermediates of the photoinduced reaction cycle were obtained by controlled illumination and increase of the temperature. As shown in **Figure 1**, the spectra of the various states of phytochrome (of the $P_r \rightarrow P_{fr}$ transformation) reveal substantial differences reflecting the changes of the chromophore configuration, conformation and interactions with the surrounding protein. In contrast to retinal proteins, however, analysis of the vibrational spectra of linear tetrapyrroles is much less advanced so that up to now only a little structural information can be extracted from the spectra. Progress in spectral interpretation can be expected by extending the studies to recombinant phytochromes including chemically modified and isotopically labelled chromophores as well as by normal analyses based on quantum chemical force field calculations.

For RR spectroscopic studies of photosynthetic pigments, the intrinsic chromophore fluorescence is again an annoyance. In addition to the techniques employed in phytochrome studies, chlorophyll-containing biomolecules allow excitation in resonance with the second (non-fluorescing) electronically excited state (~350–400 nm). In the RR spectra of chlorophylls several bands have been identified which are known to be sensitive markers for the coordination state of the central Mg ion. These bands as well as those originating from the C=O stretchings of the formyl and keto substituents provide information about chlorophyll–chlorophyll and chlorophyll–protein interactions. This information was particularly helpful for elucidating the structural motifs of the chromophore assemblies in light-harvesting pigments. RR spectroscopic studies of genetically engineered protein variants of reaction centres have contributed to the understanding of those parameters controlling redox potentials and electron-transfer processes.

In contrast to photoreceptors, haem proteins are readily accessible for RR spectroscopy inasmuch as they exhibit strong electronic transitions in the visible and near UV. Moreover, they are photostable and do not reveal any interfering fluorescence. There are numerous review articles on the RR spectroscopic results obtained for various representatives of this versatile class of chromoproteins which includes redox enzymes, electron transferring proteins and oxygen transport proteins. Interpretation of the RR spectra is backed by extensive empirical material accumulated from studies on metalloporphyrins in which spectral and structural properties were correlated. Thus, the RR spectra of haem proteins can be interpreted in terms of the kind of the porphyrin, the coordination sphere of the central iron, distortions of the haem geometry, and haem–protein interactions via the porphyrin substituents and the axial ligands. As an example, **Figure 2** shows the RR spectra of cytochrome $c_{552}$ in the oxidized and reduced state, displaying the RR bands in the region of those marker bands whose frequencies are characteristic for the oxidation, spin and coordination state of the haem iron.

In addition to those investigations which aim to characterize the structure of the haem pocket, many

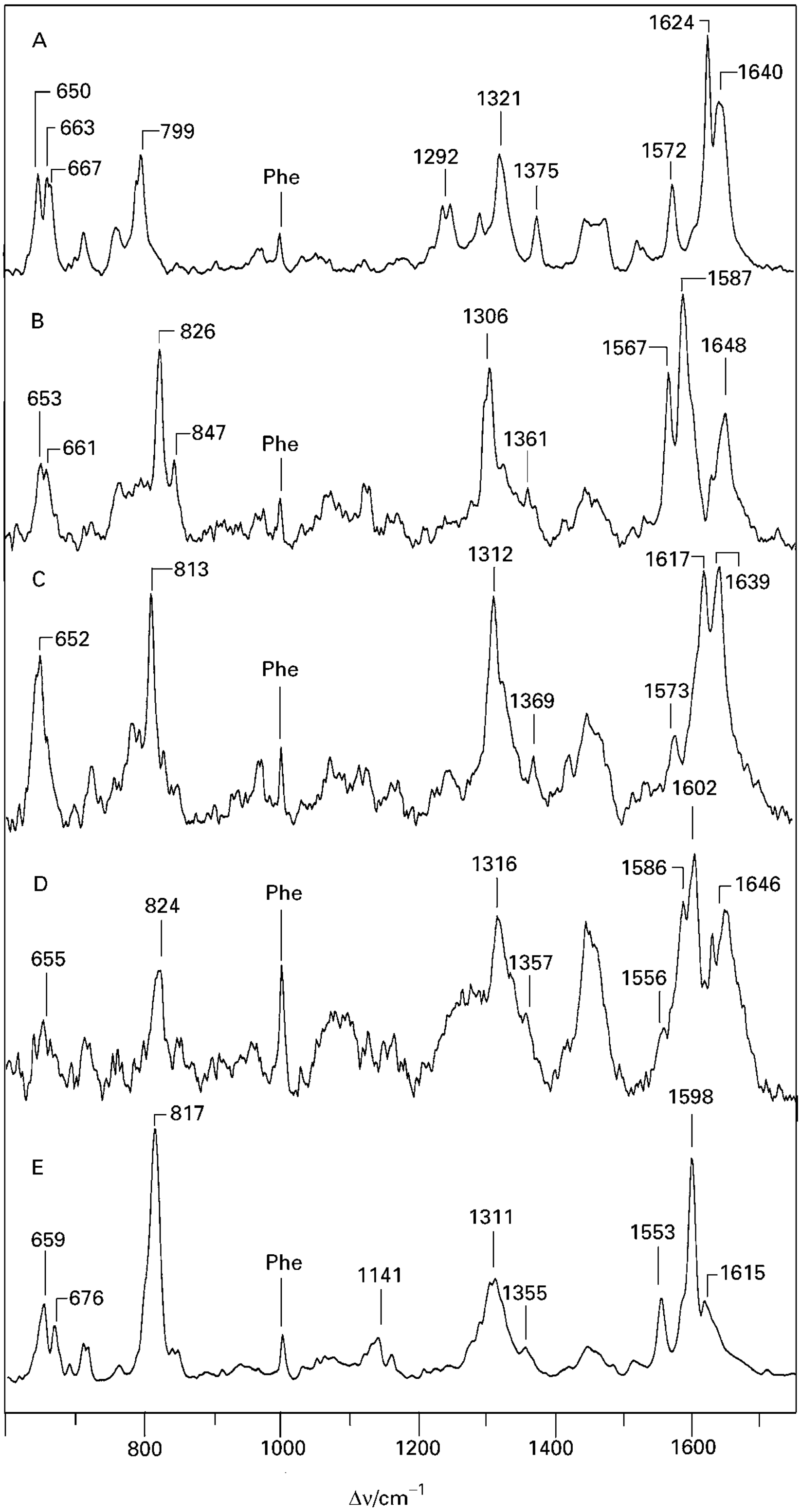



**Figure 1** Fourier-transform RR spectra (1064 nm) of various states of phytochrome trapped at low temperature. (A) $P_r$; (B) Lumi-R; (C) Meta-$R_a$; (D) Meta-$R_c$; (E) $P_{fr}$.

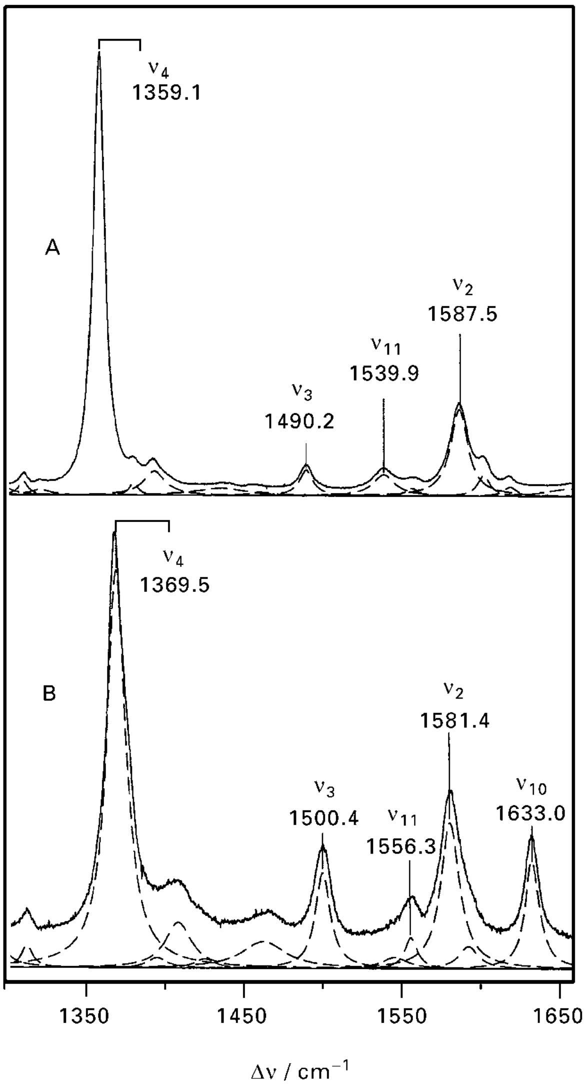



**Figure 2** RR spectra (413 nm) of cytochrome $c_{552}$ of *Thermus thermophilus* in the reduced state (A) and the oxidized state (B).

studies were directed to solve specific problems related to the function of the haem protein under consideration. Detailed studies were carried out on the cytochrome *c*–cytochrome *c* oxidase redox couple to understand the interprotein electron transfer mechanism. Specific emphasis was laid on the analysis of the protein complexes formed prior to the electron transfer in order to assess the effect of protein–protein interaction on the structure of the individual haem sites. In these experiments, dual channel Raman spectroscopy, which permits the quasi-simultaneous measurement of two spectra, is particularly useful for the detection of subtle spectral changes. In this way, RR spectroscopy provided evidence for a considerable conformational flexibility of both proteins which may be of functional relevance. Many studies of haem enzymes (e.g. cytochrome P-450) were directed to elucidate substrate-binding and substrate-specificity on a molecular level.

An interesting approach for probing molecular details of the haem pocket is the use of carbon monoxide which may occupy a vacant coordination site of the haem. The vibrations of the Fe–CO entity can be analysed in terms of the orientation of the CO ligand with respect to the haem plane, which, in turn, depends on the molecular interactions in the haem pocket. Such studies are particularly relevant for oxygen binding proteins inasmuch as CO is an appropriate analogue for molecular oxygen.

Whereas those chromoproteins discussed so far include prosthetic groups with delocalized $\pi$-electron systems, the chromophoric site in metalloproteins is constituted by a transition metal ion coordinated by various amino acid side chains. Excitation in resonance with charge-transfer transitions leads to a predominant enhancement of the metal–ligand vibrations and, to some extent, of internal ligand vibrations. Thus, the corresponding RR spectra are less complex although the interpretation is not necessarily unambiguous and straightforward in each case. Among these proteins, iron–sulfur proteins, blue-copper proteins and non-haem iron proteins have been studied extensively by RR spectroscopy.

Iron–sulfur proteins reveal a characteristic vibrational band pattern in the low frequency region which, in many cases, can be used to identify the type of the iron–sulfur cluster based on a comparison with the spectra of model compounds and related proteins of known structure. However, such studies are restricted to the oxidized form since the reduced complexes do not exhibit a charge-transfer transition in the visible region. This is also true for the blue-copper proteins for which the RR spectrum is dominated by modes including the Cu–S(Cys) stretching coordinate as well as internal modes of the Cys ligand. A substantial improvement of the understanding of these spectra was achieved by using protein variants with isotopically labelled amino acids.

The key issue for RR spectroscopists dealing with non-haem iron proteins is to determine the geometry of the oxygen-bridged binuclear complexes as well as to identify the nature of the ligands. In these studies, the RR spectra of $^{18}O$-labelled samples significantly supports the interpretation.

## Nucleic acids

Raman bands of nucleic acids originate from in-plane vibrations of the nucleic acid bases (adenine, guanine, cytosine, thymine and uracil) and from the furanose-phosphate backbone. In general, Raman spectra of DNA or RNA reveal structural

information about base stacking and interbase hydrogen bonding interactions. Whereas base stacking is reflected solely by intensity variations of purine and pyrimidine bands, changes of interbase hydrogen interactions, for instance, due to helix formation or conformational transitions of the helix, are indicated by both intensity changes and frequency shifts. Raman spectroscopy offers the possibility to determine the conformation of the nucleic acid inasmuch as there are marker bands indicative for the A-, B- and Z-forms of DNA. The most characteristic marker bands originate from the backbone since the various helical structures differ with respect to the sugar–phosphate conformation. Based on the comparison of various DNAs of known crystal structure, it was found that the phosphodiester symmetric stretching vibration is at $811 \pm 3\ cm^{-1}$ in A-DNA whereas it is upshifted to $835 \pm 5\ cm^{-1}$ in the B-form. Furthermore, some ring vibrations of nucleic acid bases, particularly of guanine and cytosine, also serve to distinguish between the various DNA structures (**Table 1**). For quantitative analysis of the DNA structure, the relative intensities of these specific marker bands can be used to assess the percentage of A-, B- and Z-conformation using the conformation-insensitive band at $1100\ cm^{-1}$ (symmetric stretching of the $PO_2$ group) as a reference.

In general, structural changes of DNA in biological processes such as transcription, replication or DNA packaging, are not global but are restricted only to a few nucleotides along the chain. Such changes can induce local melting of the secondary structure, reorientations and/or disruptions of the base stacking. Melting of RNA and DNA double helices always leads to the disappearance of the $814\ cm^{-1}$ and $835\ cm^{-1}$ bands, indicating a decrease in furanose conformational order and an increase of the backbone chain flexibility. Precise thermal melting profiles of both RNA and DNA double helical complexes have been determined based on the intensity variations of these bands. A careful analysis of the Raman spectra of nucleic acids allows the detection of subtle alterations of the helical organization. This has been demonstrated in studies of DNA packaging, which involves formation of DNA–protein complexes with a series of histone and non-histone proteins. A backbone conformation of the B-type family was consistently observed for the nucleosome DNA complexes irrespective of the histone and non-histone content. Furthermore, complex formation is reflected by the intensity decrease of the adenine and guanine ring modes at 1490 and $1580\ cm^{-1}$, respectively. Whereas the intensity attenuation of the $1580\ cm^{-1}$ band was attributed to the binding of histone proteins in the small grooves of double helical B-form DNA, the spectral changes of the $1490\ cm^{-1}$ band is related to the binding of non-histone proteins in the large grooves.

**Table 1** Conformation-sensitive non-resonant Raman bands of nucleic acid

| *Frequency ($cm^{-1}$)* | *Description* |
|---|---|
| 805–815 | phosphodiester symmetric stretching: C-3′-endo conformation (A-DNA) |
| 835 (weak) | phosphodiester symmetric stretching: C-2′-endo conformation (B-DNA) |
| 870–880 | phosphodiester symmetric stretching: C-DNA |
| 682 | guanine ring breathing: C-2′-endo-anti (B-DNA) |
| 665 | guanine ring breathing: C-3′ endo-anti (A-DNA) |
| 625 | guanine ring breathing: C-3′ endo-syn (Z-DNA) |
| 1260 (weak) | cytosine band: B-DNA |
| 1265 (strong) | cytosine band: Z-DNA |
| 1318 (moderate) | guanine band: B-DNA |
| 1318 (very strong) | cytosine band: Z-DNA |
| 1334 (moderate) | guanine band: B-DNA |
| 1355 (moderate) | guanine band: Z-DNA |
| 1362 (moderate) | guanine band: B-DNA |
| 1418 (weak) | guanine band: Z-DNA |
| 1420 (moderate) | guanine band: B-DNA |
| 1426 (weak) | guanine band: Z-DNA |

Raman spectroscopy has also been applied to the analysis of the B–Z transformation in synthetic polymers induced by drug binding. For example, it was shown that binding of *trans*-dichlorodiamine platinium to poly(dG-dC)·poly(dG-dC) appeared to inhibit the right-to-left handed transition whereas the antitumour drug *cis*-dichlorodiamine platinium was found to reduce the salt requirement for adopting a left-handed modified Z-like conformation and rendered the transition essentially non-cooperative.

A Raman spectroscopic method to obtain information about dynamic aspects of nucleic acid secondary structure monitors the deuterium exchange of the C-8 purine hydrogen. The exchange

rates can be determined from the time-dependent intensity increase of the bands characteristic for the deuterated bases. As the exchange process depends on the chemical environment and the solvent accessibility for the individual bases, these data allow differentiation of helical structures.

Pyrimidine and purine bases of nucleic acids have electronic absorptions in the range between 200 and 280 nm. Using excitation lines in this spectral region, RR spectra of nucleic acids exclusively display bands of in-plane ring modes of the bases without any interference of bands from the backbone. However, a selective enhancement of modes originating from purine or pyrimidine bases is not possible so that the practical use of UV RR spectroscopy for structural investigations of nucleic acids is limited. On the other hand, RR spectroscopy can be employed to probe drug–DNA interactions if the electronic transition of the drug is shifted towards the near-UV and visible region. In this way, intercalation of adriamycin by DNA was studied using different excitation lines to probe either the DNA Raman spectrum (364 nm) or the RR spectrum of the bound drug (457 nm).

### Membranes

Natural biological membranes represent multicomponent systems as they include a large variety of different lipids, proteins and carbohydrates. This heterogeneity is a serious obstacle to the interpretation of Raman spectra. Thus, the majority of spectroscopic studies have focused on liposomes which are taken as models for biological membranes.

Studying the fluidity and phase transition of the lipid hydrocarbon chain, a sensitive marker band originates from the longitudinal acoustic mode (LAM) at 100–200 $cm^{-1}$, which corresponds to an accordion-like stretching of the entire molecule along its long axis. The frequency of this mode depends on the all-*trans* chain length and is sensitive to *gauche* rotations of the chain. The LAM is quite strong in pure hydrocarbons. However, it is difficult to observe due to its proximity to the laser line so that in the Raman spectra of phospholipid dispersions, it could be detected only in a few cases, i.e. at low water content and low temperature.

The skeletal optical modes between 1050 and 1150 $cm^{-1}$ and the ratio of the C–H stretching bands at 2850 and 2885 $cm^{-1}$, were used to quantify membrane in order to detect membrane order perturbations induced by the incorporation of small hydrophobic molecules in the bilayer. The intensity at 1130 $cm^{-1}$ is thought to be a measure for the number of all-*trans* sections involving more than three carbon atoms. For phosphatidylcholine vesicles, it was found that the intensity of this band relative to either the 1660 $cm^{-1}$ band or the C–N choline stretching at 718 $cm^{-1}$ could be used to monitor both phase transitions and gradual melting of the lipids. Changes of a stiffening of hydrocarbon chains, e.g. by binding of divalent cations to the phosphate head groups of phosphatidylcholines, is inferred from alterations of the intensity ratio of the 1064 and 1089 $cm^{-1}$ bands.

Interactions of membranes with larger molecules are difficult to probe by Raman spectroscopy unless selectivity can be increased by taking advantage of the RR effect. This has been demonstrated in studies on drug–membrane interactions by tuning the excitation line in resonance with the electronic transition of the drug molecule, e.g. amphotericin.

## Dynamics of biomolecules

Operating in the time-resolved domain, RR spectroscopy can be employed to probe the dynamics of biological systems. Hence, such studies can provide simultaneously kinetic and structural data. In many cases, time-resolved RR spectroscopy, albeit technically demanding, may represent the method of choice for elucidating the molecular mechanisms of biological reactions.

### Photoinduced processes

Bacterial retinal proteins such as bacteriorhodopsin and halorhodopsin act as light-driven ion pumps. The ion gradient that is generated across the membrane is then converted into chemical energy (i.e. synthesis of adenosine triphosphate). The ion translocation is linked to a photoinduced reaction cycle of the retinal chromophore. The photophysical and kinetic properties as well as the stability of this class of proteins make them an ideal system for time-resolved RR spectroscopic studies. Moreover, these proteins represent suitable objects for developing and optimizing time-resolved spectroscopic techniques. The main advantage is the reversibility of the photoinduced reaction cycle, i.e. the system comes back to the parent state in a few ms after the primary photochemical event. Thus, time-resolved RR spectra can be accumulated continuously since the 'fresh sample' condition is readily established. This condition ensures that the protein is always in the same state when irradiated by the exciting laser beam.

There are two approaches for time-resolved RR spectroscopy of these retinal proteins. Using cw-excitation a time-resolution down to 100 ns can be achieved by rapidly moving the sample through the laser focus (rotating cell, capillar flow system). Pulsed laser excitation can provide a time-resolution even in the subpicosecond range depending on the

pulse width of the laser. In both methods, intermediate states are probed in dual-beam experiments with the probe beam irradiating the sample after a delay time $\delta$ with respect to the photolysis beam which initiates the reaction cycle. The systematic variation of $\delta$ allows the determination of the photocycle kinetics which, along with the structural data derived from the RR spectra, can provide a comprehensive picture of the dynamics of the protein. **Figure 3** shows a selection of time-resolved RR of halorhodopsin. The bands displayed in these spectra are diagnostic for the retinal configuration (C=C stretching) and confirm that the photoreaction includes the isomerization from the all-*trans* to the 13-*cis* configuration.

### Redox processes

Cytochrome *c* oxidase, a membrane-bound enzyme in the respiratory chain of aerobic organisms, reduces oxygen to water. This process which takes place at the binuclear metal centre constituted by a haem $a_3$ and a Cu ion runs via several intermediate states with life times in the micro- and millisecond range. Technical improvements now make it possible to monitor this reaction sequence by time-resolved RR spectroscopy using rapid flow systems. The starting point of the experiments is the thermally stable carbon monoxide complex of the reduced enzyme in oxygen-saturated solution. This complex is photodecomposed by a laser beam so that oxygen can bind to the catalytic centre constituting the time 'zero' of the reaction sequence. The various intermediates are probed by a second laser beam irradiating the (photolysed) sample volume after a delay time. The oxygen-sensitive vibrational modes which can readily be identified based on the $^{18}O/^{16}O$ isotopic shifts give insight into the nature of the various intermediates and, hence, into the molecular mechanism of the oxygen reduction. A particularly interesting technical approach for these studies is based on an artificial cardiovascular device designed to maintain a continuous enzymatic reaction. Thus, it is possible to accumulate the RR signals during a sufficiently long period of time using a minimum amount of sample.

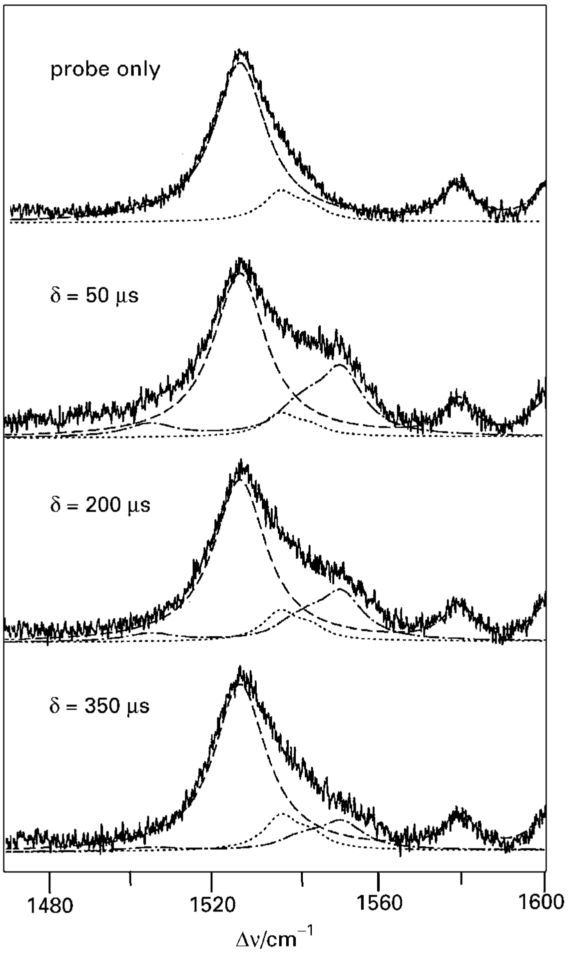



**Figure 3** Time-resolved RR spectra of halorhodopsin from *Natronobacterium pharaonis* obtained in single (probe only, 514 nm) and dual beam experiments with variable delay times δ of the probe laser (514 nm) relative to the pump laser (600 nm).

## Living cells and biomedical applications

Recent progress in laser technology and signal detection has substantially increased the sensitivity of Raman spectroscopy. This is a prerequisite for the development of Raman microspectroscopy and Raman imaging which has opened a new field for the application of Raman spectroscopy in biological and medical research. A particularly powerful method has been introduced by Greve and co-workers who combined Raman spectroscopy with confocal microscopy which allows the measurement of spatially resolved Raman spectra with a lateral resolution of less than 0.5 μm. The technique makes use of sensitive CCD detection systems by probing either the intensity of a marker band in two (spatial) dimensions (global imaging) or a complete spectrum in one dimension (linescan imaging). In these experiments, low energy excitation (>600 nm) has to be employed and the power of the tightly focused laser beam must be reduced as far as possible to avoid photoinduced degradation of the biological objects. In this way, it is possible to study single living cells and to obtain Raman spectra of the cytoplasm and the nucleus separately. The spectral information allows determination of the DNA conformation, the relative content of the individual nucleic acid bases and the

DNA/protein ratio. Furthermore, less abundant biomolecules such as carotenoids or haem enzymes or non-fluorescent drugs (e.g. cobalt octacarboxyphthalocyanine) can be detected by taking advantage of the RR effect. These studies are particularly promising for elucidating metabolic processes in cells.

This method has also been applied to medical problems related to disease diagnosis (e.g. artherosclerotic plaque) or optimization of medical treatment (e.g. bone implants). However, medical applications of Raman spectroscopy do not necessarily require the combination with microscopy. Numerous studies have indicated that based on the analysis of Raman spectra by statistical methods it is possible to differentiate between normal and pathological tissues. Despite the substantial technical improvements, the intrinsically low sensitivity of Raman spectroscopy constitutes a limit for general applicability. Fluorescent samples may impose an additional constraint, which, however, can be overcome by near-infrared (1064 nm) Fourier-transform Raman spectroscopy and microscopy. In summary, it appears that for special medical applications Raman spectroscopic techniques may become a powerful diagnostic tool in clinical situations.

*See also:* **Biofluids Studied By NMR; Chiroptical Spectroscopy, General Theory; Forensic Science, Applications of IR Spectroscopy; Industrial Applications of IR and Raman Spectroscopy; IR Spectrometers; Medical Science Applications of IR; Membranes Studied By NMR Spectroscopy; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Nucleic Acids Studied By NMR; Peptides and Proteins Studied Using Mass Spectrometry; Raman Spectrometers; Surface-Enhanced Raman Scattering (SERS), Applications.**

## Further reading

Althaus T, Eisfeld W, Lohrmann R and Stockburger M (1995) Application of Raman spectroscopy to retinal proteins. *Israel Journal of Chemistry* **35**: 227–251.

Andel F, Lagarias JC and Mathies RA (1996) Resonance Raman analysis of chromophore structure in the lumi-R photoproduct of phytochrome. *Biochemistry* **35**: 15997–16008.

Andrew CR, Han J, den Blaauwen T *et al* (1997) Cysteine ligand vibrations are responsible for the complex resonance Raman spectrum of azurin. *Journal of Bioinorganic Chemistry* **2**: 98–107.

Callender R, Deng H and Gilmanshin R (1998) Raman difference studies of protein structure and folding, enzymatic catalysis and ligand binding. *Journal of Raman Spectroscopy* **29**: 15–21.

Carey PR (1998) Raman spectroscopy in enzymology: the first 25 years. *Journal of Raman Spectroscopy* **29**:7–14.

Clark RJH and Hester RE (eds) (1986) Spectroscopy of biological systems. *Advances in Spectroscopy* **13**.

Clark RJH and Hester RE (eds) (1993) Biomolecular spectroscopy, part A. *Advances in Spectroscopy* **20**.

Clark RJH and Hester RE (eds) (1993) Biomolecular spectroscopy, part B. *Advances in Spectroscopy* **21**.

Hildebrandt P (1992) Resonance Raman spectroscopy of cytochrome P-450. In: Ruckpaul K and Rein H (eds) *Frontiers in Biotransformation*, Vol. 7, pp 166–215. Berlin/Weinheim: Akademie-Verlag/VCH.

Hildebrandt P (1995) Resonance Raman spectroscopy of cytochrome *c*. In: Scott RA and Mauk AG (eds) *Cytochrome c. A Multidisciplinary Approach*, pp 285–314. Mill Valley: University Science.

Kneip C, Mozley D, Hildebrandt P, Gärtner W, Braslavsky SE and Schaffner K (1997) Effect of chromophore exchange on the resonance Raman spectra of recombinant phytochromes. *FEBS Letters* **414**: 23–26.

Kochendoerfer GG, Wang Z, Oprian DD and Mathies RA (1997) Resonance Raman examination of wavelength regulation mechanism in human visual pigments. *Biochemistry* **36**: 6577–6587.

Manoharan R, Wang Y and Feld MS (1996) Histochemical analysis of biological tissues using Raman spectroscopy. *Spectrochimica Acta A* **52**: 215–249.

Otto C, de Grauw CJ, Duindam JJ, Sijtsema NM and Greve J (1997) Applications of micro-Raman imaging in biomedical research. *Journal of Raman Spectroscopy* **28**: 143–150.

Spiro TG (ed) (1987) *Biological Application of Raman Spectroscopy*, Vol 1: Raman spectra and the conformations of biological macromolecules. New York: John Wiley & Sons.

Spiro TG (ed) (1987) *Biological Application of Raman Spectroscopy*, Vol 2: Resonance Raman spectra of polyenes and aromatics. New York: John Wiley & Sons.

Spiro TG (ed) (1988) *Biological Application of Raman Spectroscopy*, Vol 3: Resonance Raman spectra of haem and metalloproteins. New York: John Wiley & Sons.

Tensmeyer LG and Kauffmann EW (1996) Protein structure as revealed by nonresonance Raman spectroscopy. In: Havel HA (ed) *Spectroscopic Methods for Determining Protein Structure in Solution*, pp 69–95. New York: VCH.

Thamann TJ (1996) Probing local protein structure with ultraviolet resonance Raman spectroscopy. In: Havel HA (ed) *Spectroscopic Methods for Determining Protein Structure in Solution*, pp 96–134. New York: VCH.

Twardowski J and Anzenbacher P (1994) *Raman and IR Spectroscopy in Biology and Biochemistry*. New York: Ellis Horwood.

Verma SP and Wallach DFH (1984) Raman spectroscopy of lipids and biomembranes. In: Chapman D (ed) *Biomembrane Structure and Function*, pp 167–198. Weinheim: Verlag Chemie.

Zhao X and Spiro TG (1998) Ultraviolet resonance Raman spectroscopy of hemoglobin with 200 and 212 nm excitation: H-bonds of tyrosines and prolines. *Journal of Raman Spectroscopy* **29**: 49–55.

# Biofluids Studied By NMR

**John C Lindon** and **Jeremy K Nicholson**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

## Introduction

Investigation of biofluid composition provides insight into the status of a living organism in that the composition of a particular fluid carries biochemical information on many of the modes and severity of organ dysfunction. One of the most successful approaches to biofluid analysis has been the application of NMR spectroscopy.

The complete assignment of the $^1$H NMR spectrum of most biofluids is not possible (even by use of 900 MHz NMR spectroscopy) owing to the enormous complexity of the matrix. However, the assignment problems vary considerably between biofluid types. For instance, seminal fluid and blood plasma are highly regulated with respect to metabolite composition and concentrations and the majority of the NMR signals have been assigned at 600 and 750 MHz for normal human individuals. Urine composition is much more variable because its composition is normally adjusted by the body in order to maintain homoeostasis and hence complete analysis is much more difficult. There is also enormous variation in the concentration range of NMR-detectable metabolites in urine samples. With every new increase in available spectrometer frequency the number of resonances that can be resolved in a biofluid increases and although this has the effect of solving some assignment problems, it also poses new ones. Furthermore, problems of spectral interpretation arise due to compartmentation and binding of small molecules in the organized macromolecular domains that exist in some biofluids such as blood plasma and bile.

All biological fluids have their own characteristic physicochemical properties and a summary of some of these is given in **Table 1** for normal biofluids. These partly dictate the types of NMR experiment that may be employed to extract the biochemical information from each fluid type. An illustration of the complexity of biofluid spectra, and hence the need for ultrahigh field measurements, is given in **Figure 1** which shows 800 MHz $^1$H NMR spectra of normal human urine, bile and blood plasma.

It is clear that at even the present level of technology in NMR, it is not yet possible to detect many important biochemical substances, e.g. hormones, in body fluids because of problems with sensitivity, dispersion and dynamic range and this area of research will continue to be technology limited. With this in mind, it would seem prudent to interpret quantitative $^1$H NMR measurements of intact biological materials and assignment of resonances in 1D spectra with considerable caution even when measured at ultrahigh field.

## Resonance assignment in NMR spectra of biofluids

Usually in order to assign $^1$H NMR spectra of biofluids, comparison is made with spectra of authentic materials and by standard addition to the biofluid sample. Additional confirmation of assignments is usually sought from the application of 2-dimensional (2D) NMR methods, particularly COSY and TOCSY and, increasingly, inverse-detected heteronuclear correlation methods such as HMQC and HSQC. In addition, the application of the 2D J-resolved (JRES) pulse sequence is useful for spreading out the coupling patterns from the multitude of small molecules in a biofluid. Even 2D correlation NMR spectra of complex biofluids can show much overlap of cross-peaks and further editing is often desirable. Thus simplification of 1D and 2D NMR spectra of biofluids can also be achieved using (i) spin–echo methods particularly for fluids containing high levels of macromolecules ($T_2$ editing), (ii) editing based on $T_1$ relaxation time differences, (iii) editing based on diffusion coefficient differences, (iv) multiple quantum filtering.

One major advantage of using NMR spectroscopy to study complex biomixtures is that measurements can often be made with minimal sample preparation (usually with only the addition of 5–10% $D_2O$) and a detailed analytical profile can be obtained on the whole biological sample. This in turn requires good methods for suppressing solvent resonances.

Detailed $^1$H NMR spectroscopic data for a wide range of low molecular weight metabolites found in biofluids are given in **Table 2**.

**Table 1** Normal biofluids and their physicochemical properties

| *Biofluid* | *Function* | *Water content*[a] | *Viscosity* | *Protein content*[b] | *Lipid content*[b] | *Peak overlap*[c] |
|---|---|---|---|---|---|---|
| Urine | Excretion Homoeostasis | +++ | [e] | [e] | [e] | [e] |
| Bile | Excretion Digestion | ++ | ++ | + | + | +++ |
| Blood plasma | Transport Homoeostasis Mechanical | +++ | ++ | +++ | +++ | ++ |
| Whole blood[d] | Transport Oxygenation | +++ | +++ | +++ | ++ | +++ |
| Cerebrospinal fluid | Transport Homoeostasis Mechanical | +++ | + | + | + | ++ |
| Milk | Nutrition | ++ | + | ++ | +++ | +++ |
| Saliva | Excretion Digestion | +++ | ++ | ++ | ++ | + |
| Gastric juice | Digestion | +++ | +++ | +++ | [e] | ++ |
| Pancreatic juice | Digestion | ++ | ++ | +++ | + | +++ |
| Seminal fluid | Support for spermatozoa | + | +++ | + | + | +++ |
| Prostatic fluid | Support for spermatozoa | + | ++ | + | + | ++ |
| Seminal vesicle fluid | Support for spermatozoa | + | +++ | +++ | + | +++ |
| Amniotic fluid | Protection of fetus | +++ | + | + | + | + |
| Follicular fluid | Reproduction | +++ | + | + | + | + |
| Synovial fluid | Joint protection | +++ | +++ | ++ | + | + |
| Aqueous humour | Eye function | +++ | + | ++ | [e] | + |

+++, ++, + indicate high, medium, low degree of constraint for NMR studies.
[a] Relative water intensity when compared with concentrations of metabolities of interest.
[b] In biofluids with high protein or lipid contents, spin-echo spectra must normally be employed to eliminate broad resonances.
[c] A subjective indication of spectral crowding (at 600 MHz) due to abundance of endogenous metabolites with a wide range of shifts.
[d] Presence of cells and process of cell sedimentation gives rise to magnetic field inhomogeneity problems.
[e] Not a limiting factor.

## NMR studies of dynamic interactions

Although NMR spectroscopy of biofluids is now well established for probing a wide range of biochemical problems, there are still many poorly understood physicochemical phenomena occurring in biofluids, particularly the subtle interactions occurring between small molecules and macromolecules or between organized multiphasic compartments. The understanding of these dynamic processes is of considerable importance if the full diagnostic potential of biofluid NMR spectroscopy is to be realized. Typically, it is now possible to study enzymatic reactions, chemical reactions and biofluid instability, microbiological activity in biofluids, macromolecular binding of small molecules, membrane-based compartmentation, metal complexation and chemical exchange processes.

## NMR spectroscopy of blood plasma and whole blood

### NMR peak assignments

The physicochemical complexity of plasma shows up in its $^1$H NMR spectra by the range of linewidths of the signals. This means that a number of different multiple pulse NMR experiments and/or physicochemical interventions must be applied to extract useful biochemical information. Numerous high resolution $^1$H NMR studies have been performed on the biochemistry of blood and its various cellular components and plasma. The physical properties of whole blood pose serious limitations on direct NMR investigations, but packed erythrocytes yield more useful information on cell biochemistry. Well resolved spectra are given by plasma, and $^1$H NMR

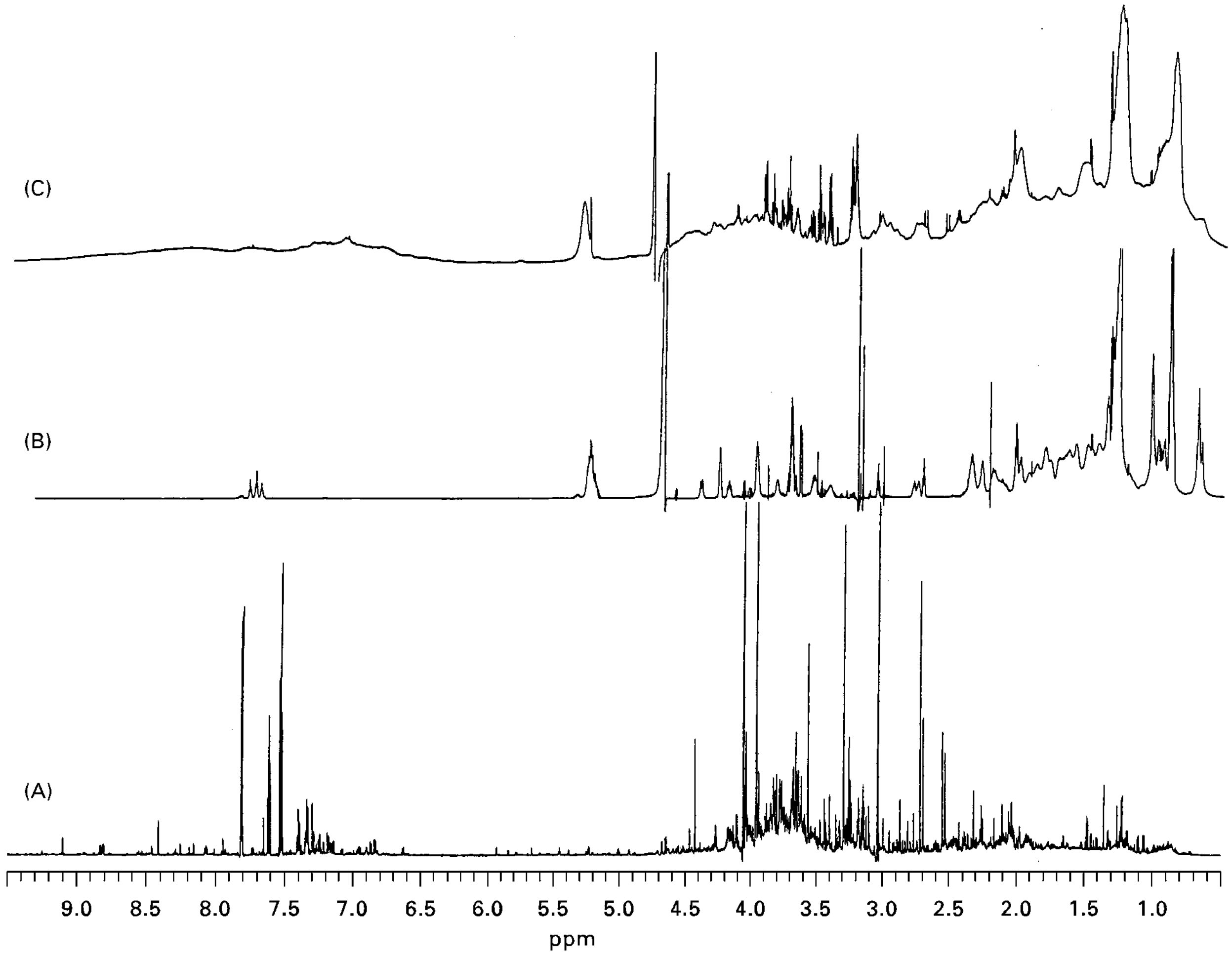



**Figure 1** 800 MHz $^1H$ NMR spectra of control human biofluids; (A) urine; (B) gall bladder bile and (C) blood plasma. Reproduced with permission of Academic Press from Lindon JC, Nicholson JK and Everett JR (1999) NMR spectroscopy of biofluids. *Annual Reports on NMR Spectroscopy* **38**: in the press.

measurements on blood serum and plasma can provide much useful biochemical information on both low molecular weight metabolites and macromolecular structure and organization.

In blood plasma, the $^1H$ NMR peaks of metabolites, proteins, lipids and lipoproteins are heavily overlapped even at 800 MHz (**Figure 1**). Most blood plasma samples are quite viscous and this gives rise to relatively short $T_1$ relaxation times for small molecules and this allows relatively short pulse repetition cycles without signal saturation. The spectral profile can be simplified by use of spin–echo experiments with an appropriate $T_2$ relaxation delay to allow signals from broad macromolecular components and compounds bound to proteins to be attenuated. By the early 1980s many metabolites had been detected in normal blood plasma although assignments were, in general, based on the observation of only one or two resonances for each metabolite. In addition, peaks from certain macromolecules such as $\alpha_1$-acid glycoprotein (*N*-acetyl neuraminic acid and related sialic acid fragments) have been assigned and used diagnostically, in particular their *N*-acetyl groups which give rise to relatively sharp resonances presumably due to less restricted molecular motion. The signals from some lipid and lipoprotein components, e.g. very low density lipoprotein (VLDL), low density lipoprotein (LDL), high density lipoprotein (HDL) and chylomicrons, have also been partially characterized.

For normal plasma at pH 7, the largest peak in the spectral region to high frequency of water is that of the $\alpha$-anomeric H1 resonance of glucose at $\delta 5.223$ (which provides a useful internal chemical shift reference). In spectra of normal human and animal plasma, there are few resonances in the chemical shift range to high frequency of $\delta 5.3$ when measured in the pH range 3 to 8.5. However, on acidification of the plasma to pH $< 2.5$, resonances from histidine and phenylalanine become detectable. Experiments

with model solutions suggested that serum albumin has a high capacity for binding aromatic amino acids and histidine at neutral pH and this is responsible for their NMR-invisibility in normal human blood plasma. Serum albumin also binds a large number of other species of both endogenous and xenobiotic origin. Even commercial 'purified' bovine serum albumin (BSA) can be shown to contain a significant amount of bound citrate and acetate which become NMR-detectable in BSA solutions at pH 2. Acidification of human plasma also renders citrate NMR-detectable in spin–echo spectra as it becomes mobilized from the protein binding sites. Through the increased spectral dispersion available from the use of 600, 750 and 800 MHz $^1$H NMR measurements and through the use of a variety of 2D methods, the assignment of resonances in blood plasma spectra in normal individuals is now extensive (see **Table 2**).

Blood plasma also has intrinsic enzymatic activities although many of these are not stable (particularly if the sample is not frozen immediately on collection). It has been noted that under certain pathological conditions, such as those following liver or kidney damage, enzymes that are present at elevated levels in the plasma because of leakage from the damaged tissue can cause NMR-detectable alterations to spin–echo spectra of plasma.

Molecular diffusion coefficients are parameters that are not related directly to NMR spectral intensities under normal conditions. However, molecular diffusion can cause NMR signal intensity changes when pulsed field gradients are applied during the NMR experiment. A number of pulse sequence developments have meant that measurement of diffusion coefficients is relatively routine. The editing of $^1$H NMR spectra of biofluids based on diffusion alone or on a combination of spin relaxation and diffusion has been demonstrated recently. This approach is complementary to the editing of $^1$H NMR spectra based on differences in $T_1$ and $T_2$. New methods for editing TOCSY NMR spectra of biofluids have been proposed based on differences in molecular diffusion coefficients and this has been termed Diffusion Edited TOCSY (DETOCSY).

Much study has been devoted to the problem of lipoprotein analysis in blood plasma using $^1$H NMR spectroscopy. This has been comprehensively reviewed recently by Ala-Korpela. Lipoproteins are complex particles that transport molecules normally insoluble in water. They are spherical with a core region of triglyceride and cholesterol ester lipids surrounded by phospholipids in which are embedded various proteins known as apolipoproteins. In addition, free cholesterol is found in both the core and surface regions. The lipoproteins are in a dynamic equilibrium with metabolic changes going on *in vivo*. Lipoproteins are usually classified into five main groups, chylomicrons, very low density lipoprotein (VLDL), low density lipoprotein (LDL), intermediate density lipoprotein (IDL) and high density lipoprotein (HDL) based on physical separation using centrifugation. Based on the measurement of $^1$H NMR spectra of the individual fractions and using lineshape fitting programs, it has been possible to identify the chemical shifts of the $CH_2$ and $CH_3$ groups of the fatty acyl side chains. Quantification can be carried out using either time-domain or frequency-domain NMR data. The usefulness of $^1$H NMR spectra for lipoprotein analysis and $^{31}$P NMR spectroscopy for phospholipid analysis in blood plasma has been explored. More recently a neural network software approach has been used to provide rapid lipoprotein analyses.

### NMR spectra of blood plasma in pathological states

A good deal of excitement was generated by a publication which reported that $^1$H NMR spectroscopy of human blood plasma could be used to discriminate between patients with malignant tumours and other groups, namely normals, patients with non-tumour disease and a group of patients with certain benign tumours. This paper stimulated many other research groups around the world to investigate this approach to cancer detection.

The test as originally published involved the measurement of the averaged line width at half height of the two composite signals at $\delta$1.2 and $\delta$0.8 in the single pulse $^1$H NMR spectrum of human blood plasma. On comparing the averaged signal width, $W$, for normal subjects with those for patients with a variety of diseases and with those for pregnant women, it was reported that several statistically significant differences existed between the groups. It was proposed that the lowering of $W$ observed in the malignant group was due to an increase in the $T_2^*$ of the lipoproteins. A mass of literature now exists on this subject and it was immediately apparent that the perfect 100% sensitivity and specificity reported in the early publications has never been repeated by any group. In addition to the presence of cancer, a number of other factors have been found to cause changes in the linewidth index, $W$. These include diet, age and sex, pregnancy and trauma as well as hyperlipidaemia. In all cases the observed changes in $W$ are caused by alterations in the plasma lipoprotein composition, especially the VLDL/HDL ratio.

The success of heart transplantation has improved recently with the chronic use of cyclosporin A to

**Table 2** $^1$H NMR assignments and chemical shifts for metabolities found in biofluids

| *Metabolite* | *Assignment* | *δ($^1$H)* | *Multiplicity* | *δ($^{13}$C)* | *Biofluid*[a] |
|---|---|---|---|---|---|
| $\alpha$-Hydroxyisovalerate | $CH_3$ | 0.81 | d | | C,U |
| $\alpha$-Hydroxy-*n*-butyrate | $CH_3$ | 0.90 | t | | C,U |
| *n*-Butyrate | $CH_3$ | 0.90 | t | | U |
| $\alpha$-Hydroxy-*n*-valerate | $CH_3$ | 0.92 | t | | U |
| Isoleucine | $\delta$-$CH_3$ | 0.94 | t | | A,C,P,S,U |
| Leucine | $\delta$-$CH_3$ | 0.96 | t | | A,C,P,U |
| Leucine | $\delta$-$CH_3$ | 0.97 | t | | A,C,P,U |
| $\alpha$-Hydroxyisovalerate | $CH_3$ | 0.97 | d | | C,U |
| Valine | $CH_3$ | 0.99 | d | 19.6 | C,P,U |
| Isoleucine | $\beta$-$CH_3$ | 1.01 | d | 14.6 | A,C,P,S,U |
| Valine | $CH_3$ | 1.04 | d | | A,C,P,U |
| Ethanol | $CH_3$ | 1.11 | t | | C,P,U |
| Propionate | $CH_3$ | 1.12 | t | | U |
| Isobutyrate | $CH_3$ | 1.13 | d | | P |
| $\beta$-Hydroxybutyrate | $CH_3$ | 1.20 | d | | C,P,S,U |
| Isoleucine | $\gamma$-$CH_2$ | 1.26 | m | | A,C,P,S,U |
| Fucose | $CH_3$ | 1.31 | d | | P |
| Lactate | $CH_3$ | 1.33 | d | 20.9 | A,C,P,S,U |
| Threonine | $CH_3$ | 1.34 | d | | A,C,P,S,U |
| $\alpha$-Hydroxyisobutyrate | $CH_3$ | 1.36 | s | | U |
| $\alpha$-Hydroxy-*n*-valerate | $\gamma$-$CH_2$ | 1.37 | m | | U |
| Alanine | $CH_3$ | 1.48 | d | 16.8 | A,C,P,S,U |
| Lysine | $\gamma$-$CH_2$ | 1.48 | m | | C,P,S,U |
| Isoleucine | $\gamma$-$CH_2$ | 1.48 | m | | C,P,S,U |
| *n*-Butyrate | $\beta$-$CH_2$ | 1.56 | d | | P,U |
| Adipate | $CH_2$ | 1.56 | m | | P,U |
| Citrulline | $\gamma$-$CH_2$ | 1.58 | m | | C,U |
| $\alpha$-Hydroxy-*n*-valerate | $\beta$-$CH_2$ | 1.64 | m | | U |
| $\alpha$-Hydroxybutyrate | $CH_2$ | 1.70 | m | | U |
| Arginine | $\gamma$-$CH_2$ | 1.70 | m | | U |
| Leucine | $CH_2$ | 1.71 | m | 40.7 | C,P,S,U |
| Lysine | $\delta$-$CH_2$ | 1.73 | m | | C,P,S,U |
| Ornithine | $\gamma$-$CH_2$ | 1.81 | m | | C,P,S,U |
| Citrulline | $\beta$-$CH_2$ | 1.88 | m | | C,P,S,U |
| *N*-Acetylglutamate | $\beta$-$CH_2$ | 1.89 | m | | C,P,S,U |
| $\gamma$-Amino-*n*-butyrate | $\beta$-$CH_2$ | 1.91 | m | | C,P,S,U |
| Lysine | $\beta$-$CH_2$ | 1.91 | m | 30.3 | C,P,S,U |
| Arginine | $\beta$-$CH_2$ | 1.93 | m | | C,P,S,U |
| Ornithine | $\beta$-$CH_2$ | 1.95 | m | | A,S,U |
| Acetate | $CH_3$ | 1.95 | s | | A,C,P,S,U |
| Isoleucine | $\beta$-CH | 1.98 | m | | C,P,U |
| Acetamide | $CH_3$ | 2.01 | s | | U |
| *N*-Acetyl groups(glycoproteins) | $CH_3$ | 2.02 | s | 23.0 | P,U |
| $\alpha$-Hydroxyisovalerate | $\beta$-CH | 2.02 | m | | U |

**Table 2** *continued*

| *Metabolite* | *Assignment* | *δ($^1$H)* | *Multiplicity* | *δ($^{13}$C)* | *Biofluid*[a] |
|---|---|---|---|---|---|
| *N*-Acetylaspartate | $CH_3$ | 2.03 | s | | C,U |
| Proline | $\gamma$-$CH_2$ | 2.01 | m | | C,S,U |
| *N*-Acetyl groups(glycoproteins) | $CH_3$ | 2.05 | s | | P,U |
| Proline | $\beta$-$CH_2$ | 2.07 | m | | C,S,U |
| *N*-Acetylglutamate | $CH_3$ | 2.04 | s | | U |
| *N*-Acetylglutamate | $\beta$-$CH_2$ | 2.06 | m | | U |
| Glutamate | $\beta$-$CH_2$ | 2.10[b] | m | 30.1 | A,C,P,S,U |
| Glutamine | $\beta$-$CH_2$ | 2.14[b] | m | | C,P,S,U |
| Methionine | S-$CH_3$ | 2.14 | s | | A,C,P,S,U |
| *n*-Butyrate | $\alpha$-$CH_2$ | 2.16 | t | | C,P,S,U |
| Methionine | $\beta$-$CH_2$ | 2.16 | m | | A,C,P,S,U |
| Adipate | $CH_2COOH$ | 2.22 | m | | U |
| *N*-Acetylglutamate | $\gamma$-$CH_2$ | 2.23 | t | | U |
| Acetone | $CH_3$ | 2.23 | s | | P,U |
| Valine | $\beta$-CH | 2.28 | m | | C,P,S,U |
| Acetoacetate | $CH_3$ | 2.29 | s | | C,P,S,U |
| $\gamma$-Amino-*n*-butyrate | $\alpha$-$CH_2$ | 2.30 | t | | C |
| $\beta$-Hydroxybutyrate | $CH_2$ | 2.31 | ABX | | C,P,S,U |
| Proline | $\beta$-$CH_2$ | 2.35 | m | | A,C,S,U |
| Glutamate[d] | $\gamma$-$CH_2$ | 2.36 | m | 34.5 | C,U |
| Oxalacetate | $CH_2$ | 2.38 | s | | C,P,U |
| Pyruvate | $CH_3$ | 2.38 | s | | C,P,U |
| Malate | $CH_2$ | 2.39 | dd | | U |
| $\beta$-Hydroxybutyrate | $CH_2$ | 2.41 | ABX | | C,P,S,U |
| Succinate | $CH_2$ | 2.43 | s | | C,P,S,U |
| Carnitine | $CH_2$(COOH) | 2.44 | dd | | C,P,S,U |
| $\alpha$-Ketoglutarate | $\gamma$-$CH_2$ | 2.45 | t | | C,P,S,U |
| Glutamine | $\gamma$-$CH_2$ | 2.46[b] | m | 31.9 | C,P,S,U |
| Glutamate[d] | $\gamma$-$CH_2$ | 2.50 | m | | A,C,P,S,U |
| *N*-Acetylaspartate | $CH_2$ | 2.51 | ABX | | C,U |
| Methylamine | $CH_3$ | 2.54 | s | | P |
| Citrate | $\frac{1}{2}CH_2$ | 2.67 | AB | | A,C,P,S,U |
| Methionine | S-$CH_2$ | 2.65 | t | | C,P,S,U |
| Aspartate | $\beta$-$CH_2$ | 2.68[c] | ABX | | C,P,S,U |
| Malate | $CH_2$ | 2.69 | ABX | | U |
| *N*-Acetylaspartate | $CH_2$ | 2.70 | ABX | | C,U |
| Dimethylamine | $CH_3$ | 2.72 | s | | C,P,S,U |
| Sarcosine | $CH_3$ | 2.74 | s | | U |
| Dimethylglycine | $CH_3$ | 2.78 | s | | P,U |
| Citrate | $\frac{1}{2}CH_2$ | 2.80 | AB | | A,C,P,S,U |
| Aspartate | $\beta$-$CH_2$ | 2.82 | ABX | | C,P,S,U |
| Methylguanidine | $CH_3$ | 2.83 | s | | U |
| Asparagine | $\beta$-$CH_2$ | 2.86 | m | | C,P,S,U |
| Trimethylamine | $CH_3$ | 2.88 | s | | U |

**Table 2** *continued*

| *Metabolite* | *Assignment* | *δ(¹H)* | *Multiplicity* | *δ(¹³C)* | *Biofluid*[a] |
|---|---|---|---|---|---|
| Asparagine | $\beta$-$CH_2$ | 2.96 | m | | C,P,S,U |
| $\alpha$-Ketoglutarate | $\beta$-$CH_2$ | 3.01 | t | | C,P,S,U |
| $\gamma$-Amino-*n*-butyrate | $\gamma$-$CH_2$ | 3.02 | t | | C |
| Lysine | $\varepsilon$-$CH_2$ | 3.03 | t | 40.3 | C,P,S,U |
| Cysteine | $CH_2$ | 3.04 | m | | C,U |
| Creatine | $CH_3$ | 3.04 | s | | A,C,P,S,U |
| Phosphocreatine | $CH_3$ | 3.05 | s | | S |
| Creatinine | $CH_3$ | 3.05 | s | | A,C,P,S,U |
| Tyrosine | $CH_2$ | 3.06 | ABX | | C,P,S,U |
| Ornithine | $\delta$-$CH_2$ | 3.06 | t | | C,S,U |
| Cysteine | $CH_2$ | 3.12 | ABX | | C,U |
| Malonate | $CH_2$ | 3.13 | s | | U |
| Phenylalanine | $\beta$-$CH_2$ | 3.13 | m | | C,P,S,U |
| Histidine | $\beta$-$CH_2$ | 3.14 | ABX | | C,P,S,U |
| Citrulline | $\alpha$-$CH_2$ | 3.15 | m | | C,U |
| *cis*-Aconitate | $CH_2$ | 3.17 | s | | U |
| Tyrosine | $CH_2$ | 3.20 | ABX | | C,P,S,U |
| Choline | $N(CH_3)_3$ | 3.21 | s | 55.0 | C,P,S,U |
| Phosphorylethanolamine | $NCH_2$ | 3.23 | t | | S |
| $\beta$-Glucose | C-H2 | 3.24 | dd | | A,C,P,S,U |
| Histidine | $\beta$-$CH_2$ | 3.25 | ABX | | C,P,S,U |
| Arginine | $\delta$-$CH_2$ | 3.25 | t | 41.3 | C,S,U |
| Trimethylamine-*N*-oxide | $N(CH_3)_3$ | 3.27 | s | | C,P,S,U |
| Taurine | $CH_2SO_3$ | 3.25 | t | | A,C,P,S,U |
| Betaine | $N(CH_3)_3$ | 3.27 | s | | C,P,S,U |
| Phenylalanine | $\beta$-$CH_2$ | 3.28 | m | | C,P,S,U |
| *Myo*-inositol | H5 | 3.28 | t | | P |
| Tryptophan | $CH_2$ | 3.31 | ABX | | P,S,U |
| Glycerophosphorylcholine | $N(CH_3)_3$ | 3.35 | s | | S |
| Proline | $\delta$-$CH_2$ | 3.33 | m | | A,P,U |
| $\beta$-Glucose | C-H4 | 3.40 | t | 70.6 | A,C,P,S,U |
| $\alpha$-Glucose | C-H4 | 3.41 | t | 70.6 | A,C,P,S,U |
| Proline | $\delta$-$CH_2$ | 3.42 | m | | A,P,U |
| Carnitine | $NCH_2$ | 3.43 | m | | C,P,S,U |
| Taurine | $NCH_2$ | 3.43 | t | | A,C,P,S,U |
| Acetoacetate | $CH_2$ | 3.45 | s | | C,P,S,U |
| $\beta$-Glucose | C-H5 | 3.47 | ddd | 76.7 | A,C,P,S,U |
| *trans*-Aconitate | $CH_2$ | 3.47 | s | | U |
| $\alpha$-Glucose | C-H3 | 3.49 | t | 76.7 | A,C,P,S,U |
| Tryptophan | $CH_2$ | 3.49 | ABX | | P,U |
| Choline | $NCH_2$ | 3.52 | m | | C,P,S,U |
| Glycerophosphorylcholine | $NCH_2$ | 3.52 | m | | S |
| $\alpha$-Glucose | C-H2 | 3.53 | dd | 72.3 | A,C,P,S,U |
| Glycerol | $CH_2$ | 3.56 | ABX | 63.5 | C,P |
| *Myo*-inositol | H1/H3 | 3.56 | dd | | P |

**Table 2** *continued*

| *Metabolite* | *Assignment* | *δ*(¹H) | *Multiplicity* | *δ*(¹³C) | *Biofluid*[a] |
|---|---|---|---|---|---|
| Glycine[d] | $CH_2$ | 3.57 | s | | C,P,S,U |
| Threonine | $\alpha$-CH | 3.59 | d | | U |
| Fructose ($\beta$-furanose) | H1 | 3.59 | d | | S |
| Fructose ($\beta$-furanose) | H1 | 3.59 | d | | S |
| Fructose ($\beta$-pyranose) | H1 | 3.60 | d | | S |
| Sarcosine | $CH_2$ | 3.61 | s | | C,P,S,U |
| Ethanol | $CH_2$ | 3.61 | d | | C,P,S,U |
| Valine | $\alpha$-CH | 3.62 | d | 64.2 | C,P,S,U |
| *Myo*-inositol | H4/H6 | 3.63 | dd | | P |
| Fructose ($\beta$-pyranose) | H6′ | 3.63 | m | | S |
| Fructose ($\beta$-pyranose) | H3 | 3.64 | m | | S |
| Glycerol | $CH_2$ | 3.65 | ABX | 63.5 | C,P |
| Isoleucine | $\alpha$-CH | 3.68 | d | | A,C,P,S,U |
| Fructose ($\beta$-furanose) | H6′ | 3.70 | m | | S |
| Fructose ($\beta$-furanose) | H1 | 3.70 | m | | S |
| $\alpha$-Glucose | C-H3 | 3.71 | t | 73.6 | C,P,S,U |
| $\beta$-Glucose | C-H6′ | 3.72 | dd | 61.6 | A,C,P,S,U |
| Leucine | $\alpha$-CH | 3.73 | t | 55.1 | A,C,P,S,U |
| $\alpha$-Glucose | C-H6′ | 3.74 | m | 61.4 | A,C,P,S,U |
| Ascorbate | $CH_2$ | 3.74 | d | | U |
| Ascorbate | $CH_2$ | 3.76 | d | | U |
| Citrulline | $\alpha$-CH | 3.76 | t | | C,U |
| Lysine | $\alpha$-CH | 3.76 | t | | C,P,S,U |
| Glutamine | $\alpha$-CH | 3.77 | t | 55.4 | C,P,S,U |
| Glutamate | $\alpha$-CH | 3.77 | t | | A,C,P,S,U |
| Arginine | $\alpha$-CH | 3.77 | t | | C,U |
| Glycerol | CH | 3.79 | ABX | 72.6 | C,P |
| Alanine | CH | 3.79 | q | | A,C,P,S,U |
| Ornithine | $\alpha$-CH | 3.79 | t | | C,P,S,U |
| Guanidoacetate | $CH_2$ | 3.80 | s | | U |
| Mannitol | CH3 | 3.82 | d | | C |
| Fructose ($\beta$-furanose) | H4 | 3.82 | m | | S |
| Fructose ($\beta$-pyranose) | H3 | 3.84 | m | | S |
| $\alpha$-Glucose | C-H6 | 3.84 | m | 61.4 | A,C,P,S,U |
| $\alpha$-Glucose | C-H5 | 3.84 | ddd | 72.3 | A,C,P,S,U |
| Fructose ($\beta$-furanose) | H6 | 3.85 | m | | S |
| $\alpha$-Hydroxyisovalerate | $\alpha$-CH | 3.85 | d | | U |
| Serine | $\alpha$-CH | 3.85 | ABX | | U |
| Methionine | $\alpha$-CH | 3.86 | t | | A,C,P,S,U |
| $\beta$-Glucose | C-H6 | 3.90 | dd | 61.6 | A,C,P,S,U |
| Fructose ($\beta$-pyranose) | H4 | 3.90 | m | | S |
| Betaine | $CH_2$ | 3.90 | s | | C,P,S,U |
| Aspartate | $\alpha$-CH | 3.91 | ABX | | C,U |
| 4-Aminohippurate | $CH_2$ | 3.93 | d | | U |
| Creatine | $CH_2$ | 3.93 | s | | P |

**Table 2** *continued*

| *Metabolite* | *Assignment* | *δ(¹H)* | *Multiplicity* | *δ(¹³C)* | *Biofluid* [a] |
|---|---|---|---|---|---|
| Glycolate | $CH_2$ | 3.94 | s | | U |
| Tyrosine | CH | 3.94 | ABX | | C,P,S,U |
| Phosphocreatine | $CH_2$ | 3.95 | s | | [e] |
| Serine | $\beta$-$CH_2$ | 3.95 | ABX | | U |
| Hippurate | $CH_2$ | 3.97 | d | | C,P,S,U |
| Histidine | $\alpha$-CH | 3.99 | ABX | | C,P,S,U |
| $\alpha$-Hydroxybutyrate | CH | 4.00 | ABX | | U |
| Asparagine | $\alpha$-CH | 4.00 | ABX | | C,S,U |
| Cysteine | CH | 4.00 | ABX | | U |
| Serine | $\beta$-$CH_2$ | 4.00 | ABX | | C,U |
| Phosphorylethanolamine | $OCH_2$ | 4.00 | m | | S |
| Phenylalanine | $\alpha$-CH | 4.00 | m | | C,P,S,U |
| Fructose ($\beta$-pyranose) | H5 | 4.01 | m | | S |
| Ascorbate | CH | 4.03 | m | | U |
| Fructose ($\beta$-pyranose) | H6 | 4.01 | m | | S |
| $\alpha$-Hydroxy-*n*-valerate | $\alpha$-CH | 4.05 | m | | U |
| Creatinine | $CH_2$ | 4.06 | s | | A,C,P,S,U |
| Tryptophan | CH | 4.06 | ABX | | S,U |
| *Myo*-Inositol | H2 | 4.06 | t | | C,S,U |
| Choline | $OCH_2$ | 4.07 | m | | C,P,S,U |
| Lactate | CH | 4.12 | q | 69.2 | A,C,P,S,U |
| Fructose ($\beta$-furanose) | H5 | 4.13 | m | | S |
| Proline | $\alpha$-CH | 4.14 | m | | A,C,P,S,U |
| $\beta$-Hydroxybutyrate | CH | 4.16 | ABX | | C,P,S,U |
| Threonine | $\beta$-CH | 4.26 | ABX | | A,C,P,S,U |
| Malate | CH | 4.31 | dd | | U |
| *N*-methylnicotinamide | $CH_3$ | 4.48 | s | | U |
| Glycerophosphorylcholine | $NCH_2$ | 3.52 | m | | S |
| *N*-Acetylaspartate | CH | 4.40 | ABX | | C,U |
| $\beta$-Galactose | C-H1 | 4.52 | d | | C |
| Ascorbate | CH | 4.52 | d | | U |
| $\beta$-Galactose | C-H1 | 4.53 | d | | P |
| $\beta$-Glucose | C-H1 | 4.64 | d | | A,C,P,S,U |
| Water | $H_2O$ | 4.79 | s | | all fluids |
| Phospho(enol)pyruvate | CH | 5.19 | t | | [e] |
| $\alpha$-Glucose | C-H1 | 5.23 | d | 92.9 | A,C,P,S,U |
| Allantoate | CH | 5.26 | s | | U |
| Phospho(enol)pyruvate | CH | 5.37 | t | | [e] |
| Allantoin | CH | 5.40 | d | | U |
| Urea | $NH_2$ | 5.78 | s | | P,U |
| Uridine | H5 | 5.80 | d | | C,S,U |
| Uridine | H1′ | 5.82 | d | | C,S,U |
| *cis*-Aconitate | CH | 5.92 | s | | U |
| Urocanate | CH(COOH) | 6.40 | d | | U |
| Fumarate | CH | 6.53 | s | | U |

**Table 2** *continued*

| *Metabolite* | *Assignment* | *δ(¹H)* | *Multiplicity* | *δ(¹³C)* | *Biofluid*[a] |
|---|---|---|---|---|---|
| *trans*-Aconitate | CH | 6.62 | s | | U |
| 4-Aminohippurate | C-H3/H5 | 6.87 | d | | U |
| 3,4-Dihydroxymandelate | cyclic H | 6.87 | d | | U |
| 3,4-Dihydroxymandelate | cyclic H | 6.90 | s | | U |
| Tyrosine | C-H3/H5 | 6.91 | d′ | 116.7 | C,P,S,U |
| 3,4-Dihydroxymandelate | cyclic H | 6.94 | d | | U |
| 3-Methylhistidine | C-H4 | 7.01 | s | | P |
| Histidine | C-H4 | 7.08 | s | | C,P,S,U |
| 1-Methylhistidine | C-H4 | 7.05 | s | | P |
| Tyrosine | C-H2/H6 | 7.20 | d | | C,P,S,U |
| Tryptophan | C-H5/H6 | 7.21 | t | | S,U |
| Indoxyl sulphate | C-H5 | 7.20 | m | | U |
| Indoxyl sulphate | C-H6 | 7.28 | m | | C,U |
| Tryptophan | C-H5/H6 | 7.29 | t | | S,U |
| Urocanate | CH(ring) | 7.31 | d | | U |
| Tryptophan | C-H2 | 7.33 | s | | S,U |
| Phenylalanine | C-H2/H6 | 7.33 | m | | C,P,S,U |
| Phenylalanine | C-H4 | 7.38 | m | | C,P,S,U |
| Urocanate | C-H5 | 7.41 | s | | U |
| Phenylalanine | C-H3/H5 | 7.43 | m | | C,P,S,U |
| Nicotinate | cyclic H | 7.53 | dd | | U |
| Hippurate | C-H3/H5 | 7.55 | t | | U |
| Tryptophan | C-H7 | 7.55 | d | | S,U |
| 3-Methylhistidine | C-H2 | 7.61 | s | | P |
| Hippurate | C-H4 | 7.64 | t | | U |
| 4-Aminohippurate | C-H2/H6 | 7.68 | d | | U |
| Tryptophan | C-H4 | 7.74 | d | | S,U |
| 1-Methylhistidine | C-H2 | 7.77 | s | | P |
| Uridine | C-H6 | 7.81 | d | | C,S,U |
| Histidine | C-H2 | 7.83 | s | | C,P,S,U |
| Hippurate | C-H2/H6 | 7.84 | d | | U |
| Urocanate | cyclic C-H3 | 7.89 | s | | U |
| *N*-methylnicotinamide | C-H5 | 8.19 | t | | U |
| Nicotinate | cyclic H | 8.26 | dt | | U |
| Formate | CH | 8.46 | s | | C,P,S,U |
| Nicotinate | cyclic H | 8.62 | dd | | U |
| *N*-methylnicotinamide | C-H4 | 8.90 | d | | U |
| Nicotinate | cyclic H | 8.95 | d | | U |
| *N*-methylnicotinamide | C-H6 | 8.97 | d | | U |
| *N*-methylnicotinamide | C-H2 | 9.28 | s | | U |

[a] Main biofluid in which compounds are found or have been observed; A, C, P, S and U refer to observation in amniotic fluid, CSF, Serum, Seminal fluid and urine, respectively.
[b] Varies according to the presence of divalent metal ions (especially Ca, Mg and Zn).
[c] Highly variable over physiological pH range.
[d] Significant shift differences due to the formation of fast exchanging complexes with divalent metal ions in biofluids.
[e] Not normally observed in biofluids because of low stability.

suppress rejection, and the better treatment of acute rejection episodes when they do occur. However, the early detection of acute cardiac graft rejection still relies on invasive and iterative right ventricular endomyocardial biopsies. This led to a new application of the measurement of the linewidth index, *W*, for the assessment of heart graft rejection after transplantation. Patients in moderate and severe rejection showed *W* values that were significantly different from those of the light rejection patients. However, as occurred in the test for cancer, the wide overlap between the values observed for each group meant that the *W* value alone could not be used to classify the patients into the four rejection grades. It has also been reported that the areas of two glycoprotein signals in the spin–echo $^1$H NMR spectra of blood plasma from heart transplant patients correlated with a standard echocardiography parameter used to monitor rejection. Although a good correlation was found in 5 patients and an acceptable correlation in 3, only a poor correlation was seen in 5 further patients. It was thought that infections and inflammatory states unrelated to rejection interfered with the correlation.

In human beings, diabetes is a relatively common condition which can have serious, complex and far-reaching effects if not treated. It is characterized by polyuria, weight loss in spite of increased appetite, high plasma and urinary levels of glucose, metabolic acidosis, ketosis and coma. The muscles and other tissues become starved of glucose, whilst highly elevated levels of glucose are found in the urine and plasma. Based on NMR spectra, there are marked elevations in the plasma levels of the ketone bodies and glucose, post-insulin withdrawal. In general, the NMR results were in good agreement with conventional assay results. The $CH_3$ and $CH_2$ resonances of the lipoproteins VLDL and chylomicrons also decreased significantly in intensity relative to the $CH_3$ signal of HDL and LDL, indicating the rapid metabolism of the mobile pool of triglycerides in VLDL and chylomicrons.

The $^1$H NMR spectra of the blood plasma from patients with chronic renal failure during dialysis, patients in the early stages of renal failure and normals have also been analysed. For patients on acetate dialysis, the method clearly showed how the acetate was accumulated and metabolized during the course of the dialysis, as well as allowing changes in the relative concentrations of endogenous plasma components to be monitored. A subsequent $^1$H, $^{13}$C and $^{14}$N NMR study of the plasma and urine from chronic renal failure patients showed that the plasma levels of trimethylamine-*N*-oxide (TMAO) correlated with those of urea and creatinine, suggesting that the presence of TMAO is closely related to the degree of renal failure.

The uraemic syndrome is associated with a complex set of biochemical and pathophysiological changes that remain poorly understood. The first application of 750 MHz NMR for studying the biochemical composition of plasma has been reported from patients on haemodialysis (HD) and peritoneal dialysis (PD). Increased plasma levels of low MW metabolites including methylhistidine, glycerol, choline, TMAO, dimethylamine and formic acid were found. The concentrations of these metabolites, and ratios to others, varied with the type of dialysis therapy. For example, the biochemical composition of plasma from patients on PD was remarkably consistent whereas pre- and post-HD significant fluctuations in the levels of TMAO, glucose, lactate, glycerol, formate and lipoproteins (VLDL, LDL and HDL) were observed. Elevated concentrations of glycoprotein fragments were also observed and may relate to the presence of high levels of *N*-acetyl glucosaminidase (NAG), and other glycoprotein cleaving enzymes, in the plasma originating from damaged kidney cells.

### $^1$H NMR spectroscopy of whole blood and red blood cells

Single pulse $^1$H NMR measurements on whole blood give very little biochemical information owing to the presence of a broad envelope of resonances from haemoglobin and plasma proteins. Spin–echo spectroscopy of whole blood can give rise to moderately well resolved signals from plasma metabolites and those present inside erythrocytes, notably glutathione, but whole blood spectra are not easily reproducible because of erythrocyte sedimentation which progressively (within a few minutes) degrades the sample field homogeneity during the course of data collection. Furthermore, there are substantial intracellular/extracellular field gradients, which give a major contribution to the $T_2$ relaxation processes for the nuclei of molecules diffusing through those gradients, and very weak spectra are obtained. Spin–echo NMR measurements on packed erythrocyte samples do give rise to well resolved signals from intracellular metabolites and a variety of transport and cellular biochemical functions can be followed by this method.

One major parameter that can be obtained from NMR is the intracellular pH, and this has been measured for erythrocytes by $^1$H NMR spectroscopy using a suitable $^1$H NMR pH indicator that has an NMR chemical shift which varies with pH in the

range desired. Candidates include the C2-H protons from histidines in haemoglobin, but an exogenous compound can be added as an indicator.

The transport of substances between the inside and outside of red cells can be monitored using NMR if the resonances from the two environments have different chemical shifts or intensities. Also, resonances outside the cell can be selectively broadened by the addition of paramagnetic species that do not cross the red cell membrane. Those used include the ferric complex of desferrioxamine, dysprosium-DTPA and the copper-cyclohexanediaminetetraacetic acid complex. To measure the rate of influx of a compound into red cells, the compound is added with the paramagnetic agent to a red cell suspension and the intensity of the resonance from the intracellular component is monitored as a function of time. This approach has been used to study the transport of glycerol, alanine lactate, choline and glycylglycine. The time scale that can be addressed covers the range of ms to hours.

Disease processes studied include the use of NMR spectroscopy of erythrocytes to investigate the effect of lithium treatment in manic depressive patients showing an increase in choline. In heavy metal poisoning a considerable proportion of the metal is found in the blood and the binding of heavy metals inside erythrocytes has received much attention. Rabenstein has reviewed much of the work in the area of red blood cell NMR spectroscopy. More recently, diffusion coefficient measurement has been used to study the binding between diphosphoglycerate and haemoglobin inside intact red blood cells.

## NMR spectroscopy of human and animal urine

### Sample preparation and NMR assignments

The composition and physical chemistry of urine is highly variable both between species and within species according to lifestyle. A wide range of organic acids and bases, simple sugars and polysaccharides, heterocycles, polyols, low molecular weight proteins and polypeptides are present together with inorganic species. Many of these moieties also interact, forming complexes some of which are amenable to NMR study, that may undergo chemical exchange reactions on a variety of different time scales. The ionic strength of urine varies considerably and may be high enough to adversely affect the tuning and matching of the RF circuits of a spectrometer probe, particularly at high field strengths. The presence of high concentrations of protein in the urine, e.g. due to renal glomerular or tubular damage, can result in the broadening of resonances from low MW compounds that may bind to these urinary proteins.

Urinary pH values may vary from 5 to 8, according to the physiological condition in the individual, but usually lie between 6.5 and 7.5. Urine samples should, therefore, be frozen as soon as possible after collection if NMR measurements cannot be made immediately. When experiments involve collections from laboratory animals housed in metabolic cages, urine samples should be collected into receptacles that are either cooled with dry ice or have a small amount of sodium azide present as a bacteriocide. However, both these procedures may inhibit or destroy urinary enzymes that may frequently be assayed by conventional biochemical methods for assessment of kidney tubular integrity in toxicological experiments. Urinary pH values are unstable because of the progressive and variable precipitation of calcium phosphates which may be present in the urine close to their solubility limits. One solution to this problem is the addition of 100–200 mM phosphate buffer in the $D_2O$ added for the lock signal followed by centrifugation to remove precipitated salts. This has the effect of normalizing the pH to a range of 6.7–7.6 which is stable for many hours during which NMR measurements can be made. Few metabolites (except for histidine and citrate) show major chemical shift variations over this pH range.

The vast majority of urinary metabolites have $^1H$ $T_1$ relaxation times of 1 to 4 s. Thus, the general rule of applying a $5 \times T_1$ relaxation delay between successive 90° transients to obtain >99% relaxation and hence quantitative accuracy cannot be applied routinely. Instead, by using 30–45° pulse angles and leaving a total $T_1$ relaxation delay of 5 s between successive pulses, spectra with good signal-to-noise ratios can usually be obtained for most metabolites in 5–10 minutes at high field with a generally low level of quantitative signal distortion. It has become standard practice to replace simple solvent resonance presaturation with pulse methods, which either induce such saturation or which leave the solvent water resonance unexcited, and one popular method involves using simply the first increment of the 2-dimensional NOESY pulse sequence.

A vast number of metabolites may appear in urine samples and problems due to signal overlap can occur in single pulse experiments. The magnitude of the signal assignment problem in urine is also apparent at ultrahigh field. There are probably >5000 resolved lines in single pulse 750 or 800 MHz $^1H$ NMR spectra of normal human urine (**Figure 1**), but there is still extensive peak overlap in certain chemical shift ranges and even at

800 MHz many signals remain to be assigned. The dispersion gain at 800 MHz even over 600 MHz spectra is particularly apparent in 2- (or 3-) dimensional experiments that can be used to aid signal assignment and simplify overlapped spectra. Many of the endogenous compounds that have been detected in NMR spectra of human and animal urines are shown in **Table 2**.

The biochemical composition of urine varies considerably from species to species and also with age as almost all species have age-related changes in renal function. Rats and other rodents have much higher levels of taurine, citrate, succinate, 2-oxoglutarate and allantoin than humans. Rat urine (and that of other rodents) is generally much more concentrated than human urine, and so NMR signal-to-noise ratios may be better for many metabolites. All animals have physiological processes that are modulated by biological rhythms. These include excretory processes and the urinary composition of an animal may vary considerably according to the time it is collected. Given these types of variation it is obviously of paramount importance to have closely matched (and in many cases timed) control samples where toxicological or disease processes are being studied. Dietary composition also affects the urinary metabolite profiles of man and animals, and it is important to distinguish these from disease-related processes in clinical or toxicological studies. For example, persons consuming large quantities of meat/poultry before urine collection may have NMR detectable levels of carnosine and anserine in their urine, consumption of cherries is associated with elevated urinary fructose and consumption of shellfish and fish are associated with high levels of betaine and trimethylamine in the urine. The 500 MHz and 600 MHz $^{1}$H NMR spectra of neonate urines are characterized by strong signals from amino acids, organic acids, amines, sugars and polyols. In particular, high concentrations of *myo*-inositol have been detected in both pre- and full-term urines. This is an important intracellular organic osmolyte in renal cells, present in high concentrations in the renal inner medulla.

## NMR spectra of urine in disease

Given the serious outcome of many inborn errors of metabolism, if not diagnosed and treated at an early stage, there has been a search for a rapid, sensitive and general method for the detection and diagnosis of inborn errors of metabolism in neonates. Conventional methods including specific enzyme assays and gas chromatography/mass spectrometry are sensitive but time consuming, involve considerable sample preparation and are not general. However, NMR spectroscopy of biofluids has been shown to be a very powerful and general method for the detection of inborn errors of metabolism and many disorders have been studied by the use of biofluid NMR.

A good example is provided by the spectra from individuals with a deficiency in methylacetoacetyl CoA thiolase (MACT). The conversion of 2-methylacetoacetyl CoA into acetyl-CoA and propionyl-CoA is inhibited and an accumulation of abnormal catabolic products is observed in the urine. Both 1D and 2D $^{1}$H NMR spectroscopy have been used to investigate the urinary metabolites of patients with this disorder. The urine spectra from the patients clearly showed the presence of both 2-methyl-3-hydroxybutyrate and tiglylglycine, which is characteristic for MACT deficiency due to the build up of metabolites close to the position of enzyme deficiency.

A similar success was observed for studies of branched chain ketoaciduria in which the second stage of the catabolism of leucine, valine and isoleucine involves an oxidative decarboxylation. In patients with branched chain ketoaciduria, this step is blocked for all three of these amino acids. The urine of these patients takes on the odour of maple syrup and hence this condition is also known as maple syrup urine disease. $^{1}$H NMR spectroscopy was used to study the urines of patients with such branched chain ketoaciduria. The spectra showed several abnormal metabolites including the amino acids leucine, isoleucine and valine and their corresponding transamination products. It was noted that 2-hydroxyisovalerate levels were very high in the urines of all the patients studied and that, as in other inborn errors of metabolism, the levels of urinary glycine were elevated.

High field $^{1}$H NMR spectroscopy has shown itself to be a very powerful tool with which to diagnose inborn errors of metabolism and to monitor the clinical response of patients to therapeutic interventions. Characteristic patterns of abnormal metabolites are observed in the patient's urine for each disease. The main advantages of using $^{1}$H NMR spectroscopy in this area are its speed, lack of sample preparation and the provision of non-selective detection for all the abnormal metabolites in the biofluid regardless of their structural type, providing only that they are present above the detection limit of the NMR experiment.

There are a number of studies of other disease processes using NMR spectroscopy of urine. These include glomerularnephritis, kidney transplant patients over a period of 14 days post-operation to investigate the biochemical effects of rejection, phenol poisoning and metabolic acidoses caused by alcohol ingestion.

### Evaluation of toxic effects of xenobiotics using NMR spectroscopy of urine

The successful application of $^1$H NMR spectroscopy of biofluids to study a variety of metabolic diseases and toxic processes has now been well established and many novel metabolic markers of organ-specific toxicity have been discovered. The method is based on the fact that the biochemical composition of a biofluid is altered when organ damage occurs. This is particularly true for NMR spectra of urine in situations where damage has occurred to the kidney or liver. It has been shown that specific and identifiable changes can be observed which distinguish the organ that is the site of a toxic lesion. Also it is possible to focus in on particular parts of an organ such as the cortex of the kidney and even in favourable cases to very localized parts of the cortex. Finally it is possible to deduce the biochemical mechanism of the xenobiotic toxicity, based on a biochemical interpretation of the changes in the urine.

A wide range of toxins has now been investigated including the kidney cortical toxins mercury chloride, *p*-aminophenol, ifosfamide, the kidney medullary toxins propylene imine and 2-bromoethanamine hydrochloride and the liver toxins hydrazine, allyl alcohol, thioacetamide and carbon tetrachloride. The testicular toxin cadmium chloride has been investigated in detail and the aldose reductase inhibitor HOE-843 has also been studied.

### Use of combined NMR spectroscopy–pattern recognition (PR) to evaluate biochemical changes in urine

The first studies using PR to classify biofluid samples used a simple scoring system to describe the levels of 18 endogenous metabolites measured by $^1$H NMR spectroscopy of urine from rats which were either in a control group or had received a specific organ toxin which affected the liver, the testes, the renal cortex or the renal medulla. The data were used to construct non-linear maps and various types of principal components (PC) scores plots in 2 or 3 dimensions. The samples were divided into two subsets, a training set and a test set. This study showed that samples corresponding to different organ toxins mapped into distinctly different regions and that in particular the renal cortical and renal medullary toxins showed good separation (**Figure 2**).

In addition, the NMR-PR approach has been used to investigate the time course of metabolic urinary changes induced by two renal toxins. In this case, toxic lesions were induced in rats by a single acute dose of the renal cortical toxin mercury(II) chloride and the medullary toxin 2-bromoethanamine. The onset, progression and recovery of the lesions were also followed using histopathology to provide a definitive classification of the toxic state relating to each urine sample. The concentrations of 20 endogenous urinary metabolites were measured at eight time points after dosing. A number of ways of presenting the data were investigated and it was possible to construct PC scores plots which showed metabolic trajectories which were quite distinct for the two toxins. These showed that the points on the plot can be related to the development of, and recovery from, the lesions. These trajectories allowed the time points at which there were maximal metabolic differences to be determined and provided visualization of the treated groups of animals.

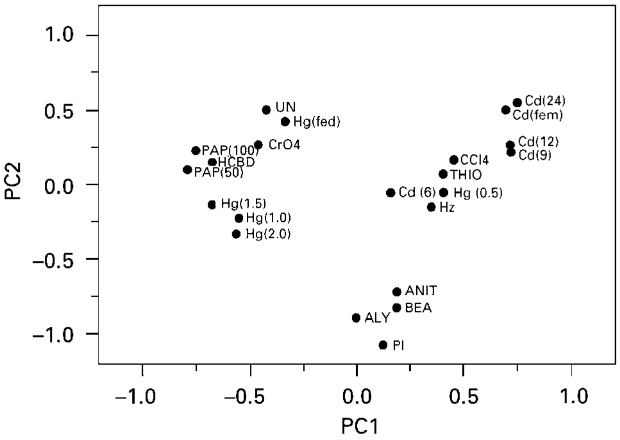



**Figure 2** A plot of PC1 versus PC2 for a set of rat urine samples using simple scored levels of 16 metabolites as the descriptor set. A cluster of renal cortical toxins is seen at the left, whilst the testicular and liver toxins are to the right. At the bottom the two renal papillary toxins (BEA and PI) appear to cluster with liver toxins (ALY and ANIT) but are well separated in the third principal component. Reproduced with permission of Academic Press from Lindon JC, Nicholson JK and Everett JR (1999) NMR spectroscopy of biofluids. *Annual Reports on NMR Spectroscopy* **38**: 1–88.

A toxicological assessment approach based on neural network software has been tested by analysing the toxin-induced changes in endogenous biochemicals in urine as measured using $^1$H NMR spectroscopy to ascertain whether the methods provide a robust approach which could lead to automatic toxin classification. The neural network approach to sample classification was in general predictive of the sample class and once the network is trained, the prediction of new samples is rapid and automatic. The principal disadvantage, common to most neural network studies, in that it is difficult to ascertain from the network which of the original sample descriptors are responsible for the classification.

Recently, studies have been published using pattern recognition to predict and classify drug toxicity effects including lesions in the liver and kidney and using supervized methods as an approach to an expert system. In this case the NMR spectra were segmented into integrated regions automatically, thereby providing standard intensity descriptors for input to the PR.

The usefulness of the NMR-PR approach to the classification of urine samples from patients with inborn errors of metabolism has also been attempted. Urine samples from adult patients with inborn errors of metabolism comprising cystinuria, oxalic aciduria, Fanconi syndrome, porphyria and 5-oxoprolinuria were studied. PCA mapping produced a reasonable separation of control urines from the clusters derived from the different types of inborn error.

The rapid and unambiguous distinction of the various clinical situations that can occur following a kidney transplant remains a problem. For example, it is necessary to distinguish rejection, which requires increased immunosuppressive therapy such as cyclosporin A (CyA), from tubular necrosis, which can be caused by too high levels of CyA. At present, kidney function is generally characterized using the level of blood plasma creatinine as a monitor. However the balance between rejection and tubular necrosis remains a difficult distinction and the only clear diagnosis is an invasive kidney biopsy.

A study of patients after renal transplantation has shown that NMR can reveal new markers of abnormality and in particular that high levels of trimethylamine-*N*-oxide (TMAO) in the urine are correlated with rejection episodes. However, the use of TMAO as a single diagnostic marker of rejection is probably not sufficiently discriminatory for clinical use because of interpatient variability. It is likely that TMAO is only one of a number of potential markers of rejection and that a combination of other metabolites present in urine will provide a new differential diagnosis of rejection and CyA toxicity. A combination of high field $^1$H NMR urinalysis, automatic data reduction and PR methods has been applied to investigate this.

NMR-PR was seen to separate the samples into three distinct groups with the greatest separation being between the CyA toxicity group and the rejection group. It was also notable that, in general, the greater the degree of allograft dysfunction (as assessed by conventional means) the further the samples plotted away from the control group. Identification of the compounds that gave rise to the significant descriptors has not been conclusively proven. However, close examination of spectra from samples from the various classes using both 600 MHz and 750 MHz $^1$H NMR spectroscopy suggested that these were 3-hydroxyisovaleric acid, *N*-acetylated glycoproteins, TMAO, hippuric acid and a molecule related to *N*-methylnicotinamide.

## NMR spectroscopy of cerebrospinal fluid (CSF)

Because of the low protein content of normal human CSF, it is possible to use single pulse $^1$H NMR experiments to obtain biochemical information without recourse to the spin–echo techniques often required for studies on blood and plasma. However, where serious cerebral damage has occurred, or in the presence of an acute infection, spectra may become dominated by broad resonances from proteins. Problems may also arise from protein binding and metal binding of some metabolites, with a consequential broadening of their proton resonances.

There are a number of $^1$H NMR studies of CSF in the literature. Also, it has been shown that high quality 2D $^1$H NMR spectra can be obtained from human CSF, and that changes in NMR patterns can be roughly related to disease states in the donor. Examination of a number of *ex vivo* control samples showed high consistency in the aliphatic region of the 1D $^1$H NMR spectra of CSF and many assignments could be made by inspection. Diseases studied using NMR spectra of CSF include lumbar disk herniations, cerebral tumours, drug overdose, diabetes, hepatic encephalitis, multiple sclerosis, AIDS dementia complex, Parkinson's disease, Creutzfeld–Jakob disease, Guillain–Barre syndrome and vitamin $B_{12}$ deficiency.

One study has reported 500 and 600 MHz $^1$H NMR data on the post mortem CSF from Alzheimer's disease (AD) patients and controls. The main differences between the spectra of the two groups were found to be in the region $\delta$ 2.4–2.9, and principal components analysis showed that separation of the two groups was possible based mainly on lower citrate levels in the AD patients. Non-matching in patient age and the time interval between death and autopsy caused a reduction in the inter-group differences but they were still significant ($p < 0.05$).

## NMR spectroscopy of seminal fluids

The 750 MHz $^1$H NMR spectra of seminal fluid obtained from a healthy individual is complex but many of the signals have been assigned (see **Table 2**). Fresh undiluted seminal fluid gives rise to NMR spectra with very broad and poorly resolved signals due to the presence of high concentrations of peptides (which are cleaved to amino acids by

endogenous peptidase activity) and the high viscosity of the matrix. The complexity of the biochemical composition of seminal fluids together with their reactivity poses a number of assignment and quantitation problems. Some metabolite signals in seminal fluids also appear to have anomalous chemical shifts when compared to simple standard solutions measured at the same pH or even when compared to other body fluids. The application of 2D NMR methods has been shown be very useful in the case of seminal fluids. An investigation of dynamic molecular processes that occur in seminal fluid has been undertaken using $^1$H NMR spectroscopy. Reactions that could be followed included hydrolysis of phosphorylcholine and nucleotides and zinc complexation.

In view of the importance of artificial insemination in farming, it is somewhat surprising that so few studies of animal seminal fluid have been reported. However there are studies on boar seminal plasma giving details of resonance assignments.

The comparison of $^1$H NMR spectra of seminal fluid from normal controls with those from patients with vasal aplasia (obstruction of the vas deferens leading to blockage of the seminal vesicles) and those with non-obstructive infertility has been reported. The $^1$H NMR spectra of the seminal fluid from patients with non-obstructive infertility were similar to those of normal subjects. However, the $^1$H NMR spectra of the seminal fluid from patients with vasal aplasia were grossly different from those of normal subjects and corresponded closely to those of prostatic secretions from normals, owing to the lack of seminal vesicle secretion into the fluid. In the $^1$H NMR spectra of the vasal aplasia patients, signals from amino acids were either absent or present at very low levels. Similarly, choline is at a low level or absent in the seminal fluid from vasal aplasia patients, as it derives (indirectly) from the seminal vesicle component. Significant differences were observed between the normal and vasal aplasia patient groups for the molar ratios of citrate:choline and spermine:choline.

Other studies of infertility include a procedure providing automatic diagnosis based on NMR spectroscopy and work on azoospermic subjects and prostate cancer. In addition $^{31}$P NMR spectroscopy has also been used to distinguish semen from healthy and infertile men.

## NMR spectroscopy of bile

### NMR spectroscopy of bile and dynamic interactions of metabolites

The single pulse NMR spectra of bile are dominated by broad resonances that arise from bile acids that are present in mixed micelles with phospholipids and cholesterol (**Figure 1**). They are broad as a result of short spin–spin ($T_2$) relaxation times reflecting constrained molecular motions within micellar particles. On lyophilization and reconstitution with water, the molecular mobility of a number of biliary metabolites changes significantly because of disruption of the micellar compartments. In particular the $T_2$ relaxation times of the aliphatic side chains of lipid moieties are increased in lyophilized bile suggesting greater mobility of these molecules. Increases in signal intensities that occur on lyophilization reflect changes in compartmentation of molecules, which is related to the disruption/reorganization of the biliary micellar compartments. Signals from $\beta$-hydroxybutyrate, valine and other branched chain amino acids do not contribute significantly to the $^1$H NMR spectra of non-lyophilized bile, but resonances from these components are clearly resolved after lyophilization. The sharper signals in bile give rise to well resolved 2D COSY spectra that allow a comprehensive assignment of the bile salt signals.

Variable temperature $^1$H NMR studies of human bile show that considerable dynamic structural information is available, particularly at high fields. The micellar cholesteryl esters that are abundant in bile appear to show liquid crystal behaviour, and it is possible to use NMR measurements to map the phase diagram for the complex biliary matrix.

A number of studies have used both $^1$H and $^{13}$C NMR spectroscopy of bile to aid characterization of its composition and structure. Thus, $^{13}$C spectra of bile from fish exposed to petroleum have been studied. $^{31}$P NMR spectra of human bile have also been investigated. $^1$H NMR spectroscopy of bile has been used to investigate the micellar cholesterol content and lipids, and both $^1$H and $^{31}$P NMR have been used to study the distribution of lecithin and cholesterol.

The use of $^1$H NMR spectroscopy of bile as a means of monitoring liver function has been proposed. The $^{31}$P NMR spectra of the bile from patients with primary biliary cirrhosis of the liver and from clinically healthy men have been compared. Also the level of lactate in bile from patients with hepatobiliary diseases including cancer has been investigated using $^1$H NMR spectroscopy.

## NMR spectroscopy of miscellaneous body fluids

### Amniotic and follicular fluids

The first study of human amniotic fluid using $^1$H NMR spectroscopy detected 18 small molecule

metabolites including glucose, leucine, isoleucine, lactate and creatinine. Following this, other studies used a combination of 1D and 2D COSY spectroscopy to assign resonances, assess NMR methods of quantitation and investigate the effects of freezing and thawing. In addition NMR results have been correlated with other clinical chemical analyses. A total of 70 samples were measured using $^{1}$H NMR at 600 MHz at different stages of gestation and with different clinical complications and significant correlations between the NMR spectral changes and maternal pre-eclampsia and foetal open spina bifida were observed. The effects of various pathological conditions in pregnancy have been investigated using $^{1}$H NMR spectroscopy of human amniotic fluid. Several studies of amniotic fluid using $^{31}$P NMR spectroscopy have also been carried out, principally to analyse the phospholipid content. One study of the metabolic profiling of ovarian follicular fluids from sheep, pigs and cows has been reported.

### Milk

Surprisingly, very little has been published on the NMR spectroscopy of milk given that it is both a biofluid and a food substance. The studies generally focus on milk as a food and Belton has also reviewed the information content of NMR spectra of milk. Very recently, $^{19}$F NMR spectroscopy has been used to detect trifluoroacetic acid in milk.

### Synovial fluid

$^{1}$H NMR has been used to measure the levels of a variety of endogenous components in the synovial fluid (SF) aspirated from the knees of patients with osteoarthritis (OA), rheumatoid arthritis (RA) and traumatic effusions (TE). The spin–echo NMR spectrum of synovial fluid shows signals from a large number of endogenous components. Many potential markers of inflammation could not be monitored because of their low concentrations or because of their slow tumbling (e.g. hyaluronic acid, a linear polysaccharide that imparts a high viscosity to synovial fluid). The low molecular weight endogenous components showed a wide patient-to-patient variability and showed no statistically significant correlation with disease state. However, correlations were reported between the disease states and the synovial fluid levels of the *N*-acetyl signals from acute phase glycoproteins. Correlations between the disease state and the levels and type of triglyceride in the synovial fluid have also been reported.

The $^{1}$H NMR spectra of the SF of a female patient with seronegative erosive RA and of another female patient with sarcoidosis and independent inflammatory OA were followed over the course of several months and standard clinical tests were performed on paired blood serum samples taken at the same time. It was found that the SF levels of triglyceride $CH_3$, $CH_2$ and CH, glycoprotein *N*-acetyl signals and creatinine all correlated well with one another, and with standard clinical measures of inflammation. The correlation of disease state with creatinine level is of particular interest, and the altered triglyceride composition and concentration in OA was suggested as a potential marker for the disease in SF.

Spin-echo $^{1}$H NMR spectroscopy has been used to detect the production of formate and a low molecular weight, *N*-acetyl-containing oligosaccharide, derived from the oxygen radical-mediated depolymerization of hyaluronate, in the SF of patients with RA, during exercise of the inflamed joint. Gamma radiolysis of rheumatoid SF and of aqueous hyaluronate solutions was also shown to produce formate and the oligosaccharide species. It has been proposed that the hyaluronate-derived oligosaccharide and formate could be novel markers of reactive oxygen radical injury during hypoxic reperfusion injury in the inflamed rheumatoid joint.

$^{13}$C NMR can be used to monitor the synovial fluids from patients with arthritis. In contrast to $^{1}$H NMR studies, signals are seen from hyaluronic acid, the main determinant of the viscoelasticity of the synovial fluid, even though the molecular weight is in the region 500 to 1600 kD. $^{13}$C NMR spectra of synovial fluids from patients with RA, OA, TE and cadaver controls have been compared with one another and with spectra of authentic hyaluronic acid, both before and after the incubation of the latter with hyaluronidase, an enzyme which depolymerizes the hyaluronic acid. Depolymerization of the hyaluronic acid was accompanied by a decrease in the half-band widths of its $^{13}$C resonances. The synovial fluid NMR spectra from the patients with rheumatoid arthritis had sharper signals for the C-1 and C-1′ carbons of hyaluronic acid than those from the osteoarthritic patients, which in turn exhibited sharper signals than those from the cadavers or the joint trauma patients. Thus the degree of polymerization of hyaluronic acid was deduced to decrease in the order: controls/joint trauma patients > osteoarthritic patients > rheumatoid arthritis patients. Since it is known that the consequence of hyaluronate depolymerization may be articular cartilage damage, it was concluded that $^{13}$C NMR spectroscopy may be a valuable method for studying these clinical relevant biophysical changes in synovial fluid.

#### Aqueous humour and vitreous humour

The first study by NMR spectroscopy on aqueous humour was on 9 samples taken during surgery for other conditions, and NMR spectra were measured at 400 MHz. A number of metabolites were detected, including acetate, acetoacetate, alanine, ascorbate, citrate, creatine, formate, glucose, glutamine or glutamate, $\beta$-hydroxybutyrate, lactate, threonine and valine. Following this, there have been a number of other studies. These include $^{1}$H NMR spectra from aqueous humour of rabbits and cod fish, $^{31}$P NMR spectra of aqueous and vitreous humour from pigs and $^{23}$Na NMR spectra of vitreous humour. Finally, the penetration of dexamethasone phosphate into the aqueous humour has been followed using $^{1}$H and $^{19}$F NMR spectroscopy.

#### Saliva

Only limited studies using NMR spectroscopy of saliva have been reported. $^{1}$H NMR spectroscopy of human saliva was used in a forensic study and following this another group reported that only parotid gland saliva gave a well resolved $^{1}$H NMR spectrum showing significant circadian effects. No age- or sex-related differences were observed for saliva from healthy subjects but marked differences were observed in cases of sialodentitis. Finally, the biochemical effects of an oral mouthwash preparation have been studied using $^{1}$H NMR spectroscopy.

#### Digestive fluids

The analysis of pancreatic juice and small bowel secretions using $^{1}$H NMR spectroscopy has been reported.

#### Pathological cyst fluid

A $^{1}$H NMR study of the fluid from the cysts of 6 patients with autosomal dominant polycystic kidney disease (ADPKD) has been reported. ADPKD in adults is characterized by the slow progressive growth of cysts in the kidney, and when these cysts reach a large size they can significantly distort the kidney and disrupt both the blood supply and renal function. ADPKD is one of the commonest causes for renal transplantation in adults. The $^{1}$H NMR spectra revealed a number of unusual features and showed the cyst fluids to be distinct from both blood plasma and urine. Unusually, the fluids from all patients contained high levels of ethanol, which was not related to consumption of alcoholic beverages or drug preparations. In general there was little variation in the composition of the cyst fluids as revealed by $^{1}$H NMR, although the protein signal intensity did vary somewhat. It was hypothesized that this constancy of composition reflected the chronic nature of the accumulation of the cyst fluid and a long turnover time of the cyst components, which thus has the effect of averaging the compositions. The unique biochemical composition of the cyst fluids was ascribed to abnormal transport processes occurring across the cyst epithelial wall.

## Drug metabolites in biofluids

Several studies have used NMR spectroscopy to determine the number and identity of drug metabolites in bile. These include the use of $^{19}$F NMR spectroscopy to study doxifluridine catabolites in human bile, $^{19}$F and $^{13}$C NMR spectroscopy of perfluorinated fatty acids in rat bile and $^{13}$C NMR for monitoring the formation of formaldehyde from demethylation of antipyrine. $^{1}$H NMR spectroscopy of rat bile has been used to monitor the excretion of paracetamol metabolites. A combination of $^{1}$H and $^{1}$H–$^{13}$C 2D NMR methods has allowed identification of 4-cyano-*N*,*N*-dimethylaniline, cefoperazone and benzyl chloride in rat bile.

NMR spectroscopy of urine combined with labelling with stable isotopes such as $^{13}$C and $^{2}$H has been used to monitor the silent process of deaceteylation and subsequent reacetylation (futile deacteylation) in the rat as this has implications for the toxicity of paracetamol and phenacetin. The metabolites of 5-fluorouracil in plasma and urine have been analysed using $^{19}$F NMR spectroscopy in patients receiving chemotherapy. $^{2}$H NMR has been employed to study the pharmacokinetics of benzoic acid in relation to liver function. The metabolism of hydrazine in rats has been probed using $^{15}$N NMR spectroscopy of urine.

Often it is advantageous to carry out a partial extraction of metabolites using a solid phase separation cartridge. In this case, the urine containing the drug metabolites is loaded on to a C18 cartridge that is then washed with acidified water to remove very polar endogenous components. The metabolites of interest can be then be selectively washed off using methanol or water/methanol mixtures.

The direct coupling of HPLC with NMR spectroscopy has required a number of technological developments to make it a feasible routine technique and now on-line NMR detection of HPLC fractions is a useful adjunct to the armoury of analytical methods.

To date, HPLC-NMR spectroscopy has been applied to the profiling and identification of the metabolites of a number of drugs and xenobiotics present in biofluids such as plasma and urine and in bile samples from rats and humans. In general the

simple stop-flow approach has predominated in these studies. The HPLC-NMR method is most useful for compounds, and their metabolites, with suitable NMR reporter groups. This includes the use of $^{19}$F NMR spectroscopy for molecules containing fluorine and diagnostic $^{1}$H NMR resonances in regions of the NMR spectrum where solvent signals do not cause interference. For example, the pattern of NMR resonances from glucuronide conjugates is particularly diagnostic.

HPLC-NMR of biofluids has been employed to identify the unusual endogenous metabolites found in rat urine after administration of compounds that induce liver enzymes leading to elevated urinary levels of a number of carbohydrates and related molecules.

A wide range of $^{1}$H and $^{19}$F HPLC-NMR studies has been carried out on human and animal biofluids, predominantly urine and bile, to characterize the metabolites of exploratory drugs and marketed pharmaceuticals. In some cases, the biofluid has been subjected to solid phase extraction to remove many of the very polar materials, which facilitates the HPLC separation by avoiding overloading the column. These studies include the identification and structuring of metabolites of antipyrine, ibuprofen, flurbiprofen, paracetamol, the anti-HIV compound GW524W91, iloperidone and tolfenamic acid. Other studies using HPLC-NMR on bile include the identification of a metabolite of 7-(4-chlorobenzyl)-7,8,13,13*a*-tetrahydroberberine chloride in rat bile and metabolites of a drug LY335979, under development to prevent multidrug resistance to cancer chemotherapy. In addition, $^{31}$P HPLC-NMR has been used to study metabolites of ifosfamide, and $^{2}$H HPLC-NMR has been employed to detect metabolites of dimethylformamide-$d_7$.

The further coupling of HPLC-NMR with mass spectroscopy has also been used to analyse drug metabolites in urine for the human excretion of paracetamol metabolites. This approach has also been used to identify metabolites of substituted anilines in rat urine.

Capillary electrophoresis (CE), and related techniques, is a relatively new technique which uses a length of fused silica capillary with an optical window to enable detection, a detector (UV, fluorescence or mass spectrometry), a high voltage source, two electrode assemblies and buffer solutions in suitable reservoirs. The technique has been shown to provide very high separation efficiencies, with hundreds of thousands of theoretical plates achievable. However, the small injection volume available to CE (a few nl) means that high sensitivity can only be achieved if concentrations of the analyte in the sample are high. Nevertheless, some results have been reported using CE-NMR and the method has been applied to the identification of paracetamol metabolites in human urine. Finally the application of capillary electrochromatography (CEC) directly coupled to NMR has been explored using the same paracetamol metabolite samples.

*See also:* **Cells Studied By NMR; Chromatography–NMR, Applications; Diffusion Studied Using NMR Spectroscopy; Drug Metabolism Studied Using NMR Spectroscopy; *In vivo* NMR, Applications, Other Nuclei; *In vivo* NMR, Applications, $^{31}$P; *In vivo* NMR, Methods; Perfused Organs Studied Using NMR Spectroscopy; Proteins Studied Using NMR Spectroscopy; Solvent Suppression Methods in NMR Spectroscopy; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals.**

## Further reading

Ala-Korpela M (1997) $^{1}$H NMR spectroscopy of human blood plasma. *Progress in NMR Spectroscopy* **27**: 475–554.

Gartland KPR, Beddell CR, Lindon JC and Nicholson JK (1991) The application of pattern recognition methods to the analysis and classification of toxicological data derived from proton NMR spectroscopy of urine. *Molecular Pharmacology* **39**: 629–642.

Lindon JC, Nicholson JK and Everett JR (1999) NMR spectroscopy of biofluids. In Webb GA (ed) *Annual Reports on NMR Spectroscopy*. Vol 38, in press. London, Academic Press.

Lindon JC, Nicholson JK, Sidelman UG and Wilson ID (1997) Directly-coupled HPLC–NMR and its application to drug metabolism. *Drug Metabolism Reviews* **29**: 705–746.

Lindon JC, Nicholson JK and Wilson ID (1996) Direct coupling of chromatographic separations to NMR spectroscopy. *Progress in NMR Spectroscopy* **29**: 1–49.

Nicholson JK and Wilson ID (1989) High resolution proton magnetic resonance spectroscopy of biofluids. *Progress in NMR Spectroscopy* **21**: 449–501.

Nicholson JK, Foxall PJD, Spraul M, Farrant RD and Lindon JC (1995) 750 MHz $^{1}$H and $^{1}$H–$^{13}$C NMR spectroscopy of human blood plasma. *Analytical Chemistry* **67**: 793–811.

Rabenstein DL (1984) $^{1}$H NMR methods for the non-invasive study of metabolism and other processes involving small molecules in intact erythrocytes. *Journal of Biochemical and Biophysical Methods* **9**: 277–306.

# Biomacromolecular Applications of Circular Dichroism and ORD

**Norma J Greenfield**, University of Medicine and Dentistry of New Jersey, Piscataway, NJ, USA

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Applications**

Most biological macromolecules, including proteins, nucleic acids and carbohydrates, are built of repeating units that are assembled into highly asymmetric structures. The specific conformations of the macromolecules cause the optical transitions of the chromophores in their repeating units (e.g. peptides, nucleotides, glycosides) to align with each other so that their electronic and magnetic transitions interact. The interactions may result in hypo or hyperchromism and in splitting of the transitions into multiple transitions. In addition the aligned transitions may exhibit a high degree of circular dichroism (CD) and optical rotatory dispersion (ORD). Because the ORD and CD spectra of macromolecules depend on their conformation, they are excellent probes for following changes in structure as a function of temperature, denaturants or ligand binding. In addition, non-optically active chromophores sometimes bind to macromolecules in highly asymmetric fashions, and generate large extrinsic CD and ORD bands in the region of absorption. These extrinsic bands can be used to follow interactions of the macromolecules with the chromophores, or to probe structural transitions of the macromolecules. The following article discusses how ORD and CD are used to study the conformation, folding and interactions of proteins, nucleic acids and carbohydrates. Specific examples are given to illustrate each application.

## Proteins and polypeptides

Proteins and polypeptides have two major classes of chromophores, the amide groups of the peptide backbone, which absorb light in the far UV (below 250 nm) and the aromatic amino acid side chains and disulfide bonds, which absorb light in both the near (320–250 nm) and far-UV. Far-UV ORD and CD are useful for studying protein structure and folding because many conformations that are common in proteins, including $\alpha$-helixes, $\beta$-pleated sheets, poly-L-proline II-like helices and turns, have characteristic spectra. **Figure 1** illustrates representative CD spectra of model polypeptides with different secondary structures. In addition, the chromophores of the aromatic amino acids of proteins are often in asymmetric environments, leading to characteristic CD spectra in the near-UV, which may serve as useful probes of the tertiary structure of proteins.

### Analysis of the secondary structure of proteins

Many methods have been developed to extract the conformation of proteins in solution from CD and ORD data. In early studies, the ORD spectra of proteins were analysed to yield structural data, but when commercial CD spectrometers became available in the late 1960s, CD became the method of choice because CD spectra have better resolution

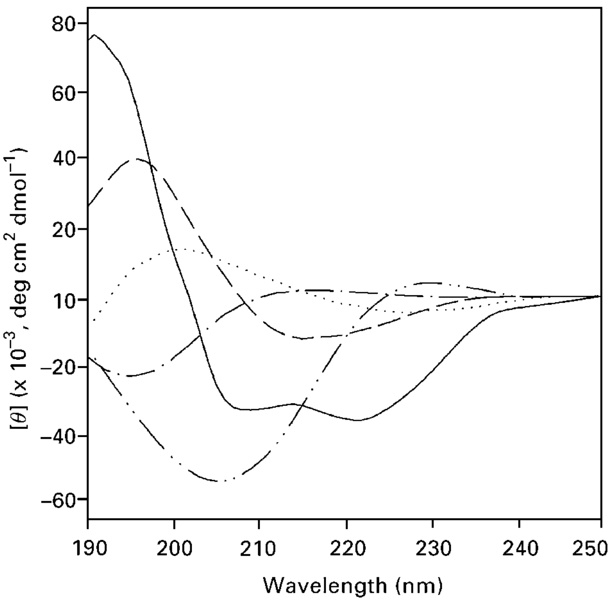



**Figure 1** The circular dichroism of polypeptides with different conformations. (—) $\alpha$-helix, poly-L-lysine in water, pH 11.1, 22°C; (---) $\beta$-sheet, poly(Lys-Leu-Lys-Leu) in 0.5 M NaF at pH 7; (·····) $\beta$-turn (Type III), *N*-acetyl-Pro-Gly-Leu-OH in trifluoroethanol at –60°C; (–·–·) random coil, Glu-Lys-Lys-Leu-Glu-Glu-Ala in 20 mM sodium phosphate, pH 2, 0°C; (–··–··) $P_{II}$, poly-L-proline in trifluoroethanol. Reprinted from data in Greenfield NJ and Fasman GD (1969) *Biochemistry* **8:** 4108–4116, Brahms S and Brahms J (1980) *J. Mol. Biol.* **138**: 149–178, by permission of Academic Press, Venyaminov SY, Baikalov IA, Shen ZM, Wu C-SC and Yang JT (1993) *Anal. Biochem.* **214**: 17–24, and Bovey FA and Hood FP (1967) *Biopolymers* **5:** 325–326 by permission of John Wiley & Sons, Inc.

than ORD spectra. Basically, all of the analysis methods assume that the spectrum of a protein can be represented by a linear combination of the spectra of the secondary structural elements and a noise term. In constrained fits the sum of all the fractional weights of each structure must be equal to one. In the first attempts to elucidate protein conformation from CD spectra, the spectra of proteins were fitted by a combination of the spectra of polypeptides with known conformations using the method of least squares (multilinear regression). Once the conformation of many proteins had been determined from X-ray crystallographic analysis, the CD spectra of several well-characterized proteins were deconvoluted into reference (basis) spectra for the $\alpha$-helix, antiparallel and parallel $\beta$-pleated sheets, $\beta$-turn and random conformations using multilinear-regression analysis. These extracted curves were then used in place of the polypeptide standard curves. Later, singular-value decomposition (SVD) and convex-constraint analysis (CCA) were used to extract the basis curves. The basis curves obtained from deconvoluting a set of protein CD spectra may vary greatly depending on the choice of reference proteins. This occurs because some proteins have unusual CD spectra in the far UV, due to aromatic amino acids, disulfide bridges, or rare conformations. To overcome these difficulties, some methods use selection procedures so that only proteins with spectral characteristics that are similar to those of the protein to be evaluated are used as standards. Methods that use selected reference data include ridge-regression, variable-selection and neural-network procedures.

The following computer programs for analysing CD data to yield the secondary structures of proteins and polypeptides are widely available: MLR is a non-constrained and G&F, LINCOMB and ESTIMATE are constrained least-squares-analysis programs; CONTIN uses the ridge-regression technique; VARSLC and SELCON use singular-value decomposition combined with variable selection to compare the structure of unknown proteins with standard; CCA uses the convex-constraint algorithm to deconvolute sets of CD spectra into a minimum number of basis curves and K2D is a neural-network recall program. **Table 1** compares the agreement between the secondary structures of 16 proteins and poly-L-glutamate, determined by analysis of the CD spectra using these computer programs, with the structures determined by X-ray crystallography.

### Determination of the thermodynamics of protein folding

The change in the CD spectra as a function of temperature can be used to determine the thermodynamics of unfolding or folding of a protein because the observed change in ellipticity, $\theta_{obs}$, is directly proportional to the change in concentration of native and denatured forms. When the protein is fully folded $\theta_{obs} = \theta_F$ and when it is fully unfolded $\theta_{obs} = \theta_U$. For a monomeric protein, the equilibrium constant of folding, $K$ = folded/unfolded. If we define $\alpha$ as the fraction folded at a given temperature, $T$, then

$$K = \frac{\alpha}{1-\alpha} \quad [1]$$

$$\Delta G = nRT \ln K \quad [2]$$

where $\Delta G$ is the free energy of folding, $R$ is the gas constant and $n$ is the number of molecules.

$$\Delta G = \Delta H(1 - T/T_M) - \Delta C_p[(T_M - T) + T(\ln(T/T_M))] \quad [3]$$

where $\Delta H$ is the Van't Hoff enthalpy of folding, $\Delta S$ is the entropy of folding, $T_M$ is the observed midpoint of the thermal transition and $\Delta C_p$ is the change in heat capacity for the transition. At $T_M$, $K = 1$; therefore $\Delta G = 0$ and $\Delta S = \Delta H/T_M$. Rearranging these equations, we obtain:

$$K = \exp(-\Delta G/RT) \quad [4]$$

$$\alpha = \frac{K}{1+K} \quad [5]$$

$$\theta_{obs} = (\theta_F - \theta_U)\,\alpha + \theta_U \quad [6]$$

To calculate the values of $\Delta H$ and $T_M$ that best describe the folding curve, initial values of $\Delta H$, $\Delta C_p$, $T_M$, $\theta_F$ and $\theta_U$ are estimated, and Equation [6] is fitted to the experimentally observed values of the change in ellipticity as a function of temperature, by a nonlinear least-squares curve-fitting routine such as the Levenberg–Marquardt algorithm. Similar equations can be used to estimate the thermodynamics of folding of proteins and peptides that undergo folded multimer to unfolded monomer transitions.

**Figure 2** illustrates how changes in ellipticity as a function of temperature and concentration have been used to determine the enthalpy of folding of a peptide derived from the coiled-coil domain of the yeast GCN4 transcription factor, which undergoes

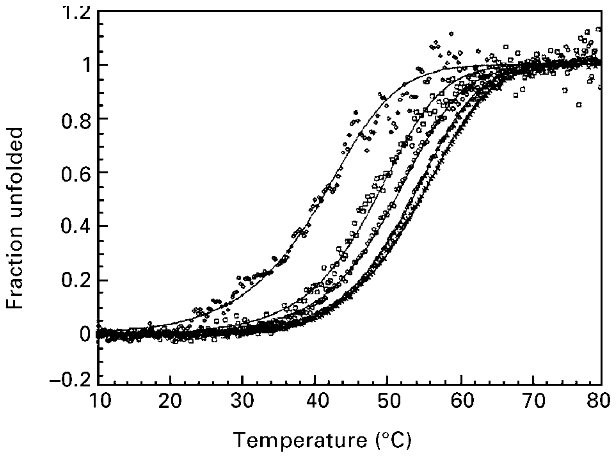



**Figure 2** The change in fraction folded of a coiled-coil leucine zipper peptide, GCN4-P1, as a function of temperature and peptide concentration. The concentrations range from 1 µM with the lowest $T_M$ to 20 µM with the highest $T_M$. The enthalpy of unfolding was 35.0 ± 1.1 kcal mol$^{-1}$ (monomer) which compared to that of 34.7 ± 0.3 kcal mol$^{-1}$ measured by calorimetry. Abstracted with permission from data in Thompson KS, Vinson CR and Freire E (1993) *Biochemistry* **32:** 5491–5496. Copyright 1993 American Chemical Society.

a two-state transition between a folded two-stranded $\alpha$-helical coiled coil and a monomeric disordered state. The values for the enthalpy of the folding transition, determined from the changes in ellipticity as a function of temperature, agree with the values obtained from scanning calorimetry.

In some cases thermal folding and unfolding studies are impractical. The protein may have a very high $T_M$, or may aggregate and precipitate upon heating. In these cases, denaturants, such as urea or guanidine-HCl, may be added to a protein or peptide to induce unfolding. *K*, the equilibrium constant of folding, is determined from the ellipticity observed at each concentration of denaturant. It is assumed that the protein is native in the absence of denaturant, and fully unfolded when there is no further change in ellipticity upon addition of higher concentrations of perturbant. At each concentration of denaturant,

$$\alpha = (\theta_{obs} - \theta_U)/(\theta_F - \theta_U) \quad [7]$$

$$K = \alpha/(n[C]^{n-1})(1 - \alpha)^n \quad [8]$$

$\alpha$, $\theta_F$ and $\theta_U$ have the same definitions as in thermal unfolding above and *n* is the number of identical subunits. [*C*] is the total concentration of protein monomers. The free energy of folding is evaluated at every concentration of denaturant using Equation [2] and is plotted as a function of denaturant. The curve is extrapolated to zero denaturant to obtain the free energy of folding of the native material. **Figure 3** depicts the determination of the free energy of folding of a coiled-coil DNA binding protein from a urea-denaturation study. The peptide forms a two-stranded coiled-coil $\alpha$-helix when folded and is a single-stranded unordered peptide when dissociated. The association constant and free energy of folding, determined from the denaturation data (**Figures 3B** and **3C**), agree with those determined from the change in ellipticity as a function of peptide concentration (**Figure 3A**).

### Analysis of conformational transitions

Some native proteins undergo conformational transitions that do not result in formation of a denatured state. Such transitions may involve changes from an $\alpha$-helical to a $\beta$-pleated sheet conformation (or vice versa). Conformational changes of proteins are sometimes followed by oligomerization and/or aggregation and precipitation reactions that may have important clinical consequences. For example, transitions from a random or $\alpha$-helical form to a $\beta$-sheet may be involved in the pathology of diseases such as scrapie, 'mad cow disease' and Alzheimer's disease. **Figure 4** illustrates how CD has been used to follow the thermal stability and conformational transitions of the scrapie amyloid (prion) protein, PrP27-30. The CD spectra of the solvent-exposed (**Figure 4A**) and rehydrated solid state PrP27-30, obtained under various conditions, were deconvoluted using the CCA algorithm and five common spectral components were identified (**Figure 4B**). Infectivity quantitatively correlated with an increasing proportion of a native, $\beta$-sheet-like secondary-structure component (**Figure 4C**), a decreasing amount of an $\alpha$-helical component, and an increasingly ordered tertiary structure.

### Detection of folding intermediates

Often when a protein or peptide folds or unfolds, the process is not a two-state transition between the native and totally disordered forms; intermediate states exist. These states may be transient, and may be observed during kinetic measurements of folding or unfolding. Alternatively, they may be stable intermediates, seen when the peptide is subject to denaturing conditions, such as high temperatures, or exposure to urea, guanidine or detergents. These states may be partially folded, with well-defined tertiary structures, or may be so called 'molten

**Table 1** Comparisons of methods of analysing protein conformation from circular dichroism data[a]

| Computer program | Standards | Wavelength range, nm | *α-helix* | | *β-sheet* | | *β-turn* | |
|---|---|---|---|---|---|---|---|---|
| | | | *P* | *σ* | *P* | *σ* | *P* | *σ* |
| *Linear regression – unconstrained fit* | | | | | | | | |
| MLR (1) | 4 peptides (1) | 240–178 | 0.91 | 0.13 | 0.43 | 0.21 | 0.07 | 0.16 |
| MLR (1) | 4 peptides (1) | 240–200 | 0.92 | 0.14 | 0.74 | 0.16 | 0.23 | 0.16 |
| *Linear regression – constrained fit* | | | | | | | | |
| G&F (2) | Poly-L-lysine (2) | 240–208 | 0.92 | 0.13 | 0.61 | 0.18 | ND | ND |
| LINCOMB (3) | 4 peptides (1) | 240–178 | 0.93 | 0.11 | 0.58 | 0.15 | 0.61 | 0.11 |
| LINCOMB (3) | 4 peptides (1) | 240–200 | 0.94 | 0.11 | 0.71 | 0.13 | 0.53 | 0.14 |
| LINCOMB (3) | 17 proteins (4) | 240–178 | 0.94 | 0.09 | 0.62 | 0.14 | 0.21 | 0.13 |
| LINCOMB (3) | 17 proteins (4) | 240–200 | 0.92 | 0.10 | 0.09 | 0.28 | 0.52 | 0.12 |
| *Singular-value decomposition* | | | | | | | | |
| SVD (4) | 17 proteins (4) | 240–178 | 0.98 | 0.05 | 0.68 | 0.12 | 0.22 | 0.10 |
| SVD (4) | 17 proteins[b] (4) | 240–200 | 0.96 | 0.07 | 0.43 | 0.14 | 0.04 | 0.13 |
| *Convex-constraint algorithm* | | | | | | | | |
| CCA (5) | 17 proteins (4) | 260–178 | 0.96 | 0.10 | 0.62 | 0.18 | 0.39 | 0.18 |
| CCA (5) | 17 proteins (4) | 240–200 | 0.97 | 0.10 | 0.42 | 0.20 | 0.52 | 0.22 |
| *Ridge regression* | | | | | | | | |
| CONTIN (6) | 17 proteins (4) | 260–178 | 0.93 | 0.11 | 0.56 | 0.15 | 0.58 | 0.08 |
| CONTIN (6) | 17 proteins (4) | 240–200 | 0.95 | 0.13 | 0.60 | 0.15 | 0.74 | 0.07 |
| *Variable selection* | | | | | | | | |
| VARSLC (7) | 17 proteins (4) | 260–178 | 0.97 | 0.07 | 0.81 | 0.10 | 0.60 | 0.07 |
| *Variable selection – self consistent method* | | | | | | | | |
| SELCON (8) | 17 proteins (4) | 260–178 | 0.95 | 0.09 | 0.84 | 0.08 | 0.77 | 0.05 |
| SELCON (8) | 17 proteins (4) | 260–190 | 0.94 | 0.09 | 0.73 | 0.09 | 0.84 | 0.05 |
| SELCON (8) | 17 proteins (4) | 240–200 | 0.93 | 0.10 | 0.73 | 0.11 | 0.71 | 0.06 |
| SELCON (8) | 33 proteins (9) | 260–178 | 0.93 | 0.09 | 0.91 | 0.07 | 0.53 | 0.09 |
| SELCON (8) | 33 proteins (9) | 260–200 | 0.88 | 0.12 | 0.86 | 0.09 | 0.46 | 0.09 |
| *Neural-network analysis* | | | | | | | | |
| K2D (10) | 18 proteins (10) | 240–200 | 0.95 | 0.09 | 0.77 | 0.10 | ND | ND |
| | | | *Mean* | *STD* | *Mean* | *STD* | *Mean* | *STD* |
| | | | 0.36 | 0.27 | 0.20 | 0.16 | 0.22 | 0.08 |

[a] The secondary structures of 16 proteins plus poly-L-glutamate were analysed using each computer program as described by Sreerama and Woody (8). *P* is the correlation coefficient between the CD-estimated and X-ray conformations and *σ* is the mean-square error between the CD-estimated and X-ray conformations. The secondary structures were assigned by the method of Kabsch and Sander (11). When protein databases were used as standards, each protein analysed was excluded from the data set used as the references. When the *σ* value is higher than the standard deviation (STD) of the mean value of each conformation found in the 17 samples which were analysed, the program does a relatively poor job of analysing that conformation.

[b] PGA was excluded from the calculation of *P* and *σ* because the fit was obviously impossible.

(1) Brahms S and Brahms J (1980) Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *Journal of Molecular Biology* **138**: 149–178.

*Continued*

**Table 1** *continued*

(2) Greenfield N and Fasman GD (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8**: 4108–4116.

(3) Perczel A, Park K and Fasman GD (1992) Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: a practical guide. *Analytical Biochemistry* **203**: 83–93.

(4) Hennessey JP and Johnson WC Jr (1981) Information content in the circular dichroism of proteins. *Biochemistry* **20**: 1085–1094.

(5) Perczel A, Hollósi M, Tusnady G and Fasman GD (1991) Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins. *Protein Engineering* **4**: 669–679.

(6) Provencher SW and Glöckner J (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**: 33–37.

(7) Manavalan P and Johnson WC Jr (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Analytical Biochemistry* **167:** 76–85.

(8) Sreerama N and Woody RW (1994) Poly(pro)II helices in globular proteins: identification and circular dichroic analysis. *Biochemistry* **33**: 10022–10025.

(9) Toumadje A, Alcorn SW and Johnson WC Jr (1992) Extending CD spectra of proteins to 168 nm improves the analysis for secondary structures. *Analytical Biochemistry* **200**: 321–331.

(10) Andrade MA, Chacón P, Merelo JJ and Morán F (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Engineering* **6**: 383–390.

(11) Kabsch W and Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**: 2577–2637.

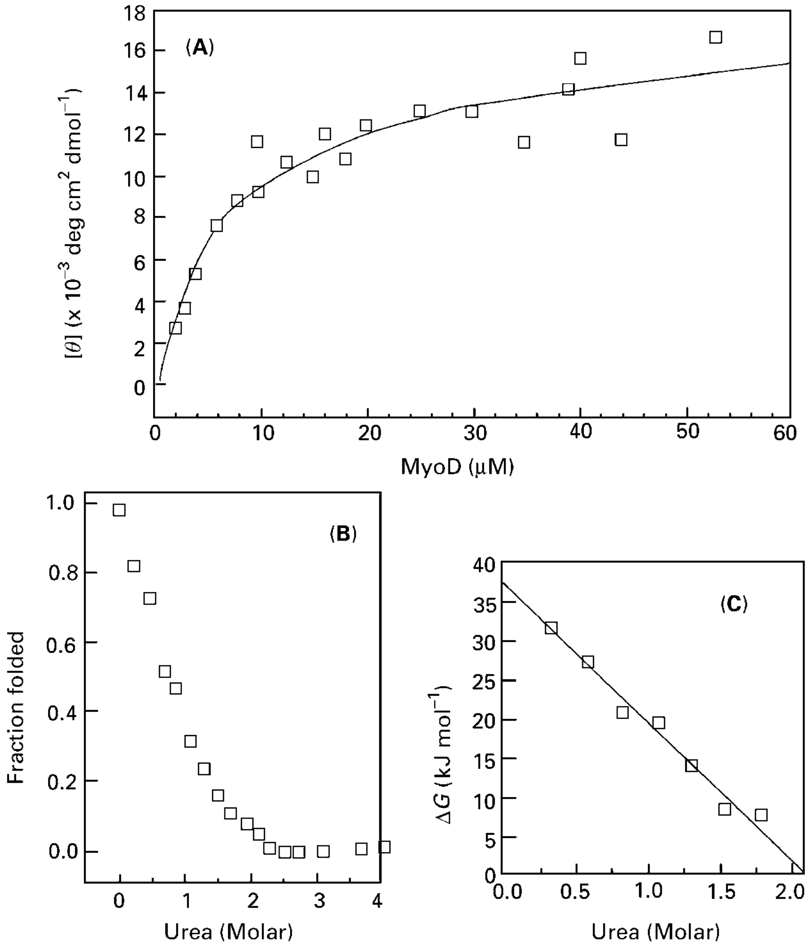



**Figure 3** The effects of concentration and urea on the folding of the helix-loop-helix DNA binding domains of the muscle-regulatory DNA transcription factors, MyoD. (A) The change in ellipticity at 222 nm of Myo D as a function of concentration. The data were used to calculate the constant of dimerization $K_{dimer}$ = 9.4 μM. (B) The change in fraction folded of MyoD as a function of urea concentration. (C) $\Delta G$ of folding calculated from the data in (B). The extrapolated free energy of folding calculated from the value at 0 urea was similar to those calculated from the dimerization constant. Redrawn with permission from data in Wendt H, Thomas RM and Ellenberger T (1998) *J. Biol. Chem.* **273**: 5735–5743. Copyright 1998. The American Society for Biochemistry & Molecular Biology.

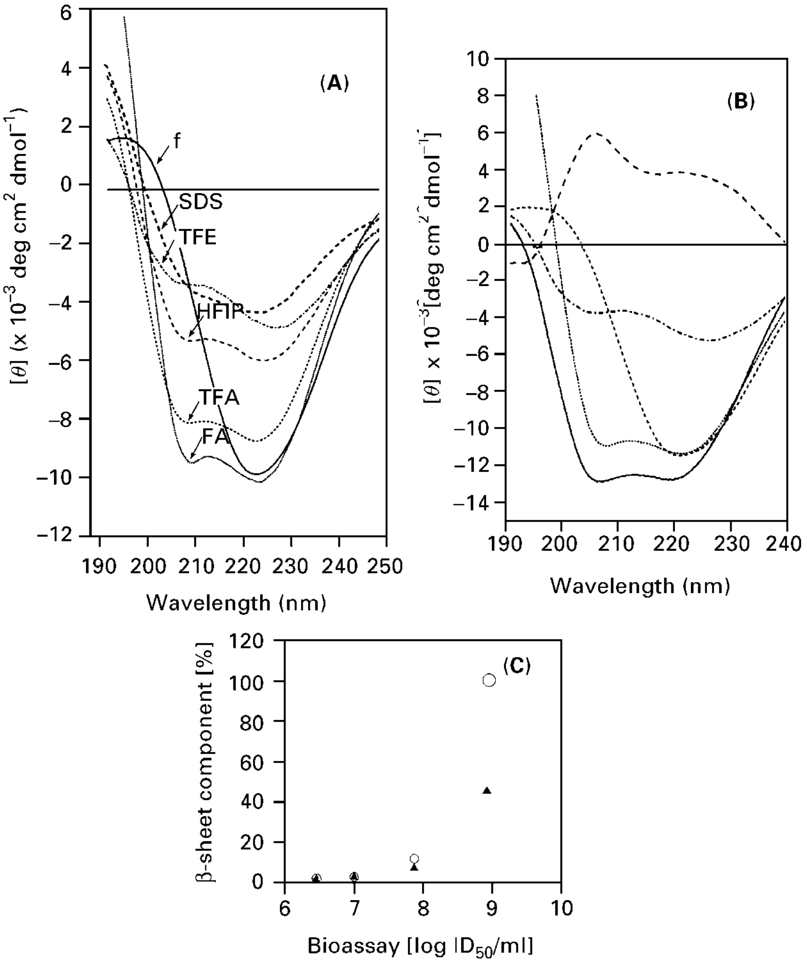



**Figure 4** (A) Solvent-induced conformational transition of PrP27-30 scrapie amyloid (prion) protein in the solid state on quartz glass. Films were exposed for 20 to 30 min to: SDS, sodium dodecyl sulfate, TFE, trifluorethanol; HFIP, hexafluorisopropanol; TFA, trifluoroacetic acid; FA, 99% formic acid. "f" is the spectrum with no treatment at 23°C. (B) Data on PrP27-30 obtained under various conditions, deconvoluted into five basis curves using the convex-constraint algorithm. (—) $\beta$-turn type I or III and/or $3_{10}$-type helix; (·····) alpha helix; (----) $\beta$-sheet; (– – –) $\beta$-turn type II; (–·–·) additional chiral contribution. (C) Correlation of the solvent-induced changes in the $\beta$-sheet component with infectivity. The data for dehydrated films are open circles and rehydrated films are closed circles. The data obtained with TFA and FA, where the proteins were highly helical and the infectivity was very low, are excluded from the fits. (o) dehydrated films (▲) rehydrated films. Redrawn with permission from data in Safar J, Roller PP, Gajdusek DC and Gibbs CJ Jr (1993) *Protein Sci.* **2**: 2206–2216. Copyright 1993 The Protein Society.

globules'. 'Molten globule' is a term describing a compact state with native-like secondary structure but slowly fluctuating tertiary structure. For example, methanol induces the formation of a molten-globule state in the globular protein, cytochrome C. **Figure** 5 illustrates the effect of increasing quantities of methanol on the CD spectra of cytochrome C in the near and far UV. A loss of ellipticity occurs in the near UV, that is present in the native state arising from the tertiary interactions of the aromatic chromophores (**Figure** 5A). The far-UV spectra (**Figure** 5B), however, show that the helical content of cytochrome C actually increases upon the addition of methanol, as shown by the increasing ellipticity at 222 and 208 nm.

## Determination of the kinetics of protein folding

Besides the study of protein folding under equilibrium conditions, CD can be used to measure rapid events in protein folding and unfolding by attaching a rapid mixing device to a CD spectropolarimeter. Stopped-flow CD can be used both to obtain the kinetic constants of folding reactions and to define the conformation of folding intermediates. Data can be collected in the far and near UV. The far-UV data give information about the secondary structure of the protein, while the near-UV data are windows that monitor the formation of tertiary structure.

Stopped-flow CD studies of highly helical two-stranded coiled coils and four-helix bundle proteins

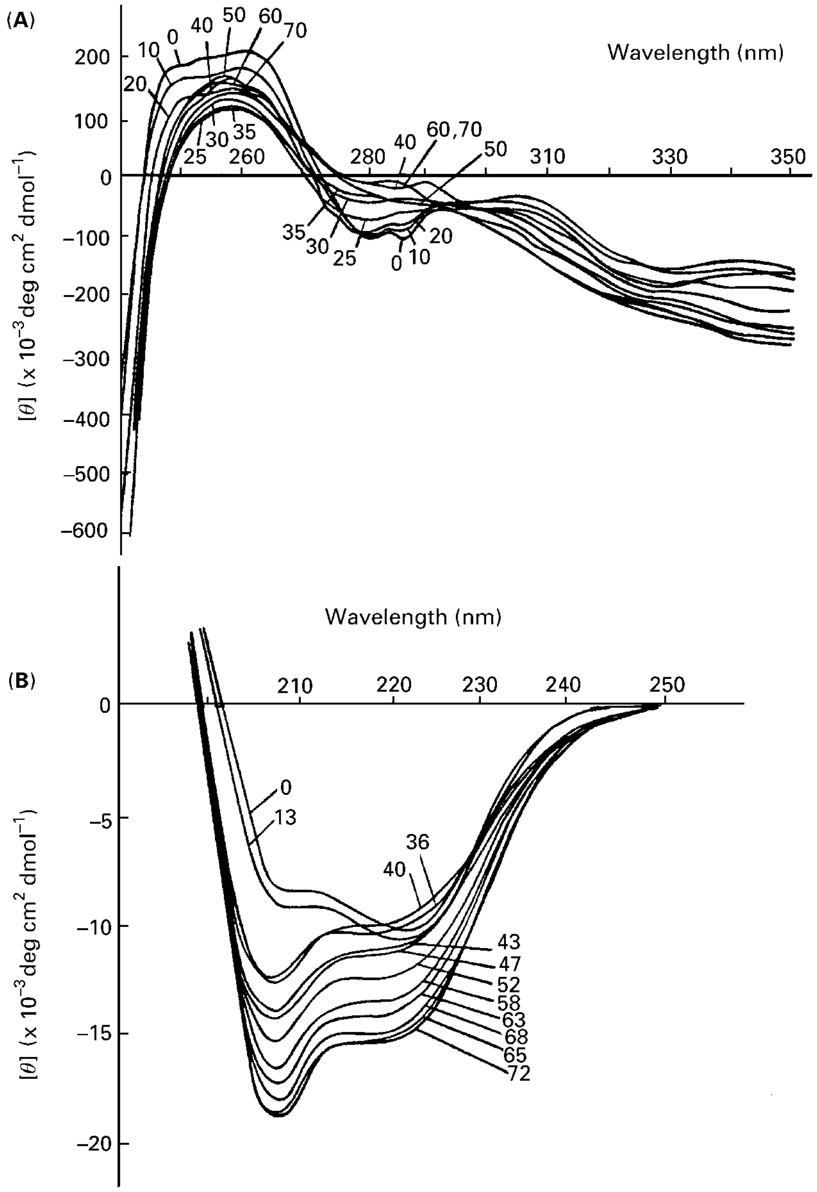



**Figure 5** (A) Near-UV CD and (B) Far-UV CD spectra of cytochrome c (pH 4.0 in 0.5 M NaCl, 24°C) at methanol concentrations shown near the curves. The negative bands in the near-UV spectrum are attributed to the side chain of the single Trp59. They decrease and almost vanish at methanol concentration > 40%. Abstracted with permission from data in Bychkova VE, Dujsekina AE, Klenin SI, Elisaveta IT, Uversky VN and Ptitsyn OB (1996) *Biochemistry* **35**: 6058–6063. Copyright 1996 American Chemical Society.

have shown them to fold in essentially two-state transitions. The kinetics of folding of globular proteins, however, can be very complex. For example, the kinetics of the unfolding and refolding of bovine beta-lactoglobulin, a predominantly $\beta$-sheet protein in the native state, are illustrated in **Figure 6**. The kinetics of unfolding (**Figure 6A, B**) show a single-phase transition between the folded and denatured form, with the loss of secondary structure (**Figure 6A**) and tertiary structure (**Figure 6B**) being simultaneous. The refolding reaction, in contrast, is a complex process composed of different kinetic phases (**Figure 6C**). In particular, a burst-phase intermediate is formed during the dead time of stopped-flow measurements that shows more intense ellipticity signals in the far UV than does the native state. This more-intense CD is due to the transient formation of a non-native alpha-helical structure (**Figure 6D**). The CD spectrum suggests the folding intermediate exists in a molten globule state, since there are no near-UV CD bands indicative of tertiary structure. Similarly, intermediate states have been shown to exist in the refolding of diverse proteins including cytochrome C, ribonuclease HI from *E. coli*, dihydrofolate

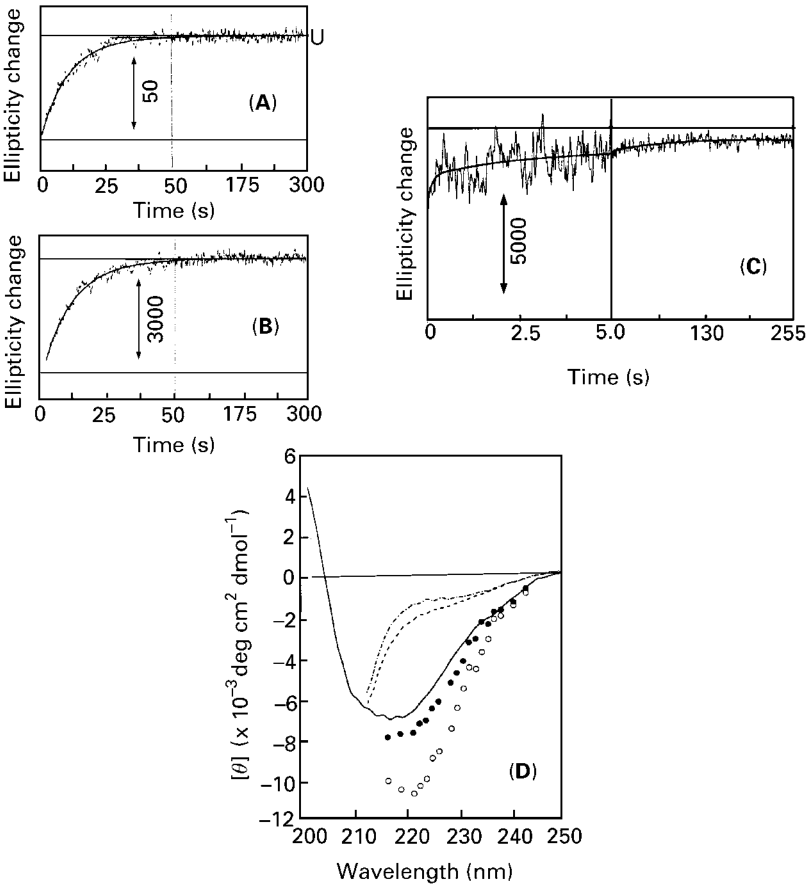



**Figure 6** (A) Unfolding curve of $\beta$-lactoglobulin ($\beta$-LGA) measured by the CD changes at 293 nm at pH 3.2 at 4.5°C. The change was initiated by jumping the guanidine HCl (GdnHCl) concentration from 0 to 4.0 M. The vertical arrows show the change in ellipticity in deg.cm$^2$.dmol$^{-1}$ (B) Same, measured at 220 nm. (C) Kinetic refolding curve of $\beta$-LGA measured by the CD change at 221 nm. Refolding was initiating by dropping the concentration of GdnHCl from 4 to 0.4 M. (D) CD spectrum of $\beta$-LGA in the far UV at pH 3.2 and 4.5°C. (—) native state; (---) unfolded state in 4 M GdnHCl; (–·–·) disulfide-cleaved and carboxymethylated material in 4 M GdnHCl; (o) ellipticity of $\beta$-LGA at 0 time after refolding; (•) ellipticity of $\beta$-LGA at 255 min after refolding. Reprinted from data in Kuwajima K, Yamaya H and Sugai S (1996) *J. Mol. Biol.* **264**: 806–822, by permission of Academic Press.

reductase, alcohol dehydrogenase, carbonic anhydrase, retinoic acid-binding protein and staphylococcal nuclease.

## Analysis of protein–ligand interactions

Circular dichroism can be used to follow the binding of ligands to proteins, peptides, nucleic acids and carbohydrates provided one of two criteria is fulfilled. Either the ligand must bind in an asymmetric fashion that induces *extrinsic* optical activity in the chromophores of the bound ligand, or the binding must result in a conformational change in the macromolecule that results in a change in its *intrinsic* CD spectrum. In the case of equivalent binding sites, the change in ellipticity due to complex formation is directly proportional to how much ligand is bound. Thus, the change in ellipticity as a function of substrate concentration can be used to estimate the binding constants.

Extrinsic ellipticity bands (called Cotton effects), developed when chromophores bind to proteins, have been used to study protein–ligand interactions in many diverse systems. **Figure 7A** illustrates the induction of extrinsic CD bands when retinol binds to interphotoreceptor retinol-binding protein, and **Figure 7B** shows the change in the ellipticity at the wavelength of maximal ellipticity as a function of ligand concentration. Extrinsic Cotton effects have been used to study the binding of many other cofactors to proteins, including the binding of nucleotides to lactic dehydrogenase, folates and antimetabolites to dihydrofolate reductase, ATP to GroEL and pyridoxal phosphate to tryptophan synthase. They have also been used to study the interactions of substrates with enzymes including tryptophan synthase and *N*-acetylglucosamyl transferase, to study the interactions of drugs with bovine serum albumin, and to examine the interactions of DNA binding proteins with DNA.

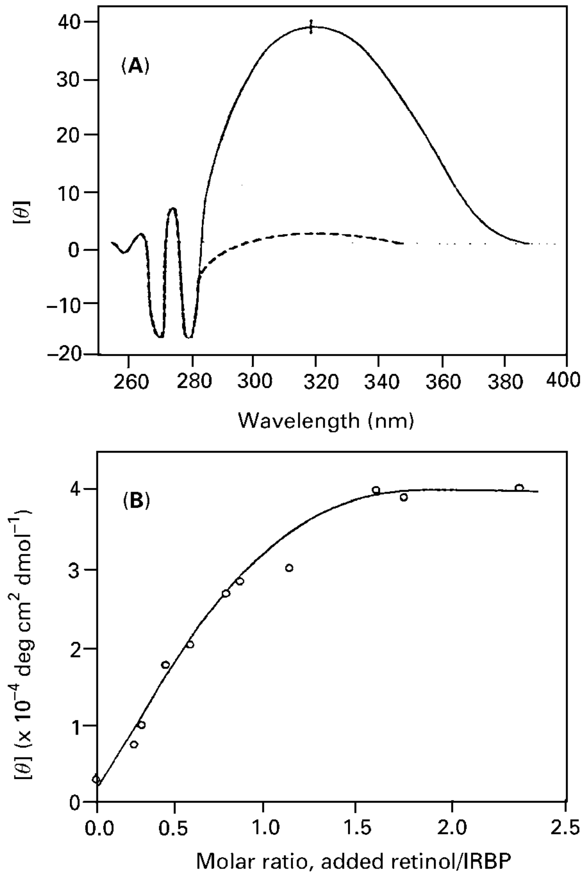



**Figure 7** (A) Near-UV mean-residue circular-dichroism ellipticity, [$\theta$], in deg.$cm^2$.$dmol^{-1}$, of interphotoreceptor retinol-binding protein (IRBP), 0.4–0.8 mg $ml^{-1}$, with and without bound retinol. (B) Circular-dichroism titration curve of apo-IRBP, 0.4 mg $ml^{-1}$, with retinol. The height of the extrinsic Cotton effect at 330 nm is calculated in ellipticity per mole of IRBP. Redrawn with permission from data in Adler AJ, Evans CD and Stafford WF (1985) *J. Biol. Chem.* **260**: 4850–4855. Copyright 1985 The American Society for Biochemistry & Molecular Biology.

The binding of ligands to proteins often results in conformational transitions. The resultant change in the intrinsic ellipticity of the backbone amides caused by the conformational change can be used to obtain the ligand-binding constants. For example, divalent cations are effectors in many biological systems, and their binding often induces large conformational changes in the target protein. Calcium-ion binding to calmodulin and troponin C induces an increase in the ellipticity of the proteins due to stabilization of $\alpha$-helical regions. Many other broad classes of protein-ligand interactions result in conformational changes. For example, DNA binding to many different DNA transcription factors results in increases in the helical content of the factors. These DNA binding proteins often contain $\alpha$-helical-coiled-coil or helix-loop-helix motifs, which serve as dimerization domains, and a relatively unstructured basic region which binds to the DNA. Upon DNA binding the basic region assumes a helical conformation.

## Nucleic acids

Nucleic acids are polymers of nucleotides that consist of purine or pyrimidine bases attached to a phosphorylated sugar. Monomeric sugars are asymmetric and they induce a small amount of ellipticity in the attached planar-symmetric base. Polymerized nucleic acids, however, are rigid and highly asymmetric. In the polymerized nucleic acids, the bases stack with one another in precise orientations depending on the conformational state. Base stacking leads to characteristic splittings of the transitions of chromophores of the bases into multiple transitions, and the rigidity leads to CD bands with high rotational strength. CD spectra of polynucleotides are sensitive both to their sequence and conformation, and can provide a good deal of structural information.

### Analysis of the secondary structure of polynucleotides

**Homopolynucleotides** Even very simple polynucleotides show multiple conformational states and complex CD and ORD spectra. Compounds as small as homodimers of ribo and deoxyribonucleotides show evidence of base stacking in solution from measurements of optical activity. For example, **Figure 8A** illustrates the circular dichroism properties of several riboadenylate compounds. AMP shows a single absorption band and a single negative CD band at 260 nm. Diadenylic acid has a single UV band at 260 nm, but a split CD spectrum with a positive band at 270 nm and a negative band at 250 nm. Theoretical calculations of the CD spectrum agree with a model where the bases are stacked in a parallel fashion in the dinucleotide. The stacking leads to hypochromicity in the absorption spectrum, and splitting of the transitions into two transitions, one perpendicular and one parallel to the helix axis with equal strength and opposite rotatory strength. Poly(rA) has a spectrum similar to that of diadenylate but higher in magnitude.

The CD spectrum of poly(rA) changes with pH, chain length and temperature. At neutral pH, poly(rA) is unprotonated and forms a single-stranded helix. At acidic pH, poly(rA) forms dimers. Depending on the pH there are two conformational forms, one where poly(rA) is fully protonated, and one where it is half-protonated. **Figure 8B** shows the

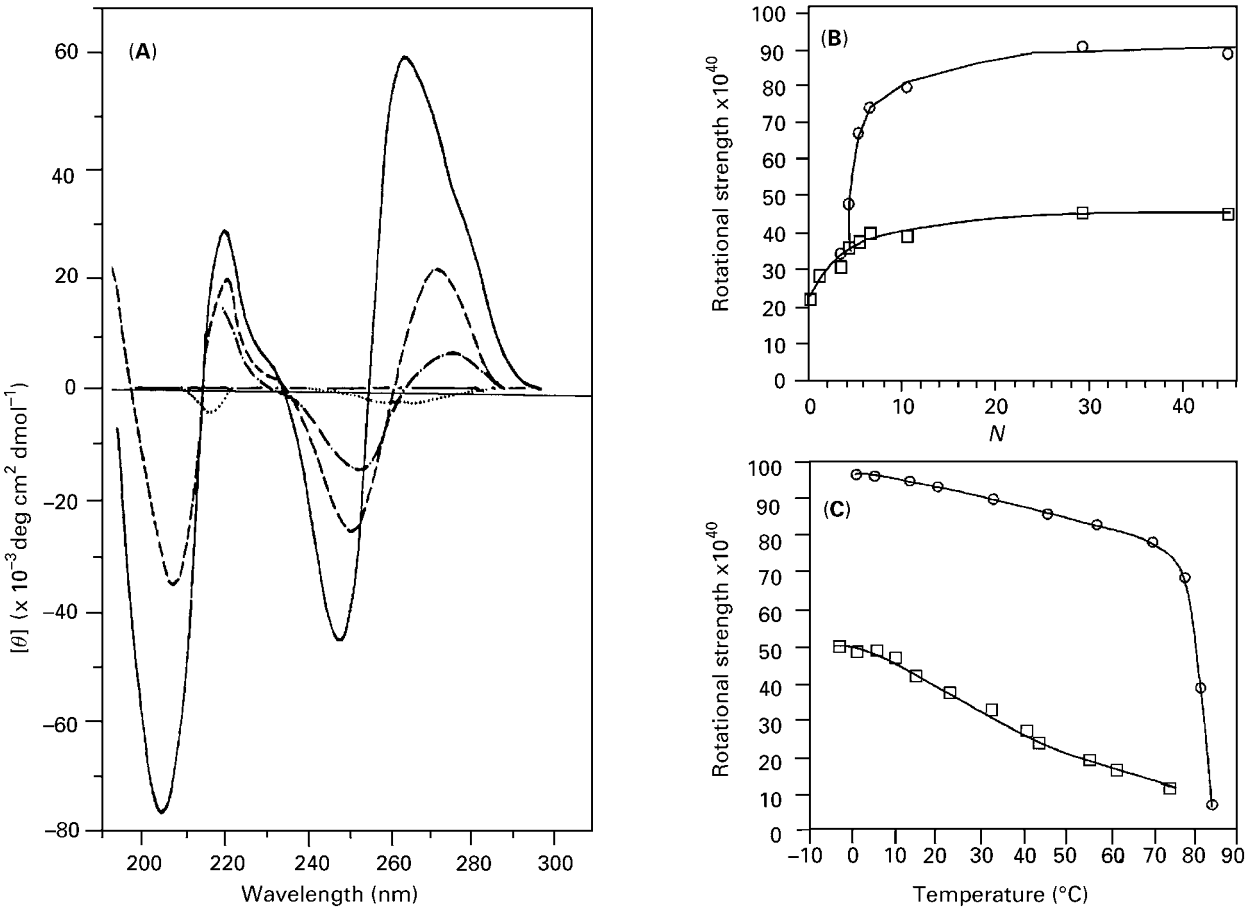



**Figure 8** (A) Circular dichroism of adenylate compounds at pH 8.5 in unbuffered 0.1 M NaF at 22°C. (—) poly (rA); (–·–·) rA-2'O-methyl-prAp; (-----) rA(3′p5)rA; (·····) rAMP-5'. Abstracted with permission from data in Adler AJ, Grossman L and Fasman GD (1969) *Biochemistry* **8**: 3846–3859. Copyright 1969 American Chemical Society. (B) Rotational strength of the positive band of adenylate oligomers at 0°C as a function of the chain length, N. (□) pH 7.0; (○) pH 4.5. (C) Rotational strength of the positive band of poly(rA) as a function of temperature at (□) pH 7.4 and (○) pH 4.5. (B) and (C) are reprinted from data in Brahms J, Michelson AM and Van Holde KE (1966) *J. Mol. Biol.* **15**: 457–488, by permission of Academic Press.

chain-length dependence, and **Figure 8C** the temperature dependence, of the rotational strength of the positive-ellipticity band of the single-stranded form at pH 7.4 and the double-stranded half-protonated form at pH 4.5. Both the chain-length dependence and the unfolding of the double-stranded forms as a function of temperature, are more cooperative than that of the single-stranded form.

**Heteropolynucleotides** Circular dichroism spectra of nucleic acids are complex, compared to those of homopolynucleotides because four different bases contribute to the CD spectra, and the spectra are sequence dependent. In addition, nucleic acids display great conformational diversity and may form single-, double- or triple-stranded helices. For example, the spectra of poly(rA), poly(rU), the dimer poly(rA)·poly(rU) and the trimer poly(rU)·poly(rA)·poly(rU) are illustrated in **Figure 9.** Complex formation results in shifts in wavelength and intensity of the bands, compared to the addition of the spectra of the unmixed components.

The various oligomerization states of the polynucleotides, moreover, can exist in multiple conformations. For example the A, B and Z forms of double-stranded nucleic acids have distinct CD spectra. **Figure 10A** illustrates the CD spectra of poly (dAdC)·poly(dGdT) in the A, B and Z conformations and **Figure 10B** illustrates CD spectra of a double-helical polynucleotide, poly(dGdC)·poly(dGdC) in the B form, the Z form and in a form in which the bases assume the Hoogsteen base-pairing conformation.

## Analysis of conformational transitions of nucleic acids

Conformational transitions of nucleic acids can be followed by examining CD and ORD spectra as a

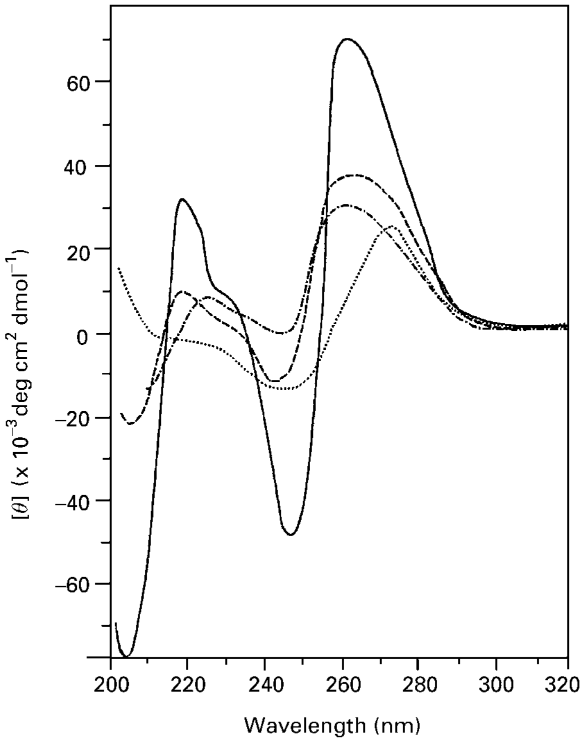



**Figure 9** The CD spectra of (——) Poly(rA); (·····) Poly(rU); (----) poly(rU)·poly(rA); (—·—) poly (rU)·poly (rA)·poly (rU) at pH 7 in sodium phosphate buffer. Redrawn with permission from data in Steely T, Gray DM and Ratiff RL (1986) *Nucleic Acid Research* **14**: 10 071–10 090. Copyright 1986 Oxford University Press.

function of temperature and denaturants, and the thermodynamic parameters of the transitions can be determined using the same equations used to determine the thermodynamics of folding of proteins. In addition, the spectra of nucleic acids, obtained under varying conditions, can be deconvoluted using the CCA and SVD methods, to determine whether there are folding intermediates, as in the case of proteins.

### Analysis of ligand binding to nucleic acids

Circular dichroism and ORD are excellent methods to follow the binding of ligands to nucleic acids. The same equations that are used to analyse protein–ligand interactions may be applied to the study of DNA–ligand interactions. For example, CD and ORD have been used to examine the binding of DNA to proteins, drugs and amines. **Figure 11** illustrates the effect of binding of spermidine to poly(dT)·poly(dA)·poly(dT). The addition of the amine has dramatic stabilizing effects. The conformation of the polynucleotide complex undergoes sequential changes from B-DNA to triplex DNA as the concentration of spermidine is increased from 0 to 50 μM. At 60 μM spermidine, the CD spectrum of triplex DNA is comparable to that of *ψ*-DNA, with a strong positive band centred around 260 nm. A negative band is also found at 295 nm. At higher concentrations of spermidine, however, the intensity of the positive band progressively decreases. The peak intensity is found at a 1:0.3 molar ratio of DNA phosphate-spermidine.

**Figure 12** illustrates an example of DNA-protein binding study. The gene 32 coded protein of bacteriophage T4 is necessary for genetic recombination and DNA replication. It denatures poly-(dAdT)·poly(dAdT)] and T4 DNA at temperatures far below their regular melting temperatures. The protein interacts with native DNA and a variety of synthetic polynucleotides. Illustrated here is the interaction of the protein with poly(dA). The CD spectra suggest that the protein keeps poly(dA) in a conformation that is equivalent to that of a single-stranded form in high salt. Similar results were

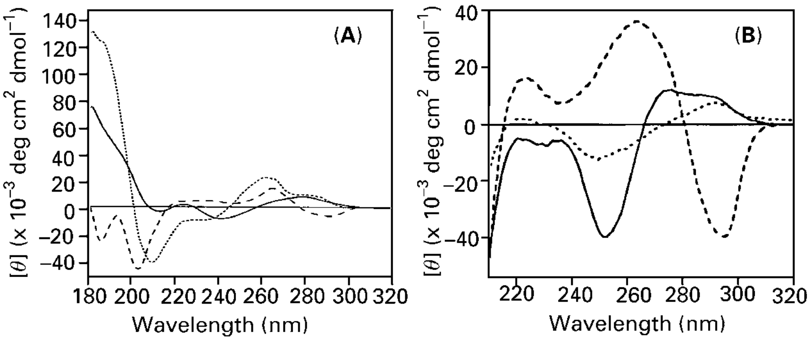



**Figure 10** (A) The CD spectra of poly (dAdC)·poly(dGdT) in the (·····) A, (—) B and (---) Z conformations. Redrawn from data in Riazance-Lawrence JH and Johnson WC Jr (1992) *Biopolymers* **32**: 271–276 by permission of John Wiley and Sons, Inc. (B) CD spectra of poly(dGdC)·poly(dGdC) in the (—) B-form, (- - -) Z-form and (····) in a form in which the bases assume the Hoogsteen base-pairing conformation. Abstracted with permission from data of Seger-Nolten GMJ, Sijtsema NM and Otto C (1997) *Biochemistry* **36**: 13 241–13 247. Copyright 1997 American Chemical Society.

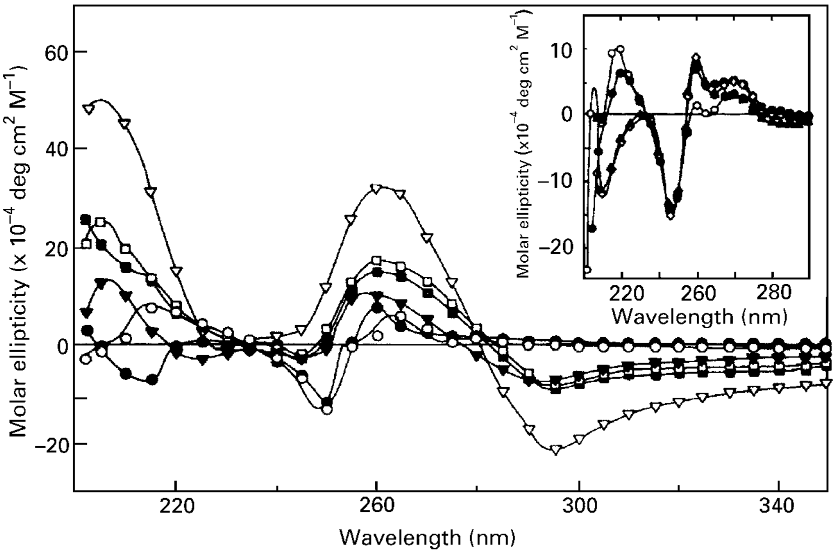



**Figure 11** CD spectra of poly(dT)·poly(dA)·poly(dT) in the presence of (○) 0; (●) 60; (▽) 70; (■) 80; (▼) 90 and (□) 100 μM spermidine. The insert shows the effects of (○) 0, (●) 5, (▼) 10, and (◆) 25 μM spermidine. Data were recorded at 25°C in 10 mM sodium cacodylate and 0.5 mM EDTA at pH 7.2. Redrawn with permission from data in Thomas TJ, Kulkarni GD, Greenfield NJ, Shirahata A and Thomas T. (1996) *Biochem J.* **319**: 591–599. Copyright (1996) Portland Press.

obtained with other dimeric DNA analogues including poly(dA)·poly(dT) and poly(dT).

## Carbohydrates

Carbohydrates are much more difficult to study using ORD and CD than proteins and nucleic acids and much less literature is available. First, carbohydrates mainly absorb below 190 nm, a region difficult to examine using commercial CD spectrometers. Second, unlike proteins and nucleic acids, there are no common sets of spectral characteristics which easily define the conformation of carbohydrates. However, ORD and CD have been used to obtain useful information about the structure of carbohydrates. In favourable circumstances they can give information about the structure and linkages in carbohydrates. In addition, they can distinguish helical from disordered conformations and be used to monitor conformational transitions induced by salts, solvents and temperature.

### Determination of carbohydrate structure and linkages

While no simple basis spectra exist that can be used to determine the conformation of carbohydrates, analysis of the CD spectra of monomeric carbohydrates has provided empirical rules for extracting some conformational information. By comparing the spectra of monomeric sugars, researchers have been able to assign the CD bands to specific transitions, and to determine quadrant rules for the effect of structure on the sign of the transitions. For example, the CD spectra of methyl-$\beta$-galactopyranoside, carrageenan and agarose are shown in **Figure 13**. X-ray films of carrageenan show that the conformation is a double helix. The CD spectrum of carrageenan is consistent with the conformation seen in the X-ray analysis. There are conflicting models for the structure of agarose in gels. In one model the chains are extended and in the other they are helical. The geometry-dependent linkage contributions to the CD of agarose were determined by subtracting the monomeric CD from the CD of the polymer. Application of the quadrant rules indicated that agarose was also helical. The difference in CD sign between carrageenan and agarose arises from a translocation of the beta-D-galactose O-2 atom from one quadrant to the neighbouring quadrant of the C-5-O-5-C-1 ether chromophore of the preceding anhydro sugar residue.

### Analysing conformational transitions of carbohydrates using CD and ORD

While the analysis of the CD spectrum of carbohydrates to obtain detailed structural information is complex, CD and ORD serve as simple methods for following order–disorder transitions in carbohydrates. Because the absorption bands of

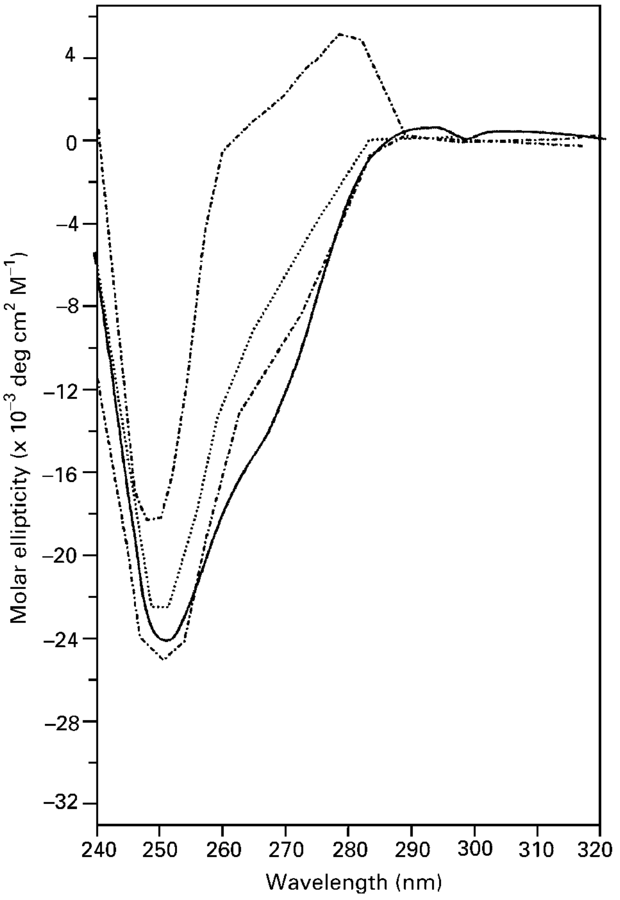



**Figure 12** CD spectra of poly(dA) complexed with the gene 32 coded protein of bacteriophage T4 at (—) 1.0, (----) 29.9 and (·····) 47.7°C. The nucleotide to protein ratio is 8.5. Also illustrated (–·–·) is the spectrum of free poly(dA) at 1°C. Redrawn with permission from Greve J, Maestre MF, Moise H and Hosoda J (1978) *Biochemistry* **17**: 887–893 Copyright (1978) American Chemical Society.

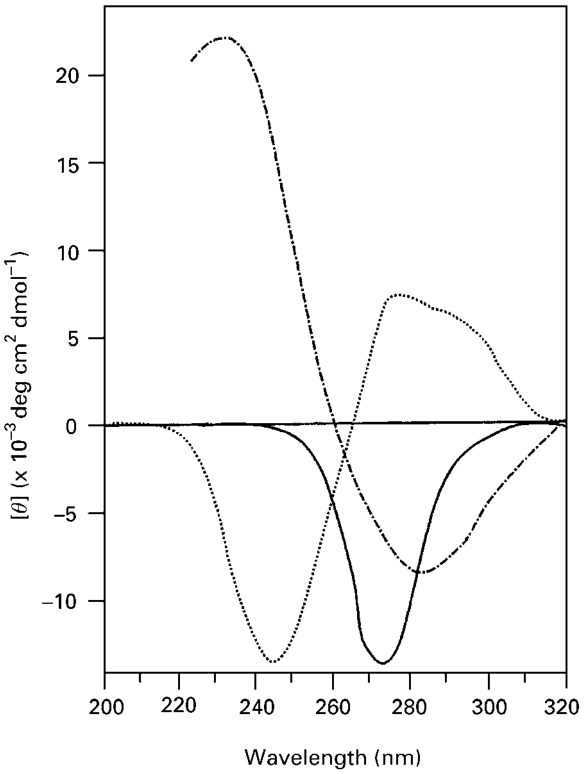



**Figure 13** CD spectrum of (—) methyl β-D-galactopyranoside, (···) carrageenan and (–·–·) agarose. Redrawn with permission from data in Arndt ER and Stevens ES (1993) *J. Am. Chem. Soc.* **115**: 7849–7853, Copyright (1993) American Chemical Society, Arndt ER and Stevens ES (1994) *Biopolymers* **34**: 1527–1534 and Stevens ES and Morris ER (1990) *Carbohydr. Polym.* **12**: 219–224.

carbohydrates are at very low wavelength, many of the studies of conformational transitions have been done using ORD, since ORD bands extend even into the visible region and are more easily accessible. For example, **Figure 14** shows how the effects of potassium-ion concentration and temperature on the order–disorder transition of xanthan, an extracellular polysaccharide produced by *Xanthomonas campestris*, have been followed by measuring the optical rotation at 365 nm. The data were used to find the enthalpy of folding of the polymer. The data could also be fitted by Zimm–Bragg helix-coil transition analysis. Similar studies of conformational transitions have been performed on other carbohydrates including acetan, carrageenan, succinoglycan and hyaluronate.

In addition to monitoring the folding of carbohydrates by examining changes in the intrinsic optical activity of the carbohydrate, structural information has been obtained by examining the CD and ORD spectra of bound dyes. For example when methylene blue binds to carrageenan, the binding induces optical activity in the dye which changes as a function of temperature. There is a progressive loss of ellipticity as a function of temperature, suggesting that the dye binds to the double-helical form of the polysaccharide, but not the coil form.

## List of symbols

$K$ = equillibrium folding constant; $n$ = number of molecules; $N$ = chain length; $P$ = correlation coefficient; $R$ = gas constant; $T$ = temperature; $\Delta C_p$ = change in heat capacity; $\Delta G$ = free energy of folding; $\Delta H$ = Van't Hoff enthalpy of folding; $\Delta S$ = entropy of folding; $\theta$ = ellipticity; $\sigma$ = mean-square error.

*See also:* **Biomacromolecular Applications of UV-Visible Absorption Spectroscopy; Carbohydrates Studied By NMR; Chiroptical Spectroscopy, General**

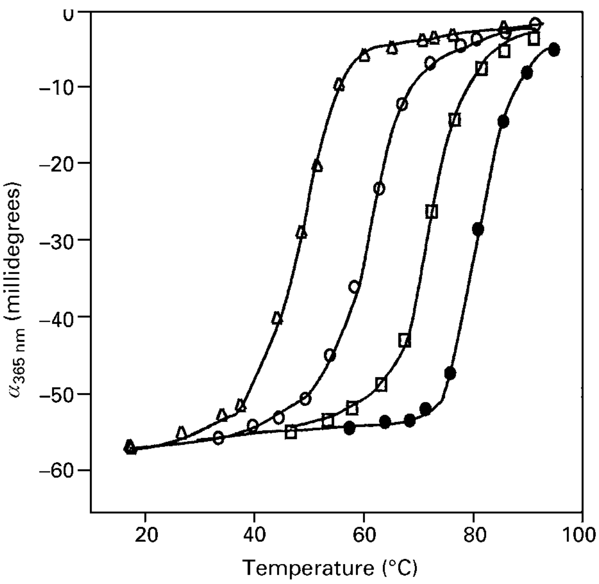



**Figure 14** The unfolding of xanthan as a function of potassium ion concentration. (Δ) 4.3, (○) 15, (□) 30 and (●) 500 mM. Redrawn with permission from data in Norton IT, Goodall DM, Frangou SA, Morris ER and Rees DA (1984) *J. Mol. Biol.* **175**: 371–394 Copyright (1984) Academic Press.

**Theory; Induced Circular Dichroism; Macromolecule–Ligand Interactions Studied By NMR; Magnetic Circular Dichroism, Theory; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Nucleic Acids Studied Using NMR; ORD and Polarimetry Instruments; Proteins Studied Using NMR Spectroscopy; Peptides and Proteins Studied Using Mass Spectrometry; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading


Adler AJ, Greenfield NJ and Fasman GD (1973) Circular dichroism and optical rotatory dispersion of proteins and polypeptides. *Methods in Enzymology* **27 part D**: 675–735.

Breslauer KJ (1987) Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions. *Methods in Enzymology* **259**: 221–245.

Eftink MR (1995) Use of multiple spectroscopic methods to monitor equilibrium unfolding of proteins. *Methods in Enzymology* **259**: 487–512.

Greenfield NJ (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Analytical Biochemistry* **235**: 1–10.

Johnson WC Jr (1996) Determination of the conformation of nucleic acids by electronic CD. In: Fasman GD (ed) *Circular Dichroism and the Conformational Analysis of Biomolecules*, pp 433–468. New York and London: Plenum Press.

Perrin JH and Hart PA (1970) Small molecule-macromolecule interactions as studied by optical rotatory dispersion-circular dichroism. *Journal of Pharmaceutical Science* **59**: 431–448.

Stevens ES (1996) Carbohydrates. In: Fasman GD (ed) *Circular Dichroism and the Conformational Analysis of Biomolecules*, pp 501–530. New York and London: Plenum Press.

Woody RW (1995) Circular dichroism. *Methods in Enzymology* **246**: 34–71.


# Biomacromolecular Applications of UV-Visible Absorption Spectroscopy


**Alison Rodger** and **Karen Sanders**, University of Warwick, Coventry, UK



Copyright © 1999 Academic Press


| ELECTRONIC SPECTROSCOPY |
| --- |
| Applications |

## Introduction

Ultraviolet (UV) or visible radiation is absorbed by a molecule when the frequency of the light is at the correct energy to cause the electrons of the molecule to rearrange (or become excited) to another, higher-energy, state of the system. Frequency, $\nu$ (measured in $\sigma^{-1}$), wavelength, $\lambda$ (usually measured in nm) and

energy, $E$ (measured in J) are related by

$$E = h\nu = hc/\lambda$$

where $h = 6.626\,196 \times 10^{-34}$ J s is the Planck constant and $c = 2.997\,925 \times 10^{17}$ nm s$^{-1}$ is the speed of light in units to match the choice of nm for $\lambda$.

Absorption can be pictorially viewed as either the electric field or the magnetic field (or both) of the radiation pushing the electron density from a starting arrangement to a higher-energy final one. The direction of net linear displacement of charge is known as the polarization of the transition. The polarization and intensity of a transition are characterized by the so-called transition moment, which is a vectorial property having a well-defined direction (the transition polarization) within each molecule and a well-defined length (which is proportional to the square root of the absorbance). The transition moment may be regarded as an antenna by which the molecule absorbs light. Each transition thus has its own antenna and the maximum probability of absorbing light is obtained when the antenna and the electric field of the light are parallel.

Absorbance is defined in terms of the intensity of incident, $I_0$, and transmitted, $I$, light:

$$A = \log_{10}\left|\frac{I_0}{I}\right|$$

In most experiments we use the Beer–Lambert law to relate the absorption of light, $A$, to the sample concentration, C:

$$A = \varepsilon Cl$$

where $l$ is the length of the sample through which the light passes and $\varepsilon$ is the molar absorption coefficient (extinction coefficient); if $l$ is measured in cm and $C$ in M = mol dm$^{-3}$, then $\varepsilon$ has units of mol$^{-1}$ dm$^3$ cm$^{-1}$. The Beer–Lambert law breaks down when the sample absorbs too high a percentage of the incident photons for the instrument to measure the emitted photons. An absorbance of 2, for example, means that 99% of the photons are absorbed. Biological samples present additional challenges to the Beer–Lambert law: if there are local high concentrations of sample (as in vesicles for example) or if molecules interact and perturb the spectroscopy of the isolated molecule, then the Beer–Lambert law becomes invalid.

The most common application of UV-visible absorption spectroscopy is to determine the concentration of a species in solution using the Beer–Lambert law. Other applications follow because the energy of UV-visible light is usually sufficient only to excite valence electrons which are the ones involved in bonding. Thus any UV-visible absorption spectrum is directly related to bonds and hence the structure of a molecule. The challenge is then to relate the plot of absorbance versus wavelength, which the spectrometer produces, to the structure of the molecules in the cuvette. With complicated systems, such as biological macromolecules and their complexes with small molecules, we usually interpret UV-visible spectra by considering changes in the spectrum as a function of a variable such as temperature, ionic strength, solvent, concentration, etc. Alternatively the absorption spectra data are used as input for interpreting other spectra such as fluorescence, circular dichroism (CD) or linear dichroism (LD).

## Biological macromolecule structure and UV spectroscopy

Proteins and DNAs are linear polymers where a limited set of residues are joined together by, respectively, the amide or phosphodiester bonds. The situation is similar for carbohydrates though the linking options are more varied.

To a first approximation the absorbance spectrum expected for a biomolecule is therefore the sum of the spectra for the component parts.

In the case of nucleic acids (DNA and RNA) the UV absorbance from 200–300 nm is due exclusively to transitions of the planar purine and pyrimidine bases (**Figure 1**). (The backbone begins to contribute at about 190 nm.) The accessible region of the spectrum (nitrogen purging is required below ~200 nm as oxygen absorption interferes with the spectrum) is therefore dominated by $\pi \rightarrow \pi^*$ transitions of the bases. The UV spectra of the bases (**Figure 2**) look as if there are two simple bands; however, each 'simple' band observed is a composite of more than one transition. This makes detailed analysis of DNA absorption difficult, but usually ensures that the absorption spectrum changes when the system is perturbed. Thus absorption spectroscopy is a useful qualitative or empirical probe of structural changes. Typical DNA UV absorption spectra are illustrated in **Figure 3**. The base transitions are significantly perturbed by the so-called $\pi$–$\pi$ stacking interactions and so both wavelength maxima and transition intensities vary depending on the base sequence and structure adopted (cf. **Table 1**).

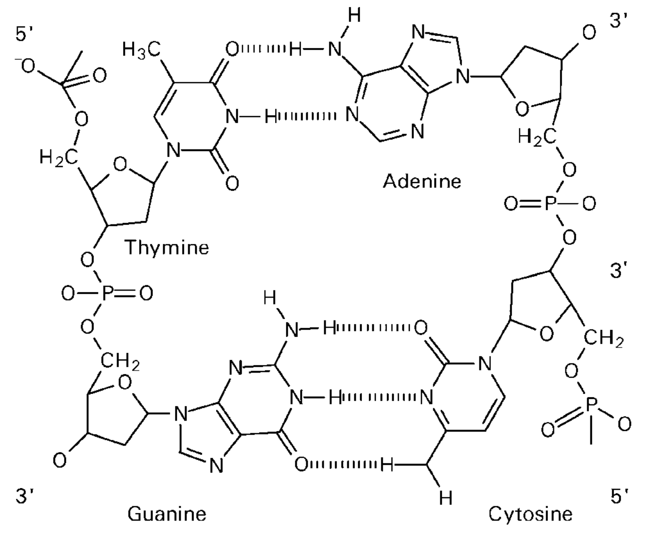



**Figure 1** Structural formula of DNA.

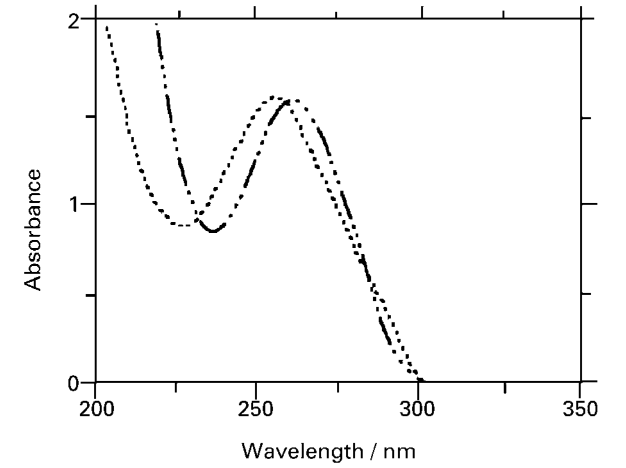



**Figure 3** Absorbance spectra in 1 cm pathlength cuvettes of 190 μM poly[d(G-C)]$_2$ (dotted line) and 240 μM poly[d(A-T)$_2$] (dashed line).

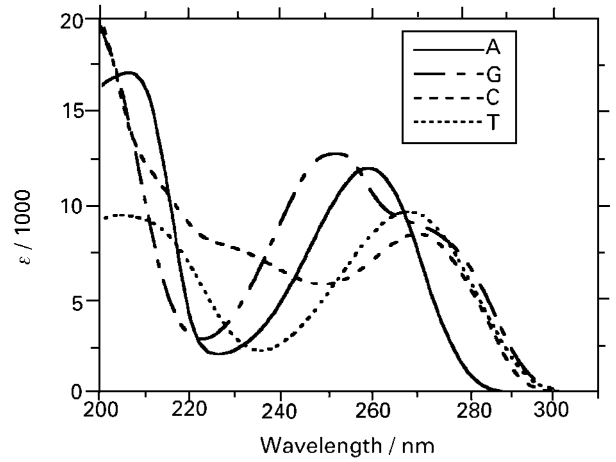



**Figure 2** UV spectra of the DNA nucleotides: deoxyadenosine 5′-monophosphate (A), deoxyguanosine 5′-monophosphate (G), deoxycytidine 5′-monophosphate (C) and thymidine 5′-monophosphate (T). The spectrum of uracil is almost indistinguishable from that of thymine.

**Table 1** Long-wavelength absorbance maxima and extinction coefficients for some DNAs. Calf thymus DNA is ~60% A-T

| *DNA* | *Wavelength of $A_{max}$(nm)* | $\varepsilon_{max}$ (mol$^{-1}$ dm$^3$ cm$^{-1}$) |
|---|---|---|
| Calf thymus | 260 | 6 600 |
| Poly[d(A-T)]$_2$ | 262 | 6 600 |
| Poly(dA).poly(dT) | 260 | 6 000 |
| Poly[d(G-C)]$_2$ | 254 | 8 400 |
| Poly(dG).poly(dC) | 253 | 7 400 |

(A = adenine, T = thymine, G = guanine and C = cytosine)

In the case of peptides and proteins the spectroscopy of the amide bonds, the side chains and any prosthetic groups (such as haems) determines the observed UV-visible absorption spectrum. However, as with DNA, intensities and wavelengths can be perturbed by the local environment of the groups. UV spectra of proteins are usually divided into the ‘near’ and ‘far’ UV regions. The near-UV in this context means 250–300 nm and is also described as the aromatic region, though transitions of disulfide bonds (cystines) also contribute to the total absorption intensity in this region. The far-UV (< 250 nm) is dominated by transitions of the peptide backbone of the protein, but transitions from some side chains also contribute to the spectrum below 250 nm.

The aromatic side chains, phenylalanine, tyrosine and tryptophan all have transitions in the near-UV region (**Figure** 4). At neutral pH, the indole of tryptophan has two or more transitions in the 240–290 nm region with total maximum extinction coefficient $\varepsilon_{max}$(279 nm) ~5000 mol$^{-1}$ dm$^3$ cm$^{-1}$; tyrosine has one transition with $\varepsilon_{max}$(274 nm) ~1400 mol$^{-1}$ dm$^3$ cm$^{-1}$; phenylalanine also has one transition with $\varepsilon_{max}$ ~190 mol$^{-1}$ dm$^3$ cm$^{-1}$; and a cystine disulfide bond absorbs from 250–270 nm with $\varepsilon_{max}$ ~300 mol$^{-1}$ dm$^3$ cm$^{-1}$. Although tryptophans have by far the most intense transitions, many proteins have few tryptophans compared with the other aromatic groups, so the near UV is not necessarily dominated by tryptophan transitions.

The peptide chromophore (**Figure** 5) which gives rise to the transitions observed in the far-UV region (180–240 nm) has non-bonding electrons on the oxygen and also on the nitrogen atoms, π-electrons which are delocalized to some extent over the carbon, oxygen and nitrogen atoms, and σ bonding electrons. The lowest energy transition of the peptide

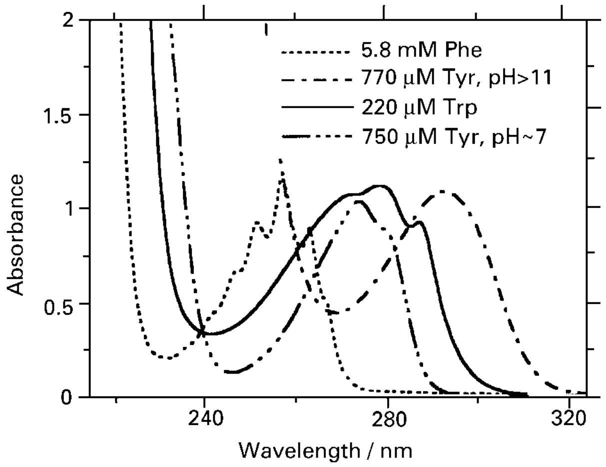



**Figure 4** Aromatic absorption spectra of tryptophan, tyrosine and phenylalanine. Note the different concentrations required to ensure absorbance maxima of 1 absorbance unit.

chromophore is an $n \rightarrow \pi^*$ transition analogous to that in ketones, and the next transition is $\pi \rightarrow \pi^*$. As in the carbonyl case, the $n \rightarrow \pi^*$ transition is predominantly of magnetic transition dipole character and is thus of low intensity ($\varepsilon$ ~100 $mol^{-1}$ $dm^3$ $cm^{-1}$), though it is not as low as for a simple ketone; it occurs at about 210–230 nm (depending mainly upon the extent of hydrogen bonding of the oxygen lone pairs) and its small electric character is polarized more or less along the carbonyl bond. The $\pi \rightarrow \pi^*$ transition ($\varepsilon$ ~7000 $mol^{-1}$ $dm^3$ $cm^{-1}$) is dominated by the carbonyl π-bond and is also affected by the involvement of the nitrogen in the π orbitals; its electric dipole transition moment is polarized somewhere near the line between the oxygen and nitrogen atoms, and it is centred at 190 nm.

In an α-helix, the coupling of the $\pi \rightarrow \pi^*$ transition moments in each amide chromophore results in a component at about 208 nm which contributes to the characteristic α-helix CD spectrum. For UV-visible spectroscopy, however, the far-UV spectroscopy is usually of little use either for concentration determination or structural analysis as the accessible region (above 200 nm) is almost a linear plot of increasing intensity with decreasing wavelength.

Carbohydrate UV-visible spectroscopy is essentially that of any substituents that have spectroscopy; a simple sugar system such as starch has no spectroscopy above 200 nm. Carbohydrates are typically derivatized with thiols or aromatic chromophores for UV spectroscopy. The spectroscopy of these compounds is largely determined by that of the derivatives. These data may be found in UV-visible spectroscopy atlases.

## Wavelength scanning

### Simple wavelength scans of biological macromolecules

Nucleic acids (DNA and RNA), proteins and peptides absorb very little light above 300 nm in the absence of ligands or prosthetic groups with chromophores (absorbing units). However, it is usually wise to collect a simple UV scan of a sample from about 350 nm. If the spectrum is not flat between 350 and 310 nm then the sample has condensed into particles whose size is of the order of the wavelength of light; therefore what is being observed is scattering of the incident light rather than absorption. UV absorbance is most commonly used to determine the concentration of a sample and also to give an indication of its purity.

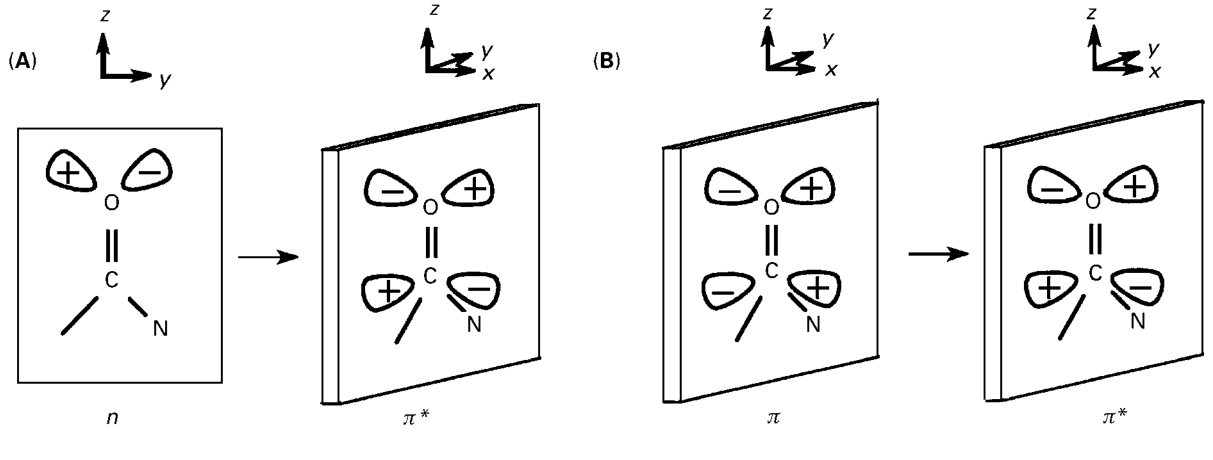



**Figure 5** Schematic illustrations of (A) $n \rightarrow \pi^*$ and (B) $\pi \rightarrow \pi^*$ transitions of peptides.

To perform a wavelength scan proceeds as follows:

1. First choose the parameters.

   The wavelength range will usually be 350–200 nm, unless the buffer cuts out the low-wavelength end (see below).

   A data interval of 0.5–1 nm is adequate as the bands are all broad. Shorter data intervals will not improve the spectrum you plot but will fill up the computer memory. (If data is being collected to aid interpretation of LD spectra, ensure that the data interval is the same in both cases.)

   The bandwidth is the wavelength range of the radiation at any specified wavelength. If data is being collected every 0.5 nm, then a bandwidth of 0.5 nm is appropriate. A larger value will lead to a greater averaging of data than the data interval suggests. Do not choose a much smaller bandwidth as the incident light intensity will be reduced and the spectral noise will increase without improving the quality of your data.

   The data averaging time determines the signal-to-noise ratio and affects the scan speed required to produce undistorted spectra. 0.033–0.1 s is usually a good choice; longer times are required if the absorbance is very small (less than 0.1 absorbance units) or small differences in absorbance are being determined (as for DNA melting curves; see below).
2. Choose the solvent or buffer for your experiment and fill a pair of matched cuvettes of the desired pathlength with it. Either switch off the baseline correction or use an air/air baseline in the instrument. Place one cuvette in the sample holder (usually the front position) and leave the reference holder empty. Run a spectrum over the desired wavelength range to check that the solvent/buffer spectrum is not significant over the wavelength range of interest. If it is then you need to change the solvent/buffer.

   In choosing your buffer note phosphate buffers are essentially spectroscopically invisible over the wavelength range usually used; however, you need to ensure that phosphate does not interact in any way with your sample. Cacodylate and ammonium acetate have a window down to ~210 nm and have the added advantage of preventing bacterial growth (whereas phosphate promotes it). Chloride ions begin absorbing just above 200 nm so high salt spectra cannot be collected at the lower-wavelength end of the spectrum. Higher buffer concentrations can be accommodated in shorter pathlength cuvettes (see Beer–Lambert law discussion above), so if it is not possible to dilute or change the buffer try a smaller pathlength cuvette. (5 mm cuvettes will stand in a normal sample holder, 1 mm cuvettes will need spacers to hold them vertical – if they are not held vertical then you will be working with a variable pathlength. Smaller pathlength cuvettes will need special holders.) UV cut-off wavelengths for solvents are readily available in liquid chromatography texts.
3. Place the matched solvent/buffer cuvettes in both the reference holder (usually the rear position) and the sample holder and perform a baseline accumulation. An alternative to having the solvent/buffer in the reference beam is to use a reference cell of water and collect a spectrum of the solvent/buffer which is subsequently subtracted from all spectra. This method may be preferable if the solvent/buffer has a significant absorbance as it is easier to determine whether an apparent peak or dip in the spectrum is due to the buffer in some way.
4. Place the sample in the sample cuvette in the sample holder and record the spectrum.

The absorbance at ~260 nm (or wherever the maximum is for a particular molecule) is generally used to determine nucleic acid concentrations using the Beer–Lambert law. As noted above, for DNA samples the linear relationship between concentration and absorbance seems to break down when the absorbance of a 1 cm pathlength solution exceeds ~1.5–2 absorbance units. Some DNA extinction coefficients are given in **Table 1.**

Protein absorbances will be dominated by tryptophan residues (if there are any) and will have a maximum at 280 nm. The other aromatic residues also absorb at 280 nm. Absorbance at 280 nm may therefore be used to give an estimate of protein concentrations. At 280 nm a 1 mg $cm^{-3}$ protein solution in a 1 cm pathlength cell often has an absorbance of ~1 absorbance unit. This is because many proteins have a similar percentage of 'aromatic' amino acid residues. However, the $A_{280}$ (1 mg $cm^{-3}$) can vary from 0.3 to 1.8. For example, the $A_{280}$ (1 mg $cm^{-3}$) for bovine serum albumin is ~0.66. In cases where the protein amino acid content and molecular weight are known then a reasonably accurate estimate of $\varepsilon$ can be made using the above $\varepsilon$ values for the residues (instead of assuming $A_{280} = 1$ for 1 mg $cm^{-3}$) and then the Beer–Lambert law applied.

The Beer–Lambert method for concentration determination of nucleic acids and proteins is based on the assumption that the samples are pure. Nucleic

acids, if present, will interfere with the protein concentration determination because they also absorb at 280 nm, and conversely. In these cases the following formula permits a rough estimate of protein concentration in the presence of nucleic acids (for a 1 cm pathlength cuvette):

$$1.55 \times A_{280} - 0.76 \times A_{260} = \text{mg protein cm}^{-3}$$

It is also common practice to report the $A_{280}/A_{260}$ ratio as an indication of purity for a given sample.

### Wavelength scans of derivatized protein samples for concentration determination

Although proteins have different percentages of the different amino acids, each residue has an amide bond linking it to the next residue in the chain. A number of concentration-determination methods have thus been developed that involve derivatizing the amides and spectroscopically determining the concentration of the derivatives. The three methods mentioned below all rely on a standard of known concentration to enable a calibration curve to be plotted. The calibration is not necessarily linear.

**Biuret method** This method is simple and reasonably specific as it depends on the reaction of copper(II) with four N atoms in the peptide bonds of proteins. Compounds containing peptide bonds give a characteristic purple colour when treated in alkaline solution with copper sulfate. This is termed the 'biuret' reaction because it is also given by the substance biuret $NH_2CONHCONH_2$. For a wide variety of proteins, 1.0 mg of protein in 2 $cm^3$ of solution results in an absorbance at 540 nm of 0.1. Many haemoproteins, for example, give spurious results due to their intrinsic absorption at 540 nm, but modifications which overcome this difficulty are known (either removal of the haem before protein estimation or destruction of the haem by hydrogen peroxide treatment). The protein content of cell fractions such as nuclei and microsomes can be estimated by this method after solubilization by detergents such as deoxycholate or sodium dodecyl sulfate.

The biuret reagent may be made by placing $CuSO_4.5H_2O$ (1.5 g) and sodium potassium tartrate.$4H_2O$ (6.0 g) into a dry 1 $dm^3$ volumetric flask and adding about 500 $cm^3$ of water. With constant swirling, NaOH solution (300 $cm^3$, 10% w/v) is added and the solution made to volume (1 $dm^3$) with water. The reagent prepared in this manner is a deep blue colour. It may be stored indefinitely if KI (1 g) is also added and the reagent is kept in a plastic container. The protein solution ($x$ $cm^3$, where $x < 1.5$) is then mixed with water [$(1.5 - x)$ $cm^3$] to make a total volume of 1.5 $cm^3$ to which 1.5 $cm^3$ of biuret reagent is added. The purple colour is developed by incubating for 20 min at 37°C. The tubes must then be cooled rapidly to room temperature and the absorbance at 540 nm determined. The colour of the solution is stable for hours.

**Folin–Ciocalteau or Lowry method** While the biuret method is sensitive in the range 0.5 to 2.5 mg protein per assay, the Lowry method is 1 to 2 orders of magnitude more sensitive (5 to 150 μg). The main disadvantage of the Lowry method is the number of interfering substances; these include ammonium sulfate, thiol reagents, sucrose, EDTA, Tris, and Triton X-100.

The final colour in the Lowry method is a result of two reactions. The first is a small contribution from the biuret reaction of protein with copper ions in alkali solution. The second results from peptide-bound copper ions facilitating the reduction of the phosphomolybdic-tungstic acid (the Folin reagent) which gives rise to a number of reduced species with a characteristic blue colour. The amino acid residues which are involved in the reaction are tryptophan and tyrosine as well as cysteine, cystine and histidine. The amount of colour produced varies slightly with different proteins. In this respect it is a less-reliable assay than the biuret method, but it is more reliable than the absorbance method since $A_{280}$ may include contribution from other species, and also the absorption of a given residue is dependent on its environment within the protein.

Two solutions are required for the Lowry method. For the alkaline copper solution, mix 50 $cm^3$ $Na_2CO_3$ (2% w/v) in NaOH (0.1 M) with 1 $cm^3$ of $CuSO_4.5H_2O$ (0.5% w/v) and 1 $cm^3$ of sodium potassium tartrate (1% w/v). This solution must be discarded after 1 day. The Folin reagent (phosphomolybdic-tungstic acid) may be made by diluting the concentrated Folin reagent obtained from e.g. Sigma with an equal volume of water so that it is 1 N (i.e. 1 M $H^+$).

To perform an assay add $x$ $cm^3$ of sample (where $x < 0.6$) containing 5–100 μg of protein as required to $(0.6 - x)$ $cm^3$ of water. Then add 3 $cm^3$ of the alkaline copper solution. The solutions must then be mixed well and allowed to stand for 10 min at room temperature. 3.0 $cm^3$ of Folin reagent is then added and after 30 min the absorbance at 600 nm is determined.

**Coomassie blue dye binding assay** This protein-determination method involves the binding of

Coomassie brilliant blue G-250 to protein. The protonated form of Coomassie blue is a pale orange-red colour whereas the unprotonated form (**Figure 6**) is blue. When proteins bind to Coomassie blue in acid solution their positive charges suppress the protonation and a blue colour results. The binding of the dye to a protein causes a shift in the absorption maximum of the dye from 465 to 595 nm and it is the increase in absorbance at 595 nm that is monitored. The assay is very reproducible and rapid with the dye binding process virtually complete in ~2 min with good colour stability.

The reagent is prepared as follows. Coomassie brilliant blue G-250 (100 mg) is dissolved in 50 $cm^3$ 95% ethanol. To this solution phosphoric acid (100 $cm^3$, 85% w/v) is added and the solution diluted to 1 $dm^3$. To perform the assay, $x$ $cm^3$ of the sample containing 5–100 μg of protein is placed in a clean, dry test tube. $(0.5 - x)$ $cm^3$ water and 5.0 $cm^3$ of diluted dye reagent are added and the solution mixed well. After a period of from 5–60 min, $A_{595}$ is determined.

The only compounds found to give excess interfering colour in the assay are relatively large amounts of detergents such as sodium dodecyl sulfate, Triton X-100 and commercial glassware detergents. Interference by small amounts of detergent may be eliminated by the use of proper controls. The assay is non-linear and requires a standard curve.

### Simple wavelength scans of macromolecules with bound ligands

When a ligand is added to, for example, a DNA solution, if it does not bind to the DNA then the UV-visible spectrum will simply be the sum of the DNA spectrum and the ligand spectrum. If the ligand binds to the DNA then the spectrum of the complex will be different from the sum spectrum. One should note that the observed spectrum is probably a complicated mixture of that due to bound and unbound ligand and free and complexed DNA.

When a planar aromatic molecule binds intercalatively (sandwiched between two base pairs) to DNA there is usually a characteristic decrease in the ligand-absorbance signal (this can be up to 50%) and a shift to the red (bathochromic shift) of between ~2 nm and 20 nm as illustrated in **Figure 7**. The DNA spectrum is also affected by any molecule such as an intercalator that causes a structural change. This makes such spectra a useful probe of DNA/drug interactions but renders absorbance useless for concentration determinations unless the perturbed extinction coefficients are known.

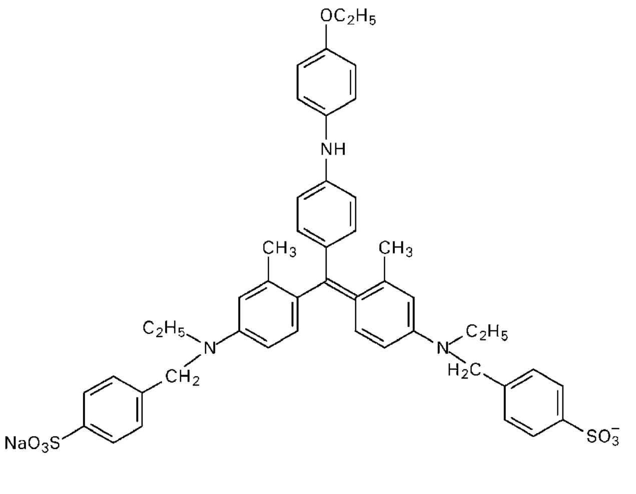



**Figure 6** Coomassie blue.

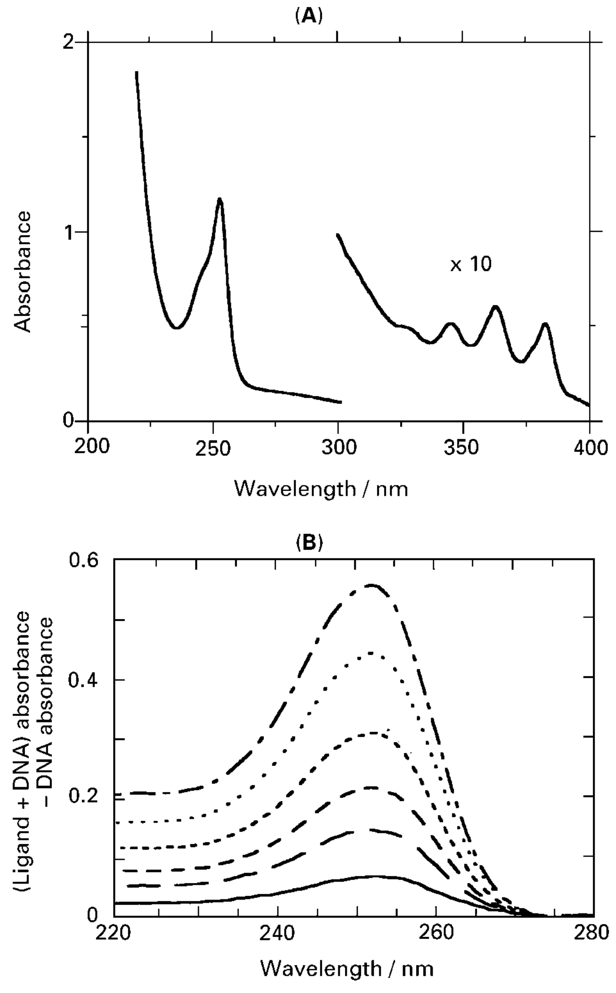



**Figure 7** (A) 5 μM anthracene-9-carbonyl-$N^1$-spermine in water. (B) 2 μM, 4 μM, 5 μM, 7 μM, 10 μM and 13 μM anthracene-9-carbonyl-$N^1$-spermine in water with 200 μM calf thymus DNA. Note broadening and magnitude decrease of 250 nm band absorbance; this molecule intercalates between DNA base pairs.

#### Titrations

The label 'titration' is used to cover experiments where spectra are collected as a function of concentration, ionic strength, pH, etc. To minimize macromolecule consumption and also (perhaps surprisingly) to minimize concentration errors, the best method is often to add solution to the cuvette. A simple way to avoid dilution effects is to proceed as follows. Consider a starting sample that has concentration $x$ M of species $X$. Each time $y$ cm$^3$ of $Y$ is added, also add $y$ cm$^3$ of a $2x$ M solution of $X$. The concentration of $X$ remains constant at $x$ M. Many variants on this theme may be derived.

If a titration series where ligand concentration is held constant while the macromolecule varies has a constant absorbance at a wavelength, $\lambda_i$, where the ligand absorbs this is called an isosbestic point. Isosbestic points only occur if two species are present in solution and those species have the same absorbance at $\lambda_i$.

#### Binding constants

If the spectral shape of a ligand spectrum remains unchanged during a titration experiment, but the magnitude changes in a manner that is proportional to the concentration of bound ligand then the spectral data can be used to determine the equilibrium binding constant, $K$. This is because in such a case the ligands are binding in one binding mode or in constant proportions in more than one mode (meaning site, orientation, sequence, etc). The data must be of very high quality for absorbance (or any other spectroscopic data) to be used to determine $K$. A simple plot of change in absorbance versus either macromolecule or ligand concentration (whichever is being varied) will probably enable the quality of the data set to be determined. It should be a smooth curve.

Consider the equilibrium

$$L_{\mathrm{f}} + S_{\mathrm{f}} \overset{K}{\rightleftharpoons} L_{\mathrm{b}}$$

where $L_{\mathrm{f}}$ is a free ligand, $L_{\mathrm{b}}$ is a bound ligand and $S_{\mathrm{f}}$ is a free site. In the simple case where the macromolecule can be treated as a series of binding sites of $n$ residues in size, then the total site concentration $S_{\mathrm{tot}} = [M]/n$, where $[M]$ is the residue concentration of the macromolecule. (For proteins it is sometimes preferable to think in terms of the concentration of molecules rather than residues. In this case $S_{\mathrm{tot}} = n'[M]$ where $n'$ is the number of binding sites per protein.) Then

$$K = \frac{nc_{\mathrm{b}}}{c_{\mathrm{f}}[M]}$$

where $c_{\mathrm{b}}$ is the concentration of the bound ligand and $c_{\mathrm{f}}$ is the concentration of the free ligand.

There are a number of methods for determining $K$ using absorbance data. The simplest is the enhancement method. This method is commonly used for fluorescence spectroscopy and may also be used to interpret absorbance data. We write

$$\begin{aligned} c_{\mathrm{tot}}A &= c_{\mathrm{f}}A_{\mathrm{f}} + c_{\mathrm{b}}A_{\mathrm{b}} \\ c_{\mathrm{tot}}A &= (c_{\mathrm{tot}} - c_{\mathrm{b}})A_{\mathrm{f}} + c_{\mathrm{b}}A_{\mathrm{b}} \\ c_{\mathrm{b}} &= \frac{c_{\mathrm{tot}}(A - A_{\mathrm{f}})}{A_{\mathrm{b}} - A_{\mathrm{f}}} \end{aligned}$$

Application of this equation requires knowledge of the absorbance of free and bound ligand. Determining the latter requires measuring an absorbance spectrum under conditions where it is known that all the ligand is bound to the macromolecules. $K$ may then be determined directly. A more accurate value of $K$ will be achieved if the data is used to perform a Scatchard plot.

The Scatchard plot is based on rewriting the equation for the equilibrium constant as:

$$\begin{aligned} \frac{r}{c_{\mathrm{f}}} &= \frac{KS_{\mathrm{f}}}{[M]} \\ &= \frac{K}{n} - rK \end{aligned}$$

where

$$r = \frac{c_{\mathrm{b}}}{[M]}$$

So, a plot of $r/c_{\mathrm{f}}$ versus $r$ has slope $-K$ and $y$-intercept $K/n$. The $x$-intercept occurs where $r = n$.

Other methods more commonly used with CD or LD data may be used with normal absorption data if the change in absorbance (the absorbance of the DNA/ligand system minus the absorbance of a free-ligand solution of the same ligand concentration) is used in the analysis.

### Macromolecule condensation

A UV spectrum may be used to follow the condensation of a macromolecule sample into particles, though, as discussed above, what is being probed is really scattering of the light rather than its absorbance. A monotonic increase in absorbance is observed above 300 nm as condensation takes place. In the case of DNA, the addition of a highly charged DNA binding ligand (such as spermine or $[Co(NH_3)_6]^{3+}$) will effect this change. Concomitantly with the increase in absorbance signal above 300 nm, a decrease in the 260 nm DNA absorbance is observed.

## DNA melting curves: absorption as a function of temperature

If there were no residue–residue interactions in a biomolecule then the UV spectrum would be independent of its geometry and the absorption spectrum would simply be the sum of the contributions from the residues are discussed above. This is particularly true for DNA, where π–π stacking interactions lower the magnitude of the absorbance at 260 nm (the hypochromic effect) and change it at most other wavelengths. The extent of this change depends on the DNA structure. In principle if the DNA is heated enough to disrupt all base structure then the spectrum would become the sum of the base spectra in appropriate proportions. A high enough temperature to achieve this cannot usually be reached in water; however, we can use the change in the UV absorbance signal at a chosen wavelength (usually 260 nm, though at ~280 nm A-T base pairs show very little change in absorbance so this wavelength may be used to probe the role of G-C relative to A-T base pairs) to follow the disruption of base stacking and hence also base-pair hydrogen bonding.

The data from such an experiment is usually illustrated as a melting curve or a derivative melting curve and summarized by the so-called melting temperature, $T_m$ (e.g. **Figure 8**). $T_m$ is the temperature where the absorbance is the average of the duplex and single-stranded DNA absorbances where 50% of the DNA has melted. Thermodynamic data relating to the stability of the duplex may also be extracted from melting curves.

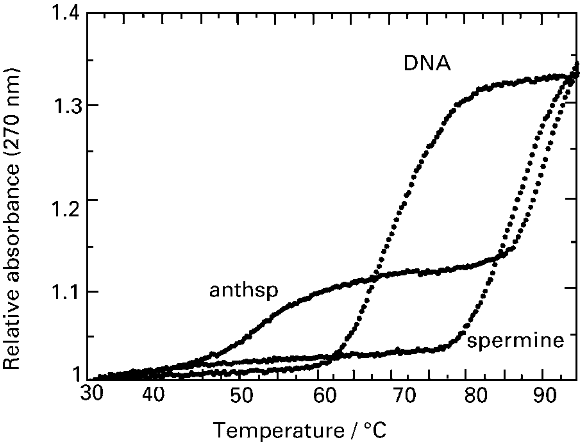



**Figure 8** UV melting curves for 200 μM calf thymus DNA in 10 mM salt (denoted DNA), and also with 20 μM spermine and 20 μM anthracene-9-carbonyl-*N*[1]-spermine (denoted anthsp). Note the premelting transition with anthracene-9-carbonyl-*N*[1]-spermine.

## List of symbols

$A$ = absorbance; $c_{b,f}$ = ligand concentration; $I$ = transmitted light intensity; $I_o$ = incident light intensity; $l$ = sample light path length; $\varepsilon$ = molar absorption coefficient; $\lambda$ = wavelength of radiation (usually in nm); $\nu$ = frequency of radiation ($s^{-1}$).

*See also:* **Biomacromolecular Applications of Circular Dichroism and ORD; Dyes and Indicators, Uses of UV-Visible Absorption Spectroscopy; Macromolecule–Ligand Interactions Studied By NMR; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Nucleic Acids Studied Using NMR; Peptides and Proteins Studied Using Mass Spectrometry; Proteins Studied Using NMR Spectroscopy.**

## Further reading


Atkins PW (1983) *Molecular Quantum Mechanics*. Oxford: Oxford University Press.

Atkins PW (1991) *Physical Chemistry*, 4th edn. Oxford: Oxford University Press.

Bradford MM (1976) *Analytical Biochemistry* **72**: 248.

Craig DP and Thirunamachandran T (1984) *Molecular Quantum Electrodynamics: An Introduction to Radiation-Molecule Interaction*. London: Academic Press.

Eriksson S, Kim SK, Kubista M and Nordén B (1993) *Biochemistry* **32**: 2987.

Gornall AG, Bardawils CJ and David MM (1949) Determination of serum proteins by means of the biuret reagent. *Journal of Biological Chemistry* **177**: 751–766.

Hiort C, Nordén B and Rodger A (1990) Enantioselective DNA binding of $[Ru(1,10\text{-phenanthroline})_3]^{2+}$ studied with linear dichroism. *Journal of the American Chemical Society* **112**: 1971.

Hollas JM (1992) *Modern Spectroscopy*, 2nd edn. Chichester: John Wiley and Sons.

Legler G *et al* (1985) *Analytical Biochemistry* **150**: 278.

Marky LA and Breslauer KJ (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**: 1601–1620.

Michl J and Thulstrup EW (1986) *Spectroscopy with Polarized Light*. New York: VCH.

Newbury SF, McClellan JA and Rodger A (1996) Spectroscopic and thermodynamic studies of conformational changes in long, natural mRNA molecules. *Analytical Communications* **33**: 117–122.

Puglisi JD and Tinoco IJ (1989) Absorbance melting curves of RNA. *Methods in Enzymology* **180**: 304–325.

Read SM and Northcliffe DH (1981) *Analytical Biochemistry* **96**: 53.

Rodger A (1993) Linear dichroism. *Methods in Enzymology* **226**: 232–258.

Rodger A and Nordén B (1997) *Circular Dichroism and Linear Dichroism*. Oxford: Oxford University Press.

Rodger A, Blagbrough IS, Adlam G and Carpenter ML (1994) DNA binding of a spermine derivative: spectroscopic studies of anthracene-9-carbonyl-N[1]-spermine with poly(dG-dC)$_2$ and poly(dA-dT)$_2$. *Biopolymers* **34**: 1583–1593.

Rodger A, Taylor S, Adlam G, Blagbrough IS and Haworth IS (1995) Multiple DNA binding modes of anthracene-9-carbonyl-N[1]-spermine. *Bioorganic and Medicinal Chemistry* **3**: 861–872.

*UV-VIS Atlas of Organic Compounds* (1992) 2nd edn. Weinheim: VCH.

# Biomedical Applications of Atomic Spectroscopy

**Andrew Taylor**, Royal Surrrey County Hospital and University of Surrey, Guildford, UK

Copyright © 1999 Academic Press

| ATOMIC SPECTROSCOPY |
| --- |
| Applications |

The techniques of flame, electrothermal and vapour generation atomic absorption, flame and inductively coupled plasma atomic emission and inductively coupled plasma mass spectrometry for the measurement of minerals and trace elements in biological specimens are described. Situations in which each of the techniques might be employed for the analysis of biomedical samples are reviewed. Interferences associated with the types of samples typically examined are mentioned with accounts of how these are removed in regular practice, either during the sample preparation or by features of the instrumentation. The advantages and disadvantages of these techniques are given with particular reference to sensitivity, single to multielement measurements and special applications such as the determination of stable isotopes. While total analyte concentrations are typically determined, examples of how speciation may be of interest are included. It is seen that for biomedical measurements, atomic spectroscopic techniques are complementary and that each may be appropriate for particular sample types or applications. Except for a few very special purposes these are the techniques of choice for measurement of minerals and trace elements in biomedical specimens.

Quantitative analytical techniques included under the general heading of atomic spectroscopy are almost always employed for the direct determination of inorganic elements. For a few biomedical applications the specimen preparation gives indirect measurements of molecular compounds, generally by using a metal-based reagent in a separation or extraction step and measurement of the metal as a surrogate for the analyte of interest. The few examples of indirect measurements are included in the extensive reviews of atomic spectrometry published annually as the *Atomic Spectrometry Updates* (ASU).

As described in the articles on 'Theory' and 'Methods & Instrumentation' in this Encyclopedia, analytical atomic spectroscopy concerns measurements of about 60 elements, generally metals, although consideration of mass spectrometry extends the range to most other elements. Biomedical applications require the measurement of almost all these elements, which are often loosely termed the minerals and trace elements. As given in **Table 1**, biological specimens contain various bulk elements which fulfill essential structural and functional roles. The trace elements, which individually account for less than 0.01% of the dry weight of the organism, are represented by a series of elements essential to health, development and well-being (**Table 2**) and others such as lead, present as contaminants from the environment. With sufficiently sensitive analytical techniques almost all elements of the periodic table can be found and are included among the trace elements.

**Figure 1** demonstrates that minerals and trace elements are relevant to a large number of disciplines which may be regarded as 'biomedical'. Some of

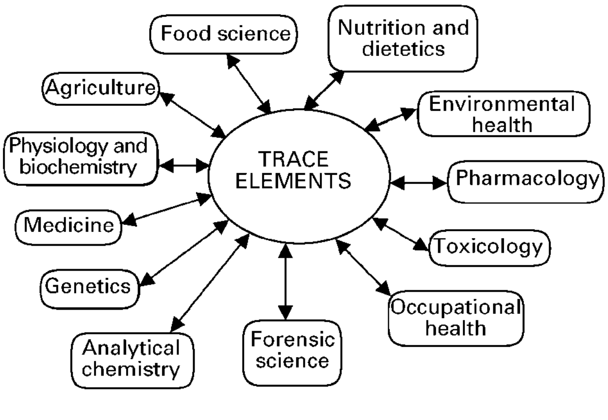



**Figure 1** The relevance of minerals and trace elements to biomedical sciences.

these are described in detail in related articles and are not considered further here.

From the preceding discussion it is evident that measurements of minerals and trace elements in biomedical samples are appropriate to investigations related to:

- the biological importance of essential trace and bulk elements;
- undue effects, i.e. the harmful consequences of a sufficiently large exposure to any element, whether essential or nonessential;
- the pharmacology of certain metals which are administered as the active principal of a pharmacological agent.

Measurements are vital to situations such as understanding mechanisms of action at cellular or molecular levels, assessing the status of, e.g. zinc, within an individual subject and monitoring the effects of exposure.

**Table 1** Bulk elements found in biological specimens

| *Minerals* | *Nonminerals* |
|---|---|
| Calcium | Carbon |
| Iron | Chlorine |
| Magnesium | Hydrogen |
| Potassium | Nitrogen |
| Sodium | Oxygen |
| | Phosphorus |
| | Sulfur |

**Table 2** Essential trace elements (not all are proven to have essential roles in man)

| | |
|---|---|
| *Minerals* | Chromium, cobalt, copper, iron, manganese, molybdenum, nickel, selenium, silicon, vanadium, zinc |
| *Nonminerals* | Fluorine, iodine |

Many analytical techniques have been applied to the determination of minerals and trace elements but the most successful are those based on atomic spectroscopy. The techniques included in the subsequent sections afford the sensitivity required to measure concentrations below 1 ppm in specimens of just a few μL or mg with almost total specificity and relatively few interferences. As situations of deficiency and toxicity are investigated, analytes may be present at very low or very high concentrations within the same specimen type. Therefore, techniques with differing sensitivities may be used, as appropriate, for the same application. Many biomedical investigations require the measurement of just one or two metals. Consequently, multi-element techniques are not necessarily as important as in some other application areas. Nevertheless, there are situations in which this facility does become important.

## Flame atomic absorption spectroscopy (FAAS)

### Biomedical samples and analytes

Flame AAS is suitable for measurement of a limited range of elements present at concentrations greater than about 1 $\mu g\ mL^{-1}$ in biological fluids, and for the analysis of solutions obtained from biological tissues at the completion of the sample preparation steps.

Typical biological fluids include blood and blood serum, blood plasma, urine and saliva. Measurement of calcium in serum was the first analysis to which the technique of AAS was applied and is an obvious example of how FAAS is useful for biomedical analysis. Other specimens e.g. dialysis fluids, intestinal contents, total parenteral nutrition solutions, may be analysed on rare occasions. Elements present at a sufficiently high concentration are lithium and gold when used to treat depression and rheumatoid arthritis respectively, and calcium, magnesium, iron, copper and zinc. Sodium and potassium can be determined by FAAS but are more usually measured by flame atomic emission spectroscopy or with ion selective electrodes. Other elements are present in fluids at too low a concentration to be measured by conventional FAAS with pneumatic nebulization. With other fluids, e.g. seminal plasma, cerebrospinal fluid, analysis may just be possible for a very few elements.

The concentrations of many metals in plant, animal or human tissues are usually much higher than in biological fluids and very often the weight of an available specimen is such that a relatively large mass of analyte is recovered into a small volume of solution, thus enhancing the concentration still

further. For the analysis of tissues (these include specimens such as hair and the cellular fractions of blood) following sample dissolution steps, FAAS may be suitable for measurement of most of the biologically important elements.

### Interferences and sample preparation

FAAS is subject to certain interferences associated with the nature of biological specimens. Mechanisms of the more important ionization, chemical and matrix interferences (**Table 3**) are discussed elsewhere.

**Ionization interferences** Falsely high results may be obtained because of the high concentrations of sodium and potassium, which are among the more easily ionized elements. As with matrix interferences, they may not be evident with all nebulizer–burner systems and the presence of possible interferences should be investigated with any new instrument. If ionization interference is observed the calibration solutions should be prepared with sodium and potassium at approximately the same concentration as in the test specimens. This applies to biological fluids and to solutions prepared from tissues.

**Chemical interferences** The only example of any consequence pertains to the measurement of calcium. The $Ca^{2+}$ ion forms a refractory phosphate molecule and atomization is thereby impaired. Diluents containing $La^{3+}$ or $Sr^{2+}$, preferentially bind to the phosphate and release the $Ca^{2+}$, or EDTA to complex the $Ca^{2+}$ in solution allowing its release in the flame, effectively eliminate this interference. An alternative strategy is to use the hotter nitrous oxide–acetylene flame. A few other chemical interferences have been recorded in atypical settings and are not relevant to biomedical samples.

**Matrix interferences** Two examples occur with biomedical specimens.

(a) Typical biological fluids contain protein and other macromolecules which increase the sample viscosity compared with simple aqueous solutions. The viscosity may reduce the aspiration rate through the narrow capillary tubing of the pneumatic nebulizer and absorbance signals will be attenuated compared with aqueous calibration solutions, giving falsely low results. Techniques to eliminate these effects include:
- Dilution of the sample with water (**Table 4**). A fivefold dilution is usually sufficient but some nebulizers may require a larger sample dilution to equalize the aspiration rates. Dilution may, however, reduce the analyte concentration to below the useful working range. If so, a lower dilution can be tolerated by matching the viscosity of the calibration and sample solutions by either addition of dilute glycerol to the calibrants or addition of agents such as butanol, propanol or Triton X-100 to the sample. Whatever measure is adopted it is important to determine the aspiration rates of calibration and test solutions to demonstrate that the corrective action has been effective.
- Precipitation of the protein with concentrated acid and removal by centrifugation. Trichloroacetic acid or nitric acid are used for this purpose.
- Use of matrix-matched calibration solutions. Either protein is added to the solutions or certified reference materials are used (if available).
- Standard additions for calibration.

(b) Samples with a high content of dissolved solids are liable to produce falsely high results because of nonatomic absorption or light scattering. If this occurs the analyte can be removed from the matrix by extraction into an organic solvent, or a background correction technique can be employed.

### Flame composition

Elements in biomedical specimens which may be measured by FAAS are determined using the air–acetylene flame. The more refractory metals, requiring a higher temperature nitrous oxide–acetylene flame for

**Table 3** Interferences in FAAS

| *Element* | *Interference* | *Remedy* |
|---|---|---|
| Na, K, Fe, Li, Cu, Zn | Ionization | Add ionization suppressant (Cs, Na, K) |
| Ca | Chemical | La, Sr, EDTA |
| | | Higher temperature (greater energy) |
| Fe, Au, | Matrix | TCA precipitation |
| Cu, Zn | | Dilution with water |
| | | Dilution with detergent, alcohol |
| | | Add glycerol to standards |

**Table 4** Typical dilution factors for measurements of minerals in serum or urine

| | |
|---|---|
| Gold | 1+1 |
| Iron | 1+1 |
| Copper, zinc | 1+1 >> 1+4 |
| Lithium | 1+9 |
| Potassium | 1+19 |
| Calcium | 1+50 |
| Magnesium | 1+50 >> 1+100 |
| Sodium | 1+200 |

atomization are at concentrations too low to be determined by flame atomization (except in a few tissue specimens or in indirect methods). Other flames which have historically been used for special applications now have no real place in the analysis of biomedical samples.

For most elements the proportion of acetylene to air in the flame has little influence on formation of the ground state atomic vapour and a large variation in flow rates can be tolerated. The important exception is calcium which, as shown in **Figure 2A**, is more efficiently atomized in a reducing, fuel-rich flame. The position of the light path in the flame is also more critical for calcium than for other elements (**Figure 2B**). Modern computer-controlled instruments are preprogrammed by the manufacturers to operate under optimal conditions.

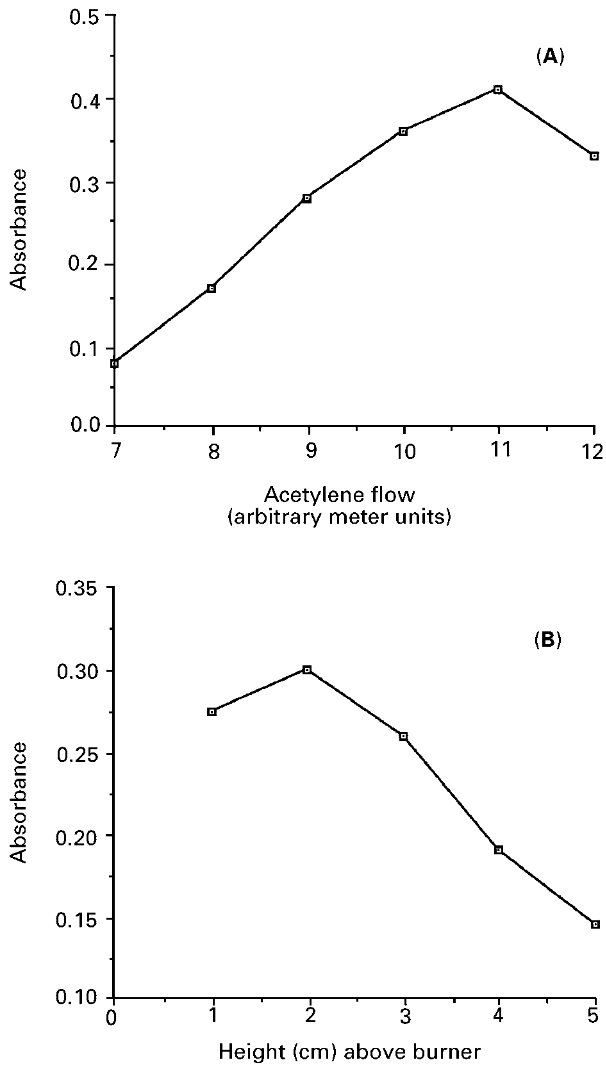



**Figure 2** (A) The effect of increasing the proportion of fuel in an air–acetylene flame on the absorbance given by a solution of calcium. (B) The effect of raising the light path above the burner head on the absorbance given by a solution of calcium.

## Nebulization and sensitivity enhancement

Pneumatic nebulizers were the original devices to introduce the sample as an aerosol into the flame gases. Despite the inefficiency of these nebulizers, coupled with depletion of the atom population in the flame, there are advantages of speed, precision and simplicity, and they have never been replaced by other devices. In practice there can be blockages associated with particulate matter in biological materials, precipitation of protein on the inner wall of the capillary tube and the viscosity effects referred to previously. However, daily cleaning and maintenance will usually obviate these difficulties. Techniques to improve sensitivity so as to allow the positive features of flame atomization to be retained, are widely applied to the analysis of biomedical specimens. These involve atom trapping, preventing dispersal of the atoms, preconcentration by liquid–liquid extraction and preconcentration using solid-phase traps.

**Atom trapping** Ground state atoms condense onto the surface of a water-cooled tube placed above the burner head and below the light path. Dilute solutions are nebulized for periods of up to several minutes to trap the analyte. Then the water is displaced by air with a rapid rise in temperature of the trap so that all the atoms are released together into the light path to give a strong atomic absorption signal. The technique has been applied to the measurement of cadmium and lead in urine.

**Preventing dispersion of the atoms** A hollow tube is mounted on the burner and the light path passes along the length of the tube. Samples are nebulized and atoms enter the tube where they are prevented from dispersing throughout and beyond the flame by the physical presence of the tube. Consequently, a more concentrated atom population is favoured with an improved atomization signal. Such devices were employed for the Delves' microcup technique for solid sample introduction directly into the flame – a procedure which is still used in a few laboratories for the measurement of lead in blood. More use is made of the slotted quartz tube and a number of different designs have been used. The tube is best suited for more volatile elements (e.g. Cd, Cu, Pb, Zn, etc.) and sensitivity enhancements of 3–10-fold are obtained. Furthermore, with the greater sensitivity, larger specimen dilutions can be made so that matrix interferences are eliminated or specimens of small size can be analysed.

Preconcentration methods are employed to enhance sensitivity for FAAS and for other atomic spectrometric techniques. These are described in a later section.

## Electrothermal atomization atomic absorption spectroscopy (ETAAS)

Electrothermal heating for sample vaporization has proven to be an extremely powerful tool for atomic absorption, atomic emission and for introduction of atoms into the sampling torch for inductively coupled plasma (ICP) mass spectrometry (MS). With the exception of a few designs modelled on the graphite rod, all commercial systems are developments of the Massman furnace, and are constructed from graphite. As explained elsewhere, ETAAS affords considerable improvements in sensitivity compared with FAAS and the higher temperatures attained permit the measurement of elements such as Al, Cr and V which do not form an appreciable population of ground state atoms within the air–acetylene flame. Because of the small sample volumes required, less than 50 μL, the technique is ideal for biomedical applications which are often characterized by limited material (as in paediatric investigations, for example).

### Biomedical specimens and analytes

A glimpse at any of the ASU reviews immediately shows that ETAAS is used for the analysis of all specimen types within the biomedical field and for the determination of virtually all applicable elements. With special furnace materials and/or linings, even the very refractory elements such as boron and the rare earths can be measured with useful sensitivity.

Although not extensively exploited, nonliquid samples can be accommodated. Tissue samples may be loaded into the furnace but this is a cumbersome procedure and because only a few mg of material can be taken it is essential that the original specimen is homogeneous – which rarely applies to biological tissues. However, it is possible to prepare suspensions of finely powdered solid samples in a thixotropic medium and the resultant slurry may then be handled as if it were a liquid and be injected into the furnace. Analyses of tissue specimens as slurries are regularly reported.

### Interferences

As explained by L'Vov, who first introduced the technique of ETAAS, electrothermal atomization is ideally accomplished under stable, isothermal conditions. However, the essential design of the Massman furnace imposes dynamic conditions with complex temperature profiles both temporally and spatially within the furnace. While this allows for simplified operational designs it does however lead to more interferences. The topic of interferences is dealt with in more detail in other articles. Of special significance to biomedical specimens, which typically contain large amounts of sodium, chloride and carbon-based biomolecules, are the following.

***Drying*** Protein-rich viscous specimens may dry unevenly with 'explosive sputtering' at the start of the ash phase, causing poor reproducibility.

***Ashing*** Volatile halides, e.g. $AlCl_3$ may form, causing preatomization loss of the analyte.

***Atomization*** Stable compounds such as carbides can be produced, giving low atomization rates (e.g. Mo); vapour phase reactions, especially molecular condensation at the cooler ends of the furnace (e.g. $Pb(g) + 2NaCl(g) \rightarrow PbCl_2(g) + 2Na(g)$), can result in nonatomic absorption of incident radiation and scattering of incident radiation by particulates (carbon, smoke, salts). The majority of developments since the introduction of ETAAS are concerned with the elimination of these interferences.

### Furnace materials and design

Commercially available furnaces are all made from electrographite, electrographite with a pyrolytic coating or total pyrolytic graphite. Devices that delay atomization until the temperature of the gas in the furnace has reached a plateau are described elsewhere. The L'Vov platform is widely used with biomedical specimens, especially in larger furnaces, which are slow to heat to the atomization temperature. A second device, the graphite probe, produces a less sensitive analytical arrangement and is not often used. Transverse heating causes less of a temperature difference between the centre and ends of the furnace compared with longitudinal heating of the conventional Massman design and, therefore, vapour phase condensation is reduced. An authorative review of materials suitable for use in furnace construction, and of recent developments in design, has been prepared by Frech (see Further reading).

### Chemical modifiers

Effective analysis of most biomedical materials requires addition of reagents that modify the behaviour of the specimen during the heating program so as to reduce the interferences described above. The chemical modifiers most commonly employed with these specimens are given in **Table 5**. Triton X-100 is used at a concentration of around 0.1% and is included in with the sample diluent. Gaseous oxygen or air are effective ashing aids but will cause rapid deterioration of the graphite furnace unless a desorption step is included before the temperature is increased for atomization. Other modifier solutions

**Table 5** Chemical modifiers used in the analysis of biomedical specimens by ETAAS

| *Modifier* | *Purpose* |
|---|---|
| Triton X-100 | To ensure smooth, even drying of protein-rich specimens and avoid the formation of a dried crust around a liquid core |
| Gaseous oxygen | To promote destruction of the organic matrix and reduce formation of smoke and other particulates which give nonatomic absorption |
| Ni, Cu, Pd | To stabilize volatile elements, e.g. Se, As, during the dry and ash phases |
| Potassium dichromate | Stabilizes Hg up to a temperature of 200°C |
| $HNO_3$ or $NH_4NO_3$ | To stabilize analyte atoms by removal of halides as HCl or $NH_4Cl$ during the ash phase |
| $Mg(NO_3)_2$ | Becomes reduced to MgO which traps the metals to reduce volatilization losses and also to delay atomization and separate the analyte signal from the background absorption |
| $NH_3H_2PO_4$ or $(NH_3)_2HPO_4$ | Usually used together with magnesium nitrate; reduces volatilization losses and also delays atomization to separate the analyte signal from the background absorption |

can be included with the sample diluent or separately added by the autosampler to the specimen inside the furnace. The choice of modifier often depends on the availability of a source material which is free from contamination.

### Background correction

Background correction is essential for almost all biomedical applications in order to correct for nonatomic absorption of the incident radiation. Nonatomic absorption is much reduced at wavelengths above 300 nm but this is relevant to just a few elements. Each of the techniques for background correction (**Table 6**) have their advantages and disadvantages. The three systems used in commercial instruments are effective although the Smith Hieftje variation is included by very few manufacturers. The Zeeman-effect technique is ideal for dealing with structured background which is of particular significance in the measurement of cadmium in urine and elements such as arsenic and selenium in iron-rich samples, i.e. blood. Without Zeeman-effect background correction these applications are difficult to carry out successfully by ETAAS.

**Table 6** Background correction (BC) techniques

| | |
|---|---|
| Measure absorbance at a nearby nonabsorbing line | |
| Deuterium BC | |
| Smith-Hieftje BC | |
| Zeeman-effect BC; | (i) applied to the light source |
| | (ii) applied to the furnace |

### Novel atomizers

Several research groups have designed atomizers with the purpose of separating the analyte from interfering species, to permit simple atomic absorption. While some appear to be effective, none are commercially available. It was shown some years ago that a 150 W tungsten filament from a light bulb could be used as an electrothermal atomizer. More recently, this concept has been used to develop very small portable instruments for onsite measurement of lead in blood. Excellent results have been reported but a commercial model is still awaited.

## Other techniques for atomic absorption spectroscopy

### Biomedical samples and analytes

The technical and instrumental features of cold vapour and hydride generation are described elsewhere. Mercury is used extensively in industry and there may be situations of environmental use, but the typical sources of exposure are from the diet, especially fish, and from dental amalgam. Depending on the nature of the exposure, e.g. inorganic salts, organomercury compounds, the metal or its vapour, it may be necessary to analyse specimens of urine, blood or tissues and foods. Many of the elements which form volatile hydrides have important biomedical properties. Selenium is an essential micronutrient, arsenic is especially recognized as toxic, while bismuth and antimony have valuable pharmacological properties. As with mercury, all types of biological specimens may require to be analysed.

### Mercury: cold vapour atomic absorption spectrometry

The concentration of mercury in urine is an excellent indicator of recent exposure to mercury vapour or inorganic compounds, and the measurement is very simple following digestion with $KMnO_4$–$H_2SO_4$ to ensure that all carbon–mercury bonds are broken. With other samples a more aggressive approach is required to not only release the mercury but to also destroy the organic matrix which inhibits vaporization of elemental mercury. These more powerful reducing conditions and heating are necessary for analysis of blood and tissues. To prevent loss of

volatile mercury either a closed digestion system or heating with a reflux condenser must be used.

Because the clinical and toxicological effects of alkylmercurials are so different from those of inorganic species it is useful to separately measure organomercury compounds. The essential reaction to the cold vapour technique is the reduction of $Hg^{2+}$ to $Hg^0$. Thus, total mercury concentration in a specimen will be determined after the disruption of carbon–mercury bonds, as described above, while if this step is omitted only the inorganic mercury is measured.

#### Hydride forming elements

Considerable interest in methods to measure arsenic and other hydride forming elements has been evident in recent years. The basic procedure involves careful digestion of the specimens to convert all the different species to a single valency form, reduction with $BH_4^-$, and vaporization to the hydride, followed by atomization in a quartz tube heated in an air–acetylene flame or with electrical thermal wire. However, measurement of total arsenic is not always entirely helpful. Fish contain large amounts of organoarsenic species, which are absorbed and excreted without further metabolism and with no adverse health effects. These species will be included in a total arsenic determination and can mask attempts to measure toxic $As^{3+}$ and metabolites. Thus, methods to measure the individual species, or related groups of compounds in urine or other samples, have been described to provide more meaningful results. These methods include separation by chromatography or solvent extraction, and pretreatment steps which transform only the species of interest into the reducible form.

Accessories from instrument manufacturers for hydride generation allow for either a ‘total consumption’ measurement when all the sample is reacted at once with the $BH_4^-$ and the gaseous hydride taken to the heated cell as a single pulse, or by flow injection to give a continuous, steady-state signal. The former gives a lower detection limit but the latter is easier to automate. The chemical hydride reaction is impaired by other hydride-forming elements and by transition metals so that careful calibration is essential. It has been shown that the hydride may be taken to a graphite furnace where the gas is trapped onto the surface. Rapid electrothermal heating then gives a very sensitive signal. Trapping is more efficient when the graphite is coated with a metal salt, e.g. Ag, Pd, Ir. Electrolysis releases nascent hydrogen and this reaction has been employed as an alternative to chemical hydride generation. With this arrangement the interferences are much less.

With careful sample preparation hydride generation is an extremely valuable analytical technique appropriate to ICP-AES and ICP-MS as well as atomic absorption spectrometry. Within the context of biomedical analysis, measurements of arsenic and selenium are currently the more important.

## Atomic emission spectroscopy

#### Flame atomic emission spectroscopy (FAES)

Of the different techniques for atomic emission spectroscopy (AES) only those which use a flame or an ICP are of any interest for analysis of biomedical specimens. Flame AES, also called flame photometry, has been an essential technique within clinical laboratories for measuring the major cations, sodium and potassium. This technique, usually with an air–propane flame, was also used to determine lithium in specimens from patients who were given this element to treat depression, and was employed by virtually all clinical laboratories throughout the world until the recent development of reliable, rapid-response ion selective electrodes. Biological fluids need only to be diluted with water and in modern equipment the diluter is an integral part of the instrument so that a specimen of plasma or urine can be introduced without any preliminary treatment.

#### Inductively coupled plasma atomic emission spectroscopy (ICP-AES)

At the temperature of the flame there is no useful emission of other biologically important elements. However, with the greater energy of the ICP, much lower detection limits, typically around 1 $\mu g\ mL^{-1}$, are obtained and many elements may be determined in solutions prepared from biological tissues. Recent developments with the optical systems and array detectors offer improvements in sensitivity and data collection. In consequence, elements such as copper, zinc and aluminium may be measured in blood plasma, while the expanded information caught by detectors is making it possible for powerful chemometric manipulation of individual signals to be undertaken.

The most valuable features of ICP-AES is the multi-element capability. In most clinical situations this is not an important requirement but for some research work and in other biomedical applications the comprehensive information derived can be of considerable interest. There are few interferences associated with ICP-AES but, as with FAAS, matrix interferences associated with uptake of sample via the nebulizer may be encountered with some systems.

## Inductively coupled plasma mass spectrometry (ICP-MS)

Other instrumental approaches to inorganic MS are available which, together with X-ray fluorescence spectroscopy, are mainly applied to the analysis of very thin solid specimens such as tissue sections. For more general application, the ICP as the ion source has provided enormous improvements to inorganic MS. These improvements are: convenience to operate, low detection limits (μg $L^{-1}$ or below), multielement analyses, and the ability to measure isotopes, thus allowing stable isotopes to be used as tracers for metabolic and other studies. However, there are important disadvantages associated with ICP-MS. Sampling times of several minutes at a nebulization rate of approximately 3 mL $min^{-1}$ are required to ensure a sufficient number of counts at each mass number and, therefore, the volume of undiluted specimen necessary to give a result can be large. As a consequence of the long sampling times and the frequent measurement of blanks, calibration specimens and quality control samples, the total number of real specimens analysed in a working day is not high. With a slow rate of analysis, the requirement for experienced staff and the very high capital and servicing expenditure, the measurement costs are extremely high.

### Interferences and sample preparation

Matrix interferences affecting the nebulizer are similar to those of FAAS and ICP-AES, and sample preparation generally involves dilution or acid digestion of liquid specimens and digestion or solubilization of tissues. In addition, some suppression of analyte signals may be caused by sodium and other dissolved inorganic components. To overcome this the salts may be removed by chelating resins or by using reference materials with a composition similar to the test specimens as calibrants. Of greater importance, ICP-MS is subject to a number of spectral interferences due to the formation of polyatomic species with the same mass number as the element being measured. For example, in the determination of zinc there are interferences due to sulfur ($^{32}S^{16}O_2$, $^{33}S^{16}O_2{}^1H$ and $^{32}S^{16}O^{18}O$), and chlorine ($^{35}Cl^{16}O_2$) on $^{64}Zn$, $^{66}Zn$ and $^{67}Zn$. Various techniques to reduce spectral interferences include the use of ion-exchange, either off-line or on-line (usually an HPLC column), or liquid–liquid extraction, to separate the analyte from the interferants or to count an isotope which is relatively free from interferences. The more recent generation of high-resolution mass spectrometers is reported to give good discrimination of analyte isotopes from other species but at a cost of higher detection limits. To improve the signal associated with isotopes of low abundance, techniques for sample introduction that avoid the use of nebulizer have been developed. These include mercury vapour and hydride generation, electrothermal vaporization and laser ablation techniques.

### Biomedical applications

The technique of ICP-MS is rarely appropriate for measurement of a single element. Exceptions are where sensitivity is unmatched by other techniques, and for isotope work. The very low detection limits for rare earth elements and the actinides have permitted a number of studies relating to the biochemistry and unusual sources of exposure to these elements. Similarly, low levels of occupational and environmental exposure to platinum and other noble metals are now being investigated. Isotope work includes: fundamental biochemical studies using isotopes of elements such as H, C, N and O; absorption and distribution of essential trace minerals, e.g. Fe, Zn, Se in situations such as childhood, pregnancy and dietary deficiencies; and the identification of sources of exposure to lead and other nonessential elements by comparisons of isotopic compositions, i.e. a ‘fingerprinting’ approach. More complex investigations require a combination of these features: e.g. determination of intestinal absorption rates of different dietary selenium species involves feeding subjects with foods previously enriched with stable isotopes, and separation of the selenium species in specimens such as blood by HPLC coupled directly to the ICP-MS, which affords both very sensitive analysis and differentiation of the selenium isotopes. Indeed, speciation work such as this represents a major area of biomedical interest to which the technique of ICP-MS is applied. A further activity involves the characterization and certification of biomedical reference materials where the high degree of accuracy required is achieved by inclusion of isotope dilution analysis into the multielement measurement.

## Sample preparation

The objectives for preparation of biomedical specimens are to (i) remove interfering components from the matrix and (ii) adjust the concentration of analyte to facilitate the actual measurement. These objectives may be realized by a number of approaches (**Table 7**) which in general are appropriate to all the techniques described in this article.

Methods for destruction of the organic matrix by simple heating or by acid digestion have been used extensively and are thoroughly validated. Microwave

**Table 7** Approaches to sample preparation

| *Procedure* | *Remarks* |
|---|---|
| Dilution, protein precipitation | Using simple off-line arrangements or flow-injection manifold |
| Dry ashing | Using a muffle furnace or a low-temperature asher |
| Acid digestion | (i) In open vessels with convection or microwave heating<br>(ii) In sealed vessels to increase the reaction pressure |
| Base dissolution | Using quaternary ammonium hydroxides |
| Chelation and solvent extraction | For analyte enhancement and removal of interferences |
| Trapping onto solid-phase media | For analyte enhancement and removal of interferences |

heating is now well established for this purpose with specifically constructed apparatus to avoid dangers of excessive pressure within reaction vessels. Although the number of specimens which can be processed is not large, microwave heating affords rapid digestion and low reagent blanks. More recent developments include continuous flow systems for automated digestion linked directly to the instrument for measurement of the analyte(s).

Preconcentration by liquid–liquid partitioning is a widely used procedure. Analyte atoms in a large volume of aqueous specimen are complexed with an appropriate agent and then extracted into a smaller volume of organic solvent. This leads to enhancement of concentration and also removes the analyte from potential/real interferences in the original matrix. It is used to measure lead in blood and metals in urine and for other applications. Preconcentration by trapping onto solid-phase media represents the area where much of the recent interest in FAAS has been focussed but is relevant to all sample-preparation work. The original work involved adsorption onto material such as charcoal or alumina but newer phases include ion-exchange resins and novel support systems to which functional groups are added to confer increased selectivity and capacity. While trapping of an analyte from a dilute sample, and elution into a small volume of release solution may be accomplished off-line, developments in flow-injection analysis provide for the assembly of simple on-line manifolds so that complete measurements may be carried through automatically. Developments in these applications involving a wide range of biomedical sample types and elements are regularly described in the ASU reviews. Biological reference materials with and without certified values are readily available and several interlaboratory comparison programmes are established for analysts to determine the accuracy of their methods.

*See also:* **Atomic Absorption, Methods and Instrumentation; Atomic Absorption, Theory; Atomic Emission, Methods and Instrumentation; Fluorescence and Emission Spectroscopy, Theory; Food and Dairy Products, Applications of Atomic Spectroscopy; Forensic Science, Applications of Atomic Spectroscopy; Inductively Coupled Plasma Mass Spectrometry, Methods; Pharmaceutical Applications of Atomic Spectroscopy; X-ray Fluorescence Spectrometers; X-ray Fluorescence Spectroscopy, Applications.**

## Further reading

Bacon J, Crain J, McMahon A and Williams J (1998) Atomic spectrometry update – Atomic mass spectrometry. *Journal of Analytical Atomic Spectrometry* **13**: 171R–208R.

Ellis A, Holmes M, Kregsamer P, Potts P, Streli C, West M and Wobrauschek P (1998) Atomic spectroscopy update – X-ray fluorescence spectrometry. *Journal of Analytical Atomic Spectrometry* **13**: 209R–232R.

Frech W (1996) Recent developments in atomizers for electrothermal atomic absorption spectrometry. *Fresenius Journal of Analytical Chemistry* **355**: 475–486.

Roelandts I (1997) Biological and environmental reference materials: update 1996. *Spectrochimica Acta, Part B* **52B**: 1073–1086.

Taylor A (1997) Applications of recent developments for trace element analysis. *Journal of Trace Elements in Medicine and Biology* **11**: 185–187.

Taylor A (1998) Atomic absorption and emission spectrometry. In: Crocker J and Burnett D (eds) *The Science of Laboratory Diagnosis*. Oxford: Isis Medical Media.

Taylor A, Branch S, Halls DJ, Owen L and White M (1999) Atomic spectrometry update – Clinical and biological materials, foods and beverages. *Journal of Analytical Atomic Spectrometry* **14**: 717–781.

# C

## $^{13}$C NMR, Methods

**Cecil Dybowski**, **Alicia Glatfelter** and **HN Cheng**,
University of Delaware Newark, DE, USA

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Methods |

The analysis of carbonaceous materials with spectroscopy takes on special importance because, without doubt, the study of the chemistry of carbon is more extensive than that of any other element. The strong sensitivity of NMR parameters to chemical state has long made NMR spectroscopy a favourite tool for analysis of organic materials. Early on, proton NMR spectroscopy was used to analyse materials, but with the development of ever-more-sensitive instrumentation, the NMR analysis of carbon in organic molecules has become the essential tool for identification of carbon-containing materials.

Several isotopes of carbon occur naturally. The most abundant isotope of naturally occurring carbon is $^{12}$C, which comprises 98.90% of all carbon. It has no magnetic moment and therefore is not active in NMR spectroscopy. Most people are familiar with the existence of $^{14}$C, a minute percentage of naturally occurring carbon, because, through measurement of its radioactive emissions, one can define the age of certain ancient objects. It is not NMR active because its spin is also zero. $^{13}$C, comprising only 1.10 % of naturally occurring carbon, is the carbon isotope on which NMR spectroscopists focus strongly because it has a spin of $\frac{1}{2}$ like the proton or the $^{19}$F nucleus.

### Properties of the $^{13}$C nucleus

The NMR properties of $^{13}$C (**Table 1**) give an indication of both the strengths and the weaknesses of analysis using it. The gyromagnetic ratio of $^{13}$C is approximately one-fourth that of the proton and, as a result, the resonance condition for $^{13}$C at a given magnetic field strength occurs at a radiofrequency approximately one-fourth that of the proton resonance. For a sample with equal numbers of spins and equal spin–lattice relaxation rates, the predicted signal-to-noise ratio of the $^{13}$C spectrum at a particular magnetic field strength would be over an order of magnitude lower than that of a proton spectrum. Additionally accounting for the concentration difference due to the natural abundance, one finds that typically the absolute receptivity of $^{13}$C relative to the proton is only $1.76 \times 10^{-4}$. Thus, there are difficulties with obtaining $^{13}$C NMR spectra compared to $^{1}$H NMR spectra. However, the utility of carbon NMR in studies of organic molecules guarantees $^{13}$C will be one of the nuclei more heavily investigated by NMR spectroscopists.

**Table 1** $^{13}$C NMR properties

| | |
|---|---|
| Abundance | 0.011 |
| Spin | $\frac{1}{2}$ |
| Gyromagnetic ratio (Hz T$^{-1}$) | $1.0705 \times 10^{7}$ |
| Typical shift range (ppm) | 300 |
| Receptivity relative to $^{1}$H | $1.59 \times 10^{-2}$ |

### Single-pulse Fourier transform NMR

The analysis of carbonaceous materials would indeed be difficult if it were not for the fact that Fourier transform NMR spectroscopy allows one to increase the signal-to-noise ratio by coadding experiments. The simplest experiment is the gathering of the response to a single pulse. The length of the pulse for maximum signal, the $\pi/2$ pulse, on most modern spectrometers is typically in the range of 5 to 20 µs. Data acquisition goes on for a time that depends on the spectrum width and resolution desired, often for times greater than a second.

In the single-pulse experiment, an important parameter is the rate at which the experiment can be repeated, determined by the time it takes to repolarize the system after the magnetization has been destroyed in the excitation and acquisition. This is a

function of the spin–lattice relaxation time, $T_1$, of the carbons in the sample. For typical small organic molecules, $T_1$ can be of the order 10 to 60 s or more at room temperature. A rule of thumb for quantitative evaluation of NMR spectra is that the repolarization time is five times the longest $T_1$ in the sample. For some samples, particularly those containing quaternary carbon atoms or carbonyl groups, this time can be substantially longer than a few minutes. Thus, a compromise is often struck, in which the repolarization time is shortened to some few tens of seconds, with the knowledge that certain resonances may be attenuated, making the resultant spectrum nonquantitative.

Another compromise involved to obtain data faster is to use a pulse width different from $\pi/2$ and a repolarization time shorter than $5T_1$. For a resonance with a relaxation time, $T_1$, and a repolarization time, $T_R$, one may calculate the angle, $\alpha_E$ (the Ernst angle), that gives the largest signal-to-noise by coaddition in a given time:

$$\cos \alpha_E = \exp(-T_R/T_1) \qquad [1]$$

The variation of the Ernst angle with $T_R/T_1$ is shown in **Figure 1**. The optimal pulse angle may be substantially less than $\pi/2$ for typical conditions encountered for carbon NMR spectroscopy where the repolarization time is short compared to $T_1$. For a fixed repetition time, the Ernst angles of carbons in a sample will not be identical because $T_1$ varies from carbon to carbon. Thus, even in defining the pulse angle, some compromise must be made, depending on the range of $T_1$ values in the sample.

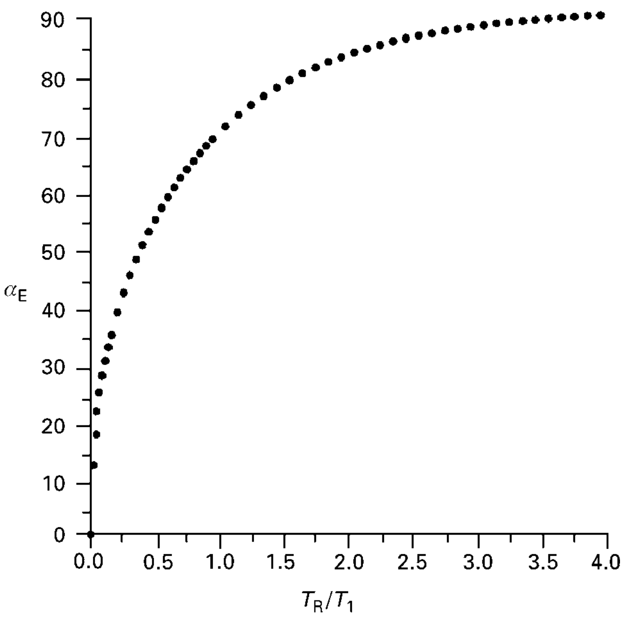



**Figure 1** The Ernst angle as a function of the ratio of the repolarization time to the spin–lattice relaxation time.

## Heteronuclear decoupling

Detection of the $^{13}$C NMR spectrum of an organic compound in solution with the single pulse experiment, as described above, is quite feasible. However, were the spectrum acquired in this manner, one would notice that the number of peaks in it exceeds the number of unique carbon environments. Most organics contain protons as well as carbon atoms. These protons are coupled to the carbons through the indirect (or scalar) *J* coupling, resulting in splitting of the carbon resonances of any carbon with attached protons.

Generally, carbons are coupled to one (AX), two ($AX_2$), or three ($AX_3$) protons. In these cases, one sees the appearance of a doublet (AX) a triplet ($AX_2$), or a quartet ($AX_3$). If the carbon is coupled to several chemically distinct protons, one may see patterns that arise from spin systems labelled AXY, $AX_2Y$, etc. These splittings contain information on the coupling constants between carbons and protons that can be of value in structure elucidation, but frequently the study of carbon NMR spectra focusses mainly on the unique carbon environments in a sample, and the splitting that the coupling brings only confuses the analysis. Thus, carbon NMR spectra are often obtained with simultaneous irradiation of the offending protons which suppresses the effects of coupling – the heteronuclear decoupled NMR experiment.

One question sometimes asked is why it is not necessary also to decouple the carbons from one another. Because of the low natural abundance of $^{13}$C, it is not very likely that a $^{13}$C will be near another spin-active $^{13}$C. Coupling requires relatively close proximity of the spins that are coupled, and it is simply quite unlikely that the spectrum has contributions from molecules containing two or more $^{13}$C atoms in a sample with the natural abundance of $^{13}$C. For sample containing substantial enrichment of the carbon in $^{13}$C (e.g. to increase signal in a single scan), such couplings may affect the spectrum, and consideration of them is an important point to keep in mind.

Another important feature of carbon spectroscopy is the fact that the spectrum in natural abundance is really a superposition of spectra from different molecules, each with substitution of $^{13}$C at different sites in the molecule. The spectrum is not that of a single molecule with $^{13}$C at each position. This fact is sometimes

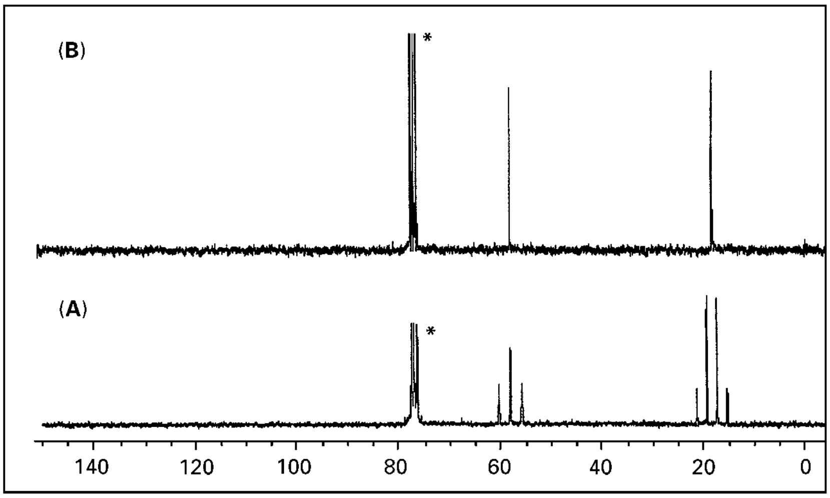



**Figure 2** (A) Ethanol: the coupled $^{13}$C NMR spectrum; (B) the decoupled $^{13}$C NMR spectrum. The solvent resonance is indicated with an asterisk.

forgotten because the analysis of the spectrum is often presented as if one were analysing a single molecule.

Heteronuclear decoupling is effected by irradiation of the protons during the acquisition of the carbon signal. While this may sound simple, the process can be quite complex. For example, exposing the protons to a single-frequency excitation will only decouple protons whose resonances are at that frequency; others will be more or less affected, depending on how close their resonances are to the irradiation frequency. Thus, NMR spectroscopists have developed ways to irradiate over a band of frequencies that encompasses all of the protons coupled to the carbons. One means to spread the irradiation over the various protons is noise modulation of the proton excitation. This random modulation can be shown to result in excitation in a band that will excite all of the protons coupled to carbons. Such a technique is known as 'noise decoupling' or 'broadband decoupling'.

The use of decoupling with pure-frequency sources has been important in assigning resonances, since the observation of collapse of a coupling pattern indicates the existence of coupling between the carbon and the particular proton resonance irradiated. Experiments providing similar information are 'off-resonance decoupling' and 'spin ticklings'. In these cases, the centre frequency of the noise irradiation or the field strength is varied to observe the effects on the resonance lines of coupling partners. From these, one gathers information on the coupling partners, the sizes of couplings, and the resonance frequencies of the coupling partners. In former times, such experiments were frequently the method of choice to assign resonances; however, the current use of two-dimensional NMR methods (see below) to determine coupling partners probably presages the decline of the use of selective-decoupling techniques in the future.

Truly random noise decoupling deposits energy in the sample over a range of frequencies near the centre irradiation frequency. Unfortunately, most of this energy falls at frequencies where there are no proton resonances. This portion of the energy does no useful work in suppressing the coupling of protons to carbons, but it does affect the sample. The energy deposited ultimately results in heating of the sample. To limit heating and to extend the range of useful decoupling, in the early 1980s, spectroscopists focussed on the possibility of using coherent pulse sequences to produce the same effect as noise decoupling – the suppression of effects due to coupling of carbons to protons. The techniques go by various acronyms such as MLEV-16 and WALTZ-16, and the concept behind them is quite simple. Coherent fast switching of a proton spin between its two spin states produces a kind of short-term cancellation of evolution due to the *J* coupling, as observed by the detected carbon magnetization. Importantly, because of the coherent nature of the process, the energy deposited in the sample goes to spin inversion, which limits heating of the sample. Most modern spectrometers implement some version of this technique when the spectrometer is performing an experiment with broadband decoupling.

The importance of heteronuclear decoupling can be seen in **Figure 2**, where the carbon spectra of ethanol obtained with and without decoupling are shown. The decoupled spectrum in **Figure 2B** shows only lines for the various unique carbons sites. The simplicity of the spectrum in **Figure 2B** indicates the power of decoupling in unravelling the spectrum of a complex material.

## Nuclear Overhauser enhancement

In the heteronuclear decoupling experiment, it has proved convenient to have noise decoupling active not just during acquisition, but also during the repolarization time between experiments. If the experiment is performed in this manner, the spectrum is modified by another effect, the nuclear Overhauser effect enhancement (NOE). The presence of the perturbing field during the repolarization period affects the return to equilibrium of the spin system by saturating the proton resonance. This saturation directly impacts the cross-relaxation between the carbons and protons, increasing the signal in the case of $^{13}$C. If the system is allowed to reach equilibrium under the perturbation during repolarization, one predicts an enhancement of the signal, $I_{eq}$, over the signal without the NOE, $I_0$:

$$I_{eq} = I_0(1+\eta) = I_0\left(1+\frac{\gamma_H}{2\gamma_C}\right) \qquad [2]$$

This increase is about 2.998, making the detection of the carbon resonance easier. Of course, for a particular repolarization time, not all $^{13}$C spins in a sample may have reached an equilibrium. The NOE is maximal if the relaxation mechanism solely originates in the direct dipole–dipole coupling between protons and carbons. This mechanism seems to be dominant for carbons directly bonded to protons, but for quaternary carbons and carbonyl carbons, such is not the case, and reduced enhancements are observed for these carbons. For this reason, a spectrum obtained with the NOE present should not be considered quantitative.

The presence of the NOE can be avoided by turning proton irradiation off during the repolarization time between experiments. If the experiment includes the activation of the proton irradiation for decoupling during acquisition, one obtains a spectrum with no enhancement and with coupling suppressed. On the other hand, if the proton irradiation is off during acquisition and the repolarization time, the spectrum has no enhancement and displays all couplings. There are four possible kinds of experiment, depending on when proton irradiation is on, as indicated in **Table 2**. Spectroscopists often use the NOE in obtaining carbon spectra as it gives a significantly larger signal for many of the carbons in a sample.

**Table 2** Effects of decoupling and NOE in carbon spectra

| *Irradiation during repolarization time* | *Irradiation during acquisition* | *Character of the spectrum* |
|---|---|---|
| No | No | Coupled, no enhancement |
| Yes | No | Coupled, with enhancement |
| No | Yes | Decoupled, no enhancement |
| Yes | Yes | Decoupled, with enhancement |

## Pulse techniques for analysis

The single-pulse NMR experiment is a powerful tool for analysing materials, especially carbon. However, there are questions and ambiguities that cannot be unravelled with this simple experiment alone. Some may be addressed with selective-decoupling techniques. However, the use of spectral-editing sequences, and ultimately 2D NMR, has aided assignment of resonances in carbon spectroscopy. As an example, knowing the number of protons to which a carbon is coupled brings additional information to the process of identifying the structure of some unknown carbon-containing material. Generally, the sequences for carrying out these investigations involve pulse excitation of the carbons and the protons in such a way that the magnetization from a particular carbon centre behaves in a predictable manner. The techniques are often referred to by spectroscopists in the acronymical shorthand so prevalent in many spectroscopies: DEPT, INEPT, APT, and many, many more. The techniques involve modulation of the spin evolution or outright transfer of polarization from one species to the other.

As an example, we discuss the so-called attached-proton test (APT). It involves the use of a spin echo, which is produced by a sequence of two pulses, as shown in **Figure 3**.

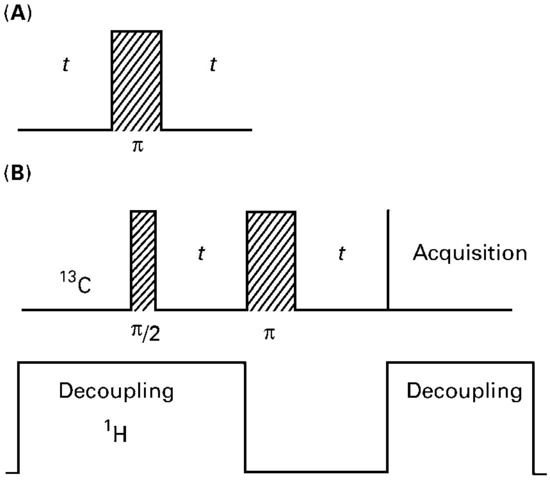



**Figure 3** (A) The single π pluse. (B) The gated-decoupling attached-proton test, showing the sequence applied to the carbons (top) and the sequence applied to the protons (bottom).

First, we discuss the simple spin echo produced by a single $\pi$ pulse (**Figure 3A**). Consider a magnetic moment, subject only to the chemical-shift interaction, that lies initially along the $x$ axis of the rotating frame. After a time $t$, it will have moved ahead of the $x$ axis of the rotating frame by an angle determined by the time $t$ and the chemical shift. The $\pi$ pulse inverts this magnetic moment so that immediately after the pulse it sits at that angle behind the $x$ axis. During the subsequent period, however, it catches up to the $x$ axis, so that at a time $t$ after the $\pi$ pulse it is aligned with the $x$ axis as it was at the beginning. In effect, at that time, the position of the magnetic moment is as if it were not subject to a chemical shift.

On the other hand, consider a spin that has no chemical shift (i.e. it sits at the frequency of the excitation) but is subject to a $J$ coupling to another spin. During the first period, the magnetic moment begins to deviate from the $x$ axis of the rotating frame, depending on its spin state and that of its coupling partner. Let us now apply a $\pi$ pulse only to the spin under observation. Once again, the magnetic moment returns to alignment with the $x$ axis in the second period.

In a third experiment, we propose to carry out the experiment of **Figure 3B** on the spins. Once again, we assume that the chemical shift is zero, but the $J$ coupling is not. During the first period, there is no evolution of the spin because the decoupling suppresses this interaction, so the magnetic moment remains along the $x$ axis in the rotating frame. The $\pi$ pulse does not change the direction of the spin, since it is along the $x$ direction. However, in the second period, the $J$ coupling causes the magnetic moment to deviate from the $x$ axis, so that a time $t$ after the $\pi$ pulse, the magnetization is moved away from the $x$ axis by an amount determined by the time $t$ and the $J$ coupling.

Were we to carry out the same experiment with a nonzero chemical shift and a nonzero $J$ coupling, we would discover that the evolution due to chemical shift would be cancelled but the evolution due to $J$ coupling would not. Thus it is possible with this experiment to label magnetic moments at time $t$ after the $\pi$ pulse according to the effects of $J$ coupling. If, at that point, an acquisition is started, the peaks in the resulting spectrum will have magnitudes corresponding to the amplitude at that time, which will be determined by the $J$ coupling.

Consider an AX system subject to this experiment. The spectrum will show no splittings, for it is taken with decoupling. The amplitude of the line will oscillate with $t$. In particular, for $t = 1/J_{CH}$, this signal will appear totally inverted. On the other hand, a carbon not coupled to an X spin (such as carbonyl or quaternary carbon) experiences no deviation and gives a positive signal under the same set of circumstances. A careful examination of the evolution of the components of the $AX_2$ and $AX_3$ species shows that, under this same experiment, the signal of the carbon in an $AX_2$ environment experiences no net deviation of the magnetic moment at a time $t$ after the pulse, whereas the $AX_3$ system experiences a net inversion. Thus, the response to such an experiment is a spectrum with the signals from AX and $AX_3$ inverted relative to those of A and $AX_2$. From such a spectrum, one can easily discriminate these kinds of environment, as one can see from the spectrum of **Figure 4**.

The utility of the APT sequence hinges on the fact that one may choose a time $t$ that is equal to $\pi/J_{CH}$ for all carbons in the sample. This will only be true if the coupling constant is similar for carbons in various environments. As it turns out, the coupling constants are frequently in the range of 120 to 150 Hz, which is a sufficiently narrow range to allow the technique to work relatively well. A typical value of $t$ that is used in these experiments is 7 ms. However, it must be emphasized that there are systems for which the coupling constants are very different from this nominal range, and lines from these carbons will not necessarily appear to be of the 'right phase'; such deviations should always be checked by additional experiments with different values of $t$.

The other commonly seen spectral editing techniques used in carbon NMR involve the transfer of polarization from protons to carbons in particular ways. A good example of this type of experiment is the so-called DEPT (distortionless enhancement by polarization transfer) sequence (**Figure 5**). By careful choice of experimental parameters, one can obtain spectra that represent only the quaternary, only the methines, only the methylenes and only the methyl groups in the sample. This is done in this case by use of the sequence shown in **Figure 5**. Four spectra are obtained with $\delta = \pi/4$, $\pi/2$, and $3\pi/4$ in the sequence

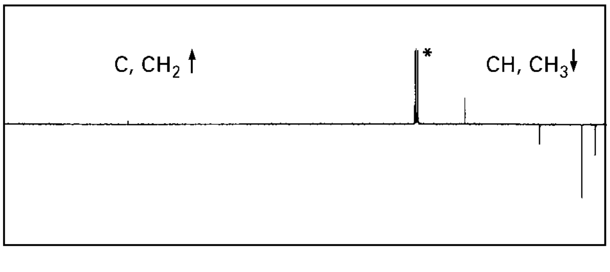



**Figure 4** The APT spectrum of ethyl isobutyrate in deuterochloroform. The solvent resonance is indicated with an asterisk.

and a single-pulse experiment. **Figure 6** shows these spectra for ethyl isobutyrate. In **Figure 6A** all the lines appear. In **Figure 6B** all the lines except the quaternary carbons appear positive. In **Figure 6C** only resonances from CH groups appear. In **Figure 6D** the CH and $CH_3$ appear positive and the $CH_2$ groups appear negative. With these spectra, one can identify the various species present.

The INEPT (insensitive nuclei enhanced by polarization transfer) experiment is similar to the DEPT experiment, except that it uses an evolution period to produce the appropriate condition for spectral selection that depends critically on the size of the coupling constant. The technique is therefore somewhat more susceptible to artefacts resulting from the variability of coupling constants than is DEPT.

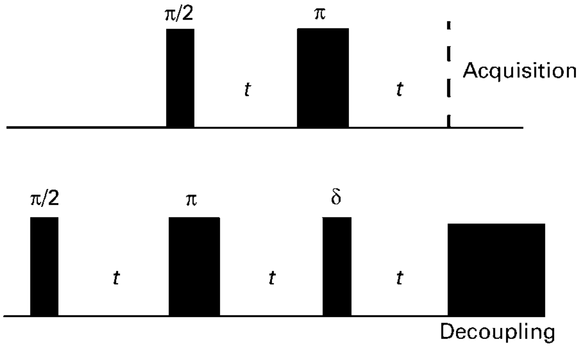



**Figure 5** The DEPT experiment. The top line is the sequence applied to the carbons and bottom is applied to the protons. The time $t$ is fixed as $1/(2J_{CH})$.

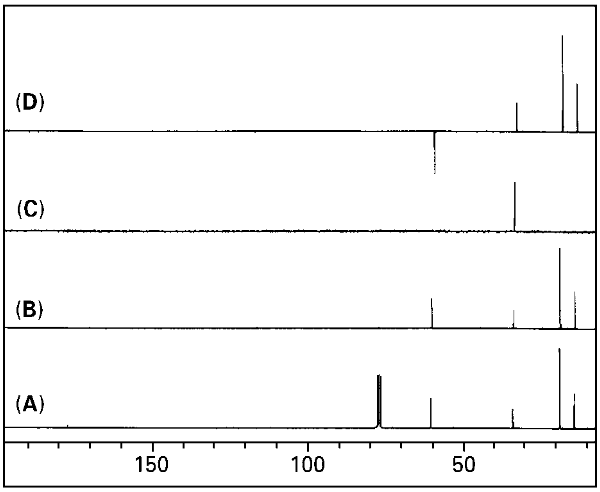



**Figure 6** DEPT spectra of ethyl isobutyrate: (A) the single-pulse spectrum; (B) DEPT-π/4; (C) DEPT-π/2; (D) DEPT-3π/4. The triplet arises from the carbon of the solvent, deuterochloroform.

## Two-dimensional carbon NMR spectroscopy

Two-dimensional (2D) NMR spectroscopy is now used routinely for many analytical problems. A convenient way to conceptualize 2D NMR spectroscopic studies is to divide the time of a 2D pulse experiment into four periods:

$$\begin{array}{l} \text{Preparation} \rightarrow \text{Evolution} \rightarrow \text{Mixing} \rightarrow \text{Detection} \\ \text{Characteristic time}: \quad t_1 \qquad\qquad\qquad t_2 \end{array}$$

During the preparation period the spin system is 'prepared' for the 2D experiment. This may entail a simple $\pi/2$ pulse or a complex series of pulses and delay times. These preparation pulses and delays are applied in such a way, depending on the experiment, as to produce a particular behaviour of the spins during the evolution period (time = $t_1$). The mixing period is needed in some 2D experiments to permit transfer of magnetization from one spin to another. The signal is finally detected in the detection period (time = $t_2$).

In a typical 2D experiment the evolution time $t_1$ is varied systematically. For every value of $t_1$, the signal as a function of $t_2$ is collected in the detection period. The complete set of these signals (as a function of $t_1$ and $t_2$) constitutes the 2D data set, $S(t_1, t_2)$. Suitable Fourier transformation of $S(t_1, t_2)$ with respect to both $t_1$ and $t_2$ then produces the 2D spectrum, $S(f_1, f_2)$, where $f_i$ denotes the frequency.

A large number of 2D experiments have been devised. A list of some simpler 2D experiments is shown in **Table 3**. These experiments can be broadly classified as either $J$-resolved or shift-correlated. In $J$-resolved experiments, the chemical shift $\delta$ and the coupling constant $J$ are plotted to form a 2D representation with $\delta$ and $J$ axes. In this way, the scalar coupling pattern for spins with different chemical shifts can be visualized clearly. In shift-correlated experiments, 2D plots with different chemical shifts can be obtained. These experiments are found to be very useful in resolving different spectral assignment problems.

An example is shown of the C–H HETCOR (heteronuclear shift correlations) experiment. The pulse sequence is given in **Figure 7**. In the preparation period, a $\pi/2$ pulse is applied to the $^1H$ nuclei (a), causing the $^1H$ magnetization to evolve with time $t_1$. Midway in the evolution period, a $\pi$ pulse is applied to the $^{13}C$ nuclei (b). The $\pi$ pulse produces refocussing of the $^{13}C$–$^1H$ coupling at the start of the mixing period (c). The mixing time $\Delta_1$ is set to $1/(2J_{CH})$ to polarize the $^1H$ magnetization. At (d), $\pi/2$ pulses at

**Table 3** Some simple 2D experiments

| *Type* | *Name* | *Function* |
|---|---|---|
| *J*-resolved | Homonuclear | Shows $^1$H−$^1$H coupling patterns; permits precise $J_{HH}$ measurements |
| | Heteronuclear | Shows $^{13}$C−$^1$H coupling patterns; permits precise $J_{CH}$ measurements |
| Shift-correlated | COSY (HH-COSY) | Permits correlation of $^1$H signals through $^1$H−$^1$H couplings |
| | NOESY | Permits correlation of $^1$H signals through NOE interactions |
| | HETCOR (CH-COSY) | Permits correlation of $^1$H and $^{13}$C signals through $^1J_{CH}$ |
| | HMQC | Gives similar information as HETCOR; indirectly detects the less receptive $^{13}$C through the more receptive $^1$H |
| | RELAY | Permits correlation of $^1$H signals through relayed coherence; $^1$H−$^1$H or $^{13}$C−$^1$H |
| | CH COLOC | Permits correlation of $^{13}$C–$^1$H signals through long-range $^{13}$C–$^1$H coupling |
| | HMBC | Gives similar information as COLOC; indirectly detects the less receptive $^{13}$C through the more receptive $^1$H |
| | 2D INADEQUATE | Gives correlation of $^{13}$C signals through direct $^{13}$C−$^{13}$C coupling |

COSY = homonuclear chemical shift correlation spectroscopy; NOESY = 2D nuclear Overhauser effect (NOE) spectroscopy; HMQC = heteronuclear mulltiple-quantum coherence; RELAY = relayed coherence transfer; COLOC = correlation via long-range coupling; HMBC = heteronuclear long-range coupling; INADEQUATE = incredible natural abundance double quantum transfer experiment.

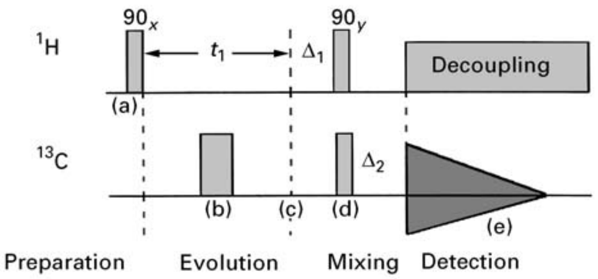



**Figure 7** The C–H HETCOR pulse sequence.

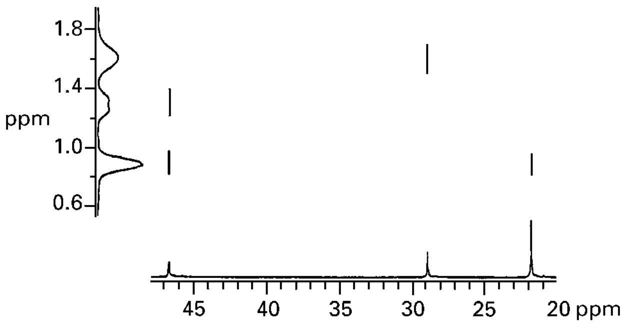



**Figure 8** The C–H HETCOR spectrum of polypropylene.

$^1$H and $^{13}$C are simultaneously applied, prompting transfer of the $^1$H polarization to $^{13}$C, which is then detected with simultaneous $^1$H decoupling to obtain the decoupled $^{13}$C spectrum.

A C–H HETCOR plot for polypropylene is given in **Figure 8**. The projection along one axis gives the $^{13}$C spectrum, and the projection along the other axis gives the $^1$H spectrum. Thus, if the $^1$H and the $^{13}$C spectra are both partially assigned, this experiment sometimes permits the information to be combined in order to derive improved assignments of both spectra.

## Spectral assignments

$^{13}$C NMR spectroscopy is ideally suited for organic structure determination. As noted above, it is a rare spin (only 1.1% in natural abundance), which means $^{13}$C–$^{13}$C couplings are not usually seen. Secondly, the $^{13}$C shift range is large (220 ppm), and neighbouring atoms (as far as $\varepsilon$ position away) affect the chemical shift of a given carbon. Thirdly, the chemical shift is less sensitive than $^1$H to inter- and intramolecular interactions, such as hydrogen bonding and self-association. As a result, in most cases a one-to-one correspondence can be made between an organic structure and the $^{13}$C NMR spectrum.

The $^{13}$C chemical shift is diagnostic of different functional groups. An abbreviated shift scale is shown in **Figure 9**. This shift scale is useful for spectral interpretation and is reminiscent of similar structure–spectral correlation charts used in infrared (IR) spectroscopy. Yet $^{13}$C NMR provides much more detailed structural information than IR. It has been known since the 1960s that the $^{13}$C shifts of alkanes are approximately additive with increasing number of substituents. Similar empirical additivity rules have been found for alkyl carbons with different functional groups and other additivity rules have been proposed for olefinic, aromatic, and carbonyl carbons. These linear additivity rules are a distinct feature of $^{13}$C NMR, and they simplify spectral interpretation.

There are literally thousands of papers using $^{13}$C NMR for structure determination of organic materials. To many NMR users, spectral assignment is an

**Table 4** Substitutent chemical shift parameters (in ppm) for $^{13}$C NMR ($k_0$ = –2.3 for tetramethylsilane reference)

| *Substituent* | | $\alpha^b$ | $\beta$ | $\gamma$ | $\delta$ |
|---|---|---|---|---|---|
| Paraffin[a] | $-\mathrm{C}\leqslant$ | 9.1 | 9.4 | –2.5 | 0.3 |
| Ether, alcohol[a] | –O– | 49.0 | 10.1 | –6.0 | 0.3 |
| Amine[a] | $-\overset{+}{\mathrm{N}}\leqslant$ | 28.3 | 11.3 | –5.1 | 0.3 |
| Ammonium[a] | $-\mathrm{N}<$ | 30.7 | 5.4 | –7.2 | –1.4 |
| Thioether, thiol[a] | –S– | 11.0 | 12.0 | –3.0 | –0.5 |
| Phenyl | $-C_6H_5$ | 22.1 | 9.3 | –2.6 | 0.3 |
| Fluoro | –F | 70.1 (1,2) | 7.8 | –6.8 | 0.0 |
| | | 69.0 (3) | | | |
| | | 66.0 (4) | | | |
| Chloro | –Cl | 31.1 (1,2) | 10.0 | –5.1 | –0.5 |
| | | 35.0 (3) | | | |
| | | 43.0 (4) | | | |
| Bromo | –Br | 18.9 (1,2) | 11.0 | –3.8 | –0.7 |
| | | 27.9 (3) | | | |
| | | 36.9 (4) | | | |
| Iodo | –I | –7.2 (1,2) | 10.9 | –1.5 | –0.9 |
| | | 3.8 (3) | | | |
| | | 20.8 (4) | | | |
| Ammonium | $-NH_3^+$ | 26.0 | 7.5 | –4.6 | –0.1 |
| Nitrile | –CN | 3.1 | 2.4 | –3.3 | –0.5 |
| Nitrate | $-NO_3$ | 62.0 | 4.4 | –4.0 | 0.0 |
| Peroxy, hydroperoxy | –OO– | 55.0 | 2.7 | –4.0 | 0.0 |
| Oxime, *syn* | –C=NOH | 11.7 | 0.6 | –1.8 | 0.0 |
| Oxime, *anti* | | 16.1 | 4.3 | –1.5 | 0.0 |
| Thiocyanate | –SCN | 21.0 | 7.2 | –4.0 | 0.3 |
| Sulfoxide | –S(O)– | 31.1 | 9.0 | –3.5 | 0.0 |
| Sulfonate | $SO_3H$ | 38.9 | 0.5 | –3.7 | 0.2 |
| Aldehyde | –CHO | 29.9 | –0.6 | –2.7 | 0.0 |
| Ketone | –C(O)– | 22.5 | 3.0 | –3.0 | 0.0 |
| Acid | –COOH | 20.1 | 2.0 | –2.8 | 0.0 |
| Carboxylate | $-COO^-$ | 24.5 | 3.5 | –2.5 | 0.0 |
| Acyl chloride | –COCl | 33.1 | 2.3 | –3.6 | 0.0 |

**Table 4** Continued

| *Substituent* | | $\alpha^b$ | $\beta$ | $\gamma$ | $\delta$ |
|---|---|---|---|---|---|
| Ester | –C(O)O– | 22.6 | 2.0 | –2.8 | 0.0 |
| | –OC(O)– | 54.5 (1,2,3) | 6.5 | –6.0 | 0.0 |
| | | 62.5 (4) | | | |
| Amide | –CONH– | 22.0 | 2.6 | –3.2 | –0.4 |
| | –NHCO– | 28.0 | 6.8 | –5.1 | 0.0 |
| Olefin[a] | C=C | 21.5 | 6.9 | –2.1 | 0.4 |
| Acetylene[a] | C≡C | 4.4 | 5.6 | –3.4 | –0.6 |

[a] Steric correction parameters (**Table 5**) apply to these substitutents.

[b] Number(s) in parentheses denotes(s) the number of non-hydrogen substituents on the carbon in question.

**Table 5** Steric correction parameters, $S(i, j)^a$

| *i* | *j = 1* | *j = 2* | *j = 3* | *j = 4* |
|---|---|---|---|---|
| Primary | 0.0 | 0.0 | −1.1 | −3.4 |
| Secondary | 0.0 | 0.0 | −2.5 | −7.5 |
| Tertiary | 0.0 | −3.7 | −9.5 | −15 |
| Quaternary | −1.5 | −8.4 | −15 | −25 |

[a] Designation $i$ = the carbon in question, and $j$ = number of nonhydrogen substituents directly attached to the $\alpha$-substituent (applicable only to $\alpha$-substituents marked with footnote *a* in **Table 4**).

A sample calculation for 2-ethyl-1-butanol is as follows:

$(CH_3CH_2)_2CH–CH_2OH$

| | $C_1$ | $C_2$ | $C_3$ | $C_4$ |
|---|---|---|---|---|
| Base | –2.3 | –2.3 | –2.3 | –2.3 |
| $\alpha$ | 9.1 | 18.2 | 27.3 | 58.1 |
| $\beta$ | 9.4 | 18.8 | 28.9 | 18.8 |
| $\gamma$ | –5.0 | –8.5 | 0 | –5.0 |
| $\delta$ | +0.6 | 0 | 0 | 0 |
| *S*(2,3) | 0 | –2.5 | 0 | –2.5 |
| *S*(3,2) | 0 | 0 | –11.1 | 0 |
| Calculated | 11.8 | 23.7 | 42.8 | 67.1 |
| Observed | 11.1 | 23.0 | 43.6 | 64.6 |

art that depends on experience and good memory. The shift scale in **Figure 9**, coupled with spectral intensities and other chemical information, often permits simple structures to be deciphered. For more complex molecules, one common method involves looking up in spectral libraries for either the compound in question or, if not available, compounds with similar structures. This process has been automated, and many computer-assisted structure-determination methods have been developed. Several research groups have been very active in advancing this important area.

Another commonly used method involves the empirical shift rules described earlier. For alkyl carbons, the $^{13}$C shifts can be approximated by a linear combination of additive terms related to the neighbouring functionalities:

$$\delta_{\text{observed}} = k_0 + n_\alpha S_\alpha + n_\beta S_\beta + n_\gamma S_\gamma + n_\delta S_\delta + n_\varepsilon S_\varepsilon + S_c \quad [3]$$

where $n_i$ refers to the number of $i$ neighbouring carbons ($i = \alpha, \beta, \gamma, \delta$, and $\varepsilon$), $S_i$ is a constant characteristic of the $i$th carbon, and $S_c$ represents steric corrective terms to be used for contiguous secondary, tertiary, and quaternary carbons. One set of additive parameters is shown in **Table 4**. The term $n_\varepsilon S_\varepsilon$ is often neglected because it is usually insignificantly small. The steric correction parameters $S_c$ are given in **Table 5**.

Computer programs have been written that automate the rather tedious arithmetic involved. A recent program provides predicted $^{13}$C shifts and spectra of organic compounds containing common functional groups by incorporating most of the reported activity rules.

In practice, spectral assignment of more complex structures or unknown compounds often requires a combination of assignment methodologies, with or without computer assistance and experimental techniques. Thus, for a given problem it is not unusual to obtain the single-pulse spectrum, the APT, INEPT, and/or DEPT spectrum of a sample. If needed, 2D spectra are also obtained. The information gathered is used, together with shift-prediction or spectral-search methods, to determine the structure of the sample in question.

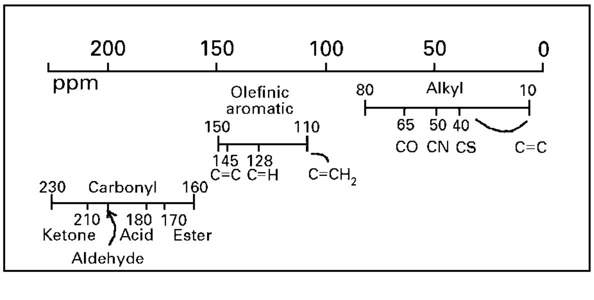



**Figure 9** A simplified chemical-shift scale for carbon.

## Summary

Carbon NMR is a valuable tool for the study of organic materials. The NMR parameters are directly correlated with the structural properties of the various carbon sites. Modern spectrometers make carbon spectroscopy on samples containing $^{13}$C feasible at natural abundance. With the development of modern NMR technology, a host of techniques may be applied to obtain a wide variety of information about the molecule under study. In particular, spectral-editing experiments can provide information on carbon–proton connectivity that aid in the assignment of resonances and the elucidation of structure.

## List of symbols

$f_i$ = frequency; $I_0$ =signal intensity without NOE; $I_{eq}$ = signal enhancement; $J$ = coupling constant; $k_0$ = base chemical shift; $n$ = number of neighbouring carbon atoms; $S$ = signal; $t$ = time; $T_1$ = spin–lattice relaxation time; $T_R$ = repolarization time; $\alpha_E$ = Ernst angle; $\gamma$ = magnetogyric ratio; $\delta$ = chemical shift.

*See also:* **$^{13}$C NMR, Parameter Survey; NMR Pulse Sequences; NMR Relaxation Rates; NMR Spectrometers; NQR, Applications; Nuclear Overhauser Effect; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals; Two-Dimensional NMR, Methods.**

## Further reading

Bax A (1982) *Two-Dimensional Nuclear Magnetic Resonance in Liquids*. Dordrecht: Delft University Press.

Becker ED (1980) *High-Resolution NMR. Theory and Chemical Applications*, 2nd edn. New York: Academic Press.

Carabedian M, Dagane I, Dubois J-E (1988) Elucidation of progressive intersection of ordered substructures for carbon-13 nuclear magnetic resonance. *Analytical Chemistry* 60: 2186–2192, and references therein.

Carbon-13 Database, Bio-Rad Sadtler Division, 3316 Spring Garden Street, Philadelphia, PA 19104, USA.

Cheng HN and Bennett MA (1991) Trends in shift rules in carbon-13 nuclear magnetic resonance spectroscopy and computer-aided shift prediction. *Analytica Chimica Acta* **242**: 43–53.

Cheng HN and Ellingsen SJ (1983) Carbon-13 nuclear magnetic resonance spectral interpretation by a computerized substituent chemical shift method. *Journal of Chemical Information and Computer Science* **23**: 197.

Cheng HN and Kasehagen LJ (1994) Integrated approach for $^{13}$C nuclear magnetic resonance shift prediction, spectral simulation and library search. *Analytica Chimica Acta* **285**: 223–235.

Clerc JT and Sommerauer H (1977) A minicomputer program based on additivity rules for the estimation of $^{13}$C NMR chemical shifts. *Analytica Chimica Acta* **95**: 33–40.

Clerc JT, Pretsch E and Sternhell S (1973) *$^{13}$C-Kernresonanz-spektroskopie*. Frankfurt: Akademische Verlagsgesellschaft.

Dodrell DM, Pegg DT and Bendall MR (1983) Correspondence between INEPT and DEPT pulse sequences for coupled spin-half nuclei. *Journal of Magnetic Resonance* **51**: 264–269.

Ernst RR (1966) Nuclear magnetic double resonance with an incoherent radio-frequency field. *Journal of Chemical Physics* **45**: 3845.

Ernst RR, Bodenhausen G and Wokaun A (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford: Clarendon Press.

Fürst A and Pretsch E (1990) A computer program for the prediction of $^{13}$C-NMR chemical shifts of organic compounds. *Analytica Chimica Acta* **229**: 17–25.

Hearmon RA (1986) Microcomputer-based $^{13}$C NMR spectral simulation of substituted alkanes. *Magnetic Resonance in Chemistry* **24**: 995–998.

INKA Database, Scientific Information Service, Larchmont, NY.

Levitt M, Freeman R and Frenkiel T (1983) Broadband decoupling in high-resolution nuclear magnetic resonance spectroscopy. In: Waugh JS (ed) *Advances in Magnetic Resonance*, Vol 11, pp 48–110. New York: Academic Press.

Lindley MR, Gray NAB, Smith DH, Djerassi C (1982) Applications of artificial intelligence for chemical inference. 40. Computerized approach to the verification of carbon-13 nuclear magnetic resonance spectral assignments. *Journal of Organic Chemistry* **47**: 1027–1035.

Martin GE and Zektzer AS (1988) *Two-Dimensional NMR Methods for Establishing Molecular Connectivity*. New York: VCH.

Morris GA and Freeman R (1979) Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* **101**: 760–762.

Munk ME and Christie BD (1989) The characterization of structure by computer. *Analytica Chimica Acta* **216**: 57–75, and references therein.

Nakanishi K (ed) (1990) *One-Dimensional and Two-Dimensional NMR Spectra by Modern Pulse Techniques*. Tokyo: Kodansha.

Noggle JH and Schirmer RS (1971) *The Nuclear Overhauser Effect*. New York: Academic Press.

Paul EG and Grant DM (1964) Carbon-13 magnetic resonance II. Chemical shift data for the alkanes. *Journal of the American Chemical Society* **86**: 2984–2990.

Pretsch E, Clerc JT, Seibl J and Simon W (1989) *Tables of Spectral Data for Structure Elucidation of Organic Compounds*, 2nd edn. Berlin: Springer.

SPECINFO Database, STN International, 2540 Olentangy River Rd., Columbus, OH 43210, USA.

Takahashi Y, Maeda S and Sasaki S (1987) Automated recognition of common geometrical patterns among a variety of three-dimensional molecular structures. *Analytica Chimica Acta* **200**: 363–377, and references therein.

# $^{13}$C NMR, Parameter Survey

**R Duncan Farrant**, GlaxoWellcome R&D, Stevenage, UK
**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

## Introduction

High-resolution $^{13}$C NMR spectroscopy is in widespread and routine use in many laboratories as a tool for the identification of organic compounds. This article is intended to provide a compilation of the chemical shifts and coupling constants available from $^{13}$C NMR spectra and to show how these are related to molecular structures. $^{13}$C NMR of organic molecules in solution is particularly powerful because $^{13}$C chemical shifts, and spin–spin coupling involving $^{13}$C nuclei generally fall into well defined ranges. Comprehensive tabulations of such values have been compiled (see Further reading) and more recently a number of computer-based predictive schemes have become available.

**Table 1** Estimation of the $^{13}$C chemical shifts in aliphatic compounds

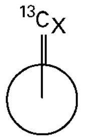
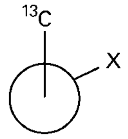
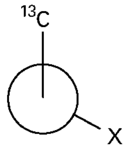
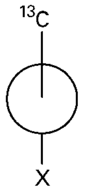
| *Substituent* | *Increment $Z_i$ for substituents in position* $\alpha$ | $\beta$ | $\gamma$ | $\delta$ |
|---|---|---|---|---|
| –H | 0.0 | 0.0 | 0.0 | 0.0 |
| –C<* | 9.1 | 9.4 | –2.5 | 0.3 |
| –C=C–* | 19.5 | 6.9 | –2.1 | 0.4 |
| –C≡C– | 4.4 | 5.6 | –3.4 | –0.6 |
| –Ph | 22.1 | 9.3 | –2.6 | 0.3 |
| –F | 70.1 | 7.8 | –6.8 | 0.0 |
| –Cl | 31.0 | 10.0 | –5.1 | –0.5 |
| –Br | 18.9 | 11.0 | –3.8 | –0.7 |
| –I | –7.2 | 10.9 | –1.5 | –0.9 |
| –O–* | 49.0 | 10.1 | –6.2 | 0.3 |
| –O–CO– | 56.5 | 6.5 | –6.0 | 0.0 |
| –O–NO– | 54.3 | 6.1 | –6.5 | –0.5 |
| –N<* | 28.3 | 11.3 | –5.1 | 0.0 |
| $-N^+$<* | 30.7 | 5.4 | –7.2 | –1.4 |
| $-NH_3^+$ | 26.0 | 7.5 | –4.6 | 0.0 |
| $-NO_2$ | 61.6 | 3.1 | –4.6 | –1.0 |
| –NC | 31.5 | 7.6 | –3.0 | 0.0 |
| –S–* | 10.6 | 11.4 | –3.6 | –0.4 |
| –S–CO– | 17.0 | 6.5 | –3.1 | 0.0 |
| –SO–* | 31.1 | 7.0 | –3.5 | 0.5 |
| $-SO_2-$* | 30.3 | 7.0 | –3.7 | 0.3 |
| $-SO_2Cl$ | 54.5 | 3.4 | –3.0 | 0.0 |
| –SCN | 23.0 | 9.7 | –3.0 | 0.0 |
| –CHO | 29.9 | –0.6 | –2.7 | 0.0 |
| –CO– | 22.5 | 3.0 | –3.0 | 0.0 |
| –COOH | 20.1 | 2.0 | –2.8 | 0.0 |
| $-COO^-$ | 24.5 | 3.5 | –2.5 | 0.0 |
| –COO– | 22.6 | 2.0 | –2.8 | 0.0 |
| –CON< | 22.0 | 2.6 | –3.2 | –0.4 |
| –COCl | 33.1 | 2.3 | –3.6 | 0.0 |
| –CS–N< | 33.1 | 7.7 | –2.5 | 0.6 |
| –C=NOH *syn*, | 11.7 | 0.6 | –1.8 | 0.0 |
| –C=NOH *anti* | 16.1 | 4.3 | –1.5 | 0.0 |
| –CN | 3.1 | 2.4 | –3.3 | –0.5 |

Steric corrections S:

| *Observed $^{13}$C centre* | *Number of substituents other than H at the $\alpha$–atom[a]* 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Primary ($CH_3$) | 0.0 | 0.0 | –1.1 | –3.4 |
| Secondary ($CH_2$) | 0.0 | 0.0 | –2.5 | –6.0 |
| Tertiary (CH) | 0.0 | –3.7 | –8.5 | –10.0 |
| Quaternary (C) | –1.5 | –8.0 | –10.0 | –12.5 |

[a] For $\alpha$-substituents marked with asterisks above.

*Conformation corrections K for $\gamma$-substitutents:*

| *Conformation* | | *K* |
|---|---|---|
| Synperiplanar |  | −4.0 |
| Synclinal |  | −1.0 |
| Anticlinal |  | 0.0 |
| Antiperiplanar |  | 2.0 |
| Not fixed | – | 0.0 |

From Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral Data for Structure Determination of Organic Compounds*, 2nd edn. Reproduced with permission from Springer-Verlag.

**Table 2** $^{13}$C Chemical shifts for methyl groups

| *Substituent X* | $\delta_{CH_3-X}$ | *Substituent X* | $\delta_{CH_3-X}$ | *Substituent X* | $\delta_{CH_3-X}$ | *Substituent X* | $\delta_{CH_3-X}$ |
|---|---|---|---|---|---|---|---|
| –H | –2.3 | –cyclopentyl | 20.5 | –CH$_2$COCH$_3$ | 7.0 | –Ocy | 55.1 |
| –CH$_3$ | 7.3 | –cy | 23.1 | –CH$_2$COOH | 9.6 | –OCH=CH$_2$ | 52.5 |
| –CH$_2$CH$_3$ | 15.4 | –CH=CH$_2$ | 18.7 | –OPh | 54.8 | –SO$_2$CH$_2$CH$_3$ | 39.3 |
| –CH(CH$_3$)$_2$ | 24.1 | –C≡CH | 3.7 | –OCOOCH$_3$ | 54.9 | –SO$_2$Cl | 52.6 |
| –C(CH$_3$)$_3$ | 31.3 | –phenyl | 21.4 | –OCOCH$_3$ | 51.5 | –SO$_3$H | 39.6 |
| –(CH$_2$)$_6$CH$_3$ | 14.1 | – $\alpha$–naphthyl | 19.1 | –OCOcy | 51.2 | –SO$_3$Na | 41.1 |
| –CH$_2$Ph | 15.7 | –$\beta$ –naphthyl | 21.5 | –OCOCH=CH$_2$ | 51.5 | –CHO | 31.2 |
| –CH$_2$F | 15.8 | –2–pyridyl | 24.2 | –OCOPh | 51.8 | –COCH$_3$ | 30.7 |
| –CH$_2$Cl | 18.7 | –3–pyridyl | 18.0 | –OSO$_2$OCH$_3$ | 59.1 | –COCH$_2$CH$_3$ | 27.5 |
| –CH$_2$Br | 19.1 | –4–pyridyl | 20.6 | –NH$_2$ | 28.3 | –COCCl$_3$ | 21.1 |
| –CH$_2$I | 20.4 | –2–furyl | 13.7 | –NH$_3^+$ | 26.5 | –COCH=CH$_2$ | 25.7 |
| –CHCl$_2$ | 31.6 | –2–thienyl | 14.7 | –NHCH$_3$ | 38.2 | –COcy | 27.6 |
| –CHBr$_2$ | 31.8 | –2–pyrrolyl | 11.8 | –NHcy | 33.5 | –COPh | 25.7 |
| –CCl$_3$ | 46.3 | –2–indolyl | 13.4 | –NHPh | 30.2 | –COOH | 21.7 |
| –CBr$_3$ | 49.4 | –3–indolyl | 9.8 | –N(CH$_3$)$_2$ | 47.5 | –COO– | 24.4 |
| –CH$_2$OH | 18.2 | –4–indolyl | 21.6 | –N–pyrrolidinyl | 42.7 | –COOCH$_3$ | 20.6 |
| –CH$_2$OCH$_3$ | 14.7 | –5–indolyl | 21.5 | –N–piperidinyl | 47.7 | –COOCOCH$_3$ | 21.8 |
| –CH$_2$OCH$_2$CH$_3$ | 15.4 | –6–indolyl | 21.7 | –N(CH$_3$)Ph | 39.9 | –CONH$_2$ | 22.3 |
| –CH$_2$OCH=CH$_2$ | 14.6 | –7–indolyl | 16.6 | –1–pyrrolyl | 35.9 | –CON(CH$_3$)$_2$ | 21.5 |
| –CH$_2$OPh | 14.9 | –F | 71.6 | –1–imidazolyl | 32.2 | –COSH | 32.6 |
| –CH$_2$OCOCH$_3$ | 14.4 | –Cl | 25.6 | –1–pyrazolyl | 38.4 | –COSCH$_3$ | 30.2 |
| –CH$_2$NH$_2$ | 19.0 | –Br | 9.6 | –1–indolyl | 32.1 | –COCOCH$_3$ | 23.2 |
| –CH$_2$NHCH$_3$ | 14.3 | –I | –24.0 | –NHCOCH$_3$ | 26.1 | –COCl | 33.6 |
| –CH$_2$N(CH$_3$)$_2$ | 12.8 | –OH | 50.2 | –N(CH$_3$)CHO | 36.5; 31.5 | –CN | 1.7 |
| –CH$_2$NO$_2$ | 12.3 | –OCH$_3$ | 60.9 | –NC | 26.8 | –SC$_8$H$_{17}$ | 15.5 |
| –CH$_2$SH | 19.7 | –OCH$_2$CH$_3$ | 57.6 | –NCS | 29.1 | –SPh | 15.6 |
| –CH$_2$SO$_2$CH$_3$ | 6.7 | –OCH(CH$_3$)$_2$ | 54.9 | –NO$_2$ | 61.2 | –SSCH$_3$ | 22.0 |
| –CH$_2$SO$_3$H | 8.0 | –OC(CH$_3$)$_3$ | 49.4 | –SH | 6.5 | –SOCH$_3$ | 40.1 |
| –CH$_2$CHO | 5.2 | –OCH$_2$CH=CH$_2$ | 57.4 | –SCH$_3$ | 19.3 | –SO$_2$CH$_3$ | 42.6 |

From Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral Data for Structure Determination of Organic Compounds*, 2nd edn. Reproduced with permission from Springer-Verlag.

**Table 3** Effect of a substituent on the $^{13}$C chemical shifts in vinyl compounds

| *Substituent X* | $z_1$ | $z_2$ | *Substituent X* | $z_1$ | $z_2$ |
|---|---|---|---|---|---|
| –H | 0.0 | 0.0 | –OCH$_3$ | 29.4 | –38.9 |
| –CH$_3$ | 12.9 | –7.4 | –OCH$_2$CH$_3$ | 28.8 | –37.1 |
| –CH$_2$CH$_3$ | 17.2 | –9.8 | –OCH$_2$CH$_2$CH$_2$CH$_3$ | 28.1 | –40.4 |
| –CH$_2$CH$_2$CH$_3$ | 15.7 | –8.8 | –OCOCH$_3$ | 18.4 | –26.7 |
| –C(CH$_3$)$_2$ | 22.7 | –12.0 | –N(CH$_3$)$_2$ | 28.0* | –32.0* |
| –CH$_2$CH$_2$CH$_2$CH$_3$ | 14.6 | –8.9 | –N$^+$(CH$_3$)$_3$ | 19.8 | –10.6 |
| –C(CH$_3$)$_3$ | 26.0 | –14.8 | –N–pyrrolidonyl | 6.5 | –29.2 |
| –CH$_2$Cl | 10.2 | –6.0 | –NO$_2$ | 22.3 | –0.9 |
| –CH$_2$Br | 10.9 | –4.5 | –NC | –3.9 | –2.7 |
| –CH$_2$I | 14.2 | –4.0 | –SCH$_2$Ph | 18.5 | –16.4 |
| –CH$_2$OH | 14.2 | –8.4 | –SO$_2$CH=CH$_2$ | 14.3 | 7.9 |
| –CH$_2$OCH$_2$CH$_3$ | 12.3 | –8.8 | –CHO | 15.3 | 14.5 |
| –CH=CH$_2$ | 13.6 | –7.0 | –COCH$_3$ | 13.8 | 4.7 |
| –C=CH | –6.0 | 5.9 | –COOH | 5.0 | 9.8 |
| –Ph | 12.5 | –11.0 | –COOCH$_2$CH$_3$ | 6.3 | 7.0 |
| –F | 24.9 | –34.3 | –COCl | 8.1 | 14.0 |
| –Cl | 2.8 | –6.1 | –CN | –15.1 | 14.2 |
| –Br | –8.6 | –0.9 | –Si(CH$_3$)$_3$ | 16.9 | 6.7 |
| –I | –38.1 | 7.0 | –SiCl$_3$ | 8.7 | 16.1 |

From Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral Data for Structure Determination of Organic Compounds*, 2nd edn. Reproduced with permission from Springer-Verlag.

# Chemical shifts

## Alkanes

$^{13}$C NMR chemical shifts are referenced to that of tetramethylsilane (TMS) added as an internal standard and taken as 0.0 ppm. Often secondary standards are used and one common approach is to use the $^{13}$C NMR resonance of the organic solvent. For the two commonest NMR solvents, these have values relative to TMS of $CDCl_3$ at 77.5 ppm and dimethyl sulfoxide-$d_6$ at 39.5 ppm.

The $^{13}$C chemical shift of methane is at –2.3 ppm relative to TMS and from this base value it is possible to calculate shifts for other alkanes (and even substituted alkanes provided that the appropriate base shifts for the substituted methane are available) using the increment data given in **Table 1** together with the equation

$$\delta C = -2.3 + \sum Z_i + \sum S_j + \sum K_k \qquad [1]$$

where $Z$ denotes the substituent effects, $s$ values are included to take into account steric effects and $K$ allows for conformations of $\gamma$-substituents.

Methyl groups can have a fairly large range of shift values and these can also be predicted using substituent effects as shown in **Table 2**. Rules have been given by Grant and co-workers for calculating $^{13}$C chemical shifts of methyl and ring carbons in cyclohexanes.

The $^{13}$C NMR chemical shifts for cyclohexanes can be calculated using similar parameters to those for the aliphatic compounds above.

Much study has been devoted to halogenated alkanes including fluorinated compounds and predictive rules have been derived.

## Alkenes

The $^{13}$C shifts for carbons of double bonds generally range from 80 to 160 ppm and they can again be estimated from empirical rules based on substituent effects. Thus the shifts of double bond carbons can be estimated from the equation

$$\delta = 123.3 + z_i \qquad [2]$$

where $z_1$ is an increment for substitution on the *ipso* carbon and $z_2$ is the effect of a substituent on the vicinal carbon. The substituent effects are given in **Table 3**.

Carbons involved in keto–enol tautomerism can experience large chemical shift differences according to the tautomer present. For example, for acetylacetone the carbonyl carbon resonates at 201.1 ppm in the keto form but is at 190.5 ppm in the enol form. Similarly, the carbon which is a methylene group in the keto form with a shift of 56.6 ppm moves to 99.0 ppm as an olefinic carbon in the enol form.

Ring strain can induce significant effects on $^{13}$C shifts. For example, the olefine carbons in cyclopropene are at 108.7 ppm, in cyclobutene they are at 137.2 ppm, in cyclopentene they appear at 130.8 ppm and in cyclohexene at 127.4 ppm.

Conjugation can also have large effects. The inner olefine carbon in butadiene is at 136.9 ppm, whilst the central carbon in allene is at 213.5 ppm with the outer carbon at 74.8 ppm.

## Alkynes

Acetylene has a $^{13}$C shift of 71.9 ppm and the effects of substituents have been evaluated. These are given in **Table 4**.

## Aromatic hydrocarbons

Benzene has a chemical shift of 128.5 ppm and the 1-, 2- and 4a-positions of naphthalene have shifts of 128.0, 126.0 and 133.7 ppm, respectively. Well

**Table 4** $^{13}$C Chemical shifts in alkynes ($H{-}C_b{\equiv}C_a{-}X$)

| *X* | *a* | *b* |
|---|---|---|
| –H | 71.9 | 71.9 |
| $-CH_3$ | 80.4 | 68.3 |
| $-CH_2CH_3$ | 85.5 | 67.1 |
| $-CH_2CH_2CH_3$ | 84.0 | 68.7 |
| $-CH_2CH_2CH_2CH_3$ | 83.0 | 66.0 |
| $-CH(CH_3)_2$ | 89.2 | 67.6 |
| $-C(CH_3)_3$ | 92.6 | 66.8 |
| –cy | 88.7 | 68.3 |
| $-CH_2OH$ | 83.0 | 73.8 |
| $-CH{=}CH_2$ | 82.8 | 80.0 |
| $-C{\equiv}C{-}CH_3$ | 68.8 | 64.7 |
| –Ph | 84.6 | 78.3 |
| $-OCH_2CH_3$ | 88.2 | 22.0 |
| $-SCH_2CH_3$ | 72.6 | 81.4 |
| –CHO | 81.8 | 83.1 |
| $-COCH_3$ | 81.9 | 78.1 |
| –COOH | 74.0 | 78.6 |
| $-COOCH_3$ | 74.8 | 75.6 |

From Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral Data for Structure Determination of Organic Compounds*, 2nd edn. Reproduced with permission from Springer-Verlag.

**Table 5** Effect of substituents on the $^{13}$C chemical shifts in monosubstituted benzenes

| *Substituent X* | $z_1$ | $z_2$ | $z_3$ | $z_4$ | *Substituent X* | $z_1$ | $z_2$ | $z_3$ | $z_4$ |
|---|---|---|---|---|---|---|---|---|---|
| –H | 0.0 | 0.0 | 0.0 | 0.0 | $–N^+(CH_3)_3$ | 19.5 | –7.3 | 2.5 | 2.4 |
| $–CH_3$ | 9.2 | 0.7 | –0.1 | –3.0 | $–NHCOCH_3$ | 9.7 | –8.1 | 0.2 | –4.4 |
| $–CH_2CH_3$ | 15.7 | –0.6 | –0.1 | –2.8 | $–NHNH_2$ | 22.8 | –16.5 | 0.5 | –9.6 |
| $–CH(CH_3)_2$ | 20.2 | –2.2 | –0.3 | –2.8 | –N=N–Ph | 24.0 | –5.8 | 0.3 | 2.2 |
| $–CH_2CH_2CH_2CH_3$ | 14.2 | –0.2 | –0.2 | –2.8 | $–N^+{\equiv}N$ | –12.7 | 6.0 | 5.7 | 16.0 |
| $–C(CH_3)_3$ | 22.4 | –3.3 | –0.4 | –3.1 | –NC | –1.8 | –2.2 | 1.4 | 0.9 |
| –cyclopropyl | 15.1 | –3.3 | –0.6 | –3.6 | –NCO | 5.1 | –3.7 | 1.1 | –2.8 |
| $–CH_2Cl$ | 9.3 | 0.3 | 0.2 | 0.0 | –NCS | 3.0 | –2.7 | 1.3 | –1.0 |
| $–CH_2Br$ | 9.5 | 0.7 | 0.3 | 0.2 | –NO | 37.4 | –7.7 | 0.8 | 7.0 |
| $–CF_3$ | 2.5 | –3.2 | 0.3 | 3.3 | $–NO_2$ | 19.9 | –4.9 | 0.9 | 6.1 |
| $–CCl_3$ | 16.3 | –1.7 | –0.1 | 1.8 | –SH | 2.1 | 0.7 | 0.3 | –3.2 |
| $–CH_2OH$ | 12.4 | –1.2 | 0.2 | –1.1 | $–SCH_3$ | 10.0 | –1.9 | 0.2 | –3.6 |
| $–CHOCH_3$ | 9.2 | –3.1 | –0.1 | –0.5 | $–SC(CH_3)_3$ | 4.5 | 9.0 | –0.3 | 0.0 |
| $–CH_2NH_2$ | 14.9 | –1.4 | –0.2 | –2.0 | –SPh | 7.3 | 2.5 | 0.6 | –1.5 |
| $–CH_2SCH_3$ | 9.8 | 0.4 | –0.1 | –1.6 | $–SOCH_3$ | 17.6 | –5.0 | 1.1 | 2.4 |
| $–CH_2SOCH_3$ | 0.8 | 1.5 | 0.4 | –0.2 | $–SO_2CH_3$ | 12.3 | –1.4 | 0.8 | 5.1 |
| $–CH_2CN$ | 1.6 | 0.5 | –0.8 | –0.7 | $–SO_2Cl$ | 15.6 | –1.7 | 1.2 | 6.8 |
| $–CH{=}CH_2$ | 8.9 | –2.3 | –0.1 | –0.8 | $–SO_3H$ | 15.0 | –2.2 | 1.3 | 3.8 |
| $–C{\equiv}CH$ | –6.2 | 3.6 | –0.4 | –0.3 | $–SO_2OCH_3$ | 6.4 | –0.6 | 1.5 | 5.9 |
| –Ph | 13.1 | –1.1 | 0.5 | –1.1 | –SCN | –3.7 | 2.5 | 2.2 | 2.2 |
| –F | 34.8 | –13.0 | 1.6 | –4.4 | –CHO | 8.2 | 1.2 | 0.5 | 5.8 |
| –Cl | 6.3 | 0.4 | 1.4 | –1.9 | $–COCH_3$ | 8.9 | 0.1 | –0.1 | 4.4 |
| –Br | –5.8 | 3.2 | 1.6 | –1.6 | $–COCF_3$ | –5.6 | 1.8 | 0.7 | 6.7 |
| –I | –34.1 | 8.9 | 1.6 | –1.1 | –COPh | 9.3 | 1.6 | –0.3 | 3.7 |
| –OH | 26.9 | –12.8 | 1.4 | –7.4 | –COOH | 2.1 | 1.6 | –0.1 | 5.2 |
| –ONa | 39.6 | –8.2 | 1.9 | –13.6 | $–COO^-$ | 9.7 | 4.6 | 2.2 | 4.6 |
| $–OCH_3$ | 31.4 | –14.4 | 1.0 | –7.7 | $–COOCH_3$ | 2.0 | 1.2 | –0.1 | 4.3 |
| $–OCH{=}CH_2$ | 28.2 | –11.5 | 0.7 | –5.8 | $–CONH_2$ | 5.0 | –1.2 | 0.1 | 3.4 |
| –OPh | 27.6 | –11.2 | –0.3 | –6.9 | $–CON(CH_3)_2$ | 8.0 | –1.5 | –0.2 | 1.0 |
| $–OCOCH_3$ | 22.4 | –7.1 | 0.4 | –3.2 | –COCl | 4.7 | 2.7 | 0.3 | 6.6 |
| $–OSi\ (CH_3)_3$ | 26.8 | –8.4 | 0.9 | –7.1 | –CSPh | 18.7 | 1.0 | –0.6 | 2.4 |
| –OCN | 25.0 | –12.7 | 2.6 | –1.0 | –CN | –15.7 | 3.6 | 0.7 | 4.3 |
| $–NH_2$ | 18.2 | –13.4 | 0.8 | –10.0 | $–P(CH_3)_2$ | 13.6 | 1.6 | –0.6 | –1.0 |
| $–NHCH_3$ | 21.4 | –16.2 | 0.8 | –11.6 | $–P(Ph)_2$ | 8.9 | 5.2 | 0.0 | 0.1 |
| $–N(CH_3)_2$ | 22.5 | –15.4 | 0.9 | –11.5 | $–PO(OCH_2CH_3)_2$ | 1.6 | 3.6 | –0.2 | 3.4 |
| –NHPh | 14.7 | –10.6 | 0.9 | –10.5 | $–SiH_3$ | –0.5 | 7.3 | –0.4 | 1.3 |
| $–N(Ph)_2$ | 19.8 | –7.0 | 0.9 | –5.6 | $–Si(CH_3)_3$ | 11.6 | 4.9 | –0.7 | 0.4 |
| $–NH_3^+$ | 0.1 | –5.8 | 2.2 | 2.2 | | | | | |

From Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral Data for Structure Determination of Organic Compounds*, 2nd edn. Reproduced with permission from Springer-Verlag.

defined substituent effects on aromatic carbon shifts in benzene derivatives have been derived. These are calculated according to the equation

$$\delta = 128.5 + \sum z_i \qquad [3]$$

and the substituent effects for *ortho*, *meta* and *para* positions are listed in **Table 5**. Similar data compilations are available for naphthalene derivatives and for pyridines. The naphthalene substituent parameters can be used with a reasonable degree of success in other condensed ring hydrocarbons provided that the base values for a particular ring system are available. These have been given for a number of ring systems, including benzofuran, benzthiophen, benzimidazole, quinoline, isoquinoline and many others (see Further reading).

The $^{13}$C shifts for purine and pyrimidine bases have also been measured and these form useful base values for studies of nucleosides, nucleotides and nucleic acids.

### Oxygen-containing compounds

Alcohols show large chemical shifts according to the proximity of the oxygen to the measured carbon. Thus *n*-propanol has shifts of 64.2, 25.9 and 10.3 ppm for the carbons α, β and γ to the OH, respectively. Ethylene glycol has a shift of 63.4 ppm, whilst glycerol has shifts of 64.5 and 73.7 ppm for the $CH_2$ and CH carbons, respectively. Trifluoroethanol, often used in protein and peptide NMR studies, has carbon shifts of 61.4 and 125.1 ppm for the $CH_2$ and $CF_3$ carbons, respectively, with $^1J_{CF} = 278$ Hz and $^2J_{CCF} = 35$ Hz (see later). Ethers show similar substituent effects to alcohols. For cyclic ethers, ring strain again induces large chemical shift changes. Thus for ethylene oxide the $CH_2$ shift is 39.5 ppm whereas for the four-membered ring analogue the shift is 72.6 ppm.

Extensive data compilations are available for monosaccharides and other sugars, including the tabulation by Bock.

### Amines

The protonation of amines causes a shielding of carbons adjacent to the nitrogen except for branched systems, where some deshielding is often seen. The effect of protonation can be about –2 ppm at an α-carbon, –3 ppm at a β-carbon and up to –1 ppm at a γ-carbon. For example, the $^{13}$C shifts of ethylamine are 36.9 ppm ($CH_2$) and 19.0 ($CH_3$) and these are shifted by –0.2 and –5.0 ppm, respectively, on protonation. Again, ring strain effects in cyclic amines are mirrored by large $^{13}$C shift effects.

### Amino acids

$^{13}$C NMR data for amino acids have been tabulated extensively as these provide base values for studies of NMR spectra of peptides and proteins. The data compilation by Pretsch and co-workers is comprehensive, as are the results shown in the book by Wuthrich. Values are very pH dependent.

### Other functional groups

Nitrile carbon shifts are in the range of 115–125 ppm whereas in isonitriles the shifts are around 155–165 ppm. In oximes, the carbon in the C=N bond appears around 150–160 ppm and in isocyanates the carbon shift is in the region of 120–125 ppm. For hydrazones with the functionality C=N–N, the $^{13}$C shifts are usually around 165 ppm. Methylene carbons adjacent to nitro groups generally appear near 70–80 ppm, and for methylene groups in nitrosoamines these appear around 45–55 ppm.

The SH group induces much smaller shift changes. The $CH_2$ resonance of carbons adjacent to a thiol group appear around 24–35 ppm according to the presence of other substituents and chain branching. Similar effects are seen in thioethers. Sulfoxides of the type RR′SO are chiral because of the pyramidal structure of the sulfoxide group and hence nonequivalence can be seen in the NMR spectra. Methylene carbons adjacent to sulfone groups appear around 55–65 ppm.

### Carbonyl carbons

For the fragment, Cβ–Cα–CO–Cα–Cβ, it is possible to estimate the carbonyl carbon chemical shift from the substituent effects given in **Table 6A** using the equation

$$\delta = 193.0 + \sum z_i \qquad [4]$$

For carboxylic acids and esters, Equation [4] and **Table 6B** apply with a base shift of 166.0 ppm. For amides, the base shift is 165.0 ppm and the substituent effects in **Table 6C** are used.

**Table 6** Additivity rules for estimating the chemical shift for carbonyl groups

*(A) Aldehydes and ketones:*

| *Substituent i* | $z_\alpha$ | $z_\beta$ |
|---|---|---|
| $-\mathrm{C}\lessgtr$ | 6.5 | 2.6 |
| $-CH{=}CH_2$ | –0.8 | 0.0 |
| $-CH{=}CH{-}CH_3$ | 0.2 | 0.0 |
| –Ph | –1.2 | 0.0 |

*(B) Carboxylic acids and esters:*

| *Substituent i* | $z_\alpha$ | $z_\beta$ | $z_\gamma$ | $z_{\alpha'}$ |
|---|---|---|---|---|
| $-\mathrm{C}\lessgtr$ | 11.0 | 3.0 | –1.0 | –5.0 |
| $-CH{=}CH_2$ | 5.0 | | | –9.0 |
| –Ph | 6.0 | 1.0 | | –8.0 |

*(C) Amides:*

| *Substituent i* | $z_\alpha$ | $z_\beta$ | $z_\gamma$ | $z_{\alpha'}$ | $z_{\beta'}$ |
|---|---|---|---|---|---|
| $-\mathrm{C}\lessgtr$ | 7.7 | 4.5 | –0.7 | –1.5 | –0.3 |
| $-CH{=}CH_2$ | 3.3 | | | | |
| –Ph | 4.7 | | | –4.5 | |

From Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral data for Structure Determination of Organic Compounds*, 2nd edn. Reproduced with permission from Springer-Verlag.

Aldehyde carbons have a similar range of chemical shifts as ketones; for example, the shift for cyclohexane-1-aldehyde is 204.7 ppm. Cyclic ketones tend to be more deshielded than alicyclic ketones. For example, methyl ethyl ketone has a carbonyl shift of 207.6 ppm but cyclohexanone is at 209.7 ppm. The carbonyl group of lactones is highly shielded relative to other carbonyl carbons, with an upfield shift of around 20 ppm. Similarly, amide and lactam carbons are in the region of 165–175 ppm. The carbonyl carbons of anhydrides are also similarly shielded.

## Coupling constants

### One-bond $J_{CH}$

An extensive review of these is now available (see Further reading). In general, these lie in the range 125–170 Hz but there are a number of values beyond this upper figure. The value can be estimated in substituted methanes of the type $CHZ_1Z_2Z_3$ using Equation [5] and the substituent effects in **Table 7**.

$$J_{CH} = 125.0 + \sum z_i \qquad [5]$$

### Two-bond $J_{CH}$

These couplings can be small and vary considerably. For example, for the fragment $^{13}$C–C–H, $J_{CH}$ is 1–6 Hz, for $^{13}$C=C–H it is 0–16 Hz and for $^{13}$C–(C=O)–H it is 20–50 Hz.

### Three-bond $J_{CH}$

These generally lie in the range 0–10 Hz but they depend on the dihedral angle in similar fashion to proton–proton coupling constants. Thus, when the $^{13}$C and $^{1}$H are in bonds which are eclipsed, the coupling is around 6 Hz, when they are antiperiplanar, the coupling is about 9 Hz and when the bonds are at right-angles, the coupling drops to close to zero.

In alkenes, the three-bond *trans* coupling is always larger than the *cis* coupling for a pair of *cis–trans* isomers. Typical values are 7–10 Hz for a *cis* coupling and 8–13 Hz for a *trans* coupling.

Three-bond couplings in aromatic molecules are generally around 7–10 Hz.

### Other couplings involving $^{13}$C

One-bond $^{13}$C–$^{13}$C couplings have a typical value of around 30–40 Hz in unstrained, unsubstituted systems. Ring strain causes a reduction in coupling constant, the value in cyclopropane, for example, being only 12.4 Hz. Couplings in conjugated systems tend to be higher, an extreme example being acetylene with a value of 171.5 Hz. $^{13}$C–$^{13}$C coupling through two or three bonds can also be measured. Values are typically 1–5 Hz.

$^{13}$C–$^{19}$F coupling constants can also often be measured. The one-bond coupling is around 280–300 Hz depending on substituents, with the two-bond coupling also being of considerable magnitude, around 20 Hz. Longer range $^{13}$C–$^{19}$F couplings are also observed, especially in aromatic and conjugated molecules. In fluorobenzene $^{1}J_{CH}$ is 245.1 Hz, $^{2}J_{CH}$ is 21.0 Hz, $^{3}J_{CH}$ is 7.8 Hz and $^{4}J_{CH}$ is 3.2 Hz.

In addition, a number of studies have made use of $^{13}$C–$^{15}$N and $^{13}$C–$^{31}$P coupling constants.

**Table 7** Additivity rule for estimating the $^{13}$C–$^{1}$H coupling constants in aliphatic compounds

| *Substituent* | *Increments $z_i$* |
|---|---|
| –H | 0.0 |
| $–CH_3$ | 1.0 |
| $–C(CH_3)_3$ | –3.0 |
| $–CH_2Cl$ | 3.0 |
| $–CH_2Br$ | 3.0 |
| $–CH_2I$ | 7.0 |
| $–CHCl_2$ | 6.0 |
| $–CCl_3$ | 9.0 |
| –C≡CH | 7.0 |
| –Ph | 1.0 |
| –F | 24.0 |
| –Cl | 27.0 |
| –Br | 27.0 |
| –I | 26.0 |
| –OH | 18.0 |
| –OPh | 18.0 |
| $–NH_2$ | 8.0 |
| $–NHCH_3$ | 7.0 |
| $–N(CH_3)_2$ | 6.0 |
| $–SOCH_3$ | 13.0 |
| –CHO | 2.0 |
| $–COCH_3$ | –1.0 |
| –COOH | 5.5 |
| –CN | 11.0 |

From Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral Data for Structure Determination of Organic Compounds*, 2nd edn. Reproduced with permission from Springer-Verlag.

## List of symbols

*J* = coupling constant; *K* = conformations of γ-substituents; *s* = steric effects; *Z* = substituent effects; δ = chemical shift.

*See also:* **Parameters in NMR Spectroscopy, Theory of; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Peptides; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals.**

## Further reading

Bock K and Pedersen C (1983) $^{13}$C NMR spectroscopy of monosaccharides. *Advances in Carbohydrate Chemistry and Biochemistry* **41**: 27–66.

Bock K, Pedersen C and Pedersen H (1984) $^{13}$C NMR spectroscopy of oligosaccharides. *Advances in Carbohydrate Chemistry and Biochemistry* **42**: 193–225.

Bruhl TS, Heilmann D and Kleinpeter E (1997) $^{13}$C NMR chemical shift calculations for some substituted pyridines. A comparative consideration. *Journal of Chemical Information and Computer Science* **37**: 726–730.

Dalling DK and Grant DM (1972) Carbon-13 magnetic resonance. XXI. Steric interactions in the methylcyclohexanes. *Journal of the American Chemical Society* **94**: 5318–5324.

Grant DM and Paul EG (1964) Carbon-13 magnetic resonance: chemical shift data for the alkanes. *Journal of the American Chemical Society* **86**: 2984–2995.

Hansen PE (1981) Carbon-hydrogen spin–spin coupling constants. *Progress in Nuclear Magnetic Spectroscopy* **14**: 175–296.

Kalinowski H-O, Berger S and Braun S (1988) *Carbon-13 NMR Spectroscopy*. New York: Wiley.

Ovenall DW and Chang JJ (1977) Carbon-13 NMR of fluorinated compounds using wide-band fluorine decoupling. *Journal of Magnetic Resonance* **25**: 361–372.

Pathre S (1996) Drawing structures and calculating $^{13}$C NMR spectra-ACD/CNMR. *Analytical Chemistry* **68**: A740–A741.

Pretsch E, Simon W, Seibl J and Clerc T (1989) *Spectral Data for Structure Determination of Organic Compounds*, 2nd edn. Berlin: Springer.

Sarneski JE, Surprenant HL, Molen FK and Reilley CN (1975) Chemical shifts and protonation shifts in carbon-13 nuclear magnetic resonance studies of aqueous amines. *Analytical Chemistry* **47**: 2116–2124.

Van Bramer S (1997) ACD/CMR and ACD/HNMR spectrum prediction software. *Concepts in Magnetic Resonance* **9**: 271–273.

Wuthrich K (1986) *NMR of Proteins and Nucleic Acids*. New York: Wiley.

# Cadmium NMR, Applications

*See* **Heteronuclear NMR Applications (Y–Cd).**

# Caesium NMR Spectroscopy

*See* **NMR Spectroscopy of Alkali Metal Nuclei in Solution.**

# Calibration and Reference Systems (Regulatory Authorities)

**C Burgess**, Burgess Consultancy, Barnard Castle, Co Durham, UK

Copyright © 1999 Academic Press

**FUNDAMENTALS OF SPECTROSCOPY**

**Methods & Instrumentation**

## Overview

Spectroscopic measurements are at the heart of many analytical procedures and processes. Proper and adequate calibration and qualification of spectrometers is an essential part of the process for ensuring data integrity. Wherever practicable, the standards used should be traceable to a recognized national or international standard. The International Organization for Standardization (ISO) 9000 series of guides are the international standard for quality systems. Quality of measurement is central to compliance: for example ISO 9001 Section 4.11 requires that:

> ...the user shall identify, calibrate and adjust all inspection, measuring and test equipment and devices that can affect product quality at prescribed intervals, or prior to use, against certified equipment having a known valid relationship to nationally recognised standards.

More detailed information on the compliance requirements for the competence of calibration and testing laboratories is contained in ISO Guide 25. This guide has been interpreted to meet national requirements in some countries: for example the UK Accreditation Service refers to it as UKAS M10. Many laboratories seek accreditation to ISO Guide 25 and operate quality and testing systems which meet the requirements. The ISO Guide 25 approach heavily focuses on good analytical practices and adequate calibration of instruments with nationally or internationally traceable standards wherever possible.

In addition to the ISO approach, environmental measurements (for example by the US Environmental Protection Agency) and the pharmaceutical industry are subject to rigorous regulatory requirements. The pharmaceutical industry has placed great emphasis on method validation in, for example, HPLC (see Further Reading section for details). However, until recently, there has been little specific regulatory requirement for assuring that the analytical instruments are working properly.

The US Food and Drug Administration (FDA) specifically requires that:

> Laboratory controls shall include: ...
>
> The calibration of instruments, apparatus, gauges, and recording devices at suitable intervals in accordance with an established written program containing specific directions, schedules, limits for accuracy and precision, and provisions for remedial action in the event accuracy and/or precision limits are not met. Instruments, apparatus, gauges, and recording devices not meeting established specifications shall not be used.

The major regulatory guidelines for Good Manufacturing Practice (GMP) and Good Laboratory Practice (GLP) are similarly vague. 'Fitness for purpose' is the phrase that is commonly used, but what does this mean in practice?

Only the Pharmacopoeias (for example the US Pharmacopoeia (1995) and the British Pharmacopoeia (1998)), and the Australian Regulatory Authority have been sufficiently worried by instrumental factors to give written requirements for instrument performance. Whilst these guidelines are not entirely consistent at least they are attempting to ensure consistency of calibration practices between laboratories.

However, there has been a resurgence of interest in the underlying data quality by regulatory authorities particularly the FDA following a major legal ruling involving Barr Laboratories, Inc. in 1993. This legal case has changed the regulatory focus and put the laboratory firmly in the spotlight in terms of the assurance of the quality of the data it produces.

Up to this point we have focused on the spectrometer and the underlying data quality. However from an analytical science perspective, process is important as illustrated in **Figure 1**. Scientific knowledge is founded upon relevant and robust information generated from reliable data.

The overall analytical process is illustrated in **Figure 2**.

Quality has to be built in at all stages of the process, for failure to ensure integrity at any level invalidates results from further processing. One of the key areas not covered by this article is the validation of the software and systems involved in instrument control and data transformation. Clearly this must be addressed as part of the overall assessment for the 'fitness for purpose' of any computerized spectrometer.

**Table 1** American Society for Testing and Materials (ASTM) standards relating to spectrometry and spectrometer performance

| *ASTM Standard* | *Title* |
|---|---|
| E 131-95 | Standard Terminology Relating to Molecular Spectroscopy |
| E 168-92 | Standard Practices for General Techniques of Infrared Quantitative Analysis |
| E 169-93 | Standard Practices for General Techniques of Ultraviolet-Visible Quantitative Analysis |
| E 275-93 | Standard Practice for Describing and Measuring Performance of Ultraviolet, Visible and Near-Infrared Spectrophotometers |
| E 386-90 | Standard Practice for Data Presentation Relating to High Resolution Nuclear Magnetic Resonance (NMR) Spectroscopy |
| E 387-84 | Standard Test Method for Estimating Stray Radiant Power Ratio of Spectrophotometers by the Opaque Filter Method |
| E 573-96 | Standard Practices for Internal Reflection Spectroscopy |
| E 578-83 | Standard Test Method for Linearity of Fluorescence Measuring Systems |
| E 579-84 | Standard Test Method for Limit of Detection of Fluorescence of Quinine Sulphate |
| E 925-94 | Standard Practice for the Periodic Calibration of Narrow Band-pass Spectrophotometers |
| E 932-93 | Standard Practice for Describing and Measuring Performance of Dispersive Infrared Spectrometers |
| E 958-93 | Standard Practice for Measuring Practical Spectral Bandwidth of Ultraviolet-Visible Spectrophotometers |
| E 1252-94 | Standard Practice for General Techniques of Qualitative Infrared Analysis |
| E 1421-94 | Standard Practice for Describing and Measuring Performance of Fourier Transform-Infrared Spectrometers; Level Zero and Level One Tests |
| E 1642-94 | Standard Practices for General Techniques of Gas Chromatography Infrared (GC-IR) Analysis |
| E 1655-94 | Standard Practices for Infrared, Multivariate Quantitative Analysis |
| E 1683-95a | Standard Practice for Testing the Performance of Scanning Raman Spectrometers |
| E 1790-96 | Standard Practice for Near Infrared Qualitative Analysis |
| E 1791-96 | Standard Practice for Transfer Standards for Reflectance Factor for Near-Infrared Instruments Using Hemispherical Geometry |
| E 1840-96 | Standard Guide for Raman Shift Standards for Spectrometer Calibration |
| E 1865-97 | Standard Guide for Open-path Fourier Transform Infrared (OP/FT-IR) Monitoring of Gases and Vapours in Air |
| E 1866-97 | Standard Guide for Establishing Spectrophotometer Performance Tests |

Note: These standards are available from ASTM, 100 Barr Harbor Drive, West Conshohocken, PA 19428, USA

**Table 2** Summary of pharmacopoeial requirements for the calibration of IR and NIR spectrometers

| *Monograph* | *Authority* | *Test* | *Standard* | *Measurement process* | *Acceptance criteria* |
|---|---|---|---|---|---|
| IR | Ph. Eur. 1997 Method 2.2.24 BP 1998 IIA | Resolution | 0.05 mm polystyrene film 0.038 mm polystyrene film | % transmittance differences at 2870 and 2871 $cm^{-1}$ and at 1589 and 1583 $cm^{-1}$ | Greater than 18 and 12 respectively |
| | Ph. Eur 1997. Method 2.2.24 | Wavenumber accuracy | 0.05 mm polystyrene film | Location of transmittance minima<br>3027.1 (± 0.3) 1583.1 (± 0.3)<br>2924 (± 2) 1181.4 (± 0.3)<br>2850.7 (± 0.3) 1154.3 (± 0.3)<br>1994 (± 1) 1069.1 (± 0.3)<br>1871.0 (± 0.3) 1028.0 (± 0.3)<br>1801.0 (± 0.3) 906.7 (± 0.3)<br>1601.4 (± 0.3) 698.9 (± 0.5) | Not specified. The uncertainties quoted in the table reflect the uncertainities in the quoted values and not the acceptance criteria for the spectrometer |
| | USP 23 (851) 1995 | Wavenumber accuracy | Polystyrene, water vapour, carbon dioxide, ammonia gas | Location of transmittance minima | Not specifed |
| NIR | Ph. Eur. 1997 Method 2.2.40 | Wavelength scale accuracy | Polystyrene or rare earth oxides | Location of absorbance or reflectance minima | Not specified |
| | | Wavelength scale repeatability | Polystyrene or rare earth oxides | Location of absorbance or reflectance minima | Standard deviation is consistent with spectrometer specification |
| | | Response scale repeatability | Reflective thermoplastic resins doped with carbon black | Measurement of reflectance values | Standard deviation is consistent with spectrometer specification |
| | | Photometric noise | Reflective thermoplastic resins doped with carbon black or reflective ceramic tile | Peak to peak or at a given wavelength | The photometric noise is consistent with spectrometer specification |

**Table 3** Summary of pharmacopoeial requirements for the calibration of UV-visible spectrometers and polarimeters

| *Monograph* | *Authority* | *Test* | *Standard* | *Measurement process* | | *Acceptance criteria* | |
|---|---|---|---|---|---|---|---|
| UV-VIS | Ph. Eur. 1997 Method 2.2.25 | Wavelength scale accuracy | 4 % holmium oxide in 1.4 M perchloric acid (Ho) or atomic lines from deuterium (D), hydrogen (H) or mercury vapour arc (Hg) | Location of absorbance maxima | | ±1 nm 200 to 400 nm ± 3 nm 401 to 700 nm | |
| | | | | 241.15 nm (Ho) | 404.66 nm (Hg) | | |
| | | | | 253.7 nm (Hg) | 435.83 nm (Hg) | | |
| | | | | 287.15 nm (Ho) | 486.0 nm (Dβ) | | |
| | | | | 302.25 nm (Hg) | 486.1 nm (Hβ) | | |
| | | | | 313.16 nm (Hg) | 536.3 nm (Ho) | | |
| | | | | 334.15 nm (Hg) | 546.07 nm (Hg) | | |
| | | | | 316.5 nm (Ho) | 576.96 nm (Hg) | | |
| | | | | 365.48 nm (Hg) | 579.07 nm (Hg) | | |
| | USP 23 ⟨851⟩ 1995 | | NIST SRM[a] 2034, holmlum oxide solution | As required by certificate | | ±1 nm | |
| | Ph. Eur. 1997 Method 2.2.25 | Absorbance scale accuracy | 57.0 to 63.0 mg of potassium dichromate per litre of 0.005 M sulfuric acid solution | **Wavelength nm** | **A (1%, 1 cm)** | **Wavelength nm** | **Tolerance** |
| | | | | 235 | 124.0 | 235 | 122.9 to 126.2 |
| | | | | 257 | 144.0 | 257 | 142.4 to 145.7 |
| | | | | 313 | 48.6 | 313 | 47.0 to 50.3 |
| | | | | 350 | 106.6 | 350 | 104.9 to 109.2 |
| | USP 23 ⟨851⟩ | Absorbance scale accuracy | NIST SRM 931 liquid filters NIST SRM 930 glass filters | As required by certificate | | Not specified but implied within NIST range of values | |
| | | Resolution | 0.02 % v/v toluene in hexane | Absorbance ratio at the maximum at 269 nm to the minimum at 266 nm | | Greater than 1.5 | |
| | | Stray light | 1.2 % w/v potassium chloride in water | Absorbance of 10 mm pathlength at 200 ± 2 nm | | Greater than 2 | |
| | BP 1998 IIB | Second derivative | 0.020% solution of toluene in methanol | Record the second derivative spectrum between 255 nm and 275 nm for a 10 mm path length | | A small negative extremum located between two large negative extrema at about 261 and 268 nm should be clearly visible | |
| Polarimeter | Ph. Eur. 1997 Method 2.2.7 | Accuracy of rotation | Quartz plate | Not specified | | Not specified | |
| | USP 23 ⟨781⟩ 1995 | | NIST traceable quartz plate | Not specified | | Not specified but implied within NIST range of values | |
| | Ph. Eur. 1997 Method 2.2.7 USP 23 ⟨781⟩ 1995 | Linearity of scale | Succrose solutions | Not specified | | Not specified | |
| | USP 23 ⟨781⟩ 1995 | | NIST SRM 17 sucrose NIST SRM 41 dextrose | Not specified | | Not specified but implied within NIST range of values | |

[a] National Institute for Science and Technology, Standard Reference Material (USA).

**Table 4** Summary of pharmacopoeial requirements for the calibration of NMR and CD spectrometer

| *Monograph* | *Authority* | *Test* | *Standard* | *Measurement process* | *Acceptance criteria* |
|---|---|---|---|---|---|
| NMR | Ph. Eur. 1997 Method 2.2.33 | Resolution | 20% v/v 1,2 dichlorobenzene in deuterated acetone or 5% v/v tetramethylsilane in deuterochloroform | Peak width at half height of the band at δ7.33 ppm or δ7.51 ppm of the symmetrical multiplet for dichlorobenzene or | Less than or equal to 0.5 Hz |
| | | | | Peak width at half height of the band at δ 0.00 ppm for tetramethylsilane. | |
| | | Signal to noise ratio | 1% v/v ethylbenzene in carbon tetrachloride | Measure the spectrum over range δ2 to 5 ppm. Measure the peak amplitude, A, of the largest peak of the methylene quartet centred at δ 2.65 ppm and the peak to peak amplitude of the baseline noise between δ 4 and δ 5 ppm. The amplitude is measured from a baseline constructed from the centre of the noise on either side of the quartet and at a distance of at least 1 ppm from its centre. | The S/N ratio is given by 2.5 ($A/H$) which has to be greater than 25:1 for a mean of five measurements |
| | | Side band amplitude | | | Amplitude of spining side bands is not greater than 2% of the sample amplitude at the rotational speed used |
| | | Integrator response repeatability | 5% v/v ethylbenzene in carbon tetrachloride | For quantitative measurements, carry out five successive scans of the protons of the phenyl and ethyl groups and determine the mean values. | No Individual value shall differ by more than 2.5% from the mean |
| CD | Ph. Eur. 1997 Method 2.2.41 | Absorbance accuracy | 1.0 mg $ml^{-1}$ isoandrosterone | Record the spectrum between 280 and 360 nm and determine the minimum at 304 nm | $\Delta\varepsilon$ is +3.3 |
| | | Linearity of modulation | 1.0 mg $ml^{-1}$ (1S)-(+)-camphorsulfonic acid | Record the spectrum between 185 and 340 nm and determine the minima at 192.5 and 290.5 nm | At 192.5 nm, $\Delta\varepsilon$ is −4.3 to −5<br>At 290.5 nm, $\Delta\varepsilon$ is +2.2 to +2.5 |

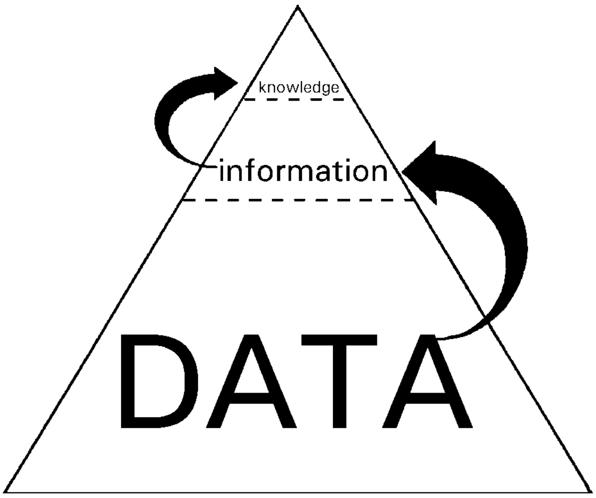



**Figure 1** Data, information and knowledge triangle.

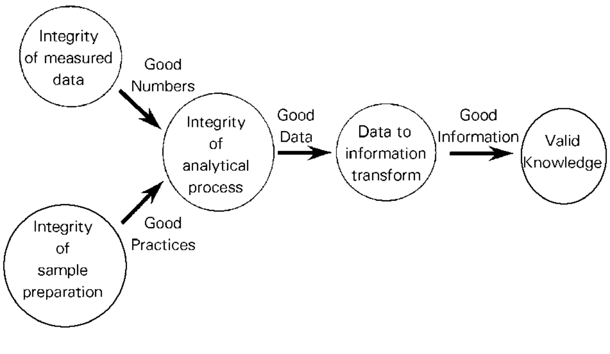



**Figure 2** Analytical process integrity.

This article is restricted to the calibration requirements for spectrometers within the regulatory context which from a practical standpoint means the major pharmacopoeias.

## Spectroscopic coverage

Current spectrometric monographs include calibration requirements for: infrared (IR), near infrared (NIR), ultraviolet and visible (UV-visible), nuclear magnetic resonance (NMR), circular dichroism and polarimetry. Monographs for atomic spectrophotometry, emission and absorption, fluorescence and X-ray fluorescence are available in the European Pharmacopoeia but do not specify calibration requirements other than that they are to be operated in accordance with the manufacturer's instructions. A series of ASTM standards are available and are listed in **Table 1**. Whilst primarily intended for use by instrument vendors and for instrumental fitness for purpose for ASTM analytical procedures, they are sometimes useful in supplementing pharmacopoeial guidance.

**Tables 2–4** summarize the current requirements of the European Pharmacopoeia (Ph. Eur.), the British Pharmacopoeia (BP) and the United States Pharmacopoeia (USP).

*See also:* **Food and Dairy Products, Applications of Atomic Spectroscopy; Food Science, Applications of Mass Spectrometry; Food Science, Applications of NMR Spectroscopy; Laboratory Information Management Systems (LIMS); Pharmaceutical Applications of Atomic Spectroscopy.**

## Further reading

NIS 45, Edition 2, May 1996, Accreditation for Chemical Laboratories.

Environmental Protection Agency, 2185-Good Automated Laboratory Practices, 1995.

International Conference on Harmonisation Note for guidance on validation of analytical procedures: methodology, 1995, CPMP/ICH/281/95.

FDA Center for Drug Evaluation and Research, Reviewer Guidance; Validation of Chromatographic Methods, November 1994.

21 CFR 211 SUBPART I LABORATORY CONTROLS §211.160 (b) (4).

e.g. United States Pharmacopoeia 23, 1995, <621> chromatography, <711> dissolution testing, <831> Refractive index, <851> spectrophotometry and light scattering.

e.g. BP 1998, Appendix II; Infrared spectrophotometry, Ultraviolet and Visible spectrophotometry and Pharm. Eur. V.6.19 Absorption Spectrophotometry.

Australian Code of GMP for Therapeutic Goods—Medicinal Products; Appendix D, November 1991; Guidelines for Laboratory Instrumentation.

DEPARTMENT OF HEALTH AND HUMAN SERVICES Food and Drug Administration [Docket No. 93N-0184] Barr Laboratories, Inc.; Refusal to Approve Certain Abbreviated Applications; Opportunity for a Hearing, Vol. 58 No. 102 Friday, May 28, 1993 p 31035 (Notice) 1/1061.

# Carbohydrates Studied By NMR

**Charles T Weller**, School of Biomedical Sciences, University of St Andrews, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

Carbohydrates, the most widely abundant biological molecules, are key components within a wide variety of biological phenomena. Polysaccharides such as glycogen, cellulose and starch have important structural or nutritional roles, but it is the mediation of specific recognition events by carbohydrates that has sparked detailed analyses of structure, conformation and function. The non-invasive, non-destructive nature of the technique is commonly mentioned as the most important reason for analysing carbohydrates by NMR, and is of considerable advantage when dealing with small amounts of precious material. This advantage is, however, offset by the relative insensitivity of the technique when compared with other methods such as mass spectrometry or enzymatic approaches that may be more appropriate for specific problems of sequence determination or compositional analysis. The value of NMR analysis is apparent in the additional data obtainable regarding composition, structure, conformation and mobility necessary for understanding recognition processes. The analysis of carbohydrates by NMR is characterized by a modification of existing techniques to address the particular problems of oligosaccharide spectra, namely poor dispersion, heterogeneity, a relatively high degree of interresidue mobility and a lack of conformational restraints.

This article will attempt to provide an introduction to a large field of study, and readers are encouraged to look to the Further reading section for additional information.

## Sample preparation

In high-field NMR, the use of high quality sample tubes can have a significant impact upon the quality of the spectra; 5 mm tubes are suitable for most work, with 10 mm sample tubes often being used for $^{13}C$ analysis.

For biologically relevant NMR analysis, experiments are generally carried out in aqueous solution. A strong water signal prevents direct analysis of proton signals lying beneath it and, owing to the problems of dynamic range, would reduce the sensitivity with which weak solute signals are detected. If, as is often the case, the exchangeable OH or NH protons are not of interest, then experiments are usually carried out in $^2H_2O$. Repeated dissolution and evaporation or lyophilization is recommended to reduce the proportion of residual protons remaining. In most cases, two dissolutions into >99.8% $^2H_2O$ followed by a final dissolution into >99.95% $^2H_2O$, preferably from a sealed ampoule, should be sufficient. Other precautions, such as pre-wetting pipettes, further rounds of dissolution and drying, or the use of a dry box may be used for greater sensitivity or quality.

Although extremes of pH or pD (outside the range δ 5–8) are to be avoided, buffer solutions are only necessary in two cases: if the molecule is charged, and the charge state is relevant to the study; and in the study of acid-labile carbohydrates, such as sialylated oligosaccharides.

The use of $^1H_2O$ as solvent is necessary in some cases, and requires appropriate water suppression techniques. Care must be taken to ensure that the water is free from impurities, as well as dissolved gases or paramagnetic species. A deuterated solvent to provide an appropriate lock signal should be added, usually 5–10% $^2H_2O$.

Oligosaccharides are not normally soluble in non-aqueous solvents, except for dimethyl sulfoxide (DMSO). Despite the biologically irrelevant environment, this permits the observation of exchangeable protons, and there is also some increase in the proton spectral dispersion. Similar precautions to those noted for the use of $^2H_2O$ should be observed, such as the use of high isotopic purity (>99.96%) solvent, and repeated dissolution to remove exchangeable protons.

To approach optimal line widths, it is often advisable to remove soluble paramagnetic components by passage through a suitable chelator. Degassing to remove dissolved oxygen is also recommended. For acceptable spectra within a reasonable time using a 500 MHz spectrometer, 10 nmol is an approximate lower limit for 1D $^1H$ spectra; for 2D $^1H$ spectra, amounts of 1 μmol are preferable, although spectra with as little as 100 nmol are possible with care.

Sample volumes vary between 0.4–0.7 mL, depending on the size of the spectrometer RF coils. Temperature should be maintained at a constant level, preferably one or two degrees above room temperature for stability. Sample spinning is not necessary, except in cases where resolution is a particular problem. $T_1$ for protons in carbohydrates is of the order of 0.1 to 0.5 s, and recycle delays of 1–1.5 s are usually sufficient.

Referencing of samples is commonly to acetone at δ 2.225 relative to DSS (5,5-dimethylsilapentanesulfonate) at 25°C, or DMSO at 2.5 ppm.

## Analysis of carbohydrate structure by NMR

The problems of carbohydrate structure addressed by NMR can be divided into four parts:

1. What is the composition of the molecule: what monosaccharides are present, are they in furanose or pyranose configurations, and α or β anomeric forms?
2. What is the nature of the linkages between the monosaccharides: at what positions do the substituents occur, and in what sequence?
3. What is the conformation of the molecule: what conformation or family of conformations are populated about the O-glycosidic bonds and hydroxymethyl rotamers?
4. What are the dynamic properties of the molecule?

Whilst the first and second questions can be satisfactorily answered by NMR, they are perhaps better approached in tandem with chemical or enzymatic methods, such as hydrolysis, followed by reduction and analysis of alditol acetates by GC-MS, or digestion with specific glycosidases. Such a concerted approach has been used to successfully determine the composition and configuration of many oligosaccharides and glycans.

The unique advantages of NMR in the analysis of carbohydrate structure are only fully apparent in consideration of points 3 and 4. In contrast to the unbranched polymeric nature of polypeptide and nucleic acid chains, carbohydrates may be branched structures, capable of substitution at several points. The monosaccharide constituents are polymerized in nature by a non-template directed, enzymatic process; the resulting oligosaccharides are often heterogeneous, differing in detail from a consensus structure. NMR is particularly efficient at investigating the solution conformations, and the dynamic properties of such molecules.

## $^1$H NMR of carbohydrates

### Primary analysis and assignment: $^1$H 1D spectra

The $^1$H NMR spectrum of the disaccharide galactopyranose β 1–4 linked to glucopyranose (Gal*p*β1-4Glc*p*, lactose, shown in **Figure 1**) is shown in **Figure 2**. Despite the relative simplicity of the molecule, with only 14 proton signals observable, the spectrum is surprisingly complex. This is due to the relatively small chemical shift dispersion of the non-exchangeable ring protons, resonating between δ 3–4. The small chemical shift dispersal combined with homonuclear spin–spin coupling gives rise in many cases to non-first-order spectra, complicating the measurement of spin-coupling values. This makes the assignment and analysis of even quite small saccharides quite difficult, and as a consequence the use of spectrometers of field strength 400 MHz or above is recommended.

Primarily owing to the electron-withdrawing effect of the ring oxygen, the anomeric (H1) protons give rise to signals lying outside this envelope of ring protons, between δ 4.25–5.5. Since equatorial protons experience a shift of approximately δ –0.5 relative to axial proton signals, α anomeric protons of D-sugars tend to be found between δ 4.9–5.5, and β protons between δ 4.3–4.7. The ring protons of each monosaccharide type give rise to characteristic patterns within the envelope of overlapping ring proton resonances. However, incorporation of a monosaccharide residue into an oligosaccharide will cause changes in these chemical shift patterns. For example, substitution at a given carbon causes the signal arising from the attached proton to change by δ 0.2–0.5.

The glucose residue of the sample shown in **Figure 2** is free to mutarotate between α and β forms. The signals arising from these are easily distinguished from those owing to the galactose H1 by the difference in area. The two glucose resonances have a combined area approximately equal to that of the galactose H1 resonance.

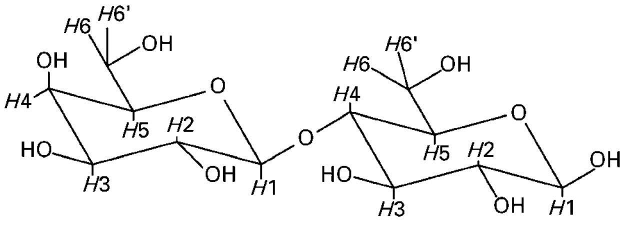



**Figure 1** Structure of the disaccharide Gal*p*β1-4Glc*p*, showing the numbering of protons within the sugar rings.

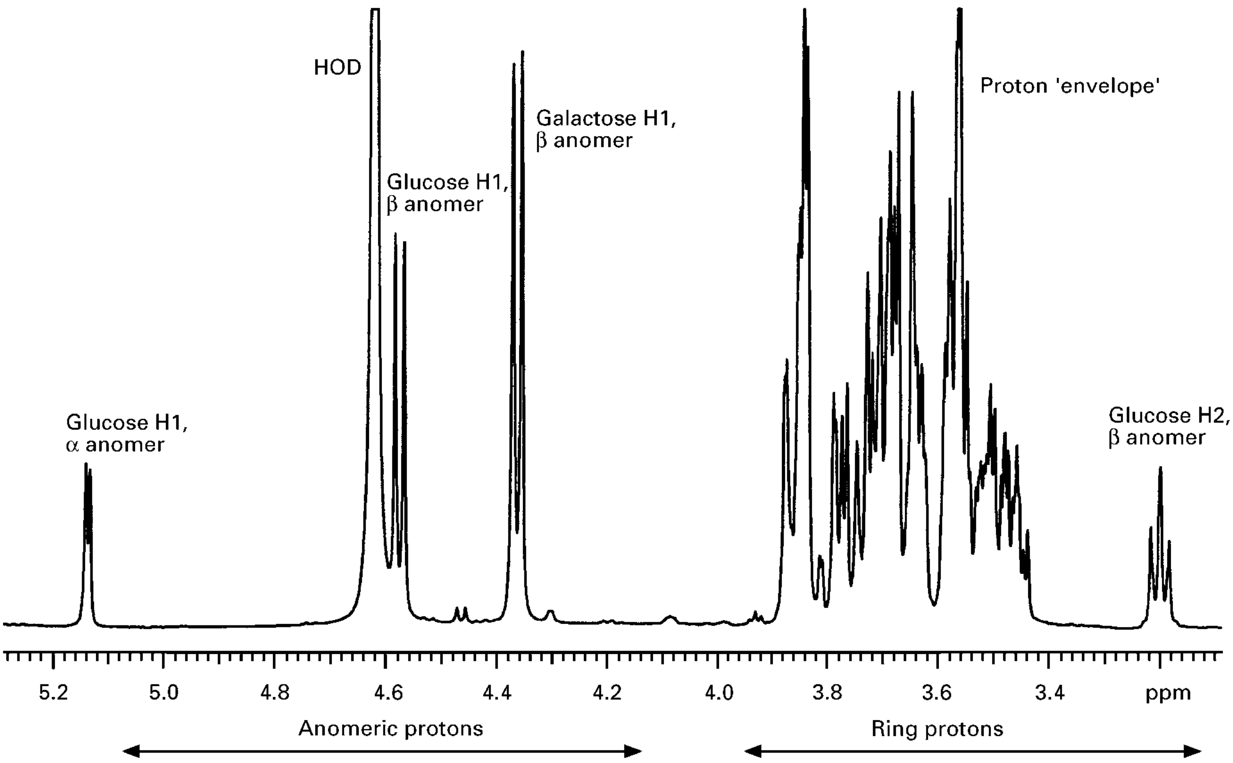



**Figure 2** $^1$H 1D spectrum of the disaccharide galactopyranose β1-4glucopyranose at 500 MHz and 303 K in $^2H_2O$. Peak assignments are given for well-resolved peaks, following the numbering of **Figure 1**. The unresolved proton 'envelope' consists of the remaining ring proton resonances. The peak marked 'HOD' arises from residual water within the sample.

Anomeric signals typically exhibit characteristic doublets, arising from the $^3J_{H\text{-}H}$ H1–H2 coupling, whereas ring protons show more complex multiplet patterns. These couplings are proportional to the dihedral angle between the two protons. Typical values for $^3J_{H\text{-}H}$ in carbohydrates are: axial–axial 7–8 Hz, axial–equatorial and equatorial–equatorial 3–4 Hz.

During the early 1980s, Vliegenthart and coworkers identified a number of 'structural reporter groups', outlined in **Table 1**. Measurement of the chemical shift values, coupling patterns and line widths of signals arising from these elements can be interpreted to aid in the assignment and interpretation of 1D $^1$H NMR spectra.

Sugar composition can thus be identified, and signals assigned on the basis of chemical shift and characteristic coupling values. The type and number of each can be established by consideration of chemical shifts, coupling patterns and relative integrals. Owing to the poor dispersion of the majority of resonances, this is usually limited to resolved anomeric protons or other structural reporter groups. Despite the development of highly efficient experiments based upon the use of selective excitation to produce edited one-dimensional spectra with more manageable information content, full proton assignment usually relies upon two- or, in some cases, three-dimensional techniques.

## Two-dimensional homonuclear analysis

The poor dispersion and strong coupling present in carbohydrate proton spectra present particular problems for the assignment process. In addition, the

**Table 1** Structural reporter groups as described by Vliegenthart and co-workers

| *Sugar type* | *Proton(s)* | *Parameter* | *Information content* |
|---|---|---|---|
| All | H1 | δ, *J* | Residue and linkage type |
| Mannose | H2, H3 | δ | Substitution within core region of branched glycan |
| Sialic acid | H3 | δ | Type and configuration of linkage |
| Fucose | H5, CH3 | δ | Type and configuration of linkage: structural environment |
| Galactose | H4, H4 | δ | Type and configuration of linkage |
| Amino sugars | N-acetyl $CH_3$ | δ | Sensitive to small structural variations |

characteristics of a given spin system differ for each oligosaccharide and are highly dependant on structure. The use of two-dimensional experiments alleviates these problems.

**Coherence transfer methods** The correlated spectroscopy (COSY) experiment reduces the degree of resonance overlap by separating resonances into two orthogonal proton dimensions. The 1D spectrum lies along the diagonal, with cross-peaks joining pairs of *J*-coupled spins. For carbohydrates, assignment can then start with the well-resolved anomeric protons, and continue by stepwise *J*-correlation to the remaining nuclei. This simple process is complicated in many cases by overlap and strong coupling between signals, even in two dimensions. To reduce these difficulties, high quality spectra are needed, with optimal digital resolution.

Sometimes absolute value mode, as opposed to pure phase spectra, can be preferable for easy assignment, despite the theoretical disadvantages of a phase-twist line shape. Carbohydrate line widths are generally narrower than those of proteins and nucleic acids, and the use of absolute value spectra does not significantly diminish the quality of spectra, as seen in **Figure 3**. The COSY spectrum requires a pseudo- echo weighting function to remove dispersive line shapes, reducing the sensitivity of the experiment. If sensitivity is an issue, then an inherently more sensitive experiment, such as the absorption mode DQ-COSY (see below) should be used.

If necessary, linkage positions can be identified using a COSY spectrum of a sample dissolved in DMSO-$d_6$. The hydroxyl protons do not exchange with solvent, and give rise to observable signals within the spectrum. Substitution removes the hydroxyl proton at that position, which can then be identified by the absence of a hydroxyl proton signal.

The COSY experiment may fail to generate the expected cross-peaks, for several reasons:

1. For strongly coupled neighbours, the cross-peak lies close to the diagonal, and may not be seen.
2. Pairs of nuclei with small scalar coupling values ($^3J_{H\text{-}H}$) will produce low intensity cross-peaks. Since the active coupling is antiphase in COSY cross-peaks, they will cancel if the value of *J* is less than the line width. This is particularly

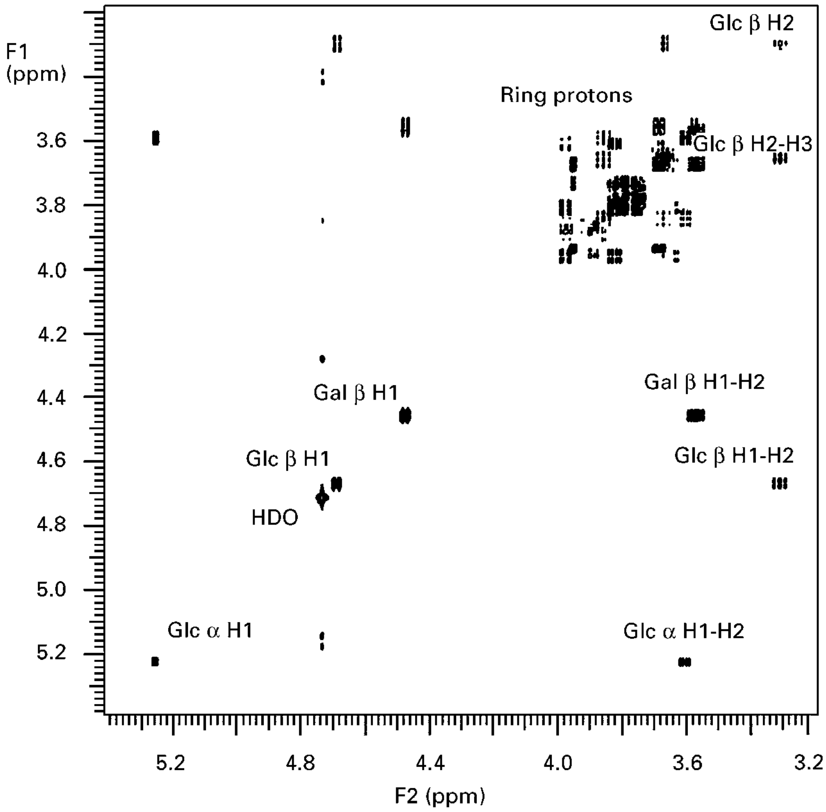



**Figure 3** $^1H$–$^1H$ 2D COSY spectrum of galactopyranose β 1-4glucopyranose at 500 MHz and 303 K in $^2H_2O$. A full proton assignment is usually possible from a spectrum like this, and peak identities are shown where space permits. The well-resolved anomeric protons provide a useful starting point for assignment, and subsequent pairs of spin-coupled nuclei are correlated through off-diagonal peaks.

noticeable in larger molecules. Common examples are: D-mannopyranose β H1–H2 ($J < 1$ Hz) and D-galactopyranose H4–H5 ($J < 1$ Hz).
3. Unrelated resonances may simply occur at the same position, and cannot be distinguished.

**Experiments incorporating isotropic mixing sequences** The homonuclear Hartmann–Hahn (HOHAHA) and the closely related total correlation spectroscopy (TOCSY) experiments are used to overcome problems of overlap and strong coupling. These experiments give correlations to all other spins within the same coupling network. The second pulse in the COSY sequence is replaced by an isotropic mixing period, such as a spin–lock sequence (TOCSY) or an efficient decoupling sequence, such as MLEV-17 (HOHAHA). The spin system becomes strongly coupled and coherence transfer occurs between all nuclei within the coupling network. This illustrates an advantage over experiments such as RELAY-COSY: values of $^3J_{H-H}$ between adjacent nuclei vary, and the use of an isotropic mixing period produces an increased efficiency of transfer. The presence of small couplings between nuclei, reducing the efficiency of transfer, still presents a problem.

Resolution in this type of experiment is better than for a COSY spectrum; however, more cross-peaks are generated, producing a more complicated spectrum. This is at first sight hard to interpret; however, proton assignments in each ring can be determined by inspection of a line lying through the frequency of the anomeric proton of the sugar residue. The cross-peaks found along this line can be distinguished by reference to the COSY spectrum, and by consideration of the multiplet structure. Pyranose rings have an effectively rigid ring geometry, and it is possible to estimate the couplings between adjacent protons based upon the dihedral angle between them. Judicious variation of the length of the isotropic mixing period can also be used to give stepwise correlations along the spin-system, since shorter periods correlate proportionately smaller fragments. One-dimensional analogues of this experiment, using selective pulses to specifically excite anomeric protons, may be quicker and more efficient, especially when the anomeric protons are well resolved.

**Multiple-quantum homonuclear experiment** Incorporation of a multiple-quantum filter into a COSY-type sequence reduces the complexity of the spectrum, and aids in the assignment process.

A DQF-COSY, incorporating a double-quantum filter transparent to two or more coupled spins, therefore passes all signals except singlets. Such a sequence has a slightly reduced sensitivity (cross-peak intensity is approximately half that of a comparable COSY), but there is no need for the sensitivity-reducing weighting functions that are necessary with COSY. One noticeable feature is the absence of a diagonal, thus cross-peaks lying close to the diagonal can be more easily identified.

Incorporation of a triple-quantum filter produces a spectrum with signals due only to three or more coupled spins with two resolvable couplings. There are no singlets or doublets, including anomeric resonances, although suppression is not absolute, due to the presence of small couplings along four or more bonds.

Under ideal conditions, one would expect to see resonances from only H5 and H6 protons, and this experiment is best used in conjunction with HOHAHA for full assignment. Sensitivity is much reduced relative to COSY spectra.

**Through-space dipolar interaction methods** Although primarily used for conformational analysis, the nuclear Overhauser effect spectroscopy (NOESY) experiment is also helpful in the assignment process. The resulting spectra show correlations between protons coupled by through-space dipolar interactions. Both inter- and intraresidue NOEs help to confirm the consistency of assignments made by the techniques described above. Long-range NOEs are not often seen in carbohydrates, and the observation of NOEs between protons on adjacent residues is often evidence of the linkage positions between the two sugars.

Through-space dipolar interactions are of most use in the conformational analysis of biomolecules, as described in the section on conformational analysis below.

At 500 MHz, moderate-sized (more than six residues) oligosaccharides lie within the spin-diffusion limit. However, for smaller molecules, as the value of the function $\omega_0\tau_c$ (where $\omega_0$ is the Larmor frequency of protons, and $\tau_c$ is the correlation time of the molecule) approaches 1 then the value of the NOE tends towards 0. Cross-peak intensities of NOESY spectra of smaller oligosaccharides (2–5 residues) may thus become too small to measure accurately. In such cases, the rotating frame Overhauser effect spectroscopy (ROESY, originally referred to as CAMELSPIN) experiment is commonly used to measure NOE values. To reduce the appearance of HOHAHA-like cross-peaks, a low power spin–lock field should be used, and the transmitter carrier offset to the low-field end of spectrum. The offset dependency of cross-peak intensities should also be removed by 90° pulses at either end of spin–lock period.

Three-dimensional experiments, often essentially hybrids of simpler sequences such as HOHAHA-COSY, have been used successfully to resolve cases of particularly difficult resonance overlap.

### Hydroxyl protons

The quest for additional conformational information has led to the investigation of hydroxyl protons in aqueous solution. Samples are dissolved in mixed methanol–water or acetone–water solvents, and analysed in capillary NMR tubes at low (–5 to –15°C) temperatures. Chemical exchange of hydroxyl protons is reduced to the point that it is possible to use them as probes of hydration and hydrogen bonding. Distance information can also be extracted from NOESY or ROESY spectra under these conditions.

## $^{13}C$ NMR of carbohydrates

### Primary analysis and assignment

Initial investigations of the $^{13}C$ spectra of carbohydrates were of natural abundance samples, although spectra of isotopically-enriched samples have become more common in recent years.

$^{13}C$ spectra are usually acquired with broad band proton decoupling. The resulting spectra are not very complex with, at natural abundance, a single sharp peak for each nucleus, there being no visible carbon–carbon coupling. The value of the carbon–proton coupling is approximately 150–180 Hz, and without proton decoupling this splitting, along with further two- and three-bond couplings, can render the spectrum quite complex.

Monosaccharides exhibit characteristic carbon chemical shift patterns. Anomeric carbons typically occur between δ 90–110 in pyranose rings, C6 (hydroxymethyl) signals between δ 60–75, and the remaining carbons resonate between δ 65–110. Substitution at a given carbon causes an approximate shift of up to δ + 10. The usefulness of this shift in identifying the position of substitution is complicated by an additional δ –1 to –2 shift of the surrounding carbon signals. In addition, the local structure of the saccharide can have a considerable effect upon the chemical shift values. Residue type and substitution patterns can thus be discovered, with care, from measurements of the carbon chemical shifts of a simple sugar. Databases of $^{13}C$ chemical shift information for large numbers of simple carbohydrates are available for comparison in the literature.

**Carbon–proton and carbon–carbon spin couplings**
A consistent estimate of the anomeric configuration is given by the value of the one-bond $^{1}J_{C-H}$ coupling; β anomers show a value of ~160 Hz, and α anomers ~170 Hz.

Three-bond carbon–proton scalar coupling values, such as $^{3}J_{CCCH}$, or $^{3}J_{COCH}$, provide useful indicators of conformation. These couplings can be used to confirm the internal conformation of a saccharide residue, using appropriate Karplus curves. Across the glycosidic link, values of the couplings H1–C1–O–Cx (where x is the aglyconic carbon) and C1–O–Cx–Hx are proportional to the glycosidic dihedral angles $\phi$ and $\psi$ respectively (see section on conformational analysis below). Selective incorporation of $^{13}C$ has been successfully used to simplify and enhance the measurement of these parameters. Linkage positions can be unambiguously determined by the identification of these couplings using experiments such as the 2D $^{13}C$–$^{1}H$ multiple-bond correlation experiment (HMBC). This experiment correlates carbon and proton resonances by means of $^{3}J_{C-H}$ couplings; $^{1}J_{C-H}$ correlations are suppressed, leaving only long-range couplings. Those couplings that span the glycosidic bond can thus be identified and measured.

### Multidimensional, heteronuclear analysis

Enrichment with $^{13}C$ allows the use of a wide range of useful experiments that greatly facilitate the investigation of carbohydrate structure. Proton detected versions of carbon–proton correlation experiments, such as that shown in **Figure 4**, overcome the sensitivity limitations of observing carbon signals. If the proton assignments are known, assignment of carbon resonances from such spectra is a relatively simple process. The use of carbon nuclei to edit spectra into an orthogonal frequency dimension to overcome overlap has led to the development of a series of useful hybrid 3D or pseudo-3D experiments, such as HCCH-COSY or HMQC-NOESY. Editing spectra in this way has helped to assign the spectra not only of complex glycans in free solution, but also of the attached carbohydrate chains of glycoproteins, both at natural abundance as well as enriched with $^{13}C$ and $^{15}N$. The use of these heteronuclei allows the carbohydrate resonances to be observed separately from those due to the protein portion of the molecule.

## Conformational analysis

Measurement of $^{3}J_{H-H}$ couplings shows that pyranose rings do not show any large degree of internal flexibility, except for pendant groups such as the hydroxymethyl in hexopyranoses. Interpretation of these uses the Haasnoot parametrization of the Karplus equation for $^{3}J_{H-H}$ couplings in HCCH

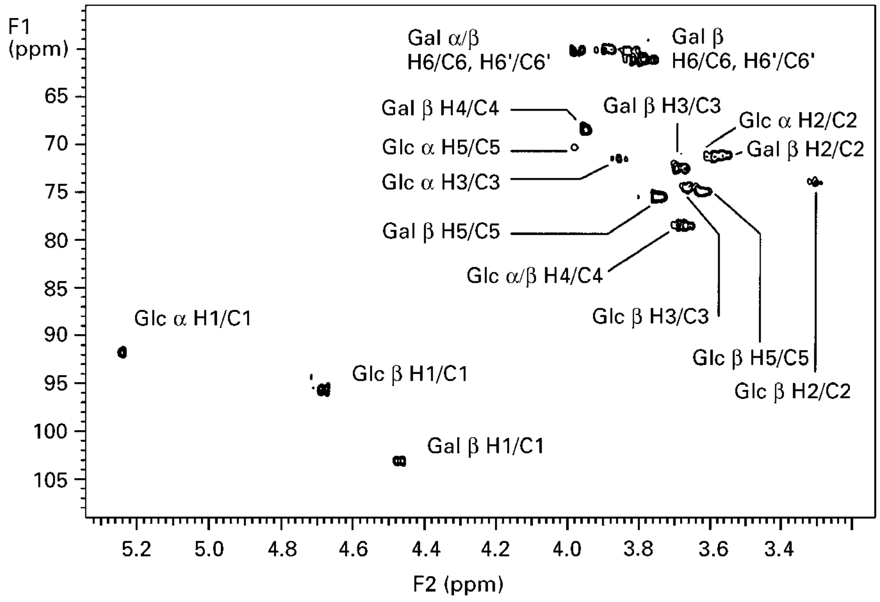



**Figure 4** HSQC $^1H$–$^{13}C$ correlation spectrum of galactopyranose β1-4glucopyranose at 500 MHz and 303 K in $^2H_2O$. Resonances are labelled with the appropriate assignment for each correlation. In practice, these can be relatively easily identified by comparison with previously assigned proton or carbon chemical shifts. Reproduced by courtesy of Homans SW and Kiddle GR, University of St. Andrews.

portions of the saccharide ring:

$$^3J_{H-H'} = 13.22\cos^2\theta - 0.99\cos\theta + \Sigma_\iota\Delta\chi_\iota[0.87 - 2.46\cos^2(\xi_\iota\theta)] \quad [1]$$

where $\theta$ is the HCCH′ torsion angle, $\Delta\chi_\iota$ are the Huggins' electronegativities of the substituents relative to protons, and $\xi$ is either +1 or −1 depending upon the orientation of the substituent. If stereo-specific assignments have been made, orientation about hydroxymethyl rotamers can be determined by similar parametrizations.

Conformational variability in pyranoside oligosaccharides is mostly owing to variations about the glycosidic torsion angles $\phi$ and $\psi$ and, for $1 \rightarrow 6$ linkages, the $\omega$ angle (**Figure 5**). NMR analysis of carbohydrate conformation thus concentrates upon determining the orientations about these dihedral angles. This is made difficult by the lack of conformational information available: even in the most favourable conditions, a maximum of around three NOEs are seen between each pair of residues in oligosaccharides, and observation of long-range correlations between different parts of the molecule is extremely unlikely.

Interresidue correlations in NOESY or ROESY spectra can be used to determine distance information by making use of the fact that the intensity of the cross-peak is proportional to the inverse sixth power of the distance between the two nuclei. The relatively inflexible ring geometry of pyranose rings allows intraresidue NOEs to be used to calibrate adjacent intraresidue correlations. When implementing such experiments it is important to ensure that the initial rate approximation holds. The motion of the molecule, both internal and overall, also affects the intensity of the NOE; thus interpretation of NOE values must take into account the motional model

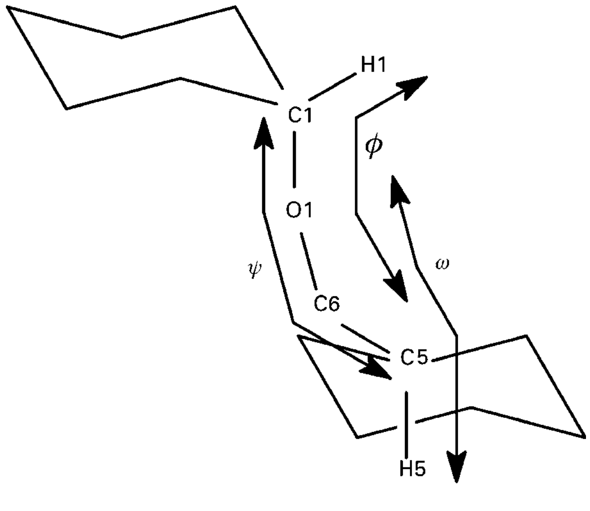



**Figure 5** A $1 \rightarrow 6$ linked disaccharide showing the dihedral angles $\phi$, $\psi$ and $\omega$, about which conformational variation is greatest.

used. Since it is now clear that the majority of carbohydrate structures are flexible to some degree, oscillating through an ensemble of related conformations, this can be a problem. Flexibility is understood to be particularly marked about $1 \to 6$ linkages owing to rotation about the additional dihedral angle, $\omega$.

The use of three-bond ${}^3J_{C\text{–}H}$ couplings as conformational probes has been described, using the Tvaroska parametrization of the Karplus relationship:

$$^3J_{(C,H)} = 5.7\cos^2\theta - 0.6\cos\theta + 0.5 \qquad [2]$$

where $\theta$ is the HCOC torsion angle. However, in a similar situation to that described for NOE restraints, these analyses have been hampered by the extent to which these values are subject to conformational averaging. Because of these problems, angular information derived from ${}^3J_{C\text{–}H}$ values is perhaps better used for evaluating the quality of conformational ensembles.

Recent work has addressed the use of $^{13}C$ within isotopically enriched compounds to derive additional conformational parameters, such as $^{13}C$–$^{13}C$ scalar couplings, and $^{13}C$–$^{1}H$ or $^{13}C$–$^{13}C$ NOEs. These approaches have yielded useful results, but have been hampered by the lack of a well-established Karplus-type relationship for $^{13}C$–$^{13}C$ scalar coupling.

Investigators have thus made considerable use of theoretical methods such as molecular mechanics to complement the limited experimental data available.

### Dynamics

The time-averaged nature of many NMR derived parameters, and the degree to which it affects oligosaccharide conformations has made it necessary to attempt to measure the degree of mobility within oligosaccharides. Since the relaxation of protonated carbons is almost entirely due to the directly attached proton it is possible to measure relaxation parameters from them that are not affected by variations in internuclear distance. Investigation of carbohydrate dynamics has been approached by molecular modelling in combination with the measurement of relaxation rates, including the $T_1$, $T_2$ of $^{13}C$ and $^{1}H$, as well as $^{1}H$–$^{1}H$ and $^{13}C$–$^{1}H$ NOE values, to define the spectral density, given by

$$J(\omega) = (2/5)\tau_c/(1 + \omega\tau_c)^2 \qquad [3]$$

where $\tau_c$ denotes the rotational correlation time. These measurements have been interpreted by the Lipari–Szabo model-free approach, in which molecular motion is separated into overall and internal motions, related to the overall and internal correlation times $\tau_O$ and $\tau_i$:

$$J(\omega) = S^2(2/5)\tau_O/(1 + \omega\tau_O)^2 + (1 - S^2)(2/5)\tau/(1 + \omega\tau)^2 \qquad [4]$$

where $1/\tau = 1/\tau_O + 1/\tau_O$. $S^2$ is a measure of the degree of internal reorientation, varying in value from 0 for isotropic reorientation to 1 for no internal motion.

These measurements can then be used to derive dynamic models of carbohydrates and oligosaccharides. It is now generally accepted that carbohydrates are dynamic molecules with internal motions that occur on a time-scale faster than the overall motion of the molecules.

## List of symbols

$J$ = coupling constant; $J(\omega)$ = spectral density function; $T_1$, $T_2$ = relaxation constants; $\Delta\chi_t$ = Huggins electronegativities; $\theta$ = torsion angle; $\tau_c$, $\tau_i$ and $\tau_O$ = rotational, internal and overall correlation times respectively; $\phi$, $\psi$ and $\omega$ = glycosidic torsion angles.

*See also:* **$^{13}C$ NMR, Methods; Labelling Studies in Biochemistry Using NMR; Nuclear Overhauser Effect; Nucleic Acids Studied By NMR; Proteins Studied Using NMR Spectroscopy; Two-Dimensional NMR, Methods.**

## Further reading


Bush CA (1996) Polysaccharides and complex oligosaccharides. In: Grant DM and Harris RK (eds) *Encyclopaedia of Nuclear Magnetic Resonance*, Vol 6, pp 3746–3750. Chichester: Wiley.

Homans SW (1993) $^{1}H$ NMR studies of oligosaccharides. In: Roberts GCK (ed) *NMR of Macromolecules: a Practical Approach*, pp 289–314. Vol 134 of Rickwood D and Hames BD (series eds) *Practical Approach Series*. Oxford: Oxford University Press.

Homans SW (1993) Conformation and dynamics of oligosaccharides in solution. *Glycobiology*, 3: 551–555.

Hounsell EF (1995) $^{1}H$ NMR in the structural and conformational analysis of oligosaccharides and glycoconjugates. *Progress in Nuclear Magnetic Resonance Spectroscopy* 27: 445–474.

Serianni AS (1992) Nuclear magnetic resonance approaches to oligosaccharide structure elucidation. In: Allen HJ

and Kisalius EC (eds) *Glycoconjugates: Composition, Structure and Function*, pp 71–102. New York: Marcel Dekker.

Tvaroska I (1990) Dependence on saccharide conformation of the one-bond and three-bond coupling constants. *Carbohydrate Research*, **206**: 55–64.

van Halbeek H (1994) NMR developments in structural studies of carbohydrates and their complexes. *Current Opinion in Structural Biology* **4**: 697–709.

van Halbeek H (1996) Carbohydrates and glycoconjugates. In: Grant DM and Harris RK (eds) *Encyclopaedia of Nuclear Magnetic Resonance*, Vol 2, pp 1107–1137. Chichester: Wiley.

Vleigenthart JFG, Dorland L and van Halbeek H (1983) High resolution spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. *Advances in Carbohydrate Chemistry and Biochemistry* **41**: 209–374.

# CD Spectroscopy of Biomacromolecules

*See* **Biomacromolecular Applications of Circular Dichroism and ORD.**

# Cells Studied By NMR

**Fátima Cruz** and **Sebastián Cerdán**, Instituto de Investigaciones Biomédicas, Madrid, Spain

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| **Applications** |

## Introduction

Nuclear magnetic resonance (NMR) methods have become available recently to study metabolism and morphology at the cellular level. The NMR methods may be classified as magnetic resonance spectroscopy (MRS), magnetic resonance imaging (MRI) and combinations of both, magnetic resonance spectroscopic imaging (MRSI). MRS provides information on the chemical composition of cells and its changes under specific circumstances, MRI yields X-ray like images of cellular anatomy and physiology and MRSI can study the spatial distribution of some metabolites within large cells or cellular aggregates. Notably, most of the nuclei participating in cellular reactions or some of their isotopes are NMR active. This allows a large variety of cellular functions to be monitored by different NMR methods. **Table 1** summarizes the magnetic properties of the nuclei most commonly used in NMR studies of cells as well as the biological information which can be obtained.

NMR studies of cells probably started in the mid 1950s analysing the dynamics of water in blood using $^{1}$H NMR. However, modern NMR studies of cellular metabolism began later with the introduction of commercially available high-field spectrometers and Fourier transform NMR techniques. Improve-ments in signal-to-noise ratios allowed, in the early 1970s, a study by $^{13}$C NMR of the metabolism of glucose in a suspension of yeast and to determine by $^{31}$P NMR the intracellular pH in erythrocyte suspensions. Since then studies of cellular metabolism by NMR methods have been used routinely in many laboratories.

The development of cellular NMR is supported by some inherent advantages of the NMR method with respect to more classical approaches. First, the noninvasive character of NMR allows repetitive, noninvasive measurements of metabolic processes as they occur in their own intracellular environment. Second, the magnetic properties of nuclei like the relaxation times $T_1$ and $T_2$ or the homonuclear and

**Table 1** NMR properties and type of biological information provided by various nuclei

| *Nuclei* | *Frequency (MHz)* | *Spin* | *Natural abundance (%)* | *Relative sensitivity* | *Information* |
|---|---|---|---|---|---|
| $^{1}H$ | 400 | 1/2 | 99.98 | 1.0 | Metabolite concentration. Cell fingerprinting. Flow through some pathways. Intra and extracellular pH. Cellular volume. Microscopic imaging. MRI |
| $^{31}P$ | 161.9 | 1/2 | 100 | $6.6 \times 10^{-2}$ | Concentration of phosphorylated metabolities. Bioenergetic status. Intra- and extracellular pH. Cellular volume. Phospholipid metabolism |
| $^{13}C$ | 100.6 | 1/2 | 1.11 | $1.6 \times 10^{-2}$ | Quantitative measurements of metabolic flow through specific pathways |
| $^{23}Na$ | 105.8 | 3/2 | 100 | $9.3 \times 10^{-2}$ | Membrance potential. Cellular volume. $Na^{+}$/H exchange |
| $^{19}F$ | 376.3 | 1/2 | 100 | 0.83 | Intra- and extracellular pH. Oxygenation state. Divalent metal ion concentration |
| $^{2}H$ | 61.8 | 1 | 0.016 | $9.7 \times 10^{-3}$ | Lipid structure. Rates of hydration–dehydration reactions. Water flow and perfusion |
| $^{39}K$ | 18.7 | 3/2 | 93.1 | $5.1 \times 10^{-4}$ | Membrane potential |
| $^{17}O$ | 54.2 | 5/2 | 0.037 | $2.9 \times 10^{-2}$ | Water transport and metabolism |
| $^{15}N$ | 40.5 | 1/2 | 0.37 | $1.0 \times 10^{-3}$ | Nitrogen metabolism |
| $^{14}N$ | 28.9 | 1 | 99.63 | $1.0 \times 10^{-3}$ | Nitrogen metabolism |

heteronuclear spin coupling patterns, contain unique information on the physiological or pathological status of the cells and on the flux through specific metabolic pathways. Third, NMR methods allow the acquisition of images of the spatial distribution of water in sufficiently large single cells or in cellular aggregates, and it seems likely that this approach will be extended to other metabolites in the near future. Despite these advantages, the NMR method is not devoid of drawbacks. In particular, NMR is a relatively insensitive technique with a metabolite detection threshold of around $10^{-1}$ mM for *in situ* cells and 50 μM in extracts for moderate-field spectrometers. Thus, to be able to obtain useful NMR spectra with an adequate signal-to-noise ratio, the cell cultures need to be grown to densities similar to those found in tissues. Even if cell extracts are used, these must be prepared from a sufficiently large number of cells to generate, in the NMR tube, metabolite concentrations in excess of the lower threshold for detection. These demanding conditions require the use of specialized perifusion systems for *in situ* NMR studies or facilities for large-scale cell culture in work with cell extracts.

In this article, we describe general procedures for cell culture compatible with NMR studies and illustrate the information that NMR methods can provide on cellular metabolism and morphology, with examples involving mainly the use of $^{31}P$, $^{13}C$ and $^{1}H$ NMR. Several reviews, some of them quoted in the Further reading section, have covered this topic previously.

## Large-scale and high-density cell culture procedures compatible with NMR studies

The use of cell cultures provides a powerful tool to understand the cellular and molecular mechanisms occurring *in vivo*. Different types of cell culture exist depending on the origin of the cells used in the preparation. Primary cultures are obtained after enzymatic, chemical or mechanical disaggregation of the original tissue. Cell lines, transformed cells or cell strains are obtained from primary cultures after several passages from transplantable tumours or by cellular cloning. The *in vivo* behaviour is more closely resembled by primary cultures, but these usually require complex preparation resulting in relatively small cellular yields. Cellular lines or strains, and cellular clones have the advantage of being homogenous cell populations that are easy to grow and maintain in large scale, but their behaviour is not as comparable to the *in vivo* situation.

**Figure 1** illustrates the general procedure used to obtain a primary culture. The tissue of interest is dissected from the appropriate organ obtained from embryonic, newborn or adult donors. The dissected tissue is then chopped and disaggregated, digested enzymatically and the resulting cells are washed and collected by sedimentation. Cell counting and assessment of viability are easily accomplished later by optical microscopy, using a haemocytometer for cell counting and the dye exclusion criteria for viability assessment. The next step consists of cell seeding.

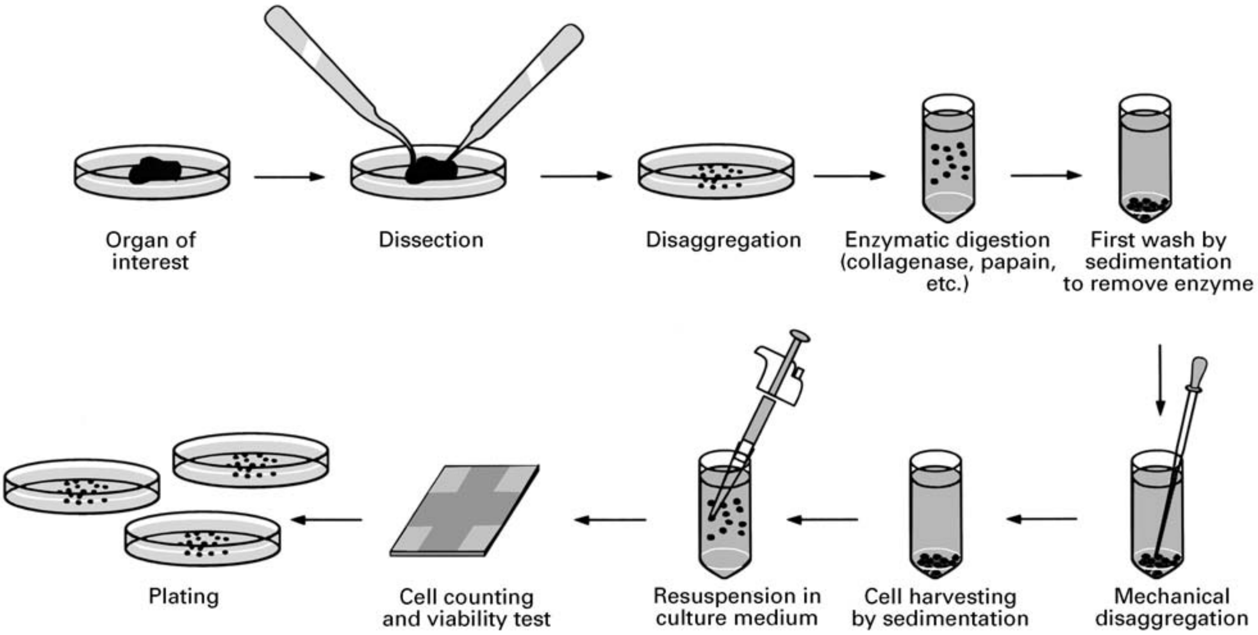



**Figure 1** General steps in the preparation of primary cell cultures.

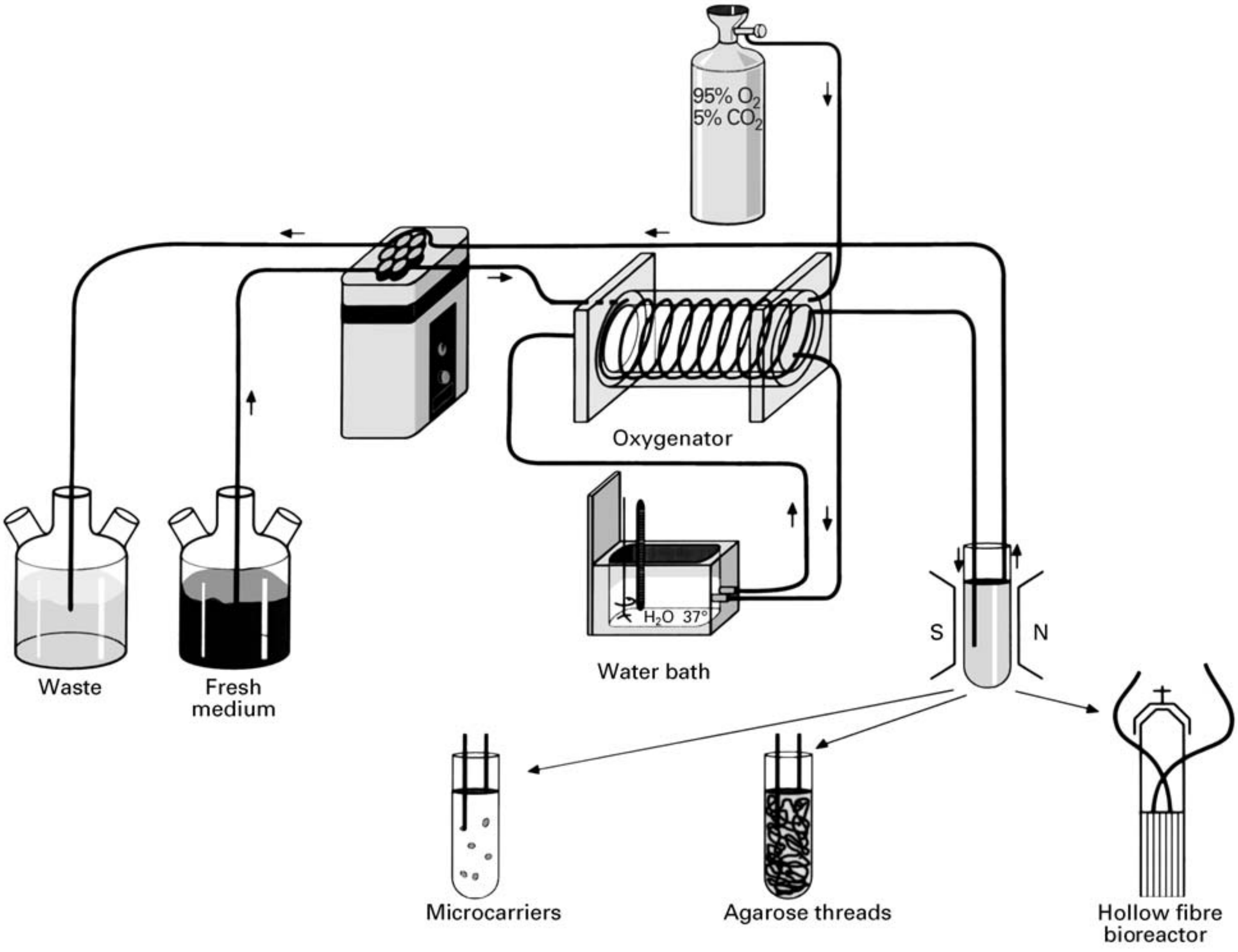



**Figure 2** High density cell culture devices compatible with *in situ* NMR spectroscopy. The tubes containing the entrapped cells can normally be accommodated in commercial high-resolution probes and placed inside the magnet (indicated by N and S).

Cells seeded normally attach to flat surfaces, grow, proliferate and differentiate until reaching confluence. Usually, NMR studies demand cell numbers in the range $10^7$–$10^8$ cells, a quantity requiring a large number of culture dishes or flasks. To further increase the surface of the culture it is possible to use rollers. These devices allow cells to grow in the inner surface of a rolling cylinder which contains a moderate amount of culture medium. Several cylinders can be maintained rolling simultaneously in a rack, increasing enormously the surface area of the culture and thus the number of confluent cells. The excellent monograph of Freshney provides a more detailed description of every cell culture procedure and optimized techniques for particular cells.

There are two different kinds of experimental design in NMR studies of cells: (i) the study can be performed inside the magnet in real time, monitoring the metabolism of cells under *in situ* conditions; or (ii) extracts can be prepared from the cells after different incubation times and examined by NMR under high-resolution conditions (**Figure 2**). In the first case special precautions need to be taken to guarantee adequate nutrient supply, good oxygenation conditions and efficient removal of waste products from the NMR tube containing the cells. Different methods have been developed to fulfil these requirements. All of them involve a perifusion system adapted to the NMR tube which contains the cells immobilized or grown in different matrices. The perifusion system consists basically of a peristaltic pump, a thermostatically stabilized gas-exchange chamber or oxygenator and the NMR tube containing the cells. The peristaltic pump delivers fresh medium to the oxygenator and NMR tube and removes the used medium to a waste reservoir. The oxygenator is constructed of gas-permeable Silastic tubing equilibrated with an atmosphere of 95% $O_2$/5% $CO_2$. The NMR tube contains the entrapped cell suspension. There are basically three types of cell entrapment protocol compatible with NMR studies: (i) polymer threads, (ii) microcarrier beads or (iii) hollow fibre bioreactors.

Cells are entrapped in polymer threads (agarose or alginate) by mixing the cell suspension with a solution of the liquid polymer (above the gelling temperature) and extruding the mixture through a Teflon tube immersed in ice. The polymer solidifies at low temperature entrapping the cells in filaments which can be easily perifused. The advantage of this method lies in its simplicity, but on the other hand, cells are placed in a nonphysiological environment. Alternatively, anchorage dependent cells can be grown on the surface of solid or porous microcarrier beads of approximately 100 μM diameter. The beads can be made of different materials and these include

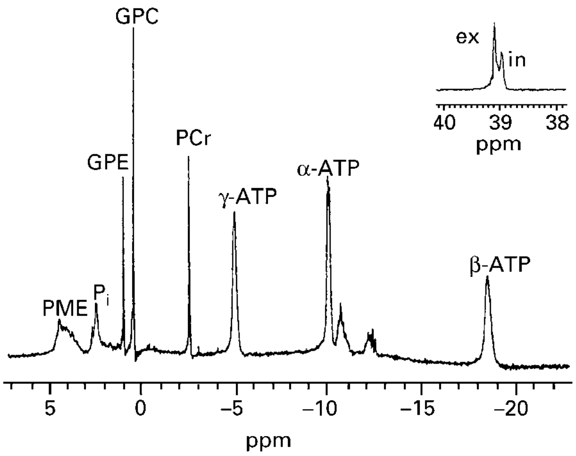



**Figure 3** $^{31}P$ NMR spectrum (161.1 MHz) of C6 glioma cells grown in a hollow fibre bioreactor. Note the important contribution of the GPC and GPE resonances. Inset: intra- (in) and extracellular (ex) resonances of DMMP. From Gillies RJ, Galons JP, McGovern KA, Scherer PG, Lien YH, Job C, Ratcliff R *et al. NMR in Biomedicine* **6**: 95–104. Copyright 1993 John Wiley & Sons Limited. Reproduced with permission.

collagen, polystyrene, polyacrylamide, gelatine, and Sephadex. This method has the advantage that cells can be treated as a suspension. However the density of cells may not be sufficiently high since most of the volume in the NMR tube is filled by the beads. The hollow fibre bioreactor approach consists of growing cells on the surface of capillary fibres which are gas and nutrient permeable. Cells are grown directly in a bioreactor adapted to the NMR tube where the experiment is performed, reaching higher density than with other methods. However, there are some technical difficulties in setting up this model properly and it is relatively expensive. For a detailed description of the bioreactor approach see the literature given in the Further reading section.

## Applications

### $^{31}P$ NMR of cells

$^{31}P$ NMR spectra of cells show a small number of resonances but contain crucial information on cellular biochemistry and physiology. **Figure 3** depicts a representative $^{31}P$ NMR spectrum obtained *in situ* from a culture of C6 glioma cells grown in a bioreactor. It is possible to distinguish clearly resonances from the β, α and γ phosphates of nucleotide triphosphates, phosphocreatine (PCr), phosphodiesters (glycerolphosphorylcholine, GPC, and glycerolphosphorylethanolamine, GPE), inorganic phosphate ($P_i$) and phosphomonoesters (PME). The resonances from the nucleotide triphosphates contain primarily (> 90%) the contribution of MgATP and are

normally referred to as ATP resonances. The PME peak is a composite resonance containing mainly contributions from phosphocholine (PCho) and phosphoethanolamine (PE), as well as from sugar phosphates in a smaller proportion.

$^{31}P$ NMR spectra provide quantitative information on the relative concentrations of phosphorylated metabolites. For spectra acquired under fully relaxed conditions, the area of the $^{31}P$ NMR resonance is proportional to the concentration of the metabolite. Absolute quantifications are more difficult to perform but can also be obtained if the concentration of one of the phosphorylated metabolites in the sample (normally ATP) is measured by an independent method, such as enzymatic analysis. Unfortunately, the resonances from the α and β phosphates of ADP overlap under *in situ* conditions with those from the α and γ phosphates of ATP, thus precluding direct quantification of ADP *in situ*. The concentration of ADP *in situ* can be estimated indirectly from the difference in area between the α ATP or γ ATP peaks and the area of the β ATP peak, which is almost exclusively derived from ATP.

Probably, the main interest in $^{31}P$ NMR spectroscopy lies in its ability to assess the bioenergetic status of the cell. This is determined by the relative rates of ATP-producing and -consuming processes according to Equation [1]

$$\mathrm{MgATP^{2-} + H_2O \rightleftharpoons MgADP^- + P_i{}^{2-} + H^+ + energy} \quad [1]$$

In most mammalian cells, ATP is produced aerobically by oxidative phosphorylation and hydrolysed to provide the energy needed to support biosynthetic activities, maintain transmembrane ion gradients and perform cellular work. The balance between production and consumption of ATP determines its steady-state level and therefore the net availability for energy. If oxygen and nutrients become limiting, ATP synthesis may proceed for a short time using PCr as a phosphate donor in reaction [2]

$$\mathrm{PCr^{2-} + MgADP^- + H^+ \rightleftharpoons MgATP^{2-} + Cr^0} \quad [2]$$

Under these conditions no significant change occurs in the ATP resonances but a net decrease in the intensity of PCr is observed. When the ATP-buffering capacity of the PCr-creatine system is exceeded, net ATP hydrolysis occurs, resulting in a net increase in $P_i$ and acidification, as described by Equation [1].

Another important aspect of $^{31}P$ NMR spectroscopy is that it allows the determination of intracellular and extracellular pH (in $P_i$-containing media) from the chemical shift of the $P_i$ resonance. This is because $P_i$ has an apparent p$K_a$ of 6.7 and its chemical shift is pH-dependent in the physiological range. Normally the chemical shift of the $P_i$ resonance is measured with respect to internal references like PCr or α-ATP, which are nontitratable in the physiological pH range, (p$K_a \approx 4.5$). The following expression can be used to measure intracellular pH from the chemical shift of $P_i(\delta)$ measured with respect to the α-ATP resonance:

$$\mathrm{pH_i} = 6.7 + \log(\delta - 11.26)/(13.38 - \delta)$$

The values 11.26 and 13.38 represent the acidic and alkaline limits of the $P_i$ titration curve chemical shifts with respect to the α-ATP resonance. Phosphonic acid derivatives, like methylene diphosphonic acid (MDPA) or dimethyl methylphosphonate (DMMP), resonating far from physiological phosphates (*ca.* 20–30 ppm) can be used as an external reference for chemical shift and concentration calibrations. Sometimes these references cross the cell membrane and two different resonances are observed for the intracellular and extracellular compounds. In these cases it is possible to determine the relative cellular volume from the relative areas of the intra- and extracellular resonances (**Figure 3** inset).

$^{31}P$ NMR spectra also contain resonances from crucial phospholipid metabolites. These include the phosphodiesters, glycerolphosphocholine and glycerolphosphoethanolamine and the phosphomonoesters, PCho and PE. Tumoural cells and *in vivo* tumours show increased PME resonances, often caused by increased PCho content. This characteristic of tumour cells has been proposed to contain diagnostic or even prognostic information on tumours. The metabolic reasons underlying the increase in PCho in tumour cells are not completely understood but seem to be caused by an increase in its synthesis rather than to an increase in the degradation of phosphatidylcholine.

## $^{13}C$ NMR of cells

$^{13}C$ NMR spectroscopy allows the detection of resonances from $^{13}C$, the stable isotope of carbon with a magnetic moment. The natural abundance for $^{13}C$ is only 1.1% of the total carbon and its magnetogyric ratio is approximately one-fourth of that of the proton. These two circumstances make $^{13}C$ NMR a relatively insensitive technique. However, this sensitivity problem can be improved by using $^{13}C$–enriched substrates. The combination of $^{13}C$ NMR detection

and substrates selectively enriched in $^{13}C$ have made it possible to follow *in vivo* and *in vitro* the activity of a large variety of metabolic pathways in cells and subcellular organelles. These include glycolysis and the pentose phosphate pathway, glycogen synthesis and degradation, gluconeogenesis, the tricarboxylic acid cycle, ketogenesis, ureogenesis and the glutamate, glutamine, γ-aminobutyric acid cycle in the brain among others.

The design of $^{13}C$ NMR experiments with selectively $^{13}C$–enriched substrates is similar to the classical radiolabelling experiments using radioactive $^{14}C$. A relevant difference is that $^{13}C$ precursors are used in substrate amounts, while $^{14}C$ substrates are normally used in tracer amounts. However, $^{13}C$ NMR presents important advantages over $^{14}C$. First, the metabolism of the $^{13}C$-labelled substrate can be followed in real time, *in situ* and noninvasively. Second, even if tissue extracts are prepared, the detection of $^{13}C$ labelling in the different carbon atoms of a metabolite does not require separation and carbon-by-carbon degradation of the compound, two normal requirements in experiments with $^{14}C$. Finally, the analysis by $^{13}C$ NMR of homonuclear spin-coupling patterns and isotope effects allows the determination of whether two or more $^{13}C$ atoms occupy contiguous positions *in the same metabolite molecule*, a possibility only available to $^{13}C$ NMR methods. As a counterpart to these advantages, $^{13}C$ NMR is significantly less sensitive than other conventional metabolic techniques like radioactive counting, mass spectrometry or optical methods. In particular, $^{13}C$ NMR studies with cells *in situ* are difficult to find and most of the $^{13}C$ NMR work has been performed with cell extracts. This allows operation under high-resolution conditions and increases significantly both the sensitivity and the resolution.

**Figures 4** and **5** illustrates the use of $^{13}C$ NMR spectroscopy to study the metabolism of (1-$^{13}C$)glucose in primary cultures of neurons and astrocytes. A simplified scheme of the metabolism of (1-$^{13}C$)glucose in neural cells is given in **Figure 4**. Briefly, (1-$^{13}C$)glucose is metabolized to (3-$^{13}C$)pyruvate through the Embden Meyerhoff glycolytic pathway. The (3-$^{13}C$)pyruvate produced can be transaminated to (3-$^{13}C$)alanine, reduced to (3-$^{13}C$)lactate or enter the tricarboxylic acid (TCA) cycle through the pyruvate dehydrogenase (PDH) or pyruvate carboxylase (PC) activities. A net increase in (3-$^{13}C$)lactate reveals increased aerobic glycolysis and is normally observed under hypoxic conditions in normal cells. If (3-$^{13}C$)pyruvate enters the TCA cycle though PDH it produces (2-$^{13}C$)acetyl-coenzyme A first, and subsequently (4-$^{13}C$)α-ketoglutarate. In contrast, if (3-$^{13}C$)pyruvate is carboxylated to (3-$^{13}C$)oxalacetate by PC, it will produce (2-$^{13}C$)α-ketoglutarate in the next turn of the cycle. The α-ketoglutarate–glutamate exchange is many times faster than the TCA cycle flux and therefore, $^{13}C$ labelling in the glutamate carbon atoms reflects accurately the labelling in the α-ketoglutarate precursor. Thus, (4-$^{13}C$)glutamate or (2-$^{13}C$)glutamate are produced from (1-$^{13}C$)glucose depending on the route of entry of the (3-$^{13}C$) pyruvate into the TCA cycle. Therefore, a comparison of intensities from the C4 and C2 glutamate resonances provides a qualitative estimation of the relative contribution of PDH and PC to the TCA cycle.

**Figure 5** shows proton decoupled $^{13}C$ NMR spectra of perchloric acid extracts obtained from primary cultures of neurons (upper panel) and astrocytes (lower panel) incubated for 24 h with 5 mM (1-$^{13}C$)glucose. The most relevant resonances are the ones derived from the C3, C4 and C2 carbons of glutamate; C3, C4 and C2 carbons of glutamine (or glutathione); C3 carbon of alanine; C3 carbon of lactate; C6 carbon of *N*-acetylaspartic acid and C2 carbon of acetate (see **Figure 5** legend for resonance assignments). The spectra of **Figure 5** reveal important differences in the metabolism of (1-$^{13}C$)glucose by neurons and astrocytes. Briefly, the lactate C3 resonance is higher in the spectrum of astrocytes than in that of the neurons. The glutamate C4 resonance is significantly larger in neurons than in astrocytes, while the C2 glutamine resonance in astrocytes is higher than in neurons. Taken together these results suggest that astrocytes metabolize glucose mainly through aerobic glycolysis having an important contribution from PC, while neurons metabolize glucose mainly through the TCA cycle, using PDH as the main route.

It is possible to obtain in cells quantitative determinations of flux through specific metabolic pathways or through specific steps of a pathway using $^{13}C$ NMR. There are two possible strategies, both based on the principles of isotopic dilution and both requiring the use of mathematical models of metabolism. The first strategy allows the determination of fluxes by solving linear systems of equations relating flux to fractional $^{13}C$ enrichments in specific carbons of precursors and products. The second strategy involves the determination of relative isotopomer populations from an analysis of the $^{13}C$–$^{13}C$ spin coupling patterns. For a more precise description of both methods the reader is referred to the article of Portais and co-workers and Künnecke and co-workers in the Further reading section.

### $^{1}H$ NMR of cells

$^{1}H$ NMR has an inherently higher sensitivity than $^{31}P$ or $^{13}C$ NMR and thus a larger number of metabolites

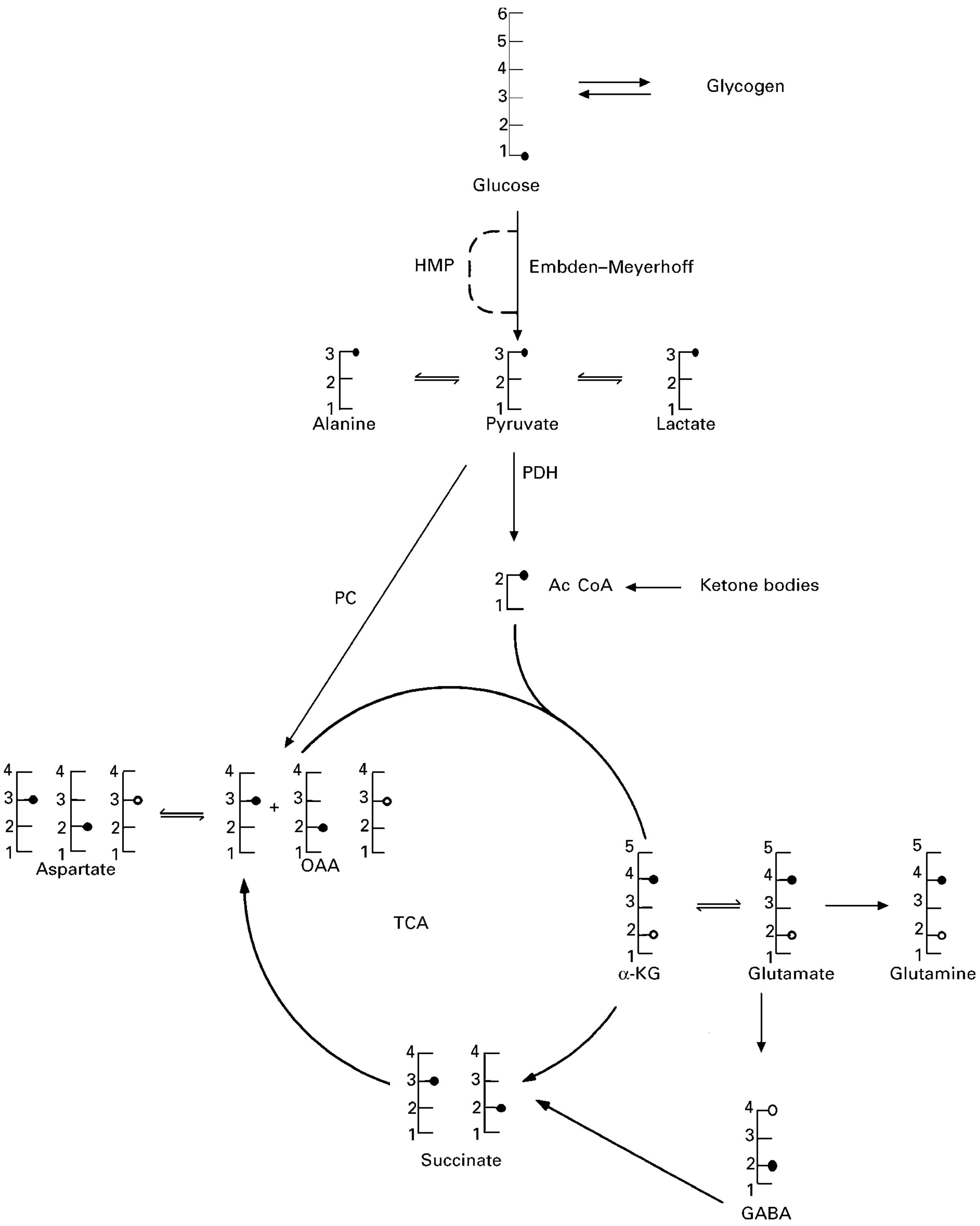



**Figure 4** Metabolism of (1-$^{13}$C) glucose in neural cells. HMP: hexose monophosphate shunt. TCA: tricarboxylic acid cycle; OAA: oxalacetate; α-KG: 2-oxoglutarate; GABA: γ-aminobutyric acid; Ac CoA: acetyl coenzyme A. Only one turn of the TCA is considered for clarity. Filled or open circles in TCA cycle intermediates indicate labelling from PDH or PC, respectively.

is potentially detectable. However, two limitations need to be considered: (i) the water signals from the cells need to be eliminated and (ii) the complex homonuclear spin coupling patterns due to overlapping proton resonances have to be resolved. Both drawbacks have been overcome in part by the

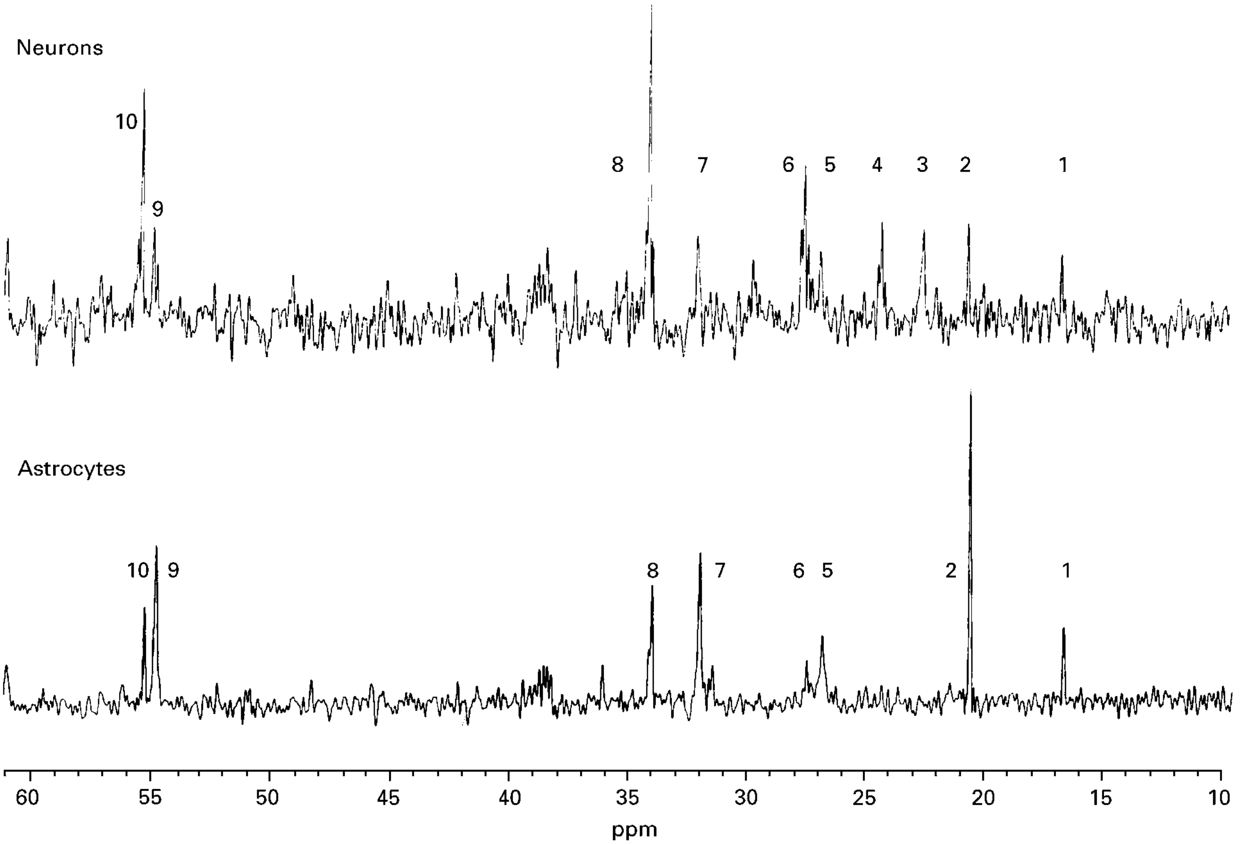



**Figure 5** Proton decoupled $^{13}C$ NMR spectrum (125.7 MHz) of neutralized perchloric acid extracts obtained after the incubation of primary cultures of neurons (upper trace) or astrocytes (lower trace) with 5 mM (1-$^{13}C$) glucose during 24 h. 1: alanine C3, 2: lactate C3, 3: *N*-acetylaspartic C6, 4: acetate C2, 5: glutamine C3, 6: glutamate C3, 7: glutamine C4, 8: glutamate C4, 9: glutamine C2, 10: glutamate C2. An artificial line broadening of 1 Hz was used in both cases.

development of water suppression techniques and different one-, two- or *n*-dimensional editing methods.

The information provided by $^{1}H$ NMR spectra of cell suspensions is illustrated in **Figure 6** which shows a representative $^{1}H$ NMR spectrum of a suspension of rat erythrocytes acquired using a simple one-pulse sequence (**Figure 6B**). The most prominent resonances derive from the H3 proton of lactate, the choline protons of ergothioneine and the glutamate and glutamine resonances from glutathione. All these resonances are located on top of a much broader resonance arising from low-mobility lipids. Therefore $^{1}H$ NMR spectroscopy allows lactate production to be followed easily. This is normally done using spin echo sequences which eliminate the broad lipid component because of its short $T_2$ (**Figure 6A**). In addition to lactate, $^{1}H$ NMR normally allows the direct quantification of creatine singlets, singlets from choline-containing compounds and many other metabolites. Sometimes, quantification of these compounds requires very high resolution conditions, high spectrometer observation frequencies and computerized deconvolution algorithms. Indeed the $^{1}H$ NMR spectrum provides a true fingerprint of the metabolites present in the cell and it appears to be characteristic for every cell type. This has been clearly shown in studies of cell typing with neural cells. Furthermore, it has been shown recently that fully transformed cells present an increase in the $^{1}H$ resonance of the PCho peak, and that the ratio PCho to GPC may serve as an indicator of multiple oncogene lesions.

A more recently developed application of $^{1}H$ NMR to cellular studies relies on the noninvasive determination of intra- and extracellular pH. The procedure is based in the use of a probe containing a reporter proton with a chemical shift sensitive to pH in the physiological range. This method is inherently more sensitive than previous $^{31}P$ NMR approaches. **Figure 6A** illustrates more clearly this procedure, by showing the spectrum of an erythrocyte suspension after the addition of the extrinsic pH probe imidazol-1-yl acetic acid methyl ester. The imidazole protons, H2 (*ca.* 8.2 ppm) and H4, H5 (*ca.* 7.25 ppm) are clearly observed in the aromatic region of the spectrum between 6 and 9 ppm (inset to **Figure 6A**). Two

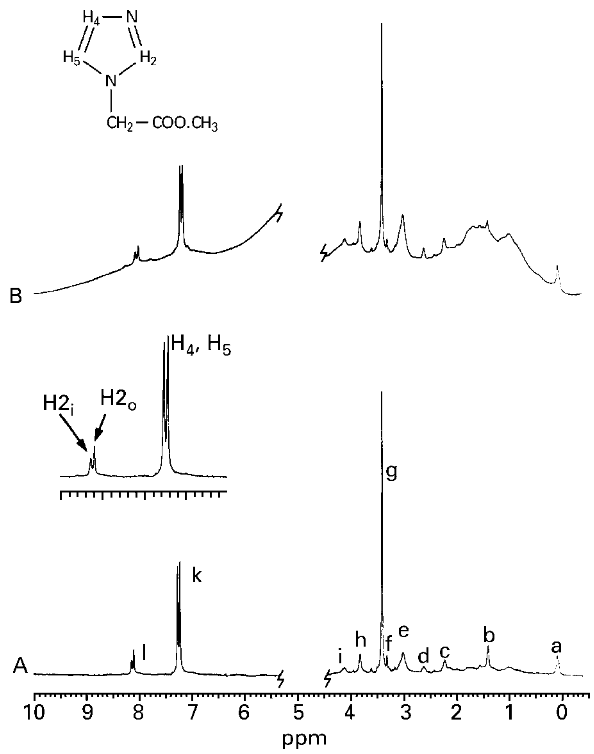



**Figure 6** $^1$H NMR spectra (360.13 MHz) of a rat erythrocyte suspension in the presence of a $^1$H NMR pH probe. The water resonance (not shown) was attenuated with a 2 s presaturating pulse. (A) Spin echo acquisition with 5 ms interpulse delay. Inset: expansion of the aromatic region. (B) Conventional one-pulse acquisition. Inset: Chemical structure of imidazole-1-yl acetic acid methyl ester, the pH probe used. a: TSP, b: lactate H3, c: glutamate H3,H3′ in glutathione, d: aspartate H3,H3′, e: creatine (methyl), f: trimethyl groups in choline, carnitine and ergothioneine, g: methanol, h: H2 of amino acids, i: lactate H2, k: H4 and H5 of the pH probe, l: H2 of the pH probe. $H2_i$: intracellular probe, $H2_o$: extracellular probe. Methanol is produced by endogenous esterases immediately after the addition of the probe.

signals for the imidazole H2 proton ($H2_i$ and $H2_o$) with different chemical shifts are observed. These signals have been shown to be derived from molecules of the probe located in the intra- ($H2_i$) and extracellular ($H2_o$) space. Thus the different chemical shift of intra- and extracellular probes allows the determination of the intra- and extracellular pH values as well as the transmembrane pH gradient.

Finally, impressive progress in NMR microscopy methods has allowed the acquisition of $^1$H NMR images of single cells. $T_2$-weighted NMR images from single cells normally show a dark cytoplasm and relatively bright nucleus. **Figure 7** provides an illustrative example of cellular NMR imaging, showing the morphology of a neuron from *Aplysia californica* before (**Figure 7A**) and after (**Figure 7B**) hypotonic shock. The hypotonic shock increased the relaxation times $T_1$ and $T_2$ of cytoplasm and nucleus,

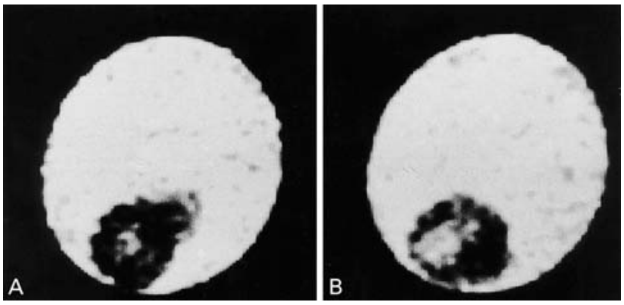



**Figure 7** $T_2$-weighted spin echo microscopic $^1$H NMR images (360.13 MHz) of an isolated neuron from *Aplysia californica* before (A) and after (B) hypotonic shock. Resolution (*x*,*y*,*z*) is 20 μm × 20 μm × 150 μm. Artificial sea water appears as a bright background surrounding the cell. $T_2$ differences between nucleus and cytoplasm cause these structures to appear as bright and dark, respectively. Reproduced with permission of the American Physiological Society from Hsu EW, Aiken NR, Blackband SJ (1996) *American Journal of Physiology* **271**: C1895–C1900.

but did not change the apparent diffusion coefficient of water. In conclusion, $^1$H NMR methods provide information on glycolysis, fingerprint cell typing, state of proliferation, intra- and extracellular pH and cellular micromorphology. Representative articles describing these $^1$H NMR applications are given in the Further reading section.

## Other nuclei

In addition to the most commonly used nuclei, such as $^{31}$P, $^{13}$C and $^1$H, other nuclei can be used in NMR studies of cellular metabolism (**Table 1**). The $^{19}$F nucleus has a magnetic moment slightly smaller than $^1$H, whereas its chemical shift range is larger by a factor of approximately 20. No contributions from tissue background signals are observed, which makes $^{19}$F a useful magnetic label for metabolic studies. Using difluoromethylalanine or fluorinated derivatives of BAPTA, $^{19}$F NMR spectroscopy has been extensively used to measure intracellular pH as well as cytosolic free $Ca^{2+}$ and free $Mg^{2+}$ in peripheral blood lymphocytes and other cells.

The $^{23}$Na nucleus is a quadrupolar spin 3/2 nucleus with approximately 100% natural abundance. Because of its quadrupolar nature, some of the $^{23}$Na is invisible to the NMR experiment. However, it has been possible to resolve intra- and extracellular $^{23}$Na resonances in yeast cells and erythrocytes and to monitor $Na^+$–$H^+$ exchange in the latter cells.

The $^{17}$O nucleus has an unfavourable gyromagnetic ratio but a large chemical shift range, close to 1000 ppm. $^{17}$O NMR has been used extensively to study water metabolism in erythrocytes and subcellular organelles.

Finally, $^2H$ is one of the most recent additions to the repertoire of cellular NMR methods. $^2H$ has a natural abundance of only 0.02% and therefore the use of $^2H$ NMR requires the use of selectively enriched substrates. $^2H$ has been mainly used in cells to investigate the structure *in situ* of choline lipids in the cell membrane.

## List of symbols

ADP = adenosine 5′-diphosphate; ATP = adenosine 5′-triphosphate; BAPTA = 1,2-bis(2-aminophenoxy) ethane-N,N,N′,N′-tetraacetic acid; CoA = coenzyme-A; CDP = cytidine 5′-diphosphate; DMMP = dimethyl methylphosphonate; GPC = glycerolphosphocholine; GPE = glycerolphosphoethanolamine; MDPA = methylene diphosphonic acid; MRI = magnetic resonance imaging; MRS = magnetic resonance spectroscopy; MRSI = magnetic resonance spectroscopic imaging; PCho = phosphocholine; PE = phosphoethanolamine; PC = pyruvate carboxylase; PCr = phosphocreatine, PDH = pyruvate dehydrogenase; $P_i$ = inorganic phosphate; PME = phosphomonoesters; TCA = tricarboxylic acid cycle; TSP = 2,2′,3,3′-tetradeutero trimethylsilyl propionate; $T_1$ and $T_2$ = relaxation times.

*See also:* **Biofluids Studied By NMR; Carbohydrates Studied By NMR; *In Vivo* NMR, Applications, $^{31}P$; *In Vivo* NMR, Methods; NMR Microscopy; Nucleic Acids Studied Using NMR; Perfused Organs Studied Using NMR Spectroscopy; Solvent Supression Methods in NMR Spectroscopy; Two-Dimensional NMR Methods.**

## Further reading

Bhakoo KK, Williams SR, Florian CL, Land H and Noble MD (1996) Immortalization and transformation are associated with specific alterations in choline metabolism. *Cancer Research* 56: 4630–4635.

Freshney RI (1994) In: Freshney RI (ed) *Culture of Animal Cells: a Manual of Basic Technique*. New York: Wiley-Liss.

Gil S, Zaderenko P, Cruz F, Cerdan S and Ballesteros P (1994) Imidazol-1-ylalkanoic acids as extrinsic $^1H$ NMR probes for the determination of intracellular pH, extracellular pH and cell volume. *BioOrganic and Medicinal Chemistry* **2**: 305–314.

Gillies RJ, Galons JP, McGovern KA, Scherer PG, Lien YH, Job C, Ratcliff R, Chapa F, Cerdan S, Dale BE (1993) Design and application of NMR-compatible bioreactor circuits for extended perifusion of high-density mammalian cell cultures. *NMR in Biomedicine* **6**: 95–104.

Hsu EW, Aiken NR, Blackband SJ (1996) Nuclear magnetic resonance microscopy of single neurons under hypotonic perturbation. *American Journal of Physiology* **271** (Cell Physiol 40): C1895–C1900.

King GF and Kuchel PW (1994) Theoretical and practical aspects of NMR studies of cells. *Immunomethods* **4**: 85–97.

Künnecke B, Cerdan S and Seelig J (1993) Cerebral metabolism of (1,2-$^{13}C_2$)glucose and (U-$^{13}C_4$)3-hydroxybutyrate in rat brain, as detected by $^{13}C$ spectroscopy. *NMR in Biomedicine* **6**: 264–277.

Lundberg P, Harmsen E, Ho C, and Vogel HJ (1990) Nuclear magnetic resonance studies of cellular metabolism. *Analytical Biochemistry* **191**: 193–222.

Portais JC, Schuster R, Merle R, Canioni P (1993) Metabolic flux determination in C6 glioma cells using carbon-13 distribution upon (1-$^{13}C$)glucose incubation. *European Journal of Biochemistry* **217**: 457–468.

Szwergold BS (1992) NMR spectroscopy of cells. *Annual Review of Physiology* **54**: 775–798.

Urenjak J, Williams SR, Gadian DG and Noble M (1993) Proton nuclear magnetic resonance spectroscopy unambiguously identifies different neural cell types. *Journal of Neuroscience* **13**: 981–989.

Vogel HJ, Brodelius P, Lilja H and Lohmeier-Vogel EM (1987) Nuclear magnetic resonance studies of immobilized cells. *In* Mosbach K (eds.), *Methods in Enzymology*, vol. 135, B: pp. 512–536. New York: Academic Press.

# Chemical Applications of EPR

**Christopher C Rowlands** and **Damien M Murphy**,
Cardiff University, UK

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

## Introduction

EPR spectroscopy is a technique that affords a means for the detection and quantification of paramagnetism, i.e. the presence of unpaired electrons. It is applicable to solids, liquids or even gases. A compound or molecular fragment containing an odd number of electrons is paramagnetic. In this article we give examples to show the power and breadth of the technique. The reader is advised, however, to read the Encyclopedia article on the theoretical aspects of EPR to appreciate fully the strength of the technique and also the excellent series *Specialist Periodical Reports on Electron Spin Resonance*, published by the Royal Society of Chemistry, which has now reached volume 16.

The technique can be applied to obtain information on the following four areas.

1. How atoms are bonded together in molecules (structure).
2. The route by which compounds are transformed from one to another, i.e. which bonds break and which are formed (mechanism).
3. The rate at which these processes occur (kinetics).
4. The molecular interactions that exist, for example, between solvent and solute (environment).

Knowledge of the $g$ values and the detailed hyperfine interactions ($A$ values) allow us to help identify radical species, and these parameters contain information about the electron distribution within the molecule. Radicals are often present as intermediates during a reaction; consequently their identification will give information concerning the reaction mechanism and measurement of how their concentration changes with time will give kinetic data.

Let us now look in greater detail at several major applications of the technique.

## Organic and organometallic free radicals; structure, kinetics and mechanism

There are many many examples of the power of the technique in this area and space does not permit us to cover fully all examples. Articles by Tabner and by Rhodes fully cover the basic principles and practical aspects for the detection and evaluation of radical anions and cations (see Further reading) while more recently Davies and Gescheidt have reviewed the literature up to 1996 in the Specialist Reports referred to earlier. In the same publication, Alberti and Hudson cover the topics of organic and organometallic free radicals up to 1995. Consequently we will select just one or two examples from these areas.

An early method, widely used for radical generation is based on the redox system $Ti^{3+}$–$H_2O_2$. This was first used by Dixon and Norman to generate $HO^\bullet$ and $HOO^\bullet$ radicals obtained by rapid mixing of aqueous solutions of $TiCl_3$ and $H_2O_2$ with a substrate immediately before the EPR cavity. It has been suggested that the main reacting species in such a system is $Ti–O–O^{\bullet 3+}$. The EPR signals obtained from such a system depend upon such factors as flow rate, temperature, pH and the ratio of the two components. Flow systems such as this have been used quite successfully to study polymerization reactions that are initiated by redox reactions. Analysis of the hyperfine interactions of the reacting species enabled the structure of the radicals to be determined.

The use of such flow and/or stopped flow systems has been used quite extensively to investigate the kinetics of radical reactions. These include the use of a stopped flow EPR system to investigate the one-electron oxidation of a range of phenols by a dimeric manganese(IV/IV) triazacyclononane complex with and without hydrogen peroxide.

Another use has been the development of a technique of photolytic time-resolved EPR that allows kinetic data to be obtained with relative ease. The method involves flowing oxygen-free solutions of the materials under investigation through a flat reaction cell placed within the EPR cavity. The flow rates of the solutions could be adjusted so that signal intensities were not dependent on their cavity dwell time. The solutions could be irradiated either continuously or intermittently and thermostatted by the use of a stream of dinitrogen. Signal averaging is used to overcome signal-to-noise difficulties.

## Spin trapping, spin labels and spin probes in polymers and biological systems

The role of free radicals in biological and medicinal systems is of considerable interest as they have been implicated in a wide range of diseases such as cancer. In such systems, the direct detection of free radicals is often not possible because of their low steady-state concentration arising from their high reactivity and transient nature. In this instance, the technique of spin trapping is used. Spin trapping involves the addition of a diamagnetic molecule, usually a nitrone or nitroso compound, which reacts with a free radical to give a stable paramagnetic spin adduct that accumulates until it becomes observable by EPR (Eqn [1]).

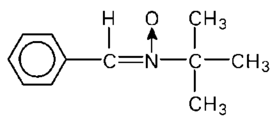
| | | | | | |
|---|---|---|---|---|---|
| $X^\bullet$ | + | ST | $\longrightarrow$ | $XST^\bullet$ | [1] |
| Transient free radical | | Spin trap | | Long-lived spin adduct | |
| *Paramagnetic* | | *Diamagnetic* | | *Paramagnetic* | |

By determination of the parameters of the spin adduct spectrum it is often possible to identify the nature of the primary trapped radical, or at least to determine the type of radical trapped. **Table 1** gives examples of the range of hyperfine coupling constants obtained for a variety of trapped species for the two most commonly used spin traps DMPO (dimethylpyridine *N*-oxide) and PBN (α-phenyl-*N*-*t*-butyl nitrone). Care has to be taken, however, because the coupling constants vary with solvent polarity (for example, in water the DMPO adduct of the *t*-butoxy radical has $a$(N) = 1.48 and $a$(H) = 1.60 mT while in toluene $a$(N) = 1.3 and $a$(H) = 0.75 mT).

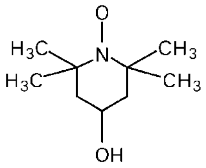
The spin trap must react at relatively fast trapping rates and the radical adduct must have a reasonable half-life; for example, DMPO has a $k_{trapping}$ for $OH^\bullet$ of $2 \times 10^9\ M^{-1}\ s^{-1}$ and half-life of about 15 min, whereas for $O_2^-$, $k_{trapping}$ is $1 \times 10^1\ M^{-1}\ s^{-1}$ and the half-life is about 90 s. Consequently the DMPO–$O_2^-$ adduct is rarely seen. This difficulty has been overcome by the use of phosphorus-substituted DMPO, in particular where one of the methyl groups in the 5 position has been replaced by a diethoxyphosphoryl group. This increases the half-life of the OOH adduct considerably in both organic and aqueous solvents, with the difference being more pronounced in aqueous solvents. One other advantage of the use of phosphorus-substituted DMPO spin traps is that the β-$^{31}P$ hyperfine coupling constant has a much larger variation (2.5–5.5 mT) than the β-H hyperfine coupling constant (0.6–2.5 mT), thus allowing an easier identification of the trapped radical.

**Table 1** Hyperfine coupling constants (mT) for a variety of spin-trapped adducts

| | *PBN* | | *DMPO* | |
|---|---|---|---|---|
| *Free radical* | *$^{14}N$* | *H* | *$^{14}N$* | *H* |
| $^\bullet H$ | 1.53 | 0.82(2H) | | |
| $^\bullet$Phenyl | 1.60 | 0.44 | 1.57 | 2.5 |
| $^\bullet CCl_3$ | 1.39 | 0.175 | | |
| $^\bullet OH$ | 1.54 | 0.27 | 1.50 | 1.50 |
| $O_2^-$ | 1.43 | 0.225 | 1.45 | 1.16 |
| $Bu^tO^\bullet$ | | | 1.48 | 1.60 |
| $SO_3^-$ | | | 1.47 | 1.60 |
| $PhS^\bullet$ | | | 1.29 | 1.415 |



PBN, *N*-*t*-butyl-α-phenylnitrone



TEMPOL, 4-hydroxy-2,2,6,6-tetramethylpiperidinooxy

**Table 2** gives a range of bimolecular rate constants with both PBN and DMPO for a variety of radical species.

In a study of heterogeneous catalysis, the spin trapping technique has been used to prove the presence of radical species on a catalyst surface. In a study of a palladium metal catalyst supported on alumina, it was shown that hydrogen is dissociatively chemisorbed by trapping hydrogen atoms with PBN. Alkyl and aromatic free radicals were also shown to be present when other feed stocks were used. In a separate study of the photodecomposition of acetaldehyde, it was shown that high-power irradiation (600 W mercury/xenon lamp) generated a different PBN adduct from that given with irradiation under direct sunlight and that the presence of oxygen played a major role in the photochemistry.

**Table 2** Second-order rate constants for radical spin trapping for the traps PBN and DMPO

| *Spin trap* | *$R^\bullet$* | *T(°C)* | *Rate constant ($M^{-1}s^{-1}$)* |
|---|---|---|---|
| PBN | $Me^\bullet$ | 25 | $4 \times 10^6$ |
| | $RC^\bullet H_2$ | 40 | $1.3 \times 10^5$ |
| | $Ph^\bullet$ | 25 | $2 \times 10^7$ |
| | $HO^\bullet$ | 25 | $2 \times 10^9$ |
| | $ROO^\bullet$ | 60 | $1 \times 10^2$ |
| DMPO | $HO^\bullet$ | 25 | $2 \times 10^9$ |
| | $O_2^-$ | 25 | $1 \times 10^1$ |
| | $RC^\bullet H_2$ | 40 | $2.6 \times 10^6$ |

Since the first reported study on nitroxide spin labelling in 1965, the technique has found widespread application to biological and chemical problems. The introduction of a small free radical, the spin label, onto the molecule of interest gives environmental information at that point. It has the ability to measure very rapid molecular motion and is usually free from interference, thus making it a very powerful technique that gives useful details about dynamic processes at the molecular level.

The types of free radical mostly used as spin probes or spin labels are based on nitroxides, such as the label known as TEMPOL (4-hydroxy-2,2,6,6-tetramethylpiperidinooxy). **Figure 1** shows the change in the EPR spectrum of a spin-labelled system going from a fully mobile isotropic system (**Figure 1C**) to a fully immobilized glass type spectrum (**Figure 1A**). Measurement of the hyperfine parameters enables information to be obtained about the mobility or rigidity of the system under study.

The spin probe technique has been used to investigate the interactions of the anionic surfactant SDS and gelatin. It was found that for simple SDS solutions the rotational correlation time of the spin probe, 16-doxylstearic acid methyl ester, increased slightly with increased surfactant concentration. In contrast, in the presence of gelatin these properties varied markedly as a function of the stoichiometric ratio of the surfactant to gelatin concentration, with the correlation time decreasing as the hyperfine coupling constant increased with increasing surfactant concentration. Such behaviour was attributed to the characteristics of the various amino acids present in the gelatin.

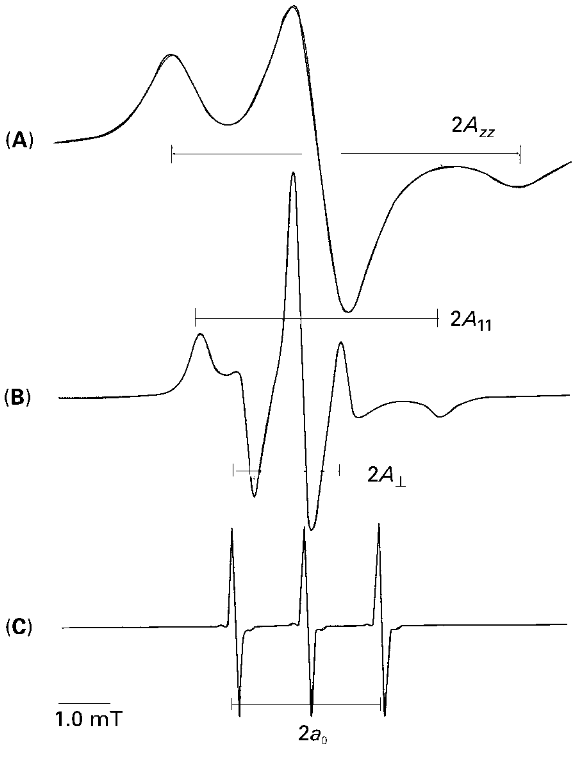



**Figure 1** EPR spectra arising from a $^{14}N$ ($I = 1$) showing (A) randomly and rigidly oriented in a frozen matrix; (B) antisotropic motion in a randomly oriented environment; (C) isotropic spectrum with no restriction on movement.

The use of spin labelling in polymer systems has enabled information to be obtained about polymer chain size and also their conformational state, with three main states being distinguished. The ability of some nitroxide radicals to diffuse easily into various polymer systems, coupled with the fact that their EPR spectra showing three lines caused by coupling to $^{14}N$ show a high degree of rotational motion in the amorphous parts of the polymer, means that studies varying the temperature through the glass transition temperature $T_g$ will enable changes to be measured in the rotational correlation time. Since the rotational correlation time obeys the Arrhenius law (Eqn [2]),

$$\tau = \tau_0 \exp(E_a/RT) \qquad [2]$$

which is valid for $T > T_g$, values for the rotational activation energy of the nitroxide free radical can be measured. Since this does not depend on radical size but only on the polymer matrix, a measure of polymer mobility is obtained.

There are several excellent reviews on the application of the method to studies in polymer systems and also to investigations of biological membranes. These are listed in the 'Further reading' section.

## Study of transition metal compounds

EPR spectroscopy is widely used to study the structure and geometry of d-transition metal ions (TMI) in inorganic complexes, in biological systems and in catalysts. Information obtained through analysis of the spectrum varies from simple identification of the metal centre to a thorough description on the electronic structure of the complex. At a simple level, the main interactions experienced by the TMI that consequently influence the EPR spectrum are (i) the electronic Zeeman effect and (ii) interaction between the electron and the nuclear spin. In the first case, the interaction is expressed via the $g$ tensor which carries information on electronic structure; in the second case, the metal hyperfine interaction is expressed via the $A$ tensor.

Although the information regarding the transition metal compound can be comprehensive, we shall

confine ourselves here to a simple illustration on the power of EPR for investigating the symmetry of the metal centre in the solid state. The simplest case arises for TMI with one d electron. For a metal in a tetragonally distorted octahedral field experiencing compression along the 4-fold axis ($D_{4h}$), the five d orbitals have the splittings shown in **Figure 2** with the nondegenerate $d_{yz}$ state lowest in energy. For a slight distortion, the low-lying 2-fold degenerate excited states ($d_{xy}$, $d_{xz}$) are close to the ground state, as depicted in **Figure 2A**. The small energy separation $\Delta_1$ results in a short relaxation time with the result that the EPR signal can only be observed at low temperatures. Greater distortions lead to longer relaxation times, so that the EPR signals become observable at increasingly higher temperatures with $g_\perp > g_\parallel$ as shown in **Figure 2A**. In the case of elongation along the 4-fold axis, the unpaired electron now occupies a doubly degenerate ground state ($d_{xz}$, $d_{yz}$) and no EPR signal is observed. However, if the ion experiences a trigonal distortion with a resulting point symmetry of $D_{3d}$ as shown in **Figure 2B**, then the single electron occupies the $d_{z^2}$ orbital so the order of the *g*-values changes or becomes 'inverted' (i.e. $g_\parallel > g_\perp$) as depicted in **Figure 2B**. This simple example illustrates the effects of the ground state of the TMI on the resulting EPR spectrum.

As shown above, the point symmetry at the metal can therefore be reflected in the **g** values, but the *A* values can also be influenced in a similar manner. Furthermore, depending on the point symmetry, the principal axes of **g** and *A* may or may not be coincident. The relationship between these tensors, the EPR spectral characteristics and the point symmetry of the metal are shown in **Table 3**. The importance of these relationships is that each type of EPR

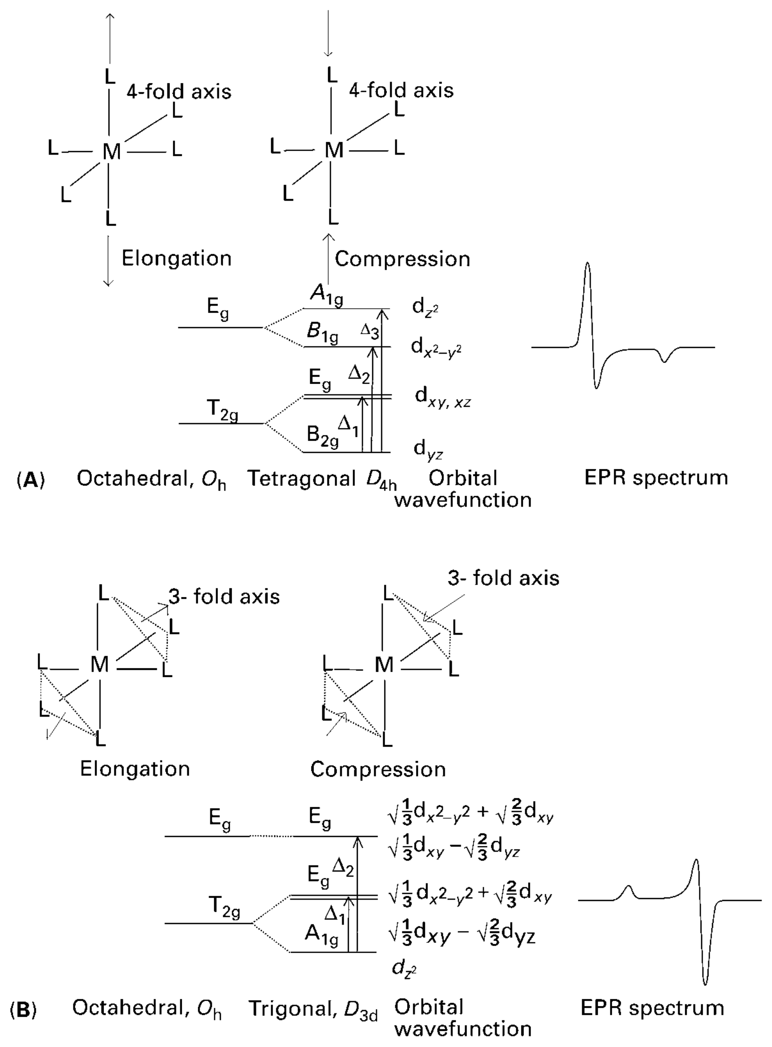



**Figure 2** Orbitals and corresponding EPR spectrum for an octahedral complex with (A) tetragonal and (B) trigonal distortion.

**Table 3** Relationship between ***g*** and ***A*** tensor, EPR symmetry and the point symmetry of the paramagnets

| *EPR symmetry* | ***g** and **A** tensors* | *Coincidence of tensor axis* | *Molecular point symmetry* |
|---|---|---|---|
| Isotropic | $g_{xx} = g_{yy} = g_{zz}$ $A_{xx} = A_{yy} = A_{zz}$ | All coincident | $O_h$, $T_d$, $O$, $T_h$, $T$ |
| Axial | $g_{xx} = g_{yy} \neq g_{zz}$ $A_{xx} = A_{yy} \neq A_{zz}$ | All coincident | $D_{4h}$, $C_{4v}$, $D_4$, $D_{2d}$, $D_{6h}$, $C_{6v}$, $D_6$, $D_{3h}$, $D_{3d}$, $C_{3v}$, $D_3$ |
| Rhombic | $g_{xx} \neq g_{yy} \neq g_{zz}$ $A_{xx} \neq A_{yy} \neq A_{zz}$ | All coincident | $D_{2h}$, $C_{2v}$, $D_2$ |
| Monoclinic | $g_{xx} \neq g_{yy} \neq g_{zz}$ $A_{xx} \neq A_{yy} \neq A_{zz}$ | One axis of *g* and *A* coincident | $C_{2h}$, $C_s$, $C_2$ |
| Triclinic | $g_{xx} \neq g_{yy} \neq g_{zz}$ $A_{xx} \neq A_{yy} \neq A_{zz}$ | Complete non-coincidence | $C_1$, $C_1$ |
| Axial non-collinear | $g_{xx} \neq g_{yy} \neq g_{zz}$ $A_{xx} = A_{yy} \neq A_{zz}$ | Only $g_{zz}$ and $A_{zz}$ coincident | $C_3$, $S_6$, $C_4$, $S_4$, $C_{4h}$, $C_6$, $C_{3h}$, $C_{6h}$ |

behaviour is associated with a restricted number of point symmetries, which places constraints upon the geometrical structures of the paramagnet. For example, if the paramagnet is known to have, say, rhombic symmetry, then the associated geometry must belong to one of the point groups $D_{2h}$, $C_{2v}$ or $D_2$ (**Table 3**). It would be incorrect to assign a structure that belongs to a more symmetric arrangement, e.g. $D_{4h}$ which is strictly axial. Therefore, for a system with unknown structure, if the EPR symmetry can be determined (**Table 3**) we have invaluable structural information since the paramagnet can only belong to a restricted range of point symmetries. As an example, consider the EPR spectrum of $H_2[V(R,R\text{-HIDPA})_2]$ diluted in the corresponding zirconium compound (**Figure 3B**). At first glance the spectrum appears to have axial symmetry, since the features associated with the *x* and *y* directions are not split. Simulation of a spectrum based purely on axial symmetry is not entirely satisfactory (**Figure 3A**). The introduction of a small anisotropy in both $g_{xx}$, $g_{yy}$ and $A_{xx}$, $A_{yy}$, however, produces an improved simulation (**Figure 3C**). Although the rhombicity is small, it is consistent with the known structure of the compound. From the crystallographic data, the highest possible point symmetry at the vanadium is $C_2$. In more complex cases, other techniques such as variable-frequency CW-EPR, ENDOR and ESEEM are required to provide a more detailed description on the electronic structure of transition metal complexes. Nevertheless, one can start to appreciate the power of CW-EPR for studies of paramagnetic transition metal complexes.

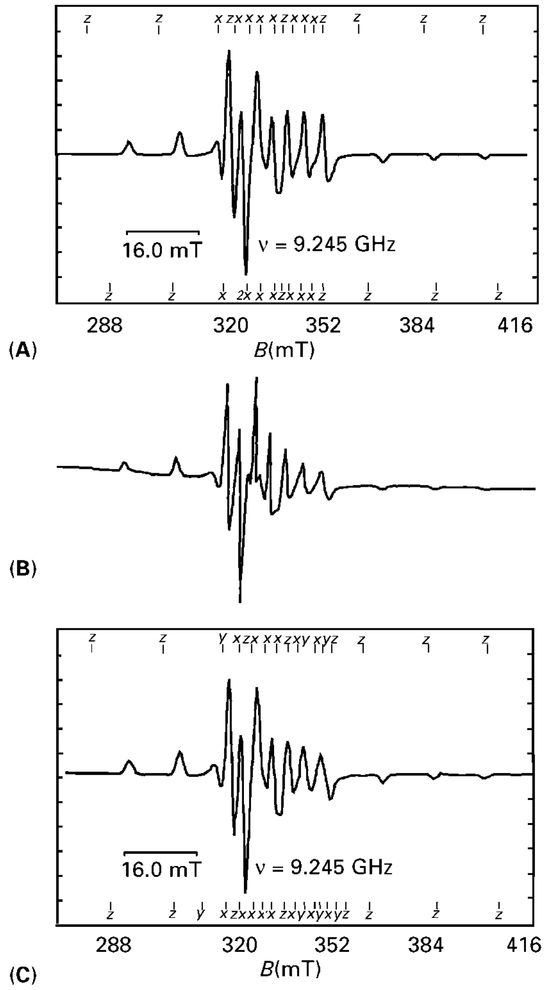



**Figure 3** Experimental and simulated X-band powder EPR spectra of $H_2[V(R,R\text{-HIDPA})_2]$ diluted in the corresponding zirconium compound, where HIDPA = hydroxyiminodipropionate. (A) Simulated spectrum with axial symmetry and $g_{\parallel} = 1.9195$, $g_{\perp} = 1.9839$, $A_{\parallel} = 17.5$ mT and $A_{\perp} = 4.9$ mT; (B) experimental spectrum; and (C) simulated spectrum with rhombic symmetry $g_{zz} = 1.9195$, $g_{xx} = 1.9848$, $g_{yy} = 1.9829$, $A_{zz} = 17.15$ mT, $A_{xx} = 4.6$ mT and $A_{yy} = 5.2$ mT. Reproduced with permission of the Royal Society of Chemistry from Mabbs FE (1993) Some aspects of the electron paramagnetic resonance spectroscopy of d-transition metal compounds. *Chemical Society Reviews* **22**: 313–324.

## Surface chemistry and catalysis

EPR has been widely applied to surface science and catalysis in order to examine a variety of surface paramagnetic species important in catalytic processes including adsorbed atoms, ions or molecules that may be intermediates in chemical reactions, intrinsic surface defects, transition metal ions supported on an oxide and spin labels interacting with a surface. Information regarding the important physicochemical characteristics of the surface can be gained

through analysis of the EPR spectrum. Properties such as surface crystal field, surface redox properties, surface group morphology, mobility of adsorbed species, coordination of surface metal ions and identification of the active site have been studied on a variety of surfaces.

In many cases, the surface itself is diamagnetic and investigations by EPR rely on indirect methods. A large amount of information regarding the state of a surface can be gained by adsorption of probe molecules (i.e. a molecule whose properties in the adsorbed state can be monitored and evaluated by EPR). These molecules may be divided roughly into two classes. The first is paramagnetic probes (or probes that become paramagnetic upon adsorption). The spectroscopic features of simple paramagnetic probes such as NO and $O_2^-$ can provide detailed information on surface electrostatic fields. For example, the *g* values of the adsorbed superoxide anion ($O_2^-$) are known to depend not only on the nominal charge of the surface cation where adsorption occurs but also on the coordination environment of the cation itself, as shown in **Figure 4**. Paramagnetic probes such as $VCl_4$ or $MoCl_5$ retain their paramagnetism during reaction with surface hydroxyl groups, allowing one to study the distribution of the surface groups themselves. The other type are diamagnetic probes that remain diamagnetic upon adsorption. These probes allow one to investigate the properties of a paramagnetic adsorption site, such as the changes in the coordination sphere of a surface transition metal ion.

Metal oxides are frequently involved in catalysis, either directly or indirectly as supports. By studying the charge transfer reactions at the gas–solid or liquid–solid interface, the nature, number and strength of the acidic or basic catalyst can be investigated. The oxidizing properties of the oxide can be studied by formation of the radical cations of aromatic molecules such as anthracene and perylene, while the reducing properties of the surface can be monitored through the formation of the radical anions of molecules with large electron affinity such as tetracyanoethylene (TCNE) or nitroaromatic compounds. The EPR spectra of the corresponding radical cations and anions yield both qualitative and quantitative information on the redox nature of the surface and the subject has been reviewed.

Finally, a molecular approach to catalysis requires a deep understanding of the coordination processes occurring at the catalyst surface. The use of EPR to investigate the various types of coordination chemistry that occur on catalytic systems involving transition metal ions and oxide matrices have been

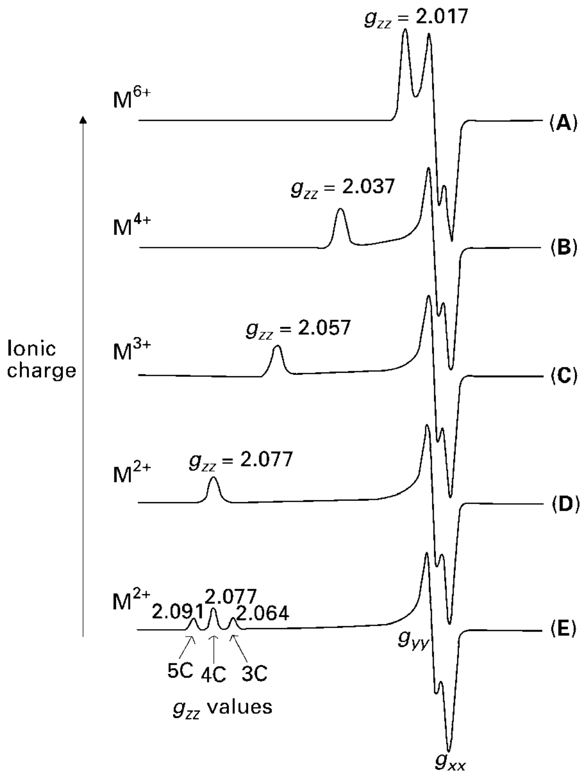



**Figure 4** Simulated EPR spectra of the superoxide ($O_2^-$) anion adsorbed on differently charged cationic sites. The $g_{zz}$ values of the radical have the form $g_{zz} = g_e + 2\lambda/\Delta$ where $\lambda$ is the spin–orbit coupling constant for oxygen and $\Delta$ is the extent of the energy splitting between the two $\pi^*$ orbitals, caused by the local electrostatic field of the cation. As the charge varies, $\Delta$ will vary and different $g_{zz}$ values are observed (A–D). The anion is also sensitive enough to distinguish between cations with the same nominal charge which, because of different coordinative environments (e.g. 5-coordinated, 4C or 3C surface cations), exert slightly different electrostatic fields (E).

reviewed. It has been shown that the reactivity and properties of the central transition metal ion in the heterogeneous system are heavily influenced by the coordinating ligands (much more than in the homogeneous counterparts) so that a variety of TMI complexes can be created and stabilized at the surface. Changes to the EPR spectroscopic features as a function of variation in the coordination sphere of a surface-supported TMI can easily be followed. An example of this is shown in **Figure 5** for a $Ni^+$ supported catalyst. Changes in the spectroscopic features of the signal as a function of differing CO pressures illustrate the sensitivity of the technique to perturbations in the local environment. This coordination chemistry approach to EPR studies of the mechanism of catalytic reactions is illustrated in the dimerization of olefins and the oxidation of methanol on the surface of oxide catalysts.

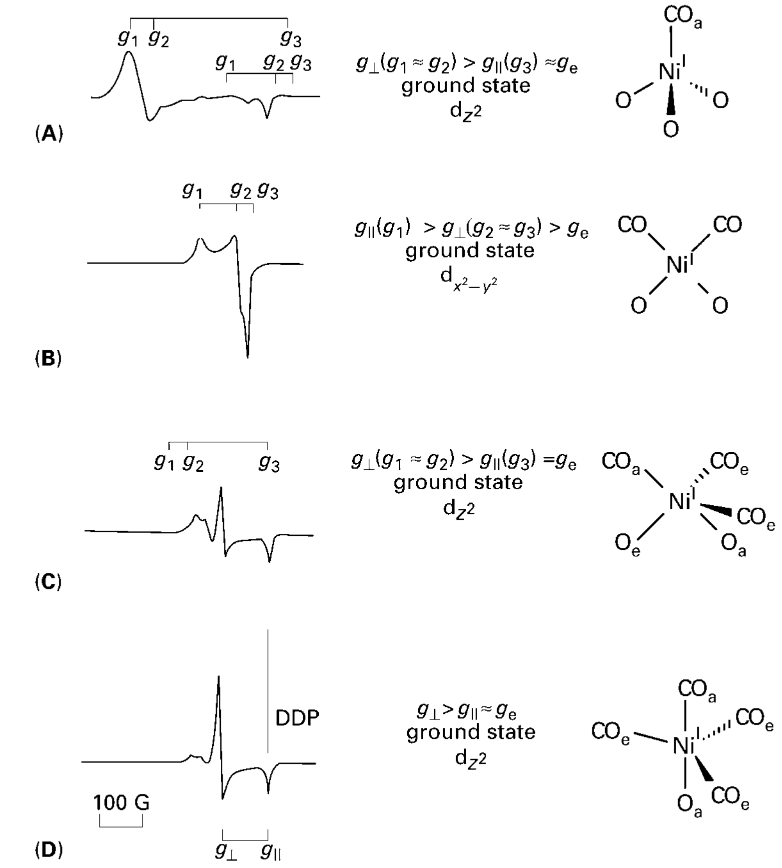



**Figure 5** X-band EPR spectra of $Ni^+$ complexes formed upon CO adsorption on reduced $Ni/SiO_2$ catalysts under a pressure of (A) 10 torr CO followed by evacuation at 340 K, (B) 10 torr, (C) 100 torr, (D) 400 torr. Subscript a represents axial and subscript e represents equitorial positions. Reprinted with permission from Dyrek K and Che M (1997) EPR as a tool to investigate the transition metal chemistry on oxide surfaces. *Chemical Reviews* **97**: 303–331 Copyright 1997 American Chemical Society.

## Applications of pulsed techniques and high-frequency EPR

The applications discussed above were performed using conventional CW (continuous wave) methods, predominantly at X-band (~9 GHz) frequency. However, in recent years EPR spectroscopy has experienced an extraordinarily rapid development in two noted areas, namely pulsed techniques and high-frequency EPR. Already both developments have proved to be invaluable tools in important areas of physics, chemistry and biology, and no article on the applications of EPR spectroscopy is complete without an explanation of the powerful attributes of these techniques.

The difference between high-frequency and low-frequency EPR is usually defined by the magnet system. Superconducting or Bitter magnets are employed for high-frequency EPR (such as those operating at 94 GHz), whereas traditional electromagnets are used for lower frequencies (e.g. Q-band (~35 GHz) to L-band (~1 GHz)). High-frequency EPR offers numerous advantages over the low-frequency spectrometers, including increased resolution of *g* factors (which is particularly important in complex or disordered systems where spectral lines are not resolved at lower fields) and increased absolute sensitivity (which is important in studies of small single crystals or any system where the number of spins is inherently low). High-frequency EPR also allows measurement of a number of integer-spin systems (such as Ni(II) species) which often have very high zero-field splittings. These systems only start to become resolved at high frequency where the energy of the photon is above the zero-field energy. The number of applications is growing rapidly, but the most compelling areas of development are in complex disordered systems prevalent in structural biology, including new insights into the detailed mechanisms associated with photosynthesis and many metalloproteins. Because EPR is only sensitive to the sites with unpaired spins, it targets only the most

important bonds that largely determine the chemical and physical properties of the metalloprotein. High-frequency EPR is also particularly appropriate for studying these complex enzymes since the active sites often involve metal ions or metal clusters with high zero-field splittings, which are EPR-silent at low frequency. Furthermore, most disordered systems are better studied at high frequencies because the improved spectral resolution allows the local structure to be inferred, especially in conjunction with other techniques such as ENDOR.

Exploitation of pulsed techniques in chemical research has been hampered in the past by expensive instrumentation and the lack of sufficiently fast digital electronics. The situation has undoubtedly changed in recent years, and the various pulsed EPR methods have found increased or exclusive use in areas such as time-resolved EPR, as tools for studying molecular motion, in electron–nuclear double resonance spectroscopy and in EPR imaging. Fourier transform EPR has proved particularly useful for studying the various spin polarization mechanisms involved in free radical chemistry, while the very high time resolution involved in echo-detected EPR makes it an ideal technique for measuring rates of chemical reactions and determining relaxation times.

The underlying basis of pulsed EPR, like NMR, requires the generation of a spin echo by a short intense microwave pulse. However, in some cases the coupling between the electron spin and the nuclear spins can cause a modulation of the echo intensity. As this electron spin echo envelope modulation (ESEEM) can only occur if anisotropic interactions are present, it is only found in the solid state. Analysis of the ESEEM spectra allows determination of very small hyperfine couplings which are too small to be measured by CW-EPR, so that information on the number, identity and distance of weakly inter-acting nuclei may be obtained. A wide variety of applications of the nuclear modulation effect have been demonstrated, including determination of ligand structure in metalloproteins, magnetic properties of triplet states and location and coordination of surface complexes. In fact, the ESEEM method ranks with EXAFS as one of the most important tools for determining the structures of active centres in metalloproteins.

## Conclusions

It is intended in this article to demonstrate the diversity of applications in which CW-EPR is traditionally used, while briefly describing some current uses of more advanced EPR techniques. However, with recent technological developments in FT-EPR, pulsed ENDOR, EPR imaging and high-frequency EPR, new and previously undreamed of prospects are likely to emerge. In many cases the spectroscopic data of interest cannot be obtained by any other method. This fact alone will ensure EPR a place as a dominant spectroscopic technique of major importance well into the future.

## List of symbols

$a$ = hyperfine splitting constant, units are mTesla; $a_0$ = isotropic hyperfine coupling constant, units are mTesla; $A_{\parallel}$ = hyperfine coupling constant parallel to a unique symmetry axis, units are mTesla; $A_{\perp}$ = hyperfine coupling constant perpendicular to a unique symmetry axis, units are mTesla; $E_a$ = rotational activation energy, in kJ mol$^{-1}$; $g$ = g factor; $g_{\parallel}$ = parallel to a unique symmetry axis; $g_{\perp}$ = perpendicular to a unique symmetry axis; $\tau$ = free radical tumbling correlational time (in seconds).

*See also:* **EPR Imaging; EPR, Methods; EPR Spectroscopy, Theory; Inorganic Chemistry, Applications of Mass Spectrometry; Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy; Surface Studies By IR Spectroscopy.**

## Further reading

Bales B, Griffiths PC, Goyffon P, Howe AM and Rowlands CC (1997) EPR insights into aqueous solutions of gelatin and sodium dodecylsulphate. *Journal of the Chemical Society, Perkin Transactions 2* 2473–2477.

Berliner LJ and Reubens J (eds) (1989) *Biological Magnetic Resonance*, vol. 8, *Spin Labelling Theory and Applications*. New York: Plenum Press.

Carley AF, Edwards HA, Hancock FE, *et al* (1994) Application of ESR to study the hydrogenation of alkenes and benzene over a supported Pd catalyst. *Journal of the Chemical Society, Faraday Transactions,* **90**: 3341–3346.

Che M and Giamello E (1987) In: Fierro JLG (ed) '*Spectroscopic Characterisation of Heterogeneous Catalysts*', vol. 57, Part B, p B265. Amsterdam: Elsevier Science Publishers.

Dixon WT and Norman ROC (1962) Free radicals formed during the oxidation and reduction of peroxides. *Nature*, **196**, 891.

Dyrek K and Che M. (1997) *Chemical Review* **97**: 305–331.

Egerton TA, Murphy DM, Jenkins CA and Rowlands CC (1997) An EPR study of spin trapped free radical intermediates formed in the heterogeneously assisted photodecomposition of acetaldehyde. *Journal of the Chemical Society, Perkin Transactions 2* 2479–2485.

Marsh D (1994) Spin labelling in biological systems. *Specialist Periodical Reports, Electron Spin Resonance* **14**: 166–202.

Möbius K (1994) EPR studies of photosynthesis. *Specialist Periodical Reports, Electron Spin Resonance* **14**: 203–245.

Ranby B and Rabek JF (1977) *ESR Spectroscopy in Polymer Research*. Berlin: Springer-Verlag (and references therein).

Rhodes CJ (1991) Organic radicals in solid matrices. *Specialist Periodical Reports, Electron Spin Resonance* **13a**: 56–103.

Schweiger A (1991) Pulsed electron spin resonance spectroscopy: basic principles, techniques and examples of applications. *Angewandte Chemie, International Edition, English* **30**: 265–292.

Tabner BJ (1991) Organic radicals in solution. *Specialist Periodical Reports, Electron Spin Resonance* **13a**: 1–55.

Wasserman AM (1996) Spin labels and spin probes in polymers. *Specialist Periodical Reports, Electron Spin Resonance* **15**: 112–151.

# Chemical Exchange Effects in NMR

**Alex D Bain**, McMaster University, Hamilton, ON, Canada

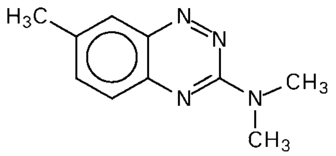
Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Theory**

## Introduction

Chemical exchange, in NMR terms, means that a nucleus moves among a set of magnetic environments. The sample is macroscopically at equilibrium, but an individual nucleus exchanges among a number of sites, so the magnetic properties of the nucleus are modulated by the exchange. Chemical exchange effects were recognized early in the development of NMR: at about the same time (and in the same laboratory) as the scalar coupling. In *N*,*N*′-dimethyl formamide and many other molecules with *N*,*N*′-dimethyl groups (**Figure 1**), the two methyl groups have different chemical shifts if the molecule is rigid. At low temperature this is effectively the case, because of restricted rotation about the C–N bond. However, as the sample is warmed, the two signals broaden, coalesce, and eventually start to sharpen up to a single line at the average chemical shift, as shown in **Figure 1**.

This behaviour is usually interpreted in terms of an 'NMR timescale'. For rotation about the C–N bond, which is slow on this timescale at low temperature, the two signals are distinct and relatively sharp. At higher temperature, the rotation is faster and all that we can detect is the average resonance frequency. The broad coalescence line shape is characteristic of the intermediate timescale, and is perhaps the most familiar and obvious manifestation of chemical exchange. In this article we will concentrate on intermediate exchange, because a thorough




3-Dimethylamino-7-methyl-1,2,4-benzotriazine

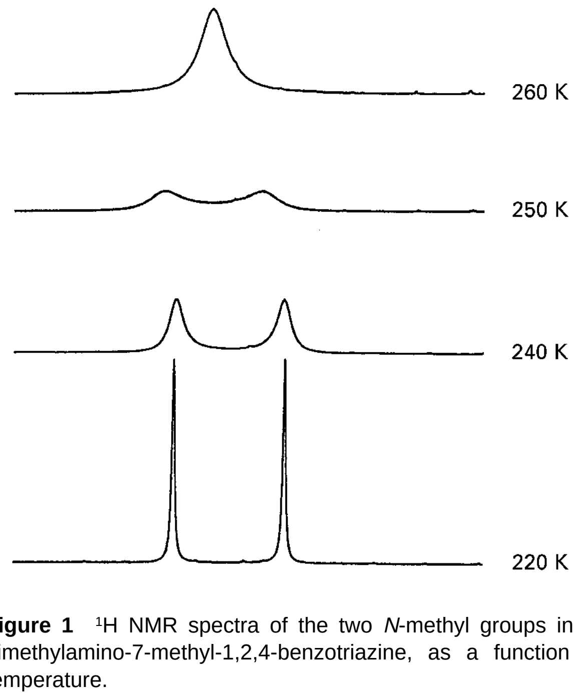



**Figure 1** $^{1}H$ NMR spectra of the two *N*-methyl groups in 3-dimethylamino-7-methyl-1,2,4-benzotriazine, as a function of temperature.

understanding of this case clarifies most other aspects of chemical exchange.

Chemical exchange in all three regimes – slow, intermediate and fast – has an effect on the NMR spectrum. Typical values of the parameters mean that processes with rates in the range of 1–$10^4$ $s^{-1}$ can be studied most easily. In other words, this means activation energies for the reaction in the range of approximately 40–80 kJ $mol^{-1}$ (10–20 kcal $mol^{-1}$). This includes conformational changes in ring structures, restricted rotations about chemical bonds, ligand rearrangements in coordination complexes, and some intermolecular transfer processes. Molecules are dynamic entities, so the effects of chemical exchange are apparent in many NMR spectra.

The theory of chemical exchange in simple; uncoupled spin systems have often been couched in terms of the Bloch equations, and this is the approach used in most of the literature. However, we feel that it is simpler and easier to consider the time domain. This time-domain method allows us to treat all exchanging systems – slow, intermediate or fast, coupled or uncoupled – in a consistent and simple way. There is also a simple physical picture of the spectroscopic transition probability that helps in this interpretation of the theory.

## The Bloch equations approach

The Bloch equations for the motion of the $x$ and $y$ magnetizations (usually called the $u$- and $v$-mode signals), in the presence of a weak radiofrequency field, $B_1$, are given in Equation [1].

$$\begin{aligned}\frac{\mathrm{d}u}{\mathrm{d}t}+\frac{u}{T_2}-(\omega_0-\omega_1)v=0\\ \frac{\mathrm{d}v}{\mathrm{d}t}+\frac{v}{T_2}-(\omega_0-\omega_1)u=\gamma B_1 M_z\end{aligned} \quad [1]$$

In this equation, $\omega_1$ is the frequency of the RF irradiation, $\omega_0$ is the Larmor frequency of the spin, $T_2$ is the spin–spin relaxation time and $M_z$ is the $z$ magnetization of the spin system. We can simplify the notation somewhat by defining a complex magnetization, $M$, as in Equation [2].

$$M=u+\mathrm{i}v \quad [2]$$

With this definition, the Bloch equations can be written as in Equation [3].

$$\frac{\mathrm{d}M}{\mathrm{d}t}+\mathrm{i}(\omega_0-\omega_1)M+\frac{1}{T_2}M=\mathrm{i}\gamma B_1 M_z \quad [3]$$

In chemical exchange, the two exchanging sites, A and B will have different Larmor frequencies, $\omega_A$ and $\omega_B$. The exchange will carry magnetization from A to B and vice versa. If we assume equal populations in the two sites, and the rate of exchange to be $k$, then we can set up two coupled Bloch equations for the two sites, as in Equation [4].

$$\begin{aligned}&\frac{\mathrm{d}M_\mathrm{A}}{\mathrm{d}t}+\mathrm{i}(\omega_\mathrm{A}-\omega_1)M_\mathrm{A}+\frac{1}{T_2}M_\mathrm{A}\\ &\qquad -kM_\mathrm{B}+kM_\mathrm{A}=\mathrm{i}\gamma B_1 M_{z\mathrm{A}}\\ &\frac{\mathrm{d}M_\mathrm{B}}{\mathrm{d}t}+\mathrm{i}(\omega_\mathrm{B}-\omega_1)M_\mathrm{B}+\frac{1}{T_2}M_\mathrm{B}\\ &\qquad -kM_\mathrm{A}+kM_\mathrm{B}=\mathrm{i}\gamma B_1 M_{z\mathrm{B}}\end{aligned} \quad [4]$$

The observable NMR signal is the imaginary part of the sum of the two steady-state magnetizations, $M_A$ and $M_B$. The steady state implies that the time derivatives are zero, and a little further calculation (and neglect of $T_2$ terms) gives the NMR spectrum of an exchanging system as in Equation [5].

$$\begin{aligned}\nu&=\frac{1}{2}\gamma B_1 M_z\\ &\times\frac{k(\omega_\mathrm{A}-\omega_\mathrm{B})^2}{(\omega_\mathrm{A}-\omega_1)^2(\omega_\mathrm{B}-\omega_1)^2+4k^2[\omega_1-(\omega_\mathrm{A}+\omega_\mathrm{B})/2]^2}\end{aligned} \quad [5]$$

This equation can be further extended to systems with unequal populations and more sites, using the same techniques.

## Chemical exchange in the time domain

If we create (by a pulse) the magnetization, $M_A$ and $M_B$, at time zero, and then turn off the $B_1$ magnetic field, Equation [4] can be simplified as in Equation [6].

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}M_\mathrm{A}\\ M_\mathrm{B}\end{pmatrix}=-L\begin{pmatrix}M_\mathrm{A}\\ M_\mathrm{B}\end{pmatrix} \quad [6]$$

$$L=\begin{pmatrix}\mathrm{i}\delta+\frac{1}{T_2}+k & -k\\ -k & -i\delta+\frac{1}{T_2}+k\end{pmatrix} \quad [7]$$

In this equation, the matrix $L$ is given by Equation [7], where we have made $\omega_1 = (\omega_A + \omega_B)/2$ and $\delta = (\omega_A - \omega_B)/2$.

$$\begin{pmatrix} M_A(t) \\ M_B(t) \end{pmatrix} = \exp(-Lt) \begin{pmatrix} M_A(0) \\ M_B(0) \end{pmatrix} \quad [8]$$

Equation [6] is a set of first-order differential equations, so its formal solution is given by Equation [8], in which exp( ) means the exponential of the matrix. In practice, we diagonalize the matrix $L$ with a matrix of eigenvectors, $U$, as in Equation [9] to give a diagonal matrix, $\Lambda$, with the eigenvalues of $L$ down the diagonal.

$$\Lambda = U^{-1}LU \quad [9]$$

Equation [8] becomes Equation [10]. The exponential of a diagonal matrix is again a diagonal matrix with exponentials of the diagonal elements, as in Equation [11].

$$\begin{pmatrix} M_A(t) \\ M_B(t) \end{pmatrix} = U\exp(-\Lambda t)U^{-1} \begin{pmatrix} M_A(0) \\ M_B(0) \end{pmatrix} \quad [10]$$

$$\begin{pmatrix} M_A(t) \\ M_B(t) \end{pmatrix} = U \begin{pmatrix} e^{-\lambda_1 t} & 0 \\ 0 & e^{-\lambda_2 t} \end{pmatrix} U^{-1} \begin{pmatrix} M_A(0) \\ M_B(0) \end{pmatrix} \quad [11]$$

As was mentioned above, the observed signal is the imaginary part of the sum of $M_A$ and $M_B$, so Equation [11] predicts that the observed signal will be the sum of two exponentials, evolving at the frequencies $\lambda_1$ and $\lambda_2$. This is the free induction decay (FID). In the limit of no exchange, the two frequencies are simply $\omega_A$ and $\omega_B$, as we would expect. When $k$ is nonzero, the mathematics becomes slightly more complicated.

If we ignore relaxation and exchange, then $L$ is a Hermitian matrix with real eigenvalues and eigenvectors. However, when the exchange is important, the Hermitian character is lost and the eigenvalues and eigenvectors have both real and imaginary parts. The eigenvalues are given by the roots of the characteristic equation (Eqn [12]).

$$\begin{vmatrix} i\delta + \dfrac{1}{T_2} + k - \lambda & -k \\ -k & -i\delta + \dfrac{1}{T_2} + k - \lambda \end{vmatrix} = 0 \quad [12]$$

The eigenvalues of Equation [7] are given in Equation [13]

$$\lambda = \left(\frac{1}{T_2} + k\right) \pm \sqrt{k^2 - \delta^2} \quad [13]$$

These eigenvalues are the (complex) frequencies of the lines in the spectrum, as in Equation [11]: the imaginary part gives the oscillation frequency and the real part gives the rate of decay. If $k < \delta$ (slow exchange) then there are two different imaginary frequencies, which become $\pm\delta$ in the limit of small $k$ (see **Figure 2**). In fast exchange, when $k$ exceeds the shift difference, $\delta$, the quantity in the square root in [13] becomes positive, so the roots are pure real. This means that the spectrum is still two lines, but they are both at the average chemical shift, and have different widths. The full expressions for these line shapes are given in the Appendix.

Because of the role of the eigenvectors in Equation [11], the factor (amplitude) multiplying the complex exponential is itself complex. The magnitude of the complex amplitude gives the intensity of the line and its phase gives the phase of the line (the mixture of absorption and dispersion). In slow exchange, the

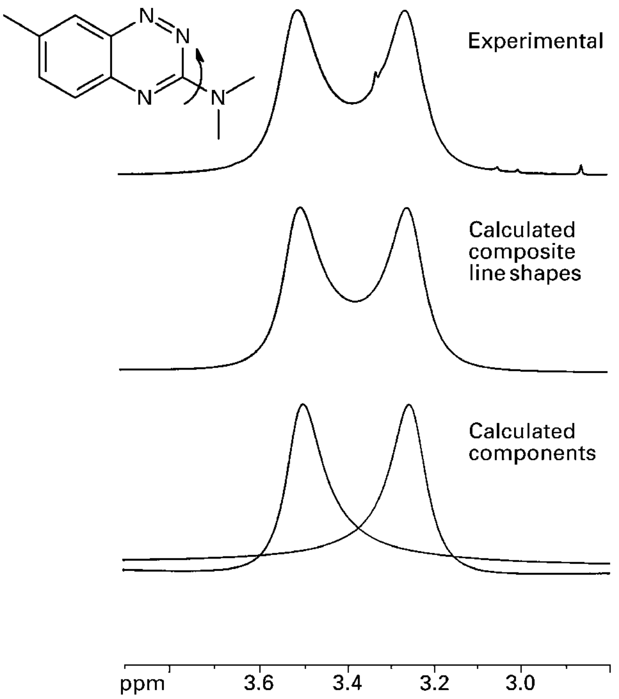



**Figure 2** Decomposition of the coalescence line shape into individual lines. Top: experimental spectrum from **Figure 1**; middle: calculated line shape to match experimental spectrum; bottom: individual lines calculated as in the Appendix.

two lines have the same real part, but the imaginary parts have opposite signs, so the phase distortion is opposite. The sum of these distorted line shapes gives the familiar coalescence spectrum, as in **Figure 2**. In fast exchange, the two lines are both in phase, but one line is negative. This negative line is very broad, and decreases in absolute intensity as the rate increases, leaving only the single, positive, in-phase line for fast exchange.

Therefore, the real and imaginary parts of the frequency, and the real and imaginary parts of the amplitude, provide the four parameters that define a line in an NMR spectrum: its intensity, its phase, its position, and its width. This gives us a time-domain picture of the chemical exchange, which can easily be converted to a spectrum by using the Fourier transform.

## Systems with scalar coupling

Scalar coupled systems are more complicated, but fundamentally no different from the uncoupled systems described in the time domain. If there is scalar coupling in the spectrum, the line shape becomes more complex. For instance, **Figure 3** shows the line shape of the diacylpyridine ligand as part of a rhenium complex. The rhenium bonds to the pyridine nitrogen and to only one of the carbonyls, lifting the symmetry of protons 3 and 5. However, the rhenium can break this latter coordination and bond to the other carbonyl, which effectively interchanges protons 3 and 5 on the pyridine ring. During this exchange, they retain their coupling to the central proton 4, and the shift of 4 does not change. Therefore, proton 4 remains as a sharp triplet even when protons 3 and 5 are very broad.

In any strongly coupled spin system, each line in the spectrum is a mixture of transitions of various nuclei, which depends on the chemical shifts and coupling constants. When a chemical exchange happens, all the spectral parameters change with it. Therefore, a magnetization (or a coherence) that was associated with a single spectral line in one site may be spread among several lines in the other site. In order to deal with these complexities, we must use the density matrix.

One way of thinking of the density matrix is that it is a list of all the observables of a spin system. For example, the magnetizations of the two exchanging sites form part of the density matrix for that system. At equilibrium, the density matrix is just the $z$ magnetization created by the static magnetic field. In a pulse Fourier transform NMR experiment, this $z$ magnetization is flipped into the $xy$ plane, and

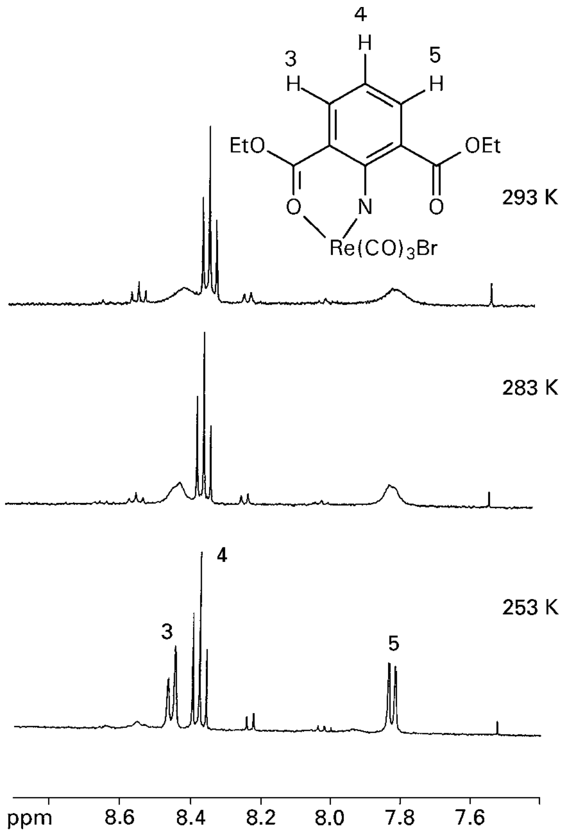



**Figure 3** Example of chemical exchange in a coupled spin system. The rhenium can bond to one or other of the carbonyl oxygen atoms, so that the symmetry of the molecule is lifted. Exchange of the rhenium between carbonyl groups broadens the signals of protons 3 and 5 (at 8.46 ppm and 7.82 ppm), but leaves 4 (8.38 ppm) unaffected.

divided among the individual lines in the spectrum, which then precess around the $z$ axis.

In order to calculate the spectrum of the system (or any other observable), we need to be able to follow the density matrix as a function of time. The equation of motion of the density matrix (ρ) is given in Equation [14], where $H$ is the Hamiltonian of the spin system.

$$\mathrm{i}\frac{h}{2\pi}\frac{\partial}{\partial t}\rho = [H, \rho] \qquad [14]$$

Some manipulation allows us to reformulate Equation [14] as Equation [15]. Mathematically, this means using superoperators in Liouville space $\boldsymbol{L}$, but the details need not concern us here. The important point is that the density matrix becomes a vector of all possible observables of the system – in this case we only deal with the $xy$ magnetizations. Anything

we do to the system, pulses, free precession, relaxation, exchange, is represented by a matrix.

$$\mathrm{i}\frac{h}{2\pi}\frac{\partial}{\partial t}\rho = \boldsymbol{L}^{\mathrm{so}}\ \rho \qquad [15]$$

If we use frequency units ($h/2\pi = 1$), then the solution to Equation [15] is given in Equation [16], which is identical to Equation [8].

$$\rho(t) = \exp(-\mathrm{i}\ \boldsymbol{L}^{\mathrm{so}}t)\ \rho(0) \qquad [16]$$

Relaxation or chemical exchange can be easily added in Liouville space, by including a Redfield matrix, $\boldsymbol{R}^{\mathrm{so}}$ (so = superoperator), for relaxation, or a kinetic matrix, $\boldsymbol{K}^{\mathrm{so}}$, to describe exchange. Both relaxation and exchange are described very conveniently by superoperators. The complete equation of motion becomes Equation [17]. Note the similarity to the equations for two uncoupled sites, but now we derived them rigorously from the equations of motion of the density matrix.

$$\begin{aligned}\rho(t) &= \exp(-\mathrm{i}\boldsymbol{L}^{\mathrm{so}} - \boldsymbol{R}^{\mathrm{so}} - \boldsymbol{K}^{\mathrm{so}})t\ \rho(0)\\ &= U\exp(-\Lambda t)U^{-1}\ \rho(0)\end{aligned} \qquad [17]$$

The details of what the density matrix is at time zero, what we observe in the receiver, and how we construct the superoperators for a general spin system are all beyond the scope of this article. However, the basic idea remains: even the most complex exchange line shape is a sum of individual lines, with complex frequencies and complex intensities, as in Equation [17]. Each of those lines has a transition probability.

## Physical interpretation of the transition probability

The standard expression for the probability of a transition from an initial state, $\psi$, to a final state, $\phi$, induced by a perturbation $I_x$ (the RF field along $x$), is given in Equation [18].

$$\text{Transition probability} \propto |<\phi\mid I_x\mid\psi>|^2 \qquad [18]$$

We can write the matrix element in Equation [18] in a rather more complicated way, as in Equation [19].

$$<\phi\mid I_x\mid\psi>=\ \mathrm{trace}\left\{\begin{pmatrix}<\phi|I_x|\phi> & <\psi|I_x|\phi>\\ <\phi|I_x|\psi> & <\psi|I_x|\psi>\end{pmatrix}\times\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\right\} \qquad [19]$$

The reason for rewriting the equation is that the right hand side can be interpreted as the dot product of the operator $I_x$ with the operator $|\phi\rangle\langle\psi|$. We have moved to Liouville space, in which all the operators are vectors. The dot product in Liouville space is defined as the trace of the product of the operators in the original spin space. One important point is that Equation [19] is independent of basis set. The matrix element in Equation [18] becomes a simple dot product of two Liouville space vectors, as in Equation [20].

$$<\phi\mid I_x\mid\psi>=\begin{pmatrix}<\phi\mid I_x\mid\phi>\\ <\psi\mid I_x\mid\phi>\\ <\phi\mid I_x\mid\psi>\\ <\psi\mid I_x\mid\psi>\end{pmatrix}\bullet\begin{pmatrix}0\\0\\1\\0\end{pmatrix} \qquad [20]$$

A dot product of two vectors gives the projection of one vector on the other, so Equation [20] defines the projection of $|\phi\rangle\langle\psi|$ onto $I_x$. In other words, this is the projection of the transition onto the total $xy$ magnetization. The standard transition probability is the square of this projection.

This leads to a physical interpretation of the transition probability. The spin system at equilibrium is represented by the total $z$ magnetization, $F_z$. This is the sum of the $z$ magnetizations of the sites, weighted by the equilibrium populations of each site. To distinguish the weighted sum, we give this spin operator a symbol different from $I_z$. A pulse flips this into the $xy$ plane, so that immediately after the pulse, the spin system is represented by $F_x$. Now each transition receives a share of the total $x$ magnetization. Its share is given by the projection of the transition (as an operator) onto the operator $F_x$. The transitions evolve independently as a function of time. However, we do not observe the transition directly, but rather the total $xy$ magnetization. The detector is simply a coil of wire, and we measure the $xy$ magnetization as a function of time to give us an FID. Fourier transforming the FID gives us a spectrum. Therefore, each transition contributes to the total signal according to its projection along $I_x$. The intensity of a transition is the product of how much

coherence it received from $F_x$ at the start, times how visible it is to the receiver.

Note that the Liouville matrix, **i*L*+*R*+*K*** may not be Hermitian, but it can still be diagonalized. Its eigenvalues and eigenvectors are not necessarily real, however, and the inverse of ***U*** may not be its complex-conjugate transpose. If we allow complex numbers in it, Equation [17] is a general result. Since $\Lambda$ is a diagonal matrix we can expand in terms of the individual eigenvalues, $\lambda_j$. We can also apply ***U*** ('backwards') to $I_x$, and obtain Equation [21].

$$\text{NMR signal} = \sum_j (UI_x)_j^* \, (U^{-1}F_x)_j \exp(-\mathrm{i}\lambda_j\, t) \quad [21]$$

Each of the terms in the sum is a transition. The NMR signal is always a sum of decaying sine waves, whose frequency and decay rate are given by the imaginary and the real parts of $\lambda_j$. The intensity is governed by the coefficients in Equation [21] of this term. This coefficient is the product of two terms. The first, $(UI_x)_j$, tells us how much the coherence overlaps with the receiver. The second term, $(U^{-1} F_x)_j$ tells us how much the coherence received from the equilibrium $z$ magnetization. The product of these terms is the generalization of the transition probability. If we do not have relaxation, the two terms are complex conjugates, and we recover the normal transition probability. However, the transition probability for a general system may have an imaginary part. If we cling to idea that each term in Equation [21] represents a transition, then an individual transition may be out of phase at $t = 0$. Immediately after the excitation pulse, the total magnetization will lie on the $y$ axis, but the individual components may not. However, our true observable is still only the FID. Whether we want to decompose it into individual transitions of individual spins depends very much on the system we are studying.

## Example of exchange of coupled spins

Provided we can construct the matrices, the description of exchange is quite simple. Again, the details are beyond our scope, but an example of the Liouvillian (no exchange) for a coupled spin system is given in Equation [22] for an AB spin system, with coupling constant $J$. The eigenvalues of this matrix are the familiar line positions of an AB system.

$$\begin{pmatrix} \mathrm{i}\omega_A & \mathrm{i}J/2 & 0 & -\mathrm{i}J/2 \\ \mathrm{i}J/2 & \mathrm{i}\omega_A & -\mathrm{i}J/2 & 0 \\ 0 & -\mathrm{i}J/2 & \mathrm{i}\omega_B & \mathrm{i}J/2 \\ -\mathrm{i}J/2 & 0 & \mathrm{i}J/2 & \mathrm{i}\omega_B \end{pmatrix} \quad [22]$$

If there is an exchange between A and B given by a rate $k$, then we set up the two blocks, as we set up the two exchanging spins in the Bloch equations. Note that we have made spin A in one block, exchange with spin B in the other. The full Liouvillian, including exchange, is given by Equation [23], in which dots replace zeroes to emphasize the form of the matrix.

$$\begin{pmatrix} \mathrm{i}\omega_A - k & \mathrm{i}J/2 & \cdot & -\mathrm{i}J/2 & k & \cdot & \cdot & \cdot \\ \mathrm{i}J/2 & \mathrm{i}\omega_A - k & -\mathrm{i}J/2 & \cdot & \cdot & k & \cdot & \cdot \\ \cdot & -\mathrm{i}J/2 & \mathrm{i}\omega_B - k & \mathrm{i}J/2 & \cdot & \cdot & k & \cdot \\ -\mathrm{i}J/2 & \cdot & \mathrm{i}J/2 & \mathrm{i}\omega_B - k & \cdot & \cdot & \cdot & k \\ k & \cdot & \cdot & \cdot & \mathrm{i}\omega_B - k & \mathrm{i}J/2 & \cdot & -\mathrm{i}J/2 \\ \cdot & k & \cdot & \cdot & \mathrm{i}J/2 & \mathrm{i}\omega_B - k & -\mathrm{i}J/2 & \cdot \\ \cdot & \cdot & k & \cdot & \cdot & -\mathrm{i}J/2 & \mathrm{i}\omega_A - k & \mathrm{i}J/2 \\ \cdot & \cdot & \cdot & k & -\mathrm{i}J/2 & \cdot & \mathrm{i}J/2 & \mathrm{i}\omega_A - k \end{pmatrix} \quad [23]$$

In this particular case, called mutual exchange, the spins simply permute themselves. We can simplify this matrix to half its original size, as in Equation [24], but for nonmutual exchange we must retain Equation [23].

$$\begin{pmatrix} \mathrm{i}\omega_A - k & \mathrm{i}J/2 & k & -\mathrm{i}J/2 \\ \mathrm{i}J/2 & \mathrm{i}\omega_A - k & -\mathrm{i}J/2 & k \\ k & -\mathrm{i}J/2 & \mathrm{i}\omega_B - k & \mathrm{i}J/2 \\ -\mathrm{i}J/2 & k & \mathrm{i}J/2 & \mathrm{i}\omega_B - k \end{pmatrix} \quad [24]$$

Except for its size, this is exactly the same form as in our previous cases. The spectrum can therefore be expressed as a sum of four lines, as in **Figure 4**.

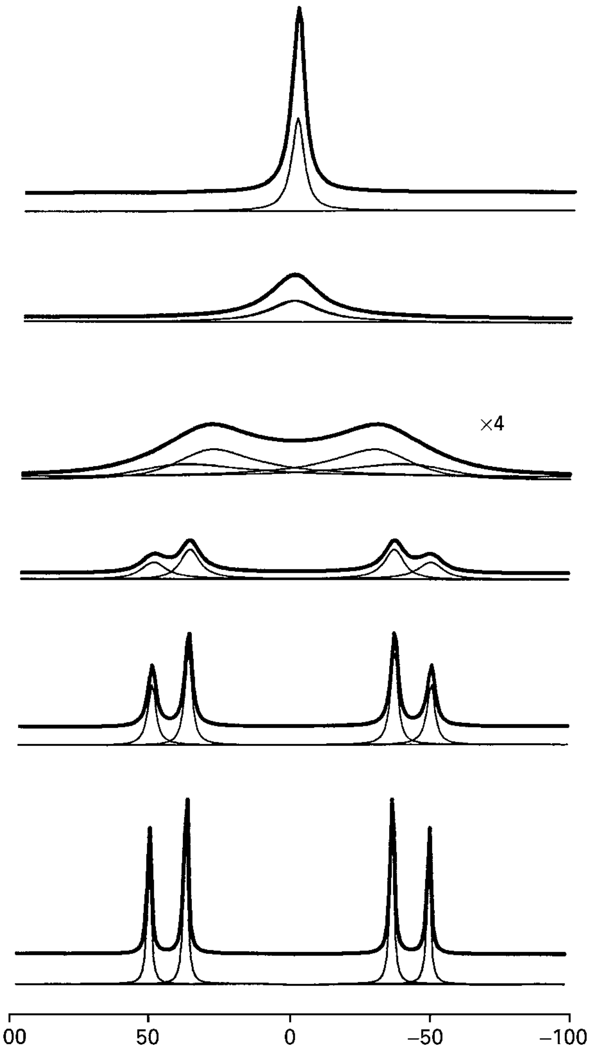



**Figure 4** Calculated spectra for mutual exchange in an AB spin system, as a function of exchange rate. The heavy traces are the total line shape and the lighter lines show the individual components. The bottom spectrum shows the typical static AB spectrum, which broadens and coalesces at higher rates of exchange.

## Slow chemical exchange

The term 'slow' in this case means that the exchange rate is much smaller than the frequency differences in the spectrum, so the lines in the spectrum are not significantly broadened. However, the exchange rate is still comparable with the spin–lattice relaxation times in the system. In this case, we can measure the rates by doing a modified spin–lattice relaxation experiment, pioneered by Hoffman and Forsen.

In the absence of exchange (and let us ignore dipolar relaxation), each $z$ magnetization will relax back to equilibrium at a rate governed by its own $T_1$ as in Equation [25].

$$\frac{\mathrm{d}}{\mathrm{d}t}[M(t) - M(\infty)] = -\frac{1}{T_1}[M(t) - M(\infty)] \qquad [25]$$

If we have two sites, A and B, then we can write the analogous equation for two sites as in Equation [26].

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} M^{\mathrm{A}}(t) - M^{\mathrm{A}}(\infty) \\ M^{\mathrm{B}}(t) - M^{\mathrm{B}}(\infty) \end{bmatrix} = \begin{pmatrix} \frac{1}{T_1^{\mathrm{A}}} & 0 \\ 0 & \frac{1}{T_1^{\mathrm{B}}} \end{pmatrix} \begin{bmatrix} M^{\mathrm{A}}(t) - M^{\mathrm{A}}(\infty) \\ M^{\mathrm{B}}(t) - M^{\mathrm{B}}(\infty) \end{bmatrix} \qquad [26]$$

If the two sites exchange with rate $k$ during the relaxation, then a spin can relax either through normal spin–lattice relaxation processes, or by exchanging with the other site. Equation [26] becomes Equation [27].

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} M^{\mathrm{A}}(t) - M^{\mathrm{A}}(\infty) \\ M^{\mathrm{B}}(t) - M^{\mathrm{B}}(\infty) \end{bmatrix} = \begin{pmatrix} \frac{1}{T_1^{\mathrm{A}}} - k & k \\ k & \frac{1}{T_1^{\mathrm{B}}} - k \end{pmatrix} \begin{bmatrix} M^{\mathrm{A}}(t) - M^{\mathrm{A}}(\infty) \\ M^{\mathrm{B}}(t) - M^{\mathrm{B}}(\infty) \end{bmatrix} \qquad [27]$$

This equation is very similar to Equations [6] and [7]. The basic situation is just as in intermediate exchange, except that now we are dealing with $z$ magnetizations rather than $xy$. The frequencies are zero, and the matrix now has pure real eigenvalues, but the approach is the same. We also stay in the time domain, since a relaxation experiment follows the $z$ magnetizations as a function of time. As before, the time dependence is obtained by diagonalizing the relaxation–exchange matrix, and calculating the magnetizations for each time at which they are sampled.

There are two main applications of slow chemical exchange: one is to determine the qualitative mechanism, and the other is to measure the rates of the processes as accurately as possible. For the first case, in which we have a spectrum in slow exchange, we need to establish the mechanism: which site is

exchanging with which. For this purpose, the homonuclear two-dimensional (2D) experiment EXSY (exchange spectroscopy) (the same pulse sequence as NOESY (2D nuclear Overhauser effect spectroscopy), but involving exchange) is by far the best technique to use. If there is no exchange, then the spectrum lies along the diagonal and there are no cross-peaks. Exchange between sites leads to a pair of symmetrical cross-peaks joining the diagonal peaks of the same site, so the mechanism is very obvious.

The EXSY pulse sequence starts with two $\pi/2$ pulses separated by the incrementable delay, $t_1$. This modulates the $z$ magnetizations, so that the relaxation that occurs during the mixing time which follows, $t_m$, is frequency-labelled. Finally the $z$ magnetizations are sampled with a third $\pi/2$ pulse. Magnetization from a different site that enters via exchange will have a different frequency label. A two-dimensional Fourier transform then produces the spectrum. However, care must be taken in choosing the mixing time if there are multiple exchange processes. If the mixing time is too long, there is a substantial probability that a spin may have exchanged twice in that time, leading to spurious cross-peaks.

For careful rate measurements, once the mechanism is established, it is our opinion that one-dimensional (1D) methods are superior to quantitative 2D ones. Apart from the fact that 1D spectra can be integrated more easily, we also have more control over the experiment. Modern spectrometers can create almost any type of selective excitation, so that we can control the conditions at the start of the relaxation. For two sites, a nonselective inversion that inverts both sites equally will mask most of the exchange effects and the relaxation will be dominated by $T_1$. However, if one site is inverted selectively, then that site can regain equilibrium by either $T_1$ processes or by exchanging with the other site that was left at equilibrium. The inverted signal will relax at roughly the sum of the exchange and spin–lattice relaxation rate, while the signal that was unperturbed at the start of the experiment shows a characteristic transient, as in **Figure 5**. For multiple sites, a wide range of initial conditions is available. Standard nonlinear least-squares methods allow us to fit these curves and derive values for the rates involved.

## Fast exchange

In fast exchange the spectrum has coalesced, usually to a single line, so the information in the spectrum is not so rich. For two equally-populated sites, for instance, we get a single Lorentzian line with a width proportional to $\delta^2/k$. We can measure the line width (or better, $T_2$) via a CPMG (Carr–Purcell pulse sequence, Meiboom–Gill modification) experiment or a $T_{1\rho}$ measurement. However, if we do not know the frequency difference, $\delta$, we cannot obtain the rate. If the exchange rate is not too fast ($< 10^4$ roughly) modifications of the $T_2$ experiments can help. Both experiments have an inherent timescale: in CPMG, it is the timing of the refocussing pulse; in $T_{1\rho}$, it is the precession frequency about the spin-locking field. If the exchange is fast with respect to the experiment,

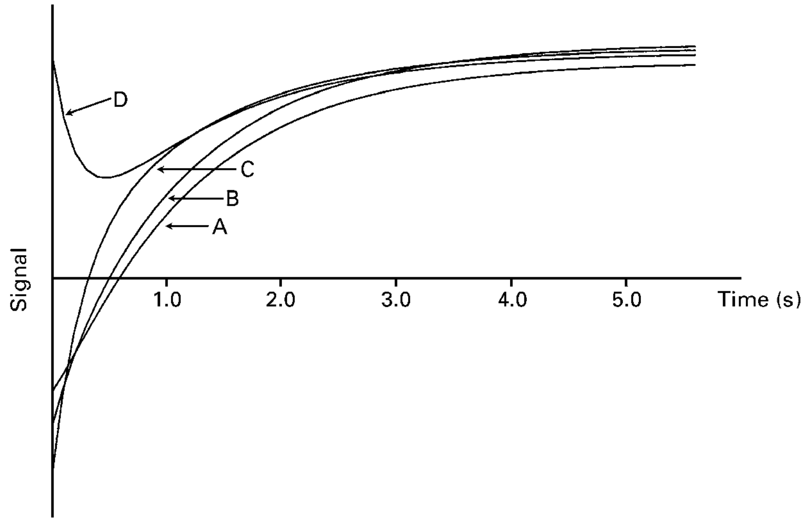



**Figure 5** Inversion–recovery curves for slow exchange between two sites. Lines A and B are the results from one experiment. They show the recovery after a nonselective inversion of both sites, showing that the two sites have slightly different $T_1$ values. C and D are obtained from a different experiment. Line C shows the recovery of a site that has been selectively inverted, and line D shows the behaviour of the line that was not perturbed in this experiment. The inverted line relaxes faster, due to a combination of spin–lattice, relaxation and exchange, and the unperturbed line shows the characteristic transient.

we measure a $T_2$ appropriate to the coalesced spectrum. If, however, the exchange is slower than the experimental timescale (but still fast with respect to the frequency difference), the apparent $T_2$ reflects the individual sites. As we change the timescale of the experiment, the apparent $T_2$ changes, and so we can obtain a value for the rate itself.

## Conclusions

All of the effects of chemical exchange can be calculated by following the appropriate magnetizations as a function of time. For slow exchange, this is the complex coupling of relaxation and exchange that leads to transient behaviour in modified spin–lattice relaxation experiments. For intermediate exchange, the rather complex line shapes can always be decomposed into a sum of individual transitions, even though the transitions are distorted in phase, intensity, position, and width by the dynamics of the system. For two uncoupled sites, this can be calculated quite easily, but for larger coupled systems the construction of the matrices may get complicated. However, the resulting lines are always governed by a transition probability, provided we extend our definition to allow probabilities with both real and imaginary parts. Working in the time domain is already familiar from Fourier transform spectroscopy, and it provides a complete and simple approach to chemical exchange.

## Appendix

### Eigenvalues and eigenvectors for two-site exchange

For two-site equally populated chemical exchange, the two magnetizations at time zero are equal, so that Equation [11] can be written as in Equation [28], where the eigenvalues are complex numbers.

$$\begin{pmatrix} M_A(t) \\ M_B(t) \end{pmatrix} = U \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} U^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad [28]$$

Since we are dealing with nonHermitian matrices, the matrix formed by the eigenvectors will not be unitary, and will have four independent complex elements. Let us simply call them $a$, $b$, $c$, and $d$, so that $U$ is given by:

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad [29]$$

Regardless of whether $U$ is unitary, its inverse is given by Equation [30], where $\Delta$ is the determinant of Equation [29].

$$U^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad [30]$$

Equation [28] then says that the signal is given by the following, regardless of slow or fast exchange.

$$\text{Signal} = \frac{(a+c)(d-b)}{\Delta} \exp(\lambda_1 t) + \frac{(b+d)(-c+a)}{\Delta} \exp(\lambda_2 t) \qquad [31]$$

The values of the eigenvectors have two forms, depending on whether $\delta > k$ (slow exchange), or $\delta < k$ (after coalescence). In the first case, the eigenvalues are given as:

$$\text{Eigenvalues} = -\left(\frac{1}{T_2} + k\right) \pm \mathrm{i}\sqrt{\delta^2 - k^2} \qquad [32]$$

and a convenient matrix of eigenvectors is given by:

$$\begin{pmatrix} k & \mathrm{i}(\sqrt{\delta^2 - k^2} + \delta) \\ -\mathrm{i}(\sqrt{\delta^2 - k^2} + \delta) & k \end{pmatrix} \qquad [33]$$

After coalescence, when the rate is greater than the frequency difference, the two transitions are both at zero frequency (i.e. the average chemical shift), but have different widths and intensities. The eigenvalues are pure real, as in Equation [34],

$$\text{Eigenvalues} = -\left(\frac{1}{T_2} + k\right) \pm \sqrt{k^2 - \delta^2} \qquad [34]$$

and the eigenvectors are similar, but reflect the fact that $k^2 - \delta^2$ is now positive.

$$\begin{pmatrix} \sqrt{k^2 - \delta^2} - \mathrm{i}\,\delta & -\sqrt{k^2 - \delta^2} - \mathrm{i}\,\delta \\ k & k \end{pmatrix} \qquad [35]$$

## List of symbols

$B_1$ = radiofrequency field; $F_z$ = total $z$ magnetization; $h$ = Planck's constant; $I_x$ = perturbation RF field along $x$; $i = \sqrt{-1}$; $k$ = rate of exchange; $K^{so}$ = kinetic superoperator matrix; $L$ = Liouville superoperator; $L^{so}$ = superoperator matrix; $M_z = z$ magnetization of the spin system; $R^{so}$ = Redfield superoperator matrix; $T_2$ = spin–spin relaxation time; $T_{1\rho} = T_1$ in rotating frame; $u = x$ magnetization of spin system; $v = y$ magnetization of spin system; $\gamma$ = magnetogyric ratio; $\delta$ = chemical shift difference; $\lambda$ = eigenvalue; $\Lambda$ = diagonal matrix; $\phi$ = final state; $\psi$ = initial state; $\omega_1$ = frequency of RF irradiation; $\omega_0$ = Larmor frequency of the spin.

*See also:* **NMR Principles; NMR Pulse Sequences; NMR Relaxation Rates; Two-Dimensional NMR, Methods.**

## Further reading

Bain AD (1988) The superspin formalism for pulse NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* **20**: 295–315.

Bain AD and Duns GJ (1996) A unified approach to dynamic nmr based on a physical interpretation of the transition probability. *Canadian Journal of Chemistry* **74**: 819–824.

Binsch G (1969) A unified theory of exchange effects on nuclear magnetic resonance lineshapes. *Journal of the American Chemical Society* **91**: 1304–1309.

Binsch G and Kessler H (1980) The kinetic and mechanistic evaluation of NMR spectra. *Angewandte Chemie, International Edition in English* **19**: 411–494.

Gutowsky HS and Holm CH (1956) Rate processes and nuclear magnetic resonance spectra. II. Hindered internal rotation of amides. *Journal of Chemical Physics* **25**: 1228–1234.

Jackman LM and Cotton FA (1975) *Dynamic Nuclear Magnetic Resonance Spectroscopy*. New York: Academic Press.

Johnson CS (1965) Chemical rate processes and magnetic resonance. *Advances in Magnetic Resonance* **1**: 33–102.

Kaplan JI and Fraenkel G (1980) *NMR of Chemically Exchanging Systems*. New York: Academic Press.

Orrell KG, Sik V and Stephenson D (1990) Quantitative investigation of molecular stereodynamics by 1D and 2D NMR methods. *Progress in Nuclear Magnetic Resonance Spectroscopy* **22**: 141–208.

Perrin CL and Dwyer T (1990) Application of two-dimensional NMR to kinetics of chemical exchange. *Chemical Review* **90**: 935–967.

Sandstrom J (1982) *Dynamic NMR Spectroscopy*. London: Academic Press.

Sorensen OW, Eich GW, Levitt MH, Bodenhausen G and Ernst RR (1983) Product operator formalism for the description of NMR pulse experiments. *Progress in Nuclear Magnetic Resonance Spectroscopy* **20**: 163–192.

# Chemical Ionization in Mass Spectrometry

**Alex G Harrison**, University of Toronto, Ontario, Canada

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Methods & Instrumentation |

## Introduction

In chemical ionization mass spectrometry (CIMS), ionization of the gaseous analyte occurs via gas-phase ion–molecule reactions rather than by direct electron impact (EI), photon impact or field ionization. EI ionization of a reagent gas (present in large excess) is usually followed by ion–molecule reactions involving the initially formed ions and the reagent gas neutrals to produce the chemical ionization (CI) reagent ion or reagent ion array. Collision of the reagent ion(s) with the analyte (usually present at ~1% of the reagent gas pressure) produces one or more ions characteristic of the analyte. These initial analyte ions may undergo fragmentation or, infrequently, react further with the reagent gas to produce a final array of ions representing the CI mass spectrum of the analyte as produced by the specific reagent gas.

To a considerable extent, the usefulness of CIMS arises from the fact that a wide variety of reagent gases and, hence, reagent ions can be used to ionize the analyte; often the reagent system can be tailored to the problem to be solved. Problems amenable to solution by CI approaches include (a) molecular

mass determination, (b) structure elucidation and (c) identification and quantification. In many instances, CI provides information that is complementary rather than supplementary to that obtained by EI, and often both approaches are used. After a brief discussion of instrumentation, the major approaches, in both positive ion and negative ion CI, to the solution of the above problems are discussed. The focus is primarily on the use of medium-pressure mass spectrometry; CI at atmospheric pressure or at much lower pressures in ion trapping instruments are discussed elsewhere.

## Instrumentation

Most commonly, CI studies have been performed at total ion source pressures of 0.1 to 1 torr using conventional sector, quadrupole or time-of-flight mass spectrometers with ion sources only slightly modified from those used for EI ionization. The essential change is that the ion source is made more gas-tight by using smaller electron beam entrance and ion beam exit slits; the latter is accomplished by having either exchangeable source volumes or a moveable ion exit slit assembly to change from EI to CI. Most manufacturers now supply instruments capable of both EI and CI and incorporating enhanced pumping to maintain adequately low pressures in the remainder of the instrument when the source is operated at elevated pressures. Provision is made for introduction of the reagent gas and for introduction of the analyte by a heated inlet system, by a solids probe, by a direct exposure probe or by interfacing a gas chromatograph to the ion source. Capillary gas chromatography–CI mass spectrometry is a particularly powerful and sensitive approach to identification and quantification of the components of complex mixtures.

## Brønsted acid chemical ionization

Gaseous Brønsted acids, $BH^+$, react with the analyte M primarily by the proton transfer reaction

$$BH^+ + M \rightarrow MH^+ + B \quad [1]$$

Other, usually minor, reactions that are possible include hydride ion abstraction (Eqn [2]) and charge exchange (Eqn [3])

$$BH^+ + M \rightarrow [M - H]^+ + BH_2(B + H_2) \quad [2]$$

$$BH^+ + M \rightarrow M^{+\bullet} + BH^\bullet \ (B + H^\bullet) \quad [3]$$

The enthalpy change for Equation [1] is given by

$$\Delta H^o = PA(B) - PA(M) \quad [4]$$

where the proton affinity of X, PA(X), is given by

$$X_{(g)} + H^+_{(g)} \rightarrow XH^+_{(g)}, \quad -\Delta H^o = PA(X) \quad [5]$$

If PA(M) > PA(B), the reaction is exothermic and normally occurs at the ion–neutral collision rate determined by ion-induced dipole and ion–dipole interactions. Conversely, if the reaction is endothermic, the rate decreases exponentially by the enthalpy of activation and the sensitivity of the ionization process decreases.

Most of the exothermicity of Equation [1] resides in $MH^+$ and is available as internal energy to promote fragmentation of $MH^+$. To maximize the probability of $MH^+$ formation, which provides molecular mass information, one uses the least exothermic reaction possible. On the other hand, structural information is derived from the fragment ions observed in the CI mass spectrum and more exothermic protonation may be desirable to promote such fragmentation. The fragmentation of the even-electron $MH^+$ ions usually involves elimination of even-electron stable molecules to form an even-electron fragment ion. Thus if the molecule contains a functional group Y, the fragmentation of $MH^+$ frequently involves elimination of the stable molecule HY. For $RYH^+$, the tendency to eliminate HY is roughly inversely related to the PA of HY, with the result that the ease of loss of Y as HY is approximately

$$NH_2 < CH_3C(=O)O < CH_3S < CH_3O < C_6H_5$$
$$\approx HC(=O)O < CN < SH < OH < I \approx Cl \approx Br$$

although the stability of the resulting ion also plays a role in competing loss of different functional groups from a common $MH^+$. Obviously, the more readily fragmentation occurs, the less likely it is that $MH^+$ ions, giving molecular mass information, will be observed, although with suitable choice of reagent ions, $MH^+$ ions are frequently formed when no $M^{+\bullet}$ ions are observed in the EI mass spectrum. In other cases, such as hydroxylic compounds, the use of Brønsted base CI (see below) may prove useful in obtaining molecular mass information.

**Table 1** lists Brønsted acid reagent systems which have seen significant use, along with the major reactant ions and the proton affinities of the conjugate bases. Most organic molecules have PAs in the range

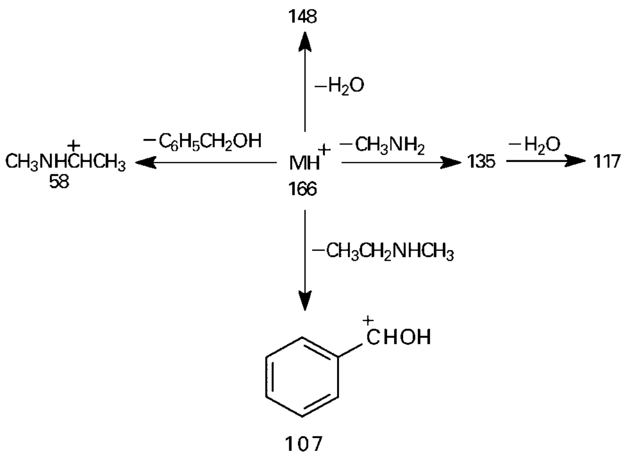
760–1000 kJ mol$^{-1}$. Thus protonation by $H_3^+$ will be strongly exothermic, while protonation in *i*-butane or ammonia CI will only be mildly exothermic or even endothermic. $CH_4$ as a reagent gas is unique in that it forms two reactant species, $CH_5^+$ and $C_2H_5^+$, in almost equal abundance; the polar reagent gases $H_2O$, $CH_3OH$ and $NH_3$ produce solvated protons with the extent of solvation depending on the partial pressure of the reagent gas. The usefulness of Brønsted acid CI in establishing molecular masses is illustrated by the comparison of the EI, $CH_4$ CI and $NH_3$ CI mass spectra of ephedrine in **Figure 1**. No molecular ion is observed in the EI mass spectra. By contrast, both the $CH_4$ CI and $NH_3$ CI mass spectra show abundant $MH^+$ (*m*/*z* 166) ions, clearly establishing the molecular mass. In line with the relative PAs of **Table 1**, the less exothermic protonation in $NH_3$ CI leads to less extensive fragmentation of the $MH^+$ ion. As is often the case, $CH_4$ CI provides not only molecular mass information ($MH^+$) but also structurally informative fragment ions; the fragmentation of protonated ephedrine is rationalized in **Scheme 1** and is seen to involve the formation of even-electron fragment ions by the elimination of small, stable, even-electron neutral molecules (the numbers in **Scheme 1** indicate the *m*/*z* of the ionic fragment). In this case the $CH_4$ CI mass spectrum provides more structural information than the EI mass spectrum.

The effect of protonation exothermicity on the extent of fragmentation is shown more dramatically by the Brønsted acid CI mass spectra of the tripeptide H-Val-Pro-Leu-OH shown in **Figure 2** and obtained using mass-selected reactant ions. With $NH_4^+$ as the reactant ion [$PA(NH_3) = 853.5$ kJ mol$^{-1}$] only $MH^+$ is observed, while the more exothermic protonation by $C_2H_5^+$ [$PA(C_2H_4) = 680.3$ kJ mol$^{-1}$] leads to significant fragmentation of $MH^+$, and the very exothermic protonation by $N_2OH^+$ [$PA(N_2O) = 580.7$ kJ mol$^{-1}$] leads to essentially complete fragmentation of $MH^+$.

**Table 1** Brønsted acid reagent systems

| *Reagent gas* | *Reactant ion* ($BH^+$) | *PA(B)* (kJ mol$^{-1}$) |
|---|---|---|
| $H_2$ | $H_3^+$ | 423.4 |
| $N_2O–H_2$ | $N_2OH^+$ | 580.7 |
| $CH_4$ | $CH_5^+$ | 550.6 |
| | $C_2H_5^+$ | 680.3 |
| $H_2O$ | $H^+(H_2O)_n$[a] | 696.0 |
| $CH_3OH$ | $H^+(CH_3OH)_n$[a] | 773.6 |
| *i*-$C_4H_{10}$ | $C_4H_9^+$ | 819.6 |
| $NH_3$ | $H^+(NH_3)_n$[a] | 853.5 |

[a] Degree of solvation depends on partial pressure of reagent gas; proton affinity given for monosolvated proton.




**Scheme 1**

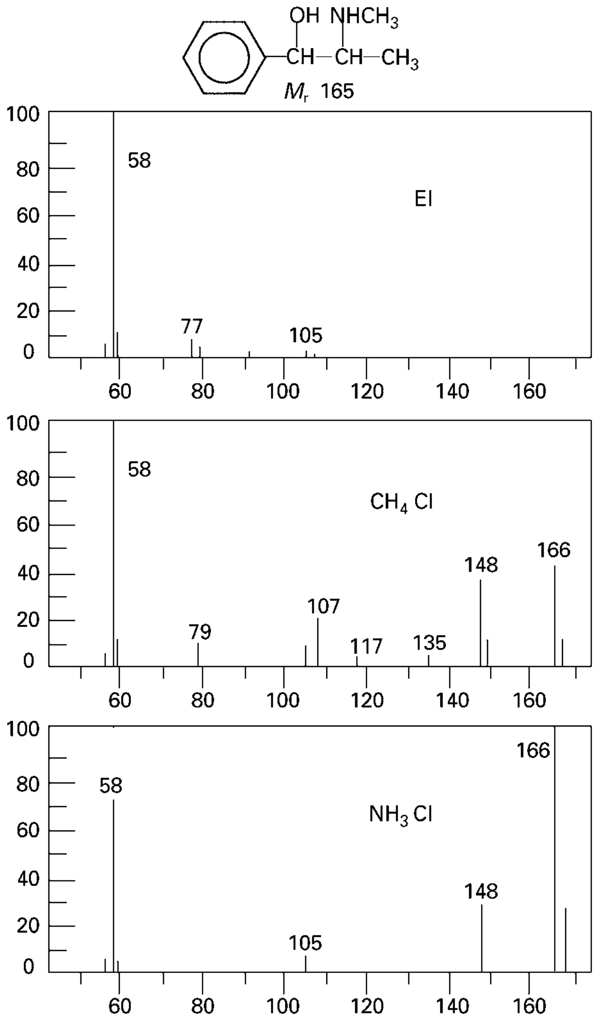



**Figure 1** EI, $CH_4$ CI and $NH_3$ CI mass spectra of ephedrine (relative intensity as a function of *m*/*z*).

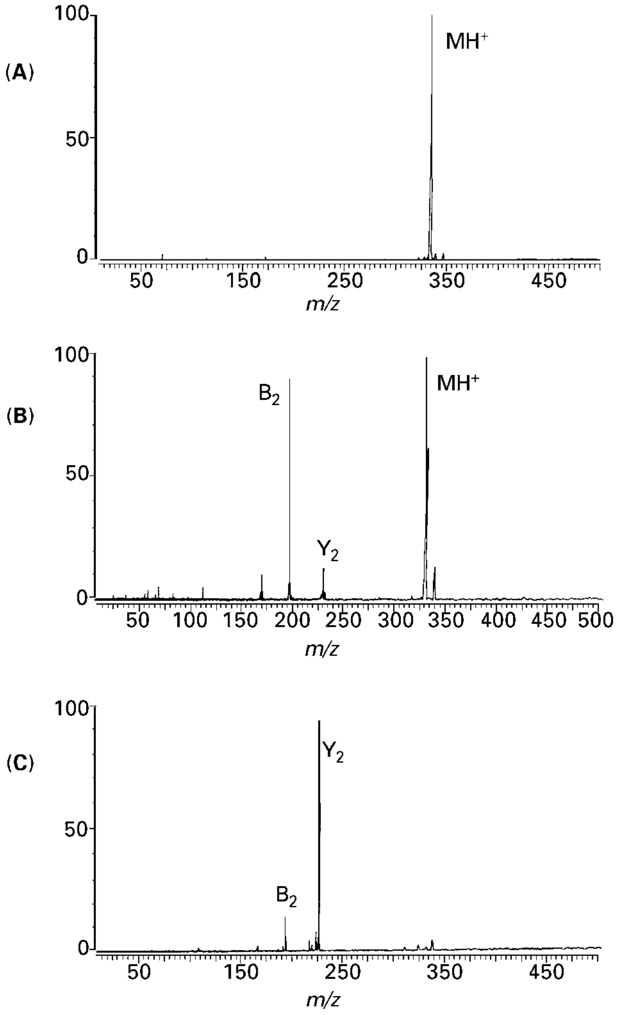



**Figure 2** Brønsted acid CI mass spectra of H-Val-Pro-Leu-OH using (A) $NH_4^+$, (B) $C_2H_5^+$ and (C) $N_2OH^+$ as reagent ions. Reprinted with permission from Speir JP, Gorman GS, Cornett DS and Amster IJ (1991) Controlling the dissociation of peptide ions using laser desorption/chemical ionization Fourier transform mass spectrometry. *Analytical Chemistry* **63**: 65–69. Copyright (1991) American Chemical Society.

Ammonia has a very high PA, with the result that $NH_4^+$ (and solvated forms thereof) will efficiently protonate only those analytes with a PA greater than about 860 kJ mol$^{-1}$; in effect, this means that only nitrogen-containing analytes (such as **Figures 1** and **2**) are efficiently protonated. Analytes with lower PAs, which contain a lone pair of electrons, frequently give $M^{\bullet}NH_4^+$ adducts.

## Charge exchange chemical ionization

In charge exchange CI, the reagent gas is chosen to produce an odd-electron species $R^{+\bullet}$ on EI ionization. This ion reacts with the analyte M by charge exchange to produce, initially, the odd-electron molecular ion (Eqn [6]), which may undergo further fragmentation (Eqn [7]).

$$R^{+\bullet} + M \rightarrow M^{+\bullet} + R \qquad [6]$$

$$M^{+\bullet} \rightarrow \text{fragments} \qquad [7]$$

Since the $M^{+\bullet}$ ion formed initially is the same as that formed initially in EI ionization, the fragmentation reactions observed will be similar to those observed in EI mass spectra. The essential difference is that, while the $M^{+\bullet}$ ions formed by electron ionization have a wide range of internal energies, the $M^{+\bullet}$ ions formed by charge exchange have an internal energy, $E(M^{+\bullet})$, given approximately by

$$E(M^{+\bullet}) = RE(R^{+\bullet}) - IE(M) \qquad [8]$$

where $RE(R^{+\bullet})$, the recombination of $R^{+\bullet}$, is defined by

$$R^{+\bullet} + e \rightarrow R, \quad -\Delta H^{o} = RE(R^{+\bullet}) \qquad [9]$$

and IE(M) is the adiabatic ionization energy of M.

Reagent systems have been developed (**Table 2**) giving reactant ions with recombination energies over the range 9.3–16 eV. By comparison, most organic molecules have ionization energies in the range 8–12 eV, so that, by suitable choice of the charge exchange reagent, a wide range of internal energies can be imparted to the molecular ion in Equation [6]. A potential use of charge exchange ionization is to enhance differences in the mass spectra of isomeric molecules compared to EI ionization; the use of charge exchange has the advantage, compared to low-energy EI, that the sensitivity of ionization is maintained. This use of charge exchange is illustrated in **Figure 3**

**Table 2** Charge exchange reagent systems

| *Reagent gas* | *Reactant ion ($R^{+\bullet}$)* | *$RE(R^{+\bullet})$(eV)* |
|---|---|---|
| $C_6H_6$ | $C_6H_6^{+\bullet}$ | 9.3 |
| $N_2$–$CS_2$ | $CS_2^{+\bullet}$ | ~10.0 |
| $CO_2$–$C_6F_6$ | $C_6F_6^{+\bullet}$ | 10.2 |
| CO–COS | $COS^{+\bullet}$ | 11.2 |
| Xe | $Xe^{+\bullet}$ | 12.1, 13.4 |
| $CO_2$ | $CO_2^{+\bullet}$ | 13.8 |
| Kr | $Kr^{+\bullet}$ | 14.0, 14.7 |
| $N_2$ | $N_2^{+\bullet}$ | 15.3 |
| Ar | $Ar^{+\bullet}$ | 15.8, 15.9 |

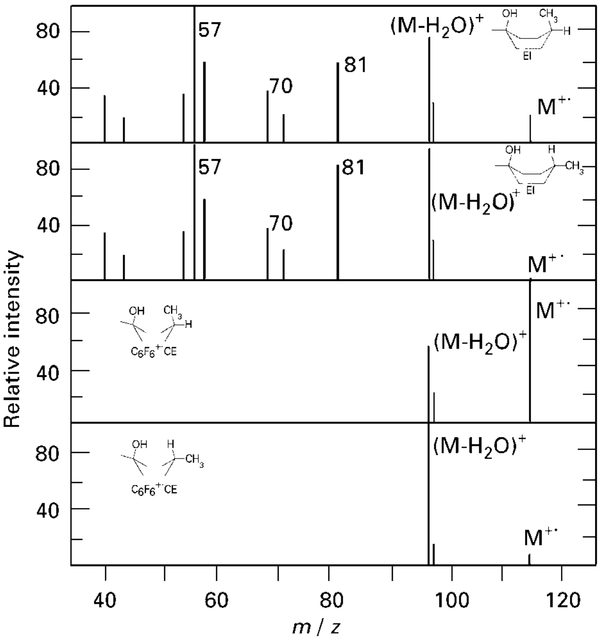



Figure 3 EI and $C_6F_6^{+\bullet}$ charge exchange (CE) mass spectra of *cis*- and *trans*-4-methylcyclohexanol. Harrison AG and Lin MS, (1984) Stereochemical applications of mass spectrometry. 3. Energy dependence of the fragmentation of stereoisomeric methyl cyclohexanols. *Organic Mass Spectrometry* **19**: 67–71. Copyright John Wiley & Sons Limited. Reproduced with permission.

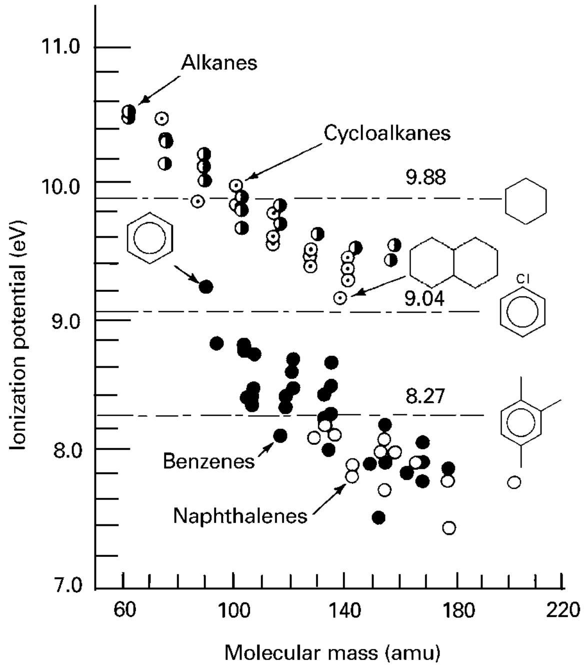



**Figure 4** Ionization potentials of hydrocarbons as a function of molecular mass. Reprinted with permission from Sieck LW (1983) Determination of molecular weight distributions of aromatic components in petroleum products with chlorobenzene as reagent gas. *Analytical Chemistry* **55**: 38–41. Copyright (1983) American Chemical Society.

which compares the $C_6F_6^{+\bullet}$ charge exchange mass spectra of *cis*- and *trans*-4-methylcyclohexanol with 70 eV mass spectra. The low-energy charge exchange spectra are much simpler; more importantly, the differences in the spectra of the two isomers are much enhanced following charge exchange ionization. A particularly useful application of charge exchange ionization is for the selective ionization of specific components of mixtures. The principle of the method is illustrated in **Figure 4** which shows the ionization energies of alkanes, cycloalkanes, alkylbenzenes and alkylnaphthalenes (components of petroleum products) plotted as a function of molecular mass. Also shown, as horizontal lines, are recombination energies of $c$-$C_6H_{12}^{+\bullet}$, $C_6H_5Cl^{+\bullet}$ and $(CH_3)_3C_6H_3^{+\bullet}$. These species should ionize only those hydrocarbons with ionization below the appropriate line. Clearly, $C_6H_5Cl^{+\bullet}$ will ionize only the benzene and naphthalene components, while $(CH_3)_3C_6H_3^{+\bullet}$ will be even more selective for the naphthalenes and higher molecular mass benzenes.

## Brønsted base chemical ionization

Gaseous Brønsted bases, $B^-$, usually react with organic analytes M by proton abstraction to produce $[M–H]^-$ as the initial product.

$$B^- + M \rightarrow [M - H]^- + BH \qquad [10]$$

Equation [10] is normally rapid provided it is exothermic, i.e. provided $\Delta H^\circ_{acid}(BH) > \Delta H^\circ_{acid}(M)$ or, alternatively, provided $PA(B^-) > PA([M-H]^-)$, where

$$BH_{(g)} \rightarrow B_{(g)}{}^- + H_{(g)}{}^+$$
$$\Delta H^\circ = \Delta H^\circ_{acid}(BH) = PA(B^-) \qquad [11]$$

In general, little fragmentation of $[M-H]^-$ is observed, with the result that Brønsted base CI often provides molecular mass information through formation of the $[M-H]^-$ ion that is 1 amu lower than the molecular mass.

**Table 3** lists possible Brønsted base reagents along with the $\Delta H^\circ_{acid}$ of the conjugate acid BH. For comparison, phenols have gas-phase acidities in the range of 1370–1465 kJ mol$^{-1}$, carboxylic acids have acidities in the range of 1350–1422 kJ mol$^{-1}$, alkynes in the range of 1520–1580 kJ mol$^{-1}$ and alkylbenzenes about 1590 kJ mol$^{-1}$. Also listed in the Table are the electron affinities of the neutral species

B; these are relevant to the tendency of $B^-$ to react by electron transfer to form $M^{-\bullet}$. These electron affinities are quite high with the exception of $O_2^{-\bullet}$ which is the only reactant likely to react by electron transfer to M.

Of the bases listed $OH^-$ has seen by far the greatest use. It can be prepared by electron bombardment of a $CH_4$–$N_2O$ (~ 90:10) mixture (at usual CI source pressures) through the reaction sequence

$$e_{therm} + N_2O \rightarrow O^{-\bullet} + N_2 \quad [12]$$

$$O^{-\bullet} + CH_4 \rightarrow OH^- + CH_3^\bullet \quad [13]$$

where the $CH_4$ also serves to thermalize the electrons. Alternatively, the addition of $CH_3ONO$ to the methane bath gas leads to formation of $CH_3O^-$. The $Cl^-$ and $Br^-$ ions can be prepared by dissociative electron capture by halogen-containing compounds such as $CH_2Cl_2$ or $CH_2Br_2$. The $F^-$ ion is usually prepared by dissociative electron capture by $NF_3$. The $OH^-$ ion is a strong base and will abstract a proton from most organic analytes to form $[M - H]^-$ Thus, $OH^-$ CI can be usefully employed to obtain molecular mass information. For example, the $OH^-$ CI mass spectra of ephedrine shows $[M - H]^-$ (*m*/*z* 164) as the dominant ion, and $OH^-$ CI could be used as readily as $CH_4$ CI or $NH_3$ CI (**Figure 1**) to establish the molecular mass. More importantly, Brønsted base CI can often be used to provide molecular mass information in cases where Brønsted acid CI fails. This is illustrated in **Figure 5** which compares the $CH_4$ CI and $OH^-$ CI mass spectra of 2,3-cyclopropyl-1,4-undecanediol. While the $CH_4$ CI mass spectrum is quite uninformative (as is the EI mass spectrum), the $OH^-$ CI mass spectrum shows an abundant $[M - H]^-$ (*m*/*z* 199) ion, as well as fragmentation by sequential loss of two water molecules from $[M - H]^-$. The hydroxyl ion can react with some RX molecules by an $S_N2$ reaction to give $X^-$ and does react with esters, R′ COOR, in part, by addition/elimination to give $R'COO^-$

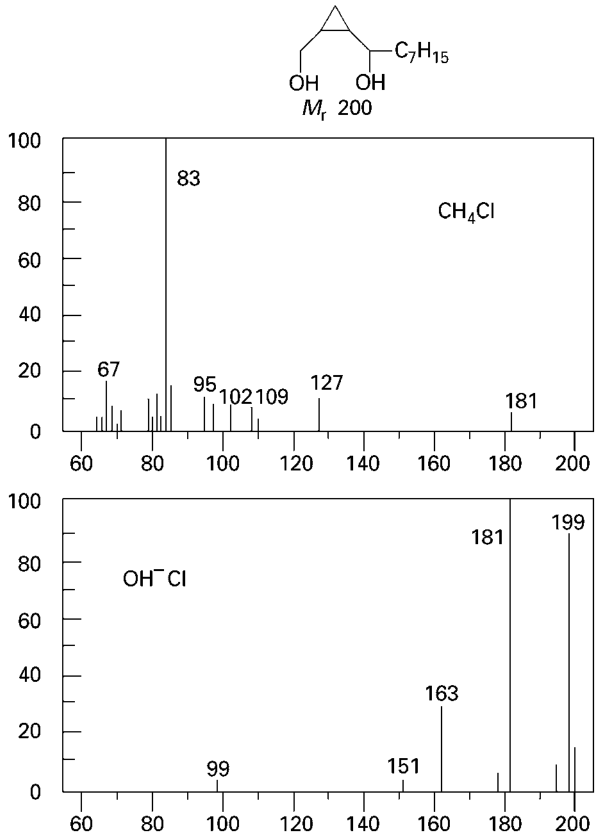



**Figure 5** $CH_4$ CI and $OH^-$ CI mass spectra of 2,3-cyclopropyl-1,4-undecanediol (relative intensity as a function of *m*/*z*).

+ ROH. The reactions of $CH_3O^-$ have not been extensively studied; since it has a similar PA to that of $OH^-$ one would expect the reactions to be similar. The $F^-$ ion is a somewhat weaker base than $OH^-$ or $CH_3O^-$ and thus is somewhat more selective in its proton abstraction reactions. $Cl^-$ is quite a weak base and will abstract only very acidic protons; in other cases $[M + Cl]^-$ adducts have been observed, which has proven useful in providing molecular mass information.

The $O^{-\bullet}$ ion, which can be produced by electron bombardment of a $N_2$–$N_2O$ (~ 9:1) mixture, is unique in that it is both a strong Brønsted base and a radical. As a result, it reacts with many organic analytes not only by proton abstraction but also by H-atom abstraction, $H_2^{+\bullet}$-abstraction and H-atom and alkyl group displacement. This variety of reactions is illustrated in **Figure 6** for $C_5$ ketones. The $[M - H_2]^{-\cdot}$ in ketones results to a significant extent by $H_2^{+\bullet}$ abstraction from one α-carbon; the resultant ion undergoes further fragmentation by loss of the opposite alkyl group (see *m*/*z* 41, **Figure 6**, for example). Alkyl group displacement results in formation of carboxylate ions for ketones. Clearly $O^{-\bullet}$ CI provides

**Table 3** Brønsted base reagents

| $B^-$ | $\Delta H^\circ_{acid}$ (BH) (kJ mol$^{-1}$) | EA(B) (kJ mol$^{-1}$) |
|---|---|---|
| $NH_2^-$ | 1690 | 75.3 |
| $OH^-$ | 1636 | 176.6 |
| $O^{-\cdot}$ | 1598 | 141.0 |
| $CH_3O^-$ | 1594 | 151.5 |
| $F^-$ | 1552 | 328.0 |
| $O_2^{-\bullet}$ | 1477 | 42.3 |
| $Cl^-$ | 1393 | 348.9 |
| $Br^-$ | 1356 | 324.7 |

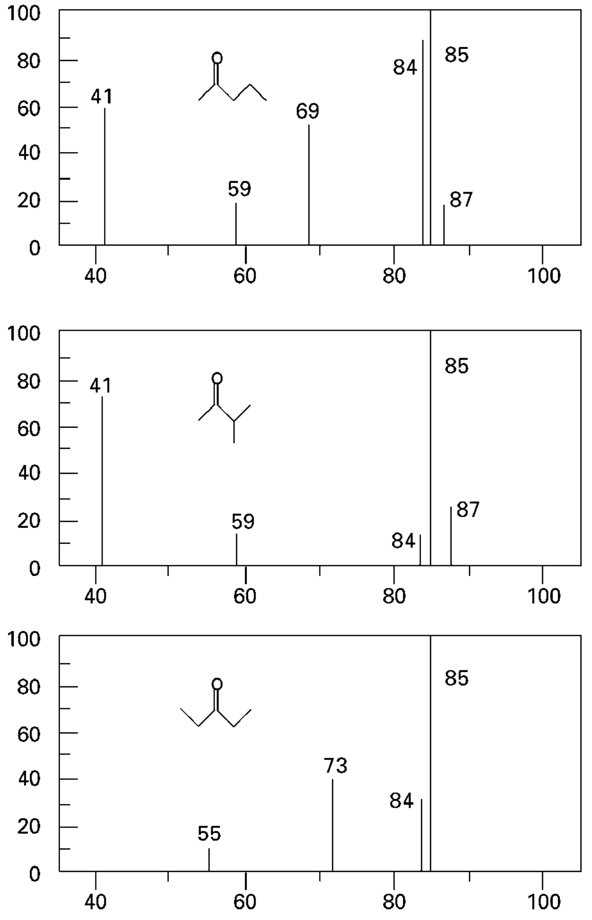



**Figure 6** $O^{-\bullet}$ CI mass spectra of $C_5$ ketones ($M_r$ 86). Data from Marshall A, Tkaczyk M and Harrison AG (1991) $O^{-\bullet}$ chemical ionization of carbonyl compounds. *Journal of the American Society of Mass Spectrometry* **2**: 292–298.

a greater possibility of isomer distinction than $OH^-$ CI which produces only $[M-H]^-$. However, $O^{-\bullet}$ CI usually has a considerably lower sensitivity than $OH^-$ CI or Brønsted acid CI.

## Electron capture chemical ionization

In electron capture CI, ionization of the analyte M occurs by electron capture (Eqn [14]) or by dissociative electron capture (Eqn [15]).

$$e_{therm} + M \rightarrow M^{-\bullet} \qquad [14]$$

$$e_{therm} + M \rightarrow F^- + N \qquad [15]$$

Both reactions are resonance processes that require electrons of near-thermal energy to occur efficiently. Typically, a high-pressure (0.1–1.0 torr) buffer gas is used to 'thermalize' the electrons (emitted from a heated filament) by inelastic scattering and positive ion ionization processes. Ideally, the buffer gas should not form stable anions, and buffer gases that have been used include He, $H_2$, $N_2$, $CH_4$, $NH_3$, $CO_2$ and $i$-$C_4H_{10}$. Polyatomic gases such as $CH_4$, $NH_3$ $i$-$C_4H_{10}$ or $CO_2$ have proven to be most efficient in promoting electron capture. The first three have seen the most use, in part because they are often used in Brønsted acid CI and are readily available.

The major attraction of electron capture chemical ionization (ECCI) is the prospect of enhanced sensitivity compared to other forms of CI or electron ionization. Whereas the maximum rate constants for the ion–molecule reactions involved in Brønsted acid or Brønsted base CI are in the range (2–4) × $10^{-9}$ $cm^3$ $molecule^{-1}$ $s^{-1}$, electron capture rate constants can be as high as $10^{-7}$ $cm^3$ $molecule^{-1}$ $s^{-1}$ because of the high mobility of the electron. Since the sensitivity of the ionization process depends, in part, on the rates of the ionization processes, in favourable circumstances ECCI can be up to 100 times more sensitive than other forms of ionization; this increase can be important when detection and quantification at trace levels is desired. At the same time the rate constants for electron capture are very compound-sensitive and can be very low, with the result that ECCI does show variable and, sometimes, very low sensitivity. It appears that only analytes with appreciable electron affinities (perhaps 0.5 eV) provide reasonable sensitivity in ECCI. The presence of halogen or nitro groups (particularly when attached to $\pi$-bonded systems), a conjugated carbonyl system or highly conjugated carbon systems leads to reasonably high electron affinities. Thus, as examples, dinitrophenols, dinitroanilines and polychlorinated aromatic compounds show good sensitivity in ECCI.

Analytes that are not suitable candidates for electron capture can often be made so by suitable derivatization. Perfluoracyl, pentafluorobenzyl, pentafluorobenzoyl and nitrobenzyl derivatives are often used not only to increase the electron capture rate constants but also to enhance chromatographic properties. However, in contrast to electron capture gas chromatography, it is not sufficient that electron capture be facile; in addition, ions characteristic of the analyte must be formed since ions characteristic only of the derivatizing agent leaves the identity of the analyte in question. This is illustrated by the data in **Table 4** for derivatization of phenols, where it can be seen that the perfluoroacyl derivatives yield ions characteristic only of the derivatizing agent, whereas the pentafluorobenzyl and pentafluorobenzoyl derivatives yield ions characteristic of the analyte. Also

**Table 4** Derivatization of phenols[a]

| Derivative | Relative intensity | | Other ions (relative intensity) |
|---|---|---|---|
| | $M^{-\bullet}$ | $ArO^-$ | |
| $-COCF_3$ | 0 | 0.2 | $CF_3COO^-$ (100) |
| $-COC_2F_5$ | 0 | 0 | $C_2F_5COO^-$(100), $C_2F_4CO^-$ (23) |
| $-COC_3F_7$ | 0 | 0 | $C_3F_7CO^-$ (100), $C_3F_6CO^-$ (6) |
| $-CH_2C_6F_5$ (1)[b] | 0 | 100 | |
| $-COC_6F_5$ (0.4)[b] | 100 | 5 | $C_6F_5^-$(8) |
| $-COC_6H_3(CF_3)_2$ (0.08)[b] | 100 | 2 | |

[a] Data from Trainor TM and Vouros P (1987) Electron capture negative ion chemical ionization mass spectrometry of derivatized chlorophenols and chloroanilines. *Analytical Chemistry* **59**: 601–610.

[b] Relative molar response.

included in the Table are the relative sensitivities (relative molar response) for the three derivatives which give ions identifying the analyte; clearly, the pentafluorobenzyl or pentafluorobenzoyl derivatives are to be preferred in this case.

In addition to variable sensitivity, two further problems that can be encountered in ECCI are nonreproducibility of the spectra and the observation of artifact peaks, i.e. ions that are not explainable simply on the basis of electron capture by the analyte. ECCI mass spectra tend to depend rather strongly on experimental conditions such as buffer gas identity and pressure, source temperature, purity of the buffer gas, amount of analyte introduced, instrument used and instrument focusing conditions. The variability of the spectra obtained is illustrated by the example in **Figure 7**. At 100°C source temperature with $CH_4$ buffer gas, nordiazepam shows $M^{-\bullet}$ (*m*/*z* 170) as the only significant ion, while at 150°C source temperature $Cl^-$ is the dominant ion, and with 1 part in 2000 of oxygen in the methane, predominant formation of $[M - H]^-$ (*m*/*z* 169) is observed. This amount of oxygen could easily be in the system if there is a small air leak. The occurrence of artifact peaks is illustrated in **Figure 8** which shows the ECCI spectra of tetracyanoquinodimethane with $CH_4$ and $CO_2$ as buffer gases. The major peaks in the spectrum with $CH_4$ buffer gas represent artifact peaks arising from reaction of the analyte with free radicals produced on electron bombardment of methane; these altered species are subsequently ionized by electron capture. By contrast, the spectrum with $CO_2$ buffer gas shows only $M^{-\bullet}$. $CO_2$ clearly is preferred in this case although when oxidation of the analyte occurs (such as formation of fluorenone from fluorene) the use of $CO_2$ as a buffer gas leads to substantial artifact peaks. It appears that the most

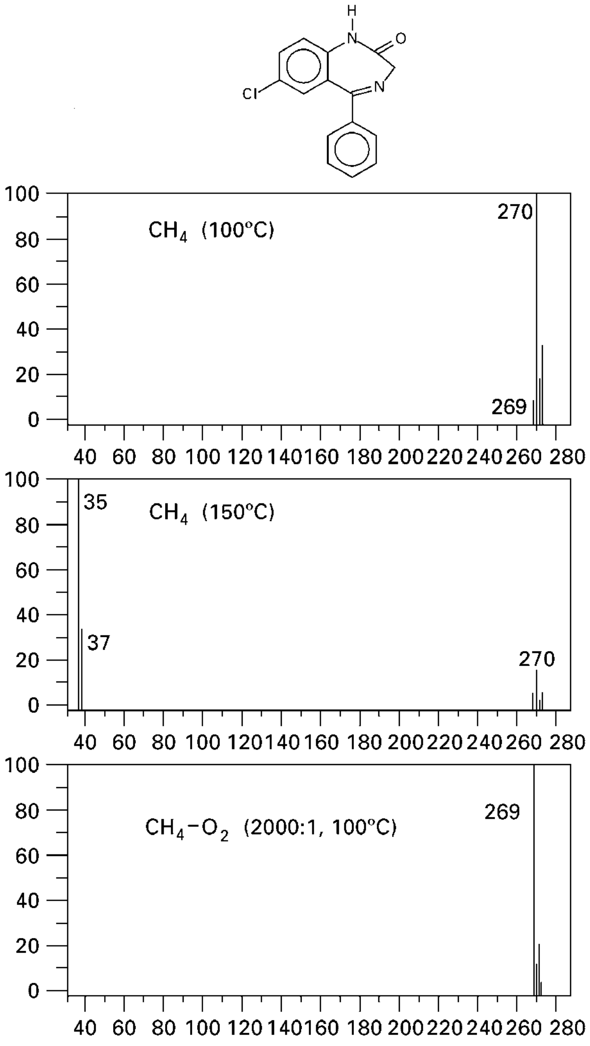



**Figure 7** Electron capture CI mass spectra of nordiazepam under different conditions (relative intensity as a function of *m*/*z*). Data from Garland WA and Miwa BJ (1983) *Biomedical and Environmental Mass Spectrometry* **10**: 126–129.

prevalent pathway to these artifacts involves a surface-catalysed reaction of the analyte with some component of the reagent gas plasma, which produces one or more species with a higher cross section for electron capture than the original analyte. Alteration of the analyte presumably occurs when other ionization methods are used; the difference is that in other modes of CI all species are ionized with similar efficiencies, while in ECCI the altered species may be preferentially ionized.

Despite these difficulties with ECCI it is often the method of choice when identification and quantification at very low concentrations and/or involving complex mixtures is necessary. In such work it is necessary to control the experimental parameters carefully since they can affect both spectra and sensitivity.

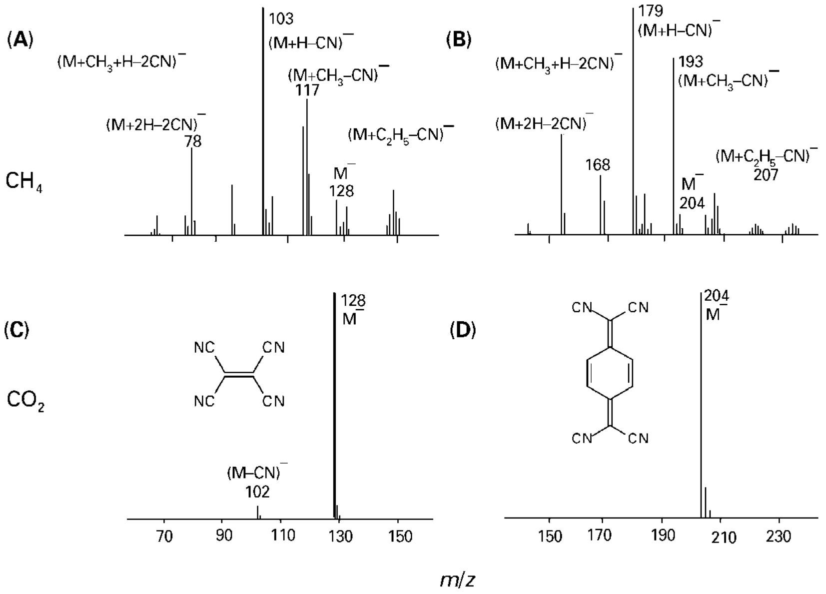



**Figure 8** Electron capture CI mass spectra of tetracyanoethylene and tetracyanoquinodimethane using $CH_4$ and $CO_2$ as buffer gases. Reprinted with permission from Sears LJ and Grimsrud EP (1989) Elimination of unexpected ions in electron capture mass spectrometry using carbon dioxide buffer gas. *Analytical Chemistry*, **61**: 2523–2528. Copyright (1989) American Chemical Society.

## List of symbols

$e_{therm}$ = electron of near-thermal energy; $E$ = energy; $m/z$ = mass to charge ratio ; $\Delta H^{o}$ = enthalpy change.

*See also:* **Ion Energetics in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry.**

## Further reading


Budzikiewicz H (1986) Negative chemical ionization (NCI) of organic compounds. *Mass Spectrometry Reviews* 5: 345–380.

Burrows EP (1995) Dimethyl ether chemical ionization mass spectrometry. *Mass Spectrometry Reviews* **14**: 107–115.

Chapman JS (1993) *Practical Organic Mass Spectrometry*, 2nd edn, Chapters 3 and 4. Chichester, UK: Wiley.

Creaser CS (1995) Chemical ionization in ion trap mass spectrometry. *In*: March RE and Todd JFJ (eds) *Practical Aspects of Ion Trap Mass Spectrometry*, Vol III, Chapter 7. Boca Raton, FL: CRC Press.

Harrison AG (1992) *Chemical Ionization Mass Spectrometry*, 2nd edn. Boca Raton, FL: CRC Press.

Munson MSB and Field FH (1996) Chemical ionization mass spectrometry I. General introduction. *Journal of the American Chemical Society* 88: 2621–2630.

Vairamani M, Mirza UA and Srinivas R (1990) Unusual positive ion reagents in chemical ionization mass spectrometry. *Mass Spectrometry Reviews* 9: 235–258.

Westmore JB and Alauddin MM (1986) Ammonia chemical ionization mass spectrometry. *Mass Spectrometry Reviews* 5: 381–465.

Winkler FJ and Splitter JS (1994) Sterochemical effects in the positive- and negative-ion chemical ionization mass spectra of stereoisomeric molecules. *In*: Splitter JS and Tureček F (eds) *Applications of Mass Spectrometry to Organic Stereochemistry*, Chapter 16. New York: VCH.

# Chemical Reactions Studied By Electronic Spectroscopy

**Salman R Salman**, University of Qatar, Doha, Qatar

Copyright © 1999 Academic Press

ELECTRONIC SPECTROSCOPY
Applications

## Introduction

This article covers applications of UV and visible electronic spectroscopy to studies of chemical reactivity. This encompasses both studies of reaction kinetics and of equilibria and chemical exchange phenomena. Chemical reactions can occur in different phases. Many of the reactions occur in the gas phase and the reactions which take place in this phase will be covered in the section on reaction kinetics.

In solution, the volume occupied by the molecules will be larger than that in the gas phase, thus the molecules will come close together and the freedom for translational motion is restricted. There is another important difference between the solution and the gas phase, which is the great proximity between the solvent molecules and the solute molecules. The solvent can act as a heat sink and energy can pass from the solvent to the solute and this affects the vibrational modes and causes the reactants to overcome activation barriers. Solvent polarity will affect reactions in solutions containing ions. Polar solvents lower the energy required for the formation of ions. In solution many processes and reactions will take place such as diffusion-controlled reactions, reactions of hydrogen and hydroxyl ions, the formation of solvated molecules and electron transfer reactions. The rate coefficients depend on many factors such as the solvent and the formation of a solvent cage, the pressure and the ionic strength of the solution.

Many reactions take place on solid surfaces, and solid materials catalyse such reactions. Surface catalysis is very important in industry, and makes important contributions to some atmospheric reactions. Molecules will be adsorbed on solid surfaces through physicosorption in which the attractive force between the reactant molecules and the surface is of the van der Waals type. Enthalpy of adsorption, $\Delta H_{ad}$, of physicosorption is relatively small, about −40 kJ mol$^{-1}$. With chemical sorption the value for the enthalpy of adsorption is large because a strong bond is formed between the chemical molecule and the surface. Catalytic reactions can be classified as homogeneous catalysis or heterogeneous catalysis.

## UV-visible spectroscopy for kinetics studies

### The problem of variable response factors

UV-visible spectroscopy has proved useful in biochemical analysis, environmental studies, in forensic science, drug kinetics, food quality, identification and quantification of chemical and biological substances but is limited as a tool for the investigation of molecular structure compared with infrared and nuclear magnetic resonance. However, it is quite useful for the study of reaction kinetics and equilibrium.

There is a problem facing the use of a UV-visible spectrum for the calculation of the concentration of different species if two species absorb at the same region, and in this case there is a need to change the reagents to obtain different colours for different species. Also, it is important to carry out a close inspection for the two bands which have the same wavelength. An example of such a system is the absorption of potassium permanganate along with the absorption of a soluble azo dye, which appears as if it is absorbing at the same region, this being near 530 nm. A close inspection and calculations show that $\lambda_{max}$ of $KMnO_4$ and the dye are 533 and 542 nm, respectively, while the extinction coefficients for $KMnO_4$ and the dye are 1.01 and $1.88 \times 10^{14}$ dm$^3$ mol$^{-1}$ cm$^{-1}$, respectively.

### Theory of UV spectroscopy

In UV-visible spectroscopy we are interested in studying the electronic energy. This energy can be calculated from the observed bands in the UV-visible region of the spectrum (195–750 nm). The observed bands in this region are due to the excitation of the electrons by light of a certain wavelength, from the ground state to the excited state. This excitation of the energy of electrons between energy states is in accord with the Frank–Condon principle which states that an electronic transition in a molecule takes place so rapidly compared with the vibrational motion of the nuclei that the internuclear distance can be regarded as fixed during the transition.

The types of electronic transitions are given in **Table 1**. In organometallic compounds a very intense colour can arise due to the $d \rightarrow d$, $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions. This is true especially in transition metals connected to an organic molecule. Several factors, such as substitution (electronic and steric effect), solvent, temperature, and pH will affect electronic transitions and band intensity and shape. In the following paragraphs a brief discussion is given.

## Sample preparation

To obtain useful and accurate information from the UV-visible spectrum it is important to fix appropriate conditions for recording the spectrum. The studied material must be stable and not undergo decomposition by itself under the experimental conditions, which include concentration, solvent and temperature. The solute must dissolve completely in the solvent. It must not be colloidal, because this will given an error in the calculation of the different species present in solution. Also, the solution must not undergo hydrolysis or decomposition in this solvent, otherwise a new species will form which may not be soluble in the solvent.

Cells for the UV region below 330 nm are made of fused silica. In the case of the visible region glass cells may be used. Cells for special measurements can be used and they are classified as sampling cells, flow cells and rectangular cells.

In all measurements it is important to work with concentrations which are within the Beer–Lambert law. There is a need to carry out calibration so that by preparing different concentrations of the sample in the solvent and plotting the concentration against the recorded absorbance we have a straight line. It is necessary to use a concentration of the sample within this calibrated curve. The usual concentration, which is used in recording a UV-visible spectrum, is $10^{-3}$–$10^{-5}$ M. The solute must not change its concentration due to decomposition or precipitation.

**Table 1** UV-visible absorption of some chromophores

| *Chromophore* | *Compound* | *Electronic transition* | $\lambda_{max}$ (nm) | $\varepsilon$ [a] ($m^2\ mol^{-1}$) |
|---|---|---|---|---|
| $C_6H_6$ | Benzene | $\pi \rightarrow \pi^*$ | 200, 255 | 800 |
| C=C | Ethylene | $\pi \rightarrow \pi^*$ | 180 | 1300 |
| C=O | Acetone | $\pi \rightarrow \pi^*$ | 185 | 95 |
| | | $n \rightarrow \pi^*$ | 277 | 2 |
| N=N | Azomethane | $n \rightarrow \pi^*$ | 347 | 1 |
| N=O | Nitrosobutane | $n \rightarrow \pi^*$ | 665 | 2 |

[a] Molar absorptivity.

## Solvent effects

It is very important to choose the appropriate solvent. The best solvent is that which is able to solubilize the solute completely, which is inert and will not interact with the solute and thus cause a perturbation of its electronic structure, and which will not absorb in the region where the sample will absorb. **Table 2** gives the cutoff wavelength for a selective range of solvents.

Sometimes it is important to use a solvent with a high solubilizing ability like water, ethanol, etc. Solvents must be of high purity because of the presence of impurities will affect the spectrum and some of those impurities have absorptions in the same region as that of the solute, while others will interact with the solute and this will change its structure. It is important to take into consideration the solvent effect on the solute. So if we take a solute dissolved in an inert solvent such as cyclohexane as our reference absorption, then some solvents will cause the absorption of the solute to shift more to the red region and this shift is denoted as bathochromic or red-shift. Other solvents may cause the band to shift to the blue region and this is called blue-shift or hypsochromic. This phenomenon is important in UV-visible spectroscopy because it is possible to study the solute–solvent interactions and hydrogen bond strength. Some solute molecules can be used as a probe for measuring the solvent properties.

**Table 2** $\lambda_{max}$ for different solvents

| *Solvent* | $\lambda_{max}$ (nm) |
|---|---|
| Water | 165 |
| Acetonitrile (spectroscopic grade) | 190 |
| Hexane | 199 |
| Heptane | 200 |
| Isoctane | 202 |
| Diethyl ether | 205 |
| Ethanol | 207 |
| Propan-2-ol | 209 |
| Methanol | 210 |
| Cyclohexane | 212 |
| Acetonitrile (chemical grade) | 213 |
| Dioxan | 216 |
| Dichloromethane | 233 |
| Tetrahydrofuran | 238 |
| Trichloromethane | 247 |
| Tetrachloromethane | 257 |
| Dimethyl sulfoxide | 270 |
| Dimethylformamide | 271 |
| Benzene | 280 |
| Pyridine | 306 |
| Propanone | 331 |

### Light absorption and colour

Colour change is very important for the determination of different species present in the media whether they are the reacting species or the species formed during the reaction. From the spectra it is easy to calculate the concentration of all species and therefore to study the kinetics of the reaction.

### Solution conditions and analysis

To obtain accurate data and concentrations for the different species present in solutions at specific times, it is important to obtain correct conditions for monitoring the species by the change of the absorbance of those species. It is important to determine the following conditions and their variation.

**Temperature and its variation** It is possible to change the temperature and this will lead to a change of absorbance, thus, it is possible to calculate the activation energy of the reaction.

**pH** This is very important for aqueous solutions. The pH can have a large affect on the UV-visible spectrum, rates of complexation or change in reaction kinetics. pH can have a pronounced effect on shifting the equilibrium in enol–keto tautomerism. A good example is the effect of pH on the equilibrium in the system dichromate/chromate.

**Ionic strength** The ionic strength of the solution is important for some reactions, especially for solutions which contain ions other than the neutral reactant molecules; interaction will occur between ions other than the reactants and this will affect the rate constant.

**Effect of pressure** For the majority of chemical reactions it is difficult to study the pressure effect. The reaction rate is insensitive to changes in pressure and a large pressure must be applied before we observe any important change. Sometimes we need to study the effect of pressure because it provides a useful insight into the mechanism of the reaction.

**Solvent polarity** In general, one should use solvents with low polarity so there will be no interference in the formation and decomposition of different species. On occasions, however, it is necessary to use solvents with high polarity to increase the solubility of the solutes or to affect the activation energy of a certain reaction.

Other factors that are important are the order of addition of different reagents, mixing and stirring to obtain solution homogeneity and choice of the proper wavelength which will be used to monitor the change of concentration during the reaction.

### Specialized techniques

**Stopped flow method** This method can be applied for the study of gases and liquids. The usual detection techniques are absorption spectrometry. High flow rates are required, to promote mixing for fast reactions.

**Flash photolysis and pulse radiolysis** In the flash photolysis technique a reactant is irradiated with an intense flash of visible or ultraviolet light. The intensity must be sufficient to produce a measurable change in chemical composition, but of short duration compared with that of the ensuing reactions, which are to be studied.

**Multidimensional spectroscopy and derivative spectroscopy** When bands of reactants and reaction products overlap in the fundamental UV-visible absorption spectra the reaction kinetics cannot be followed by the classic UV method. In many cases, the second derivative UV-visible spectrophotometry (D-2) provides an alternative method to solve the problem. Even-order derivatives are suitable to follow kinetics because the maxima in the UV-visible derivative spectrum can be associated with the minima and a low-noise online spectra is obtained which can be computed up to the 6th order derivative and even up to the 10th order with the newly developed computers. On the other hand, the first derivative does not provide the above association; and other higher odd-order derivatives are less precise, though in practice it has proved valuable to work with spectra of the 3rd and 5th order.

## Applications to chemical reactions

Many applications of UV-visible spectroscopy to the study of chemical reactions have been published and some illustrative examples which have been reported in the literature during the last two years will be covered.

### Complex formation

The kinetics of transformation of complexes in organic media can be studied. Such studies will indicate whether the complexes will form in a one-step reaction mechanism or a multistep reaction mechanism. It is possible to determine the rate constants and the stability constant of the complex. UV-visible spectroscopy can be used to study the protonation of certain complexes and it is possible to obtain information on the position of protonation.

Reduction of complexes can be studied by UV-visible spectroscopy. Reduction of *trans*-$[Pt(CN)_4X_2]^{2-}$

(X = Cl or Br) [as model compounds for antitumour-active platinum(IV) pro-drugs] to $[Pt(CN)_4]^{2-}$ by L-methionine, MeSR, has been studied at 25°C in the range 0 < pH < 12 (X = Cl) and 0 < pH < 6 (X = Br) by use of stopped-flow spectrophotometry. It was concluded that methionine-containing biomolecules may compete with thiol compounds for reduction of platinum pro-drugs under acidic conditions, and also in neutral solutions with low concentrations of thiol-containing biomolecules.

The formation kinetics and the thermodynamic stability of iron(III) complexes with a new tetradentate ligand, *N*(alpha)-salicyl-L-alaninehydroxamic acid (H-2slalh), have been investigated by the use of UV-visible and stopped-flow spectrophotometric methods. By comparing the pH-dependent absorbance change, the mechanism of the complex formation was explained.

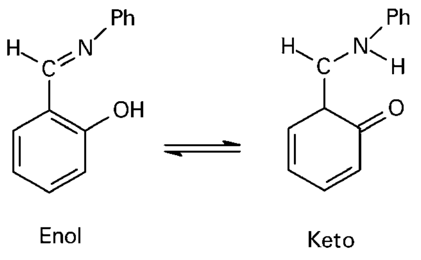
### Photochemical reactions

Photochemical behaviour of full-aromatic $\beta$-carbolines, nor-harmane, harmane and harmine in chloroform medium and the corresponding UV absorption and fluorescence emission spectra have been discussed. Irreversible electron transfer from the singlet excited $\beta$-carboline molecule to the chloroform molecule in the transient excited charge transfer complex has been proposed as the primary photochemical process initiating the mechanism of secondary reactions in this system.

In deoxygenated water, methanol, and ethanol, 4-hydroxybenzonitrile (4-HBN) is photoisomerized into 4-hydroxybenzoisonitrile (4-HBIN), which is then hydrolysed into 4-hydroxyformanilide in acidic medium. The triplet–triplet absorption of 4-HBN ($\lambda_{max}$= 300 nm) is detected by transient absorption spectroscopy. The triplet is converted into long-lived transients absorbing in the far-UV. The analysis of the kinetics of 4-HBIN formation as a function of irradiating photon flux shows that the photoisomerization of 4-HBN is a two-stage photoprocess. According to triplet-quenching studies, the first stage proceeds via the 4-HBN triplet to yield an intermediate capable of absorbing a second UV photon, which then gives 4-HBIN in the second stage. Mechanistic considerations indicate that this intermediate is likely to be an arizine.

### Tautomerism and isomerization

Tautomerism is an important phenomenon and an example of this process is the equilibrium in Schiff bases. These bases can exist as an equilibrium between two species, keto ⇌ enol [**Scheme 1**]. This process is affected by a number of factors such as solvent polarity, temperature and pH. The kinetics of tautomerism can be studied by UV-visible spectroscopy and we will deal with a few examples.




**Scheme 1**

The tautomerism between a hydroxy Schiff base and the corresponding ring-closed oxazolidine was kinetically studied in chloroform. This method indicated that this reaction is pseudo-first-order. UV examination was used to deduce the molecular species in various pH buffers. In an acid solution (e.g. pH 3.0) the Schiff base existed as the protonated Schiff base at the imine nitrogen atom, and in the alkaline region (e.g. pH 9.0) as the oxazolidine form.

The keto–enol equilibrium constants of acetylacetone, ethyl acetoacetate and ethyl benzoylacetate in water at 25°C were determined by studying the influence of surfactants on their UV-visible spectra. These measured equilibrium constants were used to obtain the reactivity of the ketones towards several nitrosating agents.

### Atmospheric, environmental reactions, gas phase and free radical kinetics

The UV absorption cross-sections of peroxyacetyl nitrate (PAN), $CH_3C(O)O_2NO_2$, have been measured as a function of temperature (298, 273 and 250 K) between 195 and 345 nm. Photolysis becomes the most important atmospheric loss process for PAN, and the OH reaction is found to unimportant throughout the troposphere.

The atmospheric chemistry of benzene oxide/oxepin, a possible intermediate in the atmospheric oxidation of aromatic hydrocarbons, has been investigated by visible and UV photolysis. The results indicated that the major atmospheric sinks of benzene oxide/oxepin are the reaction with OH radicals and photolysis and under smog chamber conditions, with high $NO_2$ concentrations, also reaction with $NO_3$.

A flash photolysis-resonance fluorescence technique was used to study the rate constant for the reaction of OH radicals dimethyl carbonate over the temperature range 252–370 K. Pulse radiolysis/transient UV absorption techniques were used to

study the ultraviolet absorption spectra and kinetics of $CH_3OC(O)OCH_2$ and $CH_3OC(O)CH_2O_2$ radicals at 296 K.

The codisposal of trace metals (e.g. Co), synthetic chelates (e.g. ethylenediaminetetraacetic acid, H(4) EDTA) and water-miscible organic solvents has occurred at some contamination sites. The reactions of Co(II)-EDTA with a redox reactive naturally occurring solid, goethite, in aqueous and semiaqueous (methanol–water, acetone–water) suspensions was studied. UV-visible spectroscopy indicated that goethite catalysed oxidation of Co(II)-EDTA to Co(III)-EDTA by dissolved $O_2$. These reactions have important implications on the fate of the redox-sensitive metal in complex, mixed waste environments.

### Photocatalytic degradation

A number of organic pollutants present in industrial and domestic wastewater resist biodegradation and they are poisonous even at low concentrations. A new method for removing toxicants from wastewater is based on the use of photocatalysis. It was found that titanium dioxide ($TiO_2$) is the best photocatalyst for the detoxification of water because it is cheap, nontoxic and easy to use and handle. The energy gap in $TiO_2$ is 3.2 eV and thus it can be activated at >400 nm.

The primary oxidant responsible for most advanced oxidation processes, i.e. the use of photochemical methods, is the hydroxyl radical, which is formed by the reduction reactions of electrophiles with water or hydroxide ions. The mechanism of the formation of the hydroxyl radical is well discussed in the literature. Heterogeneous photocatalytic oxidation of organic compounds in aqueous solution is achieved by the reactive hydroxyl radical. The photocatalytic process can remove a large number of organic hazardous compounds in water. Phenols, carboxylic acids and herbicides are among a large number of pollutants, which are destroyed in air and in water using photocatalysis.

Photodegradation of 1,2,3,4-tetrachlorodibenzo-*p*-dioxin (1,2,3,4-TCDD) in hexane solution was studied under controlled near-UV light exposure in the spectral region from 325 to 269 nm. Irradiation experiments carried out at a constant light energy (700 mJ) showed that the percentage of 1,2,3,4-TCDD left in the solution after irradiation changed from about 55 to 75%, with a minimum of 55% at 310 nm. Further irradiation experiments carried out at two wavelengths, namely 310 and 269 nm, and light energy ranging from 0 to 4000 mJ, showed that the photodegradation reaction of the TCDD always followed pseudo-first-order kinetics.

A semitransparent $TiO_2$ film with extraordinarily high photocatalytic activity was prepared on a glass substrate by sintering a $TiO_2$ sol at 450 °C. The photocatalytic properties of the film were investigated by measuring the photodegradative oxidation of gaseous acetaldehyde at various concentrations under strong and weak UV light irradiation conditions. The kinetics of acetaldehyde degradation as catalysed by the $TiO_2$ film as well as by P-25 powder were analysed in terms of the Langmuir–Hinshelwood model. It is shown that the number of adsorption sites per unit true surface area is larger with the $TiO_2$ film, as analysed in powder form, than with P-25 powder. Meanwhile, the first-order reaction rate constant is also much larger with the film than with P-25 powder. Moreover, under most experimental conditions, particularly with high concentrations of acetaldehyde and weak UV illumination intensity, the quantum efficiency was found to exceed 100% on an absorbed-photon basis, assuming that only photogenerated holes play a major role in the reaction. This leads to the conclusion that the photodegradative oxidation of acetaldehyde is not mediated solely by hydroxyl radicals, generated via hole capture by surface hydroxyl ions or water molecules, but also by photocatalytically generated superoxide ion, which can be generated by the reduction of adsorbed oxygen with photogenerated electrons.

### Solvents effects and solvolysis

Oxidations of arylalkanes by $(Bu_4NMnO_4)$-$Bu^n$ have been studied, e.g. toluene, ethylbenzene, diphenylmethane, triphenylmethane, 9,10-dihydroanthracene, xanthene and fluorene. Toluene is oxidized to benzoic acid and a small amount of benzaldehyde; other substrates give oxygenated and/or dehydrogenated products. The manganese product of all of the reactions is colloidal $MnO_2$. The kinetics of the reactions, monitored by UV-visible spectrometry, show that the initial reactions are first order in the concentrations of both $(Bu_4NMnO_4)$-$Bu^n$ and substrate. No induction periods are observed. The same rate constants for toluene oxidation are observed in neat toluene and in *o*-dichlorobenzene solvent, within experimental errors. The presence of $O_2$ increases the rate of $(Bu_4NMnO_4)$-$Bu^n$ disappearance. The reactions of toluand dihydroanthracene exhibit primary isotope effects: The rates of oxidation of substituted toluenes show only small substituent effects. In the reactions of dihydroanthracene and fluorene, the $MnO_2$ product is consumed in a subsequent reaction that appears to form a charge-transfer complex. The rate-limiting step in all of the reactions is hydrogen atom transfer from the substrate to a

permanganate oxo group. The enthalpies of activation for the different substrates are directly proportional to the $\Delta H$ for the hydrogen atom transfer step, as is typical of organic radical reactions. The ability of permanganate to abstract a hydrogen atom is explained on the basis of its ability to form an $80 \pm 3$ kcal mol$^{-1}$ bond to H–, as calculated from a thermochemical cycle.

### Polymerization

The kinetics of oxidation of a macromolecule, poly(ethylene glycol) [PEG] by ceric sulfate in sulfuric acid medium has been studied by means of UV-visible spectrophotometry. The observed difference in rates of oxidation has been explained in terms of cage formation. The oxidation of PEG proceeded without the formation of a stable intermediate complex. The order with respect to the concentrations of PEG and ceric sulfate has been found to be one and the overall order is two. The effects of acid concentration, sulfate ion concentration, ionic strength and temperature on the rate of the oxidation reaction have been studied. The thermodynamic parameters for the oxidation reaction have also been presented. Based on the experimental results, a suitable kinetic expression and a plausible mechanism have been proposed for the oxidation reaction.

Reaction of *trans*-1,3-diphenyl-1-butene (D), the *trans* ethylenic dimer of styrene, with trifluoromethanesulfonic acid in dichloromethane has been performed at temperatures lower than room temperature using a stopped-flow technique with real time UV-visible spectroscopic detection. The main product of the reaction was the dimer of D. A transient absorption at 340 nm has been assigned to 1,3-diphenylbutylium, a model for the polystyryl cation. Other absorptions at 349 nm and 505 nm have also been observed and were assigned to an allylic cation, 1,3-diphenyl-1-buten-3-ylium, resulting from hydride abstraction from D. This species was very stable at temperatures lower than −30°C. A general mechanism was proposed based on a kinetic study of the reactions involved.

### Determination of metals and molecules at low concentration

A new and simple method for measuring peroxides in a single living cell has been developed, and the generation of peroxides upon ultraviolet (UV) irradiation was measured in human and pig pidermal keratinocytes. The method was based on the fact that the nonfluorescent dye dihydrorhodamine 123 reacts in the presence of peroxides, such as $H_2O_2$, and changes into fluorescent rhodamine 123, and hence the fluorescence intensity is proportional to the amount of reacted peroxide. The epidermal kerationcytes were loaded with the dihydrorhodamine under a fluorescence microscope and exposed to UV radiation. Taking $C$ as the content of peroxides generated within the cell and $I$ as the increase influence (radiation intensity × time = photons cm$^{-2}$), the following empirical relationship was established: $C = C_s (1 - \exp^{-kI})$, where $C_s$ is the content of peroxides at the saturation state, and $k$ is a kinetic parameter. The dependence of the two parameters on wavelength in the range 280–400 nm was studied. In human keratinocytes $C_s$ had a peak at 310 nm and a small peak (shoulder) at 380 nm, while $k$ increased gradually toward shorter wavelengths. In pig keratinocytes, on the other hand, $k$ had a peak around 380 nm and a shoulder at 330 nm, while $C_s$ remained unchanged. Aminotriazole, an inhibitor of catalase, and low temperatures increased the stationary levels of peroxide generation in pig keratinocytes upon UV irradiation, indicating that the reaction used for measuring intracellular peroxides is competitive with the intrinsic reactions in scavenging peroxides.

A first simultaneous EPR and visible spectrophotometric study is reported on the interaction of the stable free radical 1,1-diphenyl-2-picrylhydrazyl (DPPH) dissolved in ethanol with thioglycolic acid ($HSCH_2CO_2H$, TGA). The results of the kinetic studies at room temperature allow us to assume 1:1 stoichiometry of the reaction between DPPH and TGA giving 1,1-diphenyl-2-picrylhydrazine (DPPH) and thioglycolic disulfide. The linear plots of EPR and UV-visible responses versus the quantity of added TGA are used to find the DPPH molar absorptivity at 520 nm to be 12350 ±3% 1 mol$^{-1}$cm$^{-1}$ which may be used as a criterion for the purity of the material itself. It was also found that the paramagnetic and optical properties of a 30-year old sample gave results suggesting that in the solid state DPPH is a fairly stable material.

### Reaction kinetics, substitution effects, structure–reactivity relationships

The kinetics of the addition of arenesulfinic acids to 4-substituted 2-nitroethenylarenes was studied by means of UV spectrophotometry. The effect of 4-substituting groups in benzenesulfinic acids, and the change in reactivity of the nitroethylene system due to typical electron-donating and electron-withdrawing groups were investigated. The substituent effect on benzenesulfinic acid fits Hammett's equation, $p$-value at 298 K being −1.12. Kinetic studies were carried out at 288–308 K, and the activating energy and the enthalpy of activation were determined.

The kinetics of the addition of unsubstituted and substituted benzenesulfinic acids to various 2-halogenonitroethenylarenes was studied by means of UV spectroscopy. The effect of 4-substituents in benzenesulfinic acids and the change in reactivity of the nitroethylene system in 2-halogeno-2-nitroethenyl-arenes in the presence of various substituents in the benzene ring and various halogens at the double bond were investigated. Kinetic studies were made at 288–308 K, and the activating energy and enthalpy of activation were determined.

### Effect of pH, pressure, temperature and ionic strength on reaction kinetics

The degradation kinetics of xanthate in homogeneous solution as a function of pH at 5°C, 20°C and 40°C was systematically studied by UV-visible spectrophotometric measurements. The results indicate that the degradation of ethyl xanthate is rapidly increased with decreasing pH at pH $< 7$. At pH 7–8, the maximum half-life of the xanthate appears. The degradation was faster at pH 9–10, but at pH $> 10$ the half-lives of xanthate once again increase. The investigations were also extended to different media other than pure water, such as, 0.1 M $NaClO_4$, 0.1 M $NaNO_3$, 0.1 M NaCl as well as in the supernatants of flotation tailings of sulfide minerals. The rate constants of xanthate degradation were calculated and presented together with half-lives and activation energies of xanthate degradation. The degradation products and reaction mechanisms are discussed based on experimental results.

The kinetics of the association reaction of the phenoxy radical with NO were investigated using a flash photolysis technique coupled to UV absorption spectrometry. Experiments were performed at atmospheric pressure, and theoretical calculations showed that the rate constant is at the high-pressure limit above 50 torr ($6.7 \times 10^3$ Pa) for temperatures below 400 K. Upon increasing the temperature, the reaction was found to be reversible, and the equilibrium kinetics have been studied at seven temperatures between 310 and 423 K.

### Validation of molecular orbital calculations

The suitability of Gaussian distribution functions to describe the shape and temperature dependence of the UV absorption continua of peroxy radicals has been investigated. The ethylperoxy radical was used as a test case. Its 298 K absorption continuum was found to be best described by a semilogarithmic Gaussian distribution function. A linear Gaussian distribution function performed less well but still adequately described the continuous absorption. The temperature dependence of the ethylperoxy radical UV absorption continuum was also well predicted. Analogous results obtained for the methylperoxy radical support these conclusions. A theoretical comparison of the semilogarithmic and linear Gaussian distribution functions is given and a potential energy diagram of the ethylperoxy radical derived. The experimentally determined absorption cross-sections of $HO_2$ have been reanalysed. It is shown that either the measurements at short wavelengths are in error or an unidentified electronic transition of $HO_2$ exists.

### Enzymatic reactions

A one-step spectrophotometric method for monitoring nucleic acid cleavage by ribonuclease. H from *E. coli* and type II restriction endonucleases has been proposed. It is based on recording the increase in the UV absorbance at 260 nm during the course of enzymatic reaction. Duplexes stable under the reaction conditions were chosen as substrates for the enzymes being studied. In order to obtain duplex dissociation following their cleavage by the enzyme appropriate temperature conditions were selected. The spectrophotometric method may be applied for rapid testing of the nuclease activity in protein preparations as well as for precise quantitative analysis of nucleic acid degradation by enzymes. This method may be successfully employed in kinetic studies of nucleic acid–protein interactions.

*See also:* **Biomacromolecular Applications of UV-Visible Absorption Spectroscopy; Colorimetry, Theory; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy.**

## Further reading

Atkins PW (1998) *Physical Chemistry*, 6th edn. Oxford: Oxford University Press.

Barrow GM (1986) *Introduction to Molecular Spectroscopy*, 17th edn. London: McGraw-Hill Book Company.

Denney RC and Sinclair R (1991) *Visible and Ultra Violet Spectroscopy*. New York: Wiley.

Kempt W (1989) *Organic Spectroscopy*, 2nd edn. London: Macmillan Education.

Ladd M (1998) *Introduction to Physical Chemistry*, 3rd edn. London: Cambridge University Press.

Pilling MJ and Seakins PW (1996) *Reaction Kinetics*, 1st edn. London: Oxford Science Publication.

Pérez-Bendito D and Silva M (1988) *Kinetics Methods in Analytical Chemistry*, 1st edn. Chichester: Ellis Horwood.

Vemulapalli K (1993) *Physical Chemistry*. New York: Prentice-Hall International.

# Chemical Shift and Relaxation Reagents in NMR

**Silvio Aime**, **Mauro Botta**, **Mauro Fasano** and **Enzo Terreno**, University of Torino, Italy

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

## Introduction

Chemical shift and relaxation reagents are substances that, when added to the systems under study, enable the spectroscopist to tackle a specific problem of spectral assignment, stereochemical determination or quantitative measurement. They are usually represented by paramagnetic compounds, and their effects may exploited by the solute and the solvent, manifested as large shifts and/or pronounced relaxation enhancement of the lines. Both effects reflect the response of the nuclear dipole to the local magnetic field of an unpaired electron in the reagent, whose magnetic moment is 658.21 times higher than that of the proton.

A knowledge of the basic theory of the paramagnetic interaction is useful to pursue a better exploitation of the effects induced by the paramagnetic perturbation.

In principle, the interaction of a nucleus with an unpaired electron (hyperfine interaction) has the same nature as a nucleus–nucleus interaction, i.e. it is determined by a direct, through-space dipole–dipole coupling ($\mathbf{H}_D$) and an indirect, through-bonds contact (scalar) coupling ($\mathbf{H}_S$):

$$\mathbf{H} = \mathbf{H}_D + \mathbf{H}_S$$

The hyperfine shift arises from components present in both terms which are invariant with time in the external magnetic field; these contributions are indicated as 'dipolar' or 'pseudocontact' shifts when resulting from dipolar interactions and 'Fermi' or 'contact' shifts when resulting from scalar interactions, respectively. The time-dependent fluctuations in either term will result in nuclear relaxation ('dipolar' and 'scalar' relaxation, respectively). Although, in theory, correct expressions for $\mathbf{H}_D$ and $\mathbf{H}_S$ can be derived, the elusive nature of the orbiting electrons and the number of their interactions make it very difficult to develop an accurate theory for shift and relaxation promoted by paramagnetic substances. In general one may conclude that, while measurements of the effects promoted by paramagnetic species are easy, interpretation of data is difficult.

According to their dominant effect on the NMR spectra of surrounding nuclei, paramagnetic reagents can be classified as chemical shift or relaxation reagents. Organic free radicals and metal ions with isotropic electron density, such as $Mn^{2+}$ (five electrons in five d orbitals) or $Gd^{3+}$ (seven electrons in seven f orbitals) do not give rise to pseudocontact shifts and are commonly regarded as relaxation probes. On the other hand, anisotropic ions with short electron relaxation times, such as all other lanthanide ions, can be referred to as chemical shift probes. Species such as $Ni^{2+}$, $Fe^{3+}$ and $Cr^{3+}$ give rise to both effects and defy classification.

## Chemical shift reagents

In 1969, Hinckley suggested that the use of paramagnetic lanthanide complexes might be very useful to simplify unresolved $^1$H, NMR resonances, thus allowing otherwise intractable spectra to be easily interpreted (**Figure 1**). Although the introduction of high-field magnets and two-dimensional methods have largely reduced the need for such qualitative use of 'lanthanide induced shift' (LIS) measurements, the exploitation of the large local fields generated by some lanthanide ions is still an item of strong interest in several areas of NMR applications.

When a substrate interacts with the lanthanide complex, it magnetically active nuclei *feel* the presence of the unpaired f electrons; their LIS is basically determined by two properties of the lanthanide complex, namely (i) its Lewis acid behaviour and (ii) the number of unpaired electrons. The commercially available lanthanide shift reagents (LSRs) are mainly tris(β-diketonate) complexes, having the 2,4-pentanedione chelate structure (**Table 1**). These neutral complexes are soluble in organic solvents and they promptly form acid–base adducts with substrates endowed with Lewis base behaviour. In fact alcohols, ethers, amines and nitriles reversibly interact with the LSR to form paramagnetic adducts whose $^1$H NMR shifts may be far away from the diamagnetic values. As the substrate B in fast exchange between the LSR-bound and free forms, the effect will be averaged on all the molecules of the substrate. The

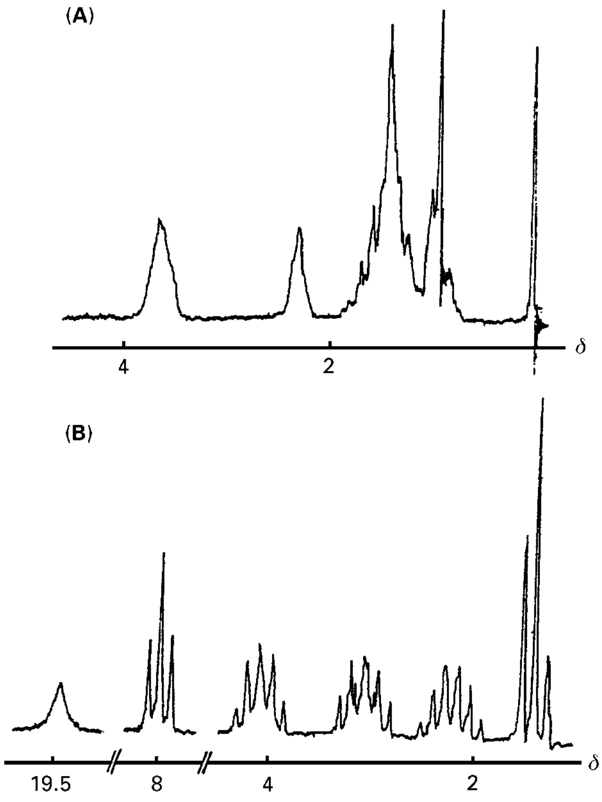
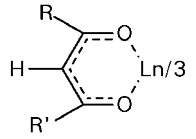



**Figure 1** Simplification of the low-field $^1H$ NMR spectrum of 1-pentanol by $Eu(thd)_3$. The clustered signals of the methylene chain are totally separated into clearly defined multiplets. (A) normal spectrum; (B) after addition of $Eu(thd)_3$

**Table 1** Lanthanide diketonate chelates used as NMR shift reagents




| *R* | *R'* | *Ln* | *Acronym* |
|---|---|---|---|
| *t*-$C_4H_9$ | *t*-$C_4H_9$ | Pr | $Ln(thd)_3$ or |
| | | Eu | $Ln(dpm)_3$ |
| | | Dy | |
| | | Ho | |
| | | Yb | |
| *t*-$C_4H_9$ | $CF_3$ | Eu | $Eu(pta)_3$ |
| *t*-$C_4H_9$ | *n*-$C_3F_7$ | Eu | $Eu(fod)_3$ |
| | | Pr | |
| *n*-$C_3F_7$ | *n*-$C_3F_7$ | Eu | $Eu(tfn)_3$ |
| $C_2F_5$ | $C_2F_5$ | Eu | $Eu(fhd)_3$ |
| $CF_3$ | *n*-$C_3F_7$ | Eu | $Eu(dfhd)_3$ |

thd = tris(2,2,6,6-tetramethyl-3,5-heptanedionato);
dpm = tris(dipivaloylmethanato);
pta = tris(1,1,1-trifluoro-5,5-dimethyl-2,4-hexanedionato);
tfn = tris(1,1,1,2,2,3,3,7,7,8,8,9,9,9-tetradecafluoro-4,6-nonanedionato);
fhd =tris(1,1,1,2,2,6,6,7,7,7-decafluoro-3,5-heptanedionato);
dfhd = tris(1,1,1,2,2,3,3,7,7,7-decafluoro-4,6-heptanedionato);
fod = tris(1,1,1,2,2,3,3-heptafluoro-7,7-dimethyl-4,6-octanedionato).

resulting shifts will depend upon the strength of the interaction, the geometry of the adduct and the LSR to substrate ratio.

The acid character of the lanthanide chelate and the donor and steric properties of the substrate define the strength of the acid–base adduct. Usually more stable complexes are found as the ionic radius of the metal decreases. Fluorinated diketonates increase the Lewis acidity of the lanthanide ion with respect to the protonated derivatives. This explains why $Eu(fod)_3$ (see **Table 1**) is the mostly commonly used LSR.

## Contributions to the lanthanide-induced shift

The shift induced by the lanthanide ion on the resonances of the organic substrate results from three contributions.

- A diamagnetic term corresponding to the coordination shift. On the $^1H$ resonances it is usually upfield and rather small.
- The contact or Fermi term, which depends on the delocalization of unpaired electron density on the substrate nuclei, and has then been related to the value of the hyperfine coupling constant.
- The dipolar or pseudocontact term, which depends on the spatial proximity of a given nucleus to the paramagnetic centre and on the anisotropy of the magnetic susceptibility tensor. In order to extract structural information from LIS data there has to be an accurate determination of the latter term.

The first step in the procedure consists of the determination of the all-bound shifts $\Delta_i^{LSR-S}$. In the presence of the equilibrium between the substrate (S) and the lanthanide shift reagent (LSR)

$$\mathrm{LSR} + \mathrm{S} \overset{K_i}{\rightleftharpoons} \mathrm{LSR{-}S}$$

the observed $\Delta\delta_i$ are proportional to the term ([LSR–S]/[S])$\Delta_i^{LSR-S}$. Thus $\Delta\delta_i$ are measured for various LSR to S ratios ($\rho$) and graphs are obtained by plotting the observed shifts $\Delta\delta_i = \delta_i^{obs} - \delta_i^{o}$(where $\delta_i^{o}$ is the chemical shift measured in the absence of the LSR or, better, in the presence of diamagnetic La or Lu analogue) as a function of $\rho$.

If an axial symmetry of the dipolar magnetic field in the LSR–S adduct is assumed, the dipolar shifts are proportional to the term $(3\cos^2\phi - 1)/r_i^3$. In this case, the principal magnetic axis of the lanthanide is taken as collinear with the lanthanide–ligand bond. $r_i$ is the distance between a given nucleus and the paramagnetic ion and $\phi$ is the angle between the $r_i$ vector and the Ln–ligand bond. Many systems have been successfully investigated on the basis of this simplifying assumption. The occurrence of axial symmetry is justified as a result of time averaging effects. This means that the fast rotation of the substrate along the coordination axis introduces an effective axial symmetry, where $\phi$ now refers to the angles between $r_i$ and the rotation axis rather than the actual symmetry axis.

In the more general case of a rhombic, nonaxial symmetry, the dipolar contribution to the (all-bound) LIS is described as follows:

$$\Delta_i^{\mathrm{LSR\ \ S}} = K_1\frac{(3\cos^2\theta_i - 1)}{r_i^3} + K_2\frac{\sin^2\theta_i\cos 2\Omega_i}{r_i^3}$$

where $r_i$, $\theta_i$ and $\Omega_i$ are the usual spherical coordinates of the $i$ nucleus i the reference frame of the principal magnetic susceptibility tensor. $K_1$ and $K_2$ are coefficients related to the anisotropy of the magnetic susceptibility tensor. Several computer programs have been written to find the molecular geometry that best fits the set of experimental $\Delta_i$ values.

A decrease in temperature causes an increase in the lanthanide-induced chemical shifts, which usually provides an enhanced resolving power of the LSR. This behaviour results (i) from an increase of the concentration of the adduct and (ii) from the temperature dependence of the paramagnetic shift (the dipolar shift shows a $1/T^2$ dependence, whereas the contact term displays a $1/T$ dependence).

Diketonates are insoluble in water. Diglycolates have been proposed as substitutes for studies in aqueous solutions. Another possibility deals with the use of lanthanide aquo-ions, and perchlorates and nitrates have been used in several cases: they are soluble at pH < 6 and provide good shifts for a number of anionic functions. Eventually, good results in aqueous solutions have been obtained by using EDTA chelates which work at pH values as high as 10.

### Chiral shift reagents

It is well established that a method to distinguish enantiomers by NMR is to convert them into diastereoisomers by means of appropriate optically active reactants. Thus it was suggested that the use of chiral LSR (LSR*) can induce shift differences in the spectra of enantiomers:

$$\mathrm{LSR}^* + \mathrm{R} \rightleftharpoons \mathrm{LSR}^*\text{–}\mathrm{R}$$
$$\mathrm{LSR}^* + \mathrm{S} \rightleftharpoons \mathrm{LSR}^*\text{–}\mathrm{S}$$

R and S give the same NMR spectrum as they are mirror images, whereas LSR*–R and LSR*–S are diastereoisomers and then are characterized by different NMR properties. In the presence of fast exchange between free and bound forms, R and S will give rise to different signals as they represent average values between the free and the bound forms. The enantiomeric shift difference between the $\Delta\delta$ LIS for the R and S configuration is usually indicated as $\Delta\Delta\delta$. The observed $\Delta\Delta\delta$ values depend on the equilibrium constants for the formation of the labile complexes between the enantiomers and the chiral reagent, and on the shift differences between the diastereomeric complexes.

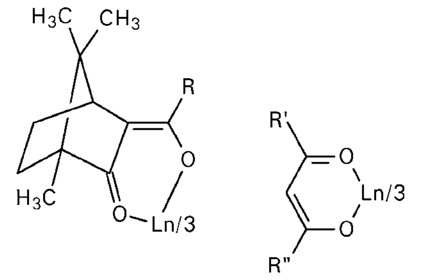
The first chiral LSR was reported by Whitesides and consisted of a camphor-based $Eu^{3+}$ complex. Since then, several chiral LSR* have been proposed and many of them are commercially available (**Figure 2** and **Table 2**).

In some cases, it has been found that enantiomeric resolution has been obtained also with the use of an achiral reagent such as $Eu(fod)_3$. This behaviour has been explained in terms of the formation of ternary 1:2 adducts which are no longer mirror images (such as LSR–R–R and LSR–R–S). In principle, another explanation may be possible on the grounds that $Eu(fod)_3$ may be considered as a racemic mixture of enantiomers.

### Lanthanide shift reagents for alkali metal ions

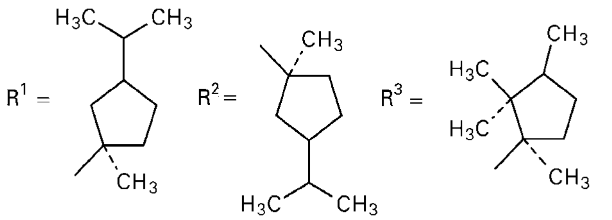
An interesting extension of the use of LSR deals with the separation of the resonances of $^{7}Li^{+}$, $^{23}Na^{+}$ and $^{39}K^{+}$ present in different biological compartments. Such NMR-active metal cations are of clinical importance, but from their routine NMR spectra it is not possible to distinguish whether they are in the intra- or the extracellular compartments. The addition of a paramagnetic complex that selectively distributes in one compartment only, can remove such signal degeneracy. In fact, the interaction with the paramagnetic species induces a shift in other resonances of the alkaline ions present in the same compartment as the LSR (**Figure 3**). Up to now in laboratory practice, four complexes have been

**Table 2** Chiral shift reagents of the β-diketonato type




(1) R = –C(CH$_3$)$_3$
(2) R = –CF$_3$
(3) R = R$^1$
(4) R = 77%R$^2$ + 23%R$^1$

(5) R' = R" = R$^1$
(6) R' = R$^1$, R" = R$_3$
(7) R' = 77%R$^2$ + 23%R$^1$; R" = R$^3$
(8) R' = R$^1$; R" = 77%R$^2$ + 23%R$^1$




studied in detail and applied to answer several questions of biomedical interest: the Dy(III) complexes of $PPP^{5-}$ (triphosphate) and $TTHA^{6-}$ (triethylenetetraaminehexaacetate) and the Dy(III) and Tm(III) complexes of $DOTP^{5-}$ (tetraazacyclododecane-*N*, *N'*, *N''*, *N'''* tetrakis (methylenephosphonate)).

The most effective reagent so far reported is the chelate $Dy(PPP)_2^{7-}$, first introduced by Gupta and Gupta in 1982. However, this metal complex has been proved to present several disadvantages that prevent its use in living animals or humans. First of all it is rather toxic, probably due to irreversible ligand dissociation promoted *in vivo*, with subsequent release of the metal ion. Furthermore, the complex is involved in several protonation equilibria and shows competition with endogenous divalent cations ($Mg^{2+}$, $Ca^{2+}$) which reduces its resolution ability. The $Dy(TTHA)^{3-}$ complex is much less toxic, as a consequence, probably, of the high stability constant of $Dy^{3+}$ with this multidentate ligand.

However, the reduced value of the negative charge on the complex and the fact that it is mainly localized on unbound carboxylic groups, away from the paramagnetic centre, makes this metal chelate less effective in removing the signal degeneracy. This implies that higher doses of the LSR are needed for an optimum resolution. The Dy(III) and Tm(III) complexes of DOTP represent an important

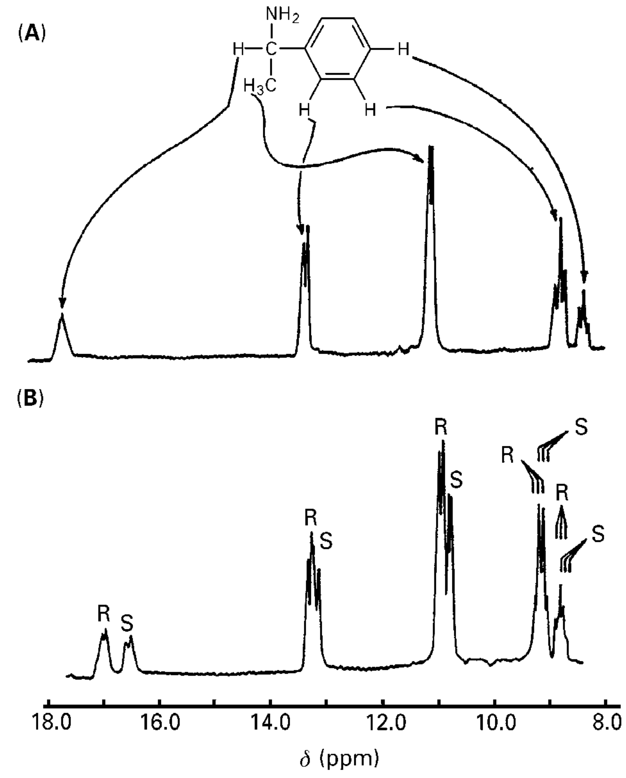



**Figure 2** $^1H$ NMR spectra of α-phenylethylamine in $CCl_4$ after addition of tris[(3-*tert*-butylhydroxymethylene)-*D*-camphorato]-europium$^{III}$. Molar ratio LSR* to substrate 0.5: (A) S-form; (B) racemic mixture.

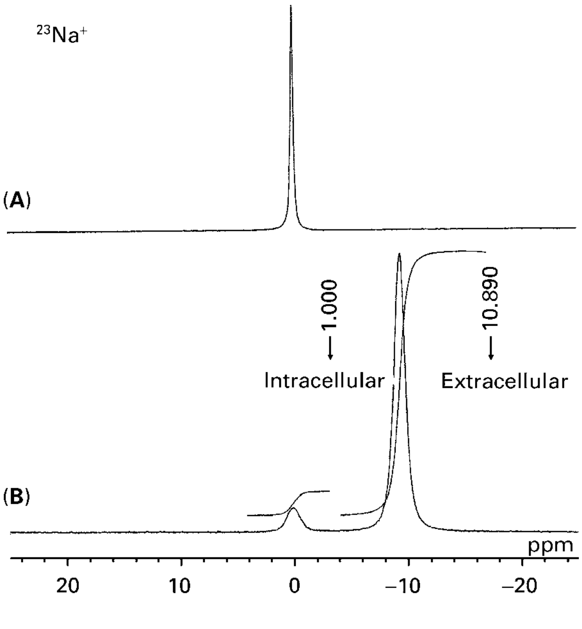



**Figure 3** $^{23}Na$ NMR spectra at 2.1 T and 39°C of whole human blood before (A) and after (B) the addition of $Dy(DOTP)^{5-}$ (5 mM). Intra- and extracellular sodium resonances are resolved.

improvement in the search for safe and effective LSRs for metal cations of biological relevance. In fact, these metal chelates are extremely resistant to dissociation processes over a wide range of pH and present in axial symmetry and interaction sites for the cations very close to their 4-fold axis of symmetry, which ensures a maximum shifting efficiency. Indeed $Dy(DOTP)^{5-}$ produces a magnitude of shift analogous to that observed for $Dy(PPP)_2^{7-}$ at the same concentration. However, the efficiency of the Dy(III) chelate has been shown to be markedly quenched by stoichiometric amounts of $Ca^{2+}$, and thus the Tm(III) complex, less sensitive to these effects, has found wider applications.

In conclusion, it is worth commenting that, although several studies have been possible by the use of the available LSRs, much remains to be done in terms of synthetic strategy and ligand design in the search for safer and more effective compounds.

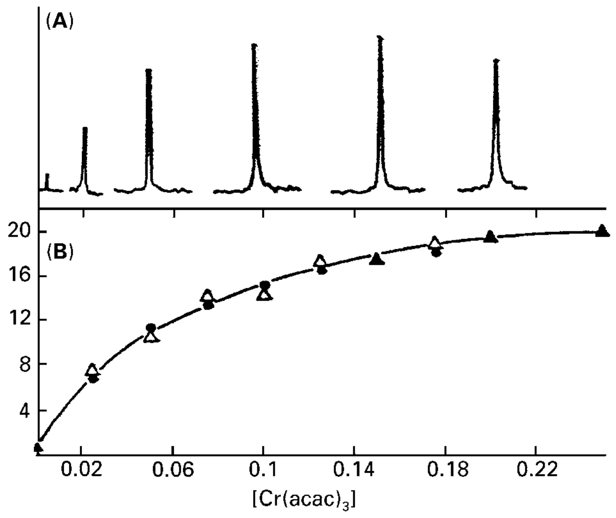



**Figure 4** $^{13}C$ NMR spectra (A) and integrated intensities (B) of a $Fe(CO)_5$ sample with different added concentrations of $Cr(acac)_3$. Different symbols refer to independent measurements.

## Relaxation reagents

An alternative exploitation of paramagnetic complexes as auxiliary NMR reagents deals with their effect on the relaxation of substrate nuclei. Indeed, in the presence of a relaxation reagent the electron–nucleus relaxation may easily become so efficient that it predominates over the intrinsic relaxation processes of a given nucleus. Such effects may be exploited either on the solute or on the solvent resonances. A number of applications based on the use of relaxation reagents have been reported, ranging from structural assignments to nuclear Overhauser effect (NOE) quenchers for correct intensity measurements, from the assessment of the binding to proteins to the separation of resonances from different compartments. A particularly important class of relaxation reagents is represented by the contrast agents for magnetic resonance imaging (MRI).

The first example dealing with the use of paramagnetic species to shorten the relaxation time of the solvent nuclei deals with the very first detection of the NMR phenomenon, as Bloch used iron(III) nitrate to avoid saturation of the water proton signal. Then in the 1970s it was rather common, in organic and organometallic chemistry, to add to the substrate solution a reagent such as $Cr(acetylacetonate)_3$ ($Cr(acac)_3$) or $Fe(acac)_3$ in order to shorten the $T_1$ of $^{13}C$ and $^{15}N$ resonances without introducing broadening of the signal. In **Figure 4** the $^{13}C$ intensities of the $Fe(CO)_5$ signal, measured at 2.1 T, in the presence of variable amounts of $Cr(acac)_3$, are reported. At this field the $T_1$ value of the $^{13}CO$ groups is very long as it is determined by the chemical shift anisotropy term; by increasing the magnetic field strength it becomes markedly shorter. Thus, the availability of high magnetic fields has limited to some extent the use of relaxation reagents for the purpose of an overall shortening of the relaxation times.

The presence of a relaxation reagent induces a dominant relaxation path for all the resonances of a given substrate present in solution and this may be exploited for quantitative NMR determinations. In fact, this method eliminates the problems associated with a correct intensity measurement when the various resonances are endowed with different relaxation times and different NOE enhancements.

Another application of relaxation reagents deals with the selective removal of a given spin-coupling pattern, i.e. the interaction with a paramagnetic reagent produces a so-called chemical decoupling mechanism. In fact, in an AX coupled system the multiplicity of each resonance may be removed when the mean lifetime of the coupled nucleus in a given state ($\alpha$ or $\beta$) becomes shorter than the inverse of the coupling constant $J_{AX}$. Therefore, the occurrence of efficient relaxation induced by the paramagnetic reagent causes the removal of the effective coupling.

### Theory of paramagnetic relaxation

The quantitative description of relaxation effects induced by paramagnetic species presents theoretical problems that are far from trivial. However, it is customary to resort to the simplified approach developed earlier by Solomon and Blombergen to

account for the relaxation enhancement of solvent nuclei of aqueous solutions of paramagnetic metal ions. On the basis of this approach the observed solvent relaxation rate is given by the sum of three contributions:

$$R_1^{obs} = R_{1p}^{is} + R_{1p}^{os} + R_1^{o}$$

where $R_1^{o}$ is the relaxation rate in the absence of the paramagnetic compound, $R_{1p}^{is}$ represents the paramagnetic contribution due to the exchange of water molecules in the inner coordination sphere with water molecules in the bulk, and $R_{1p}^{os}$ is the diffusion-controlled contribution from water molecules in the outer coordination sphere of the paramagnetic centre. In the absence of a coordinative interaction between the paramagnetic metal ion and the substrate–solvent molecules, only the latter term contributes to the observed paramagnetic relaxation enhancement. In this case the relaxation enhancement is exclusively due to the electron–nucleus dipolar interaction between the metal ion and the solvent molecules diffusing in the outer coordination sphere of the complex. This interaction is modulated by the translational diffusive motion of solute and solvent and by the electronic relaxation time. An analytical expression for this contribution has been derived by Freed and co-workers in a study of the relaxation enhancement of solvent molecules in solutions containing stable nitroxide radicals.

$$R_{1p}^{os} = C^{os}\left(\frac{1}{aD}\right)\left[7J(\omega_S) + 3J(\omega_H)\right]$$

In the expression above, $C^{os}$ is constant for a given observed nucleus, $a$ is the distance of closest approach between the paramagnetic centre and the diffusing solvent molecules and $D$ is the relative solute–solvent diffusion coefficient. The dependence on the electronic relaxation time is expressed by the nonLorentzian spectral density functions $J(\omega)$.

As far as the inner sphere contribution is concerned, $R_{1p}^{is}$ is determined by the relaxation time $T_{1M}$ of the nuclei of the substrate–solvent complex and by its lifetime $\tau_M$, weighted by the molar ratio of the bound substrate–solvent:

$$R_{1p} = \frac{[C]q}{[S]}\frac{1}{T_{1M} + \tau_M}$$

where [C] is the molar concentration of the paramagnetic reagent, $q$ is the number of interacting sites, and [S] is the concentration of the substrate–solvent complex. In the case of dilute water solutions, when the relaxation enhancement is observed on the solvent nuclei, [S] = 55.54 M.

The $T_{1M}$ value in the paramagnetic reagent–substrate adduct is determined by (i) the dipolar interaction between the observed nucleus and the unpaired electron(s) and (ii) the contact interaction between the magnetically active nucleus and the unpaired electron density localized in the position of the nucleus itself. An analytical form is given by the Solomon–Bloembergen equation (here for $T_1$; an analogous expression is derived for $T_2$):

$$\frac{1}{T_{1M}} = \frac{2}{15}\left(\frac{\mu_0}{4\pi}\right)\frac{\hbar^2\gamma_S^2\gamma_H^2}{r^6}S(S+1) \times\left[\frac{3\tau_{c1}}{1+\omega_H^2\tau_{c1}^2} + \frac{7\tau_{c2}}{1+\omega_S^2\tau_{c2}^2}\right] + \frac{2}{3}S(S+1)\left(\frac{A}{\hbar}\right)^2 \times\left(\frac{\tau_e}{1+\omega_S^2\tau_e^2}\right)$$

where $S$ is the electron spin quantum number, $\gamma_H$ and $\gamma_S$ are proton and electron magnetogyric ratios, respectively, $r$ is the distance between the metal ion and the given nucleus of the coordinated substrate, $\omega_H$ and $\omega_S$ are the proton and electron Larmor frequencies, respectively. $\tau_{c1,2}$ and $\tau_e$ are the relevant correlation times for the time-dependent dipolar and contact interactions, respectively:

$$\frac{1}{\tau_{c1,2}} = \frac{1}{\tau_R} + \frac{1}{\tau_M} + \frac{1}{\tau_{S1,2}}$$

$$\frac{1}{\tau_e} = \frac{1}{\tau_{S1}} + \frac{1}{\tau_M}$$

where $\tau_R$ is the reorientational correlation time and $\tau_{S1}$ and $\tau_{S2}$ are the longitudinal and transverse electron spin relaxation times.

The contact term is often negligible, especially when dealing with a paramagnetic ion bound to a macromolecule because the hyperfine coupling constant $A/\hbar$ is in the MHz range, whereas the coefficient of the dipolar term is one order of magnitude larger.

## The protons relaxation enhancement (PRE) method

Upon interacting with a macromolecule the relaxation induced by paramagnetic species usually displays remarkable changes, primarily related to the increase of the molecular reorientational time $\tau_R$ on going from the free to the bound form. This may

result in a strong increase of the $R_{1p}^{is}$ term, from which it is possible to assess the affinity (and the number of binding sites) between the interacting partners. Quantitatively, in the presence of a reversible interaction between the paramagnetic species and the macromolecule, the observed enhancement depends on both the molar fraction of the macromolecular adduct $\chi^b$ and the solvent–water relaxation in the all-bound limit $R_{1p}^{b}$. The increase in the relaxation rate is expressed by the enhancement factor $\varepsilon^*$:

$$\varepsilon^* = \frac{R_{1p}^*}{R_{1p}} = \frac{\left(R_1^{obs} - R_1^{o}\right)^*}{\left(R_1^{obs} - R_1^{o}\right)}$$

The asterisk indicates the presence of the macromolecule in solution. The enhancement factor ranges from 1 (no interaction) to $\varepsilon^b = R_{1p}^b/R_{1p}$ in the all-bound limit (strong excess of macromolecule).

$$\varepsilon^* - 1 = \chi^b\left(\varepsilon^b - 1\right)$$

The determination of the binding parameter $K^A$ (association constant) and $\boldsymbol{n}$ (number of independent sites characterized by a given $K_A$ value) from the mass action law is straightforward. Given the association equilibrium between the metal complex

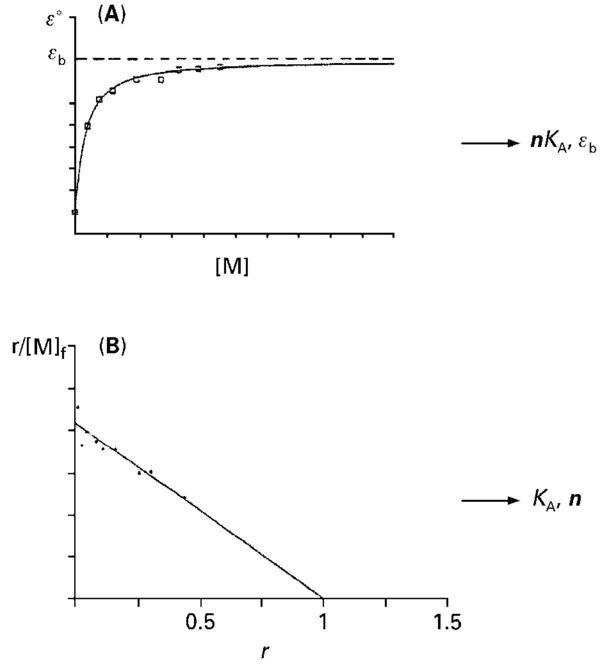



**Figure 5** Direct (A) and inverse (B) PRE titrations: by plotting $\varepsilon^*$ vs. macromolecule concentration a rectangular hyperbola is obtained, whose analysis yields the $\varepsilon^b$ value together with the $\boldsymbol{n}K_A$ product. The Scatchard's plot obtained from the inverse titration allows the determination of $\boldsymbol{n}$ and $K_A$ separately.

M and the protein (enzyme) E,

$$\mathrm{M} + \mathrm{E} \rightleftharpoons \mathrm{ME}$$

$$K_A = \frac{[\mathrm{ME}]}{[\mathrm{M}]_f[\mathrm{E}]_f} = \frac{[\mathrm{ME}]}{([\mathrm{M}]_t - [\mathrm{ME}])([\boldsymbol{n}\mathrm{E}]_t - [\mathrm{ME}])}$$

$$\varepsilon^* = \chi^b\varepsilon^b + \left(1 - \chi^b\right) = \frac{[\mathrm{ME}]}{[\mathrm{M}]_t}\varepsilon^b + \frac{[\mathrm{M}]_t - [\mathrm{ME}]}{[\mathrm{M}]_t}$$

where $n$ time [E] gives the concentration of each class of sites on the macromolecule. Superscript b and subscripts f and t indicate bound, free and total, respectively.

The experimental procedure consists of the determination of the enhancement factor $\varepsilon^*$ through two distinct titrations. In the first titration, a rectangular hyperbola describing the change of $\varepsilon^*$ as the $[\mathrm{E}]_t$ to $[\mathrm{M}]_t$ ratio increases is reported (**Figure 5**). The treatment of the obtained binding isotherm yields the value of $\varepsilon^b$ and the product $\boldsymbol{n}K_A$. In the second titration, the behaviour of $\varepsilon^*$is monitored at fixed macromolecule concentration by changing the metal complex concentration. Results are conveniently expressed under the form of a Scatchard's plot:

$$\boldsymbol{r}/[\mathrm{M}]_f = \boldsymbol{n}\,K_A - \boldsymbol{r}K_A$$

The value of $\boldsymbol{r}$ may easily be calculated once the all-bound enhancement $\varepsilon^b$ is known. In the same way, the free $[\mathrm{M}]_f$ concentration is obtained from the total $[\mathrm{M}]_t$ complex concentration.

$$\boldsymbol{r} = \frac{(\varepsilon^* - 1)[\mathrm{M}]_t}{(\varepsilon^b - 1)[\mathrm{E}]_t}$$

$$[\mathrm{M}]_f = \left(1 - \frac{\varepsilon^* - 1}{\varepsilon^b - 1}\right)[\mathrm{M}]_t$$

By plotting $r/[\mathrm{M}]_f$ vs. $r$ a straight line is obtained whose $x$-axis intercept gives the $\boldsymbol{n}$ value and whose slope is equal to $-K_A$.

### Contrast agents for magnetic resonance imaging

Magnetic resonance imaging (MRI), one of the most powerful tools in modern clinical diagnosis, is based on the topological representation of NMR parameters such as proton density and transverse and

longitudinal relaxation times. Differences in these parameters allow impressive anatomical discrimination to be made, and make it possible to distinguish pathological from healthy tissues. The potential of MRI is further strengthened by the use of suitable contrast agents (CAS), compounds that are able to alter markedly the magnetic properties of the region where they are distributed. Among them, paramagnetic $Gd^{3+}$ complexes (seven unpaired electrons) are under intense scrutiny because of their ability to enhance the proton relaxation rates of solvent water molecules (**Figure 6**). Their use provides the physicians with further physiological information to be added to the impressive anatomical resolution commonly obtained in the uncontrasted images. Thus, administration of Gd-based contrast agents has entered into the pool of diagnostic protocols and is particularly useful to assess organ perfusion and any abnormalities in the blood–brain barrier or in kidney clearance. Several other applications, primarily in the field of angiography and tumour targeting, will soon be available in clinical practice. Nowadays about 35% of MRI examinations make use of contrast agents, but this percentage is predicted to increase further following the development of more effective and specific contrast media than those currently available.

In order to be used as a CA for MRI a $Gd^{3+}$ complex must fulfil two basic requirements: (i) it must have at least one coordinated water molecule in fast exchange (on the NMR relaxation timescale) with the solvent bulk, and (ii) the $Gd^{3+}$ ion must be tightly chelated to avoid the release of the metal ion and the ligand which are both potentially harmful. **Figure 7** reports the schematic structures of four ligands whose $Gd^{3+}$ chelates are currently used as CAs for MRI. The overall PRE of water protons ($R_{1p}^{is} + R_{1p}^{os}$, see 'Theory of paramagnetic relaxation' section above) referred to a 1 mM concentration of $Gd^{3+}$ chelate is called *millimolar relaxivity* (hereafter *relaxivity*). At the magnetic field strength currently employed in MRI applications, the relaxivity of a $Gd^{3+}$ complex is roughly proportional to its molecular weight, i.e. it is determined by the value of the molecular reorientational time $\tau_R$. It follows that the search for high relaxivities has been addressed to systems endowed with slow tumbling motion in solution. This condition may be met either (i) by linking a $Gd^{3+}$ chelate to high molecular weight substrates like carboxymethylcellulose, polylysine or dendrimers, etc. or (ii) by forming noncovalent adducts with serum albumin. In the latter case the $Gd^{3+}$ chelates are designed in order to introduce on the surface of the ligand suitable functionalities able to strongly interact with the serum protein. In addition to providing high relaxivities (which, in turn, allows the administered doses of CA to be reduced) the formation of adducts between $Gd^{3+}$ complexes and albumin are of interest for the design of novel angiographic experiments for which a reduced clearance time and a better compartmentalization in the circulating blood is required.

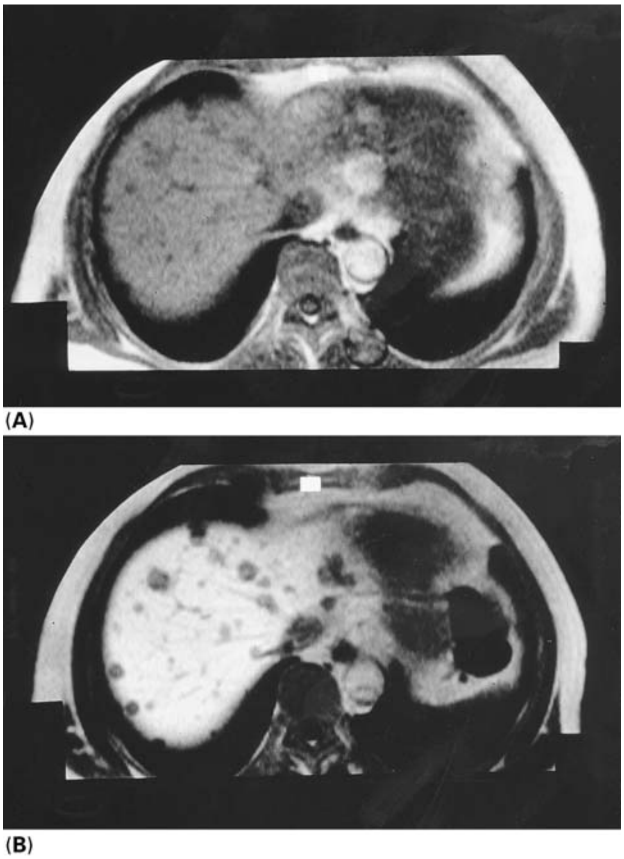



**Figure 6** Male patient, 54 year old, 67 kg with multiple hepatic metastases from a neuroendocrine tumour of the pancreas: images are taken before (A) and 90 minutes after (B) intravenous administration of 0.1 mmol $kg^{-1}$ of a liver-specific $Gd^{3+}$ chelate.

## List of symbols

$a$ = distance of closest approach between the paramagnetic centre and the diffusing solvent molecules; $D$ = relative solute–solvent diffusion coefficient; $\mathbf{H}$ = hyperfine shift; $\mathbf{H}_S$ = hyperfine shift scalar coupling; $\mathbf{H}_D$ = hyperfine shift dipole coupling; $K_i$ = interaction constant between LSR and S; $K_1$ and $K_2$ = coefficients related to anisotropy of magnetic susceptibility tensor; $K_A$ = association constant; $n$ = number of independent sites characterized by a given $K_A$ value; $q$ = number of interacting sites; $r_i$ = distance between nucleus and paramagnetic ion; $r_i$, $\theta_i$ and $\Omega_i$ = spherical coordinates of $i$ nucleus in the reference frame of the principal magnetic

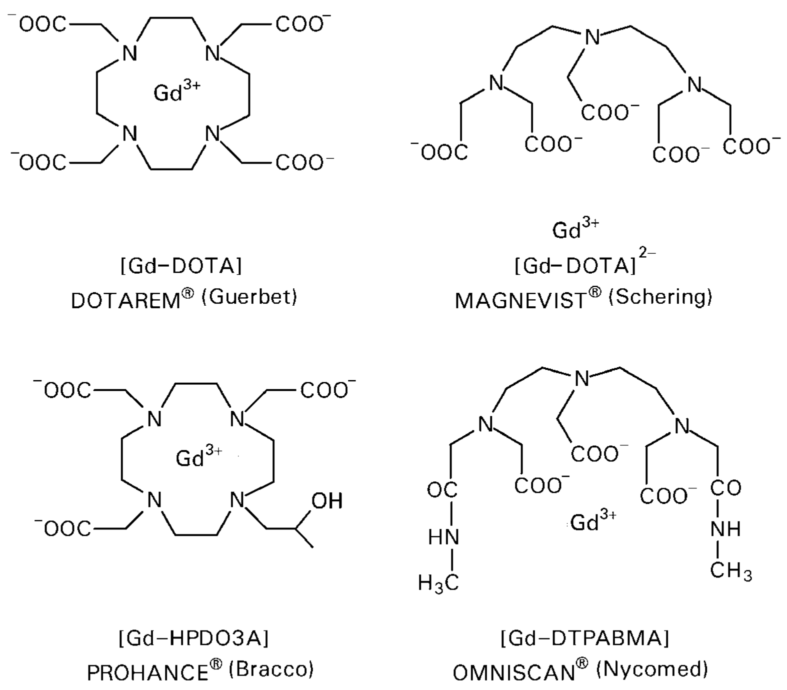



**Figure 7** Schematic structure of the $Gd^{3+}$ chelates used in clinical practice. DOTA = 1,4,7,10-tetraazacyclododecane-N,N′,N″,N‴-tetraacetic acid; DTPA = diethylenetriaminepentaacetic acid; HPDO3A = N‴-(2-(1-hydroxypropyl))-1,4,7,10-tetraazacyclododecane-N,N′N″-triacetic acid; DTPABMA = N,N″-bis(N-methylcarbamoylmethyl)-diethylenetriamine-N,N′,N″-triacetic acid.

susceptibility tensor; $R_1^{\mathrm{o}}$ = relaxation rate in absence of paramagnetic compound; $R_{1p}^{\mathrm{os}}$ = diffusion-controlled contribution to relaxation rate from water molecules in the outer coordination sphere of the paramagnetic centre; $R_{1p}^{\mathrm{is}}$ = paramagnetic contribution to relaxation rate due to exchange of water molecules in the inner coordination sphere with those in the bulk; $R_{1p}^{\mathrm{b}}$ = solvent–water relaxation in the all-bound limit; S = substrate; $T_{1M}$ = relaxation time of the protons of the water molecule in the first coordination sphere of the complex; $T_1$ = relaxation time; $\chi^{\mathrm{b}}$ = molar fraction of macromolecular adduct; $\Delta_i^{\mathrm{LSR\text{-}S}}$ = all-bound shift of the LSR-S complex; $\Delta\delta_i$ = observed LSR-induced shifts; $\Delta\Delta\delta$ = enantiomeric shift difference between the $\Delta\delta$ LIS for R and S configuration; $\varepsilon^*$ = enhancement factor; $\phi$ = angle between $r_i$ vector and Ln–ligand bond or rotation axis; $\gamma_{\mathrm{H}}$ and $\gamma_{\mathrm{S}}$ = electron and proton magnetogyric ratios; $\rho$ = LSR to S ratio; $\tau_{c1,2}$ and $\tau_e$ = relevant correlation times for time-dependent dipolar and contact interactions; $\tau_{\mathrm{M}}$ = residence lifetime of the water molecule in the first coordination sphere of the complex; $\tau_{\mathrm{R}}$ = reorientational correlation time; $\tau_{S1}$ and $\tau_{S2}$ = longitudinal and transverse electron spin relaxation times; $\omega_{\mathrm{H}}$ and $\omega_{\mathrm{S}}$ = proton and electron Larmor frequencies.

*See also:* **Contrast Mechanisms in MRI; EPR Spectroscopy, Theory; MRI Applications, Biological; MRI Applications, Clinical; MRI Applications, Clinical Flow Studies; MRI Theory; NMR Relaxation Rates; NMR Spectroscopy of Alkali Metal Nuclei in Solution; Nuclear Overhauser Effect.**

## Further reading

Aime S, Botta M, Fasano M and Terreno E (1998) Lanthanide (III) chelates for NMR biomedical applications. *Chemical Society Reviews* 27: 19–29.

Bertini I and Luchiant C (1986) *NMR of Paramagnetic Molecules in Biological Systems*. Menlo Park: Benjamin/Cummings.

Cockerill AF, Davies GLO, Harden RC and Rackham DM (1973) Lanthanide shift reagents for nuclear magnetic resonance spectroscopy. *Chemical Reviews* 73: 553–588.

Martin ML, Delpuech J-J and Martin GJ (1980) *Practical NMR Spectroscopy*, pp 377–409. London: Heyden.

Peters JA, Huskens J and Raber DJ (1996) Lanthanide induced shifts and relaxation rate enhancements. *Progress in NMR Spectroscopy* 28: 283–350.

Sherry AD and Geraldes CFGC (1989) Shift reagents in NMR spectroscopy. *In: Bünzli J-CG and Choppin GR (eds) Lanthanide Probes in Life, Chemical and Earth Science*. Amsterdam: Elsevier.

Von Ammon R and Fisher RD (1972) Shift reagents in NMR spectroscopy. *Angewandte Chemie, International Edition in English* 11: 675–692.

# Chemical Structure Information from Mass Spectrometry

**Kurt Varmuza**, Laboratory of Chemometrics, Vienna University of Technology, Austria

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Mass spectra of chemical compounds have a high information content. This article describes computer-assisted methods for extracting information about chemical structures from low-resolution mass spectra. Comparison of the measured spectrum with the spectra of a database (library search) is the most used approach for the identification of unknowns. Different similarity criteria of mass spectra as well as strategies for the evaluation of hitlists are discussed. Mass spectra interpretation based on characteristic peaks (key ions) is critically reported. The method of mass spectra classification (recognition of substructures) has interesting capabilities for a systematic structure elucidation. This article is restricted to electron impact mass spectra of organic compounds and focuses on methods rather than on currently available software products or databases.

## Introduction

Mass spectrometry (MS) is the most widely used method for the identification of organic compounds in complex mixtures at the nanogram level and below. NMR and, sometimes, also infrared spectroscopy provide better structural information than MS data but these spectroscopic methods cannot (or only with a substantial loss of sensitivity) be coupled with chromatographic separation techniques. Despite the great amount of information contained in mass spectra it is difficult to extract structural data because of the complicated relationships between MS data and chemical structures. The fragmentation processes which finally result in the measured data characterize MS as a chemical method, in contrast with NMR and IR. Chemical effects are, in general, more difficult to describe and to predict than physical ones. The occurrence of rearrangement reactions means that structure identification from mass spectra is often difficult or impossible; furthermore functional groups do not always produce the same peaks. This background explains why structure elucidation in organic chemistry is mainly based on $^{13}C$ NMR data, while H NMR, IR, UV–VIS and MS data are considered only to a much lesser extent.

The aim of spectra evaluation can be either the identification of a compound (assuming the spectrum is already known and available) or interpretation of spectral data in terms of the unknown chemical structure (when the spectrum of the unknown is not available).

Identification is performed best by library search methods based on spectral similarities; a number of MS databases and software products are offered for this purpose and are routinely used.

The more challenging problem is the interpretation of a mass spectrum, which is still a topic of research projects in chemometrics and computer chemistry. No comprehensive solutions are yet available and these methods are only rarely used in routine work. Four groups of different strategies have been applied to the complex problem of substructure recognition (or the recognition of more general structural properties) from spectral data. (1) *Knowledge-based methods* try to implement spectroscopic knowledge about spectra–structure relationships into computer programs. Because of the lack of generally applicable rules this approach was not successful in MS. However, spectroscopic knowledge has been extensively applied in other methods to guide the construction of mathematical models, and for optimizing model parameters. (2) Appropriate *interpretive library search* techniques can be used to obtain structural information if the unknown is not contained in the library. (3) *Correlation tables* containing characteristic spectral data (key ions) together with corresponding substructures have met with only limited success because a specific structural property does not always give the same spectral signals. (4) *Spectral classifiers* are algorithms based on multivariate classification methods or neural networks; they are constructed for an automatic recognition of structural properties from spectral data.

In principle, mass spectral data are fully determined by the chemical structure of the investigated

compound and the experimental spectroscopic conditions:

$$\text{mass spectral data} = F\,(\text{chemical structure, experimental spectroscopic conditions})$$

For problems of practical interest, however, the relationship $F$ can neither be formulated as a mathematical equation nor as an algorithm. Consequently, methods for the prediction of the mass spectrum from the chemical structure have had only very limited success. A similar difficult situation arises for the inverse problem, spectra interpretation, **Figure 1**. Only for selected cases can the relationship

$$\text{structural property} = f(\text{selected mass spectral data})$$

be established and applied to mass spectra of unknowns. Some relationships $f$ for different structural properties have been developed by the use of statistical methods or neural networks rather than by the influence of spectroscopic knowledge.

Because of the above-mentioned difficulties the intervention of the human expert is still required in mass spectra interpretation. However, this situation causes a number of drawbacks. Human experts are not sufficiently available and are expensive; furthermore their success rate can hardly be quantified. Automated instruments with high throughput are able to produce huge amounts of data, and analytical interest in complex mixtures (environmental chemistry, food chemistry, combinatorial synthesis) is rapidly increasing. Thus the extensive use of computer-assisted methods for data interpretation is highly desired.

## Characteristic mass spectra signals

A number of masses in the low mass region (*key ions*) are considered to be characteristic for certain substructures or classes of compounds. In addition, mass differences (*key differences*) between the molecular ion and abundant fragment ions or between abundant fragment ions are often related to functional groups. Correlation tables containing such spectral data and the corresponding chemical structures are contained in several textbooks on MS. The use of these tables is widespread and often helpful in MS; however, the capability of this approach must not be overestimated.

The following example demonstrates the potentials and drawbacks. From the NIST Mass Spectra Library a random sample containing 5000 compounds with a phenyl group (class 1) and 5000 compounds without a phenyl group (class 2) have been selected; from these data it was estimated whether the signal at *m/z* 77 (corresponding to the ion $C_6H_5^+$) can be used to predict the presence or absence of a phenyl group in the molecule. Let $n_1$ be the number of compounds from class 1 exhibiting a peak at *m/z* 77 in a given intensity interval, and $n_2$ be the corresponding number for compounds from class 2. Assuming for an unknown compound that the a priori probability for belonging to class 1 is 0.5 the probability $p(\text{phenyl} \mid I_{77})$ can be calculated from the signal at *m/z* 77 as

$$p(\text{phenyl} \mid I_{77}) = n_1/(n_1 + n_2) \qquad [1]$$

**Table 1** shows results for this example using four intensity intervals. Depending on the intensity of the peak at *m/z* 77, the probability for the presence of a phenyl group decreases or increases. If a peak at *m/z*

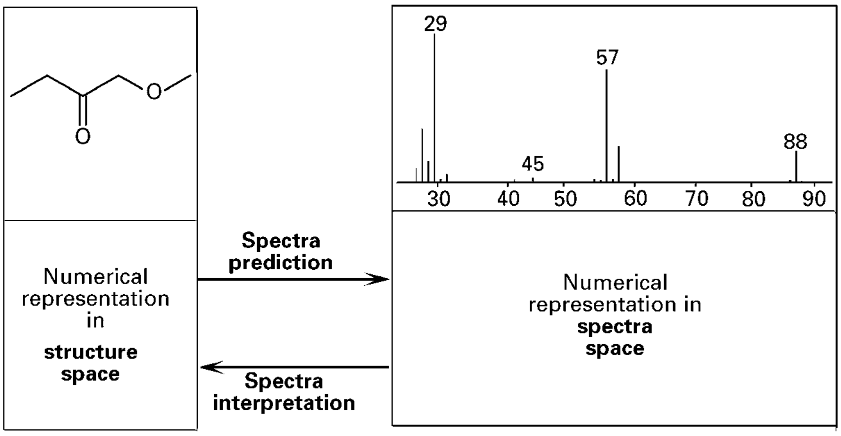



**Figure 1** Relationships between mass spectrum and chemical structure.

**Table 1** Significance of a peak at *m*/*z* 77 for the presence of a phenyl group in a molecule

| *Intensity interval (%B)* | $n_1$ | $n_2$ | $p(\text{phenyl} \mid I_{77})$ |
|---|---|---|---|
| $I \leq 1$ | 497 | 2486 | 0.17 |
| $1 < I \leq 5$ | 738 | 983 | 0.43 |
| $5 < I \leq 30$ | 2283 | 1288 | 0.64 |
| $30 < I \leq 60$ | 898 | 182 | 0.83 |
| $60 < I$ | 584 | 61 | 0.91 |
| Sum | 5000 | 5000 | |

The terms $n_1$ and $n_2$ are the number of spectra (compounds) with a peak at *m*/*z* 77 in the given intensity interval for compounds that contain a phenyl group and those that do not, respectively; $p(\text{phenyl} \mid I_{77})$ is the a posteriori probability for the presence of a phenyl group (assuming an a priori probability of 0.5)

77 is absent or has an intensity not higher than 1% of the base peak (%B) the probability for the presence of a phenyl group is still 17%. Intensities above 60%B strongly indicate a phenyl group (91% correct assignments to class 1). However, most spectra have peak intensities between these limits and therefore do not allow a reliable answer about the presence or absence of a phenyl group. A systematic examination of about 40 key ions between *m*/*z* 30 and 105 can be summarized as follows: (1) only a small number of key ions provide significant structural information, and (2) the absence of a key ion in general does not indicate the absence of the corresponding compound class.

## Mass spectra similarity search

### Overview

A basic assumption in library search systems is the hypothesis that 'similar spectra often correspond to similar structures'. A search for spectra in a library can be guided by two types of criteria: (1) a search profile is defined by some logical restrictions, for example "peak height at *m*/*z* 77 > 90%B and peak height at *m*/*z* 105 > 30%B". The result is a not-ordered hitlist containing all spectra that exactly match the search profile. (2) A numerical similarity between two spectra is used to find the reference spectra most similar to the spectrum of the unknown (**Figure 2**). In this case the resulting hitlist is ordered according to the value of the similarity criterion. Of course any type of search can be combined with additional search criteria (relative molecular mass range, restrictions to the molecular formula, name fragments, etc.). Unfortunately, the capabilities for chemical structure searches are still very limited in commercially available MS databases.

The performance and effectiveness of a spectral library search depends on the quality of the given spectrum, the content of the spectral library, the

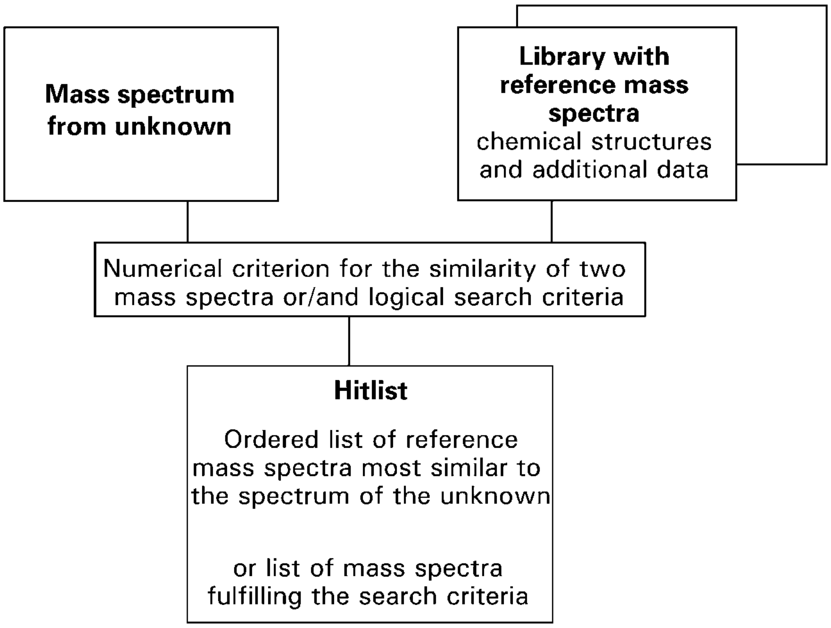



**Figure 2** Spectral library search.

**Table 2** Common quality problems with mass spectra in databases

| |
|---|
| Spectrum reduced to a small number of (large) peaks |
| Limited mass range (start mass too high) |
| Peaks present far above the relative molecular mass |
| Poor dynamics of peak intensities |
| Wrong isotope peak intensities |
| Tilted peak intensities |
| Additional peaks from impurities or instrument background |
| Small peaks missing |
| Wrong mass assignment |
| Illogical mass differences |
| Spectrum does not match with given compound |
| Missing experimental data |
| Missing structure data |
| Too many replicate spectra |

used similarity criterion and the implemented software technique. A fundamental restriction is given by the limited structural information present in mass spectra, for instance isomers can often not be distinguished by their mass spectra.

### Libraries

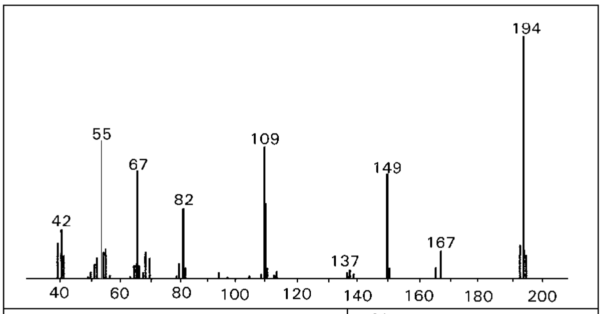
Any spectroscopic retrieval system in chemistry depends on an appropriate structural diversity of the reference data. Commercially available MS databases today contain about 40 000 up to almost 300 000 spectra. Some libraries contain a number of replicate spectra (collected from different sources), especially for common compounds. The compositions of existing libraries (except a few dedicated small spectra collections) are not systematic. Some compound classes are very well represented (for instance hydrocarbons) while others are not but may be of major interest to a particular user. Good software support for building user-libraries is therefore essential. **Table 2** lists common problems of quality deficiencies in MS reference data.

### Peak search

A simple and evident method for library search of mass spectra applies user-selected restrictions for the presence of peaks. Initially, a characteristic peak in the spectrum of the unknown is selected (typically one at a high or unusual mass number and with a not too small intensity) and an intensity interval for reference spectra is defined. The peak search software answers with the number of reference spectra containing a peak at the given mass and in the given intensity interval. This procedure is continued until a reasonable number of hits is reached (in cases where no reference spectrum fulfills all restrictions it is possible to go back one step). The resulting list of spectra can be ordered by applying an additional similarity criterion. The method allows (and requires) the interaction of the user; consequently spectroscopic knowledge (for instance about the background peaks) can be easily considered; however, automated processing of spectra is not possible. The use of appropriate sorted index files avoids time-consuming sequential searches and makes the method very fast. An example for this method is presented in **Figure 3**.

### Spectra similarity criteria

An infinite number of different similarity measures for matching two mass spectra is possible. In reality, about a dozen similarity criteria have been described and tested; however, full details of the algorithms that are actually implemented in commercial software are seldom described. Some similarity criteria are purely mathematically oriented, others are influenced by spectroscopic ideas. A single criterion cannot fit all requirements: on the one hand, a compound which is contained in the library should be retrieved even if the spectrum deviates in some parts from the reference spectrum; on the other hand, the user expects hits with chemical structures similar to the unknown even though they exhibit different spectra. In the ideal case the similarity criterion has a practicable sensitivity to chemical structures: identical compounds measured under different spectroscopic conditions appear more




| Selected peaks | | Number of reference spectra containing all given peaks |
|---|---|---|
| *m/z* | Intensity interval (%B) | |
| 194 | 70 – 100 | 421 |
| 109 | 30 – 80 | 15 |
| 82 | 10 – 60 | 4 |
| 67 | 20 – 80 | 1 |

**Figure 3** Peak search example using the NIST mass spectral database. The mass spectrum of a hypothetical unknown is from caffeine contaminated with a phthalate. Manual selection of relevant peaks easily allows the spectroscopist to consider probable contaminations (peaks at *m/z* 149, 167). The correct solution is found after the input of four peaks by excluding the typical phthalate peaks. Note the wide intensity intervals applied. Should a peak at *m/z* 149 be required the correct compound is not found.

similar than two compounds that are slightly different in their structure.

The two main types of mass spectral data, peak heights and masses, have different reliability. While the presence of a peak at a certain mass number is well reproducible the peak height often varies greatly under slightly different experimental conditions. Spectra similarity criteria have to consider this aspect.

For similarity searches an efficient ordering of the spectra in the library is not possible and, therefore, a rather time-consuming sequential search is necessary. Some speed-up is possible by a pre-search in which, for instance, reference spectra are quickly eliminated that do not have a minimum number of peaks in common with the spectrum of the unknown. Before calculating the spectra similarity the spectrum of the unknown is treated by cleaning procedures, for instance peaks at *m/z* 28, 32 and 40 (probably from $N_2$, $O_2$ and Ar respectively) and peaks from column bleeding are removed, and the peak heights are normalized to percents of the base peak. The selection of relevant peaks and weighting of peak heights is described below. Let ${}_{\mathrm{U}}A_m$ and ${}_{\mathrm{R}}A_m$ be the (weighted) abundances at mass $m$ in the spectrum of the unknown (U) and the reference spectrum (R), respectively. Summation in Equations [2] to [4] is over all selected masses $m$.

The most used similarity criterion for mass spectra is based on the *correlation coefficient*:

$$S_1 = \frac{\left[\Sigma_{\mathrm{U}}\, A_m\, {}_{\mathrm{R}}A_m\right]^2}{\sqrt{\left[\Sigma^1\, ({}_{\mathrm{U}}A_m)^2\, \Sigma\, ({}_{\mathrm{R}}A_m)^2\right]}} \qquad [2]$$

**Figure 4** explains this similarity criterion: each mass is considered as a point in a coordinate system with one axis for the abundances of the unknown and the other for the reference. For identical spectra all points are located on a straight line that passes through the origin ($S_1 = 1$). If the two spectra are different the correlation coefficient for the regression line – as given in Equation [2] – is a measure of the similarity of the spectra. The value of $S_1$ ranges from 0 to 1 because the regression line is forced to pass through the origin and all abundances are 0 or positive. The important range of $S_1$ between 0.9 and 1 can be widened by calculating $S_1^2$; for practical reasons $S_1$ or $S_1^2$ are often multiplied by 100.

A peak list with nominal (integer) masses can be considered as a point in a $q$-dimensional space, with $q$ equal to the maximum relevant mass number. The coordinates of the point are given by the abundances (weighted peak heights). Because the value of $q$ is high the $q$-dimensional space cannot be visualized directly but can be handled by rather simple mathematics. A two-dimensional simplification is used in **Figure 5** to demonstrate this view of MS data. Masses 43 and 58 have been selected for the coordinates; each spectrum (from the example in **Figure 4**) can be represented by a point or by a vector (starting at the origin). Similarity criterion $S_1$ is equal to the cosine of the angle between two spectral vectors.

Alternative similarity measures are the *Euclidean distance* $S_2$:

$$S_2 = \sqrt{\left[\Sigma\, ({}_{\mathrm{U}}A_m - {}_{\mathrm{R}}A_m)^2\right]} \qquad [3]$$

and the *city block distance* $S_3$:

$$S_3 = \Sigma \mid {}_{\mathrm{U}}A_m - {}_{\mathrm{R}}A_m \mid \qquad [4]$$

The similarity criterion used in the PBM (Probability Based Matching System) software is based on so-called uniqueness values for masses and abundances. The uniqueness of a signal is equivalent to the information content (as defined in information theory) and is given by the negative logarithm of the probability of the occurrence of this signal; this means that a low probability corresponds to a high uniqueness. Probabilities for masses and intensity intervals have been estimated from the database used. Matching of spectra is performed by using selected peaks that exhibit highest uniqueness values for mass and intensity.

The MS database system MassLib uses a composite similarity criterion containing the similarity criterion $S_1$ together with measures for the number and intensity sums of peaks that are common or not in both spectra. The algorithm has been optimized to give good results for identification as well as for interpretation.

In the STIRS (Self-training Interpretive and Retrieval System) system for 26 data classes, specific similarity criteria are defined using characteristic masses or mass differences. For each of these criteria the most similar reference spectra are searched with the aim of obtaining information about the presence of substructures.

## Weighting

Intensity scaling and mass weighting is used to consider the different significances of mass numbers and peak heights. In general, peaks in the higher mass range are more important than peaks in the lower mass range; because large peaks dominate the values of most similarity criteria, peak heights are often scaled to enhance the influence of small peaks. A

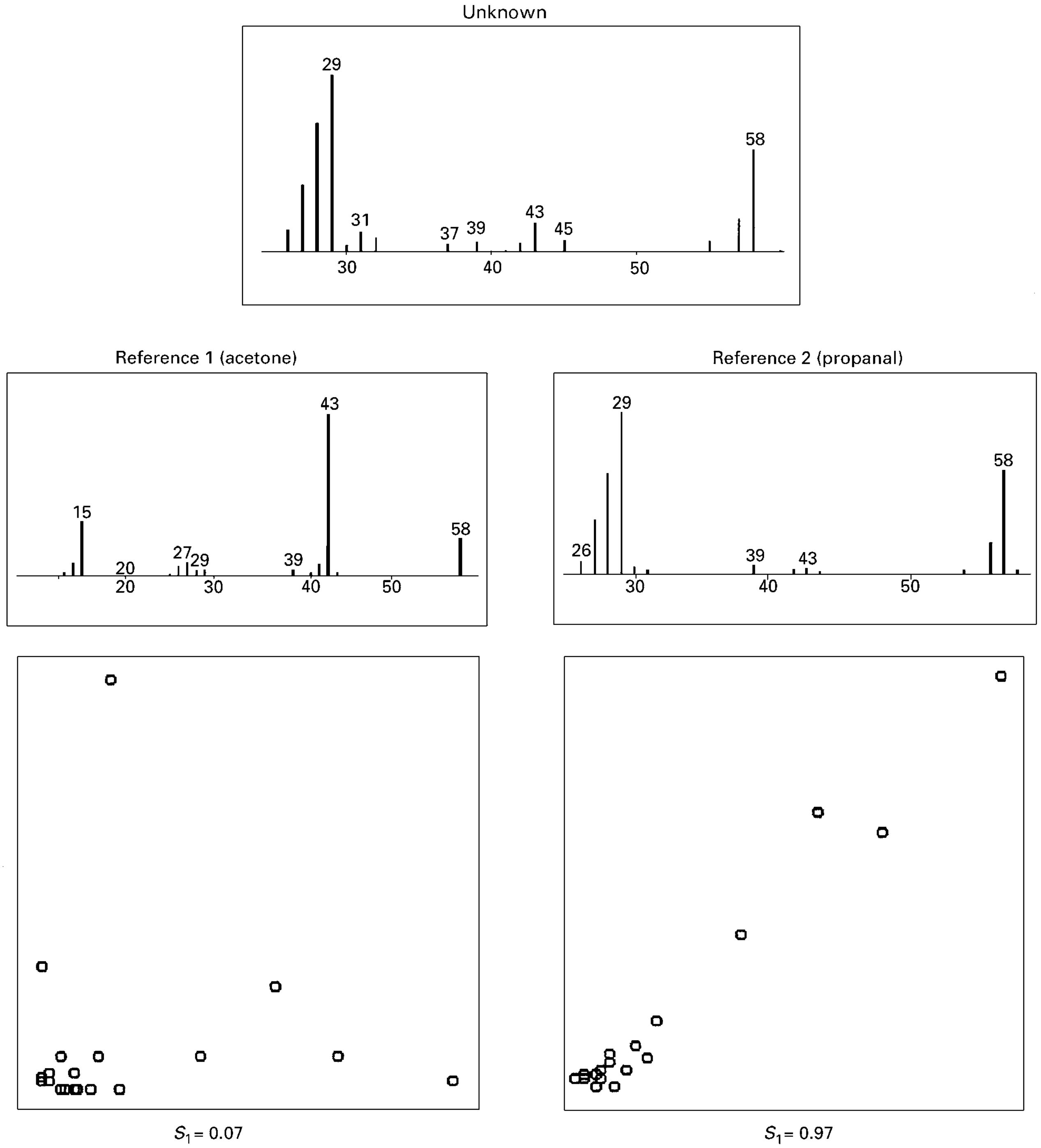



**Figure 4** Comparison of the mass spectrum from an unknown with two reference spectra using the similarity measure $S_1$ (based on the correlation coefficient). Axes in the two plots at the bottom: horizontal, peak height %B of unknown; vertical, peak height %B of reference. Each point corresponds to a mass number with a peak in at least one of the two compared spectra.

general transformation of intensities $I_m$ at mass $m$ is given by Equation [5]

$$A_m = (m)^v (I_m)^w \qquad [5]$$

Optimum values for $v$ and $w$ depend on the used similarity citerion; for $S_1$ (Eqn [1]) good results have been obtained with $v = 2$ and $w = 0.5$ [see Stein (1994) in the Further reading section].

### Selection of revelant masses

The selection of masses used to calculate a spectral similarity may greatly influence the result. The user

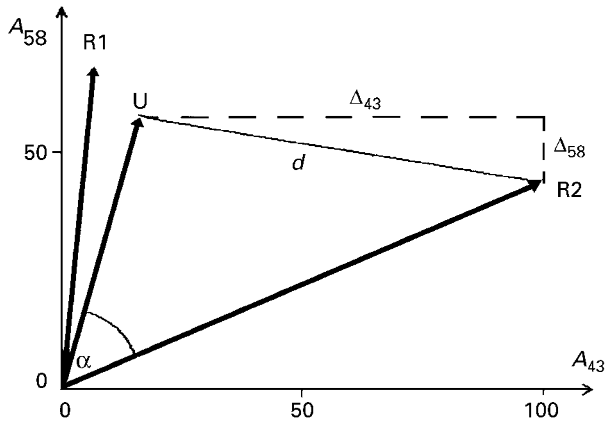



**Figure 5** Mass spectra can be considered as points or vectors in a multidimensional spectral space. For simplicity only two mass numbers (43, 58) have been selected in this example. $A_m$, abundance (peak height in %B) at mass $m$; U, spectrum from unknown; R1, reference spectrum of propanal; R2, reference spectrum of acetone. Measures for spectral similarity are the Euclidean distance ($d$), the city block distance ($\Delta_{43} + \Delta_{58}$) or the cosine of angle α (equivalent to $S_1$ in Eqn [2]).

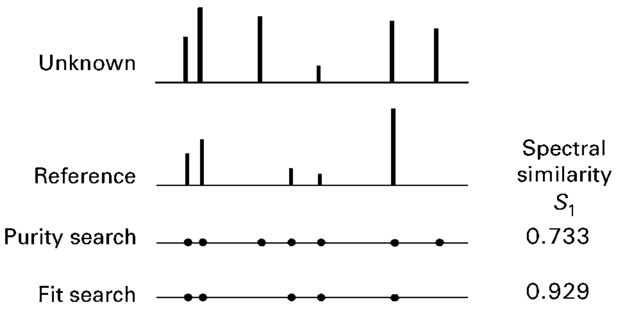



**Figure 6** Different selection of masses for the calculation of spectral similarity (schematic). The fit search considers only masses with peaks in the reference spectrum and therefore is insensitive to impurities in the unknown.

often can decide between three modes for mass selection (**Figure 6**).

(a) All masses for which at least one of the compared spectra has a peak with an intensity above a defined threshold are considered. High similarity values are obtained only if the spectra are almost identical; this approach therefore is suited for *identity searches* (sometimes also called *purity searches*).

(b) Only masses with a peak in the reference spectrum are considered. High similarity values are obtained if the reference spectrum is almost completely contained in the spectrum of the unknown. Non-matching peaks in the unknown do not influence the result; therefore this approach is routinely applied if the unknown may contain impurities or is a mixture of compounds (*fit search, reverse search*).

(c) Only masses with a peak in the spectrum of the unknown are considered. This method is not tolerant of additional peaks in the reference spectrum. From a high similarity value one may conclude that parts of the chemical structure of the reference compound are present in the unknown (*interpretive search, forward search*).

Simple selection methods of peaks such as '*k* highest peaks in the spectrum (or in given mass intervals)' have been applied to save data storage and computation time but are no longer important. A special selection of masses is used in the MassLib system: a peak at mass $m$ is considered to be significant if its intensity is higher than the average intensities at masses $m - 14$ and $m + 14$.

**Figure 7** presents an example of an MS library search in which the unknown was contained in the library.

## Mass spectra classification

### Overview

When the unknown compound is not contained in the spectral library then pure identification methods are less useful. For 'unknown unknowns' interpretive search systems or classification methods are required to obtain structural information that can be used for constructing molecular structure candidates. Such candidates are usually created manually by applying spectroscopic–chemical knowledge and intuition. A serious drawback of this strategy is that the solution is rarely complete and the procedure hardly can be documented or verified.

The most important systematic approach for structure elucidation of organic compounds is still based on the DENDRAL project, in which, for the first time, artificial intelligence principles have been applied to complex chemistry problems (**Table 3**). The central tool is an isomer generator software capable of generating an exhaustive set of isomers from a given molecular formula. The generator also has to consider structural restrictions, which are usually obtained from spectral data. Substructures which have to be present in the unknown molecular structure are collected in the so-called goodlist while forbidden substructures are put into the badlist. Mass spectrometry can contribute to this approach in various aspects: the molecular formula can be determined from high resolution data, and structural information can be derived even from low resolution data.

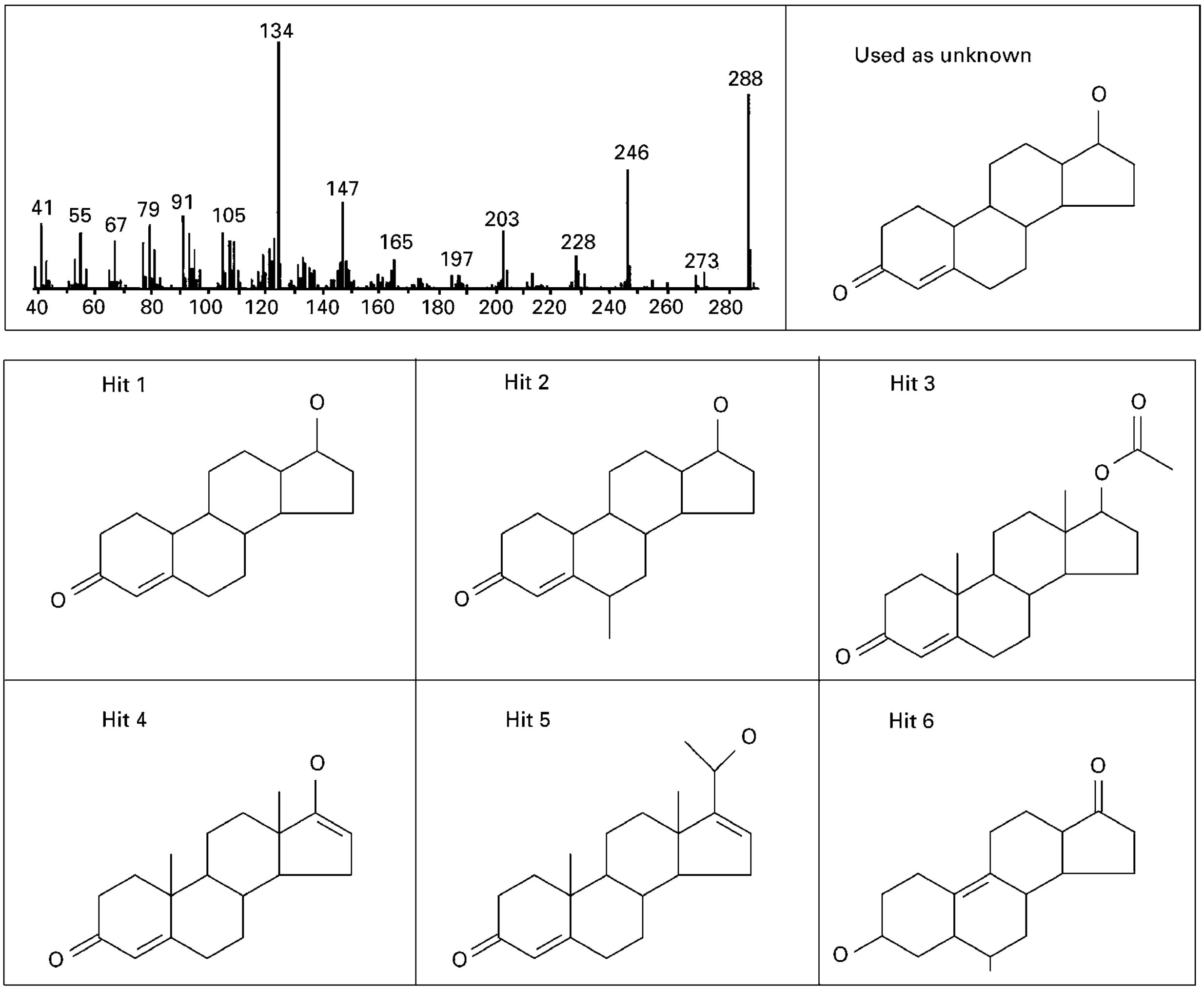



**Figure 7** Example of a library search by MassLib. A mass spectrum from testosterone has been considered as unknown and searched in a library consisting of 130 000 spectra (including duplicates). A mass spectrum from testosterone is contained in this library and has been found as the first hit (most similar spectrum). The other hits are from the same compound class and demonstrate the interpretive capabilities of this system.

The two most important computer-assisted strategies for the recognition of structural information from mass spectral data are:

(a) the structures in the hitlist from a spectra similarity search are used to estimate the probability of substructures in the unknown.
(b) Random samples of mass spectra are first characterized by a set of variables and then methods from multivariate statistics or neural networks are applied to develop spectral classifiers.

### Substructure recognition by library search

Library search systems have been developed or adjusted for the purpose of obtaining structurally similar compounds in the hitlist. In such methods an evaluation of the molecular structures of the hitlist compounds can provide useful information about the presence or absence of certain substructures in the unknown.

It is not trivial to define or select substructures that should be considered for this purpose. For the STIRS system, considerable effort has gone into the search for substructures that can be successfully classified by the implemented spectral similarity search. The MassLib system uses a predefined set of 180 binary molecular descriptors to characterize the similarity of structures. In most investigations a more or less arbitrary set of substructures, functional groups or more general structural properties (compound classes) has been considered. Self-adapting methods that automatically analyse the molecular structures in the hitlist (for instance by searching for frequent and large substructures) have not been used up to now in MS.

**Table 3** Scheme for systematic structure elucidation

(1) Determine the molecular formula of the unknown
(2) Derive structural restrictions as substructures from spectral and other experimental data and from pre-knowledge about the unknown
(3) Make consistency checks of the structural restrictions
(4) Add all substructures that are considered to be present in the unknown to the goodlist. Add all substructures that are considered to be absent in the unknown to the badlist
(5) Generate all isomers for the given molecular formula that contain all substructures from the goodlist but none from the badlist
(6) Test the generated molecular candidate structures:
   (a) by a comparison of predicted spectra with measured spectra
   (b) by a comparison of predicted properties with measured properties
   (c) by considering more complicated structural restrictions that could not be handled by the isomer generator
(7) If the number of survived molecular candidates is too large try to create additional structural restrictions and continue at step 3

The number of occurrences of a certain substructure in the hitlist is compared with the corresponding number for the library and a probability is derived for the presence of that substructure in the unknown. This classification method is a variant of the well-known '*k*-nearest neighbour classification'. Each mass spectrum is considered as a point in a multidimensional space; the neighbours nearest to the spectrum of the unknown correspond to the most similar reference spectra in library search. If the majority of *k* neighbours (*k* is typically between 1 and 10) contain a certain substructure then this substructure is predicted to be present in the unknown. A drawback of this approach is the high computational effort necessary for classifying an unknown because a full library search is required. The performance has been described by Stein (1995, see Further reading section) as 'sufficient to recommend it for routine use as a first step in structure elucidation'.

### Multivariate classification

A *mass spectral classifier* is a part of a computer program that uses the peak list of a low resolution mass spectrum as input and produces information about the chemical structure as output. For such a classification procedure a number of methods are available in multivariate statistics. Many of them have already been applied to various problems in chemistry; classification of mass spectra (with the aim of recognizing chemical compound classes) was one of the pioneering works in chemometrics.

Application of most multivariate classification methods first requires a transformation of the data to be classified (the mass spectrum) into a fixed number of variables (so-called features). Besides the simple use of the peak intensities for these variables more sophisticated transformations have been applied. Problem-relevant *numerical spectral features* $x_j$ are defined as linear or nonlinear functions of the peak intensities; the definitions are based on spectroscopic and/or mathematical concepts. **Table 4** contains a list of frequently used spectral features for mass spectra. It has been shown that such features are more closely related to chemical structures than the original peak data; an example will illustrate this. A characteristic fragment in a homologous series of compounds may not appear at the same mass but may be shifted by a multiple of 14 mass units, corresponding to $(CH_2)_n$. The often used modulo-14 summation features (also called mass periodicity spectra) consider this fact by an intensity summation in mass intervals of 14. Other spectral features are based on intensity ratios and reflect competing fragmentation pathways; autocorrelation features contain information about mass differences between abundant peaks. The stability of molecular ions and some functional groups can be characterized by features that describe the distribution of peaks across the mass range. The appropriate generation and selection of spectral features is considered to be the most essential part in the development of spectral classifiers.

The development of a classifier (**Figure 8**) for a particular substructure requires a random sample of spectra which is selected from a spectral library. Typically some 100 spectra from compounds containing the substructure (class 1) and some 100 spectra from compounds not containing the substructure (class 2) are necessary. One part of the

**Table 4** Types of numerical spectral features; a set of spectral features can be used to classify a mass spectrum as to whether a certain substructure or a more general structural property is present or absent in the molecule

| | | |
|---|---|---|
| Modulo-14 intensity sums | $x_j = \sum I_{m+14n}$ | $n$ : 0, 1, 2,…, $j$, $m$ : = 1, 2,…, 14 |
| Logarithmic intensity ratios | $x_j = \ln I_m / I_{m+\Delta m}$ | $I_m$ = max ($I_m$, 1) |
| Autocorrelation | $x_j = \sum I_m I_{m+\Delta m} / \sum I_m I_m$ | $j$, $\Delta m$ : 1, 2,…, 50 $m$ : a defined mass range |
| Distribution of peaks | $x_j = \sum I_{even} / \sum I_m$ | even : all even masses $m$ : all masses |
| Relative base peak intensity | $x_j = 100 / \sum I_m$ | $m$ : all masses |

$x_j$, spectral feature; $I_m$, intensity in %base peak at mass $m$; $\Delta m$, mass difference

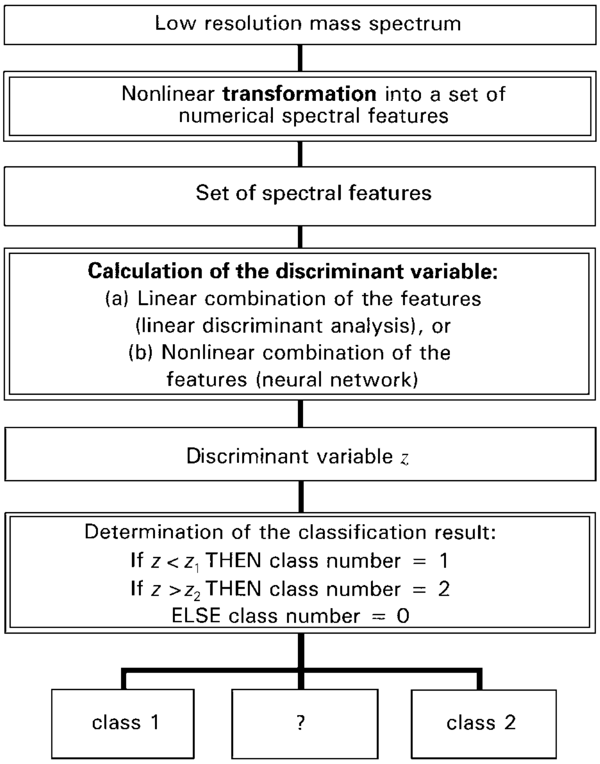



**Figure 8** Classification scheme for mass spectra. The mass spectrum is transformed into a set of typically 5 to 20 variables (spectral features). A linear or nonlinear mathematical combination of the features results in a discriminant variable. The value of the discriminant variable together with the required maximum statistical risk for wrong answers determines which of the three possible answers is given: 'class 1' (substructure present), 'class 2' (substructure absent) or 'answer not possible because error risk is too high'.

random sample is used for a training of the classifier, the other part for testing it. Classification is based on the value of a discriminant variable $z$ which is defined either as a linear function of the $q$ selected spectral features $x_j$:

$$z = b_1x_1 + b_2x_2 + \cdots + b_qx_q \qquad [6]$$

or as a nonlinear function which is usually implemented as a neural network. During the training process the parameters of the classifier can be directly calculated (as for instance by multiple linear regression or partial least-squares regression) or they have to be adjusted iteratively (if a neural network is used). The aim of the training is to obtain values for $z$ as defined by target values (for instance +1 for class 1 and −1 for class 2). The test set (which has not been used for the training) serves to estimate optimum classification thresholds $z_1$ and $z_2$ for applications of the classifier to unknowns by the scheme:

$$\begin{array}{lll} \text{IF} & z < z_1 & \text{THEN assign compound to class 1} \\ \text{IF} & z > z_2 & \text{THEN assign compound to class 2} \\ & & \text{ELSE reject classification answer} \end{array} \qquad [7]$$

Simple yes/no classifiers without the possibility to refuse the classification answer are not adequate for the recognition of substructures from mass spectra.

Today's performance of mass spectra classification by multivariate methods can be summarized as follows: (1) only a rather small number of substructures can be recognized with a low error rate. (2) Predictions of the absence of a substructure are usually more accurate than predictions of its presence. (3) Erroneous classifications cannot be avoided completely; therefore the intervention of a human expert and the parallel use of other spectra interpretation methods are advisable. (4) For small molecules a systematic and almost complete structure elucidation is sometimes possible by mass spectra classification and by application of the obtained structural restrictions in automatic isomer generation.

An example for this approach of structure elucidation is presented in **Figure 9**, where ethyl 2-(2-hydroxyphenyl)acetate has been considered as unknown. The molecular formula and the mass spectrum are given. Application of software MSclass resulted in substructures that are assumed to be present (goodlist) and others that are assumed to be absent (badlist); only substructures relevant to the molecular formula are shown. Considering these structural restriction, six isomers are possible, including the correct solution. The isomer generator used was MOLGEN. The computation time on a Pentium 233 MHz is more than 20 hours for the generation of all isomers, but only 4 s when the structural restrictions of the goodlist and badlist are considered.

## Conclusions

The best worldwide performance has been claimed for more than one commercial MS database system. However, more neutral observers state that automated spectra interpretation systems have a rather limited scope. Spectra library search systems are now widely used in MS laboratories and do a good job with routine problems. They are not so useful with complex problems or if the unknown is not contained in the library; however, current research promises considerable improvements of these methods in the future. The routine application of library

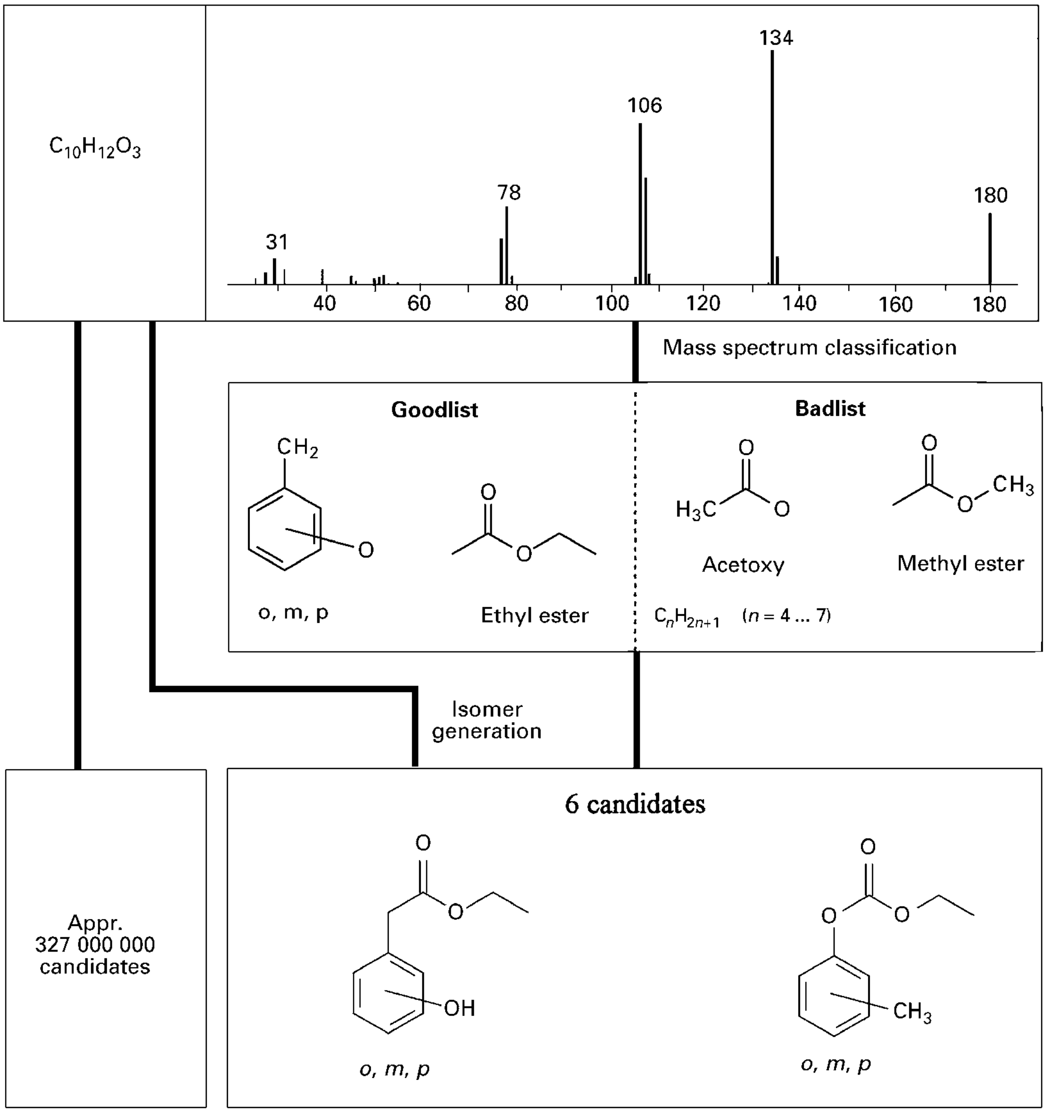



**Figure 9** Systematic structure elucidation using the molecular formula of the unknown, structural restrictions from automatic mass spectra classification and exhaustive isomer generation. Ethyl 2-(2-hydroxyphenyl)acetate is the 'unknown'.

search methods and spectra classification liberate the human spectroscopist from time-consuming, simple work or at least provides valuable preliminary suggestions for the expert.

## List of symbols

$A_m$ = abundance at mass $m$ (weighted peak intensity); $b$ = parameter of discriminant function; $d$ = distance; $I_m$ = intensity (peak height) at mass $m$ in %base peak; $m$ = mass number; $n$ = number of compounds or spectra; $p$ = probability; $q$ = number of variables (masses, spectral features); $R$ = reference spectrum; S = similarity between two mass spectra; U = spectrum from unknown; $x$ = spectral feature; $z$ = discriminant variable.

*See also:* **Chromatography-MS, Methods; Forensic Science, Applications of Mass Spectrometry; Fragmentation in Mass Spectrometry; Laboratory Information Management Systems (LIMS); Medical Applications of Mass Spectrometry; Pyrolysis Mass Spectrometry, Methods.**

## Further reading

Adams MJ (1995) *Chemometrics in Analytical Spectroscopy*. Cambridge: The Royal Society of Chemistry.

Clerc JT (1987) Automated spectra interpretation and library search systems. In Meuzelaar HLC and Isenhour TL (eds) *Computer Enhanced Analytical Spectroscopy*, Vol 1, pp 145–162. New York: Plenum Press.

Davis R and Frearson M (1987) *Mass Spectrometry (Analytical Chemistry by Open Learning)*. Chichester: Wiley.

Drablos F (1987) Symmetric distance measures for mass spectra. *Analytica Chimica Acta* **201**: 225–239.

Gasteiger J, Hanebeck W, Schulz KP, Bauerschmidt S and Höllering R (1993) Automatic analysis and simulation of mass spectra. In Wilkins CL (ed) *Computer-Enhanced Analytical Spectroscopy*, Vol 4 pp 97–133. New York: Plenum Press.

Gray NAB (1986) *Computer-Assisted Structure Elucidation*. NewYork: John Wiley.

Massart DL, Vandeginste BGM, Buydens LMC, DeJong S, Lewi PJ and Smeyers-Verbeke J (1997) *Handbook of Chemometrics and Qualimetrics*: *Part A*. Amsterdam: Elsevier.

McLafferty FW, Loh SY, and Stauffer DB (1990) Computer identification of mass spectra. In Meuzelaar HLC (ed) *Computer-Enhanced Analytical Spectroscopy*, Vol 2 pp 163–181. New York: Plenum Press.

McLafferty FW and Turecek F (1990) *Interpretation of Mass Spectra*, 4th edn. Mill Valley: University Science Books.

Owens KG (1992) Application of correlation analysis techniques to mass spectral data. *Applied Spectroscopy Reviews* **27**: 1–49.

Stein SE and Scott DR (1994) Optimization and testing of mass spectral library search algorithms for compound identification. *Journal of the American Society of Mass Spectrometry* **5**: 859–866.

Varmuza K and Werther W (1996) Mass spectral classifiers for supporting systematic structure elucidation. *Journal of Chemical Information and Computer Science* **36**: 323–333.

# Chemically Induced Dynamic Nuclear Polarization

*See* **CIDNP, Applications.**

# Chemometrics Applied to Near-IR Spectra

*See* **Computational Methods and Chemometrics in Near-IR Spectroscopy.**

# Chiroptical Spectroscopy, Emission Theory

**James P Riehl**, Michigan Technological University, Houghton, MI, USA

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| **Theory** |

All interactions involving light with chiral molecules discriminate between the two possible circular polarizations (left = L and right = R). In the absence of a perturbing static electric or magnetic field, the light emitted by a molecular chromophore will be partially circularly polarized if the emitting species is *chiral* or *optically active*. In this context these two terms have identical meaning, and describe a molecular structure in which the mirror image isomers (enantiomers) are not superimposable. The experimental technique of analysing the extent of circular polarization in the light emitted from chiral molecules has been variously referred to as circularly polarized emission (CPE), emission circular intensity differentials (ECID), or circularly polarized luminescence (CPL). In this article we will use the acronym CPL to describe this spectroscopic technique. In addition, it will not be necessary in the discussion presented here, to use the more specific terms of circularly polarized fluorescence or circularly polarized

phosphorescence. Although CPL and circular dichroism (CD) spectroscopy, i.e. the difference in absorption of left versus right circular polarization, share many common characteristics, they differ from one another in two main ways. First, due to the Franck–Condon principle, CPL probes the chiral geometry of the excited state in the same way that CD probes the ground state structure, and secondly, CPL measurements reflect molecular motions and energetics that take place between the excitation (absorption) and emission. Both of these effects have been exploited in CPL measurements, and will be a major focus of what is presented here.

## Circularly polarized luminescence transition probabilities

In CPL spectroscopy one is interested in measuring the difference in the emission intensity ($\Delta I$) of left circularly polarized light ($I_L$) versus right circularly polarized light ($I_R$). By convention this difference is defined as follows

$$\Delta I \equiv I_L - I_R \qquad [1]$$

Just as in ordinary luminescence measurements, the determination of absolute emission intensities is quite difficult, so it is customary to report CPL measurements in terms of the ratio of the difference in intensity, divided by the average total luminescence intensity.

$$g_{lum} = \frac{\Delta I}{\frac{1}{2}I} = \frac{I_L - I_R}{\frac{1}{2}(I_L + I_R)} \qquad [2]$$

$g_{lum}$ is referred to as the luminescence dissymmetry ratio (or factor). The extra factor of $\frac{1}{2}$ in Equation [2] is included to make the definition of $g_{lum}$ consistent with the previous definition of the related quantity in CD, namely, $g_{abs}$.

$$g_{abs} = \frac{\Delta \varepsilon}{\varepsilon} = \frac{\varepsilon_L - \varepsilon_R}{\frac{1}{2}(\varepsilon_L + \varepsilon_R)} \qquad [3]$$

where in this equation $\varepsilon_L$ and $\varepsilon_R$ denote, respectively, the extinction coefficients for left and right circularly polarized light.

The time dependence of the intensity of emitted light of polarization $\sigma$, at a particular wavelength, $\lambda$, for a specific transition $n \rightarrow g$ of a molecule with orientation $\Omega$ (in the laboratory frame) may be expressed in terms of the transition probability, $W^{\sigma}_{gn}$ as follows

$$I_{\sigma}(\lambda, t) = (hc/\lambda) N_n(\Omega, t) W^{\sigma}_{gn}(\Omega) f_{\sigma}(\lambda) \qquad [4]$$

where $N_n$ $(\Omega, t)$ denotes the number of molecules in the emitting state $|\,n\rangle$ with orientation $\Omega$ at time $t$, and $f_{\sigma}(\lambda)$ is a normalized lineshape function. The number of molecules in the emitting state with orientation $\Omega$ depends on the orientation at the time of excitation ($t = 0$) with respect to the polarization $\pi$ and direction of the exciting beam.

$$N_n(\Omega, t) = N^{\pi}_{e}(\Omega : \Omega_0, t)\eta_n \qquad [5]$$

$\eta_n$ denotes the fraction of molecules that end up in state $|\,n\rangle$ that were initially prepared in the intermediate state $|\,e\rangle$ by the excitation beam. This quantity is assumed to be independent of orientation. In the most general case we would also need to describe the time dependence of any conformational changes that take place between the time of excitation and emission, but for simplicity this will not be considered in the treatment presented here. The differential intensity of left minus right circularly polarized light may now be expressed as follows

$$\Delta I(\lambda, t) = (hc/\lambda) N_n(\Omega, t) \Delta W_{gn}(\Omega) f(\lambda) \qquad [6]$$

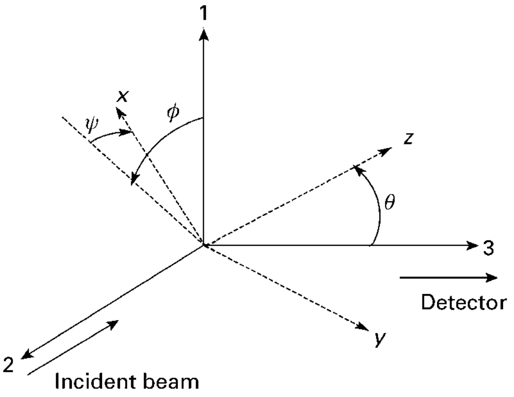
where we have introduced the differential transition probability $\Delta W_{gn}$

$$\Delta W_{gn}(\Omega) \equiv W^{L}_{gn}(\Omega) - W^{R}_{gn}(\Omega) \qquad [7]$$

The probability of emitting a right or left circularly polarized photon may be related in the usual way to molecular transition matrix elements through Fermi's Golden Rule. Under the assumption that the emitted light is being detected in the laboratory 3 direction (see **Figure 1**), and allowing for electric dipole and magnetic dipoles in the expansion of the molecule–radiation interaction Hamiltonian we obtain the following expressions

$$W^{L}_{gn}(\Omega) = K(\lambda^3)[|\mu^{gn}_1|^2 + |\mu^{gn}_2|^2 + |m^{gn}_1|^2 + |m^{gn}_2|^2 - 2i(\mu^{gn}_1 m^{gn}_1 + \mu^{gn}_2 m^{gn}_2)] \qquad [8]$$

$$W^{R}_{gn}(\Omega) = K(\lambda^3)[|\mu^{gn}_1|^2 + |\mu^{gn}_2|^2 + |m^{gn}_1|^2 + |m^{gn}_2|^2 + 2i(\mu^{gn}_1 m^{gn}_1 + \mu^{gn}_2 m^{gn}_2)] \qquad [9]$$




$$\hat{R} \equiv \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & \sin\theta\sin\phi \\ \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & -\sin\theta\cos\phi \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{pmatrix}$$

**Figure 1** Laboratory (1, 2, 3) and molecular (*x*, *y*, *z*) coordinate systems, and the Euler rotation matrix, ($R$).

where $K(\lambda^3)$ is a proportionality constant, 1 and 2 refer to laboratory axes and the electric dipole transition moment $\mu^{gn}$ and the imaginary magnetic dipole transition moment $m^{gn}$ are defined as follows.

$$\begin{aligned} \mu_1^{gn} &\equiv < g|\mu_1|n > \\ m_1^{gn} &\equiv < g|m_1|n > \end{aligned} \qquad [10]$$

The differential transition rate is, therefore,

$$\Delta W_{gn}(\Omega) = K(\lambda^3)[2i(\mu_1^{gn} m_1^{gn} + \mu_2^{gn} m_2^{gn})] \qquad [11]$$

The final connection between molecular properties and experimental observables requires knowledge of the orientational distribution of the emitting molecules with respect to the direction and polarization of the excitation light and the direction of detection, and the time dependence of this distribution.

## Circularly polarized luminescence from solutions

There have been very few attempts at measuring CPL from oriented samples due to the inherent problems associated with measurement of circular polarization in the presence of linear polarization. For this reason we restrict the discussion of the orientation dependence of CPL to randomly oriented molecular samples such as occur in liquid solutions. Of course, even if the ground state distribution is randomly oriented (isotropic), the emitting state may not be, due to the photoselection of the emitting sample by the excitation beam. We will denote an orientational distribution by brackets $\langle \ldots \rangle$. Substitution of Equations [5] and [11] into Equation [6], and allowing for an ensemble of orientations we obtain the following

$$\Delta I(\lambda, t) = \left(\frac{2ihc\eta_n}{\lambda}\right) K(\lambda^3)\langle N_e^{\pi}(\Omega : \Omega_0, t) \times [\mu_1^{gn} m_1^{gn} + \mu_2^{gn} m_2^{gn}]\rangle f(\lambda) \qquad [12]$$

We may formally separate the time dependence of the orientational distribution, $\eta(\Omega, t)$, from the number of molecules excited at time $t = 0$ by an excitation pulse as follows

$$N_e^{\pi}(\Omega : \Omega_0, t) = N_e^{\pi}(\Omega_0)\eta(\Omega, t) \qquad [13]$$

$N_e^{\pi}(\Omega_0)$ may be calculated from knowledge of the initial distribution, and direction and polarization of the excitation beam.

$$N_e^{\pi}(\Omega_0) = \kappa^2 \left|\mu^{eg} \cdot \pi\right|^2 \qquad [14]$$

where $\mu^{eg}$ is the absorption transition vector. Final expressions for the formal relationship between the time-dependent circularly polarized luminescence intensity or $g_{lum}$ require knowledge of experimental geometry, and $\eta(\Omega, t)$.

For purposes of illustration, we examine here the most common situation of random (isotropic) ground state distributions, and 90° excitation/emission geometry with unpolarized excitation. If the emission is detected in the laboratory 3 direction, then the excitation is polarized in the 13 plane. Equation [13] may then be written as

$$N_e^{\pi}(\Omega_0) = \kappa^2\left[\left|\mu_1^{eg}\right|^2 + \left|\mu_3^{eg}\right|^2\right] \qquad [15]$$

and the final formal expression (in the laboratory coordinate system) is obtained by substituting this result into Equation [12].

$$\Delta I(\lambda, t) = \left(\frac{2\mathrm{i}hc\eta_n\kappa^2}{\lambda}\right) K(\lambda^3)\Big\langle [\left|\mu_1^{eg}\right|^2 + \left|\mu_3^{eg}\right|^2]_0 [\mu_1^{gn} m_1^{gn} + \mu_2^{gn} m_2^{gn}]_t \, \eta(\Omega, t)\Big\rangle f(\lambda) \qquad [16]$$

where we have explicitly labelled the square brackets to indicate that the absorption takes place at time 0 and the emission at time $t$.

The ensemble average over molecular orientations implied in Equation [16] may be performed if $\eta(\Omega, t)$ is known. We first examine the case in which the emitting sample is 'frozen', i.e. the orientational distribution of emitting molecules is identical to that prepared by the excitation beam. In this case $\eta(\Omega, t) = \exp(-t/\tau)$, where $\tau$ denotes the emission lifetime, and, since the orientations are fixed, the subscripts on the square brackets may be eliminated. The orientational average may be performed by relating the transition moments in the molecular coordinate system to the laboratory system using the appropriate elements of the Euler matrix (see **Figure 1**). If one assumes, for example, that the absorption transition moment defines the molecular $z$-axis, then one must evaluate terms such as the following

$$\left\langle \left|\mu_1^{eg}\right|^2 \mu_1^{gn} m_1^{gn} \right\rangle = \int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi} |R_{1z}|^2 R_{1z} R_{1z} \; \mathrm{d}\theta \mathrm{d}\psi \mathrm{d}\phi \left|\mu_z^{eg}\right|^2 \mu_z^{gn} m_z^{gn} \qquad [17]$$

This photoselection of emitting distributions also affects the measurement of total luminescence, and the development of the formal expression for $I(t)$ parallels that presented above. The final result for $g_{\mathrm{lum}}(t)$ is obtained by evaluating all of the integrals implied in Equation [16], and the related expression for $I(t)$. In this particular simple model, the time-dependence is only in the decay of the emitting state, and is identical for $\Delta I$ and $I$, so that $g_{\mathrm{lum}}$ is time independent. This resulting general expression for the 'frozen' limit is

$$g_{\mathrm{lum}}(\lambda) = 4\mathrm{i}\frac{f_{\mathrm{CPL}}(\lambda)}{f_{\mathrm{TL}}(\lambda)}\left[\frac{7\mu_x^{gn} m_x^{gn} + 7\mu_y^{gn} m_y^{gn} + 6\mu_z^{gn} m_z^{gn}}{7|\mu_x^{gn}|^2 + 7|\mu_y^{gn}|^2 + 6|\mu_z^{gn}|^2}\right] \qquad [18]$$

where we have allowed for different lineshapes for CPL and TL, and an arbitrary absorption transition direction. This result simplifies under conditions in which the absorption and emission transition directions are parallel. Furthermore, the reader is referred to the 'Further reading' section concerning the results expected for other excitation/emission geometries. It should be noted that if the luminescence is partially linearly polarized, the measurement of CPL is problematic due to experimental artifacts. For this reason it is usually the case that the excitation/emission geometry and incident polarization are chosen so as to eliminate the possibility of linear polarization in the luminescence.

The other useful limiting case is the situation in which the orientation of emitting molecules is isotropic (random). This limiting case is appropriate for small molecular systems with relatively long emission lifetimes. If the distribution prepared by the excitation beam has been completely randomized by the time of emission, then Equation [16] reduces to the following

$$\Delta I(\lambda, t) = \left(\frac{2\mathrm{i}hc\eta_n\kappa^2}{\lambda}\right) K(\lambda^3)\langle \mu_1^{gn} m_1^{gn} + \mu_2^{gn} m_2^{gn} \eta(\Omega, t)\rangle f(\lambda) \qquad [19]$$

where we have assimilated the absorption strength and other parameters associated with the excitation into the constants. The orientational averaging in this equation involves the product of only two elements of the Euler matrix. Just as in the previous discussion, the time dependence of the total luminescence and the CPL for this model are identical. The final formal

relationship for this 'isotropic' limit is

$$g_{\text{lum}}(\lambda) = 4\text{i}\,\frac{f_{\text{CPL}}(\lambda)}{f_{\text{TL}}(\lambda)}\left[\frac{\mu^{gn}\cdot m^{gn}}{|\mu^{gn}|^2}\right] \qquad [20]$$

The major difference between this result and the frozen result is the fact that, in Equation [18], the different components of the emission transition vectors contribute unequally to the observable. In principle, this would allow one to investigate the chirality of molecular transitions along specific molecular directions, but to date this has not been exploited.

Equation [20] illustrates one of the operational principles of CPL (and CD) spectroscopy. In most situations one is interested in studying transitions that are formally forbidden. If the transition is allowed, the dipole strength, which is proportional to the denominator of Equation [20], is large. Since the rotatory strength represented by the numerator of Equation [20] is generally quite small, because of the magnitude of the magnetic dipole transition moment, $g_{\text{lum}}$ (or $g_{\text{abs}}$) for allowed transitions is usually very small. In CPL spectroscopy this leads to the practical result that one does not generally study highly luminescent dyes, or other strongly allowed transitions, but rather, the most interesting molecular structural information comes from the study of weakly or mildly luminescent transitions in which the inherent molecular chirality is closely connected with the emissive chromophore.

## Circularly polarized luminescence from racemic mixtures

One of the most useful applications of CPL spectroscopy has been in the study of racemic mixtures. This experiment is possible, because, even though the ground state is racemic, the emitting state can sometimes be photoprepared in an enantiomerically enriched state by use of circularly polarized excitation. If the racemization rate of the excited state is less than the emission rate, then the emission will be circularly polarized. We consider a racemic sample containing equal concentrations of R and S enantiomers. The preferential absorption of circularly polarized light is related to the CD through the absorption dissymmetry ratio $g_{\text{abs}}$ which was defined in Equation [3]. Rewriting this equation for the CD of the R enantiomer we obtain the following

$$g_{\text{abs}}^{\text{R}} = \frac{\Delta\varepsilon}{\varepsilon} = \frac{\varepsilon_{\text{L}}(R) - \varepsilon_{\text{R}}(R)}{\frac{1}{2}(\varepsilon_{\text{L}}(R) + \varepsilon_{\text{R}}(R))} \qquad [21]$$

The number of $R$ enantiomers that absorb left circularly polarized light, $N_{\text{L}}(R)$ is proportional to $\varepsilon_{\text{L}}(R)$, and similarly for $\varepsilon_{\text{R}}(R)$. Thus, we may make the following substitution

$$g_{\text{abs}}^{\text{R}} = \frac{N_{\text{L}}(R) - N_{\text{R}}(R)}{\frac{1}{2}\left(N_{\text{L}}(R) + N_{\text{R}}(R)\right)} \qquad [22]$$

since all of the proportionality constants cancel. The number of R enantiomers that would absorb right circularly polarized light is exactly equal to the number of S enantiomers that would absorb left circularly polarized light, i.e. $N_{\text{R}}(R) = N_{\text{L}}(S)$. This substitution yields the following

$$g_{\text{abs}}^{\text{R}} = \frac{N_{\text{L}}(R) - N_{\text{L}}(S)}{\frac{1}{2}(N_{\text{L}}(R) + N_{\text{L}}(S))} = \frac{2\Delta N_{\text{L}}}{N_{\text{L}}} \qquad [23]$$

$\Delta N_{\text{L}}$ is the differential population $(R - S)$ of excited enantiomers that would result from left circularly polarized excitation.

In general, an exact description of CPL from racemic mixtures requires consideration of the competition between racemization and emission. Under the assumption that racemization is much slower than emission, we may write the following expression for CPL from a racemic mixture under left circularly polarized excitation as follows

$$g_{\text{lum}}^{\text{L}}(\lambda) = 4\text{i}\,\frac{f_{\text{CPL}}(\lambda)\Delta N_{\text{L}}}{f_{\text{TL}}(\lambda)N_{\text{L}}}\left[\frac{\mu^{gn}(R)\cdot m^{gn}(R)}{|\mu^{gn}(R)|^2}\right] \qquad [24]$$

Substituting from above we obtain the following expression

$$g_{\text{lum}}^{\text{L}}(\lambda) = \tfrac{1}{2} g_{\text{abs}}^{\text{R}}(\lambda') g_{\text{lum}}^{\text{R}}(\lambda) \qquad [25]$$

where we have explicitly labelled the excitation wavelength as $\lambda'$. Examination of Equation [25] shows that the measurement of CPL from racemic mixtures depends on the product of the CD and CPL. Since the measurement of $g_{\text{lum}}$ is generally limited to magnitudes greater than $10^{-4}$, the magnitude of the intrinsic dissymmetry ratios must usually be greater than $10^{-2}$ for this unique measurement to be feasible. This restriction has limited applications of this technique to racemic lanthanide complexes, since intraconfigurational $f \leftrightarrow f$ transitions that obey magnetic dipole

transition rules (i.e. $\Delta J = 0, \pm 1$) often are associated with dissymmetry ratios greater than 0.1.

## Time-resolved circularly polarized luminescence

For the simple model enantiopure systems described above, it was concluded that the time dependence of the CPL and total luminescence were identical, and, therefore, the dissymmetry ratio contained no dynamic molecular information. This, of course, would not be the case if intramolecular geometry changes, that would effect the chirality of the molecular transitions, were occurring on the same time scale as emission. However, no such examples of this type of study have yet appeared. Time-resolved CPL measurements have been useful in the study of racemic mixtures of lanthanide complexes in which racemization or excited state quenching is occurring on the same time scale as emission.

It was assumed in the previous section, that the enantioenriched excited state population, which was prepared from a racemic ground state distribution by a circularly polarized excitation beam, was maintained until the time of emission. If this is not the case, then the differential excited state population is time-dependent. Neglecting orientational effects, we may derive the following expression for the time-dependence of the excited state population following an excitation pulse of left circularly polarized light

$$\Delta N^{\mathrm{L}}(t) = \Delta N^{\mathrm{L}}(0)\exp(-2k_{\mathrm{rac}}t - k_0 t) \qquad [26]$$

where $k_{\mathrm{rac}}$ is the racemization rate and $k_0$ is the excited state decay rate. Since the total number of emitting molecules decays as

$$N^{\mathrm{L}}(t) = N^{\mathrm{L}}(0)\exp(-k_0 t) \qquad [27]$$

one can express the time-dependence of the luminescence dissymmetry as follows

$$g^{\mathrm{L}}_{\mathrm{lum}}(t) = g^{\mathrm{L}}_{\mathrm{lum}}(0)\exp(-2k_{\mathrm{rac}}t) \qquad [28]$$

This technique has been used in a number of studies in which the decay of $g_{\mathrm{lum}}$ has been analysed to extract the racemization rate constant.

It is also possible to extract dynamic information concerning racemization rates through measurement of steady-state CPL. Again neglecting time-dependent orientational effects, one may integrate Equations [26] and [27] over long times ($0 \rightarrow \infty$) and obtain the following result

$$g^{\mathrm{L}}_{\mathrm{lum}}(\lambda) = \tfrac{1}{2} g^{\mathrm{R}}_{\mathrm{abs}}(\lambda') g^{\mathrm{R}}_{\mathrm{lum}}(\lambda)\left(\frac{k_0}{2k_{\mathrm{rac}} + k_0}\right) \qquad [29]$$

This equation has the expected limiting behaviour. If $k_{\mathrm{rac}} << k_0$ then the result of Equation [25] is obtained, and if $k_{\mathrm{rac}} >> k_0$ then the magnitude of $g_{\mathrm{lum}}$ becomes negligibly small.

The second application of time-resolved CPL is in the study of enantioselective (or more properly, diastereomer-selective) quenching. In these experiments an enantiopure quencher molecule is added to a racemic solution. The interaction of the chiral quencher with the individual enantiomers is such that the excited state of one of the enantiomers is quenched more rapidly than the other. This process is depicted schematically in **Figure 2**, where we have assumed the quencher is an R enantiomer, and $k_{\mathrm{RS}}$ and $k_{\mathrm{RR}}$ denote the quenching rate constants. This process may be described using the following simple kinetic model

$$R^* + R_{\mathrm{Q}} \underset{k'_{-\mathrm{diff}}}{\overset{k'_{\mathrm{diff}}}{\rightleftharpoons}} [R^* : R_{\mathrm{Q}}] \xrightarrow{k'_{\mathrm{ET}}} [R : R_{\mathrm{Q}}{}^*]$$
$$\xrightarrow{k'_{\mathrm{P}}} R + R_{\mathrm{Q}} \qquad [30]$$

$$S^* + R_{\mathrm{Q}} \underset{k_{-\mathrm{diff}}}{\overset{k_{\mathrm{diff}}}{\rightleftharpoons}} [S^* : R_{\mathrm{Q}}] \xrightarrow{k_{\mathrm{ET}}} [S : R_{\mathrm{Q}}{}^*]$$
$$\xrightarrow{k_{\mathrm{P}}} S + R_{\mathrm{Q}} \qquad [31]$$

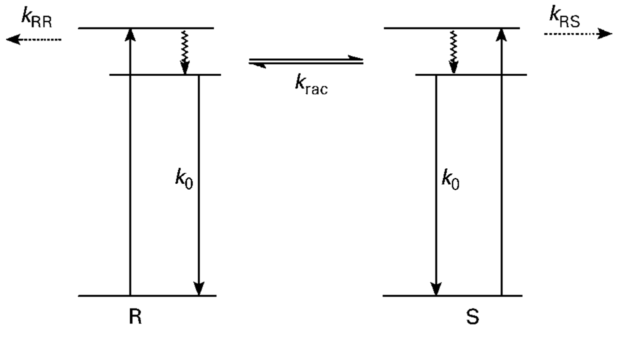



**Figure 2** Schematic energy level diagram for measurement of CPL from a racemic mixture.

These reactions are modelled in terms of a diffusional step ($k^{-}_{\rm diff}$ or $k^{+}_{\rm diff}$), resulting in the formation of a bimolecular encounter complex. This is followed by competing pathways for dissociation of the complex ($k^{-}_{\rm -diff}$ or $k^{-+}_{\rm -diff}$) or energy transfer ($k^{-}_{\rm ET}$ or $k^{+}_{\rm ET}$). Deactivation of the excited quencher is described by $k_{\rm P}$. The observed CPL is directly proportional to the difference in excited state concentrations of the two enantiomers, and this may be related to the specific rate constants introduced above. If, for example, we make the reasonable assumption that the diffusion and deactivation are independent of chirality, one can derive the following expression for $\Delta N(t)$

$$\Delta N(t) = [S^*](t) - [R^*](t) = [S]_0 e^{-k^{\rm RS}_{\rm obs}t} - [R]_0 e^{-k^{\rm RR}_{\rm obs}t} \quad [32]$$

where

$$\begin{aligned} k^{\rm RS}_{\rm obs} &= k_{\rm RS}[S] + k_0 \\ k^{\rm RR}_{\rm obs} &= k_{\rm RR}[R] + k_0 \end{aligned} \quad [33]$$

and

$$\begin{aligned} k_{\rm RR} &= \frac{k^{+}_{\rm ET}k_{\rm diff}}{k^{+}_{\rm ET} + k^{+}_{\rm -diff}} \\ k_{\rm RS} &= \frac{k^{-}_{\rm ET}k_{\rm diff}}{k^{-}_{\rm ET} + k^{-}_{\rm -diff}} \end{aligned} \quad [34]$$

If the quenching reactions are far from diffusion control (i.e. $k_{\rm ET} << k_{\rm -diff}$), then one can express the diastereomeric quenching rate constants in terms of equilibrium constants for formation of the diastereomeric ion pairs.

$$\begin{aligned} k_{\rm RR} &= K^+ k^{+}_{\rm ET} \\ k_{\rm RS} &= K^- k^{-}_{\rm ET} \end{aligned} \quad [35]$$

Thus, the time-dependence of $\Delta I(t)$ or $g_{\rm lum}(t)$ may be used as a probe of the fundamental differences in the interaction of chiral molecules.

## Spectra–structure correlations

The theoretical discussion presented above has focused on developing formal relationships between the molecular transition integrals and the experimental observables. These are, of course, of primary importance in interpreting CPL measurements in terms of inter- and intramolecular processes, including molecular reorientation, racemization and chemical reactions which affect the net molecular chirality. These processes are all well understood, and specific information is most often obtained within an appropriate model system. The more fundamental aspects of molecular chirality associated with the molecular transition matrix elements is less well developed. Although there have been a few successes in interpreting the sign and magnitude of a CPL spectrum in terms of the identity and structure of a specific enantiomer, in this regard, it is certainly not the case that the technique is as well developed as CD spectroscopy. Even in the case of fairly highly symmetric chiral lanthanide complexes, the level of computation and development of f-f intensity theory is not yet at the stage where one can associate *a priori* the sign of an individual transition with the identity of the enantiomer.

## List of symbols

$f$ = lineshape; $g_{\rm lum}$ = luminescence dissymmetry ratio; $I$ = emission intensity; $k$ = rate constant; n→g = transition; $N$ = number of molecules; $t$ = time; $W$ = transition probability; $\Delta I$ = difference in emission intensity; $\varepsilon$ = extinction coefficient; $\eta$ = fraction of molecules; $\lambda$ = wavelength; $\mu$ = absorption transition vector; $\pi$ = polarization of excitation; $\sigma$ = polarization of emission; $\Omega$ = orientation.

*See also:* **Chiroptical Spectroscopy, General Theory; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism; Luminescence, Theory.**

## Further reading

Richardson FS and Riehl JP (1977) Circularly polarized luminescence spectroscopy. *Chemical Reviews* 77: 773–792.

Riehl JP and Richardson FS (1986) Circularly polarized luminescence spectroscopy. *Chemical Reviews* **86**: 1–15.

Riehl JP and Richardson FS (1993) Circularly polarized luminescence. *Methods in Enzymology* **226**: 539–553.

Riehl JP (1994) Excited state optical activity. *Analytical Applications of Circular Dichroism.* Amsterdam: Elsevier.

Steinberg I (1975) Circular polarization of fluorescence. In: Chen RF and Edelhoch H (eds) *Biochemical Fluorescence: Concepts*. New York: Marcel Dekker.

# Chiroptical Spectroscopy, General Theory

**Hans-Georg Kuball**, **Tatiana Höfer** and **Stefan Kiesewalter**, Universität Kaiserslautern, Germany

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| Theory |

## Introduction and molecular description of absorption and circular dichroism

The interaction of light with an oriented molecule or an ensemble of oriented molecules can be treated as a scattering process with photons or semiclassically as an induction process of electric or magnetic moments by an electromagnetic field. As a result of this interaction the incident light beam undergoes changes in its state of polarization and its intensity. Using Gō's description, the complex optical rotation $\hat{\phi}$ can be given as a sum of the contributions of individual molecules:

$$\hat{\phi} = \phi + \mathrm{i}\theta = \frac{Nd}{8\pi^2 V}\int f(\alpha,\beta,\gamma) H'_{21}(\alpha,\beta,\gamma)\sin\beta\,\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma \quad [1]$$

with

$$H'_{21} = -\frac{2\pi\mathrm{i}e^2}{\hbar m^2 c\omega}\sum_n\left\{\frac{\langle p'_2\exp(-\mathrm{i}kx'_3)\rangle_{0n}\langle p'_1\exp(\mathrm{i}kx'_3)\rangle_{n0}}{\omega_{n0}-\omega-\mathrm{i}\eta} + \frac{\langle p'_1\exp(\mathrm{i}kx'_3)\rangle_{0n}\langle p'_2\exp(-\mathrm{i}kx'_3)\rangle_{n0}}{\omega_{n0}+\omega+\mathrm{i}\eta}\right\} \quad [2]$$

and

$$\langle p'_j\exp(\pm\mathrm{i}kx'_3)\rangle_{nm} = \langle n\,|\sum_\mu p'_{j\mu}\exp(\pm\mathrm{i}kx'_{3\mu})|\,m\rangle \qquad j=1,2 \quad [3]$$

Here, $f(\alpha,\beta,\gamma)$ is the orientational distribution function of an ensemble of molecules which is normalized by $1/8\pi^2\int f(\alpha,\beta,\gamma)\sin\beta\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma = 1$, $\alpha,\beta,\gamma$ are the Eulerian angles, $N$ denotes the number of molecules in the volume $V$, $d$ is the optical path length, $\sum_\mu$ represents the sum over all electrons of a molecule, $m$ and $-e$ are the mass and the charge of the electron, $p'_j$ is the operator of the linear momentum, $\omega = 2\pi c\bar{\nu}$, $c$ is the velocity of light in a vacuum, $k = \omega/c$, $\eta$ is a positive damping parameter representing the natural line shape, $\omega_{n0} = 2\pi c\bar{\nu}_{n0} = \hbar^{-1}(E_n - E_0)$, and $E_n$ and $E_0$ are the energies of the states $|n\rangle$ and $|0\rangle$. Primed symbols refer to the space-fixed and unprimed letters to the molecule-fixed coordinate system. $\phi$ is the optical rotation and $\theta$ represents the ellipicity of the light beam emerging from the sample. $H'_{21}$ is an element of the interaction matrix $H'_{ij}$ of a photon and an atom or an oriented molecule. The incident light is propagating in the $x'_3$ direction and is polarized in the $x'_1\,x'_2$ plane. Using the relation

$$\langle p'_2\exp(-\mathrm{i}kx'_3)\rangle_{0n}\langle p'_1\exp(\mathrm{i}kx'_3)\rangle_{n0} = \langle p'_1\exp(\mathrm{i}kx'_3)\rangle_{0n}\langle p'_2\exp(-\mathrm{i}kx'_3)\rangle_{n0} \quad [4]$$

where the wave functions $|j\rangle$ are chosen to be real. Expanding the exponentials of Equation [2] into a series which converges very rapidly if the extension of the molecule is small compared to the wavelength of the incident light, it follows that the complex optical rotation is given by

$$\hat{\phi} = -\frac{\mathrm{i}Nd}{2\pi\hbar cV}\int f(\alpha,\beta,\gamma)\sin\beta\,\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma \times\left\{\sum_n\frac{\omega}{\omega_{n0}^2-\omega^2-2\mathrm{i}\eta\omega}\left[\omega_{n0}D'^{0n}_{12}(\alpha,\beta,\gamma) + 2\mathrm{i}\omega R'^{0n}_{123}(\alpha,\beta,\gamma)\right]\right\} \quad [5]$$

where $D'^{0n}_{12}$ is a coordinate of the transition moment tensor. Loxsom has discussed the effect of an increasing extension-to-wavelength ratio. The third rank tensor $R'^{0n}_{123}$, antisymmetric with respect to the first two indices, can be replaced by a symmetric pseudotensor $R'^{0n}_{ij}$ of rank two which is called the tensor of rotation. This tensor of rotation can be decomposed into a magnetic dipole and an electric quadrupole contribution

$$R'^{0n}_{ij} = \frac{1}{4}\sum_{r,s}\langle\mu_r\rangle_{0n}(\varepsilon_{rsi}\langle C_{sj}\rangle_{n0} + \varepsilon_{rsj}\langle C_{si}\rangle_{n0}) \quad [6]$$

with

$$\langle C_{sj}\rangle_{n0} = -\mathrm{i}\sum_r \varepsilon_{sjr}\langle m_r\rangle_{n0} - \frac{\omega_{n0}}{c}\langle Q_{sj}\rangle_{n0} \quad [7]$$

and

$$\mu_i = -e\sum_\nu x_{i\nu} \quad Q_{ij} = -\frac{1}{2}e\sum_\nu x_{i\nu}x_{j\nu}$$
$$m_r = -\frac{e}{2mc}\sum_\nu\sum_{i,j}\varepsilon_{rij}\,x_{i\nu}\,p_{j\nu} \quad [8]$$

Here, $\langle\mu_i\rangle_{0n}$ and $\langle m_i\rangle_{n0}$ represent the coordinates of the electric and magnetic dipole transition moment vectors, respectively, $\langle Q_{ij}\rangle_{n0}$ is a coordinate of the electric quadrupole transition moment tensor, and $\varepsilon_{ijk}$ is the Levi–Civita tensor. Operators depending on the dynamic variables of the nuclei are omitted because they do not contribute to the effect in the approximation used.

The complex optical rotation $\hat{\phi}$ contains the elliptical birefringence and dichroism which can be given for small effects as a superposition of a linear and a circular birefringence and dichroism:

$$\hat{\phi} = \phi_{\mathrm{LD}} + \phi_{\mathrm{CB}} + \mathrm{i}[\theta_{\mathrm{LB}} + \theta_{\mathrm{CD}}] \quad [9]$$

with the linear birefringence (LB) and dichroism (LD) effects

$$\theta_{\mathrm{LB}} = -\pi\bar{\nu}d(n_1 - n_2) \qquad \phi_{\mathrm{LD}} = \frac{1}{4}(\ln 10)(\varepsilon_1 - \varepsilon_2)cd \quad [10]$$

and the circular birefringence (CB) and dichroism (CD) effects

$$\phi_{\mathrm{CB}} = \pi\bar{\nu}d(n_{\mathrm{L}} - n_{\mathrm{R}}), \qquad \theta_{\mathrm{CD}} = \frac{1}{4}(\ln 10)(\varepsilon_{\mathrm{L}} - \varepsilon_{\mathrm{R}})cd \quad [11]$$

Here, $\varepsilon_1$, $\varepsilon_2$ and $n_1$, $n_2$ are the molar decadic absorption coefficients and the refractive indices for a linearly polarized light beam with a polarization parallel and perpendicular to the optical axis of a uniaxial sample, and $\varepsilon_{\mathrm{L}}$, $\varepsilon_{\mathrm{R}}$ and $n_{\mathrm{L}}$, $n_{\mathrm{R}}$ are the absorption coefficients and the refractive indices for the left and right circularly polarized light beams, respectively. In the approximation used the CD/CB and LD/LB effects are additive. From the theory of the Cauchy–Rieman differential equations it can be shown that the Kramers–Kronig transforms derived with microcausality as the only presupposition are also valid for a system of oriented molecules:

$$\theta_{\mathrm{LB}}(\omega) = -\frac{2}{\pi}\oint\frac{\bar{\omega}\phi_{\mathrm{LD}}(\bar{\omega})}{\bar{\omega}^2 - \omega^2}\mathrm{d}\bar{\omega} \quad [12]$$

$$\phi_{\mathrm{CB}}(\omega) = \frac{2}{\pi}\oint\frac{\omega\theta_{\mathrm{CD}}(\bar{\omega})}{\bar{\omega}^2 - \omega^2}\mathrm{d}\bar{\omega} \quad [13]$$

The frequency dependence given in Equations [2] and [5] is based on a situation where a natural line width is found. In all cases where optical activity has been analysed, e.g. in a gas or in a liquid or solid phase, the bandwidth is determined by the broadening of rotational states – for which the theory of optical activity is not well developed and no experimental analysis exists – or by the existence of librational states etc. Therefore, broadened rotational states or librational states can be introduced for which the spectrum can be considered as a quasi-continuum. With the assumption of the Born–Oppenheimer approximation the states $|0\rangle$ and $|n\rangle$ can then be given as a product of vibronic states and librational states $\lambda_{Nn}$ and $\lambda_{Kk}$ as follows

$$|0\rangle = |Nn\rangle|\,\lambda_{Nn}(\omega)\rangle \text{ and } |n\rangle = |Kk\rangle|\lambda_{Kk}(\omega)\rangle \quad [14]$$

where $|Nn\rangle$ and $|Kk\rangle$ represent the vibronic electronic ground and excited states, respectively. $n$ and $k$ are quantum numbers for the vibrational states in the electronic states $N$ and $K$, respectively. An integration over the quasi-continuum of the librational states $\omega$ and the neglect of the natural bandwidth $\eta$ yields the spectral functions $F^{NnKk}(\bar{\nu})$ and $G^{NnKk}(\bar{\nu})$ for the absorption and the circular dichroism effect which are equal for most of the experimental conditions given. With the new line shape of the absorption band $F^{NnKk}(\bar{\nu})$ and the circular dichroism $G^{NnKk}(\bar{\nu})$, the coordinates of the absorption $\varepsilon_{ij}$ (Eqn [15]) and circular dichroism tensor $\Delta\varepsilon_{ij}$ (Eqn [16]) are given by

$$\varepsilon_{ij}(\bar{\nu}) = \frac{B\bar{\nu}}{4}\sum_n\sum_{Kk} D_{ij}^{NnKk}F^{NnKk}(\bar{\nu}) \quad [15]$$

and

$$\Delta\varepsilon_{ij}(\bar{\nu}) = B\bar{\nu}\sum_n\sum_{Kk} R_{ij}^{NnKk}G^{NnKk}(\bar{\nu}) \quad [16]$$

where for the vibronic transition $|Nn\rangle \to |Kk\rangle$ the dipole transition moment tensor is defined by

$$D_{ij}^{NnKk} = \langle \mu_i \rangle_{NnKk} \langle \mu_j \rangle_{KkNn} \qquad [17]$$

and the rotational strength tensor $R_{ij}^{NnKk}$ is given analogously to the rotational strength tensor in Equations [6] and [7] by replacing $|0\rangle \to |n\rangle$ with $|Nn\rangle \to |Kk\rangle$ and

$$B = \frac{32\pi^3 N_A}{10^3 hc \ln 10} = 7.653 \times 10^{40} \text{ cgs} \qquad [18]$$

$N_A$ is Avogadro's number, $c$ the velocity of light and $h$ is Planck's constant. Together with the Kramers–Kronig transforms (Eqns [12] and [13]) the frequency dependence of the optical rotation tensor can be obtained as follows

$$M_{ii}(\bar{\nu}) = \frac{288\pi N_A \bar{\nu}}{hc} \sum_n \sum_{Kk} R_{ii}^{NnKk} J_2^{NnKk}(\bar{\nu}) \qquad [19]$$

$$J_2^{NnKk}(\bar{\nu}) = \oint \frac{\bar{\nu}' G^{NnKk}(\bar{\nu}) \mathrm{d}\bar{\nu}'}{\bar{\nu}'^2 - \bar{\nu}^2} \qquad [20]$$

$J_2^{NnKk}(\bar{\nu})$ in Equation [19] is the principal value of the integral over the singularity in the wavenumber region of an absorption process.

## Absorption and circular dichroism in isotropic media

For isotropic solutions the orientational distribution function $f(\alpha, \beta, \gamma)$ in Equation [5] is equal to one and the absorption coefficient $\varepsilon(\bar{\nu})$, the circular dichroism $\Delta\varepsilon(\bar{\nu})$, and the molar optical rotation $[M(\bar{\nu})]$ (optical rotatory dispersion – ORD) are one third of the trace of the corresponding tensor or pseudotensor of the second rank, respectively,

$$\varepsilon(\bar{\nu}) = \frac{1}{3} \sum_i \varepsilon_{ii} = \frac{B\bar{\nu}}{12} \sum_n \sum_{Kk} D^{NnKk} F^{NnKk}(\bar{\nu}) = \sum_K \varepsilon^{NK}(\bar{\nu}) \qquad [21]$$

with the dipole strength of the transition $|Nn\rangle \to |Kk\rangle$

$$D^{NnKk} = \sum_i D_{ii}^{NnKk} = \sum_i \langle \mu_i \rangle_{NnKk} \langle \mu_i \rangle_{NnKk} = \langle \vec{\mu} \rangle_{NnKk}^2 \qquad [22]$$

and

$$\Delta\varepsilon(\bar{\nu}) = \frac{1}{3} \sum_i \Delta\varepsilon_{ii} = \frac{B\bar{\nu}}{3} \sum_n \sum_{Kk} R^{NnKk} F^{NnKk}(\bar{\nu}) = \sum_K \Delta\varepsilon^{NK}(\bar{\nu}) \qquad [23]$$

with the rotational strength $R^{NnKk}$ of the transition $|Nn\rangle \to |Kk\rangle$ derived from the rotational strength tensor

$$R_{ij}^{NnKk} = -\frac{1}{2} \mathrm{i} \sum_{r,l} \langle \mu_r \rangle_{NnKk} \cdot \langle m_l \rangle_{KkNn} \times \left[ \delta_{rl}\delta_{ij} - \frac{1}{2}\delta_{ri} \cdot \delta_{lj} - \frac{1}{2}\delta_{rj} \cdot \delta_{li} \right] - \frac{1}{4} \frac{\omega_{KkNn}}{c} \sum_{r,s} \langle \mu_r \rangle \left[ \varepsilon_{rsi} \langle Q_{sj} \rangle_{KkNn} + \varepsilon_{rsj} \langle Q_{si} \rangle_{KkNn} \right] \qquad [24]$$

$\delta_{ij}$ is the Kronecker symbol. The contribution of the electric quadrupole transition moments cancels in the isotropic solution. Then a scalar product of the electric and magnetic dipole transition moments follows for the rotational strength

$$R^{NnKk} = \sum_i R_{ii}^{NnKk} = \sum_i \mathrm{Im}\{\langle \mu_i \rangle_{NnKk} \langle m_i \rangle_{KkNn}\} = \mathrm{Im}\{\langle \vec{\mu} \rangle_{NnKk} \cdot \langle \vec{m} \rangle_{KkNn}\} \qquad [25]$$

The ORD is now given by

$$[M(\bar{\nu})] = \frac{1}{3} \sum_i M_{ii} = \frac{288\pi N_A \bar{\nu}}{3hc} \sum_n \sum_{Kk} R^{NnKk} J_2^{NnKk}(\bar{\nu}) = \sum_K [M^{NK}(\bar{\nu})] \qquad [26]$$

If vibronic states $n$, $k$ do not contribute to the ORD, there follows from Equation [26] with the dispersion

$$J_2^{NK}(\bar{\nu}) = \frac{\bar{\nu}^2}{\bar{\nu}_{NK}^2 - \bar{\nu}^2} \qquad [27]$$

the well-known Rosenfeld equation.

The right sides of Equations [23] and [26] lead to an interpretation of the CD and ORD spectra as a sum of contributions of the electronic transitions. To

every absorption band $\varepsilon^{NK}(\bar{\nu})$ belongs a $\Delta\varepsilon^{NK}(\bar{\nu})$ and a corresponding ORD curve $[M^{NK}(\bar{\nu})]$ as presented in **Figure 1**. It is evident from **Figure 1** that the optical rotation $[M^{NK}(\bar{\nu})]$ possesses a larger half bandwidth than $\Delta\varepsilon^{NK}(\bar{\nu})$. Thus $[M(\bar{\nu})]$ can only be measured as a sum of contributions of many transitions independent of the chosen wavelength even in the wavelength region far away from absorption bands. In contrast to this it is often possible to measure the CD curve of only one or an overlap of CD curves of a few transitions. The smaller bandwidth of the CD bands is the basis for the answer to the often-discussed question of why a CD measurement has advantages over an ORD measurement. In principle, one can measure either the circular dichroism or the optical rotatory dispersion because one measurable quantity can be evaluted from the other with the Kramers–Kronig relations (Eqns [12] and [13]) if the frequency region measured is sufficiently large. From the numerical point of view the calculation of an ORD curve from a CD curve is easier than vice versa because of the smaller half bandwidth of the CD curves.

The different absorption bands, to which the CD and ORD belong, can be classified by the symmetry of the molecular states involved especially by their polarization, i.e. their transition moment direction which is defined by Equations [6], [8], and [17] and visualized in **Figure 2**. Experimentally the polarization can be determined by the linear dichroism (Eqn [10]) measurement usually described in polarized spectroscopy in terms of the degree of anisotropy $R$:

$$R = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + 2\varepsilon_2} = \frac{\varepsilon_1 - \varepsilon_2}{3\varepsilon} \qquad [28]$$

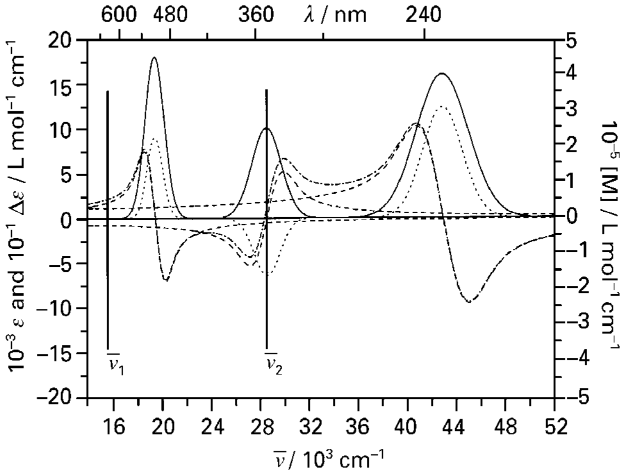



**Figure 1** UV ($\varepsilon(\bar{\nu})$, ——), CD ($\Delta\varepsilon(\bar{\nu})$, ·····) and ORD ($[M(\bar{\nu})]$, ---·) spectra of a compound as a sum of contributions $\varepsilon^{NK}(\bar{\nu})$(——), $\Delta\varepsilon^{NK}(\bar{\nu})$ (······), and $[M^{NK}(\bar{\varepsilon})]$(----) of different transitions (chromophores) $|N\rangle \rightarrow |K\rangle$.

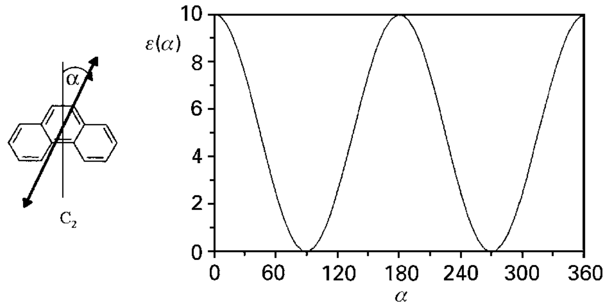



**Figure 2** Dependence of the polarized absorption on the angle α between the transition moment direction along the $C_2$ axis of phenanthrene and the plane of polarization of the light. The transition moment direction is defined by a vector $\langle\vec{\mu}\rangle_{n0}$ parallel to the direction of maximum absorption. The light beam propagates perpendicular to the plane of the molecule.

Furthermore, the dissymmetry factor

$$g(\bar{\nu}) = \frac{\Delta\varepsilon(\bar{\nu})}{\varepsilon(\bar{\nu})} \qquad [29]$$

contains information about the kind of the Cotton effect (CE) of a compound and is a hint whether an effect is measurable or not. Nowadays, the lowest limit for the sensitivity with commercial instruments is about $10^{-7}$.

From the Kramers–Kronig transformation and from the quantum-theoretical description, sum rules for the rotational strength analogous to the Kuhn–Reichert sum rule for the absorption can be derived. This means that, for the sum over the rotational strength of all transitions of the CD spectrum of a molecule, one obtains

$$\sum_{N,n,K,k} R^{NnKk} = 0 \qquad [30]$$

From this result it follows that the CD and ORD values decrease when going to wavelengths longer than those of the microwave spectral region or to wavelengths shorter than those of the vaccum UV. This can be understood from the fact that the ratio of the extension of the molecules and the wavelengths is either very much smaller or larger than for the UV-visible spectral region.

## Vibrational progressions in the CD and UV-visible spectra of an electronic transition

From Equations [21] and [23] it also follows that the UV ($\varepsilon^{NK}(\bar{\nu})$) and the CD ($\Delta\varepsilon^{NK}(\bar{\nu})$) band are built up

by a sum of different vibrational progressions

$$\varepsilon^{NK}(\bar{\nu}) = \frac{B\bar{\nu}}{12}\sum_{n}\sum_{k} D^{NnKk} F^{NnKk}(\bar{\nu}) \qquad [31]$$

and

$$\Delta\varepsilon^{NK}(\bar{\nu}) = \frac{B\bar{\nu}}{3}\sum_{n}\sum_{k} R^{NnKk} F^{NnKk}(\bar{\nu}) \qquad [32]$$

In **Figure 3** a term scheme and the corresponding CD and UV spectra for an example where two progressions contribute to the UV-visible and the CD spectrum are constructed. The vibrational progressions can contribute according to Equations [31] and [32] to the UV and CD spectra of an electronic transition with vibrations of different symmetry and thus with differently polarized bands and CD contributions of different sign. This is analogous to the electronic spectra of a molecule where transitions of different polarization, i.e. different transition moment directions, contribute to the spectra (Eqn [21]). This can be proven experimentally by a frequency-dependent degree of anisotropy (Eqn [28]). In order to eliminate the effect of the vibrational fine structure one has to integrate over the whole band of the electronic transition $|N\rangle \rightarrow |K\rangle$. The dipole strength $D^{NK}$ and the rotational strength $R^{NK}$ include information about the electronic transition from the electronic ground state $N$ to the electronic excited state $K$ and are of interest for the characterization of the transition:

$$D^{NK} = \frac{12}{B}\int \frac{\varepsilon^{NK}(\bar{\nu})}{\bar{\nu}}\mathrm{d}\bar{\nu} \qquad [33]$$

$$R^{NK} = \frac{3}{B}\int \frac{\Delta\varepsilon^{NK}(\bar{\nu})}{\bar{\nu}}\mathrm{d}\bar{\nu} \qquad [34]$$

The dimension is given, as still used even today, in cgs units. A frequency-independent dissymmetry factor $g$ can also be defined with the help of the dipole strength and rotational strength of the electronic transition

$$g = \frac{4R^{NK}}{D^{NK}} \qquad [35]$$

For an allowed electric and allowed magnetic transition $|N\rangle \rightarrow |K\rangle$ the Franck–Condon factor $\langle n|k\rangle$ can be introduced and the rotational strength (Eqn [34]) is then given by

$$\begin{aligned} R^{NK} &= \sum_{k} R^{NnKk} = \mathrm{Im}\{\langle\vec{\mu}\rangle_{NK}\cdot\langle\vec{m}\rangle_{KN}\}\sum_{k} |\langle n \mid k\rangle|^2 \\ &= \mathrm{Im}\{\langle\vec{\mu}\rangle_{NK}\cdot\langle\vec{m}\rangle_{KN}\} \qquad [36] \end{aligned}$$

$\langle\vec{\mu}\rangle_{NK}$ and $\langle\vec{m}\rangle_{KN}$ are the electric and magnetic transition moments for the 0-0 transition, respectively (a 0-0 transition means a transition where neither excited vibrational nor excited rotational states of the ground and excited electronic states are involved). Analogously, the dipole strength $D^{NK}$ can be evaluated by

$$\begin{aligned} D^{NK} &= \sum_{k} D^{NnKk} = \langle\vec{\mu}\rangle_{NK}\cdot\langle\vec{\mu}\rangle_{KN}\sum_{k} |\langle n \mid k\rangle|^2 \\ &= \langle\vec{\mu}\rangle_{NK}\cdot\langle\vec{\mu}\rangle_{KN} \qquad [37] \end{aligned}$$

For forbidden electronic transitions, intensity is stolen from other allowed electronic transitions and thus further terms have to be added in Equations [36] and [37] (Herzberg–Teller theory of vibronic coupling).

## Classification of absorption and circular dichroism spectra and of chromophores

### Classifications of absorption and CD bands

Electronic transitions $|N\rangle \rightarrow |K\rangle$ can be classified by the irreducible representations of the wave functions involved or more conveniently for a chemist by the orbitals which are involved, i.e. the non-bonding n, bonding $\sigma$, $\pi$, and antibonding n*, $\sigma^*$, and $\pi^*$ orbitals, e.g. n $\rightarrow \pi^*$, $\pi \rightarrow \pi^*$, and $\sigma \rightarrow \sigma^*$, etc. Additionally it is very convenient to classify the transitions as electrically or magnetically allowed or forbidden, i.e., $|\langle\vec{\mu}\rangle_{NK}|$ and $|\langle\vec{m}\rangle_{KN}|$ is different from or equal to zero.

More important for CD spectroscopy is the classification of chromophores with respect to symmetry operations of the second kind (i, $\sigma$, and $S_n$, with $n$>2). An inherently dissymmetric chromophore possesses no symmetry element of the second kind and thus is chiral by itself. In an inherently symmetric chromophore a symmetry element of the second kind exists and thus it is achiral by itself. In **Figure 4** the enone chromophore is shown as inherently symmetric and inherently dissymmetric. The local symmetry of a

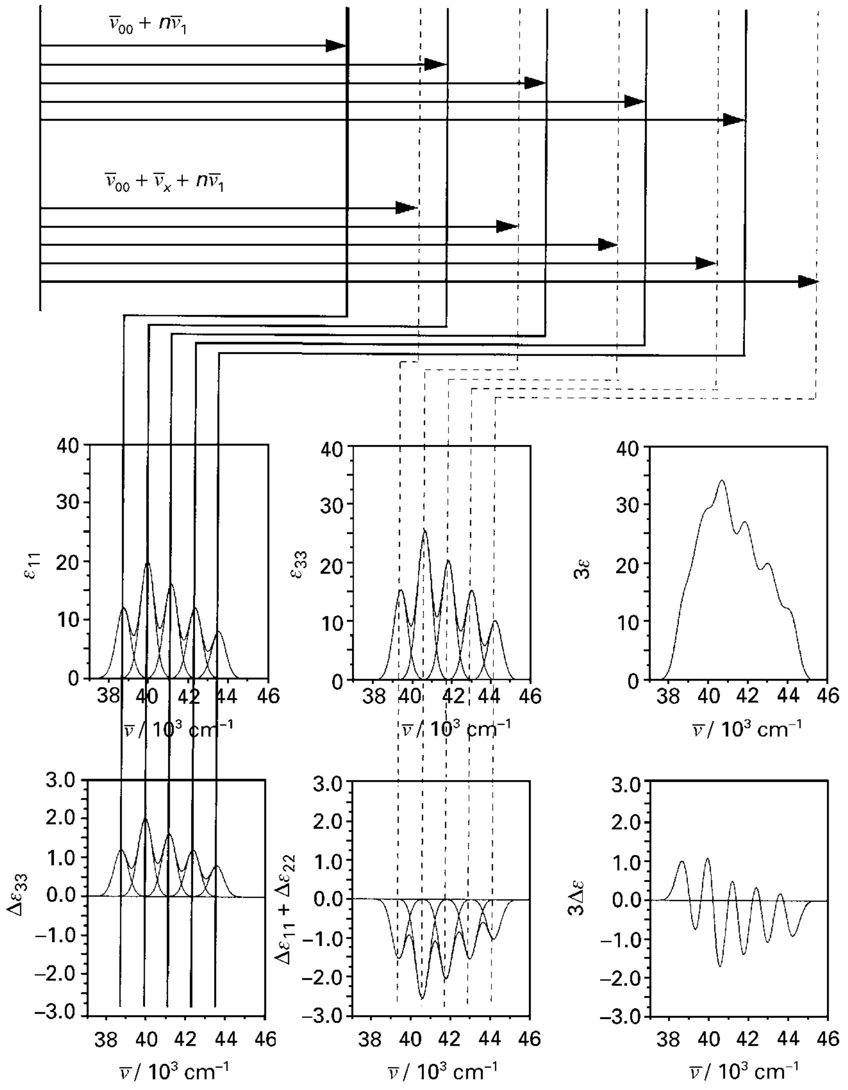



**Figure 3** Term scheme for an absorption and the corresponding CD band with contributions of one allowed ($\bar{\nu} = \bar{\nu}_{00} + n\bar{\nu}_1$) and one forbidden ($\bar{\nu} = \bar{\nu}_{00} + \bar{\nu}_x + n\bar{\nu}_1$) vibrational progression. $\bar{\nu}_{00}$ is the 0–0 transition band. $\bar{\nu}_1$ and $\bar{\nu}_x$ are a totally symmetric and a non-totally symmetric vibration, respectively. $n$ = 1,2, ....

chromophore cannot be sharply defined because the chromophore itself is only a qualitative quantity. For CD bands a characterization by the size of the dissymmetry factor $g$ is possible. For an inherently dissymmetric chromophore the dissymmetry factor $g$ is about or larger than $10^{-2}$, for an inherently symmetric chromophore $g < 10^{-2}$ (mostly $\approx 10^{-4}$) if the transition is electrically allowed and magnetically forbidden. $g > 10^{-2}$ (mostly $5 \times 10^{-3}$ to $10^{-1}$) if the transition is magnetically allowed and electrically forbidden, also for an inherently symmetric chromophore.

### What is a chromophore?

A chromophore is that part of the molecule where the absorption proceeds and where the main change of the geometry or electron density, etc. appears after the excitation process. The area of the chromophore is not very well demarcated from the residual parts of the molecule which are not involved in the absorption process. The shading in **Figure 5** demonstrates the continuous variation from the centre of the excitation to the undisturbed residue of the molecule. A carbonyl chromophore with an n$\pi^*$-transition and an olefin chromophore with a $\pi\pi^*$-transition are regarded as two distinct chromophores (**Figure 5**). If the two chromophores are located in a molecule next to each other, an influence on the position of the band and on the intensity of the absorption results. If the interaction is strong enough, the two chromophores have to be considered as one new chromophore.

The number of chromophores in a molecule is very large because of the large number of transitions

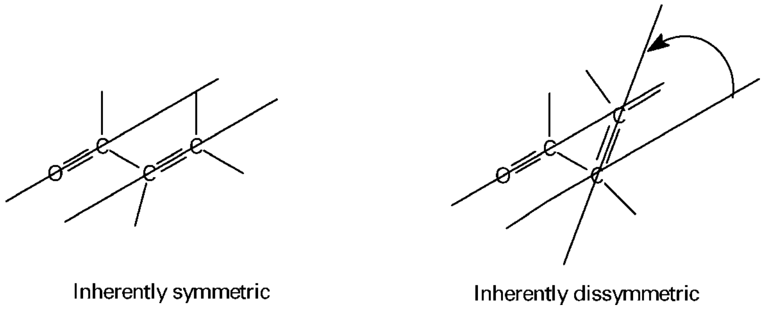



**Figure 4** The inherently symmetric (left) and inherently dissymmetric (right) enone chromophore.

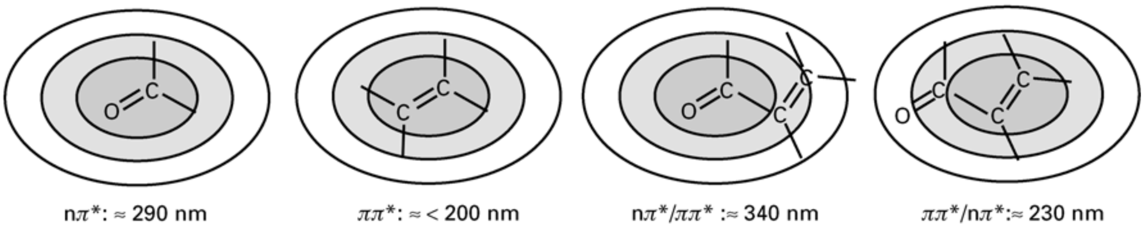



**Figure 5** The enone chromophore as an example for the interaction of two chromophores, i.e. the 'en' and 'one' chromophore.

localized in different parts of the molecule. Also in a group, like the carbonyl group, there is an infinite number of chromophores. One important chromophore for CD spectroscopy besides the n$\pi^*$-transition is a $\pi\pi^*$-transition, located in the carbonyl area and at wavelength lower than 150 nm.

According to Snatzke a molecule can be divided into several spheres (**Figure 6**). The first sphere is the chromophore itself which can be either chiral or achiral. The second sphere is the neighbourhood to the chromophore, the third sphere the area following the second sphere, and so on. The criteria by which the spheres are differentiated are the distance and the interaction between the parts of the molecules like conjugation etc. Usually changes in the second sphere can affect the sign and size of the CE whereas substitutions in spheres farther away lead to smaller effects.

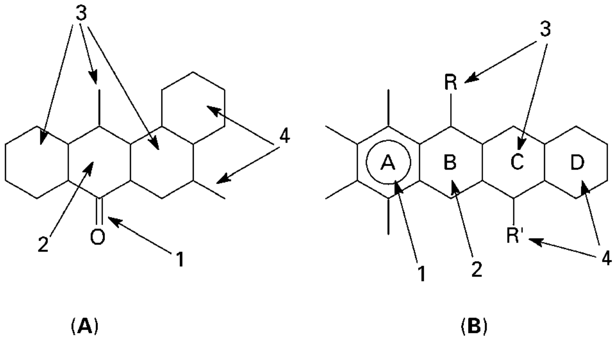



**Figure 6** Two examples for a decomposition of a molecule into spheres according to Snatzke.

## The quantitative description of the rotational strength

Methods for the description of the optical activity are either based on an estimation of the sum over the rotational strengths of all transitions weighted by a dispersion term as in the case of frequency-dependent optical rotation (ORD; denoted as $\phi$ or $[M]$) or of the rotational strength of one transition in the case of circular dichroism, denoted as $\Delta\varepsilon$ or $\theta$ (Eqn [32]). Numerical quantum-mechanical calculations of the rotational strength $R^{NK}$ for a few transitions $|N\rangle \rightarrow |K\rangle$ of a molecule lead nowadays to results of increasing reliability and thus a determination of the absolute configuration becomes possible with a CD measurement over the spectral region to which the analysed transitions belong. But this is only a development in the right direction, and further advances in computational abilities and resources are necessary. The possibility for numerical calculation is not always available in the chemist's daily work today. From this point of view there is no way to renounce the qualitative concepts and semi-qualitative methods used up to now to correlate sign and absolute configuration. Therefore, the following techniques will be discussed in more detail:

1. the quantum-mechanical calculation of the rotational strength,
2. the polarizability theory,
3. the one-electron mechanism,

4. the model of independent groups (MIG) of Tinoco,
5. the exciton model (exciton chirality method).

### Quantum mechanical calculation of the rotational strength

In order to calculate the rotational strength, reasonably good wavefunctions for all electronic states should be available. Calculation with a number of semiempirical and *ab initio* techniques have been performed with more or less success. A review of used techniques has been given by Woody in Nakanishi *et al.* As discussed there, for the calculation of the rotational strength, the electric transition moments can be either evaluated with the dipole-velocity method ($\langle N\,|p_i|\,K\rangle$, with the linear momentum operator $p_i$) which leads to an origin-independent rotational strength $R^{NK}$ but overestimates errors from inaccurate wavefunctions far away from the atomic centre, or with the dipole-length method ($\langle N\,|\mu_i|\,K\rangle$) which leads to an origin-dependent rotational strength $R^{NK}$ and is not so sensitive to other errors of the wavefunctions. Up to now the dipole-velocity method has been the method of choice for numerical quantum-mechanical calculations. More recent analyses and calculations by Grimme do not support this choice, in contrast to the leading opinion. It seems to us that one may choose the dipole-velocity method or the dipole-length method depending on the kind of molecule and the quality of the calculated wavefunctions. Furthermore, the extent of taking into account singly and doubly excited configurations is a problem when trying to achieve good results for the direction of the electric and magnetic dipole transition moments and thus for the rotational strength. A new development can also be found in the use of the density functional theory for the calculation of $R^{NK}$ by Grimme.

### The polarizability theory

Kirkwood expressed the rotatory power in terms of the polarizabilities and anisotropies of the polarizabilities of groups in the molecule. First-order perturbation theory leads in this case to a perturbation potential expressing the interaction of two transition moment dipoles $i, j$ at a distance $R_{ij}$, where this distance is large compared to the separation of charge in the dipoles. Such terms are important if the molecule contains two or more strong absorption bands near the frequency region where the rotatory dispersion is calculated originating from allowed electric dipole transitions. But even the strongest absorption bands may contribute little to the optical activity by this mechanism if the chromophoric groups are unfavourably disposed relative to each other. In the last few years there has been an improvement in the calculation of the optical rotation with Kirkwood's polarizability theory by using chirality functions as a possibility to find symmetry-adapted functions as shown by Haase and Ruch for asymmetric methane or allene derivatives. The thus obtained electric dipole transition moments are grouped together to polarizabilities or to polarizability tensor coordinates, respectively. The basis for the description is a second-order perturbation theory of the optical rotation.

### The one-electron mechanism

Historically, the one-electron mechanism was introduced by Condon and co-workers as an optical activity which originates from the excitation of a single electron in a field of suitable dissymmetry. In the MO concept, the one-electron mechanism is an excitation of one electron to a singly-excited electronic configuration of an inherently symmetric chromophore perturbed by chiral surroundings. Depending on the symmetry of the chromophore the transition is either electrically and/or magnetically dipole allowed and without the perturbation fulfills the condition that the dot product of the electric and magnetic dipole transition moments of this transition is zero. By the perturbation with groups/atoms in chiral surroundings a magnetically and/or electrically allowed dipole is induced by borrowing intensity from other transitions of the same chromophore. Often the language 'intensity is stolen from other transitions' has been used in the literature, e.g. 'contributions to the rotational strength of the $n \to \pi^*$ transitions are stolen by this mechanism from electric-dipole-allowed $\sigma \to \sigma^*$ or $\pi \to \pi^*$ transitions'. There are two possibilities for a perturbation: On the one hand electronic interactions are responsible for the intensity borrowing which is often called the dynamic coupling model. On the other hand vibrational processes change the geometry and, by taking into account the dependence of $\langle \vec{\mu} \rangle_{0n}$ and/or $\langle \vec{m} \rangle_{n0}$ on nuclear coordinates, intensity is borrowed by the transition $N \to K$ via nuclear vibrations from other allowed transitions of the chromophore. Then different vibrational progressions contribute to the UV and CD spectra as shown in **Figure 3**. This mechanism is often called the static coupling or vibronic coupling model. Depending on the type of transitions, a nomenclature system has been introduced by Moscowitz and co-workers and Weigang where the CD of a transition (a) without intensity borrowing is called a case I CD, (b) with intensity borrowing via one forbidden electronic dipole transition is called a case II CD, and (c) with intensity borrowing via

electrically and magnetically forbidden transitions is called a case III CD. Whereas in the first case only progressions with totally symmetric vibrations contribute to the CD, in cases II and III progressions with non-totally symmetric vibrations and vibrational CD bands of the same or different sign can also contribute to an electronic CD band as shown by Weigang and in Snatzke by Moscowitz.

### The model of independent groups for the calculation of the rotational strength of large molecules or polymers

Tinoco's aim with the model of independent groups (MIG method) was to calculate the total rotational strength $R^{NK}$ for a transition |*N*Ò → |*K*Ò for a large molecule or a polymer possessing a large number of equal and different groups, i.e. chromophores for which the circular dichroism is given by

$$\Delta\varepsilon(\bar{\nu}) = \frac{B\bar{\nu}}{3}\sum_K R^{NK}\sum_n\sum_k |\langle n \mid k\rangle|^2 F^{NnKk}(\bar{\nu}) = \sum_K R^{NK}F^{NK}(\bar{\nu}) \qquad [38]$$

The excited state may or may not be degenerate or accidentally degenerate. In the case of an *n*-fold degeneracy of the state |*A*Ò, the excited states |*L*Ò, linear combinations of the degenerate states |*A*Ò, contribute to the rotational strength (Eqn [39]):

$$R^{0A} = \sum_{L=1}^{n} R^{NL} = \sum_{L=1}^{n} \mathrm{Im}\{\langle\vec{\mu}\rangle_{NL}\cdot\langle\vec{m}\rangle_{LN}\} \qquad [39]$$

For the evaluation of the rotational strength the molecule is decomposed into different groups *i* which are at a distance $|\vec{R}_{ij}| = |\vec{R}_i - \vec{R}_j|$ and oriented differently with respect to each other. The operator of the interaction between two groups *i*, *j* is given by the potential $\hat{V}_{ij}$. Depending on the assumption for the interaction between the groups, different approximations for the wavefunctions of the groups are necessary. In the simplest case of completely isolated groups the wave function of the molecule can then be expressed as a product of the wavefunctions of the groups. It is assumed that the group electronic wavefunctions do not overlap, which means that there is no electron exchange between groups. Only the interaction between two groups at a time is considered. Furthermore, it is assumed that singly and doubly excited states on different groups are possible.

To calculate the total rotational strength for a transition it is necessary to describe the electric and magnetic dipole transition moment operator in an adequate way. Whereas the electric moment operator is the sum of the moments of the different groups, for the magnetic moment operator two contributions have to be taken into account. Analogously to the electric moments, there is a contribution from the magnetic moment of every group. The second contribution comes from the interaction of groups. In a simplified picture a 'moving electron' in one group leads to a magnetic moment in the other group and gives rise to a large contribution to the rotational strength, which resembles the exciton coupling and the coupled oscillator model as discussed in a later section. The rotational strength of a non-degenerate system with the assumptions given above is obtained as follows:

$$\begin{aligned}
R^{0A} = \sum_i \Bigg[ & \mathrm{Im}\langle\vec{\mu}\rangle_{i0a}\cdot\langle\vec{m}\rangle_{ia0} \\
& -2\sum_{j\neq i}\sum_{b\neq a}\frac{\mathrm{Im}\,V_{i0a,j0b}\{\langle\vec{\mu}\rangle_{i0a}\cdot\langle\vec{m}\rangle_{jb0}\bar{\nu}_a + \langle\vec{\mu}\rangle_{j0b}\langle\vec{m}\rangle_{ia0}\bar{\nu}_b\}}{h(\bar{\nu}_{b0}^2-\bar{\nu}_{a0}^2)} \\
& -\sum_{j\neq i}\sum_{b\neq a}\frac{\mathrm{Im}\,V_{iab,j00}\{\langle\vec{\mu}\rangle_{i0a}\cdot\langle\vec{m}\rangle_{ib0} + \langle\vec{\mu}\rangle_{i0b}\cdot\langle\vec{m}\rangle_{ia0}\}}{h(\bar{\nu}_{b0}-\bar{\nu}_{a0})} \\
& -\sum_{j\neq i}\sum_{b\neq a}\frac{\mathrm{Im}\,V_{i0b,j00}\{\langle\vec{\mu}\rangle_{i0a}\cdot\langle\vec{m}\rangle_{iba} + \langle\vec{\mu}\rangle_{iab}\cdot\langle\vec{m}\rangle_{ia0}\}}{h\bar{\nu}_{b0}} \\
& -\sum_{j\neq i}\frac{\mathrm{Im}\,V_{i0a,j00}\{\langle\vec{\mu}\rangle_{iaa} - \langle\vec{\mu}\rangle_{i00}\cdot\langle\vec{m}\rangle_{ia0}\}}{h\bar{\nu}_{a0}} \\
& -\frac{2\pi}{c}\sum_{j\neq i}\sum_{b\neq a}\frac{V_{i0a,j0b}\bar{\nu}_{0a}\bar{\nu}_{0b}\left\{(\langle\vec{R}_j\rangle - \langle\vec{R}_i\rangle)\cdot(\langle\vec{\mu}\rangle_{j0b}\times\langle\vec{\mu}\rangle_{i0a})\right\}}{h(\bar{\nu}_{b0}^2-\bar{\nu}_{a0}^2)}\Bigg]
\end{aligned} \qquad [40]$$

The coupling $V_{isr,jtv}$ between the groups $i$, $j$ with their singly- or doubly-excited configurations $r$, $s$ and $t$, $v$ is given by the matrix elements

$$V_{irs,jtv} = \int \varphi_{ir}\varphi_{is}\hat{V}_{ij}\varphi_{jt}\varphi_{jv}\,\mathrm{d}\tau \qquad [41]$$

The potential energy operator $\hat{V}_{ij}$ of Equation [41] stands for the electrostatic interaction between the groups $i$ and $j$. $\varphi_{ir}$, $\varphi_{is}$, ...... etc. are group wavefunctions, where the first index represents the group and the second index the electronic state of the group $r$, $s$, ... in $i$ and $t$, $v$, ... in $j$. $\langle\vec{\mu}\rangle_{i0a}$ and $\langle\vec{m}\rangle_{ia0}$ are the electric and the magnetic dipole transition moments of an excitation of a group $i$ for a transition from the ground state $|0\rangle$ to the excited state $|a\rangle$ ($|0\rangle \to |a\rangle$). $\bar{\nu}_{a0}$ and $\bar{\nu}_{b0}$ are the energies given in wavenumbers of the transitions $|0\rangle \to |a\rangle$ and $|0\rangle \to |b\rangle$, respectively.

The first term of Equation [40] contributes to the optical activity of a molecule if the transition $|N\rangle \to |K\rangle$ is electrically and magnetically allowed and the chromophore is inherently dissymmetric. This term represents the optical activity of isolated groups. In the second term the electric and magnetic dipole transition moments on different groups contribute by coupling the groups in their excited states by electrostatic interaction. The interaction of electric and magnetic dipole transition moments of different transitions of one isolated group are described with the third and the fourth term of Equation [40]. The fifth term shows the dependence on the difference of the electric dipole moment of the ground and excited state. This term is – as the fourth term – very small because of the division by $\bar{\nu}_{a0}$ instead of a difference like $\bar{\nu}_{a0}^2 - \bar{\nu}_{b0}^2$. The coupling of the electric dipole transition moments of two different groups in the sixth term represents the exciton model of two groups if their excited states are non-degenerate. This is the term which is the origin for the exciton chirality method of Harada and Nakanishi if the two excited states of the coupled electric transition moments are obtained by absorption in different spectral regions. This term corresponds to the theory of coupled oscillators. The interaction between degenerate groups does not contribute to the rotational strength of a polymer here. This effect is discussed in the last section for a system of two chromophores.

## Rules for the determination of the absolute configuration

### Chirality functions

For molecules with an inherently symmetric chromophore or molecules which can be formally decomposed into an achiral skeleton and achiral groups, so-called ligands, chirality functions have been developed which allow a quantitative description of chiral phenomena as a function of phenomenological parameters of the ligand. Additionally these functions allow the correlation of the sign of the chirality phenomenon and the absolute configuration of the molecule.

A systematic analysis of such chirality functions, taking advantage of the symmetry properties of the skeleton, has been given by Ruch and Schönhofer. The ligands, the number of which is determined by the chosen binding sites and the symmetry of the skeleton, are at positions where symmetry operations applied to the skeleton permute the ligands on the different possible binding sites and in this way create new distinguishable molecules. The binding site of a ligand is correlated with an argument for the ligands in the function, i.e. the arguments $\lambda(l_1)$, $\lambda(l_2)$, etc. correspond to the positions 1, 2, etc. in **Figure 7**. In the approximation of the so-called 'qualitative completeness', the chirality function $\chi(l_1,l_2,l_3,l_4)$ for allene derivatives with a skeleton of $D_{2d}$ symmetry and with four binding sites for the ligands which depend on two phenomenological parameters $\lambda(l_i)$, $\mu(l_i)$ for each ligand $l_i$ can be given by

$$\begin{aligned}\chi(l_1,l_2,l_3,l_4) &= \eta_1[\lambda(l_1)-\lambda(l_2)][\lambda(l_3)-\lambda(l_4)]\\ &+ \eta_2[\mu(l_1)-\mu(l_2)][\mu(l_1)-\mu(l_3)][\mu(l_1)-\mu(l_4)]\\ &\times[\mu(l_2)-\mu(l_3)][\mu(l_2)-\mu(l_4)][\mu(l_3)-\mu(l_4)]\end{aligned} \qquad [42]$$

$\eta_1$ and $\eta_2$ are two additional phenomenological parameters. For the optical rotation the phenomenological parameters for a large number of ligands have been determined. Ligand parameters of one and the same ligand are different if they are bound to achiral skeletons of different symmetry and different sites, i.e. a determined parameter for one special ligand can only be taken for compounds with identical skeletons. There are further restrictions for the

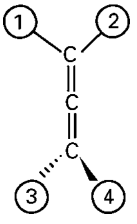



**Figure 7** Achiral skeleton with $D_{2d}$ symmetry and the four substitution positions 1 to 4 for an allene.

description of the optical rotation by chirality functions because of some experimental and theoretical reasons. One restriction is the assumption that the ligands at a binding site of the skeleton have to be invariant under the symmetry operation of the site group. This means that a ligand at a binding site must possess sufficient symmetry to make all properties invariant under the symmetry operation of the skeleton. In spite of this restriction, the chirality function allows us in an excellent way to prove whether molecular theories are suitable to describe chirality phenomena or they allow us to write these theories in a way that we can see immediately that they describe a chirality phenomenon. An example of such an application is the rewriting of Kirkwood's polarizability theory of the optical rotation of methane and allene derivatives by Haase and Ruch.

## Sector rules

In order to correlate the absolute configuration of a molecule with the sign of the Cotton effect of a transition, empirical or theoretical rules are needed. One form of rules are the so-called sector rules which are applicable to inherently symmetric skeletons/chromophores. In this case the molecules have to be decomposed into a symmetric skeleton/chromophore and perturbing atoms. The potential $V$ of the perturbing atoms is expressed as shown in Schellman's general theory for sector rules by symmetry-adapted functions as

$$V = \sum_{v,i} V_i^v \qquad [43]$$

where the $V_i^v$ belong to the $i$th row of the $v$th representation of the symmetry group of the inherently symmetric skeleton. The circular dichroism in the chromophore is then induced by the term $V_i^v$ of Equation [43] which belongs to the pseudoscalar representation of the symmetry point group of the chromophore. With this perturbation within the one-electron theory of Condon, Altar, and Erying the induced CD can be evaluated as a function of the positions of the perturbing atoms or groups. A plausible derivation according to Schellman's description for a chromophore of $C_{2v}/D_{2h}$ symmetry represented by the plane shown in **Figure** 8 can be done as follows: A perturbing atom at the position $P(x, y, z)$ (right) makes the entire system chiral with a certain handedness. The molecule is chiral as long as the perturbing atom is not positioned in one of the symmetry planes of the achiral chromophore, e.g. in $P(0, y, z)$ (middle). When the perturbing atom is shifted to the position $P(-x, y, z)$ (left), the mirror image of the system in the right part of the figure is obtained. For the resulting molecule a sign change in the Cotton effect has to be measured because this molecule is the enantiomer of the initial molecule. If the existence of chiral zero points can be excluded the sign of the CE can only be changed when the perturbing atom crosses one of the symmetry planes of the model molecule (**Figure** 8). All atoms – except fluorine – induce the same sign for the Cotton effect. This means, however, that a sector rule for this model exists which is a quadrant rule for $C_{2v}$ and an octant rule for $D_{2h}$ symmetry. The quadrant rule can change to an octant rule if it is taken into account that, e.g. a new apparent symmetry plane can be introduced into the molecule by the nodal plane of the $\pi^*$ orbital of the carbonyl group. The question whether experimentally a quadrant or an octant rule has to be used for the carbonyl chromophore has been discussed in the literature for a very long time.

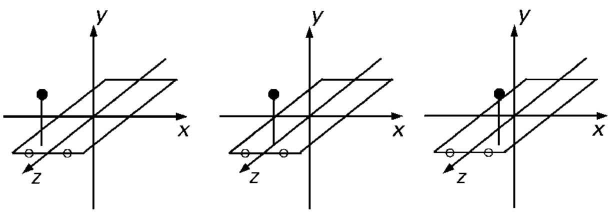


**Figure 8** Scheme for developing a sector rule.

**Figure** 9 shows six examples of inherently symmetric chromophores in which the region where atoms induce a Cotton effect of different sign are indicated as black or white. The higher the symmetry the larger is the number of sign changes when moving a perturbing atom in space. As a consequence, rules for chromophores with high symmetry are not very reliable.

For the most popular rule, the octant rule for the carbonyl chromophore, an application is

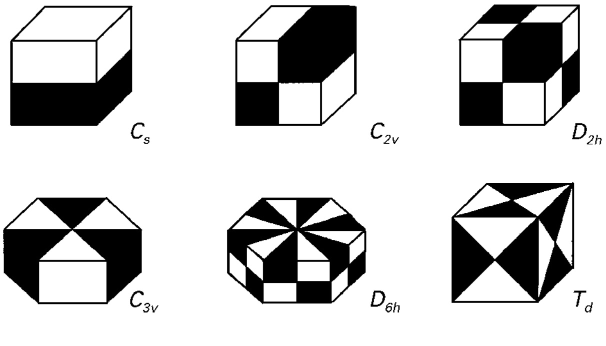


**Figure 9** Schematic representation of the sector rules for the skeletons or chromophores of $C_s$, $C_{2v}$, $D_{2h}$, $C_{3v}$,$D_{6h}$, and $T_d$ symmetry according to Schellman.

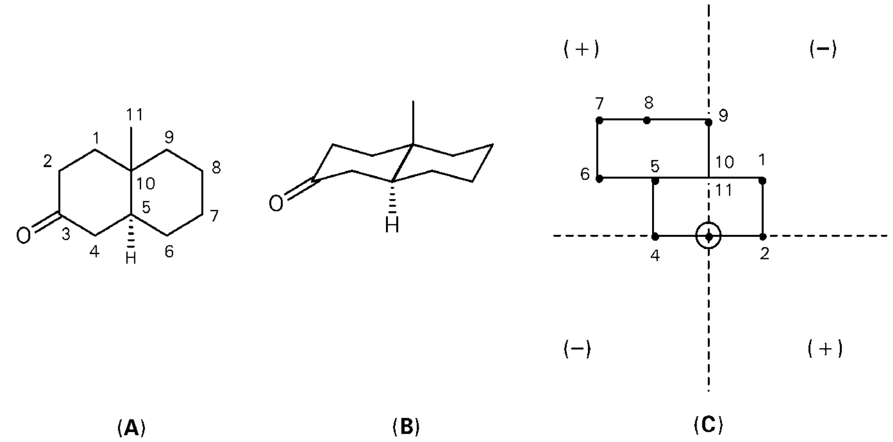



Figure 10 Octant projection of a compound with a carbonyl group showing a CE induced by the second sphere. Each atom outside the symmetry planes contributes to the CD. The contributions of atoms which are in mirror image positions compensate.

demonstrated in **Figure 10** for a cyclic ketone. To perform a sector rule the region around the chromophore is divided into sectors by symmetry or nodal planes: Atoms in a symmetry or a nodal plane do not contribute to the CD, the contributions of atoms in neighbouring sectors are always opposite in sign. Spectra of an aromatic system where a sector rule can be applied to the third sphere are given as further examples in **Figure 11**.

## Helicity rules

Sector rules fail if a chromophore is inherently dissymmetric and the rotational strength is only determined by its internal structure. From the definition of a chromophore given above it is easy to imagine that the delimitation of an inherently dissymmetric from an inherently symmetric chromophore is not straightforward. Furthermore, it is difficult to choose atoms in a molecule with which a helical arrangement can be approximately defined. If there is at least one full turn of a helix constituted by a dissymmetric arrangement of groups or atoms that are positioned exactly on a helical line, then the sense of helicity can be assigned unequivocally and a correlation of the sign of the CE can be experimentally or theoretically determined to derive a general rule. Otherwise only a helicity with atoms which are approximately positioned on a helical line can be determined. It has been pointed out in Janoschek by Snatzke that such an approximately determined helicity of a molecule – for which a helicity rule can be applied – depends on the viewing angle. In general this means, in order to find a helical order in a molecule, one has to select a direction in a chromophore and then has to choose atoms within the chromophore which lay approximately on a helical line with respect to the chosen viewing angle. But the question is which viewing angle has to be chosen, because there are many directions for the selection of an approximate helical order for which one can try to develop a helicity rule. In most cases chemical intuition from experimental experience can help to solve this problem. In some cases the CD of an oriented chromophore in an anisotropic sample (ACD) can help to decide whether there is a helical structure along a special direction or not as discussed by Kuball and Höfer.

Helicity rules have been developed which can be applied to inherently dissymmetric chromophores, such as, e.g. the enone chromophore shown in **Figures 12** and **13**, where the chosen atoms are only approximately on a helical line. Many papers have been published concerning the enone chromophore as summarized in Patai and Rappoport by Gawronski.

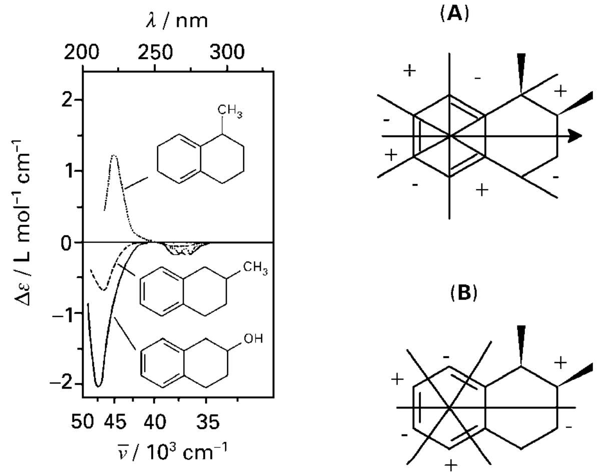



Figure 11 The benzene chromophore as an example for the application of sector rules for the third sphere. On the right the schematic representation of the sector rules for the long wavelength ${}^{1}L_{b}$ transitions (A) and for the short wavelength ${}^{1}L_{a}$ transitions (B) are given. The plane of the aromatic system is a nodal plane.

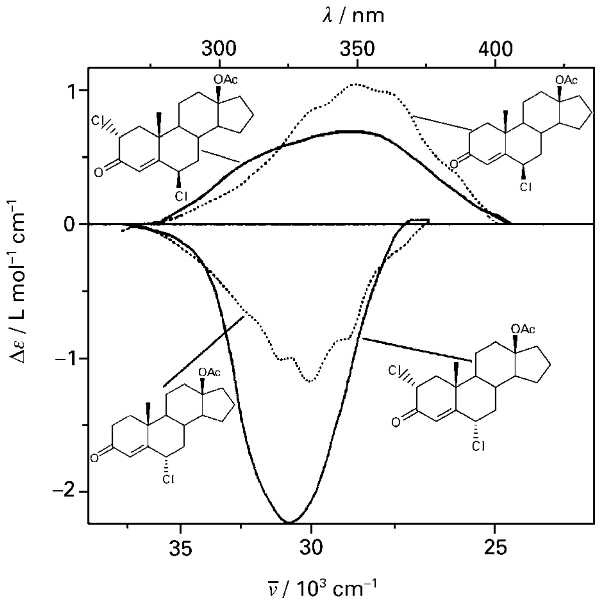



**Figure 12** The enone chromophore as an example for the application of helicity rules.

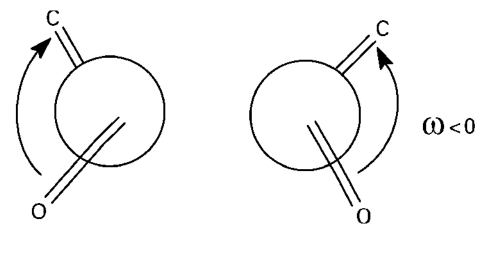



**Figure 13** Helicity of the enone chromophore.

In the simple cases the torsional angle of the enone chromophore determines the sign of the Cotton effect where for a P helix ($\omega > 0$) and for an M helix ($\omega < 0$) a positive and a negative CE for the n$\pi^*$ transition result, respectively. Substitution at the chromophore may have dramatic effects and can be responsible for the requirement of a modification of the rule.

## The exciton chirality method

Two or more chromophores located nearby in space can constitute a chiral system. An electric dipole–electric dipole interaction between the chromophores shifts and splits the energy levels of their excited states with respect to the interaction-free states. In the case of two identical chromophores, e.g. in a molecule with at least, or even approximately, the point symmetry group $C_2$, the degeneracy is removed (**Figure 14**).

As a consequence the UV spectrum is built up from two transitions with different transition

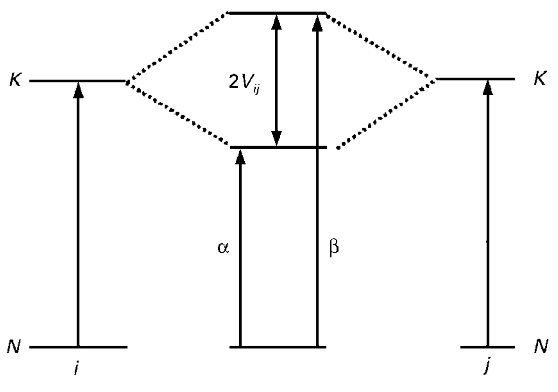



**Figure 14** Davidov splitting $2V_{ij}$ of the two degenerate excited states of two interacting identical chromophores in a compound of $C_2$ symmetry for $\vartheta < \vartheta_0$.

moment directions as shown for $\beta R$, $\beta' R$-dimethylmesobilirubin-XIII$\alpha$ (**Figure 15**) which gives rise to a frequency-dependent degree of anisotropy $R$ (**Figure 16**). In the case of a molecule of $C_2$ symmetry the transitions are polarized perpendicular to each other, one transition parallel to the $C_2$ symmetry axis and the other one in the plane perpendicular to this symmetry axis. For both of these transitions, using the shape of the UV band of the transition of the dipyrrinone fragment, the absorptions $\varepsilon^{\alpha}(\bar{\nu})$ and $\varepsilon^{\beta}(\bar{\nu})$ (**Figure 17**) for $\beta R$, $\beta' R$-dimethylmesobilirubin-XIII$\alpha$ are given by

$$\varepsilon(\bar{\nu}) = F^{NK}\left(\varepsilon^{\alpha}(\bar{\nu}) + \varepsilon^{\beta}(\bar{\nu}) - \left(\varepsilon^{\alpha}(\bar{\nu}) - \varepsilon^{\beta}(\bar{\nu})\right)\cos\vartheta\right) \quad [44]$$

where $\vartheta$ is the angle between the dipole transition moments of the fragments and $F^{NK}$ is a factor which takes into account a change of intensity of the monomer band by the interaction in the molecule (**Figure 17** left) built up of the two fragments. In the spectral regions of the absorption bands belonging to the exciton transitions, a positive and a negative CE appears as shown in **Figure 17** (right). Both CEs superpose to a couplet which can be described by bands possessing the shape of the UV spectra as follows

$$\Delta\varepsilon(\bar{\nu}) = \frac{4R^{NK\alpha}}{D^{NK\alpha}}\left(\varepsilon^{\alpha}(\bar{\nu}) - \varepsilon^{\beta}(\bar{\nu}) - (\varepsilon^{\alpha}(\bar{\nu}) + \varepsilon^{\beta}(\bar{\nu}))\cos\vartheta\right) \quad [45]$$

where $R^{NK\alpha}$ is the rotational strength of the transition $\alpha$ (**Figure 14**) and $D^{NK\alpha}$ the corresponding

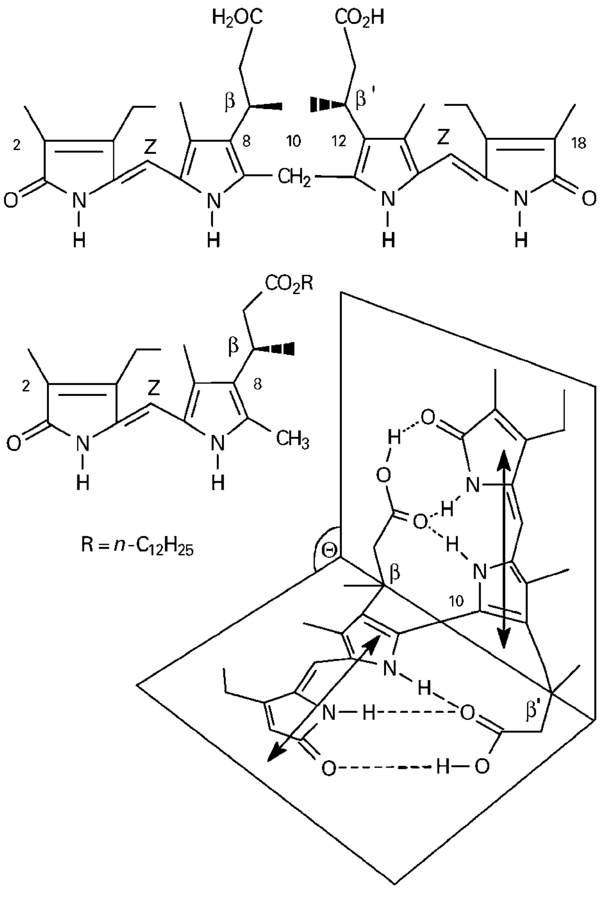



Figure 15 Linear and folded structure of $\beta R, \beta' R$-dimethylmesobilirubin-XIIIα of *M* helicity and the dipyrrinone ester as the building unit (fragment). Θ is the angle between the mean molecular planes of the dipyrrinone units. The arrows indicate the transition moment directions within the dipyrrinone units which enclose the angle $\vartheta$. The $x_2^*$ axis is chosen to be parallel to the $C_2$ symmetry axis which is the angle bisector through C10 between the mean molecular planes of the dipyrrinone units. In the solution there is an excess of the *P* conformer in $P \rightleftarrows M$.

dipole strength. The fitting of the CD and UV bands of $\beta R$, $\beta' R$-dimethylmesobilirubin-XIIIα (**Figure 15**) according to Equations [44] and [45] allows to determine the angle $\vartheta$ between the dipole transition moments of the fragments.

Using a simple version of the exciton theory, the exciton couplet built up by the CD curves of two identical chromophores $i$ and $j$ of different signs can be described assuming spectral functions of Gaussian form by

$$\Delta\varepsilon(\bar{\nu}) = \frac{2\sqrt{\pi}\bar{\nu}_0^2}{2.296 \times 10^{-39}\Delta\bar{\nu}^2}\left[\frac{\bar{\nu}_0 - \bar{\nu}}{\Delta\bar{\nu}}\right] \times \exp\left\{-\left[\frac{\bar{\nu}_0 - \bar{\nu}}{\Delta\bar{\nu}}\right]^2\right\}\vec{R}_{ij} \cdot \left(\langle\vec{\mu}\rangle_{iNK} \times \langle\vec{\mu}\rangle_{jNK}\right)V_{ij} \qquad [46]$$

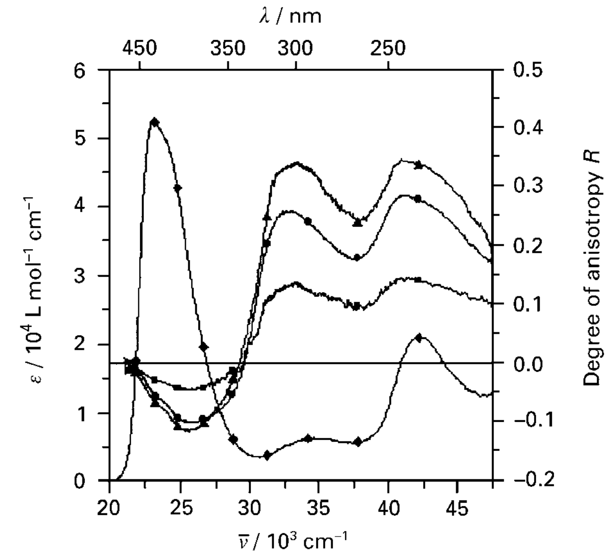



Figure 16 The UV spectrum (◆) ($T$ = 80°C) and the temperature dependent degree of anisotropy of $\beta R, \beta' R$-dimethylmesobilirubin-XIIIα in ZLI-1695 ($T$ = 28°C (▲), 43°C(●), and 65°C (■)).

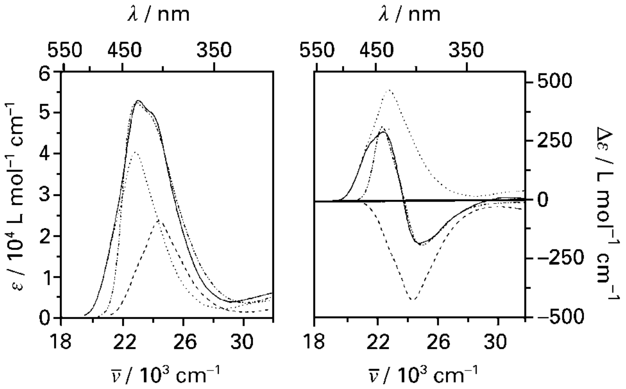



Figure 17 Exciton coupling of the two chromophores of $\beta R, \beta' R$-dimethylmesobilirubin-XIIIα dissolved in ZLI-1695 in the CD and UV spectrum. The dotted and the dashed lines are the calculated contributions of the two exciton bands. The dashed–dotted lines are the calculated CD and UV spectra.

Here, $\nu_0$ is the centre of the exciton bands and $\Delta\nu$ the half band width of the single exciton bands $\alpha$ and $\beta$. The transition moment of the transition $\alpha$ or $\beta$ can be expressed by

$$\langle\vec{\mu}\rangle^{\alpha,\beta} = \frac{1}{\sqrt{2}}\left(\langle\vec{\mu}\rangle_{iNK} \mp \langle\vec{\mu}\rangle_{jNK}\right) \qquad [47]$$

where the upper sign belongs to the $\alpha$ and the lower sign to the $\beta$ band. The pseudoscalar product $\vec{R}_{ij}\cdot(\langle\vec{\mu}\rangle_{iNK} \times \langle\vec{\mu}\rangle_{jNK})$ in Equation [46] determines the sign of the couplet. The Davidov splitting $2V_{ij}$

(**Figure 14**) between the energy states $N$, $K$ of the groups $i$ and $j$ is given by

$$V_{ij} = \frac{(\langle\vec{\mu}\rangle_{iNK}\cdot\langle\vec{\mu}\rangle_{jNK}) - 3R_{ij}^{-2}(\langle\vec{\mu}\rangle_{iNK}\cdot\vec{R}_{ij})(\langle\vec{\mu}\rangle_{jNK}\cdot\vec{R}_{ij})}{R_{ij}^{3}} \qquad [48]$$

where $\vec{R}_{ij}$ is a vector pointing from group $i$ to group $j$ and the Davidov splitting is smaller than the half-band width. For derivation of the exciton theory it is assumed that the interacting groups are far away from each other and thus the distance of the chromophores $R_{ij} = |\vec{R}_{ij}|$ can be chosen independently from the extension of the chromophores. This is a presupposition not fulfilled in most cases of application of the theory. Furthermore, one has to take into account an effect given above but often not realized in the literature: A change of sign of the couplet appears without a change of the helicity if the sequence of the states resulting from the Davidov splitting $2V_{ij}$ (Eqn [48]) is interchanged (**Figure 14**; sign change of $V_{ij}$). The sign change depends on the angle $\vartheta$ between the transition moment directions. For $V_{ij} = 0$ there is a chiral zero of the 'chirality function $\Delta\varepsilon(\bar{\nu})$' (Eqn [46]) i.e. $\Delta\varepsilon = 0$, in spite of the fact that the system is chiral. $V_{ij}$ is zero when $\vartheta = \vartheta_0$ and cos $\vartheta_0 = R_{ij}^{-2}|\langle\vec{\mu}\rangle_{iNK}|^{-1}\cdot|\langle\vec{\mu}\rangle_{jNK}|^{-1}(\langle\vec{\mu}\rangle_{iNK}\cdot\vec{R}_{ij})(\langle\vec{\mu}\rangle_{jNK}\cdot\vec{R}_{ij})$. If $\vartheta$ is about 90° and thus near a chiral-zero position only an experimental assignment of the $\alpha$ and $\beta$ transitions can help to avoid an incorrect conclusion. This is also of importance because the vector $\vec{R}_{ij}$ is strongly dependent on the origin not only in its length but also in its orientation with respect to the molecular skeleton. This fact makes it difficult to estimate the sign of the couplet because of the chiral zero ($V_{ij} = 0$) given by the uncertainty of the chosen $\vartheta$. This problem has to be considered from the point of view that in most applications of the exciton chirality method the theoretical condition that the distance between the chromophores is large compared to their extension is not well fulfilled.

Taking these discussed problems into account, the following rule holds: Positive (P) and negative (N) helicities lead always to positive (positive CE at the longer wavelength side and negative CE at the shorter wavelength side) and negative couplets of the CD as long as the angle $\vartheta$ between the transition moment direction is smaller than $\vartheta_0$ with $0 \le \vartheta \le \vartheta_0$. Here the helicity of the structure is defined as positive, when the geometrical arrangement gives rise to a clockwise rotation of the electric dipole transition moment in the foreground if looking along the axis about which the rotation is performed and vice versa. Choosing suitable values for $|\vec{R}_{ij}|$ and taking care that the dipole strength of the absorption band of the exciton couplet is approximately twice the dipole strength of the monomer's absorption which build up the inherently dissymmetric molecule, reliable conclusions for the absolute configuration can be obtained. So far, the exciton chirality method is a useful method for establishing absolute configurations and conformations of organic compounds in solution.

In order to improve the reliability of the conclusion from the CD and UV measurements of isotropic solutions, the assignment of the exciton transitions should be determined experimentally. For the evaluation of the correct dihedral angle, polarized spectroscopy can be used as a very suitable additional information for the assignment. Since, in most cases, this information has not been taken into account, the value of such an analysis is demonstrated in the following with $\beta R$, $\beta' R$-dimethylmesobilirubin-XIII$\alpha$ and its constituting unit dipyrrinone (**Figure 15**) according to Bauman and co-workers. Due to the assumed $C_2$ symmetry of $\beta R$, $\beta' R$-dimethylmesobilirubin-XIII$\alpha$ the transition moments either lie parallel to the $C_2$ symmetry axis or in a plane perpendicular to it.

From the degree of anisotropy $R$ in **Figure 14** it follows that the $\beta$ and the $\alpha$ transitions are of $A \rightarrow A$ and $A \rightarrow B$ symmetry, respectively, for $\vartheta < \vartheta_0$. From the temperature dependence of the degree of anisotropy (**Figure 16**) with the help of a computational simulation the tensor coordinates of the absorption tensor and the dipole strength tensor $D_{ii}^{*}$ with respect to the principal axes of the order tensor can be obtained. Furthermore, information about the orientational order of the molecule in the used anisotropic system as stretched polymers or ordered liquid crystal phases can be obtained too.

To determine the transition moment directions in the dipyrrinone units and of the $\alpha$-transition of $\beta R$, $\beta' R$-dimethylmesobilirubin-XIII$\alpha$ which are not determined by symmetry, the ratio $V$ (Eqn [49], derived from a simple version of the exciton theory) was used as further information to select possible mathematical solutions for the transition moment direction of the dipyrrinone unit.

$$V = \frac{D_{22}^{*\beta}}{D_{11}^{*\alpha} + D^{*\alpha}} = \frac{1+\cos\vartheta}{1-\cos\vartheta} \approx \frac{\varepsilon^{\beta}(\bar{\nu}_{\max})}{\varepsilon^{\alpha}(\bar{\nu}_{\max})} \qquad [49]$$

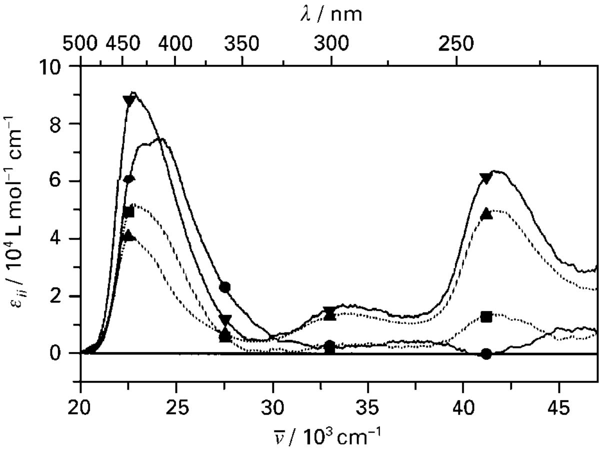



**Figure 18** The tensor coordinates $\varepsilon^*_{11}$(······■······), $\varepsilon^*_{22}$ (——●——), $\varepsilon^*_{33}$(······▲······) of the absorption tensor and the sum $\varepsilon^*_{11}+\varepsilon^*_{33}$ (——▼——) of *βR,β′R*-dimethylmesobilirubin-XIIIα in ZLI-1695. $\varepsilon^*_{11}+\varepsilon^*_{33}$ is the long wavelength ($\alpha$ band: A → B) and $\varepsilon^*_{22}$ is the short wavelength ($\beta$ band: A → A) exciton band.

This ratio determined from the $\varepsilon^{\alpha}(\bar{\nu})$ and $\varepsilon^{\beta}(\bar{\nu})$ curves of **Figure 17** and the $\varepsilon_{22}(\bar{\nu})$ and $\varepsilon_{11}(\bar{\nu})+\varepsilon_{33}(\bar{\nu})$ curves from **Figure 18** allows to check the reliability of the experimental results obtained from UV and CD spectroscopy in isotropic media and from polarized UV spectra. In the example given here, the ratio $V$ from both results is equal within 30%, which is in fair agreement with the expected accuracy of the determined reduced spectra. In **Figure 18** the reduced UV spectra for the exciton coupling are given. The analysis for this special example demonstrates how useful it can be to determine and analyse polarized spectroscopy for the use of the exciton chirality method. In particular, the angle of the chiral zero can be estimated to avoid errors resulting from interchanging the position of the $\alpha$ and $\beta$ states in the term scheme (**Figure 14**).

## List of symbols

$c$ = velocity of light; $c$ = concentration in mol L$^{-1}$; $d$ = optical path length; $D^{NnKk}$ = dipole strength for the transition $|Nn\rangle \rightarrow |Kk\rangle$; $D_{ij}^{NnKk}$ = transition moment tensor for the transition $|Nn\rangle \rightarrow |Kk\rangle$; $-e$ = charge of the electron; $E_n$ = energy of the state $|n\rangle$; $f$ = orientational distribution function; $F^{NK}$ = factor taking into account a change of intensity of the monomer bands; $F^{NnKk}$ = spectral function for the absorption band of the transition $|Nn\rangle \rightarrow |Kk\rangle$; $g$ = dissymmetry factor; $G^{NnKk}$ = spectral function of the vibronic transition $|Nn\rangle \rightarrow |Kk\rangle$; $h, \hbar = h/2\pi$ = Planck's constant; $H'_{21}$ = matrix element; $l_i$ = ligand; $\langle m_r\rangle_{n0}$ = coordinate of the magnetic dipole transition moment for the transition $|0\rangle \rightarrow |n\rangle$; $m$ = mass of the electron; $\langle \vec{m}_{ia0}\rangle$ = magnetic dipole transition moment of group $i$ for the transition $|0\rangle \rightarrow |a\rangle$; $\langle \vec{m}\rangle_{NnKk}$ = magnetic dipole transition moment for the transition $|Nn\rangle \rightarrow |Kk\rangle$; $M_{ii}$ = optical rotation tensor; $[M]$ = molar rotation; $n$, $k$ = vibrational states; $n_{1,2,\ L,R}$ = refractive indices for linearly or circularly polarized light; $N$ = number of molecules in a volume $V$; $N_A$ = Avogadro's number; $N,K$ = electronic states; $p_i$ = operator of the linear momentum; $\langle Q_{sj}\rangle_{n0}$ = electric quadrupole transition moment for the transition $|0\rangle \rightarrow |n\rangle$; $R$ = degree of anisotropy; $R^{0A}$ = rotational strength for the transition $|0\rangle \rightarrow |A\rangle$; $R^{NnKk}$ = rotational strength for the electronic transition $|Nn\rangle \rightarrow |Kk\rangle$; $\langle \vec{R}_i\rangle$ = distance vector of group $i$ from the origin of the coordinate system; $\vec{R}_{ij}$ = distance vector between two groups $i, j$; $R_{ij}^{NnKk}$ = rotational strength tensor for the transition $|Nn\rangle \rightarrow |Kk\rangle$; $V_{irs,\,jtv}$ = interaction potential between groups $i, j$; $V$ = potential of the perturbing atoms; $_{ij}$ = potential energy operator; $2V_{ij}$ = Davidov splitting; $x_i$, $x'_i$ = axes of the molecule fixed or space fixed coordinate system; $\alpha, \beta, \gamma$ = Eulerian angles; $\delta_{ij}$ = Kronecker symbol; $\Delta\varepsilon$ = CD of an isotropic system; $\Delta\varepsilon_{ii}(i = 1, 2, 3)$ = coordinates of the CD tensor; $\Delta\varepsilon_{ij}$ = circular dichroism tensor; $\varepsilon_{ij}$ = absorption tensor; $\varepsilon_{ijk}$ = Levi Civita tensor; $\varepsilon$ = molar decadic absorption coefficient; $\varepsilon_{1,2,\ L,R}$ = absorption coefficients for linearly or circularly polarized light; $\eta_1,\eta_2$ = phenomenological parameters; $\eta$ = positive damping parameter representing the natural line shape; $\theta$ = ellipticity; $\theta_{LB,CD}$ = ellipticity related to linear birefringence or circular dichroism; $\vartheta$ = angle between the electric dipole transition moments of two fragments; $\lambda$ = wavelength; $\lambda_{Nn}$, $\lambda_{Kk}$ = librational states; $\lambda(l_i)$, $\mu(l_i)$ = phenomenological parameters for ligand $l_i$; $\langle \mu_r\rangle_{0n}$ = coordinate of the electric dipole transition moment for the transition $|0\rangle \rightarrow |n\rangle$; $\langle \vec{\mu}\rangle_{NnKk}$ = electric dipole transition moment for the transition $|Nn\rangle \rightarrow |Kk\rangle$; $\langle \vec{\mu}\rangle_{i0a}$ = electric dipole transition moment of group $i$ for the transition $|0\rangle \rightarrow |a\rangle$; $\bar{\nu}$ = wavenumber; $\bar{\nu}_x$ = wavenumber for the ($k^{\text{th}}$+1) band of a progression; $\bar{\nu}_{00}$ = wavenumber for the origin of a progression; $\bar{\nu}_0$ = centre of the exciton bands; $\phi$ = optical rotation; $\phi_{LD,CB}$ = rotation related to linear dichroism or circular birefringence; $\hat{\phi}$ = complex optical rotation; $\varphi_{ij}$ = group wave functions; $\chi(l_1, l_2, l_3, l_4)$ = chirality function; $\omega = 2\pi c\bar{\nu}$ = frequency.

*See also:* **Biochemical Applications of Raman Spectroscopy; Biomacromolecular Applications of Circular Dichroism and ORD; Chiroptical Spectroscopy, Emission Theory; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems;**

**Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism; Induced Circular Dichroism; Magnetic Circular Dichroism, Theory; ORD and Polarimetry Instruments; Rotational Spectroscopy, Theory; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading


Barron LD (1982) *Molecular Light Scattering and Optical Activity*. Cambridge: Cambridge University Press.

Bauman D, Killet C, Boiadjiev SE, Lightner DA, Schönhofer A and Kuball H-G (1996) Linear and circular dichroism spectroscopic study of $\beta$, $\beta'$-dimethylmesobilirubin-XIII$\alpha$ oriented in a nematic liquid crystal. *Journal of Physical Chemistry* **100**: 11546–11558.

Charney E (1979) *The Molecular Basis of Optical Activity – Optical Rotatory Dispersion and Circular Dichroism*. New York: Wiley.

Condon EU, Altar W and Eyring H (1937) One-electron rotatory power. *The Journal of Chemical Physics* **5**: 753–775.

Grimme S (1996) Density functional calculations with configuration interaction for the excited states of molecules. *Chemical Physics Letters* **259**: 128–137.

Gō N (1965) Optical activity of anisotropic solutions I. *Journal of Chemical Physics* **43**: 1275–1289.

Haase D and Ruch E (1973) Quantenmechanische Theorie der optischen Aktivität der Methanderivate im Transparenzgebiet. *Theoretica Chimica Acta* **29**: 189–234.

Harada N and Nakanishi K (1983) *Circular Dichroic Spectroscopy – Exciton Coupling in Organic Stereochemistry*. Oxford: Oxford University Press.

Janoschek R (ed) (1991) *Chirality – From Weak Bosons to the $\alpha$-Helix*. Berlin: Springer Verlag.

Kuball H-G and Höfer T (1999) Chiroptical spectroscopy, oriented molecules and anisotropic systems. In Lindon J, Tranter G and Holmes JL (eds) *Encyclopedia of Spectroscopy and Spectrometry*. London: Academic Press.

Kuball H-G, Karstens T and Schönhofer A (1976) Optical activity of oriented molecules II. Theoretical description of the optical activity. *Chemical Physics* **12**: 1–13.

Kuball H-G, Altschuh J and Schönhofer A (1979) Optical Activity of Oriented Molecules III. The absorption process and the vibrational coupling effects for the tensor of rotation. *Chemical Physics* **43**: 67–80.

Kuball H-G, Neubrech S and Schönhofer A (1992) Optical activity of oriented molecules. $\alpha$, $\beta$-unsaturated steroid ketones and their sector rules. *Chemical Physics* **163**: 115–132.

Loxsom FM (1969) Optical rotation of helical polymers: periodic boundary conditions. *Journal of Chemical Physics* **51**: 4899–4905.

Moffit W and Moscowitz A (1959) Optical activity in absorbing media. *The Journal of Chemical Physics* **30**: 648–660.

Nakanishi K, Berova N and Woody RW (eds) (1994) *Circular Dichroism–Principles and Applications*. New York: VCH Publishers.

Patai S and Rappoport Z (eds) (1989) *The Chemistry of Enones*. New York: John Wiley & Sons.

Rodger A and Norden B (1997) *Circular Dichroism & Linear Dichroism*. Oxford: Oxford University Press.

Ruch E and Schönhofer A (1970) Theorie der Chiralitätsfunktionen. *Theoretica Chimica Acta* **19**: 225–287.

Schellman JA (1966) Symmetry rules for optical rotation. *Journal of Chemical Physics* **44**: 55–63.

Snatzke G (ed) (1967) *Optical Rotatory Dispersion and Circular Dichroism in Organic Chemistry*. London: Heyden and Son.

Snatzke G (1968) Circulardichroismus und optische Rotationsdispersion – Grundlagen und Anwendung auf die Untersuchung der Stereochemie von Naturstoffen. *Angewandte Chemie* **80**: 15–26.

Snatzke G (1979) Circulardichroismus und absolute Konformation: Anwendung der qualitativen MO-Theorie auf die chiroptischen Phänomene. *Angewandte Chemie* **91**: 380–393.

Tinoco I (1962) Theoretical aspects of optical activity part two: polymers. *Advances in Chemical Physics* **4**: 113–160.

Weigang OE (1965) Vibrational structuring in optical activity II. "Forbidden" character in circular dichroism. *Journal of Chemical Physics* **43**: 3609–3618.

# Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems

**Hans-Georg Kuball** and **Tatiana Höfer**, Fachbereich Chemie der Universität Kaiserslautern, Germany

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Theory**

## Introduction

Homochirality, such as the dominance of the left-handed α-amino acids owing to symmetry breaking arising from the nuclear weak interaction and/or other mechanisms, is the basis of biomolecular enantioselection. The 'wrong' chiral form can, for example as a compound in a drug, lead to unwanted and even dangerous effects. Thus, the phenomenon of 'chirality' itself as well as new ideas for exploiting it are of current interest.

Nowadays, chirality is often discussed from the point of view of such basic phenomena as charge conjugation, parity and time reversal (CPT symmetry). But for a 'way in' to chirality in chemistry and physics, the different occupation of space by two enantiomers, inherent in the definition of Kelvin, and its extension to statistical systems by Avnir covers most of the problems. Kelvin's definition provides a background for a hierarchy of chiral objects or structures through four levels: atoms, molecules, phases with long-range orientational and positional order, and the shape of objects. Long-range orientational and positional order in a phase can lead to a microscopic chiral structure of particles in the phase. The shape of a phase is its macroscopic appearance, e.g. the crystal habit. Chirality that is caused by long-range orientational and positional order will be called 'suprastructural phase chirality' in the following.

Methods are required to detect chirality. A so-called 'chirality observation' answers the question whether the system is chiral or not with yes or no. Then, a 'chirality measurement' yields a value and a sign for a quantity that gives information about the chirality of the molecules or the phases but, in general, gives no measure for chirality itself.

One of the most popular methods for detecting chirality is the measurement of the optical activity. A chiral object rotates the plane of polarization of linearly polarized light and generates ellipticity in the region of absorption bands because of the different refractive indices and absorption coefficients for left- and right-circularly polarized light – circular birefringence and dichroism. The main interest in circular dichroism (CD) spectroscopy is based on three applications. First, CD spectroscopy can be applied as a very specific and sensitive fingerprint method in analytical chemistry; CD spectrometer units have even been developed as detection devices in chromatography. Secondly, the determination of the absolute configuration of a molecule is of great interest in spite of the fact that the CD method is not an absolute method in contrast to Bijvoet's method. Thus, many rules, so-called sector and helicity rules, for different classes of compounds have been developed for electronic transitions in the visible and UV spectral region to correlate the CD with the absolute configuration. For vibrational transitions in the IR spectral region, the existence of such correlation is the exception, not the rule.

The circular dichroism, $\Delta\varepsilon$, of an isotropic system is one third of the trace of a pseudotensor of second rank:

$$\Delta\varepsilon = \tfrac{1}{3}(\Delta\varepsilon_{11} + \Delta\varepsilon_{22} + \Delta\varepsilon_{33}) \qquad [1]$$

To increase the quantity of information available for a molecule, the three coordinates $\Delta\varepsilon_{ii}$ ($i$ = 1,2,3) of $\Delta\varepsilon$ in Equation [1], i.e. the CD of oriented molecules in anisotropic samples (ACD), can be measured.

There is a third problem for which chirality information is of current interest: anisotropic phases are often stabilized by chiral structures. Apart from chiral structures with enantiomorphic crystals of chiral compounds, suprastructural chirality exists in liquid crystal phases built up by chiral molecules as in the cholesteric phases and the smectic C* phases. Even liquid crystalline phases with suprastructural chirality originating in achiral, so-called banana-shaped molecules, seem to be possible. Anisotropic polymer films with chiral structures have been found. It can be anticipated that chiroptical spectroscopy with anisotropic chiral systems will lead to new questions and answers.

## Birefringence and dichroism of chiral isotropic and chiral anisotropic phases

### Optical rotatory dispersion and circular dichroism in isotropic dissymmetric phases

The interaction of light with an isotropic dissymmetric solution leads to a left- and a right-circularly polarized light beam for which distinct refractive indices $n_L$ and $n_R$ as well as absorption coefficients $\varepsilon_L$ and $\varepsilon_R$ exist owing to the diastereomeric interaction between light and the chiral system. The plane of polarization of the incident linearly polarized light is rotated by an angle $\phi$ (optical rotation in rad) that is proportional to the circular birefringence $n_L - n_R$,

$$\phi = \pi \bar{\nu} d (n_L - n_R) \quad \text{(rad)} \qquad [2]$$

where $d$ is the path length of the sample in cm and $\bar{\nu}$ is the wavenumber of the incident light in $cm^{-1}$. The specific rotation $[\alpha]$ for a fluid or solid compound is defined with respect to the density $\rho$ (g $cm^{-3}$) or for a solution with respect to a concentration $q$ in g/(100 $cm^3$ solution) and the angle $\alpha$ in degree:

$$[\alpha] = \frac{10\alpha}{\rho d} = \frac{10^3\alpha}{qd} \quad \left(\frac{\text{deg cm}^3}{\text{g dm}}\right) \qquad [3]$$

The molar rotation $[M]$ can be expressed by

$$[M] = \frac{[\alpha]M}{100} = \frac{100\alpha}{cd} \quad \left(\frac{\text{deg cm}^3}{\text{mol dm}}\right) \qquad [4]$$

where $M$ is the molecular mass in g $mol^{-1}$ and $c$ (mol $L^{-1}$) the concentration. In the region of an absorption band, both circularly polarized waves superpose and the light becomes elliptically polarized with an ellipticity $\tan\theta$ given by

$$\tan\theta = \frac{10^{-\varepsilon_L cd/2} - 10^{-\varepsilon_R cd/2}}{10^{-\varepsilon_L cd/2} + 10^{-\varepsilon_R cd/2}} \cong \theta$$
$$= \frac{\ln 10}{4}(\varepsilon_L - \varepsilon_R)cd = 0.5757\,\Delta\varepsilon\, cd \qquad [5]$$

Instead of the circular dichroism $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$, often the molar ellipticity $[\theta]$ is used:

$$[\theta] = \frac{100\theta}{cd} \approx 3300\,\Delta\varepsilon \qquad [6]$$

Because there are two independent waves propagating through the sample (**Figure 2**), the Lambert–Beer law does not hold,

$$I = I_0\, 10^{-\bar{\varepsilon} cd} \cosh\left(\frac{2\,303}{2}(\varepsilon_L - \varepsilon_R)\, cd\right) \qquad [7]$$

but the deviation is in most cases negligible. $\bar{\varepsilon}$ is the average absorption of the sample:

$$\bar{\varepsilon} = \tfrac{1}{2}(\varepsilon_L + \varepsilon_R) \qquad [8]$$

The chirality measurements (**Figure 1**) circular dichroism (CD) and optical rotatory dispersion (ORD), the wavenumber dependence of the molar rotation, are designated together as the Cotton effect (CE). CD and ORD are quantitatively correlated by a Kramers–Kronig transformation. The Kramers–Kronig transformation can be used to check CD measurements by ORD measurements and vice versa by comparing calculated and experimental results.

### Anisotropic samples

The anisotropy of systems is a widespread phenomenon in nature. The description as anisotropic means that a material has different properties along different directions in space. This anisotropy, caused by long-range orientational and positional order, can be homogeneous, statistically homogeneous or periodically homogeneous. These microscopic structures, e.g. the screw axis of a $NaClO_3$ crystal or the pitch of the helical structure in a cholesteric or smectic C* phase, have to be taken into consideration when, for

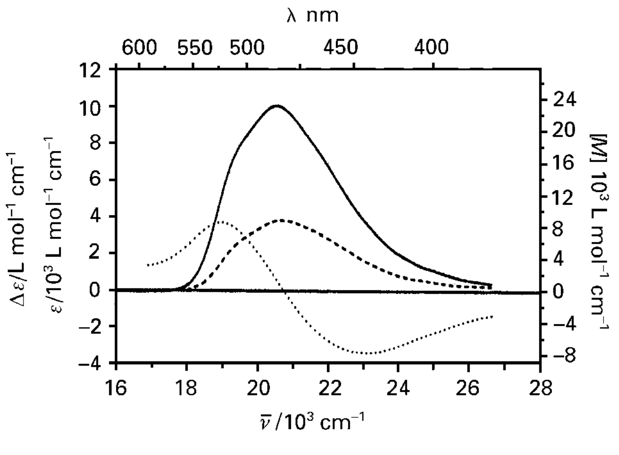



**Figure 1** A positive Cotton effect (CE), i.e. the CD (– – –) and ORD (•••), and a UV absorption band (——).

example, an internal periodicity yields internal reflection of circularly polarized light in the visible or infrared spectral region.

## The concept of eigenstates of light

The interaction of light with a material, described by the Jones or Mueller calculus, leads to the concept of eigenstates. These states are determined by the symmetry of the material that is penetrated by the light (**Table 1**). An eigenstate of light refers to light that propagates through a material without changing its state of polarization. For isotropic materials, all states of polarization, and hence also unpolarized light, are eigenstates. Inherently dissymmetric isotropic solutions possess left- and right-circularly polarized light as the two orthogonal eigenstates. For anisotropic uniaxial and biaxial materials, the eigenstates are two orthogonal linearly polarized light beams, whereas in complex systems, e.g. in chiral anisotropic systems, the eigenstates are two orthogonal elliptically polarized light beams (**Figure 2**).

The linearly, elliptically or circularly polarized eigenstates propagate with different velocities because of the different refractive indices. Different absorption of the two independent light beams leads to dichroism. Thus, each of the eigenstates has its own and distinct refractive index and molar decadic absorption coefficient (**Table 1**). The plane light waves of the two eigenstates interfere after passing the sample to form an elliptically polarized light beam with an ellipticity and an azimuth different from that of the incident light (**Figure 2**), in general. Its intensity may be less than that of the incident light beam. Linear and circular eigenstates and thus linear and circular birefringence and dichroism of a sample can be seen as two limiting cases of elliptical birefringence and elliptical dichroism.

**Table 1** Material constants for birefringence and dichroism

| *Two orthogonal eigenstates of light* | *Refractive indices (n) and absorption coefficients (ε)* | *Birefringence* | *Dichroism* | *Mean absorption* |
|---|---|---|---|---|
| Linearly polarized light | $n_1,\varepsilon_1$; $n_2,\varepsilon_2$ | $n_1-n_2$ | $\varepsilon_1-\varepsilon_2$ | $\frac{1}{2}(\varepsilon_1-\varepsilon_2)$ |
| Elliptically polarized light | $n_{EL},\varepsilon_{EL}$; $n_{ER},\varepsilon_{ER}$ | $n_{EL}-n_{ER}$ | $\varepsilon_{EL}-\varepsilon_{ER}$ | $\frac{1}{2}(\varepsilon_{EL}+\varepsilon_{ER})$ |
| Circularly polarized light | $n_L,\varepsilon_L$; $n_R,\varepsilon_R$ | $n_L-n_R$ | $\varepsilon_L-\varepsilon_R$ | $\frac{1}{2}(\varepsilon_L+\varepsilon_R)$ |

## Chiral anisotropic systems

A chiral system is a system that is not superposable on its mirror image. The number of independent measurements of different kinds needed to identify the sample as chiral depends on the knowledge of the system. If it is known, for example, that the system is an isotropic dissymmetric phase, one optical measurement is sufficient, whereas for an anisotropic uniaxial system more than one measurement is demanded. For the latter case the wavenumber dependence of the ellipticity and the rotation of the plane of polarization can be measured. Here the dispersion of the circular birefringence and dichroism can be, but needs not be, different from that of the linear birefringence and dichroism (**Figure 3**). Another possibility is the measurement of the rotation of the plane of polarization and the ellipticity for different azimuth orientations of the sample about the axis of propagation of the light beam.

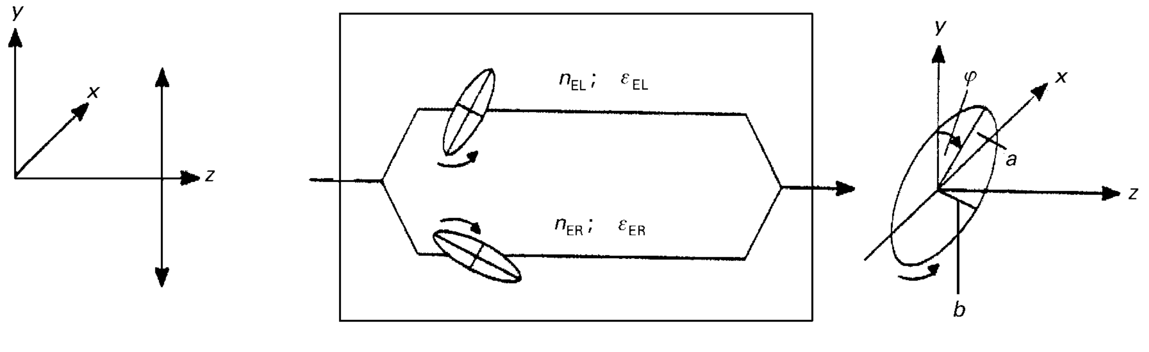



**Figure 2** Schematic representation of the two eigenstates of light (see **Table 1**) propagating through an anisotropic phase of chiral molecules with or without suprastructural chirality.

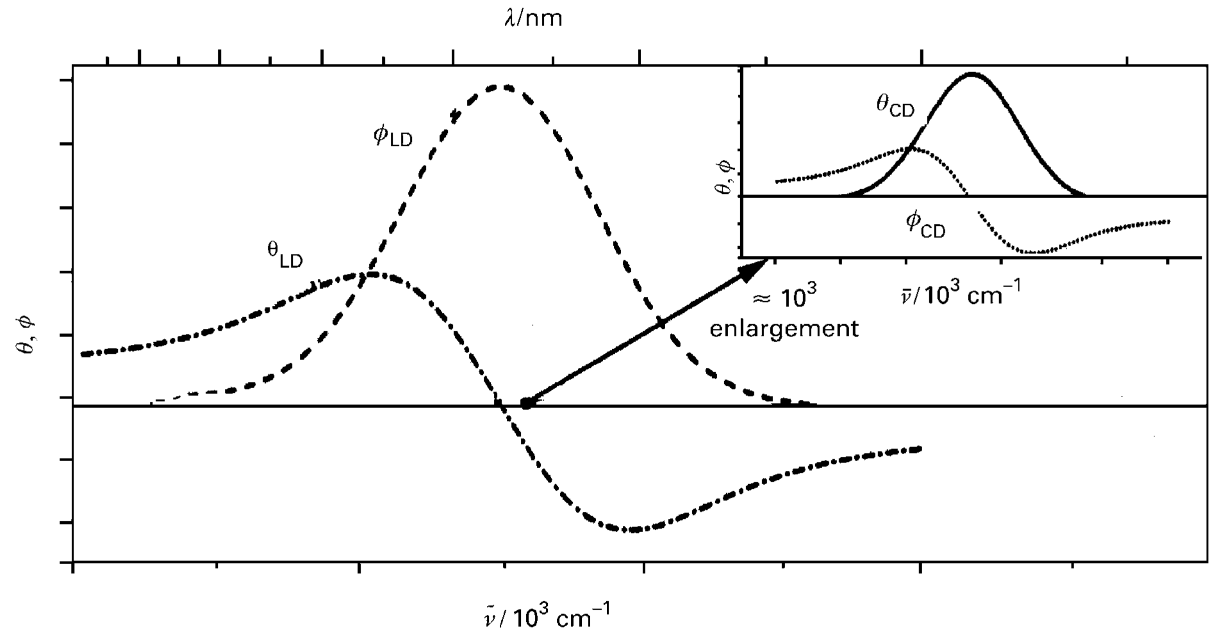



**Figure 3** Dispersion of linear and circular birefringence and dichroism, $\phi_{LD}$, $\theta_{LD}$, $\phi_{CD}$, $\theta_{CD}$ ($\Delta\varepsilon$) within an absorption band of a chiral molecule.

## The circular dichroism of oriented molecules: ACD spectroscopy

### Phenomenology of the CD of anisotropic phases (ACD) without suprastructural chirality

The two eigenstates of light emerging from a homogeneous sample superpose in accordance with their different optical path lengths. The azimuth and the ellipticity are then given by the Stokes vector $\underline{s}^{\mathrm{d}}$,

$$\underline{s}^{\mathrm{d}} = \mathbf{F} \cdot \underline{s}^{0} \qquad [9]$$

where $\mathbf{F}$ is the Mueller matrix and $\underline{s}^0$ the Stokes vector of the incident light. In $\underline{s} = \{s_0, s_1, s_2, s_3\}$, $s_0$ is the intensity whereas $\vec{s} = \{s_1, s_2, s_3\}$ describes the state of polarization (the notation { } indicates a column vector). For a nondepolarizing medium,

$$\mathbf{F} = \exp(-2nd\mathbf{A}) \quad \text{with}$$

$$\mathbf{A} = \begin{pmatrix} c' & b_1 & b_2 & b_3 \\ b_1 & c' & a_3 & -a_2 \\ b_2 & -a_3 & c' & a_1 \\ b_3 & a_2 & -a_1 & c' \end{pmatrix} = \mathbf{A}' + c'\mathbf{I} \qquad [10]$$

is obtained. $d$ is the pathlength and $n$ the number of particles in a unit volume. $\mathbf{I}$ is the identity matrix. It is convenient to rewrite the Mueller matrix in the form

$$\mathbf{F} = \exp(-2\,ndc')\boldsymbol{\Lambda}, \quad \text{with} \quad \boldsymbol{\Lambda} = \exp(-2nd\,\mathbf{A}') \qquad [11]$$

As shown by Schönhofer and co-workers $\boldsymbol{\Lambda}$ can be rewritten as a matrix polynomial of third degree:

$$\boldsymbol{\Lambda} = c_0(\mathbf{A}')^3 + c_1(\mathbf{A}')^2 + \left[(\lambda^2 - \kappa^2)c_0 + c_2\right]\mathbf{A}' + \left[(\lambda^2 - \kappa^2)c_1 + c_3\right]\mathbf{I} \qquad [12]$$

$$c_0 = -\frac{1}{\kappa^2 + \lambda^2}\left[\frac{1}{\kappa}\sinh(2nd\kappa) - \frac{1}{\lambda}\sin(2nd\lambda)\right]$$

$$c_1 = \frac{1}{\kappa^2 + \lambda^2}\left[\cosh(2nd\kappa) - \cos(2nd\lambda)\right]$$

$$c_2 = -\frac{1}{\kappa^2 + \lambda^2}\left[\kappa\sinh(2nd\kappa) + \lambda\sin(2nd\lambda)\right]$$

$$c_3 = \frac{1}{\kappa^2 + \lambda^2}\left[\kappa^2\cosh(2nd\kappa) + \lambda^2\cos(2nd\lambda)\right] \qquad [13]$$

with $^{\kappa}_{\lambda} = \{\frac{1}{2}\{((\vec{b}^2 - \vec{a}^2 + 4(\vec{a}\cdot\vec{b}^2)^{1/2} \pm (\vec{b}^2 - \vec{a}^2)]\}^{1/2} \geq 0$. For a homogeneous medium with large linear and circular birefringence and dichroism effects, all elements of the differential Mueller matrix $\mathbf{A}'$ (for the

birefringence $\vec{a} = \{a_1, a_2, a_3\}$ and for the dichroism $\vec{b} = \{b_1, b_2, b_3\}$ in Equation [12]) have to be taken to evaluate the azimuth and ellipticity of the emerging light. The elements $a_2$ and $b_2$ are responsible for the optical rotation and for the circular dichroism, respectively, and $a_1$ and $a_3$ represent the anisotropy of the refractive index of linear polarized light whereas $b_1$ and $b_3$ belong to the corresponding linear dichroism. Because there are six unknown parameters for the anisotropy (Eqn [12]), six different measurements made by rotating the sample and/or measuring with a light beam propagating in different directions within the sample are needed to determine the optical constants. Measuring the dispersion of the ACD or the optical rotatory power of an oriented ensemble (AORD) does not yield new and independent information (**Figure 3**).

The ACD measurement $\Delta\varepsilon^A$ with light propagating parallel to the optical axis of a uniaxial phase is a chirality measurement, but one has to take care to avoid artefacts. A check on the reliability of such a measurement can be obtained by measuring the CD with different sample azimuths rotated around the direction of propagation of light. The CD signal has to be independent of the orientation of the azimuth of the sample. Measurements of the optical effects with a light beam propagating perpendicularly to the optical axis are very difficult to perform because of the artefacts that may be caused by interference with the linear birefringence and dichroism. At present it does not seem possible to overcome the experimental problems with instruments working on the basis of polarization modulation. These experimental problems can be solved using other techniques, e.g. by measuring the state of polarization of the light, but solving the experimental problem does not solve the whole problem. A measurement with a light beam propagating perpendicularly to the optical axis is not a chirality measurement; that is, measuring two anisotropic phases of enantiomers with a light beam propagating perpendicularly to the optical axis does not yield values that are equal in their absolute value and different in sign, because of the interference with linear effects. The question is whether the measured quantity can be decomposed into an achiral and a chiral part ($\Delta\varepsilon_2^A$), i.e. whether chirality information can be extracted from the data obtained. For a uniaxial system, the following relation can be derived from $b_2$:

$$\Delta\varepsilon = \tfrac{1}{3}(\Delta\varepsilon^A + 2\Delta\varepsilon_2^A) \qquad [14]$$

For light polarized parallel and perpendicular to the optical axis and propagating perpendicularly to this axis, with the molar decadic absorption coefficients $\varepsilon_1$ and $\varepsilon_2$ for a uniaxial system, an equation analogous to Equation [14] is obtained:

$$\varepsilon = \tfrac{1}{3}(\varepsilon_1 + 2\varepsilon_2) \qquad [15]$$

A measurable macroscopic property of the anisotropic phase is then given by the tensor coordinates $Y_{kl}$ as given by Kuball and colleagues

$$Y_{kl} = \sum_{i,j} g_{ijkl}\, X_{ij} \qquad [16]$$

(Three tensors are of particular importance for ACD spectroscopy: the transition moment or the absorption tensor ($D_{ij}^{NK}$ or $\varepsilon_{ij}$), the rotation or the circular dichroism tensor ($R_{ij}^{NK}$ or $\Delta\varepsilon_{ij}$), and the order tensor ($g_{ij33}$). For molecules without any symmetry, the principal axes of the three tensors do not coincide: $x_i^+$ for $D_{ij}^{NK}$ or $\varepsilon_{ij}$, $x_i^{o}$ for $R_{ij}^{NK}$ or $\Delta\varepsilon_{ij}$, and $x_i^*$ for $g_{ij33}$.)

In Equation [16] $X_{ij}$ is a tensor of rank 2 that describes a molecular property. With the orientational distribution function of the molecules in the anisotropic sample $f(\alpha,\beta,\gamma)$, where $\alpha,\beta,\gamma$ are the Eulerian angles, the orientational distribution coefficients $g_{ijkl}$ (Eqn [17]) define a set of two different tensors of rank 2 with respect to the transformation of the molecule-fixed ($g_{ijkl}$: indices $ij$ for each pair of $kl$) as well as the space-fixed ($g_{ijkl}$: indices $kl$ for each pair of $ij$) coordinate system:

$$g_{ijkl} = \frac{1}{8\pi^2}\int f(\alpha,\beta,\gamma) a_{ik}\, a_{jl} \sin\beta \, d\alpha \, d\beta \, d\gamma \qquad [17]$$

$a_{ij}$ is an element of the transformation matrix from the space-fixed $x_i'$ to the molecule-fixed $x_i$ coordinate system. For a sample possessing at least a $C_3$ symmetry axis (uniaxial system) and a light beam propagating along this direction, the circular dichroism of the anisotropic sample $Y_{33} = \Delta\varepsilon^A$ is obtained:

$$\Delta\varepsilon^A = \sum_{i,j} g_{ij33}\, \Delta\varepsilon_{ij} \qquad [18]$$

$X_{ij} = \Delta\varepsilon_{ij}$ is the circular dichroism tensor of the molecules. According to Equation [14], the CD value $\Delta\varepsilon_2^A$ for light propagating perpendicularly to this axis, the optical axis, can be evaluated. Because of the interference with linear birefringence and dichroism, the present authors have not yet been able to measure $\Delta\varepsilon_2^A$.

For light polarized parallel to the optical axis and propagating perpendicularly to it follows,

$$\varepsilon_1 = \sum_{i,j} g_{ij11}\ \varepsilon_{ij} \qquad [19]$$

$X_{ij} = \varepsilon_{ij}$ is the absorption tensor of the molecules. The correlation between the second pair of indices of the orientational distribution coefficient and the numbering of the tensor coordinates of the measurable quantity gives rise to an interesting phenomenon. For the ACD ($\Delta\varepsilon^{A}$, Eqn [18]) the indices of the coordinates $g_{ijkl}$ are determined by the direction of propagation of light with respect to the molecular frame $kl = 33$. This is due to the fact that the CD effect for molecules is given by the projection of the molecular properties on the propagation direction of the light. For polarized absorption spectroscopy, the projection of the molecular properties onto the direction of polarization of the light beam is decisive, i.e. $kl = 11$, 22. This means that the numbering of the absorption tensor is obtained from the polarization direction of the light with respect to the molecular frame, whereas for the ACD the direction of propagation of the light wave is decisive.

### The experimental determination of the molecular tensors $\Delta\varepsilon_{ij}$ and $\varepsilon_{ij}$

**The equations for the ACD** The result of a CD measurement of an anisotropic sample is a weighted sum of the three quantities $\Delta\varepsilon_{ii}$ (Eqns [18] and [20]). The weighting factors are the orientational distribution coefficients. For a light beam propagating along a direction of at least a $C_3$ symmetry axis of a sample, the circular dichroism $\Delta\varepsilon^{A}$ (Eqn [18]) is a chirality measurement; this condition is fulfilled for a chiral nematic phase unwound by an electric field when the light beam propagates along the optical axis ($x'_3$), the nematic director. If the orientation of the principal axes of the order tensor $g_{ij33}$ relative to the molecular frame is temperature-dependent, one may use for the ACD:

$$\Delta\varepsilon^{A} = \sum_{i,j} a_{ij}^{2} g_{jj33}^{*}\ \Delta\varepsilon_{ii}^{o} \qquad [20]$$

The eigenvalues $\Delta\varepsilon_{ii}^{o}$ describe the molecular property $\Delta\varepsilon_{ij}$ in its system of principal axes. The terms $a_{ij}$ are the elements of the matrix that transforms the $x_i^*$ coordinates into the $x_i^o$ coordinates.

For uniaxial systems the order parameters $S^*$ and $D^*$ may be introduced instead of $g_{ij33}$

$$S^* = (3g_{3333}^{*} - 1)/2 \qquad [21]$$

$$D^* = \sqrt{3}(g_{2233}^{*} - g_{1133}^{*})/2 \qquad [22]$$

They refer to the optical axis of the sample, which is chosen as the space-fixed $x'_3$ axis, and to the molecule-fixed $x_i^*$ coordinate system. Whereas $S^*$ characterizes the order of the orientation of the molecule-fixed $x_3^*$ axis – the orientation axis – with respect to the optical axis of the uniaxial phase, the parameter $D^*$ is a measure of the deviation from a rotational symmetric distribution function for the molecules about their $x_3^*$ axis.

The principal axes of $g_{ij33}$ are fixed by symmetry for molecules with a point symmetry group different from $C_1$, $C_2$, $C_i$, $C_s$ and $C_{2h}$. For the latter point groups, the orientations of the $x_i^*$ axes with respect to the molecular skeleton have to be determined experimentally; the orientations of the $x_i^*$ axes vary, in general, with temperature. The two order parameters $S^*$ and $D^*$ can be measured, albeit with the use of some approximations, for example by IR, VIS, UV, or NMR spectroscopy.

If the temperature dependence of the orientation of the $x_i^*$ axes can be neglected, or if these axes are fixed by symmetry, $\Delta\varepsilon^{A}$ can be conveniently described as

$$\Delta\varepsilon^{A} - \Delta\varepsilon = (\Delta\varepsilon_{33}^{*} - \Delta\varepsilon)S^* + \frac{1}{\sqrt{3}}(\Delta\varepsilon_{22}^{*} - \Delta\varepsilon_{11}^{*})D^* \qquad [23]$$

From $\Delta\varepsilon^{A} - \Delta\varepsilon$ and $S^*$ and $D^*$ measured as functions of temperature, the quantities $3(\Delta\varepsilon_{ii}^{o} - \Delta\varepsilon) = 2\Delta\varepsilon_{ii}^{+} - \Delta\varepsilon\vec{s} - \Delta\varepsilon\bar{4}$ and $\Delta\varepsilon_{ii}^{NK,u} - \Delta\varepsilon\bar{\gamma}$ can be obtained. Because $\Delta\varepsilon_{2}^{A}$ is not available experimentally at present $\Delta\varepsilon$ (Eqn [1]) can be used to calculate the tensor coordinates $\Delta\varepsilon_{ii}^{*}$.

To account for the contributions of the different vibrational transitions, the anisotropic rotational strength $R^{A}$ for the electronic transition $|N\rangle \rightarrow |K\rangle$ is obtained by integrating over the spectral region of an absorption band:

$$R^{A} = \frac{3}{B}\int \frac{\Delta\varepsilon^{A}(\bar{\nu})}{\bar{\nu}}\,d\bar{\nu} = 3\sum_{i,j} g_{ij33} R_{ij}^{NK};$$

$$\frac{3}{B} = 22{\cdot}96 \times 10^{-40} \quad \text{(cgs)} \qquad [24]$$

For an isotropic solution, $g_{ij33} = \frac{1}{3}\delta_{ij}$ and $R^{A}$ becomes the rotational strength for the transition $|N\rangle \rightarrow |K\rangle$:

$$R^{NK} = \frac{3}{B}\int \frac{\Delta\varepsilon(\bar{\nu})}{\bar{\nu}}\,d\bar{\nu} = \sum_{i} R_{ii}^{NK} \qquad [25]$$

$\delta_{ij}$ is the Kronecker symbol. The tensor coordinates for the rotational strength tensor (tensor of rotation) can be evaluated experimentally as

$$R_{ij}^{NK} = \frac{1}{B}\int \frac{\Delta\varepsilon_{ij}(\bar{\nu})}{\bar{\nu}}\,\mathrm{d}\bar{\nu} \qquad [26]$$

For allowed electric dipole transitions, the contributions from vibrational coupling can be neglected. In the case of a forbidden electric or a magnetic dipole transition there should be contributions from the vibronic coupling and intensity-borrowing from other transitions via vibrations of the molecules.

The magnitude of the dissymmetry factor $g = 4R^{NK}/D^{NK}$ can be $10^{-1}$, $10^{-3}$, $10^{-4}$ for an inherently dissymmetric chromophore, an $n\pi^*$ transition and a $\pi\pi^*$ transition, where $D^{NK}$ is the dipole strength (see Eqn [31]).

**The equations for the polarized absorption** For an anisotropic sample, the molar decadic absorption coefficient for light polarized parallel ($\varepsilon_1$) and perpendicular ($\varepsilon_2$) to the optical axis of a uniaxial sample follows from Equations [19] and [15]:

$$\varepsilon_k = \sum_{i,j} g_{ijkk}\,\varepsilon_{ij} = \sum_{i,j} a_{ij}^2 g_{jjkk}^{*}\,\varepsilon_{ii}^{+}, \quad k = 1,2 \qquad [27]$$

$\varepsilon_{ii}^{+}$ are the absorption coefficients of a completely oriented system of molecules for light beams linearly polarized parallel to the $x_i^{+}$ axes (principal axes of $\varepsilon_{ij}$). The $a_{ij}$ in Equation [27] are the elements of the matrix that transforms the $x_i^{*}$ coordinates into the $x_i^{+}$ coordinates. The orientational distribution coefficients $g_{ij33}$ have to be used in the following instead of $g_{ij11}$ because, compared to the experimental situation of the CD measurement, the sample is rotated about an axis perpendicular to the propagation direction of the light beam by 90°. The $\varepsilon_{ii}^{+}$ terms are related to the absorption coefficient of the isotropic solution, analogously to $\Delta\varepsilon$, by Equation [28]:

$$\varepsilon = \tfrac{1}{3}\left(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}\right) \qquad [28]$$

If there is no temperature dependence of the $x_i^{*}$ axes, an appropriate description of the anisotropic absorption is given by

$$\varepsilon_1 - \varepsilon = (\varepsilon_{33}^{*} - \varepsilon)S^{*} + \frac{1}{\sqrt{3}}(\varepsilon_{22}^{*} - \varepsilon_{11}^{*})D^{*} \qquad [29]$$

Furthermore, for the degree of anisotropy follows,

$$R = \frac{\varepsilon_1 - \varepsilon_2}{3\varepsilon} = \tfrac{1}{2}\sum_{i,j}(3g_{ij33} - \delta_{ij})q_{ij}; \quad q_{ij} = \frac{\varepsilon_{ij}}{3\varepsilon} \qquad [30]$$

where the $q_{ij}$ are the squares of the direction cosines of the transition moment direction with respect to the $x_i$ axes in the case of a pure polarized transition. For mixed absorption bands the $q_{ij}$ are a measure of mixing. If the molecule-fixed coordinate system is arbitrarily chosen, nondiagonal elements occur in Equation [30]. The dipole strength is given by

$$D_k^{A} = \frac{12}{B}\int \frac{\varepsilon_k\,(\bar{\nu})}{\bar{\nu}}\,\mathrm{d}\bar{\nu} = 3\sum_{i,j} g_{ij33} D_{ij}^{NK} \qquad [31]$$

$k = 1, 2$ and with $g_{ij33} = \frac{1}{3}\delta_{ij}$ follows $D^{NK} = \Sigma_i\, D_{ii}^{NK}$ for an isotropic solution. The transition moment tensor

$$D_{ij}^{NK} = \frac{4}{B}\int \frac{\varepsilon_{ij}(\bar{\nu})}{\bar{\nu}}\,\mathrm{d}\bar{\nu} \qquad [32]$$

can be calculated from the coordinates of the absorption tensor (Eqn [32]).

**The tensor coordinates of the circular dichroism tensor $\Delta\varepsilon_{ij}^{*}$** The meaning of the diagonal elements of the tensor (Eqns [18] and [36]) can be visualized according to **Figure 4**. $\Delta\varepsilon_{1}^{*}$ can be obtained either by calculating a mean value of three CD measurements for completely oriented molecules with light beams polarized parallel to the three planes, schematically represented in **Figure 4**, or one CD measurement with an ensemble of molecules in a uniaxial system with $S^* = 1$ or a system with at least a $C_3$ symmetry axis with a light beam propagating parallel to the optical axis or the $C_3$ symmetry axis. The coordinates $\Delta\varepsilon_{1}^{**}$ and $\Delta\varepsilon_{22}^{*}$ have to be evaluated from the temperature-dependent $\Delta\varepsilon^{A}(T)$ and $\Delta\varepsilon$. For a direct measurement analogous to that of $\Delta\varepsilon_{11}^{*}$ a new orientational distribution is needed in which there is a symmetric rotational distribution about either the $x_1^{*}$ or $x_2^{*}$ axis. However, this procedure is nothing more than changing the numbering of the axes, because only that coordinate $\Delta\varepsilon_{ii}^{*}$ can be measured directly about which the orientational distribution function possesses a rotationally symmetric distribution.

The $\Delta\varepsilon_{ii}^{*}$ can also be different in sign and thus one may say that by 'looking' at the molecule from different directions one 'sees' $\Delta\varepsilon_{ii}^{*}$ with different sign. Changing the line of sight, the now observed $\Delta\varepsilon_{ii}^{*}$ varies strongly and can also change its sign. One might

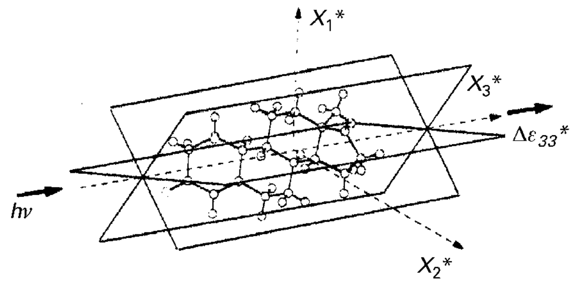



**Figure 4** Schematic representation of the measurement of $\Delta\varepsilon^*_{33}$.

loosely interpret this as 'seeing' the chirality of the molecule from different directions. This measurement must be done under conditions given in **Figure 4**.

### The molecular description of the circular dichroism and the absorption tensor

**The equations for molecules in an orientated state** From semiclassical or quantum field theory, the rotational strength $R_{ij}^{NnKk}$ for a transition from the electronic ground state $N$ with the vibrational state $n$ to the electronic excited state $K$ with the vibrational state $k$, i.e. $|Nn\rangle \rightarrow |Kk\rangle$, follows from the development of Kuball and co-workers:

$$R_{ij}^{NnKk} = \tfrac{1}{4}\sum_{r,s}\langle\mu_r\rangle_{NnKk}(\varepsilon_{rsi}\langle C_{sj}\rangle_{KkNn} + \varepsilon_{rsj}\langle C_{si}\rangle_{KkNn}) \qquad [33]$$

The tensor $\langle C_{sj}\rangle_{KkNn}$ can be decomposed into a symmetric and antisymmetric part, the magnetic dipole and the electric quadrupole moment:

$$\langle C_{sj}\rangle_{KkNn} = -\mathrm{i}\sum_r \varepsilon_{sjr}\langle m_r\rangle_{KkNn} - \frac{\omega_{KkNn}}{c}\langle Q_{sj}\rangle_{KkNn} \qquad [34]$$

$$\mu_i = -e\sum_\nu x_{i\nu}; \quad Q_{ij} = -\tfrac{1}{2}e\sum_\nu x_{i\nu}x_{j\nu};$$
$$m_r = -\frac{e}{2mc}\sum_\nu\sum_{i,j}\varepsilon_{rij}x_{i\nu}p_{j\nu} \qquad [35]$$

$\langle\mu_r\rangle_{NnKk}$ and $\langle m_r\rangle_{KkNn}$ represent the electric and the magnetic dipole transition moment, respectively, $\langle Q_{sj}\rangle_{KkNn}$ is the electric quadrupole transition moment. $-e$ is the charge, $m$ the mass of the electron. $\Sigma_\nu$ means the sum over all electrons. Operators depending on the dynamic variables of the nuclei are omitted because they do not contribute to the effect in the approximation used. $\varepsilon_{ijk}$ is the Levi–Civita tensor. $\varepsilon_{123} = \varepsilon_{312} = \varepsilon_{231} = 1$, $\varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1$; all other tensor coordinates $\varepsilon_{ijk}$ are equal to zero. With the rotational strength tensor $R_{ij}^{NnKk}$ (Eqn [33]) and the spectral function $G^{NnKk}(\bar{\nu})$ of the vibronic transition $|Nn\rangle \rightarrow |Kk\rangle$, the circular dichroism tensor $\Delta\varepsilon_{ij}$ for an oriented molecule can be expressed as

$$\Delta\varepsilon_{ij} = B\bar{\nu}\sum_n\sum_{Kk}R_{ij}^{NnKk}G^{NnKk}(\bar{\nu}) \qquad [36]$$

$$B = \frac{32\pi^3 N_A}{10^3 hc\ln 10}$$

$N^A$ is the Avogadro constant and $c$ is the velocity of light. The absorption of light by an oriented molecule requires knowledge of the electric dipole transition moment tensor (dipole strength)

$$D_{ij}^{NnKk} = \langle\mu_i\rangle_{NnKk}\langle\mu_j\rangle_{KkNn} \qquad [37]$$

and the corresponding absorption tensor $\varepsilon_{ij}$

$$\varepsilon_{ij} = \tfrac{1}{4}B\bar{\nu}\sum_n\sum_{Kk}D_{ij}^{NnKk}F^{NnKk}(\bar{\nu}) \qquad [38]$$

$F^{NnKk}(\bar{\nu})$ is the spectral function for the absorption band of the vibronic transition $|Nn\rangle \rightarrow |Kk\rangle$ which, in the approximation that is used here, is equal to $G^{NnKk}(\bar{\nu})$.

The $\Delta\varepsilon_{ij}$ term can be decomposed into an electric dipole/magnetic dipole and an electric dipole/electric quadrupole contribution:

$$\Delta\varepsilon_{ij}(\bar{\nu}) = \Delta\varepsilon_{ij}^{\mu m}(\bar{\nu}) + \Delta\varepsilon_{ij}^{\mu Q}(\bar{\nu}) \qquad [39]$$

The decomposition is, except for special circumstances, origin dependent. From the diagonal elements $\Delta\varepsilon_{ii}$ the quantities $\Delta_i(\bar{\nu})$ defined by

$$\Delta_1(\bar{\nu}) = \tfrac{1}{2}[\Delta\varepsilon_{11}(\bar{\nu}) - \Delta\varepsilon_{22}(\bar{\nu}) - \Delta\varepsilon_{33}(\bar{\nu})] \qquad [40]$$

and cyclic permutation of the indices 1, 2, 3 can be calculated. The $\Delta_i(\bar{\nu})$ terms yield information about the electric dipole ($\langle\mu_i\rangle_{NnKk}$), magnetic dipole ($\langle m_i\rangle_{KkNn}$), and electric quadrupole ($\langle Q_{ij}\rangle_{KkNn}$)

transition moments:

$$\Delta_i(\bar{\nu}) = -\tfrac{1}{2}B\bar{\nu}\sum_n\sum_{Kk}\mathrm{Im}\{\langle\mu_i\rangle_{NnKk}\langle m_i\rangle_{KkNn}\} \times G^{NnKk}(\bar{\nu}) + \Delta\varepsilon_{ii}^{\mu Q}(\bar{\nu}) \quad [41]$$

The decomposition of $R_{ii}^{NK}$ corresponding to Equation [39] yields

$$R_{11}^{\mu m} = \mathrm{Im}\{\langle\mu_2\rangle_{NK}\langle m_2\rangle_{KN} + \langle\mu_3\rangle_{NK}\langle m_3\rangle_{KN}\} \quad [42]$$

$$R_{11}^{\mu Q} = -\langle\mu_2\rangle_{NK}\langle Q_{31}\rangle_{KN} + \langle\mu_3\rangle_{NK}\langle Q_{21}\rangle_{KN} \quad [43]$$

by integration over the absorption band. $R_{22}^{NK}$ and $R_{33}^{NK}$ can be obtained by cyclic permutation of the indices. The terms $\langle\mu_i\rangle_{NK}$, $\langle m_i\rangle_{KN}$, and $\langle Q_{ij}\rangle_{KN}$ are the electric dipole magnetic dipole and electric quadrupole transition moments for the 0–0 electronic transition $|N\rangle \rightarrow |K\rangle$. By integration of Equations [40 and 41], wavenumber-independent quantities $\Delta_i$ can be obtained with the tensors of rotation given by Equations [44] and [45] and cyclic permutation of the indices 1, 2, 3:

$$\Delta_1 = \tfrac{1}{2}\left[R_{11}^{NK} - R_{22}^{NK} - R_{33}^{NK}\right] \quad [44]$$

$$\Delta_i = -\tfrac{1}{2}\mathrm{Im}\{\langle\mu_i\rangle_{NK}\langle m_i\rangle_{KN}\} + R_{ii}^{\mu Q} \quad [45]$$

**Vibronic coupling in ACD spectroscopy** Forbidden electric dipole transitions or allowed magnetic dipole transitions possess no or nearly no intensity of absorption in a spectrum. There are a number of different mechanisms by which a transition can gain intensity. One mechanism is the coupling of an electronic transition to the nuclear motions. By the weak dependence of the state functions $|Nn\rangle$, $|Kk\rangle$ on the nuclear coordinates, electronic states within the molecules are coupled. The transition moment of the forbidden transition $|N\rangle \rightarrow |K\rangle$ increases, as normally described, by borrowing intensity from other allowed electronic transitions via normal vibrations, i.e. the variation of the normal coordinates. Within the adiabatic approximation, the coordinates of the absorption tensor $\varepsilon_{ii}^*$ are intensified by the Herzberg–Teller mechanism via different normal vibrations which – depending on the irreducible representation to which the normal coordinate belongs – couple transitions with different transition moment directions. This can be demonstrated experimentally by measuring a frequency-dependent degree of anisotropy (Eqn [30]). With the Herzberg–Teller method the intensification of $\Delta\varepsilon^A$ (Eqn [46]) follows from Kuball and co-workers

$$\Delta\varepsilon^A(\bar{\nu}) = B\bar{\nu}\sum_{i,j} g_{ij33}\left(R_{ij}^{NK}G^{NK}(\bar{\nu}) + \sum_\mu R_{ij}^{NK,\mu}G^{NK,\mu}(\bar{\nu}) + \sum_{\mu,\nu}\left[R(1)_{ij}^{NK,\mu\nu}G_1^{NK,\mu\nu}(\bar{\nu}) + R(2)_{ij}^{NK,\mu\nu}G_2^{NK,\mu\nu}(\bar{\nu})\right]\right) \quad [46]$$

$R_{ij}^{NK}$ gives the contribution of the 0–0 transition to the CD band. The coefficients $R_{ij}^{NK,\mu}$, $R(1)_{ij}^{NK,\mu\nu}$ and $R(2)_{ij}^{NK,\mu\nu}$ determine the size of the intensity borrowing via the $\mu^{\text{th}}$ and $\nu^{\text{th}}$ vibrations from other electric dipole-allowed transitions with a corresponding transition moment direction. The frequency dependence of the progressions $G^{NK}(\bar{\nu})$, $G^{NK,\mu}(\bar{\nu})$, $G_1^{NK,\mu\nu}(\bar{\nu})$, and $G_2^{NK,\mu\nu}(\bar{\nu})$ are described by

$$G^{NK}(\bar{\nu}) = \sum_{n,k}|\langle n|k\rangle|^2 G^{NnKk}(\bar{\nu}) \quad [47]$$

$$G^{NK,\mu}(\bar{\nu}) = \sum_{n,k}\langle n|k\rangle\langle k|Q_\mu|n\rangle G^{NnKk}(\bar{\nu}) \quad [48]$$

$$G_1^{NK,\mu\nu}(\bar{\nu}) = \sum_{n,k}\langle n|k\rangle\langle k|Q_\mu Q_\nu|n\rangle G^{NnKk}(\bar{\nu}) \quad [49]$$

$$G_2^{NK,\mu\nu}(\bar{\nu}) = \sum_{n,k}\langle n|Q_\mu|k\rangle\langle k|Q_\nu|n\rangle G^{NnKk}(\bar{\nu}) \quad [50]$$

This means analogously to the UV absorption, each CD tensor coordinate $\Delta\varepsilon_{ii}^*$ is intensified by normal modes of different irreducible representations by a different amount, i.e. each $\Delta\varepsilon_{ii}^*$ $(i = 1,2,3)$ borrows intensity from other electronic transitions of suitable symmetry (Eqns [47] to [50]).

Because of the importance of intensity borrowing, a forbidden electric and magnetic dipole transition of a compound of symmetry $D_2$ will be analysed as an example. For the transition $|N\rangle \rightarrow |K\rangle$ between an electronic and vibrational ground and an electronic and vibrational excited state A → A, only the coefficient $R(2)_{ij}^{NK,\mu\nu}$ in Equation [46] is different from zero. With Equation [50], the frequency dependence of the CD and ACD is determined by a progression starting on false origins of one vibration $\mu = \nu = 1,2\ldots$ of an

irreducible representation $b_i$ ($i = 1,2,3$):

$$\bar{\nu}_k = \bar{\nu}_{00} + \bar{\nu}_{0\mu}(b_i) + n_k\bar{\nu}_0(a) \qquad [51]$$

$n_k = 0,1,2\ldots$ are the quantum numbers for the $k^{th}$ excited state of the involved totally symmetric vibration with the wavenumber $\bar{\nu}_0(a)$. $\bar{\nu}_k$ is the wavenumber for the $(k+1)^{th}$ band of the progression. The false origin is $\bar{\nu}_{00} + \bar{\nu}_{0\mu}(b_i)$ and every vibration of the irreducible representation $b_i$ can cause intensity borrowing from an allowed electric dipole transition A → $B_i$. The coupling scheme (**Figure 5**) for an A → $B_i$ transition contributes with four terms. The intensity borrowing for the corresponding UV band is achieved by the same progressions (Eqn [51]). The transition moment direction of the progressions is determined by the transition moment direction of the coupled electronic transitions A → $B_i$. For the A → A transition $\Delta\varepsilon_{11}$, $\varepsilon_{33}$ and $\varepsilon_{22}$ borrow intensity via vibrations of irreducible representation $b_1$ and $b_2$. $\Delta\varepsilon_{22}$, $\varepsilon_{11}$ and $\varepsilon_{33}$ borrow the intensity via vibrations of symmetry $b_3$ and $b_1$ whereas $\Delta\varepsilon_{33}$, $\varepsilon_{11}$ and $\varepsilon_{22}$ gain the intensity from the coupling with $b_3$ and $b_2$.

For each chromophore, the reduced CD spectra, the tensor coordinates $\Delta\varepsilon_{ii}$, and the reduced UV spectra, the tensor coordinates $\varepsilon_{ii}$ gain intensity (**Figure 6A** and **6B**) by a special amount of coupling. By substituting the molecule in the neighbourhood of this chromophore, another coupling is achieved and the reduced spectra $\Delta\varepsilon_{11}$, $\Delta\varepsilon_{22}$, $\Delta\varepsilon_{33}$ as well as $\varepsilon_{11}$, $\varepsilon_{22}$, $\varepsilon_{33}$ change. Thus, in both cases, $\Delta\varepsilon$, $\Delta\varepsilon^A$ $\varepsilon$, and $R$ are different. With this simple model a sign change in $\Delta\varepsilon$ and possibly also in $\Delta\varepsilon^A$ can be obtained without changing the absolute configuration, thus the sector or helicity rules for this chromophore break down, because the sector rule is now applied to the CD of vibrational progression for which the rule was not developed. Applying the rule to measurements in isotropic solution, one has to presuppose that the same progression determines $\Delta\varepsilon$ for both compounds. From ACD spectroscopy each tensor coordinate $\Delta\varepsilon_{ii}$ is obtained directly. This means that the sector rules or helicity rules can always be applied to that $\Delta\varepsilon_{ii}$ for which the rule has been developed.

## Application of ACD and polarized UV spectroscopy

In addition to the analytical application of the ACD spectroscopy as a fingerprint technique, structural differences far from the chromophore can also be detected. A complete analysis of the ACD, the CD, the UV and the polarized UV spectra requires the order tensor or at least an estimation of the order parameters. The variation of $\Delta\varepsilon^A$ with temperature for an n$\pi^*$ transition of androst-4,17-dien-3-one ([1]; **Figure 7**) and the knowledge of $g^*_{ii33}$ from $^2$H NMR allows the calculation of $\Delta\varepsilon^*_{ii}$ with respect to the principal axis of the order tensor. The tensor coordinates $\Delta\varepsilon^*_{ii}$ have different signs and are shifted with respect to each other, which indicates the vibronic effect of intensity borrowing. The UV spectrum of [1] as well as the coordinates $\Delta\varepsilon^*_{ii}$ have a low spectral resolution in comparison to the CD spectrum $\Delta\varepsilon$. The different signs of $\Delta\varepsilon^*_{ii}(\bar{\nu})$ and their shifts against each other 'burn holes' into the spectrum calculated from Eqn [1]. Thus as often found in CD spectroscopy, an apparent high spectral resolution of the CD spectrum is obtained in contrast to the UV spectrum.

As a second example, $\Delta\varepsilon^A$ for a uniformly polarized charge transfer transition of 1,4-bis($R$)-1-phenylethylamino)-9,10-anthraquinone [2] is shown in **Figure 8**. There is no vibronic coupling and therefore no difference in the frequency dependence of the tensor coordinates. Furthermore, the tensor coordinate $\Delta\varepsilon\vec{a} \approx 0$ because the transition is polarized parallel to the $x_3^*$ direction (long axis of the anthraquinone skeleton). This indicates that for [2] the $x_3^*$ axis is approximately a principal axis of the circular dichroism tensor. Of interest, but not shown for [2], is the large

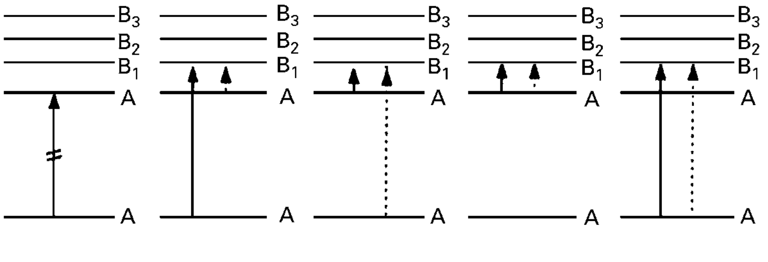



**Figure 5** Coupling scheme for intensity borrowing of the circular dichroism and absorption tensor coordinates of a forbidden A → A transition via the coefficient $R(2)\vec{b}$ with one vibration $b_1$ from an allowed electric (——) and magnetic (•••) dipole transition of A → $B_1$. Quadrupole contributions are neglected for this discussion.

contribution of the electric dipole/electric quadrupole term (Eqn [39]) to $\Delta\varepsilon^*_{ii}$ in the wavenumber region below 25 000 cm$^{-1}$.

With the steroid 3-methyl-3α-5α-3,3′,4′,5′-tetrahydrobenzo[2,3]cholest-2-en-1-one [3] and its diastereomer [4] the breakdown of a helicity rule was demonstrated by Frelek and co-workers (**Figure 9**). From X-ray analysis it is known that the helicity of the enone chromophore is opposite in sign in [3] and [4] and thus the fact that $\Delta\varepsilon < 0$ in both cases contradicts the helicity rule for cisoid enones either for [3] or [4]. From ACD spectroscopy follows that the enhancement of the $\Delta\varepsilon^*_{ii}$ by vibronic coupling for [4] yields a sign change of $\Delta\varepsilon$ without change of the absolute configuration of [4]. This is the first example where this effect has been proven.

At the end of the section on applications one must point out that it is impossible, at present, to obtain

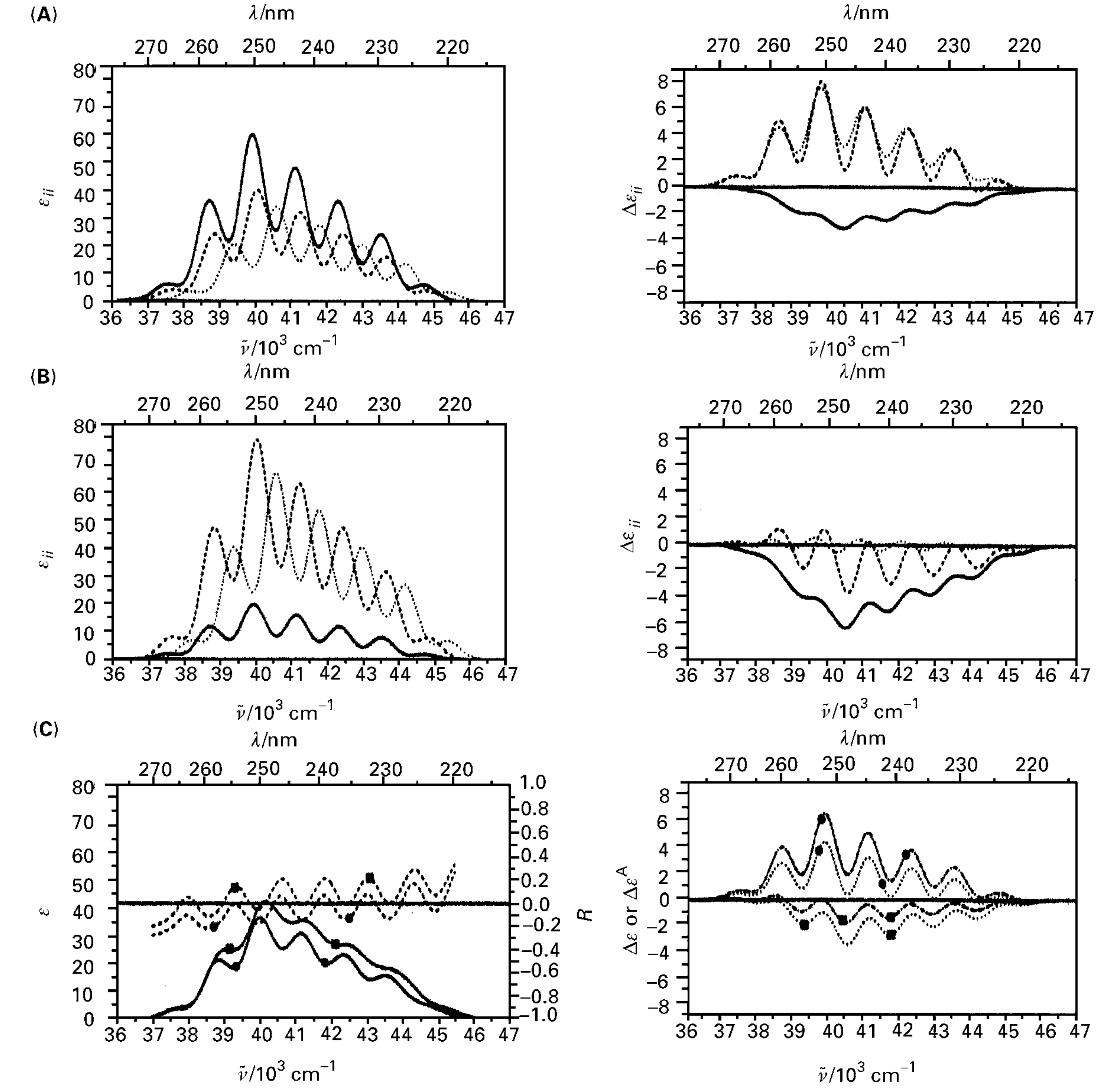



**Figure 6** Calculated UV (left) and CD (right) spectra of the A → A transition for a molecule of symmetry $D_2$ for two different strength of vibronic coupling. The intensity of the vibrational progressions is changed from $\varepsilon_{11} : \varepsilon_{22} : \varepsilon_{33} = 3 : 1 : 1$ (A left) to $\varepsilon_{11} : \varepsilon_{22} : \varepsilon_{33} = 1 : 2 : 2$ (B): ($\Delta\varepsilon_{11}$(′), $\varepsilon_{11}$(*) (——), $\Delta\varepsilon_{22}$(1), $\varepsilon_{22}$(*) (----), and $\Delta\varepsilon_{33}$(N), $\varepsilon_{33}$(A) (······). The UV spectra ε (——) and the degree of anisotropy *R* (----), (C, left), the CD spectra Δε (······) and ACD spectra $\Delta\varepsilon^A$ (---) (C, right) are given for the cases shown in (A) and (B). ((●) intensity ratio $\varepsilon_{11} : \varepsilon_{22} : \varepsilon_{33} = 3 : 1 : 1$; (■) intensity ratio $\varepsilon_{11} : \varepsilon_{22} : \varepsilon_{33} = 1 : 2 : 2$.) *R* and $\Delta\varepsilon^A$ have been calculated for the order parameter $S^* = 0.460$ and $D^* = 0.121$ where $D^*$ is calculated from Luckhurst's relation $D^* = F(S^*, \delta)$ with $\delta = 0.5$. According to Eqn [51] the assumed 0–0 transition and false origins are $\tilde{\nu}^*_{00} = 37100$ cm$^{-1}$; $\tilde{\nu}^*_{01}(b_1) = 1100$ cm$^{-1}$ $\tilde{\nu}^*_{02}(b_2) = 560$ cm$^{-1}$; $\tilde{\nu}^*_{03}(b_3) = 430$ cm$^{-1}$; $\tilde{\nu}_0(a) = 1200$ cm$^{-1}$; the intensity ratio within the progressions is chosen equally for $b_1$, $b_2$, $b_3$: 1 : 6 : 10 : 8 : 6 : 4 : 1.

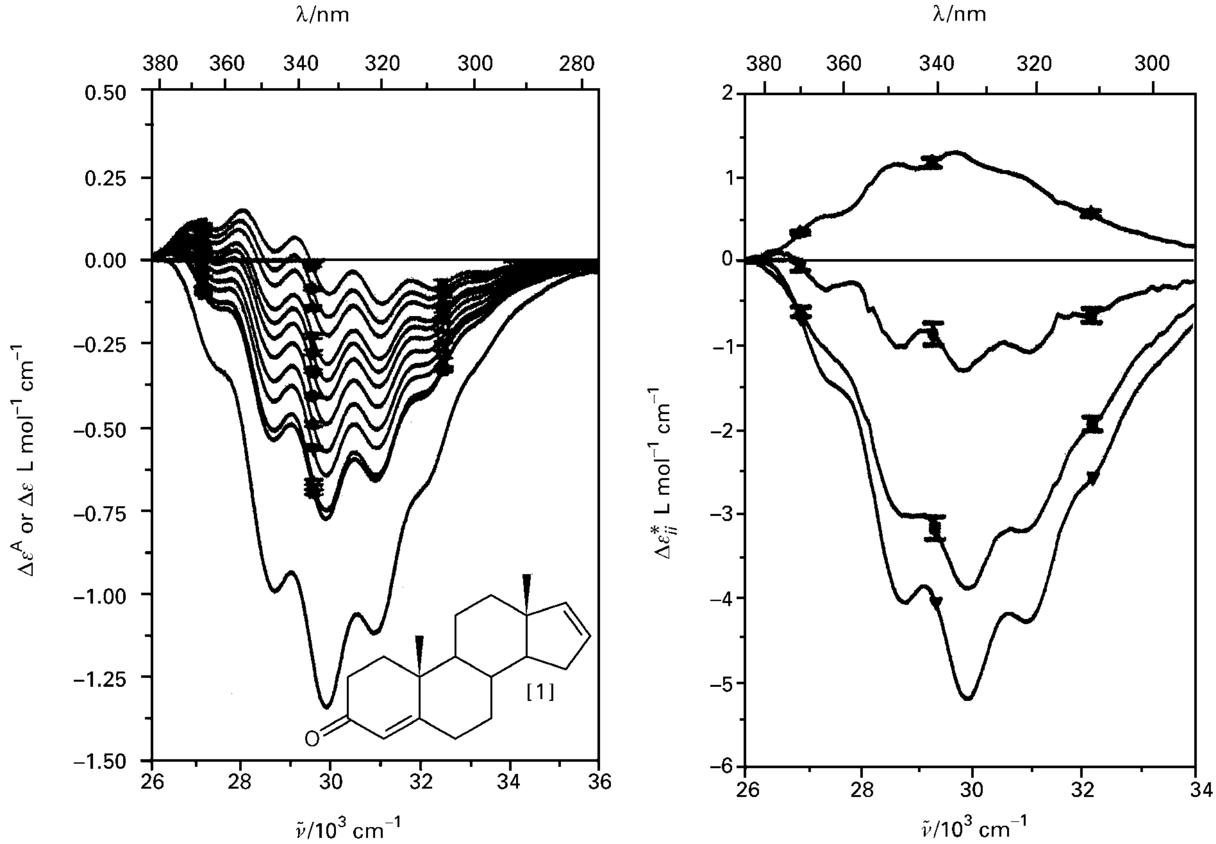



**Figure 7** $\Delta\varepsilon$ (——) ($T = 75°C$) and temperature-dependent $\Delta\varepsilon^A$ (from above $T$/°C = 31.2 to 64.0) (left) and the tensor coordinates $\Delta\varepsilon^*_{11}$ (■), $\Delta\varepsilon^*_{22}$ (●), $\Delta\varepsilon^*_{33}$ (▲) and $\Delta\varepsilon^*_{11} + \Delta\varepsilon^*_{22}$ (▼) (right) of androst-4,17-dien-3-one [1] in the 'nematic' liquid crystal phase ZLI-1695 (Merck).

experimentally the circular dichroism $\Delta\varepsilon_{ij}$ or the rotational strength tensor $R_{ij}^{NnKk}$ with respect to their own principal axes. For a molecule of $D_2$ symmetry, the orientation of the principal axes of $R_{ij}^{NnKk}$ is symmetry determined.

## Isotropic phases with suprastructural phase chirality

With ACD and CD spectroscopy, information can be obtained about suprastructural chirality of isotropic phases, e.g. cubic phases with suprastructural chirality. There is no interference with linear effects and thus $\Delta\varepsilon$ and $\Delta\varepsilon^A$ can easily be measured. A well-known example is crystalline $NaClO_3$. Another known example is that of liquid crystalline blue phases ($O^2$, $O^8$). Both with light in the region where $\lambda$ is of the order of the periodicity within the phase and with wavelengths very different from the periodicity, a CD spectrum is obtained from which a detailed analysis of the phase structure has been derived, e.g. by Hornreich and colleagues.

## Anisotropic phases with suprastructural phase chirality

### Chiral anisotropic phases of chiral molecules

Anisotropic phases are often stabilized by chiral structures. Before discussing phenomena presented by such structures, the question must be answered of how chirality information about molecular chirality and suprastructural phase chirality is to be obtained from chirality measurements of anisotropic phases. Ruch and Schönhofer defined the notion of chirality measurement for an isotropic system: 'a chirality measurement of a property is a measurement with an ensemble of non-oriented molecules and this measurement leads for enantiomers to a result of the same value but with an opposite sign'. Extending this definition to anisotropic systems one might say that 'a chirality measurement of a property with an ensemble of oriented molecules, i.e. the anisotropic phase, is a measurement with an ensemble of non-oriented supermolecules – i.e. an ensemble of "anisotropic phases" – and this measurement leads for an enantiomer of such an ensemble of supermolecules to a

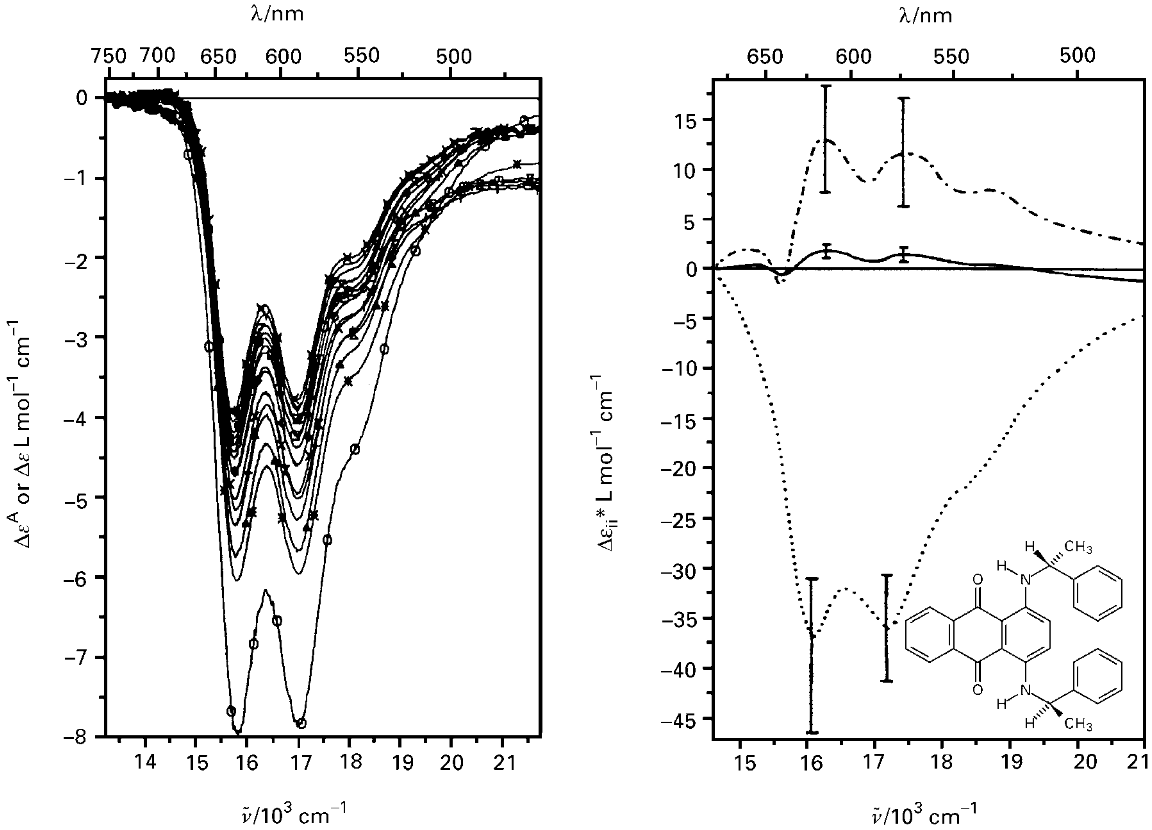



**Figure 8** $\Delta\varepsilon$ (O)($T$ = 75°C) and temperature-dependent $\Delta\varepsilon^A$ (from above $T$/°C = 28 to 68) of 1,4-bis($R$)-1-phenylethylamino)-9,10-anthraquinone [2] (left) and the tensor coordinates $\Delta\varepsilon^*_{11}$ (–•–) $\Delta\varepsilon^*_{22}$ (•••), $\Delta\varepsilon^*_{33}$(——) (right) of the charge transfer transition in the 'nematic' liquid crystal phase ZLI-1695 (Merck).

result of the same value but with an opposite sign.' Furthermore, for oriented molecules the anisotropy of chirality measurements has been demonstrated with ACD spectroscopy. In analogy to molecular properties, one must expect to find an anisotropy for any chirality measurement of a chiral anisotropic phase. Considering the chiral anisotropic phase as a supermolecule, as done before, a chirality measurement with CD spectroscopy is a measurement with the phase rotated about the direction of the propagating light beam analogously to **Figure 4**. Do such definitions hold in general? One question arises immediately from the knowledge that, for oriented achiral molecules of symmetry $C_s$, $C_{2v}$, $D_{2d}$ and $S_4$ with suitable achiral positional and orientational order, nondiagonal elements of the rotational strength tensor $R_{ij}^{Nn}$ can lead to a CD effect. This effect has been shown by Hobden using optical rotation measurements with crystals ($\bar{4}2m$ or $\bar{4}$, for example) without enantiomorphism. Is this an effect of chirality? Not really, because the CD is zero if the above definition of a chirality measurement/observation is applied.

For a homogeneous sample, the phenomenological constants $a_2$ and $b_2$ are responsible for the circular effects (Eqns [9] to [13]). But what does 'homogeneous' mean in a sample possessing positional and orientational order? Whether a phase represents a homogeneous or an inhomogeneous sample depends on the method used for the analysis. For an optical measurement, as for example with a cholesteric phase or a smectic C* phase, this depends on the chosen wavelength of the light, i.e. on whether the wavelength is close to or far from the dimension of the periodicity of the phase. In the first case Equations [9] to [13] are not valid because no 'internal reflection', originating from a periodicity of the inhomogeneity, has been taken into account. Berreman's theory and its extension for samples with very small pitch are discussed by Oldano and colleagues. The optical constants $a_1$, $a_3$, $b_1$ and $b_3$ that are responsible for local linear birefringence and dichroism are then according to the de Vries model the main origin for optical effects caused by the chirality of the sample.

Equations [9] to [13] integrated over the periodicity of a helical phase can also be applied to describe a

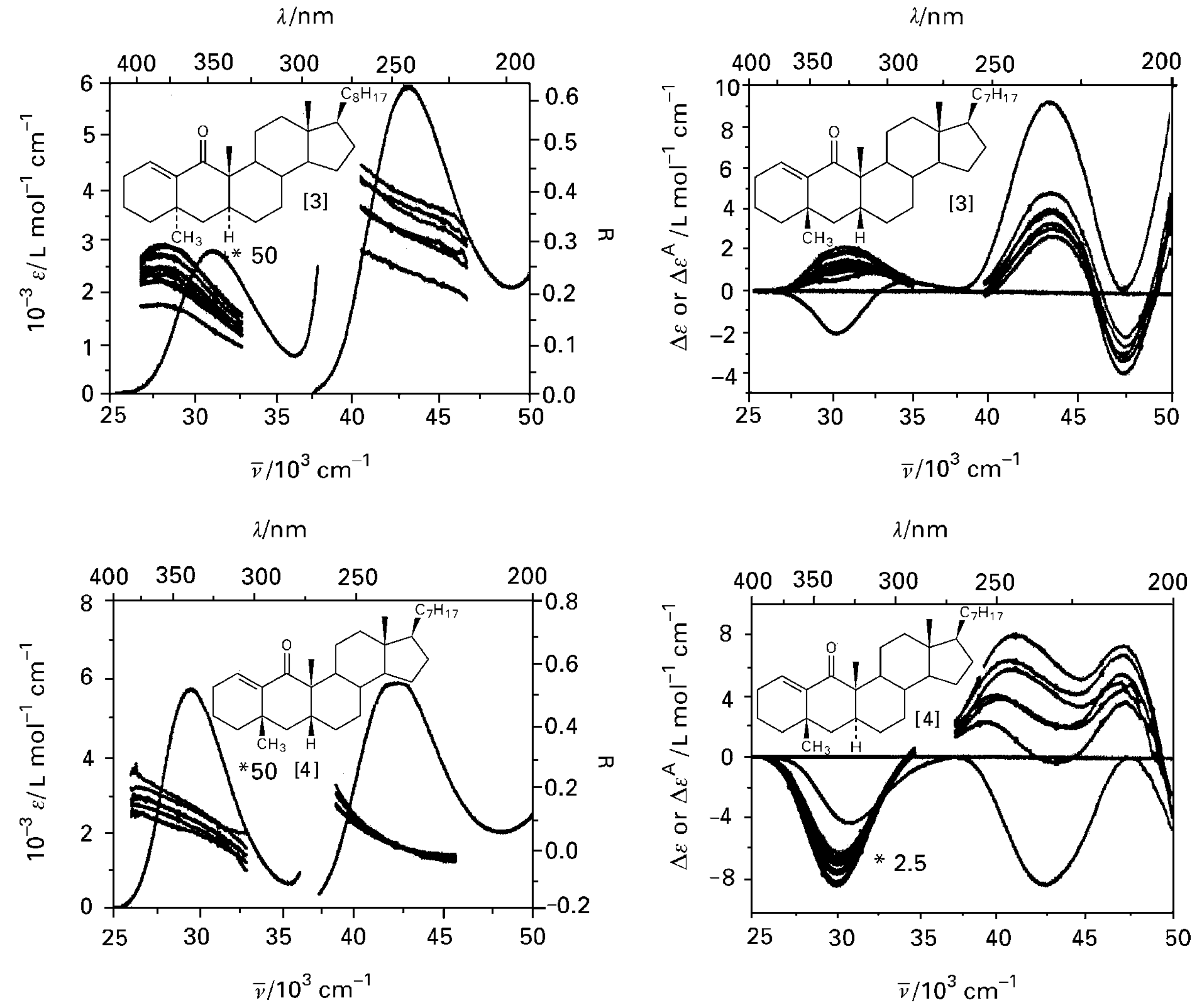



**Figure 9** UV spectra ($T = 75°C$), temperature dependent degree of anisotropy ($T/°C = 28$ to 67.5) (left), $\Delta\varepsilon$ ($T = 75°C$) and temperature dependent $\Delta\varepsilon^A$ ($T/°C = 28$ to 67.5) (right) of the cisoid enones 3-methyl-3α,5α,3,3′,4′,5′-tetrahydrobenzo[2,3]cholest-2-en-1-one [3] (above) and its 3β-diastereomer [4] (below) in the 'nematic' liquid crystal phase ZLI-1695 (Merck).

circular dichroism often termed 'liquid crystal induced circular dichroism' (LCICD) in cases where the pitch $p >> \lambda$ or $p << \lambda$. The CD originates from the helical arrangement of the absorbing molecules owing to their orientation in an anisotropic sample. The molecules and thus the chromophores dissolved in the anisotropic helical phase that are responsible for the CD can be chiral or even achiral. The LCICD is always magnitudes larger than the molecular CD in isotropic or anisotropic systems without suprastructural chirality. LCICD is found, for example, in cholesteric and chiral smectic C* liquid crystalline phases as shown by Gray. The sign of the LCICD provides information about the handedness of the phase.

Comparatively few optical investigations of chiral structures in anisotropic phases are known for two reasons. First with CD spectrometers based on polarization modulation, an artificial CD is often measured because of incorrect superposition of linear and circular effects in the signal processing. Secondly, the method for the decomposition of a measured quantity into the chiral and achiral part is not well developed. In the case of small linear and circular effects it is sufficient to expand the exponential (Eqn [11]) up to the first order. The linear approximation for F(A′) (Eqn [9]) can be applied and the contributions of linear and circular effects are additive, as for example for polymer or liquid crystalline polymer films. By measuring the state of polarization as a function of the azimuth of the sample, the CD originating from a chiral structure in the films can be evaluated. The expansion is, in general, not appropriate, because the linear effects are three to five orders of magnitude larger than the circular effects.

### Chiral anisotropic phases of achiral molecules

In inorganic salts such as $NaClO_3$ discussed above, or quartz, chiral structures are built up from achiral species. An interesting field has been opened by the analysis by Takezoe and colleagues of the so-called 'banana shaped' molecules, which build up chiral domains together with phase separation, although, as is generally believed, they consist of achiral molecules. The analysis of such chiral phases by CD or ACD spectroscopy opens up a new field for an analysis of suprastructural chirality.

## List of symbols

$\mathbf{A}$ = differential Mueller matrix; $c$ = concentration (mol $L^{-1}$), velocity of light; $d$ = path length (cm); $D^*$ = order parameter $\sqrt{3}(g^*_{2233} - g^*_{1133})/2$; $D^{NK}_{ij}$, $D^{NnKk}_{ij}$ = dipole transition moment tensor; $f(\alpha,\beta,\gamma)$ = orientational distribution function; $\mathbf{F}$ = Mueller matrix; $F^{NnKk}$, $G^{NnKk}$ = spectral functions for $|Nn\rangle \rightarrow |Kk\rangle$; $g_{ijkl}$ = orientational distribution coefficients; $h$ = Planck constant; $\mathbf{I}$ = identity matrix; I = light intensity; $[M]$ = molar rotation; $M$ = molecular mass (g $mol^{-1}$); $n$ = refractive index, number density; $N_A$ = Avogadro constant; $q$ = concentration (g/100 $cm^3$); $R^A$ = anisotropic rotational strength of a transition; $R^{NK}_{ij}$, $R^{NnKk}_{ij}$ = rotational strength tensor; $\underline{S}$ = Stokes vector; $S^*$= order parameter = $(3g^*_{3333} - 1)/2$; $X_{ij}$, $Y_{kl}$ = tensors of second rank for the molecules and the phase, respectively; $\alpha$, $\beta$, $\gamma$ = Euler angles; $[\alpha]$ = specific rotation; $\Delta\varepsilon$ = circular dichroism; $\Delta\varepsilon_{ij}$ = circular dichroism tensor; $\varepsilon$ = absorption coefficient; $\varepsilon_{ijk}$ = Levi–Civita tensor; $[\theta]$ = molar ellipticity; $\mu = \kappa$, $\lambda$ = roots of the characteristic polynomial of Det $(\mathbf{A}' - \mu\mathbf{I}) = 0$; $\bar{\nu}$ = wavenumber; $\rho$ = density (g $cm^{-3}$); $\phi$, $\alpha$ = optical rotation in rad, degree.

*See also:* **Chiroptical Spectroscopy, Emission Theory; Chiroptical Spectroscopy, General Theory; NMR in Anisotropic Systems, Theory; Raman Optical Activity, Theory; Vibrational CD, Theory.**

## Further reading

Azzam RMA and Bashara NM (1977) *Ellipsometry and Polarized Light*. Amsterdam: North-Holland.

Berreman DW (1972) Optics in stratified and anisotropic media: 4×4-matrix formulation. *Journal of the Optical Society of America* **62**: 502–510.

Demus D, Goodby J, Gray GW, Spiess H-W and Vill V (eds) (1998) *Handbook of Liquid Crystals*. Weinheim: Wiley VCH.

De Rossi U, Dähne S, Meskers SCJ and Dekkers HPJM (1996) Spontane Bildung von optischer Aktivität in J-Aggregaten mit Davydov-Aufspaltung. *Angewandte Chemie* **108**: 827–830.

Frelek J, Szczepek WJ, Weiss HP, Reiss GJ, Frank W, Brechtel J, Schultheis B and Kuball HG (1998) Chiroptical properties of the cisoid enone chromophore. *Journal of the American Chemical Society* **120**: 7010–7019.

Guo J-X and Gray DG (1995) Induced circular dichroism as a probe of handedness in chiral nematic polymer solutions. *Liquid Crystals* **18**: 571–580.

Hobden MV (1967) Optical activity in a non-enantiomorphous crystal silver gallium sulphide. *Nature* **216**: 678.

Hornreich RM and Shtrikman S (1983) Theory of light scattering in cholesteric blue phases. *Physical Review A* **28**: 1791–1807.

Katzenelson O, Zabrodsky Hel-Or H and Avnir D (1996) Chirality of large random supramolecular structures. *Chemistry – a European Journal* **2**: 174–181.

Kuball H-G, Schultheis B, Klasen M, Frelek J and Schönhofer A (1993) Circular dichroism of oriented molecules. Electric dipole/magnetic dipole and electric dipole/electric quadrupole contribution for cholest-4-en-3-one. *Tetrahedron: Asymmetry* **4**: 517–528.

Luckhurst GR and Gray GW (eds) (1979) *The Molecular Physics of Liquid Crystals*. London: Academic Press.

Mezey PG (ed) (1991) *New Developments in Molecular Chirality*. Dodrecht: Kluwer.

Michel J and Thulstrup EW (1995) *Spectroscopy with Polarized Light*. New York: VCH.

Nakanishi K, Berova N and Woody RW (eds) (1994) *Circular Dichroism – Principles and Application*. New York: VCH.

Oldano C and Rajteri M (1996) Optical activity of small-pitch helical-shaped dielectric media. *Physical Review B* **54**: 10 273–10 276.

Quack M (1989) Structure and dynamics of chiral molecules. *Angewandte Chemie International Edition in English* **28**: 571–586.

Samori B and Thulstrup EW (eds) (1988) *Polarized Spectroscopy of Ordered Systems*. NATO ASI Series, Series C: Mathematical and Physical Sciences 242. Dordrecht: Kluwer.

Schönhofer A and Kuball H-G (1987) Symmetry properties of the Mueller-matrix. *Chemical Physics* **115**: 159–167.

Schönhofer A, Kuball H-G and Puebla C (1983) Optical activity of oriented molecules IX. Phenomenological Mueller matrix description of thick samples and of optical elements. *Chemical Physics* **76**: 453–467.

Sekine T, Niori T, Watanabe J, Furukawa T, Choi SW and Takezoe H (1997) Spontaneous helix formation in smectic liquid crystals comprising achiral molecules. *Journal of Materials Chemistry* **7**: 1307–1309.

Snatzke G (ed) (1967) *Optical Rotatory Dispersion and Circular Dichroism in Organic Chemistry*. London: Heyden.

Zabrodsky H and Avnir D (1995) Continuous symmetry measures. 4. Chirality. *Journal of the American Chemical Society* **117**: 462–473.

# Chromatography-IR, Applications

**George Jalsovszky**, Hungarian Academy of Sciences, Budapest, Hungary

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

## Introduction

A chromatography–IR spectroscopy combination can be approached from two angles. The method can be seen as (i) a chromatograph with a specific IR-spectroscopic detector, to ensure selective detection or (ii) as an IR spectrometer preceded by a chromatographic device to provide samples of appropriate purity. In any case, the essence is a separation and subsequent identification of pure samples. Thus, a chromatography–IR spectroscopy combination is applicable to any analytical problem in which the components of a composite sample (mixture) are to be identified, and it can be expected that a chromatographic method can be applied successfully for separation and IR spectroscopy can give at least pieces of information for identification.

Before the introduction of FT-IR instruments, the spectroscopic measurement was always carried out off-line, separately from the chromatographic process, with the insertion of an often-tedious sample-preparation step, which limited the applicability. The multiplex and throughput advantages of FT-IR spectroscopy have significantly improved the sensitivity and reduced the time demand of IR-spectroscopic measurement, making it possible to perform chromatographic separation with simultaneous, on-line IR-spectroscopic detection and identification, opening up thereby an extremely wide field of application. The chromatographic methods that offer the possibility of on-line spectroscopic detection are gas chromatography (GC), high-performance liquid chromatography (HPLC) and supercritical-fluid chromatography (SFC). Thin-layer chromatography (TLC) can also be combined 'on-line' with IR spectroscopy insofar as spectra can be taken directly from the chromatographic plate.

Chromatographs and FT-IR spectrometers can be interfaced by two distinct approaches. In the first, the effluent from a chromatographic column is passed through a flowcell (often called 'light pipe' in GC/FT-IR) together with the chromatographic carrier, where infrared spectra are measured continuously. In the second approach the carrier is removed, the eluates are deposited on a suitable substrate (as such or isolated in a matrix), and the infrared measurements are made some time after deposition. Thus, when discussing the applications of GC/FT-IR, HPLC/FT-IR and SFC/FT-IR techniques, flowcell techniques and deposition techniques will be treated separately. TLC/FT-IR can be accomplished *in situ* or by transferring the eluates from the plate to an IR-transparent material, but these techniques will be discussed only briefly in a single section.

## GC/FT-IR – flow cell (light pipe) approach

This method is applicable to any sample that can be evaporated in the injector of the GC, separated into its components on the column, and which, mixed with the carrier gas, remains in the gas phase and suffers no decomposition in the heated transfer line and the flowcell (light pipe) of the GC/FT-IR interface. If this is not the case, HPLC or SFC should be used for separation. However, this restriction still leaves an extremely wide choice of possible applications, of which only some examples are discussed below.

### Industrial applications

In industrial processes starting materials, intermediates and end products are regularly analysed, in which GC/FT-IR may have its due share. For example, solvents are recycled after distillation or other purification steps. The purity of solvents is checked before recycling by gas chromatography, where eventual impurities, once known, are identified on the basis of retention times. In the case of unknown impurities GC/FT-IR or GC/MS methods must be used for identification, of which the former is particularly useful if aromatic isomers are to be distinguished. Another important area is monitoring reactions like catalytic oxidation or polymerization.

### Petrochemistry

The distinction of aliphatic isomers and homologues in crude oil and gasoline fractions by IR spectroscopy is a hard task. However, the method is very powerful if the identification of aromatic compounds in oil fractions is required. The discrimination ability of

IR spectroscopy is also high when the substitutional isomers of polycyclic aromatic hydrocarbons (PAH) are to be identified. The method is also applicable in the analysis of antioxidants used in oils and fuels.

### Pollutants in the environment

One of the first applications of GC/FT-IR was, in 1979, the identification of alkyl-9-fluorenones in diesel exhaust particulate extracts. Some years later in the EPA laboratory of Athens, GA, USA and then in the EPA laboratory of Las Vegas, GC/FT-IR was applied to identify toxic substances in environmental-sample extracts. In the latter laboratory, the minimum identifiable quantities of 30 priority pollutants and 25 other environmental contaminants were determined. The method was shown to be applicable for the analysis of hazardous waste and industrial wastewater extracts. The gas-phase spectra of 47 compounds of environmental importance were also taken, and their characteristic absorption frequencies and specific absorbances were published.

Important groups of pollutants are pesticides, fungicides and insecticides. Their toxicity has made it necessary to monitor their presence in the food chain, soil or water. Analytical and monitoring techniques first relied on the GC analysis of a limited number of target compounds. However, in the identification of other residues that were suspect but not among the preset target compounds, IR-spectroscopic information often proved to be vital. There are three problems making this analysis difficult: the large number of components, the presence of a number of conformers or isomers with widely varying degree of toxicity or biochemical activity, and the possibility of the formation of new degradation products. GC/FT-IR is suitable for tackling all the three, being particularly useful in the differentiation between isomers (often superior to MS) and in the identification of degradation products.

Of the other pollutants, polychloro-dibenzodioxins have attracted special attention, where a crucial point is the differentiation between isomers, since some of them are extremely poisonous. Of these compounds the vapour-phase IR spectra of 22 tetrachloro-dibenzodioxins (TCDDs) were taken to prove that the TCDD isomers have unique IR spectra. Indeed, spectra–structure correlations showed that each TCDD isomer has characteristic, individual asymmetric and symmetric COC stretching frequencies, on the basis of which they can be safely distinguished at low-microgram concentrations from other, less-toxic isomers.

GC/FT-IR' can also be used to identify pesticide degradation products, including low-volatility pesticides and their residues in biological samples. For example, phenols and 2,4-dichlorophenoxyacetic acid derivatives were analysed in body fluids of humans exposed to pesticides. More recent topics of analysis are complex marine mixtures, nitrated polycyclic aromatic hydrocarbons in environmental samples, aromatic isomers in marine samples and fatty acid methyl esters in the marine environment.

### Natural products of vegetable origin (flavours, fragrances and essential oils)

Essential oils are complex mixtures for which conventional GC techniques (retention index) are insufficient for complete identification. The need for the information supplied by the infrared spectrum justifies here the use of GC/FT-IR. In a similar manner to pesticides, particular flavours are often associated with one of the possible several isomers or conformers.

For example, the essential oil of *Anthemis nobilis L.* is a mixture of terpenoids and saturated and unsaturated $C_4$ and $C_5$ acids esterified by saturated or unsaturated $C_3$ to $C_6$ alcohols. GC/MS with electron-impact ionization fails to distinguish the angelates and tiglates of unsaturated alcohols, whereas these geometric isomers have widely different characteristic frequencies in the infrared. The same holds for the distinction between crotonates and methacrylates or between linear and branched ethers. A further interesting example is the case of α-pinene and γ-terpinene with very similar MS but sufficiently specific IR spectra. GC/FT-IR has been used to analyse a number of other essential oils, like the flavour components in kiwi, peppermint oil, lemon oil or the essential oil extracted from celery seed. An analogous task was the identification of 43 components in the headspace of coffee samples.

### Biomedical applications

GC/FT-IR can be used in biomedicine to detect and identify medicaments, toxic compounds and narcotics as well as their metabolites in biological fluids (blood, urine, bile, etc.). Of course, samples must be pre-treated prior to GC/FT-IR analysis to have relatively clean extracts. The metabolic pathways of drugs or potential drugs can also be established by analysing biological fluids for the decomposition products. An interesting example is the metabolism of trifluoromethylethyl benzhydrol, an effective hepatic enzyme inducer. Many of its metabolites are polyhydroxy compounds, the isomers of which could be distinguished only by GC/FT-IR.

In toxicology it is very important to identify the product that is responsible for fatal poisoning. A

GC/FT-IR method has been elaborated, e.g. for the identification of paraquat, a herbicide, and for the diagnosis of fatal poisoning due to the vapours of volatile solvents (acetone, diethyl ether, trichloroethylene, etc.) used by young drug addicts. In the analysis of narcotics and drugs of abuse (heroin, cannabis, LSD, cocaine, amphetamines, etc.) GC/FT-IR is essential in identifying the numerous isomers, and also applicable in the analysis of barbiturates or stimulants abused as drugs. By identifying impurities, the 'fingerprint' of drugs can be determined, which can serve as a basis of checking the geographical origin of drugs of abuse.

### Polymer analysis

GC/FT-IR is highly applicable in the study of polymers. On heating the sample the polymer is degassed or (partly or completely) decomposed, and thus gases are produced. Pyrolysis can be used to determine the structure of polymers on the basis of the resulting fragments. An example is the pyrolytic study of polybutadiene in which a number of monosubstituted vinyl compounds can be distinguished on the basis of their characteristic C–H out-of-plane frequencies. A similar approach was used in the analysis of elastomers applied in the automobile industry and of epoxy–phenolic resins. Pyrolysis-GC/FT-IR is applicable in the study of the thermal degradation of thermoresistant amide, imide and diimide polymers. Although not directly related to polymer chemistry, it is noted that substructure components of various aquatic humic substances can also be identified by this method. Thermogravimetric analysis (TGA) can also be combined with GC/FT-IR: an example is a study of the decomposition of ethylene–vinyl acetate copolymers by TGA/GC/FT-IR analysis.

### Miscellaneous applications

GC/FT-IR can be used for amino acid profiling with *tert*-butyldimethylsilyl derivatization, for the investigation of food-packaging materials, the analysis of biologically active wheat straw extracts with allopathic activity, that of *cis*- and *trans*-octadecanoic acids in margarines and of sugar units in disaccharides, for the identification of small molecular migrants extracted from γ-irradiated plastic laminates or the radiolysis products of γ-irradiated medazepam solution.

## GC/FT-IR – subambient solute trapping (direct-deposition) approach

In this direct-deposition (DD) technique, the eluates are deposited on a moving ZnSe window cooled by liquid nitrogen, and the spectra are measured continuously, a short time after deposition. The resulting spectra are very different from gas-phase spectra, being more similar to spectra of samples prepared as a KBr disk or Nujol mull; thus they may be searched against standard condensed-phase reference libraries, as was shown for the example of barbiturates (barbital, apobarbital, butabarbital, phenobarbital). This is an advantage over matrix isolation (MI) detection (see below). Detection limits for on-line measurements are almost two orders of magnitude lower than in the case of a flowcell interface, and the limits can be pushed down by post-run spectrum recording, which is easy to perform with DD interfaces.

An interesting application of this technique was the analysis of β-agonists in cattle. These compounds are a group of *N*-phenylethanolamines illegally used in the treatment of veal and cattle. In this study tissue samples were analysed for the presence of clenbuterol, mabuterol and salbutamol, all of which were trimethylsilylated before analysis. The limit of detection was ~1 to 2.5 ng $\mu L^{-1}$. Another important application is in forensic toxicology, where the compound must be identified by at least two independent methods. Here, GC/DD/FT-IR is a good second technique in addition to GC/MS. A characteristic application is the distinction between amphetamines and metamphetamines in post-mortem analyses.

## GC/FT-IR – matrix-isolation (MI) approach

The isolation and structure elucidation of complex mixtures require high-resolution chromatographic separation coupled with advanced spectroscopic techniques. This is offered by matrix isolation (MI), based on mixing the eluate with a noble gas (e.g. argon) and depositing the mixture on a metal surface cooled to a temperature of 10 K, where the eluate molecules are trapped in a matrix of the noble gas. The analysis of essential oils may require such an approach. For instance, peppermint oil contains (–)-menthol as the main component causing a cooling sensation, whereas other isomers (neomenthol, isomenthol and neoisomenthol), though structurally similar, have no such effect. Mass spectrometry is unable to differentiate between these isomers; however, the MI/FT-IR spectra show characteristic differences (see **Figure 1**). Soya bean oil is an important vegetable oil, in which the presence of unsaturated fatty acids causes instability. GC/MI/FT-IR can be applied successfully to determine the presence and structure of any unsaturated species. In essential oils used for aroma, flavour and fragrance purposes, stereochemical

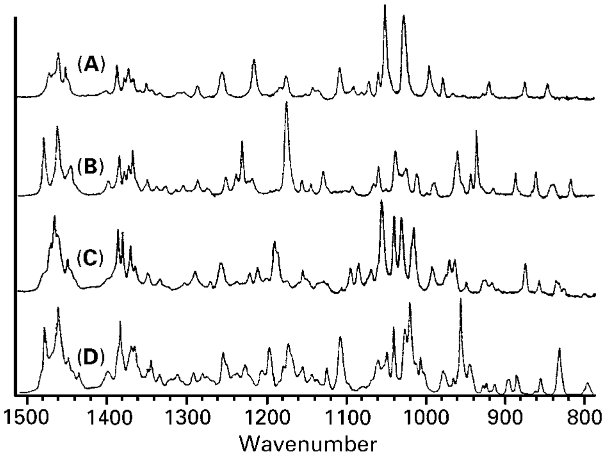



**Figure 1** MI/FT-IR spectra (fingerprint region) of (A) menthol, (B) neomenthol, (C) isomenthol and (D) neoisomenthol, 25 ng each. Reproduced with permission of the Society for Applied Spectroscopy from Coleman III WM and Gordon BM (1989) Examinations of the matrix isolation infrared spectra of organic compounds: Part XIII. *Applied Spectroscopy* **43**: 303.

conformations and configurations play an important role in sensory responses. For instance, geraniol and nerol in coriander oil differ only in the *trans* vs. *cis* conformation of the double bond, which is readily distinguished on the basis of the MI/FT-IR spectrum.

Due to the low-nanogram sensitivity of the GC/MI/FT-IR technique, minor components of natural essential oils can be detected and safely identified on the basis of spectral differences due to positional isomers. Similarly to the light pipe approach, or even more so, this technique is fitted to the identification of the isomers of polycyclic aromatic hydrocarbons (PAH), e.g. in soil samples. Picogram levels of polychlorinated biphenyls and chlorinated dibenzo-*p*-dioxins can be determined as well as nitrocresols in air-sample extracts.

## HPLC/FT-IR – flowcell approach

The first flowcell experiment was reported in 1975, and some years later in 1979 Vidrine reviewed the most important points of this technique, limited at that time to size-exclusion chromatography (SEC). The use of flowcell methods with normal- and reversed-phase chromatographic separations has proven more difficult, since for a successful coupling the spectroscopic properties of the solvent system should meet some minimum requirements. In this respect deuterated solvents seemed to be promising, and numerous applications have been reported on normal-phase, flowcell HPLC/FT-IR, in which deuterated or halogenated solvents were used. Hydrocarbons, esters, ketones, phenols, amines and azarenes were easily separated and detected. To increase chromatographic performance, small amounts of polar modifiers can be added to the mobile phase without seriously affecting IR-detection ability. However, gradient elution, or traces of water, buffer or alcohol proved to be incompatible with the flowcell method. To overcome this problem, the aqueous effluent can be extracted with an IR-compatible solvent and the resulting solution measured in the flowcell.

Using microbore LC columns, a separation method was developed for a series of weakly basic cyclic and noncyclic secondary amines, which were identified as components of coal-derived solvents. As shown in another study, using normal HPLC on silica gel with Freon-113 elution, model mixtures of aliphatic and aromatic hydrocarbons and nonpolar constituents can be separated in coal-liquefaction process solvents. Encouraging results were obtained on both semipreparative (4.6 mm i.d.) and microbore (1 mm i.d.) columns. According to further reports, carbamate pesticides, polymer additives and solvent-refined coal were analysed, and the components of bergamot oil can be identified. GPC/FT-IR can be used to detect components of cold-rolling oil and to analyse polymers, whereas SEC/FT-IR can be applied to the analysis of coal liquids and to improve detection and identification of proteins.

## HPLC/FT-IR – solvent-elimination approach

With the removal of the mobile phase, IR detection no longer restricts the selectivity of the chromatographic process through solvent-transparency requirements. Solvent elimination is more difficult with reversed-phase chromatography because of the presence of water. This problem can be overcome by extraction with organic solvent, and deposition of the organic solution in a known manner, or by a post-column reaction with 2,2-dimethoxypropane, converting water into methanol and acetone, which have sufficient volatility for the deposition technique.

In the first realization of solvent-elimination HPLC/FT-IR the solution eluting from the column is concentrated in a short heated tube, then passed dropwise onto a cup filled with KCl powder, held in a carousel. After the evaporation of solvent, the IR spectrum can be taken in diffuse reflection (DR). In another approach (called 'buffer memory') the effluent is deposited on a KBr plate as a narrow, continuous band, and IR spectra are measured in transmission. This method was illustrated for the example of polystyrene and Carbowax 6000. According to a

third technique, using microbore columns, eluates are again deposited on KCl, and IR spectra are measured in DR. The performance was investigated on a test dye solution (Butter Yellow and Indophenol Blue), and on a mixture of nitrobenzene and 4-chloronitrobenzene.

It is worth noting that over the 15 years of the history of HPLC/FT-IR only a limited number of real applications have been reported in scientific papers, most of them dealing with improvements on the technique or investigations on its performance. Of the few applications, the analysis of chloropyrene isomers and congeners, organics in water and polymer-composition distribution can be mentioned.

## SFC/FT-IR – flowcell approach

The development of SFC/FT-IR has closely paralleled that of HPLC/FT-IR. In both cases the first demonstration of the technique was the flowcell approach, and the solvent-elimination techniques (using transmission and diffuse-reflection arrangements) appeared slightly later. The flowcell approach in SFC appeared to be superior to the same in HPLC for two basic reasons: (i) the preferred SFC mobile phases transmit IR radiation over a larger frequency range than many HPLC solvents, and (ii) density programming in SFC poses fewer difficulties in solvent subtraction than the solvent-gradient technique used in HPLC analysis.

The choice of a suitable mobile phase is subject to the following conditions: (i) minimum absorption in the mid IR, (ii) adequate solvation of the molecules of interest in SFC, (iii) high solute diffusivities to permit high column efficiencies and (iv) a high degree of selectivity as a function of pressure variation. Whereas the most common candidate, supercritical carbon dioxide, fulfils many of these conditions, its polarity and solvent strength are only moderate. However, adding small amounts of more polar compounds to carbon dioxide may result in increase of solvent strength without substantially compromising IR measurement (although some authors are of different opinion). Still, even without modifiers, the absorption of carbon dioxide may cause problems, for instance in the identification of hydroxy compounds (the OH stretching region coincides with a strong carbon dioxide absorption band). Thus, some of the authors prefer using xenon as the mobile phase, which has no IR absorption.

The first demonstration of SFC/FT-IR coupling was reported in 1983 with the use of a wide-bore fused-silica capillary column, supercritical carbon dioxide and a flowcell of 1 cm pathlength. With this setup, easily identifiable spectra can be obtained for 3 µg each of anisole, acetophenone and nitrobenzene.

An interesting application is the analysis of ultraviolet-curing inks and coatings. These formulations contain oligomer, reactive diluent and photoinitiator, where the oligomers are low- to medium-molecular weight multifunctional, unsaturated materials. SFC/FT-IR is an ideal method for their analysis, since it allows high-resolution separation of mixtures at low temperatures, while the components are readily identified on the basis of their characteristic IR spectra. This is shown, e.g., by the analysis of the volatile components of citrus oil and that of model steroid mixtures, pyrethrins (naturally occurring esters with insecticidal activity). Typical identification limits, as studied with caffeine as model, are in the 2 ng range. In a commercial preparation of pyrethrin extract, the six insecticidally active components were separated by SFC, since thermal degradation was observed when attempting GC separation. IR spectra of the components taken in the flowcell clearly reveal the structural differences between them. Practical applications of SFC/FT-IR include the determination of non-volatile compounds from microwave susceptor packaging that may migrate into the heated food. Another application is the analysis of fibre finishes on fibre/textile matrices.

## SFC/FT-IR – solvent-elimination approach

In contrast to flowcell interfaces, solvent-elimination approaches lead to spectra free of solvent interferences. Various sampling techniques are possible: the sample can be deposited on a flat ZnSe plate, on a smooth metallic substrate or a thin layer of powdered alkali halide salt, whereas the spectrum can be taken in conventional transmission, external-reflection or diffuse-reflection arrangement. In one of the first applications a synthetic mixture of three quinones was separated on a microbore column packed with silica, using a mobile phase of 5% methanol in supercritical carbon dioxide. The peaks were collected on a plate on which a layer of KCl powder was deposited, and then spectra were measured by a diffuse-reflection accessory. Test measurements on acenaphthenequinone (AQ) showed later that conventional transmission spectra of samples on flat infrared-transmitting windows give the best compromise between high sensitivity, correct relative band intensities and adherence to the Lambert–Beer law.

The most up-to-date device is a real-time (on-line) SFC/FT-IR interface based on a commercial direct-deposition GC/FT-IR system, described by Norton

and Griffiths, which allows spectra of subnanogram quantities of analytes to be measured over the complete mid-IR region. It is possible to get identifiable spectra from injections as low as 600 pg for strong IR absorbers and between 1 to 10 ng for weaker IR absorbers. The interface also allows for usage of polar modifiers in the mobile phase.

## TLC/FT-IR

This combination can be approached in two ways: either the FT-IR measurement is performed *in situ*, directly on the TLC plate, or the analytes are transferred from the plate to an IR-transparent substrate prior to the FT-IR measurement. Interfaces for both methods are now commercially available. NIR and photoacoustic FT-IR detection have also been applied, although to a lesser extent, in the coupling of TLC with IR spectroscopy.

## Conclusion

It is important to note that scientific papers or reviews are not the most appropriate sources of information on the application of chromatography-IR spectroscopy combinations. These techniques, like most of the techniques of instrumental analysis, are widely applied just as an analytical tool in research laboratories and in industry. Typical accounts of these applications are progress reports, technological prescriptions or patent applications, rather than scientific papers.

When considering the application of chromatography-IR spectroscopy combination techniques, three questions should be answered: (i) is the problem the analysis of a composite sample? (ii) can the sample be separated by a chromatographic method? (iii) is there a chance that the IR spectrum of the components will help identification? If the answer is 'yes' to these questions, a chromatography-IR spectroscopy combination is a promising approach to the analytical problem.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Chromatography–IR, Methods and Instrumentation; Fourier Transformation and Sampling Theory; Industrial Applications of IR and Raman Spectroscopy; IR Spectral Group Frequencies of Organic Compounds; IR Spectrometers; IR Spectroscopy Sample Preparation Methods; Medical Science Applications of IR; Polymer Applications of IR and Raman Spectroscopy; Quantitative Analysis.**

## Further reading

Coleman III WM and Gordon BM (1994) Analysis of natural products by gas chromatography/matrix isolation/infrared spectrometry. In: Brown PR and Grushka E (eds) *Advances in Chromatography* **34**: 57–108.

Fuoco R, Pentoney L and Griffiths PR (1989) Comparison of sampling techniques for combined supercritical fluid chromatography and Fourier transform infrared spectrometry with mobile phase elimination. *Analytical Chemistry* **61**: 2212–2218.

Griffiths PR, Pentoney SL, Giorgetti A and Shafer KH (1986) Hyphenation of chromatography and FT-IR spectrometry. *Analytical Chemistry* **58**: 1349A–1366A.

Herres W (1987) *HRGC-FTIR: Capillary Gas Chromatography–Fourier Transform Infrared Spectroscopy: Theory and Applications*. Heidelberg: Hüthig Verlag.

Lacroix B, Huvenne JP and Deveaux M (1989) Gas chromatography with Fourier transform infrared spectrometry for biomedical applications. *Journal of Chromatography* **492**: 109–136.

Norton KL and Griffiths PR (1995) Performance characteristics of a real-time direct deposition supercritical fluid chromatography–Fourier transform infrared spectrophotometry system. *Journal of Chromatography A* **703**: 503–522.

Ragunathan N, Krock KA, Klavun C, Sasaki TA and Wilkins CL (1995) Review: Multispectral detection for gas chromatography. *Journal of Chromatography A* **703**: 335–382.

Somsen GW, Morden W and Wilson ID (1995) Review: Planar chromatography coupled with spectroscopic techniques. *Journal of Chromatography A* **703**: 613–665.

Taylor LT and Calvey EM (1989) Supercritical fluid chromatography with infrared detection. *Chemical Reviews* **89**: 321–330.

Vidrine DW (1979) Liquid chromatography detection using FT-IR. In: *Fourier Transform Infrared Spectroscopy – Applications to Chemical Systems, Vol. 2*, pp 129–164. New York: Academic Press.

# Chromatography-IR, Methods and Instrumentation

**Robert L White**, University of Oklahoma, OK, USA

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Methods & Instrumentation**

Linking chromatographic separation methods with IR spectroscopic detection of separated components results in powerful chemical analysis tools. As with any hyphenated analysis method, combining mixture component separation with infrared spectroscopic detection provides capabilities that surpass those of the constituent techniques. Unlike conventional chromatography detectors, IR analysis provides detailed structural information regarding complex mixture components. In favourable cases, mixture components can be unequivocally identified based on their chromatographic retention times and IR spectra. The performance characteristics of chromatography-IR combinations are largely determined by the interface employed. The ideal chromatography-IR interface would link the two techniques without sacrificing the performance of either. In practice, instrument performance tradeoffs are unavoidable. A variety of interfaces have been developed in attempts to minimize these tradeoffs when gas chromatography, liquid chromatography, supercritical fluid chromatography, and thin layer chromatography are combined with IR spectroscopic detection.

Inherent in all chromatography-IR combinations is the need for rapid IR spectrum measurement. Fourier transform IR spectrophotometry (FT-IR) provides rapid scanning and is the preferred IR spectroscopic technique for detection of species separated by chromatographic methods. Instrumental requirements for chromatography-IR interfaces depend on many factors. In particular, the physical state of matter that must be analysed determines which spectroscopic sampling techniques can be employed. For example, gas chromatography detectors must provide a means of analysing flowing vapours. Liquid and supercritical fluid chromatography detection requires analysis of flowing liquids. IR detection of thin layer chromatography solutes presents the particularly difficult problem of analysing species adsorbed on infrared opaque substrates.

## Chromatogram generation

Regardless of the type of chromatography implemented or the specific chromatography-IR interface employed, separated mixture components are ultimately represented by a plot of detector response as a function of separation time or migration distance. There are two common procedures for converting structure-specific IR spectroscopic information into chromatogram plots. The Gram–Schmidt vector orthogonalization method uses subtle variations in interferogram data acquired during FT-IR scans to detect solute elutions. The functional group chromatogram method is more computationally intensive and requires interferogram Fourier transformation and calculation of absorbance spectra, but can be used to elucidate mixture component structural features.

For Gram–Schmidt chromatogram generation, a portion of each acquired FT-IR interferogram is selected to represent a multidimensional vector. The number of interferogram data points used to form vectors determines the dimensionality of the vector space and thus the number of orthogonal axes needed to define the space. Vectors derived from interferograms obtained when no solute is present in the IR beam are similar and are confined to the same region of the vector space. By using the Gram–Schmidt orthogonalization method, a few of these background vectors are used to calculate orthogonal axes that are representative of the portion of the vector space in which background vectors are found. Chromatogram generation is accomplished by comparing the orientation of measured interferogram vectors with the background subspace. Interferograms acquired when solute is in the IR beam have vector components outside of the background subspace. The sum of vector residuals remaining after removing those components that lie in the background space are plotted as a function of elution time or migration distance to yield a chromatogram. Gram–Schmidt chromatogram intensity is therefore a measure of how much interferogram vectors differ from the background subspace, a quantity not necessarily related to IR absorbance. The signal-to-noise ratio of Gram–Schmidt generated chromatograms depends on the number of interferogram points used to define vectors, the number of orthogonal axes used to specify the background subspace, and the interferogram segment used to create vectors.

Functional group specific chromatograms are generated from chromatography-IR absorbance spectra by plotting the integrated absorbance within selected wavelength regions as a function of elution time or migration distance. Significantly more calculations are required to create functional group specific chromatograms than to generate Gram–Schmidt chromatograms. However, functional group specific chromatograms provide much more information regarding separated species structures than Gram–Schmidt chromatograms. Typically, multiple wavelength regions are integrated to produce several functional group specific chromatograms. When separated species contain functionalities that absorb within a selected wavelength region, a peak representing the elution of this substance will appear in the chromatogram. Total IR absorbance chromatograms are generated by integrating spectra over the entire mid-infrared range (e.g. 4000–400 $cm^{-1}$). The total IR absorbance chromatogram contains peaks representing all infrared absorbing solutes.

## Gas chromatography-IR

The most commonly employed chromatography-IR combination is between gas chromatography (GC) and FT-IR. Interfaces for GC-IR can be categorized as either flow cell (light pipe) or solute trapping. **Figure 1** illustrates the three GC-IR interfaces described in the following sections.

### Light pipe GC-IR

The first on-line GC-IR interface consisted of a heated flow-through gas cell or 'light pipe'. Simple and inexpensive light pipe interfaces constitute the most common method employed for combining gas chromatography with infrared spectrophotometry. In these combinations, IR radiation from an FT-IR source is directed through a light pipe, which is a glass tube with a gold-coated inner surface through which the gas chromatographic effluent flows. The light pipe contains no chromatographic stationary phase. Therefore, it does not contribute to mixture component separation. In fact, the light pipe dead volume results in degradation of chromatographic resolution. To minimize chromatographic resolution loss, small light pipe volumes are employed. Typical GC-IR light pipes used with capillary gas chromatography columns have volumes of about 100 µL. To maximize IR radiation flux and optical path length, light pipes with diameters of about 1 mm and lengths of 12–15 cm are commonly employed. Light pipe ends must be sealed in order to conduct GC effluent without loss. This is accomplished by attaching IR transparent alkali halide windows to the light pipe ends. IR transparent gases such as He and $N_2$ are employed for the gas chromatographic mobile phase so that acquired spectra represent separated mixture components only. Vapour phase IR spectra obtained by light pipe GC-IR contain sharp spectral features because solute–solute interactions are minimal. When small molecules are detected at sufficiently high spectral resolution, vibration bands may exhibit rotational fine structure.

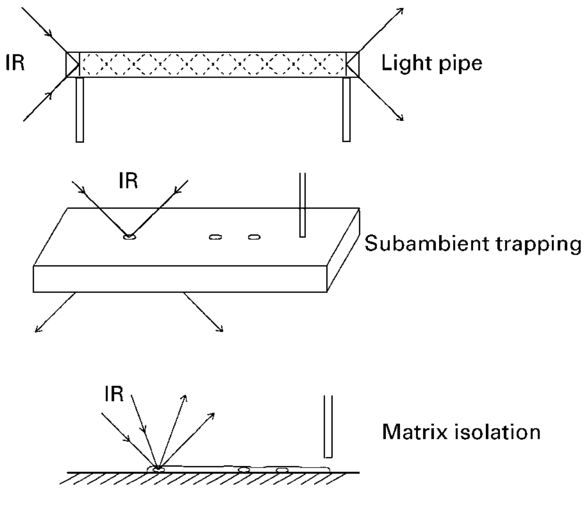



**Figure 1** Gas chromatography-IR interfaces.

### Subambient solute trapping GC-IR

Mixture components can be removed from GC effluent by trapping them on a cold surface. Subambient solute trapping GC-IR interfaces take advantage of the large differences between the volatilities of separated mixture components and GC carrier gas. When GC effluent is directed onto a surface maintained at a temperature at which mixture components condense but carrier gas does not, mixture components can be isolated from the GC carrier gas stream. In practice, mixture components are deposited on a cooled IR transparent window that is continuously moved during the chromatographic separation so that eluting solutes are deposited at different locations. IR analysis is achieved by measuring transmittance spectra of the frozen solutes. In contrast to light pipe GC-IR, trapped solute infrared measurements are not restricted to the time required for species to elute from the chromatographic column. Extended signal-averaging can be used to improve spectrum quality because IR analysis of frozen solutes is conducted off-line. Unlike light pipe GC-IR, subambient solute trapping GC-IR yields condensed phase IR spectra.

Condensed phase spectra contain broader features than vapour phase spectra and reflect significant solute–solute interactions.

### Matrix isolation GC-IR

Matrix isolation GC-IR is a variation of the subambient solute trapping technique that yields significantly different spectral features. Rather than freezing mixture components, molecules are trapped within an inert argon matrix. Interactions between solute molecules are virtually eliminated by the argon matrix. As a result, spectra contain much sharper features than those obtained by subambient solute trapping without matrix isolation. Instead of depositing solutes on a cooled infrared transparent window, matrix isolated species are formed on a reflective gold surface under vacuum and maintained near 10 K. Argon matrix gas is provided by using a pre-mixed GC carrier gas composed of helium and about 1% argon. Helium is not condensed at 10 K and is removed by the vacuum system. The argon and separated mixture components condense on the cryogenic surface. Because argon is present in the carrier gas at much higher concentration than mixture components, molecules deposited on the surface are surrounded by condensed argon. The argon matrix isolates individual solute molecules and prevents solute–solute interactions. IR spectra for separated mixture components are measured in a similar manner to that described for subambient solute trapping. However, reflectance rather than transmittance techniques are employed for IR analysis. Matrix isolation IR spectra exhibit sharp features because solute–solute interactions and molecular rotations are absent.

## Liquid chromatography-IR

Compared to GC-IR, liquid chromatography (LC)-IR measurements are often more challenging because mobile phases are not necessarily infrared transparent. Due to the fact that mobile phase concentration greatly exceeds that for mixture components in LC effluent, even mobile phases that have weak IR absorptivities can present significant problems. LC-IR interfaces can be categorized as flow cell or mobile phase elimination. **Figure 2** illustrates the three LC-IR interfaces described in the following sections.

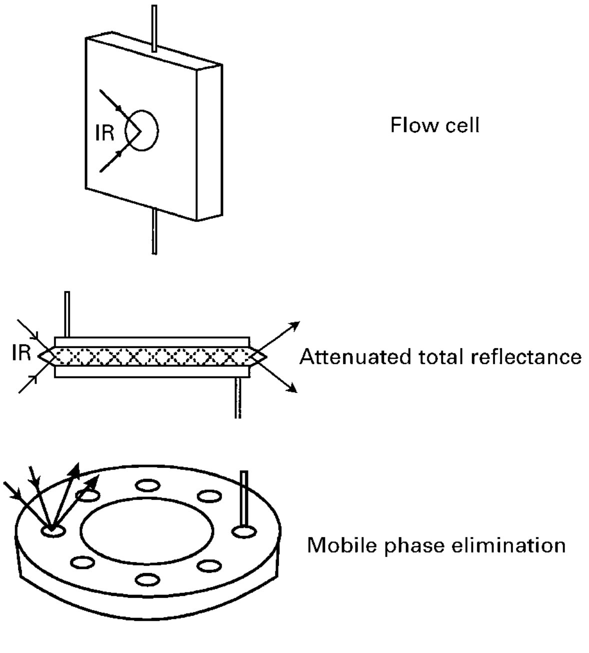



**Figure 2** Liquid chromatography-IR interfaces.

### Flow Cell LC-IR

LC-IR flow cells consist of two IR transparent windows separated by a spacer with a thickness that determines the cell path length. Alkali halide windows are commonly used with non-polar LC mobile phases whereas zinc selenide windows are used with polar mobile phases. Unlike GC-IR, optimum flow cell path lengths for LC-IR are dictated by the absorptivity of the mobile phase. Path lengths of 0.1–0.2 mm are commonly employed with moderately absorbing mobile phases. For highly absorbing mobile phases (e.g. water), path lengths of about 10 μm are often required. Mixture component IR spectra are obtained from flow cell measurements after subtracting contributions from the mobile phase.

Aqueous solution IR analysis by flow cell LC-IR is particularly difficult because of the need for extremely short flow cell path lengths. Short path length flow cells are easily clogged and internal pressure fluctuations caused by mobile phase movement through the cell can appreciably change the IR beam path length. Fortunately, attenuated total reflection (ATR) provides an alternative method for attaining short path length IR spectrum measurements. ATR flow cells consist of a high refractive index infrared transparent material through which the IR beam is passed and around which LC mobile phase flows. When the incident IR beam angle at the ATR material/mobile phase boundary exceeds the critical angle, it is totally reflected. At each internal reflection, the electric field of the IR radiation penetrates slightly into the LC mobile phase. This 'evanescent wave' can be absorbed by species dissolved in the mobile phase. The penetration depth of IR radiation at each reflection varies with wavelength and is typically on the order of μm. By using appropriate length ATR materials,

multiple reflections can be achieved to provide reproducible IR beam path lengths of 10–20 μm. Thus, by using ATR, short flow cell path lengths can be attained without the disadvantages of clogging and pressure fluctuations suffered by conventional flow cell designs.

### Mobile Phase Elimination LC-IR

One method of avoiding the problems associated with the fact that most LC mobile phases absorb IR radiation is to remove these solvents prior to IR analysis. Like subambient solute trapping GC-IR, this approach exploits the differences between solute and mobile phase volatilities. By heating the end of the LC column, the mobile phase can be vaporized and removed by aspiration. Mixture components can then be deposited onto an IR transparent substrate. In early interface designs, solutes were deposited on IR transparent potassium chloride powder contained in sample cups. In this system, multiple sample cups are contained in a rotating carousel. The carousel is rotated during LC separation to collect individual mixture components as they elute from the LC column. IR analysis is performed off-line by using diffuse reflectance optics. Like solute trapping GC-IR interfaces, there are no limits to the extent of signal averaging that can be applied to measure mobile phase elimination LC-IR mixture component spectra.

Column end heating can be an effective method for removing organic mobile phases such as *n*-hexane, but is not very effective for reverse phase separations that employ aqueous mobile phases. One method used to remove solutes from aqueous mobile phases is by extracting them into an organic solvent followed by solute deposition on an IR transparent substrate and solvent evaporation. However, on-line continuous solvent extraction greatly complicates LC-IR interfaces.

### Microbore LC-IR

Although most LC separations are conducted with analytical scale columns, microbore LC columns provide better resolution and are becoming increasingly popular. The quantity of mobile phase needed for a separation by microbore LC is much less than that required for analytical scale LC separations. In addition, optimal flow rates for microbore LC are significantly lower than for analytical scale LC. These features of microbore LC simplify both flow cell and mobile phase elimination approaches to LC-IR. In addition, because small quantities of mobile phase are employed for microbore LC separations, expensive mobile phases that may not otherwise be practical can be employed. For example, the C–H stretching region of solute IR spectra can be accessed in reverse phase flow cell LC-IR measurements by employing deuterated solvents such as $D_2O$ and $CD_3CN$ as mobile phases. Because small amounts of mobile phase are employed for microbore LC-IR, less solvent must be removed to achieve mobile phase elimination. Rather than using a carousel of sample cups, a single IR transparent window (or reflective surface) that is moved beneath the column exit can be employed to collect mixture components.

## Supercritical fluid chromatography-IR

Supercritical fluid chromatography (SFC) is often employed for separations involving non-volatile or thermally labile species that cannot be separated by GC or LC. Supercritical mobile phases (typically $CO_2$) have viscosities and solute diffusivities that are intermediate between those for gases and liquids. Supercritical $CO_2$ is created by subjecting the gas to high pressures. SFC-IR flow cells are made from high pressure UV flow cells by replacing the quartz windows in these cells with infrared transparent windows. Alternatively, light pipe flow cells similar to those employed for GC-IR can be used for SFC-IR. In either case, regions in which $CO_2$ strongly absorbs are opaque in spectra measured by using flow cell SFC-IR. Mobile phase elimination SFC-IR techniques are particularly effective because supercritical $CO_2$ mobile phase readily vaporizes when exposed to ambient conditions. The same methods used for mobile phase elimination LC-IR are applicable to mobile phase elimination SFC-IR. Matrix isolation SFC-IR can be achieved by using the same apparatus employed for GC-IR but with $CCl_4$ instead of argon as the matrix substance. To prevent $CO_2$ condensation, the matrix isolation deposition surface temperature must be maintained at 150 K for matrix isolation SFC-IR.

## Thin Layer Chromatography-IR

Two FT-IR sampling methods are commonly employed to measure IR spectra of mixture components separated by thin layer chromatography (TLC). Diffuse reflectance and photoacoustic IR spectroscopy are techniques that can be employed when sample materials are opaque.

### Diffuse reflectance TLC-IR

Diffuse reflectance (DR) IR analysis can be used to obtain IR spectra of species separated by TLC

without removing them from the TLC plate. TLC plates containing spatially separated mixture components are passed through an optical system that focuses the incident IR beam onto a small spot on the plate and collects the resulting diffusely scattered IR radiation. **Figure 3** illustrates the basic principle. When the focused beam strikes a region of the TLC plate that contains solute, the resulting diffuse reflectance spectrum is characteristic of the solute and the TLC stationary and mobile phases. Solute spectra are obtained after subtracting the IR spectral features of the stationary and mobile phases. Little solute spectral information can be obtained in wavelength regions in which the stationary phase is highly absorbing. In addition, interactions between solutes and the stationary phase often result in absorbance band wavelength and intensity differences relative to spectra measured for the same substance pressed into a KBr pellet.

### Photoacoustic TLC-IR

In contrast to most other IR analysis techniques, the photoacoustic (PA) IR spectroscopic signal is a direct measure of the amount of radiation absorbed by the sample. Consequently, samples need not be transparent or reflective to be analysed by this method. However, TLC separated solutes must be physically removed from the TLC plate prior to analysis because PA-IR requires samples to be sealed in an air-tight chamber. As with diffuse reflectance measurements of TLC separated solutes, PA-IR spectra contain features representing stationary phase and residual mobile phase in addition to the solute. Spectral subtraction is employed to isolate solute spectra from PA-IR spectra. The main disadvantage of PA-IR compared to diffuse reflectance is that the locations of mixture components on the TLC plate must be known before infrared analysis can be performed. When solutes cannot be detected visually, an additional detector must be used to track solute migrations.

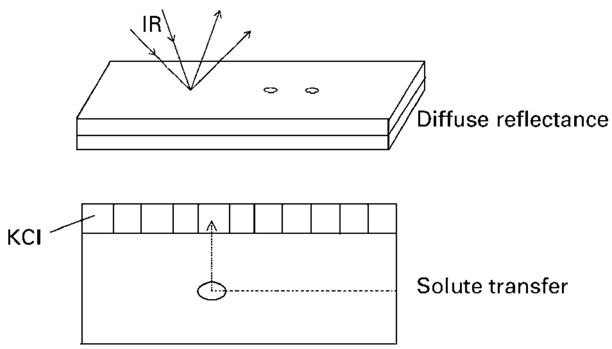



**Figure 3** Thin layer chromatography-IR interfaces.

### Solute transfer TLC-IR

IR spectra of TLC-IR solutes obtained by DR-IR and PA-IR exhibit band intensity distortions in spectral regions in which stationary and mobile phases absorb and may not resemble those obtained for the same substances by KBr pellet transmittance measurements. To avoid stationary and mobile phase interferences, solute transfer methods can be used to physically move separated mixture components from the TLC plate to an infrared transparent substrate. Solute transfer is accomplished by subjecting the developed TLC plate to a second solvent that is applied orthogonally to the direction of the chromatographic mobile phase flow. This causes separated mixture components to move to the edge of the TLC plate where they are collected on IR transparent KCl powder. **Figure 3** illustrates the basic principle. Solute spectra subsequently obtained by diffuse reflectance IR analysis are similar to those obtained by KBr pellet transmittance measurements.

## Chromatography-IR interface characteristics

Compared to mobile phase elimination techniques, flow cell chromatography-IR interfaces are simpler and less expensive. However, flow cell chromatography-IR yields optimal results only when the mobile phase is IR transparent. Detection limits in the low nanogram range can be attained by light pipe GC-IR. Due to mobile phase absorbances, detection limits for flow cell LC-IR and SFC-IR combinations are higher than for light pipe GC-IR. In fact, flow cell LC-IR detection limits vary considerably and are primarily determined by the mobile phase absorptivity.

In general, mobile phase elimination chromatography-IR methods provide lower detection limits than flow cell approaches. The primary goal of mobile phase elimination is to remove interfering absorbances from acquired IR spectra. In addition to removing chromatographic mobile phase, these interfaces concentrate solutes prior to IR analysis, which enhances solute infrared absorbances. Furthermore, off-line IR analysis by mobile phase elimination chromatography-IR facilitates increased signal averaging when compared to flow cell interfaces. When spectral noise sources are random, FT-IR spectrum signal-to-noise ratios increase by the square root of the increase in the number of scans averaged. The advantages of increased solute concentration and extended signal averaging are

evidenced by the fact that typical solute trapping GC-IR detection limits are about ten times lower than those provided by light pipe GC-IR despite the fact that GC-IR mobile phases are infrared transparent.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Fourier Transformation and Sampling Theory; IR Spectrometers; IR Spectroscopy, Theory; Matrix Isolation Studied By IR and Raman Spectroscopies; Photoacoustic Spectroscopy, Applications; Photoacoustic Spectroscopy, Theory; Quantitative Analysis; Surface Studies By IR Spectroscopy.**

## Further reading

Erickson MD (1979) Gas chromatography/Fourier transform infrared spectroscopy applications. *Applied Spectroscopy Reviews* **15**: 261–325.

Fujimoto C and Jinno K (1992) Chromatography/FT-IR spectrometry approaches to analysis. *Analytical Chemistry* **64**: 476A–481A.

Griffiths PR and deHaseth JA (1986) *Fourier Transform Infrared Spectrometry*. New York: John Wiley & Sons.

Herres W (1987) *HRGC-FTIR: Capillary Gas Chromatography–Fourier Transform Infrared Spectroscopy: Theory and Applications*. Heidelberg: Huthig.

Jinno K and Fujimoto C (1996) Microcolumn liquid chromatography in hyphenated techniques (microcolumn LC/FTIR). *Methods for Chromatography* **1**: 199–211.

Kalasinsky VF and Kalasinsky KS (1992) HPLC Detection using Fourier transform infrared spectrometry. In: Patonay G (ed) *HPLC Detection*, pp 127–161. New York: VCH.

McClure GL (1993) General methods of sample preparation for infrared hyphenated techniques. In: Coleman PB (ed) *Practical Sampling Techniques for Infrared Analysis*, pp 165–215. Boca Raton: CRC.

White RL (1987) Gas chromatography–Fourier transform infrared spectrometry. *Applied Spectroscopy Reviews* **23**: 165–245.

White RL (1990) *Chromatography/Fourier Transform Infrared Spectroscopy and its Applications*. New York: Dekker.

Wurrey CJ and Gurka DF (1990) Environmental applications of gas chromatography/Fourier transform infrared spectroscopy (GC/FT-IR). *Vibrational Spectra and Structure* **18**: 1–80.

# Chromatography-MS, Methods

**WWA Niessen,** hyphen MassSpec Consultancy, Leiden, The Netherlands

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
Methods & Instrumentation

## Introduction

Chromatography is an important tool in purification and isolation of compounds from (complex) mixtures prior to mass spectrometric analysis. Mass spectrometry can be used to detect the compounds separated by chromatography. This can be performed off-line, i.e. after fraction collection, but an even more powerful and versatile technique is achieved when the separation is coupled on-line to mass spectrometric detection. The on-line combination of a chromatographic separation method and mass spectrometry (MS) is an important and versatile tool in many areas of analytical chemistry. In principle, it enables the mass spectrometric characterization of components in complex mixtures after separation, with minimal or no sample losses. Both major and minor components in a mixture are amenable to such a technique. Furthermore, MS offers advantages in quantitative analysis, i.e. enhanced selectivity by the use of high-resolution MS or tandem MS (MS/MS) and/or the use of stable-isotopically labelled internal standards, which optimally mimic the properties of the analytes during sample pretreatment, separation and detection.

Initially, in the mid 1960s the on-line combination of gas chromatography (GC) and MS was developed, first for packed columns and soon after their introduction, also for capillary columns. The introduction of quadrupole mass analysers greatly facilitated GC–MS, because of their faster scanning and less stringent vacuum restrictions. They are now the most widely applied mass analysers for on-line chromatography-MS. The success of GC–MS initiated in the

mid 1970s, research into the on-line combination of liquid chromatography (LC) and MS. Subsequently, other separation techniques were successfully coupled to MS, e.g. supercritical fluid chromatography (SFC), thin-layer chromatography, and various electromigration techniques like capillary zone electrophoresis (CZE) and electrochromatography. The actual breakthrough of LC–MS as a routinely applicable analytical technique is taking place now in the mid 1990s.

## General instrumentation

The general instrumentation for on-line chromatography–MS consists of a conventional chromatograph, the outlet of which is connected to an interface and a mass spectrometer (see **Figure 1A**). The main function of the interface is to protect the high-vacuum system of the mass analyser from too high gas loads from the chromatograph. Simultaneously, the interface should transfer as much of the analytes eluting from the column as possible to the mass analyser. While initially interfaces were developed to keep as much of the carrier gas or mobile phase out of the high vacuum, the development of LC–MS especially resulted in changes in this function of the interface. In the current LC–MS interfaces, the mobile phase plays an essential role in analyte transfer and ionization.

### Gas chromatography–mass spectrometry

The on-line capillary GC–MS combination can be achieved in two ways, i.e. via direct coupling or via an open-split coupling.

The gas flow through a typical 0.25-mm-ID capillary GC column matches the allowable gas load of a simple single-stage MS vacuum system. Since the components eluting from the GC column are already in the vapour state, immediate introduction of the capillary column effluent into the MS ion source in electron ionization mode is possible. Although frequently applied, this direct coupling has some disadvantages. The column outlet is at high vacuum, which results in changing retention times relative to other GC detectors, such as the flame ionization detector. Furthermore, the gas load to the ion source changes during the temperature programme of the GC analysis. This may influence the tuning of the ion source parameters. The problems related to the effect of the solvent load on the lifetime of the heated filament are easily overcome by switching off the filament during the first few minutes of the analysis.

An alternative to direct coupling is the open-split coupling; a schematic diagram of such a device is shown in **Figure 2**. The column is connected with the ion source via a fixed restriction. A make-up gas flow is provided either to raise the column flow to match the restrictor throughput or to rapidly remove the excess of carrier gas. As a result, the column outlet is at atmospheric pressure, just as in a conventional GC detector. By increasing the makeup flow it is possible to efficiently divert high vapour loads from the solvent away from the MS.

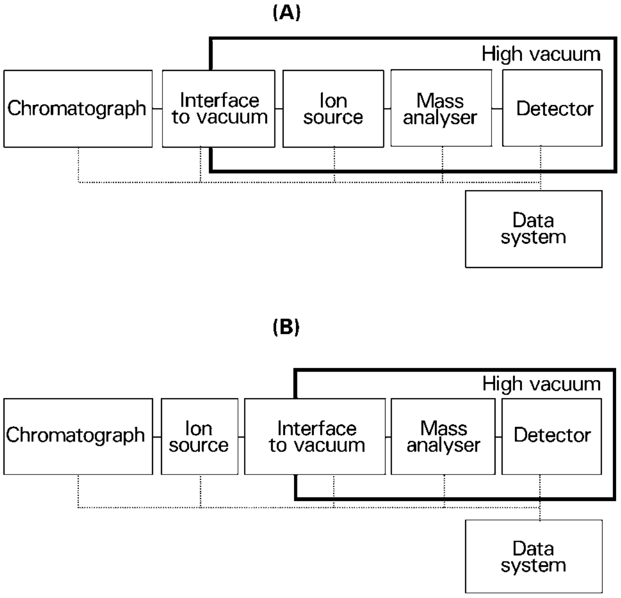



**Figure 1** Schematic diagrams of chromatography–mass spectrometry systems: (A) GC–MS system, (B) atmospheric-pressure ionization LC–MS system.

### Liquid chromatography–mass spectrometry

For a number of reasons, the on-line coupling of LC–MS appeared far more difficult to accomplish than the GC–MS coupling.

1. The gas load to the vacuum system resulting from evaporating the mobile phase in LC is generally much higher than in GC.
2. The mobile phase composition (containing non-volatile additives such as phosphate buffers) is often incompatible with MS.

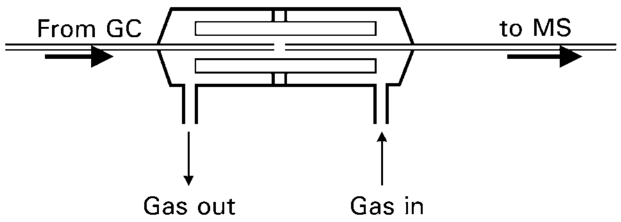



**Figure 2** Open-split coupling for GC–MS.

3. The transfer of polar and ionic analytes from the liquid to the gas phase is difficult.

In order to solve the first problem a number of LC–MS interfaces have been developed, some of which are discussed in more detail below. The second problem must be solved at the chromatography side, e.g. by changing to volatile buffers or the use of column switching. The third problem is largely solved by the advent of new liquid-based soft ionization techniques, especially thermospray and electrospray, which were actually developed in the course of LC–MS development.

A variety of interfaces was developed for on-line LC–MS coupling. A number of these found wide application when they became commercially available, e.g. the moving-belt interface, direct-liquid introduction, thermospray, particle-beam and continuous-flow fast-atom bombardment (Cf-FAB), but are hardly used any longer due to the introduction of interfaces based on atmospheric-pressure ionization (API), i.e. electrospray and atmospheric-pressure chemical ionization (APCI). Before discussing the API-based interfaces in detail, brief attention is given to three of the older LC–MS interfaces.

In a particle-beam interface (**Figure 3A**), the column effluent is pneumatically nebulized into an atmospheric-pressure desolvation chamber. This is connected to a momentum separator where the analytes are transferred to the MS ion source while the low molecular mass solvent molecules are efficiently pumped away. The analyte particles hit the heated source surface, evaporate and can be ionized by electron or chemical ionization. The evaporation step obviously limits the application range of the interface to not-too-polar analytes.

In a thermospray interface (**Figure 3B**), the column effluent is rapidly heated in a narrow bore capillary such that partial (*ca.* 90%) evaporation of the solvent is achieved inside the capillary. As a result, a mist of vapour and small droplets is formed in which the heated droplets further evaporate and ions are generated, either by the thermospray ionization process based on ion evaporation or by solvent-mediated chemical ionization initiated by electrons from a heated filament or a discharge electrode. The excess vapour is pumped away directly from the ion source.

In a Cf-FAB interface (**Figure 3C**), a small mobile-phase stream, containing glycerol as a FAB matrix, is flowing towards a FAB target where the effluent is bombarded by fast atoms or ions to achieve analyte ionization. Non-evaporated mobile phase constituents are collected in a wick, which is a compressed paper pad which has to be regularly renewed.

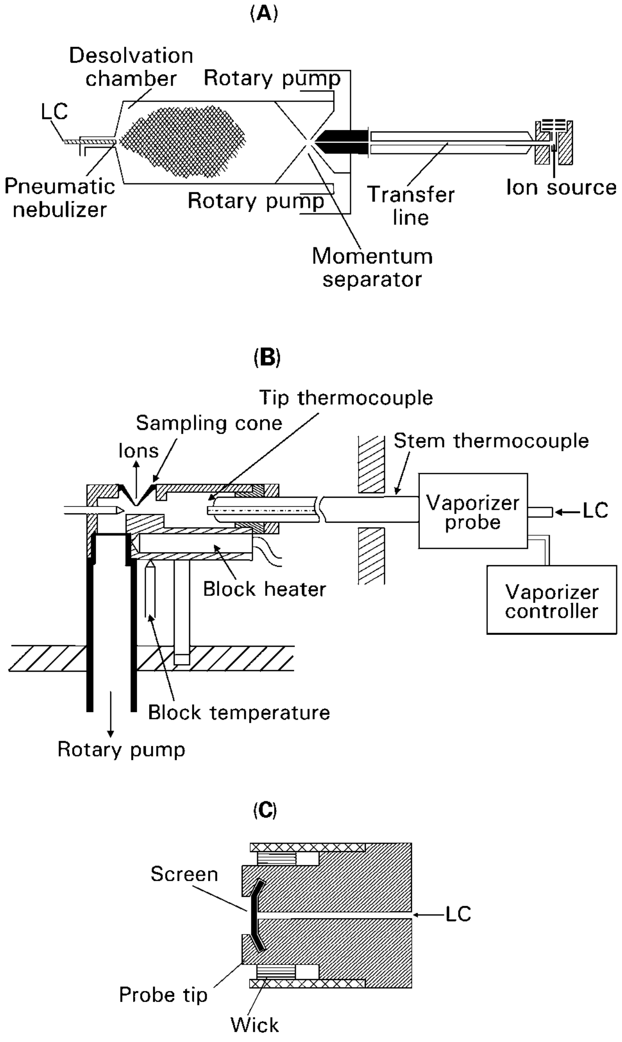



**Figure 3** Older types of LC–MS interfaces: (A) particle-beam interface, (B) thermospray interface, and (C) target area of continuous-flow fast-atom bombardment interface.

## Interfaces based on atmospheric-pressure ionization

After initial attempts in the mid 1970s, LC–MS interfaces based on API were developed in the mid 1980s. An API mass spectrometer differs from a conventional instrument because ions are generated in an atmospheric-pressure region and subsequently introduced into the high vacuum mass analyser. The difference is clearly depicted in **Figure 1B**: the ion source and the interface to vacuum are interchanged. In an API interface, the sample can be introduced in two different ways and ions can be produced in two different ways, i.e. by electrospray or by APCI. After ionization, the ions generated are transferred to the mass analyser via an API interface (**Figure 4**), which is independent of the mode of sample introduction

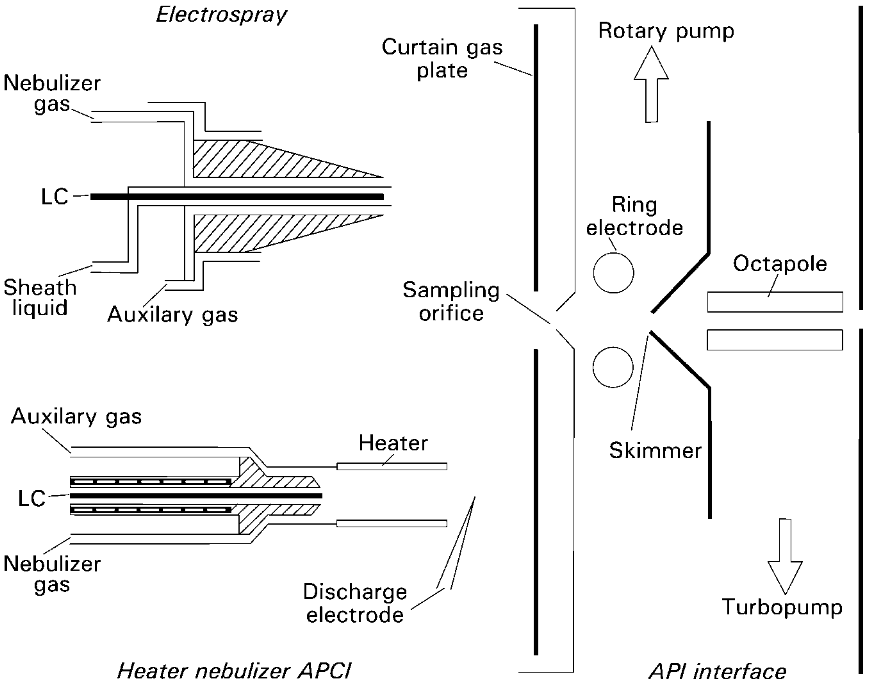



**Figure 4** Atmospheric-pressure interface for LC–MS with electrospray and heated nebulizer APCI spray probes.

or ionization. LC–MS interfaces based on API nowadays are by far the most widely used.

API interfacing thus comprises a number of steps:

1. Nebulization of the column effluent into small droplets followed by their rapid evaporuntion. Nebulization is achieved either by pneumatic nebulization in combination with a heated zone (APCI), or by (pneumatically assisted) electrospray nebulization due to the action of a high electric field on the liquid in a narrow bore needle. Schematic diagrams of the spray probes are shown in **Figure 4**.
2. Generation of ions, either by gas-phase ion–molecule reaction initiated by electrons from a corona discharge needle, where the solvent acts as the reagent gas (APCI), or by ion evaporation of preformed ions in solution (electrospray ionization).
3. Sampling of ions from the atmospheric-pressure region into a two-stage differentially pumped interface, which enables sufficient pressure reduction and transfers the ions to the mass analyser in a high-vacuum chamber.
4. Mass analysis of the ions and subsequent detection.
5. Data acquisition and handling.

## Electrospray interfacing and ionization

The electrospray spray probe typically consists of a 100 µm i.d. stainless-steel or fused-silica capillary surrounded by a coaxial tube for the nebulization gas. The nitrogen nebulization gas allows the introduction of flow rates in excess of 10 µL min$^{-1}$ (pneumatically assisted electrospray or ionspray). The spray probe is positioned in the atmospheric-pressure ion source at a short distance from a counter electrode. A 2–3 kV electric field is applied between the needle and the counter electrode. This results in the formation of a fine aerosol of charged droplets from the spray probe. These droplets are rapidly evaporating, assisted by a concurrent, countercurrent or perpendicular heated nitrogen gas stream, resulting in an increase in droplet charge at the surface. This leads to two competing processes: field-induced Rayleigh or electrohydrodynamic droplet disintegration, leading to an explosion of the initial droplet into a series of highly charged smaller offspring droplets, and field-induced ion evaporation, leading to desorption of preformed ions from the bulk solution into the gas phase. The ions generated are sampled by the mass analyser.

Although current electrospray interfaces can be used with flow rates of up to 1 mL min$^{-1}$, lower flow rates, e.g. 40–200 µL min$^{-1}$, are preferred in most LC–MS applications, thus requiring the use of either a postcolumn solvent splitter or microbore LC column.

Electrospray ionization is especially suited for the analysis of compounds that preform ions in solution, e.g. by acid–base chemistry, i.e. organic acids and

bases, peptide and proteins, oligonucleotides, etc. For analytes with a molecular mass exceeding 500 Da and multiple protonation or deprotonation sites, multiply charged ions may be generated by electrospray ionization. This enables the mass analysis of biomacromolecules by means of a quadrupole mass analyser with a limited mass range. As such, electrospray plays an important role in the characterization of peptides and proteins by MS. Given the limited sample availability in many of these applications, nano-electrospray devices have been developed which allow the introduction of 0.01–1 μL $min^{-1}$ of protein solutions into the MS.

### APCI interfacing and ionization

The APCI probe consists of a liquid introduction capillary surrounded by a coaxial tube for pneumatic nebulization. After pneumatic nebulization of the column effluent, the droplets are swept through a heated zone in order to achieve evaporation of the droplets. The solvent vapours are subsequently used as reagent gas in chemical ionization, initiated by electrons from a corona discharge electrode kept at a few kV. In positive-ion mode, a series of gas-phase ion–molecule reactions leads to the formation of protonated solvent molecules which may react with the analyte molecules having a higher proton affinity than the solvent. The protonated analytes are sampled and subsequently mass analysed.

The APCI spray probe is readily applicable to a wide flow-rate range, e.g. between 0.2 and 2 mL $min^{-1}$. It allows the ionization of analytes in a wide polarity range. It can be used with both typical reversed-phase LC mobile phases (mixtures of water and methanol or acetonitrile) and pure organic solvents. APCI is a highly efficient ionization technique.

### API interfacing in more detail

After the generation of ions in the atmospheric-pressure ion source by either electrospray or APCI, the ions are sampled into the multistage differentially pumped vacuum system. Different ion sampling devices are applied by different instrument manufacturers: a glass capillary with metallized ends, a heated stainless-steel capillary, or a sampling orifice or cone. The low-pressure side of the ion sampling device acts as a nozzle in a molecular-beam source: expansion of the gas–vapour–ion mixture takes place. The core of the expanding jet, primarily containing the constituents with the higher mass, is sampled by a skimmer. In order to preferentially sample the ions in the jet, a ring electrode is positioned between the nozzle and the skimmer. The region between the nozzle and skimmer is kept at a pressure of *ca.* 100 Pa. After sampling by the skimmer, a second expansion stage takes place in the ion optics region, kept at a pressure of 0.1–1 Pa by means of a turbomolecular pump. The ion optics region contains a quadrupole, hexapole or octapole ion guide, which efficiently transports the ions towards the mass analyser, housed in the third pumping stage, kept at a pressure below $10^{-3}$ Pa by means of a second turbomolecular pump.

The transfer of ions from an atmospheric-pressure region through a series of differentially pumped vacuum chambers can be used to collisionally induce fragmentation of these ions in the transition region, e.g. by increasing the potential difference between the nozzle and the skimmer or between the skimmer and the multipole ion guide. This in-source CID approach is successfully applied in structure elucidation of unknown compounds. In order to obtain well-interpretable mass spectra after in-source CID, either pure compounds or well-separated compounds should be used, because (in contrast to CID in an MS/MS instrument) this CID is done without prior precursor-ion selection.

### Mass analysis API interfacing

Although the mass analysis after API interfacing is in most cases performed using a (triple) quadrupole instrument, API devices have also been developed (and are commercially available) for other types of mass analysers, i.e. ion-trap, time-of-flight (TOF), magnetic sector and Fourier-transform ion-cyclotron resonance (FT-ICR) instruments as well as a number of hybrid instruments, such as quadrupole–TOF and magnetic sector–ion-trap instruments. In particular, the combination of API with ion-trap and TOF instruments is receiving considerable attention because of the favourable price to performance ratio of these instruments in a number of applications, compared to triple-quadrupole instruments. The extremely high resolution achievable with FT-ICR can especially be useful in determining charge states of multiply charged protein ions after fragmentation.

## Data acquisition and handling

The mass spectrometer has two general modes of data acquisition:

1. The full-scan acquisition mode, where a series of complete mass spectra is acquired as a function of time.

2. The selected-ion monitoring mode, where the intensity of selected ions with particular *m/z* is acquired as a function of time.

Choosing between these two modes is a matter of striking a compromise in the particular application between maximizing the measurement time of a particular *m/z* (in SIM mode), which is favourable for sensitivity, and the information content (in full-scan mode), which is favourable in identification and/or structure confirmation.

The data acquisition results in a three-dimensional data-array with ion intensity, *m/z*, and time or scan number as the three dimensions. This data-array can be handled in various ways. For the total-ion chromatogram, the ion intensities in each scan are summed irrespective of the *m/z* of the ions and the summed intensity, i.e. the total-ion current, is plotted as a function of time. In this mode the mass spectrometer acts as a universal sector. In GC–MS, the total-ion chromatogram largely resembles the chromatogram obtained via flame-ionization detection. In a base-peak chromatogram, the intensity of the base peak irrespective of the *m/z*, i.e. the most intense peak in each mass spectrum, is plotted as a function of time. This mode can be useful for peak finding in applications with high background levels, e.g. in LC–MS with API interfacing. In a mass chromatogram, the ion intensity for an ion or a series of ions with particular *m/z* is plotted as a function of time. In this mode the mass spectrometer acts as a specific or selective detector. And finally, in the mass spectrum, for each scan number, the ion intensity can be plotted as a function of the *m/z*. Summed, averaged, and/or background-corrected mass spectra over a series of scans can be made as well. The mass spectrum can be computer-searched against a mass spectral library.

## Additional and advanced techniques

The performance of on-line chromatography–MS can for some applications be greatly enhanced by the use of additional techniques. On-line chromatography–MS should be considered as an integrated and hyphenated approach, where all system parts should be optimally tuned for the complete system. The combination of a potentially high-efficiency separation and a potentially high-selectivity detection enables a number of problem-oriented analytical strategies, e.g. directed at identification of minor impurities in complex samples, requiring highly efficient separation combined with MS in full-scan mode, offering the highest information content or directed at high-throughput target compound analysis, where the separation efficiency is deliberately reduced to save analysis time and the high selectivity afforded by either high-resolution MS or tandem MS is applied selectively to detect the compounds of interest within the matrix.

### Analyte derivatization

Capillary GC–MS is an extremely powerful approach, combining the high separation efficiency of the capillary GC column with the identification power of the MS in electron-ionization mode. However the applicability range of GC is limited to relatively volatile compounds. In order to widen the applicability range, precolumn analyte derivatization strategies are often applied to enhance the volatility of the analytes. Methylation, silylation and acetylation reactions are most often applied for the analysis of compounds with amine, (poly-) hydroxy and/or carboxylic acid functional groups. Furthermore, derivatization to pentafluorobenzyl derivatives is applied to enhance the sensitivity of analytes in electron-capture negative-ion chemical ionization.

### Tandem mass spectrometry

A powerful tool to enhance the analytical potential of combined chromatography–MS especially in LC–MS, is tandem MS (MS/MS). MS/MS was first introduced as a way to induce fragmentation of ions generated in the ion source, e.g. by soft ionization methods, and after their mass selection. In such an experiment, the first mass spectrometer selects a particular precursor ion, which is (collisionally) dissociated, and the product ions are analysed with the second mass spectrometer. This is the product-ion scan mode which is primarily used to elucidate the structures of ions.

For the analytical use of MS/MS, other scan modes can be useful as well as the product-ion scan mode. The various modes are explained and summarized in **Table 1**. The precursor-ion and neutral-loss scan modes are particularly useful in screening. They are part of a powerful analytical strategy applicable in impurity, degradation product and/or metabolite profiling, as well as other applications where searching for structurally related compounds is important. After an appropriate ionization mode is found, a product-ion mass spectrum is obtained for the parent compound. In this spectrum, possible common precursor ions and/or common neutral losses have to be identified.

As an example, the screening for unknown metabolites of heptabarbital can be considered (see **Figure 5**). The fragmentation of heptabarbital under

**Table 1**

| *Mode* | *MS1* | *MS2* | *Application* |
|---|---|---|---|
| Product-ion | Selecting | Scanning | To obtain structural information by CID of ions produced in the source |
| Precursor-ion | Scanning | Selecting | To monitor compounds which in CID give an identical fragment |
| Neutral-loss | Scanning | Scanning | MS1 and MS2 are scanning at a fixed *m/z* difference; to monitor compounds that lose a common neutral species |
| Selective reaction monitoring (SRM) | Selecting | Selecting | To monitor a specific CID reaction |

CID is also shown in **Figure 5**. Knowing that most of the metabolism of this drug primarily involves the cycloheptene ring, the barbituric acid related product ion at *m/z* 157 can be used as a common precursor ion to screen for compounds characterized by metabolism of the cycloheptene ring. Alternatively, the neutral loss of 156 Da can be applied in screening.

Subsequently, these precursor-ion and/or neutral-loss scans are performed, preferentially with on-line chromatographic separation. In this way, the *m/z* for related compounds can be found. Product-ion mass spectra are obtained based on these *m/z*. Interpretation of the mass spectra leads to identification of the related compounds, impurities or metabolites. Neutral-loss scans of 176 Da and 80 Da can be applied to screen for O-glucuronide and aryl-O-sulfate conjugates of drugs, respectively.

SRM is extremely useful in quantitative environmental and bioanalyses, where interfering sample matrix components may adversely influence the achievable detection limits. The enhanced confidence level in SRM, based on the monitoring of a specific chemical reaction in the gas phase, made LC–MS/MS in SRM mode the analytical method of choice in quantitative bioanalysis during drug development and clinical trials by pharmaceutical industries. In addition, the monitoring of a specific collision-induced reaction by SRM significantly improves the selectivity, resulting in greatly improved signal-to-noise ratios and enabling a limited sample pretreatment prior to LC–MS/MS analysis, thereby increasing the sample throughput.

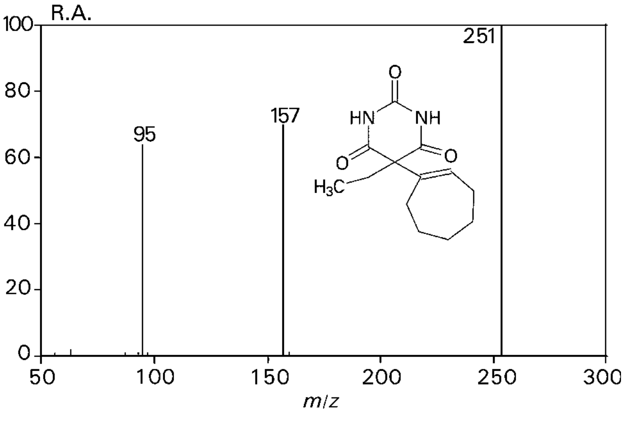



**Figure 5** Product-ion mass spectrum of heptabarbital.

### Coupled column chromatography

One of the major difficulties in LC–MS coupling is that the prolonged use of nonvolatile mobile-phase additives such as phosphate buffers and sodium dodecyl sulfate, is prohibited. In most applications, the nonvolatile additives can successfully be replaced by a volatile alternative, such as ammonium acetate and heptafluorobutyric acid. When this is not possible without significant influence on the chromatographic resolution, the use of column-switching approaches may be useful, especially in target-compound analysis. The chromatography is performed on a first column under optimized chromatographic conditions. Compounds of interest are trapped onto short trapping columns, which are switched in-line during appropriate retention time windows. These loaded trapping columns are subsequently eluted with a mobile phase and/or flow rate readily amenable to the LC–MS interface, either directly to the LC–MS/MS system (phase-system switching) or to a second column (coupled-column chromatography). These strategies may also provide (significant) improvements in selectivity. In this respect, the use of immunoaffinity chromatography as a first step can be extremely useful.

### Capillary zone electrophoresis–MS

LC–MS interfaces have successfully been applied for the coupling to MS of other liquid-based separation techniques, such as SFC, CZE and electrochromatography. The coupling of the highly efficient CZE to MS has achieved most attention because of the high promise of this combination in the analysis of ionic compounds, peptides and proteins. On-line identification of minor impurities made observable by the high separation efficiency achievable in CZE would be extremely useful. On-line CZE–MS is relatively easy to accomplish via an electrospray interface, either via the use of a coaxial sheath liquid in order to make up the sub-μL $min^{-1}$ flow rate from the CZE to the 5–10 μL $min^{-1}$ required in

conventional electrospray interfaces, or via the use of a nano-electrospray interface that can accommodate sub-μL $min^{-1}$ flow rates. However, as a result of the small injection volumes, the concentration detection limits reported are generally in the low ppm range. Various on-line strategies have been proposed to improve these detection limits, e.g. on-line preconcentration via isotachophoresis, solid-phase extraction and/or liquid–liquid electroextraction. In general, these strategies are target-compound directed and therefore appear not to be particularly useful in general impurity profiling.

## Application areas of chromatography-MS

Both GC–MS and LC–MS are widely applied in many fields. As indicated, GC–MS is limited to relatively volatile compounds. Within this restriction, GC–MS is widely applied to the identification and quantitation of compounds of environmental and pharmaceutical interest. In areas such as the characterization of oil fractions, environmental monitoring, e.g. of chlorinated pesticides like DDT and also of chlorinated dibenzodioxins, dibenzofurans and related compounds, and in a number of routine analytical procedures, e.g. in a number of clinical assays, in doping analysis in sports and for forensic purposes, GC–MS shows unsurpassed analytical possibilities, in terms of both sensitivity and identification power.

LC–MS finds wide application in the analysis of compounds that are not amenable to GC–MS, i.e. compounds that are highly polar, ionic and thermolabile, as well as (bio)macromolecules. In environmental applications, LC–MS is applied, often in combination with off-line or on-line solid-phase extraction, to identify pesticides, herbicides, surfactants and other environmental contaminants. LC–MS plays a role in the confirmation of the presence of antibiotic residues in meat, milk and other food products. Furthermore, there is a substantial role for LC–MS in the detection and identification of new compounds in extracts from natural products and the process control of fermentation broths for industrial production of such compounds, e.g. for medicinal use. LC–MS technology is also widely applied in the characterization of peptides and proteins, e.g. rapid molecular-mass determination, peptide mapping, peptide sequencing and the study of protein conformation and noncovalent interactions of drugs, peptides and other compounds with proteins and DNA. However, the most important application area of LC–MS is in the pharmaceutical drug development area, where LC-MS is involved in almost every stage of development, e.g. rapid characterization based on the molecular mass of medicinal combinatorial libraries, the detection and identification of drug metabolites, reaction by-products, and drug degradation products. LC-MS/MS in SRM mode is currently the method of choice in high-throughput quantitative bioanalyses for bioavailability studies, pharmacokinetic and pharmacodynamic studies, and during clinal trials.

*See also:* **Atmospheric Pressure Ionization in Mass Spectrometry; Biochemical Applications of Mass Spectrometry; Chemical Ionization in Mass Spectrometry; Chemical Structure Information from Mass Spectrometry; Fast Atom Bombardment Ionization in Mass Spectrometry; Forensic Science, Applications of Mass Spectrometry; Hyphenated Techniques, Applications of in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Ion Trap Mass Spectrometers; Isotopic Labelling in Mass Spectrometry; Medical Applications of Mass Spectrometry; MS–MS and MS$^n$; Peptides and Proteins Studied Using Mass Spectrometry; Quadrupoles, Use of in Mass Spectrometry; Sector Mass Spectrometers; Thermospray Ionization in Mass Spectrometry; Time of Flight Mass Spectrometers.**

## Further reading

Barceló D (ed) (1996) *Environmental Applications of LC–MS*. Amsterdam: Elsevier Science.

Cai J and Henion JD (1995) Capillary electrophoresis–mass spectrometry, *Journal of Chromatography A* **703**: 667.

Cole RB (ed) (1997) *Electrospray Ionization Mass Spectrometry*. NewYork: Wiley.

Karasek FW and Clement RE (1988) *Basic Gas Chromatography–Mass Spectrometry: Principles and Techniques*. Amsterdam: Elsevier Science.

Kitson FG, Larsen BS and McEwen CN (1996) *Gas Chromatography and Mass Spectrometry, a Practical Guide*. London: Academic Press.

Niessen WMA and Van der Greff J (1992) *Liquid Chromatography–Mass Spectrometry*, New York: Marcel Dekker (second edition, authored by WMA Niessen, will appear by the end of 1998).

Niessen WMA and Voyksner RD (ed) (1998) *Current Practice of LC–MS*, Special Issue of *Journal of Chromatography A*, Vol. 794.

Pinkston JD and Chester TL (1996) Putting opposites together: Guidelines for successful SFC–MS, *Analytical Chemistry* **67**: 650A.

# Chromatography-NMR, Applications

**JP Shockcor**, Du Pont Pharmaceuticals Co. Newark, DE, USA

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

High-performance liquid chromatography (HPLC) has become a routine tool for the separation of complex mixtures. However, structural information on substances separated using HPLC is limited by the detector system employed. The most common HPLC detectors, refractive index, UV, radiochemical, fluorescence, and electrochemical, provide little or no structural information. As a result, structure elucidation required the isolation of the analyte from the matrix, followed by off-line spectroscopic characterization. With the advent of HPLC-MS the ability to detect and identify substances at low concentrations without the need for an isolation step became possible. Although this has simplified structure elucidation to a great extent, there are often circumstances where HPLC-MS alone is insufficient for complete characterization of a compound and further studies by NMR are required. Logically the next step in instrument development would be directly coupling HPLC and NMR yielding the hyphenated technique HPLC-NMR.

## History

In the late 1970s and early 1980s, a number of attempts to couple these techniques were carried out. However, these studies suffered from the low sensitivity of the NMR spectrometer systems then available. Also, because of dynamic-range problems, there was a need to use expensive deuterated solvents for the HPLC because the solvent suppression methods in use at the time could not cope with fully protonated solvents. The reduction of HPLC-NMR to routine use was slow in developing and not practically achieved until technical improvements in electronics, higher magnetic fields strengths, advanced solvent suppression sequences, and improved instrumental sensitivity, made it feasible to interface an HPLC directly to an NMR spectrometer.

In addition, investigations into coupling other types of chromatography with NMR have been carried out. These include capillary electrophoresis, capillary electrochromatography, supercritical fluid chromatography and capillary HPLC. However there are no commercially available systems based on these separations (late 1999).

## Technical improvements

There have been many major technical improvements in NMR sensitivity in recent years. Foremost has been the development of ultra-high-field-strength magnets (by late 1999, 900 MHz). High dynamic range analog–digital converters, which give benefits in situations where large solvent peaks are present along with small analyte signals, and digital filtering techniques that allow spectral windows to be limited to include only the region of interest without the problem of signals and noise from outside this spectral region folding in have also contributed to increases in sensitivity. Another important improvement is oversampling. That is, the collection of digital data points at a rate faster than that required to satisfy the Nyquist criterion of twice the highest desired spectral frequency. In theory, for an oversampling factor of $n$, a gain in dynamic range of $\log_2(n)$ is obtained, i.e. for an oversampling of 8 times, an effective gain of 3 bits in ADC resolution of the signals results. In practice, the oversampled signal is simply averaged over the $n$ measurements to restore the same number of data points corresponding to the Nyquist criterion. This prevents folding of noise or artifacts which would have been in the extended spectral region resulting from a consequence of the oversampling being equivalent to a spectral region $n$ times wider than required to satisfy the Nyquist criterion. Thus a second consequence of oversampling is an improved signal–noise ratio from removal of folded noise when the spectral region is truncated.

## The HPLC-NMR system

A block diagram of a typical instrumental set-up for HPLC-NMR is shown in **Figure 1**. This comprises a high-resolution NMR spectrometer with its superconducting magnet into which is placed a dedicated NMR flow probe, a standard HPLC system controlled by PC-resident software and a flow control unit which enables the system to be operated in four main modes as listed below.

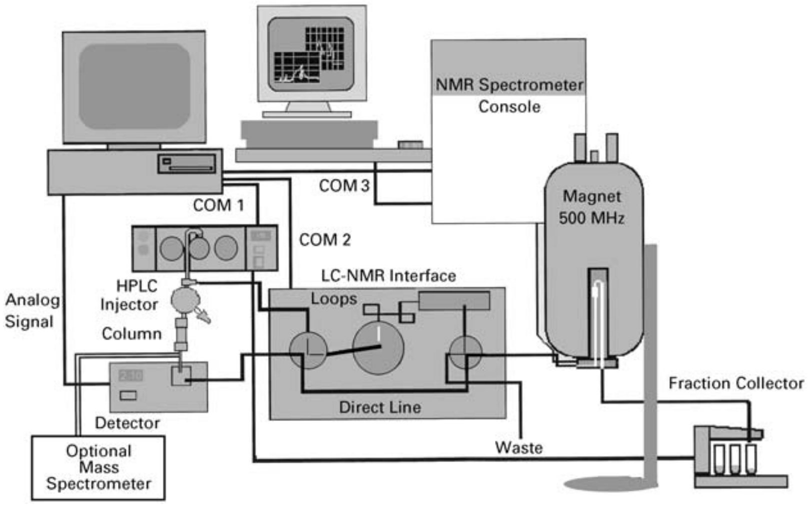



**Figure 1** Block diagram of a typical HPLC-NMR system.

- On-flow
- Stop-flow
- Time-sliced stop-flow
- Peak collection into capillary loops for post-chromatographic analysis.

The simplest method of operation is continuous-flow detection. This mode of operation is generally only feasible when using $^{1}H$ or $^{19}F$ NMR for detection unless enriched compounds are used. If on-flow detection is required during a solvent gradient elution, the NMR resonance positions of the solvent peaks will shift as the solvent proportions change. For effective solvent suppression, it is therefore necessary to determine these solvent resonance frequencies as the chromatographic run proceeds. This is accomplished by measuring a single exploratory scan as soon as a chromatographic peak is detected in real time during the chromatographic run and then applying solvent suppression irradiation at these frequencies as the peak elutes. The data from an on-flow experiment performed on a sample of human urine after dosing with paracetamol [1] is shown in **Figure 2**. The data are plotted in a pseudo-2D format with the axes being the chemical shift in ppm and the retention time of the chromatographic run. Slices extracted from the 2D plot show the 1D spectra of the glucuronide [2] and sulfate [3] conjugates of paracetamol (**Figure 3**).

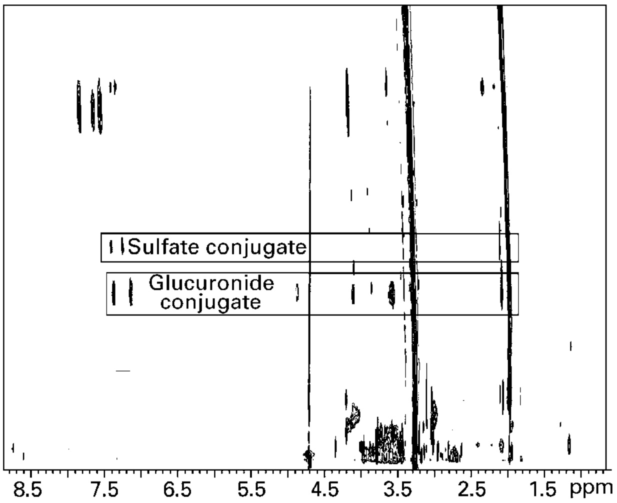
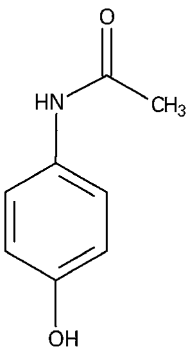
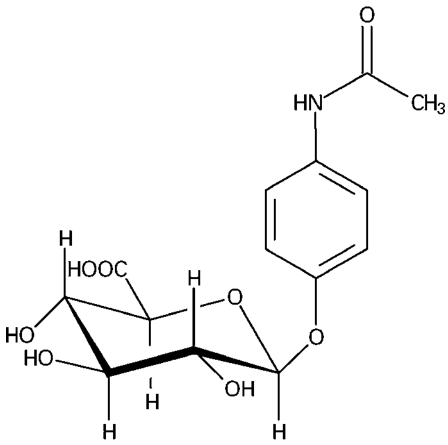



**Figure 2** Pseudo-2D plot of on-flow $^{1}H$ HPLC-NMR data obtained on human urine after dosing with paracetamol [1]. The resonance from the glucuronide [2] and sulfate [3] conjugate metabolites of paracetamol are indicated.

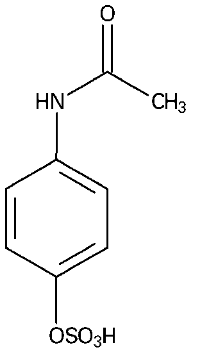
If the retention times of the compounds to be separated are known, or if they can be detected using refractive index, UV (including diode arrays), radiochemical or fluorescence detectors, stop-flow HPLC-NMR becomes an option. Upon detection the PC controlling the liquid chromatograph allows the pumps to continue running, moving the peak of the interest into the NMR probe. Once the pumps have stopped, normal high-resolution NMR is possible. It could be argued that the long length of capillary tubing connecting the HPLC system to the NMR probe will cause a significant loss of resolution in the separation. In fact this is not a problem. The NMR detection cell volume is typically 60–120 μL, and this represents the limiting factor in the chromatographic resolution. The practicality of the stop-flow approach has been amply demonstrated and although several separate stops are often made in each chromatographic run, the quality of the resulting NMR spectra are such that good structural information can be obtained. Even long 2D experiments, which provide




[1] Paracetamol




[2] Glucuronide conjugate of paracetamol




[3] Sulfate conjugate of paracetamol

correlation between NMR resonances, based on mutual spin–spin coupling such as COSY or TOCSY and heteronuclear correlation studies such as HMQC can be performed. It is in this stop-flow mode that HPLC-NMR is most commonly used.

There are two further special categories of stop-flow experiment. These comprise the ability to store eluting fractions in capillary loops for later off-line NMR study ('peak picking') and the very useful ability to stop the flow at short intervals over a chromatographic peak to 'time slice' different parts of a chromatographic run. This is analogous to the use of diode-array UV spectroscopy to obtain spectra from various positions within an eluting peak to determine peak purity. The time-slicing methods may be useful if there is poor chromatographic separation, if the compounds under study have weak or no UV chromophores or if the exact chromatographic retention time is unknown. It is also possible to time slice through an entire chromatographic run producing the equivalent of an on-flow experiment with higher signal to noise. The data from such a time slicing experiment has been referred to as a total NMR chromatogram (tNMRc).

In all the modes described above programmed gradient elution profiles can be performed and no compromise needs to be made in the chromatographic conditions, with the exception of those outlined in the following section.

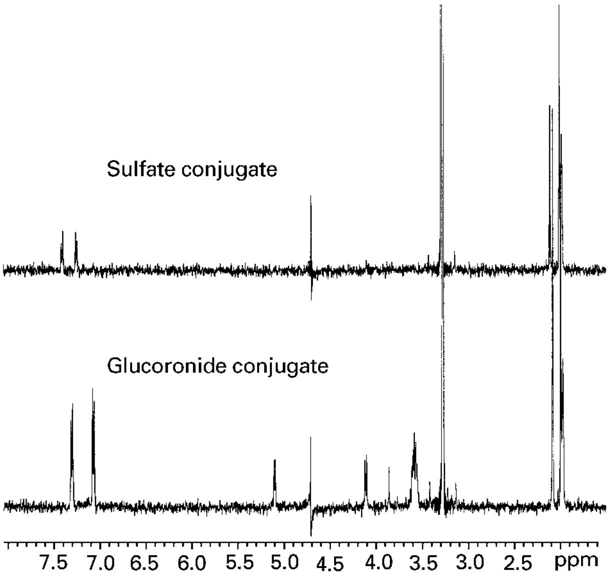



**Figure 3** 1D slices extracted from the on-flow experiment shown in **Figure 2**. The top trace is the sulfate conjugate [2] and the bottom trace is the glucuronide conjugate [3].

## HPLC-NMR solvents and chromatographic considerations

Another major problem in development of HPLC-NMR involved the use of deuterated solvents. Conventional high-resolution NMR spectroscopy routinely uses deuterated solvents. However, these solvents are thought to be prohibitively expensive if they have to be used for elution in HPLC-NMR

because of the volumes involved. It is always possible to use solvents such as $CCl_4$, which contain no protons, for normal-phase chromatography but this then requires the use of HPLC-NMR probes which have a separate external deuterium lock channel. Since most chromatographic methods are reverse-phase, they involve the use of $H_2O$ and an organic solvent such as acetonitrile or methanol. These give rise to large signals in the NMR spectrum which can obscure the spectrum of the analyte. The solution to this problem is solvent suppression which can be performed quite efficiently on modern NMR spectrometers. So efficiently, in fact, that the use of deuterated solvents is no longer a necessity. However, $D_2O$ is used rather than $H_2O$ to make the multiple solvent suppression easier and to serve as the field-lock substance. The cost of $D_2O$ is relatively modest at ~$300 per litre.

One of the most common organic solvents used in reverse-phase HPLC separations is acetonitrile, which gives rise to a singlet resonance in the $^1$H NMR spectrum at about $\delta$ 2.0. This singlet and that arising from residual $H_2O$ in $D_2O$ can be easily suppressed by presaturation. However, these suppression methods based on presaturation leave the $^{13}$C satellite peaks from the 1.1% of molecules with the naturally abundant $^{13}$C isotope at the methyl carbon. These satellite peaks are often much larger than the analyte peaks of interest in an HPLC-NMR study and so it is often necessary to suppress these also. This has been achieved in two ways. It is possible to set the suppression irradiation frequency over the central peak and the two satellite peaks in a cyclic fashion or, if an inverse-geometry probe is used, which includes a $^{13}$C coil, then $^{13}$C decoupling is possible. This will collapse the satellite peaks under the central peak, enabling conventional single-frequency suppression. The WET solvent suppression method incorporates selective RF pulses, pulse-field-gradients and optional $^{13}$C decoupling to provide a robust solvent suppression method well suited to HPLC-NMR. **Figure 4** shows an example of WET solvent suppression on the $^1$H HPLC-NMR spectrum of the desmethyl–glucuronide conjugate of naproxen [5] from human urine. The top trace shows the unsuppressed spectrum, while the bottom trace shows the same sample with WET solvent suppression.

Use of a fully deuterated solvent system, acetonitrile-$d_3$ and $D_2O$, is also an option when looking at very small quantities of material with resonances in the region near $\delta$ 2.0. Although the cost might seem to be prohibitive, at ~$1500 per litre for acetonitrile-$d_3$, the solvents for a typical HPLC run using both $D_2O$ and acetonitrile-$d_3$ cost only $12 to $15.

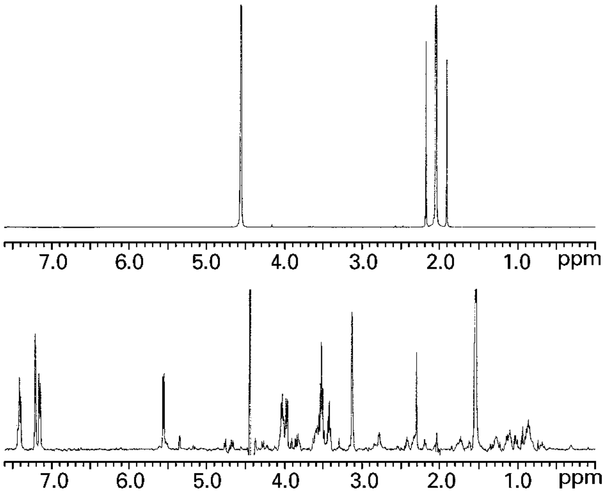


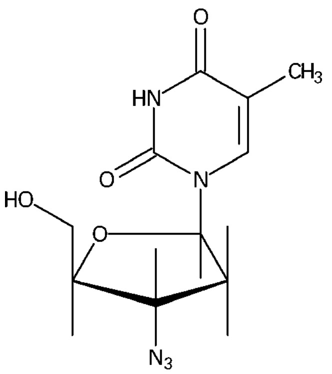

**Figure 4** An example of solvent suppression using WET. The top trace shows the data with no solvent suppression and the bottom trace shows the data with WET solvent suppression applied.

Other chromatographic considerations, outlined below, should be considered when designing an HPLC separation method.

- Use buffers that have as few $^1$H NMR resonances as possible.
  - Trifluroacetic acid.
  - Ammonium phosphate.
  - Ammonium acetate.
  - Deuterated formic acid ($130 per 5 g).
- Use ionpair reagents that have as few $^1$H NMR resonances as possible.
  - Ionpairs with *n*-butyl groups create 4 additional resonances.
  - Ionpairs with *t*-butyl groups create 1 additional resonance.
- Whenever possible use HPLC-gradient methods.
  - Run at a 5% change per min maximum.
- Use columns which will be suitable for the above criteria.
  - Reverse-phase $C_n$.
  - 2.0 to 4.6 mm diameter.

## Sensitivity of HPLC-NMR

The sensitivity of HPLC-NMR and the expected detection limits have been estimated for both 1D $^1$H NMR and 2D NMR. All NMR experiments were carried out using a Bruker Avance 500 NMR spectrometer operating at an $^1$H observation frequency




[4] AZT

of 500.13 MHz. The spectral data were measured using a dedicated $^1H$–$^{13}C$ inverse-geometry HPLC-NMR probe with a 120 µL active volume. All NMR spectra were measured at 25°C. The total probe volume including the inlet line and the outlet to waste was measured to be 800 µL. Samples were injected directly into the NMR probe and hence the tests were of probe sensitivity rather than the overall sensitivity of HPLC-NMR. The S/N ratio is defined as the peak height of a given signal multiplied by 2.5 and divided by the peak-to-peak height of the noise measured over a given spectral range of 200 Hz. Spectra were measured on a standard sample of 0.1% ethylbenzene in $CDCl_3$. The $^1H$ S/N calculation, based on the signal for the methylene quartet at 2.7 ppm, gave a value of 237:1.

A typical organic molecule 3′-deoxy-3′-azidothymidine (AZT) [4] was used to test the absolute detection limit of the same HPLC-NMR probe. Having determined the total volume of the NMR probe system to be 800 µL, a solution containing 1 µg of AZT in 120 µL was injected directly into the probe. The solvent was 80% $D_2O$–20% acetonitrile. The S/N ratio was measured on the ribose H1′ proton in the $^1H$ NMR spectrum at $\delta$ 6.2. The detection limit for this $^1H$ signal was defined as a signal height to peak-to-peak noise ratio of 3.0 (i.e. a S/N of 7.5 by the definition given earlier). Thus a detection limit of 500 ng in 64 scans was calculated.

AZT [4] was also used to test the $^1H$ sensitivity limit for a $^1H$–$^1H$ 2D TOCSY spectrum. A spectrum was obtained using a sample of 1 µg of AZT in 120 µL of 80% $D_2O$–20% acetonitrile and acquired over 4.0 hr. Examination of peak volumes shows that the detection limit of this experiment for 16 hr of acquisition (i.e. overnight) is about 500 ng.

The data acquisition parameters used were designed to ensure full $T_1$ relaxation so that any relative integral values would be quantitative. However, these are not the parameters, which give maximum S/N values. If the $^1H$ NMR relaxation times are assumed to be about 1.5 s, and if the data had been collected in a manner designed to yield maximum S/N (i.e. with a pulse repetition rate of about $1.3T_1$), then this would result in a S/N enhancement by a factor of about 1.4. In addition, because of the faster pulsing rate, more scans could be accumulated in the same experimental time leading to a further S/N enhancement factor of about 1.7. Thus the overall S/N might be expected to increase by a factor of about 2.4. In addition, for the 2-dimensional experiments, it is possible to halve the number of $T_1$ increments and extrapolate the data using linear prediction. This allows the acquisition of twice as many scans per increment for the same overall experiment time resulting in a 1.4 increase in S/N.

These detection limits will be lowered as new technical advances in NMR become available. These include higher magnetic field strengths and the future availability of HPLC-NMR cyro-probes, probes cooled to cryogenic temperatures, which reduce the noise figure and can provide as much as a 4.0 increase in S/N.

## HPLC-NMR probe design

The flow probe is the heart of the HPLC-NMR system. The main design criteria for such probes have been defined as follows:

- The geometry of the NMR cell should allow flow characteristics that give spectral resolution sufficient to allow spin-coupled multiplets to be resolved. This enables detailed structural elucidation to be carried out.
- NMR line broadening from magnetic susceptibility and other cell-wall effects should be avoided.
- The coil design and position should provide the highest possible NMR sensitivity.

The achievable signal–noise (S/N) ratio of a NMR detector is a function of a number of parameters expressed in Equation [1].

$$S/N = \gamma . N . I(I+1) . (B_0/T)^{3/2} . f . (QV_S/b)^{1/2} . n^{-1} \quad [1]$$

Therefore S/N can be increased by doing the following:

- Increase the sample volume and hence the number of detected nuclei, $N$.
- Increase the magnetic-field strength $B_0$.

- Increase the filling factor $f$ of the NMR coil (this is $V_s/V_c$ where $V_s$ is the sample volume and $V_c$ is the volume inside the detector coil).
- Increase the quality factor, $Q$, of the RF coil.
- Reduce the receiver bandwidth $b$.
- Operate at reduced temperature $T$.
- Improve preamplifiers by reducing their noise figure $n$.

All these are for a given nucleus with gyromagnetic ratio $\gamma$ and spin quantum number $I$.

For HPLC-NMR probes there is a compromise which has to be made in that the detection volume should be as small as possible to get optimum chromatographic resolution and hence this can only be compensated for by increasing the filling factor. This is achieved by fixing the RF coil directly onto the outside of the NMR cell. If this is done, it is of course not possible to spin the sample to improve magnetic-field inhomogeneities. However, in practice this is not a problem because the smaller the sample volume, the better the shimming. Additionally the use of a computer-controlled orthogonal shim sets has reduced the necessity to spin the sample by greatly improving non-spinning line shapes. One major factor in determining the sensitivity or peak heights is the observed line shape. If the peaks have wide bases, then a significant part of the signal intensity is found in this part of the peak and poor signal-to-noise (S/N) results. Thus good shimming, especially with respect to the 'hump test' (the linewidth of the chloroform $^1$H NMR peak at the height of the $^{13}$C satellite signals relative to that at 20% of the height of the $^{13}$C satellite signals) is a prerequisite for good S/N and essential to successful HPLC-NMR operation.

In superconducting magnets, the main magnetic field is parallel to the long axis of the NMR flow cell and the flow direction and the RF is applied to a Helmholtz coil on the side of the probe insert. This coil design has a lower $Q$ (quality factor) than a horizontal solenoid coil. A horizontal coil would be preferred on sensitivity grounds but shim systems are designed for vertical sample tubes, as normally samples are inserted vertically from the top of the magnet. The NMR flow cell fixed in a vertical position has the advantage of inducing laminar flow which helps remove air bubbles from the cell.

NMR flow cells have so far had much larger volumes than, for example, UV detection cells because of the lower sensitivity of NMR. The effects that large-volume NMR cells have on chromatographic peak dispersion have been investigated using a modified fluorescence detector. The selection of the detection volume of an NMR flow cell is a compromise between the needs of the NMR sensitivity and the HPLC resolution. Typically the optimum detection volume is 1–5 μL when an HPLC column of 250 × 4 mm is used. In contrast the usual detected volume in a conventional NMR tube is rarely less than 60 μL. It has been shown that the NMR resonances are broadened at different flow rates using detectors with a range of volumes. At flow rates of 1 mL min$^{-1}$ an increase of 0.14 Hz in linewidth is obtained only with flow cell volumes <120 μL. The broadening is due to an effective shortening of the spin–spin relaxation time, $T_2$, due to flow. Thus,

$$1/T_2\,(\text{obs}) = 1/T_2 + 1/t \qquad [2]$$

where $t$ is the flow rate. The selection of an HPLC-NMR probe is thus a compromise between sensitivity and resolution. The two most common flow cell volumes are 60 μL and 120 μL. In general the 60 μL flow cells are best for on-flow experiments and stop-flow using small-diameter columns with low flow rates. The 120 μL flow cell is more suited to stop-flow experiments using conventional HPLC columns and flow rates.

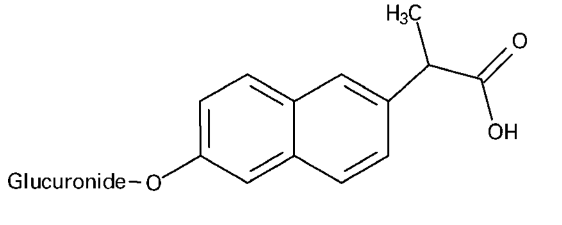
## Applications

HPLC-NMR has been applied to a wide range of analytical problems. These include:

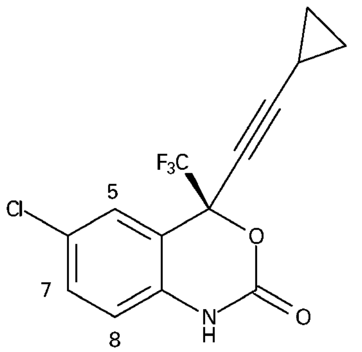
- The characterization of endogenous and xenobiotic metabolites directly from a biological matrix.
- Characterization of metabolites from *in vitro* studies.
- The dynamic study of reactive metabolites.
- Characterization of natural products from complex mixtures.
- Polymer characterization.
- Characterization of reaction impurities from small- and large-scale systems.

In order to illustrate the utility of HPLC-NMR and provide examples of typical data from HPLC-NMR experiments, the characterization of the human urinary metabolites of the anti-HIV drug efavirenz (Sustiva®) is described below.

Efavirenz [6] is a non-nucleoside inhibitor of HIV-1 reverse transcriptase. The metabolism of efavirenz in rats was studied using directly coupled HPLC-NMR-MS. This hyphenated technique utilizes an eluent splitter to also incorporate a mass spectrometer into the HPLC-NMR system in order to obtain additional information on molecular weight and to act as a detector for selection of peaks on which to perform



[5] Desmethyl - glucuronide conjugate of naproxen



[6] Efavirenz

NMR experiments. The sample for analysis was prepared by solid-phase extraction of 5 mL of urine obtained from rats after dosing with 800 mg $kg^{-1}$ of efavirenz, with 0–24 hr collection. The extract was dried and reconstituted with 100 μL of 80% $D_2O$ and 20% acetonitrile-$d_3$. A 40 μL injection was made onto a 3.9 × 150 Waters Symmetry $C_{18}$ column. A gradient elution from 80% $D_2O$ and 20% acetonitrile-$d_3$ to 50% $D_2O$ and 50% acetonitrile-$d_3$ over 20 min at a flow rate of 0.8 mL $min^{-1}$ was employed for separation. Using a splitter immediately after the column, 95% of the sample went to the UV detector and onto the NMR spectrometer while 5% went to a Finnigan LCQ ion-trap mass spectrometer equipped with an ESI probe operating in the positive-ion mode. The system was plumbed to allow the peak to reach the UV detector and the mass spectrometer at the same time.

The initial experiment, an on-flow $^{19}F$ detected HPLC-NMR-MS run, takes advantage of the $CF_3$ group present in the parent drug. Since there are no endogenous fluorinated compounds in rat urine, responses in the spectrum must arise from metabolites of efavirenz. This experiment provided the retention times of the metabolites that were of sufficient concentration to be detected. The $^{19}F$ chemical shift and the mass of the metabolites were also obtained from this single experiment. **Figure 5** shows the data as a pseudo-2D plot. The peaks of interest are labelled with their molecular mass.

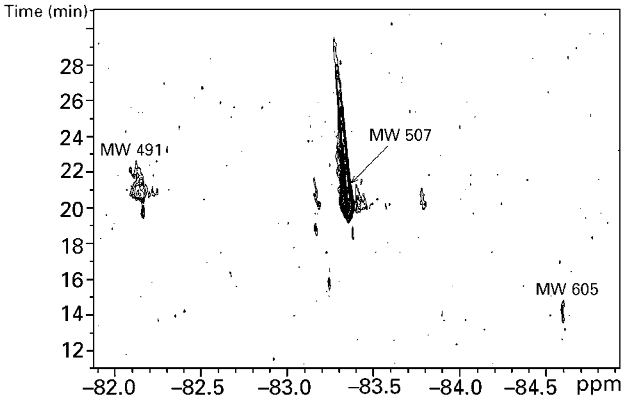
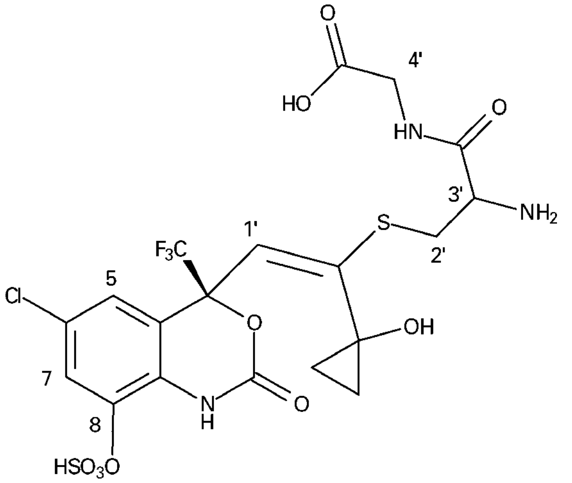



**Figure 5** Pseudo-2D plot of on-flow $^{19}H$ HPLC-NMR-MS data obtained on rat urine after dosing with efavirenz. The data supply the $^{19}F$ chemical shift, retention time, and mass of the metabolites.



[7]

The next phase of the experimental procedure is to obtain stop-flow $^{1}H$ HPLC-NMR spectra at the retention times determined in the on-flow $^{19}F$ HPLC-NMR experiment. The 1D and TOCSY spectra obtained on the minor metabolite with retention time 14.2 min, mass 605 Da and $^{19}F$ chemical shift –84.6 ppm are shown in **Figure 6**. The mass-spectral data showed a loss of 80 Da from the parent mass indicative of $SO_3$ loss. The loss of the methine proton from the cyclopropyl ring, seen clearly from the TOCSY spectrum, and the downfield shift of the remaining methylene protons indicates addition of OH to the cyclopropyl ring. The singlet at $\delta$ 6.47 is consistent with reduction of the alkyne. The spin system observed between $\delta$ 3 and $\delta$ 4 ppm is consistent with cysteinylglycine conjugation. Based on these data and the observed molecular mass of 605 Da, this peak is assigned as [7].

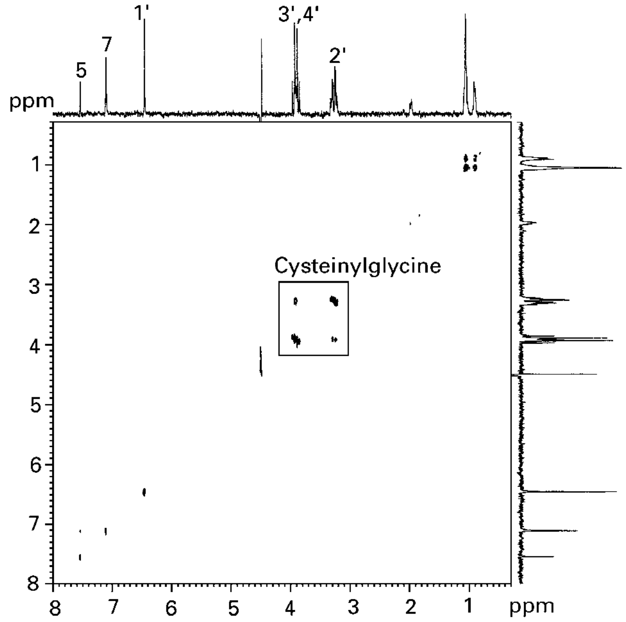
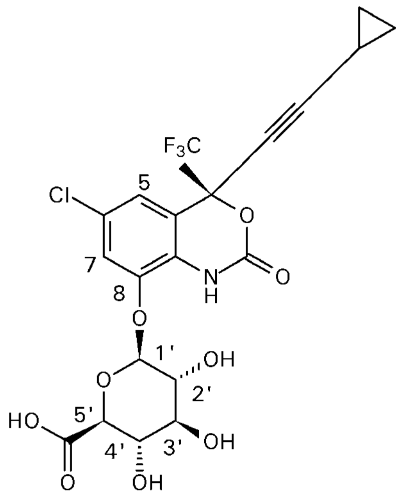



Figure 6 1D and TOCSY HPLC-NMR spectra of the 8-O-sulfate, cysteinylglycine di-conjugate of efavirenz.




[8] 8-*O*-glucuronide conjugate of efavirenz

The 1D and TOCSY data obtained on the major metabolite with retention time 20.5 min, mass 507 Da and $^{19}$F chemical shift −83.3 ppm is shown in **Figure 7**. These data clearly show the presence of a glucuronide conjugate. The observed changes in the aromatic region of the spectrum are consistent with the 8-O-glucuronide conjugate of efavirenz [8].

Co-eluting with the major metabolite is a component with mass 491 Da and $^{19}$F chemical shift −82.1 ppm. Using the mass spectrometer as a detector the stop flow was executed when the desired mass, 491 Da, was observed. While the spectrum (**Figure 8**) is dominated by the major metabolite, the 8-O-glucuronide conjugate of efavirenz, it is possible to assign several resonances to a glucuronide conjugate. The assignment of this metabolite as the *N*-glucuronide conjugate [9] is consistent with the chemical shift of H1′, $\delta$ 4.75, and with the observed molecular mass.

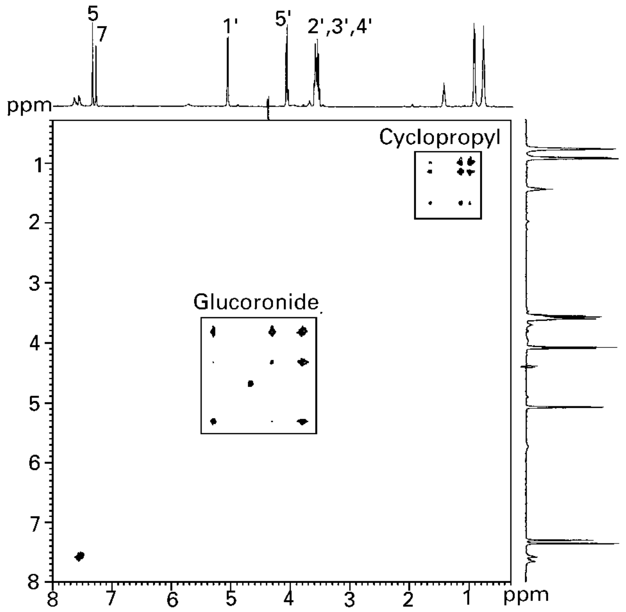
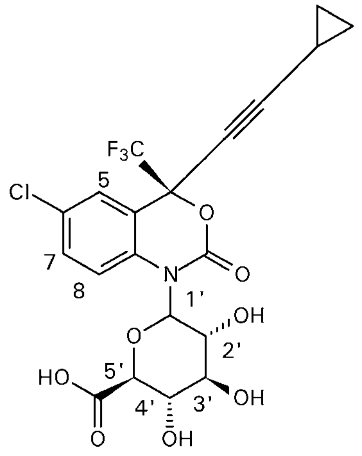



Figure 7 1D and TOCSY HPLC-NMR spectra of the 8-O-glucuronide conjugate of efavirenz.




[9] *N*-glucuronide conjugate of efavirenz

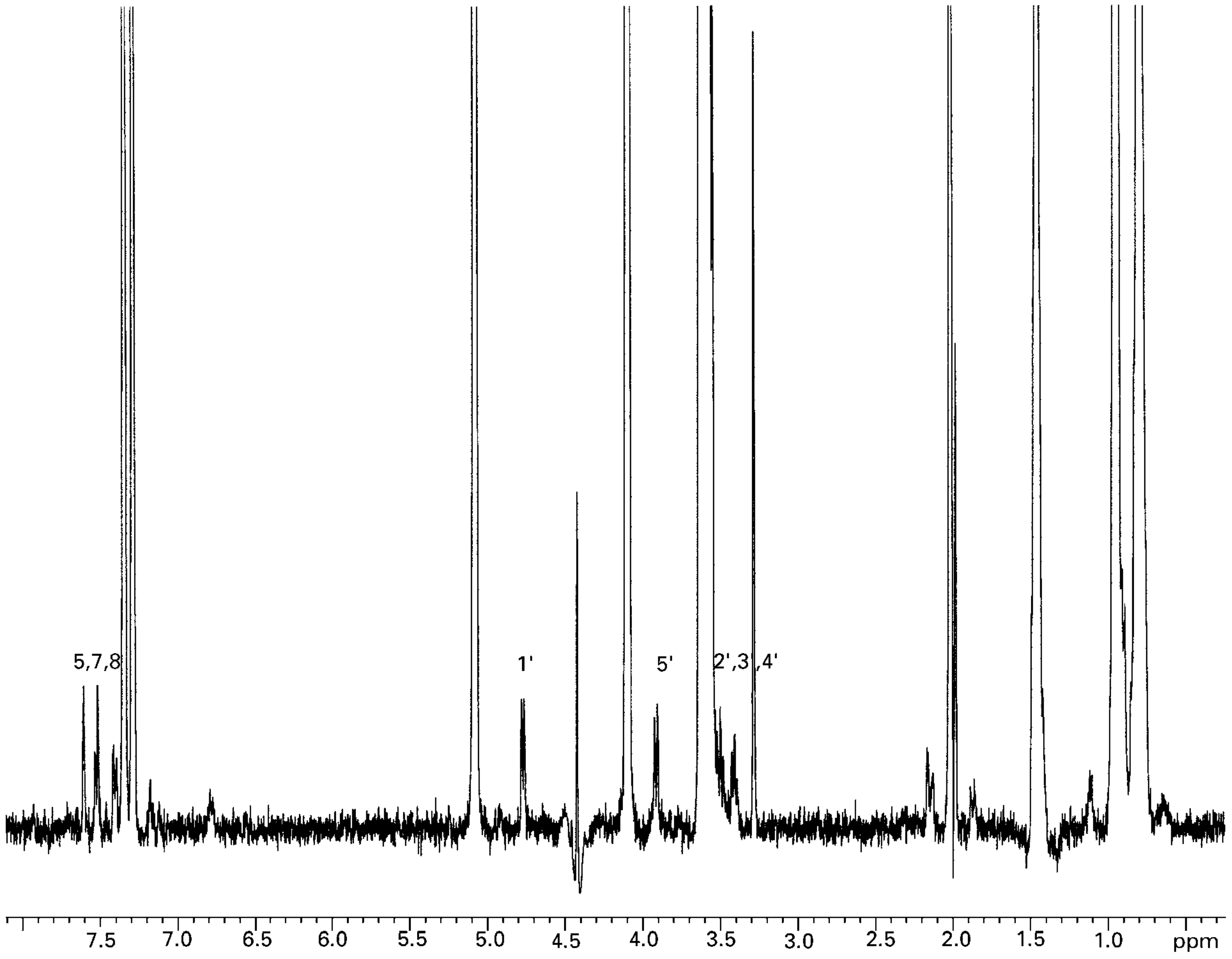



**Figure 8** The $^{1}H$ HPLC-NMR spectrum of the *N*-glucuronide conjugate of efavirenz shown as the minor component, along with the major component the 8-O-glucuronide conjugate.

## Summary

The experiments described serve to illustrate how HPLC-NMR can rapidly provide information on the structure of drug metabolites. These same methods have been employed in the study of endogenous compounds in biofluids, natural products, polymers and many other complex mixtures.

The high cost of the HPLC-NMR system is of course a factor which must be taken into consideration. Since most laboratories have seen the value of NMR in its traditional form, it is only necessary to add an HPLC system, an appropriate flow probe and a flow-control unit to an existing spectrometer to enable HPLC-NMR experiments to be performed. The cost of these accessories, or even the cost of a complete system dedicated to HPLC-NMR, is offset by the efficiency of the method. The laborious extraction of minor components from complex mixtures followed by off-line analysis or in the case of synthetic drugs the synthesis of radiochemically labelled materials with the numerous problems associated with handling and disposing of radiolabelled samples make the cost associated with HPLC-NMR less of a factor.

There are a number of limitations to the utility of HPLC-NMR. The greatest of these is the limitation on the amount of material that can be loaded on the column and consequently moved into the NMR probe. HPLC-NMR is not well suited to the less-sensitive NMR techniques like HMBC, which provides critical structural information via long-range interactions. Thus it is still necessary to isolate minor components from complex mixtures on occasion and HPLC-NMR should not be viewed as a replacement for conventional NMR. It is a tool which enhances the utility of NMR and like any tool it can be used properly and effectively or misused. Although it has experimental limitations when compared to conventional NMR, many structural problems can be solved using HPLC-NMR without resorting to tedious isolation methods.

The advantages of HPLC-NMR can be summarized as:

- Speed of analysis by elimination of the isolation step.
- Ability to obtain structural information and high-quality spectral data from complex matrices.
- The lack of need for perfect chromatographic separation.
- Synergy when directly coupled with mass spectrometry.

## List of symbols

$b$ = receiver bandwidth; $B_0$ = magnetic-field strength; $f$ = coil filling factor; $I$ = spin quantum number; $n$ = oversampling factor; $n$ = preamplifier noise figure; $N$ = number of detected nuclei; $Q$ = coil quality factor; $T$ = temperature; $T_1$ = relaxation time; $T_2$ = spin–spin relaxation time; $V_c$ = detector-coil volume; $V_s$ = sample volume; $\gamma$ = gyromagnetic ratio; $\tau$ = flow rate.

*See also:* **Biofluids Studied By NMR; Chromatography-IR, Applications; Chromatography-MS, Methods; Drug Metabolism Studied Using NMR Spectroscopy; Fourier Transformation and Sampling Theory; NMR Data Processing; NMR Pulse Sequences; Solvent Suppression Methods in NMR Spectroscopy.**

## Further reading

Albert K (1995) On-line use of NMR detection in separation chemistry. *Journal of Chromatography A* **703**: 123–147.

Albert K (1995) Direct on-line coupling of capillary electrophoresis and $^1$H NMR spectroscopy. *Angewandte Chemie International Edition in English* **34**: 641–642.

Albert K, Kunst M, Bayer E, Spraul M and Bermel W (1989) Reverse-phase high-performance liquid chromatography-nuclear magnetic resonance on-line coupling with solvent non-excitation. *Journal of Chromatography* **463**: 355.

Albert K, Braumann, U, Tseng L-H, *et al.* (1994) On-line coupling of supercritical fluid chromatography and proton high-field nuclear magnetic resonance spectroscopy. *Analytical Chemistry* **66**: 3042–3046.

Behnke B, Schlotterbeck G, Tallarek U, *et al.* (1996) Capillary HPLC-NMR coupling: high resolution NMR spectroscopy in the nanoliter scale. *Analytical Chemistry* **68**: 1110–1115.

Griffiths L (1995) Optimization of NMR and HPLC conditions for LC-NMR. *Analytical Chemistry* **67**: 4091–4095.

Laude DA Jr and Wilkins CL (1984) Direct-linked analytical scale high-performance liquid chromatography/nuclear magnetic resonance spectroscopy. *Analytical Chemistry* **56**: 2471–2475.

Lindon JC, Farrant RD, Sanderson PN, *et al.* (1995) Separation and characterization of components of peptide libraries using on-flow coupled HPLC-NMR spectroscopy. *Magnetic Resonance in Chemistry* **33**: 857–863.

Lindon JC, Nicholson JK and Wilson ID (1996) Direct coupling of chromatographic separations to NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* **29**: 1–49.

Lindon JC, Nicholson JK, Sidelmann UG and Wilson ID (1997) Directly coupled HPLC-NMR and its application to drug metabolism. *Drug Metabolism Reviews* **29**: 705–746.

Pullen FS, Swanson AG, Newman MJ and Richards DS (1995) "On-line" liquid chromatography/nuclear magnetic resonance mass spectrometry – a powerful spectroscopic tool for the analysis of mixtures of pharmaceutical interest. *Rapid Communications in Mass Spectrometry* **9**: 1003–1006.

Shockcor JP, Unger SE, Wilson ID, Foxall PJD, Nicholson JK and Lindon JC (1996) Combined HPLC, NMR spectroscopy, and ion-trap mass spectrometry with application to the detection and characterization of xenobiotic and endogenous metabolites in human urine. *Analytical Chemistry* **68**: 4431–4435.

Shockcor JP, Wurm RM, Frick LW, *et al.* (1996) HPLC-NMR identification of the human urinary metabolites of (–)-cis-5-fluoro-1-[2-(hydroxymethyl)-1,3-oxathiolan-5-yl]-cytosine, a nucleoside analogue active against human immunodeficiency virus (HIV). *Xenobiotica* **26**: 189–199.

Smallcombe SH, Patt SL and Keifer PA (1995) WET solvent suppression and its application to LC-NMR and high resolution NMR spectroscopy. *Journal of Magnetic Resonance Series A* **117**: 295–303.

Sweatman BC, Farrant RD, Sanderson PN, *et al.* (1995) Evaluation of the detection limits of directly coupled 600 MHz $^1$H AND $^1$H–$^{13}$C HPLC-NMR spectroscopy. *Journal of Magnetic Resonance Analysis* **1**: 9–12.

Wolfender J-L, Rodriguez S and Hostettmann K (1998) Liquid chromatography coupled to mass spectrometry and nuclear magnetic resonance spectroscopy for the screening of plant constituents. *Journal of Chromatography A* **794**: 299–316.

# Chromium NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# CIDNP Applications

**Tatyana V Leshina**, **Alexander I Kruppa** and
**Marc B Taraban**, Institute of Chemical Kinetics and
Combustion, Novosibirsk-90, Russia

Copyright © 1999 Academic Press



The phenomenon of chemically induced dynamic nuclear polarization (CIDNP) consists of the manifestation of unusual line intensities and/or phases of signals of radical reaction products in the NMR spectrum when reaction takes place directly in the probe of the spectrometer. These anomalous lines (enhanced absorption or emission of NMR signals), which reflect the populations of nuclear spin states deviating from the Boltzmann condition, are observed within the time range of nuclear relaxation times of the diamagnetic molecules ($T_{1N}$), which are as a rule, several seconds to tens of seconds. Subsequently, the NMR spectrum re-acquires its usual form. In 1967, two research groups in Europe (J. Bargon, H. Fischer, and U. Johnson) and the USA (H. Ward and R. Lawler) discovered independently that this phenomenon is directly associated with the free radicals involved in the process. Later on, it was shown that this also pertains to radical ions and triplet excited states of molecules.

According to currently existing theory, the non-equilibrium intensities in NMR spectra are generated as a result of electron–nuclear interactions in the so-called radical pair. These are the pairs of paramagnetic particles that may originate during the homolysis of a molecule under the action of heating, light or ionizing radiation, as well as resulting from single-electron transfer between the donor and acceptor molecules and radical encounters in the bulk preceding the recombination.

It follows from the above list that, in principle, the CIDNP phenomenon can be observed for virtually all types of radical reactions. To date, various CIDNP techniques have been employed to clarify the role of radical species in: redox reactions of organic compounds, nucleophilic and electrophilic substitution, addition and elimination reactions, isomerization processes, thermolysis and photolysis of organic peroxides and azo- and diazocompounds, a number of photosensitized processes, including the reactions of organoelement compounds of silicon, germanium, tin, mercury, aluminium, lithium, sodium, potassium, magnesium and organoxenon compounds, polymerization processes and a number of rearrangement reactions.

Note that prior to CIDNP studies, in certain cases, the involvement of paramagnetic species in the reactions had not even been anticipated. A good example is the process of sensitized *cis–trans* photoisomerization of substituted olefins. Moreover, CIDNP has brought unambiguous evidence of earlier-suggested single-electron transfer as the first stage of aliphatic nucleophilic substitution in the reaction of alkyl- and benzylhalides with alkyl derivatives of lithium.

The analysis of CIDNP effects allows information to be obtained on the structure and reactivity of active short-lived (from nanoseconds to microseconds) paramagnetic species (free radicals and radical ions, triplet excited molecules), and on molecular dynamics in the radical pair. The analysis makes it possible to distinguish between the bulk and 'in-cage' stages of complex chemical reactions. CIDNP data also provide information on the multiplicities of reacting states, which is of utmost importance for understanding photochemical processes.

CIDNP is not a direct method of detection of the structure of radical species: the information of hyperfine interaction (hfi) constants, *g*-factors, and lifetimes of paramagnetic intermediates, is obtained on the basis of the analysis of the phases and intensities of polarized lines in NMR spectra, as well as from CIDNP kinetics, and the dependence of polarization efficiency on the magnetic field strength applied to the sample during reaction. However, the simplicity and reliability of the identification of polarized lines in the NMR spectra ensure high plausibility of CIDNP-based structural and kinetic information.

Thus nowadays CIDNP is one of the powerful techniques for the investigation of elementary stages of complex radical chemical reactions.

## Nuclei demonstrating the CIDNP phenomenon

Chemical polarization is observed only for nuclei possessing a magnetic moment (spin). **Table 1** lists the nuclei for which CIDNP effects have been detected to date. At present, the majority of CIDNP experiments have been carried out for hydrogen ($^{1}H$) and fluorine ($^{19}F$) nuclei. Next in incidence are $^{13}C$ and $^{31}P$. Unless

**Table 1** Nuclei which have been used in CIDNP studies

| *Nucleus* | *Spin* |
|---|---|
| $^{1}H$ | 1/2 |
| $^{2}H$ | 1 |
| $^{13}C$ | 1/2 |
| $^{15}N$ | 1/2 |
| $^{19}F$ | 1/2 |
| $^{31}P$ | 1/2 |
| $^{119}Sn$ | 1/2 |

otherwise specified, all further consideration will employ the parameters characteristic for $^{1}H$ CIDNP.

## CIDNP observation techniques

The chemical reactions are most commonly run directly in the probe of an NMR spectrometer. Such chemical polarization formed in the magnetic fields of NMR spectrometers (with a strength of several tesla) is usually called CIDNP in high magnetic fields.

There exist other techniques of CIDNP observation, as well. In these techniques the reaction under study is carried out in a separate magnet, as a rule, with variable field strength, and afterwards the reaction mixture is transferred rapidly (with the characteristic time lower than $T_{1N}$) to the probe of the NMR spectrometer. The transfer is often performed by a flow system; however, other possibilities also exist.

It is necessary to note that CIDNP formed in high and low magnetic fields, such as the geomagnetic field, are two essentially different physical phenomena.

## Types of CIDNP effects

According to their manifestation, CIDNP effects can be subdivided into two types: net effects (enhanced absorption and emission of NMR signals), and multiplet effects, consisting of the redistribution of the line intensities between specific components of multiplet signals in the NMR spectrum (**Figure 1**).

## CIDNP effects in high magnetic fields

For the chemical application of CIDNP it is important to be aware of the fact that in high magnetic fields the nonequilibrium population of nuclear spin sublevels is formed as a result of transitions solely between singlet (S) triplet zero ($T_0$) states or a radical pair (**Figure 2**). These transitions occur due to interaction of an electron spin with spins of magnetic nuclei in the radical pair and they are also dependent on the difference in *g*-factors of radicals in the radical pair.

Theory shows that in this case the probabilities of the formation of the singlet (recombining) state of a radical pair are different for α and β projections of nuclear spins. As a result, spins with a certain orientation predominantly appear in the 'in-cage' recombination product, while spins with an opposite orientation occur in products resulting from the escape of radicals into the bulk (**Figure 3**). According to the existing Kaptein rules modified by Closs, the sign of the net CIDNP effect (*N*) in high magnetic fields is defined by the following set of parameters:

$$N = \mu \cdot \varepsilon \cdot \Delta g \cdot \mathrm{hfi} \cdot \lambda$$

where $\mu$ is the multiplicity of the precursor of the radical pair ('+' for triplet and '–' for singlet precursor), $\varepsilon$ is '+' for 'in-cage' and '–' for escape recombination products, $\Delta g$ is the sign of the difference in the *g*-factors of the radical with polarized nucleus and radical partner in the radical pair, hfi is the sign of the hyperfine constant of a given nucleus in the radical under study, and $\lambda$ is '+' for the radical pair recombining from an S-state and '–' for the radical pair recombining from a T-state (usually employed for radical ion reactions). The sign of *N* reflects the phase of the NMR signal of a given nucleus: '+' for enhanced absorption (*A*) and '–' for the emission (*E*).

There are also qualitative rules employed for the analysis of multiplet effects:

$$M = \mu \cdot \varepsilon \cdot \mathrm{hfi}_1 \cdot \mathrm{hfi}_2 \cdot J \cdot \gamma,$$

where μ, ε, are defined as in the previous case, $\mathrm{hfi}_1$ and $\mathrm{hfi}_2$ are the signs of hyperfine constants of nuclei 1 and 2 in the radical forms, *J* is the sign of the spin–spin coupling constant of these nuclei in the molecule, γ is '+' if these nuclei belong to the same radical, and '–' if they belong to different radicals in the radical pair. The sign of *M* reflects two types of multiplet effect: '+' for *E*/*A* and '–' for *A*/*E* type effects.

Analysis of the detected net and multiplet effects is also performed by means of a comparison of the experimentally observed effects with those that are calculated.

For example, the above-mentioned qualitative rules are applicable for the analysis of CIDNP effects for all nuclei from **Table 1**, except for $^{119}Sn$, which has a number of peculiarities stipulated by its anomalous magnetic resonance parameters (hyperfine constant more than 0.16 T, *g*-factor 2.0160). To analyse CIDNP effects of this nucleus one has to employ a special theoretical approach.

Thus, analysis of the experimental sign of the CIDNP effects basically allows the above enumerated parameters of the radical pair to be singled out

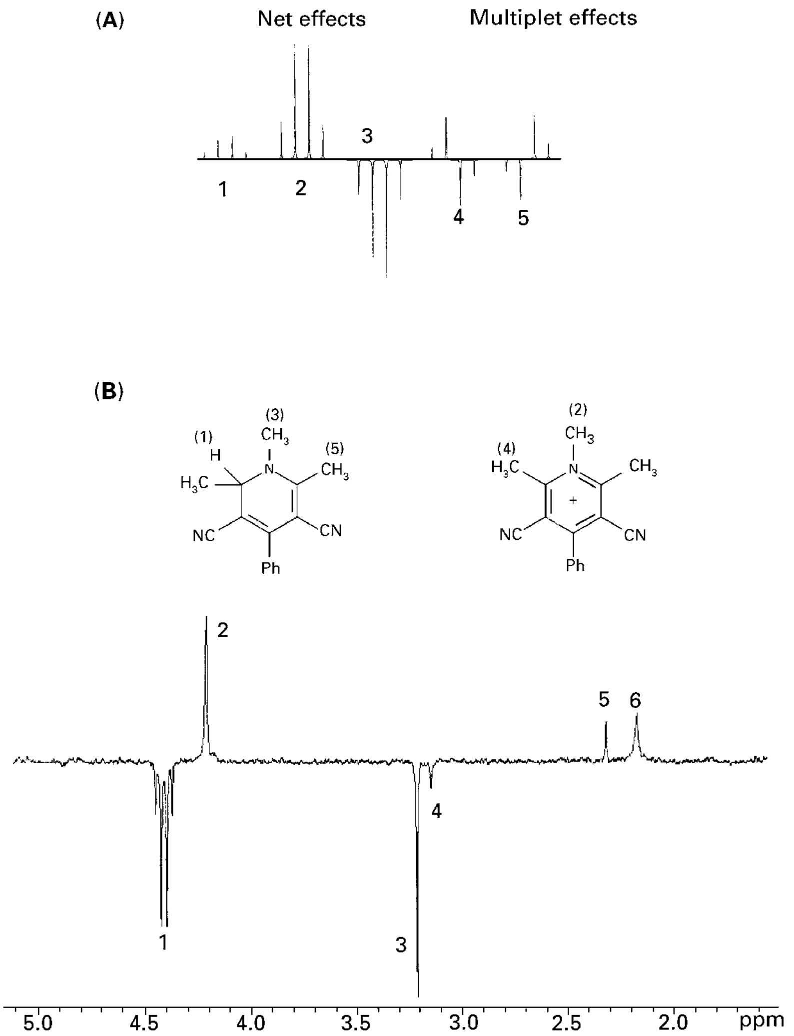



**Figure 1** (A) Schematic presentation of the types of CIDNP effects: (1) equilibrium NMR spectrum, (2) enhanced absorption (*A*), (3) emission ($E_2$), (4) multiplet effect (*A*/*E*), (5) multiplet effect (*E*/*A*). (B) Experimental example of time-resolved (100 μs after laser pulse) 250 MHz $^1$H CIDNP spectrum, detected during the photolysis of substituted 1,2-dihydropyridine in the presence of 2,5-dichlorobenzoquinone: line assignment is shown on the structures; signal (6) is attributed to HDO admixtures in $CD_3CN$. The equilibrium signals were removed using the presaturation pulse sequence preceding the laser pulse.

provided that all the others are known in advance. For example, the chemical shift of the polarized signal in an NMR spectrum and its multiplet structure often reveal the structure of the radical precursor of the polarized compound. In this case, this assumption is easily verified by substituting the magnetic resonance

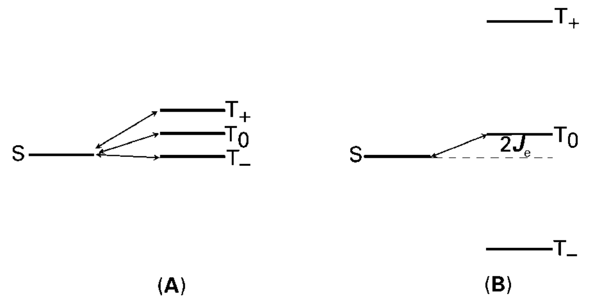



**Figure 2** Schematic presentation of the radical pair energy levels in (A) low ($H_0$ ≈ hyperfine interaction constant) and (B) high ($H_0$ >> hyperfine interaction constant) magnetic fields.

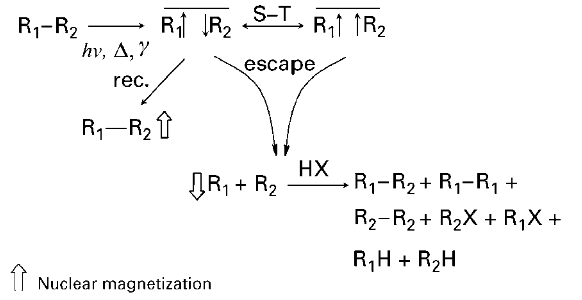



**Figure 3** Distribution of chemical polarization between geminate ('in-cage') and homogeneous products of radical reaction.

parameters of the suggested radical in the above expressions for the rules. Another possibility is the definition of the multiplicity of the reacting state for known compositions of the radical pair.

The enhancement coefficient serves as a basic quantitative feature of CIDNP. One usually distinguishes between the observed enhancement coefficient and the absolute one. The observed enhancement coefficient is defined as the ratio of the intensities of the polarized NMR signal to the equilibrium one recorded after the completion of the reaction. The time dependence of observed polarization on time ($dI^*/dt$) has two components, $(dI^*/dt)_{react}$ and $(dI^*/dt)_{relax}$, reflecting the contributions from chemical reaction and relaxation of CIDNP. As a first approximation, the relaxation of CIDNP occurs through spin–lattice relaxation, which suggests the following expression for the observed polarization:

$$\frac{dI^*}{dt} = \left(\frac{dI^*}{dt}\right)_{react} - \frac{I^*(t) - I_0(t)}{T_{1N}}$$

where $I^*$ and $I_0$ are the intensities of polarized and equilibrium signal, respectively; $(dI^*/dt)_{react} = dI_0/dt \cdot m$, where $m$ is the absolute enhancement factor of the product polarization. This expression implies that the observed polarization depends on the magnetic resonance parameters of the radical pair (hyperfine coupling, $g$-factors) through $m$, radical relaxation times in the diamagnetic state, and the kinetics of the buildup of polarized products.

The absolute enhancement coefficient of CIDNP differs from the observed one in that the former is dependent only on the radical pair parameters. At present, the absolute enhancement coefficient can be reliably measured only for the case of photochemical reactions with the employment of specially elaborated time-resolved CIDNP techniques.

From the viewpoint of experimental techniques, the major potentialities are now realized for research into CIDNP in photochemical reactions. In this case, time-resolved techniques (so-called flash-CIDNP) are widely used. These techniques utilize laser-pulse photoexcitation with pulse detection of CIDNP effects. The solution of the sensitivity versus time-resolution dilemma stipulates the following restrictions of the attainable time resolution: the duration of the detecting radiofrequency pulse should not be shorter than 100 ns, otherwise the sensitivity of the NMR method is lost. Shorter pulses can be employed only to study the processes demonstrating an extreme absolute enhancement coefficient of CIDNP. The latter are characteristic only for a limited number of model reactions, for example the photolysis of cyclic ketones.

The improvement of time resolution is possible when using both a standard NMR spectrometer and modified equipment. These techniques differ by the methods of mathematical processing of the experimental data when estimating the kinetic parameters.

Even in the simplest case of a microsecond time range, the time-resolved techniques allow one to observe homogeneous and geminate processes separately. The homogeneous processes include bimolecular recombination in the bulk, degenerate electron exchange reactions, proton exchange, and various rearrangement reactions which also occur in the bulk.

## CIDNP effects in low magnetic fields

A low magnetic field is defined as one in which the energy of the Zeeman interaction is of the same order of magnitude as the hyperfine interaction (see **Figure 2**). Therefore, a nonequilibrium population of nuclear states may arise not only via $S–T_0$ transitions, but also through $S–T_-$ and $S–T_+$ transitions between energy levels of a radical pair. The observed sign of CIDNP will predominantly depend on the difference in efficiencies of $S–T_{+,-}$ transitions.

As **Figure 2** suggests, the efficiency of $S–T_-$ and $S–T_+$ transitions also depends essentially on the magnitude and sign of the electron exchange interaction ($J_e$) in the radical pair. However, experimental studies on CIDNP in low magnetic fields show that for a radical pair comprising neutral radicals in nonviscous solutions, $J_e$ does not contribute significantly to the resulting polarization. The exceptions are radical ion pairs and biradicals.

To unravel the mechanisms of the elementary stages of radical reactions by means of the CIDNP method in low magnetic fields, one usually analyses the dependencies of CIDNP efficiencies on the magnetic field strength (field dependence of CIDNP). In this case, the data on radical pair composition are obtained on the basis of comparison of experimentally measured and theoretically calculated CIDNP field dependencies.

## Comparison of CIDNP techniques in high and low magnetic fields for investigation of elementary mechanisms of radical reactions

The basic distinction in the potential of application of these two techniques of CIDNP observation is stipulated by the difference in the mechanisms of their formation. Thus, formation of CIDNP effects

in high magnetic field is not accompanied by flips of either electron or nuclear spin (for S and $T_0$ states the projection onto the magnetic field direction is equal to zero). Here, the radical reaction actually plays the role of dispatcher, directing radicals with $\alpha$ and $\beta$ projection of nuclear spin to different products (see **Figure 3**). Thus, the formation of at least two products is a key element for CIDNP manifestation. This condition essentially narrows the range of processes open for investigation. Thus, so-called crypto-radical reactions, which are rather widespread in chemical practice, fall outside of the consideration range. To study these processes it is recommended CIDNP in a low magnetic field is used.

In the case of CIDNP in low magnetic field, the true nonequilibrium population of nuclear sublevels results from simultaneous flips of both electron and nuclear spins (see **Figure 2**). Therefore, in this case, the formation of the single product does not preclude the observation of CIDNP. A number of other reasons also stipulate the greater versatility of chemical polarization in low magnetic fields, namely, the greater observed enhancement coefficients as compared to those in a high magnetic field and the potential and to detect net polarization formed in the radical pair with identical radicals.

## Examples of CIDNP applications to the structure determination and properties of paramagnetic species (free radicals, radical ions, biradicals, carbenes, macromolecules)

### CIDNP in structural chemistry

By now the analysis of CIDNP effects has produced a large set of useful structural data on the signs of paraffin constants, *g*-factor values of the radicals, and the energies of electron exchange interaction in the radical pairs.

Of special value is the information on the so-called magnetic resonance parameters (paraffin interactions and *g*-factors) that cannot be estimated easily by other magnetic resonance techniques in the case of short-lived radicals. The data obtained allow inferences to be made on the electronic structure of paramagnetic species ($\sigma$- or $\pi$-radical) and hence about their reactivity.

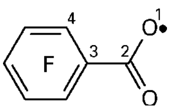
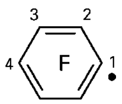
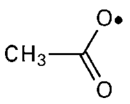
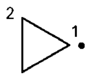
Among the main sources of such data are photolysis and thermoses reactions of various organic peroxides, demonstrating $^1$H, $^{19}$F and $^{13}$C CIDNP effects. As a rule, the mechanisms of these processes have been determined previously, and the multiplicity of the reacting state is also known. Thus, employing the rules for the analysis of multiple and net CIDNP effects in high magnetic fields (see above) one can sequential determine either the signs of the paraffin constants for the radical pairs or the sign of $\Delta g$.

To define the values of the paraffin constants and the *g*-factors of the radical pairs the entire polarized spectrum is stimulated and compared with the experimental one. This approach has been employed to obtain the values of the paraffin constants and *g*-factors listed in **Table 2**.

For example, it follows from **Table 2** that the bonds in the cyclopropyl free radical are flexed, and the proton in position 1 protrudes out of the plane of the three-membered ring, since the hyperfine constant of this proton is much lower than that observed in a methyl radical (2.3 mT).

Aside from hyperfine constants and *g*-factors, the analysis of the multiplet effect in the photolysis products of cyclopropyl peroxide and, cyclopropene has defined the sign of the vicinal proton–proton

**Table 2** Magnetic properties of free radicals defined by means of CIDNP

| *Radical* | | *Hyperfine interaction constant* | *Spin density, $\rho_\pi$* | *g-factor* |
|---|---|---|---|---|
|  | $^{19}$F:<br>$^{13}$C: | $a_4 > 0$<br>$a_2 > 0$ | $\rho_1 > 0$<br>$\rho_2 < 0$<br>$\rho_3 < 0$<br>$\rho_4 > 0$ | |
|  | $^{19}$F: | $a_2 > 0$<br>$a_3 > 0$<br>$a_4 < 0$ | | |
|  | $^1$H: | $a < 0$ (−2.3 mT) | | 2.0058 |
|  | $^1$H: | $a_1 < 0$ (−0.65 mT)<br>$a_2 > 0$ (2.34 mT) | | |

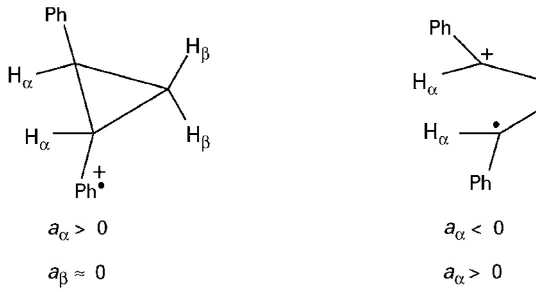



**Figure 4** Alternative structures of the 1,2-diphenylcyclopropane radical cation with the expected signs of the hyperfine interaction constants, *a*.

spin–spin coupling constant ($J_{H\text{-}H} > 0$) which was unknown earlier.

The analysis of CIDNP effects has also made possible the determination of the structures of a number of radical ions of cyclic compounds (benzonorbornadiene, trialkylcyclopropanes, *syn*- and *anti*-paracyclophanes, etc.).

For example, the structure of the radical cation of 1,2-diphenylcyclopropane has been assigned on the basis of the analysis of $^1$H CIDNP data formed in the act of photoinduced reversible electron transfer from cyclopropane to chloranil. The choice has been made between 'closed' and 'open' structures of the radical cation of 1,2-diphenylcyclopropane (**Figure 4**). The observed CIDNP effects of 1,2-diphenylcyclopropane (absorption of aromatic *ortho*-, *para*-, and $H_\alpha$, and the emission of $H_\beta$) comply only with the 'open' structure.

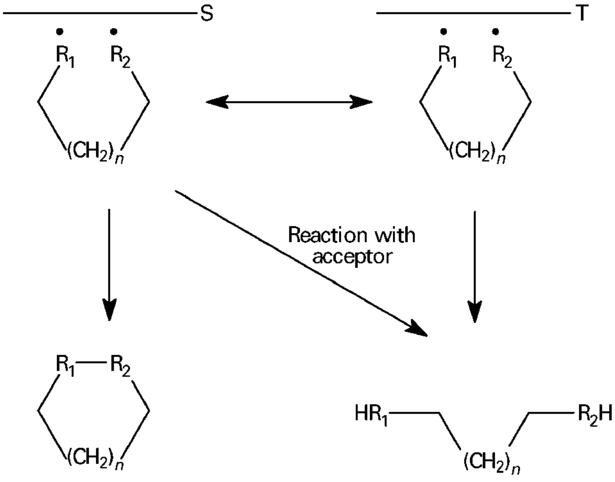
CIDNP is also a method of determining nuclear spin–lattice relaxation times of free radicals and radical ions ($T_{1R}$). This approach was used to determine the $T_{1R}$ values of the $CH_2$-protons of benzyl ($3.5 \times 10^{-4}$ s), dichloromethyl ($4.5 \times 10^{-4}$ s), and *t*-butyl ($2.4 \times 10^{-4}$ s) free radicals.

### CIDNP effects in reactions involving carbenes and biradicals

Paramagnetic species possessing two unpaired electrons (biradicals, carbenes, and their analogues) have been intensively studied using CIDNP techniques in high and low magnetic fields. The involvement of biradicals in organic and biochemical reactions is postulated in a number of cases, but the physical methods (EPR, laser spectroscopy) have allowed the detection of their formation predominantly in the photolysis of cyclic ketones. These species have also served as models for detailed studies of CIDNP peculiarities for biradicals, as well as the distinction of these effects from the case of free radicals.

A major peculiarity is the contribution of $S\text{–}T_{+,-}$ transitions to CIDNP, which is due to the electron exchange interaction between the spins of unpaired electrons (see the CIDNP effects in low magnetic fields section above).

If the processes taking place in a biradical allow the possibility of the selection from nuclear spin states, the $S\text{–}T_0$ transitions can also manifest themselves.




Reference data on biradicals of cyclic ketones point to the appearance of CIDNP effects in a wide range of magnetic field strengths. The manifestations of various mechanisms of CIDNP formation depend on the structure of the biradical.

As a rule, the field dependence of CIDNP shows up as curves with extremum, which are located in the magnetic field with the strength of about $J_e$.

Comparatively low values of $J_e$ (< 0.1 T) are characteristic of flexible biradicals with a long carbon chain ($C_8$ and more) separating the radical centres. Therefore, one might expect the manifestation of an $S\text{–}T_-$ mechanism (for $J_e > 0$, $S\text{–}T_+$ mechanism) in low magnetic fields, while in a high magnetic field the $S\text{–}T_0$ mechanism will manifest itself.

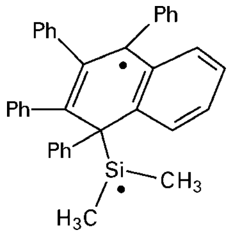
Large $J_e$ values are typical of short biradicals ($C_7$ and shorter). In this case the extremum will appear in a magnetic field similar to that of conventional NMR spectrometers, $J_e > 1$ T. Therefore, in this case only the $S\text{–}T_{+,-}$ mechanism of CIDNP formation will appear.

There are also so-called rigid biradicals whose structure does not possess a flexible carbon chain. In this case, $J_e$ may be either very large or negligibly small, the latter case reflecting the orthogonality of the orbitals of unpaired electrons. In both cases, it is reasonable to expect the manifestation of the $S\text{–}T_0$ mechanism of CIDNP formation.

In the former case, the $S\text{–}T_0$ mixing results from the existence of well-known nonradiative transitions.

Here, the energy of $J_e$ is counterbalanced by a pool of vibrational degrees of freedom.

The manifestation of an $S–T_0$ mechanism of CIDNP formation in high magnetic fields has been observed recently for biradicals containing Group 14 elements (Si, Ge).




At present, the investigation of CIDNP field dependencies for products of transformation of biradicals is the most reliable indirect method of demonstrating the involvement of biradicals in chemical reactions. The structure of the biradical can be inferred by comparison of observed with calculated field dependences. For this purpose, one might employ various approximations of the radical pair described in a number of specialized monographs.

Reference data show the examples of CIDNP manifestations of triplet (ground) and singlet (excited) states of carbenes. The triplet state is observed in the reactions of diphenylcarbene and methylene, while the singlet state has also been detected in the case of methylene. The analysis of CIDNP effects has led to the following conclusions. In contrast to the generally accepted viewpoint expressed in the chemical literature, not only the triplet but also the singlet state of carbenes reacts via hydrogen or halogen atom abstraction from the partner rather than insertion into the corresponding bond.

Recently, $^1H$ CIDNP effects have also been observed in the reactions of carbenes containing silicon and germanium with various scavengers. In this case, the radical stages have been detected in the reaction of singlet (ground) and triplet (excited) states of dimethylsilylene and -germylene.

### CIDNP as a method of investigation of spin and molecular dynamics in radical pairs

Research into spin dynamics is presented in numerous papers devoted to the definition of the range of values of $J_e$ in biradicals. This approach also allows an estimate to be made of the range of $J_e$ values in radical ion pairs.

A prominent example of the potential of the method for investigating molecular dynamics is the recent research into the process of protein folding by means of photo-CIDNP observed in stopped flow experiments. It has been shown that the use of the CIDNP technique in stopped flow experiments has certain advantages over the conventional NMR studies using flow systems. The sensitivity gain due to the CIDNP enhancement coefficient allows the folding processes to be studied on a time range two orders of magnitude shorter than those observable by mean of conventional NMR methods. In the process under study, CIDNP results from photoinduced electron transfer between a flavin additive and tyrosine and other amino acids of the protein (hen egg lysozyme). After the refolding, the CIDNP disappears, since after the conformational changes in protein structure amino acids become inaccessible to the flavin species.

At present, CIDNP is virtually a unique technique for the detection of so-called weak interactions (with the energy lower than several kT) between the reagents in chemical reactions. This includes charge transfer complexes, exciplexes and donor–acceptor complexes. As a rule, the involvement of these complexes is confirmed by the data on multiplicity of the reacting state obtained on the basis of CIDNP analysis (see corresponding sections in this article on the analysis of CIDNP effects). For example, the inconsistency between the multiplicity estimated on the basis of CIDNP data and that expected for the known set of photochemical parameters has justified the suggestion of a triplet exciplex in the photooxidation of Hantzsch ester by quinones.

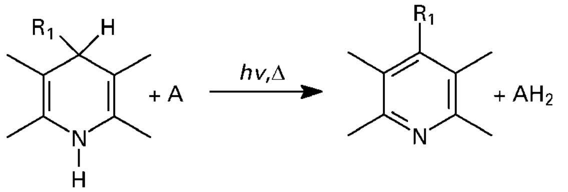
The alteration of the multiplicity in the reaction of silylene wit $CCl_4$ observed in the 7,7-dimethyl-7-silanorbornadiene derivative in the presence of triphenylphosphine has served as proof of the complexation between dimethylsilylene and $Ph_3P$.

### CIDNP in reactions involving biologically relevant molecules

The CIDNP technique has made a decisive contribution to the widely discussed question of the role of single electron transfer stages in biochemical processes. These are primarily the oxidation of organic substrates which occur under the action of UV irradiation and electron acceptors, as well as the enzymatic processes involving oxygenases.

The CIDNP technique has been used to unravel the elementary mechanisms of the oxidation of NADH coenzymes and their synthetic analogues, and substituted 1,4-dihydropyridines, in the presence of electron acceptors. The underlying mechanism of NADH transformation is the transition from the

dihydropyridine ring to pyridine.




In this case, the basic question is whether this process occurs via the single act of hydride transfer, or the redox reaction includes the sequence of several elementary stages involving paramagnetic species.

Application of $^{1}H$ and $^{15}N$ CIDNP techniques has allowed the demonstration that the process under study indeed includes the sequence of radical stages. For unsubstituted dihydropyridines, this is a chain of 'e–$H^{+}$–$H^{\bullet}$' transformations, while for substituted species and NADH, the sequence 'e–$H^{\bullet}$' has been observed.

The CIDNP technique has also been used to reveal the role of radical species in the processes of phototransformations of coenzyme $B_{12}$ model compounds and *all-trans*-retinal.

## List of symbols

$a$ = hyperfine interaction constant; $A$ = absorption; $E$ = emission; $I^*$ and $I_0$ = intensities of the polarized and equilibrium signal; $J$ = spin–spin coupling constant; $J_e$ = electron exchange interaction; $m$ = absolute enhancement factor of product polarization; $M$ = multiplet effect; $N$ = net CIDNP effect; S = singlet state; $t$ = time; $T_0$ = triplet zero state; $T_{1N}$ = nuclear relaxation time; $T_{1R}$ = nuclear spin–lattice relaxation time in a radical; $\gamma$ = parameter for location of the magnetic nuclei in the radicals; $\Delta g$ = difference in the $g$-factors of the radical with polarized nucleus and the radical partner; $\varepsilon$ = parameter for 'in-cage' or escape recombination; $\lambda$ = parameter for recombination from S- or T-state; $\mu$ = multiplicity.

*See also:* **Chemical Applications of EPR; Chemical Exchange Effects in NMR; EPR, Methods; Heteronuclear NMR Applications (Ge, Sn, Pb); NMR Pulse Sequences.**

## Further reading

Closs GL and Miller RJ (1978) Photoreduction and photodecarboxylation of pyruvic acid. Application of CIDNP to mechanistic photochemistry. *Journal of the American Chemical Society* **100**: 3483–3494.

Egorov MP, Ezhova MB, Kolesnikov SP, Nefedov OM, Taraban MB, Kruppa AI and Leshina TV (1991) $^{1}H$ CIDNP study of the addition reaction of $Me_2E$ (E = Si, Ge) to a carbon–carbon triple bond. *Mendeleev Communications* **4**: 143–145.

Goez M (1996) Optimization of flash CIDNP experiments. *Journal of Magnetic Resonance, Series A* **123**: 161–176.

Hore PJ, Winder SL, Roberts CH and Dobson CM (1997) Stopped-flow photo CIDNP observation of protein folding. *Journal of the American Chemical Society* **119**: 5049–5050.

Kruppa AI, Taraban MB, Leshina TV, Natarajan E and Grissom CB (1997) CIDNP in the photolysis of coenzyme $B_{12}$ model compounds suggesting that C–Co bond homolysis occurs from the singlet state. *Inorganic Chemistry* **36**: 758–759.

Kruppa AI, Taraban MB, Polyakov NE, Leshina TV, Lusis V, Muceniece D and Duburs G (1993) The mechanism of the oxidation of NADH analogues. 2. *N*-Methyl substituted 1,4-dihydropyridines. *Journal of Photochemistry and Photobiology A: Chemistry* **73**: 159– 63.

Kruppa AI, Taraban MB, Shokhirev NV, Svarovsky SA and Leshina TV (1996) $^{119}Sn$ CIDNP: Calculation and experiment. *Chemical Physics Letters* **258**: 316–322.

Kruppa AI, Taraban MB, Svarovsky SA and Leshina TV (1996) Paramagnetic intermediates in the photolysis of 2-methylpropanoyltripropylstannane studied by means of multinuclear CIDNP. *Journal of the Chemical Society, Perkin Transactions 2* **10**: 2151–2156.

Lawler RG and Halfon MA (1974) A flow reactor for use with an unmodified high resolution NMR spectrometer. Application to CIDNP. *Review of Scientific Instruments* **45**: 84–86.

Lepley AR and Closs GL (1973) *In*: Closs GL (ed) *Chemically Induced Magnetic Polarization*. New York: Wiley.

Muus LT, Atkins PW, McLauchlan KA and Pedersen JB (1977) *In*: Pedersen JB (ed) *Chemically Induced Magnetic Polarization*. Dordrecht: Reidel.

Polyakov NE and Leshina TV (1990) Study of CIDNP field dependencies in the photoreduction of quinones by amines. *Journal of Photochemistry and Photobiology A: Chemistry* **55**: 43–51.

Richard C and Granger P (1974) *Chemically Induced Dynamic Nuclear and Electron Polarization—CIDNP and CIDEP*. Berlin: Springer Verlag.

Roth HD and Herbertz T (1993) The structure of simple vinylcyclopropane radical cations. Evidence for conjugation between alkene and cyclopropane group. *Journal of the American Chemical Society* **115**: 9804–9805.

Salikhov KM, Molin YuN, Sagdeev RZ and Buchachenko AL (1984) *In*: Molin YuN (ed) *Spin Polarization and Magnetic Effects in Radical Reactions*. Amsterdam: Elsevier.

# Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism

**Christine L Maupin** and **James P. Riehl**, Michigan Technological University, Houghton, MI, USA

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Applications**

The measurement of the usually small net circular polarization in the luminescence from chiral molecules began with the pioneering studies of Professor L.J. Oosterhof at the University of Leiden in the late 1960s. In recent years this technique has developed into a reliable and useful spectroscopic tool for the study of a wide variety of chemical systems. The phenomenon has been referred to as circularly polarized luminescence (CPL), circularly polarized emission (CPE), and by other names. Here the most common acronym, CPL, is used to describe this experimental technique. The measurement of the differential emission intensity due to differences in the absorption of left versus right circularly polarized light is also discussed. This is commonly referred to as fluorescence-detected circular dichroism (FDCD). Whereas FDCD spectroscopy (like conventional circular dichroism (CD)) probes the chiral structure of the ground state, CPL is a probe of the chirality of the excited state. These types of experiments are not redundant, since changes of geometry do occur on electronic excitation. Furthermore, both of these experimental techniques provide information concerning molecular dynamics and energetics that occur between the time of excitation and emission.

## Instrumentation for measurement of circularly polarized luminescence

Circularly polarized luminescence spectroscopy (CPL) has been used for more than 30 years as a probe of the molecular stereochemistry of chiral luminescent molecules and complexes. Unlike CD, CPL measurements have not, as yet, become routine. This is primarily due to the lack of a commercially available instrument. The measurement of the generally small net circularly polarized luminescence, relative to the total luminescence, is difficult owing mainly to the presence of optical artefacts from various sources, and to electronic problems associated with the measurement of weak differential signals. The first CPL measurements were made using analogue phase-sensitive detection systems. Measuring the generally small CPL signal in this way is particularly difficult because of electronic problems associated with lock-in amplifiers. All of the current instruments use photomultipliers operating in photon-counting mode, and some type of gated photon counting system to determine the circular polarization. This is the type of system that will be described here.

### Steady-state CPL

Since no commercial instrumentation is available, all CPL measurements that have been published have been measured on instruments that were designed and constructed in individual laboratories. The basic configuration of a system designed to measure CPL is given in **Figure 1**. A laser system or a UV lamp used in combination with an excitation monochromator or a combination of filters can be used as the excitating source. Care should be taken in ensuring that the polarization of the excitation light is such that there will be no linear polarization in the emission since this is known to cause artefacts in the detection of circular polarization. In most luminescence experiments the emission is collected at an angle of 90.0° from the direction of the excitation beam to avoid detecting light from the excitation source. In some applications of CPL spectroscopy, it is necessary to control the excitation polarization. **Figure 1**, for example, shows the optics necessary to produce circularly polarized excitation.

The luminescence from the chiral sample is first directed through a photoelastic polarization modulator (PEM) and then a linear polarizer before it passes through any focusing lenses or filters because these optical components may be birefringent. The linear polarizer ensures that the detected light entering the monochromator will always be polarized in the same direction. This prevents artefacts associated with the emission monochromator being polarization sensitive. The PEM is composed of an isotropic, clear optical material that produces an oscillating birefringence at a fixed frequency. In CPL measurements, the PEM is operated as an electrically controlled advancing-retarding quarter-wave plate at a predetermined frequency ($f$). This introduces a periodic phase shift at frequency $f$, such that either right or

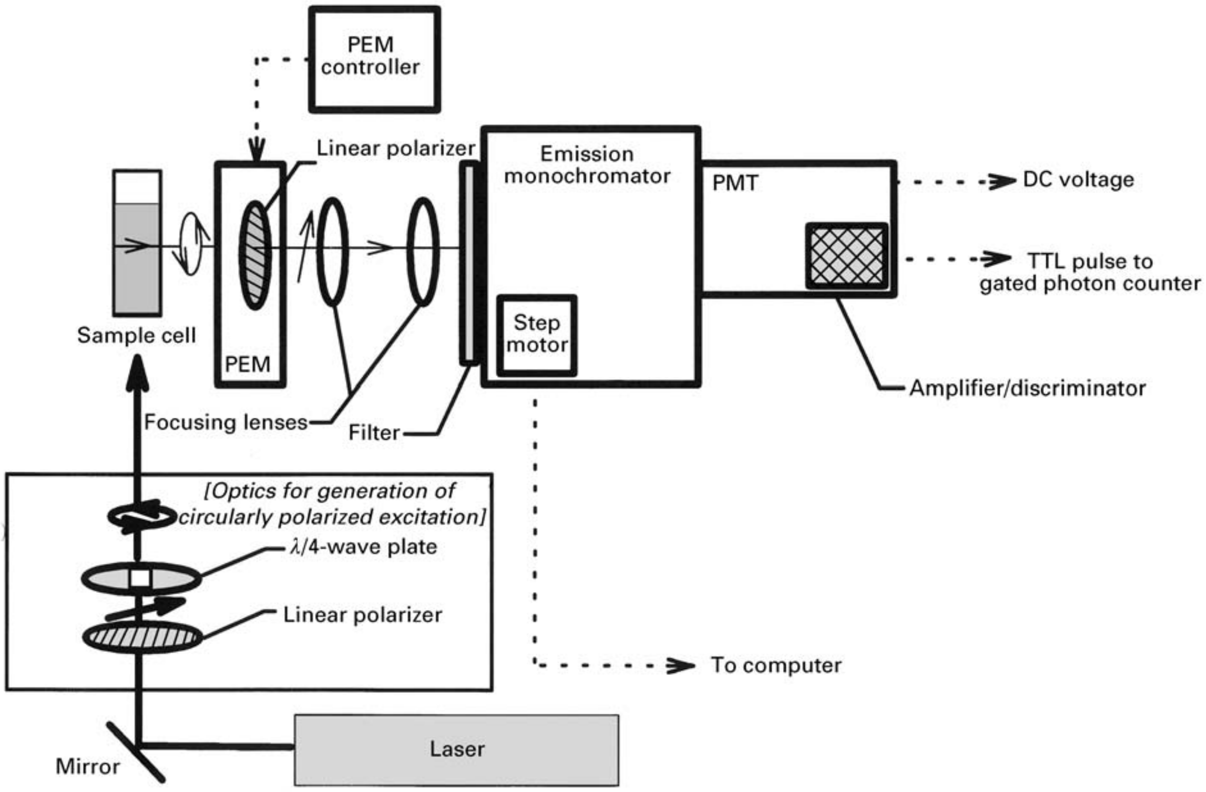



**Figure 1** Schematic diagram of an instrument designed to measure circularly polarized luminescence.

left circularly polarized emitted light is converted to linearly polarized light oriented at 45° to the PEM crystal axis. Therefore, the PEM and linear polarizer allow the passage of alternately left then right circularly polarized light. The emitted linearly polarized light is then focused onto the entrance slits of an emission monochromator. Background signal from stray excitation and room light can be reduced by placement of a long pass filter in front of the monochromator. The light exiting the monochromator is detected by a photomultiplier tube operating in photon-counting mode. The signal from the photomultiplier tube is sent into an amplifier-discriminator. These devices are capable of providing simultaneous dual output for TTL pulses that are generated for photon spikes, as well as a voltage proportional to the total emission intensity.

The TTL pulses are analysed by a photon counter using the reference signal at frequency *f* to direct the pulses into two separate counters. The time dependence of the phase shift imparted by the PEM is such that it is exactly quarter-wave retarding or advancing only at the maximum or minimum of the periodic oscillation. For this reason, it is common to select finite time-windows centred at the peak of the modulation cycle for the counting of photon pulses, and not to count the pulses in the 'dark' interval. Even when these windows are set to collect 50% of the incoming TTL pulses, only a slight error (< 5%) in the accuracy of this measurement is introduced. On alternate half-cycles of the modulation cycle, the TTL pulses are directed into two separate counters corresponding to the transmission of left ($I_L$) or right ($I_R$) circularly polarized light through the PEM. In CPL spectroscopy one commonly reports the ratio of the differential emission intensity at wavelength, $\lambda$.

$$\Delta I(\lambda) = I_L(\lambda) - I_R(\lambda) \qquad [1]$$

to the total emission intensity

$$I(\lambda) = \frac{1}{2}(I_L(\lambda) + I_R(\lambda)) \qquad [2]$$

We thus define the luminescence dissymmetry ratio, $g_{lum}$, as follows:

$$g_{lum}(\lambda) = \frac{\Delta I(\lambda)}{\frac{1}{2}I(\lambda)} = \frac{I_L(\lambda) - I_R(\lambda)}{\frac{1}{2}(I_L(\lambda) + I_R(\lambda))} \qquad [3]$$

$I$, $\Delta I$, and $g_{lum}$ may all be calculated from the values of the two counters. Because of the difficulty in

determining absolute emission intensities, it is most common to report CPL results in terms of $g_{lum}$.

### Calibration and standards

The determination of the precise state of polarization of the circularly polarized emission, and of the accuracy of a gated photon-counting system, may be accomplished by a number of methods. The simplest and most convenient approach for calibration is to use a standard to measure $g_{lum}$. For example, several groups have used the chiral and commercially available NMR chiral shift reagent tris(3-[trifluoromethylhydroxymethylene]-*d*-camphorato)europium-(III) (Eu(facam)$_3$; Aldrich) in dry dimethyl sulfoxide (DMSO). This calibration can be performed by exciting the solution of Eu(facam)$_3$ in the visible or ultraviolet, and measuring $g_{lum}$ at several different wavelengths. This complex has known $g_{lum}$ values of +0.072 at 612.8 nm for the strong emissive transition $^5D_0 \rightarrow {}^7F_2$, and –0.78 for the $^5D_0 \rightarrow {}^7F_1$ transition at a wavelength maximum of 595 nm. Calibration of the gated photon counting system could also be accomplished by passing completely left and right circularly polarized light through the PEM. This should in theory yield $g_{lum}$ values of +2 and –2, respectively. However, the value of exactly 2 (or –2) is never obtained owing to imperfect circular polarization and the inaccuracies associated with the finite time-window for the gated photon counter, as described above.

One of the main advantages of the photon-counting system over lock-in detection is the fact that, if the instrument is working properly, i.e. obeying Poisson photon statistics, then it can be shown that the distribution of measured $g_{lum}$ values should obey a Gaussian distribution. It can also be shown that the standard deviation in $g_{lum}$ is equal to the inverse square root of the total number of photon counts:

$$\sigma = \sqrt{\frac{1}{N}} \qquad [4]$$

Measuring a large number of $g_{lum}$ values and seeing whether the distribution is Gaussian is another convenient way of checking whether the instrument is working properly.

The intensity of luminescence will obviously vary greatly from one sample to the next. Some samples will be highly luminescent (i.e. give high count rate) while others will be weakly luminescent (i.e. give low count rate). For samples that are highly luminescent, accurate $g_{lum}$ measurements can be achieved in a short time. For example, if a sample is emitting $10^6$ photons s$^{-1}$, then using Equation [4], $g_{lum}$ with a standard deviation of $10^{-4}$ requires 100 s to be measured. Weakly luminescent samples require much longer collection times to achieve the same standard deviation. If the count rate is only $10^4$ photons s$^{-1}$, then more than 5 h of data collection are required.

### Time-resolved CPL

Advances in digital counting technology and high-speed computers have allowed several research groups to expand the measurement of CPL into the time domain. These advances have been led by the research groups of Professor H.P.J.M. Dekkers at the University of Leiden and Professor F.S. Richardson at the University of Virginia. The instruments used are based on modifications of steady-state instruments such as the system described above, with the major instrumental change being that the excitation source must be pulsed. The principal problem in time-resolved CPL is how to measure the decay of the polarization with an instrument based on a polarization modulator that may be oscillating much more rapidly than the lifetime of the decay. Two approaches to this problem have been successful. If the precise phase of the polarization modulator is known, then an excitation pulse can be coordinated with the periodic oscillation in such a way that the flashes occur at precise times in both halves of the modulation cycle. To ensure that the device is analysing for circular polarization at each decay point, the time between the excitation and the decay measurements is varied. Just as described above, for each finite time interval centred at time $t$ in the decay, photon pulses corresponding to $I_L(t)$ and $I_R(t)$ are accumulated, and by simple calculation, $g_{lum}(t)$ is determined. Alternatively, it is possible to completely decouple the excitation flashes from the modulation cycle and accumulate enough flash events that the complete modulation cycle is sampled for every finite time channel. This latter method has the advantage that the complete time decay for each excitation pulse contributes to the measurement.

## Applications for circularly polarized luminescence

Since, as described, there are no commercial sources of CPL instrumentation, the applications of this spectroscopic technique have been limited to the study of specific research goals of those groups who have built the instrumentation. Nevertheless, the applications have covered a fairly wide area of chemistry, including organic, inorganic, polymer and

biochemical systems. It should be mentioned that most applications of CPL have involved forbidden spectroscopic transitions. This is because, in these experiments, one is always measuring small differences in luminescence intensity. These differences are obviously easier to detect if the individual intensities themselves are not very large, i.e. if they are not associated with allowed transitions.

This section highlights examples of different kinds of applications, and the reader is referred to one of the review articles listed at the end of this article for a more detailed discussion of specific chemical systems in which CPL has been used.

### Applications to organic chromophores

Many of the early applications of CPL involved analysis of the circular polarization from $^1n\pi^*$ states of chiral ketones. Pyramidal distortion of the >C=O group upon going from the ground state to the excited/emitting state in CD spectroscopy is well known, and comparative CD and CPL studies have been used to probe these excited state geometry changes. As an example, **Figure 2** shows the CPL spectrum (upper curve) and total luminescence spectrum (lower curve) for an α-diketone, *d*-camphorquinone dissolved in ethanol. Note that the form of presentation of CPL results parallels that usually seen for CD studies. In this example, the luminescence dissymmetry ratio, $g_{lum}$, at the peak maximum for this $n \leftarrow \pi^*$ transition is approximately –0.01. $g_{abs}$ measured from the CD spectrum of this compound for the $n \rightarrow \pi^*$ transition is approximately –0.009, and the differences between these measurements reflects geometry changes within the dicarbonyl moiety upon electronic excitation. In bichromophoric molecules one may use the well-known exciton coupling model to interpret CPL results in exactly the same way that this model is applied to CD measurements. In a very few cases, for example, β,γ-enones, chirality rules have been developed that relate the sign and magnitude of the CPL signal for closely related compounds to specific aspects of chiral structure. These attempts at spectra–structure correlations are clearly not as well developed as they are for CD spectroscopy.

### Inorganic applications

As described above, CPL spectroscopy is particularly suited for the study of forbidden transitions, so the potential for applying this technique to intraconfigurational transitions in metal complexes is very high. In the case of transition metals, however, applications of CPL have been quite limited. The reasons for this are many. Certainly, the preparation of resolved chiral transition metal complexes is quite difficult,

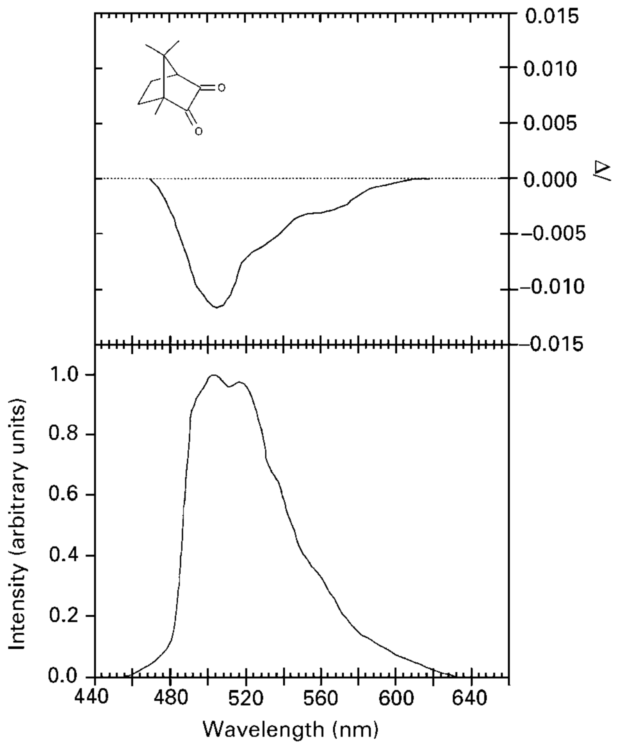



**Figure 2** CPL spectrum (upper curve) and total luminescence spectrum (lower curve) for *d*-camphorquinone in ethanol.

and often the racemization under UV or visible illumination may not be neglected. In addition, not many transition metal complexes are luminescent in liquid solution. The luminescence from 'frozen' samples can be quite high, but measurement of CPL under conditions where the sample is highly scattering and nonisotropic are fraught with experimental difficulties. There have been a number of interesting applications of CPL to chiral transition metal complexes. For example, Ru(1,10-phenanthroline)$_3^{2+}$ may easily be prepared and resolved. This compound possesses $D_3$ symmetry and, as such, exists as Δ and Λ enantiomers. The broad luminescence from this complex centred at 600nm under UV excitation is fairly intense with a measured $g_{lum}$ value of approximately 0.0007. $g_{lum}$ varies significantly in the low-wavelength shoulder of the emissive transition, indicating that the precise nature of the metal–ligand states involved are quite complicated. There have also been a few examples of CPL measurements from crystalline samples; however, the requirements here are that the sample be chiral and luminescent, and occur in a space group that is uniaxial. This combination is very restrictive.

By far the most common class of inorganic compounds that have been studied by CPL spectroscopy are luminescent lanthanide complexes. There are several reasons for this special interest. Complexes of

Tb(III), Eu(III), and to lesser extent, Dy(III), are luminescent in the visible region of the spectrum, particularly if water molecules can be excluded from the first coordination sphere. The transitions involved are isolated $f \leftrightarrow f$ transitions, and even though the absorptions are very weak, all three of these ions have absorptions in the visible, which overlap convenient strong laser lines. In many cases it is also possible to excite allowed transitions in the ligand orbitals, and populate the lanthanide ion energy levels through efficient energy transfer. Lastly, and most importantly, the states derived from the various f-electron configurations are well characterized by the total angular momentum quantum number $J$, and transitions that obey magnetic dipole selection rules, that is $\Delta J = 0, \pm 1$, are often associated with very large dissymmetry ratios.

An example of CPL from a chiral lanthanide complex is given in **Figure 3**. This spectrum was obtained by exciting the sample of the macrocyclic Tb(III) complex with 488 nm excitation from an argon-ion laser. The spectral region displayed corresponds to a transition from the $^5D_4$ excited state to the $^7F_5$ state in the ground-state manifold. This transition satisfies the magnetic dipole selection rules introduced above. The ground state for Tb(III) is the $^7F_6$ state. The various peaks shown in the spectrum correspond to different crystal field transitions within the term-to-term transition. The chiral pendant amines in the macrocyclic ligand are chiral, and this results in only one of the possible $C_4$ orientations around the central metal ion being formed. Confirmation of this latter conclusion may be obtained by varying the incident polarization from left to right circularly polarized, and the result for this system is that no effect is seen in the CPL spectrum. As can be seen in **Figure 3** $|g_{lum}|$ for magnetic dipole-allowed transitions may be very large. The values seen in **Figure 3** approach 0.4; this should be compared to the results presented in **Figure 2** (0.01), which represents approximately the maximum value that has been measured for organic chromophores.

Unlike transition metal compounds such as $Ru(phen)_3^{2+}$, which are composed of achiral ligands but can be resolved into chiral enantiomers, no such examples of the resolution of chiral lanthanide compounds with achiral ligands have appeared in the literature. This is because lanthanide complexes are generally more labile than transition metal complexes. Just like transition metal complexes, however, it is possible in some cases to perturb the ground-state equilibrium concentration through the addition of a chiral environment compound. This is sometimes referred to as the Pfeiffer effect. An example of such an experiment is presented in **Figure 4**. In this experiment, D- and L-enantiomers of glucose have been added to aqueous solutions of Tb(2,6-pyridine-dicarboxylate) ($Tb(DPA)_3^{3-}$). This complex is known to possess $D_3$ symmetry, so it occurs in solution as a mixture of Δ and Λ enantiomers.

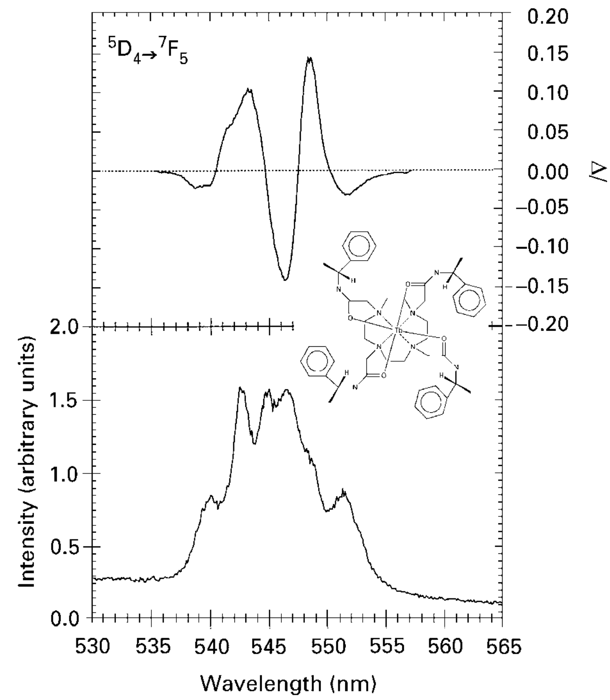



**Figure 3** CPL spectrum (upper curve) and total luminescence spectrum (lower curve) for an aqueous solution of a chiral macrocyclic Tb(III) complex.

## Racemic mixtures

Even though the ground-state distribution of a mixture of chiral molecules is racemic, in some cases it is still possible to measure the CPL if the excited state can be enriched in one of the enantiomers. If, for example, the excitation beam is circularly polarized, one of the enantiomers with be excited preferentially over the other, and the excited state will not be racemic. This experiment, which has been referred to as CPE/CPL, relies on discrimination in absorption and emission. Obviously, it is also important that the nonracemic excited-state distribution created by the excitation beam is maintained during the excited state's lifetime. An example of such an experiment is given in **Figure 5**. The CPL spectrum given in **Figure 5** was collected while exciting an aqueous solution of $Eu(DPA)_3^{3-}$ with circularly polarized 472.8 nm light from an argon-ion laser. These complexes do not racemize on the emission timescale, so one may write down the following equation relating the measured luminescence dissymmetry ratio for

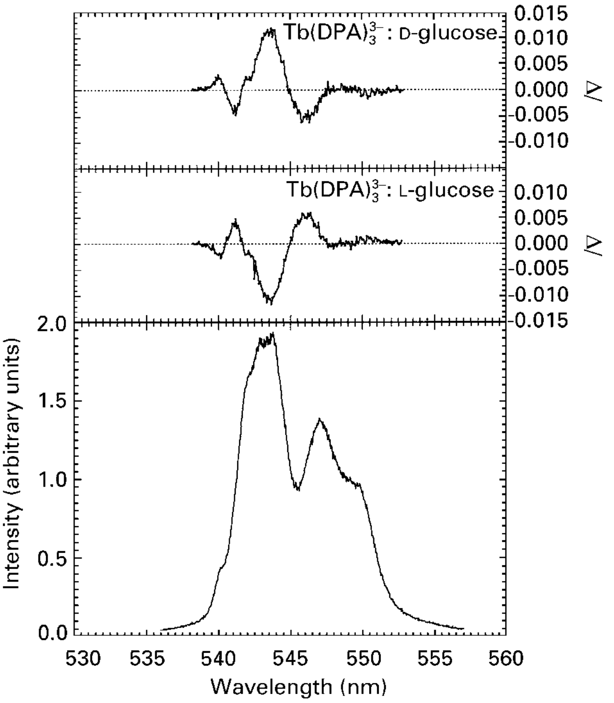



**Figure 4** CPL spectra and total luminescence (lower curve) for an aqueous solution of Tb(DPA)$_3^{3-}$ containing 0.01M D-glucose (upper curve) and L-glucose (middle curve).

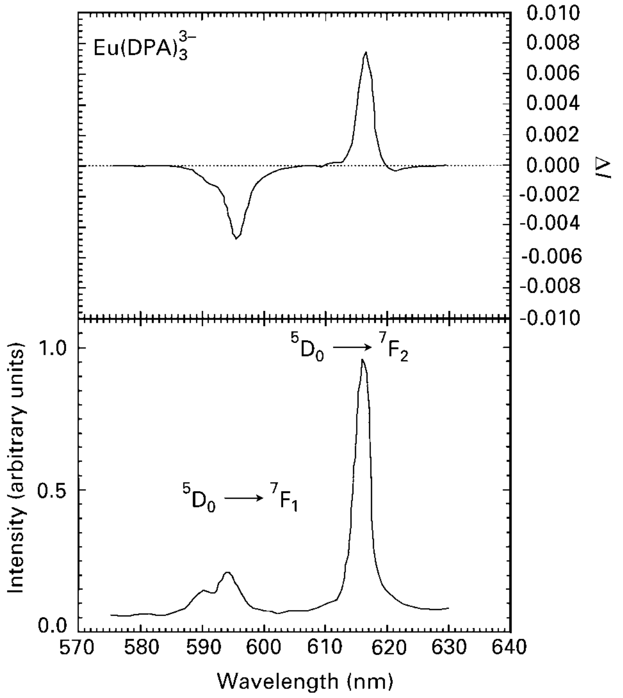



**Figure 5** CPL spectrum (upper curve) and total luminescence (lower curve) for an aqueous solution of Eu(DPA)$_3^{3-}$ using circularly polarized 472.8 nm excitation.

left circularly polarized excitation to the intrinsic absorption and emission dissymmetry ratio for one of the enantiomers as follows:

$$g_{\mathrm{lum}}^{\mathrm{L}}(\lambda) = \frac{1}{2} g_{\mathrm{abs}}^{\Delta}(\lambda') g_{\mathrm{lum}}^{\Delta}(\lambda) \qquad [5]$$

In this equation we have been careful to denote that the wavelength of excitation ($\lambda'$) is different from the wavelength of emission detection.

Measurement of CPL from racemic mixtures is not a technique that can be applied to all racemic solutions. For moderately luminescence systems $g_{\mathrm{lum}}$ values of approximately $10^{-4}$ are measurable. This means that the intrinsic absorption and emission dissymmetry ratios need to be on the order of $10^{-2}$ for this experiment to be successful. Although there have been a couple of examples using this technique in organic systems, by far the most widely studied systems are racemic lanthanide complexes, such as given in **Figure 5**, because of the large $g_{\mathrm{lum}}$ and $g_{\mathrm{abs}}$ values that may exist for certain f $\leftrightarrow$ f transitions. It is also useful to perform these experiments in a time-resolved mode and as a function of temperature to determine racemization rate constants.

There are other ways to generate enriched excited states from racemic ground states. If, for example, a chiral excited state quencher molecule is added to a racemic solution, the rate of quenching of the two enantiomers may be different. This kind of experiment has been performed both in the steady-state and time-resolved modes. At time $t = 0$ following an unpolarized excitation pulse directed at a racemic mixture, the concentration of excited enantiomers will be equal. The diastereomeric selective quenching results in a time-dependent population that becomes increasingly richer in one of the enantiomers. Results from a typical experiment of this type are given in **Figure 6**. Here $g_{\mathrm{lum}}$ is plotted as a function of time for an aqueous solution of Tb(DPA)$_3^{3-}$ into which a small amount of Δ-Ru(phen)$_3^{2+}$ has been added. The different quenching rates for the diastereomeric pairs depend in a complicated way on mutual diffusion, reorientation, electronic overlap and diastereomer structure. These types of experiments may be performed with varying donor–quencher pairs, in different solvents, and as a function of temperature and pressure to probe intimate details of this excited-state chiral discrimination.

### Biological and polymer systems

Almost all biological systems are chiral, so it is natural to expect that CPL studies involving luminescent biological material will be important. Unlike CD, where all of the absorbing chromophores contribute

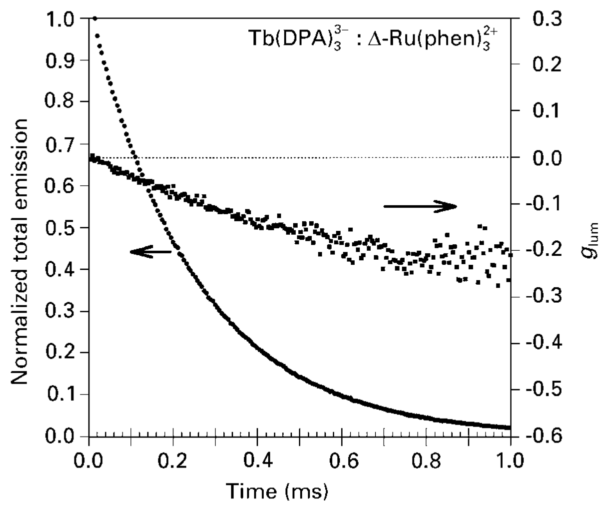



**Figure 6** Time dependence of $g_{lum}$ at 544 nm following excitation at 330 nm for an aqueous solution of Tb(DPA)$_3^{3-}$ into which a small amount of Δ-Ru(phen)$_3^{2+}$ has been added.

to the resultant CD spectrum, luminescence from protein and other biological systems is generally only seen from a very small number of emitting chromophores. Thus, CPL studies involving biological systems generally only provide specific structural information localized around the emitting chromophore, and not the longer-range information on secondary and tertiary structure that is routinely obtained from CD studies. This is true both for intrinsic protein fluorescence and in studies in which metal ion or organic fluorescence probes are used.

Some of the very early applications of CPL spectroscopy involved chiral polymeric systems. In particular, the CPL from achiral dye molecules dissolved in cholesteric liquid crystals has been used to probe chirality changes. CPL from chromophores attached to a chiral poly-amino acid may also be used to study exciton coupling between aromatic chromophores. In these polymeric systems, it is often observed that $g_{lum}$ is quite large. In some cases it is so large, in fact, that detection using static quarter-wave plates is possible.

## Instrumentation for measurement of fluorescence-detected circular dichroism

In fluorescence-detected circular dichroism (FDCD) the difference in total emission intensity is used to monitor differences in circularly polarized absorption. This spectroscopic technique was developed in the early 1970s in the laboratory of Professor I. Tinoco at the University of California, Berkeley. In the simplest application of FDCD a photomultiplier tube is placed at 90° to the direction of emission detection in a modified commercial or custom-built CD instrument and the luminescence intensity in phase with the oscillating left and right circularly polarized absorption beam is monitored. For many applications, including recent uses of FDCD as detectors in HPLC, this type of experimental setup is sufficient. The major technical challenge in the construction of an FDCD instrument is associated with the fact that it is very difficult to produce pure circularly polarized excitation without any residual polarization. Furthermore, if this linear polarization varies in phase with the incident circular polarization, serious artefacts may result. To reduce this problem, a somewhat more complicated experimental setup is required. A schematic diagram of an FDCD instrument designed to deal with the linear polarization problem is shown in **Figure 7**.

The incident (absorption) beam is wavelength-selected by an excitation monochromator and then converted to alternately left and right circular polarization by use of a linear polarizer and quarter-wave modulator. The quarter-wave modulator may be a PEM as described above for CPL measurement, or a Pockels cell. This luminescence in phase with the incident polarization modulation is collected by two photomultiplier tubes (PMT1 and PMT2) oriented at right angles to each other and the incident beam. If the excitation is linearly polarized, then the sum of

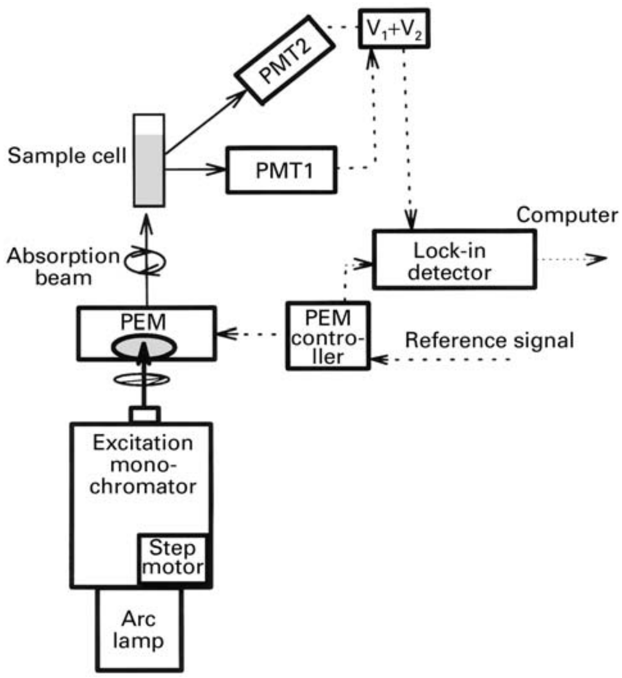



**Figure 7** Schematic diagram of FDCD instrument.

the fluorescence intensities from the two detectors due to this artefact gives a signal ($V_1 + V_2$) that is independent of the linear polarization. So even if this polarization changes as the incident polarization varies from left to right, the resultant artefact signal subtracts out of the difference in fluorescence intensity, $F_L - F_R$. The artefact signal does not subtract out of the total luminescence intensity, $F_L + F_R$, but, if present, it may be reduced by rotation of the PEM (or adjustment of the Pockels cell). The signals from the photomultiplier tubes are added and either these are input to a lock-in detector (as shown) or, if digital counting techniques are employed, the TTL pulses are analysed by a gated photon-counting system as described previously.

To relate the measurement of FDCD to conventional CD, it is necessary to take into account that, if differential absorption of the circularly polarized excitation beam is taking place, then the intensity of the absorption beam as a function of distance into the sample may be different for the two circular polarizations. The relationship between CD and FDCD measurements for isotropic systems containing only one optically active fluorophore may be expressed as

$$\Delta\varepsilon = \varepsilon_{\rm L} - \varepsilon_{\rm R} = \frac{g_{\rm F}\left(1 - 10^{-A}\right)}{cd\, 10^{-A} \ln 10} \qquad [6]$$

where $A$ denotes the absorbance of the sample, $c$ is the concentration, and $d$ is the path length. Processing of FDCD signals under conditions where there are more than one chiral chromophore is more complicated. Calibration of an FDCD instrument may be accomplished by measuring the signal for an achiral fluorescent molecule to ensure that a value of zero is obtained, and then by measuring a standard compound with a known CD such as *d*-camphorsulfonic acid.

## Applications of fluorescence-detected circular dichroism

FDCD measurements have the potential for providing valuable spectroscopic and structural information for systems in which conventional CD measurements are problematic. Of particular interest are complex systems containing multiple absorption chromophores but only one fluorescent chromophore. It is also possible to use the intrinsically higher sensitivity of luminescence detection to determine the CD of samples for which only very small quantities are available. FDCD has also shown promise in high-pressure liquid chromatographic detection of luminescent chiral molecules, although no commercial applications of this technology are yet available.

Although interesting and important applications for the FDCD technique have been undertaken, the technical problems associated with the generation of exactly equal intensities of circularly polarized exciting light, and the artefacts associated with residual linear polarization, have severely limited the widespread use of this technique. These problems are not insurmountable, and it is anticipated that modification by the manufacturers of commercial CD instruments to include FDCD attachments should result in considerably more applications of this technique to interesting and important chemical problems.

## List of symbols

$A$ = absorbance; $c$ = concentration; $d$ = path length; $f$ = frequency; $g_{\rm abs}$ = absorption dissymmetry ratio; $g_{\rm lum}$ = luminescence dissymmetry ratio; $I$ = light intensity; $J$ = total angular quantum number; $t$ = time; $\varepsilon$ = extinction coefficient; $\lambda$ = wavelength; $\sigma$ = standard deviation.

*See also:* **Biochemical Applications of Fluorescence Spectroscopy; Biomacromolecular Applications of Circular Dichroism and ORD; Chiroptical Spectroscopy, Emission Theory; Induced Circular Dichroism; Luminescence, Theory; ORD and Polarimetry Instruments; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading

Richardson FS and Riehl JP (1977) Circularly polarized luminescence spectroscopy. *Chemical Reviews* 77: 773–792.

Riehl JP (1994) Excited state optical activity. In: Purdy N and Brittain HG (eds) *Analytical Applications of Circular Dichroism*. Amsterdam: Elsevier.

Riehl JP and Richardson FS (1986) Circularly polarized luminescence spectroscopy. *Chemical Reviews* **86**: 1–15.

Riehl JP and Richardson FS (1993) Circularly polarized luminescence. *Methods in Enzymology* **226**: 539–553.

Steinberg I (1975) Circular polarization of fluorescence. In: Chen RF and Edelhoch H (eds) *Biochemical Fluorescence: Concepts*. New York: Marcel Dekker.

# Cluster Ions Measured Using Mass Spectrometry

**O Echt,** University of New Hampshire, Durham, NH, USA
**TD Märk,** Leopold Franzens Universität, Innsbruck, Austria
Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
| --- |
| **Applications** |

## Introduction

Clusters are defined as homogeneous aggregates of atoms ($A_n$) or molecules ($M_n$), although mixed (heterogeneous) aggregates ($X_nY_m$ or $X_nY^+$) are also considered as clusters. The size range of interest starts at the dimer ($n = 2$) and reaches up to $10^6$ to $10^8$ constituents per cluster. Depending on the nature of the building blocks, the binding energy between the constituents can be rather weak, as in the case of van der Waals bound rare gas clusters (10–100 meV), or rather large, as in the case of covalently bonded carbon clusters (several eV). Gas-phase clusters can be produced either by fragmentation of solid material (laser desorption or sputtering into vacuum) or by aggregation of single particles (supersonic jet expansions, gas aggregation); the clusters may be either neutral or ionized. Owing to the statistical nature of these methods, the cluster sample (beam) usually comprises a broad distribution of different cluster sizes (a notable exception being the formation of $C_{60}$). Thus virtually all experiments – in particular those concerned with the properties of a single cluster size, or with the evolution of cluster properties as a function of size – rely on mass spectrometry as a major tool for investigating this unique form of matter.

Besides high mass resolution for the proper identification of a specific cluster ion (**Figure 1**), one of the most important requirements for cluster studies is a large mass range (see **Figure 2** which shows a photoionization time-of-flight mass spectrum of caesium oxide clusters). This has led to the sophisticated and improved use of established techniques and to the innovative development of new concepts in mass spectrometry (electrospray ion sources, reflection type mass spectrometers, etc.), which nowadays are also of great use in other fields such as medicine and biology for the investigation of large biomolecules. Besides identification of cluster ions and cluster ion distributions, the other major use of mass spectrometry in cluster science is the selection of mass-analysed monodisperse cluster ions in order to allow the

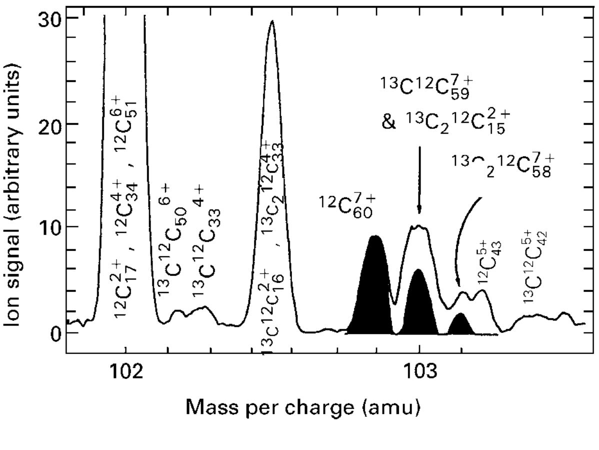



**Figure 1** Section of high resolution $C_{60}$ mass spectrum obtained by multistep electron impact ionization. The black peaks are calculated isotopomers ofthe septuply charged $C_{60}^{7+}$ ion normalized to the first isotopomer of the experimental run. After Scheier P and Märk TD (1994) *Physical Review Letters* **73**: 54.

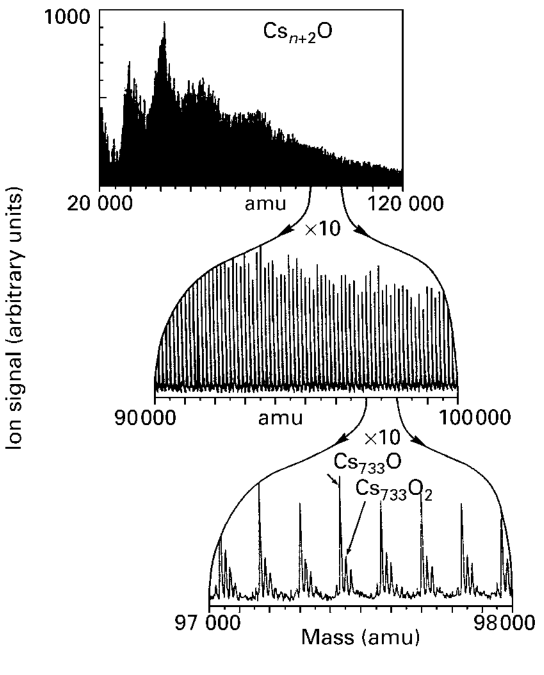



**Figure 2** Section of mass spectra for Cs/$O_2$ clusters using photoionization time-of-flight mass spectrometry shown on three different mass scales. After Martin TP, Bergmann T, Näher U, Schaber H and Zimmermann U (1994) *Nuclear Instruments and Methods* **B88**: 1.

investigation of cluster ion properties such as the structure, energetics, stability and reactivity. While size-selected neutral cluster targets are usually unavailable, mass spectrometry plays a crucial role in their investigation as well. For example, the information gained from photoelectron spectroscopy of mass-selected metal cluster anions pertains to the neutral species. Likewise, the sizes of neutral clusters that are interrogated by charge transfer reactions or resonance-enhanced ionization can be determined in a mass spectrometer, provided fragmentation does not occur.

## Atomic structure

The geometric arrangement of atoms is of paramount interest in cluster science. Mass spectrometry has provided essential clues to cluster structure and morphology, even though the approach is indirect. In short, one attempts to relate anomalies in the size distribution of clusters to anomalies in their properties, such as stability, reactivity or ionization energy.

**Figure 3** shows a mass spectrum of krypton clusters. Neutral clusters are grown by expansion of the gas through a nozzle into vacuum. Several abundance anomalies, often called 'magic numbers', are observed. Two questions arise: (1) Assuming the numbers indicate enhanced stability, is there a growth sequence that can consistently explain them? (2) Do the anomalies exist before ionization, or are they caused by dissociation following ionization?

Atoms in rare gas solids arrange in a close-packed structure, each atom being surrounded by 12 nearest neighbours. Clusters of rare gases are expected to assume compact shapes. Hence, it is reasonable to assume that clusters grow layer by layer around a core comprising 13 atoms, and that closed shells are particularly stable. The abundance of these clusters will be enhanced if clusters are warm enough to evaporate atoms.

However, the 13-atom core may be either a cuboctahedron cut out of the fcc crystalline solid, or an icosahedron (**Figure 4**). Either way, the total number of atoms in complete-shell clusters would be 13, 55, 147, 309, 561; these cluster sizes are, indeed, particularly abundant in the mass spectra. **Figure 3** reveals several additional magic numbers. They indicate closure of subshells that occurs whenever a facet of the cluster is covered by a layer of atoms. A cuboctahedral cluster has 14 facets, the icosahedral one has 20, leading to distinct subshell closures. For argon through xenon in this size range, only the

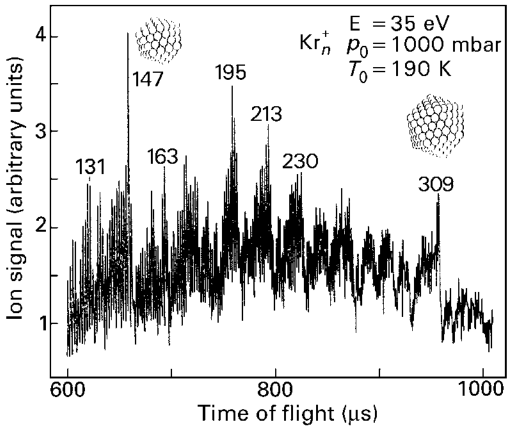



**Figure 3** Mass spectrum of krypton clusters, grown in a supersonic expansion and ionized by electron impact. The cluster size and assumed cluster structure of particularly prominent mass peaks is indicated.

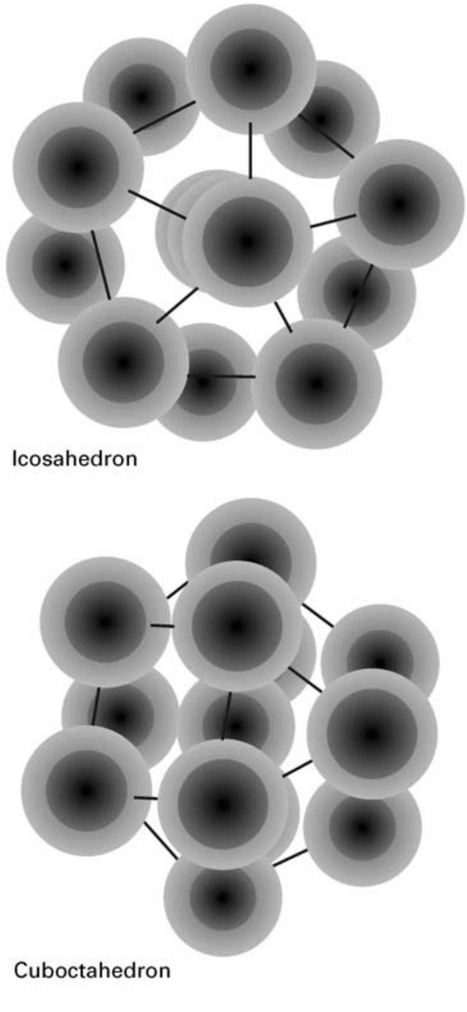



**Figure 4** Schematic view of a 13-atom icosahedron and a 13-atom cuboctahedron.

icosahedral growth sequence can explain the observed anomalies.

Concerning ionization, theory and a large number of experiments indicate that it is essentially impossible to 'softly' ionize a rare gas cluster or, in fact, any cluster with van der Waals bonding. Dissociation will lead to enhanced intensities for particularly stable charged clusters; metastable decay will cause the relative intensities to become time dependent (**Figure 5**). Likewise, mass spectra of small clusters formed in the charged state, e.g. by sputtering, may deviate considerably from spectra of clusters formed in the neutral state. For larger clusters, however, neither the system (Ar, Kr, Xe) nor the cluster formation/ionization process affects the anomalies.

Magic numbers in spectra of calcium and strontium clusters also indicate icosahedral growth, while aluminium clusters appear to adopt an fcc structure with octahedral morphology. Magic numbers observed when metal atoms are grown around a cluster of known geometry, $C_{60}$, nicely confirm the power of the closed-shell or closed-subshell concept.

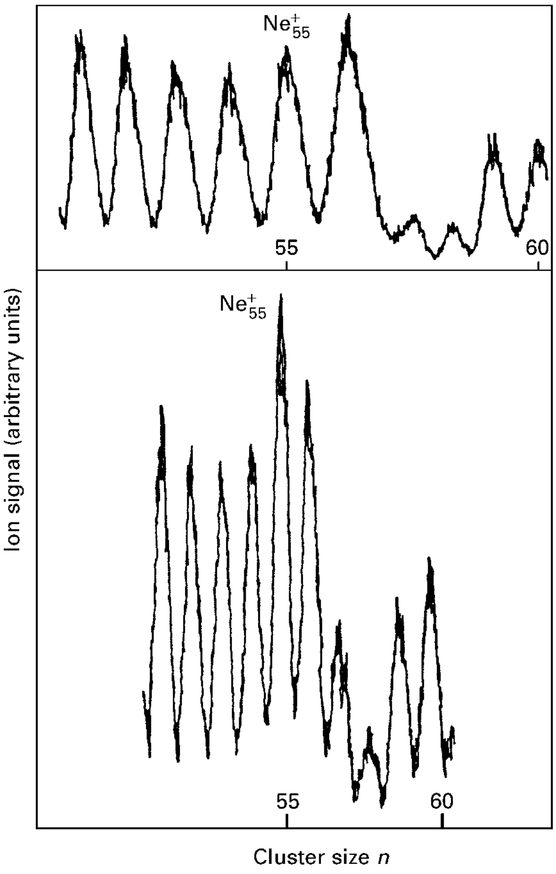



**Figure 5** Section of a $Ne^+_n$ cluster ion mass spectrum in the vicinity of the magic number size $n = 55$ measured approximately 68 μs (upper trace) and 118 μs (lower trace) after the primary ionization event. After Märk TD and Scheier P (1987) *Chemical Physics Letters*, **137**: 245.

Mass spectra of transition metal clusters do not exhibit any abundance anomalies. However, if they are mixed with reactive gases prior to expansion into vacuum, the number of molecules being absorbed exhibits pronounced size effects. Mass spectra may directly reveal the number of reactive sites on the cluster surface; $Co_{19}$, for example, offers six highly reactive sites towards ammonia. This number tends to increase with cluster size, but local minima are observed at certain cluster sizes, indicating closed shells or subshells in agreement with an icosahedral growth sequence.

The stoichiometry of NaCl cluster cations is $(NaCl)_nNa^+$. A pronounced sequence of magic numbers is observed if $n = 13, 22, 37, 62, 87,\ldots$. Assuming the ionic arrangement in the cluster is that of the crystal (rock salt, fcc lattice), compact shapes can be obtained by cutting cuboids out of the crystal. The number of ions in such a cuboid is odd only if the number of ions along each edge is odd. The numbers $n$ listed above correspond to cuboids with 3×3×3, 3×3×5, 3×5×5, 5×5×5 and 5×5×7 ions. Several other alkali halide clusters show the same pattern.

A magic number in mass spectra of carbon clusters has ultimately led to the discovery of a new form of crystalline carbon. **Figure 6** displays spectra of carbon clusters formed by laser vaporization into helium gas with subsequent expansion into vacuum. Only even-sized clusters are observed above $n = 30$. Smaller clusters, which are known to form chains and rings, occur for odd and even sizes. A structural transition takes place near $n = 30$. Ion chromatography has revealed that various structural isomers may co-exist. The abundance anomalies at $n = 60$ and 70 become more pronounced if the helium gas density is raised (**Figure 6**, upper spectrum). The structure of carbon clusters in this size range can be described as that of a graphitic sheet which is curved and closed upon itself by replacing 12 hexagons with pentagons. $C_{60}$ is the smallest of these so-called fullerenes in which all pentagons can be separated by hexagons; it has icosahedral symmetry. $C_{70}$ would be the next larger fullerene with no adjacent pentagons. The unusually large enhancements of $C_{60}$ and $C_{70}$ in **Figure 6** reflect their inertness towards growth. Later research has led to the synthesis of macroscopic amounts of $C_{60}$ and larger fullerenes. They are stable in air and, unlike other clusters, do not react with each other.

Fullerenes are hollow; $C_{60}$ can cage atoms as large as xenon. The 'endohedral' nature of a complex may be demonstrated by its stability. $(HeC_{60})^+$, formed by collisions of $C^+_{60}$ with helium gas, survives

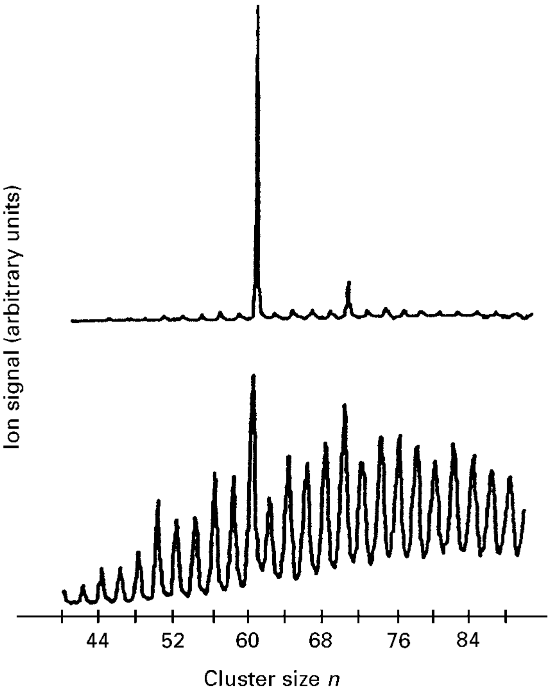



**Figure 6** Mass spectrum of carbon clusters formed by laser vaporization into helium gas of high pressure (upper spectrum) and low pressure (lower spectrum). Reproduced with permission of Macmillan Magazines Ltd from Kroto HW, Heath JR, O'Brien SC, Curl RF and Smalley RE (1985) *Nature* **318**: 162.

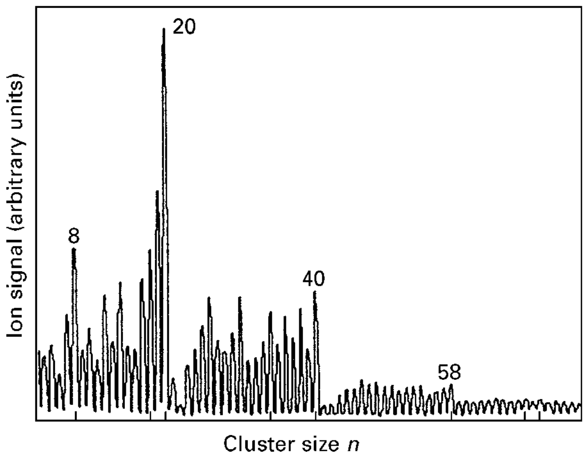



**Figure 7** Mass spectrum of sodium clusters, 'softly' ionized by low-intensity light from a UV lamp. After Knight WD, Clemenger K, deHeer WA, Saunders WA, Chou MY and Cohen ML (1984) *Physical Review Letters* **52**: 2141.

neutralization and re-ionization. The number of carbon atoms required to cage an atom depends on the size of the latter: photoexcited $(C_nK)^+$ will lose $C_2$ units until, below $n = 44$, potassium escapes. $(C_nCs)^+$, however, requires at least 48 carbon atoms.

Abundance anomalies as discussed above relate to unique atomic structures. Molten clusters will not exhibit magic numbers. This has been utilized to determine melting temperatures of clusters in the gas phase. If sodium clusters are thermalized in warm helium and then allowed to effuse into vacuum, no magic numbers related to atomic structure are observed. If the gas is cooled to 307 K (83% of the bulk melting temperature), magic numbers appear for clusters of size $n \geq 6000$. They extend to smaller clusters if the temperature is lowered further. Smaller clusters melt at lower temperatures. Molten sodium clusters show a different sequence of magic numbers as explained in the following section.

## Electronic structure

Key features of the electronic structure of some non-transition metal clusters were first revealed through mass spectrometry. **Figure** 7 shows a mass spectrum of sodium clusters, grown from vapour in cold helium gas and ionized by light from a UV lamp. Abundance anomalies are observed at $n = 8$, 20, 40, 58. If, instead, alkali clusters are ionized by an intense pulsed laser in the visible or near UV, abundance anomalies are observed at $n = 9$, 21, 41, 59. It appears that cluster stability is enhanced whenever the number of valence electrons, $n_e$, is 8, 20, 40, 58, .... The mass spectrum in **Figure** 7 is believed to reflect the abundance and, indirectly, the stability distribution of neutral clusters. On the other hand, multiphoton excitation leads to excessive dissociation of charged cluster ions, thus revealing the stability of cluster cations. Copper and silver clusters directly formed in the charged state by sputtering do indeed show the same magic numbers 9, 21, 41,.... Experiments involving other mono-, di- and trivalent metals provide further support for this hypothesis. For example, $Cu_{13}^-$ and $Ag_{43}^{3+}$ and $Al_{13}^-$ (also see **Figure** 8) form local abundance maxima in their respective distributions.

The observed 'magic numbers' find an explanation in the 'jellium model', which assumes the valence electrons to move freely throughout the cluster, only confined by an effective potential that is, thanks to the shielding by the other free electrons, approximately constant throughout the cluster. This quasi-free electron approximation may seem naive, but it successfully explains basic properties of many metallic solids. Making the approximation that the effective potential is spherically symmetric, one immediately predicts electronic shell structure: spherical symmetry gives rise to a degeneracy of the single-electron quantum states, the energy does not depend on the magnetic quantum number. Whenever a 'shell' of electrons is filled, addition of a monovalent atom

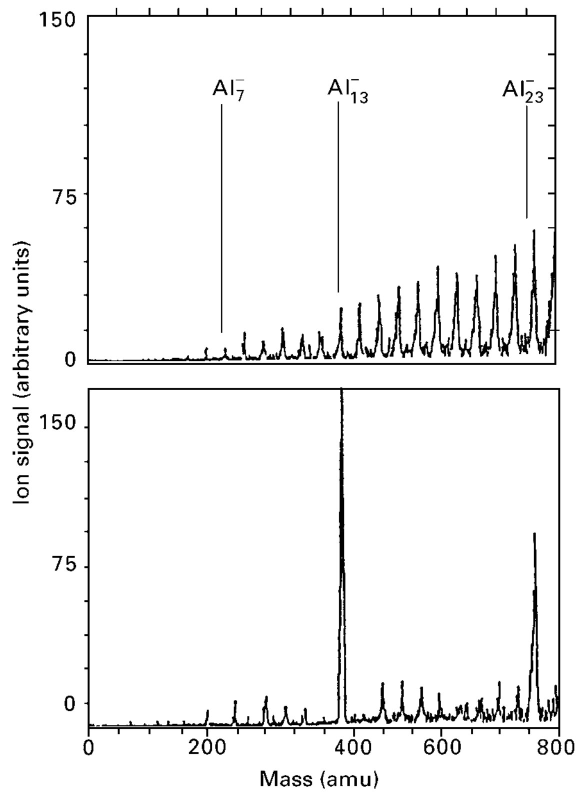



**Figure 8** Mass spectrum of aluminium cluster anions before (upper) and after (lower) reaction with oxygen. $Al_{13}^-$ is least reactive. After Leuchtner RE, Harms AC and Castleman Jr. AW (1989) *Journal of Chemical Physics* **91**: 2753.

implies that its valence electron has to occupy a higher state, making this extra atom less tightly bound. The energetic order of the first few shell closures is consistently predicted to occur at $n_e = 2$ (1s state), 8 (1p, in the notation used in nuclear physics), 18 (1d), 20 (2s), 34 (1f), 40 (2p), 58 (1g), etc. for the simple square-well potential as well as for more sophisticated potentials calculated self-consistently. Some of these gaps, like that between the 1d and 2s states, may be quite small, explaining the absence of an abundance anomaly at $n_e = 18$.

Many related effects have been studied in detail. Dissociation energies have been derived from metastable decay rates of alkali cluster ions. The gaps between electronic shells have been probed by measuring the size dependence of the ionization energy. Reduced ionization energies are, indeed, found for systems with 9, 21, 41, etc. valence electrons. The anomalies are much less than predicted, but refined effective (spherically nonsymmetric) potentials can explain this trend as well as the observation of some additional fine structure in the size distribution and in the ionization energy. Additional information has been obtained from photoelectron spectroscopy of neutral and negatively charged clusters. Again, the basic features of these spectra and their size dependence can be rationalized within the jellium model.

Classically, the ionization energy $E_{ion}(R)$ of a metallic sphere of radius $R$ is expected to converge to the bulk value $W$ (the work function) as $E_{ion}(R) = W + \text{const}\ R^{-1}$ (for a compact cluster of size $n$, the radius will scale as $n^{1/3}$). This behaviour is indeed observed for most metal clusters, except for the local anomalies mentioned above. Mercury, however, behaves differently (**Figure 9**). For small sizes, the ionization energies are significantly larger than expected, but a transition to the expected asymptotic behaviour occurs between $n \cong 20$ and 70 ($R \cong 5$–8 Å). This phenomenon has been assigned to a transition from an insulator to a metallic system.

The jellium model also provided a motivation to develop high-resolution mass spectrometry of very large clusters. Although the gaps between electronic states will gradually disappear with increasing level density, i.e. with increasing cluster size, they were predicted to decrease in a non-monotonic fashion owing to interference effects between different types of semiclassical orbits of the quasi-free electrons in the clusters. These 'supershells', originally predicted for nucleonic systems much larger than naturally occurring nuclei, have been confirmed for clusters of sodium, lithium and gallium containing thousands of valence electrons.

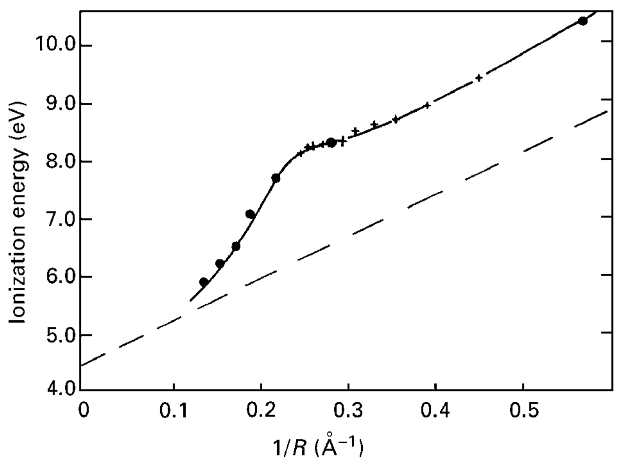



**Figure 9** Ionization energy of mercury clusters (solid circles, crosses) versus (estimated) reciprocal cluster radius. The slow convergence to the behaviour expected classically for a metallic sphere (dashed line) is attributed to the occurrence of an insulator-to-metal transition in the size range between 20 and 70 atoms. After Rademann K, Kaiser B, Even U, and Hensel F (1987) *Physical Review Letters* **59**: 2319.

## Cluster ion reactivity

Two categories of reactions have been studied that require the temporal evolution of the cluster ion composition to be known. The first comprises reactions of (neutral or ionized) clusters with neutral or ionized atoms, molecules, clusters or surfaces (not to mention reactions of photons or electrons with the clusters which, however, are usually discussed in the context of ionization and dissociation reactions); the second deals with reactions inside (excited) cluster ions. In the first case the mass spectrometer in various combinations serves as the chemical reactor, whereas in the second the cluster itself may be considered to be the chemical laboratory.

### Cluster ion reactions

One of the interesting questions addressed in this field is the influence that solvation, the degree of aggregation and the proximity of solute complexes have on the reactions and properties of reactive species. Huge changes (by many orders of magnitude) of the corresponding reaction rate constants may occur when going from the gas phase to the cluster environment; or very specific selectivity may be observed in reaction channels for metal cluster ions reacting with ligands (see **Figure 8**). Moreover, research on the reactivity of naked metal cluster ions is of special interest as these systems are considered to be potential catalysts since, in general, they are highly unsaturated. $Fe_4^+$ ions, produced by sputtering, thermalized and trapped in the storage cell of a FT-ICR mass spectrometer, were shown to react sequentially with ethylene, forming benzene precursors. The existence of these precursors and the occurrence of a complete catalytic cycle has been shown in a $MS^5$ experiment, where the various intermediates are successively identified.

Reactions between clusters may lead to fusion-like reactions, such as $C_{60}^+ + C_{60} \rightarrow C_{120}^+$, where fusion occurs only for collision energies exceeding a barrier of about 70 eV. On the other hand, they may cause fragmentation and even multifragmentation reactions, the latter showing an activation energy of about 70 eV for $C_{60}$. Studies involving various types of tandem mass spectrometer combinations have, in addition, dealt with the rapidly growing topic of cluster ion/surface reactions. Besides reflection, surface-induced dissociation, charge exchange and surface-induced reactions have been observed, thus allowing study of the reactivity and, in addition, the structure of these ions. For instance, collisions of size-selected stoichiometric and nonstoichiometric acetone ions with a stainless steel surface covered by hydrocarbon always give, besides typical fragmentation patterns expected for weakly interacting van der Waals cluster ions, the protonated acetone ion as the major secondary ion. Using deuterated acetone it is possible to distinguish two completely different reaction routes leading to this major product ion, one involving an H abstraction (pick-up) reaction with the surface adsorbates and one involving an ion molecule reaction within the cluster ion (intracluster reactions). The branching ratio for these two reactions depends strongly on the cluster size. Again, the ultimate goal is to study new reactions in the novel non-equilibrium environment present under the 'nano-shock' conditions shortly after cluster impact on the surface.

### Reactions in cluster ions

One of the most fascinating aspects of cluster reactivity is given by those unimolecular reactions that occur inside of an energized cluster (ion) after excitation/ionization by, for example, electrons or photons. Whereas inelastic interactions of free electrons with gas-phase atoms or molecules result in changes of the electronic configuration and/or nuclear motion, interactions of electrons with clusters may, in addition, (i) comprise multiple collisions of the incoming (scattered) electron(s) and (ii) involve subsequent intra- and intermolecular reactions within the 'cluster reactor'. This sequence of events leads to the formation of excited cluster ions (which may differ in composition and stoichiometry from the initial neutral precursor), involving the deposition of excess energy into various degrees of freedom. Energy and charge exchange between different sites in the cluster may, in turn, cause additional spontaneous reactions not present in ordinary gas-phase molecules. These spontaneous reactions, occurring on time scales ranging from a few vibrational oscillations up to the millisecond time regime, will contribute strongly to the final fragmentation pattern observed in mass spectrometry. They are also responsible for the fact that measured cluster mass spectral distributions are often quite different from the distribution of clusters in the probed neutral beam and are usually only snapshots of the changing ion compositions after a primary ionization or excitation event (see **Figure 5**). Here, from the large number of different types of reactions, only four examples will be discussed, namely two internal ion molecule reactions and two unimolecular cluster decay reactions.

Reactions in the cluster may lead to the production of ions not produced in the case of the monomer, and vice versa. For instance, electron impact ionization of $CO_2$ does not give $O_2^+$ fragment ions,

whereas electron impact ionization of $CO_2$ clusters yields $(CO_2)_mO_2^+$ ions by the following reaction sequence: via dissociative ionization of a single $CO_2$ molecule within the cluster, the incoming electron produces an $O^+$ fragment ion. In a second step, this fragment ion, via an ion molecule reaction with a neighbouring $CO_2$ molecule, forms $O_2^+$ plus CO.

In contrast, for polyatomic molecules, fragment ion abundances are usually much smaller if the ionization and subsequent fragmentation occur within a cluster. For instance, for a propane cluster, the relative abundance of the parent molecular ion $C_3H_8^+$ and the abundances of the first-generation fragment ions $C_3H_m^+$, $C_3H_6^+$, $C_2H_5^+$, $C_2H_4^+$, are significantly enhanced; those of the second-generation fragment ions $C_3H_5^+$, $C_3H_4^+$, $C_2H_3^+$, $C_2H_2^+$ are diminished in comparison to the situation with the monomer. This is due to energy transfer from the originally excited propane molecular ion, via quenching reactions, to the cluster constituents.

The common unimolecular (spontaneous) decay in the metastable time regime for (excited) cluster ions is monomer evaporation, although in certain cases dimer and trimer ejection have also been observed. Whereas electronic predissociation and barrier penetration are likely mechanisms for the decay of small cluster ions, vibrational predissociation is the dominant metastable dissociation mechanism for larger cluster ions, leading in some cases — owing to the large amount of energy stored — to sequential decay series. Moreover, because cluster ions of a specific size produced in the ion source of a mass spectrometer by electron impact of a neutral cluster beam will comprise a broad range of internal energies (owing to the statistical nature of the complicated ionization process and the possibility of cascading from different neutral precursors), the decay of such an ion ensemble will not be governed by a single decay constant $k$ as predicted by RRKM theory for a polyatomic ion excited to a specific energy. The experimental results show, in accordance with predictions using the concept of an evaporative ensemble (EEM), a nonexponential function for the time dependence of the apparent metastable decay rate (**Figure 10**). Moreover, this EEM model can be also used to describe the experimentally observed increase of the apparent decay rate with cluster size. Deviations observed from the prediction are due to the special nature of 'magic number' cluster sizes. It is thus not surprising that these anomalously small or large decay rates agree in a mirror-like fashion with enhanced or depleted ion abundances in the mass spectrum.

In addition to the single monomer evaporation based on the statistical predissociation mechanism,

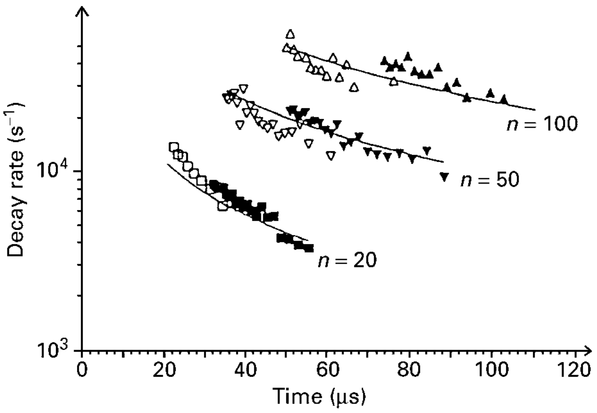



**Figure 10** Measured and predicted (solid line) apparent metastable decay rates (metastable fraction divided by the experimental time window) for three different cluster ion sizes and two different rare gases (open symbols, argon; filled symbols, krypton) versus time since electron ionization. After Ji Y, Foltin M, Liao CH and Märk TD (1992) *Journal of Chemical Physics* **96**: 3624.

several unusual metastable fragmentation reactions involving the ejection of more than one monomer have been observed. It has been demonstrated by mass spectrometric techniques that these decay reactions are due to the decay of isolated excited states within the cluster ion, including vibrational states, excimers and isomers.

Hot surfaces may 'boil off' electrons if the corresponding activation energy is not much higher than that for the competing reaction, evaporation of atoms. For energetic reasons, this 'thermionic emission' process is not observed for ordinary molecules unless they are negatively charged (autodetachment). The cohesive energy per building block tends to increase with increasing cluster size, while the ionization energy decreases. Also, a large system can store excitation energy more effectively than a small molecule; the cluster can serve as its own heat bath. Hence, delayed electron emission from neutral clusters excited by photons or by collisions is a viable reaction. This process produces asymmetric peaks in conventional time-of-flight (TOF) mass spectra, because the arrival time of an ion is usually measured with respect to the time at which the parent molecule was excited. **Figure 11** (top) shows a TOF mass spectrum of a $C_{60}/C_{70}$ mixture, using multiphoton excitation at 266 nm. To select the time of ionization separately from the time of flight through the instrument, the extraction field is modulated in time such that only ions formed within a certain time interval (4–5 μs in this case) reach the detector, thus giving rise to narrow mass peaks that reflect the rate of formation within the chosen time interval (**Figure 11**, lower). It

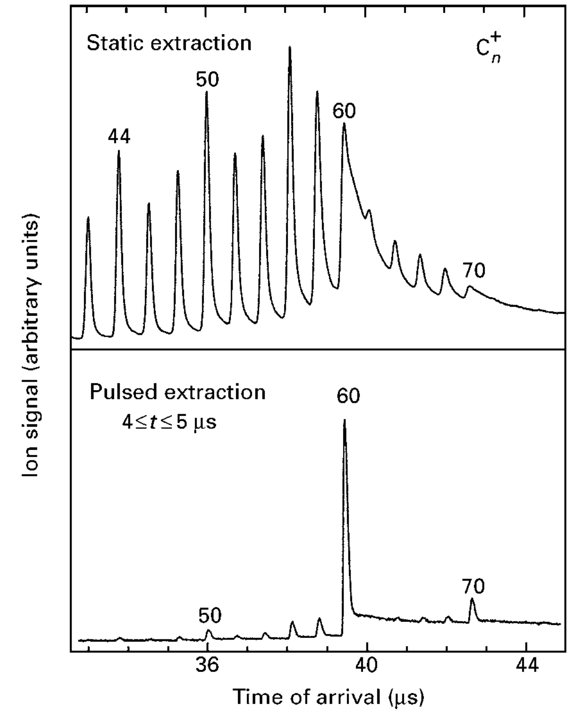



**Figure 11** Time-of-flight mass spectrum of multiphoton-ionized fullerenes. Upper spectrum: static extraction, delayed ions give rise to asymmetric peaks. Lower spectrum: pulsed extraction, delayed ions formed within a chosen time window are bunched into narrow peaks.

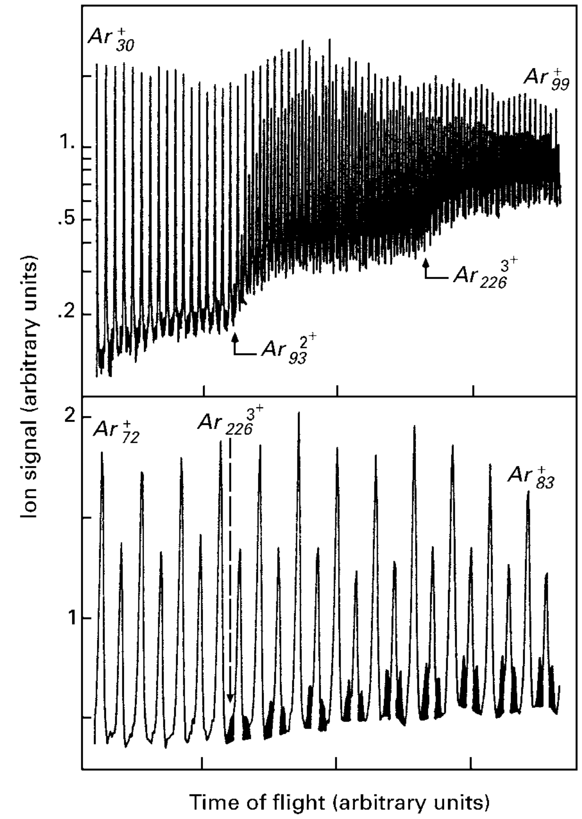



**Figure 12** Mass spectrum of argon clusters ionized by electron impact. Doubly and triply charged clusters are observed above characteristic appearance sizes.

becomes apparent that $C_{58}^+$ and other fragment ions also have delayed components. $C_{60}^+$ may form hundreds of microseconds after excitation. Delayed ionization has been observed for neutral refractory metal clusters containing as few as five atoms, for some metal oxides and for metallocarbohedrenes such as $Ti_8C_{12}$. The phenomenon is ubiquitous for negatively charged clusters owing to their reduced activation energy for electron emission.

## Multiply charged clusters

Systems with a high charge density become unstable with respect to dissociation into charged fragments (fission). This effect limits the net charge of molecules, small clusters, and droplets of a given size, as well as the size of natural and man-made atomic nuclei. Multiply charged cluster ions are easily identified by their non-integer size-to-charge ratios (**Figures** 1 and 12). However, for each charge state $z$ there is a characteristic appearance size $n_z$ below which the species becomes undetectable, as shown in **Figure** 12 for argon. For more tightly bound metal clusters, $n_{z=2}$ may be as small as 2, but the concept of an appearance size is still meaningful for higher charge states.

The effect has been observed for various atomic and molecular van der Waals clusters, hydrogen-bonded clusters and metal clusters. Seven-fold sodium clusters, for example, are not observed below $n_7 \cong 445$. Under typical experimental conditions, multiply charged clusters are, like their singly charged counterparts, thermally excited; they are found to evaporate atoms or molecules in a statistical, unimolecular process if their size is well above the appearance size. In the vicinity of the latter, however, fission competes with evaporation. Below the appearance size, the rate of fission greatly exceeds the time needed for mass spectrometric detection. The appearance size can be somewhat reduced if colder multiply charged clusters are generated. Ultimately, however, the clusters become unstable and the fission barrier vanishes.

The size distribution of fission fragments can be determined mass spectrometrically. Distributions tend to be asymmetric, but there are considerable variations in the details. Several studies have been devoted to fullerenes. $C_{60}^{z+}$ is observed with charge states up to $z = 9$. Spontaneous (unimolecular) dissociation of $C_{60}^{3+}$, as well as producing $C_{58}^{3+}$ and

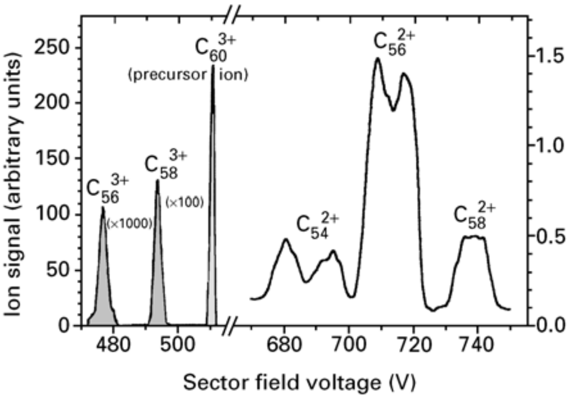



**Figure 13** MIKE scan of $C_{60}^{3+}$ in a double focusing sector field mass spectrometer, showing in addition to the parent ion spontaneous fission into different fragment ions. The broadening and splitting for the charge separation reactions arises from the large kinetic energy release in these reactions. Reproduced with permission of Taylor & Francis from Scheier P, Dünser B, Wörgötter R *et al* (1996) *International Review of Physical Chemistry* **15**: 93.

smaller fragments, produces $C_{54}^{2+}$, $C_{56}^{2+}$ and $C_{58}^{2+}$ via a 'charge separation' reaction (**Figure 13**). These MIKE spectra (scan of the electric sector field voltage) reveal the energy released upon fissioning. $C_{54}^{2+}$ shows the greatest broadening because it is related to the heaviest fission partner, $C_6^+$. Detection of the light fission fragments is difficult because of their large recoil velocity, but $C_6^+$ has been directly identified.

## Cluster anions

Often a single molecule cannot bind an excess electron, although it is quite possible to trap

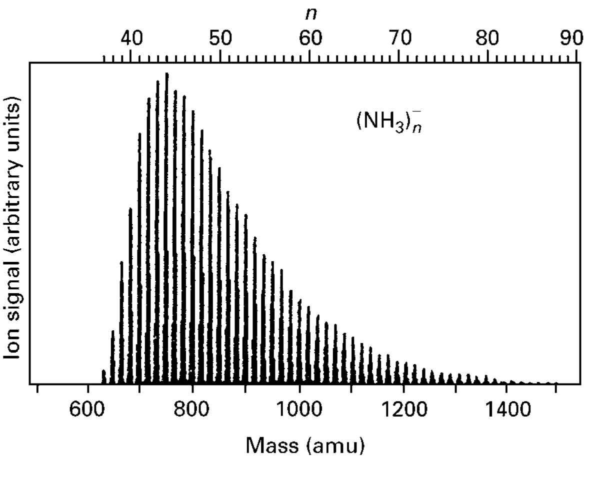



**Figure 14** Mass spectrum of negatively charged $(NH_3)_n$ clusters showing a sharp onset around the minimum appearance size $n_{min}$ = 35. After Haberland H, Schindler HG and Worsnop PR (1984) *Berichte der Bunsengesellschaft für Physikalische Chemie* **88**: 270.

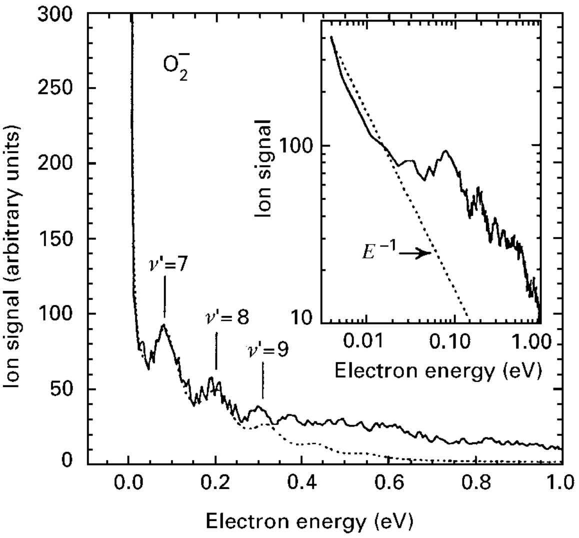



**Figure 15** Measured (full line) and calculated (broken line) $O_2^+$ anion signal produced by electron attachment to an oxygen cluster beam as a function of electron energy. Inset: Same data on a log–log scale demonstrating the $E^{-1}$ dependence predicted by s-wave scattering theory (dashed line). After Matejcik S, Kiendler A, Stampfli P, Stamatovic A and Märk TD (1996) *Physical Review Letters* **77**: 3771.

electrons in liquids and solids of these molecules. Negatively charged clusters of these substances are of particular interest since it is possible to study details of this trapping as a function of aggregation. It has been shown by mass spectrometry that three mechanisms are responsible for the occurrence of negatively charged stoichiometric cluster anions in these cases: (i) stabilization by collision if accommodation of the electron affinity leads to autodetachment for a single isolated molecule (e.g. $O_2$); (ii) stabilization by solvation if Franck–Condon problems lead to dissociation (e.g. $CO_2$); and (iii) generation of a solvated electron (polaron) if the electron affinity of the monomer is zero (e.g. $H_2O$). Whereas in the first two cases attachment of an electron to the cluster leads to the production of monomer anions and higher analogues, in the last case only cluster anions above a certain minimum size $n_{min}$ can be observed (see **Figure 14** showing $n_{min} = 35$ for ammonia clusters). It turns out that this minimum size in the observed mass spectrum may depend on the method of anion production. Cluster anions of volatile compounds can be formed in the gas phase either by injection of free electrons into the expansion zone of a supersonic jet or by attachment of electrons to a beam of pre-existing neutral clusters. For water clusters one obtains $n_{min} = 2$ in the former and $n_{min} \cong 11$ in the latter mode. Calculations show that the electron is localized in a

'surface state' for the very small water cluster anions ($n = 2–8$); vertical detachment energies measured by photodetachment methods are, in fact, very low. Thus long-lived anions below $n \cong 11$ can only be produced in the special (cold) environment of the expanding jet.

As well as production of stoichiometric cluster anions, nonstoichiometric fragment cluster anions have also been observed (e.g. $(O_2)_mO^-$ for oxygen). Whereas these fragment anions can usually be produced by attaching electrons to pre-existing clusters at energies similar to the known resonances for the monomer (e.g. at around 6 eV and higher for ($O^-/O_2$), the stoichiometric anions are, in addition, often produced with very high probability at energies close to zero. **Figure 15** shows the region of the 'zero-energy' resonance of oxygen clusters as measured with a high-resolution spectrometer. In addition to the zero-energy peak ascribed to s-wave capture by the entire cluster, further peaks are observed; these are attributed to vertical attachment of the incident electron to a single molecule within the cluster, thus allowing probing of the vibrational levels of a single oxygen anion within the cluster.

*See also:* **Ionization Theory; Thermospray Ionization in Mass Spectrometry.**

## Further reading

Andersen HH (1997) *Small Particles and Inorganic Clusters*. Berlin: Springer-Verlag.

Bernstein ER (1990) *Atomic and Molecular Clusters*. Amsterdam: Elsevier.

Bréchignac C and Cahuzac Ph (eds) (1995) *Comments on Atomic and Molecular Physics* **31**: 137–477.

Castleman Jr. AW and Bowen KH (1996) Clusters: structure, energetics, and dynamics of intermediate states of matter. *Journal of Physical Chemistry* **100**: 12911–12944.

de Heer WA (1993) The physics of simple metal clusters: experimental aspects and simple models. *Review of Modern Physics* **65**: 611–676.

Deng R and Echt O (1998) Efficiency of thermionic emission from $C_{60}$. *Journal of Physical Chemistry A* **102**: 2533–2539.

Dresselhaus MS, Dresselhaus G and Eklund PC (1996) *Science of Fullerenes and Carbon Nanotubes*. San Diego: Academic Press.

Haberland H (1994) *Clusters of Atoms and Molecules: Theory, Experiments, and Clusters of Atoms*. Springer Series in Chemical Physics 52. Berlin: Springer-Verlag.

Haberland H (1994) *Clusters of Atoms and Molecules II: Solvation and Chemistry of Free Clusters, and Embeded, Supported and Compressed Clusters*. Springer Series in Chemical Physics 56. Berlin: Springer-Verlag.

Ingólfsson O, Weik F and Illenberger E (1996) The reactivity of slow electrons with molecules at different degrees of aggregation: gas phase, clusters and condensed phase. *International Journal of Mass Spectrometry and Ion Processes* **155**: 1–68.

Jortner J (1992) Cluster size effects. *Zeitschrift für Physik* D **24**: 247–275.

Märk TD (1995) Mass spectrometry of clusters. In: Cornides I, Horváth G and Vékey K (eds) *Advances in Mass Spectrometry*, Vol 13, pp 71–94. Chichester: Wiley.

Martin TP (1996) Shells of atoms. *Physics Reports* **273**: 199–241.

Näher U, Bjørnholm S, Frauendorf S, Garcias F and Guet C (1997) Fission of metal clusters. *Physics Reports* **285**: 245–320.

Ng CY, Baer T and Powis I (1993) *Cluster Ions*. Chichester: Wiley.

# Cobalt NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# Colorimetry, Theory

**Alison Gilchrist** and **Jim Nobbs**, University of Leeds, UK
Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Theory**

Colorimetry is the science of the measurement of colour. It involves the replacement of subjective responses, such as ‘light blue’, ‘rich dark purple’, ‘bright gold’, with an objective numerical system. This study began with the work of Young, Helmholtz and Maxwell in the early nineteenth century, who recognized the principles of additive and subtractive colour mixing, and proposed the trichromatic nature of human colour vision. The science began to be formalized in 1931, when the Commission International de l’Eclairage (CIE) recommended a system of colour specification based on the three tristimulus values $X$, $Y$ and $Z$. The current CIE reference document relating to colorimetry is CIE Publication 15.2-1986 (Edition 3 is in preparation); for more practical information on colorimetric measurement, ASTM document E308-96 should be consulted.

## Illumination, object, observer

Three things contribute to our perception of the colour of an object: the nature of the illumination, the optical properties of the object itself and the response of the human eye (**Figure 1**). The nature of the illumination can be characterized by the spectral power distribution $S(\lambda)$ of the light source, the relative intensity of the illumination at each wavelength in the visible spectrum. The object reflects a certain fraction of the incident light and this can be characterized by the reflectance spectrum $R(\lambda)$. The intensity of light entering the eye, $I(\lambda)$, is the product of these terms. Thus in order to measure colour, and to specify colour by numbers, it is necessary to specify each of these three components of the colour trio. In 1931, the CIE recommended standard illuminants for use in colorimetry, published data representing the standard observer and recommended standard optical geometries for use in colour-measuring instruments.

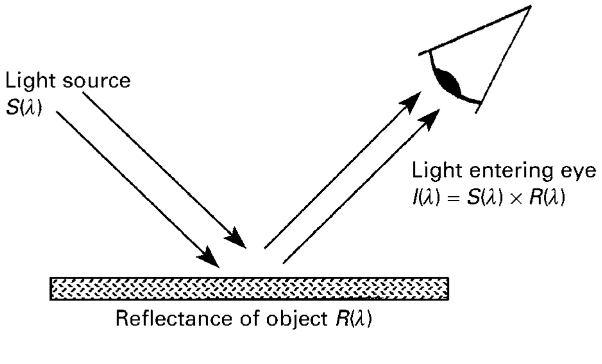



**Figure 1** How colour reaches the eye.

## Standard illuminants (A, C, D65, F)

In 1931 the CIE recommended three standard sources and Illuminants (A, B and C) where a source is a physically realizable light source and an illuminant is a table of data (spectral power distribution) representing a source. A is intended to represent an incandescent tungsten filament lamp (colour temperature 2856 K); B and C may be produced by filtering the light from source A to produce a representation of noon sunlight (B) or average daylight (C). More recently, the CIE has recommended the D series as illuminants only, representing daylight in various phases. D75 (colour temperature 7500 K) represents north sky daylight, D65 (colour temperature 6500 K) represents average daylight and D55 (colour temperature 5500 K) represents sunlight plus skylight. D65 is now the reference illuminant for use in colorimetry (although the graphic arts industry tends to prefer D55) and so should be included in any colour specification. Unfortunately, illuminant D65 is not realizable in practice, and so colour-matching cabinets generally contain a daylight simulator (a filtered fluorescent lamp). A further series of florescent lamps, the F series, were also defined, although none is specified as a standard illuminant. There are three types: normal, broad-band and three-band. The most important are F2 (normal, known as ‘cool white’), F7 (broad-band, known as ‘artificial daylight’ and commonly used for colour matching) and F12 (three-band, with high energy efficiency, and so often used in shops). **Figure 2** shows the relative spectral power distribution of the more commonly used illuminants.

## Standard observers (2-degree, 10-degree)

The experiments that established the basis of colorimetry were carried out separately by J. Guild at the

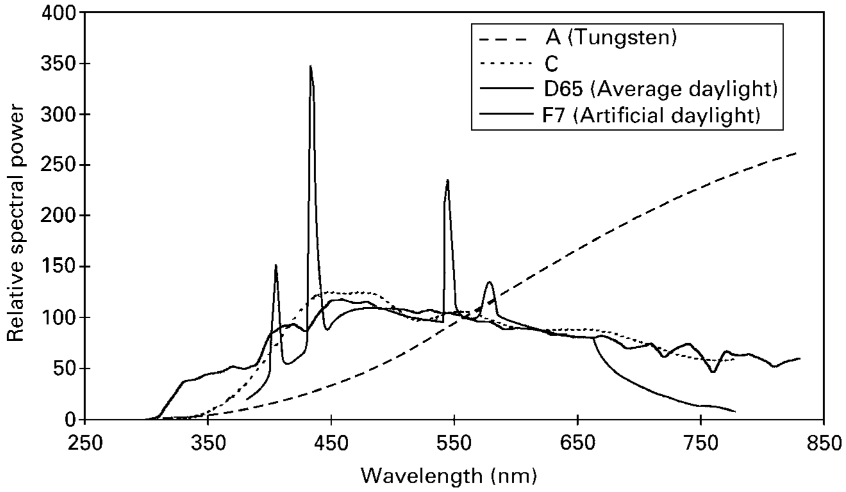



**Figure 2** Relative spectral power distributions of common CIE illuminants.

National Physical Laboratory (UK) and W.D. Wright at Imperial College London in a famous series of experiments in the late 1920s. These experiments were dedicated to characterizing the response of the human eye to different wavelengths of light, and were carried out by colour matching using red, green and blue lights. A set of red, green and blue light sources are often called primary sources since the appearance of any one of the sources cannot be matched by blending together light from the other two. A possible experimental arrangement is illustrated in **Figure 3**. Each worker used several observers (Guild 7 and Wright 10) who adjusted the mixture of light from three primaries until it matched the appearance of the monochromatic light.

These early experiments were carried out using the central area of the retina, the fovea. This region contains the highest concentration of colour-sensitive cells, and occupies the central 2° field (based on the

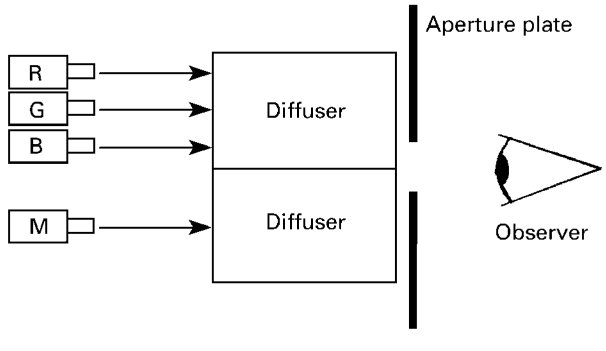



**Figure 3** Possible arrangements of Guild's and Wright's experiments.

angle to the line of sight, drawn from the centre of the lens to the foveal pit). The standard observer derived from Guild's and Wright's experiments is therefore known as the 2° standard observer (1931). The colour matching functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$ represent the amounts of light from the red, green and blue primaries, in tristimulus units, needed to match unit intensity of light with a narrow band of wavelengths centred on $\lambda$. Tristimulus units are defined such that a mixture of equal units of the red, green and blue primaries will match an equal-energy white. The values of the colour-matching functions are normalized so that there is an equal area under each of the spectral curves. The values are plotted against wavelength in **Figure 4**.

It was found that not all the colours produced by monochromatic light could be matched by a simple additive mixture of light from three primaries. For some wavelengths it was necessary to add light from one of the primaries to the monochromatic light, reducing the colour saturation of the monochromatic light. The appearance of the desaturated mixture could then be matched by a combination of the other two primaries. When this occurs, the amount of the primary added to the monochromatic light appears as a negative number in the colour specification.

Later workers observed that although the fovea is the region most sensitive to colour, most everyday objects occupy a larger area of the retina. The 10° standard observer (1964) was thus derived in a similar manner and is now the preferred observer function. The true cone sensitivities were not measured directly until 1979, but correspond very well to those expected from the colour-matching functions found by Guild and Wright.

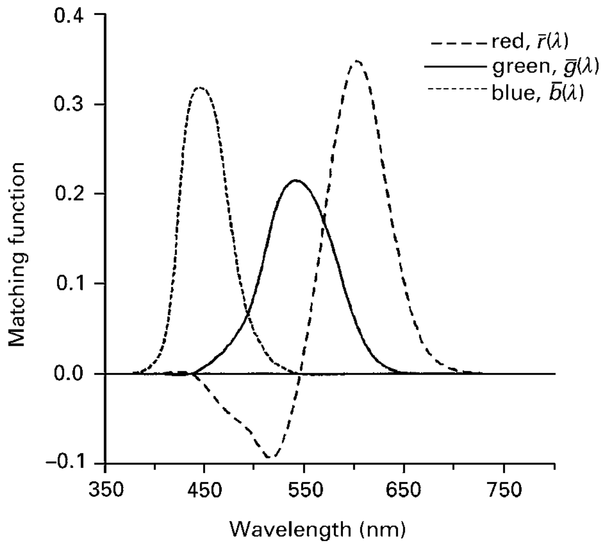



**Figure 4** Guild's and Wright's colour-matching functions.

## RGB specification

The RGB system of colour specification was based on the trichromatic theory of colour vision and the colour-matching experiments carried out by Guild and Wright. It had been suggested by Maxwell in 1860 that if the specifications of the red, green and blue lamps used in such experiments were known, then the amounts of each light required to match a sample colour would provide a numerical specification of that colour. These amounts may be found either by passing the light from the sample through filters matching the RGB primaries, or by calculation from the reflectance values of the sample at a series of wavelengths through the visible spectrum:

$$R = k[I(\lambda_1)\,\bar{r}(\lambda_1) + I(\lambda_2)\,\bar{r}(\lambda_2) + \cdots]$$
$$\text{or} \quad R = k\sum_{\lambda=380}^{\lambda=780} I(\lambda)\,\bar{r}(\lambda)$$
$$G = k[I(\lambda_1)\,\bar{g}(\lambda_1) + I(\lambda_2)\,\bar{g}(\lambda_2) + \cdots]$$
$$\text{or} \quad G = k\sum_{\lambda=380}^{\lambda=780} I(\lambda)\,\bar{g}(\lambda)$$
$$B = k[I(\lambda_1)\,\bar{b}(\lambda_1) + I(\lambda_2)\,\bar{b}(\lambda_2) + \cdots]$$
$$\text{or} \quad B = k\sum_{\lambda=380}^{\lambda=780} I(\lambda)\,\bar{b}(\lambda) \qquad [1]$$

Although RGB values are a valid representation of the colour of an object viewed under specified conditions, for practical reasons these values are rarely used. For a significant proportion of colours one of the three values is negative, and this was considered a distinct disadvantage.

## *XYZ* specification (1931)

The *XYZ* system of colour specification was recommended by the CIE in 1931 and was closely based on the RGB system. The difference lies in the fact that *R*, *G* and *B* were real light sources of known specifications. *X*, *Y* and *Z* are theoretical sources which are more saturated than any real light source, and allow matching of any real colour using positive amounts of the three primary sources. The *X*, *Y* and *Z* stimuli are defined so that:

- The *XYZ* tristimulus values of all real colours are positive.
- The *Y* tristimulus value is proportional to the luminance.
- A mixture of equal amounts of *X*,*Y* and *Z* has the same colour appearance as an equal-energy white.
- The *X*, *Y*, *Z* values of the visible colours have the widest possible range of values.
- The *X* stimulus represents a red more saturated than any spectral red.
- The *Y* stimulus represents a green more saturated than any spectral green.
- The *Z* stimulus represents a blue more saturated than any spectral blue.

*X*, *Y*, *Z* values for the 2° observer may be calculated from RGB values using the equations

$$X = 0.49R + 0.31G + 0.20B$$
$$Y = 0.17697R + 0.81240G + 0.01063B$$
$$Z = 0.00R + 0.01G + 0.99B \qquad [2]$$

Alternatively, *XYZ* values may be calculated from reflectance values using similar equations to those for RGB, but using weighting functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ in place of $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$. The CIE nomenclature is to denote the 10° observer function with a subscripted '10' (Eqn (3)). The *XYZ* system is an improvement over the RGB system in that it eliminates negative numbers in colour specification. However, it is still based on the mixing of coloured lights and not on the visual appreciation of colour. In particular, the visual colour difference between two samples is not linearly related to the difference in *XYZ* values:

$$X_{10} = k[I(\lambda_1)\,\bar{x}_{10}(\lambda_1) + I(\lambda_2)\,\bar{x}_{10}(\lambda_2) + \cdots]$$
$$\text{or} \quad X_{10} = k\sum_{\lambda=380}^{\lambda=780} I(\lambda)\,\bar{x}_{10}(\lambda)$$
$$Y_{10} = k[I(\lambda_1)\,\bar{y}_{10}(\lambda_1) + I(\lambda_2)\,\bar{y}_{10}(\lambda_2) + \cdots]$$
$$\text{or} \quad Y_{10} = k\sum_{\lambda=380}^{\lambda=780} I(\lambda)\,\bar{y}_{10}(\lambda)$$
$$Z_{10} = k[I(\lambda_1)\,\bar{z}_{10}(\lambda_1) + I(\lambda_2)\,\bar{z}_{10}(\lambda_2) + \cdots]$$
$$\text{or} \quad Z_{10} = k\sum_{\lambda=380}^{\lambda=780} I(\lambda)\,\bar{z}_{10}(\lambda) \qquad [3]$$

## *L*a*b** colour space (1976)

CIE $L^*a^*b^*$ colour space is based on the opponent theory of colour vision, which arose from the ideas of Hering. Hering pointed out that of the thousands of words used to describe hue, only four are unique; red, green, yellow and blue. They are unique because they cannot be described without using their own colour name, yet any other hue can be described using one or more words from this set. When taken together with white and black, they form a group of six basic colour properties that can be grouped into three opponent pairs, white/black, red/green and yellow/blue. The concept of opponency arises from the observation that no colour could be described as having attributes of both redness and greenness, or of both yellowness and blueness. A red shade of green just does not exist.

It seems sensible to associate these three opponent pairs with the three axes of a three dimensional colour space. On a colour chart, the amount of redness or greenness of a colour sensation could be represented by the distance along a single axis, with pure red lying at one extremity and pure green lying at the other. In a similar way yellow and blue are opposite extremities of a second axis that could be placed perpendicular to the red/green direction. The third axis would go from white to black and lie in the plane normal to the other two. CIE $L^*a^*b^*$ (1976) space uses three terms $L^*$, $a^*$ and $b^*$ to represent colour, as shown in **Figure 5** and described below.

$L^*$ The vertical axis represents lightness; 100 represents a perfect white sample and 0 a perfect black.

$a^*$ The axis in the plane normal to $L^*$ represents the redness–greenness quality of the colour. Positive values denote redness and negative values denote greenness.

$b^*$ The axis normal to both $L^*$ and $a^*$ represents the yellowness–blueness quality of the colour. Positive values denote yellowness and negative values denote blueness.

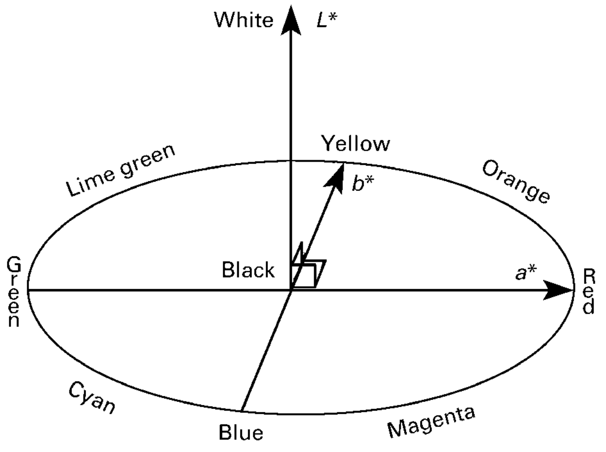



**Figure 5** CIE $L^*a^*b^*$ colour space.

$L^*a^*b^*$ values are calculated from *XYZ* values using the following equations:

$$F_X = \left(\frac{X}{X_0}\right)^{1/3} \qquad [4a]$$

unless

$$\left(\frac{X}{X_0}\right) \leqslant 0.008856$$

when

$$F_X = 7.787\left(\frac{X}{X_0}\right) + \frac{16}{116} \qquad [4b]$$

where $X_0$ is the $X$ tristimulus value of a perfect white sample under the chosen illuminant. Values of $F_Y$ and $F_Z$ are calculated in a similar manner. Then:

$$L^* = 116F_Y - 16$$
$$a^* = 500[F_X - F_Y]$$
$$b^* = 200[F_Y - F_Z] \qquad [5]$$

Since $L^*a^*b^*$ values for a sample may change for a different illuminant and observer combination, any colour specification must always include the illuminant and observer for which the values were calculated.

## CIE $L^*C^*h°$ coordinates

The coordinates of the points representing the colour of a sample can also be expressed in cylindrical coordinates of lightness $L^*$, chroma $C^*$ and hue $h°$, as illustrated in **Figure 6**. This system corresponds more closely to the natural description of colour.

$L^*$ Lightness (see above)
$C^*$ Chroma:

$$C^* = \sqrt{\left(a^{*2} + b^{*2}\right)}$$

$h°$ Hue angle:

$$h° = \tan^{-1}\left(\frac{b^*}{a^*}\right) \qquad [6]$$

Neutral grey samples would have $C^*$ values between 0 and about 5 units. The hue angle is always measured anticlockwise from the positive $a^*$ axis. It is traditional to associate one of the psychological primary hues with each axis direction, as shown in the diagram, although in reality the angles of the pure hues do not lie exactly along the axis directions.

## Colour difference

As the CIE $L^*a^*b^*$ system was designed to be visually uniform, it is fairly simple to devise an equation for the total colour difference between a trial and standard sample. The total colour difference is the distance between the two points representing those colours in the colour space, as shown in **Figure 7**. The distance, expressed as $\Delta E^*_{ab}$, is determined using the laws of right-angle triangles:

$$\Delta E^*_{ab} = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \qquad [7]$$

where

$$\Delta L^* = L^*_{\mathrm{trl}} - L^*_{\mathrm{std}}$$
$$\Delta a^* = a^*_{\mathrm{trl}} - a^*_{\mathrm{std}}$$
$$\Delta b^* = b^*_{\mathrm{trl}} - b^*_{\mathrm{std}} \qquad [8]$$

The overall colour difference may also be split into lightness, chroma and hue terms, which again correspond more closely to the natural description of colour:

$$\Delta C^* = C^*_{\mathrm{trl}} - C^*_{\mathrm{std}} \qquad [9]$$

$$\Delta H^* = \sqrt{\Delta E^{*\,2}_{ab} - \Delta L^{*2} - \Delta C^{*2}} \qquad [10]$$

It is worth remembering that the scaling of the colour space was set so that a value of $\Delta E^*_{ab} = 1$ between the colour of two samples should be just visually perceptible. This means that if a large

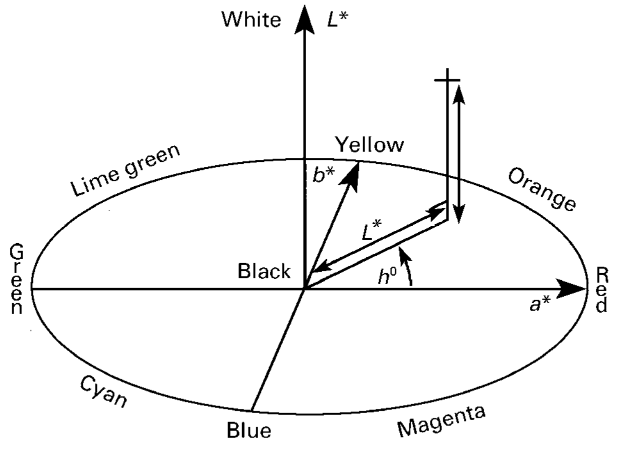



**Figure 6** CIE $L^*C^*h°$ colour coordinates.

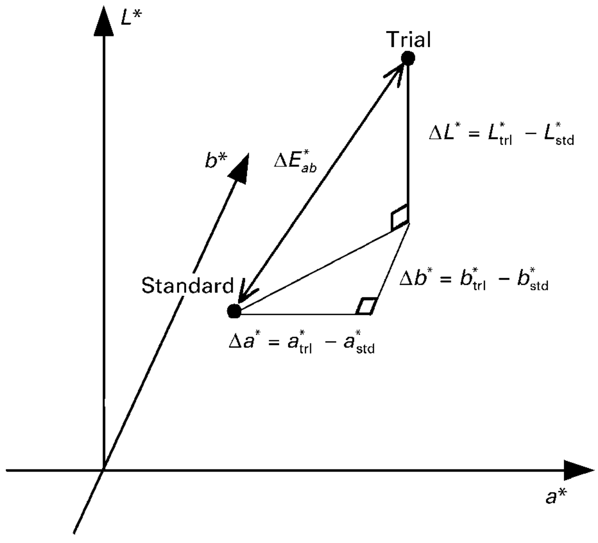



**Figure 7** The CIE $L^*a^*b^*$ colour difference calculation.

number of people were asked to judge the colour difference between the two samples, about 50% of observers would say that there was a difference in colour, the other 50% would say the two colours matched. This is the point of maximum argument.

Comparisons between the majority visual decision and a numerical assessment, based on the condition that $\Delta E^*_{ab}$ must be less that 1 for a match, show that the instrument decision will disagree with the majority visual decision about 19% of the time. On average, an individual colourist will disagree with the majority decision about 17% of the time when judging pairs of panels with fairly small colour differences. Pass/fail decisions based on the CIE $L^*a^*b^*$ (1976) equation with a limit of $\Delta E^*_{ab} = 1$ are, on average, slightly less reliable than visual judgements made by a single colourist. However, a colour-measuring instrument has the great advantage of giving much more repeatable decisions.

A number of scientists have devised more reliable or optimized colour difference equations. The simplest method of improvement is to assign individual tolerances to $\Delta L^*$, $\Delta a^*$ and $\Delta b^*$. Alternatively, individual tolerances can be assigned to $\Delta L^*$, $\Delta C^*$ and $\Delta H^*$ (differences in lightness, chroma and hue). The most successful equations are based on the $\Delta L^*$, $\Delta C^*$, $\Delta H^*$ system of colour splitting and the CMC (*l*, *c*) and CIE 94 equations are of this type. The CMC equation was developed by the Colour Measurement Committee of the Society of Dyers and Colourists (Bradford, UK) and has been adopted as the ISO standard for small colour difference assessment of textile materials. In 1994 the CIE suggested the simpler CIE 94 equation for evaluation. The published evidence collected so far by the CIE suggests that pass/fail decisions made using either of the optimized colour difference equations are considerably more reliable than those based on $\Delta E^*_{ab}$ or on the judgement of a single human colourist. Comparison with decisions made by panels of observers show that the equations disagree with the majority decision about 13% of the time. Neither equation has yet been adopted as a CIE standard as it has been shown that further improvements can still be made.

## Colour and appearance

The systems of colorimetry described above can only specify the colour of an isolated sample viewed against a neutral grey background. However, almost all colours are viewed against different backgrounds, with different fields of view, viewing environments, and states of adaptation to that environment. The surrounding colours or patterns will affect the perceived colour of the sample in ways which cannot be predicted using the CIE $L^*a^*b^*$ system. More complicated models have been developed that take these factors into account; the CIE have recently recommended the CIECAM97s model for further testing. Colour appearance models are becoming increasingly important in these days of digital images and global communication. The faithful reproduction of an image from computer screen to printed hard copy requires an effective colour appearance model at the heart of the colour management system.

## Summary

The use of colour-measuring instruments allows specification and communication of colour by means of international standard terms such as CIE $L^*a^*b^*C^*h°$. Colour difference values can also be determined instrumentally and, if optimized colour difference equations are used, can out-perform a single trained colourist. However, the visual appreciation of colour and colour difference is still a subjective response, affected by many factors. Care must be taken to include measurement conditions in the interpretation of colour measurement data. The gold standard answer must always be that which agrees with the majority of a group of human observers. Whether this can be achieved by optimized equations based on CIE $L^*a^*b^*$ colour space or by colour appearance models remains to be established.

## List of symbols

$a^*$ = redness/greenness; $b^*$ = yellowness/blueness; $c^*$ = chroma; $F_XF_YF_Z$ = functions of $XYZ$ values used in deriving $L^*a^*b^*$ values; $h°$ = line angle; $I(\lambda)$ = intensity of light entering the eye; $k$ = normalization constant; $L^*$ = lightness; $L^*, a^*, b^*$ = values in the CIE $L^*a^*b^*$ system; $L^*, C^*, h°$ = values in the CIE $L^*C^*h°$ system; $r(\lambda)$, $g(\lambda)$, $b(\lambda)$ = Guild/Wright colour-matching functions; $R(\lambda)$ = reflectance; $R,G,B$ = values in the RGB system; $S(\lambda)$ = spectral power distribution of light source; $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ = weighting functions replacing $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ in the $XYZ$ system; $X,Y,Z$ = tristimulus values in the $XYZ$ system; $\Delta E^*_{ab}$ = distance function in the CIE $L^*a^*b^*$ system.

*See also:* **Art Works Studied Using IR and Raman Spectroscopy; Colorimetry Theory; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy.**

## Further reading

Billmeyer FW Jr and Saltzman M (1981) *Principles of Colour Technology*, 2nd edn. New York: Wiley.

CIE Publication 15.2 (1986) *Colorimetry*. Vienna: CIE.
Hunt RWG (1987) *Measuring Colour*. Chichester: Ellis Horwood.
Hunter RS (1975) *The Measurement of Appearance*. New York: Wiley-Interscience.
Judd DB and Wyszecki G (1975) *Colour in Business, Science and Industry*, 3rd edn. New York: Wiley.
Kuehni RG (1997) Color: *An Introduction to Practice and Principles*. New York: Wiley-Interscience.
McDonald R (ed) (1997) *Colour Physics for Industry*, 2nd edn. Bradford: Society of Dyers and Colourists.
Nassau K (ed) (1998) *Color for Science, Art and Technology*. Amsterdam: Elsevier.
Wright WD (1969) *The Measurement of Colour*, 4th edn. London: Hilger.
Wyszecki G and Stiles WS (1982) *Color Science Concepts and Methods, Quantitative Data and Formulae*, 2nd edn. New York: Wiley.

# Computational Methods and Chemometrics in Near-IR Spectroscopy

**Paul Geladi**, Umeå University, Sweden
**Eigil Dåbakk**, Foss Sverige, Sollentuna, Sweden

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
Methods & Instrumentation

## Introduction

In near-infrared and related spectroscopic methods, spectra consisting of spectroscopic measures (absorbance, transmittance, reflectance) at a large number of wavelengths, frequencies, energies or wavenumbers (also called by the general term 'variables') are produced. A few hundred to a few thousand variables are not uncommon, even with fairly simple, medium-priced instruments. For most uses it is not the study of the spectra itself (spectroscopy) but extracting information about the samples (spectrometry) that is important. This requires a multivariate treatment. Some of the multivariate techniques used for this are presented here. Much of the explanation in this text will be based on figures showing spectra or plots visualizing analysis parameters and results. Since space is limited, equations are avoided in most instances and the reader is referred to the quoted literature. Some lesser used multivariate techniques will be given just as literature references. Literature on the near-infrared technique and the related infrared, Raman and Fourier transform techniques is easily found in this Encyclopedia.

## Raw spectra

Near-infrared (NIR) spectroscopy/spectrometry is the study of spectra recorded on instruments in the region 780–2500 nm. Many instruments add the visual region of 400–780 nm to this, so that colour information in the spectra can also be studied. Other instruments are more restrictive, e.g. between 850 and 1050 nm. Technical details about instrumentation can be found elsewhere.

It is important to obtain a visual impression of the measured spectra and this is done in line plots. In these plots the horizontal axis is wavelength or wavenumber and the vertical axis can be transmission, reflectance or some type of calculated absorbance. Sometimes it may be useful to study a scatter plot of two raw spectra. Many researchers and industries prefer to work with transformed spectra and smoothing and derivation are used for that purpose (see **Figure 1**). Although raw spectra can show a great deal about classes of samples and outliers, some of the finer details cannot be found in them. Therefore, a more thorough, often multivariate analysis is needed. **Figure 2** shows a typical data matrix obtained from NIR measurement. Some instruments only use 19 well selected wavelength bands and filters to give only 19 (or even fewer) variables. At the other extreme, some instruments give several thousands of wavelength variables.

## Exploration and classification

The first multivariate operation to be done on a data matrix (see **Figure 2**) is exploration of the data. Exploration of multivariate space is done by factor analysis methods, of which principal component analysis (PCA) is the easiest choice. The data are projected to a matrix with less and more meaningful variables called scores (latent variables) (see

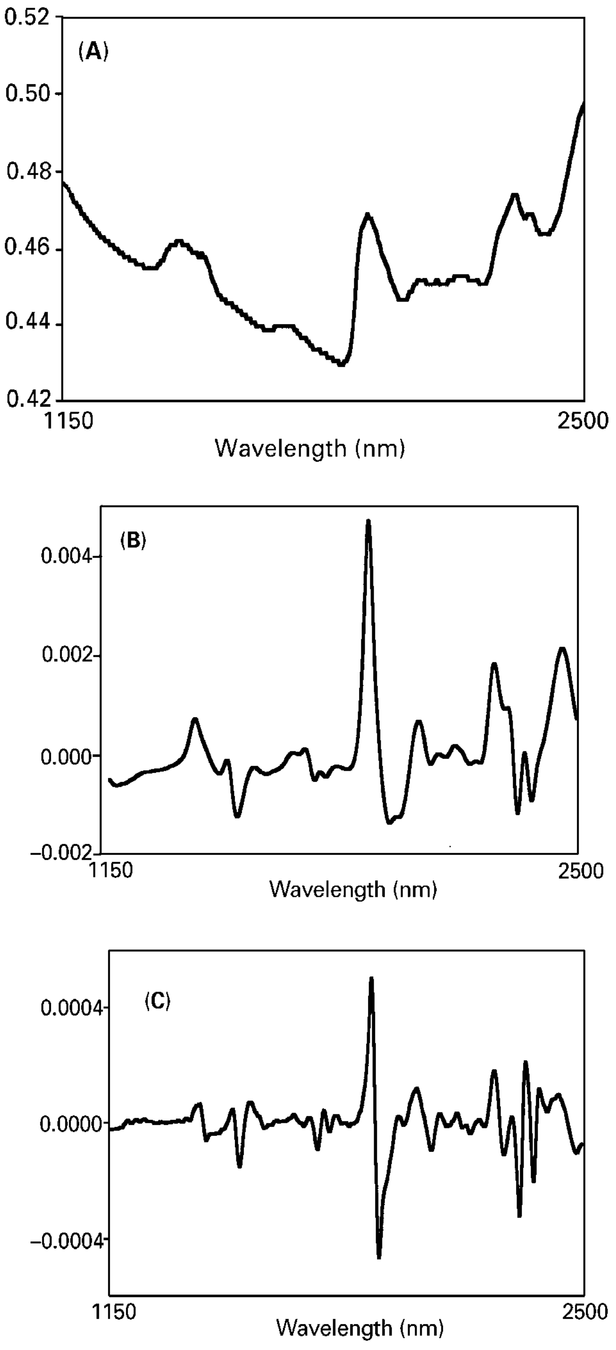



**Figure 1** (A) A typical NIR spectrum and its (B) first and (C) second derivatives. The raw spectrum is in absorbance units.

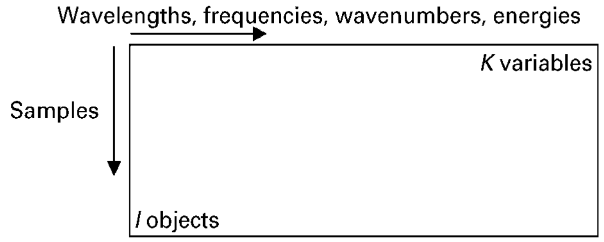



**Figure 2** A typical data matrix from an NIR study. The number of objects *I* may be a few tens to a few hundreds; *K* may be a few tens to several thousand.

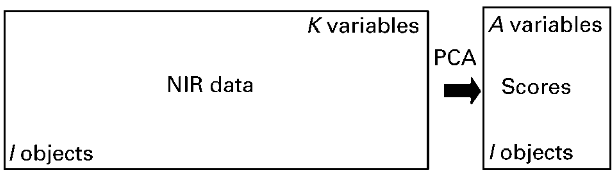



**Figure 3** Principal component analysis (PCA) projects the *K* variables into *A* principal component scores containing all meaningful information about the samples.

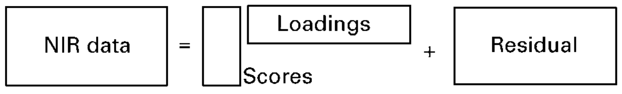



**Figure 4** PCA produces a score matrix and a loading matrix for *A* components and a residual matrix containing mainly noise. This is a data reduction.

**Figure 3**). In **Figure 4**, a more complete view of PCA is given. Scatter plots of the scores are called score plots. Score plots are 2D windows in multivariate space. The points in the plots are the objects. In this way, outliers may be detected and groupings of objects may be found. The loadings explain how the original variables contribute to the scores. They can be used to study spectral properties, or to select wavelengths. Loadings are usually shown as line plots of one loading vector or as scatter plots of two loading vectors.

Any analysis of NIR data should start with the detection of outliers in score plots. The outliers should be identified and removed. After this, grouping of objects may be observed. Distinct groups, overlapping groups and gradients between groups may be found. Grouping may be supervised, meaning known in advance, or it may be unsupervised i.e. found only after the analysis. With supervised grouping, one uses the score plot to check whether the expected (*a priori* known) groups really occur in the NIR spectra. Two different types of coffee are expected to have different spectra and if no difference can be found, the spectroscopic technique chosen may not be appropriate. Unsupervised grouping may show storage or sample treatment differences. A single type of coffee may give grouped spectra because different storage conditions have resulted in different humidity or because the grinding was performed in different machines.

Studies of score plots may be used for studying grouping, ageing, deterioration, erroneous processing

or counterfeiting of all kinds of organic products, e.g. food and feed products, cellulose, paper and textiles, wood products, fossil fuels and biofuels, seeds and growing crops, pharmaceuticals, packaging materials and environmental samples. An industrial application is to keep the spectra of satisfactory batches or products in a file and to compare new spectra to see if they still fit the 'satisfactory' group. A large collection of classification methods (sometimes called 'pattern recognition') is found in the literature, but many of these methods are less useful for NIR data because of the huge number of variables compared with objects. Also in this respect, a data reduction to scores is useful, allowing the use of the classical clustering and pattern recognition techniques on a limited number of score vectors. Statistical tests can be applied to the residual matrix.

## Calibration

One of the main purposes of measuring NIR data is the determination of chemical composition or physical properties in a quantitative way. It was found very early that NIR spectra contain information on the water, protein, carbohydrate and fat content of food products. Measuring a NIR spectrum is quick and cheap (30–60 samples per hour for solids) while the wet chemical methods for measuring water, fat, protein, and carbohydrates are slow and expensive. Hence if a function $y = f(\boldsymbol{x})$ can be found where $\boldsymbol{x}$ is the NIR spectrum and $y$ is the desired chemical concentration, then time and money are saved. As an extra advantage, the NIR spectral measurement can be automated to function continuously on a process stream. Constructing and testing the functions $y = f(\boldsymbol{x})$ is calibration. Using the function $f(\ )$ to find a value of $y$ with only $\boldsymbol{x}$ known is prediction. Many methods are available for finding $f(\ )$ and many diagnostics are available for checking its quality. Generally, for the calibration step, accurate and precise measures of $y$ and $\boldsymbol{x}$ are required. However, noise can never be avoided. Therefore, it is impossible to find an exact function $f(\ )$ and approximations have to be found. Linear approximations are easy to calculate, more easily interpreted and robust. Therefore, a linear model $y = \boldsymbol{xb} + f$ is often adequate. One important aspect to point out is that in the equation $y = \boldsymbol{xb} + f$ both $\boldsymbol{x}$ (the spectra) and $y$ (the response) have systematic and random noise.

There are many different techniques for building transfer functions between spectra and quantities (responses) measured by a reference method. Several techniques are found in the literature, but not all of them are frequently applied. The equation given above can be extended to $\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{f}$ or even to $\mathbf{Y} = \mathbf{XB} + \boldsymbol{F}$ (see **Figure 5**), with $\boldsymbol{y}$ $(I \times 1)$ a vector of response values for many samples, $\mathbf{Y}$ $(I \times M)$ a matrix of response values for many responses, $\boldsymbol{X}$ $(I \times K)$ the spectral data for the samples, $\boldsymbol{b}$ $(K \times 1)$ a vector of regression coefficients, $\mathbf{B}$ $(K \times M)$ a matrix of regression coefficients for the $M$ responses, $\boldsymbol{f}$ $(I \times 1)$ a vector of residuals and $\boldsymbol{F}$ $(I \times M)$ a vector of residuals for many responses. One may consider percentages of fat, water and protein as three different responses $(M = 3)$. The traditional least-squares solution for $\boldsymbol{b}$ $(\boldsymbol{B})$ is called multiple linear regression or ordinary least squares. The method was used in the pioneering days of NIR, when there were only few wavelength bands available.

The selected linear regression methods differ in the way the regression coefficients are estimated. They also have different properties for outlier detection and other diagnostics. In the early days it was found that wavelengths produced by spectrophotometers caused problems in a regression situation because of collinearity between the variables. There is also the problem of finding a least-squares solution to the equation above when the number of samples is less than the number of variables. Some selected methods are outlined below.

### Stepwise multiple linear regressions (SMLR)

SMLR is a way of cleverly reducing the number of wavelengths in order to find a least-squares solution, based on the meaningful assumption that repeated and useless information may be present in some wavelengths. In stepwise multiple linear regression, one starts with selecting the wavelength amongst all available ones, that correlates the best to the measured response. When the first wavelength is found, another wavelength is selected that increases the degree of explanation maximally and so on, until a stop criterion is fulfilled.

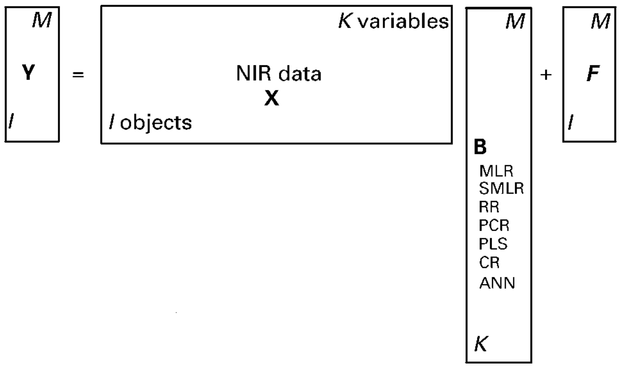



**Figure 5** A regression model is built between the matrix **X** with NIR spectra and the responses **Y**. The regression coefficients **B** have to be calculated by MLR, SMLR, PCR, PLS, RR, ANN, etc. ***F*** is the residual containing the noise.

### Principal component regression (PCR)

PCR uses the scores from a principal component analysis to avoid the problems of many variables and collinear variables described earlier. The PCA part also gets rid of some of the noise in the spectral data as an extra advantage. The regression solution becomes easier to calculate and more stable, and one obtains predictions that are more reliable. In addition, the PCA scores give the possibility of outlier detection, both for samples in the calibration data and for later samples for prediction.

### Partial least-squares regression (PLS)

PLS is similar to PCR, in that regression is done on scores. The difference is that the responses $y$ (Y) are used to find scores that have a large covariance between $X$ and $y$ (Y). One advantage of PLS over PCR is that the number of required components is reduced. Since PLS is used on scores, these can be used for the detection of outliers and groupings, as was explained for PCA. In this way, PLS automatically gives access to a number of diagnostic tools for outlier detection (in $X$ and $y$). PCR and PLS add an important parameter to the regression model: the number of components used in the model, A. This is an extra choice to be made. Too few components give a bad model and too many components give a model that is sensitive to noise. PLS, PCR and MLR variants are the most commonly used regression methods for quantification of multivariate spectral data. Other methods are ridge regression (RR) and continuum regression (CR).

### Artificial neural networks (ANNs)

The use of ANNs is increasing. ANNs are especially used for pattern recognition and nonlinear calibration. There are many types of ANN. However, it seems that the back-propagation ANNs are the most useful for calibration purposes. In linear calibrations, the traditional linear regression techniques will more often give stable and robust results as compared with ANNs. However, in nonlinear situations and with large heterogeneous data sets, ANNs work very well.

### Locally weighted regression (LWR)

LWR is based on the idea that even in a nonlinear situation there exist regions where a local model is close to linear. In LWR the neighbours of a spectrum to be used for prediction are used to build a local regression model. This model can be determined by PCR or PLS. There are different techniques for selecting appropriate neighbours, using different distance criteria and weighting (see **Figure 6**).

In addition to the building of regression models, testing them is also important. Therefore, the data are split into a training set (size $I$) on which the model is built and a test set (size $J$) that is used for testing the model in a simulated 'real' situation (see **Figure 7**). Checking how well this works is called

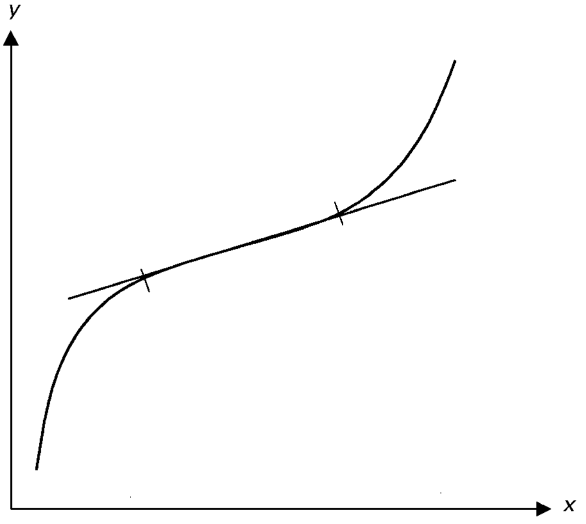



**Figure 6** Non-linear relationship between the variables $x$ and $y$. However, when the range is narrowed, local linear models can be used.

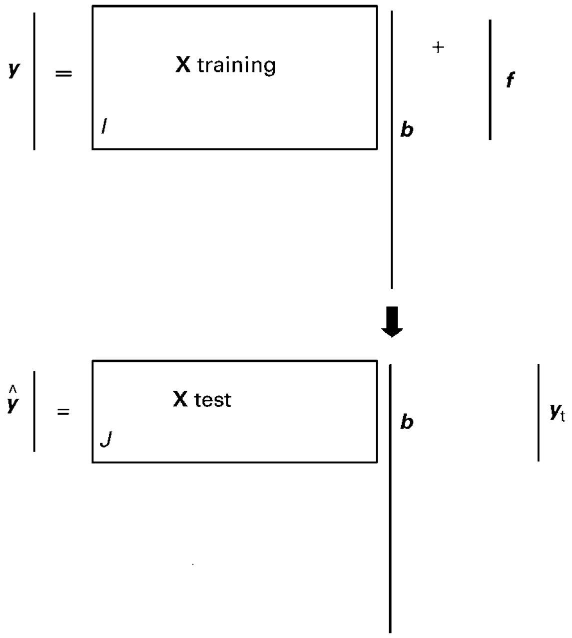



**Figure 7** The training set ($I$ objects is used to find the regression coefficients vector $\boldsymbol{b}$. With a test set ( $J$ objects), the predicted responses can be calculated using the same $\boldsymbol{b}$. Usually the responses for the test set are kept in the background for diagnostic checking.

validation and is explained in its own section below. Calibration is very important for NIR data since the classical interpretation of the spectra (as is done for IR and UV spectra) is nearly impossible. Because of this, one may state that there is a very close relationship between an NIR instrument and the calibration model, making an instrument useless if no calibration model is available. A complete flow chart for NIR data analysis is given in **Figure 8**.

## Validation

It is very important to test the quality of a regression model constructed by PCR, PLS or any other regression method. The model for the training set (see **Figure 7**) can be improved by using more PCR or PLS components, but some of these will only describe noise. With an appropriate test set, a check can be made whether the model also has predictive properties. A revealing test is to plot the predicted values for the test set against the measured values in a scatter plot. In this plot, all points should ideally fall on the diagonal. Outliers, systematic errors, groupings and noisy behaviour are all detected easily. Sometimes numbers are needed. The predicted responses; $\hat{y}$ for the test set can be compared with the measured values $y_t$. This is done by constructing a sum of squares PRESS = $(y_t - \hat{y})'(y_t - \hat{y})$ and the corresponding standard deviation RMSEP = $[\mathrm{PRESS}/J]^{1/2}$. A useful plot is that of PRESS (prediction residual sum of squares) or RMSEP (root mean square error of prediction) against the number of components in the model (see **Figure 9**). When $I$ and $J$ are very small because of lack of samples, cross validation may be used advantageously. **Figure 10** shows a typical predicted versus observed plot.

## Calibration transfer

The building of a good, robust calibration model with excellent prediction properties is extremely important and sometimes very expensive and time consuming. In many uses of NIR instruments and calibration models, long series of measurements are required and one has to be sure that the calibration model works well over long periods of time. This is not easy since instrument ageing, parts replacement and repairs may create the need for a new calibration model. Another situation arises when different instruments are used to do the same task in parallel. In order to avoid repeating the calibration, the techniques of calibration transfer have been developed. Some calibration transfer techniques are based on keeping individual instrument differences small or

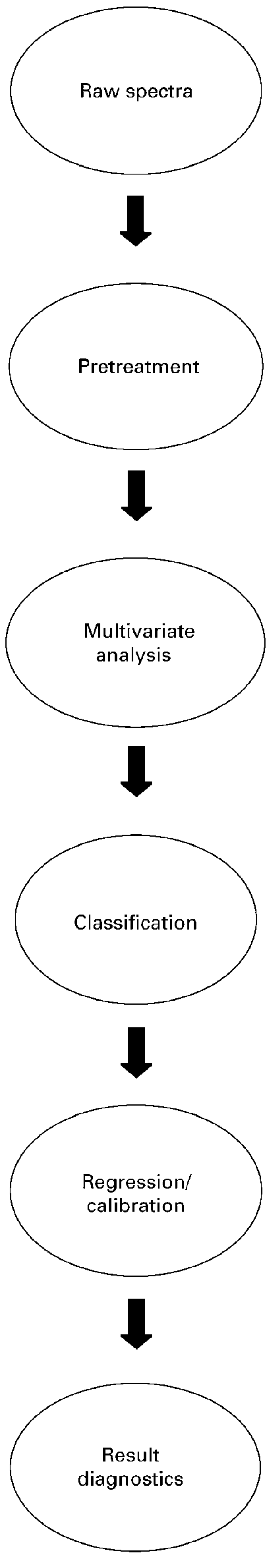



**Figure 8** A flow chart for a complete NIR data analysis.

compensating in the instrument hardware itself. Others work by calculations for making the spectra from one instrument (the field instrument) look like those of another one (the master instrument). In this way, the calibration model from the master instrument can still be used successfully.

## Pretreatment of the spectra

Several pretreatments are used in the processing of spectral data. The usual mean centring, scaling and nonlinear transformations are part of normal procedure related to classification and regression. Some other techniques were developed specifically with the nature of the NIR measurement in mind. The main purpose is always to improve the performance of a classification or calibration. Pretreatments can also be used to compress spectral information and ease computation on large sets of data, as is the case with Fourier and wavelet transforms. Multiplicative signal correction (MSC) was originally developed for NIR reflectance data in order to correct for different pathlengths and scatter arising from different particle sizes in the samples. This is done by adjusting baselines and slopes. The basis for MSC is simple. Each spectrum is compared with a reference spectrum, usually the mean spectrum of all good calibration samples. The sample spectrum is assigned the same offset and average slope as the reference spectrum. Orthogonal signal correction (OSC) works in a similar way to MSC but uses the responses for orthogonalization, thus ensuring that all information related to the responses stays in the transformed spectra. Derivatives (see **Figure 1**) encompass basically the same thing as MSC. A second derivative of a spectrum removes baseline offsets and linear slopes. Derivation (or differentiation) is a widely used, simple and intuitive data pretreatment. For digitized spectra, different smoothing algorithms can also be applied, e.g. using different widths of derivation.

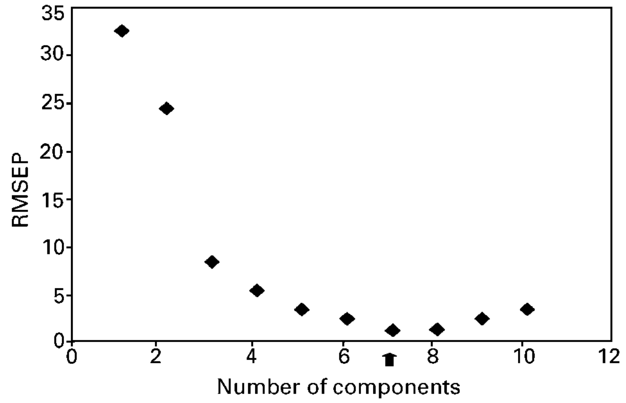



**Figure 9** A plot of RMSEP against the number of components used in the model gives a minimum (arrow) indicating how many components should be in the model and how low the prediction error can be made.

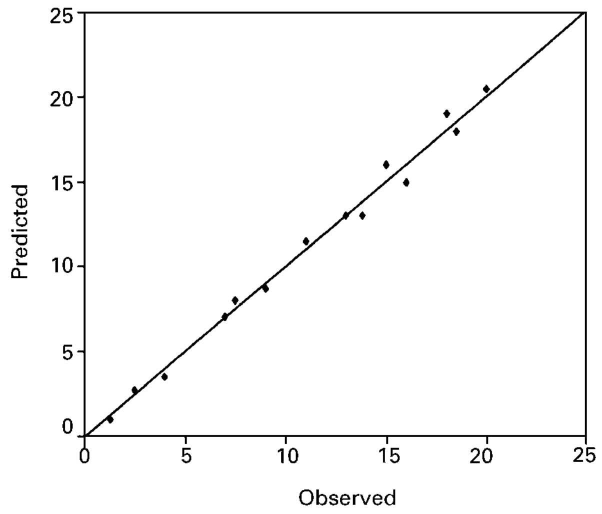



**Figure 10** Plotting predicted values of the response against observed values give an idea of prediction quality. All points should fall close to the diagonal. Individual outliers or deviating groups can easily be detected.

## List of symbols

$b$ = vector of regression coefficients; **B** = matrix of regression coefficients for many responses; $f$ = vector of residuals; $F$ = vector of residuals for many responses; PRESS = prediction residual sum of squares; RMSEP = root mean square error of prediction; $X$ = spectral data for samples; $y$ = vector of response values for many samples; **Y** = matrix of response values for many responses; $y_t$ = measured value; $\hat{y}$ = predicted response.

*See also:* **Calibration and Reference Systems (Regulatory Authorities); FT-Raman Spectroscopy, Applications; IR Spectrometers; IR Spectroscopy, Theory; Near-IR Spectrometers; Raman Spectrometers.**

## Further reading

Beale R and Jackson T (1990) *Neural Computing: an Introduction*. Bristol: Institute of Physics Publishing.

Beebe K, Pell R and Seasholtz M-B (1998) *Chemometrics. A Practical Guide*. New York: Wiley.

Brown P (1993) *Measurement, Regression and Calibration*. Oxford: Clarendon Press.

Davies A and Williams P (eds) (1996) *Near Infrared Spectroscopy: The Future Waves*. Chichester: NIR Publications.
Geladi P and Dåbakk E (1995) An overview of chemometrics applications in NIR spectrometry. *Journal of Near-Infrared Spectroscopy* 3: 119–132.
Johnson R and Wichern D (1982) *Applied Multivariate Statistical Analysis*. Englewood Cliffs, NJ: Prentice Hall.
Jackson J (1991) *A User's Guide to Principal Components*. New York: Wiley.
Joliffe J (1986) *Principal Component Analysis*. Berlin: Springer.
Malinowski E (1991) *Factor Analysis in Chemistry*, 2nd edn. New York: Wiley.
Mardia K, Kent J and Bibby J (1979) *Multivariate Analysis*. London: Academic Press.
Martens H and Naes T (1989) *Multivariate Calibration*. Chichester: Wiley.
Wold S, Esbensen K and Geladi P (1987) Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 2: 37–52.

# Contrast Mechanisms in MRI

**IR Young**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCES**
**Theory**

## Introduction

The key factor which has led to the huge growth and clinical acceptance of magnetic resonance imaging (MRI) is really quite fortuitous. This is that tissues have quite different values of $T_1$ and $T_2$, and that, more importantly, the relaxation time constants of a great many diseased tissues are quite markedly different from those of the normal structures surrounding them. Clinical MRI seeks to highlight these differences at the expense of all other considerations. As a result, all the methods used in traditional NMR have been investigated to find out if they can deliver maximal differentiation between normal and abnormal tissue. The manipulations are performed through complex sets of instructions issued to the various components of a machine at the correct times which are known as 'sequences'. In the following sections we discuss the most important of these, and their variants.

Thereafter a number of phenomena are described which are both artefacts of images if care is not taken, and sources of desirable contrast under other circumstances. Included among these are such things as the effect of flow, a matter of great concern in medical imaging where data about the cardiovascular system is of substantial importance in many clinical conditions. Some of the implications of apparently recherché observations such as that of diffusion are very significant in the diagnosis of unexpected disease situations.

An important, though unfortunately brief, note discusses the single most significant artefact in MRI, arising from the physiological measurements of the patient. The problem of controlling the effect of respiratory motion in particular has delayed the application of MRI to the imaging of much of the thorax and abdomen very substantially.

## Standard sequences

Most clinical MRI revolves around the contrast generated by three very standard sequences, partial saturation, spin echo and inversion recovery, which are familiar to those working on traditional high-resolution spectroscopy, but having a resonance and significance in clinical MRI not found elsewhere. The notation used in this article for times and other sequence features is that proposed by the American College of Radiology, and generally used in human MR studies.

### (Partial) saturation recovery sequence

This is the classic $\alpha$–TR–$\alpha$ sequence of NMR in imaging form. The sequence form, showing how the various gradients and RF pulses are manipulated for a spin-warp data acquisition, is sketched in **Figure 1A**. The dotted lines indicate that the phase encoding gradient is varied from one acquisition to the next. The theoretical relationship for the sequence is:

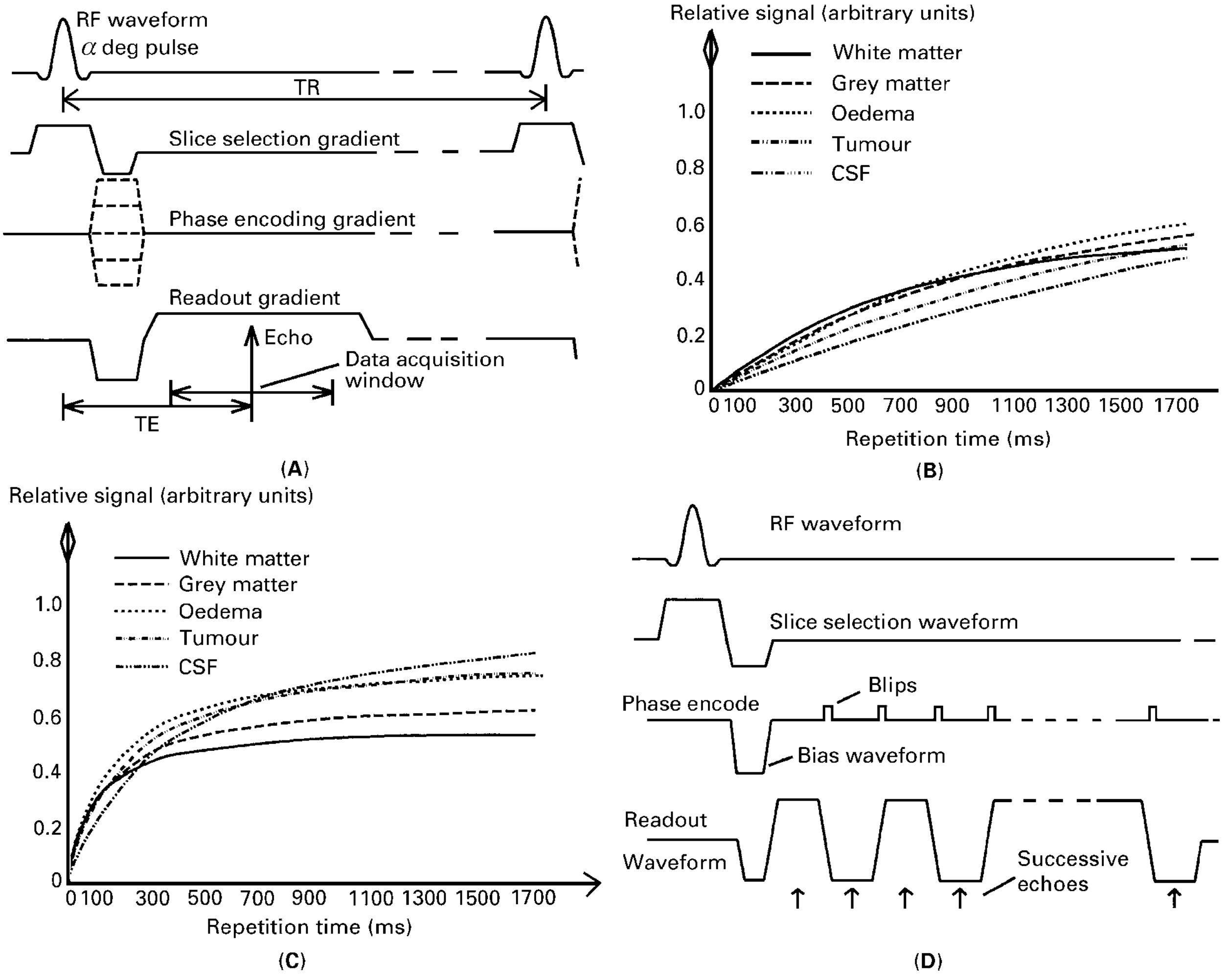



**Figure 1** (Partial) saturation recovery sequence. (A) Sequence configuration. A flip angle of $\alpha$ is assumed in this diagram. (B) Calculated signal characterization for the tissues, the parameters for which are given in **Table 1**. $\alpha$ in this case is 90°. (TE = 20 ms for the purposes of this plot.) (C) The same characteristics as in **Figure 1B**, but with $\alpha$ = 30°. (TE = 20 ms for the purposes of this plot.) (D) Blipped echo planar sequence. An echo is formed during each excursion of the readout gradient, with a line of *k*-space being acquired at the same time. The bias pulse in the phase encoded waveform is used to control the starting point in *k*-space of the image acquisition.

$$S = \frac{S_0[1 - \exp(-\mathrm{TR}/T_1)]\exp(-\mathrm{TE}/T_2^*)\sin\alpha}{1 - \exp(-\mathrm{TR}/T_1)\cos\alpha} \quad [1]$$

where $S$ is the observed signal, and $S_0$ is that which is available after full relaxation and immediately prior to excitation. In this sequence TR is the time between successive excitations, while TE is the time from the centre of the excitation to the centre of the echo formed by the refocusing effects of the gradients. $\alpha$ is the actual flip angle. Note that this relationship is much simplified from that derived by solution (see below) of the Bloch equations, by assuming $T_1 >> T_2$ and that off-resonance effects and spatial selection effects are ignored.

What this means in practice in terms of signal differences and behaviour is shown in **Figure 1B**. This

**Table 1** Tissues, and their parameters, used in the description of the main sequences

| *Tissue* | $T_1$ (ms) | $T_2$ (ms) | Relative available proton density |
|---|---|---|---|
| White matter | 670 | 85 | 0.7 |
| Grey matter | 920 | 95 | 0.8 |
| Oedema | 1060 | 150 | 0.9 |
| Tumour | 1410 | 200 | 0.9 |
| CSF | 2500 | 500 | 1.00 |
| Fat | 265 | 50 | 0.6 |

The relative available proton density column indicates the relative peak signal observable from the tissue. Cerebrospinal fluid (CSF) which has a signal near to that of pure water in normal subjects is taken as unity.

(and other similar figures to follow) uses the values in **Table 1** for typical normal and abnormal brain tissue. It is the differences between these components that decide the efficacy of MRI. The values in **Table 1** are typical of those at 1.0 T. The coefficient, $\nu$, relating field and relaxation time constant variation for grey matter and white matter can be taken to be ~0.3, for blood and oedema as ~0.2, for tumour as ~0.1 and for cerebrospinal fluid (CSF) as being ~0.0. Note that this implies a tendency for relative contrast levels to fall as the main field ($B_o$) increases. (Note: $\nu$ is the coefficient in the term ($T_1 \alpha B_o^{\nu}$).)

$$S = \frac{S\exp(-\mathrm{TE}/T_2)[1-\exp(-\mathrm{TR}/T_1)]\{\sin\alpha\cos\beta\exp(-\mathrm{TE}/2T_1)-\sin\beta[1-\exp(-\mathrm{TE}/2T_1)(1-\delta\cos\alpha)\}}{1+\exp(-\mathrm{TR}/T_1)(\cos\alpha\cos\beta+\sin\alpha\sin\beta)} \quad [2]$$

**Figure 1C** shows how the contrasts obtained are radically affected by reducing $\alpha$ from 90° (the value in **Figure 1B**).

Echo planar imaging is an extreme example of this sequence. In this procedure, after a single excitation, one of the gradients (the readout gradient) is switched to form a series of echoes during each of which one line of *k*-space is acquired, while the equivalent of the phase encoding gradient is incremental in some way. Though the original form of the sequence was rather different, the sequence form in current use ('blipped' echo planar) is shown in **Figure 1D**.

### Spin-echo sequence

The partial-saturation sequence is typically used with short values of TE (except in special circumstances) and is thus primarily dependent for its contrast on differences in $T_1$ (i.e. it has '$T_1$- weighting'). The spin-echo sequence, though it is poor from the point-of view of discriminating between normal anatomical structures, has proved to be extraordinarily effective in identifying a very wide range of pathological conditions, once it had been successfully implemented using relatively long TR and TE to give '$T_2$-weighting'. Initially applied to normal anatomy by the current author more than 2 years before its successful application clinically, it had seemed more or less completely uninteresting for MRI. Unlike inversion recovery, which generated spectacular contrast between normal tissues such as grey and white matter (see below), an inappropriate choice of the echo time (TE) can result in very little contrast indeed in normal organs (as **Figure 2B** suggests).

**Figure 2A** shows a typical sequence structure for a spin-echo acquisition. The simplified relationship governing the sequence is:

and

$$\delta = \frac{1-\exp(-\mathrm{TR}/T_1)}{1+\exp(-\mathrm{TR}/T_1)(\cos\alpha\cos\beta+\sin\alpha\sin\beta)}$$

where $\alpha$ and $\beta$ are respectively the excitation and refocusing pulse flip angles. When these angles are respectively 90° and 180° (and $\alpha$ in Equation [1] is 90°) the relationships for the partial saturation and spin-echo sequences look misleadingly similar, with the difference between $T_2$* and $T_2$ being the only apparent change. Spin echo was very popular in the early days of clinical MRI because it was robust, and tolerant of many machine defects. In practice, while still regarded as an essential screening tool, it is now seen as being a relatively blunt instrument in terms of the subtlety of contrast that can be obtained with it.

**Figures 2C** and **2D** show contrast characteristics for a value of TR less than that of typical normal tissue, and a value of TR substantially longer, respectively. The characteristics suggest how choice of sequence parameters can emphasize, or eliminate, contrast between any two tissues.

The fast spin-echo (rapid acquisition relaxation enhanced (RARE)) sequence is the spin-echo equivalent of echo planar imaging, and its structure is shown in **Figure 2E**. Contrast is manipulated by choosing at which point in the echo train the central point of *k*-space is acquired.

### Inversion recovery sequence

Although two of the very early developers of MRI placed substantial weight on it, the inversion recovery sequence was regarded through much of the development of MRI in the 1980s as being a difficult sequence to operate, and saddled with long acquisition times. The sequence is shown schematically in **Figure 3A**, and is governed by the relationship:

$$S = S_o\frac{\{[1-\exp(-(\mathrm{TR}-\mathrm{TI})/T_1)]\exp(-\mathrm{TI}/\mathrm{T}_1)\cos\beta\sin\alpha-[-\exp(-\mathrm{TI}/T_1)][\sin\alpha+\exp(-(\mathrm{TR}-\mathrm{TI})/T_1)]\}\exp(-\mathrm{TE}/T_2)}{1-\cos\beta\cos\alpha\exp(-\mathrm{TR}/T_1)+\sin\alpha\sin\beta\exp(-(\mathrm{TR}-\mathrm{TI})/T_1)\exp(-\mathrm{TI}/T_2)} \quad [3]$$

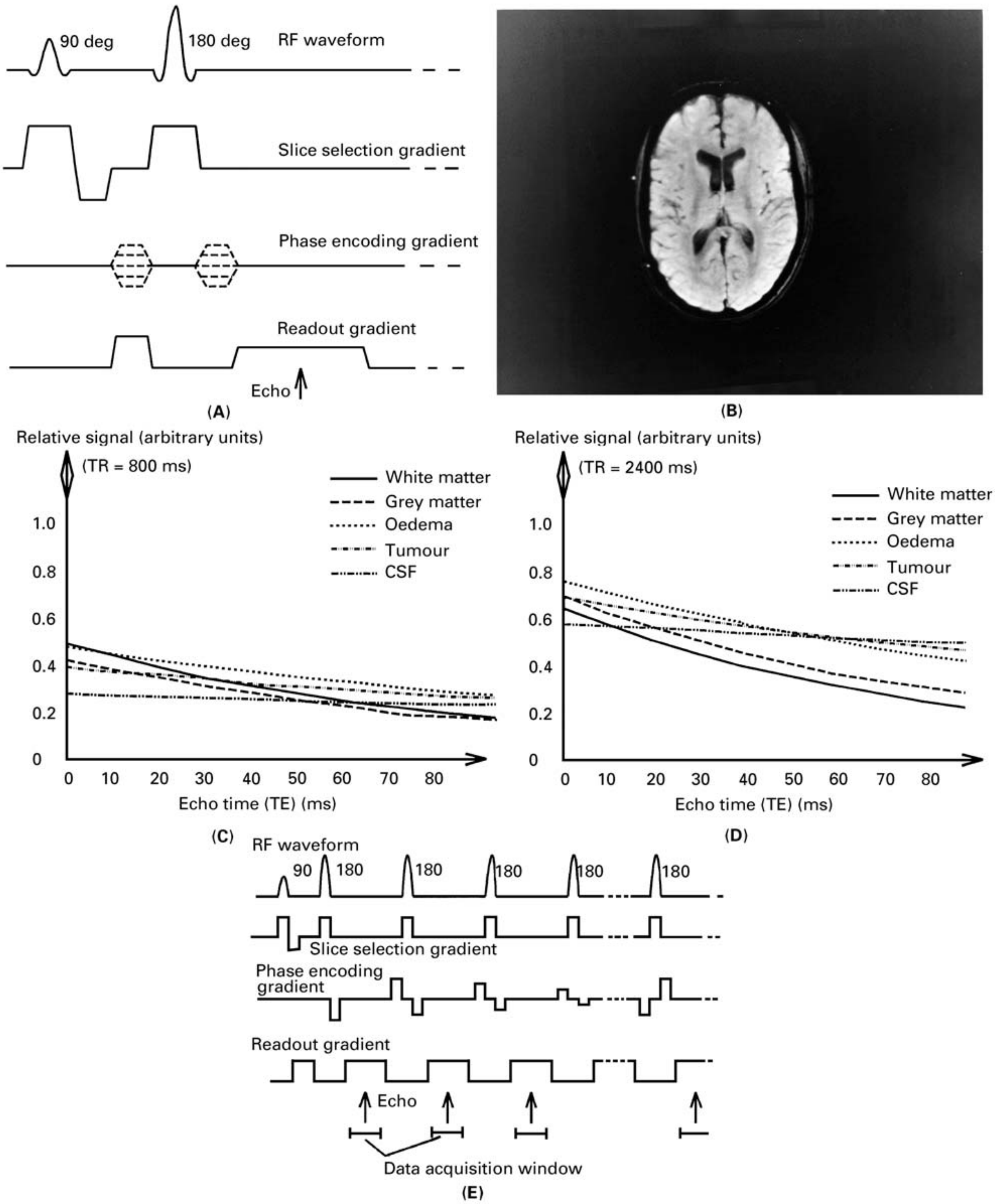



**Figure 2** Spin echo sequence. (A) Sequence configuration. Both 90 and 180° pulses are shown as being selective. The splitting of the phase encoding gradient (one half being in the opposite sense to the other) is quite normal, as is the placing of the warping gradient before the 180° pulse. These help a little with artefact generation. (B) Spin-echo image taken with TR 860 ms and TE 20 ms, showing how small the contrast between normal tissues can become if an inappropriate choice of parameters is made. (Slice 4 mm; matrix 256 × 192; field of view 22 cm.) (C) Signal characteristics for the spin-echo sequence, for the tissues given in **Table 1**. $\alpha$ = 90° and $\beta$ = 180°; in this case TR (800 ms) was about half that of the mean of the $T_1$ values of white and grey matter. (D) Signal characteristics as in (C), but with TR being three times that of the mean of the $T_1$ values of grey and white matter at 2400 ms. (E) Schematic drawing of the RARE sequence (shown in this case for an echo train of 8: the letters 'ETL' in a sequence description refer to this value). Actual sequences are much more complex and subtle.

(where TI is inversion time) which simplifies to the familiar form:

$$S = S_0\left[1 - 2\exp\left(-\frac{\mathrm{TI}}{T_1}\right) + \exp\left(-\frac{\mathrm{TR}}{T_1}\right)\right]\exp\left(-\frac{\mathrm{TE}}{T_2}\right) \quad [4]$$

when $\alpha = 90°$ and $\beta = 180°$. A gradient-recalled echo data acquisition is assumed for simplicity (as the equation, in its spin-echo form, becomes very cumbersome).

Tissue characteristics for the sequence, with a TR assumed to be long compared with the $T_1$ of both grey and white matter, are shown in **Figure 3B** for the same tissues as before, but with the addition of fat.

This sequence is unique in having both positive and negative signals and exploitation of this leads to its versatility. **Figure 3C** shows an image acquired with a typical TI in the brain using real reconstruction (i.e. retaining the sense of the signal signs). **Figure 3D** shows the same data reconstructed as a magnitude map. The TI was such that grey and white matter on the one hand, and CSF on the other, have signals of alternate sign. Regions where two tissues with alternate magnetization overlap show reduced intensities as their signals oppose each other. The black rims round the ventricles are very obvious, but the unwary investigator can readily mistake more subtle differences, due to this effect, as having clinical significance.

The special property of inversion recovery, that a value of TI can be chosen which results in no signal from anything with the appropriate $T_1$, has turned out to be a surprisingly powerful tool. From Equation [4], when TR > $3T_1$ (or thereabouts), in effect the null point is obtained when $\mathrm{TI}/T_1 \approx \ln 0.5$. The use of this form of cancellation was first applied to fat (in the short tau inversion recovery (STIR) sequence). Because, generally speaking, tissues with long $T_1$ times also have substantial $T_2$ times (practically all of which are longer than that of fat), in practice this means that tissue contrasts dependent on $T_1$ and $T_2$ are cumulative for practically all components, giving the approach additional diagnostic power. At the other extreme, cancellation of CSF might not seem to be as obviously beneficial. However, CSF behaves not unlike pure water, and its $T_1$ and $T_2$ are very long. Artefacts from its very large signals (due to tissue pulsatility, in the main) contaminate very many of the spin-echo acquisitions, particularly those with longer TEs, and elimination of the CSF signals has turned out to have substantial diagnostic benefits, particularly in the cervical spine.

Two inversions (with the cancellation of signals from two tissues) have also been studied using the sequence shown in **Figure 3E**, which results in the relationship:

$$S = S_0\left[1 - 2\exp\left(-\frac{T_\mathrm{B}}{T_1}\right) + 2\exp\left(-\frac{T_\mathrm{A}+T_\mathrm{B}}{T_1}\right) - \exp\left(-\frac{\mathrm{TR}}{T_1}\right)\right]\exp(-\mathrm{TE}_2) \quad [5]$$

For simplicity all flip angles are assumed to be either 90° or 180° as appropriate. The sequence has had limited clinical application. An example of its use, to cancel the signals for white matter and CSF while using a long TE data acquisition to attenuate the fat signal, is shown in **Figure 3F**. The residual tissue is predominantly grey matter. Attempts to extend the method to use more inversions still are unproductive as the signal-to-noise ratio of the images becomes unacceptable.

## Artefacts

Radiologists are trained to be suspicious of even slightly unusual results, and to examine them very carefully for the possibility that they are artefacts. Spectroscopists (and other more traditional investigators using NMR) are less aware of the possibility, though more conscious that there may be problems in multidimensional Fourier transform NMR. Nevertheless, such terms as 'F1 noise' used to describe something which is actually artefact is indicative of users' perceptions.

The utility of the majority of clinical MRI scans is, however, not limited at all by signal-to-noise ratio concerns, even in quite low fields, but by artefact. Since this is a function of signal, arising, in most instances, from any instability at all in the data set, in effect the value of MRI images obtained from a study is quite substantially independent of field level, but dictated by the performance of the machine in use, and the skill of the operators in positioning the patient, choosing the sequences to use, and in persuading the patient to remain still.

Artefact analysis is a huge topic in its own right, and no more than a very cursory glance at it is possible here. Two different forms of artefact will be considered in very broad outline, to illustrate two quite different errors that can occur. These are those arising from slice selection problems, and those due to

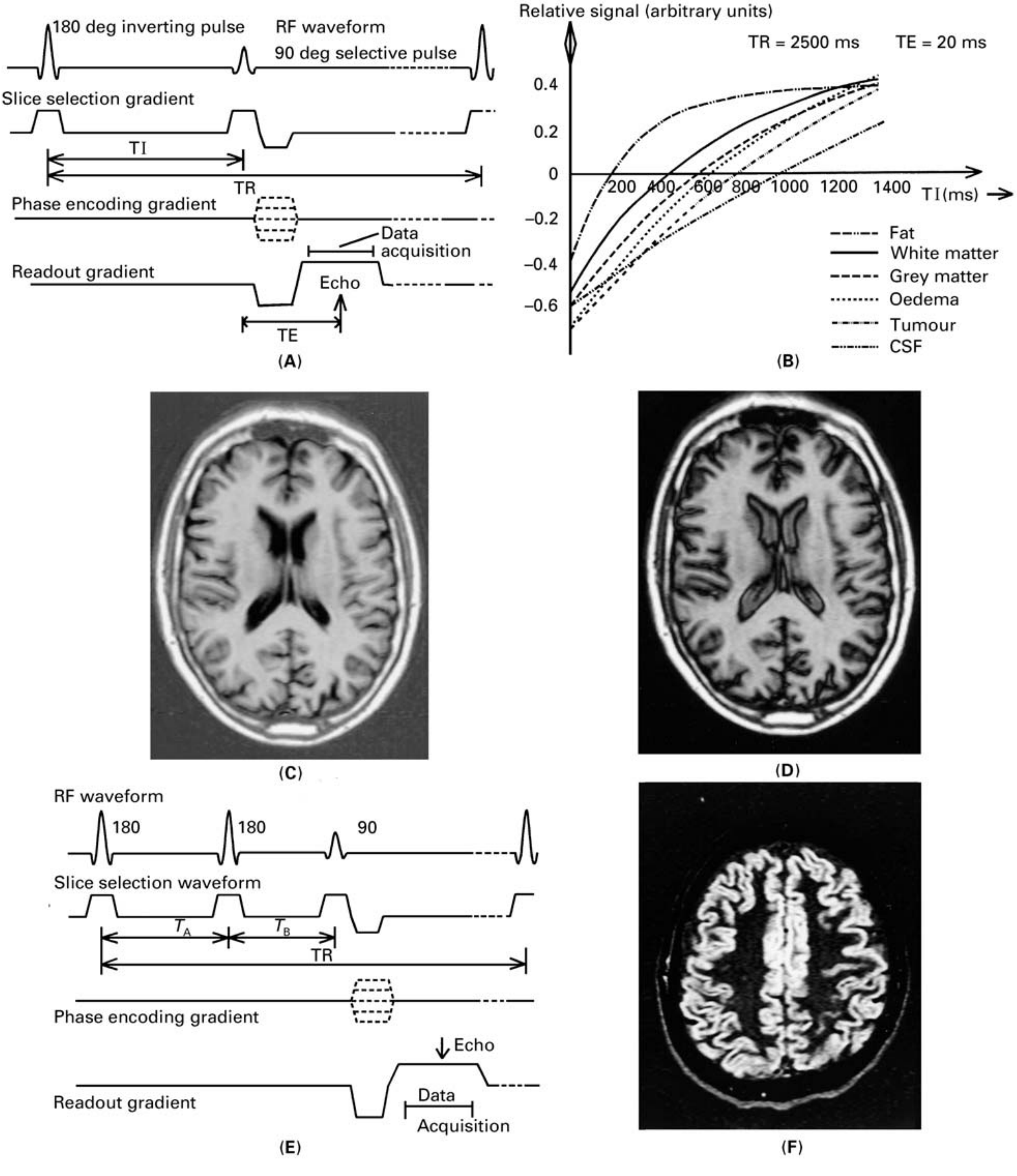



**Figure 3** The inversion recovery sequence. (A) The form of the inversion recovery sequence. The initial inverting 180° pulse is normally selective (as shown) but in some situations this is omitted. The data acquisition shown is a gradient recalled echo, but, very frequently, a spin echo is used. (B) Characteristics of the inversion recovery for different values of TI at fixed, relatively long, TR (greater than twice the mean value of the $T_1$ values of grey and white matter). Because of its importance in some inversion recovery sequences the fat characteristic is also included. (C) Inversion recovery image using real reconstruction (TR 2720 ms, TI 600 ms, TE 30 ms) (128 × 256 matrix, 25 cm field of view; 2 acquisitions; 6 mm slice). (D) Magnitude reconstruction of the same inversion recovery image data as in (C). The sequence was designed so that grey and white matter and CSF have opposite signs. (E) Sequence for a double inversion recovery acquisition. The terms on the figure are those used in Equation [5]. A gradient-recalled echo data recovery is shown, though spin echoes are very frequently used. (F) Double inversion recovery image of the brain of a volunteer with long TE (80 ms), leaving the cortex of the brain as the sole remaining tissue with substantial signal. (TR 6000 ms; $T_A$ 2300 ms; $T_B$ 245 ms; 3 mm slice; 128 × 256 matrix.)

motion artefact (which, ranging from the gross to the microscopic, is the most pervasive of all).

To illustrate slice selection problems, it is useful to begin with the fuller expression for selective excitation [2] for a partial saturation sequence:

$$S = S_o \frac{\exp(-\mathrm{TE}/T_2)[1 - \exp(-\mathrm{TR}/T_1)]\sin\alpha}{1 - \exp(-\mathrm{TR}/T_1)\cos\alpha - [\exp(-\mathrm{TR}/T_1) - \cos\alpha]\exp(-\mathrm{TR}/T_2)} \quad [6]$$

Note that this ignores effects due to the spatial localization process. Assuming the rather benign shape for the RF spectrum shown in **Figure 4A** (which has relatively sharp edges and little overshoot) **Figures 4B** and **4C** show, respectively, the predicted slice shapes for two tissues, one with TR = $2T_1$, the other with TR = $0.2T_1$. In a typical head image in which the range of $T_1$ values is ~12 (from fat at 250 ms to CSF at 3000 ms), the slice shape changes quite markedly from place to place across the image. Regions in which partial volume effects mean that a single voxel can include two tissues with dramatically different $T_1$ values result in very different shapes at either side of a single slice. It should be borne in mind that in slice selective MRI, the slice thickness dimension is usually the largest one in a voxel by a substantial margin (at least a factor of 2, and more usually 5–10). Measurements of tissue parameters made from such data must be treated with great care, and the appreciation that the signals may well be badly contaminated by unwanted material.

Motion artefact has proved to be one of MRI's greatest challenges, and the most intractable of problems. It is generally associated with abdominal motion, but can occur anywhere due to either voluntary or involuntary motion of the patient, or regular physiological effects such as blood pulsatility. Motion can be very large, or very small.

The periodicity of the effects is given by (for motion in the phase encoding direction; other directions have similar forms of relationship):

$$\frac{\Delta y}{y_{\mathrm{f}}} = FN\mathrm{TR} \quad [7]$$

where $\Delta y$ is the spacing in the $y$ (phase encoding) direction, $y_f$ is the size of the field of view in that direction, $F$ is the frequency of the motion, $N$ the number of averages of each phase encoding step which is acquired and TR is the repetition time.

Their amplitude is given by:

$$S_m = \left| F\left[ J_m\left( \frac{\pi A \Delta G_y}{y_{\mathrm{v}} G_{y\max}} \right) \right] \right| \quad [8]$$

for the $m$th ghost. $F$ is the Fourier transform function, $J_m$ the $m$th-order Bessel function of the first kind, $\Delta G_y$, is the gradient movement, $G_{y\max}$ the maximum value of the gradient, $A$ the amplitude of the motion and $y_v$ the length of a voxel in the $y$-direction. This assumes regular motion, which holds only to a limited extent. Irregular behaviour is largely unmodellable and probably not productively so in general.

A variety of methods are used to attempt to control the effects, ranging from rapid imaging with breathholding, to the relatively sophisticated probability assessments used in the respiratory ordered phase encoding method, or attempts to use internal markers to detect motion (as in the Navigator methods). These techniques are aimed primarily at controlling effects in the phase encoding direction(s). In the read-out or slice selection directions, methods such as motion artefact suppression techniques, which rely on the use of multiple gradient pulses to

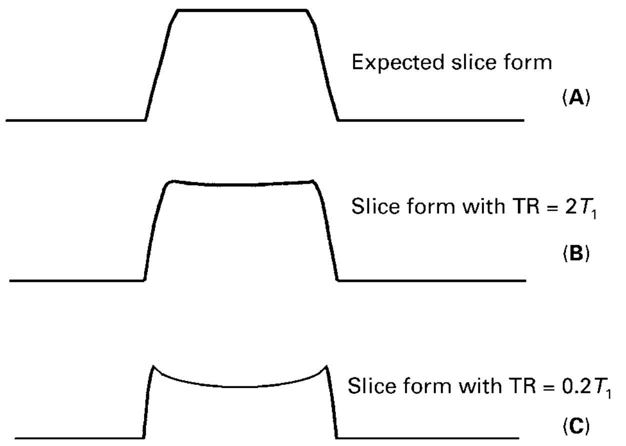



**Figure 4** Slice selection artefact. (A) Target slice shape in a partial saturation data recovery (equivalent to the spectral distribution of the RF pulse). This is rather more rounded than is achieved in the best formulations, but has less in the way of undershoot and oscillation at the edges, than is typically achieved. (B) Predicted slice shape for a tissue with TR = $2T_1$. Note that $T_2$ effects are ignored. The pulses in (A) and (B) are scaled in the same way. (C) Predicted slice shape for a tissue with TR = $0.2T_1$. This pulse is scaled at twice that of those above.

refocus tissues moving with first, second, third and potentially more motion orders, are quite effective.

## Susceptibility effects

Susceptibility effects have a significance for MRI that goes far beyond the concerns high-resolution spectrosocpists have about them as a potential source of line broadening and loss of resolution in their spectra. This is because, as with so many things in MRI, they are also a means of generating contrast, and of recognizing disease patterns. Apart from effects in tissue adjacent to air (as in the sinuses) and, to some extent, trabecular bone, both components having significantly different susceptibilities to that of tissue, the important applications of susceptibility relate to the presence of blood. While oxygenated blood is diamagnetic (possibly more so than normal tissue), deoxygenated blood is mildly paramagnetic, while the degradation products of blood (found in haemorrhagic deposits) become increasingly strongly paramagnetic as they change from methaemoglobin, through haemosiderin to ferritin.

Susceptibility effects cause changes to both expected signal intensities and their phases, both of which can be observed in clinical MRI. The phase changes given by the mean field deviations in voxels with sides $x_m$, $y_m$, $z_m$ and field deviation $\Delta B_{0xyz}$ at a point $x$, $y$, $z$ inside them are:

$$\varphi = \frac{\gamma \mathrm{TE} \int_{x=0}^{x_m} \int_{y=0}^{y_m} \int_{z=0}^{z_m} \Delta B_{0xyz} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z}{x_m y_m z_m} \qquad [9]$$

while the amplitude differences are:

$$S = \left[1 - \exp\left(-\frac{\mathrm{TR}}{T_1}\right)\right] \exp\left(-\frac{\mathrm{TE}}{T_2}\right) \times \int_{x=0}^{x_m} \int_{y=0}^{y_m} \int_{z=0}^{z_m} S_{xyz} \exp(\gamma \Delta B_{0xyz} \mathrm{TE}) \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \qquad [10]$$

where $S_{xyz}$ is the available signal from location d$z$, d$y$, d$x$ and $\Delta B_{0xyz}$ is the local field deviation. TE is the echo time as usual. Relaxation effects have been assumed to be constant over the whole voxel.

For a linear gradient in the slice selection direction, for example, which is, as noted before, the largest dimension of typical voxels, the result is:

$$S = [S_o \operatorname{sinc} \theta]_{-\theta_0}^{+\theta_0} \qquad [11]$$

where $\theta = \gamma G_s s \mathrm{TE}$, with $G_s$ being the erroneous gradient in the slice selection direction and $\theta_o$ is the value of $\theta$ at the voxel boundary.

Susceptibility effects have recently taken on a great significance in MRI, in the form of the suggested basis of the blood oxygen level-dependent approach to the observation of brain function. The experiment is modelled as arising from the diffusion of water molecules in regions with changing concentrations of paramagnetic deoxyhaemoglobin.

## Flow effects

Apart from its tendency to generate artefacts, flow is very well visualized by MRI. The approaches used to measure flow are variants of the gradient method and have been developed as a quantitative method by Bryant *et al.* who used the flow sequence shown in **Figure 5**. Using the parameters shown in the figure the velocity of flow ($v$) is measured as a phase difference ($\varphi$) and given by

$$\varphi = \gamma G \delta \Delta v \qquad [12]$$

Flow velocity is then linearly related to phase, though care has to been taken with the choice of sequence parameters to ensure that the velocities being measured do not exceed the available range (as the unrestrained signal pattern is a repetitive saw-tooth, with consequent ‘wrap round’, and potential ambiguity).

Flow is used in a wide range of applications in MRI of which the most important is angiography. Basically there are two strategies used for this. One is ‘phase contrast’ based on a strategy derivative from Equation [12] but using very short echo partial saturation (or gradient recalled echo) sequences. The alternative method is the ‘time of flight’ approach, in which spins outside a region to be studied are tagged in some way, and their future progress monitored.

## Diffusion effects

Diffusion effects have long been studied in NMR. The pulse sequence used is identical to that of **Figure 5**, though the gradient amplitudes are much larger. The relationship for the signal appropriate for an experiment involving tissue diffusion is:

$$S = S_0\left[1 - \exp\left(-\frac{\mathrm{TR}}{T_1}\right)\right] \exp\left[-\frac{\mathrm{TE}}{T_2} - D\gamma^2 G^2 \delta^2 \left(\delta - \frac{\Delta}{3}\right)\right] \qquad [13]$$

which can be reformulated as:

$$S = S' \exp\left(-\frac{\mathrm{TE}}{T_2} - bD^*\right) \qquad [14]$$

where

$$b = \gamma^2 G^2 \delta^2 \left(\delta - \frac{\Delta}{3}\right)$$

$D^*$ (the apparent diffusion coefficient) can be calculated from the results of two experiments which are identical except for the values of $b$. (Note: the qualification by which this experiment is assumed to measure $D^*$ and not $D$, the true diffusion coefficient, is a reflection of the uncertainty about the real diffusion process in tissue, as is further suggested in the next paragraph.)

However, the model for diffusion on which this theory is based assumes a uniform isotropic medium. In actuality, as evaluated in human subjects, diffusion in tissue is anything but isotropic. The coefficient can appear to be quite different depending on the direction of sensitization. This topic is much too large to be covered adequately here, but the anisotropy is of surprising clinical importance, particularly in the study of stroke and the development of myelination in the infant brain.

## Magnetization transfer contrast (MTC)

Another method which MRI has adopted from traditional NMR is MTC. This seeks to investigate chemical and magnetization exchange between what are typically NMR visible and invisible spin pools. Normally the system is modelled as having two pools (A, visible, and B, invisible in the experiment being performed), though three or more might be necessary to explain what really happens. This concept was applied to MRI, and it is possible to map the different behaviours of various tissues. The model is shown in **Figure 6A**, and the form of sequence used to implement the technique in **Figure 6B**. The prolonged irradiation that is used is applied off-resonance, and in the traditional experiment the amplitudes of the residual signal are plotted against the offset frequencies (at constant RF amplitude).

In MRI, useful contrast improvements are achieved by MTC in angiography and in haemorrhage studies. The time needed to achieve saturation of the pool with the broad-line has to be minimized in MRI (for thermal deposition reasons), and the effective time constant which applies is that appropriate to the

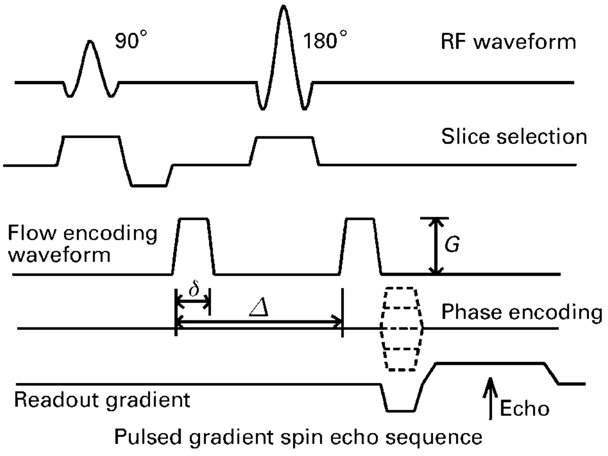



**Figure 5** Sequence for flow measurement. The parameters used in Equation [12] are marked on the figure. Note that the sensitizing pair of gradient pulses can be included in any of the gradients (or any combination of them).

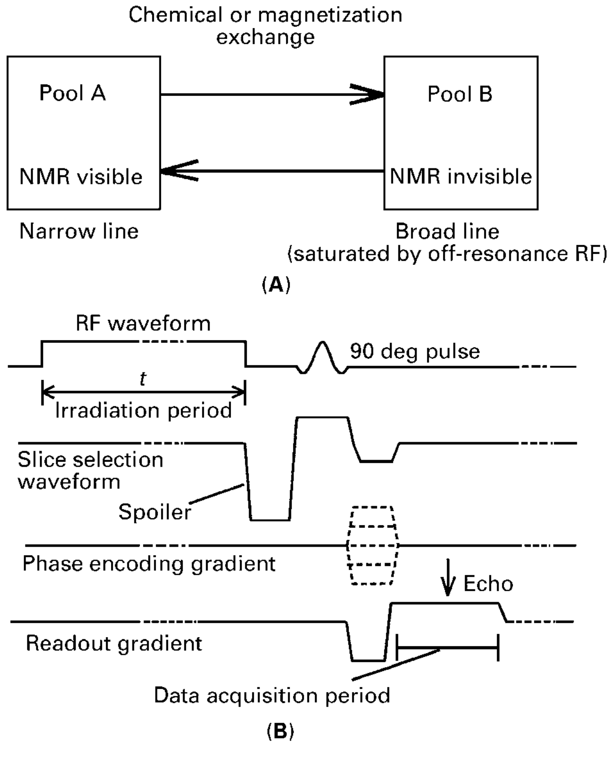



**Figure 6** Magnetization transfer technique. (A) Two-pool model used for the evaluation. This is certainly a simplification of reality, although imaging data quality does not justify further elaboration. (B) MTC imaging sequence. A gradient-recalled echo data acquisition is used once more. The spoiler shown is used to dephase residual transverse magnetization after the end of the off-resonance irradiation period (of length $t$). Practical sequences of all types use pulses like this to eliminate the unwanted signals from spins pursuing complex patterns of excitation and recovery.

signal relationship:

$$S_t = S_{0\mathrm{ASAT}} + (S_{\mathrm{o}} - S_{0\mathrm{ASAT}}) \exp\left(-\frac{t}{T_{1\mathrm{ASAT}}}\right) \quad [15]$$

$S_t$ is the signal after a period of irradiation $t$, $S_{0\mathrm{ASAT}}$ is the residual signal after prolonged irradiation, and $T_{1\mathrm{ASAT}}$ the time constant which defines the development of saturation. $T_{1\mathrm{A}}$ the time constant of the free pool of spins (in the absence of any exchange) is given by:

$$\frac{S_{\mathrm{o}}}{S_{0\mathrm{ASAT}}} = \frac{T_{1\mathrm{A}}}{T_{1\mathrm{OBS}}} \quad [16]$$

where $T_{1\mathrm{OBS}}$ is the value of $T_1$ measured in the absence of any RF irradiation. The experiment needed to derive the data is one in which the signals are measured after various durations of irradiation and the effects in tissue can be substantial.

A common factor of both diffusion weighting and MTC can on the one hand be a highly desirable effect but can on the other hand be an artefact of other measurements, affecting, for example, the accuracy of measured time constants if not carefully monitored.

## Conclusion

This has been a very brief review of a large, complex and still quite fast growing area. Contrast is a matter of obsessive importance to radiologists and as a result will continue to be investigated as long as new possibilities for its generation can be found. For anything outside the scope of this article, readers seeking more information are referred to the burgeoning literature. Useful major sources are given in the Further reading section, and there are a great many more books discussing clinical contrast for particular tissue systems and structures.

## List of symbols

$A$ = amplitude of motion; $B_{\mathrm{o}}$ = main electromagnetic field; $D$ = time diffusion coefficient; $D^*$ = apparent diffusion coefficient; $F$ = frequency of motion or Fourier transform function; $G_{\mathrm{s}}$ = gradient in the slice selection direction; $G_{y\mathrm{max}}$ = maximum gradient value; $J_m$ = $m$th-order Bessel function of the first kind; $S$ = observed signal; $S_m$ = signal of the $m$th order artefact; $S_{\mathrm{o}}$ = signal after full relaxation; $S_{0\mathrm{ASAT}}$ = residual $T_{1\mathrm{A}}$ signal after prolonged irradiation; $S_t$ = signal after irradiation period $t$; $T_1$ = longitudinal relaxation time; $T_{21}$ = spin–spin relaxation time; $T_{\mathrm{A}}$ = first inversion period in a double inversion recovery sequence; $T_{1\mathrm{A}}$ = time constant of free pool of spins; $T_{1\mathrm{ASAT}}$ = time constant that defines development of saturation; $T_{1\mathrm{OBS}}$ = $T_1$ in the absence of any RF irradiation; $T_{\mathrm{B}}$ = second inversion period in a double inversion recovery sequence; $y_{\mathrm{f}}$ = field-of-view size in $f$-direction; $y_{\mathrm{v}}$ = voxel length in $y$-direction; $\alpha$ = excitation pulse flip angle; $\beta$ = refocussing pulse flip angle; $\gamma$ = gyromagnetic ratio; $\delta$ = pulse duration; $\Delta G_y$ = gradient movement; $\Delta y$ = spacing in $y$-direction; $\theta$ = spin phase angle; $\theta_{\mathrm{o}}$ = value of $\theta$ at the voxel boundary; $v$ = velocity of flow; $\varphi$ = phase change.

*See also:* **Diffusion Studied Using NMR Spectroscopy; MRI Applications, Biological; MRI Applications, Clinical; MRI Applications, Clinical Flow Studies; MRI Instrumentation; MRI of Oil/Water in Rocks; MRI Theory; MRI Using Stray Fields.**

## Further reading

American College of Radiology (ACR) (1983) *American College of Radiology Glossary of NMR Terms*. Chicago: ACR.

Bryant DJ, Payne JA, Firmin DN and Longmore DB (1984) Measurement of flow with NMR imaging using a gradient pulse and phase difference technique. *Journal of Computer Assisted Tomography* 8: 588–593.

Collins AG, Bryant DJ, Young IR, Gill SS, Thomas DGT and Bydder GM (1988) Analysis of the magnitude of susceptibility effects in diseases of the brain. *Journal of Computer Assisted Tomography* **12**: 775–777.

Foster MA and Hutchinson JMS (eds) (1987) *Practical NMR Imaging*. Oxford: IRL Press.

Grant DM and Harris RK (eds) (1996) *Encyclopedia of NMR*. Chichester, UK: Wiley.

Hajnal JV, Doran M, Hall AS, Collins AG, Oatridge A, Pennock JM, Young IR and Bydder GM (1991) Magnetic resonance imaging of anisotropically restricted diffusion of water in the nervous system: technical, anatomic and pathologic considerations. *Journal of Computer Assisted Tomography* **15**: 1–18.

Le Bihan D (1995) *Diffusion and Perfusion Magnetic Resonance Imaging*. New York: Raven Press.

Mansfield M (1977) Multiplanar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics* **10**: L55–58.

Ogawa S, Lee T-M, Nayak AS and Glynn P (1990) Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine* **14**: 68–78.

Partain CL, Price RR, Patten JA, Kulkarni MV and James AE (eds) (1988) *Magnetic Resonance Imaging*, 2nd edn. Philadelphia: W.B. Saunders.

Pattany PM, Philips JJ, Chiu LC, Lipcamon JD, Duerk JL, McNally JM and Mohapatra SN (1987) Motion artifact suppression technique (MAST) for MR imaging. *Journal of Computer Assisted Tomography* **11**: 369–377.

Stark DD and Bradley W (1992) *Magnetic Resonance Imaging*, 2nd edn. St. Louis MO: Mosby Year Book (two volumes).

# Copper NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# Cosmochemical Applications Using Mass Spectrometry

**JR De Laeter,** Curtin University of Technology, Perth, Australia

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Applications**

Cosmochemistry is the field of science which investigates the composition and evolution of material in the universe, and in particular in the Solar System. Cosmochemistry is intimately related to nuclear astrophysics, because almost all the chemical elements were synthesized by nuclear reactions in the interior of stars. The distribution of the abundances of the elements provides the basis for an understanding of the chemical and nuclear processes which determined their formation and evolution. The discipline of cosmochemistry therefore embraces the universe both in space and time. Cosmochemistry requires a diverse range of observations and measurements of the elements, and in particular their isotopes, in order to understand the processes and events that shaped their past. Variations in isotope abundances are a vast information source, and their retrieval from mass spectra are of fundamental importance to cosmochemistry because of their unique isotopic signatures. Mass spectrometric isotopic abundance measurements are therefore crucial to unravelling our cosmochemical history. The role of mass spectrometry in providing isotopic information in areas such as 'cosmic' abundances, isotopic anomalies, cosmochronology and planetary science, which are essential to our understanding of cosmochemistry, is the subject of this article.

## Cosmic abundances

Cosmic abundance tables of the chemical elements provide the basic data for theories of cosmochemistry and nuclear astrophysics. The abundances are based on a variety of terrestrial, meteoritic, solar, stellar and cosmic ray data. **Figure 1** shows a schematic curve of the elemental abundances (normalized to Si equal to $10^6$ atoms) as a function of mass number $A$. This curve is often referred to as the 'cosmic' abundance curve, but in reality the data is primarily of Solar System origin, and the curve is therefore representative of main sequence stars of similar age and mass as the Sun. However, since the overwhelming majority of stars have essentially solar compositions, the term 'cosmic' is of some relevance. Thus, first we need to determine the distribution of the elements and then proceed to an understanding of the processes which produced the relative abundances of these elements in the Solar System – and by inference in other parts of the universe. So the task of cosmochemistry, and its associated discipline nuclear astrophysics, is one of immense scope and importance.

Since the 1950s, when the elemental abundance distribution depicted in **Figure 1** became available, significant improvements have occurred in the accuracy with which the Solar System abundances are known. In part this is owing to the realization that Type I carbonaceous chondrites – made up of low-temperature condensates that escaped the many alteration processes which affected other classes of meteorites – closely approximate the condensable fraction of primordial Solar System material, and so represent our best opportunity of obtaining an abundance distribution which approximates reality. Isotope dilution mass spectrometry (IDMS) is an effective analytical method for determining trace element concentrations, because of the excellent sensitivity, accuracy and precision of the technique, and because a quantitative recovery of the element concerned is not required. The abundance of the

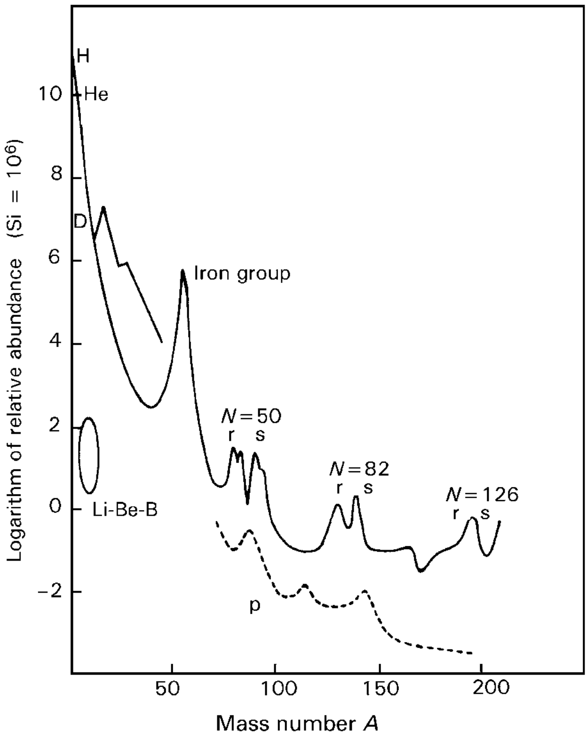



**Figure 1** A schematic curve of atomic aboundance (relative to Si = $10^6$) versus mass number *A* for the Sun and similar main sequence stars. The symbols s, r and p stand for the slow and rapid neutron capture processes and the proton capture process, respectively.

element is determined from the change produced in its isotopic composition by the addition of a known quantity of a tracer of that element whose isotopic composition is different from that of the naturally occurring element. IDMS has been used extensively in refining the Solar System abundance curve because it is ideally suited to analysing small samples – such as the Type C-I carbonaceous chondrites. Meteoritical elemental abundances can now be determined by IDMS at the 3% uncertainty level, although most of the Solar System abundances have an uncertainty of 10% or greater.

The canonical model of nucleosynthesis, which was developed to explain the elemental abundance distribution, is based on the assumption that presently constituted matter is derived from a mixture of numerous components synthesized by a number of processes with a variety of nuclear histories. The chemical elements are synthesized through a continuous process in which a fraction of the material is cycled through successive generations of stars. Apart from hydrogen, most of the helium and some lithium and boron, which were produced during the high temperature, high density stage of the universe which is referred to as the 'big bang', the remaining elements were synthesized in the interior of stars by a combination of thermonuclear reactions (for the 'light' elements with $Z < 20$), nuclear statistical equilibrium processes (for the iron peak group of elements), and the slow (s) and rapid (r) neutron capture processes (for the 'heavy' elements with $Z > 31$).

The products of nucleosynthesis are returned to the interstellar medium by supernovae and by mass loss from red giants and novae. The processed and expelled material, isotopically enriched to varying degrees, is then incorporated in later generations of stars, where further nuclear processing takes place. Thus the Solar System was formed from material that has been processed by a number of nuclear processes in a variety of stellar environments over the lifetime of the galaxy.

## Isotopic anomalies

The conventional wisdom of the late 1960s, based on the apparent isotopic homogeneity of Solar System materials, was that the memory of interstellar chemistry had been erased by a hot gaseous homogenization of pre-existing material from which the Solar System condensed. Technological advances in mass spectrometry, particularly in ion microprobe analysis, have enabled the isotopic composition of extremely small meteoritic samples to be measured with high precision in relation to 'standard' samples, and these analyses have revealed the presence of isotopic anomalies, some of which can be correlated with nucleosynthetic mechanisms occurring in stellar sites. Isotopically anomalous material provides compelling evidence that the Solar System is derived from compositionally different and imperfectly mixed nucleosynthetic reservoirs. Whilst considerable isotopic homogenization undoubtedly took place during the formation of the solar nebula and in planetary formation, some remnants of primordial material have been preserved in refractory inclusions and prestellar grains found in primitive meteorites.

Isotopic abundances are said to be anomalous if the measured isotopic ratios cannot be related to the terrestrial (and by inference solar) isotopic composition of elements through a mass fractionation relationship resulting from a natural physiochemical process, and/or the mass spectrometer measurement itself. The only variations in the isotope abundances in terrestrial materials occur by radioactive decay and by the physiochemical processes mentioned above. Thus terrestrial samples serve as a good example of homogenized Solar System material against which anomalous isotopic material can be compared.

The first isotopic anomaly was discovered in 1960 in the noble gas Xe. The anomalies were of a dual nature – the so-called 'special' anomaly in $^{129}$Xe which was shown to be caused by the presence of the extinct radionuclide $^{129}$I, and the 'general' anomalies which occurred in most of the other isotopes (**Figure 2**). The pattern of enrichment in the heavy isotopes $^{136,134,132,131,130}$Xe, which is particularly pronounced in carbonaceous chondrites, became known as carbonaceous chondrite fission xenon, because it was hypothesized that these anomalies were produced by the spontaneous fission of a super-heavy element. Isotopic anomalies in Xe, and also Ne, made little impression on the cosmochemical community because the systematics of the noble gases were poorly understood and other explanations for the anomalies were forthcoming. It was not until isotopic variations were found in the abundant element oxygen in 1973 that the reality of isotopically anomalous material in meteorites was fully accepted.

## Refractory inclusions

The fall of the carbonaceous meteorite Allende in 1969 was fortuitous in that it contained refractory inclusions – a logical location for the survival of presolar material – in sufficient quantity to enable extensive mass spectrometric investigations to be carried out on a wide range of elements. These inclusions condensed from the cooling gas of solar composition and comprise refractory oxides such as spinel, perovskite, pyroxene and melilite and may contain isotopically anomalous material since they probably condensed before isotopic homogenization was complete. Oxygen consists of three isotopes, and thus in principle it is possible to distinguish between physiochemical and nuclear effects. The former produces fractionation that is linearly proportional to the relative mass difference $\Delta m/m$ and thus the measured $\delta^{17}$O values would be approximately half the $\delta^{18}$O values. Terrestrial samples fall on this fractionation line of slope 0.5 (**Figure 3**), whereas meteoritic materials yield values that define a line of slope unity. This unexpected result was explained by arguing that a component of almost pure $^{16}$O, produced by nucleosynthesis, was incorporated in meteorites. The δ notation is extensively used in mass spectrometric measurements to distinguish isotopic variations in a sample from a terrestrial standard. In the case of oxygen the terrestrial standard is called 'standard mean ocean water' or SMOW. The δ values are customarily given as per mil deviations from the standard.

The elements of the iron peak group provide information on the late evolution of massive stars by nuclear statistical equilibrium processes. Any isotopic anomalies found in these elements would reflect variations in the quasi-equilibrium conditions of stellar interiors, provided that the isotopic

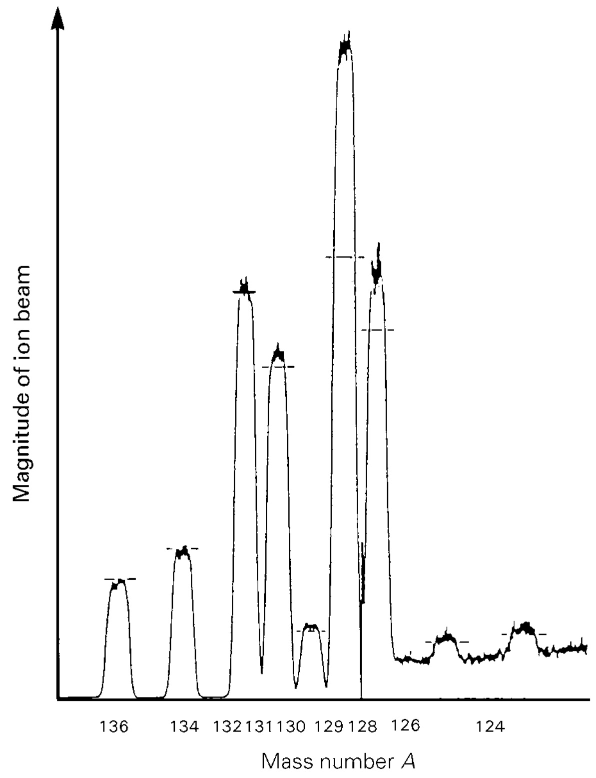



**Figure 2** The mass spectrum of Xe extracted from the Richardton meteorite. The horizontal lines show the comparison spectrum of terrestrial Xe. The isotopes 124, 126 and 128 have been measured at 10 times the sensitivity of the remaining Xe isotopes.

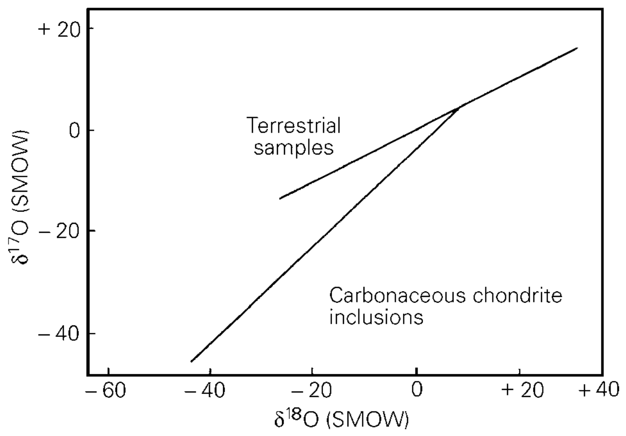



**Figure 3** The isotopic composition of O in inclusions from carbonaceous chondrites which define a line of slope approximately unity in contrast to terrestrial samples which fall on a fractionation line of slope 0.5. SMOW is a terrestrial standard against which the per mil deviation δ of the oxygen isotopes is measured.

abundances were not affected during their injection into the solar nebula. High-precision thermal ionization mass spectrometry (TIMS) has revealed correlated anomalies in the neutron-rich isotopes of Ca, Ti, Cr, Ni and Zn from refractory element-rich inclusions in carbonaceous chondrites. The magnitudes of these anomalies in Ca–Al-rich inclusions from the Allende meteorite are of the order of ≥ 1‰. No single set of nuclear statistical equilibrium conditions can be found to explain the relative excesses in all the elements in these inclusions. However, a multi-zone mixing model has been proposed in which material from a number of stellar zones, with a variety of neutron enrichments, is mixed during the ejection of that material from a star. The magnitudes of the anomalies are affected by dilution with normal Solar System material, and by chemical fractionation processes that occur after nucleosynthesis has terminated.

**Figure 4** shows the isotopic anomalies of the iron peak elements predicted by the multi-zone mixing model as compared with the average excesses as observed in Ca–Al-rich inclusions. The match between the two data sets is impressive, except for Fe and Zn. In the case of Fe no significant anomalies have been measured, but the multi-zone mixing model only predicts a $^{58}$Fe excess of approximately 1 part in $10^4$, which is at the limit of present mass spectrometric capability. In the case of Zn, the excess in $^{66}$Zn is approximately an order of magnitude less than that expected. This can be explained in terms of the volatility of Zn with respect to the other iron peak elements, as it would be the last of these elements to condense from the expelled stellar material. The correlation between anomalies in neutron-rich isotopes in the iron peak elements can therefore be explained in terms of the nuclear statistical equilibrium processes, which took place at a late stage in the evolution of massive stars.

The isotopic anomalies detected in the iron peak elements are extremely small (**Figure 4**), since the anomalous material is diluted with material with a terrestrial isotopic composition. This contamination results from the fact that the meteorite inclusion is taken into solution before chemically extracting the relevant element in a form suitable for conventional TIMS analysis. However, ion microprobe mass spectrometry can be used to analyse small meteoritic inclusions in situ without the need of chemical processing. This enables single inclusions to be analysed for a variety of elements, whilst maintaining the petrographic context of the sample. The carbonaceous chondrites Murchison and Murray also contain refractory inclusions such as corundum and hibonite, but they are invariably small and difficult to analyse by TIMS. Ion microprobe mass spectrometry revealed that Ti anomalies in these hibonite inclusions were an order of magnitude greater than those found in Allende, presumably because hibonite has a higher condensation temperature than the assemblages of the Allende inclusions. Both excesses and deficits were measured in $^{50}$Ti and these anomalies were correlated with anomalies in $^{48}$Ca (**Figure 5**). A number of morphological, chemical and isotopic features show marked correlations in Murchison refractory inclusions, which provides strong evidence that nucleosynthetic anomalies have survived through discrete inclusion-forming events in the early Solar System.

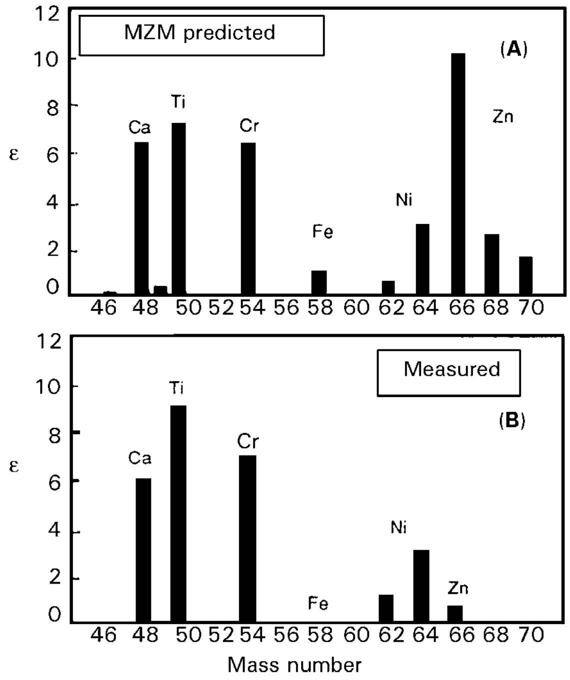



**Figure 4** (A) Isotopic excesses of several iron peak elements as predicted by the multi-zone mixing model. They are given as ε units (parts in $10^4$) relative to normal terrestrial isotopic composition. (B) Average excesses observed in normal Allende inclusions are displayed.

## Presolar grains

The discovery of presolar grains was made possible by the development of chemical procedures in which carbonaceous meteorites were subjected to a stringent acid digestion regime. Carbon compounds such as diamonds, SiC and graphite were isolated in this manner and identified through their distinctive $\delta^{13}$C pattern and anomalous noble gas component. These carbonaceous phases are samples of interstellar matter which provide a window into the prehistory of the Solar System.

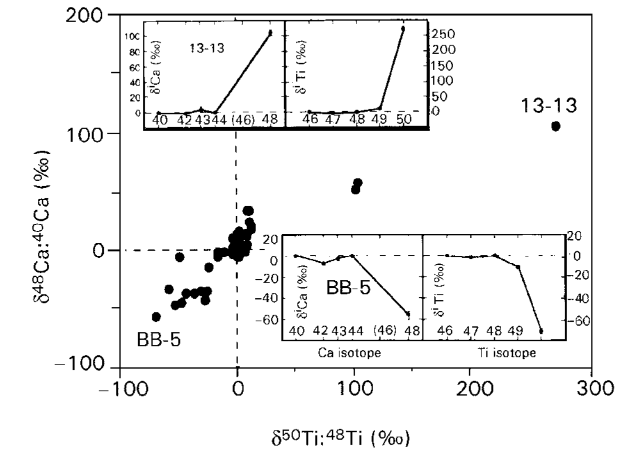



**Figure 5** Ca and Ti isotopic compositions in hibonite ($Ca[AlMgTi]_{12}O_{19}$) inclusions displayed a large range from excesses in the most neutron-rich isotopes to deficits, while the other isotopes are close to solar proportions (defined as deviations, $\delta^i$-Ca or $\delta^i$-Ti, in ‰). The most anomalous grains are both from the Murchison meteorite, 13-13 and BB-5 which have large excesses and large depletions respectively in $^{48}$Ca and $^{50}$Ti (insets). The $^{48}$Ca and $^{50}$Ti anomalies are clearly correlated and are inferred to come from neutron-rich supernova ejecta.

Although diamonds have the highest abundance of prestellar grains (400 ppm by weight of the bulk meteorite), their average size is only 2 nm, which is currently too small for microanalysis. The diamonds exhibit a characteristic enrichment in the heavy and light isotopes of Xe (Xe-HL) as shown in **Figure 6**. The diamonds may have been produced by supernova shock waves passing through molecular clouds or from chemical vapour deposition. Individual grains of SiC may be up to 20 nm in size, which enables ion microprobe analysis of individual grains. These presolar grains contain measurable concentrations of a number of trace elements which has enabled isotopic anomalies in Si, C, N, Mg, Ti, Ba, Nd and Sm to be identified. Most of the heavy elements show evidence of s-process nucleosynthesis (**Figure 6**), whilst $^{12}$C and $^{15}$N enrichments show evidence of a mixture of CNO-cycle material with isotopically normal C and N. A minor fraction of the SiC grains have excesses in the decay products of short-lived radionuclides, indicating that they may represent C-rich zone material from the outer layers of a supernova. The third carbonaceous presolar grain material, graphite, represents less than 2 ppm by weight of these primitive meteorites, and only a fraction is isotopically anomalous. Graphite spherules show a large range in anomalous C, and also contain small inclusions of refractory carbides of Si, Ti, Zr, Mo and W. Some s-process anomalies occur in Zr, suggesting that some of the graphite grains were produced in Asymptotic Giant Branch stars.

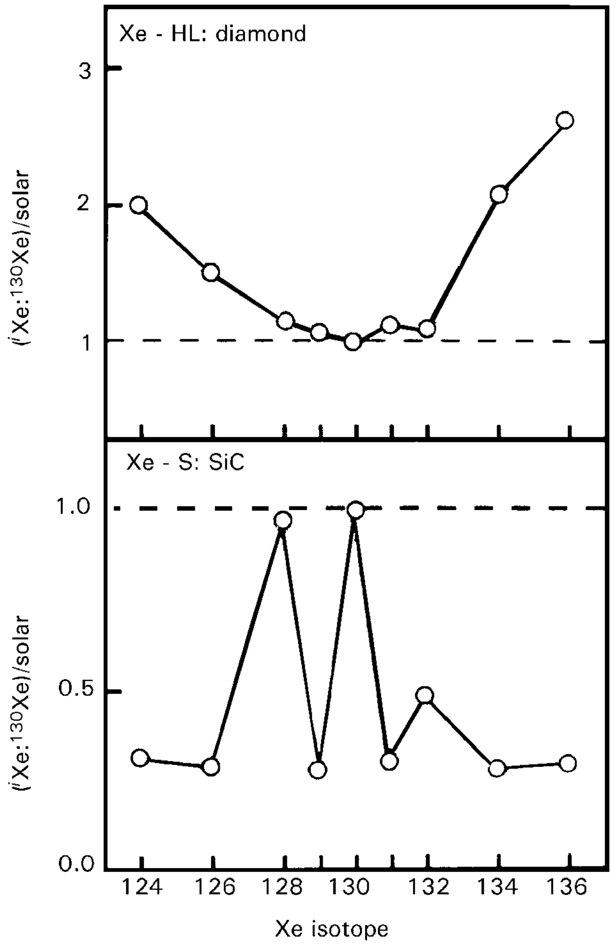



**Figure 6** Progressive enrichment of isotopically exotic Xe components led to the discovery of interstellar diamond and SiC. The light and heavy components of Xe-HL cannot be produced in the same nucleosynthetic event and are probably the result of mixing. The Xe-S component from SiC reflects a mixture between the composition produced in s-process nucleosynthesis and a near-normal component of Xe.

It was fortuitous that the resistance to acid digestion of presolar carbonaceous material led to their discovery. In the digestive procedures adopted, silicates and oxides are destroyed. It is hypothesized that most interstellar grains in the solar nebula are oxides, but apart from corundum ($Al_2O_3$) grains it has not been possible to isolate interstellar oxides. Furthermore the abundance of corundum grains is very small, although they do show isotopic anomalies in $^{26}$Mg, $^{17}$O and s-process Ti.

## Cosmochronology

One of the most challenging tasks in cosmochemistry is to place a time scale on the events that have occurred in the formation and evolution of the Solar

System and for the nucleosynthesis of the chemical elements. **Figure** 7 is a schematic diagram of the major epochs in the history of the cosmos, as described in the figure caption.

The dating of terrestrial materials (geochronology) does not provide any significant information on cosmochronology, but as early as 1953 an estimate for the age of formation of the Solar System was obtained from U–Pb isotopic analyses on meteoritic lead. Subsequently meteorites were dated by the U–Pb technique to give an age of $4.551 \pm 0.004 \times 10^9$ y, and the fine-scale separation of events that took place during planetary formation can be resolved down to a few million years.

Numerous chronological investigations of lunar materials were carried out in the 1970s resulting in the chronology depicted in **Figure** 8. Crystallization ages of ~$4.6 \times 10^9$ y have been obtained from lunar rocks, so that there is a good correspondence between the meteoritic and lunar data for the time of planetary formation. The melt rocks derived from impact metamorphism gave ages of 3.85–$4.05 \times 10^9$ y. This is interpreted as the result of a major bombardment of the Moon by small planetary bodies, which created the lunar basins and implies widespread bombardment of the Solar System during this time. After the termination of this bombardment phase there was decreasing activity until approximately $3 \times 10^9$ y ago, after which the lunar surface appears quiescent.

To obtain accurate dating it is necessary to select a decay scheme whose half-life is of similar magnitude as the age of the event to be measured. Fortunately nature has provided a number of radionuclides with half-lives from $10^5$ to $10^8$ y which have given us the

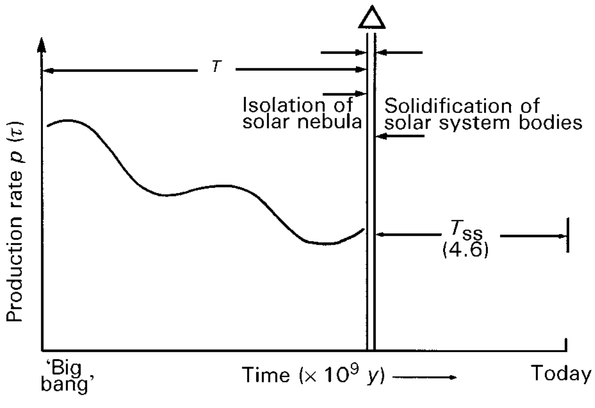



**Figure 7** Schematic diagram of the history of the universe. The first epoch *T* represents the time period from the 'big bang' to the isolation of the solar nebula from galactic nucleosynthesis and the solidification of Solar System bodies, while the third epoch $T_{ss}$ represents the age of the Solar System. The shape of the production rate for nucleosynthesis is for illustrative purposes only.

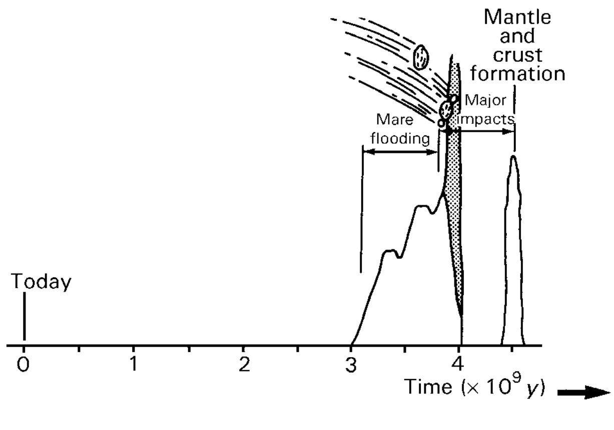



**Figure 8** Schematic diagram showing the chronology of major lunar events.

ability to place tight radiometric constraints on the chronology of nebular and planetary events in the early Solar System, and even on nucleosynthetic events that influenced the presolar molecular cloud before its collapse (**Table 1**). The radionuclides listed in **Table 1** are extinct in that they cannot be directly observed at the present time. Their prior presence must be inferred from an excess in their daughter product linked to the magnitude of the parent:daughter ratio.

The first mass spectrometric evidence of the presence of an extinct radionuclide in meteorites was $^{129}$I, which, with a half-life of $16 \times 10^6$ y, produced

**Table 1** Short-lived radionuclides

| *Parent isotope* | *Daughter isotope* | *$10^{-6}$ Half-life (y)* | *Early solar system abundance* |
|---|---|---|---|
| $^{26}$Al | $^{26}$Mg | 0.74 | $^{26}$Al:$^{27}$Al $\approx 5 \times 10^{-5}$ |
| $^{53}$Mn | $^{53}$Cr | 3.7 | $^{53}$Mn:$^{55}$Mn $\approx 0.1$–$6.7 \times 10^{-5}$ |
| $^{60}$Fe | $^{60}$Ni | 1.5 | $^{60}$Fe:$^{56}$Fe $\approx 4 \times 10^{-9}$ |
| $^{107}$Pd | $^{107}$Ag | 6.5 | $^{107}$Pd:$^{108}$Pd $\approx 2 \times 10^{-5}$ |
| $^{129}$I | $^{129}$Xe | 16 | $^{129}$I:$^{127}$I $\approx 1 \times 10^{-4}$ |
| $^{146}$Sm | $^{142}$Nd | 103 | $^{146}$Sm:$^{144}$Sm $\approx 0.005$–$0.015$ |
| $^{244}$Pu | Xe fission | 81 | $^{244}$Pu:$^{238}$U $\approx 0.004$–$0.007$ |
| $^{41}$Ca | $^{41}$K | 0.10 | $^{41}$Ca:$^{40}$Ca $\approx 1.5 \times 10^{-8}$ |
| $^{182}$Hf | $^{182}$W | 9 | $^{182}$Hf:$^{180}$Hf $\approx 2 \times 10^{-4}$ |
| $^{36}$Cl | $^{36}$Ar | 0.30 | $^{36}$Cl:$^{35}$Cl $\approx 1.4 \times 10^{-6}$ |
| $^{92}$Nb | $^{92}$Zr | 35 | $^{92}$Nb:$^{93}$Nb $\approx 2 \times 10^{-5}$ |
| $^{99}$Tc | $^{99}$Ru | 0.21 | $^{99}$Tc:$^{99}$Ru $\approx 1 \times 10^{-4}$ |
| $^{205}$Pb | $^{205}$Tl | 15 | $^{205}$Pb:$^{204}$Pb $\approx 3 \times 10^{-4}$ |

an excess of $^{129}$Xe (**Figure 2**). In the 1970s a study of refractory inclusions from Allende gave an excess in $^{26}$Mg, the magnitude of which was correlated with the abundance of Al in the samples (as shown in **Figure 9**). A plausible explanation is that $^{26}$Al, with a half-life of $0.74 \times 10^6$ y, was present in these minerals and that the decay was in situ. **Figure 9** is a typical isochron diagram, so useful in geochronology, and the slope of the linear array enables a formation age $\Delta$ of ~$10^6$ y to be calculated. Another important discovery was that of excess $^{107}$Pd in iron meteorites, resulting from the decay of $^{107}$Pd (with a half-life of $6.5 \times 10^6$ y), which gave a value for $\Delta$ of ~$10^7$ y.

**Table 1** lists the extensive isotopic evidence for these extinct radionuclides, which demonstrates that the time between the cessation of nucleosynthesis and the formation of meteorites was so short that it supports the hypothesis that a 'last minute' injection of nucleosynthetic products may have accompanied a supernova shock wave which triggered the collapse of the solar nebula. In this case it is likely that some of the nucleosynthetic products may not have been well mixed in the molecular cloud from which the Solar System evolved.

Unlike the situation with extinct radionuclides, there are few long-lived radionuclides which are appropriate to measure the period of nucleosynthesis in that part of the interstellar medium from which our Solar System formed about $4.6 \times 10^9$ y ago. Various models have been developed to measure the mean age of nucleosynthesis $T$, and $^{232}$Th, $^{235}$U, $^{238}$U and $^{187}$Re have been used to calculate $T$, but the estimation of nuclear production rates is a difficult undertaking for each of these r-process nuclides. If one adds the age of the Solar System $T_{ss}$ to estimates of the mean age of nucleosynthesis $T$, taking into account big bang nucleosynthesis and stellar and galactic evolution, an age of the universe of ~$14 \pm 4 \times 10^9$ y can be obtained.

## Planetary science

Mass spectrometry is ideally suited to the investigation of planetary atmospheres and cometary material in terms of both elemental and isotopic abundances. Mass spectrometers have been an integral component of space probe instrumentation because they are mechanically robust, have a low power consumption, yet are sensitive and versatile.

Dust particle impact mass spectrometers were carried on space probes that approached close to Halley's comet in 1986. These are time-of-flight mass spectrometers of the design shown in **Figure 10**. The data confirm that the dust from the comet is essentially solar in both elemental and isotopic composition. A double focussing mass spectrometer was also used to determine the composition, density and velocity of the neutral gases and of low-energy cometary ions. This mass spectrometer showed that the neutral gas composition of the coma of Halley's comet was dominated by water vapour and other light element constituents. The third mass spectrometer carried on the Halley space probes was designed to determine the composition, density, energy and angular distribution of ions in

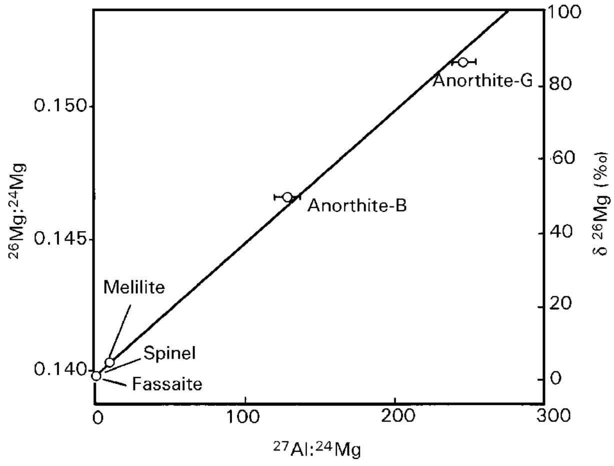



**Figure 9** An 'isochron' diagram of $^{26}$Mg:$^{24}$Mg versus $^{27}$Al:$^{24}$Mg for minerals from an Allende inclusion. The linear dependence of the magnitude of the anomalous $^{26}$Mg on the Al:Mg ratio can be interpreted as resulting from the decay of the radioactive nuclide $^{26}$Al.

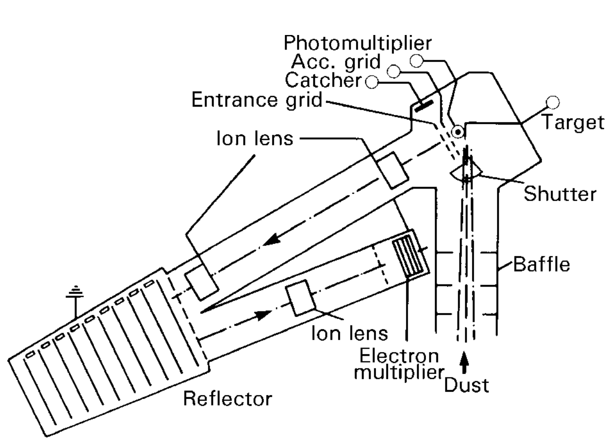



**Figure 10** Schematic diagram of the time-of-flight dust-particle impact mass spectrometer which was mounted in a space probe to examine the mass spectra of dust particles from Halley's comet.

the solar wind and cometary plasma. This ion mass spectrometer consists of two independent sensors – a high-energy range spectrometer and a high-intensity spectrometer. As the space probe Giotto approached comet Halley, $H^+$, $C^+$, $H_2O^+$, $CO^+$ and $S^+$ were found in a diffuse shell-like distribution out to $3 \times 10^5$ km from the comet. As Giotto penetrated the atmosphere to within 1300 km to the nucleus of the comet, the main ion species were $H^+$, $H_2{}^+$, $C^+$, $OH^+$, $H_2O^+$, $H_3O^+$, $CO^+$ and $S^+$. The gas and dust stream emitted from comet Halley contains a higher proportion of volatiles than is found in meteorites, which is consistent with the comet's original location in the cooler regions of the Solar System. Isotopic measurements show that O, S, C, Mg, Si and Fe in cometary and stratospheric dust particles are typical of those found in other Solar System material, implying that the dust has the same nucleosynthetic reservoir as the rest of the Solar System. On the other hand, isotopic anomalies in H and Mg confirm meteoritic evidence that some solid particles survived the homogenization processes that occurred in the early history of the Solar System.

The particle flux that is part of the expanding corona of the Sun is known as the solar wind. The elemental and isotopic composition of the solar wind was studied by various space craft in the period 1965–1975, all of which carried electrostatic positive ion analysers. An ideal opportunity for further study was afforded by the Apollo mission to the Moon, in that aluminium foil collectors were exposed on the lunar surface in the Apollo 11–16 missions which were then analysed for the noble gases by gas source mass spectrometry on the returned Al foil.

The lunar soil is an ideal repository for implanted solar wind elements, as are certain gas-rich meteorites. Deuterium is depleted relative to the terrestrial standard in these materials, the D:H ratio of $<3 \times 10^{-6}$ being consistent with the hypothesis that D is converted into $^3$He in the proto-Sun. Ion probe mass spectrometry has been used to study Mg, P, Ti, Cr and Fe which are present to enhanced levels in lunar minerals, indicating an exposure age of approximately $6 \times 10^4$ y. The isotopic data indicate that the light isotopes of a number of elements have been preferentially lost from lunar material because of volatilization by micrometeorites or solar wind bombardment. There is some indication, from a study of Ne in gas-rich meteorites, of a large solar flare irradiation during the early history of the Solar System, perhaps related to the T-Tauri phase of the Sun.

Isotopic studies of Gd, Sm and Cd from lunar samples show depletions in those isotopes which possess large thermal neutron capture cross sections. This has not only enabled the integrated neutron flux and neutron energy spectrum in the lunar samples to be determined, but has provided information on the stability of the lunar surface by analysing these elements and the noble gases in lunar core material.

Interplanetary dust particles (IDPs) can be collected on adhesive surfaces by U-2 aircraft at altitudes of 20 km, and a fraction of these have been identified as of non-terrestrial origin. Some of these IDPs have been analysed by gas source mass spectrometry using a double focussing instrument employing the Mattauch–Herzog geometry (**Figure 11**). The measured $^3$He:$^4$He and $^{20}$Ne:$^{22}$Ne ratios are lower than those observed in the solar wind. The sensitivity of static gas source mass spectrometry is such that stepwise heating of these IDPs has the potential to elucidate the thermal history of the stratospheric particles, which enables cometary and asteroidal origins to be differentiated.

Mass spectrometers were an integral component in the scientific payload of Viking Landers 1 and 2 and were designed to measure the composition and structure of Mars' upper atmosphere. Carbon dioxide is the major constituent of the atmosphere, whilst the isotopic composition of C and O in the Martian atmosphere is similar to that of the terrestrial atmosphere. However, $^{15}$N is enriched in Mars' atmosphere by approximately a factor of 1.6. Certain meteorites have been identified as samples from the Martian crust. These SNC meteorites, including the orthopyroxenite ALH84001, on which claims for the evidence of life on Mars have been based, have produced a large quantity of isotopic evidence which has been used in conjunction with the Viking Lander data. Since meteorites are delivered to our doorstep,

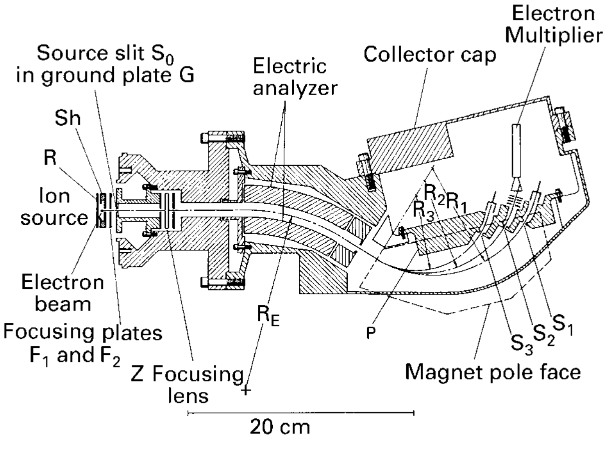



**Figure 11** Schematic diagram of the high-performance, double focussing mass spectrometer used to measure the isotopic composition of the noble gases on IDPs.

free of charge as it were, it is important to acquire and classify new samples because of their potential value to planetary science.

Meteorites have provided the bulk of the data on which the 'cosmic' abundance distribution has been established, which led directly to the canonical theory of nucleosynthesis. A range of nucleosynthetic sites is now available for laboratory study and nucleosynthetic models can be compared with measured isotopic compositions. Isotopic data on refractory inclusions and prestellar grains from primitive meteorites are a new source of astrophysical data, giving fresh insights into nuclear processes in stars. Isotope abundance measurements have also been invaluable in giving a cosmochronological time scale to the major events in the history of the universe. Mass spectrometers have been an essential component of the payload of space probes to the planets and Halley's comet, and technological advances have even enabled the noble gases in IDPs to be measured isotopically.

## List of symbols

H = high; L = low; $m$ = mass; p = proton capture process; r = rapid neutron capture process; s = slow neutron capture process; $T$ = time period from big bang to isolation of the solar nebula; $T_{ss}$ = the age of the Solar System; $\delta$ = deviation from terrestrial composition(‰); $\Delta$ = time period from isolation of solar nebula to solidification of Solar System bodies; $\Delta m$ = difference in mass; $\varepsilon$ = 1 part of $10^4$.

*See also:* **Fluorescent Molecular Probes; Geology and Mineralogy, Applications of Atomic Spectroscopy; Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; Interstellar Molecules, Spectroscopy of; Isotope Ratio Studies Using Mass Spectrometry; Isotopic Labelling in Mass Spectrometry; Labelling Studies in Biochemistry Using NMR; MRI of Oil/Water in Rocks; Time of Flight Mass Spectrometers.**

## Further reading

Anders E and Grevesse N (1989) Abundances of the elements: meteoritic and solar. *Geochimica et Cosmochimica Acta* **53**: 197–214.

Anders E and Zinner E (1993) Interstellar grains in primitive meteorites: diamond, silicon carbide and graphite. *Meteoritics* **28**: 490–514.

Bernatowicz TJ and Zinner E (eds) (1996) *Astrophysical Implications of the Laboratory Study of Presolar Materials*, Proceedings no. 402. New York: American Institute of Physics.

Cowley CR (1995) *An Introduction to Cosmochemistry*. Cambridge: Cambridge University Press.

De Laeter JR (1990) Mass spectrometry in cosmochemistry. *Mass Spectrometry Reviews* **9**: 453–497.

De Laeter JR (1994) Role of isotope mass spectrometry in cosmic abundance studies. *Mass Spectrometry Reviews* **13**: 3–22.

Ireland TR (1995) Ion microprobe mass spectrometry. In Hyman M and Rowe M (eds) *Techniques and Applications in Cosmochemistry, Geochemistry and Geochronology in Advances in Analytical Chemistry*, Volume II, pp 51–118. Greenwich (UK): JAI Press.

Krankowsky D and Eberhardt P (1990). In: Mason JW and Horwood E (eds) *Comet Halley Investigations, Results and Interpretations*, Volume 1, pp. 273–296. New York: Simon and Schuster.

Nier AO (1996) Planetary atmospheres with mass spectrometers carried on high speed probes or satellites. Rarefied gas dynamics. *Progress in Astronautics and Aeronautics* **51**: 1255–1275.

# D

## Dairy Products, Applications of Atomic Spectroscopy

*See* **Food and Dairy Products, Applications of Atomic Spectroscopy.**

## Data Sampling Theory

*See* **Fourier Transformation and Sampling Theory.**

## Diffusion Studied Using NMR Spectroscopy

**Peter Stilbs**, Royal Institute of Technology, Stockholm, Sweden

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| Applications |

### The magnetic field gradient spin-echo method

Under the influence of a non-uniform magnetic field, otherwise equivalent nuclei at different locations give rise to different NMR frequencies. A linear magnetic field gradient along one sample axis will thus quantitatively map NMR frequency with a location along this axis. This is the basic tool used for NMR imaging and microscopy. However, it can also be used for the detection of transport phenomena such as diffusion and flow. The traditional and most widespread NMR method for measuring diffusion is based on the Hahn spin-echo experiment in such a field gradient. Under its influence a phase dispersion results in the rotating frame after application of the initial 90° RF pulse. However, if the NMR frequency remains the same during the whole experiment sequence a second RF pulse refocuses the spin dispersion and leads to the formation of a spin-echo. Normally a 180° pulse is used (applied at time $t$ after the first RF pulse), and the spin-echo will thus appear at a time $2t$ (**Figure 1**). The effect of diffusion on this experiment is easy to work out – the random molecular movement from self-diffusion will partially change the location and NMR frequency of all nuclei in the sample in a random fashion, leading to an incomplete refocusing of the echo. Originally these concepts and experiments were developed and performed in static magnetic field gradients (hence the notation SGSE – static gradient spin-echo). For SGSE experiments the echo attenuation is given by

$$A(2\tau) = A(0)\exp(-2\tau/T_2)\exp(-\gamma^2 g^2 \tau^3 D/3) \qquad [1]$$

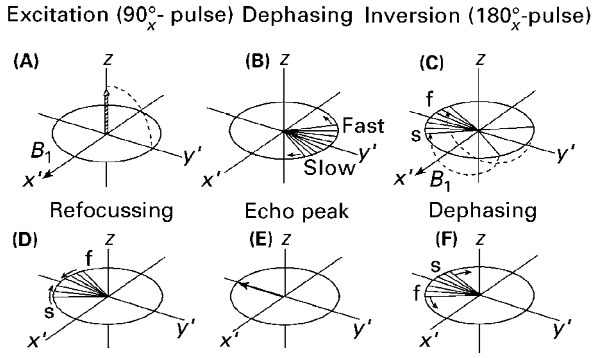



**Figure 1** The basic $90^\circ_x$–$180^\circ_x$ spin-echo experiment. In the case of diffusion measurements the cause of 'fast' and 'slow' spin vectors is that nuclei in different locations precess at different frequencies under the influence of the imposed magnetic field gradient. A useful supplementary picture of the phenomenon is to visualize a formation of a helical magnetization pattern in the sample along the gradient direction during the first gradient pulse, and the unwinding of this magnetization helix during the second gradient pulse. At the time of the echo peak, the net transverse magnetization is sampled and becomes the NMR signal.

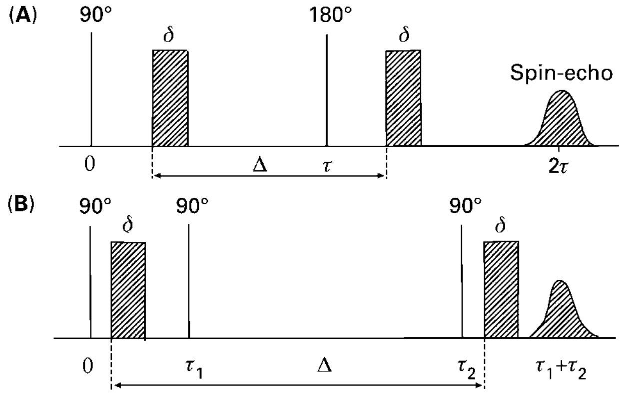



**Figure 2** The pulsed-gradient spin-echo (PGSE) experiment (A), and its stimulated echo analogue (B). Radiofrequency pulse phase cycling is required to isolate the stimulated echo from four other echo peaks that occur as a result of three RF pulses (at times $2\tau_1$, $2\tau_2 - 2\tau_1$, $2\tau_2 - 2\tau_1$ and $2\tau_2$).

where $T_2$ represents the irreversible spin–spin relaxation time, g the strength of the applied gradient, D the self-diffusion coefficient of the molecule in question and $\gamma$ the magnetogyric ratio of the monitored nucleus. It is evident that the echo attenuation effect is greatest for protons, but experiments on other nuclei are feasible.

Numerous developments of these procedures have been made since their initial discovery by Hahn in the early 1950s. Around 1965 the pulsed-gradient spin-echo (PGSE) experiment was suggested and tested. In its basic form it relies on the use of two linear magnetic field gradient pulses (of amplitude 'g', duration '$\delta$' and separation '$\Delta$'; Figure 2A. Under these conditions the amplitude of the spin-echo attenuates from its full value according to the so-called Stejskal–Tanner relation:

$$A(2\tau) = A(0)\exp(-2\tau/T_2) \times \exp[-D(\gamma g\delta)^2(\Delta - \delta/3)] \quad [2]$$

The experiment is always made at a fixed value of $\tau$, so as to keep the $T_2$-attenuation effects constant. The main reasons for using pulsed gradients at the time were primarily to allow a variation of the observation timescale of the experiment [here = (D − $\delta$/3)], to provide more experimental flexibility (in that $\Delta$, $\delta$ and g can readily be varied over a large range of values), and to detect the echo in a relatively homogeneous magnetic field, thereby avoiding the necessity of using the (noisy) broadband spectrometer electronics needed for the detection of very sharp echoes. Fourier transform techniques in NMR that were developed some years later made it possible to make a frequency separation of the components of the echo (FT-PGSE), so as to provide a direct pathway to multicomponent self-diffusion measurements (Figure 3A). Of course, the homogeneous magnetic field during the PGSE detection period is an essential component in this context.

The $T_2$ relaxation is the main limiting factor of the magnetic field gradient experiment, especially when studying slow diffusion. Unfortunately there is an intrinsic correlation between short $T_2$ values and low D values in a given system. For this reason, the PGSE technique is not normally applicable to solids, and it is often also inapplicable to most heteronuclei. The overwhelming majority of studies are made on protons in solution.

In macromolecular systems in solution (where $T_1$ is often much longer than $T_2$) one can often lessen the effects of $T_2$ relaxation by instead applying the three RF-pulse stimulated echo (STE) experiment (normally performed with three 90° pulses, at times 0, $\tau_1$ and $\tau_2$ – the stimulated echo appears at time $\tau_1 + \tau_2$; Figure 2B. Here the echo attenuation is given by

$$A(\tau_1 + \tau_2) = \tfrac{1}{2} A(0)\exp[-(\tau_2 - \tau_1)/T_1] \times \exp(-2\tau_1/T_2)\exp[-D(\gamma g\delta)^2(\Delta - \delta/3)] \quad [3]$$

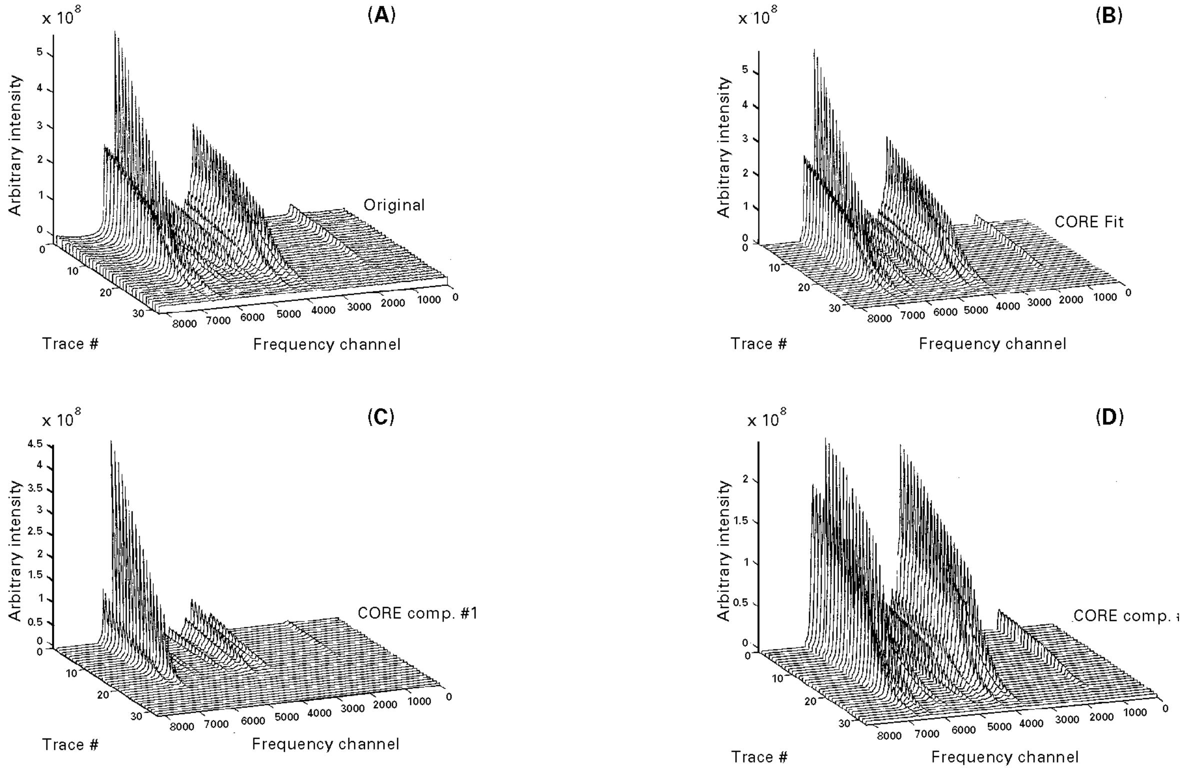



**Figure 3** (A) A typical proton FT-PGSE data set from a complex solution sample (gelatin and two surfactants in $D_2O$, at 300 MHz). The experiment was based on a stimulated-echo PGSE sequence and made at fixed values of $\tau$, $\Delta$ and $\delta$, while varying the gradient current. The residual proton water peak is suppressed already at the lowest gradient setting. Figures (B)–(D) illustrate the application of a CORE analysis of this data set. CORE stands for COmponent REsolved spectroscopy. It is a computational approach that applies a global least-squares analysis to the whole data set, using the prior knowledge that component band shapes remain constant (see text). Band shape regions that are below 2% of the maximum intensity were omitted from the analysis. Two band shape components (gelatin and a common surfactant one) are found, which suggests that both surfactants diffuse at the same rate in this system and indicates the formation of a common aggregate.

One should note that $T_2$ relaxation can be effectively suppressed, since it only occurs between the two first RF pulses (the spacing of which only needs to be slightly longer than the duration of the gradient pulse) and after the third. However, there are further advantages in using the STE experiment, especially in homonuclearly coupled spin systems. By using a short first RF pulse interval one will also suppress effects of J-modulation. This is most often very signal-enhancing since for complex spin systems J-modulation has effectively the same attenuation effect as very efficient $T_2$ relaxation. In addition, short-$\tau_1$ FT-PGSTE spectra resemble normal absorption spectra, since the period during which J-modulation evolves is made small compared with 1/J, where J represents the magnitude of the homonuclear spin coupling.

PGSE methods are the overwhelmingly dominant modes for measuring self-diffusion by NMR. It should be mentioned, however, that an RF field gradient analogue of the magnetic field gradient experiment has also been developed, but its use has not yet become widespread.

Other NMR techniques may be useful in certain situations, but are normally less powerful, or suffer from problems with regard to interpretation. Relaxation by paramagnetic species and intermolecular spin relaxation of protons, for example, are significantly dependent on translational diffusion rates. Separating these effects from other relaxation paths and translating them into self-diffusion data of the components is never straightforward.

Direct imaging of relaxation of the diffusional profile of a concentration gradient by NMR microscopy methods is sometimes a useful alternative approach to PGSE. This is especially true in heterogeneous systems, where basic PGSE techniques may fail owing to rapid $T_2$ relaxation. One should note, however, that only one component can normally be imaged in a single experiment and that the quantity measured is mutual diffusion, rather than self-diffusion. The mutual diffusion coefficient is not directly translatable into a molecular displacement in space – rather it relates to the relative motions of the different components in the system.

## Recent methodological developments

In just a decade or so, the performance level of PGSE instrumentation has increased by roughly two orders of magnitude with regard to most aspects, and the actual hardware has become commercially available almost to the level of a standard spectrometer option. Consequently, PGSE methodological development and experimental application have greatly increased. To a large extent this is coupled to parallel developments with regard to field gradient equipment in NMR microscopy and imaging, and to the development of pulsed field gradient methods for coherence selection and solvent suppression in high-resolution NMR.

Required equipment for high-performance PGSE on modern pulsed NMR spectrometers is a probe with self-shielded field gradient coils (i.e. having an auxiliary coil of opposite polarity – wound in such a way so as to cancel out the pulsed field gradients outside the actual sample volume) and a high-performance current-regulated gradient driver that can produce subsequent gradient pulses that match at the ppm level with regard to area/shape. Present-day PGSE instrumentation is often capable of producing high-resolution FT-PGSE spectra at maximum gradient settings of 100–1000 Gauss $cm^{-1}$ (1–10 T $m^{-1}$). This is typically made at maximum gradient current levels of 10 – 40 A, applied during 1–10 ms.

A general problem with PGSE measurements is that rapid switching of high gradient currents create eddy currents in surrounding metal (probe housing, wires, supercon magnet bore) near the gradient coils. These effects, which last of the order of milliseconds, create disruptive time-dependent magnetic field gradients shortly after gradient switching. They can greatly disturb the experiment, but are dramatically suppressed by the above mentioned use of self-shielded gradient coils. Further suppression can be achieved by selecting a larger-bore supercon magnet type, or probe housing metals other than aluminium (e.g. non-magnetic stainless steel). Notably, eddy current problems are almost completely absent in PGSE measurements made in resistive iron magnet geometries.

Eddy current suppression can also be achieved by using shaped (ramped) gradient pulse shapes (i.e. with much lower rise/fall times than for ‘rectangular’ pulses). One should then note that the functional form of the Equations [2] and [3] assumes the use of rectangular gradient pulses and that the appropriate actual analytical expressions may be considerably more complicated if this is not the case. However, when the separation of the gradient pulses ($\Delta$) is very large compared with their duration ($\delta$) the expressions converge to a situation where only the pulse area matters. It can then merely be substituted as a ‘corrected’ quantity into, for example, Equations [2] and [3]. Two other remedies for the eddy current problem have also been suggested and used: (i) the so-called LED (longitudinal eddy delay) pulse sequence, where a further 90° pulse is appended to the experiment, so as to ‘store away’ the magnetization in the z direction (where it is unaffected by magnetic field gradients) until the magnetic field gradients from eddy currents have decayed. A further 90° pulse then reads back the longitudinal magnetization to create the echo in the transverse plane (ii) the ‘bipolar pulse pair’, where two gradient pulses of equal strength, but of opposite polarity, are sandwiched closely around a 180° degree RF pulse. For large gradient pulse separations this pair (for which eddy current effects cancel out to first approximation) effectively corresponds to a single (monopolar) gradient pulse of twice the length. Both approaches require appropriate RF phase cycling to work. For the LED pulse appendix to the stimulated echo sequence, no less than 32 cycling steps are needed. In the author’s experience neither of these two latter methods are normally even needed nor recommendable if one uses a properly designed shielded gradient coil probe in a wide-bore magnet.

Considerable progress has also recently been made with regard to adequate data evaluation in the FT-PGSE experiment. The key point is that the entire band shape of a given component attenuates exactly the same amount with regard to the pertinent field gradient parameters in Equations [2] and [3], provided the experiment is made under conditions of a constant RF pulse interval setting. Consequently, a global minimization approach using this prior knowledge will result in direct extraction of the respective band shapes of components in the composite FT-PGSE data set [Figure 4A–4D]. The main advantages of this approach is that all the information in the data set is accounted for. Therefore the effective signal-to-noise ratio of the experiment increases an order of magnitude as compared with an evaluation based on peak heights alone. It also allows confident application of the FT-PGSE method to multicomponent systems with extensive NMR signal overlap and closely similar component self-diffusion rates. Within the powerful computing environment of present-day NMR spectrometers this appears to be the natural way to proceed, even for routine work.

It has recently become popular to present results in a 2D manner, with spectral information on the x and z axes (frequency and intensity, respectively) and diffusion information on the y axis. This display mode for FT-PGSE results has been named DOSY (diffusion-ordered spectroscopy). Its introduction

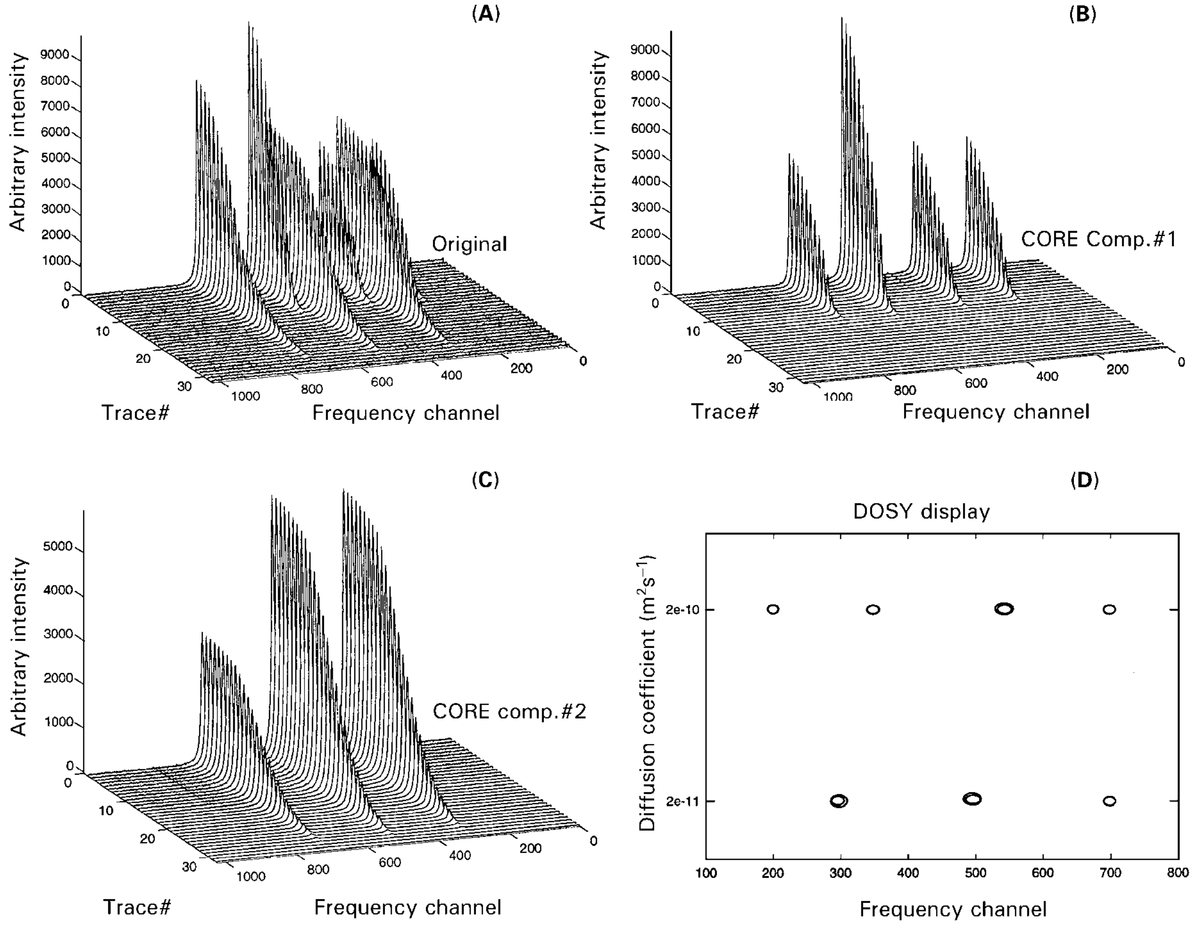



**Figure 4** (A) A synthetic FT-PGSE data set from a sample with two components that diffuse at rates of $2.0\times10^{-10}$ and $2.0\times10^{-11}$ $m^2$ $s^{-1}$ and (B,C) its CORE analysis. (D) The corresponding DOSY display, as based on the CORE analysis. In the diffusion direction, the width of the DOSY peak will reflect either or both of the error limits of the *D* value and the width of the *D* value distribution, as a result of, for example, the polydispersity of the component in question.

has, unfortunately, caused considerable confusion in the literature and the unjustified implication of new types of experiments. One should note that a 'DOSY-display' is most often just based on interpolated individual single-frequency FT-PGSE data (or variants thereof) that are evaluated without taking account of prior knowledge about constant component band shapes with echo attenuation level. Nonetheless, although the DOSY display mode does not add any new information, it is appealing to the eye and has therefore undoubtedly stimulated the use and development of FT-PGSE-based techniques.

Finally one should note that two new variants of the original (static gradient) spin-echo experiment have recently been found to be quite useful, especially for systems where very slow diffusion and very short spin–spin relaxation rates prevail. One is based on the fact that commonly available superconducting magnets produce a dipolar-like stray field that has a very strong, stable and linear field gradient (at a lower than nominal, but still quite strong, magnetic field) a few 10 cm above or below the normal sample position. Thus, by pulling the probe out to the appropriate location one can monitor diffusion through the echo attenuation. In principle, this is possible on most modern supercon spectrometers. The notation STRAFI is often used for this type of stray field setup. The second variant is still quite unusual, and is based on sandwiching two supercon magnets in an 'anti-Helmholz' fashion. A preferred notation for both experimental setups is supercon-fringe field steady-gradient spin-echo (SFF-SGSE). One should note that this technique is not limited to protons–multinuclear measurements (e.g. protons and $^7Li$) are straightforward in the same multinuclear probe. The intrinsic problems with SGSE experiments are (a) under the influence of a strong magnetic field gradient the RF pulses only excite a thin slice of the sample – the echo will therefore be quite weak and (b) actual echo detection is made in an inhomogeneous magnetic field, therefore the signal is very sharp and Fourier transformation to separate its different components

is not possible. In addition, the spectrometer must have broadband detection electronics, data acquisition filters disabled and a fast ADC to properly detect and sample the echo spike.

## Applications of NMR self-diffusion measurements

Self-diffusion data is intrinsically a very direct source of information. For unrestricted diffusion during a selected time span, the characteristic self-diffusion coefficient (D) simply translates into a mean square displacement in space ($\langle \Delta r^2 \rangle$) during the observation time ($\Delta t$) through the Einstein relation: $\langle \Delta r^2 \rangle = 6D\Delta t$. One should note that under normal measurement conditions of PGSE techniques the mean square displacement of even large macromolecules during the characteristic time span of the experiment is generally very much larger than the average macromolecular diameter. Therefore the quantity ($\langle \Delta r^2 \rangle$) normally requires no further modelling and D values thus become very easy to interpret.

### Diffusion in simple liquids

The NMR FT-PGSE approach is the only method that can provide quantitative information on multicomponent self-diffusion. Such data are central as a reference for, for example, theories of simple liquids.

### Binding studies

In binding studies the self-diffusion approach typically relies on a relative comparison of time-averaged self-diffusion rates between a 'bound' and 'free' state, meaning that in the (simplest two-state situation the effective self-diffusion coefficient D(obs) will be given by D(obs) = p D(bound) + (1 – p) D(free) where the degree of binding (p) may assume values in the range $0 < p < 1$. This equation can be rewritten into the form p = [D(free) – D(obs)]/[D(free) – D(bound)]. Hence, such experiments do provide the information needed to obtain binding isotherm information, for example, by a simple comparison of experimentally available self-diffusion coefficients. It is evident that the method is most suited for binding/aggregation of small molecules to larger entities and is most accurate at intermediate p values. Numerous such studies have been presented.

### Non-Gaussian diffusion in macromolecular systems

One should note that the Einstein relation $\langle \Delta r^2 \rangle = 6D\Delta t$ only holds in the limit of large $\Delta t$. 'Large', in this context, has a different meaning for small molecules and macromolecules. Uhlenbeck and Ornstein derived a generalized relation, in their classical treatment of the subject back in the 1930s:

$$\langle \Delta r^2 \rangle = 6D\{\Delta t - \tau_c(1 - \exp[1 - \Delta t/\tau_c])\}$$

where the correlation time $\tau_c = m/f$, f being the coefficient of friction and m the mass of the Brownian particle. For small molecules in liquids the correlation time is of the order of a nanosecond, while the shortest imaginable NMR spin-echo experiments will exceed a fraction of a millisecond, but are more typically tens or hundreds of milliseconds. For macromolecules, however, this correlation time may be such that it affects (and thus becomes measurable by) the typical NMR PGSE experiment, when using appropriately selected measurement parameters (strength, duration and separation of gradient pulses). Only a few studies along these lines have appeared to date.

### Non-Gaussian diffusion in heterogeneous systems

In porous and heterogeneous systems, penetrant diffusion is not necessarily isotropic or Gaussian on a given timescale. PGSE techniques have been used since the early 1960s for characterizing such aspects of system morphology through the effects of non-Gaussian diffusion on the echo attenuation in the PGSE experiment. Restricted diffusion manifests itself in a variation of the apparent diffusion coefficient with the observation time (i.e. $\Delta - \delta/3$ in the PGSE experiment). A large number of system geometries and experimental conditions have been considered and analysed. Several complex pulse sequences have also been suggested for eliminating the effect of local magnetic field gradients that arise from susceptibility variations in the sample.

Another important recent development of experimental PGSE techniques applied to heterogeneous systems has been the discovery and utilization of so-called diffraction effects with regard to the attenuation of the echo, i.e. weak echo level maxima and minima that occur at high attenuation levels. In a system of spherical pores, for example, the system geometry manifests itself in a first echo maximum when the inverse of the pore diameter matches the gradient setting $\gamma g\delta/2\pi$. Mainly methodological work has been presented so far. Notably, the technique has also found use for the investigation of cell suspensions such as, for example, erythrocytes.

## List of symbols

D = self-diffusion coefficient; g = applied gradient strength, amplitude; J = magnitude of homonuclear spin coupling; $T_1$ = longitudinal relaxation time; $T_2$ = irreversible spin–spin relaxation time, transverse relaxation time; $\delta$ = duration of magnetic field gradient pulse; $\Delta$ = separation of magnetic field gradient pulses; $\gamma$ = magnetogyric ratio; $\tau$ = time delay between RF pulses.

*See also:* **Magnetic Field Gradients in High Resolution NMR; MRI Applications, Biological; MRI Applications, Clinical; MRI Theory; MRI Using Stray Fields; NMR Microscopy; NMR Pulse Sequences.**

## Further reading

Callaghan PT (1991) Principles of Nuclear Magnetic Resonance Microscopy. Oxford: Clarendon Press.

Callaghan PT and Coy A (1994) PGSE and molecular translational motion in porous media. In: Tycko P (ed) NMR Probes of Molecular Dynamics, pp 489–523. Dordrecht: Kluwer Academic.

Canet D (1997) Radiofrequency field gradient experiments. Progress in Nuclear Magnetic Resonance Spectroscopy 30: 101–135.

Hahn EL (1950) Spin echoes. Physics Review 80: 580–594.

Johnson CS (Jr) (1994) Transport ordered 2D-NMR spectroscopy. In: Tycko R (ed) NMR Probes of Molecular Dynamics, pp 455–488. Dordrecht: Kluwer Academic.

Kärger J, Pfeifer H and Heink W (1988) Principles and application of self-diffusion measurements by nuclear magnetic. In: Waugh JS (ed) Advances in Magnetic Resonance, pp 2–89. San Diego: Academic Press.

Kimmich R (1997) NMR Tomography Diffusometry Relaxometry. Berlin, Heidelberg, New York: Springer.

Lindblom G and Orädd G (1994) NMR studies of translational diffusion in lyotropic liquid crystals and lipid membranes. Progress in Nuclear Magnetic Resonance Spectroscopy 26: 483–515.

Norwood NJ (1994) Magnetic field gradients in NMR – friend or foe. Chemical Society Reviews 23: 59–66.

Price WS (1996) Gradient NMR. In: Annual Reports on NMR Spectroscopy, pp 51–142. New York: Academic Press.

Price WS (1997) Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion. Part 1. Basic theory. Concepts in Magnetic Resonance 9: 299–336, and Part 2 Experimental aspects. Concepts in Magnetic Resonance 10: 197–237.

Söderman O and Stilbs P (1994) NMR studies of complex surfactant systems. Progress in Nuclear Magnetic Resonance Spectroscopy 26: 445–482.

Stilbs P (1987) Fourier transform pulsed-gradient spin-echo studies of molecular diffusion. Progress in Nuclear Magnetic Resonance Spectroscopy 19: 1–45.

Waldeck AR, Kuchel PW, Lennon AJ and Chapman BE (1997) NMR diffusion measurements to characterize membrane transport and solute binding. Progress in Nuclear Magnetic Resonance Spectroscopy 30: 39–68.

# Drug Metabolism Studied Using NMR Spectroscopy

**Myriam Malet-Martino** and **Robert Martino**,
Université Paul Sabatier, Toulouse, France

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

The purpose of this article is to briefly review the scope and potential of NMR as a non-invasive technique for studying drug metabolism. NMR has been used for the study of drug metabolism in vitro (biofluids, tissue extracts, biopsies, intact cells, perfused organs) and in vivo (animal models and human subjects). NMR spectroscopy is unique in its ability to permit both these kinds of studies. We shall see that the advantages and limitations of NMR as a tool for drug metabolism and disposition studies are different depending on the approach chosen (in vitro or in vivo studies). The first part of this article deals with the various isotopes used in such studies. The second part is a general discussion of the value and difficulties encountered with NMR in investigations on drug metabolism. In the third part, we have chosen some specific examples to emphasize the interest of NMR.

## Isotopes used in the study of the metabolism of drugs by NMR

There are several NMR active nuclei (Table 1) that can be routinely used in drug metabolism and disposition. The merits and disadvantages of some of these nuclei are discussed below.

### Fluorine-19 ($^{19}F$)

The majority of studies in the literature concern $^{19}F$ NMR mainly because of the favourable NMR characteristics of this nucleus, including a nuclear spin of $\frac{1}{2}$, relatively narrow lines, 100% natural abundance, high sensitivity (83% that of proton), large chemical shift range and short longitudinal relaxation times ($T_1$) which permit rapid pulsing with a corresponding improvement in the signal-to-noise ratio per unit time. Moreover, the negligible level of endogenous fluorine eliminates interfering background signals and dynamic range problems. As a number of fluorinated drugs are currently in clinical use, $^{19}F$ NMR offers a powerful method of monitoring their pharmacokinetics and metabolism either in vitro or in vivo. The main requirement is that drugs and metabolites should be present in minimal concentrations, near $10^{-3}$ mM, for studies in biofluids or $5 \times 10^{-2}$ mM for in vivo studies.

### Proton ($^1H$)

The $^1H$ nucleus is present in all drugs, and has the highest NMR sensitivity of any stable nucleus. Moreover, it is the most abundant (99.98% natural abundance) of the two natural isotopes ($^1H$ and $^2H$ or deuterium) of the hydrogen atom. However, the small chemical shift range and the extensive multiplicity due to homonuclear J-coupling sometimes make the observation and quantification of drug metabolites difficult. Moreover, $^1H$ NMR spectra of biofluids or tissues contain many intense resonances from water, proteins and lipids that must be reduced or eliminated. A simple method for water suppression (which nevertheless implies a treatment of the biological sample) consists in freeze-drying the sample and redissolving it in $D_2O$. Several NMR techniques are used to achieve the suppression of water and macromolecule resonances. The form of water suppression most commonly used is presaturation of the $H_2O$ resonance using a secondary irradiation field at the solvent resonance frequency. A second approach takes advantage of the short $T_2$ (transverse relaxation time) of tissue water, proteins and some fats, and consists in applying a spin-echo sequence that will suppress the broad resonances of water and macromolecules. Presaturation and spin-echo sequences can also be combined. Despite these drawbacks, several studies have shown that this nucleus can be exploited for the detection and measurement of drugs and metabolites in biofluids, i.e. for in vitro studies. The main requirements are that metabolites should have suitable protons and be present at concentrations 0.01–0.1 mM. As $^1H$ is a poor nucleus for the in vivo monitoring of drugs and metabolites, in vivo $^1H$ NMR spectroscopy has so far had little application. The reasons for this are manifold. Because of factors such as tissue heterogeneity, restricted molecular mobility, magnetic field inhomogeneity over the relatively large sample volume and the use of lower magnetic field strength spectrometers, in vivo signals linewidths are substantially broader than those obtained in vitro, in biofluids for example. There is thus excessive signal overlap due to the very large number of signals that occupy a small chemical shift range and the signals from endogenous metabolites may obscure the detection and quantification of the drug and/or its metabolites. Moreover, the replacement of $H_2O$ by $D_2O$ becomes impractical for studies carried out on whole animals or humans since high concentrations of $D_2O$ are toxic and $D_2O$ is very expensive. More elaborate NMR techniques than those reported above are thus required for solvent suppression and spectral editing.

### Carbon-13 ($^{13}C$)

One advantage of $^{13}C$ NMR is that its large chemical shift range enables reliable structural identification. However, although the carbon nucleus is found in all drugs, the magnetically active isotope ($^{13}C$) is neither abundant nor particularly sensitive. These drawbacks combined with the low concentrations of drugs and their metabolites in biological systems make this nucleus in its natural abundance rather hard to detect. Improved sensitivity can be attained with polarization transfer experiments which partly transfer the favourable properties of the proton to $^{13}C$. Furthermore, $^{13}C$-enriched drugs can be synthesized to increase the NMR sensitivity. However, to obtain a full picture of the chemical changes taking place, several key positions usually need to be labelled. This may be chemically difficult, presupposes an a priori knowledge of the metabolism of the particular drug and is rather expensive in practice. It should be stressed that the label still needs to be present at millimolar concentrations, rather than the trace quantities that are typically required for radiolabelling studies. Because of these difficulties, the use of $^{13}C$ for in vitro or in vivo NMR studies of drug metabolism and disposition has been very limited.

**Table 1** NMR isotopes for drug metabolism studies

| *Isotope* | *Spin* | *Natural abundance (%)* | *NMR receptivity*[a] | *Chemical shift range* (ppm) | *Minimum concentration* (mM) *for in vitro studies*[b] | *Minimum concentration* (μmol g$^{-1}$) *for in vivo studies* |
|---|---|---|---|---|---|---|
| $^{1}H$ | 1/2 | 99.98 | 100 | 15 | 0.01 | 0.1 |
| $^{2}H$ | 1 | 0.016 | 0.00015 | 15 | 1.2 | 1.5 |
| $^{7}Li$ | 3/2 | 92.58 | 27 | 12 | | 0.1 |
| $^{10}B$ | 3 | 19.58 | 0.389 | 200 | | 2[c] |
| $^{11}B$ | 3/2 | 80.42 | 13.3 | 200 | | 5–10 |
| $^{13}C$ | 1/2 | 1.11 | 0.018 | 250 | 0.1–1[d] | 50[d] |
| $^{14}N$ | 1 | 99.63 | 0.1 | 350 | 1[e] | |
| $^{15}N$ | 1/2 | 0.37 | 0.00038 | 350 | 1[d] | |
| $^{17}O$ | 5/2 | 0.037 | 0.0011 | 2500 | | |
| $^{19}F$ | 1/2 | 100 | 83.3 | 500[f] | 0.001–0.005 | 0.05–5 |
| $^{31}P$ | 1/2 | 100 | 6.6 | 800 (30)[g] | 0.01 | 0.2–1 |
| $^{195}Pt$ | 1/2 | 33.8 | 0.34 | 10 000 | 10 | |

[a] Receptivity ($D$) relative to proton (H) is a relative sensitivity figure used to compare signal areas theoretically obtainable at a given magnetic field strength: $D_X = [\gamma_X{}^3 N_X I_X (I_X+1) / [\gamma_H{}^3 N_H I_H (I_H+1)] \times 100$ where $\gamma$ is the magnetogyric ratio, $N$ the natural abundance and $I$ the spin quantum number.
[b] In biofluids.
[c] With an enriched drug and if an indirect method of observing protons coupled to the $^{10}B$ nuclei is employed.
[d] With an enriched drug.
[e] Only nitrogen centres with relatively symmetrical electron distributions can be studied.
[f] For organofluorine compounds.
[g] Chemical shift range of endogenous phosphorus compounds.

### Phosphorous-31 ($^{31}P$)

The $^{31}P$ nucleus is readily detected as it has a 100% natural abundance and is relatively sensitive. The presence of endogenous phosphates and derivatives (phosphomonoesters, phosphodiesters, etc.) may interfere with signals from phosphorated drugs and their metabolites. In practice, this is not a large obstacle as there are relatively few endogenous compounds that produce detectable signals. In the studies carried out to date, no interference in NMR signals between phosphorus-containing drugs or metabolites with endogenous compounds, apart from inorganic phosphate, have been described. As phosphorus-containing drugs are fairly rare, only a few $^{31}P$ NMR drug studies have so far been carried out.

### Lithium-7 ($^{7}Li$)

The major lithium isotope, $^{7}Li$, has a high NMR sensitivity. Since $^{7}Li$ is not a spin $I = \frac{1}{2}$ but a spin $I = \frac{3}{2}$ nucleus, it possesses a quadrupole moment, which gives rise to broad spectral lines. $^{7}Li$ NMR is not hampered by interference with endogenous signals. The only studies to date involve the in vivo determination of the pharmacokinetic profile of lithium salts.

### Boron-10 and boron-11 ($^{10}B$, $^{11}B$)

NMR is a potentially valuable technique for evaluation of boron neutron capture therapy (BCNT) agents since both boron isotopes are magnetically active. The $^{11}B$ isotope has better sensitivity (16.5 versus 2% relative to $^{1}H$) and higher natural abundance (80.4 versus 19.6%) than $^{10}B$. Even though $^{10}B$ is the active nucleus for neutron capture in BCNT, $^{11}B$ is more appropriate for NMR studies except if an indirect method of observing protons coupled to the $^{10}B$ nuclei is employed.

### Other nuclei

Deuterium ($^{2}H$), tritium ($^{3}H$), $^{17}O$, $^{15}N$, $^{14}N$ and $^{195}Pt$ suffer from several problems which cannot be overcome easily, including low sensitivity ($^{17}O$, $^{15}N$, $^{14}N$, $^{2}H$, $^{195}Pt$), poor resolution and broad signals ($^{17}O$, $^{14}N$, $^{2}H$), and the radioactivity ($^{3}H$) associated with the relatively high concentrations required for NMR studies. The sensitivity of detection can be improved by isotopic enrichment combined with enhancement by polarization transfer from proton. Nevertheless, all these problems are reflected in the limited use of these nuclei in NMR studies both in vitro or in vivo.

## Advantages and limitations of NMR for drug metabolism studies

### *In vitro* studies

Many analytical techniques, especially chromatographic methods, are in routine use for the analysis

of biofluids. The reasons for the additional use of NMR spectroscopy are manifold and lie in the unique properties of NMR spectroscopy (Table 2).

NMR is non-destructive and enables the direct study of any intact biofluid without resort to treatment. Problems of extraction recovery and chemical derivatization, and those that may be encountered with pH-sensitive metabolites, are consequently avoided. NMR spectroscopy is thus particularly suited to the study of delicate samples.

The method is non-specific and unexpected substances are not overlooked during the investigation, since all metabolites (provided they bear the nucleus under observation and they are present at sufficient concentrations) are detected simultaneously in a single analysis. This contrasts with chromatography which usually requires some prior knowledge of the structures of metabolites to optimize sample preparation and/or detection. Novel metabolites may therefore be missed. NMR also avoids the use of a number of different chromatographic techniques, which is sometimes necessary when metabolites have different chemical structures. This is an important attribute of NMR in the search for novel metabolites when, often, the analyst will have no idea of the type of molecule to look for.

The NMR spectrum contains much structural information, and the observed chemical shifts and spin–spin coupling patterns can provide valuable clues about the structure of novel metabolites. Even though metabolites must still be extracted and purified for unequivocal elucidation of the structure, the NMR 'behaviour' nevertheless can give a good estimate of the structures of unknown compounds.

Although optimization of the quantification of drugs and metabolites in biofluids by NMR is somewhat tedious [choice of a suitable standard, determination of the $T_1$, checking that excitation is uniform over the whole frequency range when a large spectral width is observed (often the case in $^{19}F$ and $^{13}C$ NMR), taking into account the heteronuclear nuclear Overhauser effect when proton decoupling is applied, validation of the assay, etc.] it is nevertheless fairly easy to obtain quantitative data, and NMR can be used routinely in the same way as HPLC. Moreover, given a good signal-to-noise ratio, NMR can quantify substances accurately and reproducibly. Quantification in plasma and bile may nevertheless be affected by the binding of drugs or metabolites to macromolecules. This results in substantial signal broadening and gives a reduced signal-to-noise ratio or even leads to NMR-invisibility. Some sample pretreatment may be required in such cases.

Set against these advantages, however, are a number of disadvantages which need to be taken into account.

NMR spectrometers are expensive instruments, with the cost rising steeply with the field strength of the magnet used; higher field strengths provide greater spectral dispersion and yield better sensitivity. However, the magnets have long useful lifetimes (~ 15 years) and the electronic components may be upgraded continuously at reasonable cost.

Although there is little or no spectral interference from endogenous molecules using $^{19}F$, $^{31}P$, $^{7}Li$ and $^{13}C$ (since $^{13}C$-enriched drugs are required) NMR, the problem of peak overlap is critical in $^{1}H$ NMR since the signals of a large number of endogenous compounds are present in a relatively narrow chemical shift range. Moreover, suppression of the intense signal from water is a prerequisite for $^{1}H$ NMR spectroscopy of biofluids. Resonances of the drug and metabolites must occur in spectral regions which are free of signals from endogenous metabolites. With a spectrometer operating at a field $\geq$ 9.4 T, and taking the necessary precautions, drug metabolite resonances may be measured on the entire spectrum, except between 3.5 and 4.1 ppm as this range contains many overlapped resonances from endogenous molecules. In most biofluids, the region of the

**Table 2** Advantages and drawbacks of *in vitro* NMR for drug metabolism studies

| *Advantages* | *Drawbacks* |
|---|---|
| No destruction of the sample | NMR spectrometers are expensive instruments |
| No sample pretreatment necessary | Critical problem of peak overlap in $^{1}H$ NMR. Resonances of the drug and metabolites must occur in spectral regions which are free of signals from endogeneous metabolites |
| No degradation of unstable metabolites due to sample treatment | |
| All metabolites are detected simultaneously in a single analysis | |
| Does not require preselection of conditions based on a knowledge of the compound to be analysed | The volume of biofluid needed for NMR analysis may be large |
| An analytical technique does not have to be designed for each newapplication | NMR is relatively insensitive; this limits its use to the observation of drugs that are present at relatively high concentrations |
| NMR gives both structural and quantitative information | |
| Little or no spectral interference from endogenous molecules in $^{19}F$, $^{31}P$, $^{7}Li$ and $^{13}C$ NMR | |

spectrum deshielded from the water signal is relatively low in interfering resonances, so that aromatic or heterocyclic drugs can often be studied with relative ease. The positive aspect of this limitation is that information can be obtained on alterations in composition of the biofluid after drug administration since the $^{1}H$ NMR spectrum of a biofluid gives a characteristic pattern of resonances for a range of important endogenous metabolites.

The volume of biofluid needed for NMR analysis ranges between 0.3 to 0.4 or 1.5 to 2.5 mL with 5 or 10 mm tubes respectively. This can be a hindrance for pharmacokinetic studies which require numerous plasma samples or for difficult-to-obtain biofluids such as bile, cerebrospinal fluid (CSF) or those from neonates. Recent progress in the use of flow-injection NMR probes will alleviate this disadvantage.

Compared with most chromatographic and other spectroscopic techniques, NMR is relatively insensitive, which represents the principal drawback of the technique and limits its use to the observation of those drugs that are present at relatively high concentrations. It is difficult to give a precise figure for the minimum concentration requirement because there are many factors that influence the signal-to-noise ratio. The theoretical detection limit for $^{1}H$ NMR is around $10^{-3}$ mM for 600 MHz spectrometers using a 5 mm probe, but in practice it is often difficult to accurately quantify substances found at concentrations below 0.1 mM. With one currently available spectrometer (7 T, corresponding to a proton resonance frequency of 300 MHz) the detection threshold in a 10 mm diameter tube is $5 \times 10^{-3}$ mM and $10^{-2}$ mM for 10 to 15 h of $^{19}F$ NMR or $^{31}P$ NMR recording. With a 11.7 T spectrometer (corresponding to a proton resonance frequency of 500 MHz) a detection threshold of 1 to $2 \times 10^{-3}$ mM has been achieved in 12 h of $^{19}F$ NMR data acquisition in 10 mm tubes. However, the accuracy and reproducibility of the NMR concentration determinations are generally $\leq 10\%$ for concentrations $\geq 5 \times 10^{-2}$ mM for $^{19}F$ and $^{31}P$ NMR. For lower concentrations, the accuracy and reproducibility decrease and depend on the measurement time and the resulting signal-to-noise ratio in the spectra.

However, in spite of these difficulties, NMR spectroscopy has now become a technique of demonstrated utility and of unique ability in the analysis of biofluids. Moreover, the recent development of using separation methods in conjunction with NMR spectroscopy has given rise to a radical new approach in biofluid metabolite analysis. The coupling of HPLC to NMR in 1992 in the form of a commercially available instrument has been shown to be of great use in separating and determining the structure of drug metabolites in biofluids. The sensitivity of conventional HPLC-NMR provides a detection limit of about 150 ng of analyte for an overnight data accumulation. This technique has been extended to other chromatographic techniques such as supercritical fluid chromatography and to the use of nuclei other than $^{1}H$ or $^{19}F$ which form the basis of most studies so far because of their high NMR sensitivity. Another major recent development has been the combination of both NMR and mass spectrometry to an HPLC system, which will be of great help to elucidate the structures of drug metabolites in biofluids.

The advantages and limitations of NMR for biofluids also apply to the analysis of excised tissue samples or intact cells packed in NMR tubes. For quantitative studies, the metabolites of interest should not be degraded by enzymes even at 4 °C. For the analysis of cells, one must make sure that the density of the cellular suspension does not change with time, i.e. that there is no sedimentation in the NMR tube. This can be checked by acquisition of successive data blocks. The spectra are then examined individually, and if there is little or no change in metabolite levels, spectra can be added together to improve the signal-to-noise ratio. In such samples, there is often considerable signal broadening owing to inhomogeneities in the medium analysed and signals from compounds with similar chemical shifts may merge into a single peak. The $^{19}F$ NMR detection threshold at 470 MHz was found to be about 5 nmol $g^{-1}$ of tissue or cellular suspension for 1–2 h measurement on 0.5–1 g samples. Recently, it has been shown that it is possible to obtain high-resolution $^{1}H$ NMR spectra of biopsy tissues simply by spinning at the magic angle ($\theta = 54.7°$) at modest speeds (around 3 kHz). This could open up new routes of monitoring drug metabolism.

## *In vivo* studies (Table 3)

The main interest of in vivo NMR is obvious. Measuring drug levels in plasma does not always reflect the drug concentration at the receptor sites, which are generally located in the tissue cells of the target organ. Consequently, a method that allows one to measure the concentration of drugs and their metabolites in situ may be extremely useful.

In vivo NMR spectroscopy can be used repetitively on the same patient since it is noninvasive. The ability to monitor the pharmacokinetic of a drug and to evaluate the potential modifications in drug metabolism could prove extremely valuable in patient management.

However, there are three main obstacles to the use of NMR in vivo. The first and major limitation is

**Table 3** Advantages and drawbacks of *in vivo* NMR for drug metabolism studies

| *Advantages* | *Drawbacks* |
|---|---|
| Measurement of drug and /or metabolites levels in target organ | Low sensitivity. Tissue concentration necessary for *in vivo* detection of a drug is 1–10 times that necessary *in vitro* |
| Can be used repetitively on the same patient | Only molecules with unrestricted mobility are visible |
| | Prolonged data acquisition may be required |
| | The observed signals may have various anatomical origins |
| | Poor resolution. Chemically similar compounds may not be readily distinguished |
| | Difficult in making accurate quantitative measurements |

the lack of sensitivity. The minimum detectable concentration (Table 1) depends on many factors, including magnetic field strength, intrinsic sensitivity, presence of background signal, data collection time and sample size. Tissue concentration necessary for in vivo detection of a drug is typically 1–10 times that necessary in vitro. Order-of-magnitude detection limits for endogeneous metabolites are 0.2 μmol $g^{-1}$ for $^{31}P$ and 0.1 μmol $g^{-1}$ for $^{1}H$. The in vivo $^{19}F$ NMR detection limit of 5-fluorouracil (FU) and its metabolites has been estimated at around 0.05 μmol $g^{-1}$ with a 7 T vertical bore magnet for a 1 h measurement and 0.1 μmol $g^{-1}$ with a 1.5 T whole-body spectrometer for total measurement time that is tolerable (30 min) for patients. It should not be forgotten that these detection limits are for molecules with unrestricted mobility, which is not always the case since drug and/or metabolites are often strongly bound to macromolecules or cell membranes.

The insensitivity of NMR has two consequences. First, prolonged data acquisition may be required to obtain satisfactory spectra. This is because the confident detection of resonances requires a good signal-to-noise ratio, which is proportional to the square root of the number of scans performed. The kinetics of drug metabolism or elimination must therefore not be too rapid. Second, the surface coils used for the detection of drugs and metabolites in situ are relatively large, particularly in humans. The observed signals may thus have various anatomical origins (e.g tumour tissue and normal tissue). To get a specific pharmacokinetic profile from one tissue, spatially localized spectroscopy using field gradients to define the volume of interest is useful. However, the relatively low concentrations of metabolites encountered have so far limited the use of this method for drug metabolism studies.

The second limitation is the lack of resolution. NMR signals observed in vivo are broad, and chemically similar compounds may not be readily distinguished.

The third limitation of in vivo NMR is the difficulty in making accurate quantitative measurements. The only quantitative studies of drug metabolism carried out so far have concerned fluorine- or lithium-containing compounds with $^{19}F$ or $^{7}Li$ NMR. For quantitative determinations, signal areas are normalized relative to that of an external reference – a capillary or a sphere filled with a standard of known concentration and volume. Within a particular experiment this method can estimate the relative concentrations of metabolites with an accuracy of about 10%, rising to around 40% for the estimation of absolute concentrations. The absolute concentration in human brain of fluorinated neuroleptics or lithium salts has been determined by performing an additional separate calibration experiment (immediately before and after each experiment) with a phantom of known concentration and volume placed in the position of the tissue. However, since the tissue volume is not readily measured and since the loading on the coil and hence the nutation angle are assumed to be the same for the phantom as for the tissue, the method tends to be rather inaccurate. Another method consists in using a calibrated internal standard such as tissue water.

## Some recent applications of NMR in drug metabolism studies

Most of the $^{19}F$ NMR studies are related to anaesthetic, psychoactive and antineoplastic drugs. The distribution of anaesthetics in the brain and the pharmacokinetics of their elimination are still a subject of controversy. The feasibility of in vivo $^{19}F$ NMR studies of halothane in humans has been recently demonstrated. Resonances attributable to this compound were observed up to 90 min after withdrawal of the anaesthetic agent in eight patients.

Fluoxetine (Prozac®) is a widely used antidepressant drug. In patients receiving daily doses of 20 or 40 mg, brain concentration continued to increase well after the clinical effects were evident (2 weeks), seemed to level off after about 6 to 10 months of treatment and was about 20 times the concentration in plasma. It thus appears that there is no correlation between brain concentration and clinical response. A combined in vivo and in vitro study on several patients showed that the single $^{19}F$ resonance observed in vivo contains, in roughly equal proportion, both the drug and its active metabolite

(norfluoxetine) (Figure 1). The comparison of in vivo and in vitro quantifications (1–11 μg of fluoxetine + norfluoxetine per mL of tissue in vivo versus 12–19 μg $mL^{-1}$ of tissue in extracts) suggested that a substantial fraction of the drug and its metabolite is not visible in vivo.

Most of the $^{19}F$ NMR studies on fluorinated drugs have concerned FU. This drug is an ideal case for $^{19}F$ NMR studies for three reasons. High doses are administered to patients. The fluorine atom remains intact during biotransformation. Its metabolism, and more especially its catabolic pathway, leads to compounds that are structurally different from the parent compound. The $^{19}F$ NMR signals are therefore displayed over a large spectral width with no peak overlap (except for fluoronucleotides in vivo). The number of in vitro and in vivo $^{19}F$ NMR studies published reflects the contribution of the technique to the knowledge of FU metabolism. The following two examples emphasize the interests in $^{19}F$ NMR.

FU injected at high doses by continuous intravenous perfusion during 4 to 5 days provoked a cardiac toxicity. For commercial purposes, FU is solubilized in a solution of sodium hydroxide at two different pH (8.6 or 9.2). The $^{19}F$ NMR analysis of these solutions revealed the presence of about a hundred signals of fluorinated 'impurities', among which fluoroacetaldehyde and fluoromalonic acid semialdehyde were identified (Figure 2). These products come from the degradation of FU and are formed with time in the basic medium which is indispensable to the solubilization of this antitumour agent. These two compounds are highly cardiotoxic on the isolated perfused rabbit heart model since fluoroacetaldehyde is metabolized into fluoroacetate which is a violent cardiotoxic and neurotoxic poison. The intensity of the cardiotoxicity of FU commercial solutions, tested on the isolated perfused rabbit heart model, is a function of the pH of the solution, and the preparations at pH 9.2 are less cardiotoxic than the preparations at pH 8.6. Nevertheless, the non-negligible frequency of cardiotoxic accidents observed after injection of the commercial solutions at pH 9.2 was too significant to be explained by the levels of fluoroacetaldehyde and fluoromalonic acid semi-aldehyde found in these preparations. The possibility of an eventual metabolism of FU itself into fluoroacetate has thus been explored. It has long been postulated, without ever being demonstrated, that the main catabolite of FU, $\alpha$-fluoro-$\beta$-alanine, could be transformed into fluoroacetate. Using in vitro $^{19}F$ NMR and the isolated perfused rat liver model, it was found that the last catabolite of FU is not $\alpha$-fluoro-$\beta$-alanine. The metabolism continues and leads to two new catabolites, fluoroacetate and fluoromalonic acid semi-alcohol (Figure 3). The conclusion of this study was twofold. To limit the cardiotoxicity of FU, it would be better to use in clinical treatments a new dosing form of FU which is devoid of degradation products (lyophilized material for example). However, to suppress this toxicity it would be necessary to act on the degradation pathway of FU by blocking it before the formation of $\alpha$-fluoro-$\beta$-alanine.

In vivo $^{19}F$ NMR has been used to monitor FU metabolism in the liver metastases of colorectal cancer patients. The patients were treated with a continuous low dose intravenous infusion of FU until the point of refractory disease, at which time interferon-$\alpha$ was added with the objective of modulating FU activity. Figure 4 shows the spectra obtained in a 1.5 T whole body spectrometer with a 16 cm diameter surface coil positioned over hepatic metastases of a patient treated with 300 mg of FU $m^{-2}$ $day^{-1}$. The first NMR scan (Figure 4A) was performed after 1 week of FU infusion. A large signal of catabolite ($\alpha$-fluoro-$\beta$-alanine) and a small resonance from FU were observed. The patient went on to have an objective partial response but later developed progressive disease. At this time (Figure 4B), the FU peak had disappeared and the catabolite signal had reduced. The next scan (Figure 4C), after a week on interferon-$\alpha$, showed a new anabolite peak. A week later, the last scan (Figure 4D) showed FU and catabolite signals. At this stage the patient had a symptom and tumour marker response and stabilization of tumour size.

The $^{1}H$ NMR spectrum of a human urine sample from a volunteer who had ingested the anti-inflammatory drug flurbiprofen is shown in Figure 5A. The complexity of the spectrum precluded any detailed structural or quantitative analysis. For example, the ability to discern resonances from the aromatic protons of flurbiprofen metabolites is severely limited. However, a stopped-flow HPLC-NMR experiment at 30.5 min afforded the $^{1}H$ spectrum shown in Figure 5B which corresponds to the two diasteromeric forms of the $\beta$-D-glucuronic acid conjugate of 4'-hydroxyflurbiprofen. The resonances of the para substituted phenyl ring appear at 6.91 and 7.42 ppm and the other three aromatic resonances as a complex multiplet at 7.19 ppm. The $\beta$-D-glucuronic acid H1 proton appears as two resonances at 5.49 and 5.52 ppm. The remaining $\beta$-D-glucuronic acid resonances are located between 3.4 and 3.6 ppm. The methyl resonances of the propionic acid sidechain of flurbiprofen are two doublets at 1.49 and 1.51 ppm. The two methine signals are observed at about 3.94 ppm as overlapped multiplets. This example emphasizes the interest in HPLC coupled to NMR spectroscopy for the identification of drug metabolites.

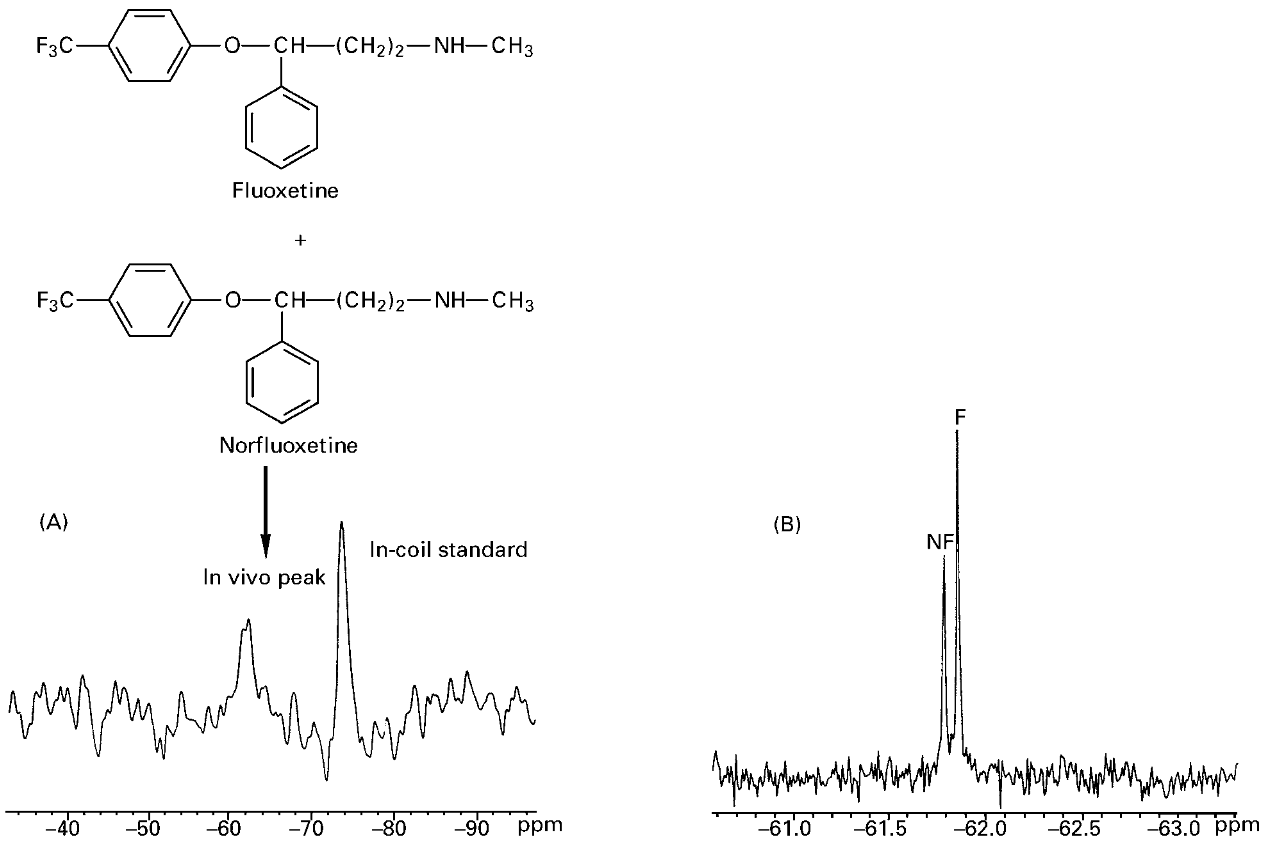



**Figure 1** The $^{19}F$ NMR spectra (A) from the head of a patient on 20 mg per day of fluoxetine for 1 month, (B) of the extract of a small section of temporal cortex from a deceased patient on 40 mg per day of fluoxetine for 19 weeks. F = fluoxetine, NF = norfluoxetine. Adapted with permission of Williams and Wilkins from Komoroski *et al* (1994) *In vivo* $^{19}F$ spin relaxation and localized spectroscopy of fluoxetine in human brain. *Magnetic Resonance in Medicine* **31**: 204–211.

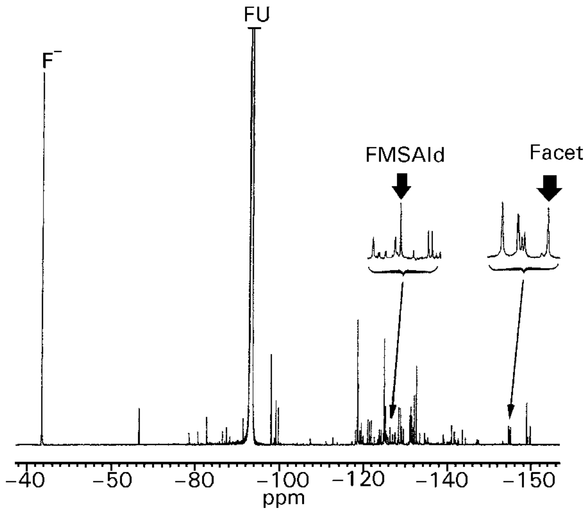



**Figure 2** The $^{19}F$ NMR spectrum of a commercial solution of 5-fluorouracil (pH 9.2). $F^-$, fluoride ion; FU, 5-fluorouracil; FMASAld, fluoromalonic acid semi-aldehyde; Facet, fluoroacetaldehyde.

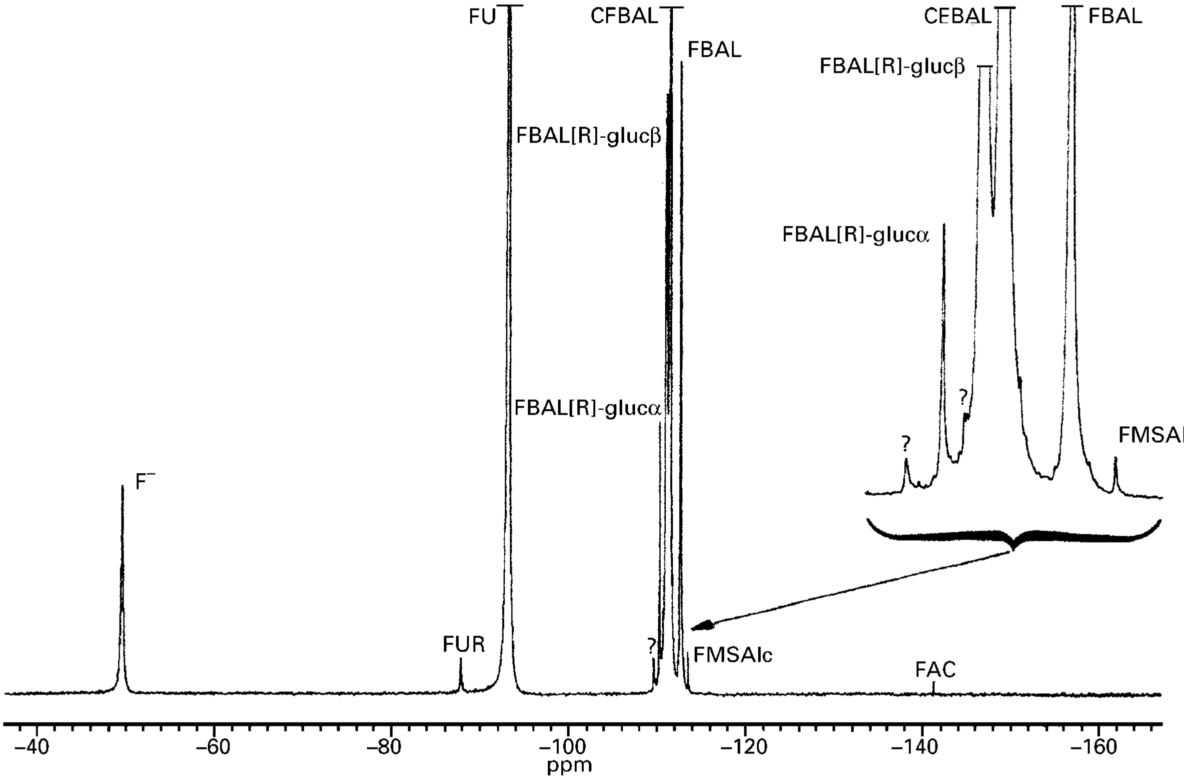



**Figure 3** The $^{19}F$ NMR spectrum of a perfusate (concentrated 20-fold by lyophilization) from an isolated perfused rat liver treated with 5-fluorouracil at a dose of 45 mg per kg b.w. for 3 h. $F^-$, fluoride ion; FUR, 5-fluorouridine; FU, 5-fluorouracil; FBAL, $\alpha$-fluoro-$\beta$-alanine; CFBAL, N-carboxy-$\alpha$-fluoro-$\beta$-alanine; FBAL[R]-gluc$\alpha$, FBAL[R]-gluc$\beta$, adducts of $\alpha$-fluoro-$\beta$-alanine with $\alpha$-glucose or $\beta$-glucose; FMASAlc, fluoromalonic acid semi-alcohol; FAC, fluoroacetate.

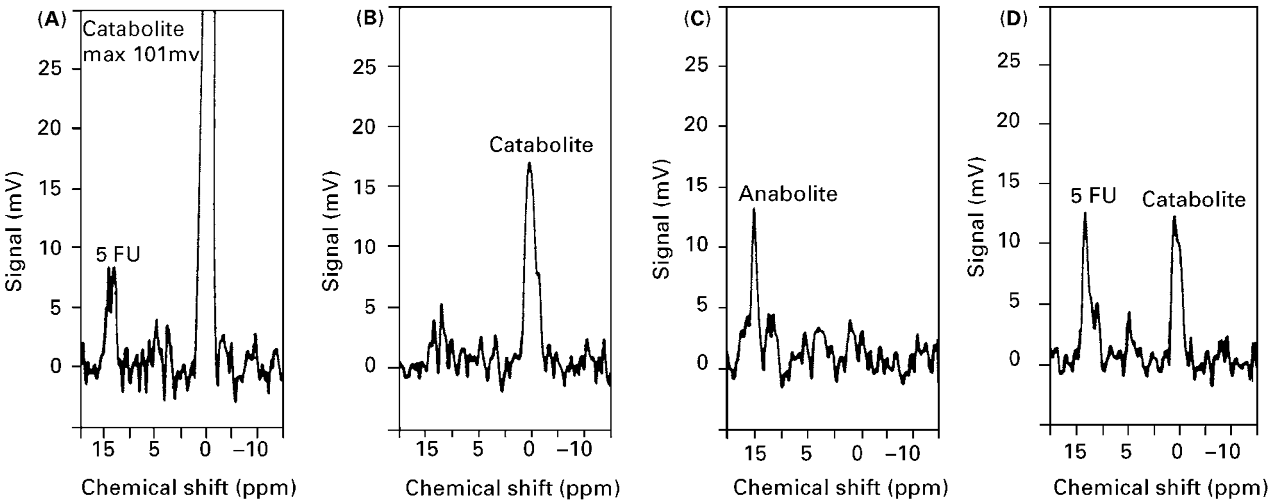



**Figure 4** A series of $^{19}F$ NMR spectra of liver metastases from a colorectal cancer patient treated with 5-fluorouracil at a dose of 300 mg $m^{-2}$ $day^{-1}$ (A, B) and 5-fluorouracil + interferon-$\alpha$ (C, D). 5FU, 5-fluorouracil. mV: millivolts (arbitrary values). Reproduced with permission of Kluwer Academic from Findlay *et al* (1993) The non-invasive monitoring of low dose, infusional 5-fluorouracil and its modulation by interferon $\alpha$ using *in vivo* $^{19}F$ MRS in patients with colorectal cancer. *Annals of Oncology,* **4**: 597–602.

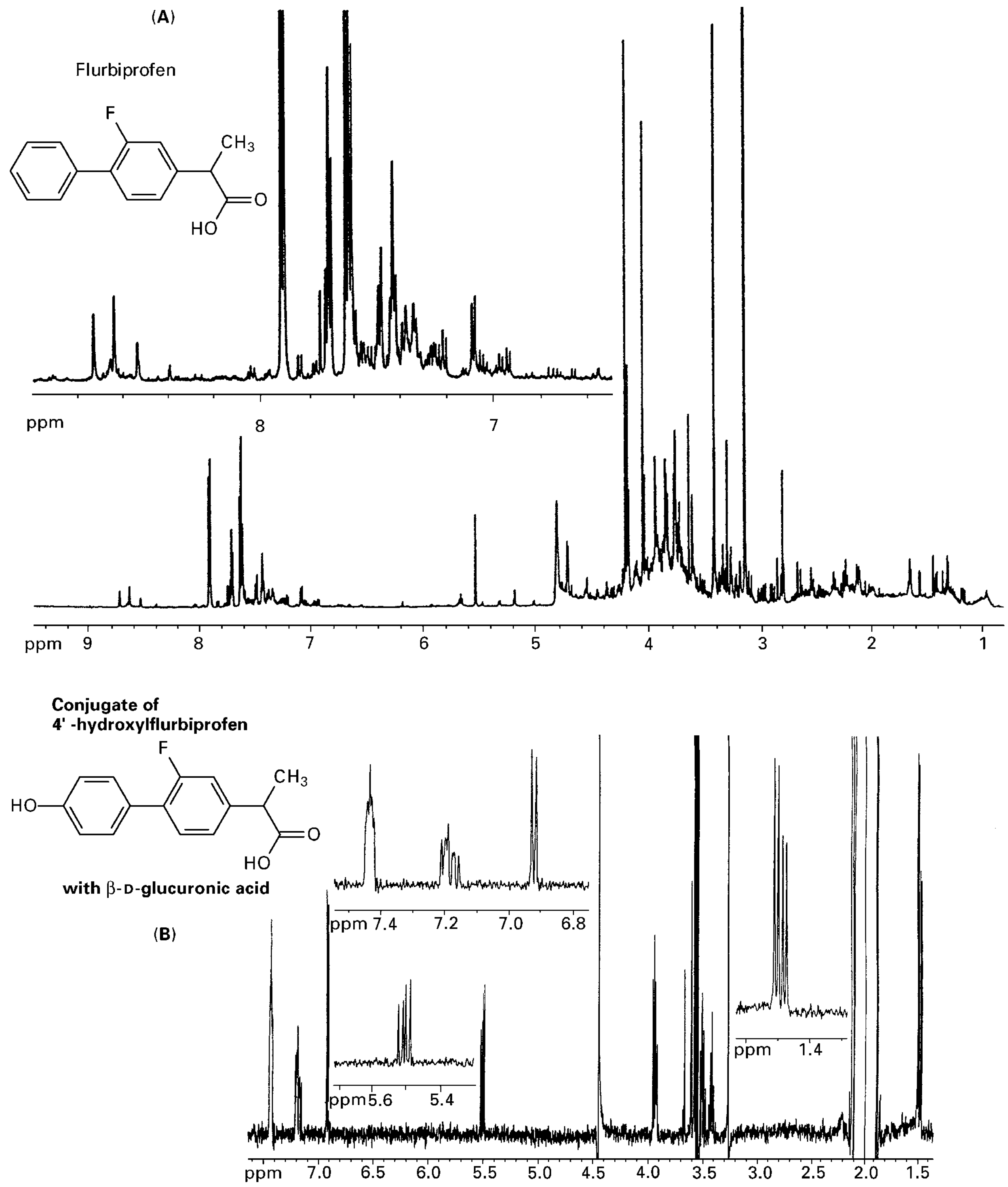



**Figure 5** The 600 MHz $^{1}H$ NMR spectra of (A) a sample of human urine collected after the oral ingestion of 200 mg flurbiprofen, (B) the 30.5 min retention time species corresponding to the $\beta$-D-glucuronic acid conjugate of 4′-hydroxyflurbiprofen. Reprinted from (A) Spraul *et al* (1993). Coupling of HPLC with $^{19}F$ and $^{1}H$ NMR spectroscopy to investigate the human urinary extraction of flurbiprofen metabolites. *Journal of Pharmaceutical and Biomedical Analysis*, **11** (10), 1009–1015; (B) Lindon *et al* (1996) Direct coupling of chromatographic separations to NMR spectroscopy. *Progress in NMR Spectroscopy*, **29**, 1–49, with kind permission from Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

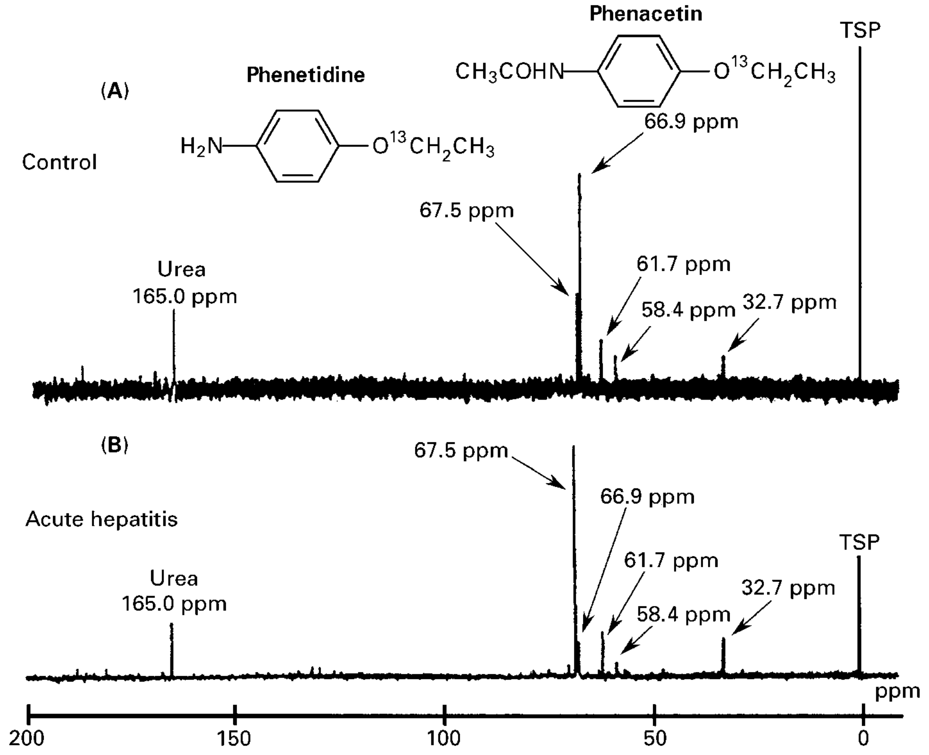



**Figure 6** The $^{13}C$ NMR spectra of human urine taken 1 h after oral administration of ([1-$^{13}C$]ethoxy)phenacetin to a healthy subject (A) and to a patient with acute hepatitis (B). TSP, 3-(trimethylsilyl)propionate (internal standard). Adapted with permission of The Pharmaceutical Society of Japan from Kajiwara *et al* (1996) Studies on $^{13}C$ phenacetin metabolism II. A combination of breath test and urine test of *in vivo* metabolites in the diagnosis of liver disease. *Chemical Pharmaceutical Bulletin* **44**: 1258–1260.

$^{13}C$-labelled phenacetin has been used for the diagnosis of liver disease. In the urine of healthy subjects, the $-OCH_2-$ signal of phenacetin was higher than that of its metabolite, phenetidine, whereas in patients with acute hepatitis the situation was the reverse (Figure 6). A return to the normal profile was observed during convalescence. The anticancer agent temozolomide labelled with $^{13}C$ was noninvasively detected in mice tumours by a selective cross-polarization transfer $^{13}C$ NMR method which has the advantage of higher sensitivity with lower power deposition. The resonance of temozolomide was observed but not the signal of its active metabolite, 5-(3-methyltriazen-1-yl)imidazole-4-carboxamide (Figure 7).

$^{31}P$ NMR spectroscopy has been used to analyse biofluids from patients treated with the antineoplastic drug ifosfamide, widely used in the treatment of soft-tissue sarcomas and paediatric malignancies. Its metabolism is complex (Figure 8). The spectrum in Figure 9 emphasizes the wealth of information on drug metabolism that can be gained from in vitro studies. About 20 phosphorus-containing compounds were detected. In addition to the already known metabolites (unchanged ifosfamide, carboxyifosfamide, 2-dechloroethylifosfamide, 3-dechloroethylifosfamide, isophosphoramide mustard,

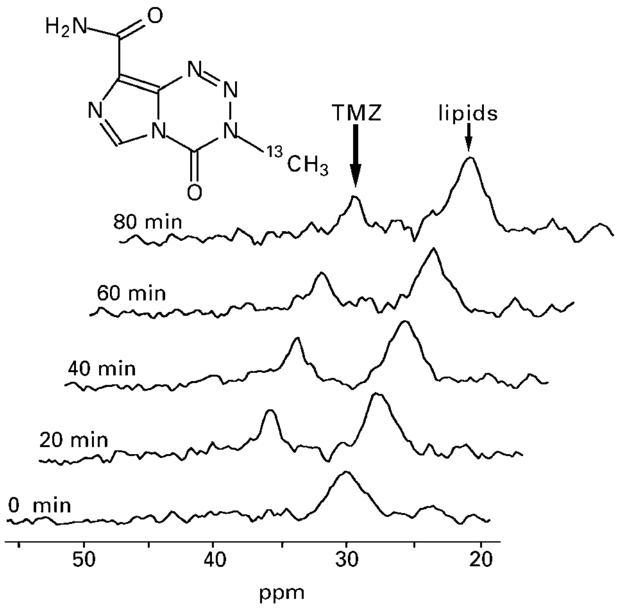



**Figure 7** *In vivo* $^{13}C$ NMR spectra of a RIF-1 tumour before and after an intraperitoneal injection of $^{13}C$-temozolomide (TMZ). Adapted with permission of Williams and Wilkins from Artemov *et al* (1995) Pharmacokinetics of the $^{13}C$-labelled anticancer agent temozolomide detected *in vivo* by selective cross polarization transfer. *Magnetic Resonance in Medicine,* **34**: 338–342.

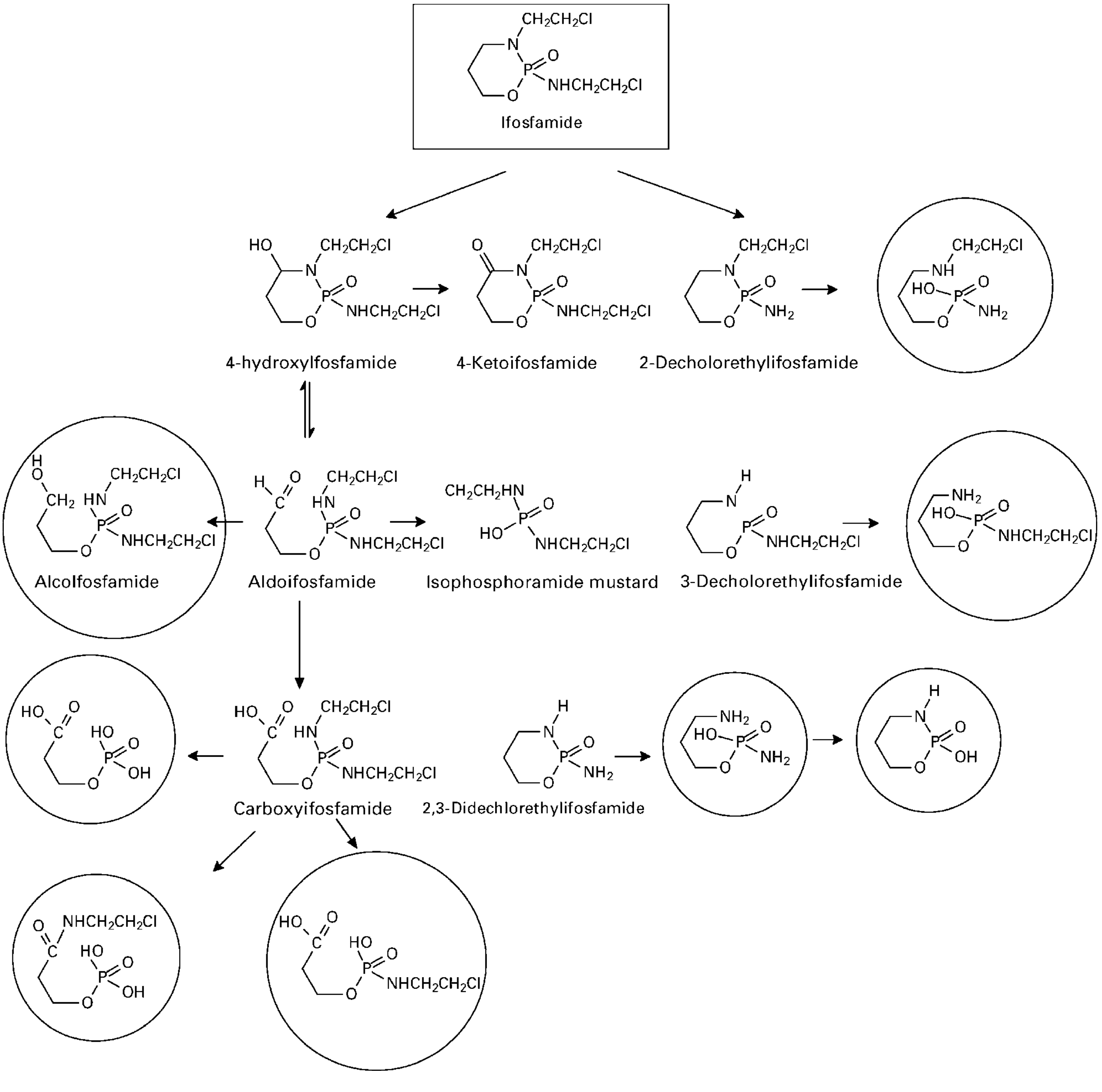



**Figure 8** Metabolic pathways of ifosfamide. Compounds ringed are phosphorated metabolites that have been identified using $^{31}P$ NMR.

ketoifosfamide) eight other compounds were identified. Furthermore $^{31}P$ NMR spectroscopy has been used to detect and quantify ifosfamide in vivo in rat tumours. In addition to the resonances of the phosphorus-containing endogenous metabolites, only one broad signal could be detected for ifosfamide and, probably, also some of its metabolites. Carbogen breathing increased the amount of ifosfamide taken up by the tumour by 50% (Figure 10).

The concentration of Li in serum is not an adequate measure of Li efficacy. Localized spectroscopy that permits the measurement of Li in the brain may be a better approach of efficacy and/or neurotoxicity. Lithium was detected noninvasively in vivo in the brain and in the calf muscle of normal volunteers and of psychiatric patients. The lithium levels in both muscle and the brain rose more slowly than in serum. The concentration of Li in the brain is typically 0.1–0.6 mM. Therapeutic serum levels (ranging between 0.4–1.0 mM) are typically reached after 1 or 2 days of lithium treatment whereas clinical efficacy is not observed for several more days. The $^{7}Li$ NMR data strongly suggested that this delay may be due to the relatively slow accumulation of lithium in the brain (after 3–7 days of treatment). After the lithium treatment was stopped, the elimination half-time was slower in the brain (32–48 h) than in the serum or muscle (~24 h).

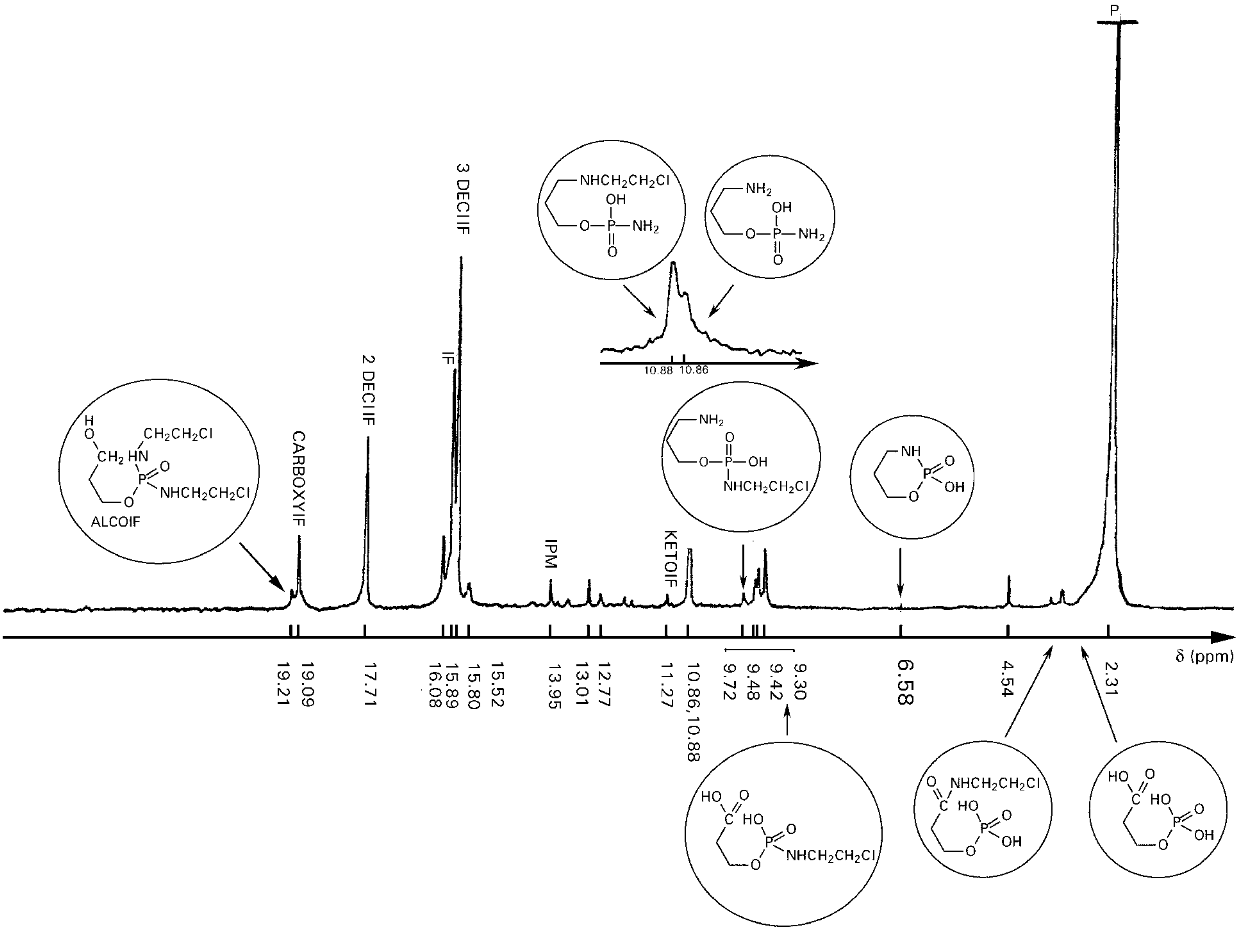



Figure 9 The $^{31}P$ NMR spectrum of a urine sample from a patient treated with ifosfamide (3 g $m^{-2}$ injected intravenously over 3 h). The urine sample was collected 8–16 h after the start of the infusion, frozen immediately after collection and concentrated 17-fold by lyophilization. Chemical shifts are related to external 85% $H_3PO_4$. ALCOIF, alcoifosfamide; CARBOXYIF, carboxyifosfamide; 2DEC1IF, 2-dechloroethylifosfamide; IF, ifosfamide; 3DEC1IF, 3-dechloroethylifosfamide; IPM, isophosphoramide mustard; KETOIF, ketoifosfamide. Compounds ringed are phosphorated metabolites that have been identified using $^{31}P$ NMR.

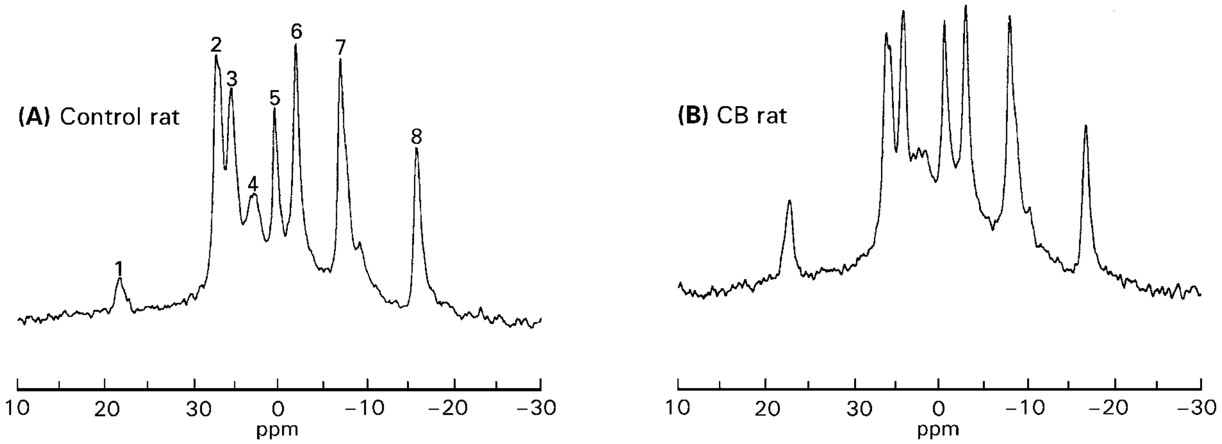



Figure 10 The $^{31}P$ NMR spectra of a GH3 prolactinoma. (A) Ifosfamide administered after 10 min breathing air. (B) Ifosfamide administered during tenth minute of breathing carbogen (CB). The spectra were acquired 15 min after intravenous ifosfamide. Peak assignments: 1, ifosfamide; 2, phosphomonoesters; 3, $P_i$; 4, phosphodiesters; 5, phosphocreatine; 6, $\gamma$-phosphate of NTP; 7, $\alpha$-phosphate of NTP; 8, $\beta$-phosphate of NTP. Reproduced with permission of Churchill Livingstone from Rodrigues *et al* (1997) *In vivo* detection of ifosfamide by $^{31}P$ MRS in rat tumours: increased uptake and cytoxicity induced by carbogen breathing in GH3 prolactinoma. *British Journal of Cancer* **75**: 62–68.

## Conclusion

The increasing number of NMR studies of drug metabolism and pharmacokinetics testifies to the growing interest of the method. For in vitro studies (biofluids in particular), NMR can now be considered as complementary to, and even in some cases replacing, other analytical techniques such as chromatography. In vivo experimental and clinical applications of NMR are still at an early stage of development, and the techniques will need to be improved before localized spectra and accurate absolute quantification are routinely available.

*See also:* **Biofluids Studied by NMR; Cells Studied by NMR; Chromatography-NMR, Applications; *In vivo* NMR, Applications, $^{31}$P; *In vivo* NMR, Applications, Other Nuclei; $^{31}$P NMR; Perfused Organs Studied Using NMR Spectroscopy; Solvent Suppression Methods in NMR Spectroscopy.**

## Further reading


Bell JD, Gadian DG and Preece NE (1990) NMR studies of drug metabolism and disposition. European Journal of Drug Metabolism and Pharmacokinetics 15: 127–133.

Everett JR (1991) High-resolution nuclear magnetic resonance spectroscopy of biofluids and applications in drug metabolism. Analytical Proceedings 28: 181–183.

Findlay MPN and Leach MO (1994) In vivo monitoring of fluoropyrimidine metabolites: magnetic resonance spectroscopy in the evaluation of 5-fluorouracil. Anti-Cancer Drugs 5: 260–280.

Heimberg C, Komoroski RA, Newton JEO and Karson C (1995) $^{19}$F-MRS: a new tool for psychopharmacology. Progress in Psychiatry 47: 213–234.

Komoroski RA (1994) In vivo NMR of drugs. Analytical Chemistry 66: 1024A–1033A.

Kuesel AC, Kroft T and Smith ICP (1991) Nuclear magnetic resonance spectroscopy. Analytical Chemistry 63: 237R–246R.

Lindon JC and Nicholson JK (1997) Recent advances in high-resolution NMR spectroscopic methods in bioanalytical chemistry. Trends in Analytical Chemistry 16: 190–200.

Lindon JC, Nicholson JK and Wilson ID (1996) Direct coupling of chromatographic separations to NMR spectroscopy. Progress in NMR Spectroscopy 29: 1–49.

Malet-Martino MC and Martino R (1992) Magnetic resonance spectroscopy: a powerful tool for drug metabolism studies. Biochimie 74: 785–800.

Preece NE and Timbrell JA (1990) Use of NMR spectroscopy in drug metabolism studies: recent advances. Progress in Drug Metabolism 12: 147–203.


# Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy

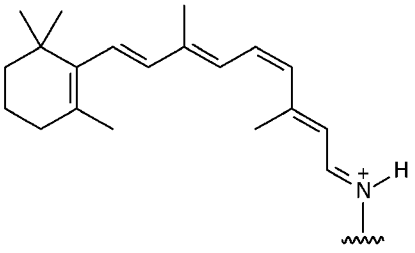

**Volker Buss**, University of Duisberg, Duisberg, Germany
**Lutz Eggers**, Gerhard-Mercator-Universität, Duisburg, Germany


Copyright © 1999 Academic Press

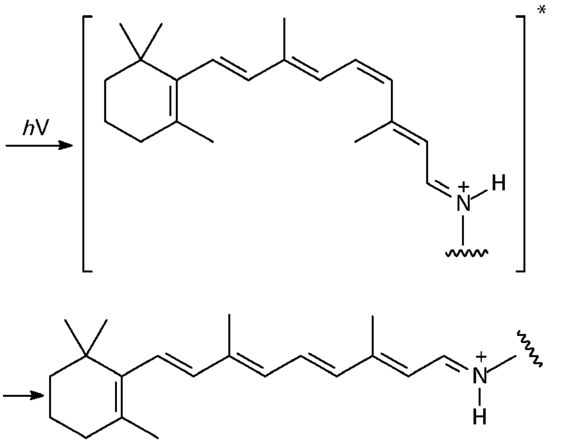
ELECTRONIC SPECTROSCOPY
Applications

It seems appropriate to start an article on UV-visible absorption spectroscopy and dyes with a consideration of the fundamental processes underlying vision, for reasons which will presently become evident.

When electromagnetic radiation with wavelengths between 800 and 400 nm hits the human retina, a chain of events is initiated which eventually results in the excitation of the visual nerve and the perception of light or colour in the brain. The molecular basis for this process is the photoreceptor protein rhodopsin, which forms part of the membrane layers inside the cone and rod cells of the retina. The chromophore of rhodopsin, i.e. the coloured part of an otherwise colourless protein, is a highly unsaturated entity, 11-cis-retinal protonated Schiff base (Scheme 1). Through the action of light the Schiff base is transformed into an electronically excited state from which it returns, in a very fast reaction, to the ground state with the chromophore in the all-trans configuration (Scheme 2). Though there are many different visual pigments found in vertebrate eyes, with absorption maxima ranging from 415 nm in chicken to 575 in humans, nature uses the same chromophore, protonated retinal Schiff base,




**Scheme 1**




**Scheme 2**

throughout. In addition, different pigments are found in the same species, a structural prerequisite for colour vision much like the light-sensitive layers of a colour film coated with different dyes.

How these spectral changes are brought about is not completely understood. The chromophore without the attached protein absorbs at 440 nm in ethanol. Embedding the chromophore in the protein changes the position of the absorption maximum, a consequence of the different environment. This environment consists of the amino acid residues of the protein, differing from species to species, which may be charged or uncharged. There are other counterions and water molecules which might induce solvatochromic effects. An amino acid placed in a particular position might induce deprotonation. For nature, this presents a perfect example of tuning the spectral characteristics of the visual pigment according to its needs. Conversely, one might consider the chromophore an extremely sensitive indicator of its molecular environment.

All molecular properties are a function of the environment. Often these changes are subtle and hidden to the human eye. Sometimes, especially when dyes are involved, they are striking. Thus, dyes may be used as indicators of solvent polarity, of the pH or of the redox potential. Dyes will change their colours when they are adsorbed on surfaces or when they form complexes with other species or with themselves, i.e. when they aggregate. Spectral change, which is what we perceive as the colour change of a dye, always has a molecular basis. It provides yet another window for studying molecular structure and molecular interactions.

## Spectral characteristics of dye molecules

When light travels through a homogenous isotropic solution of an absorbing species, the fraction of light that is absorbed is proportional to the number of molecules in the light path according to the Beer–Lambert equation

$$\log_{10}(I_0/I) = \varepsilon \cdot c \cdot l$$

in which $I_0$ is the intensity of the incident and I of the emergent radiation. With the concentration c in moles per litre and the path length l in cm, the proportionality constant ε is the molar absorptivity or molar extinction coefficient. While ε is a measure of the intensity of the absorption, the wavelength λ at which the absorption occurs, or better its inverse, the wavenumber $\bar{\nu}$, is a measure of the energy which is absorbed, according to the well-known relation

$$E = hc/\lambda = hc\bar{\nu}$$

in which h is Planck's constant and c the speed of light. The range of UV-visible absorption extends from 800 to 200 nm (12 500 to 50 000 $cm^{-1}$) which corresponds to energies from ca. 35 to 145 kcal $mol^{-1}$ (1 kcal = 4.184 kJ). In molecules absorbing in this wavelength range these energies are sufficient to create short-lived electronically excited states. In such a state the electronic wavefunction does not correspond to the most stable arrangement of the electrons in the field of the nuclei; as a consequence the molecule returns, within $10^{-7}$ to $10^{-9}$ seconds, to the electronic ground state. Factors that affect the two states differently, e.g. chemical substitution, protonation or change of the solvent, will induce spectral shifts either to longer ('bathochromic shift') or shorter ('hypsochromic shift') wavelengths.

Empirical rules are widely employed to correlate chemical structure changes with spectral shifts. The spectra of short conjugated double bond systems can be predicted based on the topology and the presence of substituents (Woodward rules); linear free-energy relationships have been developed to correlate

substituent and spectral changes in closely related dye molecules. Other rules originate from basic quantum-mechanical principles, e.g. the linear relation between the absorption maxima of cyanine dyes and the length of the conjugated chain; the perturbation of a conjugated carbon chain by substituents of different electronegativity (perturbational molecular orbital theory); or the effects of steric crowding and, as a consequence, non-planar distortion of the molecule.

For a more quantitative description, theories may be applied which range from the purely π–electron, but highly successful PPP-theory and the semi-empirical all-valence electron theories such as CNDO/S and AM1 to ab initio methods. Only the latter allow the calculation of the energies without requiring recourse to experimental data. While it is possible today to calculate molecular ground state energies with an accuracy of a few kilocalories, depending on the size of the molecule and the computational effort expended, the uncertainty is much larger in the excited state. The ab initio calculation of absorption wavelengths which come within 50 nm of the experimental value must still be considered successful.

Calculated energies refer to the isolated molecule in the (dilute) gas phase. Every solvent will change these data to a degree. Solvent shifts can be striking when there is significant charge rearrangement between the ground and excited state. For computing these effects, one may allow specific interactions, e.g. through hydrogen bonds, of the dye with the solvent, though this approach is limited in the number of the solvent molecules that can be considered. In other approaches the dye molecule is specifically solvated and then embedded into a solvent continuum which in its response to the solute is treated as a polarizable dielectric.

A measure of the absorption intensity is the oscillator strength f, a dimensionless quantity which is obtained by integrating over the absorption band:

$$f = 4.315 \times 10^{-9} \int \varepsilon \mathrm{d}\bar{\nu}$$

f can also be calculated from a theoretical quantity, the transition dipole moment *M*,

$$f = (8\pi m_e \bar{\nu}/3he^2)|\boldsymbol{M}|^2$$

where $m_e$ and e are, respectively, the mass and charge of the electron and $\bar{\nu}$ is the absorption wavenumber. *M*, a vector quantity, is a measure of the charge displacement caused by the electronic excitation. When the absorption can be attributed to the excitation of an electron from one molecular orbital to another with an energy sufficiently different from all other excitations of the molecule the magnitude and direction of *M* can be estimated from multiplying the two orbitals which are involved in the excitation (Figure 1). This procedure gives atom-centered transition charge densities which are used to calculate the transition dipole moment. These 'back-of-an-envelope' calculations readily explain why the oscillator strength increases with the size of the chromophore, or why it is larger in the stretched conformation of a molecule than when it is coiled. The fixed orientation of *M* with respect to the molecular frame is also the physical basis for dichroism, i.e. the orientation dependence of the extinction coefficient of an ordered ensemble of dye molecules.

Electronic excitation in the UV-visible range is accompanied by changes in the vibrational and rotational state of the molecule as well. Since the spacing between rotational and vibrational levels is small relative to the spacing between electronic

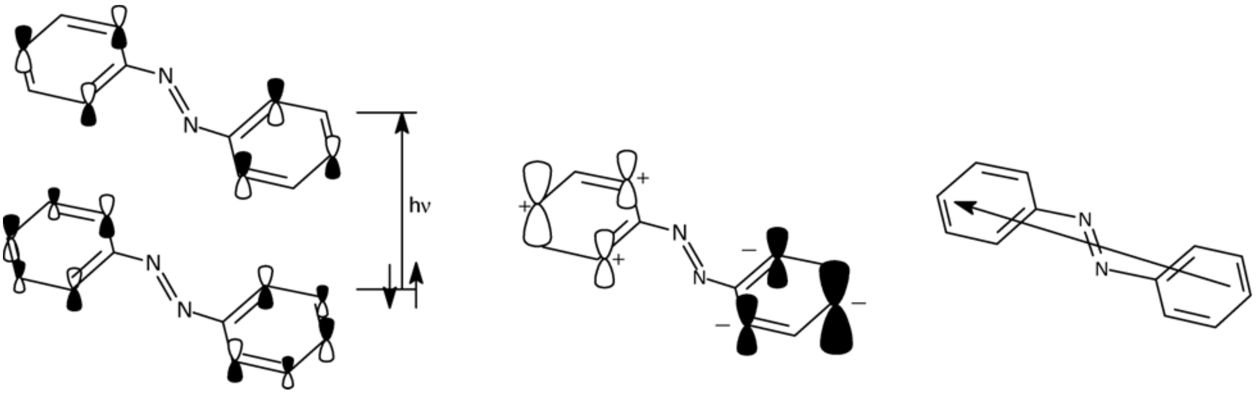



**Figure 1** Highest occupied and lowest unoccupied molecular orbitals of azobenzene (schematic, left), their product, giving the transition density (middle) and the resulting transition moment vector (right).

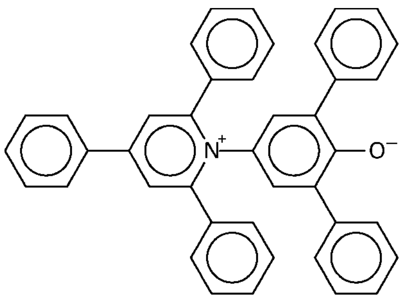
states, a typical UV-visual spectrum will appear as a band, or an envelope curve, of close-lying 'rovibronic states', the band shape being determined by the relative intensities of those states. Strong vibrations of the excited state may give rise to 'vibrational fine structure' of an absorption band, which may change as a function of the solvent and temperature.

## Dyes as indicators of solvent polarity

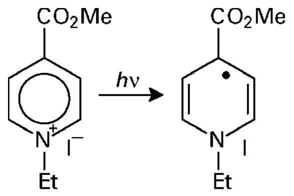
Any solvent will change the wavelength, intensity and shape of an absorption band compared to the spectrum in the gas phase to a degree, a consequence of the different perturbation of the electronic ground and excited states of the solute by the solvent. The term 'Solvatochromism' has been coined to describe the change in position of an UV-visible absorption band due to a change in the polarity of the medium. A hypsochromic or blue shift with increasing solvent polarity is called negative solvatochromism, whereas a bathochromic or red shift is called positive solvatochromism. It is unreasonable to expect any macroscopic parameter, such as the dipole moment of a solvent or its refractive index, to correlate quantitatively with solvent effects which depend in addition on interactions at the molecular level. Therefore, empirical scales have been devised to correlate solvent polarity and spectral shifts.

One of the most popular and successful scales has been developed by Dimroth and Reichardt. It is based on the pyridinium-N-phenoxide betaine [3], which exhibits one of the largest solvatochromic effects ever observed. The solvatochromism of this dye is negative since its ground state is considerably more polar than the excited state and is stabilized by polar solvents. Thus, in diphenylether the dye absorbs at 810 nm and appears blue-green, whereas in water the absorption is centred at 453 nm, giving an orange impression. The transition energy, expressed in kcal mol $^{-1}$, is the so-called $E_T(30)$ value of the solvent. $E_T(30)$ values have been determined and tabulated for more than 270 pure solvents and many different solvent mixtures. Protonation converts the dye (Scheme 3) into a phenol; as a consequence, $E_T(30)$ values cannot be measured for acidic solvents, such as carboxylic acids.

Another solvent polarity scale, Kosower's Z-scale, is based on the highly solvent-dependent charge transfer absorption of the para-substituted pyridinium salt (Scheme 4). The ground state of this dye is best described as a highly polar tight ion-pair, whereas the less polar charge-transfer form is more important in the excited state. Thus, the solvatochromism of the pyridinium salt is negative, like Reichardt's dye. In fact, the $E_T(30)$- and the Z-scale




**Scheme 3**




**Scheme 4**

show a very good correlation, and similar anomalies with respect to theoretical polarity functions, which is not surprising because both use similar molecular models to 'measure' solvent polarity.

For converting experimental transition energies $\bar{\nu}$ (in $cm^{-1}$) into $E_T(30)$ and Z-values, the following equation may be used:

$$E_T(30) = Z = hcN_A\bar{\nu}$$
$$= 2.8591 \times 10^{-3}\,\bar{\nu}$$

in which $N_A$ is Avogadro's constant and both $E_T(30)$ and Z have the dimensions kcal $mol^{-1}$. The use of normalized, dimensionless values '$E_T^N$' has been recommended in order to avoid the non-SI unit. In this scale, water and trimethylsilane (TMS) are taken as reference solvents, with assigned values of $E_T^N$ of 1.00 and 0.00, respectively.

$E_T(30)$ values show a good, often linear, correlation with a large number of other solvent sensitive processes, such as reaction rates and shifts of chemical equilibria. The betaine dye (Scheme 3) and specially designed derivatives are useful molecular probes in the study of micellar interfaces, microemulsions and phospholipid bilayers, of rigid rod-like isocyanide polymers, and the retention behaviour in reversed-phase chromatography. In addition to its solvatochromic behaviour, the dye is sensitive to temperature ('thermosolvatochromism') and pressure changes ('piezosolvatochromism') and also to the presence of electrolytes ('halosolvatochromism').

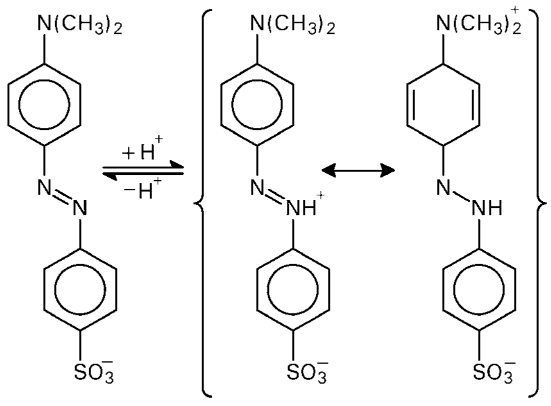
## Dyes as indicators of chemical equilibria

For a dye to act as indicator of a chemical equilibrium, it has to meet two requirements; first, it must be able to participate in the reaction. That is, in an acid/base equilibrium the dye itself must have acidic or basic properties; in a redox reaction, the dye must exist in different oxidation states; in a complexation reaction involving e.g. metal ions, the dye must have complexing properties. Second, the different forms of the dye involved must absorb at different wavelengths, i.e. they must have different colours. Applications of dyes as indicators are of enormous importance in analytical chemistry.

One of the oldest applications of UV-visible spectroscopy is found in acid–base titrations and the determination of pK values. Several types of organic dyes are very sensitive to changes in the concentration of $H_3O^+$ (or $OH^-$) ions, e.g. azo dyes, phthaleines or triphenylmethane dyes. As a typical example, consider the protonation of methyl orange (Scheme 5), which occurs at a position on the aza bridge which allows the delocalization of the positive charge into the p-dimethylamino group. As a consequence, one of the aromatic rings becomes part of a cyanine type chromophore (eight π-electrons delocalized over seven atomic centres), with a concomitant bathochromic shift of the absorption spectrum and a colour change from yellow to red.

The intense red, dianionic form of phenolphthalein (Scheme 6), which is stable in weakly basic solution, can be formulated as a resonance hybrid but is really another example of a cyanine-type chromophore, with oxygen instead of nitrogen. In acidic or strongly basic solutions, the negative charges are localized and the compound is colourless.

During titration, the sharp change of the pH near the end point is indicated by the visible colour change




**Scheme 5**

as the added indicator dye is converted from the basic into the acidic form (or vice versa). Since the end point of the titration does not necessarily coincide with neutrality of the solution, an indicator should be employed with a pK value that corresponds most closely to the pH of the end point. The concentrations (or better activities) of both the acidic and the basic form of the indicator are then the same and true mixed colour is observed. For a determination of the indicator pK, the extinction coefficient ε of the dye is measured in buffered solutions at different pH values. According to the relationship

$$pK = pH + \frac{\varepsilon - \varepsilon_{\text{base}}}{\varepsilon_{\text{acid}} - \varepsilon}$$

the pK can be calculated from ε and the extinction coefficients of the indicator in strongly alkaline ($\varepsilon_{\text{base}}$) and strongly acidic solutions ($\varepsilon_{\text{acid}}$). All measurements should be performed in dilute solutions, and at wavelengths sufficiently distant from the isosbestic point where $\varepsilon_{\text{base}} = \varepsilon_{\text{acid}}$ and the extinction does not change with the pH.

Similar considerations apply to redox titrations, where the indicator dye should possess differently coloured oxidation states and a redox potential which is appropriate for the reaction to be studied. A large number of dyes are known which are suitable for this purpose. Ferroin (Scheme 7), a complex between iron(II) and phenanthrolin with a deeply red colour, upon oxidation forms a pale blue iron(III) complex. Another redox indicator is diphenylamine (Scheme 8), which is colourless in the reduced state. It is oxidized in acidic solution irreversibly to diphenylbenzidine (Scheme 9) which can be further oxidized reversibly to the intensely coloured diphenylbenzidine violet (Scheme 10).

Comparable to an acid/base titration the end point of a redox titration is reached when the concentrations of the reduced and the oxidized form of the dye are the same. The latter situation is somewhat more complex, however, because most redox reactions involve the exchange of protons, i.e. the equilibria are pH dependent.

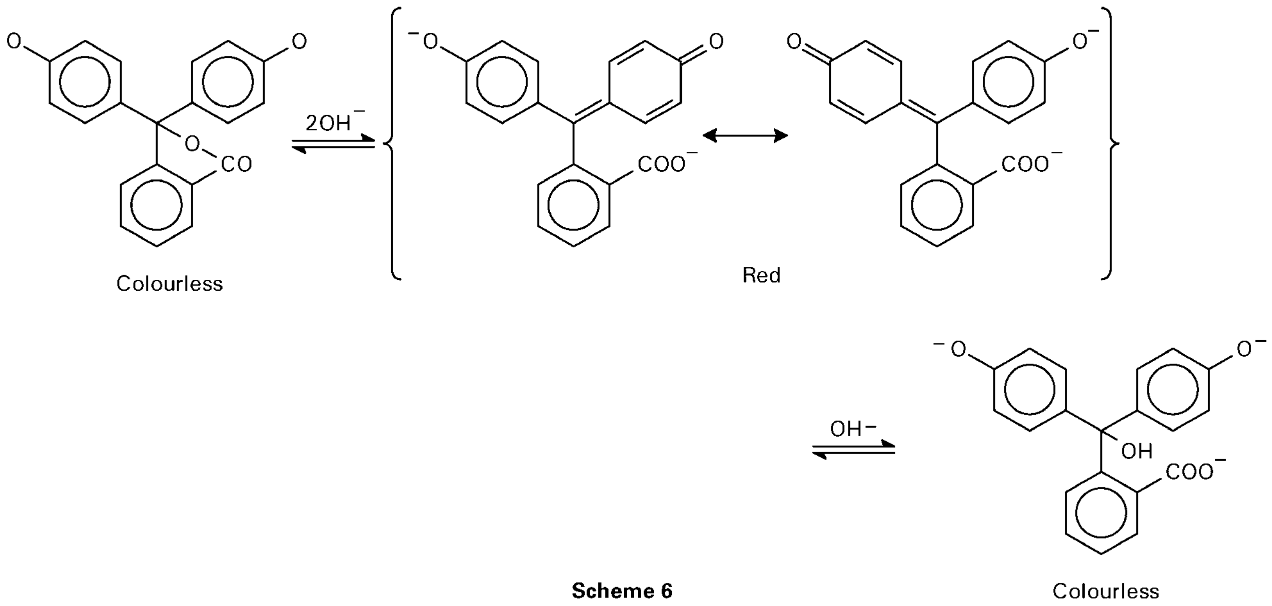
In the determination of metal ions by ligand binding with ethylenediaminetetraacetic acid, EDTA, which has become known as complexometric or chelatometric titration, the end point may be indicated by so-called metal indicators. For this method to work, the metal to be titrated must form a less stable complex with the dye than with the complexing agent. Titration with the complexing agent binds all free metal in solution, until finally the dye is demetallated, too, which is manifest in the colour change.

A very popular indicator is Eriochrome Black T (Scheme 11), which forms complexes with many different metals (the chromium complex is also used to dye protein and polyamide fibres). Between pH values of 7 and 11, the dye changes its colour from blue in the uncomplexed to red in the complexed form.

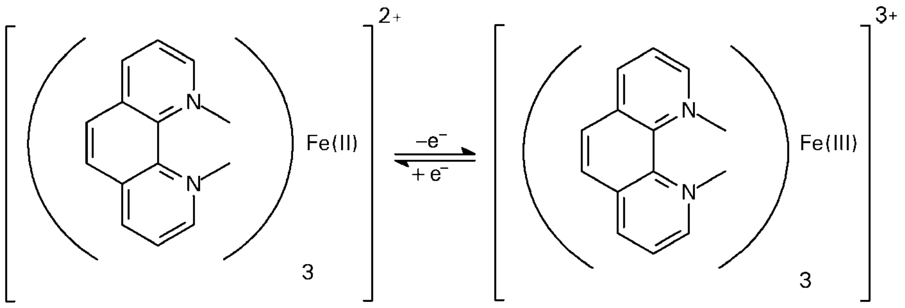
The end point of a titration may be indicated, besides the chemical transformation of a dye indicator, by physical processes, such as the adsorption of eosin from solution on the surface of a silver halide precipitate.

## Dyes as indicators of molecular environment

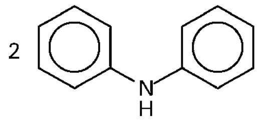
Any medium other than the vacuum will affect the absorption spectrum of a dye, through its bulk properties as a dielectric, which does not imply a fixed geometry or stochiometry, and through specific interactions, such as electrostatic interaction, π–π stacking, or hydrogen bonding. The combination of any of these between a dye and an organic substrate may lead to the formation of an entity with a more or less well-defined supramolecular geometry, which exhibits characteristic spectral changes. As an example, consider the complex formed between p-nitrophenol and α-cyclodextrin (Figure 2). Cyclodextrins are water-soluble cyclic oligomers of amylose with a torus-like structure and a strongly hydrophobic inner lining. While they form complexes with many different guests, the complexation of dyes can be quantitatively followed by absorption spectroscopy. Usually, the low resolution of the spectra is not sufficient to develop a detailed structural model of these complexes, and other methods have to be used to implement the UV-visual spectroscopic data. The most accurate information is obtained from an analysis of X-ray data (Figure 2) which displays the orientation of the dye inside the cyclodextrin host. In contrast to the solid state the condition in solution is dynamic and more complex, since more species are involved, including the water molecules which are equilibrating between the solvent, the solute and the pore. However, the formation of a



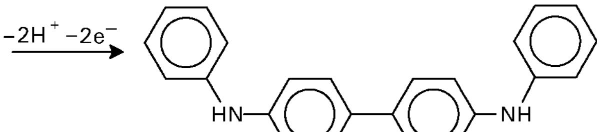

**Scheme 6**



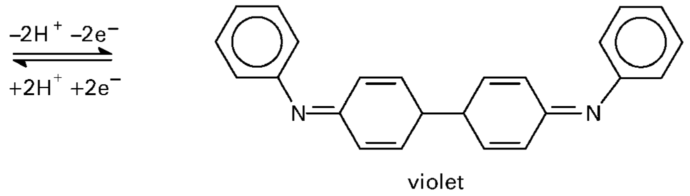

**Scheme 7**




**Scheme 8**




**Scheme 9**

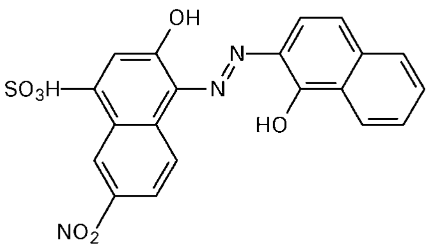



**Scheme 10**

structured host–guest complex is supported by other spectroscopic methods, such as 2-dimensional NMR spectroscopy.

The situation depicted above is typical for many cases where the complex formation between dyes and organic substrates is studied as a model for intermolecular interaction, the dye functioning at the same time as the complexed ligand and the indicator of this interaction. The literature dealing with this topic is vast and ranges from histochemistry, where tissue is stained for diagnostic purposes, and biochemistry, where the complexation with dyes is used to model enzyme–ligand interactions and study possible binding sites, to organic polymers where dyes act as indicators of macromolecular conformations or of phase transitions.

Dyes, and their absorption changes, can be used to sense chirality and differentiate between two enantiomers 'with the naked eye', i.e. without the use of a polarimeter or any other device that measures optical activity. In principle every diastereomeric interaction between a chiral medium and the two enantiomers of a chiral dye must lead to different spectral characteristics. In practice, however, the effects are




**Scheme 11**

usually too small to be observed. The situation is different when diastereomeric supramolecular structures are involved. Thus, the colour of a phenol based bowl shaped agent (a 'calixarene') with two indophenol chromophores and a chiral binaphthyl subunit (Figure 3) turns blue-violet when it forms a 1:1 complex with (R)-phenyl-glycinol, while the corresponding complex with the (S)-isomer has a much smaller

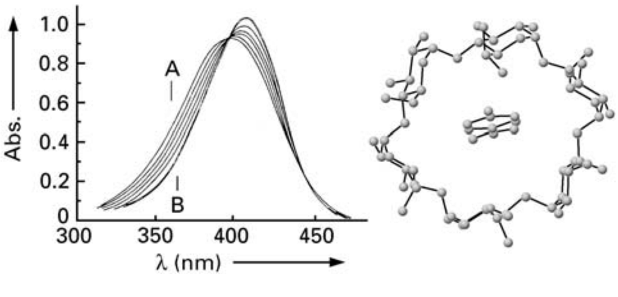



**Figure 2** Spectral change, A to B, upon addition of α-cyclodextrin to an aqueous solution of *p*-nitrophenolate (left; after Saenger W (1980). *Angewandte Chemie* **92**: 343). X-ray structure of the α-cyclodextrin-*p*-nitrophenol complex (right: after Harata K (1977) *Bulletin of the Chemical Society of Japan* **50**: 1416).

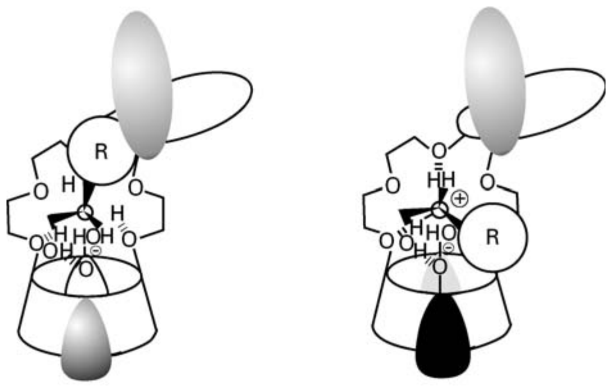



**Figure 3** Possible structures of diastereomeric complexes formed between the bowl-shaped chiral calixarene and an (*S*)-(left) and an (*R*)-configurated aminoalcohol. The shaded entities at the bottom represent the indophenol chromophores, the ones on top the binaphthyl unit. R is a phenyl group of the chiral amine (Reproduced with permission from Kubo Y, Hirota N, Maeda S, Tokita S (1998) *Analytical Sciences: the International Journal of the Japan Society for Analytical Chemistry* **14**: 183).

association constant and exhibits mainly the spectrum of the uncomplexed chiral sensor, which is red. As a consequence, it is possible to detect chirality by simple visible inspection. Similarly, supramolecular structures involving different kinds of non-bonded interactions have been invoked to account for the ability of chiral phases to discriminate between enantiomers in chromatographic separation techniques.

For a study of the molecular microenvironment, but also as promising new optical materials in solar energy conversion, for dye lasers, as optical switches or in non-linear optics, silicate glasses doped with dye molecules have been prepared by the sol–gel method. The formation of the porous solid from the initially liquid organic solution has been followed by UV-visible spectroscopy, and the acid/base reaction of the dyes with the inorganic matrix has been studied. Solvatochromic dyes have been used to determine the $E_T(30)$ values of different sol–gel materials. One advantage of these so-called xerogels is the immobilization of the embedded dye in a rigid cage, which prevents aggregation processes which are typical for concentrated dye solutions (see next section). Another aspect is the increased photochemical stability and fluorescence efficiency of the dyes. For the development of biochemical sensors, whole enzymes have been encapsulated in xerogels and their activity shown by the formation of dyes inside the glass from coupled enzyme reactions.

## Dyes as indicators of aggregation and molecular order

Dyes have a natural tendency to aggregate, a consequence of their typically flat geometries which allows for maximum contact when the aggregation takes place side by side, in a sandwich manner. The aggregation is effected by non-bonding interactions, including $\pi$–$\pi$ stacking, and is supported by hydrophobic effects when the solvent-accessable molecular surface is reduced as a consequence of aggregation. Aggregation is promoted by increasing the dye concentration in a solution or lowering the temperature. It is one of the most common causes for deviation of a solution from the Beer–Lambert law. Often a change of the solvent from polar to less polar or vice versa will be sufficient to effect dye aggregation or deaggregation. Aggregates may form spontaneously in solution or in the presence of polyelectrolytes, such as polyphosphates, or biological macromolecules, which may influence or even control the formation of the aggregate. Aggregates may form on solid surfaces, such as mica, or they may be assembled as monomolecular layers on glassy surfaces. Dye aggregates grown on the surface of silver halide grains absorb light of suitable wavelength leading to latent images in the halide; this so-called spectral sensitization is the basis for photography. Dyes tend to concentrate at liquid interfaces and interact with surfactants leading to distinct spectral changes. Finally, dyes may adjust to the highly ordered environment present in liquid crystals, an area of huge technical importance.

Conspicuous colour changes take place when dyes aggregate. The spectral shifts caused by aggregation can be to longer or to shorter wavelengths relative to the monomeric dye absorption. When dyes dimerize, a splitting of the absorption band may be observed. Formation of trimers, tetramers or larger oligomers is accompanied by increasing spectral shifts. The limiting value corresponding to an aggregate of ‘infinite’ size is called a ‘J-band’ characteristic of a J-aggregate when lying on the long wavelength side of the monometer; if it appears on the high-energy side, it is called an ‘H-band’ (Figure 4).

These spectral changes are the result of the individual dye molecules interacting with each other in the close contact of the aggregate. Theoretical models based on the Coulomb interaction of the transition moments have been put forward to correlate the spectral shifts with certain geometrical parameters of the aggregate. Thus the hypsochromic shift observed in H-aggregates results when the dye contacts are mainly side by side. A lateral shift of the dye molecules against each other or a stacking into a brickstone work arrangement will lead to the bathochromic shift of a J-aggregate. While the H-band is usually sharp and intense due to the high order of the aggregate, the J-band may appear broad and featureless due to the loose and more irregular structure, but exceptions are well documented.

For studying dyes adsorbed on solids by UV-visible spectroscopy transparent solutions are necessary.

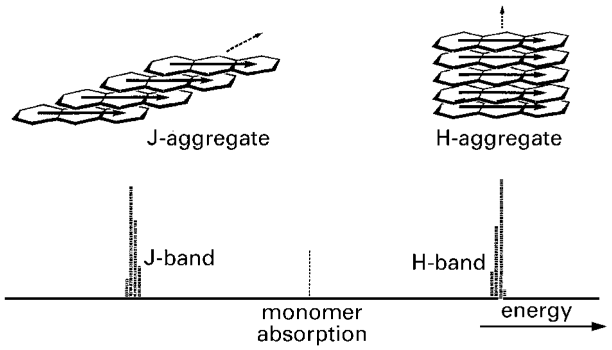



**Figure 4** Stacking of dye molecules (with transition moment vectors) which leads to a J-aggregate with red-shifted absorption relative to the monomer (left) and to an H-aggregate with blue-shifted absorption (right).

There are many inorganic colloids suitable for this purpose, such as silicas, charcoal, clays, zeolites or metal oxides, but also inorganic and organic polymers and even aggregates, such as micelles. Flocculation or aggregation of smaller particles may occur due to the adsorption of a charged organic dye on a polyelectrolyte which reduces the elecrostatic repulsion by charge compensation.

The number of publications dealing with this topic is increasing rapidly. For studying the interaction of dyes with an enzyme the adsorption on solid surfaces with different binding characteristics may serve as a model. In this case the dyes are used in their classical role as indicators. Dyes are well suited for this purpose because of their strong tendency towards chemi- and physisorption. When the UV-visible properties of the adsorbed species change sufficiently the immediate comparison with dye molecules dissolved in a homogeneous solution is possible. Also, dynamic aspects of the adsorption of molecules on these surfaces can be examined by electronic absorption techniques.

The colour change accompanying the self-aggregation of dyes on solid surfaces is called metachromasy. Other properties of adsorbed dyes can also change allowing new or improved applications, such as enhanced stability against environmental effects such as oxygen or water or higher quantum yields for a photochemical conversion. New technologies are developing based on such processes. Solar energy conversion in photovoltaic cells, where charge carriers are generated by photoexcitation of dyes deposited on titanium dioxide surfaces, is a good example.

Complex equilibria between the dye and its surroundings may take place in heterogenous liquid mixtures, involving self-association, binding with ionic surfactants, micellar aggregation and micellar solubilization. These equilibria are concentration and temperature dependent, and involve entities with sometimes unspecific spectroscopic properties which leave plenty of room for discussion. Still, much can be learned from studying how dyes distribute in heterogenous systems since this can serve as models for the biodistribution of drugs in tissues and membranes and their behaviour at the lipophilic barrier. Also, the photosensitizing properties of a dye may increase when brought into a lipophilic environment.

H-aggregation of methyl orange has been postulated to account for the hypsochromic shift of the dye in a water-in-oil microemulsion, with the dye molecules in a parallel orientation at the interface beween the water and the oil phase. The observation of salt-induced strongly red-shifted J-bands in aqueous solutions of Rose Bengal, an anionic xanthene dye, in the presence of zwitterionic surfactants was explained on the basis of head-to-tail aggregates of the dye molecules in novel premicellar aggregates.

Highly ordered phases, so-called liquid crystals, are formed by a variety of rod-like, rigid dipolar organic molecules. Liquid crystals are also called the fourth state of matter since they combine the long-range order of a crystalline solid with the fluidity of a liquid. When a dye is dissolved in a liquid crystal the molecules will form ordered aggregates, with their long axes preferentially aligned with the liquid crystal molecules. These crystals are dichroic, i.e. they have different absorptivites for light polarized parallel and perpendicular to that axis, from which the orientation of the transition moment vector relative to the molecular framework can be determined.

The dichroism is the physical basis for the use of dyes in LCDs or liquid crystal displays: a small amount (0.1–1%) of the dichroic dye (the 'guest') is dissolved in a nematic or cholesteric liquid crystal (the 'host'). Different orientations of the dye aggregates are achieved by application of an electric field through transparent electrodes. The dye will for example align with its long axis parallel to the long axis of a nematic liquid crystal. If this solution is aligned in a parallel mode of the electrodes and has a positive dielectric anisotropy, the cell will be coloured in the off position, when the dye absorbs strongly, and become colourless upon application of an electric field, when the alignment is changed.

## List of symbols

c = molar concentration; $e_e$ = charge of electron; $E_T(30)$ = solvent transition energy; f = oscillator strength; I = intensity; l = path length; $m_e$ = mass of electron; $M$ = transition dipole moment; ε = molar extinction coefficient.

*See also:* **Biomacromolecular Applications of UV-Visible Absorption Spectroscopy; Biomacromolecular Applications of Circular Dichroism and ORD; Colorimetry, Theory; ORD and Polarimetry Instruments**.

## Further reading

Fabian J and Hartmann H (1980) Light Absorption of Organic Colorants. Berlin: Springer.

Rao CNR (1975) Ultraviolet and Visible Spectroscopy. London: Butterworth.

Reichardt C (1995) Solvatochromic dyes as solvent polarity indicators. Chemical Reviews 94: 2319–2358.

Sommer L (1989) Analytical Absorption Spectrophotometry in the Visible and Ultraviolet. Oxford: Elsevier.

Suppan P and Ghoneim N (1997) Solvatochromism. Cambridge: The Royal Society of Chemistry.

Zollinger H (1991) Color Chemistry. Weinheim: VCH.

# E

# EELS, Applications

*See* **High Resolution Electron Energy Loss Spectroscopy, Applications.**

# Electromagnetic Radiation

**David L Andrews**, University of East Anglia, Norwich, UK

Copyright © 1999 Academic Press

| FUNDAMENTALS OF SPECTROSCOPY |
|---|
| Theory |

Electromagnetic radiation is the technical term for light; not just visible light, but any light, ranging right through from radio frequencies to gamma rays. As the name suggests, light of all kinds is radiated through conjoined electric and magnetic fields, as shown by Maxwell in the mid-nineteenth century. To fully appreciate the nature of electromagnetic radiation, however, we first have to consider both its wave-like and photonic aspects.

## Waves and photons

In the case of monochromatic (single-frequency) radiation, light propagates as a wave with a well-defined repeat or wavelength $\lambda$ (**Figure 1**). Travelling at the speed of light, $c$, these waves oscillate at a characteristic frequency $\nu$ given by $c/\lambda$. The electric field, the magnetic field and the direction of propagation are mutually perpendicular, and for convenience may be chosen to define a set of Cartesian axes ($x$, $y$, $z$), respectively. With propagation in the $z$ direction, the electric field vector points in the $x$ direction and oscillates in the $xz$ plane:

$$E(z,t) = E_0 \sin(kz - \omega t) \qquad [1]$$

where $E_0$ is the field amplitude and where also for conciseness we introduce $k = 2\pi/\lambda$ and $\omega = ck = 2\pi\nu$. The magnetic field vector accordingly points in the $y$ direction and oscillates in the $yz$ plane:

$$B(z,t) = B_0 \sin(kz - \omega t) \qquad [2]$$

with amplitude $B_0$. Given that the electric and magnetic fields oscillate in phase and in mutually perpendicular planes, equations [1] and [2] together satisfy Maxwell's equations provided the amplitudes are related by $E_0 = cB_0$. In fact most spectroscopy involves the interaction of matter with the electric rather than the magnetic field of the radiation, because of its stronger coupling – usually by a factor of a thousand or more – with atomic and molecular charge distributions.

The speed of travel of an electromagnetic wave as described above can be understood in terms of the motion of any part of the waveform, such as the crest. The interval $\Delta t$ between the arrival of two successive crests at any given point is given by $\omega\Delta t = 2\pi$, while at any time their spatial separation $\Delta z$ also satisfies $k\Delta z = 2\pi$. So the speed of propagation $\Delta z/\Delta t = \omega/k = c$, which *in vacuo* takes the value $c_0 = 2.9979 \times 10^8$ m s$^{-1}$.

On the radiation travelling through any substance, the electronic influence of the atoms or molecules traversed by the light reduces the propagation speed to a value less than $c_0$; then we have $c = c_0/n_\lambda$ where

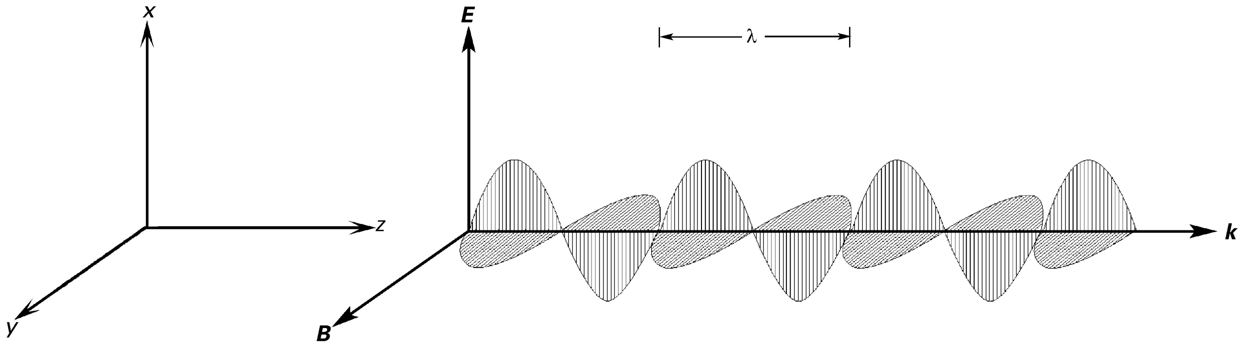



**Figure 1** Oscillating electric field E and magnetic field B associated with mononchromatic radiation.

$n_\lambda$ is the refractive index of the medium for wavelength λ; accordingly, $k = 2\pi n_\lambda/\lambda$. The significance of the refractive correction is greatest in the solid or liquid phase, especially at wavelengths close to an optical absorption band of the medium.

A fundamental paradox in the nature of electromagnetic radiation, to some extent apparent even in the earliest scientific studies by Newton and others, is that it exhibits not just wave-like but also particle-like (corpuscular) properties, and both prove to be of key importance in spectroscopy. In particular, it is only through the association of discrete units of energy with electromagnetic radiation of any given frequency that we can properly understand atomic and molecular transitions and the appearance of spectra.

In the modern quantum representation of light developed by Planck, Einstein, Dirac and others, we now understand the twin wave and particle attributes through a description in terms of *photons* (a term in fact first introduced by the thermodynamicist Lewis). As such, electromagnetic radiation of a given frequency ν is seen to propagate as discrete units of energy $E = h\nu$, where $h$ is the Planck constant.

With the key concepts in place, we can now take a look at some of the more detailed aspects of electromagnetic radiation in two stages — first by more fully enumerating the properties of photons in general terms, and then by examining those more specific features that relate to particular wavelength or frequency regions of the electromagnetic spectrum.

## Photon properties

**Mass** Photons are elementary particles with zero rest mass – necessarily so, since, from special relativity theory, no particle with a finite mass can move at the speed of light.

**Velocity** The speed of light is normally quoted as speed *in vacuo*, $c_0$, with refractive corrections applied as appropriate; the free propagation of any photon also has a well-defined direction, usually denoted by the unit vector $\hat{\mathbf{k}}$.

**Energy** Photon energy is linked to optical frequency $\nu$ through the relation $E = h\nu$ (where $h = 6.6261 \times 10^{-34}$ J s). Each photon essentially conveys an energy $E$ from one piece of matter to another, for example from a television screen to a human retina.

**Frequency** The optical frequency ν expresses the number of wave cycles per unit time. Also commonly used in quantum mechanics is the circular frequency $\omega = 2\pi\nu$ (radians per unit time), in terms of which the photon energy is $E = \hbar\omega$ where $\hbar = h/2\pi$. The lower the optical frequency, the more photons we have for a given amount of energy; and the larger the number of photons, the more their behaviour approaches that of a classical wave (this is one instance of the 'large numbers' hypothesis of quantum mechanics). It is for this reason that electromagnetic radiation becomes increasingly wave-like at low frequencies, and why we tend to think of radiofrequency and microwave radiation primarily in terms of waves rather than particles.

**Wavelength** The wavelength $\lambda$ of the electric and magnetic waves is given by $\lambda = c/\nu$. In spectroscopy, common reference is made to its inverse, the *wavenumber* $\bar{\nu} = 1/\lambda$, usually expressed in $\text{cm}^{-1}$.

**Momentum** Each photon carries a linear momentum $p$, a vector quantity of magnitude $h/\lambda = \hbar k$ pointing in the direction of propagation. It is then convenient to define a *wave vector or propagation vector* $\mathbf{k} = k\hat{\mathbf{k}}$ such that $\mathbf{p} = \hbar\mathbf{k}$. Since the photon momentum is proportional to frequency, photons of high frequency have high momenta and so exhibit the most particle-like behaviour. X-rays and gamma rays, for example, have many clearly ballistic properties not evident in electromagnetic radiation of lower frequencies.

**Electromagnetic fields** The electric and magnetic fields, $E$ and $B$ respectively, associated with a photon are vector quantities oriented such that the unit vectors ($\hat{E}$, $\hat{B}$, $\hat{k}$) form a right-handed orthogonal set.

**Polarization** For plane-polarized (also called linearly polarized) photons, the plane within which the electric field vector oscillates can sit at any angle to a reference plane containing the wave vector, as shown in **Figures 2A** and **2B**. Other polarization states are also possible: in the right- and left-handed circular polarizations depicted in **Figures 2C** and **2D**, the electric field vector sweeps out a helix about the direction of propagation. Elliptical polarization states are of an intermediate nature, between linear and circular. Together, the wave vector and polarization of a photon determine its *mode*.

**Spin** Many of the key properties of photons as elementary particles relate to the fact that they have an intrinsic spin $S = 1$, and so are classified as *bosons* (particles with integer spin as opposed to half-integer spin particles of matter such as electrons). As such, photons collectively display a behaviour properly described by a Bose–Einstein distribution. At simplest, this means that it is possible for their oscillating electromagnetic fields to keep in step as they propagate. Through this, coherent beams of highly monochromatic and unidirectional light can be produced; this is of course the basis for laser action.

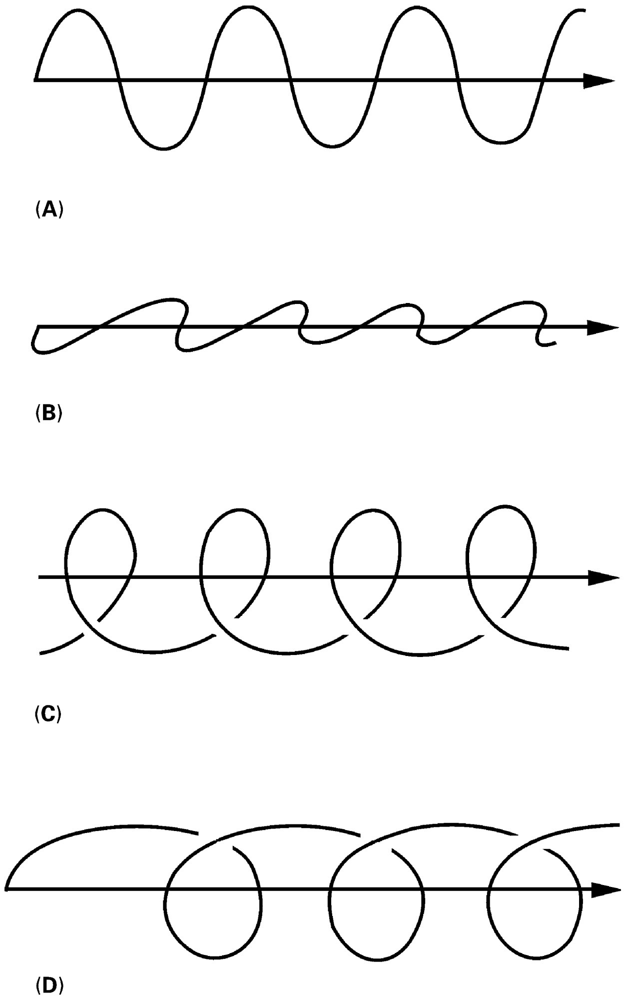


Figure 2 Polarization states: (A) and (B) plane; (C) and (D) circular.

**Angular momentum** The intrinsic spin of each photon is associated with an angular momentum, a feature that plays an important role in the selection rules for many spectroscopic processes. Circularly polarized photons have the special property of quantum angular momentum: the two circular polarization states, left- and right-handed, respectively carry +1 or –1 unit of angular momentum, $\hbar$.

## Regions of the spectrum

We can now delineate the properties of the various spectral regions shown in **Figure 3**. Looking across the electromagnetic spectrum, it is immediately apparent that there is enormous difference in scale between the extremes of wavelength (or frequency), and it is not surprising to find that they encompass an enormous range of characteristics.

**Gamma-rays** This is a region of highly penetrating radiation with the highest optical frequencies (exceeding $3 \times 10^{19}$ Hz), and an upper bound of 10 pm on the wavelength. Spectroscopy in this region primarily relates to nuclear decay process, as in the Mössbauer effect.

**X-rays** This region encompasses both 'hard' X-rays (wavelengths down to 10 pm) and less penetrating 'soft' X-rays (wavelengths up to 10 nm) with optical frequencies lying between $3 \times 10^{19}$ and $3 \times 10^{16}$ Hz. With the capacity to produce ionization by electron detachment, photon energies in this region are often reported in electronvolts (eV), and run from around $10^5$ eV down to 100 eV (1 eV = $1.602 \times 10^{-19}$ J). X-ray absorption and fluorescence spectra as such mostly relate to atomic core electronic transitions.

**Ultraviolet** With UV wavelengths running from 10 nm up to the violet end of the visible range at around 380 nm, this region is commonly divided at a wavelength of 200 nm into the 'far-UV' region (wavelengths below 200 nm, often referred to as 'vacuum UV' because oxygen absorbs here) and the near-UV (200–380 nm). With photon energies from 100 eV down to less than 10 eV – the latter being typical of the lowest atomic or molecular ionization

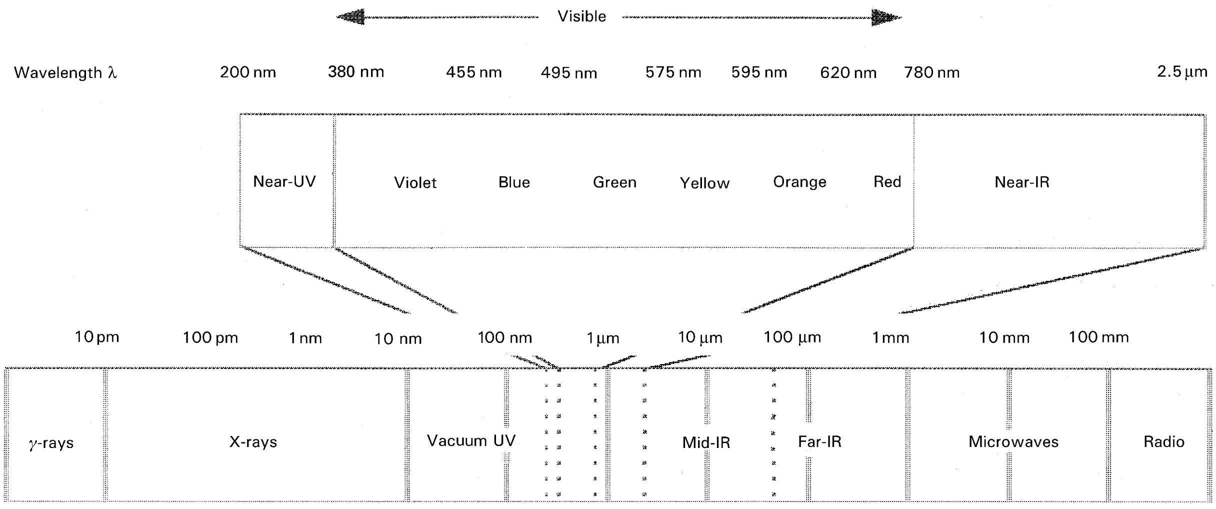



**Figure 3** The electromagnetic spectrum.

energy – photoionization processes are associated with many of the techniques of ultraviolet spectroscopy. At the lower end of this region, photon energies are comparable with valence bond energies. As such they are commonly scaled by Avogadro's number and reported in units of kJ $mol^{-1}$; for example, the wavelength of 240 nm corresponds to 500 kJ $mol^{-1}$, a typical bond energy. One other form of division often applied to the near-UV region relates to its photobiological effects and applies principally to solar radiation; UV-A (320–380 nm) is the relatively safe region closest to the visible range; UV-B (280–320 nm) signifies radiation that can produce extensive tissue damage; UV-C radiation with wavelengths below 280 nm is potentially more damaging but is mostly filtered out by atmospheric gases that absorb here, the most important being ozone.

**Visible** The visible range extends from approximately 380 nm (violet) to 780 nm (red) (**Table 1**). The precise divisions are a little arbitrary and depend on individual perceptions, but the wavelengths given in **Table 1** are a reasonable guide. Note that the link with perceived colour is not 1:1 – for example, light containing an equal mixture of red and green wavelengths appears yellow, although no yellow wavelengths are present. The visible range spans near enough an octave of frequencies and is a region in which photon energies are comparable with the bond energies of some of the weaker chemical bonds, running up from around 150 kJ $mol^{-1}$ at the red end to over 300 kJ $mol^{-1}$ at the other. Most of the spectroscopy in this range of wavelengths relates to electronic transitions unaccompanied by chemical change, and of course all absorption or fluorescence processes responsible for colour.

**Table 1** The visible spectrum

| *Colour* | *Wavelengths* (nm) |
|---|---|
| Red | 780–620 |
| Orange | 620–595 |
| Yellow | 595–575 |
| Green | 575–495 |
| Blue | 495–455 |
| Violet | 455–380 |

**Infrared** This is another region commonly subdivided, in this case into the near IR (wavelengths running from the red end of the visible spectrum at 780 nm out to 2.5 μm), the mid-IR (from 2.5 to 50 μm) and the far IR (50 μm to 1 mm). The absorption of near-IR radiation is commonly associated with the excitation of low-lying electronic excited states and overtones or combinations of molecular vibrations. In spectroscopic connections, the unqualified term 'infrared' generally refers to the mid-range, where spectral positions are usually cited by reference to wave-numbers $\bar{\nu} = 1/\lambda$. In the mid-IR range between 200 and 4000 $cm^{-1}$, it is principally vibrational transitions that accompany the absorption of radiation, with the corresponding nuclear motions less localized at the lower-wavenumber end of the scale. The far-IR region beyond relates mostly to low-frequency molecular vibrations and inversions, or rotations in small molecules.

**Microwave** With wavelengths in the 1–100 mm range and frequencies on the GHz ($10^9$ Hz) scale, spectroscopy in the microwave region relates

primarily to transitions involving molecular rotation, or others involving states of different electron spin orientation (electron spin resonance).

**Radio** With radiation of wavelengths exceeding 100 mm we finally run into the radio wave region, where frequencies are commonly reported in MHz ($10^6$ Hz). Photon energies here are too small to lead to transitions associated with electronic or nuclear movement, but they can produce transitions between spin states (nuclear magnetic resonance).

As a caveat by way of conclusion, it should be pointed out that the nature of transitions studied by a particular spectroscopic technique is not always so obviously linked with a particular optical frequency or wavelength region if the elementary interaction involved in the spectroscopy entails more than one photon. For example, Raman spectroscopy allows molecular vibrational transitions to be studied with visible or ultraviolet light, while multiphoton absorption of IR radiation can lead to electronic excitations.

## List of symbols

$B$ = magnetic field vector; $B_0$ = magnetic field amplitude; $c$ = speed of light; $c_0$ = speed of light *in vacuo*; $E$ = photon energy; $E$ = electric field vector; $E_0$ = electric field amplitude; $h$ = Planck constant; $\hbar$ = Planck constant/$2\pi$; $k$ = wave vector; $n_\lambda$ = refractive index at wavelength $\lambda$; $p$ = photon momentum; $S$ = spin; $t$ = time; $\lambda$ = wavelength; $\nu$ = frequency; $\bar{\nu}$ = wavenumber; $\omega$ = circular frequency.

*See also:* **Light Sources and Optics.**

## Further reading

Ditchburn RW (1976) *Light*, pp 631–640. London: Academic Press.

Fishbane PM, Gasiorowicz S and Thornton ST (1993) *Physics for Scientists and Engineers*, extended version, pp 1009–1047. New Jersey: Prentice-Hall.

Goldin E (1982) *Waves and Photons*, pp 117–134. New York: Wiley.

Hakfoort C (1988) Newton's optics: the changing spectrum of science. In: Fauvel P, Flood R, Shortland M and Wilson R (eds) *Let Newton Be!*, pp 81–99. Oxford: Oxford University Press.

Sheppard N (1990) Chemical applications of molecular spectroscopy – a developing perspective. In: Andrews DL (ed) *Perspectives in Modern Chemical Spectroscopy*, pp 1–41. Berlin: Springer-Verlag.

# Electron Magnetic Resonance Imaging

*See* **EPR Imaging.**

# Electronic Components, Applications of Atomic Spectroscopy

**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

| ATOMIC SPECTROSCOPY |
| --- |
| Applications |

Atomic absorption spectroscopy and atomic emission spectroscopy have found application in many areas of materials science. The electronics industry requires materials of high purity and hence there is a need to monitor trace impurity levels in materials used for electronic components. As a consequence, various techniques of atomic spectroscopy have been applied to the analysis of materials used in the manufacture of electronic components. A number of recent studies are listed in the Further reading section and these therefore provide an entry point to the scientific literature. The reader should also consult other articles in this encyclopedia for related topics.

For example, the analysis of electronic grade silicon has been studied using atomic emission spectroscopy from inductively coupled plasma and the level of aluminium and phosphorus in such material has also been determined using an indirect atomic

absorption method. The growth of europium on palladium surfaces has also been investigated using atomic emission spectroscopy.

The use of inductively coupled plasma–atomic emission spectroscopy (ICP-AES) has been applied to the analysis of traces of certain rare earth metals in highly pure rare earth matrices and to the levels of approximately 25 impurities in sintered electronic ceramics.

The impurity levels and profiles from the surface into the bulk sample have been probed using a combination of secondary-ion mass spectrometry and electrothermal atomic absorption spectroscopy in composite materials such as CdZnTe.

*See also:* **Atomic Absorption, Methods and Instrumentation; Atomic Absorption, Theory; Atomic Emission, Methods and Instrumentation; Atomic Spectroscopy, Historical Perspective; Inductively Coupled Plasma Mass Spectrometry, Methods; Materials Science Applications of X-Ray Diffraction; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Bertran F, Gourieux T, Krill G, Alnot M, Ehrhardt JJ and Felsch W (1991) Growth of Eu on Pd(111) – AES, photoemission and RHEED studies. *Surface Science* **245**: L163–L169.

Daskalova N, Velichkov S, Krasnobaeva N and Slavova P (1992) Spectral interferences in the determination of traces of scandium, yttrium and rare earth elements in pure rare earth matrices by inductively coupled plasma atomic emission spectrometry. *Spectrochimica Acta* **B47**: E1595–E1620.

Gerardi C, Milella E, Campanella F and Bernadi S (1996) SIMS-ETAAS characterization of background impurities in CdZnTe bulk samples. *Material Science Forum* **203**: 273–278.

Gupta PK and Ramchandran R (1991) Indirect atomic absorption spectrometric determination of phosphorus in high purity electronic grade silicon using bismuth phosphomolybdate complex. *Microchemical Journal* **44**: 34–38.

Morvan D, Amouroux J and Claraz P (1984) Analysis of electronic and solar grade silicon by atomic emission spectroscopy from inductively coupled plasma. *Progress in Crystal Growth and Characterization* **8**: 175–180.

Uwamino Y, Morikawa H, Tsuge A, Nakane K, lida Y and Ishizuka T (1994) Determination of impurities in sintered electronic ceramics by inductively coupled plasma atomic emission spectrometry. *Microchemical Journal* **49**: 173–182.

# Ellipsometry

**GE Jellison Jr**, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Applications**

Ellipsometry is a technique often used to measure the thickness of a thin film. Generally speaking, the measurement is performed by polarizing the incident light beam, reflecting it off a smooth sample surface at a large oblique angle and then re-polarizing the light beam before the intensity of the light beam is measured. Since the process of reflecting light off a smooth sample surface generally changes linearly polarized light into elliptically polarized light, the technique has been called ‘ellipsometry’.

The earliest ellipsometry measurements (ca 1890) were used to determine the optical functions (refractive index $n$ and extinction coefficient $k$, or equivalently, absorption coefficient $\alpha$) for several materials. In the 1940s it was discovered that a single-wavelength nulling ellipsometry measurement could be used to determine the thickness of certain thin films very accurately. Since that time, single-wavelength ellipsometry has evolved to be the standard of thickness measurement for several industries, including the semiconductor industry.

Ellipsometry experiments produce values that are not useful by themselves: computers must be used to obtain useful quantities such as thin-film thickness or the optical functions of materials. The advent of modern computers has resulted in the invention of several spectroscopic ellipsometers and the creation of more realistic analysis programs required to understand spectroscopic ellipsometry data.

Any ellipsometer (see **Figure 1**) consists of five elements: (1) a light source, (2) a polarization state generator (PSG), (3) a sample, (4) a polarization state detector (PSD), and (5) a light detector. The light source can be a monochromatic source, such as from a laser, or a white light source, such as from a xenon or mercury arc lamp. The PSG and

PSD are optical instruments that change the polarization state of a light beam passing through them and they contain optical elements such as polarizers, retarders and photoelastic modulators. In most ellipsometry experiments, light from the PSG is reflected from the sample surface at a large angle of incidence $\phi$. Spectroscopic ellipsometers use a white light source and a monochromator (either before the PSG or after the PSD) to select out specific wavelengths. Some spectroscopic ellipsometers image the white light from the PSD onto a detector array, thereby allowing the whole spectrum to be collected simultaneously.

Any ellipsometer will only measure characteristics of the light reflected from or light transmitted through the sample. Ellipsometers do not measure film thicknesses or optical functions of materials, although these parameters can often be inferred very accurately from the ellipsometry measurements. Data analysis is an essential part of any ellipsometry experiment.

## Polarization optics

Several optical elements will change the polarization state of a light beam interacting with them. Linear polarizers (see **Figure 2A**) transmit only one polarization of the light beam; the azimuthal angle of the polarizer will determine the azimuthal angle of the light polarization. The best linear polarizers in the visible part of the spectrum are prism polarizers made from birefringent crystals such as calcite, magnesium fluoride or quartz, which have refractive indices that depend upon the direction in the material.

Retarders or compensators (see **Figure 2B**) are another common type of optical element used in ellipsometers. In combination with a polarizer, a retarder can produce circularly or elliptically polarized light if the direction of the linear polarization is not parallel to a major axis in the retarder. Static retarders, such as quarter-wave or half-wave plates produces a fixed amount of retardation at a fixed wavelength. Circularly polarized light can be produced from a polarizer–quarter-wave retarder pair if the light polarization is ±45° with respect to the fast axis of a retarder.

Another common retarding element used in spectroscopic ellipsometers is the photoelastic modulator (PEM), shown schematically in **Figure 2C**. The drive element is a crystal quartz bar, which is cut so that an a.c. voltage $V$ applied to the front and back faces causes the bar to vibrate along its long axis at the frequency of $V$. The optical element is a bar made of fused quartz that is placed in intimate contact with the crystal quartz bar so that it vibrates at the same frequency. Both bars are sized so that the vibrations are resonant, making the frequency of the PEM extremely stable. A polarizer–PEM pair produces dynamically elliptically polarized light if the polarizer

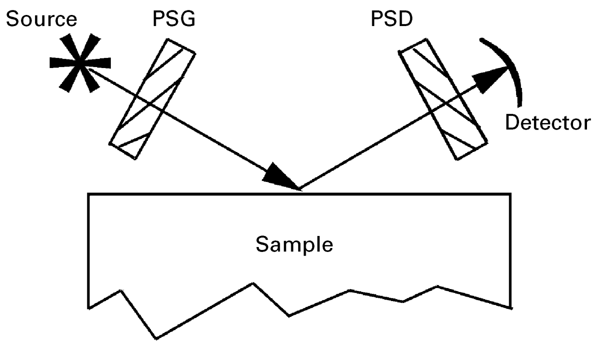



**Figure 1** Schematic diagram of an ellipsometer. The PSG is the polarization state generator and the PSD is the polarization state detector.

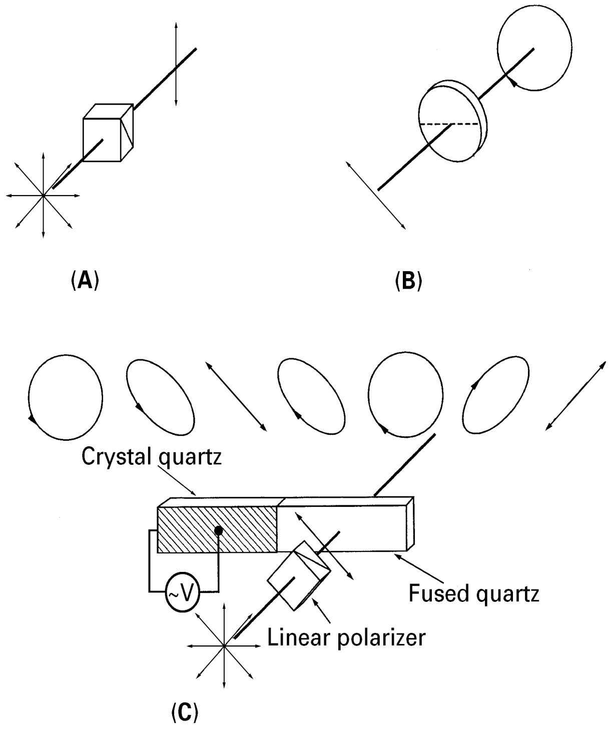



**Figure 2** Several optical elements used in ellipsometers that alter the polarization state of a light beam passing through them. (A) A linear polarizer, which transforms any light beam (polarized or unpolarized) to linearly polarized light. (B) A retarder, which transforms linearly polarized light to elliptically polarized light. In certain cases, the light can be circularly polarized. (C) A photoelastic modulator, which changes linearly polarized light into dynamically elliptically polarized light.

is not lined up with the vibration axis of the PEM; generally, the polarization state of the emergent light beam will cycle through linearly polarized, circularly polarized and elliptically polarized states.

The Stokes vector representation of light polarization has often been used for ellipsometry measurements. (Another representation is the Jones vector representation, of which details can be found in the book by Azzam and Bashara listed under Further reading). In the Stokes representation, the polarization state of a light beam is given by its four-element Stokes vector,

$$\boldsymbol{S} = \begin{bmatrix} I_0 \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} I_0 \\ I_0 - I_{90} \\ I_{45} - I_{-45} \\ I_{\mathrm{rc}} - I_{\mathrm{lc}} \end{bmatrix} \qquad [1]$$

All the elements of the Stokes vector are intensities and therefore are real. The total intensity is $I_0$, while $I_0$, $I_{90}$, $I_{45}$, and $I_{-45}$, are the intensities of linearly polarized light at 0°, 90°, 45°, and −45°, respectively. The fourth element of the Stokes vector $V$ is the difference between the intensities of right circularly polarized light $I_{\mathrm{rc}}$, and left circularly polarized light $I_{\mathrm{lc}}$. A linear polarizer can only transform a light beam into a linear polarization state; therefore the fourth element of the Stokes vector for linearly polarized light will be $V = 0$. Generally, a retarding optical element is needed to transform the polarization state of a light beam to one where $V \neq 0$. The Stokes representation can also represent partially polarized light. In general,

$$I_0 \geq \sqrt{Q^2 + U^2 + V^2} \qquad [2]$$

where the equality holds when the light beam is totally polarized.

Each optical element must transform the polarization state from one four-element vector to another four-element vector; therefore, optical elements are represented by 4 × 4 Mueller matrices, where all the elements are real. The intensity of the light beam passing through an ellipsometer can therefore be written using matrix notation as

$$I = \boldsymbol{S}^T_{\mathrm{PSD}}\, \mathrm{M}\, \boldsymbol{S}_{\mathrm{PSG}} \qquad [3]$$

where $S_{\mathrm{PSG}}$ is the Stokes vector representing the polarization state for light coming from the PSG, **M** is the Mueller matrix for the light interaction with the sample, and $S^{\mathrm{T}}_{\mathrm{PSD}}$ is the transpose of the Stokes vector representing the effect of the PSD. In the most general case, 16 elements are required to describe the light interaction with a sample. Fortunately, many of these elements are normally 0 or equal to another element in **M**. Since detectors can only measure the light intensity coming from an instrument, Equation [3] emphasizes the fact that ellipsometers can only measure elements of **M**.

## Ellipsometry data analysis

Ellipsometry measurements are not useful by themselves but can be extremely useful if the measurements are interpreted with an appropriate model. Ultimately, ellipsometry results are always model dependent. Fortunately, the physics of light reflection from surfaces is well understood, and very detailed and accurate models can be made using classical electromagnetic theory based on the Maxwell equations.

**Figure 3** shows a schematic of light reflection from a sample surface. The light beam is incident upon the sample surface at an angle of incidence $\phi$. The specularly reflected beam comes from the sample surface, also at angle $\phi$. The incident and reflected beams define a plane, called the plane of incidence. This in turn defines two polarization directions: p for the light polarization parallel to the plane of incidence (in the plane of the paper), and s for the light polarization perpendicular to the plane of incidence (perpendicular to the plane of the paper). All azimuthal angles are defined with respect to the plane of incidence, where positive rotations are defined as clockwise rotations looking from the light source to the detector.

If the sample surface is isotropic and has no film or other overlayer ($d = 0$ in **Figure 3**), then one can use the Maxwell equations to calculate the complex reflection coefficients:

$$r_{\mathrm{p}} = \frac{N_{\mathrm{s}} \cos(\phi_0) - N_0 \cos(\phi_{\mathrm{s}})}{N_{\mathrm{s}} \cos(\phi_0) + N_0 \cos(\phi_{\mathrm{s}})} \qquad [4a]$$

$$r_{\mathrm{s}} = \frac{N_0 \cos(\phi_0) - N_{\mathrm{s}} \cos(\phi_{\mathrm{s}})}{N_0 \cos(\phi_0) + N_{\mathrm{s}} \cos(\phi_{\mathrm{s}})} \qquad [4b]$$

In Equations [4], $N_0$ and $N_{\mathrm{s}}$ ($= n_{\mathrm{s}} + \mathrm{i}k_{\mathrm{s}}$, where $n_{\mathrm{s}}$ is the refractive index and $k_{\mathrm{s}}$ is the extinction coefficient) are the complex indices of refraction for the ambient (usually air, where $N_0 = 1$) and the

substrate, respectively. The quantities $\phi_0$ and $\phi_s$ are complex angles, determined from the Snell law [$N_0$ sin ($\phi_0$) = $N_s$ sin ($\phi_s$)]. The reflection ratios $r_p$ and $r_s$ are complex, indicating that light reflected from a surface will generally undergo a phase shift.

If the sample near-surface region consists of a single film (see **Figure 3**), the composite reflection coefficients can be calculated from the Airy formula:

$$r_{s,p} = \frac{r_{1s,p} + r_{2s,p}e^{-2ib}}{1 + r_{1s,p}\, r_{2s,p}\, e^{-2ib}} \qquad b = \frac{2\pi d N_f \cos(\phi_f)}{\lambda} \qquad [5]$$

where $d$ is the thickness of the film, $N_f$ is the complex refractive index of the film, and $\phi_f$ is the complex angle within the film defined by the Snell law. The quantities $r_{1s,p}$ and $r_{2s,p}$ are the complex reflection coefficients calculated using Equations [4a] and [4b] for the air–film interface and the film–substrate interface, respectively.

Reflections from more complicated layer structures can be calculated using matrix methods. If all the media in the calculation are isotropic, the matrix formulation of Abelés (using $2 \times 2$ complex matrices) can be used to calculate the composite $r_s$ and $r_p$.

If any of the media are birefringent, many of the implicit assumptions made above are no longer valid. For example, for isotropic media the s and p polarization states represent eigenmodes of the reflection; that is, if the incoming light is pure s or p polarized, then the reflected light will be pure s or p polarized. If any of the media in the sample are birefringent, this assumption is no longer generally valid. As a result, one must also calculate the cross-polarization coefficients $r_{sp}$ and $r_{ps}$ as well as $r_{ss} \equiv r_s$ and $r_{pp} \equiv r_p$. If $r_{sp} \neq 0$ or $r_{ps} \neq 0$, the $2 \times 2$ matrix methods must be replaced by more complicated $4 \times 4$ matrix methods.

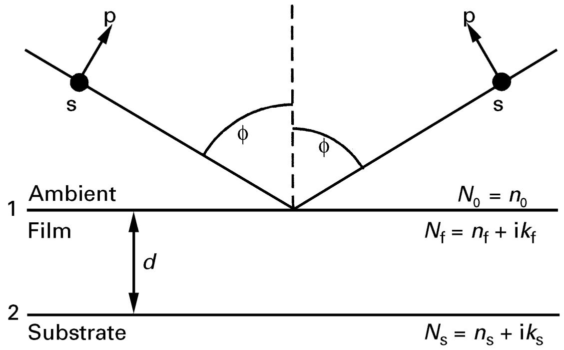



**Figure 3** A schematic diagram of light reflecting from a sample surface. The s and p vectors indicate the direction of s and p polarized light, as defined by the plane of incidence. The angle of incidence is $\phi$, and $N_0$, $N_f$ and $N_s$ are the complex refractive indices of the ambient, film and substrate, respectively.

Up to this point, a clear distinction has been drawn between the parameters measured by an ellipsometer (the sample Mueller matrix **M**) and the calculated parameters obtained from classical electromagnetic theory, which can be expressed as elements of the sample Jones matrix:

$$\mathbf{J} = \begin{bmatrix} r_{pp} & r_{ps} \\ r_{sp} & t_{ss} \end{bmatrix} \qquad [6]$$

Fortunately, the two representations are related, so long as it can be assumed that the sample does not depolarize the incident light beam. In this case, a Mueller–Jones matrix $\mathbf{M}_{MJ}$ can be calculated from the elements of the Jones matrix (Eqn [6]),

$$\mathbf{M}_{MJ} = \mathbf{A}\,(\mathbf{J} \otimes \mathbf{J}^*)\,\mathbf{A}^{-1} \qquad [7a]$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix} \qquad [7b]$$

For isotropic samples ($r_{sp} = r_{ps} = 0$), the normalized sample Mueller–Jones matrix is given by

$$\mathbf{M} = \mathbf{M}_{MJ} = \begin{bmatrix} 1 & -N & 0 & 0 \\ -N & 1 & 0 & 0 \\ 0 & 0 & C & S \\ 0 & 0 & -S & C \end{bmatrix} \qquad [8a]$$

where

$$N = \cos(2\psi) \qquad [8b]$$

$$S = \sin(2\psi)\sin(\Delta) \qquad [8c]$$

$$C = \sin(2\psi)\cos(\Delta) \qquad [8d]$$

$$\rho = \frac{r_p}{r_s} = \tan(\psi)e^{i\Delta} = \frac{C + iS}{1 + N} = \rho_r + i\rho_i \qquad [8e]$$

The angles $\psi$ and $\Delta$ are the traditional ellipsometry angles, and are the naturally measured parameters for nulling ellipsometers (see below). The matrix **M** simplifies considerably when the sample is isotropic: **M** is block-diagonal, with eight elements equal to 0, and

only two parameters (such as $\rho = \rho_r + i\,\rho_i$ or $\psi$ and $\Delta$) are needed to specify **M**, since $N^2 + S^2 + C^2 = 1$.

For birefringent samples, $r_{sp} \neq r_{ps} \neq 0$, and the off-block diagonal elements of **M** are no longer 0, nor are the on-block diagonal elements of **M** so simply defined. However, Equation [7a] is still valid as long as the sample is nondepolarizing. In this case, six parameters are required to specify all 16 elements of the normalized sample Mueller matrix, given by Equation [8e] and

$$\rho_{ps} = \frac{r_{ps}}{r_{ss}} = \tan(\psi_{ps})e^{i\Delta_{ps}} = \frac{C_{ps} + iS_{ps}}{1+N} \qquad [8f]$$

$$\rho_{sp} = \frac{r_{sp}}{r_{ss}} = \tan(\psi_{sp})e^{i\Delta_{sp}} = \frac{C_{sp} + iS_{sp}}{1+N} \qquad [8g]$$

Up to now, we have assumed that no optical element depolarizes the light. Usually, this is true, but there are some cases where depolarization must be considered. (1) If the input light beam illuminates an area of the sample where the film thickness(es) is (are) not uniform, quasi-depolarization can occur. (2) If the sample substrate is transparent, then light reflecting from the back surface will contribute an intensity component to the light beam entering the PSD that is not phase-related to the light reflected from the front face and the light beam will be quasi-depolarized. (3) If the sample is very rough, then some of the light reaching the PSD will not have an identifiable polarization state or cross-polarization can occur in nominally isotropic systems. All of these effects invalidate the Mueller–Jones matrix representation of the sample surface shown in Equation [7a].

## Types of ellipsometers

There are many different kinds of ellipsometers, some of which are shown schematically in **Figure 4**. All ellipsometers measure one or more quantities that can be related to the complex reflection coefficient ratio $\rho$ shown in Equations [8].

### Nulling ellipsometer

The oldest and most common type of ellipsometer is the nulling ellipsometer, shown schematically in **Figure 4A**. The light source for a modern nulling ellipsometer is usually a small laser, but other monochromatic sources can also be used. The PSG is a polarizer–retarder pair, and the PSD is a linear polarizer, historically called the analyser. If the quarter-wave plate is oriented at 45° with respect to the plane of incidence, the intensity of the light beam incident upon the detector is given by

$$I = I_0[1 - N\cos(2\theta_a) + \sin(2\theta_a)(C\sin(2\theta_p) + S\cos(2\theta_p))] \qquad [9]$$

where $\theta_p$ and $\theta_a$ are the azimuthal angles of the polarizer and analyser, respectively, and $N$, $S$ and $C$ are the quantities given in Equations [8]. The intensity given in Equation [9] has a null at two sets of angles: $(2\theta_p, \theta_a) = (270° - \Delta, \psi)$, $(90° - \Delta, 180° - \psi)$. Therefore, nulling ellipsometry measurements are made by rotating the PSG polarizer and the PSD polarizer (analyser) until the light intensity reaching the detector is a minimum. The angles $\psi$ and $\Delta$ are determined from the azimuthal angles of the PSG and PSD polarizers.

The nulling ellipsometer is one of the simplest ellipsometers and it is capable of very accurate measurements with proper calibration. However, it normally uses a single-wavelength source and it is not easily made spectroscopic. Moreover, measurement times are slow, so this type of instrument is not useful for spectroscopic ellipsometry or for fast time-resolved measurements.

### Rotating analyser ellipsometer (RAE)

The most common type of spectroscopic ellipsometer is the rotating analyser (or rotating polarizer) ellipsometer, shown schematically in **Figure 4B**. The PSG is a linear polarizer, as is the PSD, and one of the optical elements (the analyser in **Figure 4B**) is physically rotated, making the light intensity at the detector a periodic function of time. For an analyser (PSD) rotating at an angular frequency of $\omega$ and with the polarizer (PSG) set at an azimuthal angle of $\theta_p$, the intensity is given by

$$I(t) = I_0[1 + \alpha\cos(2\omega t + \delta_c) + \beta\sin(2\omega t + \delta_c)] \qquad [10a]$$

where

$$\alpha = \frac{\cos(2\theta_p) - N}{1 - N\cos(2\theta_p)} \qquad [10b]$$

$$\beta = \frac{C\sin(2\theta_p)}{1 - N\cos(2\theta_p)} \qquad [10c]$$

The quantity $\delta_c$ is a constant phase shift. The coefficients $\alpha$ and $\beta$ can be determined using demodulation techniques or Fourier analysis. Normally, $\theta_p = 45°$, where $\alpha = -N$, and $\beta = C$.

In its normal form, the RAE includes only polarizers, although a retarder can be added to either the PSG or the PSD. If only polarizers are used, this instrument is very easy to make spectroscopic, since no optical component other than the sample is wavelength dependent. Furthermore, the measurement time depends principally on the rotation speed of the rotating element. For a rotation speed of 100 Hz, measurements can be completed in 40 ms. In some configurations of the rotating polarizer ellipsometer, a spectrograph is placed after the PSD and the light is detected by a photodiode array. This ellipsometer can collect an entire spectrum with a single series of measurements, typically taking about 1s to collect and analyse each spectrum.

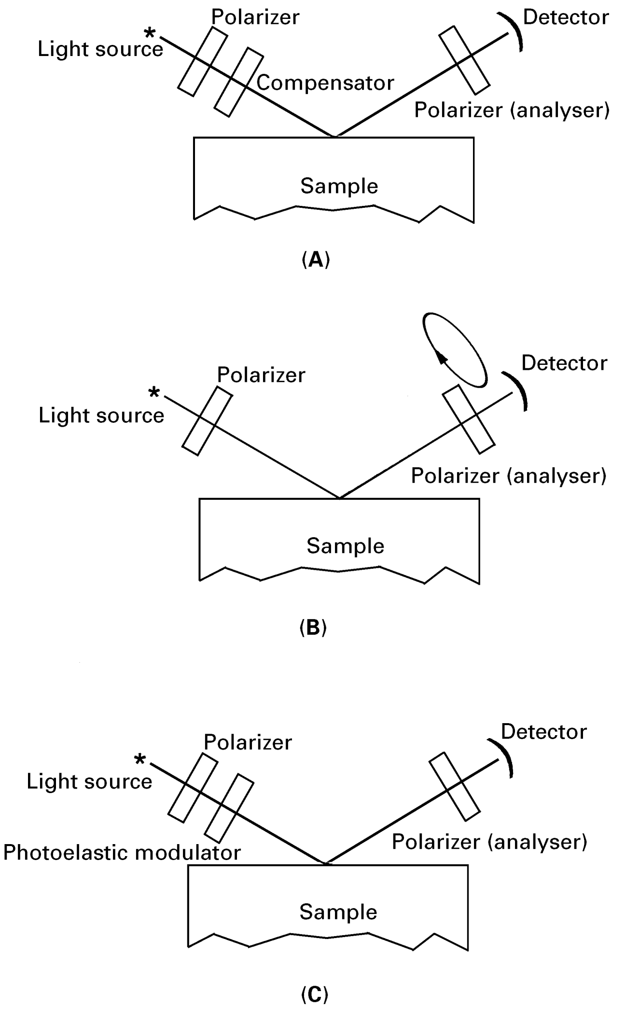



**Figure 4** Schematic diagrams of three ellipsometers in common use. (A) The nulling ellipsometer, which normally uses monochromatic light. (B) The rotating analyser ellipsometer. (C) The polarization modulation ellipsometer.

The RAE measures two quantities, $N$ and a linear combination of $C$ and $S$. If there is no retarder, the RAE measures $N$ and $C$, and will therefore be subject to large errors in $\Delta$ if $\Delta$ is near 0° or 180° and an uncertainty of the sign of $\Delta$. The use of a quarter-wave retarder shifts the inaccuracy in $\Delta$ to the regions near ±90°.

### Photoelastic modulator ellipsometer (PME)

Another common spectroscopic ellipsometer is based on the photoelastic modulator (PEM), shown in **Figure 4C**. Normally, the PSG consists of a polarizer–PEM pair, where the polarizer is oriented at 45° with respect to the vibration axis of the PEM (see **Figure 2C**). The intensity of the light arriving at the detector is a function of time given by

$$I(t) = I_0[I_{dc} + I_X \sin(A\sin(\omega t)) + I_Y \cos(A\sin(\omega t))] \quad [11a]$$

where

$$I_{dc} = 1 - N\cos(2\theta_a) \quad [11b]$$

$$I_X = S\sin(2\theta_a) \quad [11c]$$

$$I_Y = \sin(2\theta_m)\ [\cos(2\theta_m) - N] - C\cos(2\theta_m)\sin(2\theta_a) \quad [11d]$$

The quantity $A$ is the Bessel angle of the PEM modulation, and is proportional to the amplitude of modulation; normally, this is set at 2.4048 radians. The frequency of the PEM is denoted by $2\pi\omega$, and is normally ~50 kHz. The quantities $\theta_m$ and $\theta_a$ are the azimuthal angles of the modulator and the PSD polarizer, respectively.

The time dependence of the light intensity from the PME is considerably more complicated than that from the RAE, since the basis functions are $\sin(A\sin(\omega t))$ and $\cos(A\sin(\omega t))$, rather than $\sin(\omega t)$ and $\cos(\omega t)$ as for the RAE. However, these functions can be expressed in terms of a Fourier–Bessel infinite series:

$$\sin(A\sin(\omega t)) = 2\sum_{k=1}^{\infty} J_{2k-1}(A)\sin((2k-1)\omega t) \quad [12a]$$

$$\cos(A(\sin(\omega t)) = J_0(A) + 2\sum_{k=1}^{\infty} J_{2k}(A)\sin(2k\omega t) \quad [12b]$$

Therefore, a demodulation or Fourier analysis of the waveform of Equation [11a] can also be used to get the coefficients $I_X$ and $I_Y$.

The PME is more complicated than the RAE, since the modulation amplitude of the PEM must be calibrated and its small but significant static retardation measured at all wavelengths used. Furthermore, the frequency of the PEM is about 500 times faster than the rotation speed of the RAE; this improves the time resolution to ~1 ms, but requires faster electronics for data collection. In its normal configuration discussed above, the PME can measure *S* and either *N* or *C*. However, if the PSD polarizer is replaced with a Wollaston prism polarizer and both channels are detected (measuring a total of four quantities), then it is possible to measure *N*, *S*, and *C* simultaneously. This implementation is called the two-channel spectroscopic polarization modulation ellipsometer (2C-SPME).

### Other configurations

The ellipsometer configurations discussed above are particularly useful if the sample is isotropic. If the sample is anisotropic, then many of the simplifications used above are not valid, and many more measurements must usually be made, since the normalized sample Mueller matrix contains six independent quantities. One such implementation is the two-modulator generalized ellipsometer (2-MGE) which uses two PEMs operating at different frequencies. The 2-MGE is spectroscopic and is capable of measuring all six independent elements of the sample Mueller matrix simultaneously. Other configurations are discussed in the review article by Hauge.

## Application: Silicon dioxide films on silicon

One of the most useful applications of ellipsometry has been the routine measurement of silicon dioxide ($SiO_2$) film thicknesses grown on silicon. Since the refractive index of $SiO_2$ is well known, and does not depend significantly on film deposition technique, nulling ellipsometry measurements are usually sufficient to determine film thickness.

**Figure 5** shows a plot of $\psi$ versus $\Delta$ for thin-film $SiO_2$ grown on silicon for a wavelength of 633 nm (the wavelength of a HeNe laser). The angle of incidence used for the calculation was 70°, the film refractive index $n_f = 1.46$, and the complex refractive index of silicon $N_s = n_s + ik_s = 3.86 + i\,0.018$. This plot shows the trajectory that $\psi$–$\Delta$ follows as the film thickness increases from zero to ~280 nm, where the position along the trajectory for several thicknesses is noted in the figure. For zero film thickness, the value of $\Delta$ is very close to 0° or 180°, owing to the very small value of $k_s$. If other film refractive indices were included in this plot, it would be seen that they would converge to this point. It is therefore very difficult to determine the refractive index and the thickness of a very thin film from ellipsometry measurements. Note that the curve comes around on itself near 283 nm. Therefore, any ellipsometry measurement of $SiO_2$/Si can only be used to determine the film thickness modulo the repeat thickness (283 nm). For film thicknesses from 140 to 160 nm, the measurements are very sensitive to film thickness, crossing the $\Delta = 180°$ point for a thickness of 141.5 nm (half the repeat thickness).

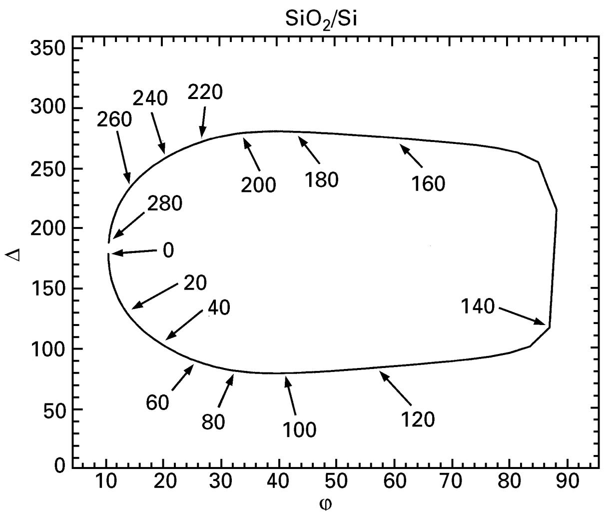



**Figure 5** The $\psi$–$\Delta$ trajectory for a thin film of $SiO_2$ on silicon. The angle of incidence $\phi = 70°$, $n_f = 1.46$, $N_s = 3.86 + i0.018$, and the wavelength $\lambda = 633$ nm (HeNe laser).

The same data are plotted in a different manner in **Figure 6**. Here, the complex $\rho$ (= $\tan(\psi)\,e^{i\Delta}$, see Equation [8e]) is plotted versus film thickness. Near zero thickness, Im($\rho$) (which is proportional to sin($\Delta$)) is very small. As the thickness approaches 140 nm, –Re($\rho$) gets very large and Im($\rho$) changes sign (Im($\rho$) = 0 at half the repeat thickness); this corresponds to the region of large $\psi$ in **Figure 5**, with the change in the sign of Im($\rho$) corresponding to $\Delta$ going through 180°. Again, the value of $\rho$ repeats itself after 283 nm.

**Figures 7A–C** show the measured values of of $\rho$ for three thicknesses of $SiO_2$ films grown on silicon

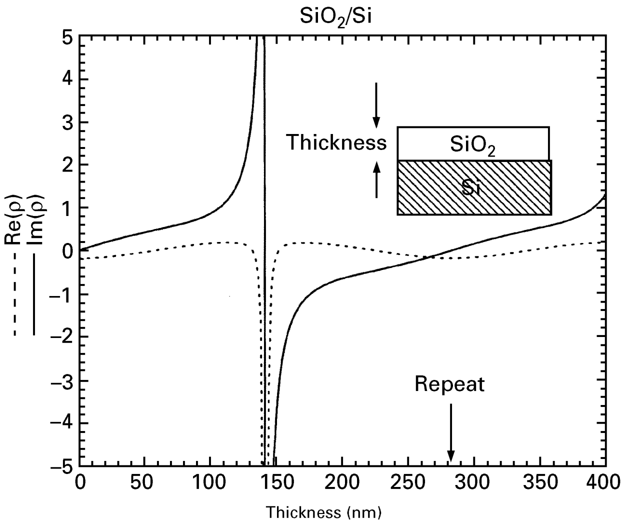



**Figure 6** The complex reflection ratio $\rho$ for the system described in **Figure 5**.

(66 nm, 190 nm and 321 nm) at an angle of incidence of 65°. There are several points where $\mathrm{Im}(\rho) = 0$ and where $-\mathrm{Re}(\rho)$ is a maximum: ~335 nm for the 66 nm film; ~305 nm for the 190 nm film; and ~315 nm and ~505 nm for the 321 nm film. Using spectroscopic ellipsometry for this system is similar to sampling several points along the $\psi$–$\Delta$ trajectory shown in **Figure 5**, although the precise values of the null points of $\mathrm{Im}(\rho)$ will not scale, since $N_f$ and $N_s$ are both functions of wavelength.

## Spectroscopic ellipsometry data analysis

It is easy to see that there are more than 100 data points in each of **Figures 7A–C**, but only a few parameters to be determined from the data; the problem is overdetermined. Therefore, fitting procedures must be used to determine film thicknesses and other properties of thin films from spectroscopic ellipsometry. This process has three steps: (1) model the near-surface region of the sample; (2) parametrize or determine the optical functions of each layer; (3) fit the data using a realistic figure of merit to measure the 'goodness of fit'. For the case of $SiO_2$ grown on silicon, a four-media model is used, consisting of air/$SiO_2$/interface/silicon. The optical functions of the $SiO_2$ layer are assumed to follow the Sellmeier approximation:

$$n^2 = 1 + \frac{A\lambda^2}{\lambda^2 - \lambda_0^2} \qquad [13]$$

**Table 1** The results of the fitting procedure on the spectroscopic ellipsometry data taken on several thin-film $SiO_2$/Si samples, shown in **Figures 7A–C**. The refractive index is calculated from Equation [13] assuming $\lambda_0 = 92.3$ nm

| *Sample* | *Film thickness (nm)* | *A* | *Refractive index (600 nm)* | *Interface thickness (nm)* | $\chi^2$ |
|---|---|---|---|---|---|
| a | 66.4 ± 0.1 | 1.149 ± 0.002 | 1.475 | 0.0 ± 0.2 | 0.18 |
| b | 188.9 ± 0.3 | 1.115 ± 0.003 | 1.464 | 1.4 ± 0.2 | 1.21 |
| c | 321.5 ± 0.4 | 1.139 ± 0.003 | 1.472 | 1.0 ± 0.2 | 0.84 |
| Fused silica | – | 1.099 | 1.458 | – | – |

where the two possible fitting parameters are the amplitude $A$ and the resonance wavelength $\lambda_0$. The interface is assumed to be a thin layer whose optical functions are those of a Bruggeman effective medium, consisting of 50% $SiO_2$ and 50% Si, and the optical functions of silicon are taken from the literature. Using this model, three parameters are determined in the fitting procedure (listed in **Table 1**): the film thickness $d_f$, $A$ from Equation [12] and the thickness of the interface $d_{\text{interface}}$; the resonant wavelength $\lambda_0 = 92.3$ nm. The third step of the analysis involves the actual fitting of the calculated values of ρ with the measured values of ρ. This can be done using a Levenberg–Marquardt algorithm to solve for the fitted parameters $d_{\text{interface}}$, $d_f$ and $A$ using the reduced $\chi^2$ as the figure of merit:

$$\chi^2 = \frac{1}{N - m - 1} \sum_{i=1}^{N} \frac{(\rho_{\text{exp}}(\lambda_i) - \rho_{\text{calc}}(\lambda_i, z))^2}{\delta\rho_{\text{exp}}^2(\lambda_i)} \qquad [14]$$

The sum in Equation [13] is taken over $N$ wavelength points ($\lambda_i$), $z$ is a vector of the fitted parameters (in this case, $d_{\text{interface}}$, $d_f$ and $A$), and $m$ is the dimensionality of $z$ (i.e. 3). The quantities in the denominator ($\delta\rho_{\text{exp}}^2(\lambda_i)$) are the errors in the experimental data. The $\chi^2$ is a good metric for the 'goodness of fit' to ellipsometry data. If $\chi^2$ is near 1 the fit is a good fit, but if it is much greater than 1 the model does not fit the data. After the parameters are determined, it is also necessary to calculate the error limits as well as the correlation coefficients of all the fitted parameters.

The results of the fits to the data presented in **Figures 7A–C** are shown in **Table 1**, as well as the errors of each of the fitted parameters and the final $\chi^2$ of the fit. All $\chi^2$ values are near 1, indicating that the model fits the data. The errors listed for the thickness are all small, showing that the film thickness is a very well determined parameter for these data sets. The parameter $A$ is a measure of the $SiO_2$ refractive index,

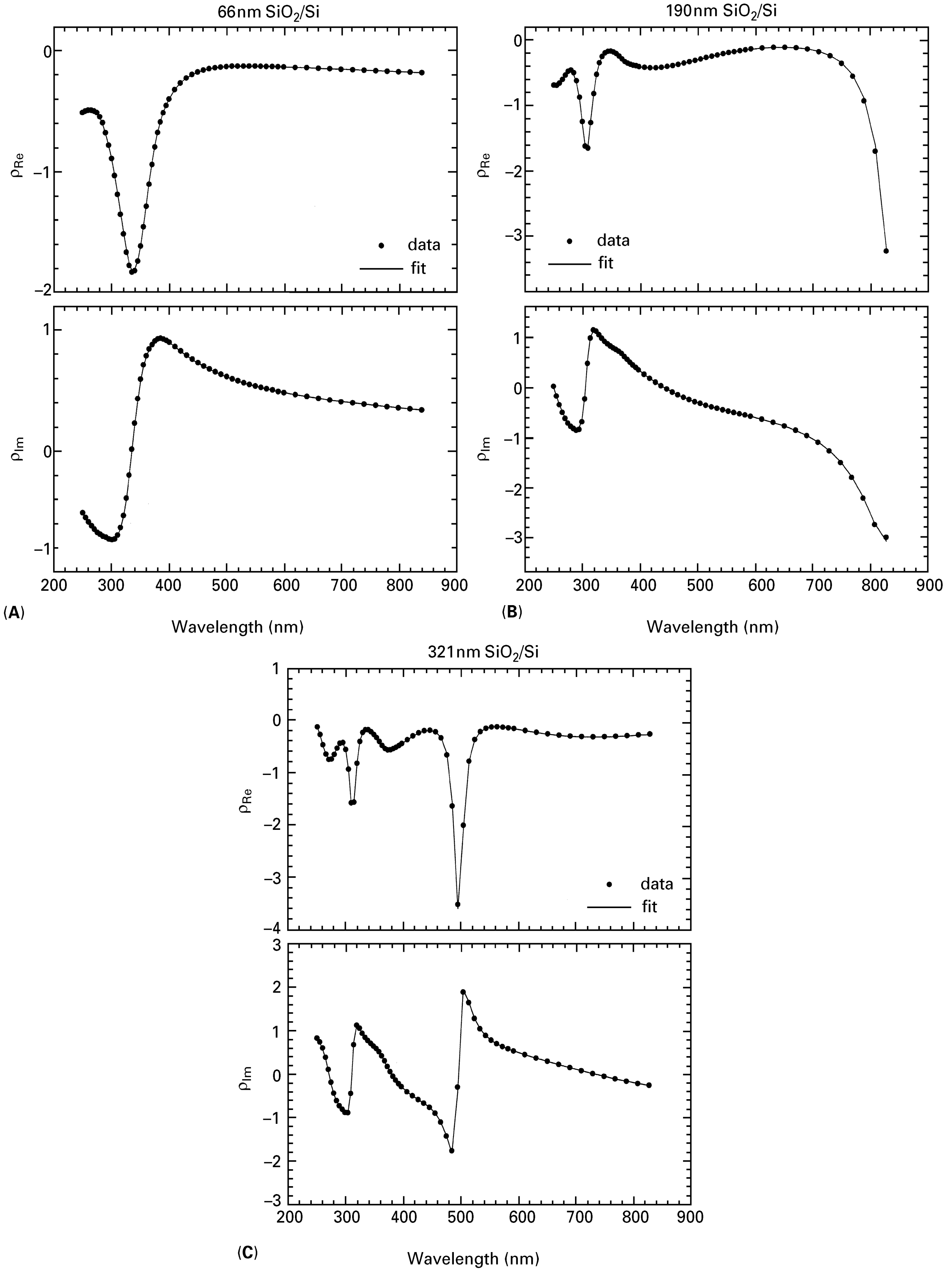



**Figure 7** The spectroscopic values of the complex reflection ratio ρ for three samples of $SiO_2$ films grown on silicon as measured by spectroscopic ellipsometry. The three different thicknesses are (A) 66nm, (B) 190nm, and (C) 321nm.

showing that thin-film $SiO_2$ can have a slightly larger refractive index than bulk fused silica.

## List of symbols

$A$ = Bessel angle of PEM modulation, amplitude; $C = \sin(2\psi)\cos(\Delta)$; $d$ = film thickness; $I_0$ = total intensity element of Stokes vector; $I_n$ = element of Stokes vector, intensity of light linearly polarized at $n$ degrees; $J_n$ = Bessel functions; **J** = Jones matrix; $k$ = extinction coefficient; $m$ = dimensionality of $z$; **M** = Mueller matrix, normalized Mueller–Jones matrix; $\mathbf{M}_{MJ}$ = Mueller–Jones matrix; $n$ = refractive index; $N_0$, $N_f$, $N_s$ = complex refractive indices; $r_p$, $r_s$, etc. = complex reflection coefficients; $S = \sin(2\psi)\sin(\Delta)$; $S$ = Stokes vector; $V$ = voltage, $I_{rc}$–$I_{lc}$ element of Stokes vector; $z$ = vector of fitted parameters (Eqn [14]); $\alpha$ = absorption coefficient; $\delta_c$ = constant phase shift; $\Delta$ = ellipsometry angle; $\theta$ = azimuthal angle; $\lambda$ = wavelength; $\rho$ = complex reflection ratio; $\phi$ = angle of incidence; $\phi_f$, $\phi_s$ = complex angles; $\chi^2$ = figure of merit; $\psi$ = ellipsometry angle; $\omega$ = angular frequency.

*See also:* **Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Fibres and Films Studied Using X-Ray Diffraction; Fluorescence Polarization and Anisotropy; Light Sources and Optics; Linear Dichroism Theory.**

## Further reading

Azzam RMA (1995) *Ellipsometry*, In: Bass M (ed) *Handbook of Optics II*, Chapter 27. New York: McGraw-Hill.

Azzam RMA and Bashara NM (1987) *Ellipsometry and Polarized Light*. Amsterdam: Elsevier Science.

Chipman RA (1995) Polarimetry. In: Bass M (ed) *Handbook of Optics II*, Chapter 22. New York: McGraw-Hill.

Collins RW (1990) Automatic rotating element ellipsometers: calibration, operation, and real-time applications. *Review of Scientific Instruments* **61**: 2029–2062.

Drévillon B (1993) Phase modulated ellipsometry from the ultraviolet to the infrared: *in-situ* application to the growth of semiconductors. *Progress in Crystal Growth and Characterization of Materials* **27**: 1–87.

Hauge PS (1980) Recent developments in instrumentation in ellipsometry. *Surface Science* **96**: 108–140.

Jellison GE Jr (1993) Characterization of the near-surface region using polarization-sensitive optical techniques. In: Exharos GJ (ed) *Characterization of Optical Materials*, pp 27–47. Stoneham, MA: Butterworth-Heinemann.

Jellison GE (1998) Spectroscopic ellipsometry data analysis: measured vs. calculated quantities. *Thin Solid Films* **313–314**: 33–39.

Kliger DS, Lewis JW and Randall CE (1990) *Polarized Light in Optics and Spectroscopy*. New York: Academic Press.

Tompkins HG (1993) *A User's Guide to Ellipsometry*. New York: Academic Press.

# Enantiomeric Purity Studied Using NMR

**Thomas J Wenzel**, Bates College, Lewiston, ME, USA

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

## Introduction

Living organisms have a remarkable ability to distinguish enantiomers. Enzymes and cell surface receptors are handed, such that enantiomers often exhibit different properties in living systems. In extreme cases one enantiomer might be beneficial to a living organism, whereas the other might be harmful. Therefore, it is critical that methods exist to determine the enantiomeric purity and assign absolute configurations of chemical compounds. Various types of chiroptical spectroscopy are used for this. NMR spectroscopy is one of the most common and powerful methods for the determination of enantiomeric excess (typically to 1% of the minor enantiomer) and absolute configurations. There are generally considered to be three broad categories of reagents suitable for chiral resolution in NMR spectroscopy: chiral derivatizing agents, chiral solvating agents, and lanthanide shift reagents.

Chiral derivatizing agents are optically pure compounds that undergo reactions with the enantiomers being analysed. Most derivatization reactions involve the formation of covalent bonds; however, some examples, most notably with carboxylic acids and amines, rely on the formation of soluble salts.

The resulting complexes are diastereoisomers, which may then exhibit different chemical shifts in their NMR spectra. Derivatization agents must react qu antitatively with both enantiomers, with complete retention of configuration of both the substrate and resolving agent or with complete stereospecificity of any perturbations that occur. Useful chiral derivatizing agents usually have a discrete and specific NMR signal (e.g. a methyl group, fluorine, or phosphorus singlet in the $^{1}H$, $^{19}F$, or $^{31}P$ NMR spectrum) that is conveniently monitored for enantiomeric resolution. Most chiral derivatizing agents function with compounds that are readily modified, such as amines, alcohols, and carboxylic acids. In many instances it is possible to assign absolute configurations either by observing trends in the data on similar families of compounds, or through a detailed analysis of the geometric differences that lead to distinctions in the chemical shift values of the diastereoisomers.

Chiral solvating agents are additives that react *in situ* with the compound being analysed. Rather than forming covalent bonds or ion pairs, chiral solvating agents associate with the substrate through combinations of van der Waals forces. Hydrogen bonding, charge-transfer complexation of electron-rich and electron-deficient aromatic rings, and steric effects are important interactions. The change in chemical shifts that occur in the spectrum of a compound in the presence of a chiral solvating agent are denoted as $\Delta\delta$. The extent of enantiomeric resolution that might then be observed for a shifted resonance is denoted as $\Delta\Delta\delta$.

Chiral solvating agents function by two possible mechanisms: Equations [1] and [2] represent the association of a pair of enantiomers with a chiral solvating agent (CSA), in which $K_R$ and $K_S$ denote the association constants for the *R* and *S* enantiomers

$$R + \mathrm{CSA} \rightarrow R\text{-}\mathrm{CSA} \qquad K_R \qquad [1]$$

$$S + \mathrm{CSA} \rightarrow S\text{-}\mathrm{CSA} \qquad K_S \qquad [2]$$

respectively. These reactions should occur rapidly on the NMR time scale such that the spectrum of the substrate is a time average of the bound and unbound form. A typical observation is that $K_R$ and $K_S$ are not equal, so that one enantiomer spends more time free in the bulk solvent than the other. The different time-averaged solvation environments may result in differences in the NMR spectrum. In addition, the two complexes (*R*-CSA and *S*-CSA) are diastereoisomers and therefore can have different chemical shifts irrespective of the relative magnitude of the association constants. In many cases, it is reasonable to assume that both mechanisms partially contribute to the enantiomeric resolution. It is sometimes possible to determine which of the two mechanisms predominates. With certain chiral solvating agents, the mechanism that leads to the enantiodiscrimination is well understood and absolute configurations can be assigned on the basis of the relative shifts.

Enantiomeric resolution with chiral solvating agents is dependent on the relative concentrations of the reagents. Usually, enantiomeric resolution is best when the concentration of solvating agent is greater than or equal to that of the enantiomers. Temperature is also an important variable. Lowering the temperature usually, but not always, increases the association of solvating agent and substrate, thereby increasing the enantiomeric resolution.

Lanthanide shift reagents are in actuality chiral solvating agents; however, their behaviour is in some ways unique and they are typically considered a separate category of chiral resolving agents. Chirality is achieved by complexing the lanthanide ion with a chiral ligand. Electron-rich hard Lewis bases (typically oxygen- and nitrogen-containing compounds) interact through a donor–acceptor process with the positive lanthanide ion. Binuclear lanthanide–silver complexes have been developed for use with soft Lewis bases such as olefins, aromatics, phosphines, and halogenated compounds. With the binuclear reagents, the soft Lewis base associates with the silver ion.

Paramagnetic lanthanide ions induce substantially larger shifts, and often larger enantiomeric resolution, than is observed with other chiral derivatizing or solvating agents. Enantiodiscrimination can occur because of differences in association constants and/or because the resulting complexes are diastereoisomers. An advantage of lanthanide shift reagents is the broad range of compounds for which they are potentially effective. A disadvantage is that the chiral ligands in the lanthanide complexes are not geometrically rigid but instead are conformationally mobile. As a substrate binds, the chiral ligands alter their spatial position to accommodate the additional ligand. This effectively eliminates the ability to understand the basis of the enantiodiscrimination at the molecular level, so that assigning absolute configurations is difficult and only accomplished by observing trends on similar compounds with known configurations.

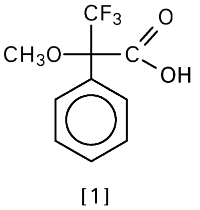
## Chiral derivatizing agents

Perhaps no reagent for enantiomeric resolution is more widely known than α-methoxy-α-trifluoromethylphenylacetic acid (MTPA), otherwise known as Mosher's reagent [1]. Either the acid or acid

chloride form is suitable for derivation of alcohols into esters and amines into amides. MTPA possesses all the classical features of an effective chiral derivatizing agent. The methoxy singlet in the $^{1}H$ NMR spectrum and the trifluoromethyl singlet in the $^{19}F$ NMR spectrum often show enantiomeric resolution. Enantiomeric distinction is usually observed only when the site of chirality is near the hydroxy or amine group at which derivatization occurs. In addition, by analysing the geometry of the resulting derivatives, it is often possible to explain the order of shifts in the $^{1}H$ NMR spectrum. An example is shown in **Figure 1** for the MTPA derivatives of secondary carbinols and primary amines.

A variety of chiral resolving reagents that are structurally similar to Mosher's reagent have been reported. These include examples in which the phenyl group of [1] has been replaced with naphthyl or anthryl groups; the methoxy group with acetoxy, methyl, ethyl, or hydroxy groups or a fluorine atom; or the trifluoromethyl group with a cyano group or a hydrogen atom. Collectively, these reagents are effective for amines, alcohols, β-amino alcohols, and thiols. Several have been used to resolve the spectra of amines through the formation of diastereoisomeric salts. The compounds α-methoxy-α-(1-naphthyl) acetic acid and α-methoxy-α-(9-anthryl) acetic acid, which possess a more sterically hindered group than [1], often provide a larger degree of enantiomeric resolution in the NMR spectra and can be used to analyse compounds with more remotely disposed chiral centres. The $^{19}F$ NMR spectrum of derivatives of α-cyano-α-fluoro-phenylacetic acid, in which the fluorine atom is closer to the chiral centre than in MTPA, show enantiodiscrimination in compounds with remotely disposed chiral centres. Enantiomeric excesses of primary alcohols that are highly hindered or have the asymmetric centre at the β carbon have been determined using a sterically hindered compound such as α-methoxy-α-(9-anthryl)-acetic acid.

The choice of solvent is critical when diastereoisomeric salts are formed. Polar solvents solvate the ions, thereby inhibiting ion pair formation and reducing enantiomeric resolution. Non-polar solvents such as chloroform or benzene are preferable. A similar observation occurs as well with many chiral solvating agents and lanthanide shift reagents, and non-polar solvents are preferred.



[1]

Axially chiral carboxylic acids such as [2] and [3] can be used to prepare diastereoisomers of secondary alcohols. The NMR spectra of the diastereoisomers show enantiomeric resolution and it is possible to assign absolute configurations. Monocyclic, chiral carboxylic acid compounds such as tetrahydro-5-oxo-2-furancarboxylic acid and 2,2-diphenylcyclopropane-carboxylic acid have been used to derivatize secondary alcohols for the purpose of enantiomeric resolution. Similarly, bicyclic derivatizing agents such as camphanic acid, camphanic acid chloride, or camphor sulfonyl chloride have been used to derivatize amines or alcohols to their respective amides or esters for the purpose of enantiomeric resolution.

By using optically pure alcohols such as methylmandelate, α-phenylethanol, 2-(trifluoromethyl)benzhydrol, 1,1′-binaphthalene-8,8′-diol, or ethyl 2-(9-anthryl)-2-hydroxyacetate, or optically pure amines such as methylbenzylamine, (1-naphthyl)ethylamine, 9-(1-aminoethyl)anthracene,

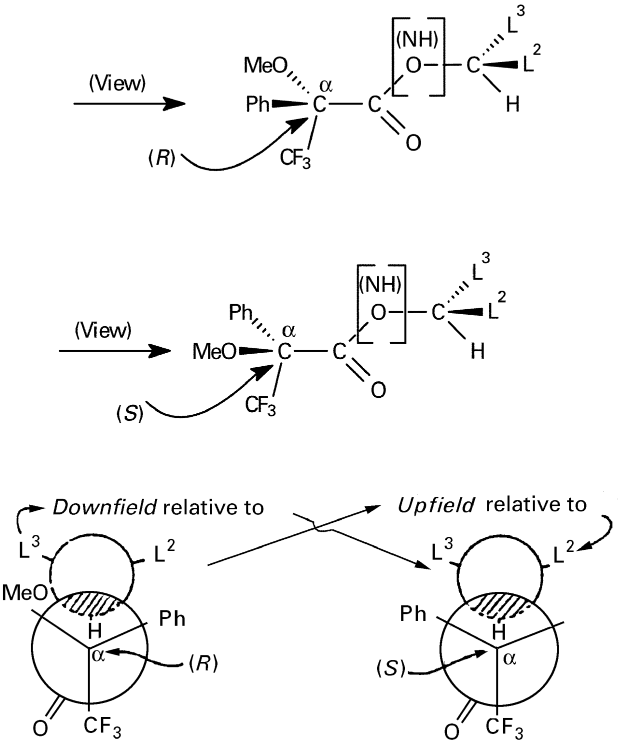
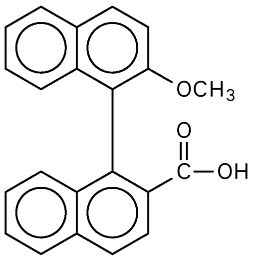
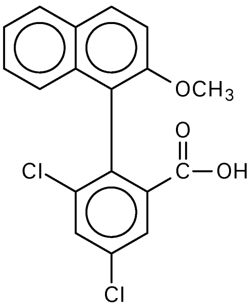



**Figure 1** Configurational correlation model for (*R*)-MTPA and (*S*)-MTPA derivatives of secondary carbinols and amines. Reprinted with permission from Dale JA and Mosher HS (1973) Nuclear magnetic resonance enantiomer reagents. Configuration correlations via nuclear magnetic resonance chemical shifts of diastereomeric mandelate, *O*-methylmandelate, and α-methyl-α-trifluoromethylphenylacetate (MTPA) esters. *Journal of the American Chemical Society* **95**: 512–519. Copyright (1973) American Chemical Society.




[2]




[3]

2-(diphenylmethyl)pyrrolidine, 1,2-diphenyl-1,2-diaminoethane, or 9-(1-amino-2,2-dimethylpropyl)-9,10-dihydroanthracene as the chiral derivatizing agent, it is possible to assess the enantiomeric purity of chiral carboxylic acids. Amine derivatives of the naturally occurring monoterpene (+)-3-carene have been used for this purpose as well. In certain cases it is possible to assign absolute configurations with these reagents.

Whereas there are a wide variety of effective chiral derivitizing agents for amines, alcohols, and carboxylic acids, there are considerably fewer options for ketones. The absolute configuration of six-membered cyclic ketones can be determined using butane-2,3-diol or the related 1,4-dimethoxybutane-2,3-diol. Enantiomeric resolution is observed in the $^{13}$C or $^{1}$H NMR spectra of the resulting butane-2,3-diol acetal. Absolute configurations can be assigned according to the model in **Figure 2**, provided the cyclohexane ring of the acetal has a preference for the chair form. The model does not work for acyclic ketones, cyclopentanones, or cycloheptanones.

The $^{31}$P nucleus is especially well suited for observing enantiodiscrimination in diastereoisomers and a variety of phosphorus-containing chiral reagents (**Figure 3**) have been developed [4]–[11]. Certain of these reagents [5, 6, 8] function with chiral primary alcohols. A particularly interesting example is the use of [8] to resolve the $^{2}$H and $^{31}$P spectra of derivatives of [1-$^{2}$H] ethanol ($CH_3CHDOH$). Other of the phosphorus-containing reagents function with amines, secondary or tertiary alcohols, and thiols. The phosphoric acid derivative [7] can be used to resolve amines through the formation of diastereomeric salts.

The $^{77}$Se nucleus has a spin of $\frac{1}{2}$, reasonable natural abundance (7.5%), adequate sensitivity ($6.98 \times 10^{-3}$ compared with $^{1}$H), a large chemical shift range, and is extremely sensitive to its electronic environment. Compound [12] is an effective chiral derivatizing agent for chiral carboxylic acids. Compounds such as 5-methylheptanoic acid and lipoic

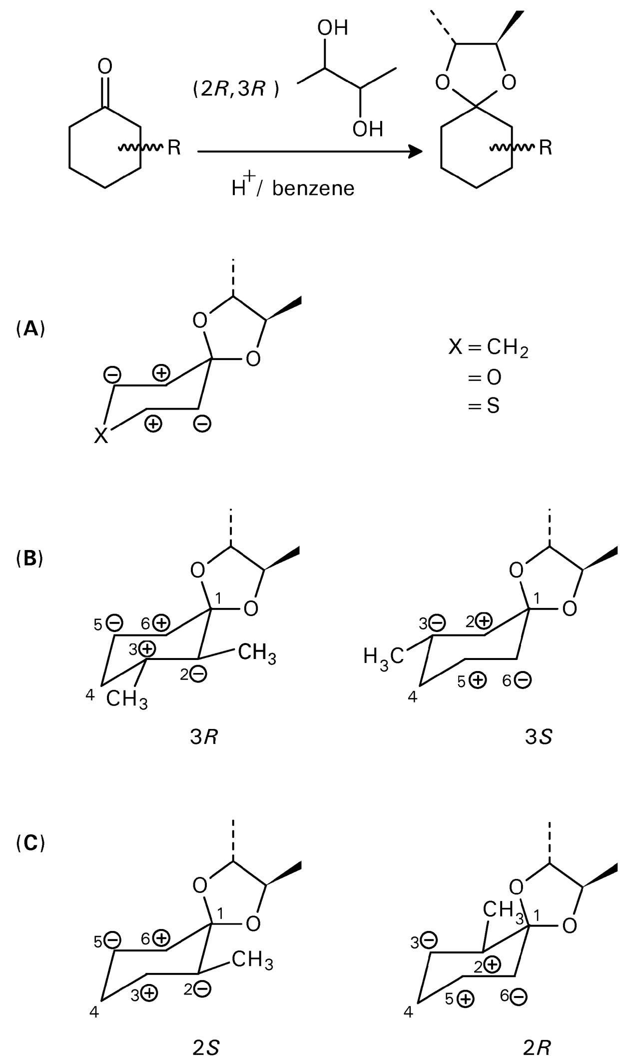


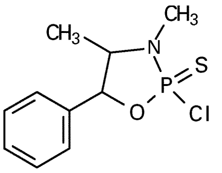

**Figure 2** (A) The model, (B) both enantiomers of 3-methylcyclohexanone acetalysed with (2*R*, 3*R*)-butane-2,3-diol and fitted in the model, and (C) both enantiomers of 2-methylcyclohexanone acetalysed with (2*R*, 3*R*)-butane-2,3-diol and fitted in the model. Reprinted with permission from Lemiere GL, Dommisse RA, Lepoivre JA *et al* (1987) Determination of the absolute configuration of six-membered-ring ketones by $^{13}$C NMR. *Journal of the American Chemical Society* **109**: 1363–1370. Copyright (1987) American Chemical Society.

acid, which have remotely disposed chiral centres, and α-deuterophenylacetic acid, which is enantiomeric by virtue of replacement of a hydrogen atom with deuterium, show enantiomeric resolution in the $^{77}$Se NMR spectrum upon derivatization with [12].

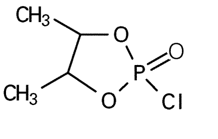
Metal complexes other than the lanthanides can sometimes be used to effect chiral resolution in the NMR spectra of enantiomers. Zinc porphyrins with chiral ligands have been used to distinguish N-substituted, anionic amino acids and amino acid esters. Techniques such as NOESY and ROESY can be used to determine relative interatomic distances. In metal

complexes containing a chiral ligand of fixed geometry, such techniques can be used to definitively establish the absolute configuration of the ligand.

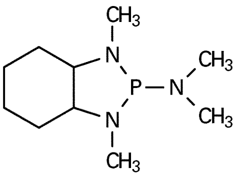
Chiral metal cations such as $[Ru(bipy)_3]^{2+}$ and $[Ru(phen)_3]^{2+}$ can be enantiomerically resolved through the formation of diastereoisomeric salts with [13]. Both the metal complex and [13] possess $D_3$ symmetry, which presumably accounts for the discriminating interactions.



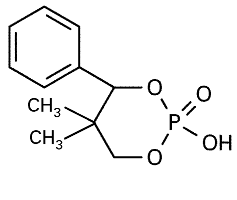
[4]



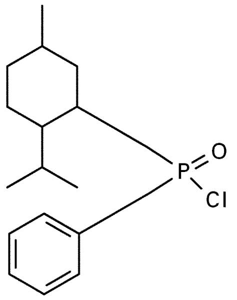
[5]



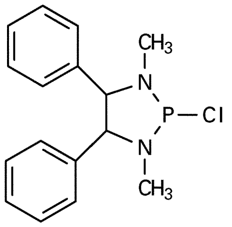
[6]



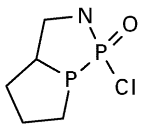
[7]



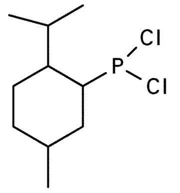
[8]



[9]

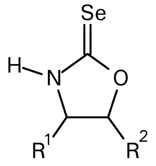


[10]

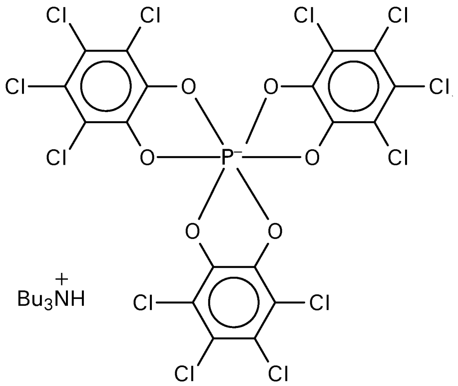


[11]

**Figure 3** Phosphorus-containing chiral derivatizing agents. [4] 2-chloro-3,4-dimethyl-5-phenyl-1,3,2-oxazophospholidine-2-sulfide; [5] 2-chloro-4,5-dimethyl-1,3,2-dioxaphospholane-2-oxide; [6] 2-dimethylamino-1,3-dimethyloctahydro-1H-1,3,2-benzodiazaphosphole; [7] 5,5-dimethyl-2-hydroxy-4-phenyl-1,3,2-dioxaphosphonan-2-oxide; [8] L-menthyl(phenyl)phosphorylchloride; [9] 2-chloro-1,3-dimenthyl-4,5-diphenyl-1,3,2-diazaphosphole; [10] dichloro(menthyl)phosphine.



[12]

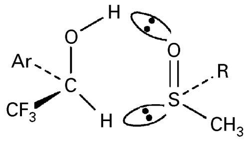


[13]

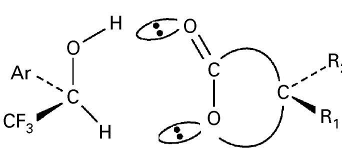
## Chiral solvating agents

The first application of a chiral solvating agent involved the observation of distinct $^{19}F$ NMR resonances for 2,2,2-trifluoro-1-phenyl ethanol [14] dissolved in (*R*)-α-phenylethylamine. Optically pure [14] has since been applied more widely as a chiral solvating agent for a range of compounds, including oxaziridines, *N*,*N*-dialkyl arylamine oxides, and sulfoxides. (*R*)-(1)-(9-Anthryl)-2,2,2-trifluoroethanol [15], which is similar to [14], is often a better chiral solvating agent than [14]. Compound [15] is also effective with γ-lactones. **Figure 4** shows the association of racemic sulfoxides and lactones with [14] and [15]. Shielding of the aryl group causes the group *cis* to it to occur at higher field than the enantiomer with the same group in the *trans* position. Similar structural considerations explain the relative shift orders in the spectra of oxiridines and *N*,*N*-dialkylarylamine oxides in the presence of [14] and [15].

3,5-Dinitrobenzoyl derivatives of leucine [16] and phenylglycine, and 1-(1-naphthyl)ethylurea

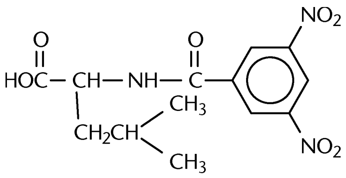


[14]

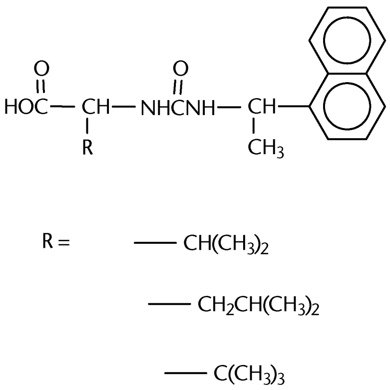


[15]

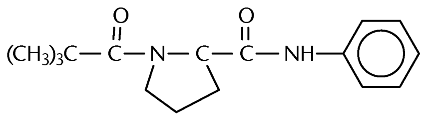
**Figure 4** Association of racemic sulfoxides and lactones with 2,2,2-trifluoro-1-phenyl ethanol [14] and 1-(9-anthryl)-2,2,2-trifluoroethanol [15].



[16]



[17]



[18]

derivatives of amino acids such as valine, leucine, and *t*-leucine [17] were originally developed and exploited in the liquid chromatographic separation of enantiomers. The chromatographic phases are potentially suitable for separating thousands of different enantiomers. Organic-soluble analogues, achieved by the conversion of the acid functionality into an ester, are useful NMR chiral solvating agents for a wide range of substrates. The related 3,5-dinitrobenzoyl derivative of phenylethylamine is also useful as a chiral solvating agent. The dinitrobenzoyl ring is especially suited to charge-transfer complexation with aromatic residues of the substrate, the amide functionalities provide sites of partial negative and positive charges, and the R group of the amino acid provides a site for steric hindrance. The naphthyl ring of [17] and the aromatic moiety of the proline-containing reagent [18] will form charge-transfer associations with compounds that have been converted into their dinitrobenzoyl derivatives.

**Figure 5** shows the interactions of aromatic rings and hydrogen bonding groups that occur between [16] and [18].

Aromatic diols such as [19] and [20] are axially chiral. Compound [20] is an especially effective chiral solvating agent for sulfoxides, phosphine oxides, alcohols, lactams, and amines. With sulfoxides, it is possible to assign absolute configurations. Quinine has been used as a chiral solvating agent for hemiacetals and β-hydroxyesters. The compound *t*-butylphenylphosphinothioic acid [21] is a useful chiral solvating agent for a broad range of compounds, including tertiary amine oxides, alcohols, thiols, and amines.

Other classes of chiral solvating agents function through the formation of host–guest complexes. Most notable among these are chiral crown ethers and cyclodextrins, although host–guest chemistry is an active area of research and reports describing the synthesis of specialized hosts designed to effect specific molecular recognition are increasingly common. Compounds [22]–[25] are examples of chiral crown ethers that induce enantiodiscrimination in the NMR spectra of chiral substrates. Substituents on the crown are responsible for the chirality. The 18-crown-6 unit in [22] and [23] is especially well suited for inclusion complexation of protonated primary amine functionalities.

The chirality in [22] is imparted by the axial chirality of the binaphthyl group; however, steric hindrance from the substituent phenyl groups promotes a greater degree of enantiomeric selectivity. The

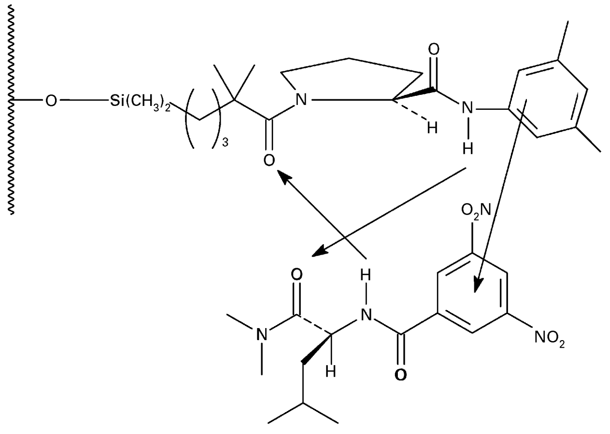
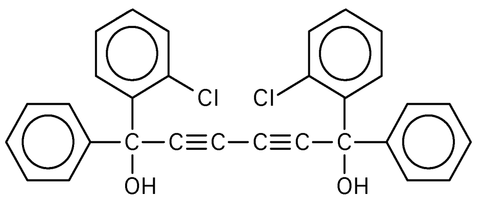
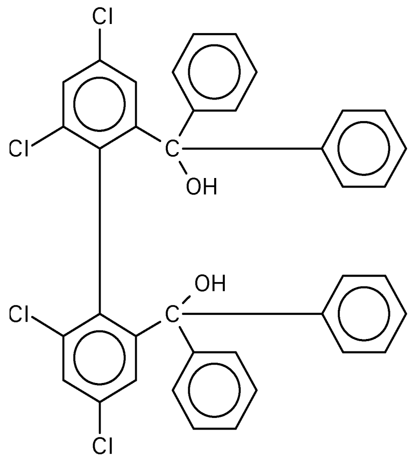
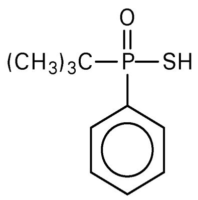


**Figure 5** Interactions proposed to account for the retention of the more highly associating enantiomer in [16] and [18]. Reprinted with permission from Pirkle WH, Murray PG, Rausch DJ and McKenna ST (1996) Intermolecular $^1H–^1H$ two-dimensional nuclear Overhauser enhancements in the characterization of a rationally designed chiral recognition system. *Journal of Organic Chemistry* **61**: 4769–4774. Copyright (1996) American Chemical Society.



[19]





[21]

[20]

mannitol unit is responsible for the chirality of [23]; however, enantiomeric discrimination is promoted through steric effects of the *t*-butyl group. Crown ethers can be used as chiral solvating agents in a variety of solvents, including chloroform, acetonitrile, and methanol. Unlike many other chiral solvating agents, enantiodiscrimination with crown ethers is often enhanced in more polar solvents. The phenolic crown ethers [24] and [25], which each have two sites of chirality from *cis*-cyclohexane-1,2-diol or 1-phenylethane-1,2-diol units, are effective chiral hosts for neutral amines or aminoethanols.

Cyclodextrins are cyclic oligomers, the most common of which contain six (α), seven (β), and eight (γ) D-glucose units (**Figure 6**). The molecules have a tapered cavity with the secondary hydroxyl groups positioned at the larger opening and the primary hydroxyl groups at the smaller. Chiral discrimination most often occurs through differences in hydrogen bonding of the substrate at the secondary opening. Underivatized cyclodextrins are water soluble. Organic-soluble cyclodextrins can be obtained by alkylation of the hydroxyl groups.

The underivatized cyclodextrins are rather unusual in their ability to function as water soluble chiral solvating agents. Enantiomeric resolution is observed in the NMR spectra of a wide range of water-soluble cationic and anionic substrates. Organic-soluble cyclodextrins are one of only a few reagents that can be used to enantiomerically resolve the spectra of hydrocarbons such as trisubstituted allenes, α-pinene, and aromatic hydrocarbons. Amines, alcohols, and carboxylic acids can also be resolved with organic-soluble cyclodextrins.

A set of host compounds of considerable current interest are calix(*n*)arenes [26] and the related resorcarenes [27]. Both sets of compounds form cone-shaped molecules. Calixarenes contain *para* substituted phenol units connected by methylene bridges. The most common have four, six, or eight phenolic units, and by varying the nature of the substituent group or derivatizing the hydroxyl groups

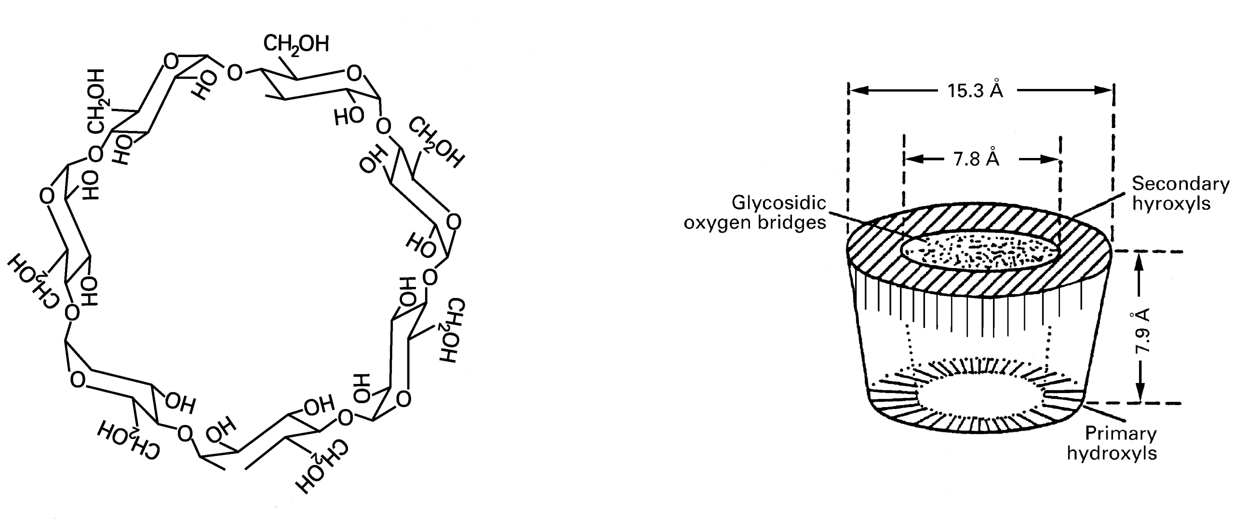
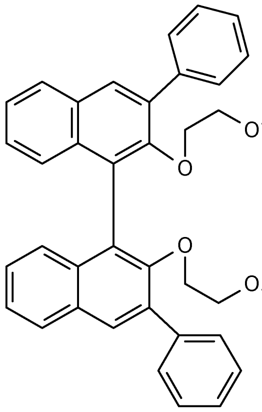
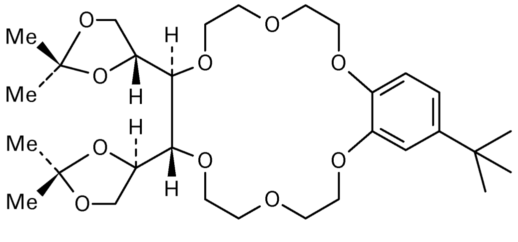


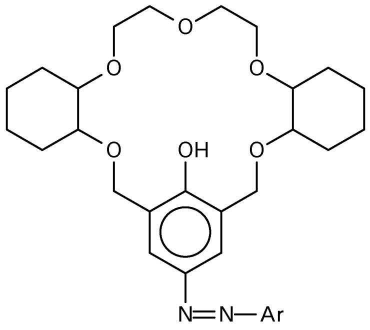

**Figure 6** Structure of β-cyclodextrin. Reprinted with permission from Li S and Purdy WC (1992). Cyclodextrins and their applications in analytical chemistry. *Chemical Reviews* **92**: 1457–1470. Copyright (1992) American Chemical Society.



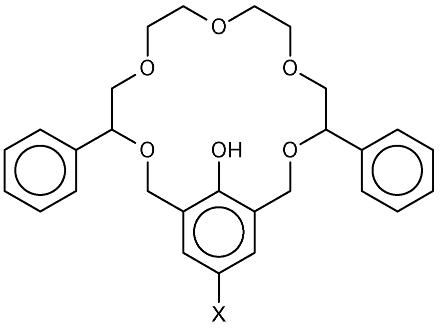
[22]



[23]




[24]



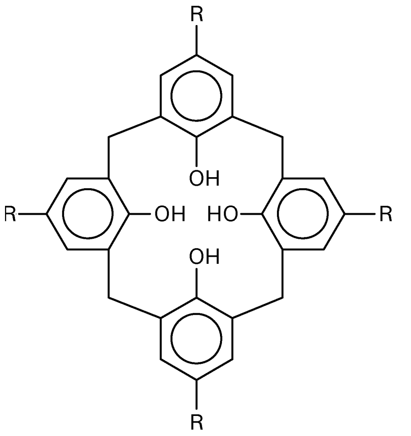

[25]

an endless variety of calixarenes can be prepared. Chirality can be achieved by one of two means. Inherently chiral calixarenes can be obtained by replacing one of the ring hydrogens on one of the phenol units with a substituent group or by synthesizing a compound in which each of the phenol rings has a different R group. Alternatively, one can attach optically pure R groups or derivatize the hydroxyl groups with an optically pure reagent. The calix resorcarene [27], in which chirality is achieved by attaching optically pure proline units to the aromatic ring, is suitable for enantiomeric resolution of certain water-soluble aromatic alcohols, amino acid esters, and carboxylic acids in water. Given the current interest in this family of compounds, further developments in this area can be expected.

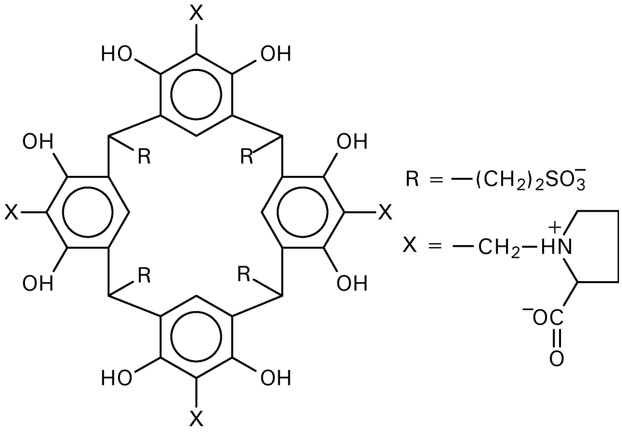
Chiral liquid crystals consisting of poly-γ-benzyl-L-glutamate have been used in various organic solvents to enantiomerically resolve the NMR spectra of alcohols, amines, carboxylic acids, esters, ethers, epoxides, tosylates, halides, and hydrocarbons. Enantiomeric pairs exhibit different ordering properties in the liquid crystal, leading to different chemical shifts in the $^{1}$H NMR spectra. A more sensitive way to probe distinctions among enantiomers in these systems is to examine deuterium quadrupolar splittings. The disadvantage is that deuterium labelling of the compound under study is then required, although procedures are available for conveniently incorporating deuterium-containing substituents into amines, amino acids, and alcohols.

## Lanthanide shift reagents

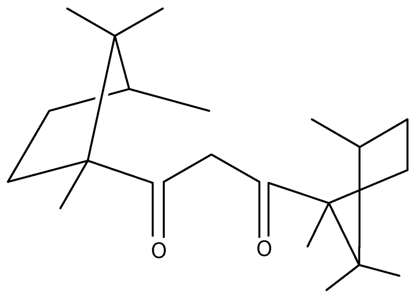
Chiral lanthanide shift reagents are noteworthy because they are effective at inducing enantiomeric resolution in the NMR spectra of such a wide range of compounds. Organic soluble lanthanide shift reagents for hard Lewis bases (primarily oxygen- and nitrogen-containing compounds) consist of tris-β-diketonate complexes in which the ligands are chiral. The most common metals used in such studies are Eu(III), Pr(III), and Yb(III). The lanthanide complex with 3-(*t*-butylhydroxymethylene)-D-camphor was the first one to be studied. A large number of chiral lanthanide shift reagents have since been evaluated,




[26]




[27]




[28]

and those that are generally most effective and commercially available include complexes with 3-trifluoroacetyl)-D-camphor (facam or tfc), 3-heptafluorobutyryl-D-camphor (hfbc or hfc), and D,D-dicampholylmethane (dcm) [28]. The shifts in the spectra of substrates are usually larger with chelates of hfbc than with facam; however, the degree of enantiomeric resolution exhibits no consistent pattern. Chelates with dcm are particularly effective at causing enantiomeric resolution. This may be owing, in part, to the steric bulk of the ligand which results in enhanced enantiodistinction of binding substrates.

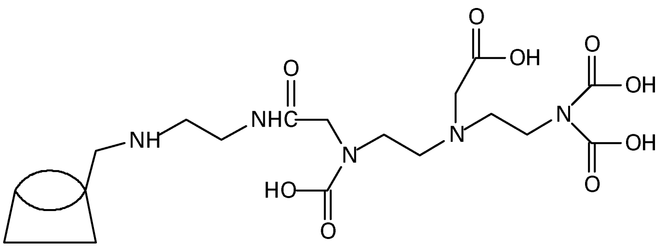
Lanthanide tris-β-diketonates rely on association of a lone pair of electrons with the tripositive lanthanide ion, and substantially larger enantiomeric resolution is observed in non-coordinating solvents. In addition, the reagents are hygroscopic and so the presence of water will significantly reduce the binding of substrates. Examples of the classes of compounds for which lanthanide tris-β-diketonates have been utilized for enantiomeric resolution include alcohols, ketones, esters, lactones, aldehydes, carboxylic acids, ethers, amines, nitrogen heterocycles, amides, lactams, sulfoxides, sulfines, sulfones, sulfoximes, sulfilimines, sulfenamides, thiocarbonyls, phosphorus oxides, and chiral metal complexes that contain ligands with functional groups suitable for direct association with the lanthanide ions.

Chiral lanthanide shift reagents for soft Lewis bases have been prepared by mixing a chiral lanthanide tris-β-diketonate such as $Ln(facam)_3$ or $Ln(hfbc)_3$ with a silver β-diketonate such as Ag(fod) (fod = 2,2-dimethyl-6,6,7,7,8,8,8-heptafluorooctane-3,5 dione). Addition of the two reagents leads to the formation of a binuclear reagent as illustrated in Equation [**3**]. The binuclear reagent consists of an ion pair between a lanthanide tetrakis chelate anion and a silver cation.

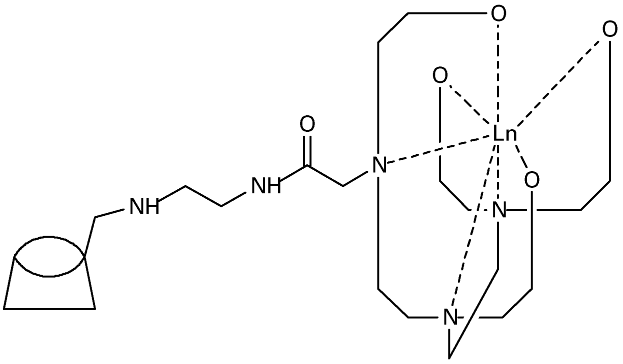
$$\begin{aligned} &\mathrm{Ln}(\beta\text{-dik})_3 + \mathrm{Ag}(\beta\text{-dik}) \\ &\qquad \rightarrow \mathrm{Ag}[\mathrm{Ln}(\beta\text{-dik})_4] \end{aligned} \qquad [3]$$

The silver ion coordinates with the soft Lewis base and the lanthanide induces the shifts. Chiral binuclear reagents have been applied to chiral olefins and aromatics. Binuclear reagents formed using the dcm ligand are not effective for chiral resolution, presumably because this ligand is too sterically hindered to permit formation of the tetrakis chelate anion. Experimentation is often needed to find the correct combination of lanthanide and silver β-diketonate to cause enantiomeric resolution.

Tetrakis chelate anions formed when either a silver or potassium β-diketonate is mixed with a lanthanide tris β-diketonate are effective shift reagents for organic-soluble salts. Adding an organo-halide or fluoroborate salt leads to precipitation of silver halide or potassium fluoroborate, resulting in an organic-soluble ion pair. These reagents have been used to enantiomerically resolve the spectra of chiral isothiouronium and sulfonium ions.




[29]




Complex with [29]

Water-soluble, chiral lanthanide shift reagents with ligands such as (*S*)-[(carboxymethyl)oxy]succinic acid, (*R*)-propylene-1,2-diaminetetraacetic acid, (*S*,*S*)-ethylenediamine-*N*,*N*′-disuccinic acid, *N*, *N*, *N*″,*N*′-tetrakis-(2-pyridylmethyl)-(*R*)-propylenediamine, and (*S*,*S*)-((1*S*,2S)-diaminophenylethylene)-*N*,*N*′-disuccinic acid have been developed. These are especially effective for resolving underivatized α-amino or carboxylic acids. In some cases, consistent trends in relative shift magnitudes are observed which allows the assignment of absolute configurations.

Achiral lanthanide shift reagents such as $Eu(fod)_3$ have been used in a variety of instances, including with MTPA esters, to enhance the resolution in the spectra of diastereoisomers formed from derivitization with an optically pure reagent. The relative shifts in the spectra of the diastereoisomers often show characteristic trends, enabling the assignment of absolute configurations. Achiral lanthanide shift reagents can also be added to solutions of chiral solvating agents and substrates to enhance enantiomeric resolution. Two mechanisms can explain the enhancement.

The first is observed when the chiral solvating agent has a high association with the achiral lanthanide complex. Bonding of the CSA to the lanthanide (Ln) effectively creates a chiral lanthanide shift reagent (Ln–CSA) (Eqn [4]). This species then interacts with the chiral substrate (S) to cause enantiomeric resolution (Eqns [5] and [6]). The chiral substrate can directly associate with either the lanthanide ion (Eqn [5]), or, as observed with certain optically pure chiral carboxylate species, the lanthanide-bonded chiral solvating agent (Eqn [6]).

$$\mathrm{Ln + CSA = Ln\text{-}CSA} \qquad [4]$$

$$\mathrm{Ln\text{-}CSA + S = S\text{-}Ln\text{-}CSA} \qquad [5]$$

$$\mathrm{Ln\text{-}CSA + S = Ln\text{-}CSA\text{-}S} \qquad [6]$$

The second mechanism is favoured when the substrate has a much stronger association with the lanthanide complex than the chiral solvating agent. If the enantiomers have different association constants with the chiral solvating agent, the enantiomer less favourably associated with the solvating agent is more available for complexation with the lanthanide and exhibits larger lanthanide-induced shifts (Eqns [7] and [8]).

$$\mathrm{S + CSA = CSA\text{-}S} \qquad [7]$$

$$\mathrm{S + Ln = S\text{-}Ln} \qquad [8]$$

This mechanism has been observed with perfluoroalkyl carbinols [14] and [15], dinitrobenzoyl derivatives of amino acids [16], 1-(1-naphthyl)-ethylurea derivatives of amino acids [17], and chiral crown ethers [22] and [23]. The lanthanide shift reagent is usually $Ln(fod)_3$. With crown ether–ammonium salt combinations, lanthanide tetrakis chelate anions of the formula $Ln(fod)_4^-$ are used.

Lanthanide ions have been covalently attached to either the primary or secondary side of cyclodextrins through a diethylenetriaminepentaacetic acid (DTPA) ligand [29]. Adding Dy(III) to the cyclodextrin-DTPA derivatives can enhance the enantiomeric resolution in the spectra of substrates that form host–guest complexes with the cyclodextrin. Enhancements in enantiomeric resolution are larger with the secondary derivative, which is consistent with the observation that enantiodiscrimination with cyclodextrins usually involves interactions with the hydroxyl groups on the secondary opening of the cavity.

## Conclusions

The number of available chiral derivatizing agents, solvating agents, and lanthanide shift reagents provides the possibility of enantiomerically resolving the NMR spectra of a wide variety of compound classes. Furthermore, in many instances the mechanism that leads to the enantiomeric resolution is well enough understood that it is possible to assign absolute configurations based on the relative shift order. Since none of the reagents is ever 100% effective, experimentation with several resolving agents may be necessary. Given the increasing interest in knowing and controlling the optical purity of chemicals, a better understanding of existing systems and the development of additional reagents for enantiomeric discrimination in NMR spectroscopy seems to be assured.

*See also:* **Chemical Shift and Relaxation Reagents in NMR; Chiroptical Spectroscopy, General Theory; Heteronuclear NMR Applications (O, S, Se, Te); Liquid Crystals and Liquid Crystal Solutions Studied By NMR; NMR Relaxation Rates; Nuclear Overhauser Effect; $^{31}$P NMR; Parameters in NMR Spectroscopy, Theory of; Vibrational CD, Applications.**

## Further reading

Lehn JM (1995) *Supramolecular Chemistry*. Weinheim: VCH Verlagsgesellschaft.

Morrison JD (ed) (1983) *Asymmetric Synthesis*, Volume 1. New York: Academic Press.

Parker D (1991) NMR determination of enantiomeric purity. *Chemical Reviews* **91**: 1441–1457.

Pirkle WH and Hoover DJ (1982). In Eliel EL and Wilen SH (eds) *NMR Chiral Solvating Agents*, Topics in Stereochemistry, Volume 1, pp 263–331. New York: Wiley-Interscience.

Stoddart JF (1988). In Allinger NL, Eliel EL and Wilen SH (eds), *Chiral Crown Ethers*, Topics in Stereochemistry, Volume 17, pp 207–288. New York: Wiley-Interscience.

Sullivan GR (1978) In Elien EL and Wilen SH (eds) *Chiral Lanthanide Shift Reagents*. Topics in Stereochemistry, Volume 10, pp 287–329. New York: Wiley-Interscience.

Webb TH and Wilcox CS (1993) Enantioselective and diastereoselective molecular recognition of neutral molecules. *Chemical Society Reviews* **22**: 383–395.

Wenzel TJ (1987) *NMR Shift Reagents*. Boca Raton (FL): CRC Press.

# Environmental and Agricultural Applications of Atomic Spectroscopy

**Michael Thompson**, Birkbeck College, London, UK
**Michael H Ramsey**, Imperial College of Science Technology and Medicine, London, UK

Copyright © 1999 Academic Press

| ATOMIC SPECTROSCOPY |
| --- |
| Applications |

We consider the determination of the concentration of elements in various materials studied in agricultural and environmental applications, by the use of the following methods: atomic absorption spectroscopy (AAS) using a flame (FAAS) or a graphite furnace (GFAAS) as an atom cell; inductively coupled plasma atomic emission spectroscopy (ICPAES); inductively coupled plasma mass spectrometry (ICPMS) and X-ray fluorescence (XRF). The analytical characteristics of the methods as normally practised are compared with the requirements of fitness for purpose in the examination of soils and sediments, waters, dusts and air particulates, and animal and plant tissue. However, there are numerous specialized techniques that cannot be included here.

## Introduction

Environmental and agricultural studies involve a number of different materials that need to be analysed. The principal test materials encountered by the analyst in this sector are soils, sediments, air, dusts, water, plant material and animal tissue. For completely practical reasons, we need to know the composition of these media over wide concentration ranges, from major constituents to elements at ultratrace levels. One is especially interested in the way elements are transferred between media, their biological role and metabolic pathways, and their ultimate fate.

Almost invariably in these applications we can be satisfied if the analytical result obtained is within a range of 90–110% of the true result. This modest fitness-for-purpose requirement is an outcome of the high uncertainties in sampling of environmental materials. In these circumstances, the total uncertainty (sampling plus analytical) is not improved by the use of high-accuracy analytical methods. Overall, therefore, it is a more effective use of resources to analyse a larger number of samples with only moderate accuracy.

Atomic spectroscopy is almost exclusively the method of choice in the determination of concentrations of elements in agricultural and environmental studies. There are several reasons for this: (a) for the most part no separations are needed before the measurement, only a straightforward dissolution procedure; (b) test materials can be analysed in rapid succession; (c) elements can be determined at suitably low concentrations; (d) apart from atomic absorption, atomic spectroscopy methods are suitable for determining many elements simultaneously; (e) modern atomic spectroscopy methods can easily fulfil the accuracy requirement specified above. No other methods can match this range of performance. Where speciation studies require a preliminary separation of the analytes, atomic spectroscopy is often the method used in 'hyphenated' methods to provide a very sensitive and virtually specific detector for almost any element.

In the following sections we consider the key features of the atomic spectroscopy methods, how they relate to user requirements, and how they are applied to specific types of test materials. Although only laboratory-based methods are considered here, it must be recalled that field methods sometimes find worthwhile applications: XRF is one atomic spectroscopy method that is available as a truly portable instrument.

### Detection limits

The detection limit is (informally) the lowest concentration of the analyte that can be reliably detected, and is a reflection of the precision of the instrumental response obtained by the method when the concentration of the analyte is zero. Obviously, if the uncertainty range of the measurement (at some specified level of probability) includes zero, we are not sure that the analyte has been detected. However, detection limits as supplied in the literature and manufacturers' brochures can be misleading to the unwary. Their magnitude depends critically on the conditions under which the precision was estimated. Manufacturers and method developers usually quote 'instrumental detection limits', where the precision is estimated on the pure solvent (or other matrix) in the shortest possible time, with no adjustments at all to the instrument.

Such detection limits usually need to be adjusted before they can be applied to the analysis of real materials, *inter alia* to allow for the 'dilution' of the analytes. For instance, in the analysis of a soil, a 0.1 g test portion may be treated with reagents and the analytes liberated into solution may be made up to a volume of 10.00 mL for presentation to the instrument. The analytes have been diluted by a factor of 100, in that an element at a concentration of 100 ppm in the soil will be presented to the instrument at a concentration of 1 ppm in the test solution. In addition, in 'real analysis', many features other than short-term instrumental variation play a part in the variability of results. Thus, detection limits estimated under repeatability or reproducibility conditions may be considerably higher than simple dilution-adjusted instrumental detection limits, by as much as an order of magnitude. We here consider instrumental detection limits, adjusted for dilution and with an extra factor of 5 to convert them roughly to repeatability detection limits.

The detection limit also depends critically on details of the instrumentation. For example, magnetic sector ICPMS provides detection limits lower than those of quadrupole instruments by between one and two orders of magnitude.

Among the atomic spectroscopy methods, flame and furnace methods have very variable detection limits, because they are affected by variations among the elements in terms of efficiency of atomization and efficiency of ionization. ICPAES and ICPMS show a smaller range of detection limits, because atomization is uniformly close to 100% efficient, and the efficiency of ionization accounts for much of the remaining differences. Thus, in ICPMS, elements more difficult to ionize (e.g. nonmetals) have higher detection limits. The outcome is that FAAS, being fast, cheap and robust, is used for a limited range of more abundant elements. GFAAS, being slow and requiring much care, but with much lower detection limits for some elements, is used on a much more restricted scale, typically for low trace elements such as lead and cadmium in food and drink. ICPAES is fast and robust and moderately priced, and is widely used as the workhorse method for major and trace elements, with instrumental detection limits typically 1–50 ppb (parts per $10^9$). ICPMS is unique in the very low detection limits obtainable in simultaneous analysis, typically 0.1–10 ppt (parts per $10^{12}$), but it requires a high dilution factor (typically 500–1000), is more expensive, and is more prone to downtime. In XRF methods, the detection limit is poor for light elements but roughly constant at a few ppm for elements heavier than iron, and the method usually requires only small dilution factors.

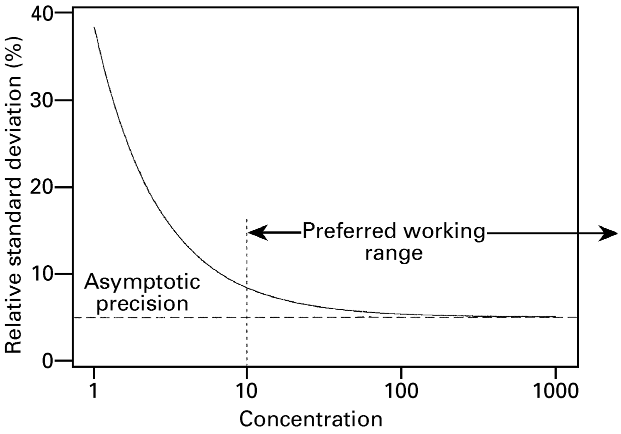



**Figure 1** Relationship between precision of measurement (as relative standard deviation) and concentration of analyte (expressed in units of multiples of detection limit). The graph shows the sharp fall in relative standard deviation just above the detection limit, and the comparatively steady values in the preferred working range.

## Asymptotic precision

The other aspect of precision that affects the judgement of instrumental analysis can be called 'asymptotic precision', which is precision estimated at a concentration orders of magnitude above the detection limit. Under that condition we usually find that relative precision is concentration invariant, i.e. the relative standard deviation (RSD) is constant. Thus, in general, with increasing concentration the asymptotic precision always increases, but the relative standard deviation falls sharply, from about 30% at the detection limit, to close to the asymptotic level at 20–100 times the detection limit (**Figure 1**). Usually we try to arrange matters so that we make measurements at these near-asymptotic precisions.

Asymptotic precisions for atomic spectrometry methods depend on the type of instrumentation, but also on other factors such as the protocols used for calibration and measurement, and whether internal standardization has been used. Typical between-run values as RSDs are: FAAS, 1.5%; GFAAS, 5%; ICPAES, 1%; ICPMS, 3%; XRF, 0.5–1%.

## Abundance–detection limit (ADL) diagrams

A convenient way of summarizing the applicability of the various methods to various types of test materials, over the range of relevant elements, is the use of abundance–detection limit (ADL) diagrams. For each element in these diagrams, the detection limit of the method ($d$), is plotted against the median concentration found in the sampling medium in question (the abundance, $a$). (The detection limit is adjusted

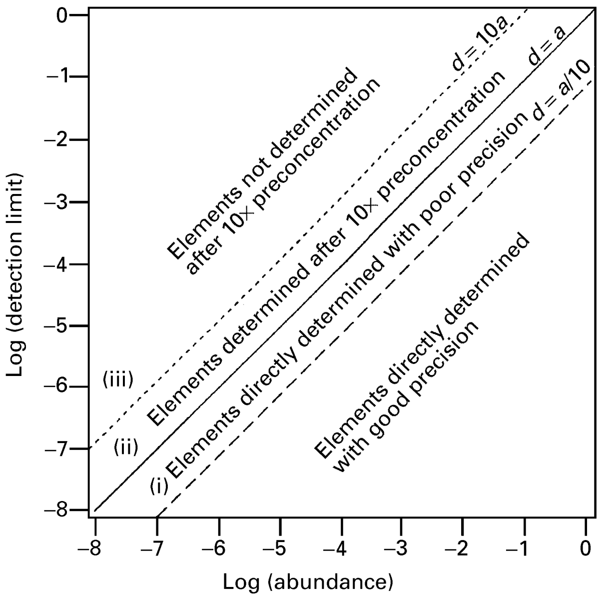



**Figure 2** The layout of an ADL (abundance–detection limit) diagram, showing the for interpretation for elements that plot in the various regions of the diagram. The concentration units used for both axes are mass fractions.

for a likely dilution factor and with an extra allowance to convert instrumental detection limits into between-run detection limits.) Both variables are expressed as mass fractions (e.g., 1 ppm ≡ $10^{-6}$) and plotted on $\log_{10}$ scales. (Frequency distributions of concentrations of trace elements in randomly selected agricultural and environmental media are usually strongly skewed to the positive tail, so lower concentrations are well represented by the median. In any event, information on percentiles other than the median is difficult to compile.)

ADL diagrams are divided into four zones (**Figure 2**) by three lines: (i) $d = a/10$; (ii) $d = a$; and (iii) $d = 10a$. Elements plotting to the right of and below line (ii) can be determined with at least moderately good precision at normal concentrations. Line (i) represents concentrations at 10× the detection limit, at which the elements can be determined with a relative precision of about 5%. Elements plotting to the right of and below this line can therefore be determined with completely acceptable precision, and those plotting between lines (i) and (ii) with moderate precision. Elements between lines (ii) and (iii) can be determined with moderate precision only after a 10× preconcentration, while elements above or to the left of line (iii) can be determined only if preconcentrated to a degree greater than tenfold. These diagrams will be used to compare the capabilities of the various methods.

### Interference effects

Another figure of merit that describes the quality of the data produced in atomic spectroscopy is 'trueness', which, in the absence of analytical blunders, is synonymous with lack of interference. Interference is defined as the influence on the analytical signal relating to an analyte caused by constituents of the test materials other than the analyte (the concomitants). Broadly, atomic spectroscopy is affected by interference to a remarkably low extent. However, interference is not absent and usually has to be taken seriously. Fortunately, there are well-established analytical techniques for obviating the effects of interference that are eminently applicable to atomic spectroscopy.

Interference effects are often spectral in origin. In atomic emission and absorption they result when a concomitant gives rise either to photons close to the same wavelength as the analyte (spectral overlap) or to photons of another wavelength that inadvertently arrive at the detector (stray light). Methods for overcoming such spectral problems usually involve either a separate estimation of the background to a spectral line, or estimation of the concomitant concentration at another region of the spectrum and the calculation therefrom of an appropriate correction term. These features are built into modern instruments and their software, but the correction procedure may degrade other aspects of performance such as the detection limit. Comparable problems occur in mass spectrometry.

Nonspectral interferences are called matrix effects. In free-atom methods, matrix effects often reflect changes in the efficiency of atomization of the analyte or of excitation of the separate atoms produced in the atom cell. In FAAS and ICPAES, changes in the physical characteristics of the test solutions (e.g., surface tension, viscosity, density) may additionally affect the efficiency of the nebulizer. However, matrix effects seldom cause errors of greater than ±5% and can often be ignored in agricultural and environmental studies. In GFAAS, by contrast, the composition of the matrix is crucial in determining the atomization efficiency, and needs to be carefully controlled. Likewise, in XRF the gross composition of the pellet or bead may affect the intensity of fluorescent X-rays.

There are a number of convenient and effective ways of handling matrix effects. Where the gross composition of the test materials is effectively constant, it is sufficient to calibrate the instrument with matrix-matched calibrators. Even when test materials are also variable, they may sometimes be matched with calibrators by swamping the native

matrix with a much larger quantity of a matrix modifier. This technique is commonly used in FAAS, GFAAS and XRF methods. If matrix matching is not possible, the method of analyte additions is always applicable, if somewhat laborious, to methods with linear calibration functions. In XRF a set of complex correction equations based on the major composition is used to correct for matrix effects.

### Performance enhancement by chemical pretreatment

The detecting power of all of these methods can be improved by preconcentration after the initial decomposition of the test material. This procedure, effected by methods such as ion-exchange, solvent extraction or solid-phase extraction, is very effective. It often has the additional advantage of removing the analytes from their original matrix and thereby considerably reducing matrix effects. Concentration by a factor of 10 is straightforward, and by 100 is usually feasible. For the atomization methods it is convenient, after the initial separation, to return the analytes to a small volume of aqueous medium. Trace elements can also be analysed directly in organic solvents. This is not a popular option, although there are no inordinate difficulties about the procedure. Different options for presenting the separated analytes exist in XRF, for instance trace analytes can be concentrated by passing the test solution through an ion-exchange membrane and analysing the membrane in a specially designed holder.

Another widely used method of improving detection limits is the injection of certain elements in the gas phase. The elements Ge, Sn, Pb, As, Sb, Bi, Se and Te all readily form hydrides in acidic aqueous media on reduction with sodium tetrahydroborate. As the hydrides are not very soluble in water, they can be readily transferred into the gas phase and thus completely separated from the matrix. The efficiency of hydride formation is high (often approaching 100%), so that detection limits for these elements can be improved by a factor of 100 without problems.

### Speed of analysis

Speed of analysis is one of the attractive features of the atomic spectroscopy methods described here. An important discrimination can be made, however, between methods that determine elements sequentially and those, clearly much faster overall, that can determine a whole suite of elements simultaneously. In methods involving nebulization, the critical determinant is the time taken for the system to ‘wash out’, so that material from one test solution is completely replaced by its successor in the analytical system and a new stable signal established. It usually takes only a few seconds subsequently to make a measurement of the signal intensity. Such sample changeovers can be as short as 15 s, but are usually longer (1–2 min) because of poor attention to design in this aspect of commercial instrumentation. In GFAAS the speed of analysis is largely controlled by the time needed for the furnace to cool sufficiently before introduction of the next sample. XRF exhibits different properties. It may require a substantial time for the signal to be collected (minutes), whereas the exchange of one test material for the next takes only seconds (because only solids in special holders are moved).

The question of speed is related to the reliability of the instrumentation. Reliable methods with low operating costs can readily be automated and can left unattended, e.g. overnight. In those circumstances, speed may be a secondary consideration. XRF methods come into this category. For methods that are more prone to problems during running (and so cannot be left unattended) or that use expensive consumables (e.g. argon for ICP methods), speed is usually of the essence.

### Instrumental reliability and stability

Reliability is also a key feature in the comparison of methods. Some methods are inherently reliable and stable, e.g. XRF. The signal obtained from an analyte in a particular material will vary only slightly from run to run, so that calibrations have to be updated only infrequently. The situation is quite different in methods requiring nebulization and/or atomization, because there the stability of the signal depends on such variables as power input, gas flows, etc. In ICP methods particularly, the efficiency of atomization can occasionally vary suddenly due to partial blockages in the nebulizer or in the injector tip of the plasma torch. Vigilance is required for early detection of such problems. In addition, in ICPMS, drift may be caused by gradual occlusion of the orifice in the sampling cone (part of the plasma–MS interface) by material passing through the plasma. This problem must be minimized by the use of high dilutions (i.e. at least 1:500) or by the use of flow injection to deliver the test solutions in short slugs, or its effects may be alleviated by the employment of an internal standard.

### Ease of calibration

Two features contribute to ease of calibration in analytical methods, (a) linearity over a wide concentration range, and (b) use of solutions (rather than solids) to present the analyte to the instrument.

Linearity makes calibration simple by limiting to two the number of calibrators necessary to estimate the calibration function for a single analyte. A nearly (but not necessarily completely) linear calibration over a considerable concentration range (e.g. more than four orders of magnitude) is also essential in multielement calibration, where major and trace constituents are determined simultaneously.

Preparation of calibrators in solution can be effected by simple mixing of the correct amounts of the constituents in an appropriate matrix. There is little problem with either homogeneity or speciation of the analytes in that instance. This consideration applies directly to conventional ways of operating AAS, ICPAES and ICPMS. It also applies to XRF methods that employ fused beads, where the test material is homogenized in a flux. (The end-product is, of course, a solid solution). Where solid calibrators are used, more problems are usually encountered. The most common of such are problems of creating a sufficiently homogeneous mixture of the analytes in a realistic matrix, and problems of speciation of the analyte. Fortunately, neither of these is a serious problem in routine XRF, although they may be in specialist applications of ICPAES and ICPMS, such as laser ablation work.

There are special problems associated with multielement calibration when more than a few elements are required. First, it is necessary to provide a set of calibrators that span the range of the test materials adequately. This is simple if there are no constraints on the design of the calibrator set. Often there are such constraints. For instance, for major constituents, it is not realistic (and sometimes not even possible) to have all of the major constituents at their maximum concentration in a single calibrator. In consequence, methods such as serial dilution of a single concentrated solution cannot be undertaken, and more complex designs are required. Constraints due to minor interferences apply to methods such as ICPAES. In some styles of calibration for major and trace elements simultaneously, there are minor deviations from linearity and small interference effects from major constituents that have to be taken into account. Finally, the calibration set must be designed to avoid major matrix mismatches with the test materials.

In XRF it is common for practitioners to calibrate with reference materials. As these are natural materials, designs cannot be made optimal. It would in fact be easier and more accurate to use pure stoichiometric compounds to synthesize calibrators, just as in liquid solution work, and this strategy would allow optimal designs.

## Soils and sediments

Soil is the key medium in environmental and agricultural studies of the distribution of the elements. Soils and sediments are the main repositories of trace metals. For conventional analysis by AAS and ICP methods, it is necessary to bring the analytes into solution. This is usually effected by treatment of the test material with strong mineral acid mixtures. In many instances a complete dissolution of the soil is not required. Most elements of agricultural or environmental import can be sufficiently solubilized by evaporation with nitric and perchloric acids, followed by a dissolution with hydrochloric acid. Any elements remaining undissolved after this type of treatment will hardly be of significance to living organisms. An attempt at studying the bioavailability of significant metals can be made by the use of ‘selective extractants’. These are designed to enable experimenters to distinguish between categories such as ‘readily exchangeable’, ‘organic bound’, etc.

By making some reasonable assumptions about dilution factors and the relationship between instrumental and repeatability detection limits, it is possible to construct ADL diagrams for soils analysed by the various atomic spectroscopy methods. These are shown in **Figures 3–6** for a selection of key elements. A surprisingly uniform picture emerges, with only ICPMS being suitable for the direct determination of elements such as As, Cd, Hg, Mo, Se and U, a range of important essential or toxic elements of low abundance. However, preconcentration or other sensitivity-enhancing methods can usually render these elements capable of determination by AES or AAS methods. Of the methods considered here, perhaps the least applicable is FAAS, which is also probably the most laborious, being a single-element method.

## Water

Fresh water is in many ways an ideal medium for analysis by the atomization methods. It is ready for direct analysis with virtually no pretreatment, and no dilution is required. Matrix effects are usually negligible or easily handled. The method *par excellence* for water analysis is ICPMS (**Figure 7**), which can be used to determine most of the elements directly. Another important method here is GFAAS which, although undoubtedly slow, can be used very effectively for a range of useful trace elements. The main problem with the remaining methods (**Figure 8**) is that many important elements are present in fresh waters at very low concentrations, below the

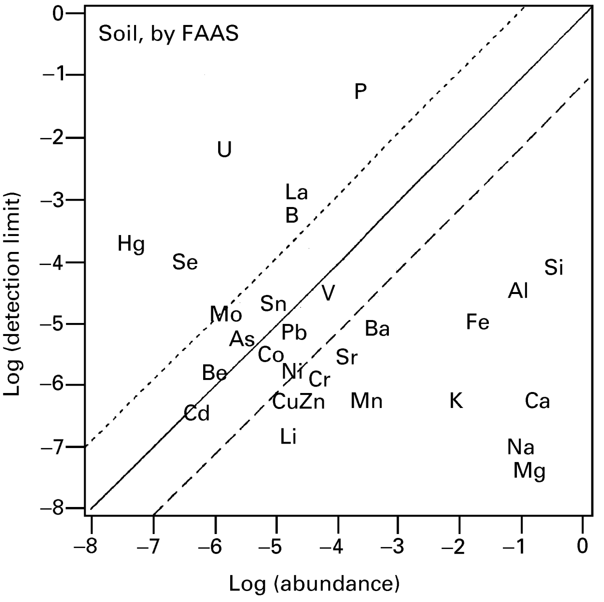



Figure 3 ADL diagram for soil analysed by flame atomic absorption spectroscopy. (For explanation see **Figure 2**.)

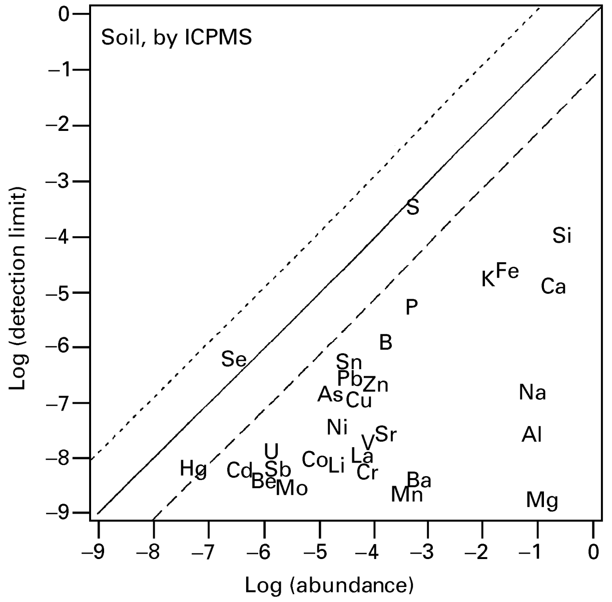



Figure 5 ADL diagram for soil analysed by inductively coupled plasma mass spectrometry. (For explanation see **Figure 2**.)

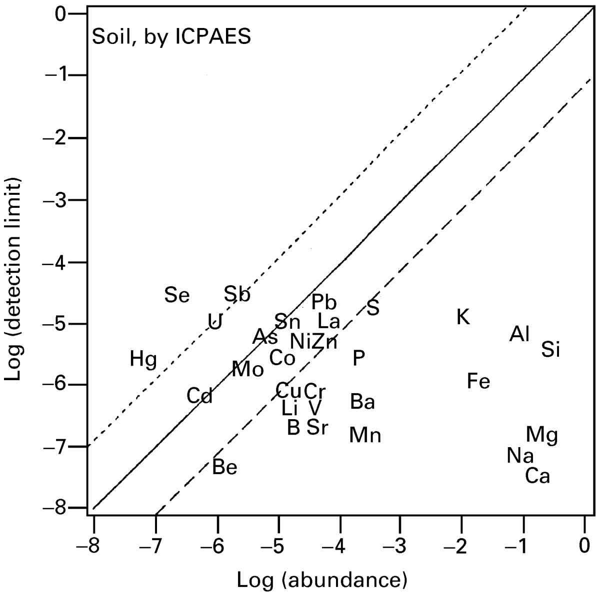



Figure 4 ADL diagram for soil analysed by inductively coupled plasma atomic emission spectroscopy. (For explanation see **Figure 2**.)

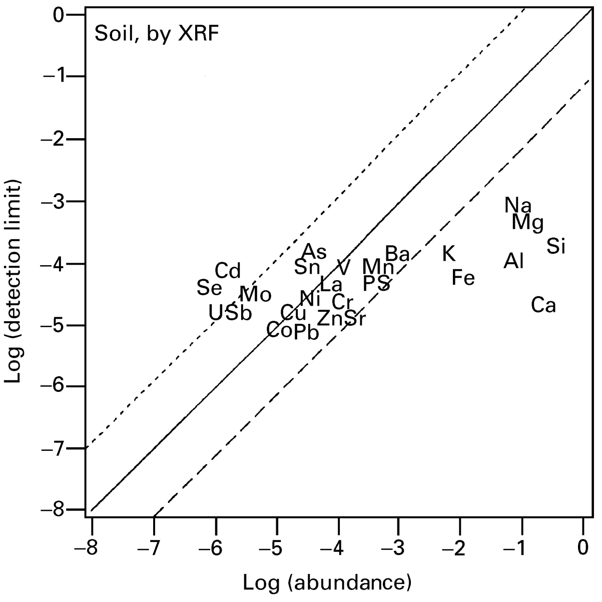



Figure 6 ADL diagram for soil analysed by X-ray fluorescence spectroscopy. (For explanation see **Figure 2**.)

detection limits for direct determination by most AAS and AES methods. For most elements this drawback can readily be overcome by preconcentration. Simple evaporation is effective for 10–20-fold preconcentration, without causing serious matrix effects. Matrix separations are also simple and useful. For example, elements such as Cd, Cu, Fe, Mn, Pb and Zn can be determined by FAAS or ICPAES after solvent extraction or ion exchange.

Sea water and other saline waters pose some problems for the atomization methods. Nebulizers for the ICP tend to become blocked with salt encrustations with a disastrous effect on sensitivity. This can be overcome in direct analysis by flow injection techniques, or by the use of 'high-solids' nebulizers. Dilution also helps here but, of course, degrades the detection limits. The sampling cone in ICPMS equipment is also prone to be gradually occluded when

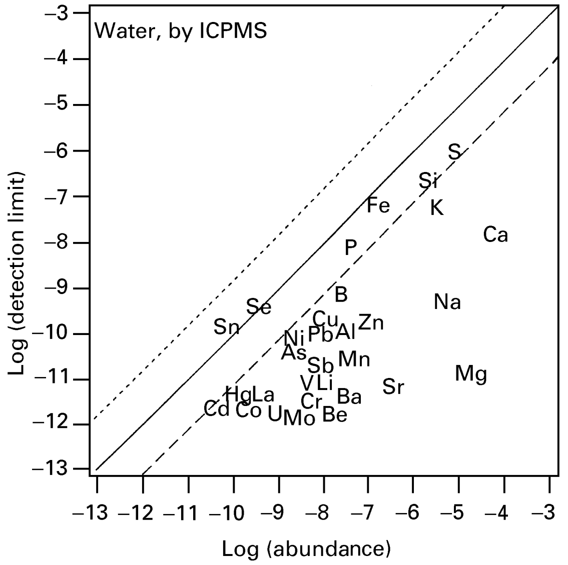



**Figure 7** ADL diagram for water analysed by inductively coupled plasma mass spectrometry. (For explanation see **Figure 2**.)

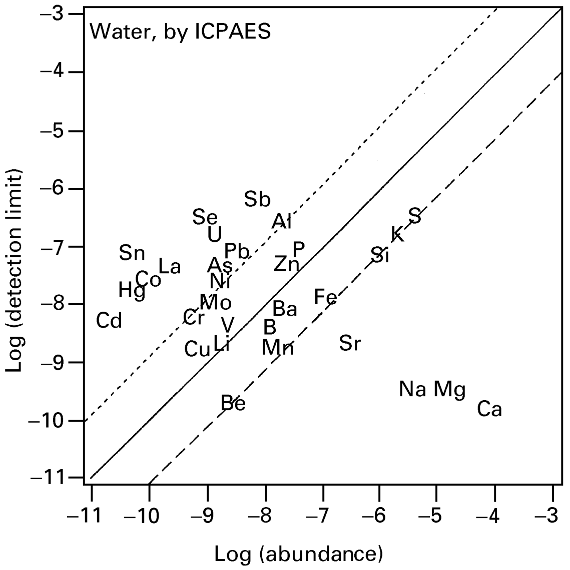



**Figure 8** ADL diagram for water analysed by inductively coupled plasma atomic emission spectroscopy. (For explanation see **Figure 2**.)

high-solids solutions are sprayed. Again, preconcentration techniques that provide a separation of the analytes from the matrix are very useful in this context.

The speciation of certain elements in water is important, as it throws light on the transport of the elements between other media, e.g. the soil–plant interface. In this context, atomic spectrometry has the role of providing a sensitive and virtually specific detector for the individual elements, after some

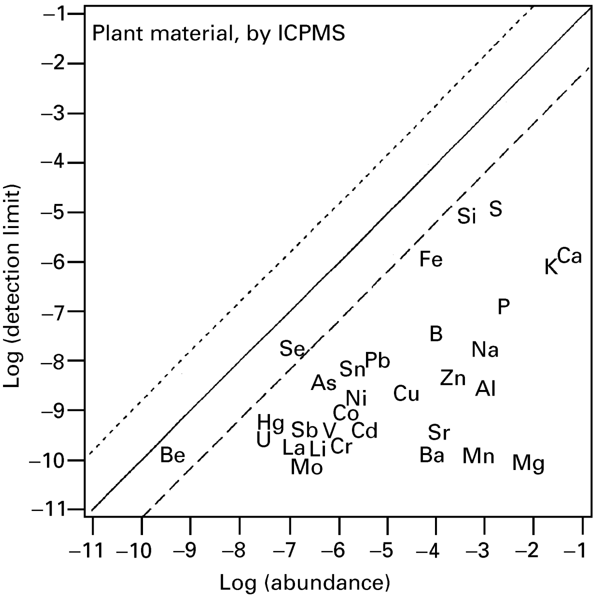



**Figure 9** ADL diagram for plant material analysed by inductively coupled plasma mass spectrometry. (For explanation see **Figure 2**.)

preliminary process such as chemical separation or chromatography.

## Air and dust

Metallic elements are present in air mostly as particulate material, which can be removed *inter alia* by filtration. Dusts are simply air particulates that have collected naturally by gravitation. Atomic spectrometry methods are well adapted to the analysis of used air filters. For instance, filters of the cellulose acetate type can conveniently be destroyed by oxidation with nitric/perchloric acid while the metallic elements are thereby solubilized. The resulting aqueous solution, of small volume, can be analysed simultaneously for many elements by ICPAES or ICPMS. The former is quite capable of determining most elements at their threshold concentration after the filtration of (say) 10 litres of air. Alternatively, air filters have a suitable form for direct analysis by XRF in special filter holders.

## Biological tissues

Animal and plant tissues are usually prepared for elemental analysis by destruction of the organic matrix by oxidation, either in a furnace (dry ashing) or by the use of oxidizing acids (wet ashing). The two methods have similar outcomes in most respects, except that there is a tendency for volatile elements

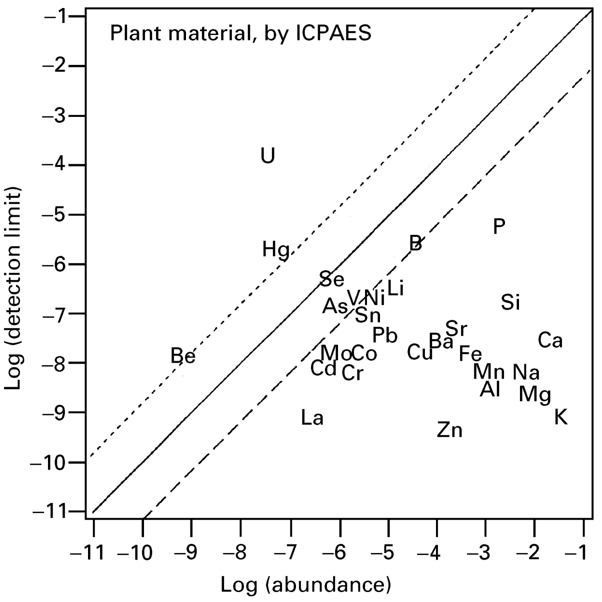



**Figure 10** ADL diagram for plant material analysed by inductively coupled plasma atomic emission spectroscopy. (For explanation see **Figure 2**.)

such as As and Se and even Cd to be lost during dry ashing, especially if the temperature is allowed to rise above 500°C. The normal outcome is for the elements to be nebulized from a dilute solution of hydrochloric acid. Matrix effects are not usually troublesome in this sector, although some materials high in calcium, potassium or phosphorus may require special attention. **Figures 9** and **10** show the applicabilities of ICPMS and ICPAES for plant material. Concentrations of elements in animal tissues are broadly similar, although there are contrasts in certain cases.

*See also:* **Atomic Absorption, Methods and Instrumentation; Atomic Absorption, Theory; Atomic Emission, Methods and Instrumentation; Atomic Spectroscopy, Historical Perspective; Calibration and Reference Systems (Regulatory Authorities); Environmental Applications of Electronic Spectroscopy; Food & Dairy Products, Applications of Atomic Spectroscopy; Food Science, Applications of Mass Spectrometry; Food Science, Applications of NMR Spectroscopy; Inductively Coupled Plasma Mass Spectrometry, Methods; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Alloway BJ and Ayres DC (1997) *Chemical Principles of Environmental Pollution*, 2nd edn. London: Blackie.

Boumans PWJM (ed) (1987) *Inductively Coupled Plasma Emission Spectrometry*. New York: Wiley.

Fifield FW and Haines PJ (1995) *Environmental Analytical Chemistry*. London: Blackie.

Gill R (ed) (1997) *Modern Analytical Geochemistry*. Harlow: Longman.

Haswell SJ (ed) (1991) *Atomic Absorption Spectrometry: Theory, Design and Applications*. Amsterdam: Elsevier.

Thompson M and Walsh JN (1989) *Handbook of Inductively Coupled Plasma Spectrometry*, 2nd edn. Glasgow: Blackie.

# Environmental Applications of Electronic Spectroscopy

**John W Farley**, University of Nevada, Las Vegas, NV, USA
**William C Brumley**, National Exposure Research Laboratory, US EPA, Las Vegas, NV, USA
**DeLyle Eastwood**, Air Force Institute of Technology, Wright-Patterson AFB, OH, USA

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| Applications |

## Introduction

Electronic spectroscopy is widely used to detect environmental contamination. Environmental applications of electronic spectroscopy involve challenging analytical problems. Both qualitative identification and quantitative determination are performed for analytes which may occur at concentrations ranging from parts per hundred to parts per quadrillion. The analytes occur in a very wide range of matrices. Electronic spectroscopy consists of monitoring the absorption of light by the sample or monitoring the emission of light, often after excitation of the sample by an appropriate light source or laser beam. Both inorganic and organic analytes can be detected. Real-world samples are often complex mixtures, and therefore environmental applications may involve coupling a separation technique with electronic spectroscopy. Separation techniques include liquid chromatography (LC), gas chromatography (GC), high-performance liquid chromatography (HPLC), capillary electrophoresis (CE), and immu-nochemical analysis. Emerging techniques include the increasing use of hybrid techniques and the development of highly parallel detection instruments.

## The analytical problem

The environmental analytical problem generally presents two questions: what substances are present in an environmental sample and how much of each of the substances is there? Sometimes these questions can be answered by spectroscopic techniques in a direct manner with little sample handling. However, the range of needs encompasses a vast array of matrices and levels of determination. To indicate some scope to the problem consider a list of potential matrices: soil, sediment, water, plants, animals, fly ash, sludge, waste water, leachates, food, blood, urine, hair (fur), drinking water, commercial pesticide formulations, air, dust, automobile and truck exhaust, and rain. The various matrices require extraction techniques, cleanup techniques, and often a final separation or detection technique. Contaminant levels range from percent levels (part per hundred) to parts per quadrillion (pg $kg^{-1}$). To give some realism to these numbers, for example, pesticides in formulations may be present at percentage levels. In another context, contaminants at Superfund sites are often found in the parts per million (µg $g^{-1}$) level. (A Superfund site is an identified site under the authorities of the Comprehensive Environmental Response, Compensation and Liability Act where releases of hazardous substances had already occurred or might occur and posed a serious threat to human, health, welfare or the environment.) Near the other extreme, the European Union regulates phenoxyacid herbicides at the 100 parts per trillion (pg $g^{-1}$) level in drinking water.

The applications, therefore, fall into two broad categories: qualitative identification and quantitative determination of substances or classes of compounds in environmental matrices. The qualitative identification is usually accomplished by acquiring a full UV-visible absorption or emission spectrum of the compound of interest. This spectrum is then compared with libraries of spectra for potential matches. Usually, the specificity of this match does not produce the same level of certainty as could be obtained from a mass spectrum that has been properly matched to a library spectrum. Nevertheless, within the context of a limited number of compounds and a particular analysis, some degree of certainty is obtained. In some favourable cases (e.g. fluorescence of oils), UV-visible identification has prevailed in legal proceedings and shown similar degrees of certainty to GC-MS. In UV-visible identifications, comparison is often made to site-specific standards, rather than against a library. The certainty of identification by UV-visible can be enhanced by the association of a retention time or migration time achieved by employing one or more separation techniques in

combination with the spectroscopic data obtained on-the-fly.

The identification problem or the specificity of the determination can be a complicated matter because of the large number of analytes, matrix components, and the variability of environmental matrices depending upon location. There are now over 11 million organic compounds recognized by the *Chemical Abstracts Registry*, with over 30 000 compounds considered as chemicals of commerce, and perhaps 500 to 1000 routinely measured as members of industrial solvents, disinfection by-products, insecticides, PNAs, herbicides, organochlorine pesticides, PCBs, phenols, anilines, benzidines, and potential endocrine-disrupting compounds.

The quantitative determination is usually based on the absorption of light at some fixed wavelength or wavelengths, the selection of a wavelength for quantitation from the diode array detection (DAD) data, or, in fluorescence detection, the emission of light at some fixed wavelength (or band) while exciting at a fixed wavelength. For even better quantitation, one monitors the area under a peak or area of the whole spectrum. Thus one may monitor absorption at 214 nm for the sensitive detection of many compounds. Alternatively, one may monitor emission at 520 nm with excitation at 488 nm for the sensitive fluorescence detection of fluorescein and fluorescein-related compounds.

## Inorganic compounds

Electronic fluorescence techniques may be used to detect inorganic compounds, which exist as ions in solution, using one of several methods. Photoluminescence (or simply luminescence) refers to the photoexcitation of an analyte in solution, caused by absorption of visible or ultraviolet radiation, followed by emission at a longer wavelength. Luminescence is classified as either fluorescence or phosphorescence. Fluorescence is an allowed radiative transition from the first excited singlet state, while phosphorescence is a spin-forbidden transition from the first excited triplet state. Typical fluorescence lifetimes are nanoseconds to sub-microseconds, while typical phosphorescence lifetimes are microseconds to seconds. For many inorganic substances, fluorimetric methods are the best analytical methods.

(1) If the inorganic ion luminesces in solution it may be detected simply by placing the ion in an appropriate solution. This technique works primarily for rare earths and uranyl ($UO_2^+$) compounds. When combined with an inorganic reagent (e.g. HCl, HBr, etc.), the technique works for some group 3–5 elements.

Since the rare earths and uranyl compounds luminesce in solution, these ions may be detected by ‘native fluorescence’. Even though the bare ion fluoresces in these cases, the luminescence intensities can be greatly enhanced by appropriate complexation. The spectrum may be broad or narrow, depending upon the element and the nature of the electronic transition involved: the fluorescence spectrum is broad for Ce(III), Pr(III), and Nd(III), because they result from electron transitions from the 5d to the 4f shell. For example, cerium salts fluoresce in a broad band from 330 to 402 nm. In contrast, the luminescence spectra of Sm(III), Eu(III), Td(III), and Dy(III) are narrow, consisting of a few characteristic bands because they result from transitions of the 4f electron, and the f electron is little perturbed by its environment. For example, Gd has a single bright, narrow band at 311 nm. Sensitive fluorimetric methods have been developed for Pr, Nd, Sm, Eu, Gd, Tb, Dy, Er, and Tm. Detection limits range from 50 picograms to a microgram depending on the rare earth. Uranyl ($UO_2^+$) ions phosphoresce in several bands in the 460–620 nm range. The uranyl ion is typically detected in acid solution. The fluorescence efficiency depends on pH and on impurities. Many organic compounds can quench the luminescence from uranyl. In phosphoric acid or phosphate solutions uranyl can be detected at 5 parts per trillion.

(2) If the inorganic ion does not luminesce in solution, the ion may be combined with an organic ligand to yield a metal chelate. In favourable cases, the isolated ligand does not fluoresce, but the metal chelate does fluoresce, allowing for sensitive detection of the ion. In other cases where the complexing agent luminesces, the metal complex is still shifted enough in peak position to be detected even in the presence of the free agent. Over 40 different metals have been determined by this technique. When this method is applicable it is very sensitive and specific. For example, gold may be detected by its formation of a complex with Rhodamine B in 0.4 M HCl. Using excitation and detection wavelengths of 550 and 575 nm, respectively, as little as 20 ppb of Au can be detected, a level 25 times more sensitive than atomic absorption. In another example, cyanide ($CN^-$) can be detected at levels of 0.2 μg by complexing with *p*-benzoquinone to yield an intense green

fluorescence. Sulfide, thiosulfate, thiocyanate, ferrocyanide, and 26 other ions had no effect.

Several kinds of reaction between inorganic analyte and the organic reagent are possible: binary or ternary complex formation, substitution, redox, catalytic, and enzymatic. This technique can sometimes be applied to both cations and anions. Examples of commonly used reagents are 8-hydroxyquinoline, azo reagents, Schiff's bases, hydrozones, hydroxyflavones and chromone derivatives, and anthraquinones.

(3) If neither of the above approaches works, 'indirect detection' may be employed. The inorganic ion may quench the fluorescence of a second compound, in which case the analyte is detected as a decrease in the fluorescence of the second compound. For example, most assays for fluoride ($F^-$) are based on the decrease in fluorescence of an Al or Zr chelate due to subsequent fluoride complex formation ($AlF_6^-$ or $ZrF_6^-$). Quenching is generally less sensitive and less selective than fluorescence. Alternatively, the analyte may cause the release of a ligand which then reacts to form a fluorescent product.

Most fluorimetric determinations of inorganic species are equilibrium methods in which the reaction goes to completion. However, in some cases kinetic methods are used, in which the initial reaction rate is linearly related to the initial analyte concentration. For example, platinum(IV) can be detected at 0.2–0.6 ppm levels using a reaction with di-2-pyridylketone hydrozone, yielding a blue autooxidation product, whose rate of appearance is measured at 435 nm with an excitation wavelength of 359 nm.

## Organic compounds

Organic compounds are often of major concern in environmental applications. In this section we discuss the detection of organic compounds by electronic absorption and luminescence spectroscopy.

The technique of UV-visible absorption spectroscopy is a mature technology which is widely used for detection. A beam of light whose wavelength is in the visible or ultraviolet range traverses a sample containing the sample. The signal is detected as a decrease in the intensity of the light. UV-visible absorption spectroscopy has a moderate sensitivity and a low to moderate specificity. The specificity is low because many compounds absorb in the same general wavelength range. Because of the low specificity, this technique can be used for real-world samples of low to moderate complexity. The technique works best for unsaturated compounds (aromatic, polycyclic aromatic or heterocyclic). In these cases, often there is enough vibrational structure for relatively good specificity. It works for dyes and for colorimetric reaction products.

The second technique is UV-visible luminescence, meaning either fluorescence or phosphorescence. When this technique is applicable it can be very sensitive, especially with laser excitation. In aqueous solution, the sensitivity can be parts per billion to parts per trillion. Luminescence is more sensitive than absorption because in the former case one is detecting an increase from a very small background, while in the latter case one is detecting a decrease from a large background. The technique is applicable to aromatics, most polycyclic aromatics and their derivatives, and some heterocycles. For many polycyclic aromatic hydrocarbons, the vibrational structure is sufficient to greatly enhance specificity. It can be applied to many other analytes, using fluorimetric reagents. The selectivity can be enhanced by varying the wavelength of excitation and the wavelength of detection, or by using time or phase resolution.

One of the problems that arises in luminescence studies of large molecules is that the spectrum may be relatively broad and unstructured. Accordingly, in a real-world sample there can be interferences caused by other compounds. One technique to improve the specificity is 'synchronous luminescence', in which both the excitation wavelength and detection are scanned, usually with a fixed difference between the two wavelengths. This tends to sharpen the spectra, reducing the problem of interferences and thereby improving the specificity of the technique. On theoretical grounds, it makes more sense to scan the two wavelengths with a fixed difference in the photon energy, rather than a fixed difference in the wavelength.

Luminescence is applicable to polyaromatic compounds, fluorescent dyes, fluorimetric reaction products, polychlorinated biphenyls (PCBs), phenols, 50% of pesticides, many semivolatiles, many nonvolatiles, and petroleum oils. For luminescence to work, the compound must fluoresce. However, many large molecules absorb but do not fluoresce in the visible or UV. For such non-fluorescing compounds, absorption spectroscopy must be used instead (with much less sensitivity) or, alternatively, the organics can sometimes be complexed.

In practical applications, these spectroscopic techniques are often used for field screening, in which samples are screened in the field rather than being transported to the laboratory for traditional analysis, which can mean a delay of up to a month in obtaining the results of analysis. Recent technical advances include: more compact lasers, miniaturized optical hardware, new types of detectors, increased

use of fibre optics, and better computer software for spectral data processing and pattern recognition.

## The problem of spectral congestion

Many real-world samples consist of a large number of substances whose spectra are relatively broad and relatively featureless. This limits the ability to discriminate between compounds with similar or substantially overlapping spectra. There are two techniques that can be used to discriminate between compounds that are spectrally overlapping. One technique is discrimination based on fluorescent lifetime. Two compounds that spectrally overlap can be distinguished by their fluorescence lifetimes, provided that their fluorescent lifetimes are sufficiently different. The second technique is chromatographic discrimination; i.e. the use of a separation technique.

## Separations coupled with electronic spectroscopy

In some cases, e.g. forensic oil identification, the spectra of mixtures of fluorescent compounds may themselves be of interest for fingerprinting purposes, especially when combined with chemometric or pattern recognition techniques. Many real-world environmental problems involve large numbers of compounds, and the spectra of individual compounds are needed for identification and quantitation. Detection techniques are often preceded by a separation step. Historically, the workhorse separation technique of environmental analysis has been capillary GC combined with MS or with specialized GC detectors such as the flame ionization detector and the electron capture detector, or more recently combined with the liquid chromatography or capillary electrophoresis (LC-MS or CE-MS). In recent decades, the introduction into commerce and the environment of increasingly polar and nonvolatile compounds has resulted in new interest in liquid separations for environmental problems, resulting in the growth of liquid separation methods of analysis, and it is especially useful for polar or thermally labile compounds or for nonvolatile high relative molecular mass compounds. The following discussion is organized by type of separation technique or other analytical technique rather than by compound type.

The most common application of electronic spectroscopy has been as the detection technique for liquid separations where absorption spectroscopy is the workhorse technique. Absorption spectroscopy is well suited for liquid separations because most of the solvents of choice exhibit large UV transparent regions that enable the sensitive detection of compounds of interest. This is in contrast to other types of detectors that might universally detect both the solvent and the analyte of interest. For example, infrared detection is often compromised by either the presence of water or other solvents because of strong absorptions that obscure key portions of the useful wavelength range. Even mass spectrometry requires special interfaces to remove the solvent preferentially before introduction into the ion source even though solvent molecular masses are generally below 100 Da.

The detection limits for absorption spectroscopy are approximately $1 \times 10^{-5}$ absorption units (AU), varying with the extinction coefficient as a function of wavelength. (The absorption measured in AU is defined as the common logarithm of the initial to the final light intensity.) This results in detection limits in the ppm to low ppb range in favourable cases, providing sensitivity adequate for environmental analysis. In fluorescence detection, the possibility of detecting a single molecule exists. For environmental samples, often fluorescing interferences limit the sensitivity and in these cases pretreatments or liquid separations remove these interferences. In practical terms, fluorescence detection for appropriate molecules can be 100 to 1000 times more sensitive than absorption detection. As a result, detection limits in the low parts per trillion range can be achieved for selected analytes, limited only by excitation source intensities or fluorescence from capillary or solvent background fluorescence.

## HPLC

The most widely used detection technique for HPLC involves absorption at either a fixed wavelength or using a DAD that allows for wavelength selection and continuous or selected scans. In the filter photometric format, a wavelength such as 220 nm is selected by means of a filter placed in front of the deuterium lamp. **Figure 1** illustrates the detection of a series of polynuclear aromatic (PNA) compounds using their absorption at 220 nm. Common solvents used in reversed-phase separations (i.e. a hydrophobic adsorbent with water/organic solvent as mobile phase) have UV transparence from ~200 nm on up. A similar situation holds for normal phase separation (i.e. a polar adsorbent with hexane/polar solvent as mobile phase) where UV transparency is counted on. When the solvent choice includes toluene or some other absorber then a different wavelength or different detection technique must be chosen.

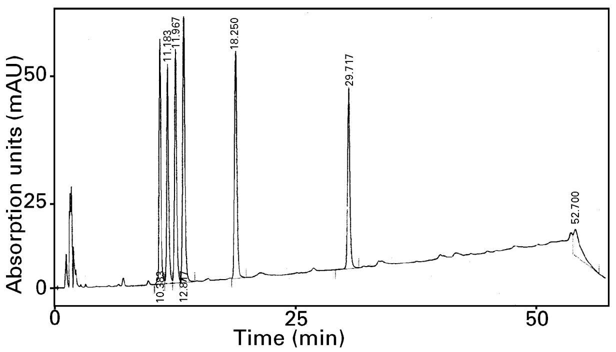



**Figure 1** HPLC separation of PNAs with detection by absorption spectroscopy at 220 nm.

When the DAD is used, a selection of wavelengths may be available for individual monitoring as well as the selection of the best wavelength for quantitation in consideration of background noise that may be largely chemical noise or interferences from the matrix. In the case of PNAs, the full absorption spectrum for benzo[*a*]pyrene is given in **Figure 2**.

Molecules that do not possess the appropriate chromophore may be detected when derivatized with a chromophore with high extinction coefficients at peaks in the visible spectrum to facilitate detection. For example, carboxylic acids can be derivatized with 7-methoxy-4-bromomethylcoumarin to give them a chromophore. Specialized books may be consulted for further information concerning derivatization. Other techniques include resorting to wavelengths below 200 nm under special conditions. In some cases, indirect detection can be used based on the presence of a background UV absorber in the mobile phase. The analyte displaces some of this absorber under certain conditions, resulting in a decrease in the absorption level in the detector cell which can be measured in the same manner as an increase in absorption. Further information on this technique can be obtained from specialized references on the subject.

When the analyte of interest possesses fluorescence character then an extremely sensitive detection can be designed based on the absorption/emission properties of the molecule. In **Figure 3**, the fluorescence detection of benzo[*a*]pyrene is illustrated using a excitation of 325 nm with an emission at 425 nm. In this manner, low pg quantities of compound can be detected, making this one of the most sensitive detections in analytical chemistry. Increasingly, charged coupled devices (CCDs) are being utilized to give even higher sensitivity especially with laser excitation, and whole spectra rather than single wavelengths are being recorded. Techniques such as fluorescence lifetime detection with phase or frequency modulation can also be used to separate peaks that would otherwise appear to overlap.

## Capillary electrophoresis (CE)

CE has emerged as a pre-eminent separation technique for ionic analytes based on free zone electrophoresis. The separation is based on differing mobilities among the ions. In analogy to HPLC, an absorption detection is a fairly universal detector. The DAD response for a group of phenols is illustrated in **Figure 4**. The influence of ionization on the absorption spectrum of phenols is evident in the shift to longer wavelengths that results from the anionic character. One of the chief drawbacks of CE applications using absorption detection has been the relatively high concentration detection limit (i.e. ppm range rather than ppb range of HPLC). This is a direct result of the capillary format of the separation. In capillary separations, normally tens of nanolitres of sample are injected on-column. This is a 1000-fold smaller volume than in HPLC. In addition, the capillary format usually demands on-column detection to avoid band broadening from post-column detection approaches. The diameter of the capillary is about 75 μm, resulting in a very short optical path. There are specialized techniques for overcoming these limitations and more specialized references should be consulted.

The concentration detection limit problem can be overcome for molecules possessing fluorescence. In the case of phenols, illustrated in **Figure 5**, a frequency-doubled laser operating at 244 nm will enable sensitive detection using laser-induced fluorescence (LIF). In this case, a filter arrangement allows detection of emission light in the 355 to 425 nm range. The use of lasers in this context improves detection limits because of the light intensity as well as the intrinsic narrow beam which allows efficient focussing of the light onto the capillary bore carrying the sample.

## Immunochemical analysis

Immunochemical analysis is a convenient approach for environmental analysis provided appropriate antibodies and format have been developed. Often, the assay depends on a colorimetric determination for its quantitative result. The advantage of immunochemical analysis is its specificity as well as its speed. Specific references should be consulted on applications.

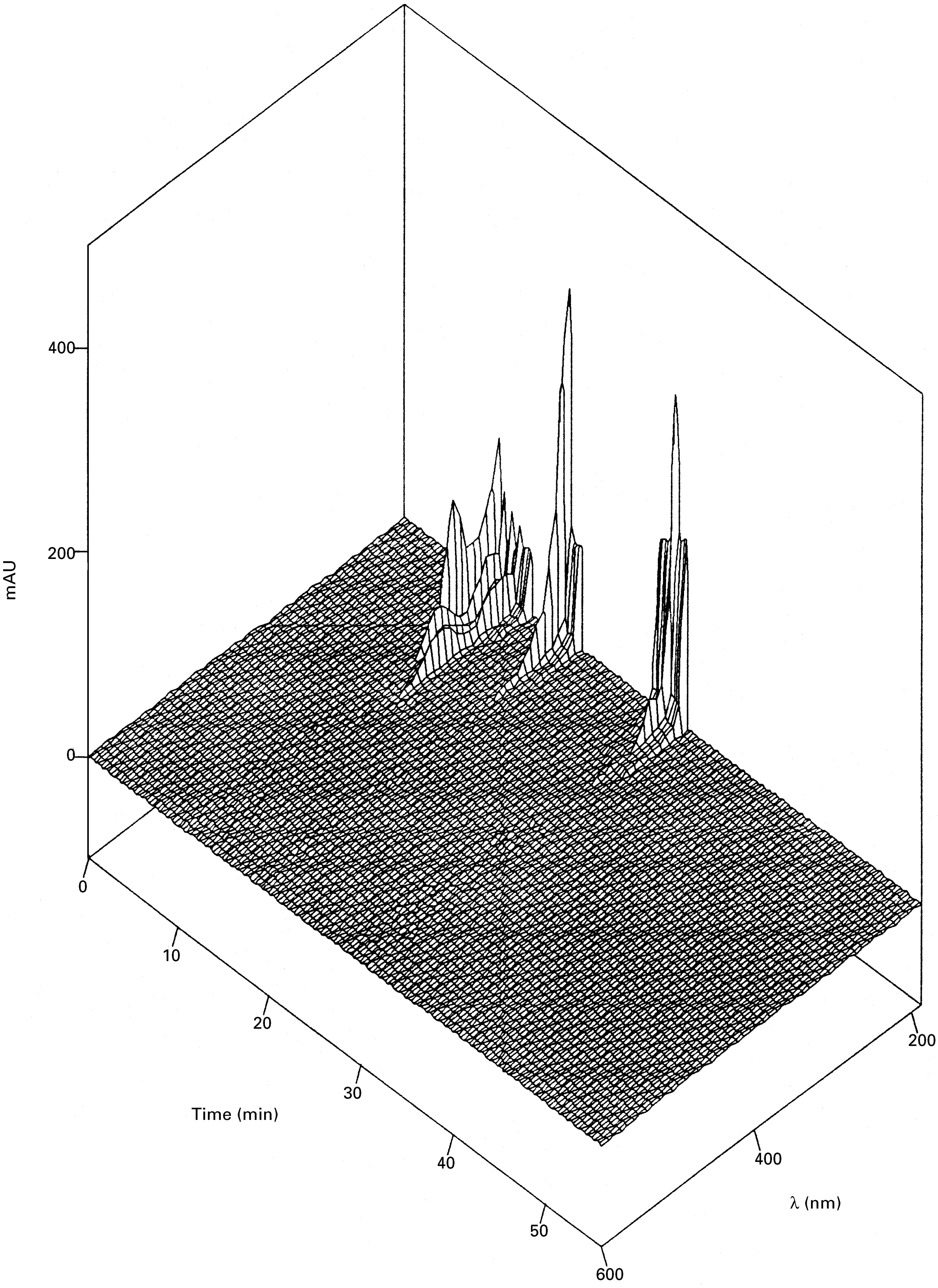



**Figure 2** HPLC separation of selected pesticides with DAD with a full absorption spectrum (200 to 600 nm) illustrated for each in a three-dimensional format.

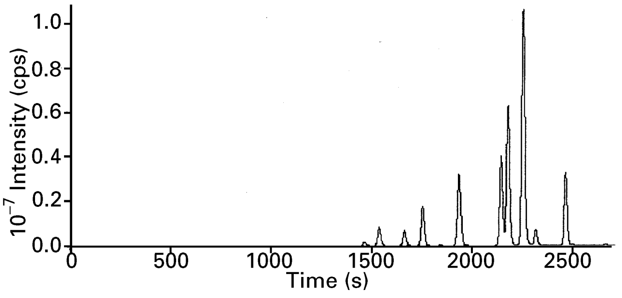



**Figure 3** HPLC separation of selected PNAs using fluorescence detection with excitation at 325 nm (xenon lamp) and emission at 425 nm.

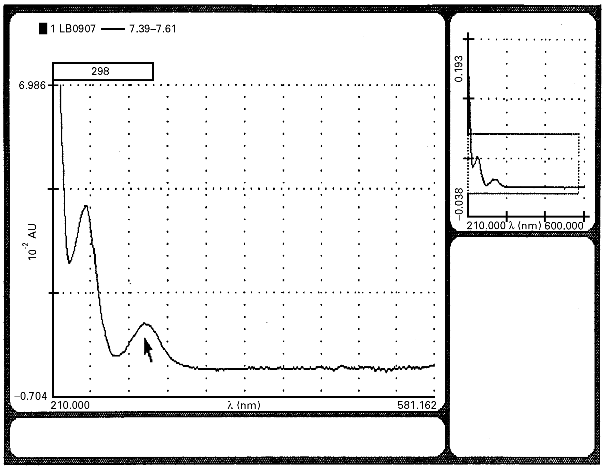



**Figure 4** DAD detection of phenols in capillary electrophoresis. The arrow indicates the wavelength position of 298 nm.

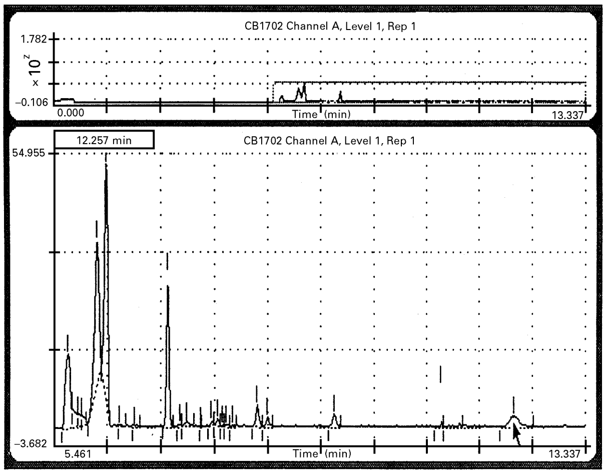



**Figure 5** CE-LIF detection of phenols using a frequency-doubled laser. Excitation at 244 nm and emission at 355–425 nm. The arrow indicates the migration time position of 12.257 min.

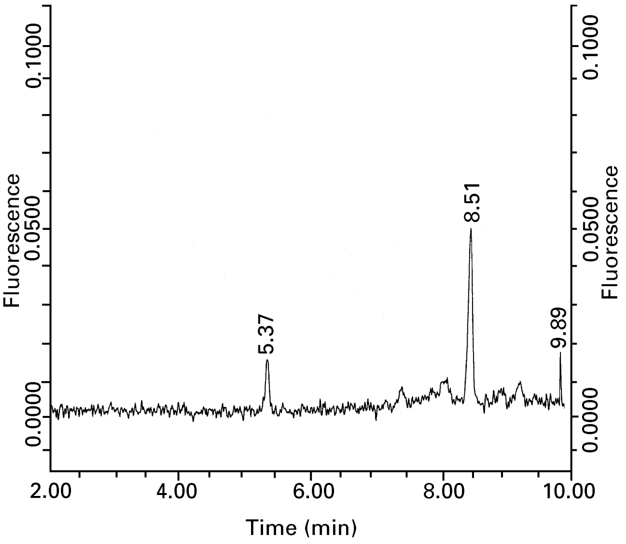



**Figure 6** CE-LIF detection of probe peak (fluorescein-labelled 2,4-dichlorophenoxyacetic acid) in an immunoassay format of herbicide analysis. Excitation at 488 nm and emission at 510–530 nm.

An emerging format for the assay depends on CE-LIF determination of a fluorophore-labelled analyte (probe) in either a competitive or noncompetitive assay with the unlabelled analyte as they find binding sites with the antibody. **Figure 6** illustrates the concept for a fluorescein-labelled analogue of 2,4-dichlorophenoxyacetic acid (2,4-D) where the change in the amount of the probe peak indicates whether unlabelled 2,4-D was present in the sample.

## Gas chromatography

One detection technique for GC, called atomic emission detector (AED), is of fairly common occurrence and makes use of spectroscopic properties for detection. The elements of a compound that elutes from the column are determined in a He plasma based on the emission spectra of the individual elements (e.g. C, H, O, N, S, Cl, etc.). Reference to more specialized works should be made. The AED technique is mentioned for the sake of completeness, although since AED relies on atomic spectroscopy emission, it is outside the scope of this article, which is mostly restricted to molecular electronic spectroscopy.

## Thin-layer chromatography

Thin-layer chromatography (TLC) is usually coupled to either absorption or fluorescent detection of complexed species and is common in many fields of analysis. The applications of TLC to environmental analysis have been less frequent than with other

forms of separation. It has recently been demonstrated that UV-visible absorption, near-IR absorption, fluorescence, and phosphorescence can be accomplished for hundreds of spots on a single TLC plate by using CCD detection.

## Emerging techniques

The application of molecular electronic spectroscopy as a detector for separations is expected to continue to grow. Optical techniques are a primary detection approach for the hybrid technique of capillary electrochromatography which combines aspects of CE and HPLC, especially capillary LC. To increase throughput of sample handling, arrays of capillaries have been built with optical detection via CCD cameras. Using CE-LIF approaches, hundreds of capillaries can operate simultaneously on samples with simultaneous detection. Further advances are expected in the areas of remote sensing, immunochemical assays, capillary arrays, and detector sensitivity.

*See also:* **Chromatography-NMR, Applications; Chromatography-MS, Methods; Environmental and Agricultural Applications of Atomic Spectroscopy; Luminescence, Theory.**

## Further reading

Acree Jr WE, Tucker SA and Fetzer JC (1991) Fluorescence emission properties of polycyclic aromatic compounds in review. In: *Polycyclic Aromatic Compounds*, Vol 2, 75–105. Lausanne, Switzerland: Gordon and Breach.

*Analytical Chemistry*. This journal is generally useful as a survey of the field. In odd years, they review applications, while on even years they review fundamentals (i.e. techniques).

Camalleri P (ed) (1993) *Capillary Electrophoresis: Theory and Practice*. Boca Raton, FL: CRC Press.

Fernandez-Gutierrez A and Munoz de la Pena (1985) Determination of inorganic substances by luminescence methods. In: Schulman SG (ed) *Molecular Luminescence Spectroscopy Methods and Applications*, Vol I, Chapter 4. New York: Wiley Interscience. Later volumes in the series include Vol II (1987) and Vol III (1993).

Grosser ZA, Ryan JF and Dong MW (1993) Environmental chromatography methods and regulations in the United States of America. *Journal of Chromatography*. **642**: 75–87.

Guilbault GG (1990) Inorganic substances. In: Guilbault GG (ed) *Practical Fluorescence*, Chapter 5. New York: Marcel Dekker.

Karcher W, Devilliers J, Garrigues Ph and Jacob J *Spectral Atlas of Polycyclic Aromatic Compounds*. Vol I (1983), II (1988), and III (1991). Boston: Kluwer Academic.

Lakowicz JR (1991) *Topics in Fluorescence Spectroscopy*, Vol I. Plenum Press. Later volumes include Vol II Principles (1992), Vol II *Biochemical Applications* (1993), and Vol IV *Probe Design and Chemical Sensing* (1994).

Lawrence JF (ed) (1984) *Liquid Chromatography in Environmental Analysis*. Clifton, NJ: Humana Press.

Sweedler JV, Ratzlaff KL and Bonner Denton M (1994) *Charge Transfer Devices and Spectroscopy*. New York: VCH Press.

Vo-Dinh T and Eastwood D (ed) (1990) *Laser Techniques in Luminescence Spectroscopy*. ASTM STP 1066. Philadelphia: American Society for Testing and Materials.

Wehry EL (ed) (1981) *Modern Fluorescence Spectroscopy*. New York: Plenum Press.

# EPR Imaging

**LH Sutcliffe**, Institute of Food Research, Norwich, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

The basic requirement for an EMRI experiment is that a species having unpaired electrons be present. Although free radicals and some transition metal compounds fulfil this requirement, there are not many circumstances in which these are present naturally in sufficient concentration to give a suitable signal; this is in contrast to MRI where the ubiquitous proton can be used to study most materials. Fortunately, the high sensitivity of EPR compared with NMR (arising from the difference in electron and nuclear magnetogyric ratios) enables paramagnetic material to be introduced in low concentrations into systems under investigation with minimum interference. However, NMR has an advantage over

EPR in that pulse techniques can be used to boost the signal-to-noise ratio: both the irradiating frequency and the applied magnetic field gradients can be pulsed. The main reason why most EPR experiments are carried out in the continuous-wave (CW) mode is because spin-lattice relaxation times for paramagnetic materials are of the order of microseconds (compared with hundreds of milliseconds for proton NMR) and this causes considerable difficulties in achieving the short times required for the pulse experiment. A further difficulty is the spectral bandwidth of an EPR spectrum: this, coupled with a large line width (three orders of magnitude greater than in NMR) means that very large magnetic field gradients (100-1000 times greater than in MRI) have to be applied. However, for the few systems not presenting these difficulties, FT-EMRI will become the method of choice. The ability to employ pulse techniques for both the irradiating frequency and for the field gradients also allows a much wider variety of experiments to be carried out with MRI compared with EMRI.

## Spin probe

The paramagnetic materials added to systems to obtain signals in EMRI are referred to as spin probes. These can be transition metal compounds such as those of manganese(II) or vanadium(IV) but they are more likely to be organic free radicals. In addition to having high thermal, chemical and metabolic stability (for biological applications), these radicals should preferably have spectra with very narrow line widths in order to maximize the signal-to-noise ratio. The peak-to-peak line width of a first-derivative spectral line can be represented by the equation:

$$\Delta B_{PP}(m_I) = A + Bm_I + C(m_I)^2$$

where A, B and C are constants and $m_I$ is the magnetic quantum number of a given line. The constant A is dependent on the g factor anisotropy and experimental factors, B upon the g factor and the nuclear hyperfine interaction anisotropies and C upon the nuclear hyperfine interaction anisotropy. EPR spectra are normally recorded in the first-derivative mode but, obviously, the absorption mode is more convenient for EMRI. The most common stable free radicals used as spin probes are nitroxyls and there is both g factor and hyperfine interaction anisotropy of the $^{14}N$ nucleus. A typical set of values at X-band for a nitroxyl radical having unresolved proton hyperfine structure is A = 0.116, B = −0.010 and C = 0.006 mT. Thus the observed line width is of the order of 0.1 mT and there is little scope for reducing the line width further for imaging at X-band: a nitroxyl radical in a nitrogen-purged low-viscosity solvent can show a resolution of proton hyperfine couplings with observed line widths smaller than 0.02 mT. The three lines of a nitroxyl radical, in addition to having a large bandwidth, also represent a ‘waste’ of signal since it is more practical to image a single line. Some improvement can be made by substituting $^{14}N$ with $^{15}N$ and/or by reducing the proton hyperfine interaction by deuteration. However, there is a soluble commercial single-line free radical based upon the $^{\bullet}C(Ar)_3$ structure which has a line width of only 0.006 mT in a deoxygenated solvent: the sharp line is observed because of the small g factor and hyperfine anisotropies for a carbon-centred radical. One of the main aims in biological studies is to spin-label drugs and to image their organ compartmentation, and hence there are some difficult chemical problems to be solved. For some studies, it is possible to use solid particles as spin probes in which case very narrow single lines are attainable. An example is lithium phthalocyanine which has a line width of 0.002 mT in a deoxygenated system. Other single-line solids are special forms of coal, such as fusinite, and carbohydrate chars.

## Theory

Resonance imaging depends on the application of magnetic field gradients to a sample. If two samples are placed at different positions of the spectrometer magnetic field then only one of them will be at resonance, the second one can be brought to resonance by changing the spectrometer field or by applying a magnetic field gradient. Since field gradients can be applied in three dimensions, it is possible to determine the relative positions of both the samples, i.e. to produce an image. Mathematical methods are of central importance in any type of image reconstruction and they were developed some time ago for radiographic tomography. The necessity of using CW techniques in EMRI limits the number of imaging strategies that can be employed.

### Spatial imaging

Two-dimensional images of spin density in the xz or yz planes (z axis is parallel to the main magnetic field) can be obtained: for a $n \times n$ image, $n\pi/4$ projections are required. Most spatial images of paramagnetic centres with multiline spectra have been constructed by selecting a single line, using a limited

field gradient to avoid overlap with other lines. The resolution can be enhanced by deconvolution of the line shape.

### Spectral-spatial imaging

The advantages of spectroscopic imaging compared with spatial imaging methods are:

(i) the problems of resolution and gradient magnitudes owing to the presence of hyperfine lines can be overcome;
(ii) separating spectral and spatial components by assigning them to different imaging axes reduces the resolution degradation arising from the finite width of the spectral lines;
(iii) asymmetric and multiline spectra can be imaged whereas only symmetric line shapes should be used for spatial imaging

Images are obtained as a function of sample position by recording spectra at a series of magnetic field gradients which correspond to projections in the spectral-spatial plane. A pseudo-object of length B is examined in the spectral dimension and length L in the spatial dimension. An angle (α) defines the orientation of a projection relative to the spectral axis. The maximum field gradient, $G_{max}$, is related to B and L by

$$G_{\max} = \tan(\alpha_{\max})B/L$$

The maximum magnetic field swept for each projection is given by

$$\text{Sweep width} = 2^{1/2}\,B\cos\alpha$$

For given values of $G_{max}$ and B the ratio $\tan(\alpha_{max})/L$ is fixed. Most image reconstruction algorithms require projections over 180° with equal angular increments. If a complete set of projections is used for spectral-spatial imaging, the number of projections determines the values of $\alpha_{max}$ and $G_{max}$. Thus for 64 projections, L = 0.8 cm and B = 3.2 mT, $\alpha_{max}$ = 88.59°, $G_{max}$ = 162.5 mT and the sweep width required is 183.9 mT. Spectral-spatial images can be reconstructed from incomplete sets of projections and this has the advantage of reducing both the maximum field and sweep width requirements. Assuming experimental data were obtained for 60 out of 64 projections and estimates were made for the 'missing' projections, then $G_{max}$ and the sweep width requirements are reduced to 32.4 and 37.0 mT, respectively.

### Projection reconstruction methods

Important image reconstruction procedures are:

(i) back projection;
(ii) simultaneous, algebraic and iterative least-squares;
(iii) two-dimensional and filtered back projection;
(iv) use of incomplete data sets from which it is also possible to obtain slices from the image.

### Modulated magnetic field gradients

These gradients, $B_{GM}$, provide fields in addition to the main magnetic field that have both time- and space-dependent amplitudes. The time-independent spectrum can only be recorded in the z plane where the amplitude of $B_{GM}$ is zero. Images can be obtained without the use of difficult mathematical deconvolution by scanning the plane of the field gradient across the sample.

## X-Band imaging

Imaging at about 9 GHz very convenient since most commercial EPR spectrometers operate at this frequency and the imaging system can be fitted as an accessory. Gradient coils can be fitted to the outside or to the inside of a standard resonant cavity. **Figure 1** shows the arrangement for the former: to generate sufficiently large gradients, ~1 kW of power is needed and hence provision for water cooling has to be made. **Figure 2** shows how small coils, used for microscopy, can be placed inside the cavity. X-Band frequencies can be used for non-polar samples up to 10 mm in diameter but polar samples are restricted to 1.5 mm diameter because of the absorption of the microwave energy. An estimate of the achievable resolution is a pixel of 10 × 10 × 10 μm that contains 1 mM spins. Note that the latter concentration cannot be increased appreciably without causing line broadening from spin-spin interactions.

An example of two-dimensional spectral-spatial imaging is shown in **Figure 3**. This depicts the projection-reconstruction image obtained from four different-sized specks of DPPH (diphenylpicrylhydrazyl), each of which gives an exchange-narrowed single spectral line. **Figure 4** is an example of a two-dimensional spectral-spatial image, showing a phantom that contains solutions of three types of free radicals which each have some hyperfine structure.

X-Band EMRI has been applied to the study of the diffusion and distribution of oxygen in models of biological systems with the aim of studying diffusion

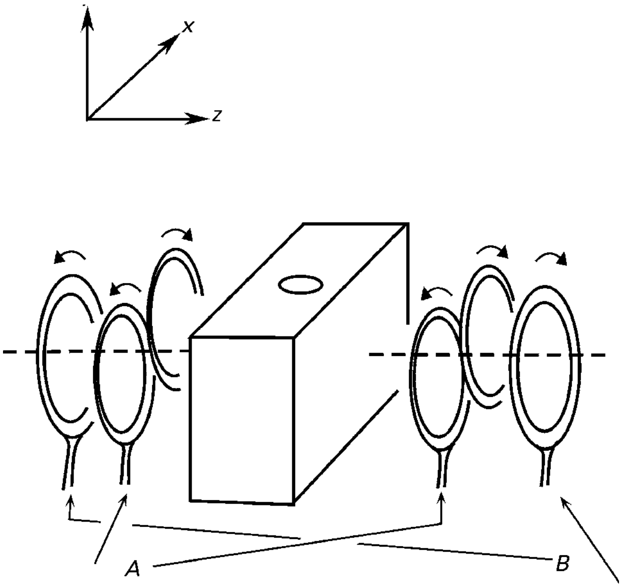



**Figure 1** The arrangement of figure-of-eight (A) and anti-Helmholtz (B) coils fitted to the outside of a rectangular resonant cavity, to provide gradients $\delta B_z/\delta_x$ and $\delta B_z/\delta_z$ respectively.

and metabolism in vivo. The method is based on the fact that oxygen causes paramagnetic broadening of the spin probe lines. Other applications have been (i) for the determination of depth profiles of E' defects in X-irradiated silica, (ii) for radiation chemistry, where a dosimeter has been designed using crystalline L-α-alanine which can be used to check the spatial variation of the radiation, (iii) a variation of this theme to study the distribution of alkyl and alkoxy radicals in γ-irradiated polypropylene, (iv) to study diffusion and transport of nitroxyl radicals, for example into zeolites, into solid polymers and into the caryopsis of wheat seeds and (v) to study paramagnetic humic substances found in soil. **Figure 5** shows an image obtained from the soaking of these into a wheat grain.

There are several strategies for EMRI microscopy, and an example is shown in **Figure 2**. Arrangements of this type are capable of giving a resolution of 50 μm. EMRI microscopy can be used for (i) small-scale dosimetry of materials exposed to high-energy radiation, (ii) ceramic and semiconductor technology

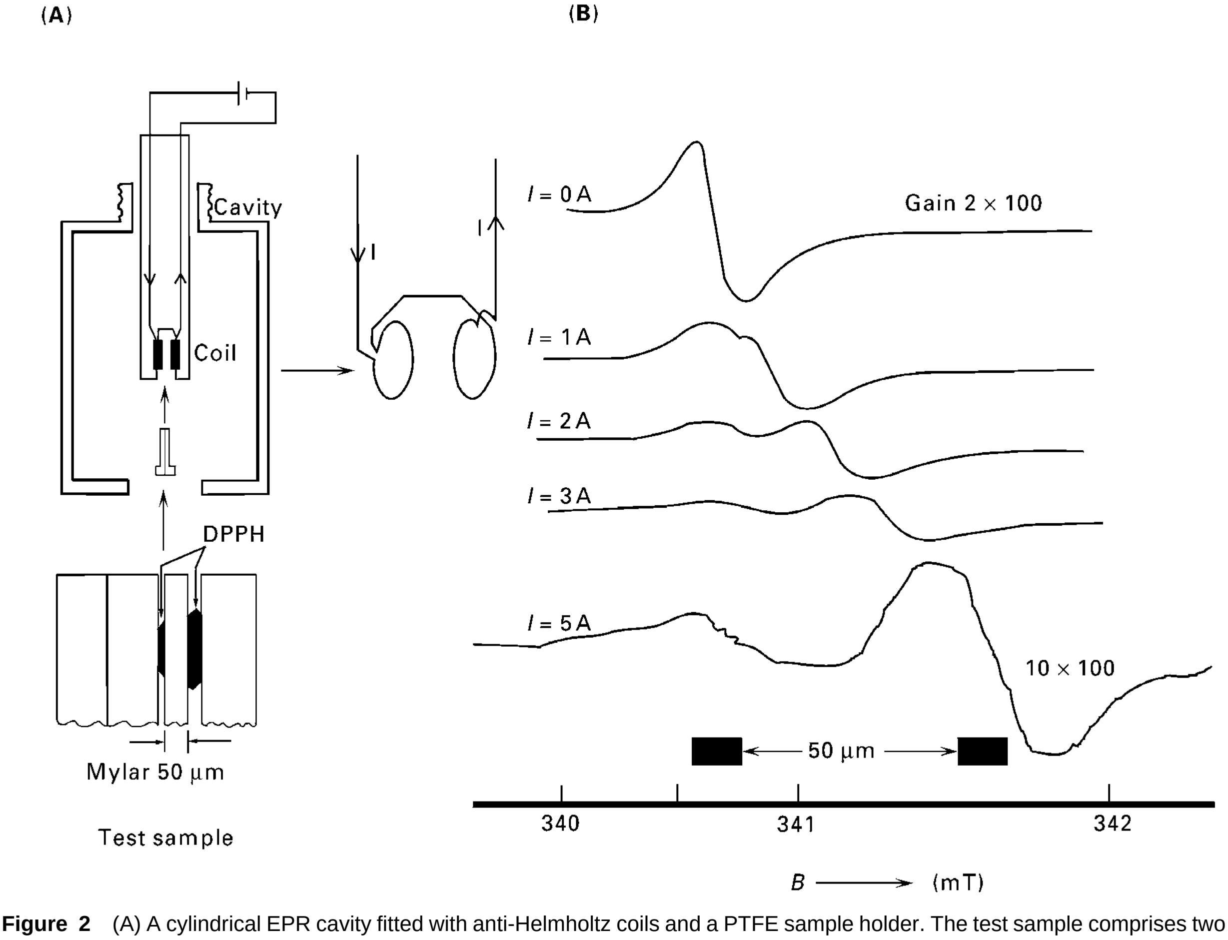



**Figure 2** (A) A cylindrical EPR cavity fitted with anti-Helmholtz coils and a PTFE sample holder. The test sample comprises two portions of DPPH powder separated by a 50 μm mylar film. (B) EPR spectra of DPPH for currents ranging from zero to 5 A in the gradient coils. The signal separation of about 1 mT corresponds to a field gradient of 20 T $m^{-1}$. Reproduced with permission from Ikeya M and Miki T (1987). *Japanese Journal of Applied Physics* **26**: L929.

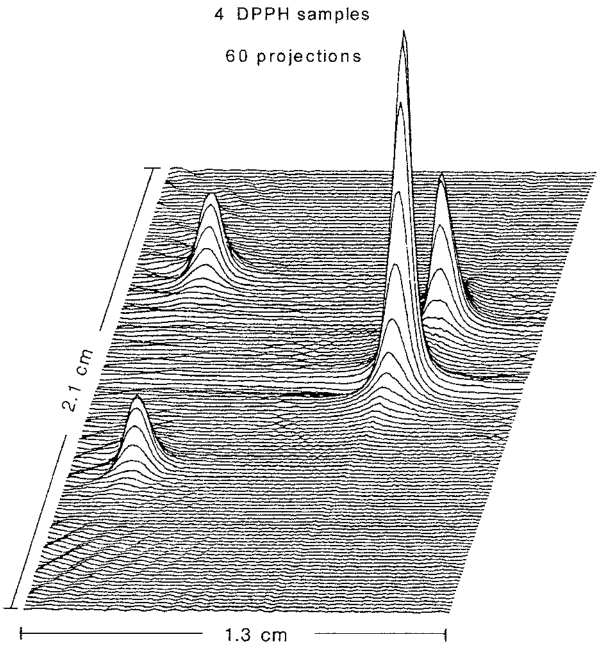



**Figure 3** A 2D spatial X-band image of four specks of solid DPPH. Sixty EPR spectra were obtained with maximum field gradients $\delta Bz/\delta z$ = 10 mT $cm^{-1}$ and $\delta Bz/\delta y$ = 6 mT $cm^{-1}$. The spectra constitute a full set of projections over 180°. Reproduced with permission from Eaton GR and Eaton SS (1995) Introduction to EPR imaging using magnetic-field gradients. *Concepts in Magnetic Resonance* **7**: 49.

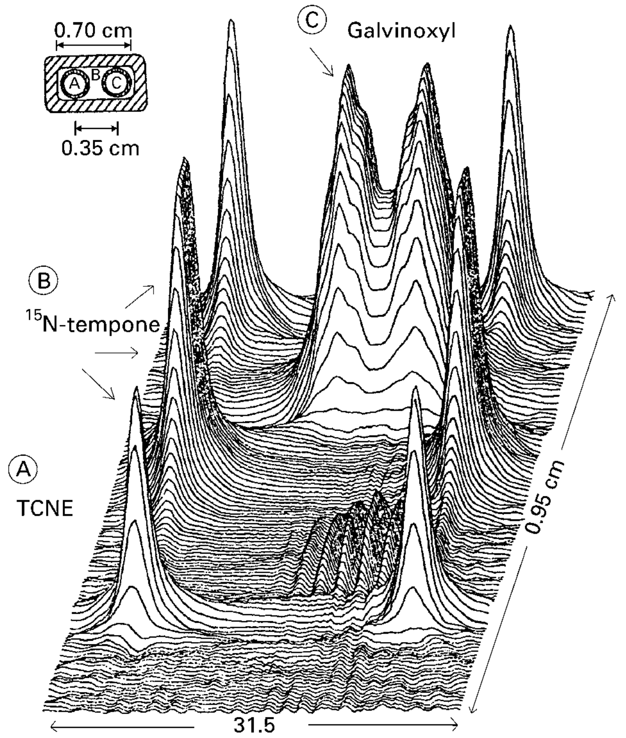



**Figure 4** Spectral-spatial X-band image of a composite sample of the tetracyanoethylene (TCNE) radical in tube A, the galvinoxyl radical tube C, and $^{15}N$-labelled tempone in region B. The maximum magnetic field gradient was 30 mT $cm^{-1}$. The image was reconstructed iteratively from 60 experimental projections and four missing projections. Reproduced with permission from Eaton GR and Eaton SS (1995) Introduction to EPR imaging using magnetic-field gradients. *Concepts in Magnetic Resonance* **7**: 49.

and (iii) the distribution of nitrogen centres in diamonds.

## L-Band imaging

Most current biological imaging systems operate at L-band frequencies (~1 GHz) which allows larger and more 'lossy' samples to be studied (microwaves of this energy can penetrate about 6 mm into a biological specimen) than is possible at X-band. A fairly standard spectrometer can be adapted for this work but loop-gap resonators, external surface coils and re-entrant cavities are more useful than a standard resonant cavity.

A non-biological application of L-band EMRI has been the study of the oxidation of cylindrical coal samples at 150°C where it was found that there is a homogeneous distribution of radicals across the samples. An important biological application is the study of oxygen concentrations ('oximetry') in a perfused rat heart using implanted solid paramagnetic particles: this rapid (1-5 min) multisite in vivo spectroscopy can be applied to a wide range of problems in experimental physiology and medicine. Other biological applications involve the imaging of soluble free-radical distribution: some examples include (i) a three-dimensional image of the head of a living rat, showing that the nitroxyl-deficient part corresponds to the brain, and (ii) accurate images of the vessel lumen and wall of a perfused rabbit aorta. When a glucose char (having a single EPR line) is introduced into an isolated biological organ, the resolution obtainable is 0.2 mm (**Figure 6**) whereas a soluble nitroxyl spin probe allows a resolution of only 1-2 mm. The single-line soluble probe $^{\bullet}C(Ar_3)$ has allowed images of a rat kidney to be visualized with a resolution of 100 μm for samples up to 25 mm. As in MRI, data acquisition can be gated to overcome the problems associated with rapidly-moving organs such as a beating heart.

A biologically important naturally occurring free radical is nitric oxide but, unfortunately, it has very broad EPR lines and thus cannot be imaged. However, its generation in rat hearts subjected to global ischaemia has been mapped by using the spin trap bis(N-methyl-D-glucamine dithiocarbamate)iron(II). The latter forms a stable nitroxyl free-radical with nitric oxide that can be imaged.

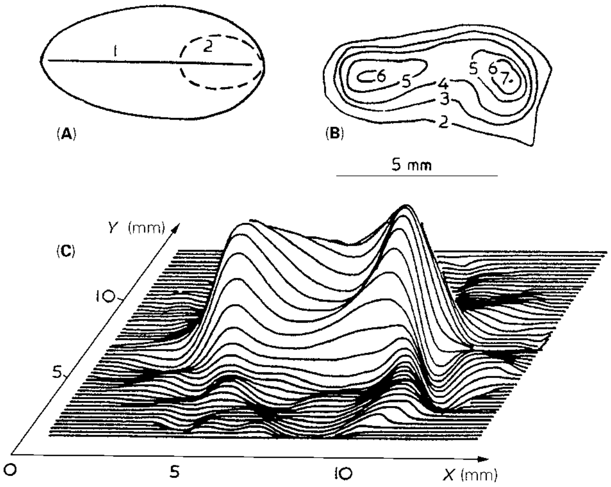



**Figure 5** (A) The schematic structure of a wheat grain (1 is the vascular bundle and 2 is the germ). (B) The spatial distribution of the humic substance signal in the wheat grain. (C) The relief surface of the spin density distribution in the wheat grain. Reproduced with permission of Academic Press from Smimov AI, Yakimchenko, Golovina OE, Bekova SK and Lebedev YS (1991) *Journal of Magnetic Resonance* **91**: 386.

Four-dimensional spectral-spatial EMRI has been developed (1000 projections were used to construct 32 × 32 × 32 × 32 pixel images) and, while considerable data acquisition times are needed, it is feasible to obtain spectral information for a biological structure such as a rat heart.

## High-field imaging

A high-field imager using a superconducting magnet operating at 5 T (140 GHz) has been constructed that derives its very large field gradients from ferromagnetic discs and wedges. This type of instrument is aimed at studying processes that have spatial ranges of less than 10 μm such as diffusion, defect formation in irradiated samples and polymer fracture under stretching and loading. High spectral resolution can be achieved because of the high absolute point sensitivity. A commercial EPR spectrometer is available for W-band (95 GHz) that can be adapted for high-field EMRI.

## Radiofrequency imaging

It has been stated above that it is difficult to obtain images at L-band frequencies of lossy objects larger than 25 mm. Thus to image large samples, such as live animals, lower frequencies have to be used. There is also an added incentive to work at radiofrequencies because of the low cost of components compared with those used at microwave frequencies.

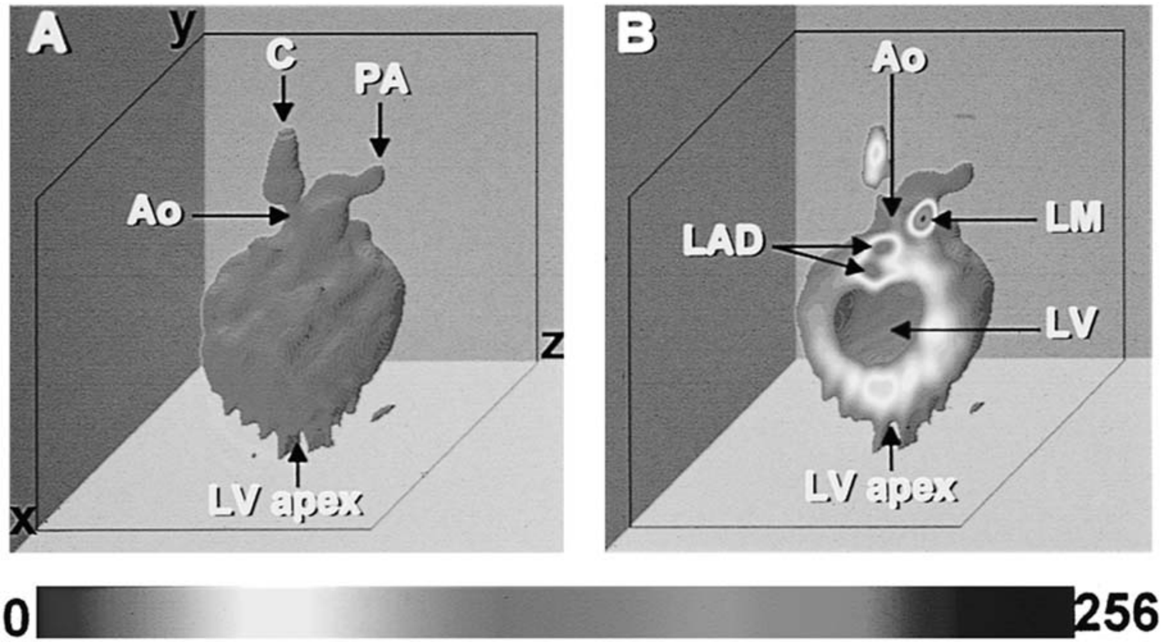



**Figure 6** Three-dimensional images (25 × 25 × 25 mm$^3$) of an ischaemic rat heart infused with a glucose char suspension. (A) Full view of the heart. (B) A longitudinal cut out showing the internal structure of the heart. C is the cannula; Ao the aorta; PA the pulmonary artery; LM the left main coronary artery; LAD the left anterior descending artery; and LV the left ventricular cavity. Reproduced with permission from Kuppusamy P, Wang P and Zweir JL (1995). *Magnetic Resonance in Medicine* **34**: 99. (See Colour Plate 18)

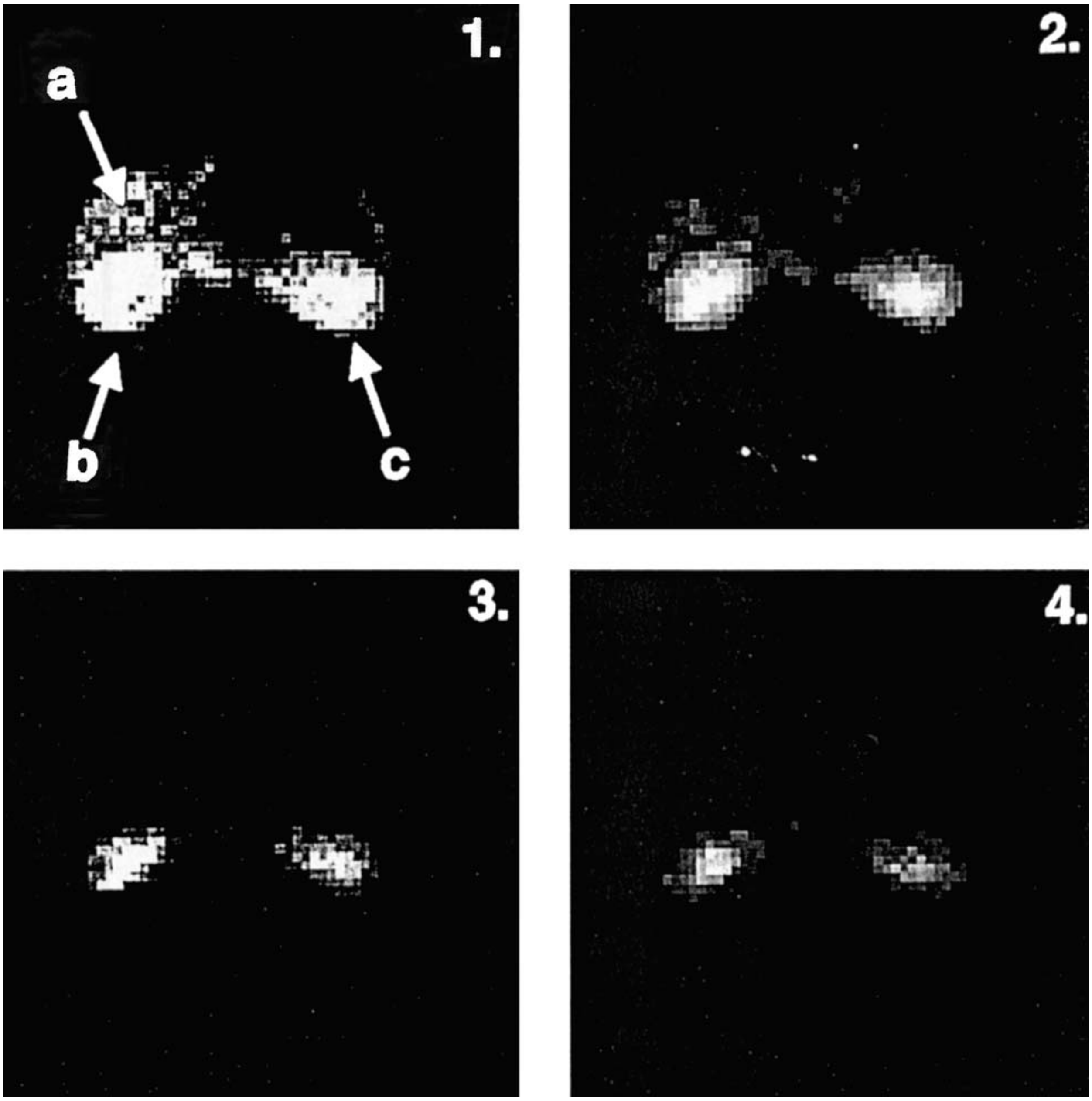



**Figure 7** Set of proton Overhauser images from the abdomen of a supine, 350 g, anaesthetized rat, following the administration of a solution of hydrogencarbonate-buffered 3-carboxy-2,2,5,5-tetramethylpyrrolidine-I-oxyl (PCA). Image matrix size 64 × 64; field-of-view, 10 × 10 cm; slice thickness, 2 cm; field strength, 10 mT; EPR Irradiation frequency, 238.7 MHz. Image no. 1 was collected 1 min after administration of PCA and the remaining images were collected at 6, 11 and 16 min post-injection respectively. The labels are (a) liver; (b) right kidney (c) left kidney. Reproduce with permission of Wiley from Lurie DJ (1995). In Grant DM and Harris RK (eds) *The Encyclopedia of Nuclear Magnetic Resonance.*

The consensus of opinion is that the most suitable frequency is about 300 MHz, which has a penetration depth of about 7 cm in systems of high relative permittivity. Theory suggests that the signal-to-noise ratio and resolution of EMRI are very much worse than they are for MRI. However, the high point sensitivity of EMRI compared with MRI provides advantages for probe experiments and thus EMRI should include, at some level, a spectral dimension. At radiofrequencies the line widths of nitroxyl spin probes are similar to those at X- and L-bands since the dominant factor for these probes is the anisotropic hyperfine interaction which is field independent. It should be noted that, for nitroxyls, account has to be taken of the Breit- Rabi effect. As might be expected, the experimental approach has features in

common with MRI; for example, ‘bird-cage’ resonators can be used. Other useful types of resonators are surface coils, re-entrant cavities and loop-gap resonators. The main magnetic field required for 300 MHz is about 11 mT, which can be obtained from a simple air-cored Helmholtz design. Because a non-ferrous magnet is used, the main field windings can also be used for modulation at a frequency of about 4 kHz. The x, y and z gradient coils can be wound on concentric cylinders and an advantage of the bird-cage resonator is that it can be mounted concentrically inside the cylinder, whereas a loop-gap resonator would have to be mounted across the cylinder making sample access more difficult. A bird-cage resonator has been used to implement longitudinally detected EPR, which has the advantage of giving less noise from animal motion. The latter can also be minimized by automatic frequency control of the radiofrequency. A multipole magnet has been constructed which allows the main field and two of the gradient fields to be generated; the third gradient is obtained by inserting an air-cored coil into the magnet. In another design, two Helmholtz loops are splayed to produce the z direction field gradient.

A variety of studies using RF EMRI have been carried out on living mice. These include (i) oximetry in tumour tissues and the response to perfluorocarbons for their potential use in radiotherapy, (ii) the detection of hydroxyl radicals (by spin trapping) in $\gamma$-irradiated tumours, (iii) the distribution of nitroxyl radicals with distinct pharmacological abdominal compartment affinities and (iv) quantitative measurement of the viscosity of a non-vascular compartment.

## Overhauser imaging

The broad lines encountered in EPR lead to poor spatial resolution compared with MRI and hence it is unlikely that a whole-body human-size imager is possible using the techniques described above. There is, however, an imaging strategy that could lead to a large-sample imager by taking advantage of the Overhauser Effect - techniques using the effect are also known as PEDRI (proton- electron double resonance imaging) and DNP (dynamic nuclear polarization) imaging. This is done by using a double resonance technique in which low-field proton MRI is combined with EPR. The EPR spectrum of a free radical is irradiated during the collection of a proton NMR image. The Overhauser effect produces an enhancement (by a factor of about 100) of the NMR signal in the parts of the sample that contain free radicals. The big advantage is that MRI resolution is achieved without having to use large field gradients but the disadvantage is that spectral information cannot be obtained. To overcome this drawback, an instrument has been constructed that combines Overhauser imaging with EMRI. In Overhauser imaging, an EPR frequency lower than 300 MHz has to be used to avoid heating of a biological sample. When nitroxyl spin probes are being used it is possible to irradiate each of the three $^{14}N$ manifolds simultaneously to achieve a maximum Overhauser enhancement. Resonators are similar to those used in radiofrequency imaging, namely loop-gap or bird-cage structures. Normally, a MRI image is obtained and subtracted from an Overhauser MRI image for the slice under study. This is done by using an interleaved pulse sequence, each alternate NMR excitation being preceded by EPR irradiation. An example of images obtained in this way is shown in **Figure 7**.

By cycling the main magnetic field, large samples can be imaged by reducing the EPR frequency required to about 60 MHz without impairing the NMR signal-to-noise ratio. Most of the applications of Overhauser imaging are similar to those given above for radiofrequency imaging except that it is possible to study much larger samples. Oximetry, however, is carried out differently since it is possible to measure quantitatively the reduced Overhauser enhancement caused by the oxygen. An interesting aspect of Overhauser imaging is that it can be carried out at fields as low as that of the earth, i.e. 0.06 mT. The reason is that the enhancement is still high for magnetic fields lower than the hyperfine coupling of the spin probe, ~1.5 mT (70 MHz) for nitroxyl radicals. Thus it is possible to build an imager having a main field of about 0.3 mT that is regarded as biologically safe, is protected against magnetic perturbations and is of low cost.

*See also:* **EPR, Methods; Magnetic Field Gradients in High Resolution NMR; MRI Instrumentation; MRI Theory; Nuclear Overhauser Effect; Radio Frequency Field Gradients in NMR, Theory.**

## Further reading

Eaton GR, Eaton SS and Ohno K (eds) (1991) EPR Imaging and In Vivo EPR. Boca Raton, FL: CRC Press.

Eaton GR and Eaton SS (1995) Introduction to EPR imaging using magnetic-field gradients. Concepts in Magnetic Resonance **7**: 49.

Fairhurst SA, Gillies DG and Sutcliffe LH (1990) Electron spin resonance imaging. Spectroscopy World **2**: 14.

Ikeya M and Miki T (1987) Japanese Journal of Applied Physics **26**: L929.

Kuppusamy P, Wang P and Zweier JL (1995) Magnetic Resonance in Medicine **34**: 99.

Lurie DJ (1995) Imaging using the electronic Overhauser effect. In: Grant DM and Harris RK (eds) The

*Encyclopedia of Nuclear Magnetic Resonance*. Chichester: Wiley.
Lurie DJ (1996) Commentary: electron spin resonance imaging studies of biological systems. *British Journal of Radiology* **69**: 983.
Smirnov AI, Yakimchenko OE, Golovina HA, Bekova SK and Lebedev YS (1991) *Journal of Magnetic Resonance* **91**: 386.
Sutcliffe LH (in the press) *Physics and Medicine and Biology*.

# EPR Spectroscopy, Theory

**Christopher C Rowlands** and **Damien M Murphy**, Cardiff University, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Theory |

This article will discuss the concept and outline the theory of electron paramagnetic resonance (EPR), sometimes known as electron spin resonance (ESR), spectroscopy. Paramagnetism arises as a consequence of unpaired electrons present within an atom or molecule. It can be said that EPR is the most direct and sensitive technique to investigate paramagnetic materials. It will not be possible within this article to fully expound on all the theoretical aspects of EPR and so the reader is encouraged to refer to the excellent texts listed in the Further reading section.

Consequently this work is aimed at the professional scientists who would like to expand their knowledge of spectroscopic techniques and also at undergraduates with the intention of allowing them to fully appreciate the scope, versatility and power of the technique.

## Basic principles of the EPR experiment

The electron is a negatively charged particle that moves in an orbital around the nucleus and so consequently has orbital angular momentum. It also spins about its own axis and therefore has a spin angular momentum given by

$$\text{spin angular momentum} = h/2\pi[S(S+1)]^{1/2} \quad [1]$$

where $S$ is called the spin quantum number, with a value of $\frac{1}{2}$, and $h$ = Planck's constant ($6.626 \times 10^{-34}$ Js). If we restrict ourselves to one specified direction, the $z$ direction, then the component of the spin angular momentum can only assume two values and consequently $S_z = M_S\, h/2\pi$. The term $M_S$ can have $(2S + 1)$ different values of $+S$, $(S - 1)$, $(S - 2)\ldots(-S)$. If the possible values of $M_S$ differ by one and range from $-S$ to $+S$ then the only two possible values for $M_S$ are $\frac{1}{2}$ and $-\frac{1}{2}$.

Associated with the electron spin angular momentum is a magnetic moment $\mu$ and it is related to the spin as follows

$$\mu = g_e h/4\pi m_e\, [S(S+1)]^{1/2} \quad [2]$$

where $m_e$ = mass of electron, e = electron charge, $S$ = spin quantum number and $g_e$ is the spectroscopic factor known as the $g$ factor (which has a value of 2.0023 for a free electron). The above equation can be rewritten as

$$\mu = -g_e\mu_B\, S/\hbar \quad [3]$$

where $\mu_B = eh/2m_e$ and $\hbar = h/2\pi$. The $z$ axis component $\mu_z$ will then be

$$\mu_z = -g_e\mu_B M_S \quad [4]$$

$\mu_B$ is the Bohr magneton and has the value $9.2740 \times 10^{-24}$ J T$^{-1}$. The negative sign arises because of the negative charge on the electron which results in the magnetic dipole being in the opposite direction to the angular momentum vector. The electron is restricted to certain fixed angular positions that correspond to the $z$ component of the spin

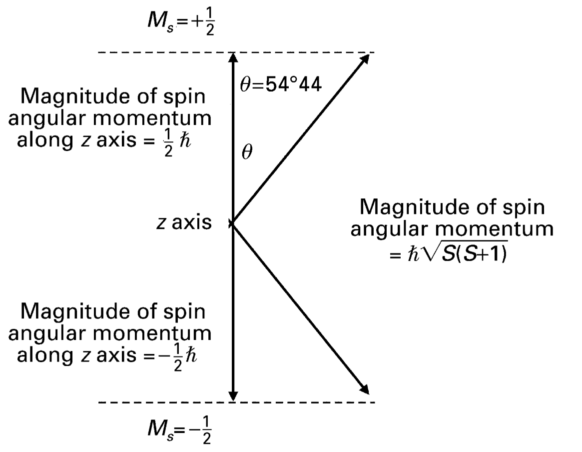



**Figure 1** Allowed values of the magnitude of the spin angular momentum along the *z* axis for $S = \frac{1}{2}$.

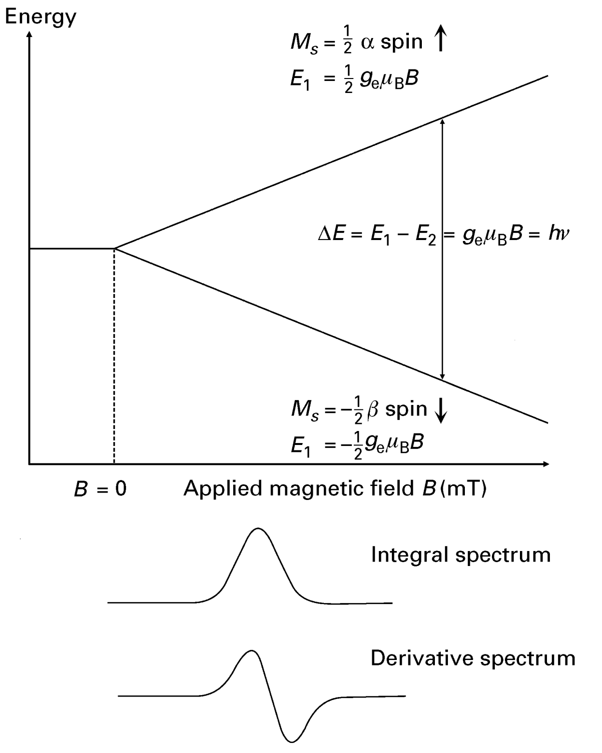



**Figure 2** The Zeeman energy levels of an electron ($S = \frac{1}{2}$) in an applied magnetic field for a fixed microwave frequency.

angular moment, $S_z$ is a half integral number ($M_S$) of $h/2\pi$, units. Since $S = \frac{1}{2}$ then $[S(S+1)]^{1/2}$ cannot be $\frac{1}{2}$ and consequently the spin angular momentum vector $S$ and the magnetic moment vector $\mu$ can never be aligned exactly in the field direction and they maintain a constant angle $\theta$ with the $z$ axis and hence the applied field (**Figure 1**).

Hence, for a single unpaired electron, where $S = \frac{1}{2}$, the magnitude of the spin angular momentum along the $z$ axis will be given by $+\hbar/2$ and $-\hbar/2$. For a free electron the spin angular momentum can have two possible orientations and these give rise to two magnetic moments or spin states of opposite polarity. In the absence of an external magnetic field the two spin states are degenerate. However, if an external magnetic field is applied then the degeneracy is lifted, resulting in two states of different energy. The energy of the interaction between the electron magnetic moment and the external magnetic field is given by

$$E = -\mu_z B \qquad [5]$$

where $B$ = the strength of external magnetic field and $\mu_z$ = magnetic dipole along the $z$ axis. In quantum mechanics the $\mu$ vector is replaced by the corresponding operator leading to the following Hamiltonian

$$H = -g_e \mu_B \boldsymbol{B}\boldsymbol{S}_z \qquad [6]$$

so that the energy is given by $E = -g_e\, \boldsymbol{\mu} B M_S B$. For a single unpaired electron $M_S = \pm\frac{1}{2}$ and this gives rise to two energy levels known as Zeeman levels. The difference between these two energy levels is known as the Zeeman splitting (see **Figure 2**):

$$E_1 = \tfrac{1}{2} g_e \mu_B B \qquad M_S = +\tfrac{1}{2} \qquad \alpha\,\text{spin}\uparrow \qquad [7]$$

$$E_2 = -\tfrac{1}{2} g_e \mu_B B \qquad M_S = -\tfrac{1}{2} \qquad \beta\,\text{spin}\downarrow \qquad [8]$$

Since $\Delta E \propto B$, the difference between the two energy levels is directly proportional to the external applied magnetic field. Transitions between the two Zeeman levels can be induced by the absorption of a photon of energy, $h\nu$, equal to the energy difference between the two levels:

$$\Delta E = g_e \mu_B B = h\nu \qquad [9]$$

where $h$ = Planck's constant and $\nu$ = frequency of electromagnetic radiation. The existence of two Zeeman levels, and the possibility of inducing transitions from the lower energy level to the higher energy level, is the very basis of EPR spectroscopy. We can see, therefore, that the position of absorption varies directly with the applied magnetic field. EPR spectrometers may operate at different fields and frequencies (two common frequencies being ~9.5

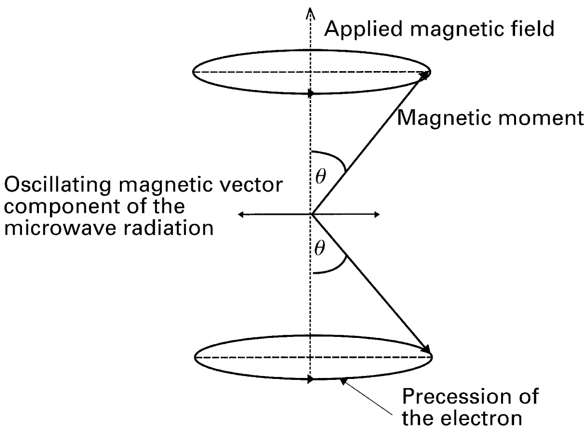



**Figure 3** Two allowed electron spin orientations and the corresponding Larmor precession for $S = \frac{1}{2}$.

and ~35 GHz, known as X and Q band respectively). It is therefore far more convenient to refer to the absorption in terms of its $g$ value:

$$g = \Delta E/\mu_{\rm B}\, B = h\nu/\mu_{\rm B}\, B \qquad [10]$$

## Larmor precession, spin populations and relaxation effects

It was noted above that the two spin states were at a constant angle to the $z$ axis. However, in the presence of a magnetic field, the magnetic moment of the electron cannot align with the applied field and a turning couple is set up which results in a precession of the spin states around the $z$ axis (**Figure 3**). This gives a physical mechanism by which spins can interact with electromagnetic radiation. When the magnetic component of the microwave radiation, which is orthogonal to the applied field, is at a frequency equal to that of the *Larmor precession* then the two are said to be in resonance, and energy can be transferred to the electron, inducing $\Delta M_S = \pm\frac{1}{2}$ transitions.

The incident radiation induces transitions not only from the lower to the higher energy states but can also induce emission with equal probability. Consequently the extent of absorption will be proportional to the population difference between the two states. At thermal equilibrium the relative *spin populations* of the two Zeeman levels is given by a Boltzmann equation:

$$N_{\rm U}/N_{\rm L} = e^{-\Delta E/kT} \qquad [11]$$

where $k$ = Boltzmann constant, $T$ = absolute temperature (K), $N_{\rm L}$ = number of electrons in the lower level, $N_{\rm U}$ = number of electrons in the higher level and $\Delta E$ = difference in energy between the two energy levels. At room temperature (300 K) and in a magnetic field of 300 mT the populations of the two Zeeman levels are almost equal, but a slight excess exists in the lower level and it is this that gives rise to a net absorption. However, this would very quickly lead to the disappearance of the EPR signal as the absorption of energy would equalize these two states. For the experiment to proceed there has to be a process by which energy is lost from the system. Such processes are known as *relaxation processes*.

To maintain a population excess, electrons in the upper level must be able to return to their low energy state. Therefore they must be able to transfer their excess spin energy either to other species or to the surrounding lattice as thermal energy. The time taken for the spin system to lose $1/e$ of its excess energy is called the relaxation time. Two such relaxation processes are:

(a) spin–lattice relaxation ($T_{1e}$): this process is due to lattice motions such as molecular tumblings in solids, liquids and gases which have a frequency comparable to that of the Larmor precession and this provides a pathway which allows the electron's excess energy to be transferred to the surroundings.
(b) spin–spin relaxation ($T_{2e}$): the excess spin energy is transferred between paramagnetic centres either through dipolar or exchange coupling, from one molecule to another. This mode of relaxation is important when the concentration of paramagnetic species is high (spins are close together). If the relaxation time is too fast then the electrons will only remain in the upper state for a very short period of time and give rise to a broadening of the spectral line width as a consequence of Heinsenberg's uncertainty principle.

In general $T_{1e} > T_{2e}$ and the line width depends only on the spin–spin interactions ($T_{2e}$). However, if, in certain circumstances, both spin–spin and spin–lattice relaxations contribute to the EPR line width ($\Delta H$), then the resonance line width can be simply written as

$$\Delta H \propto 1/T_{1e} + 1/T_{2e} \qquad [12]$$

When $T_{1e}$ becomes very short, below ~$10^{-7}$ s, its effects on the lifetime of the species in a given energy level makes an important contribution to the linewidth. In some cases the EPR lines are broadened beyond detection. The term $T_{1e}$ may also be inversely proportional to the absolute temperature ($T_{1e} \propto T^{-n}$) with $n$ depending on the precise relaxation

mechanism. In such cases, cooling the sample increases $T_{1e}$ and usually leads to detectable lines. For this reason, EPR experiments are frequently recorded at liquid nitrogen (77 K) or helium (4 K) temperatures.

## Basic interaction of the electron with its environment

An expression for the spin state energies of an electron whose only interaction was with an applied magnetic field (a free electron) was derived above (Eqn [6]). This represents an idealized model as in the real situation the electron suffers a variety of electrostatic and magnetic interactions which complicate the resonances. Consequently we must be able to account for these deviations. Magnetic resonance spectroscopists seek to characterize and interpret EPR spectra and these deviations quantitatively using a device known as a *spin Hamiltonian*. The EPR spectrum is essentially interpreted as the allowed transitions between the eigenvalues of this spin Hamiltonian. The Hamiltonian contains terms which reflect the interactions of the spins of electrons and nuclei with the applied magnetic field and with each other, e.g.

$$\boldsymbol{H} = \mu_{\mathrm{B}}\boldsymbol{BgS} + \Sigma_j \boldsymbol{IAS} + \Sigma_j g_{\mathrm{n}}\mu_{\mathrm{N}}\boldsymbol{BI} \quad [13]$$

Analysis of a spectrum amounts to identifying which interactions are involved. The origin of the various terms in this spin Hamiltonian will be examined below. For example, changes in the *g* factor are shown for a variety of paramagnetic species in **Table 1**.

**Table 1** Isotropic *g* value variations for a series $S = \frac{1}{2}$ paramagnetic species. For organic radicals the *g* value deviations (from $g_e$) are usually small, less than 1%, but for systems that contain heavier atoms, such as transition metal ions, the variations can be much larger

| *Species* | *g value* |
|---|---|
| $e^-$ | 2.0023 |
| $CH_3^{\bullet}$ | 2.0026 |
| Anthracene radical cation | 2.0028 |
| Anthracene radical anion | 2.0029 |
| 1,4-Benzoquinone radical anion | 2.0047 |
| $SO_2^-$ | 2.0056 |
| $HO_2^{\bullet}$ | 2.014 |
| $\{(CH_3)_3C\}_2NO$ | 2.0063 |
| Copper(acetonyl acetate) | 2.13 |
| VO(acetonyl acetate) | 1.968 |
| Cyclopentadienyl $TiCl_2AlCl_2$ | 1.975 |

### The *g* tensor: significance and origin

The main reason for the deviation in *g* values comes from a spin–orbit coupling that results in an orbital contribution to the magnetic moment. This arises from the effect of the orbital angular momentum *L* which is non-zero in the case of orbitals exhibiting p, d or f character. In this case the spin is no longer exactly quantized along the direction of the external field and the *g* value cannot be expressed by a scalar quantity but becomes a tensor. The orbital angular momentum *L* is associated with a magnetic momentum given by

$$\mu_{\mathrm{L}} = \mu_{\mathrm{B}}\boldsymbol{L} \quad [14]$$

For a system with a doublet ($S = \frac{1}{2}$) non-degenerate electronic ground state, the interaction with the external magnetic field can be expressed in terms of a perturbation of the general Hamiltonian by the following three terms

$$H = g_{\mathrm{e}}\mu_{\mathrm{B}}\boldsymbol{BS} + \mu_{\mathrm{B}}\boldsymbol{BL} + \lambda\boldsymbol{LS} \quad [15]$$

The first and second terms correspond, respectively, to the electron Zeeman and orbital Zeeman energies. The third one represents the energy of the spin–orbit coupling where $\lambda$ is the spin–orbit coupling constant which mixes the ground state wavefunctions with the excited states. The extent of the interaction between *L* and *S* mainly depends on the nature of the system considered and in many instances this interaction is stronger than that between the magnetic field and the orbital angular momentum. Depending on the strength of the molecular electric field, two limiting cases can be distinguished:

(a) Strong fields: *L* must align itself along the field so that only *S* can orient itself with respect to the external magnetic field and contribute to the paramagnetism. In this case $g = g_e$. Many organic radicals or systems in which the unpaired electron is in a molecular orbital delocalized over a large molecule experience this situation.
(b) Weak fields: *L* is no longer under the constraint of this weak field and the spin–orbit coupling $L + S$ can take place, giving a resultant total angular momentum:

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S} \quad [16]$$

associated with the magnetic moment:

$$\mu_{\mathrm{J}} = -g_{\mathrm{J}}\mu_{\mathrm{B}}\boldsymbol{J} \quad [17]$$

$g_{\mathrm{J}}$ is called the Landé $g$ factor. This situation occurs in the rare earth elements

When intermediate fields are present in the paramagnetic centre, $L$ is only partially blocked by the molecular field (transition metal ions, inorganic radicals). The system must be treated in terms of the perturbation Hamiltonian in Equation [6] which can be written as

$$H = \mu_{\mathrm{B}}\boldsymbol{BgS} \quad [18]$$

The $g_{\mathrm{e}}$ scalar value reported in Equation [6] is now replaced by $\boldsymbol{g}$, a second rank tensor which represents the anisotropy of the interaction between the unpaired electron and the external magnetic field and also outlines the fact that the orbital contribution to the electronic magnetic momentum may be different along different molecular axes.

The $g$ tensor may be depicted as an ellipsoid whose characteristic (principal) values ($g_{xx}$, $g_{yy}$, $g_{zz}$) depend upon the orientation of the symmetry axes of the paramagnetic entity with respect to the applied magnetic field (**Figure 4**). The most general consequence of the anisotropy of $g$, from an experimental point of view, is therefore that the resonance field of a paramagnetic species, for a given frequency, depends on the orientation of the paramagnetic centre in the field itself. The $g$ value for a given orientation depends on $\phi$ and $\theta$ values according to the relation $g^2 = g_{xx}{}^2\cos^2\phi\sin^2\theta + g_{yy}{}^2\sin^2\phi\cos^2\theta + g_{zz}{}^2\cos^2\theta$. Accordingly, the Zeeman resonance will occur at field values given by

$$B_{\mathrm{res}} = h\nu/\mu_{\mathrm{B}}(g_{xx}{}^2\cos^2\phi\ \sin^2\theta + g_{yy}{}^2\sin^2\phi\ \cos^2\theta + g_{zz}{}^2\cos^2\theta)^{-1/2} \quad [19]$$

In the most general case, the resonance observed for a paramagnetic centre in a single crystal is obtained at distinct field values of $B_x$, $B_y$ and $B_z$ when the magnetic field is parallel to the $x$, $y$ or $z$ crystal axis respectively. The $g$ values corresponding to these three orientations ($g_{xx}$, $g_{yy}$ and $g_{zz}$) are the principal (diagonal) elements of the $g$ tensor. When the principal axes of the crystal are not known, the $g$ tensors can be labelled $g_1$, $g_2$ and $g_3$. Absolute determination of the $g$ values may in principle be carried out by independent and simultaneous measurement of $B$ and $\nu$ using a gaussmeter and a frequency meter, respectively, following the equation

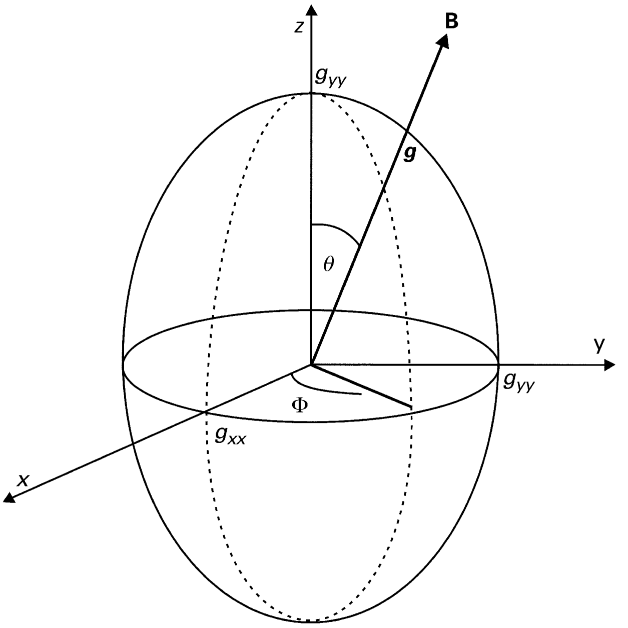



**Figure 4** Orientation of the magnetic field $\boldsymbol{B}$ with respect to the $\boldsymbol{g}$ tensor ellipsoid in the crystallographic frame $x$, $y$, $z$. The characteristic angles $\theta$ and $\phi$ define the orientation of $\boldsymbol{B}$ which leads to the $\boldsymbol{g}$ ($\theta$,$\phi$).

$$g = h\nu/\mu_{\mathrm{B}}\ \boldsymbol{B} \quad [20]$$

## The *A* tensor: significance and origin

The source of this splitting is the magnetic interaction between the electron spin and neighbouring nuclear spins, which gives rise to hyperfine structure in the spectra. It arises because a nearby magnetic nucleus gives rise to a local field, $B_{\mathrm{local}}$, which must be compounded with the applied magnetic field $B$ to satisfy the EPR condition $h\nu = g_{\mathrm{e}}\mu_{\mathrm{B}}B$. We must now rewrite this equation as

$$h\nu = g_{\mathrm{e}}\ \mu_{\mathrm{B}}(\boldsymbol{B} + \boldsymbol{B}_{\mathrm{local}}) \quad [21]$$

and clearly the value of $B$ required to achieve resonance will depend on $B_{local}$ (i.e. the magnetic moments of the nuclei). Several nuclei possess spin and corresponding magnetic moments. The nuclear spin quantum number ($I$) of a given nucleus can assume integral or half-integral values in the range 0–6. The magnetic moment $\mu_n$ associated to a nucleus is collinear with the spin vector $\boldsymbol{I}$ according to the relation

$$\mu_n = g_n \mu_N \boldsymbol{I} \quad [22]$$

which is similar to Equation [4]; $g_n$ is the nuclear $g$ factor and $\mu_N$ the nuclear magneton which is smaller than the Bohr magneton by a factor of 1838, i.e. the ratio of the mass of a proton to that of an electron.

When the paramagnetic centre contains one or more nuclei with non-zero nuclear spin ($I \neq 0$), the interaction between the unpaired electron and the nucleus with $I \neq 0$ produces further splittings of the Zeeman energies and consequently there are new transitions, which are responsible for the so-called hyperfine structure of the EPR spectrum.

Let us consider the proton as a case in point; it is a spinning positive charge and has its own associated magnetic field and consequently can interact with both the external magnetic field (cf. NMR) and that of the electron, i.e. it has a nuclear spin $I = \frac{1}{2}$ and $M_I = \pm\frac{1}{2}$. This leads to a situation whereby the energy of the system can either be lowered or raised depending on whether the two spins are parallel or antiparallel. The energy of an electron with spin quantum number $M_S$ and a nucleus with spin quantum number $M_I$ is given by

$$E(M_I, M_S) = g_e \mu_B B M_S - g_N \mu_N B M_I + h A_0 M_S M_I \quad [23]$$

where $A_0$ is the isotropic coupling constant. The first term in the equation gives the contribution due to the interaction of an electron with an applied field, giving rise to two electron Zeeman levels. The second term is the contribution due to the interaction of the nucleus with the applied magnetic field, the nuclear Zeeman levels (**Figure 5**). The final term is the energy of interaction between the unpaired electron and the magnetic nucleus. If we now substitute the values for $M_S$ and $M_I$ then the interaction, between an unpaired electron ($S = \frac{1}{2}$) and a single proton ($I = \frac{1}{2}$), gives rise to four energy levels, $E_1$ to $E_4$. These are calculated to be

| | $M_S$ | $M_I$ | |
|---|---|---|---|
| $E_1 = \frac{1}{2}g_e\mu_B B - \frac{1}{2}g_N\mu_B B + \frac{1}{4}hA_0$ | $+\frac{1}{2}$ | $+\frac{1}{2}$ | [24] |
| $E_2 = \frac{1}{2}g_e\mu_B B + \frac{1}{2}g_N\mu_N B - \frac{1}{4}hA_0$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | [25] |
| $E_3 = -\frac{1}{2}g_e\mu_B B + \frac{1}{2}g_N\mu_N B + \frac{1}{4}hA_0$ | $-\frac{1}{2}$ | $-\frac{1}{2}$ | [26] |
| $E_4 = -\frac{1}{2}g_e\mu_B B - \frac{1}{2}g_N\mu_N B - \frac{1}{4}hA_0$ | $-\frac{1}{2}$ | $+\frac{1}{2}$ | [27] |

The selection rules for the allowed transitions between the energy levels are

$$\Delta M_I = 0 \quad \text{and} \quad \Delta M_S = \pm 1 \quad [28]$$

Thus two resonance transitions can occur at

$$\Delta E_A = E_1 - E_4 = g_e \mu_B B + \tfrac{1}{2} h A_0 \quad [29]$$

$$\Delta E_B = E_2 - E_3 = g_e \mu_B B - \tfrac{1}{2} h A_0 \quad [30]$$

These two possible transitions give rise to two absorption peaks which, at constant applied microwave frequency, occur at magnetic fields of values

$$B_1 = \frac{h\nu}{g_e\mu_B} - \frac{hA_0}{2g_e\mu_B} = \frac{h\nu}{g_e\mu_B} - \frac{a}{2} \quad [31]$$

$$B_2 = \frac{h\nu}{g_e\mu_B} + \frac{hA_0}{2g_e\mu_B} = \frac{h\nu}{g_e\mu_B} + \frac{a}{2} \quad [32]$$

where $a = hA_0/g_e\mu B$ and is the hyperfine space coupling constant (HFCC) given in mT. The spectrum obtained for such a system would be a doublet centred at the $g$ value and separated by the hyperfine coupling constant (hfc) $a$. If we extend the number of equivalent nuclei then the complexity of the spectrum increases in a manner described in **Table 2**. A further example is given in **Figure 6** for the radical anion of naphthalene.

It is now necessary to examine the mechanisms by which the hyperfine interactions arise. Two types of electron-spin–nuclear-spin interactions must be considered, of isotropic and anisotropic nature.

**The isotropic interaction** This concerns exclusively s-type orbitals or orbitals with partial s character (such as hybrid orbitals constructed from s-type

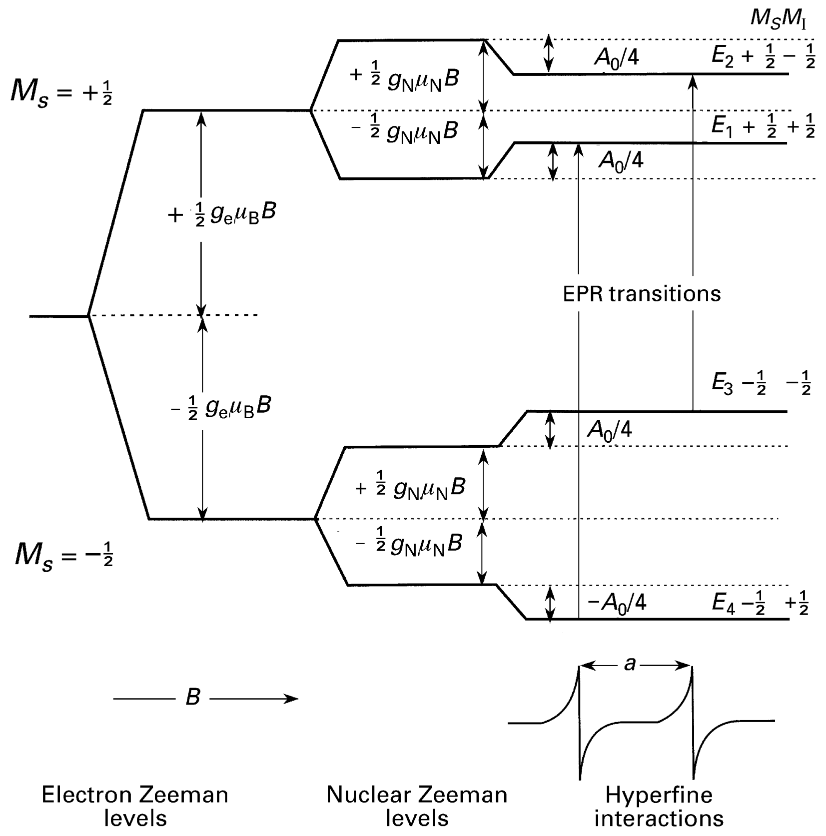



**Figure 5** Energy level manifold in a high magnetic field, resulting from the interaction of an unpaired electron ($S = \frac{1}{2}$) with a nucleus of $I = \frac{1}{2}$. Note that a > 0. $|a/2| < V_n$.

**Table 2** Line intensities for a series of nuclei with different values of $I$. For the case of $I = \frac{1}{2}$ the line intensities can be described by the expansion of $(x + 1)^2$ and hence follow Pascal's triangle

| *Nuclei* | *Intensity of lines* |
|---|---|
| $I = \frac{1}{2}$, e.g. $^1H$ | |
| Number of equivalent nuclei | |
| 1 | 1 1 |
| 2 | 1 2 1 |
| 3 | 1 3 3 1 |
| 4 | 1 4 6 4 1 |
| $I = 1$, e.g. $^{14}N$ | |
| Number of equivalent nuclei | |
| 1 | 1 1 1 |
| 2 | 1 2 3 2 1 |
| 3 | 1 3 6 7 3 1 |
| 4 | 1 4 10 16 19 16 10 4 1 |
| $I = \frac{3}{2}$, e.g. $^{63,65}Cu$ | |
| Number of equivalent nuclei | |
| 1 | 1 1 1 1 |
| 2 | 1 2 3 4 3 2 1 |

orbitals) because only these orbitals have a finite probability density at the nucleus. This mechanism is a quantum interaction related to the finite probability of the unpaired electron at the nucleus and is termed the Fermi contact interaction. The corresponding isotropic coupling constant $a_{iso}$ is given by

$$a_{iso} = (8\pi/3\,g_e g_n \mu_B \mu_N)\ |\ \psi(0)\ |^2 \qquad [33]$$

where $g_n$ and $\mu_N$ are the nuclear analogues of $g_e$ and $\mu_B$, respectively, and $|\psi(0)|^2$ is the square of the absolute value of the wavefunction of the unpaired electron evaluated at the nucleus.

Since s orbitals have a high electron density at the nucleus, the hyperfine coupling constant will be large and since s orbitals are also symmetrical it will be independent of direction. The spherical symmetry of s orbitals accounts for the isotropic nature of the contact interaction.

Thus in the case of the hydrogen atom ($M_I = \frac{1}{2}$) with 100% spin density in the 1s orbital, the EPR spectrum is composed of two resonance lines

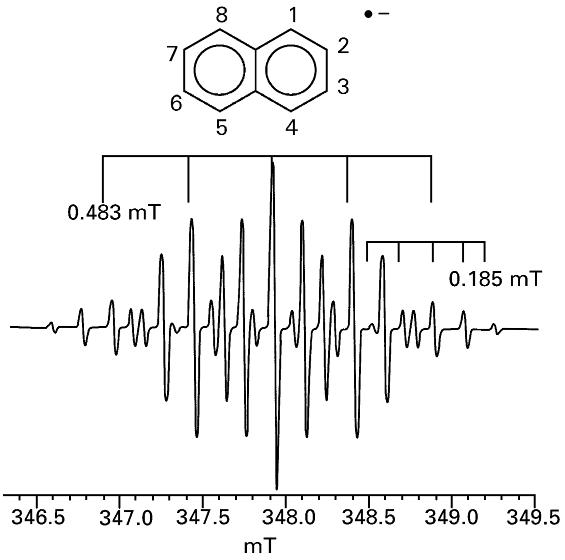



**Figure 6** EPR spectrum of the radical anion of naphthalene in tetrahydrofuran as solvent. The radical was formed by dissolving naphthalene in sodium-dried tetrahydrofuran under vacuum and passing the resultant solution over a sodium metal film, again under vacuum conditions. The number of lines is given by $2nI + 1$ where $I$ equals the nuclear spin and $n$ the number of nuclei. HFCC for 1,4,5,8 = 0.483 mT and for 2,3,6,7 = 0.185 mT.

separated by approximately 50.8 mT. In the case of naphthalene, the smaller spin density associated with the $^1$H s orbital results in a smaller isotropic coupling.

**The anisotropic interaction** In non-spherically symmetric orbitals (p, d or f) there is no electron density at the nucleus and the electron is to be found some distance away. The interaction between it and the nucleus will be as two magnetic dipoles and consequently the interaction will be small and dependent on the direction of the orbital with respect to the applied magnetic field as well as to their separation. The coupling then arises from the classical dipolar interaction between magnetic momenta whose energy is given by

$$E = \mu_s \mu_n / r^3 - 3(\mu_s \boldsymbol{r})(\mu_n \boldsymbol{r})/r^5 \quad [34]$$

where $\boldsymbol{r}$ is the vector relating the electron and nuclear moments and $r$ is the distance between the two spins. The above equation gives the energy of interaction of two bar magnets if their size is small compared with the distance between them. Moving to quantum mechanics we construct the Hamiltonian by replacing the magnetic moments by the appropriate operators. The quantum mechanics analogue of Equation [34] is then obtained by replacing $\mu_s$ and $\mu_n$ by their expressions given in Equations [4] and [22]:

$$H_{\text{aniso}} = -g_e \mu_B g_n \mu_B (\boldsymbol{I}{\cdot}\boldsymbol{S}/r^3 - 3(\boldsymbol{I}{\cdot}\boldsymbol{r})(\boldsymbol{S}{\cdot}\boldsymbol{r})/r^5) \quad [35]$$

Equation [35] must be averaged over the entire probability of the spin distribution. $H_{\text{aniso}}$ is averaged out to zero when the electron cloud is spherical (s orbital) and comes to a finite value in the case of axially symmetric orbitals (p orbitals, for instance). In addition, in the case of very rapid tumbling of the paramagnetic species (as occurs in a low viscosity solution) the anisotropic term of the hyperfine interaction is averaged to zero and the isotropic term is the only one observed.

**Combination of isotropy and anisotropy** Since any orbital may be considered as a hybrid of suitable combinations of s, p or d orbitals, so also may a hyperfine coupling be divided into a contribution arising from p or d orbitals (anisotropic) and s orbitals (isotropic). In general both isotropic and anisotropic hyperfine couplings occur when one or more nuclei with $I \neq 0$ are present in the system. The whole interaction is therefore dependent on orientation and must be expressed by the second rank $A$ tensor as given in Equation [13]. The $A$ tensor may be split into an anisotropic and isotropic part as follows

$$\mathbf{A} = \begin{vmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{vmatrix} = a_{\text{iso}} + \begin{vmatrix} T_1 & 0 & 0 \\ 0 & T_2 & 0 \\ 0 & 0 & T_3 \end{vmatrix} \quad [36]$$

with $a_{\text{iso}} = (A_1 + A_2 + A_3)/3$. In a number of cases, the second term matrix of Equation [36] is a traceless tensor ($T_1 + T_2 + T_3 = 0$) and has the form ($-T$, $-T$, $+2T$). The anisotropic part of the $A$ tensor corresponds to the dipolar interaction as expressed by the Hamiltonian in Equation [13]. The s and p character of the orbital hosting the unpaired electrons are given by the following relations:

$$c_S{}^2 = a_{\text{iso}}/a_0 \quad \text{and} \quad c_P{}^2 = b/b_0 \quad [37]$$

where $a_0$ and $b_0$ are the theoretical hyperfine coupling constants (assuming pure s and p orbitals for the elements under consideration). The terms $a_{\text{iso}}$ and $b$ are the experimental isotropic and anisotropic coupling constants respectively.

### The *D* tensor: significance and origin

The spin Hamiltonian described in Equation [13] applies to the case where a single electron ($S = \frac{1}{2}$) interacts with the applied magnetic field and with surrounding nuclei. However, if two or more electrons are present in the system ($S > \frac{1}{2}$), a new term must be added to the spin Hamiltonian (Eqn [13]) to account for the interaction between the electrons. At small distances, two unpaired electrons will experience a strong dipole–dipole interaction, analogous to the interaction between electron and magnetic dipoles, which gives rise to anisotropic hyperfine interactions. The electron–electron interaction is described by the spin–spin Hamiltonian given in Equation [38]:

$$H_{ss} = \boldsymbol{SDS} \qquad [38]$$

where $D$ is a second rank tensor (the zero-field parameter) with a trace of zero. As with the $g$ and $A$ tensors, the $D$ tensor can also be diagonalized so that $D_{xx} + D_{yy} + D_{zz} = 0$. Equation [38] can be added to Equation [18] to obtain the correct spin Hamiltonian for an $S > \frac{1}{2}$ system (in the absence of any interacting nucleus):

$$H = \mu_B \boldsymbol{BgS} + \boldsymbol{SDS} \qquad [39]$$

Since the trace of $D$ is zero, calculation of the energy state for a system with $S = 1$ requires only two independent parameters which are designated $D$ and $E$. The spin coupling is direct in the case of organic molecules in the triplet state and biradicals, but occurs through the orbital angular momentum in the case of transition metal ions. In the latter case, the $D$ and $E$ terms depend on the symmetry of the crystal field acting on the ions:

$$H = D\left(S_z^2 - \frac{S^2}{3}\right) + E(S_x^2 - S_y^2) \qquad [40]$$

For axially symmetric molecules, the calculated shape of the $\Delta M_S = 1$ lines are given in **Figure 7A**. The separation of the outer lines is $2D'$ (where $D' = D/g\mu_B$) while that of the inner lines is $D$ ($E$ is zero in this case). The theoretical line shape for a randomly oriented triplet system with $E \neq 0$ is shown in **Figure 7B**. The separation of the outermost lines is again $2D'$ whereas that of the intermediate and inner pairs is $D' + 3E'/2$ and $D' - 3E'/2$ respectively. As the zero-field interactions become comparable to, and larger than, the microwave energy, the line shape exhibits severe distortions from the simulated case in **Figure 7**. Well known examples of $S = 1$ include the excited triplet state of naphthalene with $D/hc = 0.1003\ \text{cm}^{-1}$, $E/hc = -0.0137\ \text{cm}^{-1}$ and $g = 2.0030$.

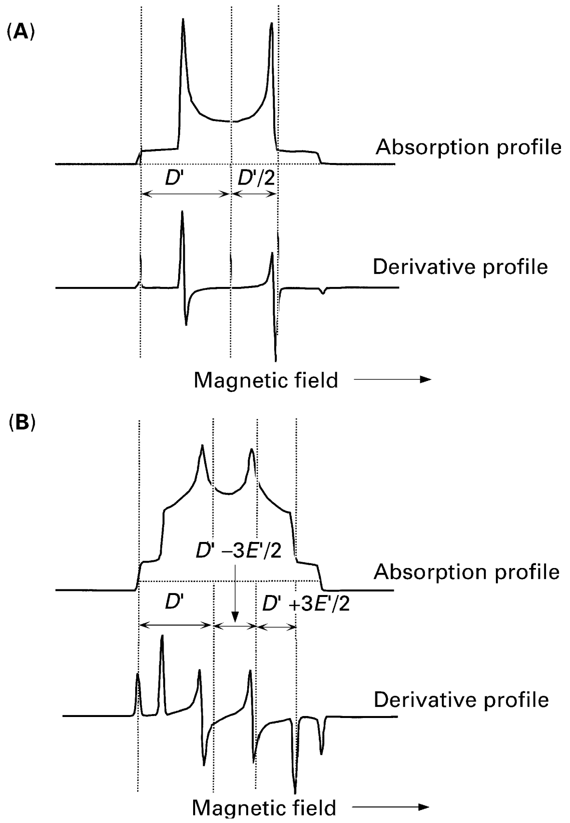



**Figure 7** Theoretical absorption and first derivative spectra of the $\Delta M_S = 1$ region of a randomly oriented triplet. (A) $S = 1$ for a given value of $D$ ($E = 0$) and isotropic $g$. (B) $S = 1$ where $D$ is 6.5 times larger than $E$ and $g$ is isotropic.

## Line shapes and symmetry considerations in EPR spectra

An important consideration in EPR is the effect of symmetry on the line shapes. In many cases the samples investigated by EPR are polycrystalline materials, composed of numerous small crystallites that are randomly oriented in space. The resultant *powder* EPR spectrum is the envelope of spectra corresponding to all possible orientations of the paramagnetic species with respect to the magnetic field; provided the resolution is adequate, the magnitude of the $g$ and $A$ tensor components can be extracted from the

powder spectra whereas no information can be obtained on the orientation of the tensor principal axes.

The profile of the powder spectrum is determined by several parameters, including the symmetry of the *g* tensor, the actual values of its components, and the line shape and the line width of the resonance. Concerning the symmetry of the *g* tensor, three possible cases can be identified.

### Isotropy of *g*

In this case, the *g* tensor is characterized by $g_{xx} = g_{yy} = g_{zz} = g_{\text{iso}}$ and a single symmetrical line is observed. This simple situation is not very often encountered in powders except for some solid state defects and transition metal ions in highly symmetric environments such as $O_h$ or $T_d$.

### Axial symmetry of *g*

Paramagnetic species isolated in single crystals exhibit resonances at typical magnetic fields which depend on their orientation and are given by Equation [19]. In the particular case of axial symmetry of the system, e.g. ($C_{4v}$, $D_{4h}$), if *z* is the principal symmetry axis of the species and $\theta$ the angle between *z* and the magnetic field, the *x* and *y* directions are equivalent and the angle $\phi$ becomes meaningless (**Figure 4**). Equation [19] reduces to

$$B_{\text{res}} = h\nu/\mu_{\text{B}}(g_{\perp}{}^2 \sin^2\theta + g_{\|}{}^2 \cos^2\theta)^{-1/2} \qquad [41]$$

where $g_{\|} = g_{zz}$ and $g_{\perp} = g_{yy} = g_{xx}$ are the *g* values measured when the axis of the paramagnetic species is, respectively, parallel and perpendicular to the applied magnetic field.

The powder spectrum is the envelope of the individual lines corresponding to all possible orientations in the whole range of $\phi$. Assuming the microcrystals are randomly distributed, simple considerations show, however, that the absorption intensity, which is proportional to the number of microcrystals at resonance for a given $\theta$ value, is at a maximum when $\theta = \pi/2$ ($B_{\perp}$) and a minimum for $\theta = 0$ ($B_{\|}$); this allows the extraction of the $g_{\|}$ and $g_{\perp}$ values, which correspond to the turning points of the spectrum. **Figure 8** gives a schematic representation of the absorption curve and of its first derivative for a polycrystalline sample containing a paramagnetic centre in axial symmetry. The variation in the spectral profile for the same species as $M_I$ varies is also shown in **Figure 9**.

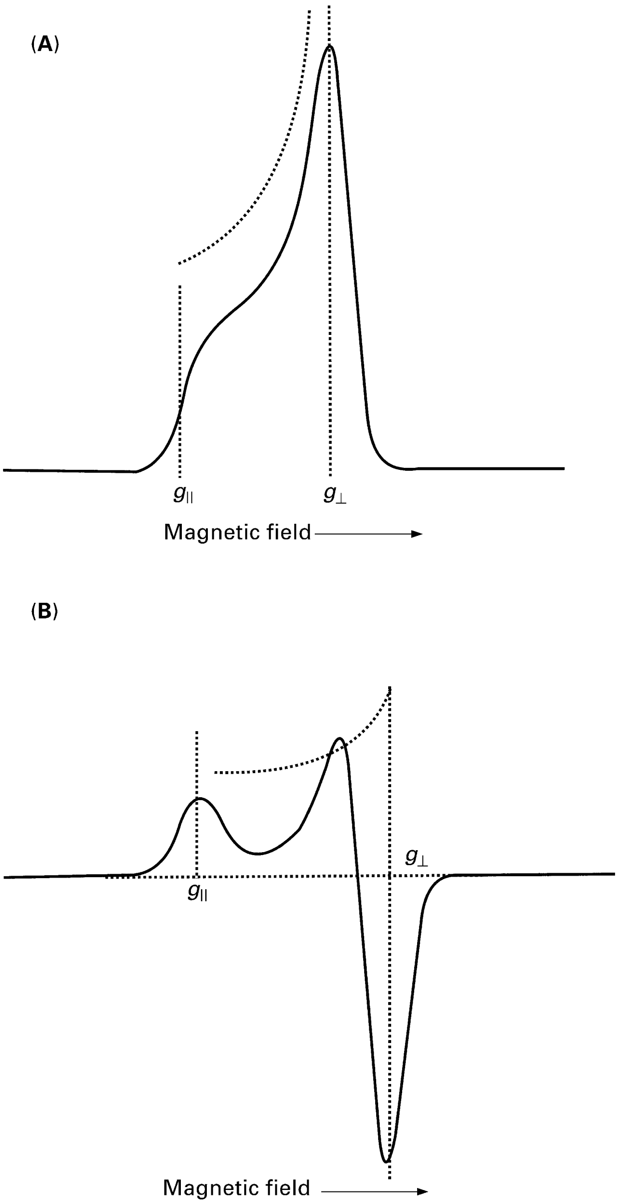



**Figure 8** Powder EPR spectrum of a paramagnetic species with $I = 0$ in axial symmetry. (A) absorption profile, (B) first derivative profile. The dotted lines have been calculated by assuming a zero line width whereas the solid lines correspond to a finite line width.

### Orthorhombic symmetry of *g*

Three distinct principal components are expected in the case of a system with orthorhombic symmetry, e.g. ($D_{2h}$, $C_{2v}$). For polycrystalline samples, the absorption curve and its first derivative exhibit three singular points, corresponding to $g_1$, $g_2$ and $g_3$ (**Figure 10**). For powder spectra the assignments of $g_1$, $g_2$ and $g_3$ to the components $g_{xx}$, $g_{yy}$ and $g_{zz}$ related to the molecular axes of the paramagnetic centre is not straightforward and must be based on

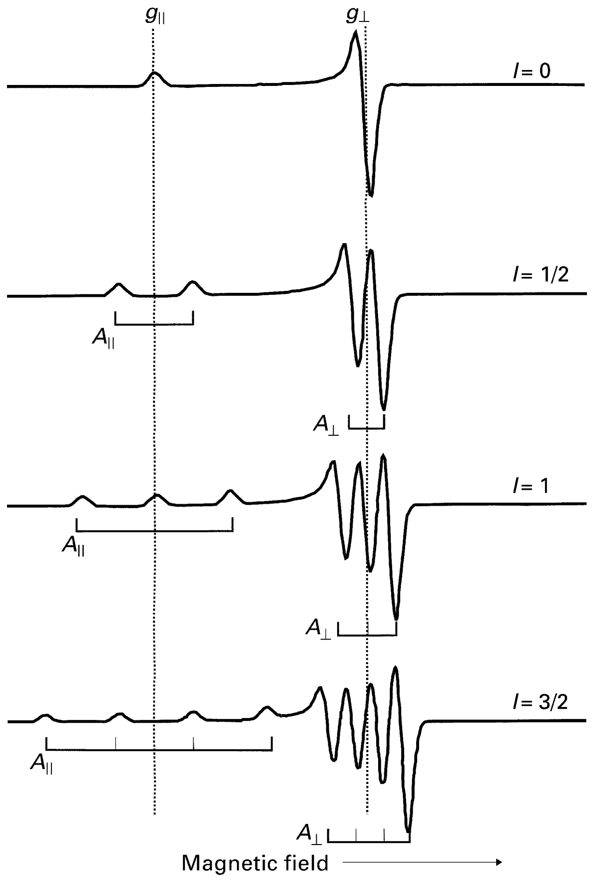



**Figure 9** Effects of nuclear spin on the calculated powder EPR spectra for axial symmetry. The value of $I$ varies from 0 to $\frac{3}{2}$. Note that the $g$ values remain constant (illustrated by dotted lines). $\boldsymbol{g}_{\parallel}$ is split by $\boldsymbol{A}_{\parallel}$ and $\boldsymbol{g}_{\perp}$ is split by $\boldsymbol{A}_{\perp}$. The $\boldsymbol{A}_{\parallel}$ and $\boldsymbol{A}_{\perp}$ values remain constant.

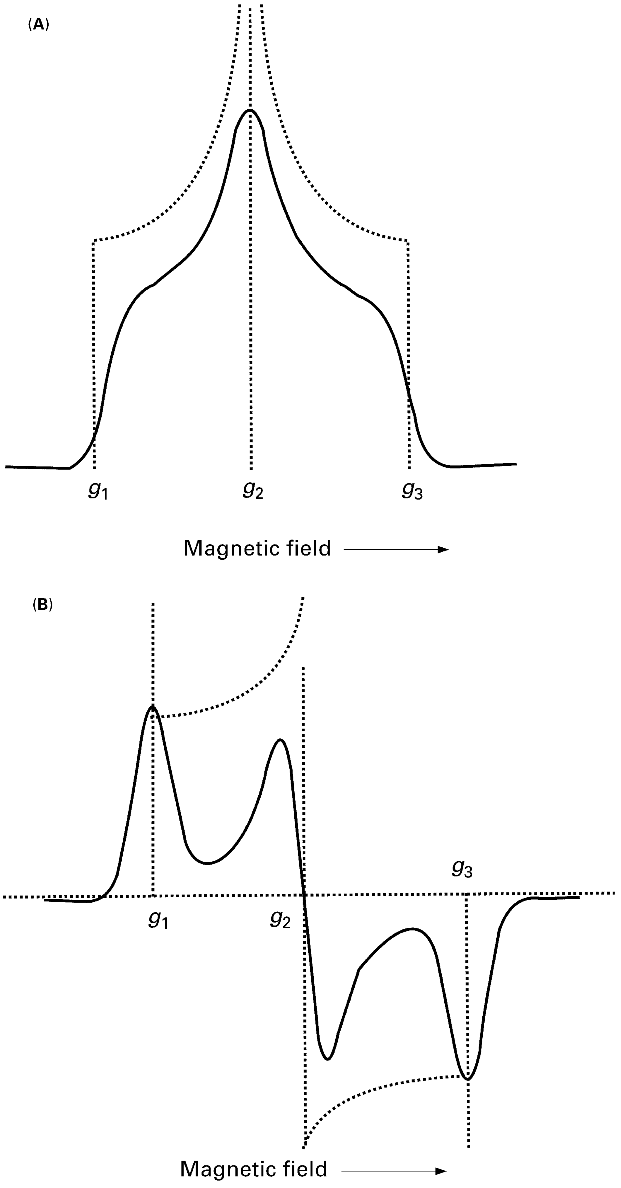



**Figure 10** Powder EPR spectrum of a paramagnetic species in orthorhombic symmetry. (A) Absorption profile, (B) first derivative profile. The dotted and solid lines have the same significance as described in **Figure 7**.

theoretical grounds or deduced from measurements of the same paramagnetic species in a single crystal. The situation becomes much more complicated in the presence of nuclei with $I \neq 0$. In several cases an initial interpretation of the experimental data can be achieved by using a simple analysis for the $g$ tensor and neglecting second-order effects since the $A$ tensor has the same type of angular dependence as the $g$ tensor. Provided the principal axes are the same, then each of the three possible lines of the $g$ tensor ($g_1$, $g_2$ and $g_3$) will be split into a number of lines that depend on the nuclear spin with the spacing corresponding to the appropriate component of the $A$ tensor ($A_1$, $A_2$ or $A_3$); $g_1$ is split by $A_1$, $g_2$ is split by $A_2$ and $g_3$ is split by $A_3$ (**Figure 11**).

## Summary

The basic principles of EPR spectroscopy have been outlined. Although far from being an exhaustive treatment of the subject, it provides the reader with essential knowledge of the technique and is a useful stepping stone for the articles on methods and instrumentation and applications.

## List of symbols

$a$ = hyperfine coupling constant; $A_0$ = theorectical isotropic hyperfine coupling constant; $A$ = hyperfine tensor; $g$ = $g$ factor; $g_e$ = $g$ factor of the free electron; $g_J$ = Landé $g$ factor; $g_N$ = nuclear $g$ factor; $g_{\parallel}$ and $g_{\perp}$ = refer to the parallel and perpendicular values

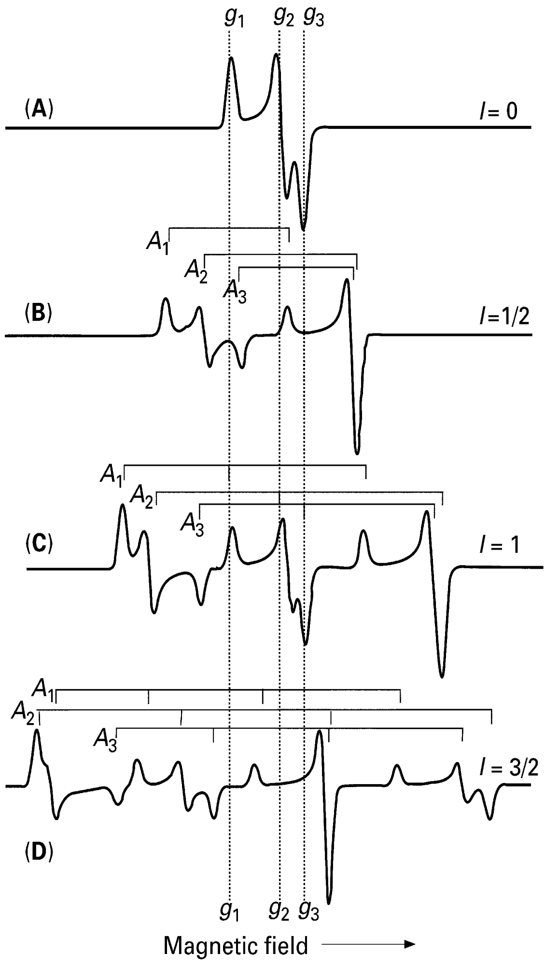



**Figure 11** Effects of nuclear spin on the calculated powder EPR spectrum of a species with orthorhombic symmetry. The spectrum is anisotropic both in $\boldsymbol{g}$ and $A$ ($\boldsymbol{g}_1$ is split by $A_1$, $\boldsymbol{g}_2$ is split by $A_2$ and $\boldsymbol{g}_3$ by $A_3$). The value of $I$ varies from 0 to $\frac{3}{2}$. It should be noted that when $I$ assumes an integer value (as in C), the features of the original $I$ = 0 spectrum (A) are maintained in the centre of the spectrum. When $I$ has half-integer values the hyperfine values are symmetrically disposed about the centre.

with respect to axis of symmetry; $g_{xx}$, $g_{yy}$, $g_{zz}$ = refer to values in the respective crystallographic directions; $h$ = Planck's constant; $\hbar = h/2\pi$; $I$ = nuclear spin quantum number; $J$ = total angular momentum; $k$ = Boltzman constant; $L$ = orbital angular momentum vector; $m_e$ = mass of the electron; $N_L$ and $N_U$ = electron populations in the lower and upper energy levels respectively; $r$ = distance between the spins; $\boldsymbol{r}$ = vector relating electron and nuclear moments; $S$ = electron spin quantum number; $S_z$ = vector of the electron spin quantum number in the $z$ direction; $T_{1e}$ and $T_{2e}$ = spin–lattice and spin–spin relaxation times, respectively; $\alpha$ and $\beta$ = refer to electron spin; $\mu$ = magnetic moment; $\mu_B$ = Bohr magneton; $\mu_N$ = nuclear magneton; $\nu$ = frequency.

*See also:* **Chemical Applications of EPR; EPR Imaging; EPR, Methods; NMR in Anisotropic Systems, Theory; NMR Principles; NMR Relaxation Rates.**

## Further reading

Abragam A and Bleaney B (1970) *Electron Paramagnetic Resonance of Transition Ions* Oxford: Oxford University Press.

Atherton NM (1993) *Principles of Electron Spin Resonance*, PTR Prentice Hall Physical Chemistry Series. Englewood Cliff, NJ: Prentice-Hall.

Bolton JR, Weil JA and Wertz JE (1994) *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications.* New York: Wiley-Interscience.

Che M and Giamello E (1987) In: Fieno JLG (ed) *Spectroscopic Characterization of Heterogeneous Catalysts. Part B: Electron Paramagnetic Resonance*, Vol 57. Amsterdam: Elsevier Science Publishers.

Mabbs FE and Collison D (1992) *Electron Paramagnetic Resonance of d-Transition Metal Compounds*, Vol 16 in Studies in Inorganic Chemistry. Amsterdam: Elsevier.

Pilbrow JR (1993) *Transition Ion Electron Paramagnetic Resonance.* Oxford: Clarendon Press.

Symons MRC (1978) Chemical and biological aspects of electron spin resonance spectroscopy. *Chemical Aspects of Physical Techniques.* New York: Van Nostrand Reinhold.

# EPR, Methods

**Richard Cammack**, King's College, London, UK

Copyright © 1999 Academic Press

MAGNETIC RESONANCE

Methods & Instrumentation

## Introduction

Electron paramagnetic resonance (EPR) is also known as electron spin resonance (ESR) or electron magnetic resonance (EMR) spectroscopy; no single name has become generally accepted. It is a means of investigating materials that are paramagnetic, i.e. having unpaired electrons. These include organic free radicals, compounds of transition metal ions and defects in solids. Since the development of the X-band continuous-wave spectrometer in the 1950s and 1960s, the basic design of the EPR spectrometer has remained relatively unchanged, while a wide variety of applications in physics, chemistry and biochemistry have been devised. Examples are the observation of radical intermediates in chemical reactions; the determination of electron density distributions on radicals; the observation of rapid motion of spin labels in solution; and the determination of coordination geometry of metal ion sites. Recent advances in solid-state microwave electronics have led to instrument extensions in many different ways, for specialist applications. Examples are ENDOR and TRIPLE and ESEEM, to measure hyperfine interactions, and low-frequency EPR imaging.

The principles of typical applications will be described, with reference to the relevant equipment and sample requirements.

## Characteristics of the EPR spectrum

EPR involves the interaction of electromagnetic radiation, usually in the microwave region, with the paramagnetic material in a magnetic field. Because paramagnetic centres are less common than nuclear spins, the method is more selective than the analogous technique of NMR. There are some differences in practice between NMR and EPR that result from the much greater magnetic moment of the electron compared with the nuclear spins. The bandwidth of the spectrum is much greater than the parts-per-million scale of NMR. It is impossible to adjust the frequency over such a wide range, owing to the geometry of microwave waveguides, so it is the usual practice to observe the EPR spectrum at fixed frequency by sweeping the magnetic field. The relaxation rates of electron spins are faster than for nuclear spins, so that in some cases (particularly transition metals such as iron) it is necessary to cool the sample to cryogenic temperatures to observe the spectrum.

The spectrum is characterized by a number of parameters which give information about the nature of the paramagnetic centres and their surroundings.

### Zeeman interaction

The $g$ factor defines the energy of the Zeeman interaction. This splits the energy levels of the paramagnet (**Figure 1**). At the resonance condition the energy $\Delta E$ required to reverse the direction of the electron spin in a magnetic field $B_0$ is equal to the energy of the microwave quantum $h\nu$, given by

$$\Delta E = h\nu = g\mu_B B_0$$

where $h$ is Planck's constant and $\mu_B$ the Bohr magneton. In an EPR spectrum the signals are obtained by increasing the $B_0$ field. The $g$ factor is inversely proportional to $B_0$.

For a free electron in a vacuum, $g$ is a constant that is very precisely known ($g_e$ = 2.002 319 3044…). In a paramagnetic molecule, $g$ varies from this, under the influence of spin–orbit interactions. These increase considerably with atomic number, and are particularly important for lanthanides. In the spin Hamiltonian formalism the effects of spin–orbit coupling are described by treating $g$ as a spectroscopic variable, the 'spectroscopic splitting factor', which is characteristic of the paramagnetic centre. Therefore the EPR spectrum may be considered as a spectrum of $g$ factors.

### Hyperfine interaction

The interaction between the magnetic moment of the electron and nuclear spins is a valuable feature of EPR spectra. It introduces splittings of the electron energy levels (**Figure 1B**). Each nuclear spin $I$ induces a splitting into $[2I + 1]$ levels. The magnitude of the splitting is defined by the hyperfine coupling constant, $A$, which is expressed in energy units (usually MHz). Where there are several nuclei, for example protons in a radical, each level is further split by each nuclear hyperfine interaction. The

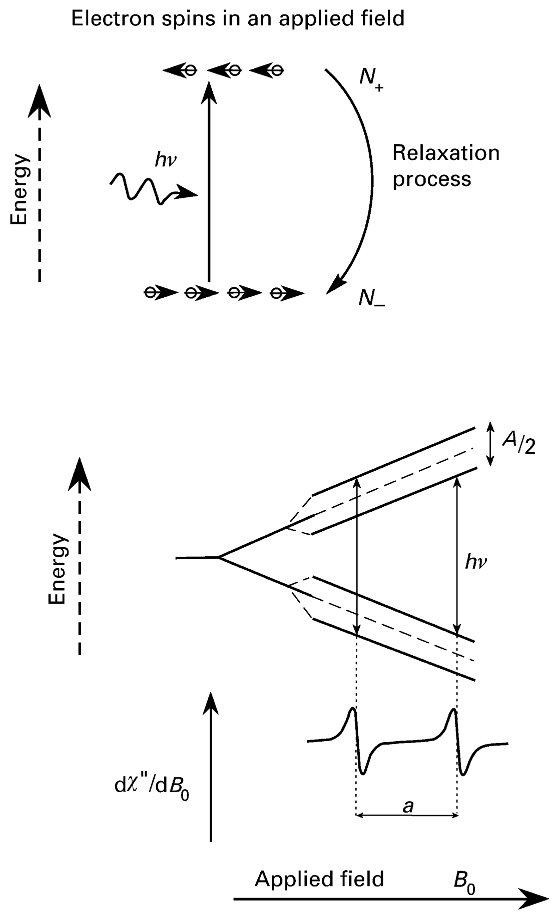



**Figure 1** (A) Principle of electron paramagnetic resonance. (B) Effect of hyperfine splitting with a nucleus, $I = \frac{1}{2}$.

hyperfine coupling can be observed as splitting of the EPR spectrum, if the coupling is greater than the line width. The magnitude of the splitting is given the symbol $a$, expressed in magnetic field units.

The value of $A$ depends on the nuclear moment, and the extent of interaction of the unpaired electron spin density with the nucleus; A has a sign as well as magnitude. If $A > 0$, the state in which electron and nuclear spins align antiparallel is of lower energy. Measurements of the magnitude and sign of hyperfine couplings provide some of the most detailed experimental evidence for electronic structures of molecules. They have been used to verify the results of molecular orbital calculations.

In EPR spectra of transition ions, the hyperfine interaction gives rise to characteristic splittings, for example for copper ($I = \frac{3}{2}$, 4 lines) and manganese ($I = \frac{5}{2}$, 6 lines) (**Figure 2B**). It can be used to identify paramagnetic centres and determine the electron density distribution over radicals. Isotopic substitution can be used (**Table 1**) to identify specific metals, for example $^{57}$Fe for natural iron.

**Table 1** Some nuclei with magnetic moments

| *Isotope* | *% Natural abundance* | *Nuclear spin, I* | $\mu_N$ |
|---|---|---|---|
| $^{1}$H | 99.985 | $\frac{1}{2}$ | + 2.793 |
| $^{2}$H | 0.015 | 1 | + 0.857 |
| $^{13}$C | 1.11 | $\frac{1}{2}$ | +0.702 |
| $^{14}$N | 99.63 | 1 | +0.404 |
| $^{15}$N | 0.37 | $\frac{1}{2}$ | −0.283 |
| $^{17}$O | 0.037 | $\frac{5}{2}$ | −1.89 |
| $^{19}$F | 100 | $\frac{1}{2}$ | +2.629 |
| $^{31}$P | 100 | $\frac{1}{2}$ | +1.132 |
| $^{33}$S | 0.76 | $\frac{3}{2}$ | +0.643 |
| $^{55}$Mn | 100 | $\frac{5}{2}$ | +3.444 |
| $^{57}$Fe | 2.19 | $\frac{1}{2}$ | +0.090 |
| $^{59}$Co | 100 | $\frac{7}{2}$ | +4.649 |
| $^{63}$Cu | 69.09 | $\frac{3}{2}$ | +2.226 |
| $^{65}$Cu | 30.91 | $\frac{3}{2}$ | +2.385 |
| $^{95}$Mo | 15.72 | $\frac{5}{2}$ | −0.913 |
| $^{97}$Mo | 9.46 | $\frac{5}{2}$ | −0.933 |

Interactions with nuclei of ligands, sometimes called superhyperfine splittings, allow the determination of the electron density distribution in transition metal complexes. Nuclei with spins greater than $S = \frac{1}{2}$ have a quadrupole moment and this gives rise to further splitting of the energy levels.

## Interactions between electrons (fine structure interactions)

The magnetic moment of the electron is 658 times that of the proton, so that interactions between electrons are of greater energy than those between nuclear spins, and are detectable over greater distances. This has a number of consequences. For example, when two or more electron spins are present in a paramagnetic centre such as a transition ion with spin $S > \frac{1}{2}$ or an organic triplet state ($S = 1$), coupling of the electron spins gives rise to fine structure splittings (known as zero-field splittings) which strongly influence the apparent g factor. If the spin is an integral value, the splitting between the energy levels may be larger than the X-band microwave quantum, and no EPR signal is detected. However, for an odd number of electron spins there are always energy levels that can only be split by the magnetic field. As a consequence paramagnets with half-integral spin, such as high-spin $Fe^{III}$ and $Co^{II}$, are

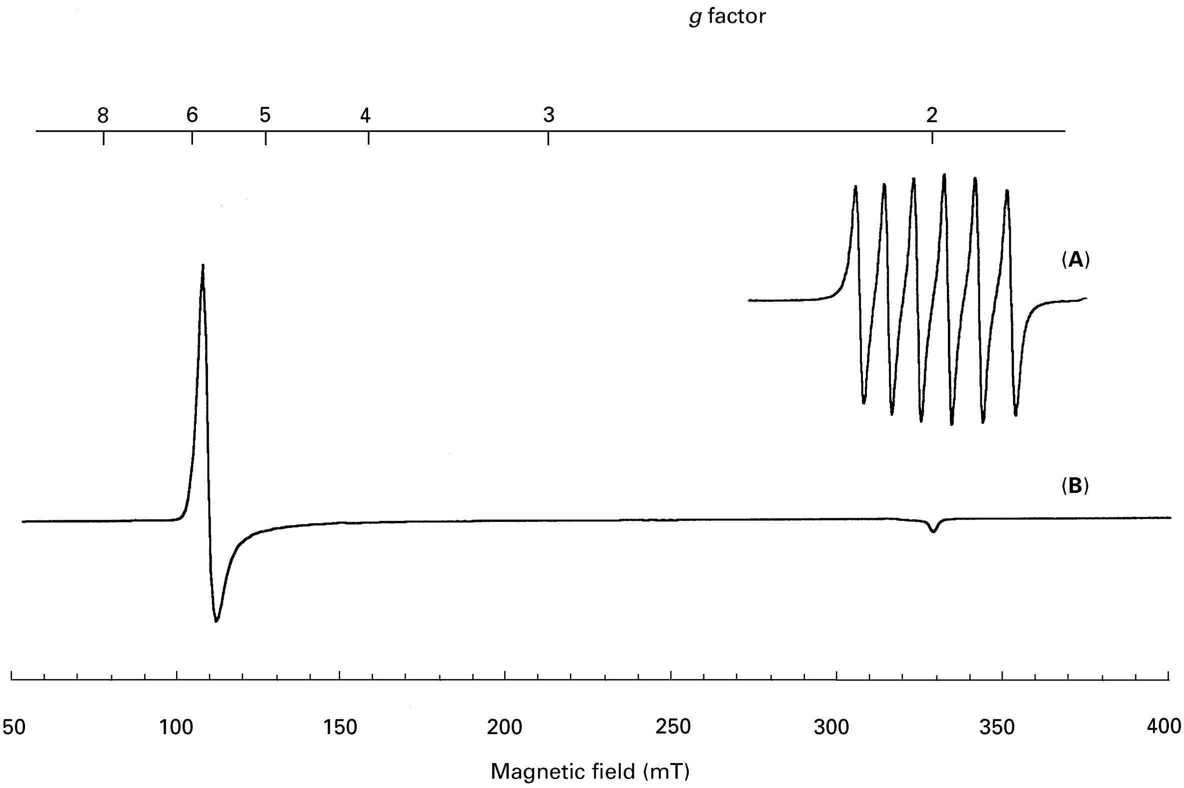



**Figure 2** EPR spectra of transition ions (A) manganese ions in aqueous solution; note the six-line hyperfine splitting owing to the $^{55}$Mn nucleus ($I = \frac{5}{2}$); (B) high-spin $Fe^{III}$ in metmyoglobin, frozen at 77 K; note the *g* factors at 6.0 and 2.0.

EPR-detectable. For the latter ions, with $S = \frac{5}{2}$, *g* factors may extend from 10 to 0. EPR spectra of transition ions are often measured in the solid state, so that the spectrum is broadened by anisotropy of the *g* factor.

Electron spin–spin interactions occur between paramagnetic centres in solution and in the solid state. This causes splitting or broadening of the EPR signal. For solid paramagnetic materials the broadening may be so great as to render the X-band EPR spectrum undetectable. For EPR spectra in the solid state, the material must usually be *magnetically dilute*, for example doped into a diamagnetic host material. Paramagnets in biological materials such as proteins are usually fixed within the protein matrix and are thus detectable. Spin–spin interactions give rise to characteristic changes in line shape, and enhancement of the electron-spin relaxation rate, from which it is possible to estimate distances between the paramagnetic centres.

### Anisotropy

The Zeeman, dipolar hyperfine and fine structure interactions are all anisotropic, meaning that the energy of interaction depends on the angle between the paramagnetic centre and the applied magnetic field. In solution spectra of small molecules, the anisotropy is averaged, and spectral lines are narrow. Note, however, that the isotropic hyperfine interaction does not average to zero and so there is still hyperfine splitting (e.g. **Figure 2A** and **Figure 3A**). For randomly oriented samples in the solid state, (so-called powder spectra) the distribution of molecules in different orientations gives rise to characteristic broadening of the line shape (**Figure 2B** and **Figure 3C**).

### Effects of temperature, electron-spin relaxation and spectral linewidths

Temperature has a number of effects on the EPR spectrum. The populations of the electron energy levels, $N_-$ and $N_+$ (**Figure 1A**) differ by a small amount because of the Boltzmann distribution:

$$\frac{N_+}{N_-} = \exp\left(\frac{\Delta E}{k_B T}\right) = \exp\left(\frac{g\mu_B B_0}{k_B T}\right)$$

where $k_B$ is the Boltzmann constant and $T$ the absolute temperature. At room temperature and a magnetic field of 0.32 T, the difference is about 0.07% of the unpaired electrons. It increases at

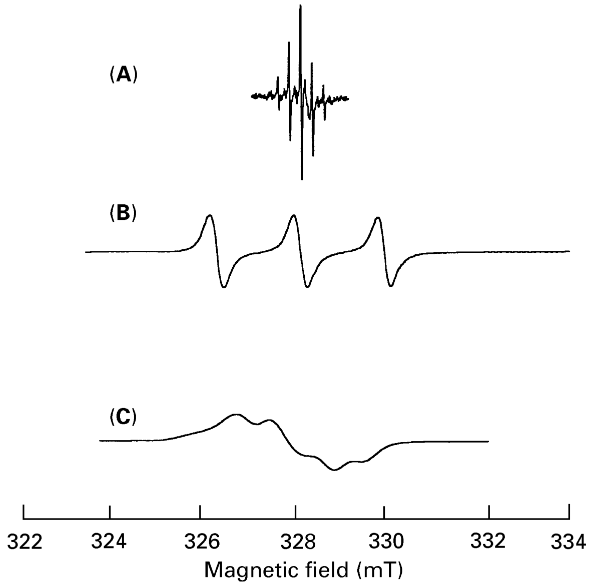



**Figure 3** EPR spectra of radicals (A) benzoquinone anion radical in solution; (B) the nitroxide TEMPOL in solution of photosynthetic reaction centres.

higher microwave frequencies (hence higher fields), and at lower temperatures.

The spin–lattice relaxation, with a characteristic time $T_1$, is responsible for maintaining the population difference between levels, $N_-$ and $N_+$. The spin–spin relaxation time $T_2$ reflects the lifetime of the excited state and its effect on the line width. If the electron-spin relaxation rate is too rapid, the lifetime of the excited state is short and the EPR spectrum becomes broadened. At high temperatures the spectrum may become too broad for detection, hence the use of cryogenic temperatures for some transition ions. However, if the spin–lattice relaxation is too slow, the population difference $N_- - N_+$ cannot be maintained, and the amplitude of the signal is attenuated, a situation known as microwave power saturation. Electron-spin relaxation times may be estimated by measuring the amplitude of the signal as a function of applied microwave power.

## Operation of the spectrometer

The outline of a conventional continuous-wave EPR spectrometer is shown in **Figure 4**. The term 'continuous-wave (CW)' refers to the fact that microwaves are applied throughout the measurement. It is designed to optimize the weak EPR signals originating from the sample, against a background of noise from the source. The microwave source is a reflex klystron (a type of thermionic valve) or a Gunn diode (a semiconductor device). The detector is a microwave diode. The operating frequency is typically in the X-band (~9.5 GHz). This frequency is a compromise; higher frequencies increase the ratio $N_-$:$N_+$ (and hence sensitivity) but require smaller cavities and thus smaller sample size (and hence number of spins being measured). A typical sample volume is of the order of 0.1 $cm^3$.

The microwave circuit is known as the microwave bridge, to emphasize that it is a balanced circuit. The source is tuned to the resonant frequency of the cavity. Microwaves pass through waveguides to the cavity via a circulator, which ensures that signal arising from the sample in the cavity passes to the detector. The reference arm serves to balance the signal with the input microwave power. The circuit is optimally tuned by matching its impedance with the cavity using the iris, to a condition known as critical coupling. In this situation the intensity of the microwave field $B_1$ at the sample is enhanced by the $Q$-factor of the cavity, typically several thousand-fold. When electron paramagnetic resonance occurs in the sample, a small amount of power is transmitted to the detector and is amplified and recorded.

The homogeneous $B_0$ field is provided by an electromagnet. EPR spectra are obtained by modulation of the field by a small amount, at a frequency of 100 kHz. A lock-in amplifier detects the EPR signal in phase with the modulation. Modulation and phase-sensitive detection is a standard method of noise rejection and enhances the signal-to-noise ratio, and hence the sensitivity of the technique, by as much as $10^5$. It gives rise to the characteristic first-harmonic (first-derivative) line shape. Although the absorption spectrum could be derived from this by integration, spectra are usually retained as the derivative form as it emphasizes narrow hyperfine splittings and other features.

### Temperature control

The form of the paramagnetic material may be solid, liquid or gas, and so appropriate temperature control of the sample is needed. The temperature of the sample may be controlled by flowing gas through a quartz vacuum jacket. For low temperatures, helium is used. The sample may be held in a Dewar of liquid nitrogen or helium. For smaller resonator designs such as loop-gaps, the whole resonator is immersed in a cryostat containing helium or nitrogen.

### Resonators

The quality, or $Q$ factor, of the microwave cavity or resonator is a measure of its efficiency at concentrating microwave energy. For a resonant a.c. circuit it is given by

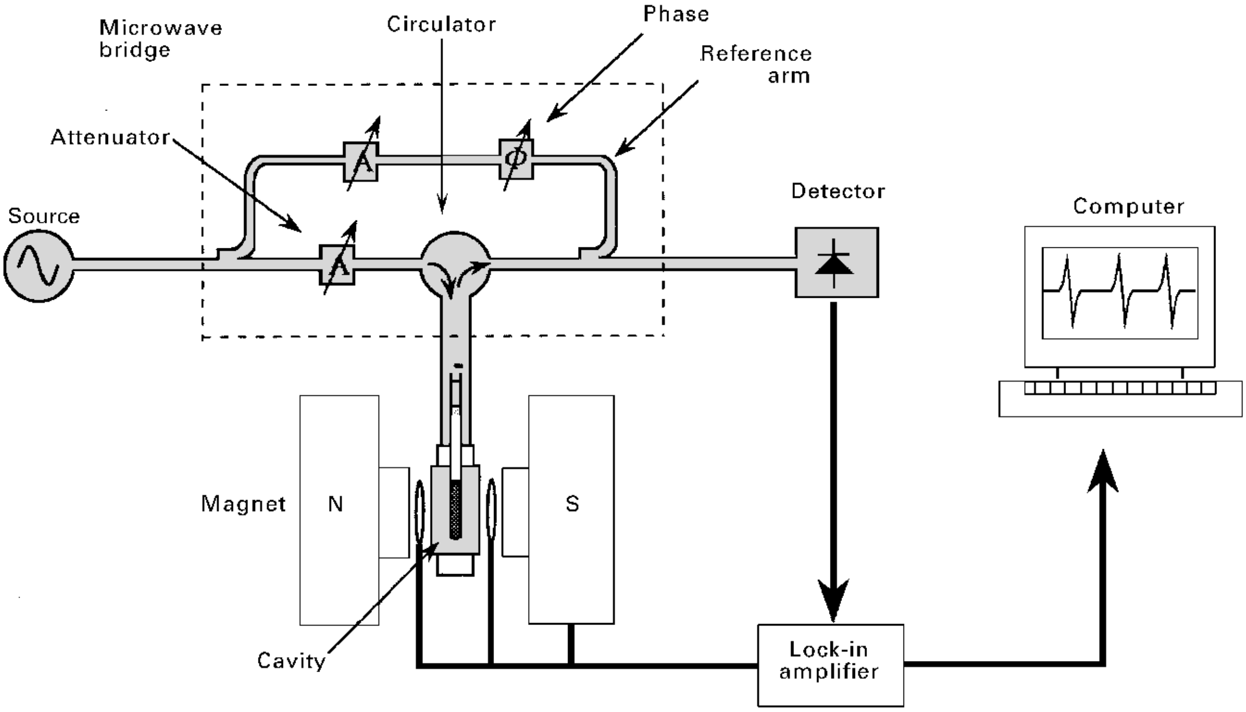



**Figure 4** Outline of a typical continuous-wave EPR spectrometer.

$$Q = \frac{2\pi(\text{Energy stored})}{(\text{Energy dissipated per cycle})}$$

The filling factor $\eta$ is the fraction of the microwave magnetic field $B_1$ which bisects the sample.

Two examples of resonators used in EPR are illustrated in **Figure 5**. Rectangular cavities [**Figure 5(A)**] are of metal which is coated with silver and gold for high conductivity. The sample is positioned at the maximum of the microwave magnetic field, $B_1$. It may be irradiated with ultraviolet or visible light, in this example through a waveguide of length $\lambda$ which prevents microwave leakage. The applied magnetic field $B_0$ is modulated with Helmholtz coils in the walls of the cavity. The $B_0$ and $B_1$ fields are oriented at right angles, which gives a maximum magnetic resonance interactions for $S = \frac{1}{2}$ paramagnets. For paramagnets with an integer spin, the resonant transition is forbidden in this geometry. Parallel mode cavities are used for this situation, in which $B_1$ is parallel to $B_0$.

In the loop-gap resonator (**Figure 5B**), the coupling loop acts as an antenna to transmit microwave power to and from the bridge to the split-ring resonator element. The number of gaps is variable, depending on the sample size. The microwave magnetic field $B_1$ circulates in the ring, while the electric component is mostly confined to the gaps. Magnetic field modulation is also by external coils (not shown). Loop-gap resonators have better filling factors than rectangular cavities, so that the resonator can be smaller and the sample tube larger. They are particularly effective for lossy samples such as aquesous solutions.

Other designs of cavity are possible, including cylindrical resonators constructed of an inert dielectric material such as sapphire (alumina).

### Computer control and data acquisition

Most operations of the instrument are under computer control. The spectrum is recorded as a series of data points, along a linear axis of magnetic field. The signal-to-noise ratio can be improved by repetitive scanning and signal averaging. Fourier transform techniques are normally only used for analysis of pulsed EPR experiments. There is a range of different software for analysis of the spectroscopic data, in particular for simulation.

### Sample holders and sample preparation

The holders for EPR samples are constructed of a diamagnetic, nonconductive material such as quartz. Electrically conductive materials such as metals cannot be used as they reflect microwaves and prevent penetration of microwaves into the sample.

It is not possible to measure EPR spectra of large aqueous samples at frequencies above 1 GHz

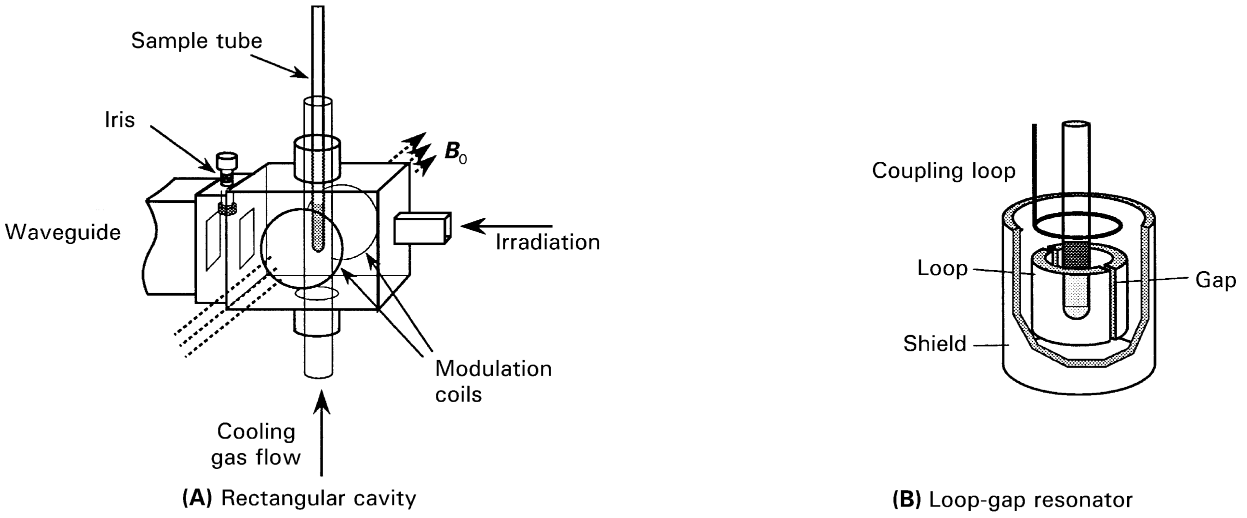



**Figure 5** Resonators for EPR spectroscopy. (A) Rectangular cavity; (B) loop-gap resonator.

because of microwave power losses. For measurements in an X-band rectangular cavity, specially constructed flat cells are used, with a thickness of about 0.3 mm (**Figure 6**). They are placed in the centre of the cavity, which is a nodal plane of the electric vector, hence minimizing dielectric losses.

Oxidation–reduction reactions are used to generate organic or inorganic compounds in their paramagnetic oxidation states, for example quinones in their semiquinone states, or transition metal ions in their paramagnetic states (**Table 2**). Electrolytically generated radicals may be obtained by means of electrodes placed in the cell above and below the cavity.

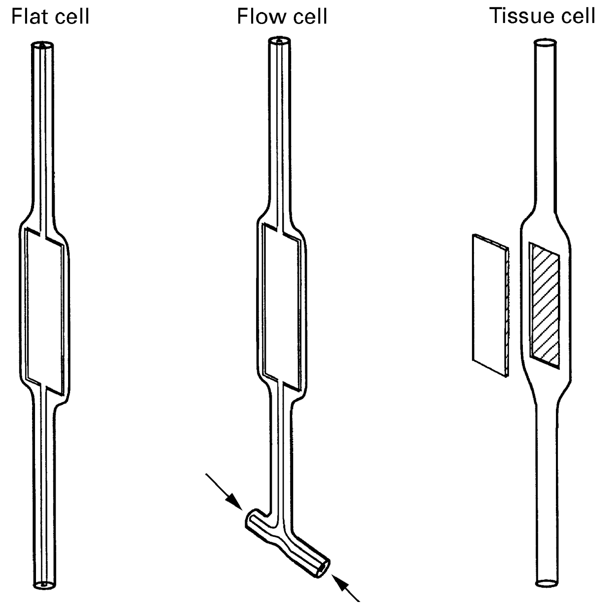



**Figure 6** Sample holders for aqueous samples.

**Table 2** Examples of paramagnetic transition ions

| *Metal ion* | *Paramagnetic oxidation states* | *Other states* |
|---|---|---|
| Vanadium | $V^{IV}$ | $V^{V}$,$V^{III}$ |
| Chromium | $Cr^{III}$, $Cr^{VI}$ | |
| Manganese | $Mn^{II}$, $MN^{IV}$ | $Mn^{III}$ |
| Iron | $Fe^{III}$ | $Fe^{II}$ |
| Cobalt | $Co^{II}$ | $Co^{I}$, $Co^{III}$ |
| Nickel | $Ni^{I}$, $Ni^{III}$ | $Ni^{II}$ |
| Copper | $Cu^{II}$ | $Cu^{I}$ |
| Molybdenum | $Mo^{V}$ | $Mo^{IV}$, $Mo^{VI}$ |

Note – other oxidation states and isotopes exist for these elements.

For kinetic measurements, paramagnets may be generated by mixing solutions and flowing them into a cell inside the cavity (**Figure 6B**). For short-lived states of transition metals in solution, samples may be obtained by rapid freezing, in which samples are mixed and squirted into a bath of isopentane at 140 K.

## Multifrequency EPR

Although the majority of EPR spectra are still taken at X-band frequency, there are advantages for some applications to use lower or higher microwave frequencies, or a range of frequencies. To operate at each frequency it is usually necessary to use a different microwave bridge and cavity. These frequencies are identified by their frequency bands, as listed in **Table 3**.

In high-frequency spectrometers, superconducting magnets are used. Up to 100 GHz, the microwaves

**Table 3** Microwave wavebands and frequencies

| Band | Frequency (GHz) (app.) | λ (mm) | Field for g = 2 (mT) | Typical applications |
|---|---|---|---|---|
| Radiofrequency | 0.25 | 1200 | 8.93 | EPR imaging |
| L | 1 | 300 | 35.7 | Aqueous sample EPR |
| S | 4 | 75 | 143 | Hyperfine splittings |
| X | 9.5 | 31.6 | 339 | Routine spectroscopy |
| K | 24 | 12.5 | 857 | |
| Q | 35 | 8.6 | 1250 | Electron spin-spin interactions |
| V | 65 | 4.6 | 2322 | |
| W | 96 | 3.12 | 3430 | *g* Factor anisotropy in radicals |
| D | 140 | 2.14 | 5000 | Large zero-field splittings |
| G | 250 | 1.2 | 8931 | Even spin system |

The definitions of bands follow EPR convention and the frequencies are approximate.

may be obtained by mixing with lower-frequency radiation in a heterodyne arrangement. For higher frequencies, the radiation may be regarded as far infrared, and Fabry–Pérot resonators are used which consist of two parabolic mirrors to concentrate the radiation. In low-frequency spectrometers the magnetic field is produced by Helmholtz coils and the resonator is a coil similar to that used in NMR.

Some features of the EPR spectrum increase with magnetic field, such as the separation between $g$ factors. Others, such as the separation between hyperfine splittings, do not change, to first order. By altering the microwave frequency it is possible to resolve the effects of multiple electron–nuclear and electron–electron interactions. The EPR spectra of organic radicals at X-band (9 GHz) are dominated by hyperfine splittings by protons and other nuclei. At W-band (96 GHz) the spectra are dominated by the anisotropy of the $g$ factor.

## Electron–nuclear double resonance (ENDOR) and TRIPLE resonance

The hyperfine interactions with nuclei adjacent to the paramagnetic centre are often too small to be observed because they are smaller than the EPR line width. Electron–nuclear double resonance is a technique which detects these interactions. In addition to the microwave field $B_1$ a radiofrequency field, $B_2$, is applied at right angles. A commercial design uses a coil wound around the sample tube in a special cavity.

The energy levels and transitions relevant to the ENDOR measurement are illustrated in **Figure 7**. The microwave field, at the EPR resonant frequency, partially saturates the signal, equalizing the populations of the $m_S = \pm\frac{1}{2}$ levels. The radiofrequency is swept, and when the latter is at the resonant frequency of the interacting nuclear spins this provides further pathways for electron spin relaxation and the EPR signal is enhanced. The resulting spectrum consists of a pair of lines. The positions of the lines depend on the relative magnitudes of the hyperfine coupling, $A$, and the nuclear Zeeman frequency, $\nu_n$ (the NMR frequency of the nucleus) (**Figure 7**). In the case where $\nu_n > A/2$, as for protons, which have a large nuclear moment, the lines will be centred at $\nu_n$ and separated by $A$. When $A/2 > \nu_n$, as in the case of strong couplings to other nuclei, the ENDOR spectrum consists of two lines centred at $A/2$ and separated by $2\nu_n$.

Hyperfine interactions of the order of a few megahertz can be resolved and provide information about nuclei at distances up to 0.5 nm. An additional advantage of ENDOR is that the number of lines is additive, rather than multiplicative. If there are numerous nuclear spins in the paramagnetic molecule, each type of hyperfine splitting will give just one pair of lines and so the ENDOR spectrum is simpler than the EPR spectrum. This feature is particularly useful in the interpretation of the complex spectra of radicals.

TRIPLE resonance is another method which resolves further a spectrum containing multiple hyperfine couplings. It is analogous to the Overhauser effect in NMR. By continuously applying radiofrequency radiation at the resonant frequency of one line in the spectrum, other lines in the ENDOR spectrum are suppressed or enhanced if they are connected to the same transition. This method makes it possible to determine the sign of $A$, which is used in molecular orbital calculations.

## Pulsed EPR

Pulsed EPR spectrometers are now commercially available and have some of the same electronic features as NMR spectrometers. The layout of the microwave bridge is similar to that of **Figure 4**, but in addition a pulse programmer controls the

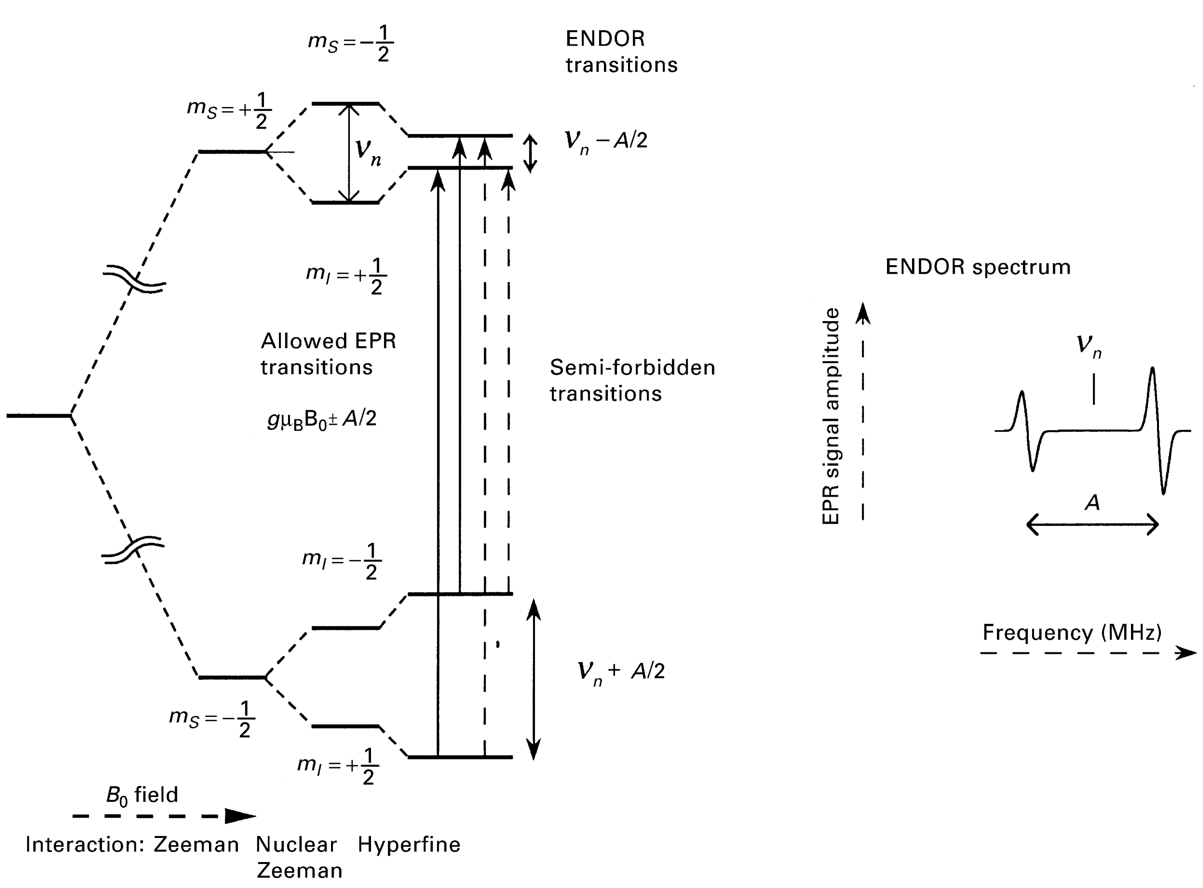



**Figure 7** Energy levels for an electron interacting with an $I = \frac{1}{2}$ nucleus, showing the transitions relevant to ENDOR and electron-spin echo envelope modulation (ESEEM).

microwave supply through diode switches. The microwave pulses, of duration 5–50 ns, are amplified by a travelling-wave tube amplifier, to a power up to 1 kW. The cavity has a lower $Q$-factor than for CW EPR to allow the microwave field to build up more quickly. The main differences between pulsed EPR and pulsed NMR arise from the broad bandwidth of the EPR spectrum.

### Flash photolysis

For the spectra of radicals that are dispersed over a narrow range of frequency, it is possible to detect a free-induction decay (FID). This makes it possible to derive a spectrum by Fourier transformation and observe the spectrum of a transient species induced, for example, by a laser flash which is synchronized with the microwave pulse programmer (**Figure 8A**).

Normally, however, the spectrum is too broad, or the FID decays too rapidly. In this case it is possible to detect an electron-spin echo after a two-pulse ($\pi/2$–$\pi$–echo) sequence. The spectrum of the flash-induced species is derived by repeating the measurement over a range of magnetic field strength. The lifetime (which should be a minimum of about 100 ns) is measured by repeating the electron-spin echo sequence, while the interval between the flash and the pulses is systematically incremented.

### Electron-spin relaxation rates

The values of $T_1$ and $T_2$ are measured in different experiments which observe the decay of the electron-spin echo with time; $T_1$ is measured with an inversion-recovery experiment as in NMR, while $T_2$ is observed by a two-pulse echo sequence, as in **Figure 8A**. The amplitude of the echo decreases with increasing delay time $\tau$, owing to loss of 'phase memory'. The spin–spin relaxation is affected by various factors, including molecular motion and spin–spin interactions between paramagnets.

### Hyperfine couplings

Electron-spin echo envelope modulation (ESEEM) spectroscopy is a pulsed technique that detects weak hyperfine and quadrupolar couplings in the solid state. It is complementary to ENDOR in that it gives

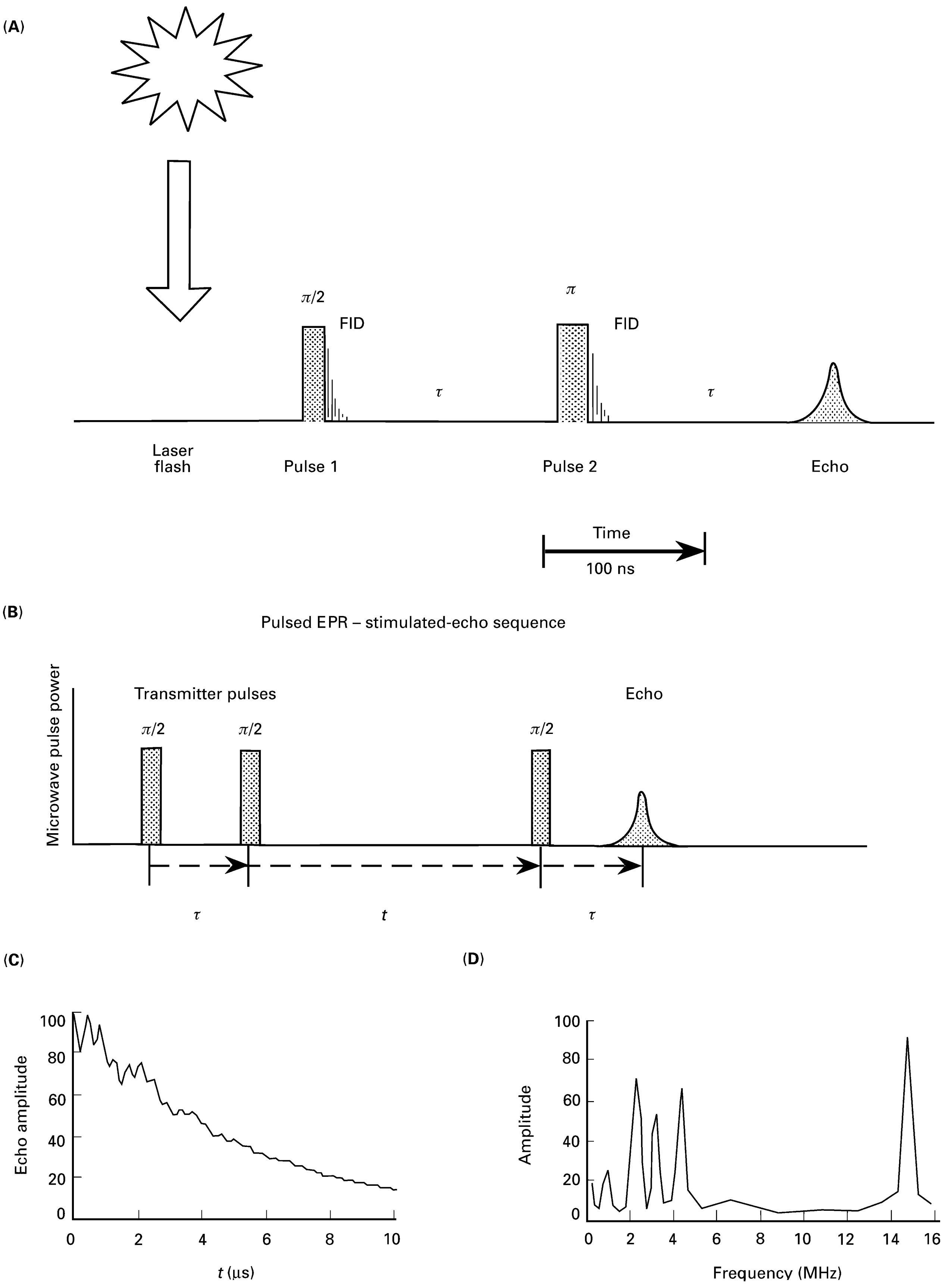



**Figure 8** Pulsed EPR experiments. (A) flash photolysis, (B)–(D) ESEEM.

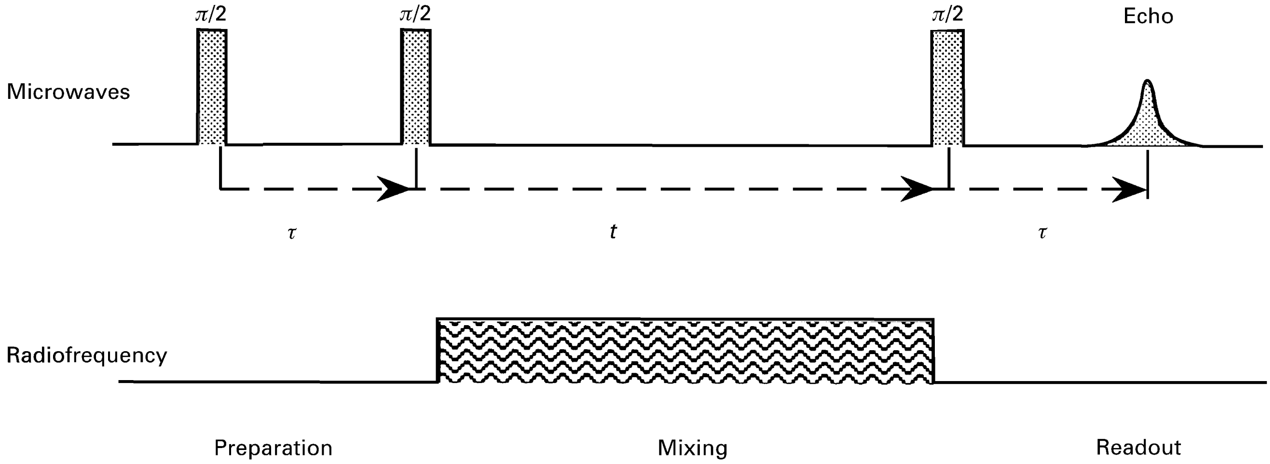



**Figure 9** Principle of Mims pulsed ENDOR.

the best results in the intermediate regime of hyperfine couplings where $A \approx \nu_N$; in addition, it is particularly sensitive to couplings with quadrupolar nuclei such as $^2H$ and $^{14}N$. The Mims three-pulse echo sequence is similar to the two-pulse sequence (**Figure 8B**) except that the $\pi$ pulse is split into two $\pi/2$ pulses separated by time $T$. The decay of the echo is now determined by the slower spin–lattice relaxation time $T_1$ allowing echoes to be detected over a longer period of time. During the evolution of the spin echo, hyperfine coupling frequencies cause dephasing of the electron spins, and so the echo amplitude is modulated with time (**Figure 8C**). This can be considered as being due to an interference between the frequencies of the allowed transitions, $\Delta m_S = \pm 1$, with the 'semi-forbidden' transitions, $\Delta m_S = \pm 1$, $\Delta m_I = \pm 1$, which are allowed owing to quadrupolar interactions and anisotropy of the hyperfine interaction. Fourier transformation of the time-domain spectrum yields the spectrum of hyperfine frequencies (**Figure 8D**).

Two-dimensional pulsed EPR experiments are also possible, such as HYSCORE (hyperfine sublevel correlation) spectroscopy. This uses a pulse sequence as in **Figure 8B**, separated by an inversion pulse into two variable time intervals. Fourier transformation in the two time-domains gives a 2D spectrum, in which the effects of multiple hyperfine couplings can be resolved.

Pulsed ENDOR is an alternative technique, in which radiofrequency pulses at NMR frequencies are used to perturb the spin-echo sequence (**Figure 9**). Hyperfine interactions lead to loss of coherence of the electron spins during the mixing period, so the amplitude of the echo is decreased. A plot of echo amplitude against RF frequency yields a spectrum analogous to that of a CW ENDOR spectrometer.

## Types of material studied

### Free radicals

Organic free radicals can be produced by one-electron oxidation or reduction reactions, by irradiation processes or by homolytic cleavage of a chemical bond. They may occur in the solid, liquid or gaseous states. Organic radicals in solution give complex EPR spectra. The hyperfine structure of their spectra provides information about the type of radical and the distribution of the unpaired electron in it. Free radicals tend to be unstable and to decay by recombination processes. However, some radicals, particularly those in the solid state, are stable. If they have a half-life of at least a few seconds they can readily be examined by conventional EPR spectroscopy, using rapid-scan coils to sweep the $B_0$ field. If they are more short-lived, the radical may be stabilized by cooling to low temperatures, or chemically trapped with a spin trap which reacts with the unstable species to produce a more stable radical.

### Spin traps

Ideally, a spin trap is a diamagnetic compound, which is introduced into a biological system without perturbing it and which reacts rapidly and specifically with a transient free radical to form a stable radical that can be identified from its EPR signal. These requirements are somewhat contradictory. The two most commonly-used spin traps for organic radicals are nitroso compounds and nitrones (**Table 4**). These react with radicals to form nitroxides. The nitroso compounds give an adduct in which the radical added is bonded directly to the nitroxide nitrogen, so that the hyperfine splittings in the EPR signal are more diagnostic of the original radical; however,

the adducts are less stable and may decay within minutes. Nitrones tend to give more stable adducts which are more difficult to identify.

### Spin labels

Spin labels are stable paramagnetic molecules that are introduced into a biological sample, and extend the use of EPR to situations where no endogenous radicals exist. The commonest type of spin labels are nitroxides (aminoxyl) radicals, though other types of radicals and transition metal ions have been used. The advantage of the specific labelling approach is that it makes it possible, through judicious labelling, to investigate the local environment and motion of defined parts of the system under study.

### Transition metals

Transition metals are paramagnetic in certain oxidation states (**Table 2**). Although this paramagnetism is stable, the sample may have to be cooled to low temperatures for the EPR signal to be observed.

### Defects in solids

Examples of defects are cation vacancies (V centres), anion vacancies (F centres) and impurities in crystals. Defects can be generated in various ways, such as irradiation with ultraviolet or ionizing radiation, or by imperfect crystallization. An example is a defect (latent image) generated in photographic emulsion by light irradiation. In addition, finely divided solid materials commonly show radical centres due to homolytic cleavage of chemical bonds.

## Information from EPR spectroscopy

### Identification of the paramagnet

The type of paramagnetic species may be recognized from the characteristic features of its spectrum such as $g$ factors and hyperfine splittings.

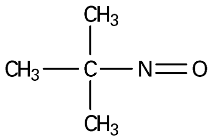
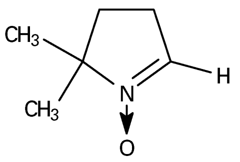
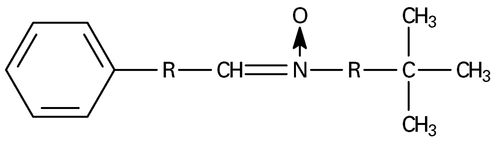
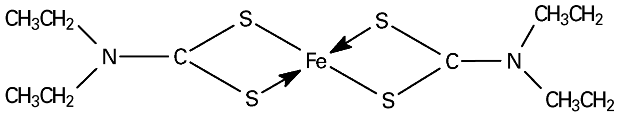
For free radicals in solution, examination of the pattern of hyperfine splittings can provide information about the number of interacting nuclei and hence the structure of the radical. For example, in the benzoquinone radical (**Figure 3A**), the splitting by four equivalent protons gives rise to five lines with amplitudes in the ratio 1:4:6:4:1.

A powerful technique is to substitute atoms in the molecule with nuclei having different nuclear spins, for example deuterium $^{2}H$ ($I$ = 1) for $^{1}H$ or $^{13}C$ ($I = \frac{1}{2}$) for $^{12}C$. Examples of nuclear spins and moments are given in **Table 1**. The observed change in the hyperfine splitting pattern is conclusive evidence of interaction with the nucleus.

### Quantification

The concentration of unpaired electrons can be determined from the integrated intensity of the EPR spectrum by double integration or by numerical simulation of the spectrum. For paramagnets with $S = \frac{1}{2}$, measured under non-saturating conditions, the

**Table 4** Types of spin traps

| *Type / spin trap* | *Structure* | *Suitable application* |
|---|---|---|
| 2-Methyl-2-nitrosopropane (MNP) |  | C,S-centred radicals |
| 5,5-Dimethyl-1-pyrroline-*N*-oxide (DMPO) |  | C,S,O-centred radicals |
| *N-tert*-Butyl-$\alpha$-phenylnitrone (PBN) |  | C,S,O-centred radicals |
| $Fe^{II}$(diethyldithiocarbamate)$_2$ |  | Nitric oxide |

integrated intensity for a given concentration is the same, regardless of the nature of the paramagnetic species. Hence the concentration can be referred to a standard material such as $Cu^{II}$.

### Spin Hamiltonian parameters

The parameters of the EPR spectrum are described formally by a spin Hamiltonian equation, which summarizes the energies of the different types of interaction. For example, the Hamiltonian for a spin $S$ interacting with a nuclear spin $I$ is

$$\mathcal{H} = g\mu_B B_0 S + AIS$$

where the operators $g$ and $A$ are shorthand notation for $3 \times 3$ matrices or tensors. The EPR spectrum may be interpreted by computer simulation of the energy of interaction with the magnetic field $B_0$. This provides information about the electronic state of the paramagnet. For radicals, the information can be used in detailed molecular orbital calculations of electronic structure. For transition metal ions, EPR provides information about the coordination geometry, spin state and types of ligands.

### Redox and spin states

Changes in oxidation states of paramagnets, by adding or removing unpaired electrons, cause the EPR signals to appear and disappear. This can be used to monitor oxidation–reduction reactions, and to measure reduction potentials. Generally the EPR spectrum is a useful indication of the oxidation–reduction state of the species observed. Often the presence of an EPR signal is conclusive, though in the case of semiquinones they may be anion radicals (formed by electron addition) or cation radicals (formed by removal of an electron). The EPR spectra of $Ni^{I}$ and $Ni^{III}$ are similar. In these cases further oxidation–reduction experiments are needed to distinguish them.

For transition ions such as $Fe^{III}$, the EPR spectrum is a good indicator of spin state, in this case high spin ($S = \frac{5}{2}$) or low spin ($S = \frac{1}{2}$). The spectra of the high-spin states are influenced by zero-field splittings, in this case giving $g$ factors of up to 10.

### Distances

The hyperfine coupling between an electron spin and a nucleus or between electron spins consists of two components: an exchange interaction (acting through chemical bonds, and generally isotropic), and a dipolar interaction. The dipolar component is anisotropic and its magnitude is inversely proportional to the $(\text{distance})^3$, and this can be used to estimate distances. In favourable cases, hyperfine couplings from ENDOR can be used to determine distances of protons up to 0.5 nm. Electron spin–spin interactions can be observed as splittings in the EPR spectrum over distances up to 1.5 nm, and relaxation effects at distances up to 4 nm.

### Orientation

If the molecules of a paramagnet are orientated, as in a single crystal, the EPR signal consists of narrow lines, which shift as the crystal is rotated in the magnetic field. The angular dependence can be measured precisely using a goniometer. Some ordering is seen in fibres and layered materials, though less precisely than in crystals. From this orientation dependence it is possible to estimate the spin Hamiltonian parameters, yielding detailed information about the axes of the $g$ and $A$ matrices, and hence the orbital states.

### Motion

The line width of the EPR spectrum of a radical such as a nitroxide in solution is sensitive to motion in the range of correlation times $10^{-7}$–$10^{-11}$ s. This is due to the anisotropy of the $^{14}N$ hyperfine splitting. If the motion is rapid enough to average this anisotropy, the spectrum shows narrow lines. For slower motion the spectrum broadens out. The rate of motion is determined by simulation, or by comparison with labelled molecules of known correlation time. By means of an experimental technique known as saturation-transfer EPR the spectra may be made to be sensitive to correlation times up to $10^{-3}$ s, as found, for example, in membrane-bound proteins.

### EPR dosimetry

The build-up of radicals in irradiated solids forms a method of estimating the absorbed dose of ionizing radiation. A standard used in EPR dosimeters is the amino acid alanine, in pellet form, which gives a linear relationship between radical signal amplitude and absorbed dose over several orders of magnitude. Where there has been accidental radiation exposure, the EPR of radicals in hydroxyapatite in tooth enamel has been used to estimate the absorbed dose. Irradiation of foods such as meat may be detected by radicals formed in bone.

### EPR dating

Buried archaeological samples that contain crystalline materials show radical EPR signals due to the effects of the radiation in the soil. The time of burial

can be estimated by taking measurements of the radioactive isotopes in the surrounding soil and observing the signal amplitude after applying an equivalent dose of radiation to the sample. The decay of the radical signal over time depends on the stability of the defects in the material. Dating works best for crystalline material such as silica; bone is unreliable because of recrystallization. The technique has been used for samples dated between $10^5$ and $10^6$ years, which is between the ranges possible for radiocarbon and potassium dating.

### Medical applications

Oxygen radicals have been widely investigated in biomedical systems and are of great interest because of their cytotoxic effects and involvement in processes such as inflammation. They may be observed by spin traps (**Table 4**). Nitric oxide is also generated in conditions such as inflammation, and can be trapped by iron complexes such as iron diethyldithiocarbamate.

Oximetry exploits another feature of the spectra of radicals in solution: that they are broadened by paramagnetic interactions with oxygen in solution. With narrow-line radicals it is possible to measure oxygen in human tissues at concentrations below $10^{-5}$ mol $dm^{-3}$.

## List of symbols

$A$ = hyperfine coupling constant; $B_0$ = static magnetic field; $B_1$ = microwave magnetic field; $B_2$ = radiofrequency magnetic field; $g$ = electron $g$ factor; $h$ = Planck's constant; $H$ = Hamiltonian operator; $I$ = nuclear spin operator; $k$ = Boltzmann constant; $m_I$ = nuclear spin angular momentum; $m_S$ = electron spin angular momentum; $N_-$, $N_+$ = populations of electron energy levels separated by a magnetic field; $Q$ = quality factor of a resonator; $S$ = electron spin operator; $T$ = absolute temperature; $T_1$ = longitudinal (spin–lattice) relaxation time; $T_2$ = transverse (spin–spin) relaxation time; $\eta$ = filling factor of a resonator; $\lambda$ = microwave wavelength; $\mu_B$ = Bohr magneton. In an EPR spectrum the signals are obtained by increasing the $B_0$ field at fixed frequency; $\nu$ = microwave frequency; $\nu_n$ = nuclear Larmor frequency; $\tau$ = delay in pulse sequence.

*See also:* **Chemical Application of EPR; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; EPR Imaging; EPR Spectroscopy, Theory; Fluorescence Polarization and Anisotropy; Fourier Transformation and Sampling Theory; Laboratory Information Management Systems (LIMS); NMR in Anisotropic Systems, Theory; NMR Pulse Sequences; NMR Relaxation Rates; Nuclear Overhauser Effect; Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy.**

## Further reading

Abragam A and Bleaney B (1970) *Electron Paramagnetic Resonance of Transition Ions*. Oxford: Oxford University Press.

Atherton NM (1993) *Principles of Electron Spin Resonance*. Chichester: Ellis Horwood.

Bencini A and Gatteschi D (1989) *EPR of Exchange Coupled Systems*. Berlin: Springer-Verlag.

Berliner LJ and Reuben J (1993) *EMR of Paramagnetic Molecules: Biological Magnetic Resonance*, Vol 13. New York: Plenum.

Czoch R and Francik A (1989) *Instrumental Effects in Homodyne Electron Paramagnetic Resonance Spectrometers*. Chichester: Ellis Horwood.

Dikanov SA and Tsvetkov YD (1992) *Electron Spin Echo Envelope Modulation (ESEEM) Spectroscopy*. Boca Raton, FL: CRC Press.

Eaton GR, Eaton SS and Salikhov KM (1997) *Foundations of Modern ESR*. Singapore: World Scientific Publishing.

Eaton GR, Eaton SS and Ohno K (1991) *ESR Imaging and In Vivo ESR*. Boca Raton, FL: CRC Press.

Hoff AJ (ed) (1989) *Advanced ESR: Applications in Biology and Biochemistry*. Amsterdam: Elsevier.

Keijzers CP, Reijerse EJ and Schmidt J (1989) *Pulsed ESR – A New Field of Applications*. Amsterdam: North Holland.

Kurreck H, Kirste B and Lubitz W (1988) *Electron Nuclear Double Resonance Spectroscopy of Radicals in Solution*. Weinheim: VCH Publishers.

Lowe DJ (1995) *ENDOR and ESR of Metalloproteins*. Berlin: Springer.

Mabbs FE and Collison D (1992) *Electron Paramagnetic Resonance of d Transition Metal Compounds*. Amsterdam: Elsevier.

Pilbrow JR (1990) *Transition Ion Electron Paramagnetic Resonance*. Oxford: Oxford University Press.

Poole CPJ (1983) *Electron Spin Resonance*, 2nd edn. New York: Wiley.

*Specialist Periodical Reports on ESR Spectroscopy*. Cambridge: Royal Society of Chemistry.

Weil JA, Bolton JR and Wertz JE (1994) *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications*. New York: Wiley.

# ESR Imaging

*See* **EPR Imaging.**

# ESR Methods

*See* **EPR, Methods.**

# ESR Spectroscopy, Applications in Chemistry

*See* **Chemical Applications of EPR.**

# ESR Spectroscopy, Theory

*See* **EPR Spectroscopy, Theory**.

# Exciton Coupling

**Nina Berova**, Columbia University, New York, NY, USA
**Nobuyuki Harada**, Tohoku University, Sendai, Japan
**Koji Nakanishi**, Columbia University, New York, NY, USA

Copyright © 1999 Academic Press

ELECTRONIC SPECTROSCOPY
**Theory & Applications**

Circular dichroic spectroscopy (CD) is a powerful method for determining the absolute configurations and confirmations of chiral molecules. In particular, the exciton-coupled circular dichroism arising from through-space chromophore- chromophore interactions, called exciton coupling, is one of the most sensitive spectroscopic tools for determining the molecular chirality or molecules in solution. The theoretical analysis of the exciton coupling phenomena is based on the coupled oscillator theory and molecular exciton theory. Since exciton-coupled CD is basically a nonempirical method, the derived conclusions are

nonambiguous. It is a microscale technique of great versatility applicable to compounds ranging from small molecules to various types of ligand- receptor complexes. However, the extent of its potential has yet to be explored. Here the theory and applications of exciton-coupled CD are outlined.

When two chromophores with an identical or similar excitation energy are in close spatial proximity to one another and chirally disposed, the individuality of a monomer chromophore transition is lost. The chromophore excited state can delocalize over all chromophores within the system and becomes an exciton. The excitons interact and couple, thus giving rise to a characteristic pair of intense CD bands with opposite signs and comparable band areas, called a CD exciton couplet. The coupling of two identical chromophores is illustrated in **Figure 1** by a steroidal 2,3-bis-*para*-substituted benzoate. The intense intramolecular charge-transfer (CT) bands stemming from the $^1L_a$ transition of *para*-substituted benzoates give rise to the main electronic absorption band; it is the through-space coupling between these CT bands which is used for determining the absolute configuration between two hydroxyl groups. Although there is some free rotation around the C- O bonds, ester bonds are known to adopt the *s-trans* confirmation. Furthermore, X-ray crystallographic data, energy calculations, etc., show that the ester C=O is *syn* with respect to the methine hydrogens. Thus, the intramolecular CT transition polarized along the long axis of the benzoate chromophore is almost parallel to the alcoholic C- O bond. Hence the exciton CD data which represent the absolute chirality between the transition moments also represent the chirality between any two hydroxyl groups, provided that their spatial distance permits the two chromophores to couple. When the absolute sense of chirality between the C(2)- O and C(3)- O bonds constitute a clockwise twist as shown by the Newman projection in **Figure 1**, it is defined as positive, and vice versa.

The outline of exciton-coupled CD in a system consisting of two identical chromophores (a degenerate case) is as follows:

(i) Chromophores exhibit intense $\pi$- $\pi^*$ absorptions.
(ii) The excited state of the system is split into two energy levels, $\alpha$ and $\beta$, by exciton interaction between the two chromophores, and the energy gap $2V_{ij}$ is called the Davydov splitting.
(iii) In the UV-visible spectra, this coupling gives rise to two-component spectra of the same sign

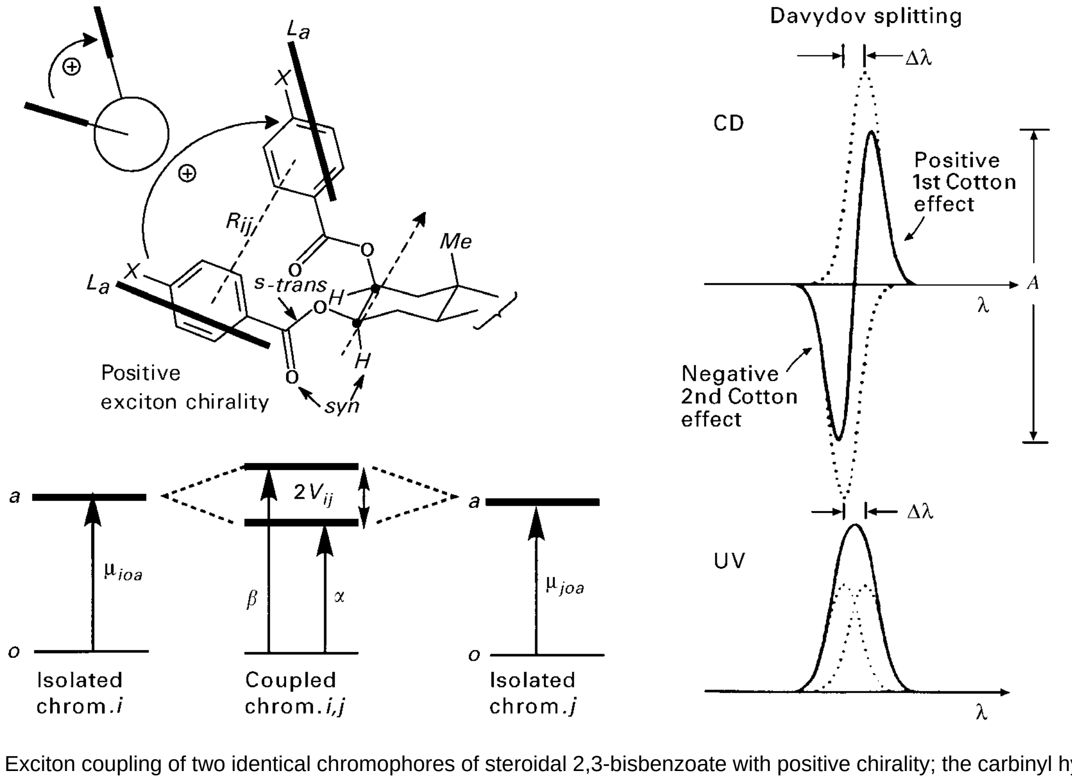
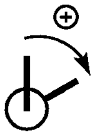
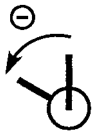



**Figure 1** Exciton coupling of two identical chromophores of steroidal 2,3-bisbenzoate with positive chirality; the carbinyl hydrogens are represented *syn* to the ester carbonyl and the ester groups is in an *s-trans* conformation. Also shown are the splitting of the excited states of isolated chromophores *i* and *j* into $\alpha$ and $\beta$ states with the Davydov splitting $2V_{ij}$, summation CD and UV curves (solid line) of two component curves (dotted line).

**Table 1** Definition of exciton chirality for a binary system

| | *Qualitative definition* | *Quantitative definition* | *Cotton effects* |
|---|---|---|---|
| **Positive exciton chirality** |  | $\boldsymbol{R}_{ij}\cdot(\mu_{ioa}\times\mu_{joa})\,V_{ij}>0$ | **Positive first and negative second Cotton effects** |
| **Negative exciton chirality** |  | $\boldsymbol{R}_{ij}\cdot(\mu_{ioa}\times\mu_{joa})\,V_{ij}<0)$ | **Negative first and positive second Cotton effects** |

which usually appears as a single absorption maximum with double intensity. In contrast, in CD, this excitation to two split energy levels gives rise to two CEs of opposite signs; summation of the two curves of opposite signs separated by energy gap $\Delta\lambda$ results in a CD with two extrema, a couplet (commonly called a 'split CD'). The extrema at longer and shorter wavelengths are called the first CE (Cotton effect) and second CE, respectively.

(iv) The amplitude of exciton split CE is defined as the *A* value ($A = \Delta\varepsilon_1 - \Delta\varepsilon_2$), where $\Delta\varepsilon_1$ and $\Delta\varepsilon_2$ are intensities of the first and second CEs respectively.

(v) If the electric transition moments $\mu$ of the two chromophores constitute a right-handed screw, a positive couplet with positive first and negative second CEs are observed, and vice versa (Table 1).

(vi) The sign and intensity of exciton split CEs are governed by 'exciton chirality', which is theoretically defined as

$$\boldsymbol{R}_{ij}\cdot(\mu_{ioa}\times\mu_{joa})V_{ij}$$

where $R_{ij}$ is interchromophoric distance vector from *i* to *j*, $\mu_{ioa}$ and $\mu_{joa}$ are electric transition dipole moments of excitation $o \rightarrow a$ of groups *i* and *j*, respectively, and $V_{ij}$ is interaction energy between two groups *i* and *j*.

(vii) The absolute configuration of two chromophores can be determined, provided that the direction of the transition moment in the chromophore is established.

The cage compound [1] with two anthracene moieties, (6*R*,15*R*)-(+)-6,15-dihydro-6,15-ethanonaphtho[2,3-*c*]pentaphene (Figure 2B), provides a typical example of exciton-coupled CD. Of the four absorption bands of anthracene in the UV-visible region, the band at 252 nm ($\varepsilon = 204\,000$, $^1B_b$ transition) is the most intense and is polarized along the long axis of the chromophore. The UV of cage molecule [1] (Figure 2B) resembles that of anthracene (Figure 2A) with the expected bathochromic shift ($\lambda_{max}$ 267.2 nm, $\varepsilon = 268\,600$). The CD of (+)-[1] shows two intense CEs of opposite signs at the $^1B_b$ transition: a first positive CE, $\lambda_{ext} = 268.0$ nm, $\Delta\varepsilon = +931.3$, and a second negative CE, $\lambda_{ext} = 249.7$ nm, $\Delta\varepsilon = -720.8$; $A = +1652.1$ (these CEs are one of the strongest ever observed!). The positive *A* value indicates that the two $^1B_b$ transition moments along the long axes of the anthracene moieties constitute a clockwise screwsense. This leads to the 6*R*, 15*R* absolute configuration shown for (+)-[1].

Both exciton-coupled CD and X-ray techniques are nonempirical methods for determining the absolute configuration of chiral compounds. Although these two methods differ in that CD is for solutions whereas X-ray is for crystals, both methods naturally give the same absolute configuration.

The application of CD exciton coupling formalism to structural analysis started as the 'dibenzoate chirality rule', which became the 'exciton chirality method' when it was expanded to include the coupling between a variety of chromophores for determining the absolute configurations and/or confirmations. As shown schematically in Figure 1, 'exciton coupling' is also present in UV-visible spectra, except that the electronic spectrum simply appears as a single maximum with double intensity; only in rare cases when the energy gap $2V_{ij}$ between the $\alpha$ and $\beta$ levels is substantial will the spectra exhibit double maxima. Figure 3A (plotted in $cm^{-1}$) shows the broad UV-visible band with $\lambda_{max} \approx 390$ nm of Schiff base [1]; upon protonation this is converted into the cyanine dye [2] with delocalized electrons which give rise to a red-shifted band with narrow half-band width ($\lambda_{max} \approx 515$ nm). The coupling of two such cyanine dye chromophores

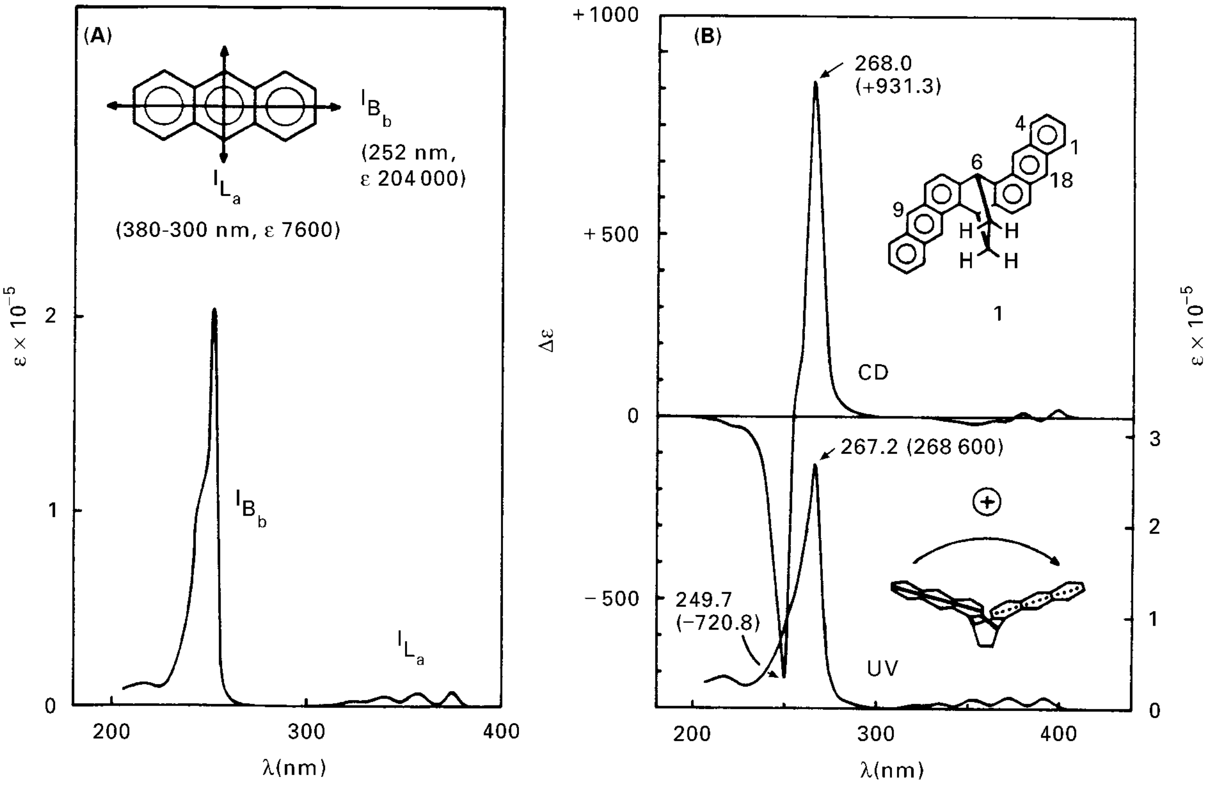



**Figure 2** (A) UV spectrum of anthracene in EtOH and polarization of the UV transitions. (B) CD and UV spectra of (6*R*,15*R*) -(+)-6,15-dihydro-6,15-ethanonaphtho[2,3-*c*]pentaphene [1] in 0.18% dioxane–EtOH. Adapted from Harada N *et al. Journal of the American Chemical Society* **98**: 5408.

give rise to an electronic spectrum with a rare double maximum. The CD spectrum exhibits an intense negative couplet with A = 462. That the integrated areas or rotational strength of the negative and positive CEs are comparable is a manifestation of the 'sum rule', which states that summation of the positive and negative rotational strengths over the entire CD spectrum should be zero. The 'sum rule' is one of the requirements of the exciton coupling mechanism. The observed regional sum rule, instead of a sum rule covering the entire CD range, is due to the well-separated cyanine dye maxima. In order to emphasize the similarity between electronic and chiroptical spectroscopy, the term 'exciton-coupled CD' rather than 'exciton chirality CD' is used in the following.

## Chromophores for exciton-coupled CD

Some aspects of practical importance are summarized as follows:

(1) Sample amount: μg scale.
(2) Solvent: acetonitrile, methylcyclohexane. Alcoholic solvents should be used with caution when dealing with acylated substrates because of the possibility of ester exchange.
(3) Any chromophore with strong intensity and known direction of μ can be used.
(4) The sign of the couplet or 'split CD' is determined by the absolute sense of twist of μ.
(5) Amplitude A is proportional to $\varepsilon^2$, or to a first approximation A and ε have a linear relation.
(6) Amplitude A is inversely proportional to the square of R, the interchromophoric distance.
(7) Additivity relation holds for A values in case of identical chromophores, and for the entire CD curves when the interacting chromophores are different (see below).
(8) There is no exciton coupling when the angles of interacting μs are 0° or 180°; in 1,2-bisacylates, the A value is maximum when the projection angle of interacting μ is ~70°.
(9) The interacting chromophores do not have to be identical. Moreover, they do no have to be in the same molecule, e.g. a ligand and its receptor could also interact and give exciton-coupled CD. Chromophores are introduced into chiral molecules lacking suitable chromophores. When only one chromophore is present in the

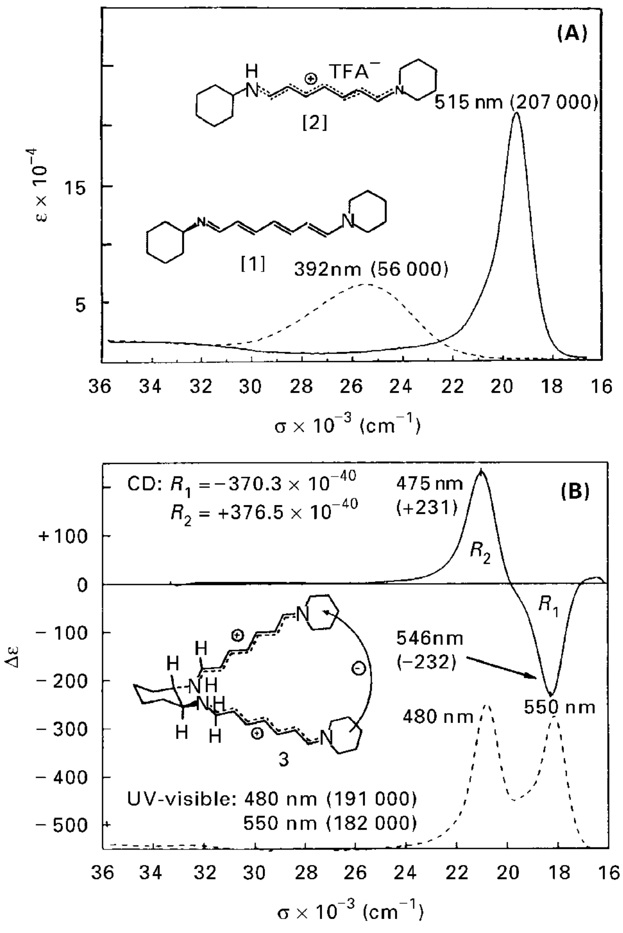



**Figure 3** (A) UV-visible spectra of mono-Schiff base [1] and monocynanine dye [2] in $CH_2Cl_2$; (B) UV-visible and CD spectra and rotational strength *R* in cgs units of biscyanine dye derivative of (1*S*, 2*S*)-*trans*-1,2-cyclohexanediamine [3] in $CH_2Cl_2$–TFA.

substrate molecule, the newly introduced chromophore is selected so that its absorption maximum will be close to that of the preexisting chromophore.

The chromophores used for determining absolute configuration or conformation by exciton-coupled CD should have intense $\pi \rightarrow \pi^*$ transitions with a known direction of the coupled transition moments; thus chromophores with high symmetry are more suited. Some chromophores referred to in this section are summarized in Table 2.

The UV spectra of para-substituted benzoates and the direction of the transition moments are shown in Figure 1. Among the para-substituted benzoates, p-dimethylaminobenzoate (Table 2, [3]), is highly recommended because of the red-shifted maximum and large ε value. Ortho- and meta-substituted benzoates are not suited because their transition moments are not parallel to the alcoholic C-O bond; consequently, the exciton chirality will depend on the conformation around the C-O bond. para-Substituted benzoyl chromophores can be used for derivatizing primary amino groups, e.g. the intramolecular CT band of the benzamido group, $\lambda_{max} \approx 225$ nm ($\varepsilon = 11\,200$), is again polarized along the long axis of the chromophore; however, for sec-amino groups the benzamido chromophores cannot be used unless the conformation around the C-N bond is established. The chromophore of 2,3-naphthalenedicarboximide [12] exhibits an intense $^1B_b$ transition around 260 nm, which is polarized along the long axis of the chromophore. This makes it ideally suited for exciton coupled CD because the long axis-polarized $^1B_b$ transition moment is exactly parallel to the C-N bond of the amine moiety; the intense ε and its fluorescence are further attributes (see below). Since the maxima of benzamido-2,3-naphthalenedicarboximido [12] and benzoate chromophores [1]-[4] are close to one another, combination of these chromophores is useful for determining the absolute configuration of amino alcohols.

### Dienes, enones, ene-esters, ene-lactones, diene-esters, other benzenoids

The transition moments of the π-π* band or conjugated dienes, enones, etc., are almost parallel to the long axis of chromophores (Figure 4). Phenylacetylene and benzonitrile chromophores with their intramolecular CT bands around 230 nm are also ideally suited because their CT transitions are polarized exactly parallel to the long axis of chromophores.

### Polyacenes

As depicted in Figure 2A, the $^1B_b$ transition of polyacene chromophores is ideally suitable for observing exciton coupled CD. See below for a description of the fluorescent 2-anthroate chromophore.

### Examples of exciton-coupled CD and natural products

The CD of the plant hormone abscisic acid exhibits a positive couplet arising from the enone-dienoic acid coupling (directions of electric transition moments are shown by thick lines). The positive twist between these two moieties shown in Figure 4 directly indicates that the absolute configuration of the tert-OH is pointing down, or α. Similarly, the negative couplet of dendryphiellin F shows the dienoate side-chain to be β whereas that of manumycin shows the triene amide chain to be α. In the case of lithospermic

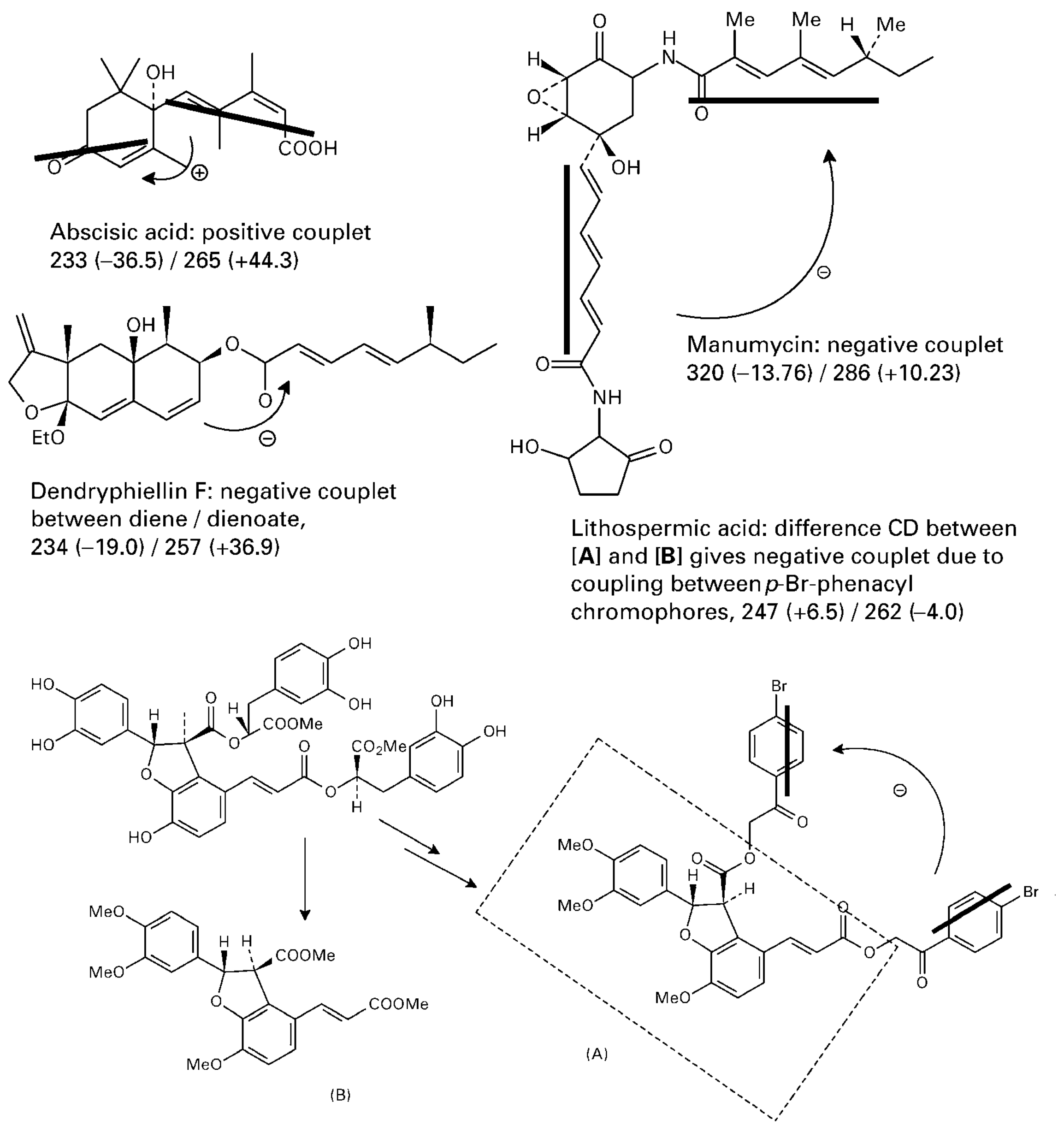



**Figure 4** Examples of natural products exhibiting an exciton-coupled CD.

acid, hydrolysis of its heptamethyl ether followed by p-bromophenacylation yielded derivative A, while methanolysis of the same heptamethyl ether gave fragment B. Subtraction of the CD of B from that of A gave a negative couplet; moiety B within the structure of A is indicated by the dotted rectangle. This difference CD shows that the coupling between the two bromophenacyl groups is negative. This defines the helicity between the two ester side-chains in lithospermic acid and hence its absolute configuration. As in any other spectroscopy, difference CD offers a powerful tool for extracting the necessary information hidden in a far more complex structure or a ligand- receptor complex.

### Chromophores with intense absorptions

The A values of split CD curves are proportional to the absorption intensity ε of the chromophore. Hence the stronger the absorption the larger is the A value or the more sensitive is the measurement. The 2-anthorate chromophore (Table 2, [8]) with an exceptionally intense absorption is a powerful chromophore for exciton-coupled CD measurements. The 2-anthroate and 2-naphthoate [6] chromophores are of lower symmetry than para-substituted benzoates [1]- [4]; however, in both cases the orientation of the strong electric transition moment, $^1B_b$, is almost parallel to the long axis of the chromophore and

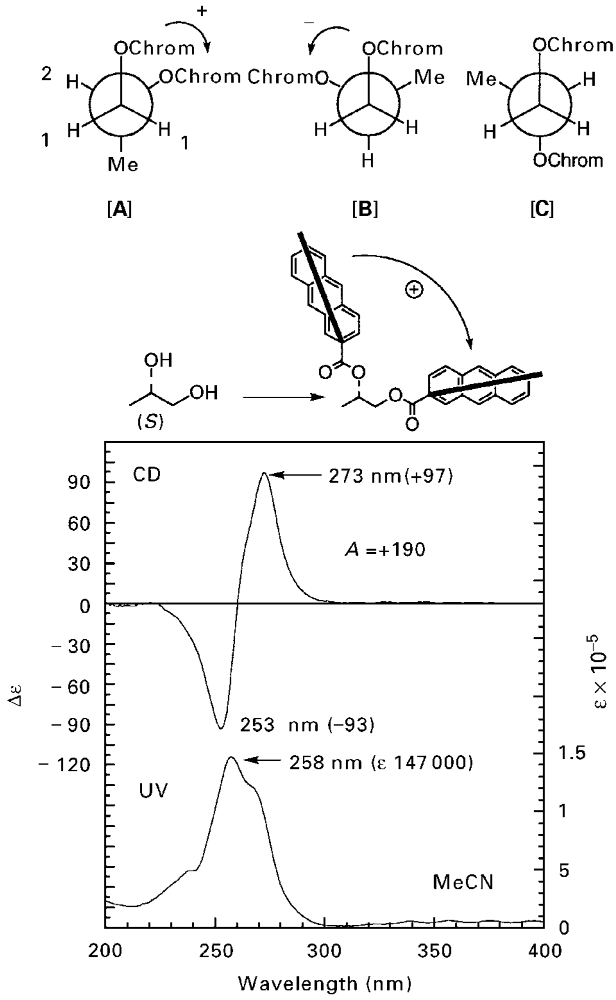



**Figure 5** Conformations and UV and CD spectra (in MeCN) of bis-2-anthroyl-(2*S*)-1,2-propanediol

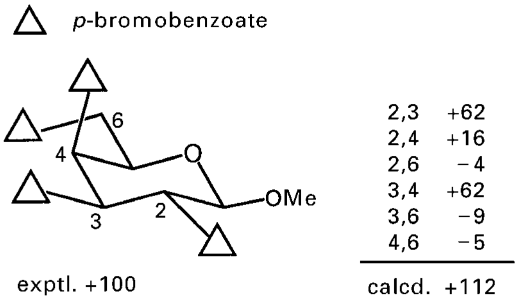



**Figure 6** Amplitudes and methyl 2,3,4,6-tetrakis[*O*-(*p*-bromobenzoyl)]-β-D-galactopyranoside, experimental (in MeOH) and calculated.

therefore almost parallel to the C-O bond (Figure 5). Moreover, its strong fluorescence not only facilitates microscale manipulations but also should make it attractive for application for fluorescence CD (FDCD). The terminal acyclic 1,2-diol is a common moiety in many natural products and synthetic intermediates and hence assignment of the absolute configuration of the chiral centre is important. Derivatization of (2S)-1,2-propanediol with anthroylimidazole furnishes the bisanthroate with a positive CD bisignate curve (MeCN) at 263 nm with A = +190. Three distinct conformations A, B and C are conceivable for the bisanthroate. A and B lead to opposite CD exciton couplets, whereas exciton coupling in conformer C should be almost zero and therefore this conformer can only influence the amplitude and not the sign of the bisignate curve. Of the remaining two, B is not expected to contribute much because of the extra steric interactions between the methyl and the chromophores. That conformer A is the major form is supported by NMR coupling constants of $J_{1,2}$ = 6.8 Hz (trans) and 3.7 Hz (gauche). The very intense UV is reflected in the large A value, which is almost 10-fold that of bis-p-bromobenzoate [2], i.e. +21 (in EtOH).

## Pairwise additivity relation in exciton-coupled CD

An aspect of fundamental importance in exciton-coupled CD is the principle of pairwise additivity. Although Brewster and Kauzmann proposed the 'principle of pairwise interactions as a basis for an empirical theory of optical rotatory power' in 1961, it was not possible to secure experimental data to prove this since they were addressed to rotations at 589 nm (Na D-line), where values are not only extremely small but also represent summation of contributions of all stereogenic centres. However, subsequent experimental results have established that the additivity relation holds for both monochromatic systems and bichromophoric systems.

Pairwise additivity in systems containing identical chromophores The amplitude A of the split CD curve for a compound containing two or more identical chromophores can be approximated by the summation of each interacting basis pair. Thus, in a trischromophoric system containing three identical chromophores I, II and III, the observed amplitude can be approximated by the sum of three interacting bischromophoric pairs (basis pairs): $A_{total} = A_{I,II} + A_{II,III} + A_{I,III}$. This is illustrated for a sugar acylate containing four p-bromobenzoates in Figure 6. Six interacting pairs of benzoates are present in this derivative, 2,3-, 2,4-, 2,6-, etc. The A values of the six pairs of p-bromobenzoates are shown in the table in Figure 6; the summation gives a value of +112, which is in excellent agreement with the experimental value of +100. The A values (i.e. constants) of all 12 conceivable pairs for hexopyranosides have been measured and tested against 42 cases of bis-, tris- and

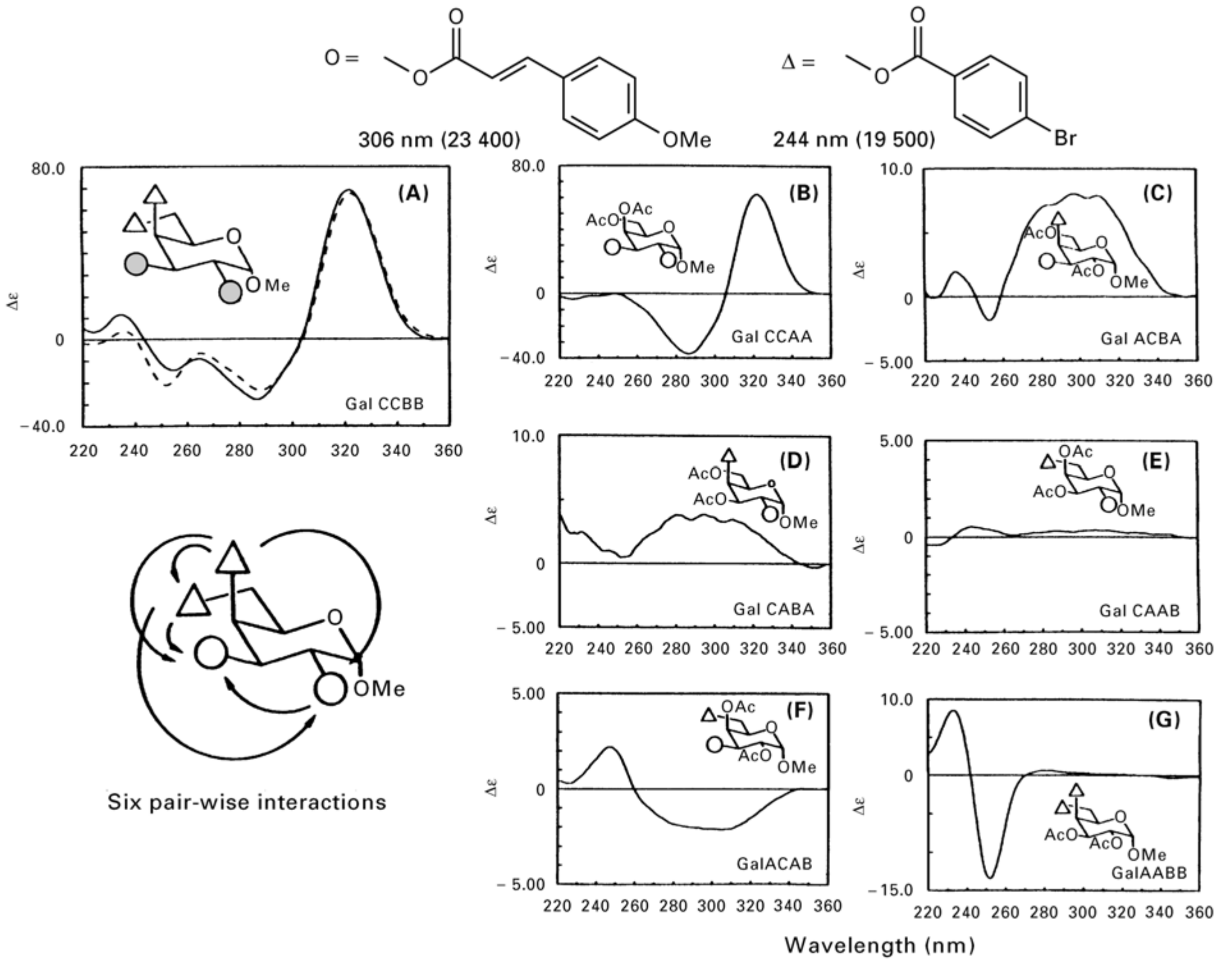



**Figure 7** (A) Experimental (solid line) and calculated sum CD spectra (dashed line) of α-methyl galactyopyranoside 2,3-bis-*p*-methoxycinnamate 4,6-bis-*p*-bormobenzoate (GalCCBB). (B)–(G) six experimental pairwise exciton-coupled CDs of four interacting chromophores in GaLCCBB.

tetrakis-acylates and in all cases excellent agreement is seen. This additivity relation is also corroborated by theoretical calculations. The A values for O-acylates and N-acylates are identical, and hence they are applicable to all hexopyranosides (but not to N-acetyl sugars since the confirmation of the acylated N-acetyl bond is not known). The additivity principle also leads to a simple microscale method for determining the branching point in oligosaccharides, namely the sugars are permethylated, hydrolysed and the liberated hydroxyls which are tags of the linkage point are bromobenzoylated; each sugar benzoate is then identified from the A value (microgram level).

The pairwise additivity relation also holds for p-phenylbenzylates [11], thus showing that the conformation of the benzylate closely resembles that of the benzoate; the benzylates can readily be oxidized to the corresponding p-phenylbenzoates [5], which give rise to another set of data. The advantage of benzylates is that they can derivatize hindered OH groups, including tertiary.

Pairwise additivity in systems containing two different chromophores  Two different chromophores can still exhibit exciton coupling when the absorption maxima differ by ~100 nm. Furthermore, the pairwise additivity relation has been shown to hold for systems containing two different chromophores, as exemplified in Figure 7 for the α-methyl galactopyranoside Gal CCBB substituted with p-bromobenzoate ($\lambda_{max}$ = 245 nm) at C-2,3 and p-methoxycinnamate ($\lambda_{max}$ = 311 nm) at C-4,6. The four-symbol descriptor designates the substituents at positions 2, 3, 4 and 6, their sequence corresponding to the above order of locants: A = O-acetyl, C = O-(4-methoxycinnamoyl) and B = O(4-bromobenzoyl). The CD of the six galactosides containing the interacting basis sets are also shown in Figure 7. As shown in Figure 7A, there is excellent overlap between the experimental and calculated curves. This agreement holds for all permutations of the galactose, glucose, mannose series, desoxy sugars and amino sugars with primary amino groups, a

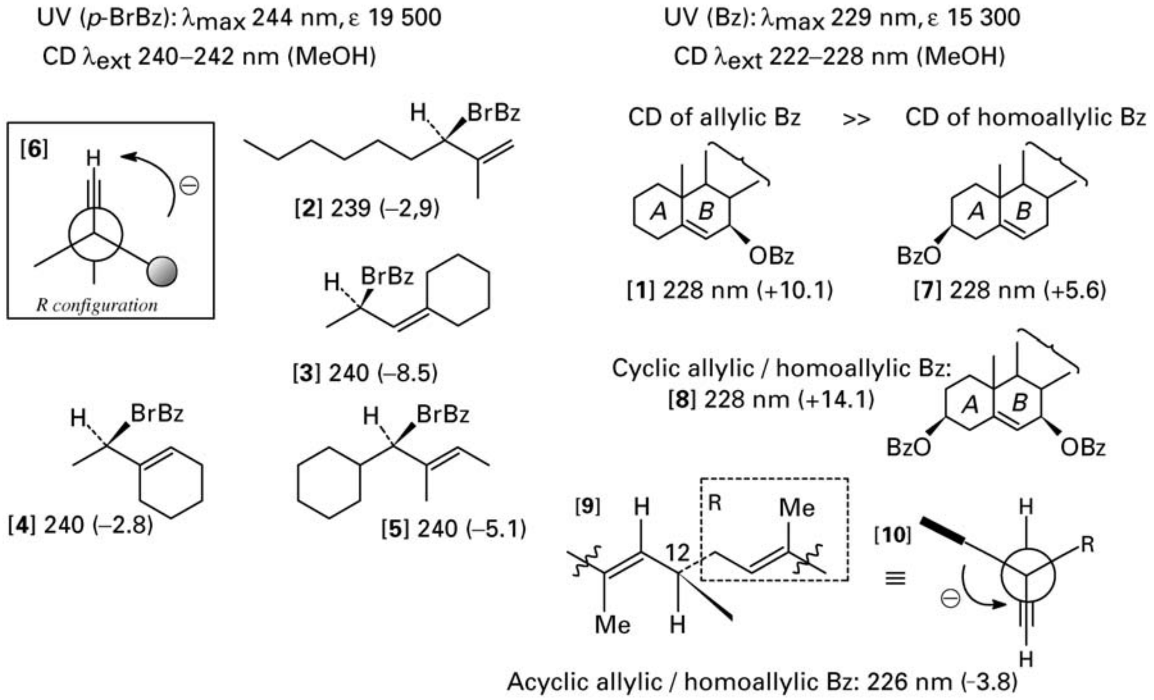



**Figure 8** Absolute configurations of allylic and homoallylic alcohols, cyclic and acyclic.

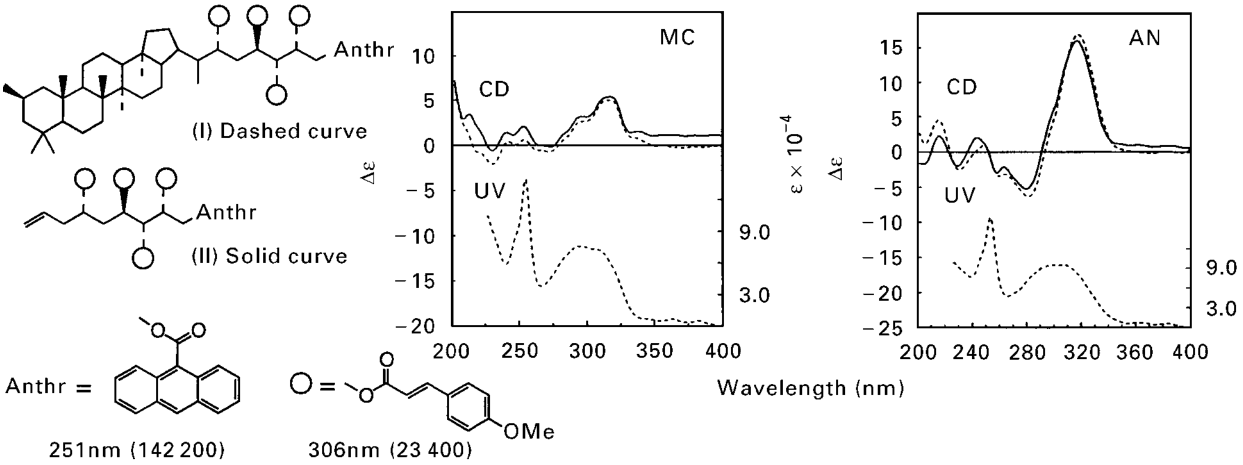



**Figure 9** Reference CD spectra for an acyclic 1,3,-polyol and a bacteriohopanepentol in methylcyclohexane (MC) and acetonitrile (AN).

total of 150 cases. These 150 characteristic curves could provide valuable data for the development of an alternative means to conventional methylation analysis of oligosaccharides by tagging free hydroxyls as bromobenzoates or phenylbenzylates, and then derivatizing the hydroxyl groups liberated upon methanolysis. This approach requires no reference sample and, moreover, would also establish the absolute stereochemistry of the individual sugar moieties. As the additivity relation holds for other coupled systems involving various multichromophoric systems, including cage structures, its proper application greatly expands the utility of exciton-coupled CD.

### Allylic and homoallylic alcohols, cyclic and acyclic

As chromophores with a difference in $\lambda_{max}$ of ~100 nm can still couple, this means that the observation of only one branch of the couplet suffices to determine the absolute configuration. Such a case is the allylic benzoates, where the sign of the first CE ascribable to the benzoate, $\lambda_{max} = 229$ nm [1] represents the chirality between the benzoate

chromophore and the double bond; it is not necessary to observe the CE arising from the 195 nm double bond absorption. Thus the positive CE of steroid [1] shown in Figure 8 reflects the positive twist between the 7β-benzoate and the 5,6-ene. The CD of a series of acyclic allylic alcohols with established configurations, [2]- [5], show that the relation can be extended to acyclic cases. In this case, the signs of CEs are rationalized by the most stable conformer of the allylic acylates in which the allylic hydrogen adopts a syn relation with the double bond [6] so that the sign of the 'couplet' is determined by the twist between the benzoate and the double bond, as in cyclic allylic benzoates. Unsubstituted benzoates are preferred in these studies since the maxima are closest to the ene maxima.

The CE of homoallylic benzoate [7] is also in accord with the chirality, although its intensity is lower than that of allylic benzoate [1]. The additivity again holds, as seen for steroid derivative [8]. Application of this additivity between allylic and homoallylic benzoates was used to determine the configuration of moiety [9] in an acyclic natural product.

### Acyclic 1,2-, 1,3- and mixed 1,2-/1,3-acyclic polyols and aminopolyols

Acyclic 1,2-, 1,3- and mixed 1,2-/1,3-polyols and aminopolyols are frequently present in natural products, including polyene macrolides. Because of the lack of general methods for determining their relative and absolute configurations, the stereostructures of most of them remain undetermined. However, a library of such polyols up to pentaols is now available for all permutations of stereorelations; in this library, the primary hydroxyl or amino groups are derivatized by the anthroyl chromophore [10] while the other hydroxyl or amino groups are derivatized by [6]; as in other cases, the CD properties of O and N derivatives are very similar so that the same set of data can be used. Figure 9 shows that the CD of the mixed polyol moiety present in the hopanoid [I] is superimposable on that of the reference mixed polyol derivative [II] in two solvents, the nonpolar methylcyclohexane (MC) and more polar acetonitrile (AN); this determines the absolute configuration of the side-chain on a microgram scale.

### Exciton-coupled CD due to pre-existing chromophores

The well-known bitter principle of the wood of Quassia amara L., quassin, has two enone chromophores within its structure shown in Figure 10. Accordingly, its CD exhibits a typical couplet at 266 nm (Δε = +10.4) / 242 nm (Δε = −9.5) arising

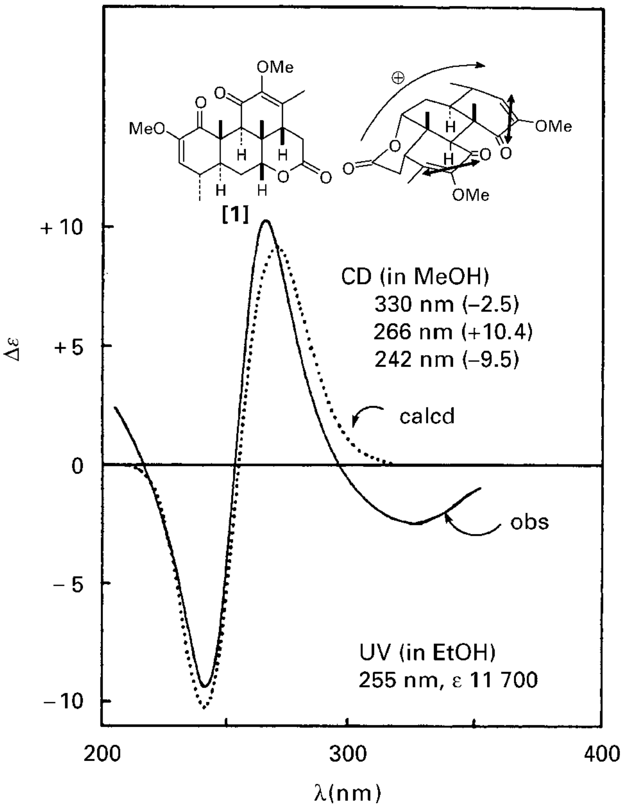



**Figure 10** Absolute configuration of quassin.

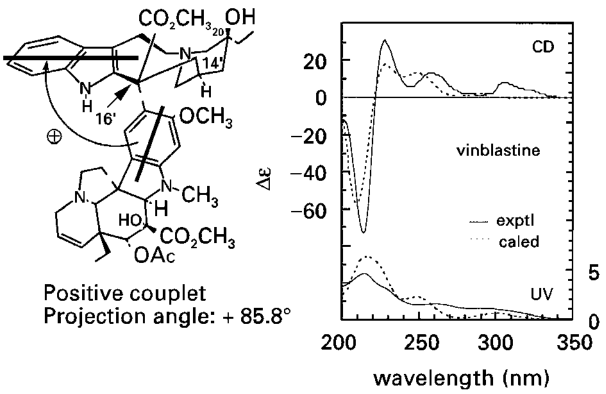



**Figure 11** Experimental (in MeCN) and calculated CD spectra of vinblastine.

from the twist between the two chromophores, thus directly establishing its absolute configuration as depicted. The excellent agreement between the CD calculated from the UV and geometry of the molecule and observed CD not only establishes the absolute configuration of quassin but also proves that the observed CD of quassin around 250 nm is due to exciton coupling. This conclusion is consistent with the absolute configurational assignment of quassin

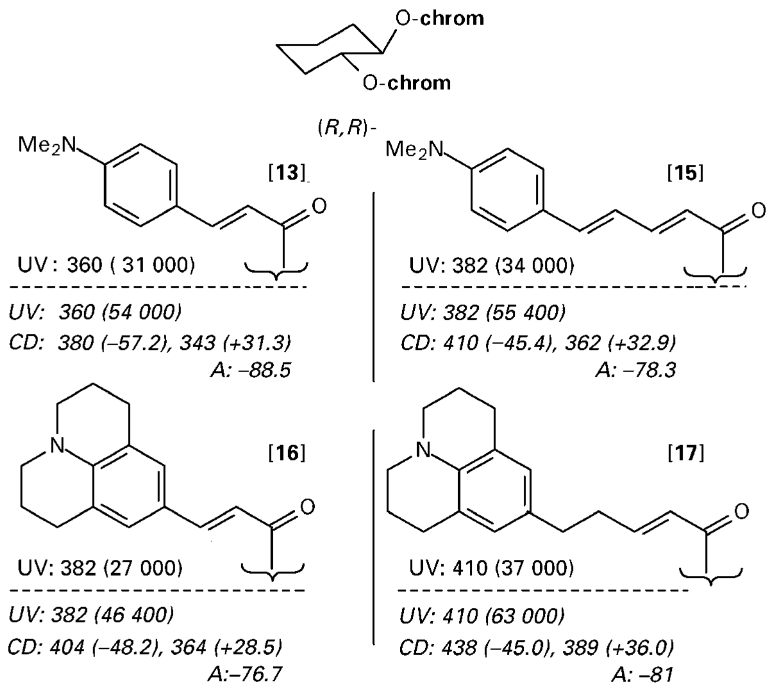



**Figure 12** UV-visible spectra of chromophores [13], [15], [16], [17] and CD spectra of diesters formed from 1(*R*), 2(*R*)-cyclohexanediol, in MeCN.

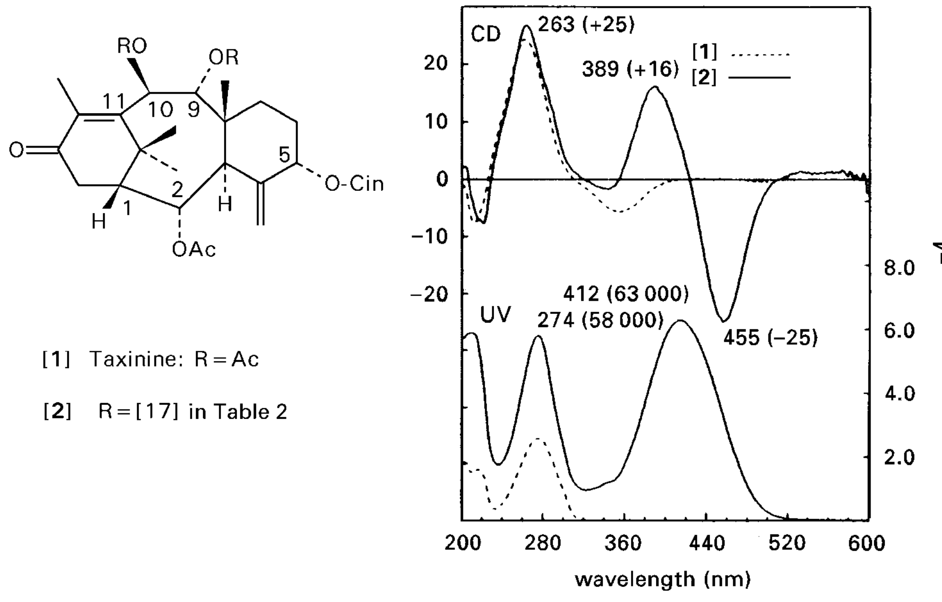
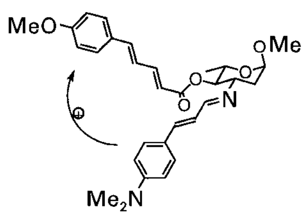



**Figure 13** UV-visible and CD spectra of taxinine [1] and its bis-ester formed with chromophore [17] in MeCN.

based on biosynthetic considerations and X-ray analysis of biogenetically related compounds. The negative CE at 330 nm is due to the $\pi \rightarrow \pi^*$ transition of the enone chromophores. It should be noted that despite the negative 330 nm CE, the optical rotation at the sodium D-line (589 nm) is positive: $[\alpha]_D + 34.5°$ (c 5.09, $CHCl_3$). Therefore, the D-line rotation, which is simply the reading at 589 nm taken in the optical rotatory dispersion (ORD) mode, is governed by the intense positive Cotton effect at 266 nm. The CD of abscisic acid described in Figure 4 is very similar to that of quassin.

The clinically important anticancer drug vinblastine also incorporates two chromophores, the indole and indoline moieties, which are coupled through space. The biologically active molecule is




Acosamine derivative

CD neutral: 370 (+62) / 330 (−48), *A* +110
CD Protonated: 469 (+35) / 350 (−25), *A* +60

**Figure 14** CD of acosamine derivatized with chromophores [14] and [20]/[21], in MeCN.

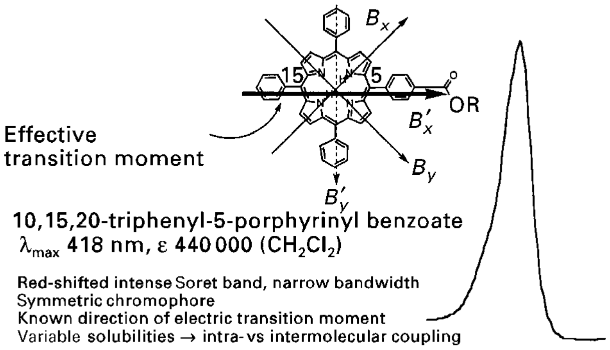



**Figure 15** Tetraarylporphyrins.

that depicted in Figure 11, while its biologically inactive diastereomer at C-16′ has a different CD. The UV spectrum has an intense band at 214 nm ($\varepsilon = 46\ 200$) with a shoulder at 228 nm which arises from the $^1B_a$ transitions of the indoline (217 nm) and indole (225 nm) chromophores; this is accompanied by the intense positive CD couplet in this region. The weaker 260 nm ($\varepsilon = 15\ 000$) band is due to the $^1L_a$ transition of the indoline moiety, while the overlapping bands at 293 nm ($\varepsilon = 10\ 500$) are from the $^1L_a$ and $^1L_b$ transitions of the indole chromophore and the $^1L_b$ transition of the indoline chromophore. Theoretical calculations of the UV and CD of the strong bands centred around 220 nm using atomic coordinates based on MM2 and SCF-CI-DV MO are in excellent agreement with the observed CD couplet, thus proving that the couplet, which is used empirically to differentiate the bioactive isomer from the inactive isomer, is indeed due to exciton coupling between the two moieties.

### Red-shifted chromophores for hydroxyl groups

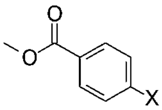
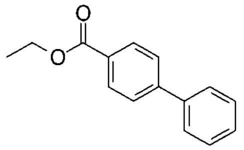
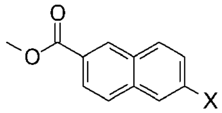
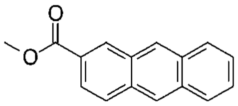
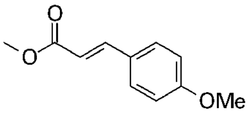
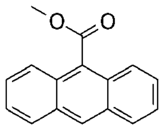
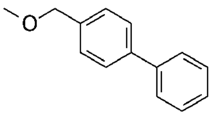
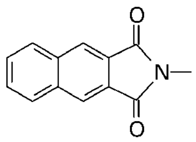
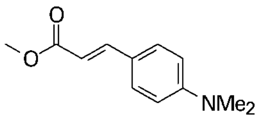
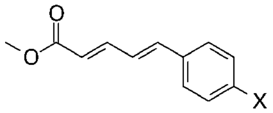
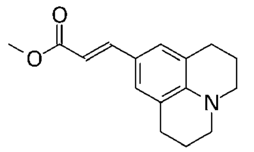
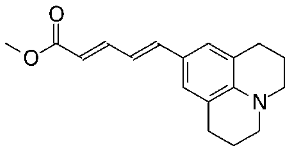
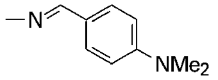
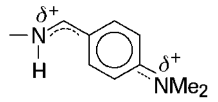
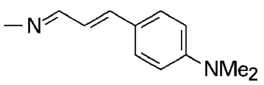
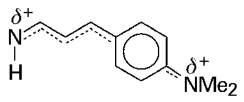
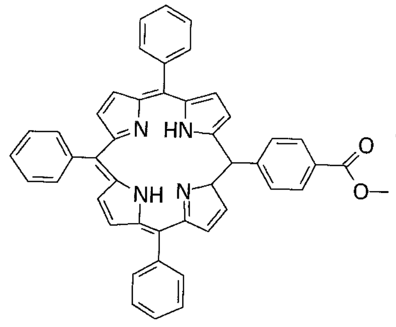
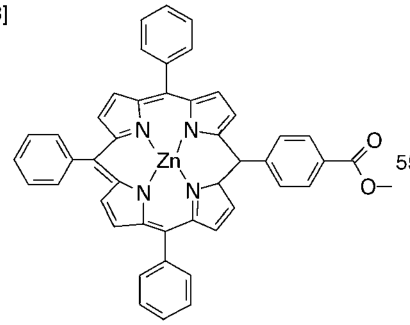
Since the natural substrate itself frequently contains a chromophore or may be contaminated with some UV-absorbing substance, the use of chromophores with absorbance shifted to the red is advantageous in order to avoid complications due to unnecessary interactions. Chromophores [13]-[17] and [18]-[21] (Table 2) with red-shifted intense maxima can be used for derivatizing hydroxyl and amino groups, respectively. Figure 12 lists the UV-visible and CD data for derivatives of 1(R),2(R)-cyclohexanediol; in all cases the A values are intense and the spectra appear in regions that usually are transparent.

Figure 13 demonstrates the advantage of using such chromophores. The yew tree constituent taxinine shares the same skeleton as taxol and taxotere, prominent anticancer compounds. The highly strained enone group of the skeleton shows a strong CE at 262 nm arising from a $\pi \rightarrow \pi^*$ transition and a weaker CE at 354 nm from the $\pi \rightarrow \pi^*$ transition. The 262 nm band overlaps with conventional chromophores. However, [17] with its maximum at 410 nm is completely removed from the enone absorptions. Hence the C-9 and C-10 bis-esters formed from this chromophore show typical exciton couplet in a region remote from enone absorption. The negative couplet agrees with the absolute configuration of the taxane skeleton deduced by other means, including exciton-coupled CD of a derivative using a more conventional chromophore, the interpretation of which was not as clearcut.

### Red-shifted chromophores for amino groups

Chromophores [18] and [19] (Table 2) are simply imines or Schiff bases formed in high yield by reacting primary amines with p-dimethylaminobenzaldehyde and p-dimethylaminocinnamaldehyde under mild conditions and in the presence of unprotected hydroxyl groups. Addition of a drop of trifluoroacetic acid to the UV or CD cell converts them into protonated Schiff bases, which exhibit intense and very sharp absorptions in the red due to their cyanine dye structure; moreover, addition of a small amount of water will hydrolyse the protonated Schiff base bond, leading to recovery of the starting amine. The derivative of acosamine is prepared by derivatizing the amino group and then the hydroxyl group with the respective chromophores [20] and [14] (Figure 14). The exciton couplet is intense with an amplitude of 110. Acidification shifts the couplet to longer wavelength because of the shift of the imine band from 360 to 460 nm; however, the intensity decreases owing to the diminished overlap between the chromophores, from 330 / 360 nm in the neutral form to 330 / 460 nm in the protonated derivative. If the coupling is between two protonated Schiff base moieties, i.e. between two cyanine dyes, then the couplet becomes well-separated and is greatly intensified (see Figure 3).

**Table 2** Absorption maxima and intensities of some chromophores used for derivatizations of hydroxyl and amino groups

| *Chromophore* | *Type* | | |
|---|---|---|---|
| | *Primary and secondary OH* | | |
| [1] |  | X = H | 229 nm<br>15 300 (EtOH) |
| [2] | X = Br | | 244 nm<br>19 500 (EtOH) |
| [3] | X = $NMe_2$ | | 307 nm<br>28 200 (MeCN) |
| [4] | X = OMe | | 257 nm<br>20 400 (EtOH) |
| [5] |  | | 270 nm<br>20 700 ($CH_3CN$) |
| [6] |  | X = H | 234 nm<br>58 000 ($CH_3CN$) |
| [7] | X = OMe | | 237 nm<br>48 000 ($CH_3CN$) |
| [8] |  | | 258 nm<br>93 000 ($CH_3CN$) |
| [9] |  | | 306 nm<br>23 400 (MeCN) |
| | *Selective for primary OH* | | |
| [10] |  | | 251 nm<br>142 200 ($CH_3CN$) |
| | *OH including tertiary and hindered OH* | | |
| [11] |  | | 253 nm<br>20 300 ($CH_3CN$) |
| | *Selective for primary $NH_2$* | | |
| [12] |  | | 258 nm<br>64 000 ($CH_3CN$) |
| | *Red-shifted for OH* | | |
| [13] |  | | 360 nm<br>31 000 (MeCN) |
| [14] |  | X = OMe | 333 nm<br>40 400 (MeOH) |
| [15] | X = $NMe_2$ | | 382 nm<br>34 000 (MeCN) |
| [16] |  | | 382 nm<br>27 000 (MeCN) |
| [17] |  | | 410 nm<br>37 000 (MeCN) |
| | *Red-shifted, selective for primary $NH_2$ (neutral and protonated)* | | |
| [18] |  | | 305 nm<br>24 300 (MeCN) |
| ↓ *TFA* | | | |
| [19] |  | | 395 nm<br>51 700 (MeCN) |
| [20] |  | | 361 nm<br>37 000 (MeCN) |
| ↓ *TFA* | | | |
| [21] |  | | 460 nm<br>64 500 (MeCN) |
| [22] | *Porphyrins for OH*<br> | | 418 nm<br>440,000 ($CH_2Cl_2$) |
| [23] |  | | 419 nm<br>550 000 ($CH_2Cl_2$) |

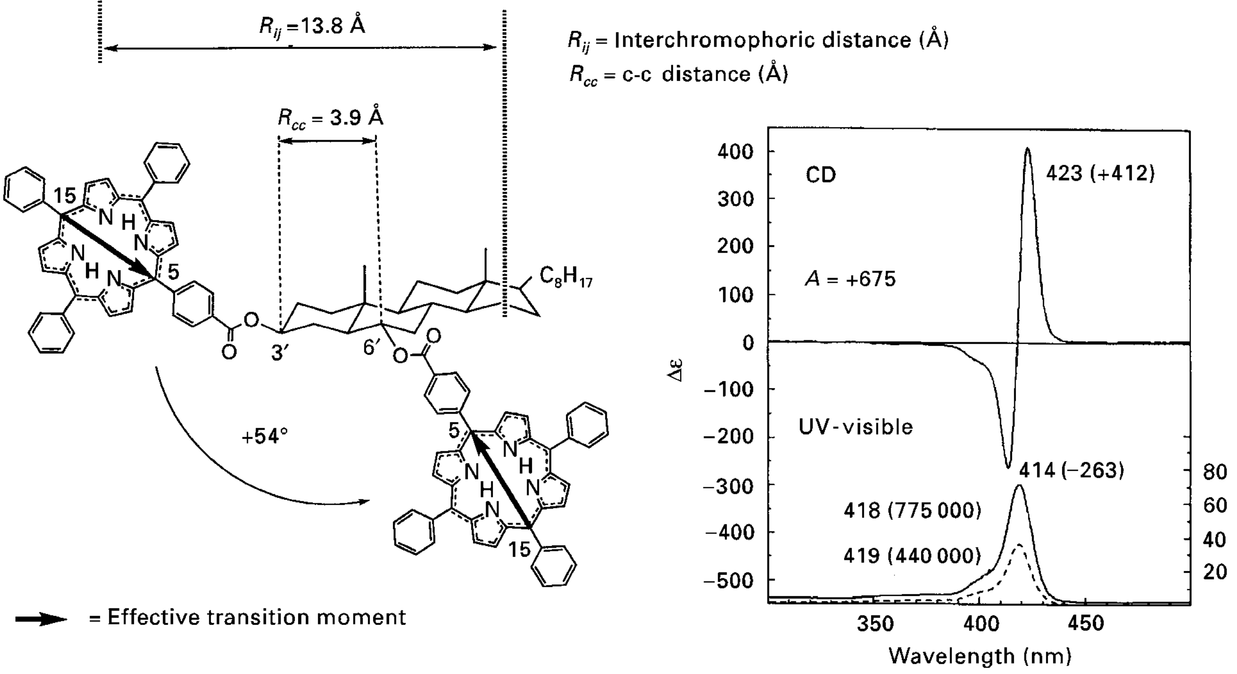



**Figure 16** Porphyrin/porphyrin exciton coupling and UV-visible and CD spectra in $CH_2Cl_2$ ($c$ = 1.0 μM) of 5α-androstane -3α,17β-diol bis[*p*-(10′,5′,20′-triphenyl-5′-porphyrinyl)benzoate] (solid line) and UV-visible spectrum of 5-(4′-carboxyphenyl)-10,15,20-triphenylporphyn in $CH_2Cl_2$ (dashed line).

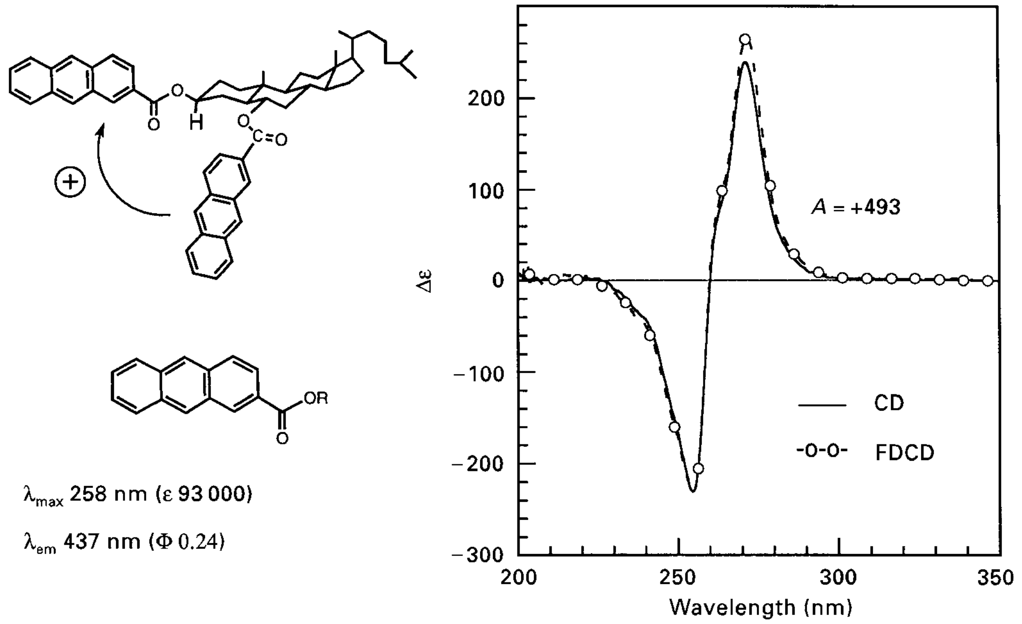



**Figure 17** Comparison of conventional CD (solid line) with the CD converted from FDCD (dashed line) for 5α-cholestane-3β,6α-diol bis(2-anthroate) in MeCN.

## Porphyrins, a very promising chromophore

The porphyrin chromophores exemplified by [22] and [23] are powerful chromophores for exciton-coupled CD because of the very intense ε of the red-shifted and sharp Soret band at 418 nm (Figure 15). The effective direction of the electric transition moment can be regarded as running parallel to the C- O bond of the derivatized alcohol as in any other para-substituted benzoates. The solubilities

can be modified over a wide range through the appended aromatic rings at C-10, -15 and -20; the triphenyl derivative depicted is hydrophobic whereas a trimethylpyridinium derivative is hydrophilic. Thus the pairing of a hydrophobic or hydrophilic porphyrin derivative with the polarity of the solvent will lead either to intramolecular coupling or to intermolecular stacking. This results in great versatility in the use of porphyrins in exciton-coupled CD. Attachment of trimethylpyridinium porphyrin to a brevetoxin B derivative, a molecule with a rigid skeletal structure, shows intramolecular exciton coupling even over a distance of ~40- 50 Å (in methanol). The very efficient intramolecular coupling arising from the intense ε is illustrated in Figure 16. The amplitude of the positive couplet of the 3,6-bis-substituted steroid is +675 at an interchromophoric distance of 13.8 Å; the amplitude of the 3,6-bis-p-dimethylaminobenzoate [3], in contrast, is +89. Zinc porphyrins (e.g. [23]) are even more promising because of their more intense absorptions; furthermore, the ability of the central metal atom to coordinate with amino groups adds further versatility to their use in exciton-coupled CD.

### Exciton-coupled fluorescence CD

This field is still fairly new. However, exciton-coupled CD detected either by absorption (conventional CD) or by emission both give the same CD curve, as shown in Figure 17. The greatest advantage of fluorescence-based CD exciton coupling is that the sensitivity is increased 50- 100-fold under favourable conditions involving two identical fluorophores, and that it selectively records the coupling between fluorophores and is not perturbed by nonfluorescent chromophores. However, this new field requires further studies, from the viewpoint of both theory and application.

## Quantum mechanical theory of exciton coupled CD

### CD spectra and rotational strength of Cotton effect

The rotational strength R, a theoretical parameter representing the sign and intensity of a CD Cotton effect, is expressed in Equation [1] and is obtained from the observed CD spectra.

$$R = 2.296 \times 10^{-39} \int \Delta\varepsilon(\sigma)/\sigma \mathrm{d}\sigma \quad \text{(cgs units)} \qquad [1]$$

where σ is wavenumber.

The rotational strength R is also formulated by the Rosenfeld equation:

$$R = \mathrm{Im}\{\langle o|\mu|a\rangle \cdot \langle a|\boldsymbol{M}|o\rangle\} \qquad [2]$$

where Im denotes the imaginary part of the terms in brackets, ⟨ ⟩ denotes the integration over configuration space and μ and M are operators of electric and magnetic moment vectors, respectively. The dot · stands for scalar product of two vectors and o and a are wavefunctions of ground and excited states, respectively. Thus the rotational strength R is equal to the imaginary part of the scalar product of electric and magnetic transition moments.

Provided a CD Cotton effect is approximated by the Gaussian distribution, Equation [3] is obtained:

$$\Delta\varepsilon(\sigma) = \Delta\varepsilon_{\max} \exp\{-[(\sigma - \sigma_o)/\Delta\sigma]^2\} \qquad [3]$$

where $\Delta\varepsilon_{max}$ is the maximum intensity of the Cotton effect, $\sigma_o$ is the central wavenumber of the Cotton effect, and Δσ is the standard deviation of the Gaussian distribution. From Equations [1] and [3]:

$$R = 2.296 \times 10^{-39} \sqrt{\pi} \Delta\varepsilon_{\max} \Delta\sigma/\sigma_o \qquad [4]$$

From Equations [3] and [4], the calculated CD curve is formulated as

$$\Delta\varepsilon(\sigma) = [\sigma_o/(2.296 \times 10^{-39} \sqrt{\pi} \Delta\sigma)] \times R \exp\{-[(\sigma - \sigma_o)/\Delta\sigma]^2\} \qquad [5]$$

where Δσ can be evaluated from observed UV-visible spectra. Provided the rotational strength R is calculated by Equation [2], the CD spectrum can be reproduced by theoretical calculation.

### UV-visible spectra and dipole strength

The dipole strength D representing the intensity of UV-visible bands is estimated from the observed spectra:

$$D = 9.184 \times 10^{-39} \int \varepsilon(\sigma)/\sigma \, \mathrm{d}\sigma \qquad [6]$$

The dipole strength is formulated as

$$D = \{\langle o|\mu|a\rangle\}^2 \qquad [7]$$

Hence, the dipole strength D is equal to the square of the electric transition moment $\mu$.

### Molecular exciton theory of a binary system with two chromophores.

In the exciton coupling system composed of two identical chromophores i and j, exciton wavefunctions are expressed as Equations [8] and [9], where each chromophore undergoes excitation o $\rightarrow$ a

$$\text{ground state}: \quad \phi_{io}, \quad \phi_{jo} \qquad [8]$$

$$\text{excited state}: \quad \phi_{ia}, \quad \phi_{ja} \qquad [9]$$

The Hamiltonian operator of the coupling system is formulated as:

$$H = H_i + J_j + H_{ij} \qquad [10]$$

where $H_i$ and $H_j$ are the Hamiltonian of groups i and j, respectively, and $H_{ij}$ is the interaction energy term between two groups i and j.

### Wavefunction and energy of ground state

The ground-state wavefunction and energy of a binary system are

$$\Psi_o = \phi_{io}\ \phi_{jo} \qquad [11]$$

$$E_o = 0 \qquad [12]$$

### Wavefunction and energy of singly excited state

The singly excited state of the binary system splits into two energy levels, $\alpha$ and $\beta$ states. For the $\alpha$-state,

$$E^{\alpha} = E_a - V_{ij}, \Psi_a{}^{\alpha} = (1/\sqrt{2})\{\ \phi_{ia}\ \phi_{jo} -\ \phi_{io}\ \phi_{ja}\} \qquad [13]$$

where $V_{ij}$ is the interaction energy between two groups i and j and is approximated by a point dipole approximation method:

$$V_{ij} = \mu_{ioa}\mu_{joa}R_{ij}^{-3}\{\boldsymbol{e}_i \cdot \boldsymbol{e}_j - 3(\boldsymbol{e}_i \cdot \boldsymbol{e}_{ij})(\boldsymbol{e}_j \cdot \boldsymbol{e}_{ij}) \qquad [14]$$

where $\mu_{ioa}$, $\mu_{joa}$ and $R_{ij}$ are absolute values of vectors $\mu_{ioa}$, $\mu_{joa}$ and $R_{ij}$, respectively; $e_i$, $e_j$ and $e_{ij}$ are unit vectors of $\mu_{ioa}$, $\mu_{joa}$ and $R_{ij}$, respectively. For $\mu_{ioa}$, $\mu_{joa}$ and $R_{ij}$, see Equations [17] and [29].

For the $\beta$-state,

$$E^{\beta} = E_a + V_{ij}, \Psi_a\beta = (1/\sqrt{2})\{\ \phi_{ia}\ \ \phi_{jo} +\ \phi_{io}\ \ \phi_{ja}\} \qquad [15]$$

These equations indicate that the binary system has two electronic transitions o$\rightarrow\alpha$ and o$\rightarrow\beta$, in the UV-visible spectrum. If $V_{ij} > 0$, the $\alpha$-state is lower in energy than the $\beta$-state, and therefore the transition o$\rightarrow\alpha$ locates at longer wavelengths and the transition o$\rightarrow\beta$ at shorter wavelengths.

### Electric transition moment of a binary system

The electric dipole moment operator $\mu$ of a whole system is defined as

$$\mu = \Sigma\mu_i = \Sigma\,\Sigma e\boldsymbol{r}_{is} \qquad [16]$$

where $\mu_i$ is the electric dipole moment operator of group i, e is the elementary charge and $r_{is}$ is the distance vector of electron s in group i from the origin. The electric transition moment $\langle a|\mu|o\rangle\alpha$ of the transition o$\rightarrow\alpha$ is expressed as

$$\langle o|\mu|a\rangle^{\alpha} = \mu_{oa}{}^{\alpha} = \int \Psi_o\mu\Psi_a{}^{\alpha}d\tau$$
$$= (1/\sqrt{2})(\mu_{ioa} - \mu_{joa}) \qquad [17]$$

where $\mu_{ioa} = \int \phi_{io}\mu_i\phi_{ia}d\tau_i$ and $\mu_{joa} = \int \phi_{jo}\mu_j\phi_{ja}d\tau_j$. These are electric transition moments of the transition o$\rightarrow$a in groups i and j, respectively. The electric transition moment of the transition o$\rightarrow\beta$ is similarly expressed as

$$\langle o|\mu|a\rangle^{\beta} = \mu_{oa}{}^{\beta} = \int \Psi_o\mu\Psi_a\beta d\tau$$
$$= (1/\sqrt{2})(\mu_{ioa} + \mu_{joa}) \qquad [18]$$

### Magnetic transition moment of a binary system

The magnetic moment operator M of a whole system is formulated as

$$\boldsymbol{M} = \Sigma\boldsymbol{M}_i = (e/2mc)\Sigma\ \Sigma\boldsymbol{r}_{is} \times \boldsymbol{p}_{is} \qquad [19]$$

where m is the mass of an electron, c is the velocity of light, $p_{is}$ is the linear momentum of electron s in group i, and × represents the vector product of two vectors.

The magnetic moment operator is further changed as

$$\boldsymbol{M} = (e/2mc)\Sigma\, \boldsymbol{R}_i \times \boldsymbol{p}_i + \Sigma \boldsymbol{m}_i \quad [20]$$

where $R_i$ is distance vector of group i from the origin and $p_i$ and $m_i$ are linear momentum and internal magnetic moment operators of group i, respectively. The magnetic transition moment $\langle a|M|o\rangle^\alpha$ of the excitation o→α is calculated as

$$\begin{aligned}\langle a|\boldsymbol{M}|o\rangle^\alpha &= \int \Psi_a{}^\alpha \boldsymbol{M}\Psi_o \mathrm{d}\tau \\ &= (1/\sqrt{2})\{(e/2mc)\boldsymbol{R}_i \times \boldsymbol{p}_{iao} + \boldsymbol{m}_{iao} \\ &\quad - (e/2mc)\boldsymbol{R}_j \times \boldsymbol{p}_{jao} - \boldsymbol{m}_{jao}\} \quad [21]\end{aligned}$$

where $p_{iao}$ and $m_{iao}$ are linear momentum and internal magnetic moment of group i, respectively.

For the linear momentum of a group, the next equation is very useful:

$$\boldsymbol{p}_{oa} = -(2\pi i mc/e)\sigma_o \mu_{oa} \quad [22]$$

where i is the symbol for 'imaginary', $\sigma_o$ is excitation energy expressed in wavenumber units and $\mu_{oa}$ is electric transition moment of transition o→a. Applications of this equation to group i:

$$\boldsymbol{p}_{ioa} = (2\pi i mc/e)\sigma_o \mu_{ioa} \quad [23]$$

Accordingly,

$$\begin{aligned}\langle a|\boldsymbol{M}|o\rangle^\alpha &= (1/\sqrt{2})\{i\pi\sigma_o \boldsymbol{R}_i \times \mu_{ioa} - i\pi\sigma_o \boldsymbol{R}_j \\ &\quad \times \mu_{joa} + \boldsymbol{m}_{iao} - \boldsymbol{m}_{jao}\} \quad [24]\end{aligned}$$

In a similar way, the magnetic transition moment of excitation o→β is calculated as

$$\begin{aligned}\langle a|\boldsymbol{M}|o\rangle^\beta &= (1/\sqrt{2})\{i\pi\sigma_o \boldsymbol{R}_i \times \mu_{ioa} - i\pi\sigma_o \boldsymbol{R}_j \\ &\quad \times \mu_{joa} + \boldsymbol{m}_{iao} - \boldsymbol{m}_{jao}\} \quad [25]\end{aligned}$$

### Dipole strength a binary system

The dipole strength $D^\alpha$ for the excitation o→α is expressed as

$$D^\alpha = (1/2)(\mu_{ioa} - \mu_{joa})^2 \quad [26]$$

Similarly,

$$D^\beta = (1/2)(\mu_{ioa} + \mu_{joa})^2 \quad [27]$$

### Rotational strength a binary system

The rotational strength $R^\alpha$ of the α–state is calculated from Equations [2], [17] and [24]:

$$\begin{aligned}R^\alpha &= \mathrm{Im}\{\langle o|\mu|a\rangle^\alpha \cdot \langle a|\boldsymbol{M}|o\rangle^\alpha\} \\ &= (1/2)\mathrm{Im}\{(\mu_{ioa} - \mu_{joa}) \cdot (\boldsymbol{m}_{iao} - \boldsymbol{m}_{jao})\} \\ &\quad + (1/2)\pi\sigma_o \boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa}) \quad [28]\end{aligned}$$

where $R_{ij}$ is the interchromophoric distance vector from group i to group j, and is defined as

$$\boldsymbol{R}_{ij} = \boldsymbol{R}_j - \boldsymbol{R}_i \quad [29]$$

Similarly,

$$\begin{aligned}R^\beta &= \mathrm{Im}\{\langle o|\mu|a\rangle^\beta \cdot \langle a|\boldsymbol{M}|o\rangle^\beta\} \\ &= (1/2)\mathrm{Im}\{(\mu_{ioa} + \mu_{joa}) \cdot (\boldsymbol{m}_{iao} + \boldsymbol{m}_{jao})\} \\ &\quad - (1/2)\pi\sigma_o \boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa}) \quad [30]\end{aligned}$$

In the case of π→π* transitions of common molecules, internal magnetic transition moments $m_{iao}$ and $m_{jao}$ are negligible. Therefore, rotational strengths are approximated as

$$R^\alpha = +(1/2)\pi\sigma_o \boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa}) \quad [31]$$

$$R^\beta = -(1/2)\pi\sigma_o \boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa}) \quad [32]$$

These equations indicate that the Cotton effects of α- and β–states have equal intensity but are of opposite signs, and therefore the exciton CD satisfies the sum rule:

$$\Sigma R_\mathrm{A} = R^\alpha + R^\beta = 0 \quad [33]$$

The rotational strength is proportional to the triple product of interchromophoric distance and electric transition moments of groups i and j. Therefore, provided chromophores exhibiting intense $\pi$-$\pi^*$ transitions are employed, intense exciton CD Cotton effects are observable. The rotational strengths of exciton CD satisfy the origin independence as shown in Equation [32].

### CD spectra of *N*-mer and dimer: quantitative definition of CD exciton coupling

The exciton theory is also applied to UV-visible and CD spectra of N-mers having N identical chromophores. When N chromophores undergoing intense $\pi$-$\pi^*$ transitions (o→a) interact with one another, the excited state splits into N energy levels. The wavenumber $\sigma_k$ of the kth excitation is expressed as

$$\sigma_k - \sigma_o = 2\Sigma\,\Sigma\, C_{ik}\, C_{jk}\, V_{ij} \qquad [34]$$

where the coefficients $C_{ik}$ and $C_{jk}$ can be obtained by solving the secular equation of Nth order. The rotational strength of $R^k$ is similarly formulated as

$$R^k = -\pi\sigma_o\Sigma\Sigma\, C_{ik}\, C_{jk}\, \boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa}) \qquad [35]$$

For the N-mer, the CD curve is formulated as

$$\Delta\varepsilon(\sigma) = \{\sigma_o/(2.296 \times 10^{-39}\sqrt{\pi}\Delta\sigma)\}\Sigma\, R^k \times \exp\{-((\sigma - \sigma_k)/\Delta\sigma)^2\} \qquad [36]$$

The Taylor expansion of Equation [36] against $\sigma_k/\Delta\sigma$ around $\sigma_o/\Delta\sigma$ yields the second term of the expansion as

$$\Delta\varepsilon(\sigma) = \{2\sigma_o/(2.296 \times 10^{-39}\sqrt{\pi}\Delta\sigma)\} \times \exp\{-((\sigma - \sigma_o)/\Delta\sigma)^2\}\{(\sigma - \sigma_o)/\Delta\sigma\} \times \Sigma R^k\{(\sigma_k - \sigma_o)/\Delta\sigma\} \qquad [37]$$

From Equations [34], [35] and [37], the CD equation of the N-mer is obtained:

$$\Delta\varepsilon(\sigma) = \{4\sqrt{\pi}\sigma_o{}^2/(2.296 \times 10^{-39}\Delta\sigma^2)\} \times \{(\sigma_o - \sigma)/\Delta\sigma\}\exp\{-((\sigma_o - \sigma)/\Delta\sigma)^2 \times \Sigma\{\Sigma\,\Sigma\, C_{ik}\, C_{jk}\, \boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa})\} \times \{\Sigma\,\Sigma\, C_{ik}C_{jk}V_{ij}\} \qquad [38]$$

In the case of a binary system, since the coefficients for the $\alpha$-state are $1/\sqrt{2}$ and $-1/\sqrt{2}$, and for the $\beta$-state, $1/\sqrt{2}$ and $-1/\sqrt{2}$, Equation [38] is simplified to

$$\Delta\varepsilon(\sigma) = \{2\sqrt{\pi}\sigma_o{}^2/(2.296 \times 10^{-39}\Delta\sigma^2)\} \times \{(\sigma_o - \sigma)/\Delta\sigma\}\exp\{-((\sigma_o - \sigma)/\Delta\sigma)^2 \times \boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa})V_{ij} \qquad [39]$$

This is the CD equation of a binary system. The next term of Equation [39],

$$\{2\sqrt{\pi}\sigma_o{}^2/(2.296 \times 10^{-39}\Delta\sigma^2)\} \times \{(\sigma_o - \sigma)/\Delta\sigma\}\exp\{-((\sigma_o - \sigma)/\Delta\sigma)^2$$

represents an anomalous dispersion curve with positive and negative extrema. The sign and intensity of exciton CD depend on the quadruple term, $R_{ij} \cdot (\mu_{ioa} \times \mu_{joa})\, V_{ij}$. Therefore, the term $R_{ij} \cdot (\mu_{ioa} \times \mu_{joa})\, V_{ij}$ is adopted as the quantitative definition of exciton chirality. This term is changed to

$$\boldsymbol{R}_{ij} \cdot (\mu_{ioa} \times \mu_{joa})V_{ij} = D_{ioa}D_{joa}R_{ij}{}^{-2}\boldsymbol{e}_{ij} \cdot (\boldsymbol{e}_i \times \boldsymbol{e}_j) \times \{\boldsymbol{e}_i \cdot \boldsymbol{e}_j - 3(\boldsymbol{e}_i \cdot \boldsymbol{e}_{ij})(\boldsymbol{e}_j \cdot \boldsymbol{e}_{ij})\} \qquad [40]$$

where $D_{ioa}$ and $D_{joa}$ are transition dipole strengths of groups i and j, respectively. This equation indicates that the exciton CD amplitude is inversely proportional to the square of the interchromophoric distance $R_{ij}$.

## List of symbols

A = amplitude of splitting; c = speed of light; D = dipole strength; $E_0$ = ground-state energy; H = Hamiltonian operator; J = NMR coupling constant; m = mass of an electron; m = internal

magnetic transition moment. M = operator of magnetic moment vector; $p_{is}$ = linear momentum of electron s in group i; $r_{is}$ = distance vector of electrons in group i; R = rotational strength; $R_{ij}$ = interchromophoric distance vector; $V_{ij}$ = interaction energy; $2V_{ij}$ = Davydov splitting; $\varepsilon$ = molar absorptivity; $\lambda$ = wavelength; $\mu$ = electric transition moment; $\sigma$ = wavenumber; $\Psi_0$ = ground-state wavefunction.

*See also:* **Biomacromolecular Applications of Circular Dichroism and ORD; Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism; Induced Circular Dichroism; Magnetic Circular Dichroism, Theory; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading


Berova N, Gargiulo D, Derguini F, Nakanishi K and Harada N (1993) Unique UV-Vis absorption and circular dichroic exciton-split spectra of a chiral biscyanine dye: origin and nature. Journal of the American Chemical Society 115: 4769-4775.

Davydov AS (1962) Theory of Molecular Exciton. Kasha M and Oppenheimer M, Jr (translators). New York: McGraw-Hill.

Harada N, Chen SL and Nakanishi K (1974) Quantitative definition of the exciton chirality and the distant effect in the exciton chirality method. Journal of the American Chemical Society 97: 5345-5352.

Harada N and Nakanishi K (1972) The exciton chirality method and its application to configurational and conformational studies of natural products. Accounts of Chemical Research 5: 257-263.

Harada N and Nakanishi K (1983) Circular Dichroic Spectroscopy - Exciton Coupling in Organic Stereochemistry. Mill Valley, CA: University Science Books.

Heyn MP (1975) Dependence of exciton circular dichroism amplitudes on oscillator strength. Journal of Physical Chemistry 79: 2424-2426.

Kasha M, Rawls HR and Ashraf EL-Bayoumi M (1965) The exciton model in molecular spectroscopy. Pure and Applied Chemistry 11: 371-392.

Kirkwood JG (1937) On the theory of optical rotatory power. Chemical Physics 5: 479-491.

Mason SF (1979) General models of optical activity. In Mason SF (ed) Optical Activity and Chiral Discrimination, pp 1-22. Dordrecht: Reidel.

Moffitt W, Fitts RD and Kirkwood JG (1957) Proceedings of the National Academy of Sciences of United States of America 43: 723-730.

Nakanishi K and Berova N (1994) The exciton chirality method. In Nakanishi K, Berova N and Woody R (Eds) Circular Dichroism – Principles and Applications, pp 361-398. New York: VCH.

# F

# $^{19}$F NMR, Applications, Solution State

**Claudio Pettinari** and **Giovanni Rafaiani**,
Università di Camerino, Italy

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

Fluorine-19 is spin 1/2, 100% abundant, and has a sensitivity relative to the proton of 0.83. It is a very easy nucleus to observe, and even prior to 1970 a great deal of information was obtained. $^{19}$F NMR has found widespread application in chemistry and biology: for example, the high-resolution $^{19}$F NMR study of polymers in solution furnishes relevant chemical information concerning polymer structures; furthermore, fluorine NMR can be useful for studying the binding of fluorinated molecules and macromolecules and membranes.

One reason for this wide application is that the wide range and the high sensitivity of the fluorine-19 chemical shift makes the resonance a very sensitive means of monitoring conformational, electronic and solvent effects and that it is possible to study fluorine bonded to nearly every element in the periodic table.

There is a sizeable literature on fluorine, and in particular several reviews with extensive compilations of $^{19}$F chemical shifts (solution state) and coupling constants, both in organic and in inorganic chemistry.

The nuclear properties of $^{19}$F are summarized in **Table 1**: the gyromagnetic ratio, 0.94 that of $^{1}$H, indicates a basic resonance frequency of 188.1 MHz on a spectrometer which operates at a field strength of 4.7 T with $^{1}$H NMR at 200 MHz.

**Table 1** NMR properties of fluorine-19

| | |
|---|---|
| Spin | 1/2 |
| NA | 100 |
| $R^{H}$ | 0.83 |
| $R^{C}$ | $4.73 \times 10^{3}$ |
| $\gamma$ ($10^{7}$ rad $T^{-1}$ $s^{-1}$) | 25.1815 |
| $\Xi$, MHz | 94.094003 |

NA is the natural abundance (%); $R^{H}$ and $R^{C}$ are the receptivity relative to that of $^{1}$H and $^{13}$C, respectively; $\gamma$ is the gyromagnetic ratio; $\Xi$ is the resonance frequency for the reference compound $CFCl_3$ in a magnetic field for which $(Me)_4Si$ has a proton resonance of 100 MHz.

## Experimental aspects

In $^{19}$F NMR experiments, a problem exists in the choice of chemical shift reference since the reactivity of some fluorine compounds has often precluded their use as internal standard. However, the standard reference compound most used is internal $CFCl_3$, a very volatile and unreactive derivative, which unfortunately, like $C_6F_6$, has a sizeable temperature and solvent dependence: medium effects are, in fact, greater in $^{19}$F NMR spectroscopy than in that of $^{1}$H NMR, the signal of several compounds being shifted by 2–10 ppm relative to the gas phase when the spectra are recorded in solvents such as *n*-heptane, $CHCl_3$ or $CH_3I$.

Unlike $^{1}$H NMR, in $^{19}$F NMR experiments it is not very important to consider the bulk susceptibility contribution to the shielding, the relative error being negligible due to the larger shift for $^{19}$F than for $^{1}$H with the same bulk susceptibility contribution. On the other hand, there are great difficulties in application of broad band decoupling owing to the large $^{2}J(^{19}F,^{1}H)$ coupling and the proximity of $^{1}$H and $^{19}$F chemical shift resonance frequencies.

$^{19}$F chemical information (for example, the magnitude and sign of $^{n}J(^{19}F, X)$ coupling constants) can be efficiently obtained from two-dimensional NMR

methods such as $^{1}H$–$^{19}F$, $^{13}C$–$^{19}F$, and $^{19}F$–$^{19}F$ correlation spectroscopy (COSY). These methods are helpful in determining small long-range coupling constants and in spectral assignment in polyfluoro derivatives.

$^{19}$F NMR chemical shifts cover such a wide frequency range that problems can occur with low digital resolution. For example, to cover a spectral width of approximately 400 ppm using a spectrometer which operates at 564 MHz for $^{19}$F (600 MHz for $^{1}$H NMR) requires a data set comprising 512 k points to give a reasonable digital resolution. With modern computer systems this is not a problem. Alternatively, the use of digital filtering or selective pulse methods can negate the need for such large data sets.

## Chemical shift

Fluorine chemical shifts cover a very large range (more than 1300 ppm, about two orders greater than $^{1}$H chemical shift) when organic, inorganic, and organometallic compounds are examined. **Table 2** exhibits selected $^{19}$F data all referenced to $CFCl_3$, a negative sign indicating a higher shielding or a shift to lower frequency. In several experiments, other compounds have been used as a chemical shift reference, so that to draw straightforward conclusions it is often necessary to convert the chemical shift values to the $\delta$ scale with $CFCl_3$ as reference, taking into account whether an external or internal standard has been used.

**Table 2** Selected $^{19}$F chemical shift in ppm relative to $CFCl_3$

| *Compound* | *δ(ppm)* | *Compound* | *δ(ppm)* |
|---|---|---|---|
| HF | 40 | $C_6H_5F$ | –116 |
| F–C≡N | –156 | $C_6F_6$ | –163 |
| F–C≡C–F | –95 | 1,2-$C_6F_4$ | –139 |
| $F_2C{=}CF_2$ | –135 | 1,3-$C_6F_4$ | –104 |
| $CH_2{=}CHF$ | –114 | 1,4-$C_6F_4$ | –113 |
| $CH_2{=}CF_2$ | –81 | $F_2$ | 429 |
| *cyclo*-$C_6F_{12}$ | –133 | $F_2CS$ | 108 |
| *cyclo*-$C_5F_{10}$ | –133 | $FSiH_3$ | –217 |
| *cyclo*-$C_4F_8$ | –135 | $F_3SiH$ | –109 |
| *cyclo*-$C_3F_6$ | –151 | $F_2SiH_2$ | –151 |
| $XeF_2$ | 258 | $CF_4$ | –63 |
| $XeF_4$ | 438 | $CH_3F$ | –272 |
| $XeF_6$ | 550 | $CH_2F_2$ | –144 |
| $F_2O_2$ | 865 | $CHF_3$ | –79 |
| ClF | –448 | $F_2O$ | 249 |
| $CF_3C({=}O)CF_3$ | –85 | $FClO_3$ | 287 |
| $CFBr_3$ | 7 | $PhSO_2F$ | 66 |
| $F^-K^+$ | –125 | $(CF_3)_3N$ | –59 |
| $(CF_3)_4C$ | –63 | $SiFBr_3$ | –77 |
| $F_3CCl$ | –33 | $SiF_2Br_2$ | –95 |
| $F_3CBr$ | –21 | $SiF_3Br$ | –125 |
| $CF_2Br_2$ | 7 | $SiF_4$ | –164 |
| $NF_3$ | 145 | $SeF_6$ | 55 |
| CHF=CHF(*E*) | –183 | $BF_3$ | –131 |
| CHF=CHF(*Z*) | –165 | $AsF_5$ | –66 |
| CFBr=CFBr (*E*) | –113 | $(AsF_6)^-$ | –70 |
| CFBr=CFBr (*Z*) | –95 | $IF_7$ | 170 |
| CFCl=CFCl (*E*) | –120 | $CF_3COOH$ | –78 |
| CFCl=CFCl (*Z*) | –105 | 1-Fluoronaphthalene | –123 |

## Isotope effects

Secondary isotope effects are relevant for $^{19}$F chemical shifts: the main known isotope effects have been observed with the standard $CFCl_3$ (**Figure 1**), its signal being split into four components of 21:21:7:1 intensity, corresponding to $CF^{35}Cl_3$, $CF^{35}Cl_2{}^{37}Cl$, $CF^{35}Cl^{37}Cl_2$ and $CF^{37}Cl_3$ compounds, due to the isotope effect of the $^{35}$Cl and $^{37}$Cl.

In $F(H_2O)_x$–HF–$HF_2^-$ species, which generally are obtained when KF is dissolved in $H_2O$, the substitution of H with D produces a significant high-field $^{19}$F chemical shift. For this reason it is possible to use the $^{19}$F chemical shift changes of KF to determine the percentage of $D_2O$ in $H_2O/D_2O$ mixtures.

## Solvent and temperature dependence of fluorine chemical shifts

Fluorine chemical shifts are very sensitive to temperature and solvent even in the absence of chemical changes (e.g. electrostatic solute–solvent interactions such as hydrogen bond formation, dissociation and change in the structure of the compounds examined). Moreover, variable temperature $^{19}$F NMR spectra can be employed to distinguish between species undergoing internal molecular dynamics between conformations: for example, at –130°C, the two

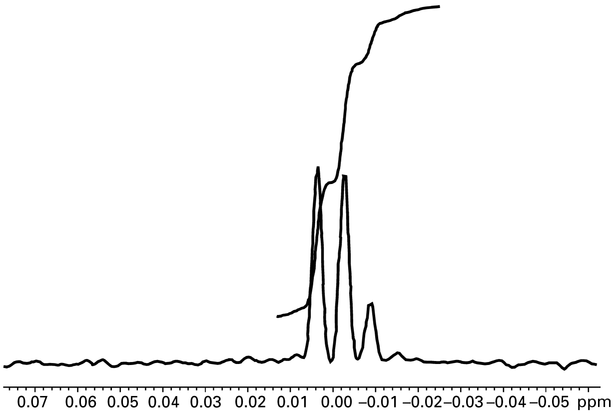



**Figure 1** $^{19}$F NMR spectrum of the standard $CFCl_3$.

rotamers of the $CFBr_2$–$CF_2Br$ molecule exhibit two different $^{19}$F chemical shifts and their ratio can be determined.

As an example of the scale of the temperature dependence of $^{19}$F chemical shifts, **Figure 2** shows the $^{19}$F resonance of 3-trifluoromethylpyrazole at different temperatures.

## Effective nuclear charge dependence of fluorine chemical shifts

A good correlation between $^{19}$F chemical shifts and electronegativity has been observed for the shielding in binary fluorides: it was found that for $EF_n$ compounds (E = Se, As, Te, Sb; $n = 5$ or 6) $^{19}$F resonance frequencies increase with increasing electronegativity. The effect of electronegativity can also be illustrated by the changes of $^{19}$F chemical shifts along the series $CH_nF_{4-n}$ and $SiH_nF_{4-n}$ as the value of $n$ changes from 1 to 3 (**Figure 3**).

An opposite trend (i.e. the shielding increases with increasing electronegativity) has been found in the $CF_nCl_{4-n}$ series: this is probably due to the greater +M (mesomeric) effect of F with respect to Cl.

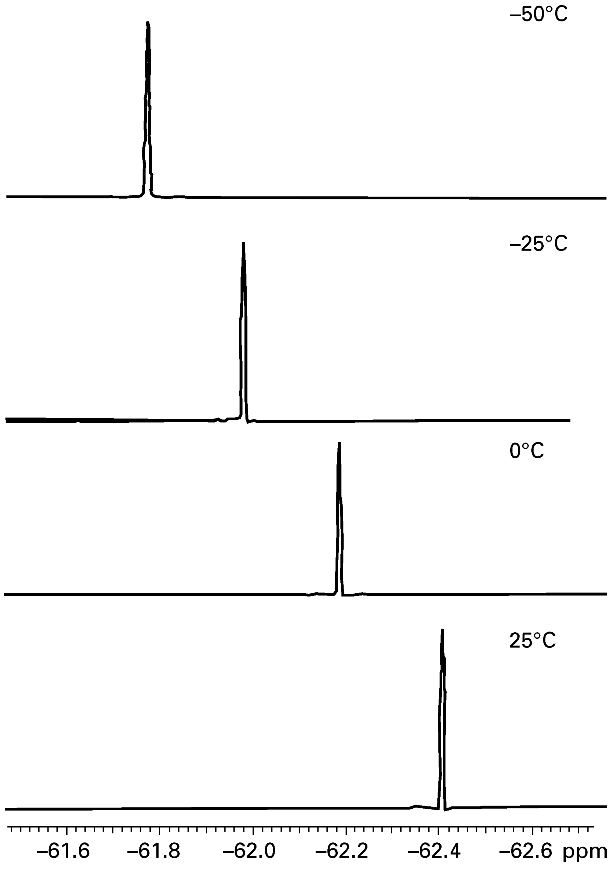



**Figure 2** $^{19}$F NMR resonances of the 3-trifluoromethylpyrazole at different temperatures.

Analogous simple correlations have also been reported between the $^{19}$F chemical shifts of mixed silicon halides of the type $SiF_nY_{4-n}$ (Y = Cl or Br; $n = 1$, 2 or 3) and the sum of the electronegativities of the substituents (**Figure 4**).

A similar pattern has been reported for the following pairs of compounds:

$$(CF_3)_2O \quad \delta(^{19}F) = -62 \text{ ppm} \qquad (CF_3)_2S \quad \delta(^{19}F) = -39 \text{ ppm}$$

$$CF_3CH_2Cl \quad \delta(^{19}F) = -71 \text{ ppm} \qquad CF_3CH_2Br \quad \delta(^{19}F) = -68 \text{ ppm}$$

$$CF_2ClCH_3 \quad \delta(^{19}F) = -45 \text{ ppm} \qquad CF_2BrCH_3 \quad \delta(^{19}F) = -37 \text{ ppm}$$

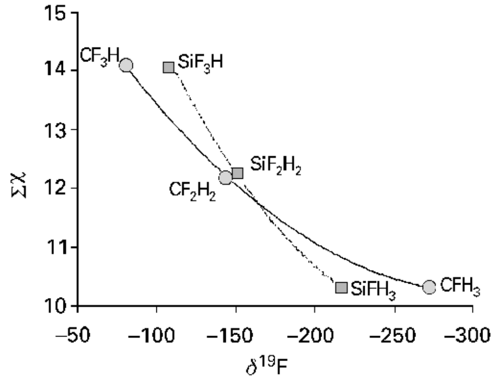



**Figure 3** Plot of $\delta(^{19}F)$ against $\Sigma\chi$ (electronegativity of the substituents) for a series of monomeric compounds $CH_{3-n}F_n$ and $SiH_{3-n}F_n$.

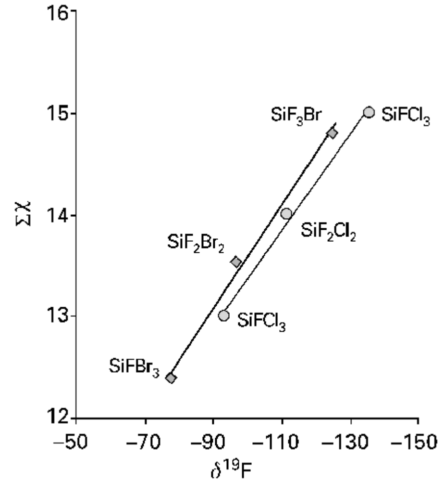



**Figure 4** Plot of $\delta(^{19}F)$ against $\Sigma\chi$ (electronegativity of the substituents) for a series of monomeric compounds $SiCl_{3-n}F_n$.

It should always be borne in mind that $^{19}$F shielding is controlled by paramagnetic and anisotropic terms to a greater extent with respect to the -I (inductive) effect of substituents.

The $^{19}$F chemical shift is also dependent on the total charge: $^{19}$F shielding decreases from the neutral to the anionic species; for example, $\sigma(MF_4) > \sigma(MF_6)^{2-}$ (M = Si, Ge, Sn) and also from neutral to cationic species $\sigma(NF_3) > \sigma(NF_4)^+$. In general, the shielding decreases with increasing number of fluorine atoms in the compound.

In poly- and perfluoroalkanes, the following simple relation (with some exceptions) has been found

$$\delta CF_3 > \delta CF_2 > \delta CF$$

which is completely the opposite of that found for $^1$H NMR spectra of alkane derivatives.

It is worth noting that the order of $^{19}$F chemical shifts in FC≡CCF$_3$ corresponds perfectly to the relative electron density of the two groups: the anomaly due to the anisotropy of the triple bond found in the $^1$H NMR chemical shift of acetylene is not detected in fluorine resonance.

## Substituent and structure dependence of fluorine chemical shifts

The chemical shift of the $CF_3$ group in a series of 1,1,1,4,4,4-hexafluorobut-2-enes, such as those in **Figure 5**, clearly indicates that in compounds containing *trans*-$CF_3$ groups, $\delta(^{19}F)$ is generally shifted to higher field with respect to isomers containing *cis*-$CF_3$ groups.

*Geminal*, and *cis* and *trans vicinal* fluorine in molecules such as $CF_2$=CFH also possess very different chemical shifts, the *geminal* (–185 ppm) being often the more shielded with respect to *cis* (–102 ppm) and *trans* (–129 ppm) *vicinal* ones.

Aromatic fluorine nuclei are more shielded with respect to perfluoroethylene derivatives, and shielding generally increases with increasing substitution of H with F: for example, $\delta(^{19}F)$ is –116 ppm for $C_6H_5F$, –139 ppm for 1,2-$C_6H_4F_2$, and –163 ppm for $C_6F_6$.

The $^{19}$F chemical shift values of fluoroaromatic compounds can be used as a measure of the electronic interaction of ring substituents: for example, it has been shown that the 4-fluorine shift can be related to the $\pi$-electron donating or withdrawing properties of the substituent group X in pentafluoroaromatics.

Several studies carried out on monofluorophenylplatinum complexes like that in **Figure 6** gave information on the electronic character of the bond between the platinum and the anionic X ligand. By using a simple relationship between the shift of the 4-fluorine atom and the coupling constant between the 2- and 4-fluorine atoms it has been possible to determine empirically the extent of $\pi$-interaction in organometallic compounds.

The difference in the chemical shift of the 3- and 5-fluorine atoms and that of 4-fluorine in *trans*-$[(Et_3P)_2Pt(C_6F_5)X]$ (where X = Me, Cl, Br, I,

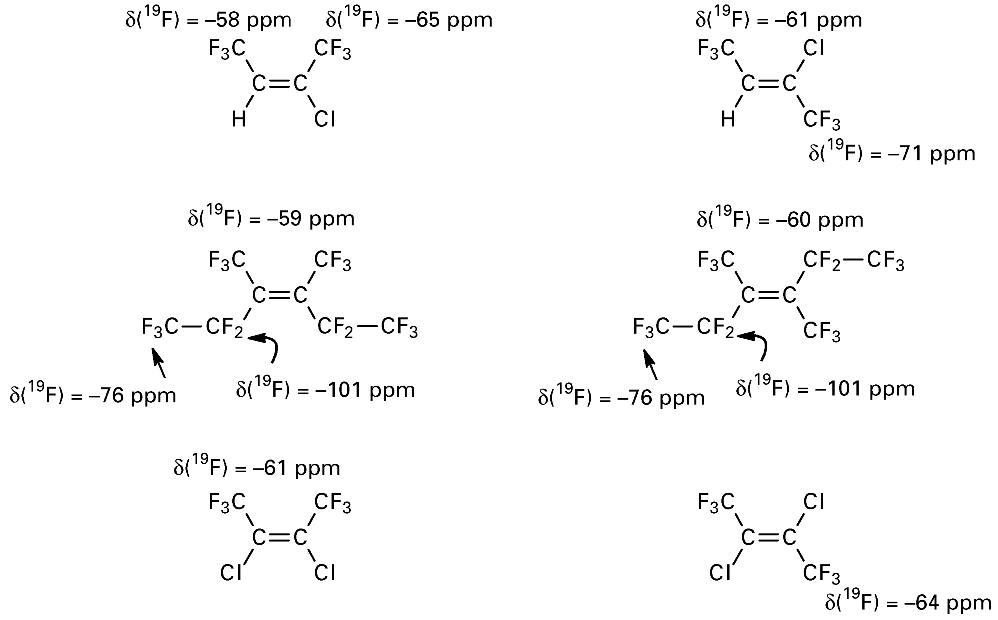
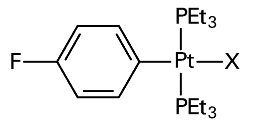



**Figure 5** $\delta(^{19}F)$ for a series of hexafluorobut-2-enes.




X = Cl, CN, Ph, *p*–$FC_6H_4$

**Figure 6** Monofluorophenylplatinum(II) complexes.

$NO_2$, NCS, CN and $ONO_2$) derivatives has been used as a criterion of $\pi$-acceptor interaction. The $\pi$-bonding sequence resulted in the following order sequence, in accordance with the *trans* effect:

$$\mathrm{Me} < \mathrm{Cl} < \mathrm{Br} < \mathrm{Br} < \mathrm{I} < \mathrm{NO_2} < \mathrm{NCS} < \mathrm{CN} < \mathrm{ONO_2}$$

The $^{19}$F chemical shift can be also employed to investigate halide-exchange reactions: from the reaction of $Me_3SbF_2$ ($\delta = -106.1$ ppm) with $Me_3SbCl_2$ the mixed halide $Me_3SbClF$ has been obtained ($\delta = -110.9$ ppm). Its $^{19}$F chemical shift value is significantly different from that found for the starting reagent.

## Anisotropy of fluorine chemical shifts

In anisotropic media the fluorine chemical shift is strongly dependent on the orientation of the molecule with respect to magnetic field. This fact is very important because many systems widely investigated by $^{19}$F NMR, for example biological matrices, are anisotropic media.

## Coupling constants

Coupling constants are very useful for discussing the structure and nature of chemical bonding in compounds containing fluorine. However, to date only a few relative signs of couplings have been reported, and it is not possible in all cases to establish with certainty the effect of the substituents and of the structure on the element–fluorine bond.

Coupling constants involving fluorine atoms are solvent and temperature dependent: for example, for the derivative ClF a difference of ~85 Hz has been found for $^1J(^{35}\mathrm{Cl},^{19}\mathrm{F})$ between the liquid and the gasphase. The differences are, in general, of the order of 0.1–10%. A dependence on temperature has been observed even when no conformational changes or exchange averaging processes are present. Several investigations, carried out on temperature coefficients of fluorine spin–spin coupling, indicated that these coefficients can have both signs and can even be as large as 2.3–2.4 Hz per °C.

## One-bond couplings

Typical values of one-bond spin–spin coupling constants between $^{19}$F and other nuclei $^1J(\mathrm{X},^{19}\mathrm{F})$ are reported in **Table 3**.

In binary fluorides, it has been shown that the reduced one-bond coupling constant $^1J(\mathrm{X},^{19}\mathrm{F})$ follows a periodic table trend; the value of $J(\mathrm{X},^{19}\mathrm{F})$ generally increases with increasing atomic number of X. The one-bond coupling constants between fluorine and metal atoms vary within a very large range; for example, the coupling between tungsten and fluorine is about 60–70 Hz, whereas that between platinum

**Table 3** Magnitudes of some representative one-bond coupling constants, $^1J(\mathrm{X},^{19}\mathrm{F})$

| *Compound* | *X* | $^1J(X,^{19}F)/Hz$ |
|---|---|---|
| HF | $^1$H | 530 |
| $Br_2C{=}CFBr$ | $^{13}$C | 324 |
| $Cl_2C{=}CFCl$ | $^{13}$C | 303 |
| $Br_2C{=}CF_2$ | $^{13}$C | 290 |
| $Cl_2C{=}CF_2$ | $^{13}$C | 289 |
| *cis*-BrFC=CBrF | $^{13}$C | 325 |
| *cis*-ClFC=CClF | $^{13}$C | 299 |
| *trans*-BrFC=CBrF | $^{13}$C | 355 |
| *trans*-ClFC=CClF | $^{13}$C | 290 |
| $CF_2HCF_2SiF_3$ | $^{29}$Si | 278 |
| $CH_2ClSiF_3$ | $^{29}$Si | 267 |
| $CHCl_2SiF_3$ | $^{29}$Si | 267 |
| $CCl_3SiF_3$ | $^{29}$Si | 264 |
| $CH_3SPF_4$ | $^{31}$P | 1032 |
| $C_2H_5SPF_4$ | $^{31}$P | 1045 |
| $C_6H_5SPF_4$ | $^{31}$P | 1060 |
| $(CH_3S)CH_3PF_3$ | $^{31}$P | 925 [$^1J(^{31}\mathrm{P}–^{19}\mathrm{F_{ax}})$] |
| | | 1062 [$^1J(^{31}\mathrm{P}–^{19}\mathrm{F_{eq}})$] |
| $(C_2H_5NH)_2PF_3$ | $^{31}$P | 694 [$^1J(^{31}\mathrm{P}–^{19}\mathrm{F_{ax}})$] |
| | | 875 [$^1J(^{31}\mathrm{P}–^{19}\mathrm{F_{eq}})$] |
| $(CH_3)_3PF_2$ | $^{31}$P | 552 |
| $(C_2H_5)_3PF_2$ | $^{31}$P | 580 |
| $(C_6H_5)_3PF_2$ | $^{31}$P | 660 |
| $(CH_3)_2P(O)F$ | $^{31}$P | 980 |
| $(CH_3)_2P(S)F$ | $^{31}$P | 985 |
| $(CH_3)P(O)F_2$ | $^{31}$P | 1104 |
| $(CH_3)P(S)F_2$ | $^{31}$P | 1147 |
| $(C_6H_5)_2PF$ | $^{31}$P | 905 |
| $(CH_3O)_2PF$ | $^{31}$P | 1210 |
| $WF_6$ | $^{183}$W | 44 |
| $WF_5Cl$ | $^{183}$W | 25 |
| *trans*-$WF_4Cl_2$ | $^{183}$W | 20 |

and fluoride groups ranges from 1100 to 2100 Hz. $^1J(\mathrm{M},^{19}\mathrm{F})$ depends strongly on the structure and coordination number of the metal atom M.

$^1J(^{13}\mathrm{C},^{19}\mathrm{F})$, generally negative, has a range centred at about 300 Hz and depends strongly on the nature and electronic properties of substituents. The one-bond coupling constants between phosphorus and fluorine have values between 500 and 1500 Hz. It is interesting to note that the $^1J(^{31}\mathrm{P},^{19}\mathrm{F})$ of $RPF_2$ groups generally decreases upon coordination to a metal: for example, in $CF_3PF_2$ the $^1J(^{31}\mathrm{P},^{19}\mathrm{F})$ is 1245 Hz, whereas in the complex $(CF_3PF_2)_2Mo(CO)_4$ it is 1172 Hz. The $^1J(^{31}\mathrm{P},^{19}\mathrm{F})$ coupling constant can be used to distinguish axial and equatorial fluorinated groups in octahedral derivatives: for example, in the tetrafluoro(pentane-2,4-dionato)phosphorus (V) derivative (**Figure 7**) which contains two sets of magnetically different fluorine atoms, the two coupling constants are 824 (axial) and 741 (equatorial) Hz, respectively.

The $^1J(^{29}\mathrm{Si},^{19}\mathrm{F})$ in silicon fluorides is dependent not only on temperature but also on the time-varying electric field, i.e. dispersion interactions, which alters the net electronic state of derivatives. The values of $^1J(^{29}\mathrm{Si},\ ^{19}\mathrm{F})$ increase with increasing intermolecular dispersion interactions.

$^1J(^{11}\mathrm{B},^{19}\mathrm{F})$ coupling constants are generally very small and dependent on concentration and the nature of the solvent: for example, in acetonitrile the $^1J(^{11}\mathrm{B},^{19}\mathrm{F})$ for the $BF_4$ ion is 0.39 ± 0.07 Hz, whereas in water it is 1.07 Hz due to ion-pair formation.

## Two-bond couplings

Typical values of two-bond spin–spin coupling $^2J(\mathrm{E}–\mathrm{X}–^{19}\mathrm{F})$ are reported in **Table 4**. *Geminal* coupling constants $^2J(\mathrm{E}–\mathrm{X}–^{19}\mathrm{F})$ are sensitive not only to the factors reported for the one-bond coupling, i.e. electronegativity, but also to the nature of the intervening atom X, to the F–X–E bond angle, and finally on the stereochemistry of the compound.

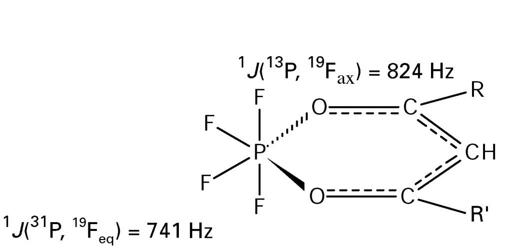



**Figure 7** Fluorine coupling constants in tetrafluoro(pentane-2,4-dionato)phosphorus(V).

**Table 4** Magnitudes of some slected geminal coupling constants $^2J(\mathrm{X},^{19}\mathrm{F})$

| *Compound* | *X* | *$^2J(X,^{19}F)$/Hz* |
|---|---|---|
| $Ph_2CF–CH_2F$ | $^1$H | 48 |
| $Ph_2CF–CHF_2$ | $^1$H | 52 |
| $HCF_2–CF_2OF$ | $^1$H | 56 |
| $CH_2F–C(CH_3)OHCO_2Me$ | $^1$H | 48 |
| $CH_3–CHF–CH(OH)–CO_2Et$ | $^1$H | 47 |
| *cis*-CHF=CHF | $^1$H | 73 |
| *trans*-CHF=CHF | $^1$H | 74 |
| $CF_3–CH_2–CFHI$ | $^1$H | 51 |
| $CF_3–CFH–CH_2I$ | $^1$H | 45 |
| $CF_3–CH_2–CFHBr$ | $^1$H | 50 |
| $CH_2F–O–CH_2F$ | $^1$H | 54 |
| $CH_2F–CHF–CO_2Et$ | $^1$H | 47, 48 |
| $CH_2F–CH_2–CO_2Et$ | $^1$H | 55 |
| $PhCHF–CCl_3$ | $^1$H | 42 |
| $CHF_2–CF_2–SiMe_3$ | $^1$H | 54 |
| $F_2HPBH_3$ | $^1$H | 55 |
| $(MeS)Ph_2PF_2$ | $^{19}$F | 28 |
| $WF_5Cl$ | $^{19}$F | 73 |
| $CF_3PF_2$ | $^{31}$P | 87 |
| $CH_2F–CF_2–PCl_2$ | $^{31}$P | 99 |
| $CHF_2–CF_2–PCl_2$ | $^{31}$P | 81 |
| $CHF_2–CHF–PCl_2$ | $^{31}$P | 83 |

### $^2J(^1\mathrm{H}–\mathrm{X}–^{19}\mathrm{F})$ and $^2J(^{13}\mathrm{C}–\mathrm{X}–^{19}\mathrm{F})$

Most of the data reported are for X = C. The $^2J(^1\mathrm{H}–\mathrm{X}–^{19}\mathrm{F})$ are generally in the range 50–80 Hz. It is very interesting to note that no significant difference has been observed between $^2J(^1\mathrm{H}–\mathrm{C}–^{19}\mathrm{F})$ and $^2J(^1\mathrm{H}–\mathrm{N}–^{19}\mathrm{F})$.

In methanes, HCFA′ A″ (A′ and A″ = alkyl, aryl, halides, etc.) a simple relationship (Eqn [1]) between $^2J(^1\mathrm{H}–\mathrm{C}–^{19}\mathrm{F})$ and the electronegativities $E_{\mathrm{A}'}$ and $E_{\mathrm{A}''}$ of A′ and A″ substituents has been found on the basis of a large set of experimental data.

$$^2J(^1\mathrm{H–C–}^{19}\mathrm{F}) = 78.76 + 8.45E_{\mathrm{A}'}E_{\mathrm{A}''} - 16.73\,(E_{\mathrm{A}'} + E_{\mathrm{A}''}) \qquad [1]$$

It has been possible to distinguish *cis*- and *trans*-1,2-difluoroethylene on the basis of their different *geminal* coupling constant values, the $^2J(^1\mathrm{H}–\mathrm{C}–^{19}\mathrm{F})$ for the *trans* being greater than that for the *cis* isomer.

The $^2J(^{13}\mathrm{C}–\mathrm{X}–^{19}\mathrm{F})$ are generally in the range 8–15 Hz and positive.

### $^2J(^{19}F–X–^{19}F)$

*Geminal* coupling constants $^2J(^{19}F–X–^{19}F)$ are very sensitive to hybridization of the intervening atom X and to the electronegativity of its substituents: the isotropic $^2J(^{19}F–C–^{19}F)$ can range from < 5 to 110 Hz for sp$^2$ carbon, and from ~100 to ~350 Hz for saturated carbon. In X–$CF_2$–$CH_2$F and X–$CF_2$–CFYZ derivatives (X, Y, Z = halide ≠ F), a relationship [2] has been found:

$$|^2J(^{19}F–X–^{19}F)| = 135.3 - 5.95\ \Sigma\chi_i \qquad [2]$$

where $\chi_i$ = electronegativity.

Also, in olefinic compounds XYC=$CF_2$, the $^2J(^{19}F–X–^{19}F)$ coupling constants are dependent on the nature of substituents. The magnitude of $^2J(^{19}F–X–^{19}F)$ is ~50 Hz for X and Y = H or C, whereas values of more than 100 Hz have been observed when X and Y are metals such as Pt or Pd.

In compounds of the type $F_2C$=CXY, a linear correlation between F–F *geminal* coupling constants and fluorine chemical shift has been found. It has been ascribed to a change in the density of the nuclear spin information carrying electrons to the intervening carbon atom due to valence bond resonance structures such as $F_2C^+$–CF=CF–$O^-$.

The dependence of F–F *geminal* coupling constants on the F–C–F angle is very different from that of *geminal* H–H couplings: in fact, the latter increase monotonically with increasing H–C–H angle, whereas the former change markedly with F–C–F angles and have a minimum near 110°. This is due to the fact that the Fermi contact term (FC) is sensitive to F–C–F angle, whereas both the spin dipolar (SD) and orbital (OB) terms seem to be insensitive to the change of F–C–F angle.

## Vicinal couplings

Typical values of three-bond spin–spin coupling $^3J(E–X–X–^{19}F)$ are reported in **Table 5**. They are dependent on the electronegativity of substituents, on the bond angle E–X–X–$^{19}$F, and also on the dihedral angle E–X–X–$^{19}$F.

From a comparison of proton–fluorine coupling constants and bond angles in a series of cyclic organic molecules, it has been found that $^3J(^1H,^{19}F)$ generally decreases as the H–C–C–F bond angle increases. In addition, the dihedral angle dependence of $^3J(^1H,^{19}F)$ is very similar to the dependence of $^3J(^1H,^1H)$, i.e. ($^1$H,$^{19}$F) has a maximum value when the dihedral angle $\phi$ between H and F is 0° and 180°, and a minimum value when the dihedral angle is 90°. In olefinic compounds (for example the last six derivatives in **Table 5**) it has been found that $^3J(^1H,^{19}F)_{cis}$ is always smaller than $^3J(^1H,^{19}F)_{trans}$.

**Table 5** Magnitudes of some slected vicinal coupling constants $^3J(X,^{19}F)$

| *Compound* | *X* | *$^3J$ (X,$^{19}$F)/Hz* |
|---|---|---|
| $Ph_2$CF–$CH_2$F | $^1$H | 20 |
| | $^{19}$F | 20 |
| $Ph_2$CF–$CHF_2$ | $^1$H | 12 |
| | $^{19}$F | 12 |
| $(CH_3)_2$CF–CH(OH)$CO_2$Me | $^1$H | 14 (CF–CH), 21(CF–$CH_3$) |
| $CH_3$–$CF_2$–$CO_2$Et | $^1$H | 19 |
| $CF_3$–$CH_2$–$CO_2$Et | $^1$H | 11 |
| *cis*-CHF=CHF | $^1$H | 4 |
| *trans*-CHF=CHF | $^1$H | 20 |
| *cis*-CHF=CFMn$(CO)_5$ | $^1$H | 10 |
| *trans*-CHF=CFMn$(CO)_5$ | $^1$H | 25 |
| *cis*-$(C_6H_5)$CF=$CHCH_3$ | $^1$H | 22 |
| *trans*-$(C_6H_5)$CF=$CHCH_3$ | $^1$H | 36 |
| $CF_3$–CFH–$CH_2$I | $^{19}$F | 11 |
| $CF_3$–CFI–$CH_2$–$CF_3$ | $^{19}$F | 11 |
| $CF_3$–CFBr–$CO_2$Et | $^{19}$F | 9 |
| $CF_3$–CFCl–$CO_2$Et | $^{19}$F | 7 |
| $CF_2$=C=CFCl | $^{19}$F | 30 |
| $ClCF_2$–$CF_2$–C(=O)F | $^{19}$F | 8($CF_2$–C(=O)F) 5(ClCF–$CF_2$) |
| $CHF_2$–$CH_2$–$PH_2$ | $^{31}$P | 8 |
| $CHF_2$–$CF_2$–$PCl_2$ | $^{31}$P | 27 |
| $CHF_2$–$CH_2$–$PCl_2$ | $^{31}$P | 13 |
| $CHF_2$–$CF_2$–$PCl_2$ | $^{31}$P | 49 |

The angular dependence of *vicinal* F–F couplings on the $^{19}$F–C–C–$^{19}$F dihedral angle is very different from that reported for *vicinal* H–H couplings. Some theoretical studies have suggested that in *vicinal* F–F couplings, although the spin dipolar and orbital terms are a simple function of the dihedral angle $\phi$, $^{19}$F–C–C–$^{19}$F, the Fermi contact term is a function which is strongly dependent both on substituents and $\phi$. Therefore, for total *vicinal* F–F couplings (to which the Fermi contact term makes the most important contribution), it is very difficult to find clear relationships involving dihedral angles.

## Long-range couplings

Long-range coupling constants have been found to be structurally and stereochemically dependent. $^4J(^{19}F,^{19}F)$ coupling constants can be much larger in magnitude than $^3J(^{19}F,^{19}F)$ and the values range from 0 to 200 Hz. The values of $^4J(^{19}F,^{19}F)$ are diagnostic in several saturated cyclic, unsaturated, and aromatic

molecules. For example, from the spectra of *cis*- and *trans*-4H-perfluoromethylcyclohexane, both having an equatorial $CF_3$ group, the following $^4J(^{19}F,^{19}F)$ coupling constant values have been deduced:

$$^4J(^{19}F_{eq},^{19}F_{eq}) \approx 9\ \text{Hz};\ ^4J(^{19}F_{eq},^{19}F_{ax}) \approx 1\ \text{Hz};$$
$$^4J(^{19}F_{ax},^{19}F_{ax}) \approx 27\ \text{Hz}$$

$^4J(^{19}F_{ax},^{19}F_{ax})$ values can be very large due to the through-space interactions of sterically close diaxial fluorine atoms. On the other hand, in aromatic fluorobenzene derivatives the $^4J(^{19}F,^{19}F)$ coupling constant provides only limited structural information: the $^3J(^{19}F,^{19}F)$, $^4J(^{19}F,^{19}F)$ and $^5J(^{19}F,^{19}F)$ coupling constants are often very similar and they can range from 18 to 20, 0 to 20 and 5 to 18 Hz, respectively.

A single large five-bond constant value (~75 Hz) has been observed for the conformationally gauche form of bis(ditrifluoromethyl)tetrachlorethane and it has been ascribed to the close proximity of the fluorines on the two trifluoromethyl groups (**Figure 8**).

In other cases, it has been proposed that some unusually large long range couplings with the nuclei not in close spatial proximity are due to the connection of the fluorine by an intervening π network.

The $^nJ(^{19}F,^{19}F)$ coupling constant between the two fluorine atoms in compound I in **Figure 9** is exceptionally large. It is the highest value reported to date and it also exceeds three- and four-bond coupling constants. This has been attributed to the overlapping nonbonding electron distribution for two F nuclei in close spatial proximity which allows correlation not only of electron-pair bond to electron pair, but also of electron spin.

The $^nJ(^{19}F,^{19}F)$ coupling constant determined for other molecules containing the F–C=C–C=C–F moiety, in which coupling occurs through the s electrons, are nearly two orders of magnitude smaller; for example, $^5J(^{19}F,^{19}F)$ in $CF_2$=CH–CH=$CF_2$ is 35.7 Hz, and in compounds II and III in **Figure 9** it is 16.5 and 7.0 Hz, respectively.

In fluorofuranoses (for example compound IV in **Figure 9**), the $^{19}$F chemical shift and the coupling constants often provide a means of assigning the stereochemistry: for example, 1$\alpha$-F resonates to higher field with respect to 1β-F, the *cis* coupling to H2 is much smaller (4–6 Hz) than the corresponding *trans* coupling (20 Hz), and finally the *trans* coupling to H4 (4–8 Hz) is much larger than the *cis* one (1–2 Hz).

## $^{19}$F relaxation

Very few reports have been published on fluorine relaxation times: in most cases the results reported for several different molecules indicate that the spin rotation mechanism is dominant at higher temperatures whereas the intermolecular dipole–dipole mechanism is not negligible at lower ones. The chemical shift anisotropy contribution can also be an important factor, whereas the intramolecular dipole–dipole mechanism and scalar coupling contribution seem to be negligible in $^{19}$F relaxation. Some fluorine relaxation times ($T_1$, $T_2$ and $T_{1\rho}$) have been used for establishing dynamics in a copolymer of tetrafluoroethene and hexafluoropropene.

## Applications

The ever-increasing capabilities of NMR instruments and the high NMR sensitivity allow analytical applications of fluorine NMR to small and big molecules and relevant important information to be gained on chemicals, polymers, membranes, and intracellular environments. Such information is not readily available by other methods. For example, $^{19}$F

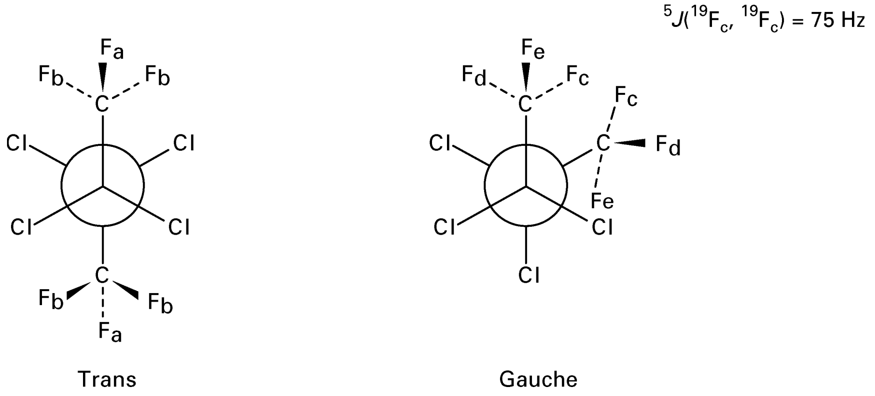



**Figure 8** The two conformational isomers of bis(ditrifluoromethyl)tetrachloroethane.

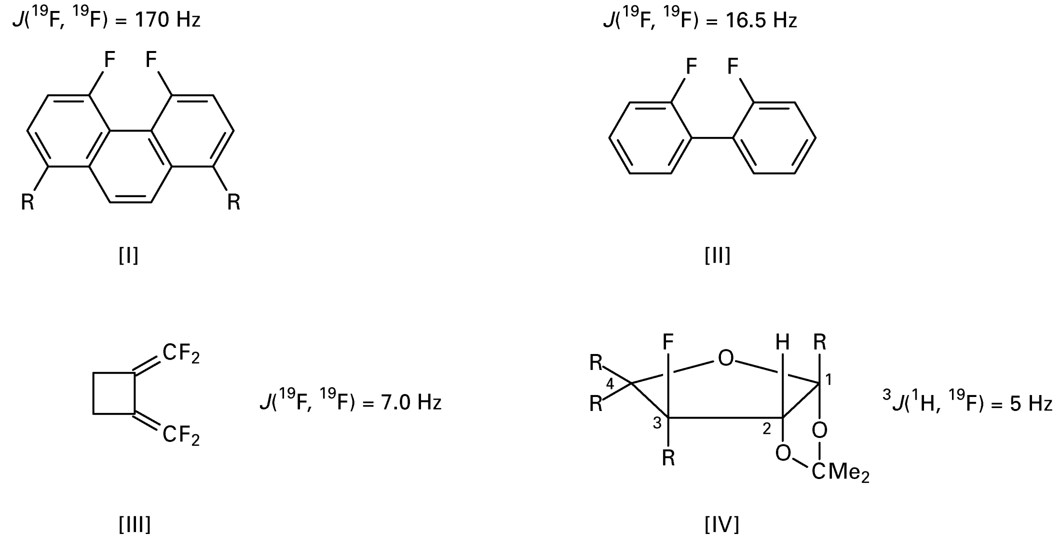



**Figure 9** Long-range coupling constants in different compounds containing fluorine.

spectroscopy has proved to be useful within a large range of applications, from the study of various random and alternating copolymers of fluorosiloxane and hybrid fluorocarbon fluorosiloxane to the study of fluoro complexes of aluminium in aqueous solution and in zeolites. Very recently, $^{19}$F NMR has been applied to study the effect of plasticizers on $LiCF_3SO_3$-containing polymer electrolytes and also to investigate the molecular structure of hydrogen-bonded clusters between $F^-$ and $(HF)_n$.

$^{19}$F NMR spectroscopy can be successfully applied to study relevant catalytic reactions: for example, the oxidative addition of $C_6Cl_2F_3I$ to Pd(0) and the *cis*-to-*trans* isomerization (involved in the Stille reaction and other Pd-catalysed syntheses) have been investigated, and it has been shown that the isomerization of *trans*-$[Pd(C_6Cl_2F_3)I(PPh_3)_2]$ reveals a first-order law. A four-pathway mechanism for this isomerization has also been proposed on the basis of $^{19}$F NMR kinetic study.

Several $^{19}$F NMR studies have been reported on ligand exchange reactions of metal complexes (i.e. the $Al(Ttfac)_3$-$Al(bzac)_3$ system) and also on relative polarity and reactivity of N–H, N–Hg and N–Au bonds in intermolecular exchange reactions of fluorine-containing ligands (for example, 2-(4-fluorophenyl)benzimidazole) and their PhHg and $Ph_3PAu$ derivatives.

Another interesting chemical application is the determination of the enantiomeric excesses of chiral acids by $^{19}$F NMR of their esters derived from fluorine compounds.

Molecules containing fluorine that can be transported into cells can be employed as indicators for monitoring intracellular environments. Recently, $^{19}$F NMR has been used to characterize drug–protein conjugates (immunogens) and to probe drug–protein binding. $^{19}$F NMR spectroscopy has been shown to be a powerful technique for identifying metabolites of fluorine-containing drugs and for evaluating changes in the metabolism of fluoropyrimidines after the use of a biochemical modulator. This technique often allows a correlation between improved therapeutic response and the biochemical effects generated in tissues. For example, in the monitoring of *in vivo* tumour metabolism, after biochemical modulation of 5-fluorouracil by the uridine phosphorylase inhibitor 5-benzylacyclouridine, analysis of the NMR data revealed an increased formation and retention of fluorouracil nucleotides and fluorouridine in colon tumours treated with the regimen containing 5-benzylacyclouridine, and a reduction in 5-fluorouracil catabolites.

*See also:* **Chemical Exchange Effects in NMR; Macromolecule–Ligand Interactions Studied By NMR; NMR in Anisotropic Systems, Theory; NMR Principles; NMR Relaxation Rates; Nucleic Acids Studied Using NMR; Parameters in NMR Spectroscopy, Theory of; Structural Chemistry using NMR Spectroscopy, Organic molecules.**

## Further reading

Dungan CH and Van Wazer JR (1970) *Compilation of Reported 19F Chemical Shifts.* New York: Wiley.

Emsley JW and Phillips L (1971) Fluorine chemical shifts. *Progress in NMR Spectroscopy* **7**: 1–523.

Emsley JW, Phillips L and Wray V (1976) Fluorine coupling constants. *Progress in NMR Spectroscopy* **10**: 83–756.

Fields R (1972) Fluorine-19 nuclear magnetic resonance spectroscopy. *Annual Reports on NMR Spectroscopy* **5A**: 99–304.
Fields R (1977) Fluorine-19 nuclear magnetic resonance spectroscopy. *Annual Reports on NMR Spectroscopy* **7**: 1–117.
Jameson CJ (1987) Fluorine. In: Mason J (ed) *Multinuclear NMR*, pp 305–333, New York: Plenum Press.
Mooney EF and Winson PH (1967) Fluorine-19 nuclear magnetic resonance spectroscopy. *Annual Reports on NMR Spectroscopy* **1**: 243–311.
Sohár P (1983) *Nuclear Magnetic Resonance Spectroscopy*, vol II, pp 261–268. Boca Raton, Florida: CRC Press, Inc.
Wray V (1980) Fluorine-19 nuclear magnetic resonance spectroscopy (1976–1978). *Annual Reports on NMR Spectroscopy* **10B**: 1–507.
Wray V (1983) Fluorine-19 nuclear magnetic resonance spectroscopy (1976–1978). *Annual Reports on NMR Spectroscopy* **14**: 1–406.

# Far-IR Spectroscopy, Applications

**James R Durig**, University of Missouri–Kansas City, MO, USA

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
**Applications**

## Introduction and history

The far-infrared spectral region is the range of wavenumbers where one finds many of the large-amplitude anharmonic vibrations. These include both the symmetric and asymmetric internal torsional modes of many small organic and organometallic molecules, ring puckering vibrations of four- and five-membered rings, and the heavy-atom skeletal bending modes. Additionally, this is the spectral region where one finds lattice vibrations from which information is obtained on intermolecular forces that are of considerable interest to both chemists and the physicists.

Historically, this region has been defined from 200 to 10 $cm^{-1}$, which is the region below the cutoff of CsI, where it can no longer be used as a dispersive material. Initial investigations in this spectral region began before the turn of the twentieth century in Europe, and it has always been more popular in Western Europe and Japan than in the United States. Even before the use of Fourier transform interferometers became common for this spectral region, there were several laboratories in these other countries where the far-infrared spectral region was studied almost exclusively. The first chemical applications of Fourier transform infrared (FT-IR) spectroscopy were made in the far-infrared spectral region because the interferometric instrumentation required to obtain spectra in this region is much simpler than that needed for the mid- or near-infrared regions. The instrumental advantages in the far-infrared spectral region included a lower tolerance for the mirror drive of a Michelson interferometer, a smaller dynamic range of the interferogram and a longer sampling time, which results in a reduced number of data points.

The far-infrared spectral region initially had its inherent difficulties, such as low-energy sources, poor detectors, lack of suitable optical materials and spectral interference of water vapour. Research workers in this wavenumber area have overcome many of these difficulties and, therefore, spectral investigations in the far-infrared region have become more common. This is particularly true with the commercial availability of Fourier transform spectrometers, which routinely provide spectral data to 30 $cm^{-1}$ with excellent signal-to-noise ratios. However, because of the absorption of water vapour in this spectral region (see **Figure 1**), it is highly desirable to have a vacuum interferometer bench.

Many of the far-infrared spectral studies for chemical information have emphasized the gas phase. This emphasis should not be interpreted as meaning that far-infrared spectral studies of solids are not important. Physicists have put interferometers to use in the study of the solid state to determine optical constants such as the index of refraction, complex indices, phase angle transmission coefficients and the electronic processes in insulation crystals, as well as the intermolecular vibrations of molecular crystals. These studies have been very important in the development of interpretative theories of solids.

With the increased ability to utilize *ab initio* calculations for predicting relative conformational

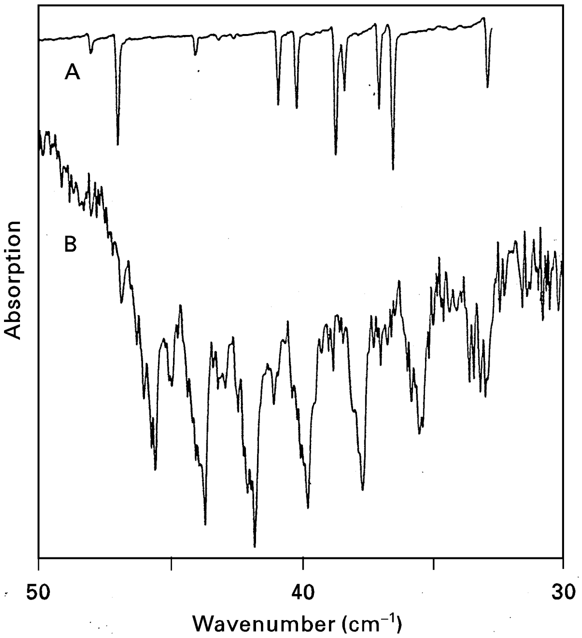



**Figure 1** Far-infrared spectra of (A) water vapour and (B) trifluoroacetyl fluoride.

stabilities, barriers to internal rotation, small ring inversions and low-wavenumber vibrations, the theoretical predictions are frequently compared with the experimental results in this spectral region. The use of either an incomplete or an incorrect experimental database in comparisons with theoretical results, may lead to erroneous conclusions with regard to the reliability and accuracy of the theoretical results.

Poor experimental databases with regard to the experimental determination of potential barriers and conformational equilibria arise for three basic reasons. First, there is frequently insufficient structural information in many cases. Numerous potential functions governing internal rotation of asymmetric tops (conformational interchange) have been published without the knowledge of the structures of both conformers, particularly of the high-energy conformer. It is extremely important, for the accurate experimental determination of the energy levels and potential functions, that some reasonable structural information be known about all of the exhibited conformations. Furthermore, in the absence of reliable structures determined from experimental techniques, such as those determined from microwave spectroscopy or analysis of the electron diffraction pattern, the reliability of optimized geometries determined from theory cannot be ascertained.

Secondly, there is frequently a lack of complementary experimental data. Owing to symmetry and the resulting selection rules, both infrared and Raman spectroscopic results are often needed to identify the high-energy and low-energy conformers. The Raman spectrum of the gas is almost always neglected, and it is not unusual to find vibrational conformational studies for which spectra are not reported for all three phases.

Thirdly, there is often a lack of high-resolution data. The highest possible resolution available should be used to record the low-frequency spectral data. Many torsional bands are extremely sharp and, therefore, transitions may not be observed at the lower resolution. Additionally, insufficient resolution may result in misleading relative intensities of the Q branches from different energy levels.

Low-wavenumber FT-IR spectroscopy is among the most generally applicable methods used in the study of the conformers for certain types of small molecules with few substituents. Infrared spectra can, and variable-temperature experiments may, be investigated in all phases. The gas-phase band contours observed for the infrared spectra, along with Raman depolarization data, provide considerable information on the molecular symmetry of the conformers. The limitation of the vibrational spectroscopic technique is that it is best applied to relatively simple molecules that contain at least one element of symmetry, and one or perhaps no more than two portions of the molecule capable of producing different conformations upon internal rotation.

One of the earlier uses of the far-infrared spectral region was to determine the barrier to internal rotation of symmetric-top rotors. The most frequently studied molecules were those with one threefold internal rotation, and many of those studies were for methyl tops. The first indication that the rotation around single bonds was not free came from thermodynamic data in the mid-1930s. Barriers were calculated from the thermodynamic data by relating the difference in the observed and statistical entropies by tables involving the barrier height and the reciprocal of the partition function for free rotation. Later, barrier values were obtained from investigations of the microwave spectra of small molecules where the observed perturbations on the pure rotational transitions were correlated with the torsional barrier heights by either the splitting or intensity methods. The splitting method is the most exact method since it depends only on the height of the barrier and usually gives barrier values to a few percent. The microwave intensity method is frequently used for the determination of the fundamental frequencies of asymmetric rotors.

One wishes to obtain the fundamental wavenumber for the torsional mode for the gaseous molecule

so that the barrier height in the isolated molecule can be ascertained. However, the dipole moment change associated with the torsional mode may be quite small, and consequently the resulting infrared band intensity may be weak. Therefore, assignments of the torsional modes of molecules in the gas were frequently in error in the initial investigations and the use of isotopic substitution is often necessary to verify the torsional assignments. Also, torsional modes that give rise to B-type infrared band contours may have bands that are very broad and ill-defined because of unresolved excited-state transitions. In favourable cases where the torsional mode gives rise to A-type or C-type band contours with relatively strong Q branches, several excited state transitions may be resolved and not only can the barrier height be obtained but also the detailed shape of the potential well may be ascertained.

When a molecule has a far-infrared spectrum too complicated for quantitative interpretation, estimates of the barrier heights may be obtained by studying the infrared and Raman spectra of the sample in the crystalline state. At liquid-nitrogen temperature, the upper vibrational states are effectively depopulated and only the $1 \leftarrow 0$ torsional transition is observed. The barrier height may be estimated from this single experimental datum, but a detailed analysis of the shape of the potential well is not possible. Additionally, it is sometimes possible to observe the torsional mode in the solid state when it is forbidden for the gaseous molecule. Barriers obtained in the solid state are usually 10–15% higher than those for the gaseous molecules.

## Internal rotors

Internal rotors fall into two categories: symmetric and asymmetric tops. For symmetric tops, a rotation about the top-frame bond of $2\pi/n$ (where $n$ is an integer) will bring the top to a position symmetrically equivalent to, or indistinguishable from, the original configuration. Therefore, it is usual to speak of the 'foldness' of the rotor in terms of $n$. For example, a perfluoromethyl group ($CF_3$) is a threefold symmetric top (local $C_{3v}$ symmetry), and twofold rotors include $–NO_2$, $–BF_2$ and phenyl groups (local $C_{2v}$ symmetry). When a rotation of 360° (i.e. when $n = 1$) is the only operation that results in a symmetrically equivalent position for the top, the rotor is known as an asymmetric top. For the case of a symmetric frame, the top with the highest degree of symmetry prevails and, when two tops of different foldness are bonded directly to one another, the resultant foldness is the product of the two individual tops' foldness. For instance, $CH_3BH_2$ would be classified as a sixfold internal rotor.

The energy minima and maxima for a symmetrical threefold group ($CF_3$) are 60° apart. The simplest mathematical function that will reproduce such a potential variation upon rotation is a cosine function. If the problem is assumed to be one-dimensional, the quantum-mechanical energy solutions are readily obtainable. The model employed is a rigid symmetric top ($CF_3$ group) attached to a rigid frame, which may be completely asymmetric. There are four rotational degrees of freedom, three for overall rotation and one for the hindered rotation of the two groups. The axis of internal rotation is usually assumed to coincide with the unique axis of the symmetric top. Since the top has a threefold symmetry axis, the potential energy hindering rotation may be expressed by a Fourier expansion:

$$V(\phi) = \frac{V_3}{2}(1 - \cos 3\phi) + \frac{V_6}{2}(1 - \cos 6\phi) + \cdots$$

where $V_3$ is the height of the threefold barrier; $V_6$ the sixfold; and so forth. A positive $V_6$ makes the minima narrower and the maxima broader, which results in the energy levels near the bottom of the well becoming somewhat more widely separated. A negative $V_6$ term has the opposite effect. Experimentally it has been found that $0 \le V_6/V_3 < 0.05$ and $V_6 >> V_9 >>$ higher-order terms.

An example of the use of far-infrared spectral data for the determination of the barrier to internal rotation is given by trifluoroacetyl fluoride, $CF_3CFO$ (**Figure 1**). Seven torsional transitions are clearly observed, beginning with the fundamental at 45.65 $cm^{-1}$ and continuing to the $7 \leftarrow 6$ transition at 33.40 $cm^{-1}$. Utilizing an $F$ number ($h^2/8\pi^2 I_r$ where $I_r$ is the reduced moment of inertia of the top) of 0.5970 $cm^{-1}$ along with the wavenumbers of these seven transitions, the $V_3$ value is $382 \pm 2$ $cm^{-1}$ and $V_6 = 8 \pm 1$ $cm^{-1}$. The intensity of the $4 \leftarrow 3$ transition and higher excited state transitions are significantly higher than expected on the basis of the Boltzmann factors because of the increased anharmonicity in the higher excited states. The statistical uncertainty is very low for both the $V_3$ and $V_6$ terms and the $V_6$ term is very small. This molecule is too heavy for the barrier to be determined by the microwave splitting method. Molecules with two or three equivalent rotors have also been extensively investigated, but the spectral data are much more complex and frequently there is not sufficient data to obtain values for all of the potential constants. Good barrier determination from far-infrared data requires that the internal

torsional mode is not mixed (potential energy distribution) with other low-wavenumber bending modes.

The asymmetric rotor is the other type of torsional motion frequently studied. These rotors usually lead to two or more stable conformers. One of the major goals of conformational analysis is the calculation of the energy difference between the two conformers and the energy necessary for interconversion. Four types of information are required to characterize an asymmetric potential function: (1) the approximate dihedral (torsional) angle of each conformer, because the number of torsional energy levels is directly related to the number of potential minima; (2) the approximate relative enthalpy difference between the high- and low-energy conformers, since this is one of the constraints defining the potential functions; (3) the change in molecular kinetic energy as a function of torsional angle; and (4) accurate observation of torsional transition frequencies. A typical spectrum for such a rotor is shown in **Figure 2** where the fundamental torsional mode of the *trans* conformer of $FCH_3CFO$ is observed at 118.9 $cm^{-1}$ with the excited torsional modes falling to lower wavenumbers to the 13 $\leftarrow$ 12 transition at 67.0 $cm^{-1}$. The fundamental torsional mode for the *cis* conformer is assigned at 52.1 $cm^{-1}$ with five successive excited states falling to lower wavenumber with the last one observed at 43.7 $cm^{-1}$. These data give the following potential constants: $V_1 = 86 \pm 11$, $V_2 = 946 \pm 33$, $V_3 = 407 \pm 4$, $V_4 = 138 \pm 20$, and $V_6 = -14 \pm 7\,cm^{-1}$ with a *trans* to *cis* barrier of $1297 \pm 26$ $cm^{-1}$.

## Small ring molecules

Another type of large-amplitude anharmonic vibration is that of the ring puckering modes of small compounds. From the analysis of the frequencies for these ring-bending motions it is possible to obtain the potential surfaces for the interconversion of the different conformers. The equilibrium conformation of four-membered and unsaturated five-membered rings is determined by two major opposing forces. The first is the ring strain, which tends to keep the ring skeleton planar. Puckering the ring decreases the already strained angles, thereby increasing the angle strain. The second is the torsional forces. The torsional repulsions of adjacent groups are at a maximum for a planar ring since the groups are eclipsed; therefore, bending the ring out of the plane will reduce these repulsive forces. It is the delicate balance between these two large forces that determines the ground-state structure of the molecule.

A single substituent on the cyclobutane ring leads to conformational isomers and introduces asymmetry into the potential function for ring inversion. As the ring inverts, the substituent goes from the axial conformation to the equatorial, or vice versa. These two conformations will of course have different energies and the spectra can be interpreted using a potential function of the form

$$V(x) = ax^4 - bx^2 - cx^3$$

where $x$ is the coordinate of the ring inversion. The cubic term is added to the usual quartic-harmonic potential function because of the asymmetry.

The results for chlorocyclobutane will be used as an example of the utility of far-infrared spectra of the gas for the investigation of ring puckering motions. The far-infrared and Raman spectra of gaseous chlorocyclobutane in the region of the ring puckering fundamental are shown in **Figure 3**. From the far-infrared spectrum, this fundamental is observed as well-defined Q branches occurring at 157.55, 149.29, 139.81, 130.12 and 128.08 $cm^{-1}$ with additional, weaker Q branches at 116.46, 115.69, 111.46 and 110.88 $cm^{-1}$. The first four pronounced Q branches appear to form a reasonable series and are assigned as the first four transitions of the ring puckering mode for the more stable

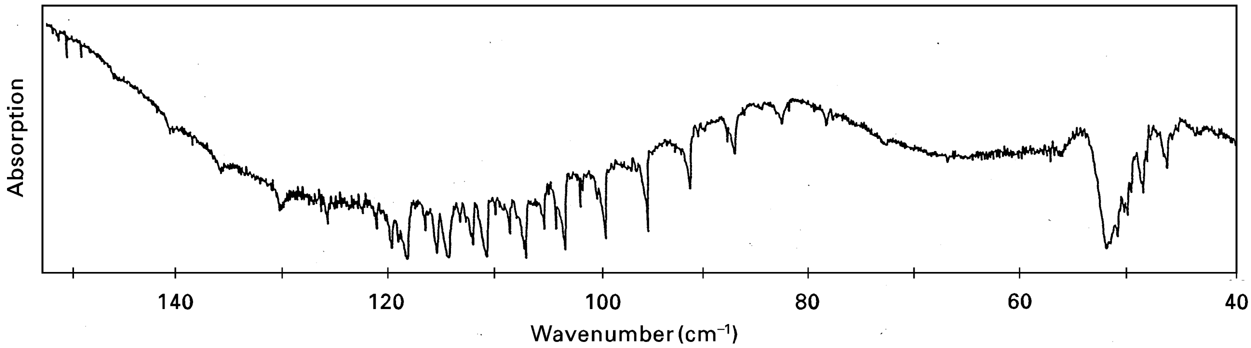



**Figure 2** Far-infrared spectrum of fluoroacetyl fluoride, $FCH_2CFO$.

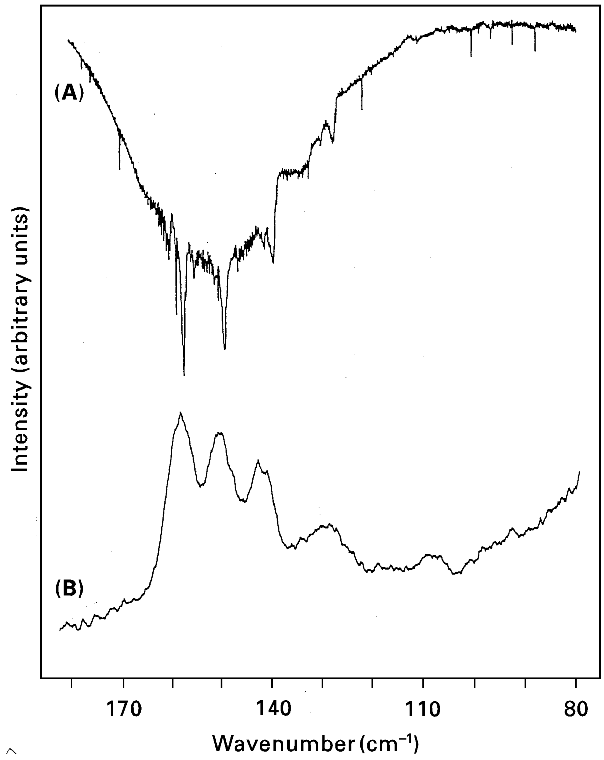



**Figure 3** (A) Far-infrared transmission spectrum of chlorocyclopropane; (B) Raman spectrum of gaseous chlorocyclobutane.

equatorial conformer. From the Raman spectrum, the ring puckering fundamental is observed as a weak series of Q branches beginning at ~158 cm$^{-1}$. Although the fundamental for the axial form can be assigned at 128 cm$^{-1}$ in this Raman spectrum, the 128.08 cm$^{-1}$ Q branch observed in the infrared spectrum is far more definitive.

With these assignments, the potential function governing ring inversion can be calculated with the aforementioned function. The function calculated using this assignment is shown in **Figure 4**. This potential is consistent with an energy difference between the equatorial and axial conformers of 449 cm$^{-1}$ (1.28 kcal mol$^{-1}$) and a barrier to ring inversion of 827 cm$^{-1}$ (2.36 kcal mol$^{-1}$). Although the uncertainty in the puckering angles determined at the minima is directly related to the uncertainty in the assumed reduced mass (198 amu), as well as the nature of the molecular motion involved, this potential gives puckering angles of 22° for the equatorial form and 17° for the axial conformer. From the *ab initio* calculations utilizing the 6-31G* basis set, the values for these angles are 25.1° and 20.3°, respectively, indicating that the difference in the two angles should be about 5°, which is the value obtained from the potential function.

In five-membered ring molecules, there are two low-frequency out-of-plane ring motions. These are usually qualitatively described as the ring-twisting (radial mode) and the ring-puckering (pseudorotational mode) motions. Initially, in order to handle the interpretation of the low-frequency far-infrared data, assumptions were made about the forms of these normal vibrations. If the five-membered ring contains an endocyclic double bond, it has usually been assumed that there is no interaction between the 'high'-frequency ring-twisting mode, which falls around 400 cm$^{-1}$ and the 'low'-frequency ring-puckering mode, which is near 100 cm$^{-1}$. Therefore, the two out-of-plane ring modes are handled as two one-dimensional problems with the anharmonic, low-frequency ring-puckering transitions interpreted in terms of a one-dimensional potential function of the form $V(x) = ax^4 \pm bx^2$. In recent years it has become increasingly clear that while a one-dimensional model of the ring-puckering motion yields reasonably good barrier height values, it does not allow for the interactions with other vibrational modes, which may often be significant.

Some of the most recent advances in the determination of potential surfaces governing ring inversions in four- and five-membered ring molecules have been in the development of two- and three-dimensional models. Such models then allow for the interaction of the ring puckering mode with other vibrational modes. These interactions have been shown to alter the puckering levels in the excited states of interacting motions and, if neglected, can result in poorly calculated barrier heights and misinterpretation of the far-infrared and Raman spectral features arising from the modes involved.

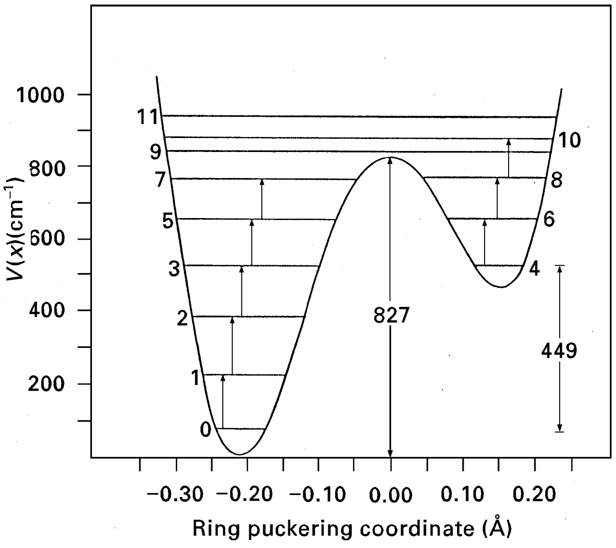



**Figure 4** Potential function governing the conformational interchange for axial and equatorial chlorocyclobutane.

## Skeletal bending nodes

The third type of low-frequency vibration is heavy-atom skeletal bending modes. Because the vibrations are of very low wavenumbers, there are frequently many excited states populated. Accordingly, the infrared contours in the gas phase can be very complex. In fact, many of these vibrations show a large number of excited states and the band contours are difficult to analyse because of the many vibrational–rotational transitions falling on top of each other. Attention to these types of vibrational modes has been fairly limited.

## Gases, liquids and solids

Another type of absorption in the low-wavenumber range is due to the pure rotation of small molecules. The intensity of this type of absorption is frequently $10^4$ times that of a low-wavenumber vibration, so very small amounts of HF, HCl, HBr, $H_2O$ (see **Figure 1**) or other similar impurities result in relatively strong absorption from these molecules. This type of far-infrared absorption was studied extensively from the 1930s to the 1960s. The high-resolution infrared data obtained for these molecules made it possible to determine very accurate centrifugal distortional constants.

Only limited studies have been carried out on the far-infrared absorption of molecules in the liquid state, which results in generalized absorption in the region of 30–80 $cm^{-1}$. However, if the molecule contains hydrogen bonding, a significant amount of information can be obtained for the low-wavenumber vibrations resulting from the absorption of dimers of such molecules. These include the wavenumber for the X···H–Y stretching and bending.

Physicists have primarily carried out the spectroscopic study of solids in the far-infrared spectral region. A major exception to this generalization is the study of the lattice modes of molecular crystals. To account for the lattice (phonon) modes of solids, the Bravais space cell is used by molecular spectroscopists to obtain the irreducible representation for the lattice vibrations. The crystallographic unit cell may be identical with the Bravais cell or it may be larger by a multiple of 2, 3 or 4. The relationship between the Bravais cell and crystallographic cell can be obtained from the Hermann–Mauguin X-ray symbol that is used to designate the crystal symmetry. For all crystal structures designated by a symbol *P* (primitive), the crystallographic unit cell and the Bravais unit cell are identical. Crystal structures designated with capital letters A, B, C or I are double primitive and, therefore, the crystallographic unit cells contain two Bravais cells. Crystal structures designated with the letters R or F are triply and quadruply primitive, respectively, and the crystallographic unit cells contain three and four Bravais cells, respectively.

The potential energy of a crystal can be considered to be made up of the following terms: $V_{\mathrm{Total}} = \Sigma_j V_j + \Sigma_i \Sigma_j V_{ij} + V_{\mathrm{L}} + V_{\mathrm{L}j}$ where $V_j$ is the potential due to the internal coordinates, $V_{ij}$ is the potential due to the correlation field, $V_{\mathrm{L}}$ is the potential for the external degrees of freedom and $V_{\mathrm{L}j}$ represents the potential due to the interaction of the internal modes with the lattice modes. Three symmetries must be considered when studying the vibrational spectrum of a crystal: the molecular symmetry, the site symmetry, and the factor group symmetry. The factor group is not a point group but it is isomorphous with the space group, which means there is a one-to-one correspondence between the two.

Oxamide $(OCNH_2)_2$ will be used as an example of the determination of the frequency of lattice modes. This molecule has molecular symmetry *trans* $C_{2h}$ with space group $P1$. Oxamide has one molecule per unit cell so there are three acoustical translations, which are inactive, and three optical librations, which are only Raman active. The three lattice bands are very prominent features in the Raman spectrum and occur at 106, 134, 157 $cm^{-1}$ (**Figure 5**) with corresponding bands at 100, 137, 156 $cm^{-1}$ for the deuterated compound. An estimate of the range for the force constant of the librations can be calculated associating the highest torsional wavenumber with the largest moment of inertia and the lowest wavenumber with the smallest moment of inertia. The widest possible range for the force constant is found to be 0.14 to 0.79 mdyn $Å^{-1}$. These values are relatively large for librational force constants and they reflect a high degree of hydrogen bonding for this compound.

For molecular crystals in which the molecules are nearly spherical there will frequently be a high-temperature crystal phase of which the site symmetry is higher than the molecular symmetry. These crystals are referred to as plastic crystals and they have unusual properties such as very high temperatures of melting and low anisotropic properties similar to those of liquids. The theory of lattice vibrations of orientationally disordered solids has been addressed, and the results indicate that such disorder should lead to a broadening of the vibrational bands and that all modes may be active in both the infrared and Raman spectra. Spectra of this type are referred to as density-of-states spectra, since the bands correspond to the flat points in the dispersion curve.

Ionic crystals are the other type of solid that has been studied extensively in the low-wavenumber

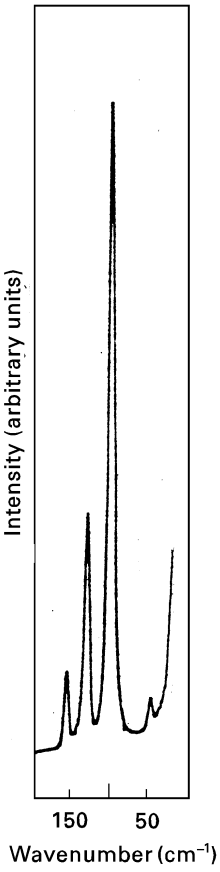



**Figure 5** Optical librational modes of oxamide.

spectral range. Of particular interests have been ferroelectrics, where one of the modes can shift from 150 cm$^{-1}$ to nearly zero as the transition temperature is approached.

## List of symbols

$h$ = Planck constant; $I_r$ = reduced moment of inertia; $V(x)$ = potential for conformational change; $V_n$ = height of $n$-fold rotational barrier; $V(\phi)$ = potential for rotation; $x$ = coordinate of ring inversion; $\phi$ = rotation angle.

*See also:* **IR Spectrometers**; **IR Spectroscopy, Theory.**

## Further reading

Carreira LA, Lord RC and Malloy TB Jr (1979) In: Eliel EL and Allinger NL (eds) *Topics in Current Chemistry*, Vol 82, pp 1–95. Berlin: Springer-Verlag.

Chantry GW (1971) *Submillimetre Spectroscopy; A Guide to the Theoretical and Experimental Physics of the Far Infrared*. London: Academic Press.

Compton DAC (1981) In: Durig JR (ed) *Vibrational Spectra and Structure*, Vol 9, pp **255**–404. Amsterdam: Elsevier.

Durig JR and Cox AW (1978) In: Ferraro JR and Basite LJ (eds) *Fourier Transform Infrared Spectroscopy*, Vol 1, pp 215–274. London: Academic Press.

Groner P, Sullivan JF and Durig JR (1981) In: Durig JR (ed) *Vibrational Spectra and Structure*, Vol 9, pp 405–487. Amsterdam: Elsevier.

Hehre WJ, Radom L, Schleyer PVR and Pople JA (1986) *Ab Initio Molecular Orbital Theory*. New York: Wiley.

Killough PM, Irwin RM and Laane JJ (1982) *Journal of Chemical Physics* **76**: 3890.

Martin DH (ed) *Spectroscopic Techniques for Far Infrared, Submillimetre and Millimetre Waves*. Amsterdam: North-Holland.

Mitra SS and Nudelman S (eds), *Far Infrared Properties of Solids*. New York: Plenum Press.

Möller KD and Rothschild WG (1971) *Far Infrared Spectroscopy*. New York: Wiley-Interscience.

Palik ED (1960) A far infrared bibliography. *Journal of the Optical Society of America* **50**: 1329.

Rounds TC and Strauss HL (1978) In: Durig JR (ed) *Vibrational Spectra and Structure*, Vol 7, pp 238–269. Amsterdam: Elsevier.

Wurrey CJ, Durig JR and Carreira LA (1976) In: Durig JR (ed) *Vibrational Spectra and Structure*, Vol 5, pp 176–220. Amsterdam: Elsevier.

# Fast Atom Bombardment Ionization in Mass Spectrometry

**Magda Claeys** and **Jan Claereboudt**,
University of Antwerp (UIA), Belgium

Copyright © 1999 Academic Press

MASS SPECTROMETRY
**Methods & Instrumentation**

## Introduction

The mass spectrometry community uses 1981 as the starting point of fast atom bombardment (FAB) when Barber and co-workers published their first paper. The basis of FAB was laid down in the mid-1970s through pioneering research on static secondary ion mass spectrometry (SIMS) by Benninghoven. At the time of its discovery FAB meant a significant breakthrough in the analysis of polar biomolecules, such as peptides. Until then other ionization techniques such as field desorption, SIMS using ion beams and $^{252}$Cf radiation, and laser desorption of ions from surfaces had been applied to the analysis of polar biomolecules with varying degrees of success and experimental difficulty. Since the development of FAB new powerful ionization techniques have emerged, namely matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI), which are routinely used nowadays and have greatly extended the applicability of mass spectrometry to polar biomolecules of high molecular mass ($M_r > 200\,000$). Despite the enormous successes of these last two ionization techniques, it is fair to state that FAB still enjoys an important place in bioanalytical laboratories involved in the analysis of molecules of medium polarity, including, for example, most natural products. A survey of the literature in a representative journal on natural product research, *The Journal of Natural Products*, for the years 1996 and 1997 revealed that of the articles in which mass spectrometric analysis was reported, FAB was used in 44% of the studies. The gas-phase ionization methods, electron impact (EI) and chemical ionization (CI), are still employed in 65% and 15% of the studies respectively, while the use of ESI and MALDI was rather limited (i.e. 5% and 0.7% respectively).

In FAB the sample ions are formed by bombardment of the sample in a liquid matrix with a high-energy beam of atoms (xenon or argon) or ions (caesium). The defining attribute of FAB is the use of a viscous liquid matrix to obtain long-lasting spectra and to 'soften' the sputtering process. The liquid matrix is able to provide continuous surface renewal so that intense primary beams may be used. The overall result is that secondary ion beams with a useful intensity for scanning mass spectrometers may be prolonged to periods of 20 min or more. Interestingly, the importance of the liquid matrix in FAB was not realized by Barber and co-workers when they discovered the technique. In their first article emphasis was laid on the use of fast atoms (argon) instead of ions as energetic bombarding particles and the term 'fast atom bombardment' was introduced to describe the new technique. However, similar spectra to those obtained using a fast neutral atom beam were achieved with better sensitivity using caesium or mercury ions as bombarding particles. The latter technique is known as 'liquid secondary-ion mass spectrometry' (LSIMS), but very often and also in this review the acronym FAB is used to refer to LSIMS.

## Instrumentation

The experimental set-up for FAB MS analysis is basically very simple and is illustrated in **Figure 1**. A beam of fast atoms is produced in a saddle field discharge source, called a 'FAB gun', by first ionizing an inert gas (Xe or Ar) to generate ions ($Xe^+$ or $Ar^+$) and accelerating these ions into an appropriate medium where the fast ions can capture an electron, thereby being converted from fast ions (having a kinetic energy due to acceleration) to fast atoms with energies as high as 10 keV. The beam of bombarding particles contains neutral atoms but also ions in various charge states. As illustrated in **Figure 1** the beam of bombarding particles is usually maintained at a large angle relative to the axis of the beam of secondary ions extracted by the ion optics. The intersection of the primary atom–ion beam and the secondary-ion beam is at the focal point of the ion optics. The sample film is mounted on a clean metal tip of the FAB probe which is introduced into the mass spectrometer through a vacuum-lock assembly so that the sample surface rests at the focal point of the instrument.

The experimental set-up for LSIMS analysis is very similar; the only difference is the replacement

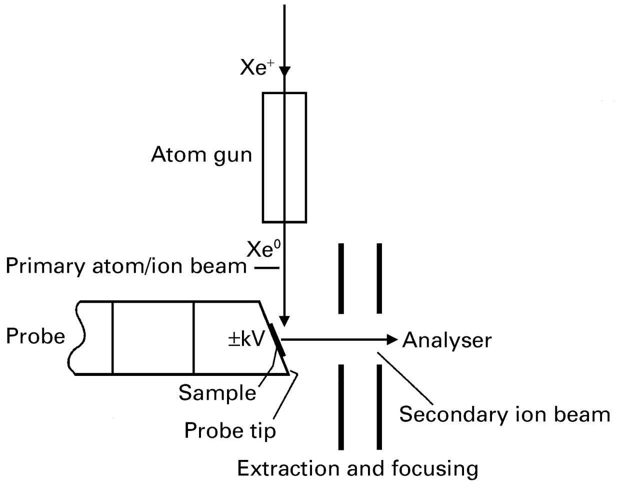



**Figure 1** Experimental set-up for FAB.

of the fast atom gun by a caesium ion source which produces a focused beam of fast $Cs^+$ ions (up to 35 keV). The caesium ions are thermionically emitted from an alkali aluminosilicate solid maintained at high temperature (approximately 1000°C). It is worth mentioning that caesium ion impact has largely replaced FAB in modern instruments. Advantages of the caesium ion source compared to a fast atom gun are that the energy distribution of the beam is narrow, beam focusing can be accomplished and the very low gas load introduced into the ion source.

## Mechanism of ion formation

It should be noted that whilst most of the particle-induced desorption techniques are simple to perform, they are only partially understood in theory, i.e. there exists no clear understanding of the mechanisms involved in ion formation from organic molecules in the condensed phase. It is, however, striking that different desorption ionization (DI) techniques [including FAB and LSIMS but also plasma desorption (PD) and laser desorption (LD)] produce reasonably similar mass spectra from nonvolatile organic molecules. Similarities in the spectra produced with incident beams of keV atoms or ions, or MeV particles and photons, can only be explained by similar ion formation and ion dissociation processes occurring after the initially very different physical excitation process. A general mechanism for the DI process can be presented, at least schematically, by the following sequence of events:

$$\begin{array}{lll} \text{Energy} & & \\ \downarrow & & \\ \text{Sample} & \rightarrow & \text{Intermediate} \\ \text{(condensed phase)} & & \text{(selvedge)} \\ \rightarrow \quad \text{Molecular ion species} + \text{fragment ions} & & \\ \qquad \text{(gas phase)} & & \end{array}$$

It is evident that the primary processes involved in the energy deposition step are dependent on the type of physical excitation used, and some of the properties of the intermediate phase must depend on the nature of these primary processes. However, the final step, leading to the production of molecular ions and fragment ions, appears to be independent of the primary processes occurring in the condensed phase. It is reasonable to propose that in the processes of ion formation and stabilization, chemical reactions strongly influence the end-products. These reactions are believed to take place in the vibrationally excited, disturbed surface layer, known as the 'selvedge' region. Several models describing the mechanism of FAB are available in the literature. It can be stated that there are almost as many models as there are theoreticians studying the problem.

### A unified model for desorption ionization

A comprehensive qualitative model for the DI process, as exemplified in FAB, LSIMS, PD and LD, has been proposed. The main features of this unified model are: (1) irrespective of its origin, the energy deposited at the sample surface is converted from its original form into vibrational energy; (2) desorption of intact molecules and preformed ions which can be described as a vibrational or thermal process (although no equilibrium is implied); (3) ion–molecule reactions (e.g. protonation, cationization and cluster ion formation) and EI occurring in the selvedge region; and (4) dissociation of energetic (metastable) ions well removed from the surface (i.e. in the vacuum region). A schematic illustration of this unified model is given in **Figure 2**.

In the unified model the processes of desorption and ionization are considered separately. Another characteristic of the model is that desorption is followed by chemical reactions of two types occurring in two distinct regions. First, in the selvedge region, fast atom(ion)–molecule reactions and EI can take place. Secondly, in the free vacuum, unimolecular dissociations occur which are governed by the internal energy

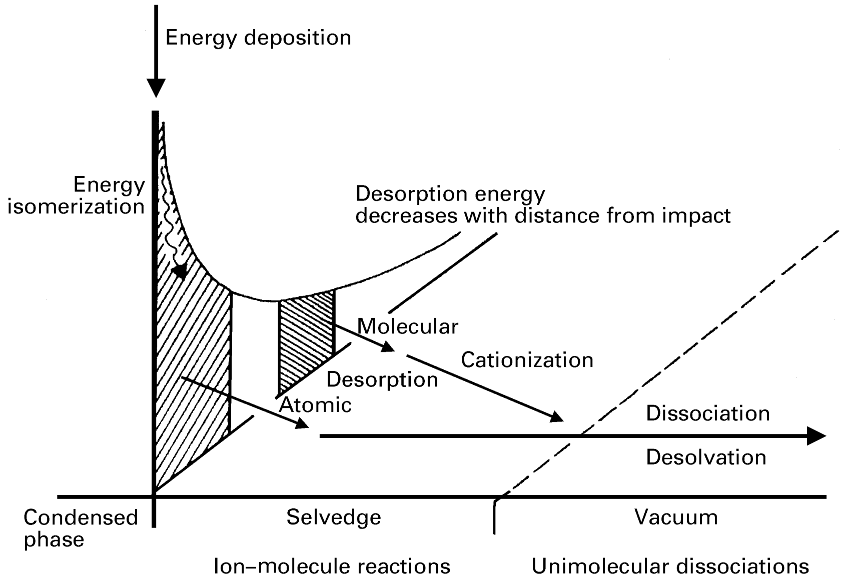



**Figure 2** Unified model for desorption ionization. Reprinted with permission from Cooks RG and Busch KL (1983) Matrix effects, internal energies and MS/MS spectra of molecular ions sputtered from surfaces. *International Journal of Mass Spectrometry and Ion Physics* **53**: 111–124. Copyright Elsevier Science.

of the parent ion, the characteristic timescale of the instrument and the structures of the gas phase ions.

## Ion formation mechanisms

The four major mechanisms for the formation of ions operating during FAB are: (1) direct desorption of preformed ions; (2) cationization (including protonation) and anionization in which a neutral analyte (M) is observed as adduct with a cation, i.e. $[M+Cat]^+$ or with an anion, i.e. $[M+An]^-$, (3) ion-beam induced processes leading to $[M–H]^+$ and $M^{+\bullet}$ ions; and (4) cluster ion formation. These four mechanisms cover most of the processes observed in FAB and largely describe the majority of ions detected in FAB spectra. Molecular-ion like species which are commonly encountered in FAB spectra are listed in **Table 1**.

Possible mechanisms for the formation of positive ions are given in **Table 2**; analogous mechanisms can be formulated for negative ions. That ion formation is the difficult, energetically demanding step in FAB is indicated by the observation that production of ions in the gas phase is much easier from samples that exist as preformed ions in the condensed phase than it is for neutral molecules. The mechanism by which preformed ions are converted into gaseous ions is conceptually the most straightforward: the molecular agitation generated by the bombarding particle, a process called 'sputtering', releases preformed ions into the gas phase. The types of compound that comprise preformed ions are inorganic salts, organic salts, and strong acids and bases. An artist's conception of the violent activity at the surface of a matrix solution of the analyte is presented in **Figure 3**.

The next type of ion for which the formation can be easily rationalized is the $[M+Cat]^+$ ions; $[M+Na]^+$ is used as a typical example. However, attachment of other metal ions, especially alkali ions, also readily occurs if these ions are present in the sample. $Na^+$ may be present in the sample as a salt impurity or it may be added deliberately by the analyst; it is a preformed ion which can attach to a neutral molecule (M) either in the gas phase or in the condensed phase. The $[M+H]^+$ ions can be generated by a mechanism analogous to that proposed for $[M+Na]^+$; in this case the preformed ion is the proton. Free

**Table 1** Molecular-ion like species in FAB spectra

| *Positive ions* | *Negative ions* |
|---|---|
| $Cat^+$ from salt $Cat^+$ $An^-$ | $An^-$ from salt $Cat^+$ $An^-$ |
| $[M + H]^+$ | $[M - H]^-$ |
| $[M + Cat]^+$, e.g. $[M + Na]^+$, $[M + K]^+$ | $[M + An]^-$, e.g. $[M + Cl]^-$ |
| $[M + 2Na - H]^+$ and analogues | $[M + Na - 2H]^-$ and analogues |
| $M^{+\bullet}$ | $M^{-\bullet}$ |
| $[M - H]^+$ | $[M + H]^-$ |
| Clusters: | Clusters: |
| $[Cat_{n+1} An_n]^+$ | $[Cat_n An_{n+1}]^-$ |
| $[M(M + H)]^+$, e.g. sample–sample or sample–matrix clusters | $[M(M - H)]^-$ |

**Table 2** Possible ion formation mechanisms in FAB

| *Ions formed* | *Mechanism* |
|---|---|
| Preformed ions | Sputtering |
| e.g. $Cat^+$ | $Cat^+An^-$(cond.) → $Cat^+$(gas) + $An^-$(gas) |
| $[M + Na]^+$ | Attachment of $Na^+$ in the gas phase |
| | $Na^+$ (cond.) → $Na^+$(gas) |
| | M(cond.) → M(gas) |
| | M(gas) + $Na^+$(gas) → $[M + Na]^+$(gas) |
| | Attachment of $Na^+$ in the condensed phase followed by sputtering |
| | M(cond.) + $Na^+$ (cond.) → $[M + Na]^+$(cond.) |
| | $[M + Na]^+$ (cond.) → $[M + Na]^+$(gas) |
| $[M + H]^+$ | Attachment of $H^+$ in the gas phase analogous to $Na^+$ attachment |
| | Attachment of $H^+$ in the condensed phase followed by sputtering, analogous to $Na^+$ attachment |
| | Disproportionation |
| | RO(H)••• HO(R) (cond.) → $ROH_2^+$ (gas) + $RO^-$ (gas) |
| $M^{+\bullet}$ | Disproportionation |
| | 2M(cond.) → $M^{+\bullet}$(cond.) + $M^{-\bullet}$(cond.) |
| | $M^{+\bullet}$(cond.) → $M^{+\bullet}$(gas) |
| | Fast atom–ion beam-induced reactions |
| | M(cond.) – $H^\bullet$ → $[M–H]^\bullet$ (cond.) |
| | $[M-H]^\bullet$(cond.) + $H^+$ (cond.) → $M^{+\bullet}$ (cond.) |
| | $M^{+\bullet}$(cond.) → $M^{+\bullet}$(gas) |
| $[M – H]^+$ | Fast atom–ion beam-induced reactions |
| | M(cond.) – $H_2$ → $[M–H_2]$ (cond.) |
| | $[M–H_2]$ (cond.) + $H^+$ (cond.) → $[M–H]^+$ (cond.) |
| | $[M–H]^+$(cond.) → $[M–H]^+$ (gas) |

cond. = condensed phase.

protons are probably produced during particle bombardment at the impact site, or can be added deliberately in the sample in the form of a strong acid. Another mechanism for the formation of protonated molecules which has been considered is a disproportionation reaction, i.e. perturbation in a hydrogen-bonded system caused by the bombarding particle can result in the sputtering of a $[M+H]^+$ ion and its complementary $[M-H]^-$ anion. True odd-electron molecular ions ($M^{+\bullet}$) are also formed, especially from compounds with low ionization potentials. $M^{+\bullet}$ ions are usually obtained for polyunsaturated molecules with conjugated systems, for example the fullerenes $C_{60}$ and $C_{70}$. The study of the mechanism of ion formation in FAB is still a research topic of interest. There is no firm evidence that $M^{+\bullet}$ ions are produced by direct EI. Current mechanistic studies focus on molecular hydrogen and hydrogen radical loss from protonated M which yield $[M–H]^+$ and $M^{+\bullet}$ ions respectively. $[M–H]^+$ and odd-electron $M^{+\bullet}$ ion formation in FAB was formerly attributed to gas-phase ionization-like processes but recent studies have demonstrated that fast atom–ion beam induced processes should be considered.

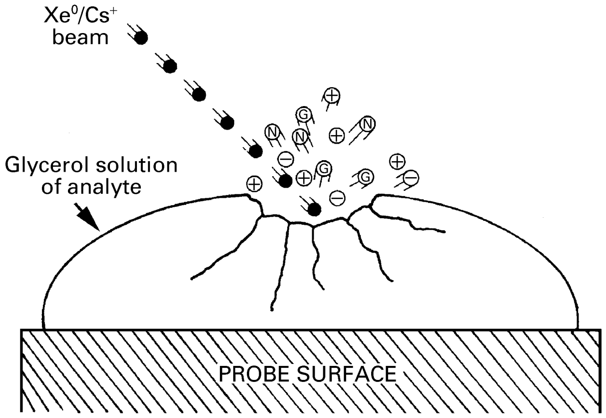



**Figure 3** Artist's conception of the violent activity taking place at the surface of a matrix solution during FAB.

The ion species of the types $[M+H]^+$, $[M–H]^-$ and $[M+Cat]^+$ are very useful for molecular mass determination, which is of key importance in the structure elucidation of natural products. A literature survey in *The Journal of Natural Products* for the years 1996 and 1997 revealed that in 52% of the studies in which FAB was used, additional accurate mass measurement at high resolution was performed on molecular ion species to obtain the precise molecular composition.

### Energy deposition

FAB is generally regarded as a 'soft' ionization technique but the question may be asked: what is meant by 'soft'? Molecular-ion like species formed during FAB show little fragmentation, suggesting that they are formed with a low internal energy. Energy deposition during FAB has been addressed by several workers and it has been established that ions are formed with energies varying between 1 and 4 eV, revealing a maximum at 1 eV and a high-energy tail. The classification 'soft' should, therefore, be used with some caution.

## Matrix selection and properties

The defining attribute of FAB is the use of a viscous liquid matrix. Consequently, many studies have been devoted to selection of suitable matrices in FAB. General matrix requirements concerning the solvent properties of the matrix are summarized as follows:

(1) the samples must be soluble in the matrix;
(2) only low vapour pressure solvents can be considered in the vacuum of the mass spectrometer;

(3) the matrix must be electrically conductive to avoid charging of the surface layer;
(4) ions from the matrix itself must not interfere with analyte ions in the FAB mass spectrum;
(5) the matrix must be chemically inert.

Lists of useful matrices for FAB are available and their physical and chemical properties have been compiled. In the authors' laboratory where FAB is still routinely applied to the analysis of plant secondary metabolites such as saponins, flavonoid glycosides, fatty acid derivatives and small synthetic peptides ($M_r < 3000$), two matrices are mainly used: glycerol, in the analysis of polar hydrophilic compounds and *m*-nitrobenzylalcohol (*m*-NBA) for lipophilic compounds. When glycerol is selected the sample is first dissolved in a cosolvent, methanol or a methanol–water mixture, while in the case of *m*-NBA dichloromethane is employed as cosolvent to facilitate addition of the sample to the matrix. It is a misconception that FAB can only be applied to the analysis of polar analytes; lipophilic compounds such as fatty acids and their derivatives are well amenable to FAB analysis if a lipophilic matrix is selected. Other matrices that have often been employed in peptide analysis include thioglycerol and a eutectic mixture of dithiothreitol and dithioerythritol (3:1, w/w), known as 'magic bullet'. For negative ion FAB the basic matrices di- and triethanolamine have also been used.

Evaporation of the liquid matrix in the vacuum of the mass spectrometer has to be considered because it can result in significant changes of the physical state of the sample solution. The effect of matrix evaporation on secondary ion formation has been investigated for samples dissolved in glycerol with and without a cosolvent. Depending on the residence time of the sample solution in the vacuum, the bulk and surface concentration of the analyte in the matrix can become supersaturated, analyte molecules can precipitate, and as such the conditions for secondary ion formation can be altered. Fortunately, the solid layers formed by precipitation of analyte molecules at the surface of sample solutions can be removed by sputtering, thereby making the underlying layer amenable to FAB analysis. In other models where evaporation of the liquid matrix was not taken into account, replenishment of the surface layer has been rationalized by diffusion of analyte molecules from the bulk to the surface layer. It is not likely that diffusion, which is not a rapid mechanism, is as important as previously thought to obtaining long-lasting analyte signals in FAB.

## Sample preparation for FAB analysis

Impure samples of biological origin cannot be directly submitted to FAB analysis. Polar hydrophilic samples are usually isolated from a biological matrix containing alkali salts by chromatographic procedures employing buffers. When buffers are required for the isolation, preference is given to systems composed of volatile salts, acids and bases. Alkali salts can be eliminated from samples by resorting to simple desalting procedures based on the use of reversed phase cartridges. In the presence of $Na^+$, for example, acidic compounds containing carboxylic or phenolic groups give rise to multiple sodiated molecular species (i.e. $[M + Na]^+$, $[M - H + 2Na]^+$ and $[M - 2H + 3Na]^+$) leading to decreased detection sensitivity. In order to improve the detection of peptides it is common practice to add a volatile strong acid such as trifluoroacetic acid to the analyte–matrix mixture whereby neutralization of the carboxylate part of the peptide zwitterionic structure results in an enhanced $[M + H]^+$ ion formation. Derivatization strategies to increase the solubility of the analyte in the liquid matrix or to convert the analyte into a salt form with better desorption properties may also be considered. Molecules containing labile groups which are prone to hydrolysis in FAB can be stabilized by using a $Li^+$-containing matrix.

## Continuous-flow FAB

FAB is the basis for an effective coupling technique for liquid chromatography (LC), namely, continuous-flow FAB. A set-up for continuous-flow FAB is given in **Figure 4**. In this technique, a liquid flow of typically 5–10 μL $min^{-1}$ obtained by splitting the LC flow, is introduced into a heated FAB source via a narrow-bore fused silica capillary. The glycerol matrix is added to the LC effluent to a concentration of approximately 0.5% and subsequently the mixture is directed into the FAB source through a metallic frit. The liquid flow ensures that there is a continuous flow on the probe tip in which previously eluted sample is continuously removed from the area where sputtering occurs. In other designs there is no metal frit but then a cotton wick is used to disperse the solvent and to obtain stable operating conditions. In addition to coupling to LC, continuous-flow FAB has been shown to be a useful technique for the introduction of flow-injected samples and effluents of capillary electrophoresis and microdialysis.

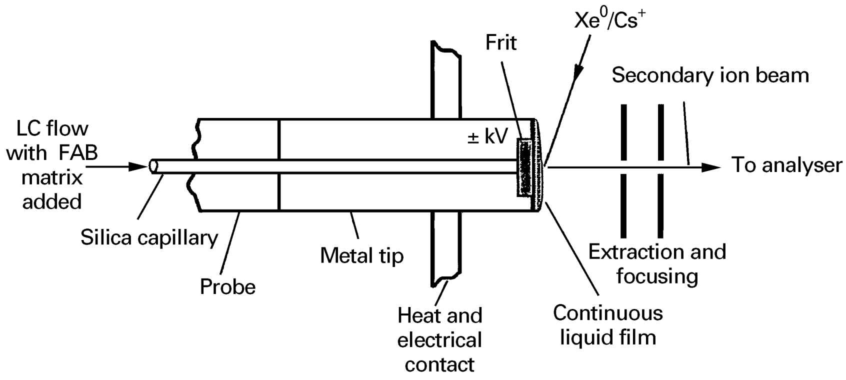



**Figure 4** Experimental set-up for continuous-flow FAB.

## Combination of FAB with collision-induced dissociation and tandem mass spectrometry

In order to obtain structural information on biomolecules it is common practice to use FAB in combination with collision-induced dissociation (CID) and tandem mass spectrometry. As discussed above FAB is a soft ionization technique producing mainly molecular-ion like species with a low internal energy. These species can be fragmented by CID during which the internal energy of the ions is increased. CID can be performed at low and high collision energies generally leading to different and complementary structural information. The FAB spectrum of compounds of biological origin is usually quite complex for several reasons: peaks due to the FAB matrix are always present but, in addition, impurities and salt forms of the molecular ions are often encountered making the interpretation of a FAB mass spectrum particularly difficult. Using CID and tandem mass spectrometry it is possible to obtain a spectrum of one well-defined molecular ion species allowing the corresponding fragment ions to be determined unambiguously. **Figure 5** illustrates first-order FAB spectra using glycerol as liquid matrix obtained for a flavonoid glycoside, kaempferol-3-*O*-rutinoside, which was isolated from the leaves of an African medicinal plant *Morinda morindoides*, with and without subsequent desalting. Without desalting several molecular ion species of kaempferol-3-*O*-rutinoside are observed with low intensity, namely, $[M + H]^+$, $[M + Na]^+$, $[M - H + 2Na]^+$ and $[M - 2H + 3Na]^+$. The presence of $Na^+$ in the sample is also evident from the matrix peaks including abundant sodiated species. Using desalting a sufficiently intense $[M + H]^+$ signal could be obtained which was amenable to CID and tandem mass spectrometry. This methodology showed that the ions at $m/z$ 287, 449 and 461 are fragment ions of the protonated molecule.

## Selected applications

A number of reviews and overviews are available in the literature. For comprehensive surveys of the literature the biennial reviews of mass spectrometry in the journal *Analytical Chemistry* can be consulted. As it is impossible to review all applications of FAB only selected applications will be mentioned here. FAB has been particularly useful for the analysis of peptides and proteins, providing molecular mass information for large peptides with an upper mass limit of approximately 24 000 Da. It can be stated that in the field of peptides and proteins FAB has now largely been overtaken, but not entirely replaced by, MALDI and ESI. Other biomolecules that have been successfully analysed using FAB are glycoconjugates, including glycopeptides, nucleotides, terpenoid and flavonoid glycosides, and complex lipids such as glycosphingolipids, gangliosides, phospholipids and steroids. The relatively frequent use of FAB compared to other soft ionization techniques (i.e. CI, ESI and MALDI) in natural product research can be attributed to the reliable molecular mass and composition that can be obtained which is of key importance in the structure characterization of unknown compounds. FAB has also been applied in areas that are not of direct biochemical interest, such as the analysis of dyestuffs, organometallics, quaternary ammonium salts, triphenyl phosphonium salts, surfactants and chiral complexes. With regard to this last application FAB has been found to be more suitable than ESI for studying enantioselective intermolecular interactions.

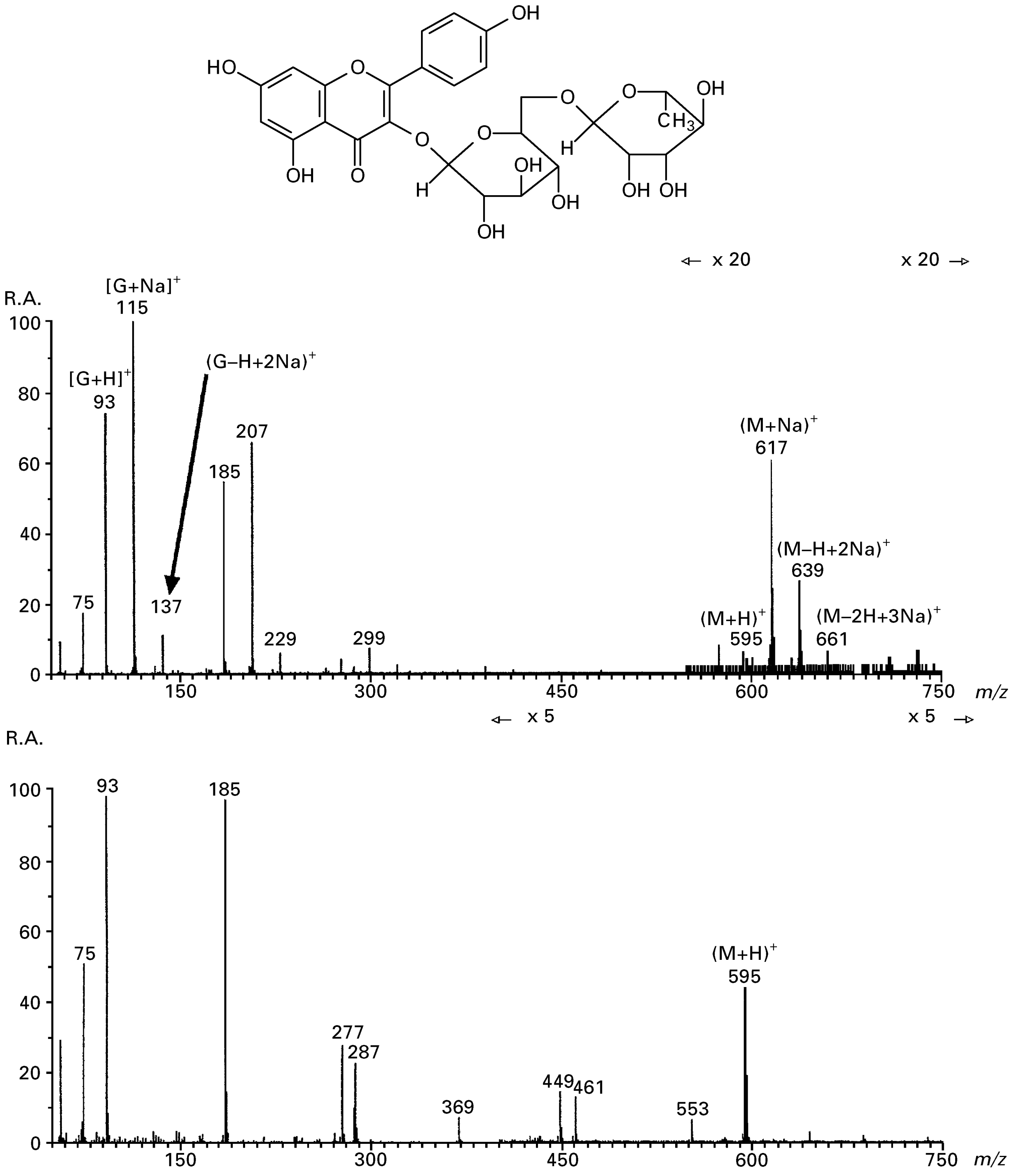



**Figure 5** First-order FAB spectra obtained using glycerol (G) as the liquid matrix for kaempferol-3-*O*-rutinoside, isolated from a medicinal plant, with and without subsequent desalting. Reprinted with permission from Li QM, Dillen L and Claeys M (1992) Positive ion FAB analysis of flavonoid glycosides. Simple procedures for desalting and control of sodium salt contamination. *Biological Mass Spectrometry* **23**: 408–410. Copyright John Wiley and Sons Ltd.

*See also:* **Biochemical Applications of Mass Spectrometry; Fragmentation in Mass Spectrometry; Ion Energetics in Mass Spectrometry; Ionization Theory; IR Spectroscopy Sample Preparation Methods; MS-MS and MS*n*; Organometallics Studied Using Mass Spectrometry; Peptides and Proteins Studied Using Mass Spectrometry; Plasma Desorption Ionization in Mass Spectrometry; Spectroscopy of Ions.**

## Further reading

Barber M, Bordoli RS, Sedgwick RD and Tyler AN (1981) Fast atom bombardment of solids (F.A.B.): a new ion source for mass spectrometry. *Journal of the Chemical Society, Chemical Communications* 325-327.

Barber M, Bordoli RS, Elliott GJ, Sedgwick RD and Tyler AN (1982) Fast atom bombardment mass spectrometry. *Analytical Chemistry* **54**: 645A-657A.

Caprioli RM and Suter MJF (1992) Continuous-flow fast atom bombardment: recent advances and applications. *International Journal of Mass Spectrometry and Ion Processes* **118/119**: 449-476.

Claeys M, Li QM, Van den Heuvel H and Dillen L (1996) Mass spectrometric studies on flavonoid glycosides. In: Newton RP and Walton TJ (eds) *Applications of Modern Mass Spectrometry in Plant Science Research*, pp 182-194. Oxford: Oxford Science Publications.

Cook KD, Todd PJ and Friar DH (1989) Physical properties of matrices used for fast atom bombardment. *Biomedical and Environmental Mass Spectrometry* **18**: 492-497.

Cooks RG and Busch KL (1983) Matrix effects, internal energies and MS/MS spectra of molecular ions sputtered from surfaces. *International Journal of Mass Spectrometry and Ion Physics* **53**: 111-124.

De Pauw E (1986) Liquid matrices for secondary ion mass spectrometry. *Mass Spectrometry Reviews* **5**: 191-212.

De Pauw E, Agnello A and Derwa F (1991) Liquid matrices for liquid secondary ion mass spectrometry-fast atom bombardment: an update. *Mass Spectrometry Reviews* **10**: 283-301.

Gower JL (1985) Matrix compounds for fast atom bombardment mass spectrometry. *Biomedical Mass Spectrometry* **12**: 191-196.

Hemling M (1987) Fast atom bombardment mass spectrometry and its application to the analysis of some peptides and proteins. *Pharmaceutical Research* **4**: 5-5.

Junker E, Wirth KP and Röllgen FW (1992) Effects of matrix evaporation during continuous sputtering in fast atom bombardment. *International Journal of Mass Spectrometry and Ion Processes* **122**: 3-23.

Li QM, Dillen L and Claeys M (1992) Positive ion FAB analysis of flavonoid glycosides. Simple procedures for desalting and control of sodium salt contamination. *Biological Mass Spectrometry* **23**: 408-410.

Sawada M (1997) Chiral recognition detected by fast atom bombardment mass spectrometry. *Mass Spectrometry Reviews* **16**: 73-90.

# Fibre Optic Probes in Optical Spectroscopy, Clinical Applications

**Urs Utzinger** and **Rebecca R Richards-Kortum**,
The University of Texas at Austin, TX, USA

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Applications**

## Introduction

In the clinical environment, optical techniques have been used for decades (microscope, colposcope, ophthalmoscope, endoscope, laparascope). Integration of spectroscopic devices into existing procedures is an obvious task. Fibre optic cables provide a flexible solution for adequate optical interfacing between the optical and spectroscopic device and the sample to be interrogated *in situ*. Fibre optic probes can be inserted into cavities and tubular structures, put in contact with epithelial surfaces and also inserted into structures that can be punctuated by rigid devices such as needles. Fibre optic devices for optical spectroscopy can be manufactured as flexible catheters with an outer diameter not exceeding 0.5 mm. We will present in this article fibre optic solutions for fluorescence and both elastic and inelastic scattering detection. After describing the fibre optic interface, we will present probes for fluorescence spectroscopy followed by probes for reflectance measurements and side looking. Diffuser tips, refocusing and designs for Raman spectroscopy are discussed towards the end of the article.

## The fibre optic interface

A spectroscopic system incorporates a light source, an optical analyser with detector, and a light transport conduit which in many cases is made of fibre optic cables. A separate illumination and collection channel minimizes background signals produced in the illumination fibre (**Figure 1A**). The excitation or illumination light source is typically a laser or a filtered

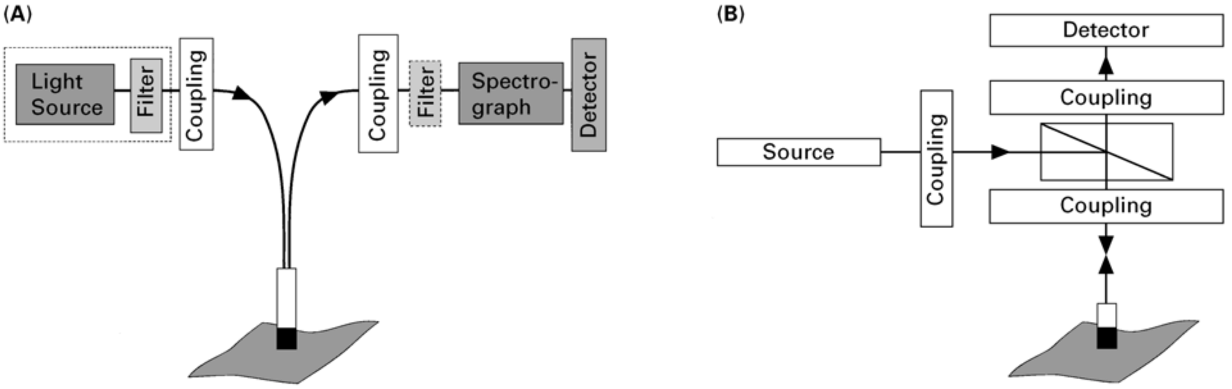



**Figure 1** (A) A fibre optic spectroscopy system with separate illumination and collection path is based on an excitation source, which is a laser or a white light source (reflectometry) or a monochromator filtered arc lamp (fluorescence). Optics couple the excitation light into the flexible probe. A probe collects the emitted light. Coupling optics adapt the numerical aperture of the probe to the spectrograph or filter system. An optical detector (charge coupled device (CCD), photodiode array, photomultiplier tube) is read out and digitized. (B) A fibre optic spectroscopy system with a probe that incorporates one optical fibre needs a dichroic beam splitter and well aligned optics to separate excitation and fluorescence light. Reproduced with permission of Optical Society of America Inc. from Greek LS, Schulze HG, Blades MW, Haynes CA, Klein K-F and Turner RFB (1998) Fiber-optic probes with improved excitation and collection efficiency for deep-UV Raman and resonance Raman spectroscopy. *Applied Optics* **37**(1).

white light source such as xenon or mercury lamps. Dielectric bandpass filters (Omega Optical, Inc., Brattleboro, VT, USA), monochromators or double monochromators (ISA Edison, NJ, USA) can be used as filters according to the need for spectral purity. For fluorescence spectroscopy, xenon lamps (ORC, Azusa, CA, USA | Hamamatsu Corp. Bridgewater, NJ, USA) with more than 150 W power consumption need additional cold and hot mirrors to protect the optical parts. The coupling optics adapt the f-number of the light source to the numerical aperture of the fibre and guarantee optimal irradiance into the fibre. The probe transports the collected light to the spectroscopic system. New techniques such as holographic transmission gratings (Holospec, Kaiser Optical Systems Inc., Ann Arbor, MI, USA) and back-illuminated thinned CCDs with high quantum efficiencies (Princeton Instruments, Inc., Trenton, NJ, USA) allow short integration times and sufficient spectral and spatial resolution. Additional filter stages reduce the influence of stray light in the spectrograph. Most fluorescence applications require longpass (Schott Glass Technologies Inc., Duryea, PA, USA) and Raman spectroscopy devices notch filters (Holographic notch filters, Kaiser Optical System Inc).

To achieve smallest probe diameters, single fibre solutions are used in combination with a dichroic beam splitter and with well-aligned coupling optics (**Figure 1B**). Single fibre solutions are limited by the difficulty of reducing back-scattered excitation light and the suppression of parasitic light induced by the illumination path of the probe.

## Fibre optic cables

An optical fiber for spectroscopy consists of a silica core, a doped cladding and a protective jacket. Light is transmitted based on the principle of total internal reflection. Half of the maximal angle ($\alpha$) a fibre can accept light is characterized by the numerical aperture (NA), which is defined by the difference in the refractive indices (n) of core and cladding (**Figure 2**):

$$\mathrm{NA} = n_{\mathrm{media}} \sin\alpha = \sqrt{n_{\mathrm{core}}^2 - n_{\mathrm{cladding}}^2}$$

Improvements in the preform manufacturing (Heraeus Amersil Inc., Duluth, GA, USA) allow

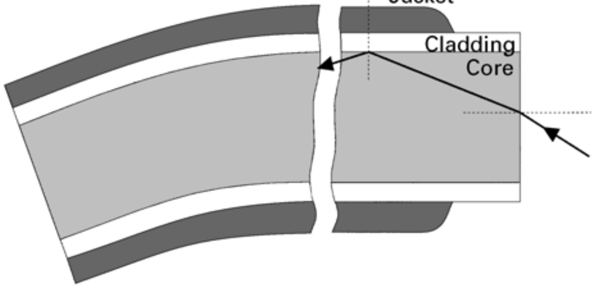



**Figure 2** A fibre optic cable for spectroscopy consists of a core material (quartz) and a cladding with a lower refractive index (doped quartz) and a rugged supportive jacket. Light is transported by total internal reflection.

transmittance from 200 nm (solarization-resistant-UV grade fibre, Polymicro Technologies, Inc, Phoenix, AZ, USA) up to 2500 nm (low-OH fibre, Fiberguide Industries, Inc., Stirling, NJ, USA | Ceramoptec, Bonn, Germany). This opens the application of fibre optic probes to UV resonance Raman (UVRR) and IR Raman spectroscopy. Due to bending of the fibre and scattering centres in the fibre, light may exit into the cladding and hit the jacket. Most plastic claddings like Nylon and Tefzel® (DuPont, Wilmington, NJ, USA) produce significant autofluorescence when irradiated with UV light. Polyimide (Thermocoat, Fiberguide Industries Inc.) and metallic coated fibres (gold, aluminium) exhibit minimal fluorescence. Colour centres may be produced during UV radiation in quartz and the fibres fluoresce in a broad band at 450 and 650 nm. This process is reversible but depends on the pulse repetition rate and the pulse energy. It is further known that silica produces Raman intrinsic signals at NIR excitation, which interfere with in vivo Raman spectroscopy such that the dynamic range of the detector becomes critical.

## Fluorescence probes

A growing number of clinical studies have demonstrated that fluorescence spectroscopy can be used to distinguish normal and abnormal human tissues in vivo in the skin, head and neck, genito-urinary tract, gastrointestinal tract, breast and brain. It is well known that fluorescence intensity and lineshape are a function of both the excitation and emission wavelength in samples containing multiple chromophores, such as human tissue. Major fluorescence contributors are structural proteins, NADH, FAD, tryptophane and porphyrins.

### Single pixel measurements

The classic fibre optic probe to measure fluorescence consists of at least one excitation and one collection fibre (**Figure 3A**). A quartz shield placed at the distal end of the fibres allows the illuminated and probed areas to overlap. The fraction of overlapping increases with an enlargement of the numerical aperture of the fibres and the thickness of the shield. A larger shield depth requires a larger diameter shield. The collection efficiency $\beta_t$ for an isotropic fluorescence source can be described as:

$$\beta_t = \frac{\beta_0}{(z + \mathrm{shield}_t)^2}$$

where z is the position along the optical axis of the probe and $\beta_0$ a constant which includes the detector efficiency. Previous studies performed in arterial tissue have demonstrated empirically that tissue does not emit isotropically, as a result of high forward scattering. For arterial specimens, tissue fluorescence power decreases with the detector-sample separation distance, R, as $1/R^n$, where $n = 1.1$. A typical thickness of the shield is 1.7-6.4 mm. However, scattered fluorescence could also be detected without a shield. A typical probe consists of excitation fibres, collection fibres, carbon filled epoxy and tubing. A rigid type uses metallic tubing, a flexible type uses shrinking tubing (Zeus Industrial Products Inc., Orangeburg, SC, USA). The shield and sleeve are detachable to allow disinfection of the probe. A probe based on seven 200 µm fibres has a diameter of at least 1.5 mm. The probe with an outer diameter of 4 mm presented in **Figure 3B** has been successfully used by Ramanujam and the authors in studies with more than 100 patients.

InnovaQuartz (InnovaQuartz, Phoenix, AZ, USA) has successfully manufactured fibre optic tips, using $CO_2$ laser micro machining (**Figure 3C**). Melting and compressing the fibres eliminates the dead space between them. An increase of power density and almost complete overlap between illumination and collection areas has been reported. However, this technique is limited to fibre bundles with a few fibres only.

### Multi-pixel measurements

Pitris, Agrawal and co-workers have successfully adopted the principle of single pixel fluorescence measurements into a design of 31 simultaneous measurement locations (**Figure 4A**, **Figure 4B**) without sacrificing spectral resolution. Modern imaging spectrographs for fluorescence applications (Instruments SA, Edison, NJ, USA | Chromex, Albuquerque, NM, USA) allow the simultaneous spectral dispersion of up to 30 input channels with minimal cross talk. Aberration corrected spectrographs project the input slit onto CCD devices with a focal plane size of $30 \times 12$ mm (TRIAX, Spectrum One MCR, Instruments SA). Spot scanning and imaging devices increase the spatial resolution further at the cost of spectral resolution. **Figure 4A** illustrates the hexagonal arrangement of the fibre pairs. Adding excitation fibres can increase the illumination intensity. With lamp based excitation sources the focal spot size at the coupling site is determined by the arc size, which is typically in the order of 0.5-3 mm. A 300 W xenon lamp with parabolic reflector (Compact

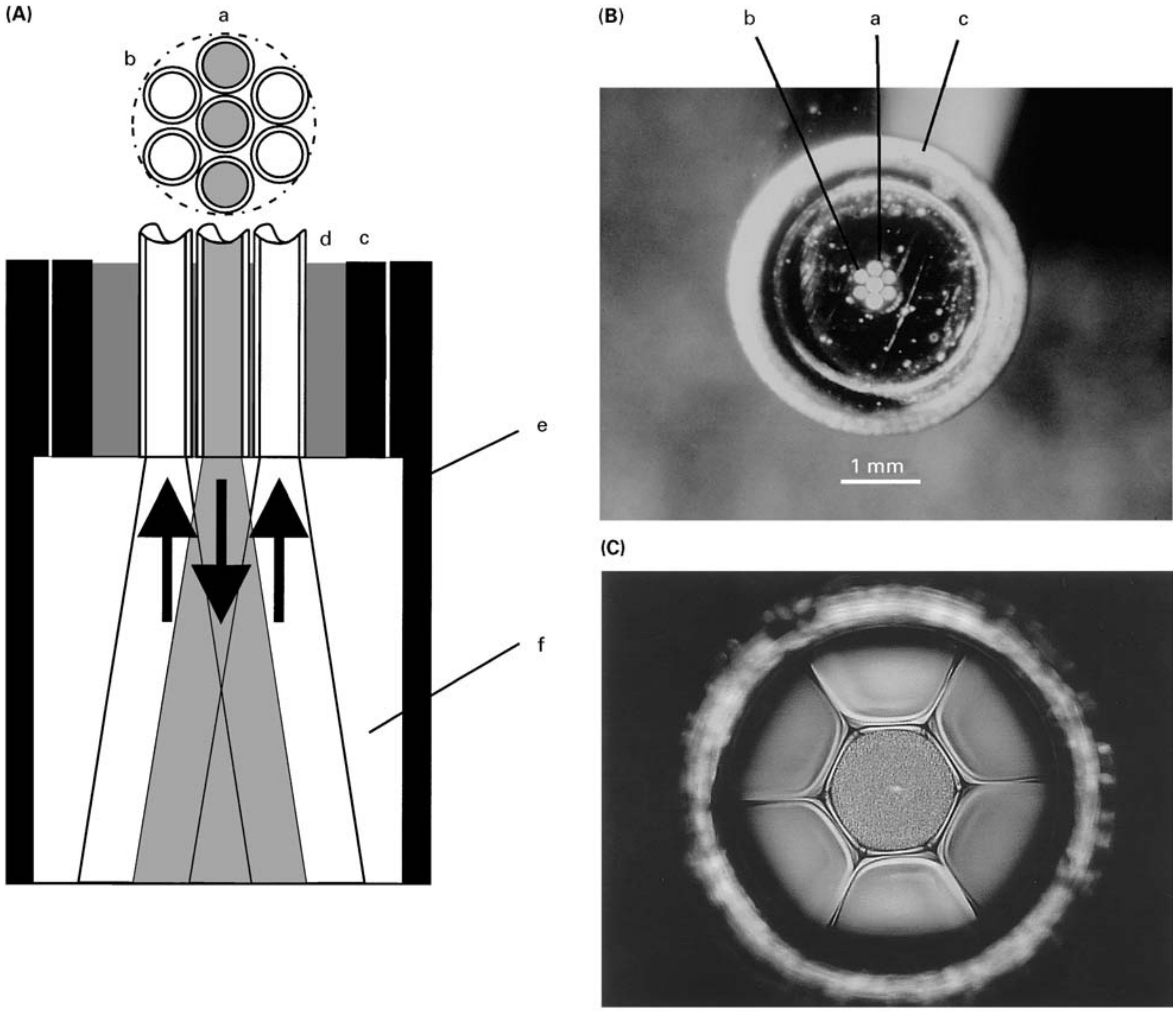



**Figure 3** (A) The classic fibre optic probe for fluorescence spectroscopy is based on two or more fibres. Hexagonal packing is a dense arrangement and allows the selection of multiple fibres for different excitation sources and collection channels. A quartz shields permits the overlap of excitation and collection areas. The parts are: (a) excitation fibres; (b) collection fibres; (c) tubing; (d) carbon filled glue; (e) sleeve; (f) shield. (B) A typical fluorescence probe for single point sampling; (a) excitation fibres; (b) collection fibres; (c) sleeve. (C) A hexagonal fibre bundle with a melted and compressed arrangement allows excellent illumination and interrogation overlap. Reproduced by permission of Innova Quartz Inc.

Illuminator 6000CI, ORC) produces a spot diameter of approximately 7 mm after a lens with a focal length of 50 mm. According to the formula for hexagonal packing:

$$n_{\mathrm{fiber}} = 1 + \sum_{k=0}^{m} 6k$$

where m is the total number of rings, 547 fibres with a core diameter of 200 μm can be successfully located in this spot which would lead up to 17 excitation fibres per fluorescence collection fibre. **Figure 4B** shows a multi-pixel probe, which is used by the authors for fluorescence distribution studies. The outer diameter is 25 mm and the shield thickness 6.4 mm.

## Separation of illumination and detection spot

Keijzer and co-workers described scattering of fluorescence outside the illumination spot experimentally and theoretically (**Figure 5**) for arterial tissue. $\beta$ and $\alpha$ absorption bands of oxygenated haemoglobin alter the fluorescence spectra. This effect is evident when the illumination and collection

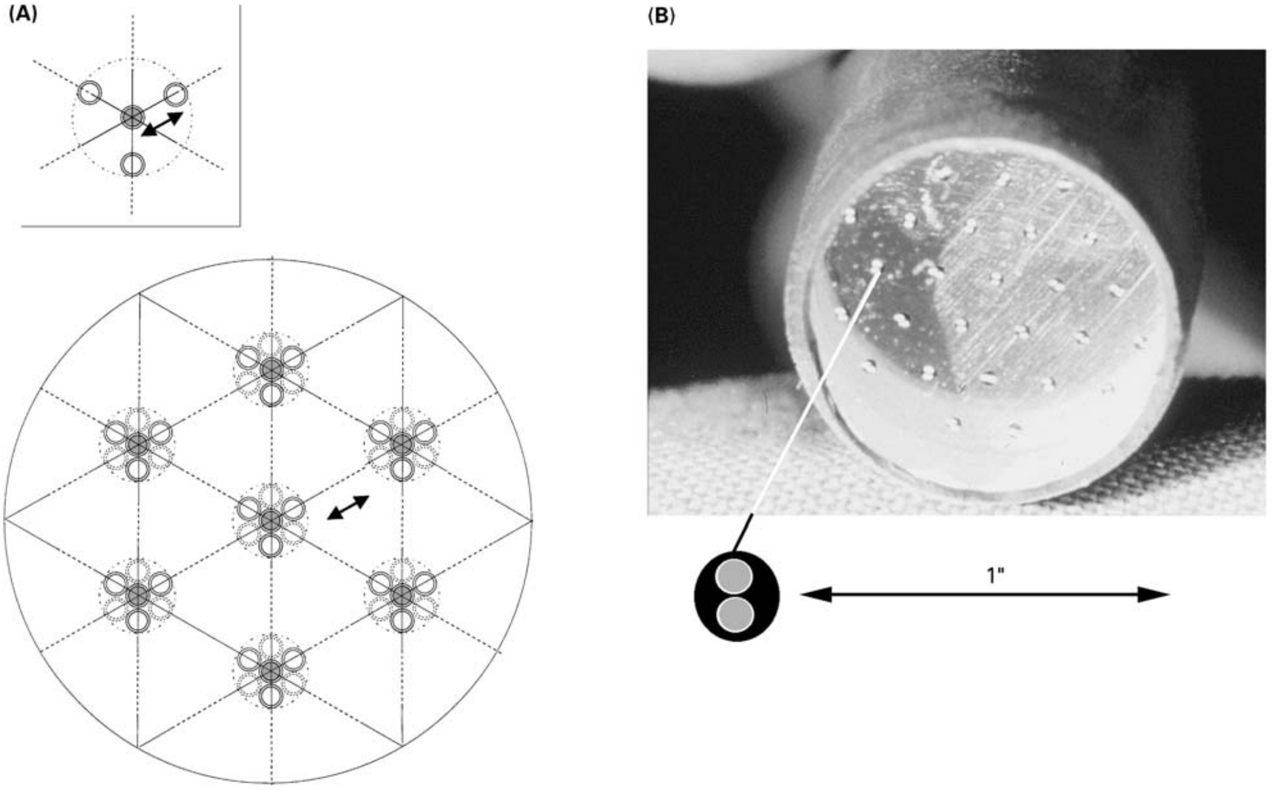



**Figure 4** (A) A multi-pixel fibre optic design based on the design of **Figure 3A**. At each sample spot one collection fibre is surrounded by one to six illumination fibres. Signal tunneling from in between sample spots can be influenced by the separation of the collection and emission fibres and the shied thickness. (B) A typical fluorescence probe for simultaneous multi-pixel measurements. Black shrinking tubing holds the quartz shield. Emission and collection fibre pairs are arranged hexagonally.

spot do not overlap so the spacing of the collection spots and the shield thickness have to be chosen carefully. In order to enhance the absorption influence on fluorescence spectra, the excitation and collection fibres need to be further separated (**Figure 4A**).

## Reflectance probes

The multiply scattered light that escapes the sampling volume is called reflectance. The transport mean free path length of a photon is defined as:

$$\mathrm{mfp}' = \frac{1}{(\mu_a + \mu_s)}$$

where $\mu_a$ is the absorption coefficient and $\mu_s'$ is the reduced scattering coefficient which is the isotropic equivalent of the anisotropic scattering medium:

$$\mu_s' = \mu_s(1-g)$$

where g is the average cosine of the scattering angle and $\mu_s$ the scattering coefficient. The optical properties $\mu_s$, $\mu_a$ and g depend on the chemical and architectural composition of the sample. The chemical composition of tissue depends on the blood supply, the metabolic state and the tissue type. The architecture of the sample is reflected in the shape and size distribution of the scattering particles. All these characteristics vary spatially and are wavelength dependent. Several recent studies have suggested that differences in the optical properties, assessed using diffuse reflectance spectroscopy, can be used to discriminate normal and abnormal human tissues in vivo in the urinary bladder, pancreas and the skin.

The goal of fibre optic probes for reflectometry is to measure the reflectance distribution and to derive the optical properties by fitting the data to analytic expressions based on diffusion theory or iterative algorithms based on Monte Carlo modelling. Therefore fibre optic probes consist of a single excitation source and several spatially distributed collection fibres (**Figure 6A** and **Figure 6B**). A probe with a linear alignment constructed by Wang and co-workers

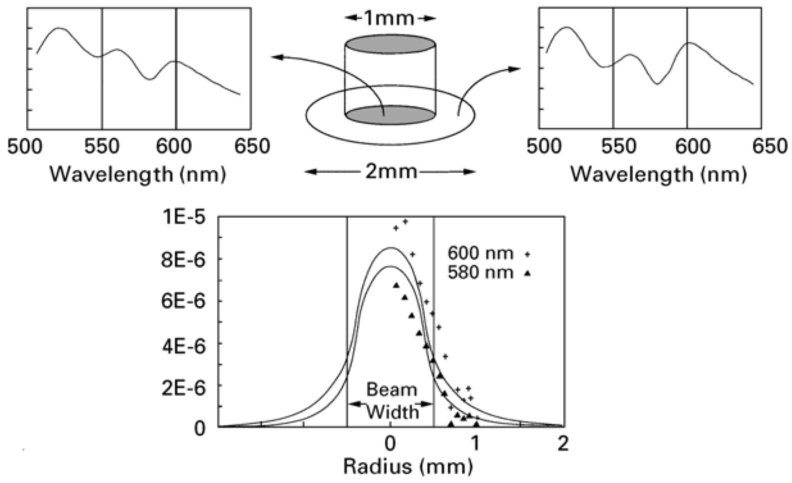



**Figure 5** Cross-talk. Fluorescence generated in a 1 mm spot area is detectable from an outside area. It spreads by scattering processes and will be reshaped by absorption processes. Monte Carlo simulations and measured data show that haemoglobin absorption in tissue affects emission intensities at 580 nm more than at 600 nm. Adapted with permission from Keijser M, Richards-Kortum RR, Jacques SL and Feld MS (1989) Fluorescence spectroscopy of turbid media: Autofluorescence of the human aorta. *Applied Optics* **28**(20): 4286–4292.

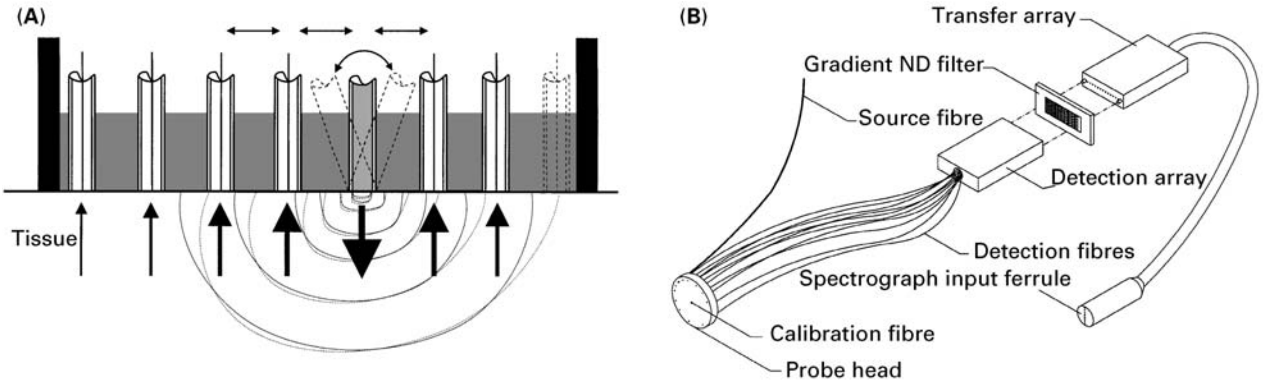



**Figure 6** (A) A fibre optic probe for reflectometry. The light scattered from a single excitation fibre is detected by a linear array of collection fibres. Tilting the excitation fibre shifts the profile along the surface by an amount that is determined by the sample's optical properties. (B) A fibre optic probe with a circular fibre arrangement for reflectometry. A 2 cm probe head consist of a central calibration fibre, an excitation fibre and non-equally spaced fibres. Neutral density (ND) filters adapt the light intensity in a transfer array to a smaller dynamic range.

places the detection fibres over a range of ±1 to 10 mfp′ (**Figure 6A**). Nichols and co-workers use a circular arrangement with a similar range (**Figure 6B**). This results in probes with a diameter of approximately 2 cm and an assumption that the optical properties do not vary over this range. However the optical properties on currently investigated target areas (e.g. cervix, ovaries, oral cavity) vary over orders of millimetres (lesion size) and their size is similar to the diameter of the proposed probes. Therefore an accurate measurement of optical properties from small, premalignant lesions is difficult.

Wang and co-workers modified their linear probe design with an oblique incident source fibre and measured the shift of the reflectance profile. This removes the necessity of measuring in absolute units. The circular fibre arrangement of Nichols and others allows a simple calibration of the system by placing a source fibre in the centre of all fibres (**Figure 6B**). Measuring the spectrally resolved reflectance of all fibres simultaneously requires a dynamic range of 4 orders of magnitude. Neutral density filters or other techniques reduce the dynamic range required.

### Combined probes for fluorescence and reflectance spectroscopy

It is well known that the absorption and scattering properties of tissues in vivo affect both the intensity

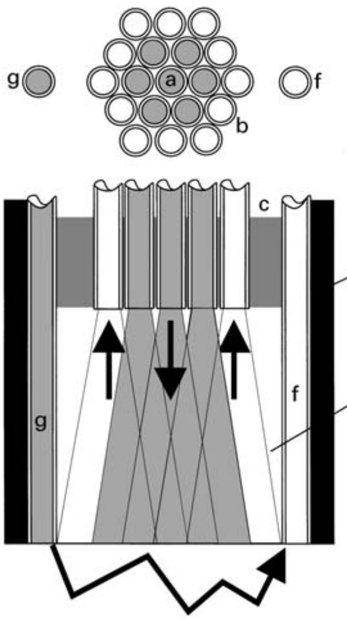



**Figure 7** A fibre optic probe that combines fluorescence and reflectance measurements. White light tunnels through the fluoresced sample volume; (a) excitation fibres; (b) fluorescence collection fibres; (c) carbon filled glue; (d) outer tubing; (e) quartz shield; (f) white light illumination; (g) reflectance collection.

and the lineshape of measured fluorescence spectra. Furthermore, measuring both fluorescence and diffuse reflectance spectra may provide additional information of diagnostic value.

A probe that combines the measurement of fluorescence and optical properties was described by Durkin. Scattered light is collected with a white light 'transmission' measurement which travels through the same tissue volume that was excited for fluorescence measurements.

The probe, illustrated in **Figure 7**, consists of a total of 21 optical fibres (200 µm diameter, NA = 0.2) arranged in concentric bundles. The centre bundle contains seven fluorescence excitation fibres and twelve fluorescence collection fibres. At the distal end of the probe the fibres which excite and collect fluorescence are sealed with a quartz shield. This shield is placed in contact with the sample surface, and ensures that the area from which fluorescence is collected is the same as that directly illuminated. The reflectance fibres are flush with the tip of the central shield. Similar attempts have been previously described in the literature.

## Side-looking probes and their application

Oblique polishing of individual fibres deflects the output of the fibre with respect to the fibre axis (**Figure 8A**). If the critical angle for total internal reflection is reached, the light will leave the fibre through the cylindrical side. The shape of the fibre wall focuses the beam in an angular direction resulting in an elliptical focal spot close to the fibre surface. The critical angles for the silica- water and the silica- air interface are 66° and 43.3° respectively. In order to permit the light to leave the fibre sideways at the tip, the jacket needs to be stripped. Depending on the jacket material, it can be dissolved in acid, mechanically abraded or burned away with a lighter.

Illumination analysis (ASAP, Breault Research Organization Inc., Tucson, AZ, USA) shows maximal irradiance increase at a polishing angle of 40° for water and air as surrounding media. The focal position is at a distance of 1.3 times the cladding radius positioned away from the side of the fibre in air and 3.17 times the radius in water. The increase of irradiance is 1.6 in water and 2.36 in air (**Figure 8B**).

Many probe designs are based on the oblique polishing of fibres (InnovaQuartz | Gaser, Visonex Inc., Warner Robins, GA, USA). **Figure 8C** (left) shows a design where the sampling volume of a single excitation fibre overlaps with concentrically arranged fibres. To construct the tip, which will be used in fluids, the fibre bundle is glued together. After its conical polishing, the middle fiber is polished flat. In order to measure on a sample surface the tip can also be enclosed at the distal end with a quartz or sapphire window (**Figure 8C** right).

Enclosing the oblique polished fibre in a glass capillary tube allows the production of minimal diameter deflecting probes. These off the shelf products (MicroQuartz) are used for vaporization and coagulation of tissue, but they could also be used for spectroscopic applications with a single or double fibre system (**Figure 8D**).

If separate illumination and collection fibres are needed, their alignment will be a tedious task. An example of separate illumination and collection fibres is shown in **Figure 8E**. In order to produce a ring probe, fibre pairs are mounted in a cylindrical tube with grooves. A simultaneous investigation of sites along the circumference of the probe is possible (ring probe). The fibre pairs need to be aligned so that the illumination and collection spot overlap. Shrinking tubing or elastic bands could assist in this task.

## Diffuser tips

With the medical approval of photosensitive drugs such as Photofrin® (QLT PhotoTherapeutics Inc., Vancouver, BC, Canada) by the US Food and Drug Administration and the plans for clinical trials with

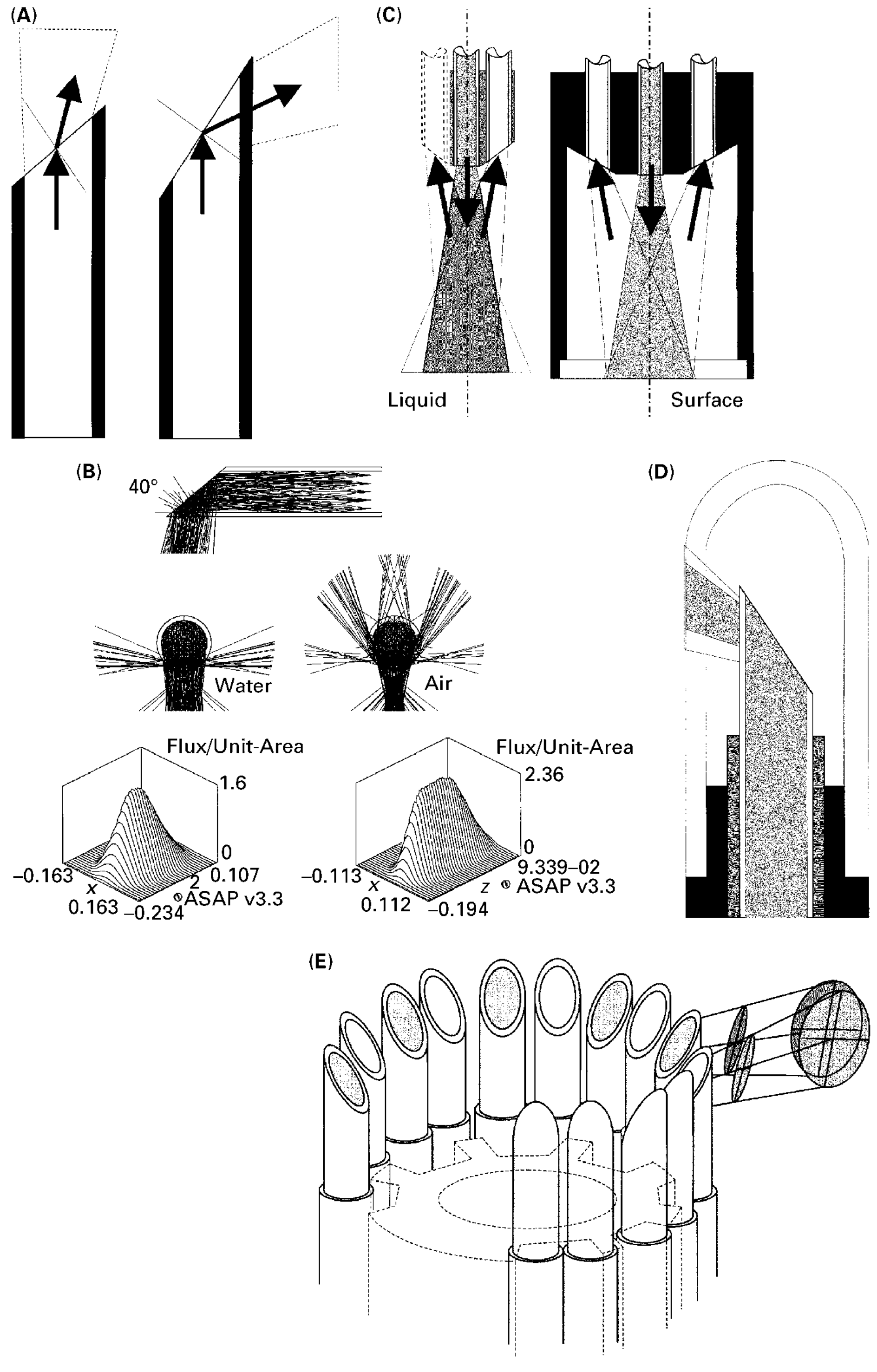



**Figure 8** (A) Oblique polished fibres deflect the output beam or focus it at the sidewall of the fibre. (B) Illumination analysis with a uniform spatial and angular light distribution inside the fibre show a focusing in air and aqueous surrounding media. Irradiance increases to a factor of 2 in the focal spot. (C) Two probe designs for the interrogation of liquids (left) and surfaces (right). A central fibre illuminates the sampling volume or the surface. Six surrounding fibre collect the emitted light. Collection and emission are interchangeable. Housing with a thin shield (quartz, sapphire) permits a constant sampling distance for surface measurements. (D) Oblique polished fibres are enclosed in quartz capillary tubing for minimal diameter side-looking probes. A two-fibre arrangement for separate illumination and collection paths can also be manufactured (InnovaQuartz). (E) A circular arrangement of oblique polished fibres allows the fabrication of a ring probe with separate illumination and excitation channels.

Texaphyrin, the market for probes with a homogeneous illumination of large areas in canals and on surfaces is growing (OPTIGUIDE® QLT Photo Therapeutics | LIGHTSTIC® Rare Earth Medical, Yarmouth, MA, USA). Diffusely scattering volumes mounted at the end of fibres or fibre bundles distribute light over large areas. Beside therapeutic applications such as photodynamic therapy (Lambda Plus Photodynamic Laser, Coherent Palo Alto, CA, USA) and coagulation these diffuse light delivery systems can also be used for illumination of spectral measurements.

Scattering particles are titanium ($TiO_2$) or aluminium oxide ($Al_2O_3$) which are embedded in a transparent matrix such as optical glue (Epo-tek 305, Epoxy Technology, Inc., Billerica, MA, USA). Flexible containment is produced with micro tubing based on fluoropolymers (PTFE, FEP from Zeus, Orangeburg, SC, USA). Rigid tips are based on ceramic or materials similar to Spectralon (Labsphere, North Sutton, NH, USA). Using UV radiation for fluorescence excitation the autofluorescence and transmission of these materials is of importance. FEP tubing (0.3 mm wall thickness) shows good transmission from 250 nm up to 2 μm (40% at 250 nm, VIS 80–90%) while autofluorescence is at least an order of magnitude lower than tissue fluorescence. $Al_2O_3$ does not fluoresce in the UV-visible range.

Lightstics® show uniform light distribution along the optical axis: a reflector mounted at the end of the cylindrical scattering volume reflects light that would otherwise leave the probe backwards. The length of the probe determines the concentration of scatterers (**Figure 9A** and **Figure 9B**).

A reflective foil (aluminium, gold) placed around a diffuser tip restricts the angular illumination. The foil is attached with a light wall shrinking tubing (**Figure 9C**) and the result of a 180° illumination is illustrated in **Figure 9D**. Oblique polished fibres probe the light emitted by the sample area. The jacket of the fibres needs to be removed in order to pass the illumination light. A design based on this principle was proposed by the authors for fluorescence measurements in the endocervical canal.

## A sapphire tip based ring probe

Sampling along the circumference of the probe is useful to identify the angular position of obstructions in tubes (e.g. arteriosclerosis). A probe initially designed for radial ablation and simultaneous spectroscopic analysis is illustrated in **Figure 10A**. The fibres are glued onto a metallic ring (outer diameter 2.4 mm, inner diameter 2.05 mm). A hollow plug attaches internal flexible tubing for flushing and a guide wire. A custom made sapphire prism (Swiss Jewel, Tenero, Switzerland) is attached to the plug. The output of the fibres (NA = 0.22) is totally deflected at the back surface of the prism. NA is the numerical aperture, whose value is sine of half the angle × index of refraction. An outer flexible silicone tube isolates the tip from the liquid environment. Two different outer diameters were realized: 1.4 and 2.8 mm. **Figure 10B** shows the individual parts partially assembled. In the upper figure the outer tubing is not completely mounted. For spectroscopic measurements, smaller 100 μm fibres were integrated in between the ablation fibres (lower figure). The sealing of such a tip is a tedious task since only small amounts of glue can be used.

Similar tips can be produced with vee-, ring-, orifice-, cup- and chisel jewels (Swiss Jewel, Tenero, Switzerland; Sapphire Engineering Inc., Pocasset, MA, USA; SWIP, Biel, Switzerland).

## Refocusing

In order to increase the irradiance or the view of field on the sample area while keeping the distance to the fibres, applications may require converging the excitation beam. Focusing the beam reduces the measurement spot size. If the specimen is put outside the focus, illumination intensities drop while larger areas are illuminated.

### Spherical fibre tips

Melting the end of the fibre shapes the exit surface of optical fibres. An almost spherical surface will be created by the surface tension of liquid quartz. Its form is determined by the volume of melted quartz, which is equivalent to the amount of absorbed energy. Several techniques have been used to achieve this deformation; micro furnace, Bunsen microburner, electrical arc and a $CO_2$ laser beam.

The smallest possible surface curvature is a hemisphere and theoretical results are shown in water and air. The beam exciting a hemispherical fibre tip in air and water is shown in **Figure 11A**. Because the refractive index of silica is 1.46, the focusing power is limited in aqueous media. **Figure 11B** illustrates the focusing of a fibre with a 200 μm core diameter and an NA of 0.22 in air and water. The beam in water is reduced to a quasi-parallel beam. A local concentration of rays can be found at a distance 3 times the distal radius of curvature from the vertex. The maximal irradiance was found to be 3 times the irradiance within the fibre.

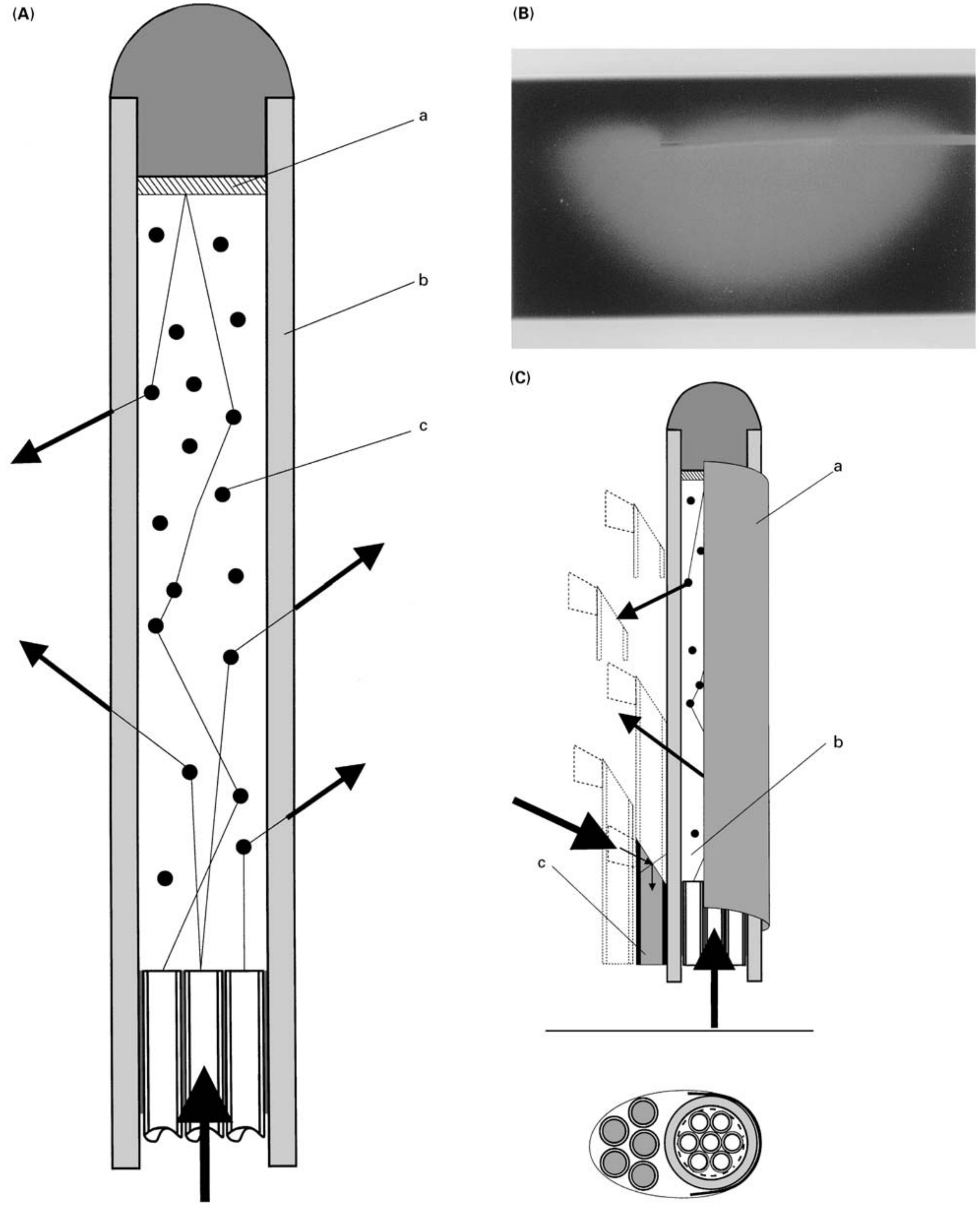



**Figure 9** (A) A diffuser mounted as a tip on a fibre bundle allows uniform illumination along the probe axis (Lightstick®, Rare Earth Medical, Inc.). A (metallic) reflector (a) at the diffuser end and the local concentration of scattering particles (c) determine the intensity profile. Tubing (b) made of fluoropolymers has a high transmission and good heat resistance. (B) A Lightstick® with a length of 2 cm and a reflector allows axial illumination with a 180° field of view. (Courtesy of Rare Earth Medical, Inc., West Yarmouth, MA, USA). (C) This line-sampling probe is based on a combination of a diffuser illuminator (b) and oblique polished fibres (c). With the jacket removed from the collection fibre, the light emitted by the diffuser passes and interacts with tubular surfaces. A metallic foil (a) restricts the illumination to a reduced angle.

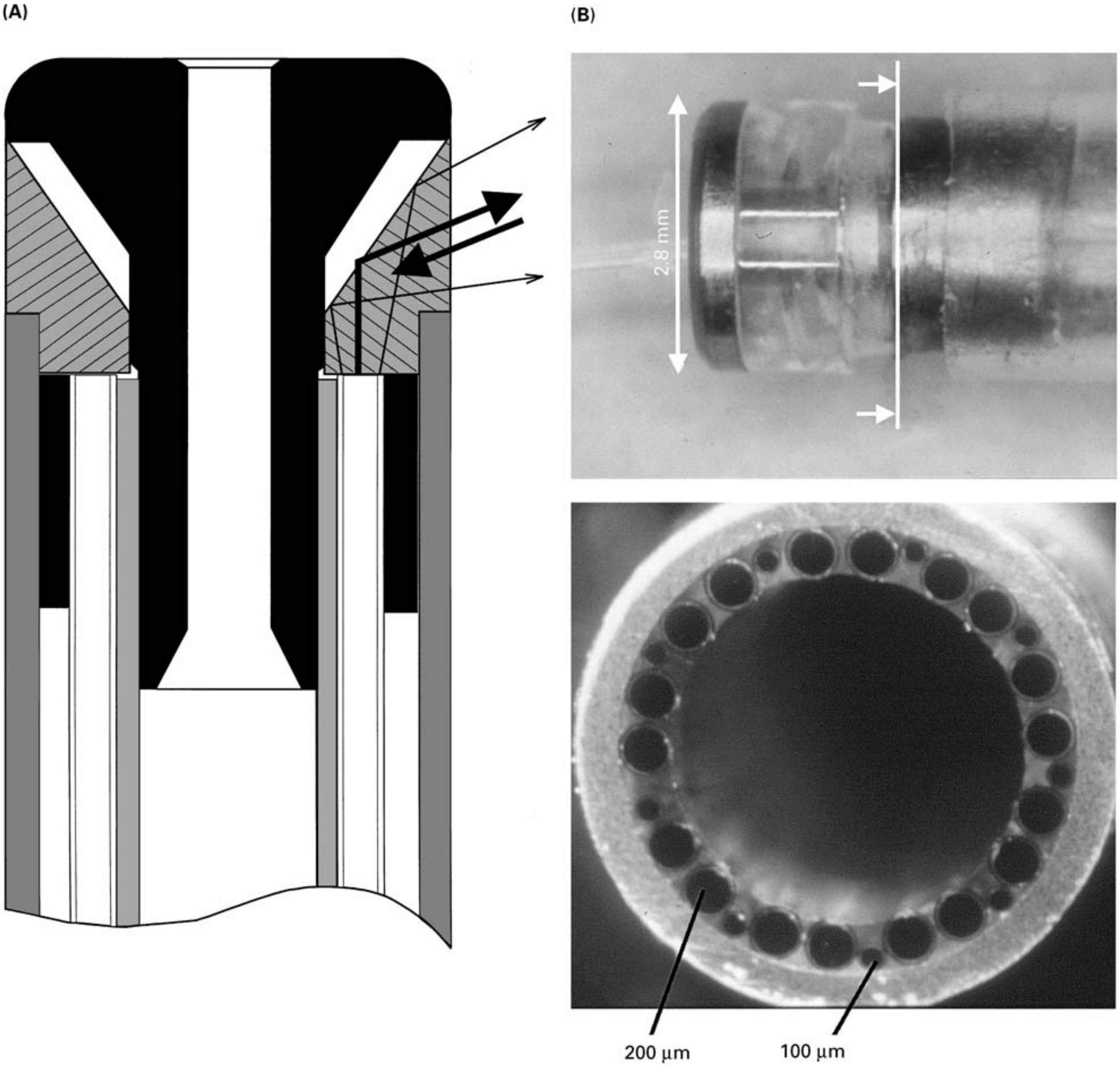



**Figure 10** (A) A sapphire prism deflects the output of a fibre ring into a radial direction. The probe consists of a plug holding the prism and inner tubing. Fibres are glued into a metallic ring. Outer tubing and glue seal the fibre tip. (B) In the upper picture the sapphire prism and the fibres are partially assembled. The lower figure shows the arrangement of illumination and smaller collection fibres in a metallic ring.

## Ball lenses

Spherical ball lenses have been used as collimators in fibre optic connectors. Sapphire spheres with a high refractive power are industrially produced in a wide range of diameters (Sandoz SA, Cugy, Switzerland; Rubis Precis, Charquemont, France). Light emitted from a fibre can be focused with this lens type onto the sample site (**Figure 12A**). Rol found a maximal performance when the ratio of the back surface distance to the fibre and the radius of the sphere is between 1.8 and 2.8 in water and air. Illumination analysis (ASAP) shows a maximal increase of the irradiance by a factor of 4.8 in water and 5.7 in air (**Figure 12B**). If the fibre diameter is increased, spherical aberrations will become important.

A fibre optic probe (**Figure 12A**) with a 200 μm core diameter fibre and a sphere with a radius of 0.4 mm will not exceed an outer diameter of 0.9 mm. The sphere and fibre can be held in a cylinder with a conic inner shape. For side-looking

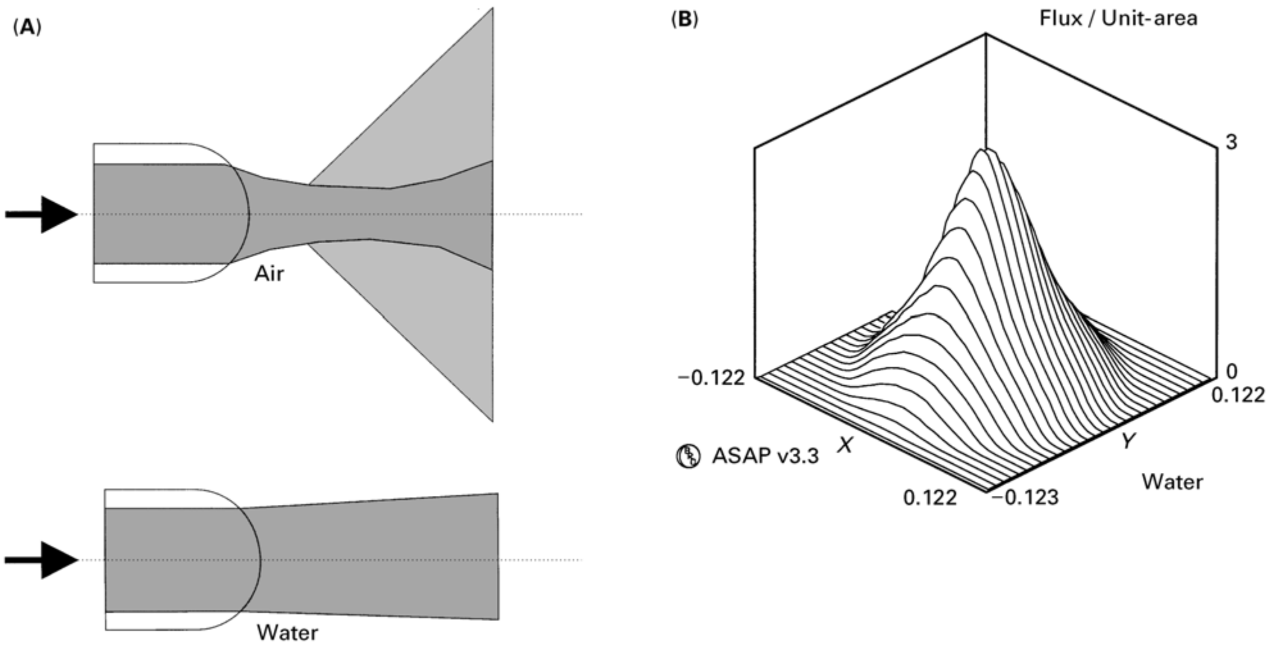



**Figure 11** (A) Hemispherically shaped ends of fibres focus the output beam in air and collimate it in water. Adapted with permission from Rol P and Niederer P (1991) High-power laser transmission through optical fibres: Applications to ophthalmology. In: Wolbarsht ML (ed) *Laser Applications in Medicine and Biology*, Vol 5, pp 141–198. New York: Plenum. (B) Maximal irradiance is found at a distance of 3 times the fibre radius in water for quartz fibres with an NA of 0.2. The irradiance increases by a factor of 3.

applications the same principle is used with a half sphere. If the plane is inclined by 45° and the step in the refractive index is equivalent to sapphire- air, the beam will be deflected with total internal reflection at the interface and leave the probe at 90° with respect to the probe axis. A micro quartz tube encloses the half sphere and provides air as the refractive medium.

### Spherical and parabolic reflectors

Mirrors can concentrate light. The complex refractive index of dielectrics defines the absorption coefficient and the normal incidence reflectance of a mirror:

$$\hat{n} = n + i\kappa$$

$$\mu_a = \frac{4\pi\kappa}{\lambda}$$

$$R_{\perp} = \frac{(n-1)^2 + \kappa^2}{(n+1)^2 + \kappa^2}$$

where $\hat{n}$ is the complex index of refraction, n is the index of refraction, $\kappa$ is the extinction coefficient, $\lambda$ is the wavelength and $R_{\perp}$ is the normal incidence reflection.

For side-looking applications, parabolic or spherically polished mirrors focus the output of several fibres onto the same area (**Figure 13A**). Best performance is shown by a parabolic surface. The manufacture of such mirrors requires a high-precision numerically controlled lathe. Surfaces with optical quality can be manufactured with single-point-diamond-cutting machines.

An example of a probe illuminating a ring is shown in **Figure 13B**. A fibre ring bundle is glued into a metallic ring. A plug with the optimized polished surface is centred into this ring and deflects the axial fibre output in a radial direction. A quartz tube shields the reflective surface. Illumination simulations (ASAP) show that the output of 5 fibres overlap (**Figure 13C**). The rotationally symmetric design spreads the light from the fibres in a quasi-elliptical spot around the circumference. The inner and outer concentric rings of fibres overlap at the cylindrical sample surface. This design confines the sampling volume close to the probe surface because outside the focal spot the light diverges rapidly.

## Probes for NIR and UVR Raman spectroscopy

Near-infrared Raman scattering can be used as a tool to perform in situ histochemical analysis. In biological applications approximately $10^{-10}$ of the incident

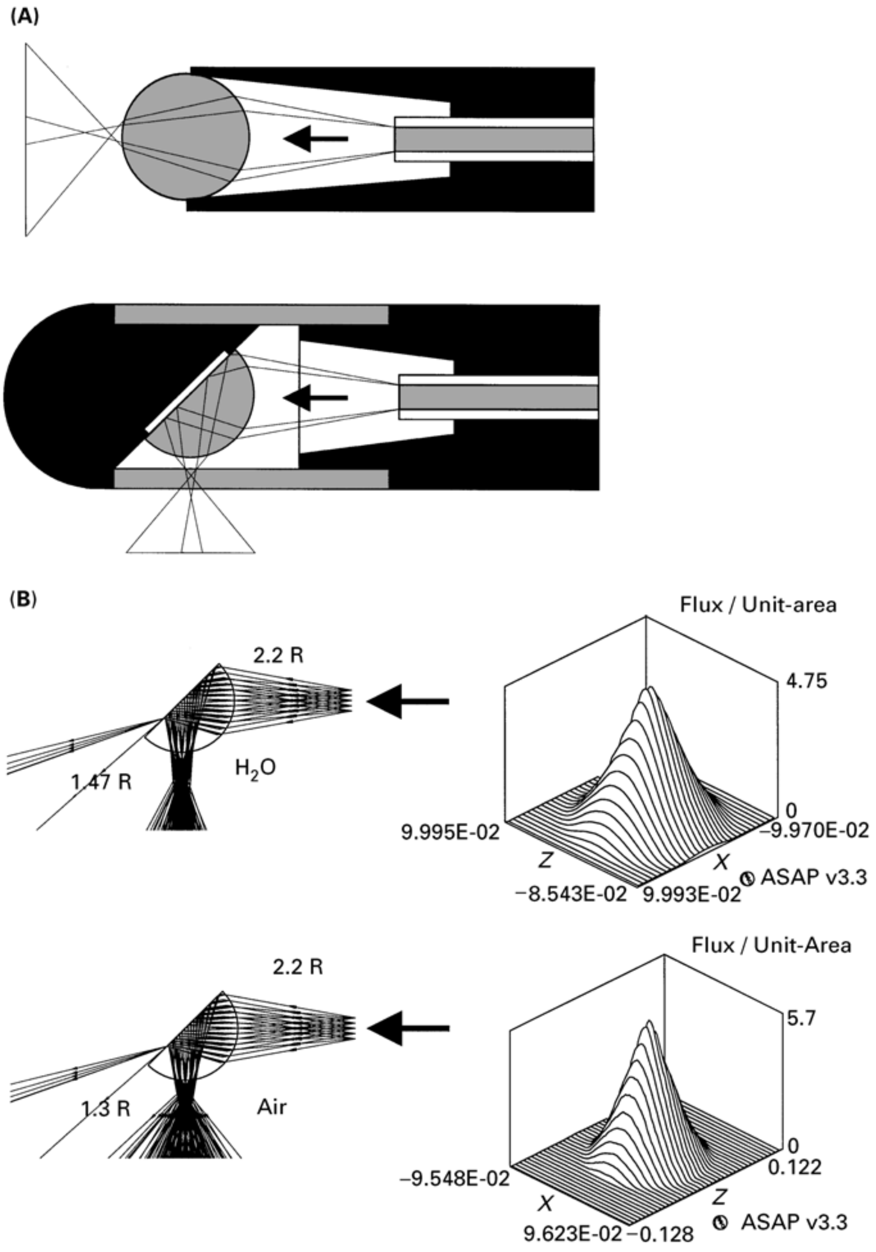



**Figure 12** (A) Ball lenses image the output of a fibre and focus it. A deflecting surface is integrated with a half sphere. On this surface light is internally reflected when using a sapphire–air boundary. The deflecting probe is housed in a quartz tubing, while the axial probe will also focus in aqueous media surrounding the exit surface of the sphere. Adapted with permission from Rol P, Utzinger U, Beck D and Niederer PF (1995) Fiber beam shaping and ophthalmic applications. *SPIE Lasers in Ophthalmology II*: **2330**: 56–62. (B) Illumination analysis demonstrates a five-fold increase of irradiance in the focal spot. (B) Illumination analysis demonstrates a five-fold increase of irradiance in the focal spot.

light is Raman scattered and the Raman signal is normally 6 orders of magnitude weaker than typical fluorescence signals. The amount of light that can be delivered to the sample area is limited due to overheating and American National Standard for Safe Use of Lasers (ANSI) safety standards. The design of a fibre optic probe is therefore driven by the need for maximal light collection. Additional background signal originating from the laser source, the fibres and all optical components is crucial in filling the dynamic range of the detector and producing additional Raman signals. These signals must be reduced with filters for sensitive in vivo measurements.

A design developed by Myrick and Angel (Figure 14A) is based on GRIN lenses (NSG, Tokyo, Japan). Filters are placed in the excitation and

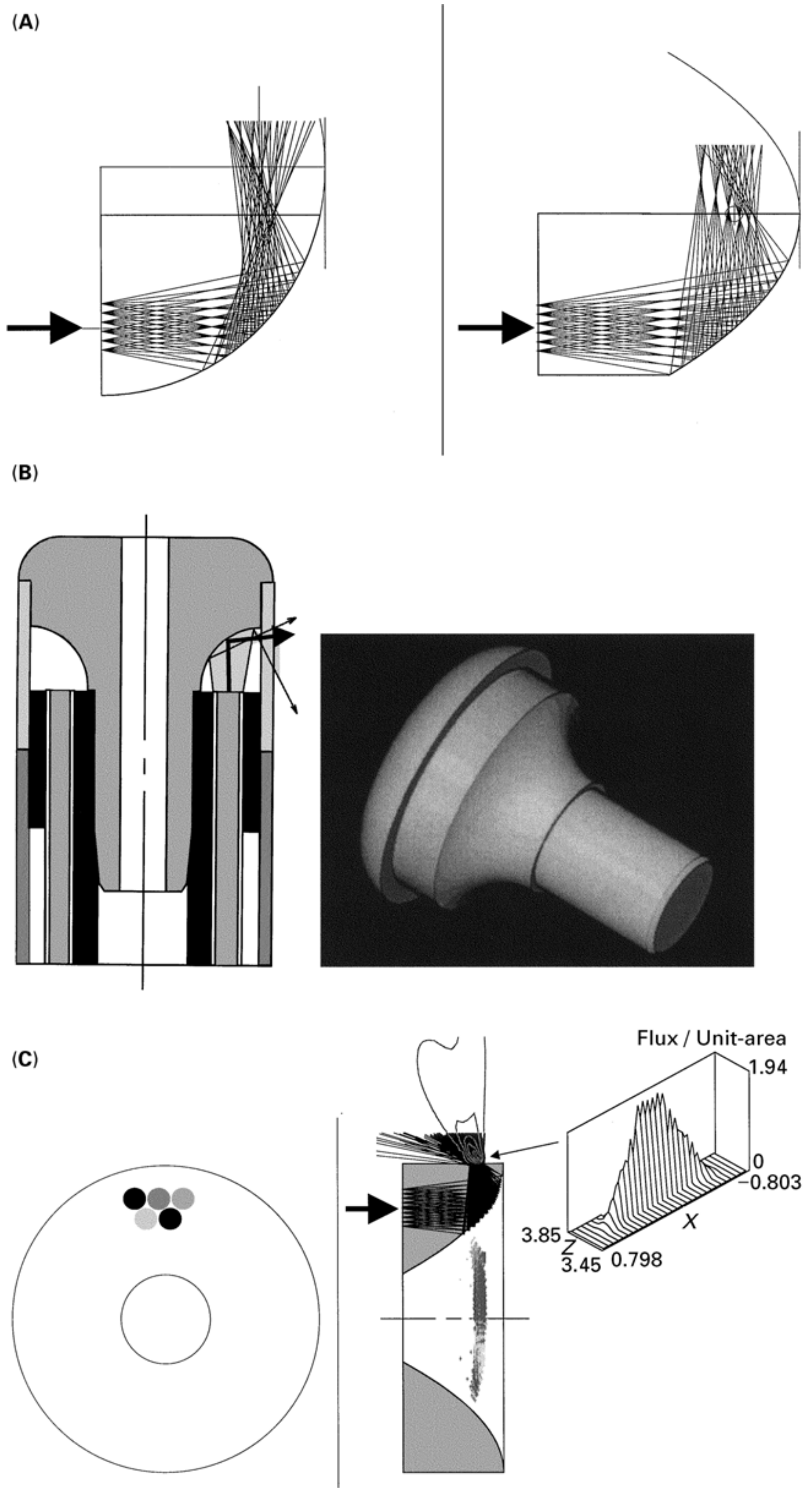



**Figure 13** (A) Parabolic or spherical mirrors focus and deflect the fibre output. Rotationally symmetric mirrors with optical surface quality can be manufactured with numerically controlled diamond turning machines. (B) A fibre optic ring probe with a parabolic reflector. As in **Figure 10A** the plug is mounted into the ring of fibres and flexible tubing. A short piece of quartz tube isolates the mirror from the probe environment. (C) Illumination analysis of the output from 5 fibres shows a good overlap of the fibres placed closer and further out from the axis. In the circumferential direction the measurement spot spreads out.

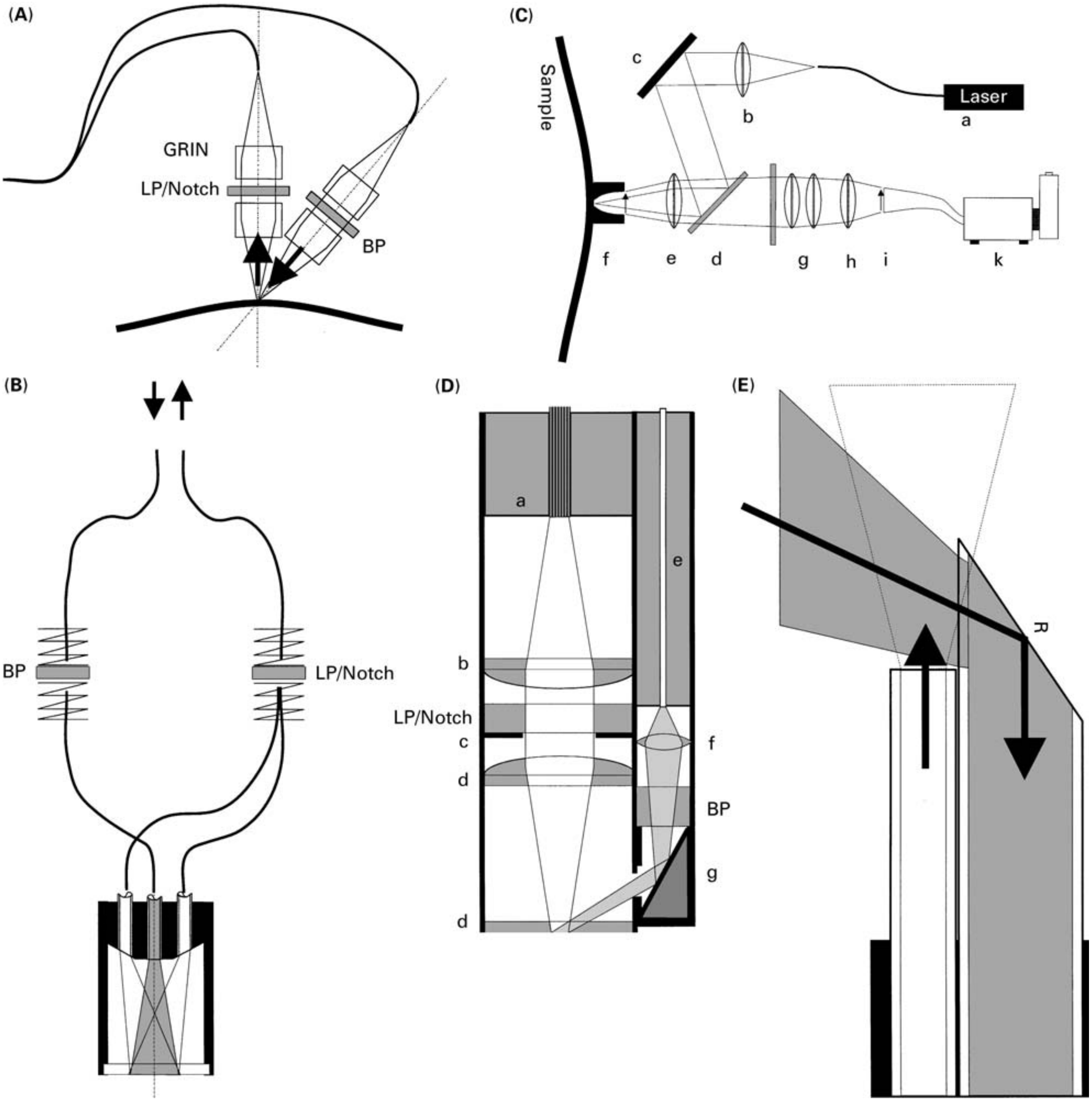



**Figure 14** (A) A fibre optic probe for Raman spectroscopy based on separate illumination and collection fibres. The output of the fibres is collimated and filtered and then projected onto the same area on the sample. Bandpass filters are used for the excitation fibres and longpass filters for the collection fibres. GRIN lenses are assembled into the probe in order to maintain small diameters. Adapted with permission from Myrick ML and Angle SM (1990) Elimination of background in fiber-optic Raman measurements. *Applied Spectroscopy* **44(4)**: 565–570. (B) A fibre optic probe based on the design in **Figure 8C**. In-line filters are mounted into spring loaded SMA connectors. The fibre length between the tip and the filters is kept short. Adapted with permission from Nave S, O'Rourke P and Toole W (1995) Sampling probes enhance remote chemical analysis. *Laser Focus World* December: 83–88. (C) A CPC Raman collection system consists of: (a) NIR diode laser; (b) collimating lens; (c) reflective mirror; (d) dichroic mirror; (e) focusing lens; (f) parabolic concentrator; (g) field lenses to prevent vignetting; (h) a focusing lens to couple the Raman signal into the collection fibre bundle (i) and a high efficiency spectrograph with a liquid nitrogen cooled CCD (k). (D) Raman probe. The excitation light is transported through a single fibre (e) and imaged with a biconvex lens (f) through a dielectric bandpass filter via a mirror (g) onto the sample area. The inelastic scattered light is collected behind a quartz window (d) and imaged by lenses (d) and (b) through an aperture stop and a holographic notch filter onto a flexible fibre bundle (a). (E) The UVRR spectroscopy probe incorporates a smaller diameter solarization-resistant-UV-grade quartz fibre for excitation and an oblique polished fibre to collect scattered light close to the excitation fibre: Reproduced with permission from Schulze HG, Bludes MW, Haynes CA, Klein KF and Turner RFB (1998) Fiber-optic probes with improved excitation and collection efficiency for deep UV Raman and resonance spectroscopy. *Applied Optics* **37**(1): 170–180.

emission paths and allow remote analysis with fibre optic cables. The light source is bandpassed to eliminate signals produced in the fibres and a longpass filter reduces Fresnel and elastic scattered light that enters into the collection fibres.

A similar approach is shown in Figure 14B. A probe developed by Savannah River Technology Center (Aiken, USA) consists of dielectric filters placed between fibres which are held with spring-loaded SMA connectors. Six collection fibres are arranged around an excitation fibre. They are arranged as shown in Figure 8C. A sapphire window isolates the probe from the surrounding environment.

Berger, Tanaka, and other workers have developed a novel design for improved signal collection to allow spectral acquisition in a short time. A compound parabolic concentrator (CPC) is used at the distal end of the probe. They produced a hollow shell, based on a mandrel, electrolytically. The mandrel followed an optimized parabolic form, which was cut with a numerically controlled lathe. Small CPC dimensions with an input aperture of 0.57 mm, exit aperture of 2.1 mm and length of 4.1 mm were achieved (Figure 14C). The probe design with the CPC improved the collection efficiency six-fold. The system incorporates a 500 mW tunable 830 nm diode laser, which is coupled with a mirror and a dichroic beam splitter onto the sample site. Inelastic scattered light exciting the CPC is further collimated with the first lens and filtered by the dichroic mirror and additional filters. Field lenses prevent vignetting of the signal. The last lens images the output onto 200 100 μm fibres, which illuminate a f/1.8 spectrograph. The signal is detected with a back illuminated, liquid nitrogen cooled charge coupled device (CCD). Spectral resolution is 13 $cm^{-1}$.

Mahadevan-Jansen and the authors have successfully measured Raman spectra on the cervix with a fibre optic probe in less than 3 minutes (Figure 14D). A diode laser is coupled into a single 200 μm fibre. A small diameter dielectric filter (3- 4 mm) rejects out-of-band light (optical density (OD)5) and a gold mirror deflects the focused beam onto the specimen site. The beam has to pass a quartz window, which is a part of the housing. Scattered light from the same spot area is imaged with biconvex lenses on a fibre bundle. Elastic scattered and Fresnel reflected light are rejected with a beam aperture stop and a holographic notch filter (OD6). The optics in the detection arm are 8 mm in diameter and the whole probe diameter is less than 2 cm. A similar spectroscopic detection system is used as described in Figure 14C. The spectrograph with the holographic transmission grating (Holospec, Kaiseroptics) and the deep depletion, back illuminated, liquid nitrogen cooled CCD allow optimal system performance (Princeton Instruments).

All the Raman probes described above include imaging optics to enhance the collection efficiency. This makes it difficult to correct the spectral throughput of the system with standardized radiators such as the tungsten filament lamp. The lamp needs to be imaged onto the location of the measurement site and the NA of the probe has to be filled with the light source. The intensity of the source needs to be reduced by orders of magnitude in order to prevent detector saturation.

The scattering cross-section for Raman events can be increased dramatically if excitation is effected using photons in resonance with the electronic transitions of the chromophore involved in the fibration. For many biologically relevant chromophores these transitions are in the UV. Resonance enhancements are in the order of $10^4$- $10^5$ for UV absorption bands. In order to measure these phenomena with fibre optic probes the problems posed by inherent signals produced in the delivery and collection system had to be overcome. Turner and co-workers have successfully built a fibre optic probe for UV resonance Raman spectroscopy (Figure 14E). UV-grade and solarization-resistant-UV-grade optical fibres (Polymicro) are used for the collection and excitation path respectively. Their design incorporates an excitation fibre and an oblique polished fibre with a larger diameter to collect the scattered light (600 μm). This configuration minimizes inner filtering (sample self-absorbance). Because solutions are investigated a reflective surface based on a metallic coating is needed on the collection fibre. Aluminium reflects more than 90% at 250 nm.

## Summary

A large variety of optical designs allow optimal illumination and light collection for spectroscopic applications. Specific solutions for fluorescence, reflectance and Raman spectroscopy are available and the possibility to combine them in a single probe is feasible. Future research will lead towards integrated, multifunctional and highly optimized fibre optic probes.

*See also:* **Light Sources and Optics; Near-IR Spectrometers; Raman Spectrometers.**

## Further reading

ANSI (1993) American National Standards for Safe Use of Laser. Orlando FL: American National Safety Institute.

Bigio JJ and Mourant JR (1997) Ultraviolet and visible spectroscopies for tissue diagnostics: fluorescence spectroscopy and elastic-scattering spectroscopy. Physics in Medicine and Biology 42(5): 803-814.

Brennan J, Wang Y, Dasari RR and Feld MS (1997) Near-infrared Raman spectrometer system for human tissue studies. Applied Spectroscopy 51(2): 201-208.

Durkin AJ (1995) Quantitative fluorescence spectroscopy of turbid samples, PhD Dissertation, University of Texas at Austin.

Farell TJ, Patterson MS and Wilson BC (1992) A diffusion theory model of spatially resolved, steady state diffuse reflectance for the non-invasive determination of optical properties in vivo. Medical Physics 19: 879-888.

Keijzer M, Richards-Kortum RR, Jacques SL and Feld MS (1989) Fluorescence spectroscopy of turbid media: Autofluorescence of the human aorta. Applied Optics 28(20): 4286-4292.

Mahadevan-Jansen A and Richards-Kortum RR (1996) Raman spectroscopy for the detection of cancers and precancers. Journal of Biomedical Optics 1(1): 31-70.

Myrick ML and Angel SM (1990) Elimination of background in fiber-optic Raman measurements. Applied Spectroscopy 44(4): 565-570.

Nave S, O'Rourke P and Toole W (1995) Sampling probes enhance remote chemical analysis. Laser Focus World, December: 83-88.

Nichols MG, Hull EL and Foster T (1997) Design and testing of a white light, steady state diffuse reflectance spectrometer for determination of optical properties of highly scattering systems. Applied Optics 36: 93-104.

Nishioka NS (1994) Laser-induced fluorescence spectroscopy. Gastrointestinal Endoscopy Clinics of North America 4(2): 313-326.

Ramanujam N, Mitchell MF, Mahadevan-Jansen A et al (1996) Cervical precancer detection using a multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths. Photochemistry and Photobiology 64(4): 720-735.

Richards-Kortum R (1995) Fluorescence spectroscopy of turbid media. In: Welch AJ and van Gemert MJC (eds) Optical-Thermal Response of Laser-Irradiated Tissue. New York: Plenum.

Richards-Kortum R and Sevick-Muraca E (1996) Quantitative optical spectroscopy for tissue diagnosis. Annual Review of Physical Chemistry 47: 555-606.

Richards-Kortum RR, Mitchell MF, Ramanujam N, Mahadevan A and Thomsen S (1994) In vivo fluorescence spectroscopy: potential for non-invasive, automated diagnosis of cervical intraepithelial neoplasia and use as a surrogate endpoint biomarker. Journal of Cellular Biochemistry (Suppl) 19: 111-119.

Rol P (1992) Optics for transscleral laser applications. PhD Dissertation Swiss Federal Institute of Technology, Zurich.

Rol P and Niederer P (1991) High-power laser transmission through optical fibers: Applications to Ophthalmology. In: Wolbarsht ML (ed) Laser Applications in Medicine and Biology Vol 5, pp 141-198. New York: Plenum.

Rol P, Utzinger U, Beck D and Niederer PF (1995) Fiber beam shaping and ophthalmic applications. SPIE Lasers in Ophthalmology II: 2330.

Roy AS (1979) Concentrating Collectors, Solar Energy Conversion. An Introductory Course, Dixon AE and Leslie JD (eds) pp 185-252. Oxford: Pergamon.

Shane Greek L, Schulze HG, Blades MW, Haynes CA, Klein KF and Turner RFB (1998) Fiber-optic probes with improved excitation and collection efficiency for deep-UV Raman and resonance Raman spectroscopy. Applied Optics 37(1): 170-180.

Tanaka K, Pacheco M, Brennan J, Itzkan I, Berger A, Dasari R and Feld MS (1996) Compound parabolic concentrator probe for efficient light collection in spectroscopy of biological tissue. Applied Optics 35(4): 758-763.

Utzinger U (1995) Selective coronary excimer laser angioplasty. Dissertation Swiss Federal Institute of Technology, Zurich.

Wang LH and Jacques SL (1992) Monte Carlo Modeling of Light Transport in Multi-layered Tissues in Standard C, developed at University of Texas M.D. Anderson Cancer Center. Available through lwang@tamu.edu.

Wang LH and Jacques SL (1995) Use of a laser beam with an oblique angle of incidence to measure the reduced scattering coefficient of a turbid medium. Applied Optics 34: 2362-2366.

Ward H (1987) Molding of laser energy by shaped optic fiber tips. Lasers in Surgery and Medicine 7: 405-413.

Welch AJ, van Gemert MJC, Star WM and Wilson BC (1995) Definition and overview of tissue optics. In: Welch AJ and van Gemert MJC (eds) Optical-Thermal Response of Laser-Irradiated Tissue. New York: Plenum.

# Fibres and Films Studied Using X-Ray Diffraction

**Watson Fuller** and **Arumugam Mahendrasingam**,
Keele University, Stoke-on-Trent, UK

Copyright © 1999 Academic Press

HIGH ENERGY SPECTROSCOPY
Applications

The fibrous state is a natural habit for those polymer molecules which, because of a high degree of regularity in their monomer sequence, tend to assume helical conformations rather than, for example, the less regular, although nevertheless highly organized, globular structures characteristic of many proteins. Such fibres typically contain regions in which the polymer molecules are arranged with their chain axes parallel to the fibre axis. This orientation is often associated with order in the side-by-side packing of the polymer chains so that the fibre can be regarded as consisting of an array of microcrystallites separated by less ordered regions. Generally there is no preferred orientation of these microcrystallites about the crystal axis parallel to the fibre axis. Consequently an X-ray diffraction pattern from such a crystalline fibre is similar to that which would be obtained from a single crystal if it were rotated through 360° about one of its principal axes during data collection. Although such rotation inevitably results in some loss of 3D information on the crystal structure, this loss is frequently much less significant than might be expected. Indeed the major expansion over the past few years in the use of powder diffraction, where the loss of information because of the complete absence of a preferred crystallite orientation is very much more extensive than that in oriented fibres, provides abundant evidence of the value of the information which can be obtained from studying specimens with a lower degree of order than that of single crystals. In this context it is important to emphasize that the degree of crystallinity in oriented fibres, as indicated by the upper limit in the angle of diffraction for which sharp Bragg diffraction peaks are observed, is often comparable with that observed for single crystals of proteins and other macromolecules. Many polymer materials are manufactured as films with a thickness of a few hundred micrometres. Since crystallite sizes in polymer films are typically a few tens of nanometres or less in lateral dimensions (i.e. ~ ten thousand times less than the film thickness) they can be analysed using similar methods to those used to study fibres.

## The relevance of X-ray diffraction studies of fibres and films

Fibre diffraction is one of the most powerful and generally applicable techniques available for the study of macromolecular structure. The validity of this claim may be demonstrated by the following points:

(1) For many biological structures such as muscle and connective tissue, the fibrous state is functionally important and fibre diffraction techniques therefore allow both the structure of the macromolecules themselves and their organization in higher order structures to be investigated with the minimum of disturbance to the *in vivo* structure.
(2) The technological value of many synthetic polymer materials can be related to their molecular conformation and arrangement, typically expressed in terms of the degree of polymer orientation and crystallinity. Fibre diffraction provides the opportunity to investigate these structures in functionally important situations, e.g. in commercially available sheets and fibres.
(3) Because of the high degree of regularity in helical structures and the availability of information on covalent bond geometry from X-ray single crystal studies of model compounds, it is possible to describe the conformation of many fibrous polymers in terms of a relatively small number of parameters, i.e. the torsional angles about the various covalent single bonds. The use of such descriptions can have a dramatic effect on the ratio of observed data to parameters to be determined and hence on the accuracy with which fibrous structures can be determined.
(4) The fibrous state is particularly well suited to the investigation of structural transitions in the conformation of polymer molecules and in their arrangement. Such investigations are of increasing importance in the study of both naturally occurring and synthetic polymers.

(5) While many biological macromolecules can undergo conformational transitions, these are typically between well defined structures. In many cases these structures persist in a wide variety of molecular environments, e.g. the B form of the deoxyribonucleic acid (DNA) double-helix determined from X-ray fibre diffraction is observed with relatively little conformational variation in solution, in association with a number of different proteins and in fibres in a range of crystalline and semicrystalline arrays. Such observations comprise just one aspect of the large body of data which bears testimony to the biological relevance of structural information obtained from diffraction studies of macromolecules in fibres, films and single crystals.

Biological evolution exhibits a wealth of examples of the exploitation of polymer structure, with muscle a particularly impressive and intriguing example. The recognition that these biological structures have been refined over hundreds of millions of years in response to competitive pressures has prompted an increasing interest in using them as a source of ideas for the rational design of man-made materials and with it increasing emphasis on efforts to relate the structure of polymer materials to their physical properties.

## The impact of X-ray synchrotron radiation sources

The advent of dedicated X-ray synchrotron radiation sources, beginning with the Daresbury Laboratory Synchrotron Source (SRS) in 1981, dramatically extended the power and scope of X-ray diffraction studies of fibres and films. The high flux per unit area of the beam at the specimen position produced by such sources allowed diffraction patterns to be recorded with exposure times two or three orders of magnitude shorter than those required for a conventional rotating anode X-ray source. As a consequence it became possible to record within minutes complete 2D high angle X-ray fibre diffraction patterns which had previously required exposure times of hours. As third generation sources such as the European Synchrotron Radiation Facility (ESRF), Grenoble, together with corresponding improvements in beam line instrumentation, became available during the 1990s exposure times in the millisecond range became possible with the prospect of recording in real time changes in the X-ray diffraction pattern corresponding to changes in polymer conformation and organization. Such data is of central importance in investigating the relationship between structure and function in biological macromolecules and in relating the structure of industrial polymers to their bulk physical properties. The very high degree of collimation in radiation from sources such as the ESRF together with innovations in beam line optics has made it possible to reduce the diameter of the incident X-ray beam by two orders of magnitude from that typical of rotating anode sources and hence to allow structural variation in polymer materials to be recorded with a spatial resolution of ~1 micrometre. Since the radiation emitted from synchrotron sources comprises a continuous spectrum from X-rays to the far infrared it is possible to select the wavelength used in any particular experiment to match specific features in the absorption spectrum of the material being studied. In particular this ‘wavelength tunability’ offers the possibility of undertaking anomalous scattering experiments in which the wavelength of the incident X-rays is varied in the region of an atomic absorption edge. This approach has considerable potential but, to date, its application in the study of polymer materials has been limited because of technical difficulties.

Two regimes are distinguished in X-ray diffraction studies of fibres and films according to whether they are in the domain of wide angle X-ray scattering (WAXS) or small angle X-ray scattering (SAXS). In WAXS, X-rays scattered through angles greater than ~1° to the incident X-ray beam are recorded. This typically provides information on structural features with dimensions in the range 5 nanometres to 0.1 nanometres, e.g. the detailed conformation and packing of polymer chains. In SAXS, X-rays scattered through angles less than ~1° to the incident X-ray beam are recorded. This typically provides information on structural features with dimensions in the range 100 nm to 5 nm, e.g. the overall folding and organization of polymer chains and the shape, size and distribution of crystallites. This distinction between the information which can be obtained from WAXS and that which can be obtained from SAXS illustrates the general principle in scattering theory that spacings in a diffraction pattern (i.e. the angle of scatter) are reciprocally related to the dimensions of features in the scattering object to which they relate.

The information which can be obtained in X-ray diffraction studies on fibres and films by the exploitation of synchrotron radiation sources is illustrated in this article by studies of both biological and man-made materials. All the studies were made at either the SRS or the ESRF but the techniques described are typical of those used at the rapidly increasing number of synchrotron sources available throughout the world. Examples of the data which can be obtained from highly ordered fibrous structures are

illustrated for a WAXS pattern from DNA in **Figure 1** and a SAXS pattern from muscle in **Figure 2**. The possibility of recording WAXS and SAXS data simultaneously in real-time from the same specimen offers the possibility of relating changes in local polymer structure as revealed by WAXS to larger scale changes in polymer organization as revealed by SAXS. Increasingly, the information obtained from SAXS and WAXS on molecular structure and organization is complemented by the use of other techniques in which the recording of observations is synchronized to SAXS and WAXS measurements, e.g. stress/strain and calorimetric measurements in studies of the response of polymer materials to mechanical and thermal stress.

## Experimental facilities and techniques

The radiation emitted from the electron storage rings at synchrotron radiation sources is typically chanelled through a number of beam lines into some 30 or so work stations (**Figure 3**). The work stations vary in having their beam line optics designed to produce a beam with the cross-sectional area and degree of parallelism at the specimen which best matches the demands of the type of experiments to which they are dedicated. For X-ray diffraction studies of fibres and films the most common requirements are a highly monochromatic beam with a wavelength of ~0.1 nm and the highest possible X-ray intensity within a beam diameter ~100 µm. On third generation sources such as the ESRF, microbeams with diameters of ~10 µm or less are produced without unacceptable loss of intensity through collimation or a loss of parallelism because of focusing elements. Loss of parallelism is of particular concern for SAXS studies because the more the incident X-ray beam is focused on the specimen, the more divergent will be the unscattered X-rays passing through the specimen and hence the larger the minimum angle at which scattered X-rays can be recorded and the lower the upper limit in the dimensions of structural features in the specimen on which information can be obtained. Monochromatization and focusing of the main beam are typically achieved using a variety of crystal monochromators and curved mirrors which reflect X-rays at glancing incidence often from surfaces coated with thin metal layers. The increased use of insertion devices such as undulators in straight sections of the electron storage ring in more recent synchrotron sources allows the radiation fed to particular work stations to be matched more closely to specific applications in terms of intensity as a function of wavelength. Such developments are of particular importance for diffraction studies where there is often a need to maximize the intensity in a very narrow wavelength range.

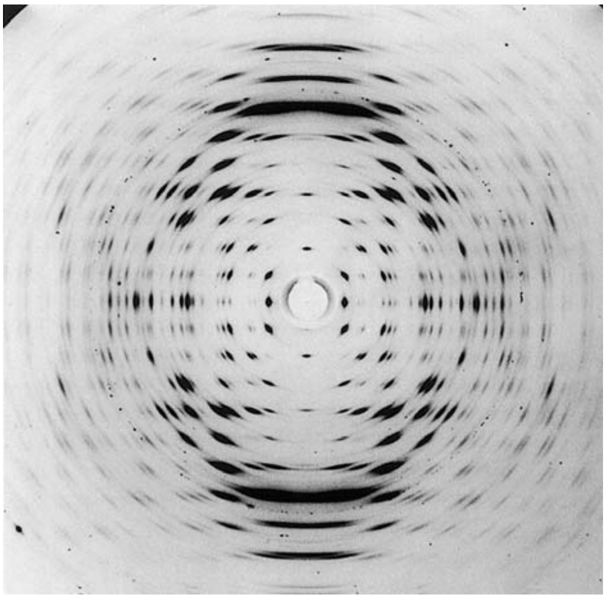


**Figure 1** Wide angle X-ray fibre diffraction pattern recorded from the D form of the DNA double-helix.

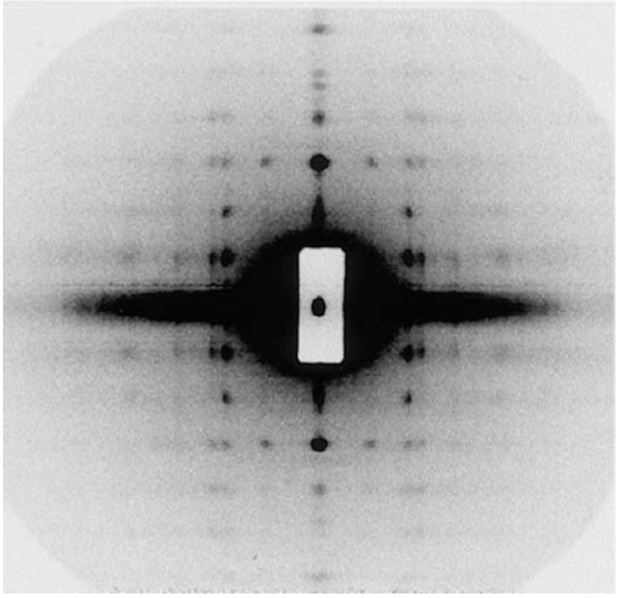


**Figure 2** Small angle X-ray fibre diffraction pattern recorded from plaice fish muscle.

A variety of X-ray cameras have been developed for use on synchrotron radiation beam lines to provide a controlled environment in which the material to be studied can, for example, be subjected to particular conditions of relative humidity, thermal and mechanical stress. The capacity to control relative

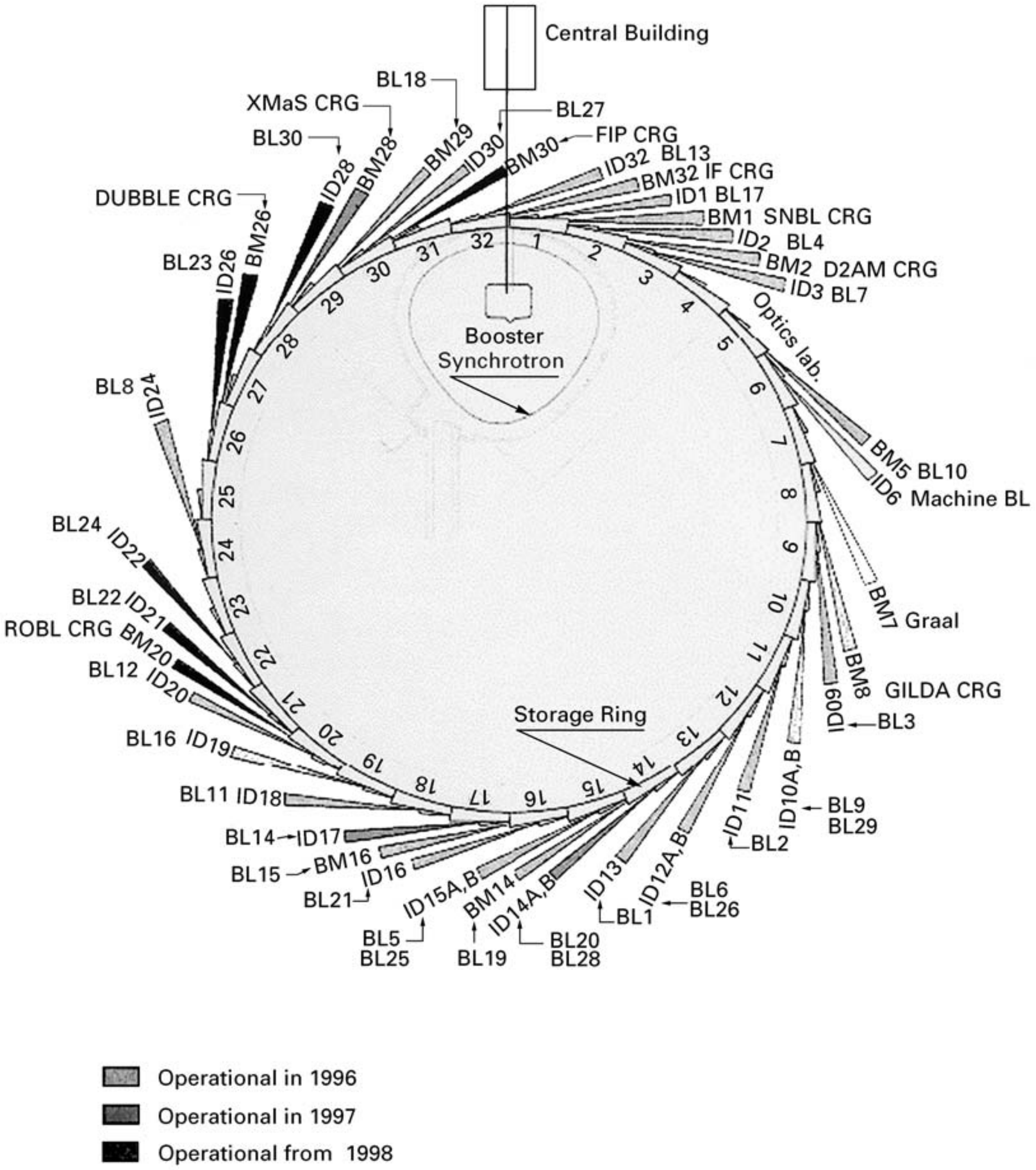



**Figure 3** Schematic representation of beam lines and work station at the ESRF.

humidity is of particular importance for biological materials since *in vivo* they are commonly extensively hydrated. The processing of polymer materials typically involves exposing them to thermal and mechanical stress and cameras for studying them under industrial processing conditions typically incorporate drawing rigs able to apply various patterns of mechanical stress at well defined temperatures. The front face of an X-ray camera normally accommodates a final collimating system defining the size of the X-ray beam falling on the specimen mounted close to the final aperture of the collimator. In some cameras the specimen mount incorporates a goniometer driven remotely which allows data to be collected as a function of the tilt of the fibre with respect to the incident X-ray beam. For the collection of WAXS data the distance between the front and rear faces of the camera is ~10 cm. For the collection of SAXS data this distance may be 1 m or even more if data corresponding to periodicities in the specimen of greater than 50 nm are to be collected. The detector on which the diffraction pattern is recorded is placed close to the rear face of the camera which is made from a material with low X-ray absorption such as a 10 μm thick sheet of the plastic Melinex on which a lead backstop is mounted to absorb the component of the incident beam not scattered by the specimen. To reduce degradation of the data due to

scattering by air of the X-ray beam not scattered or absorbed by the specimen, the camera is evacuated or the air replaced by helium gas which is a much weaker scatterer of X-rays. Until relatively recently X-ray diffraction data from fibres was commonly recorded on photographic film. However, during the 1990s a variety of electronic area detectors have largely replaced film at synchrotron sources. One of the most important advantages of electronic area detectors of the gas filled multiwire and CCD types is that they allow the diffraction pattern to be displayed in real time so that experiments can be genuinely interactive. In contrast, it may take 10 or 15 min to process film. This time does not of course include scanning of the processed film as a necessary preliminary to accurate determination of the position and intensity of diffraction spots. In contrast, data recorded electronically can be immediately downloaded to an online computer for data reduction and analysis. In deciding between different types of detector it is important to compare parameters such as:

(1) *spatial resolution* – the grain size on X-ray film is typically less than 10 μm and although electronic area detectors generally available have significantly worse resolution than this, they are adequate for recording fibre diffraction data;
(2) *quantum efficiency* – depends critically on the wavelength of the X-rays, therefore while comparisons with film are complex it is fair to say that state of the art electronic area detectors can show efficiency comparable to film;
(3) *dynamic range* – the limited dynamic range of film is one of its major limitations and an area where its performance is generally inferior to electronic area detectors.

Image plates have a spatial resolution comparable to film with a much larger dynamic range. However, they still suffer from the need for scanning which, even with ingenious online carousels and dedicated scanners, still makes it impossible to conduct genuinely real-time interactive experiments. The examples of diffraction data included in this review illustrate the use of the various types of detector.

## Examples of the exploitation of X-ray diffraction using synchrotron radiation in the study of fibres and films at high temporal and spatial resolution

These examples are chosen to illustrate the impact of both WAXS and SAXS studies on the characterization of naturally occurring and man-made materials. These studies exploit both the greatly increased intensity and much reduced divergence in the X-ray beam available from synchrotron sources as compared with that from laboratory rotating anode X-ray sources.

### Conformational transitions in the DNA double-helix

The original proposal of a double helical structure for DNA was based on wide angle X-ray fibre diffraction data. Subsequently, a number of distinct conformations of the double-helix have been observed and designated A, B, C, D and Z. The sensitivity of the conformation of the DNA double-helix to its degree of hydration allows conformational transitions to be induced by changing the relative humidity of the fibre environment. The availability of X-ray synchrotron radiation sources offers the possibility of observing the variation in diffraction data during such transitions with the possibility of describing the stereochemical pathway followed during the transition. In an early application of the Daresbury SRS the reversible transition between the B and D conformations was followed through a number of cycles. Selected fibre diffraction patterns from this transition are in **Figure 4**. The transition between the D and B conformations can be followed through the appearance in the original highly crystalline D pattern of reflections corresponding to a semicrystalline intermediate between the D and B forms. As the transition proceeds, these reflections become more intense, corresponding to an increasing fraction of the DNA assuming the semicrystalline phase, and gradually change in position corresponding to a continuous change in helix pitch from the 2.4 nm characteristic of the crystalline D form to the 3.4 nm characteristic of the semicrystalline B form. The diffraction data on which this analysis was based were recorded on photographic film with exposure times of a few minutes. In the more than 10 years since these studies were made major advances in both synchrotron performance and detectors have dramatically decreased the exposure times required for such time-resolved studies.

### Changes in molecular organization and interactions during muscle contraction

Muscle is a complex assembly of proteins embracing the metabolic capabilities characteristic of globular proteins and the structural capabilities of fibrous proteins and is thus well suited to perform the dual roles of converting chemical energy into mechanical energy and sustaining tension. Over 40 years ago a combination of electron microscopy and X-ray fibre

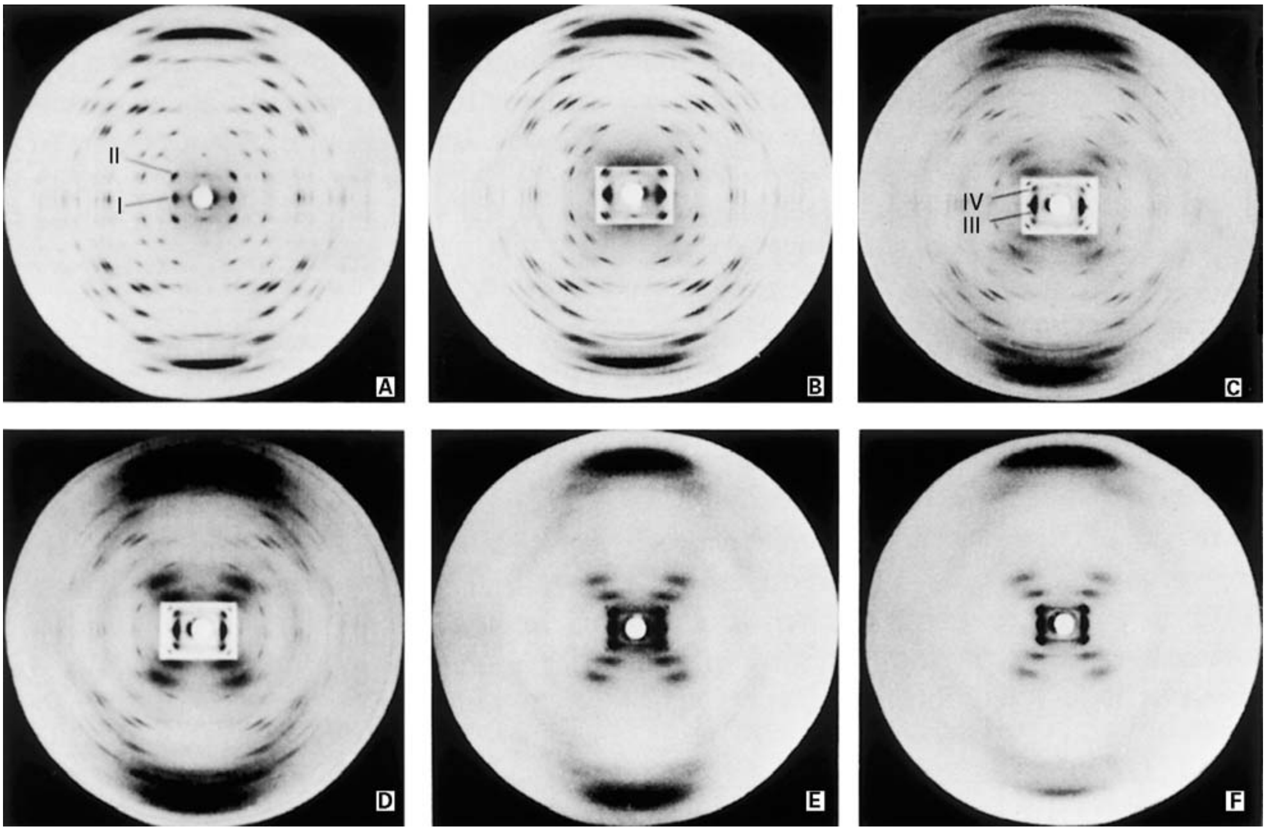



**Figure 4** Diffraction patterns illustrating typical stages in the D to B transition in the DNA double-helix. In (A) the reflection marked I is related to the lateral intermolecular separation and that marked II to the helical pitch of molecules in the D form. In (C) the reflection marked III is related to the lateral intermolecular separation and that marked IV to the helical pitch of molecules in the semicrystalline intermediate form.

diffraction showed that muscle consisted of a regular array of actin filaments interpenetrating a regular array of myosin filaments and that during muscular contraction one family of filaments slid over the other. The dimensions of these arrays and the periodicities along the length of the actin and myosin filaments are within the 10 to 100 nm range and X-ray fibre diffraction information on these structures has largely come from the small angle regime. A particularly well defined example of such SAXS data is in **Figure 2**. The high brilliance of synchrotron sources together with recent advances in the efficiency of multiwire area detectors have allowed the variation in diffraction during the contraction cycle, in which a dissected muscle is repeatedly stimulated by an electrical voltage, to be recorded with a time resolution of a few milliseconds. This data is being used to model changes in the interaction between the actin filaments and the heads of the myosin molecules making up the myosin filaments since it is believed that changes in the stereochemistry of these interactions is the origin of the mechanical stress which results in shortening in the length of the muscle against a load. This modelling is an excellent example of complementarity between an X-ray structure determination of a protein single crystal (in this case the structure of isolated myosin heads) and a small angle X-ray fibre diffraction study of the whole muscle assembly.

### Strain-induced crystallization in poly(ethylene terephthalate) (PET)

One of the most challenging issues in the crystallization of polymers is the nature and kinetics of the strain-induced crystallization that takes place when polymers such as PET are mechanically drawn close to the glass transition temperature. From recent experiments at the ESRF and SRS, it has been possible to monitor crystallization in real-time when amorphous PET was drawn at rates comparable to those used in industrial processing. The X-ray camera used in these studies consisted of a 150 mm × 150 mm × 150 mm oven of which three sides were interchangeable so that it was possible to exchange sides depending on the particular application. This, for

example, allowed the viewing port in which a video camera was mounted to be varied. The temperature could be controlled to within 1°C and the maximum temperature attainable was 350°C. A strip of polymer material could be clamped between two jaws attached to stepper motors which allowed uniaxial bidirectional drawing at rates up to 72 000% per minute. Diffraction patterns were recorded with exposure times of 40 milliseconds. The variation in WAXS in a typical experiment is shown in **Figure 5**. In this sequence of patterns the sharp crystalline peaks can be seen to progressively emerge above a diffuse background scatter. Simultaneously recorded video camera observations on changes in the position of reference marks on the specimen are also shown in **Figure 5** and allow the precise draw ratio

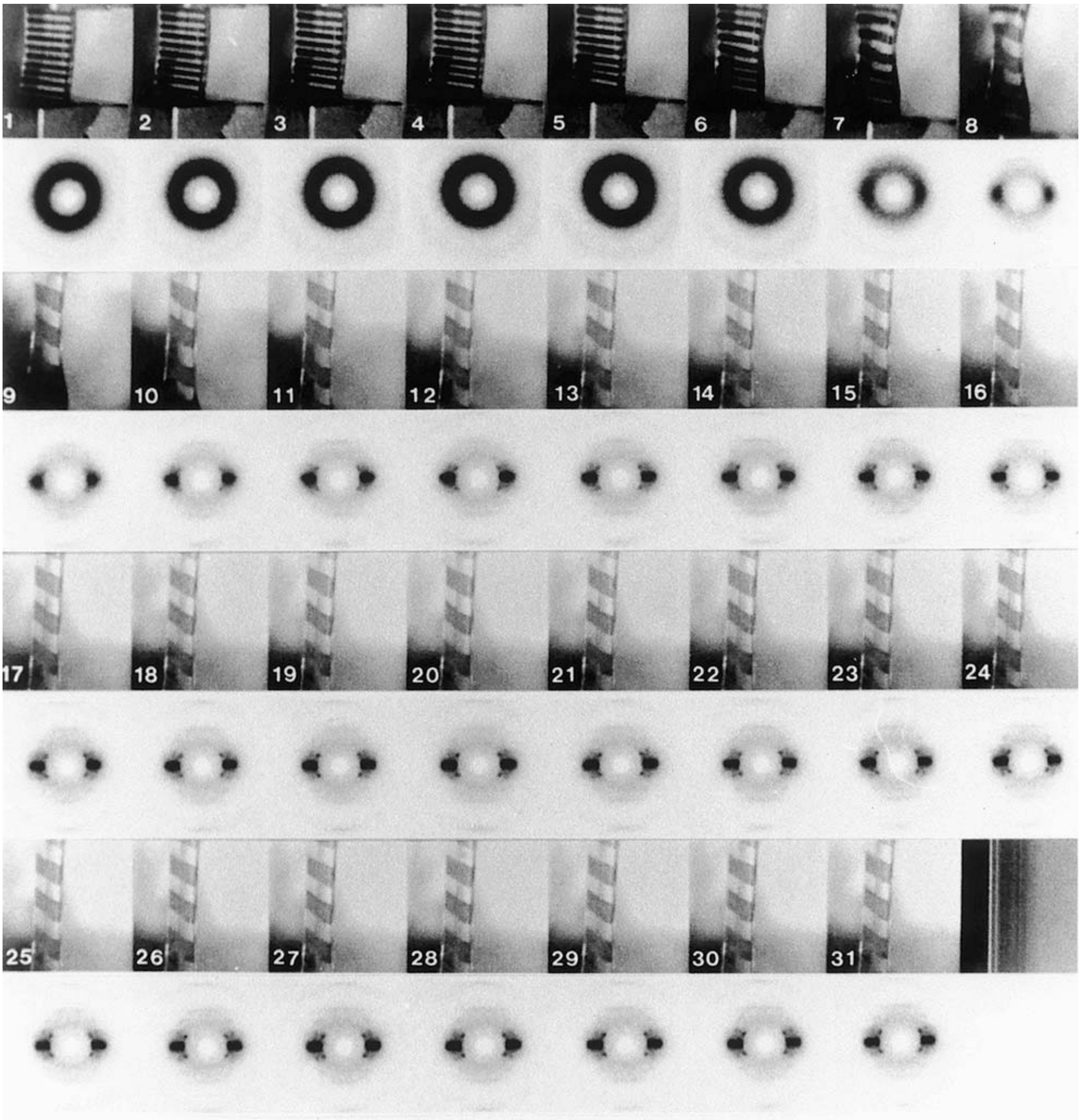



**Figure 5** WAXS patterns with corresponding video image of sample of PET recorded during drawing at 90°C at an overall draw rate of 72 000%/min. Each of the 31 frames was recorded with an exposure time of 40 ms. There was essentially no dead time between frames so that the drawing experiment was completed in ~1.2 s.

and draw rate to be calculated at the point in the specimen from which X-ray diffraction data was recorded. Measurements on these patterns allowed the development of polymer orientation and crystallinity to be quantified and hence the kinetics of crystallization to be determined for a number of temperatures and for a variety of draw rates and draw ratios. An important conclusion from these studies, which is at variance with earlier conclusions drawn from studies on samples quenched at various stages during drawing and annealing, is that, at the fast draw rates typical of industrial processing, significant crystallization does not begin until the conclusion of drawing. When the rate constants determined from these experiments were related to theoretical models for the relaxation of deformed chains within entangled networks it became clear that at fast draw rates the onset of crystallization does not occur until the end of draw because the draw rate exceeds the rate of chain relaxation.

## Relation between long and short range order during deformation of polyethylene

Instrumentation for simultaneous recording of SAXS and WAXS data is illustrated schematically in **Figure 6**. A hole, ~10 mm diameter, in the centre of the 15 μm thick aluminium foil through which diffraction beams exit the camera is covered with a circular glass cover slip (13 mm in diameter; and 15 μm thick) so that the SAXS and the main beam not scattered by the specimen pass wholly through the glass cover slip. This avoids diffraction of the main beam by the aluminium resulting in spurious sharp diffraction peaks in the WAXS pattern. Both SAXS and WAXS data were recorded using Photonic Science CCD detectors. An evacuated beam pipe was positioned between the SAXS detector and the glass cover slip mounted on the aluminium foil to minimize scattering by air along the long path length. The WAXS detector was positioned close to the aluminium foil window at the rear of the camera so that SAXS data could still pass unobstructed for collection at the second CCD detector.

A specimen, 10 mm wide and 20 mm long, was cut from a 1 mm thick sample of high density polyethylene and mounted within the camera in jaws attached to stepper motors. The specimen was drawn at a rate of 18 000%/minute at a temperature of 35°C. Strain data, for the region of the specimens from which the SAXS/WAXS data was collected, was calculated from an accurately synchronized continuously recorded video image of the specimen. **Figure 7** shows selected SAXS and WAXS frames recorded from the specimen. In the initial WAXS frame two rings of intensity are clearly visible. These may be indexed as reflections from the orthorhombic form of polyethylene. The initial SAXS frame consists of a largely isotropic ring indicating that the undeformed sample is spherulitic with a lamellar spacing of ~35.0 nm. From the sequence of WAXS patterns it is apparent that increasing strain results in increasing orientation in both of the prominent orthorhombic reflections together with the transient appearance of a monoclinic phase. The SAXS patterns show striking variation in intensity which can be attributed to the formation of microvoids.

## Variation in crystallite order and orientation in spherulites of Biopol

X-ray diffraction provides one of the most powerful techniques for the investigation of the spatial

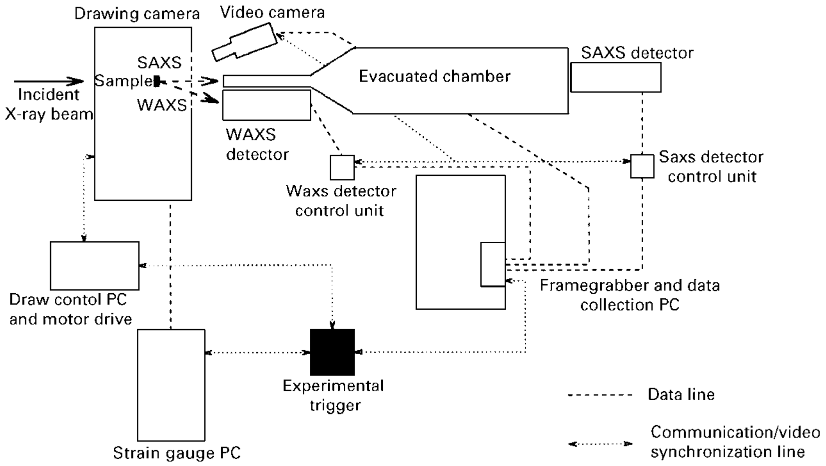



**Figure 6** Schematic representation of the experimental arrangement for simultaneous recording of SAXS/WAXS and strain data during mechanical deformation of a sheet of polymer material.

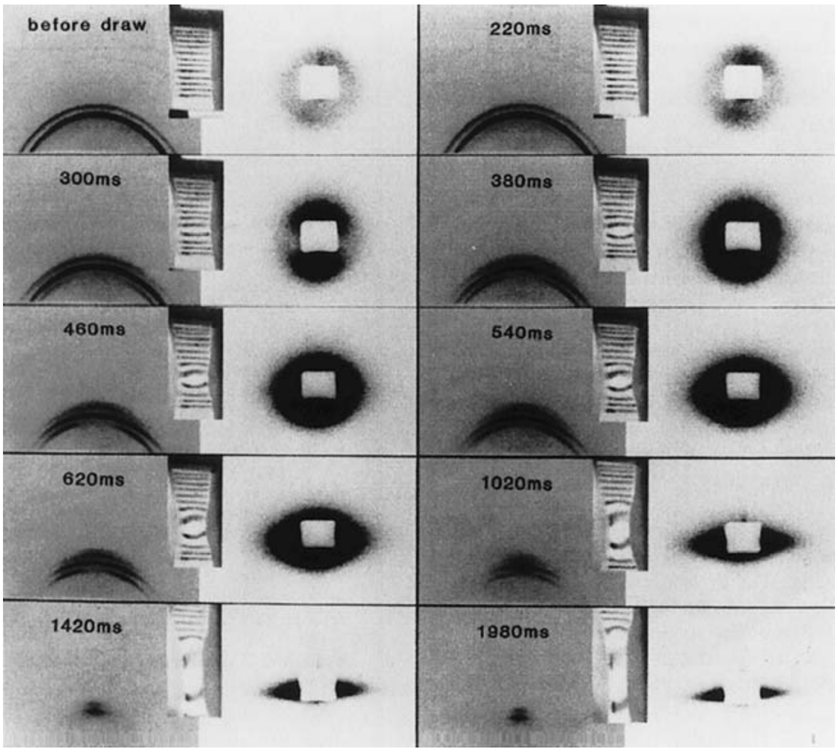



**Figure 7** Selected pairs of simultaneous WAXS (left) and SAXS (right) frames with inserts of corresponding synchronously recorded video images which allow strain data to be related to changes in polymer structure and organization of high density polyethylene. Each frame was recorded in 40 ms.

variation of polymer conformation and organization within partially ordered materials. From optical microscopy it is clear that molecular orientation within many polymer materials varies over dimensions of less than 100 μm. Characterization of this variation by X-ray diffraction has been achieved using the microfocus capability of the ESRF which can provide an intense highly monochromatic X-ray beam with a wavelength of 0.092 nm in a beam diameter at the specimen of ~10 μm. Crucial to these experiments was the availability of a computer-controlled X/Y stage which allowed the specimen to be tracked in two dimensions perpendicular to the X-ray beam in steps as small as 1 μm (**Figure 8**). WAXS data were recorded using a Photonics Science CCD detector from spherulites of the optically active thermoplastic polymer Biopol. This material is produced by a variety of microorganisms, which use it for energy and carbon storage. The biological rather than petroleum origins of Biopol, its biodegradability, and its ability to form films with the properties of conventional thermoplastics give it technological importance as well as fundamental interest. The purity of the polymer leads to an absence of heterogeneous nuclei and gives it the capacity to form large spherulites of high crystallinity. X-ray diffraction patterns with exposure times of 1 s were recorded as a spherulite of Biopol was stepped across a square net of area 150 μm × 150 μm centred on the centre of the spherulite, in horizontal and vertical steps of 10 μm. The 15 × 15 diffraction patterns recorded in this way are assembled as a 2D array in **Figure 9**. This array of patterns vividly demonstrates that all the crystallites have their *a*-axes oriented along a radius of the spherulite.

### Variation in polymer orientation and crystallinity across a plastic bottle wall

Poly(ethylene terephthalate) (PET) is used extensively in the manufacture of fibres, bottles and films. Physical properties, such as strength, transparency and gas-transport properties depend on the degree of polymer orientation and crystallinity and their variation in the artefact.

During blowing of a bottle, the inner surface of a preform is cooled by the expansion of the compressed air used for inflation, the outer surface is cooled by contact with the cold mould whilst the internal thermal energy is raised by strain heating. In addition, the contour shape of the container and the complex blowing process itself produces considerable variation in wall thickness and direction of stretching at different positions in the container. The

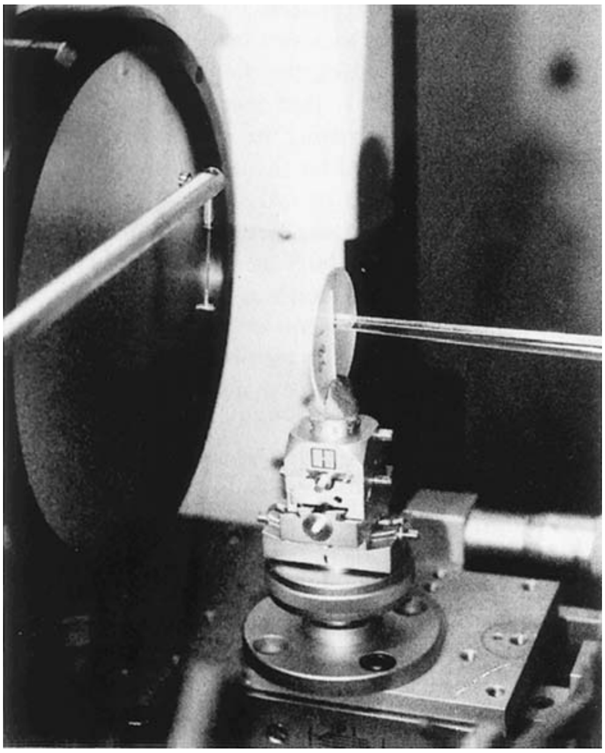


**Figure 8** Experimental arrangement for recording WAXS data with high spatial resolution. The specimen is mounted on a goniometer on an X/Y stage which allows tracking across the X-ray beam leaving the glass focusing collimator. Data is recorded on the CCD detector. A backstop to absorb unscattered X-rays is suspended immediately before the detector.

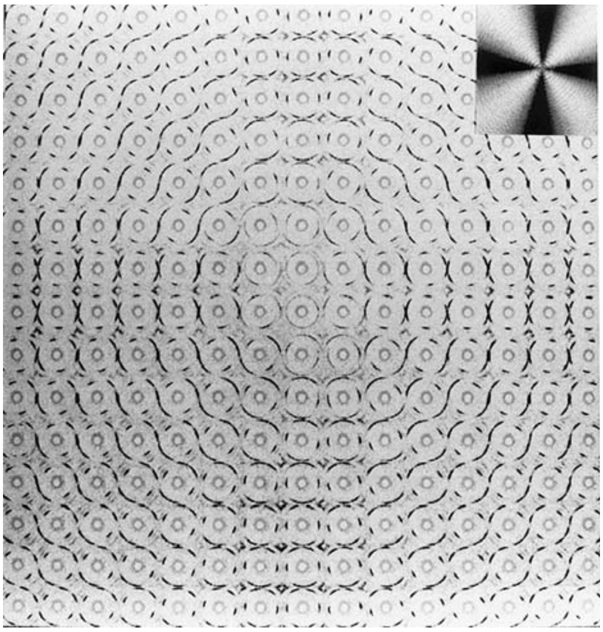


**Figure 9** X-ray fibre diffraction patterns from the region of the spherulite of Biopol illustrated in the inserted polarizing microscope image. Each pattern was recorded from an area 10 μm in diameter with its position in this figure corresponding to the position in the specimen from which it was recorded. Each crystallite is oriented with its *a*-axis radial; the degree of orientation decreases towards the centre of the spherulite.

polymer therefore experiences a wide variation in draw temperature, draw ratio and draw direction at different positions of the container and at different depths in the wall, resulting in a wide variation of polymer orientation and crystallinity.

The sample for X-ray microbeam examination was prepared by cutting a 1 mm square piece from the container and mounting it on an X/Y stage so that it could be tracked across the X-ray beam. **Figure 10** shows selected frames from a dataset comprising 32 frames. The patterns show a marked variation in crystallinity through the thickness of the wall. At the inside edge there are strong oriented crystalline reflections superimposed on a fairly isotropic amorphous halo. The intensification of the crystal reflections around the equator is consistent with preferential chain alignment along the longitudinal axis of the bottle. The intensity of the crystal reflections diminishes towards the centre of the wall thickness leaving an almost isotropic amorphous diffraction halo. Nearer the outer edge of the wall there is an intensification of the halo on the equator indicating an increasing degree of orientation in the disordered chains. These observations can be related to variation in the local draw rate, draw ratio and temperature across the wall during fabrication and provide a basis for refining the manufacturing process to produce bottles of improved performance.

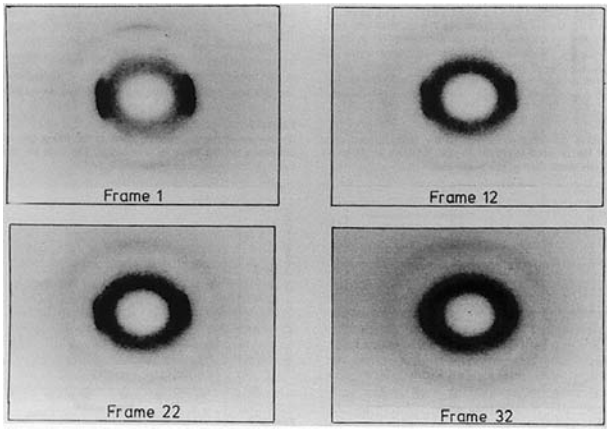



**Figure 10** Variation of crystallite size, crystallinity and polymer orientation as a function of position in the wall of a PET bottle. The X-ray diffraction patterns were selected from the 32 frames recorded during the experiment. Frame 1 was collected at the inside face of the bottle wall, whilst frame 32 was collected at the outside face.

*See also:* **Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; Materials Science Applications of X-Ray Diffraction; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Nucleic Acids Studied Using NMR; Polymer Applications of IR and Raman Spectroscopy; Powder X-Ray Diffraction, Applications; Small Molecule Applications of X-Ray Diffraction.**

## Further reading

Blundell DJ *et al.* (1996) Characterization of strain-induced crystallization of poly(ethylene terephthalate) at fast draw rates using synchrotron radiation. *Polymer* **37**: 3303.

Donald AM and Windle AH (1992) *Liquid Crystalline Polymers*. Cambridge: Cambridge University Press.

Harford J and Squire J (1997) Time-resolved diffraction studies of muscle using synchrotron radiation. *Reports of Progress in Physics* **60**: 1723–1787.

Hasnain SS (ed) (1990) *Synchrotron Radiation and Biophysics*. Chichester: Ellis Horwood Limited.

Hughes DJ *et al.* (1997) A simultaneous SAXS/WAXS and stress–strain study of polyethylene deformation at high strain rates. *Polymer* **38**: 6427–6430.

Mahendrasingam A *et al.* (1995) Microfocus X-ray diffraction of spherulites of poly-3-hydroxybutyrate. *Journal of Synchrotron Radiation* **2**: 308–311.

Mahendrasingam A *et al.* (1986) Time-resolved X-ray diffraction studies of the B to D structural transition in the DNA double helix. *Science* **233**: 195–197.

Martin C *et al.* (1997) Investigation of the variation in orientation and crystallinity in poly(ethylene terephthalate) containers using microfocus X-ray diffraction. *Journal of Synchrotron Radiation* **4**: 223–227.

Michette A and Pfauntsch S (eds) (1996) *X-rays: The First Hundred Years*. New York: Wiley. See in particular articles by Franks A (The First Hundred Years), Fuller W (X-ray Diffraction), Munro I (Synchrotron Radiation) and Miller A (The Next Hundred Years).

Ryan AJ and Wilkinson N (1997) *Polymer Processing and Structure Development*. Kluwer Academic Publishers.

# Field Ionization Kinetics in Mass Spectrometry

**Nico MM Nibbering**, University of Amsterdam, The Netherlands

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| **Methods & Instrumentation** |

## Introduction

The mass spectra of molecules, ionized by electron ionization (EI), field ionization (FI) or photoionization (PI) in the gas phase, usually show many peaks with varying intensities. These peaks result from fragment ions generated by a series of competing and consecutive unimolecular dissociation reactions of the molecular ions provided that the latter have obtained during their formation an amount of energy in excess of the ionization energy of the corresponding molecules. The resulting mass spectrum therefore is determined by the relative rates of these unimolecular dissociation reactions, which in their turn are related to the relative abundances of the fragment ions observed. It will be clear that studying the kinetics of ion dissociation can provide an invaluable insight into the foundation of a mass spectrum.

Although trapped-ion mass spectrometry has been used to investigate the extent of fragmentation of ions as a function of time in the time range of μs to ms, another and unique method is that of field ionization kinetics (FIK). The roots of this method date from the work of Beckey at the University of Bonn, Germany, who interpreted (as early as 1961) the broadening of peaks in field ionization (FI) mass spectra, obtained with a single-focusing mass spectrometer, as a consequence of ion lifetimes. Picosecond resolution of these ion lifetimes was achieved in 1971 by Derrick and Robertson at King's College in London, United Kingdom on the basis of a thorough calibration of the time-scale following FI. Since then the development of the experimental techniques to measure such ion lifetimes and of the necessary theory of the corresponding ion decomposition rates have led to the use of the name FIK. With this method the rates of gas phase unimolecular decompositions of ions, occurring at times as short as picoseconds and over a time range of seven orders of magnitude ($10^{-12}$–$10^{-5}$ s), and corresponding ion lifetimes can be measured.

Following a description of the principle of the FIK method, some illustrative examples of its application in gas-phase ion chemistry studies will be presented.

## Principle of the FIK method

FIK experiments are performed on double-focusing mass spectrometers having either the conventional (the electric sector precedes the magnetic sector) or the reversed (the magnetic sector precedes the electric sector) geometry. A schematic diagram of a field ionization double-focusing mass spectrometer with the conventional geometry is given in **Figure 1**.

Ions are generated in a very narrow region close to an anode (in the case of positive ions). This region becomes less narrow as the anode is changed from a tip, to a blade to a 1 μm bare tungsten wire or to a 10 μm activated tungsten wire (**Figure 2**) at a potential $V_0$ (~8-10 kV) with respect to a slotted cathode at ground potential and at a distance that is increased from 0.25 to 1.5 mm.

If these ions are stable for $\geq 10^{-5}$ s they will acquire a kinetic energy $qV_0$ during acceleration from the ionization region (where to a good approximation the potential is equal to the anode potential $V_0$) to the slotted cathode and will pass the electric sector set to transmit ions with a kinetic energy $qV_0$. Consequently, they will be mass analysed at their correct $m/z$ values if the magnetic field is scanned. In this way, a FI spectrum is obtained which contains peaks that correspond to the ions generated very near the anode, i.e. within approximately $10^{-11}$ s. Fragment ions $m^+$ (see Eqn [1]), generated in unimolecular dissociation reactions that occur between the anode and cathode (region I in **Figure 1**) at a potential $V_x$, will not be transmitted through the electric sector because they have insufficient kinetic energy $[q(m/M)(V_0 - V_x) + qV_x]$ to be focused by this sector.

$$M^{\bullet +} \rightarrow m^+ + (M - m)^\bullet \qquad [1]$$

However, these ions can be transmitted through the electric sector by either decreasing its potential at a constant anode potential $V_0$ or by increasing the anode potential at a constant electric sector potential set to transmit ions with a kinetic energy $qV_0$. The most widely used method involves increasing the anode potential.

The determination of the amount, $\Delta V$, with which the anode potential $V_0$ has to be increased to satisfy Equation [2] permits a calculation of the potential $V_x$ at which reaction [1] has occurred.

$$qV_0 = q(m/M)\,(V_0 + \Delta V - V_x) + q\,V_x \qquad [2]$$

It is then possible to calculate the ion lifetime, i.e. the time elapsed between formation of the molecular ion $M^{\cdot +}$ and its subsequent fragmentation at potential $V_x$, if the potential distribution between the anode and cathode is known. Thus, a scan of the anode potential achieved by increasing $\Delta V$ allows one to monitor reaction [1], following FI of M over a continuous time range of ~ $10^{-11}$- $10^{-9}$ s (region I in **Figure 1**). The required, time-consuming data processing can be speeded up by use of an online computer, which also enables computer controlled data acquisition and multiscan averaging to enhance the signal-to-noise ratio of the usually weak FIK signals.

Decompositions of $M^{\cdot +}$ ions at times longer than $10^{-9}$ s can be investigated in the focusing region II

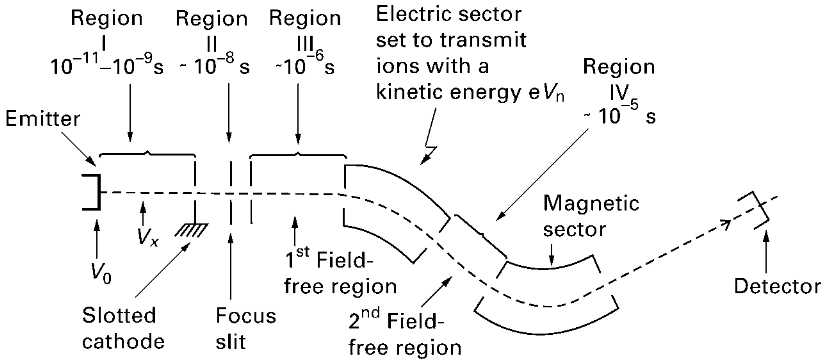



**Figure 1** Schematic diagram of a double-focusing field ionization mass spectrometer of conventional geometry (not drawn to scale), showing the four regions in which field ionization kinetic measurements can be made. The stated times refer to an ion source with a blade emitter. Reproduced with permission of Wiley from Nibbering NMM (1984) Mechanistic studies by field ionization kinetics. *Mass Spectrometry Reviews* **3**: 445–477.

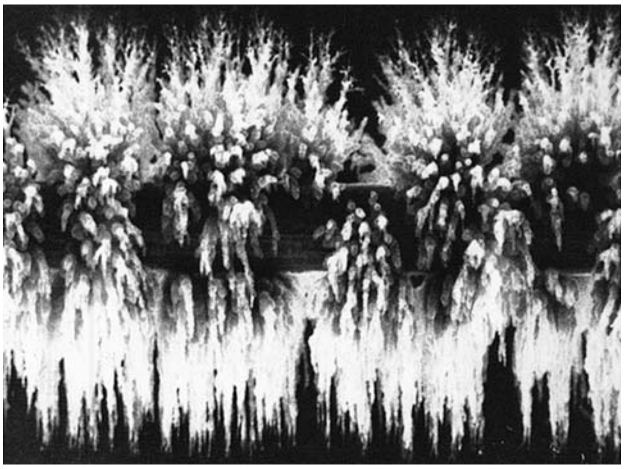


**Figure 2** Scanning electron micrograph of a 10 µm tungsten wire covered with carbon microneedles grown by activation.

and in the first and second field-free regions III and IV (Figure 1) by using metastable ion methods.

The experimental results take the form of curves of product ion current $I_m(t)$ versus kinetic energy, which are transformed according to the principles outlined above into curves of product ion current $I_m(t)$ versus reaction time $t$. The product ion current $I_m(t)$ is related to the rate of reaction, $dI_m(t)/dt$, by Equation [3], where $\Delta t$ is a small interval within which decomposing $M^{\bullet+}$ ions contribute to the product ion current $I_m(t)$ at time $t$:

$$\frac{dI_m(t)}{dt} = \frac{I_m(t)}{\Delta t} \quad [3]$$

This interval is determined by the energy resolution of the electric sector and depends on the molecular and product ion masses and on the reaction time $t$, i.e. this time window interval will decrease with increasing mass difference between the molecular and product ions (not unlimited because of the difficulties in increasing the anode potential due to arcing problems) and at shorter decomposition times of the molecular ions because of the increase of the electric field gradient with decreasing distance from the anode.

Division of Equation [3] by the measured mass-resolved molecular ion current $I_M$ at an anode potential $V_0$, i.e. in the normal FI spectrum, yields the so-called normalized rate $k(t)$:

$$k(t) = I_m(t)/I_M\Delta t \quad [4]$$

Plots of $k(t)$ versus $t$ obtained with different mass spectrometers, different anodes and different ion lifetime calculations have been shown to be reproducible to within 25%.

The normalized rate $k(t)$ is related to the time-independent rate constant $k(E)$ as defined by the quasi-equilibrium theory (QET) in the following way. Suppose that the molecular ions $M^{\bullet+}$ decompose via a single channel to fragment ions $m^+$. If the $M^{\bullet+}$ ions have acquired an internal energy $E$ upon FI, their abundance $M(t)$ at time $t$ will be given by

$$M(t) = M_0 e^{-k(E)t} \quad [5]$$

where $M_0$ is the abundance of the molecular ions at the time of their formation. The rate of formation of the fragment ions $m^+$ is then given by

$$dm^+/dt = k(E)M(t) = k(E)M_0 e^{-k(E)t} \quad [6]$$

The molecular ions $M^{\bullet+}$, however, will have acquired a distribution of internal energies, $P(E)$, upon FI. In this case all individual rates of formation of $m^+$ given by Equation [6] have to be summed with respect to the internal energy $E$. This leads to Equation [7] with $P(E)$ as the weighting function and $I_m(t)$ as the product ion current at time $t$ (see above):

$$\frac{dI_m(t)}{dt} = \int_0^\infty P(E)k(E)M_0 e^{-k(E)t} dE \quad [7]$$

Similarly, the molecular ion current $I_M$ is obtained by integration of Equation [5] with respect to the internal energy $E$, taking $P(E)$ as the weighting function:

$$I_M = \int_0^\infty P(E)M_0 e^{-k(E)t}\, dE \quad [8]$$

Division of Equation [7] by Equation [8] then gives the relation between the normalized rate $k(t)$ and the time-independent rate constant $k(E)$:

$$k(t) = \frac{\int_0^\infty P(E)k(E)e^{-k(E)t}\, dE}{\int_0^\infty P(E)e^{-k(E)t}\, dE} \quad [9]$$

Usually a maximum is observed in the $k(t)$ versus $t$ plots. The maximum has been suggested to represent

$k_{max}(E)$ in the distribution of $k(E)$s and generally occurs at longer times for rearrangement reactions than for simple cleavage reactions. These reactions usually compete with each other continuously over the time range of $10^{-11}$ - $10^{-5}$ s.

An example of a logarithmic plot of $k(t)$ versus $t$ for the loss of a methyl radical from the molecular ion of pent-3-en-2-ol, $CH_3CH{=}CHCH(OH)CH_3$, is given in Figure 3.

The figure shows two maxima. The first maximum at a molecular ion lifetime of $10^{-10.5}$ s corresponds to the loss of a methyl radical from the unrearranged molecular ion via a direct cleavage reaction, whereas the second maximum at a molecular ion lifetime of $10^{-10.1}$ s is owing to the loss of a methyl radical following a rearrangement reaction of the molecular ion.

The strength of FIK in mechanistic gas-phase ion dissociation studies compared with the conventional EI method is that it gives a time-resolved view of the chemistry and isomerization reactions of ions over a time range of $10^{-12}$ - $10^{-5}$ s. The details of these processes are largely lost in the EI method which integrates all the events up to $10^{-6}$ s following ionization, a very broad time range compared with the vibrational periods of chemical bonds ($10^{-14}$ - $10^{-13}$ s).

The strength of the FIK method in mechanistic ion dissociation studies is shown to full advantage when it is used in combination with stable isotopic labelling. In that case, the fragment ions from competing reactions are nearly equal in mass so that the ratios of their ion currents measured at any time $t$ are practically equal to the ratios of the normalized rates $k(t)$ (see above and Eqn [4]). Most of the mechanistic ion dissociation studies discussed below are therefore based upon changes in relative rates as a function of the ion lifetimes.

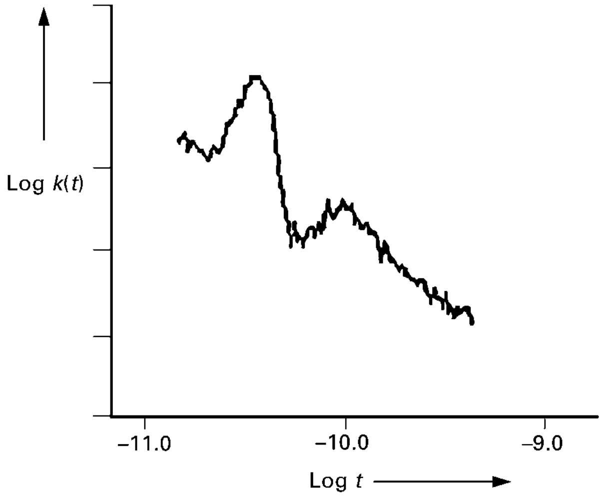



**Figure 3** Normalized rate constant $k(t)$ for the loss of a methyl radical from the molecular ions of pent-3-en-2-ol, $CH_3CH{=}CHCH(OH)CH_3$. The vertical axis covers log $k(t)$ values in the range of 8–12. Reproduced with permission from Zappey HW, Ingemann S and Nibbering NMM (1991) Study of the rate of methyl loss and the structure of the product ion as a function of the lifetime of the molecular ion of penten-3-en-2-ol. *Organic Mass Spectrometry* **26**: 241–246.

## Mechanistic ion dissociation studies

Four examples of the application of FIK to mechanistic ion dissociation studies will be presented.

The first deals with the elimination of ethylene from the molecular ion of cyclohexene, the second and third with the dynamics of stereochemically controlled dissociation reactions and the fourth with time-resolved MS-MS to identify the structures of product ions generated at different molecular ion lifetimes.

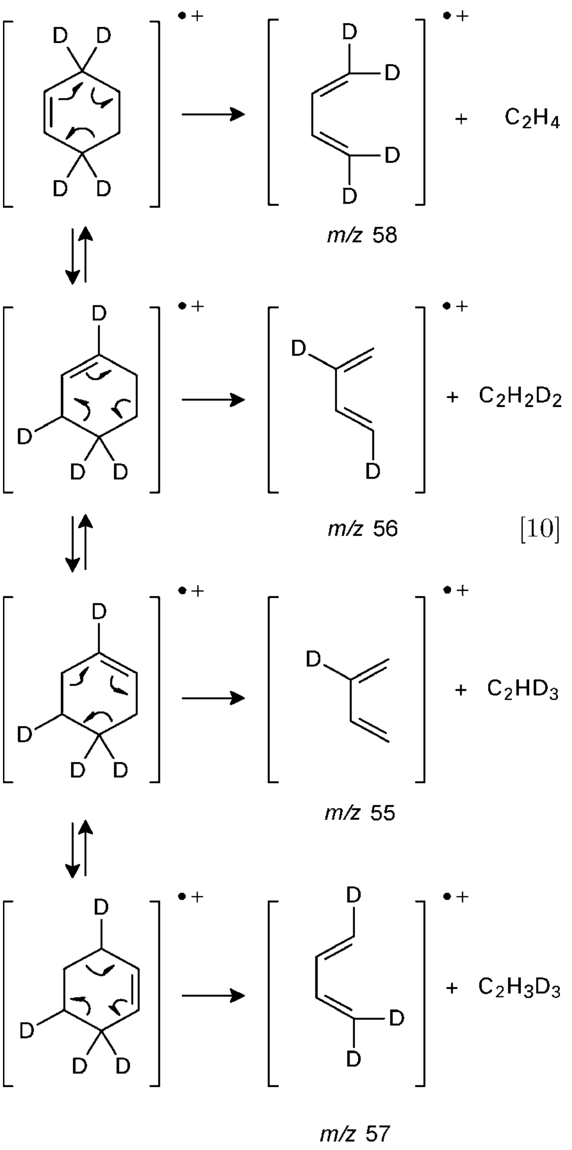
### Elimination of ethylene from the molecular ion of cyclohexene

In both linear and cyclic alkene EI-generated molecular ions, the double bond shifts before or during fragmentation, cause the mass spectra of isomers to be very similar, if not indistinguishable. The same is true for the EI generated molecular ion of cyclohexene, in which specific D-labelling has shown a practically statistical distribution of H and D atoms over the carbon skeleton before or during fragmentation.

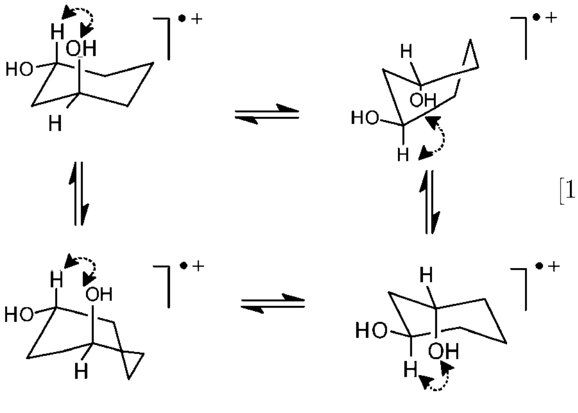
Upon FI, four variously deuterated ethylene molecules, $C_2H_4$, $C_2H_3D$, $C_2H_2D_2$ and $C_2HD_3$ are eliminated from cyclohexene-3,3,6,6-$d_4$.

At $10^{-11}$ s, more than 50% is lost as $C_2H_4$, but at longer times a rapid increase of the loss of the deuterated ethylene molecules is observed until at $10^{-9}$ s a statistical distribution of the H and D atoms has occurred before decomposition. The specific loss of $C_2H_4$ at $10^{-11}$ s can readily be explained by a retro-Diels- Alder reaction of the molecular ion to give the ion $m/z$ 58 (Eqn [10]). The interesting observation made, however, is that the maximum rates of the variously labelled ethylene eliminations occur at different times; the order with increasing time is $C_2H_4 < C_2H_2D_2 < C_2HD_3 < C_2H_3D$. This has been rationalized by invoking 1,3-allylic H/D shifts in the molecular ions before the retro-Diels- Alder reaction (Eqn [10]).

As can be seen from Equation [10], the formation of $m/z$ 58, 56, 55 and 57 requires 0, 1, 2 and 3 allylic 1,3-H/D shifts, respectively, which coincides with the shift of the maximum rate of the corresponding ethylene eliminations to longer times.




[10]

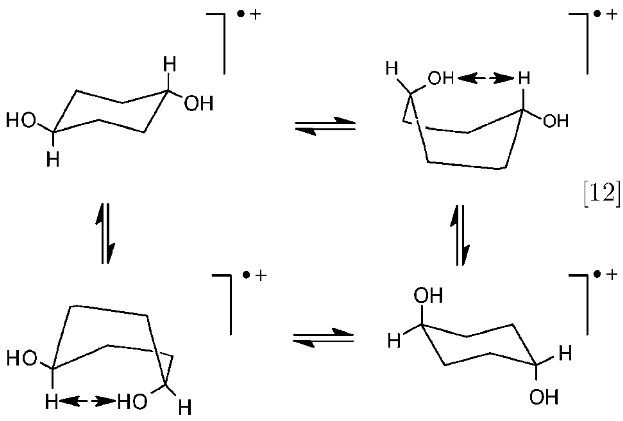


[11]

molecular ions predominantly lose water by a transannular 1,3- and 1,4-elimination, respectively, as determined by using FIK. In both cases the normalized rate for this process reaches a maximum value at an ion lifetime of about 110 ps, but the rate is also significantly higher for the *trans*-1,3-diol than for the *trans*-1,4-diol at shorter ion lifetimes. The explanation for these observations is the following: both in the molecular ion of the *trans*-1,3-diol and in that of the *trans*-1,4-diol the hydroxyl group has to come in close proximity of the hydrogen atom to be transferred to effect the water elimination. This is the case for the two identical ground-state chair conformations of the *trans*-1,3-diol, which have a distance of 2.5 Å between the oxygen atom of the hydroxyl group and the hydrogen atom to be transferred, and for the corresponding boat conformation, as shown in Equation [11].

However, for the *trans*-1,4-diol, only the boat conformations have the appropriate geometry for the transannular 1,4-elimination of water, as shown in Equation [12].

Similar arguments apply to the enhanced transannular loss of water from *cis*-cyclohexane-1,3-diol versus that from the *cis*-1,4-diol. The departing water now contains both hydroxyl hydrogen atoms,



[12]

### Loss of a molecule of water from the molecular ions of stereoisomeric cyclohexane diols

An FIK study of the water loss from stereoisomeric cyclohexane diols has shown how elegantly this method can give an insight into the dynamics of unimolecular decompositions of organic ions. This is especially true for the *trans*- and *cis*-1,3- and 1,4-diols, which generate relatively abundant $(M-H_2O)^{\cdot+}$ ions upon FI, in contrast with the *trans*- and *cis*-1,2-diols. The latter compounds yield low-abundant $(M-H_2O)^{\cdot+}$ ions, which is probably owing to a fast cleavage of the $C_1$-$C_2$ bond of the molecular ions.

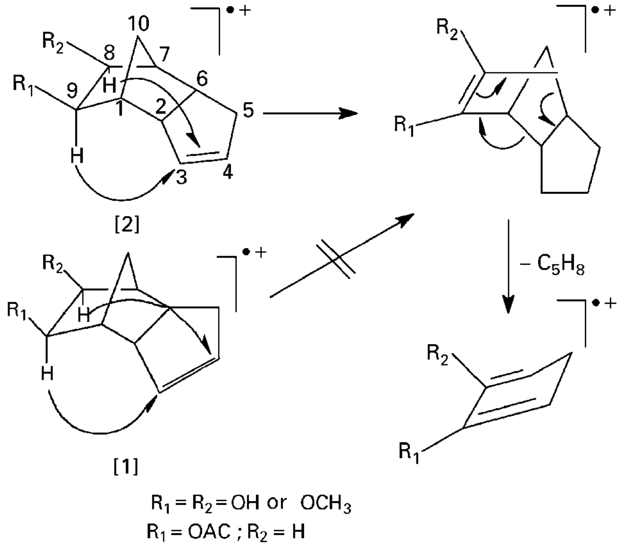
The 1,3-$d_2$ analogue of *trans*-cyclohexane-1,3-diol and the 1,4-$d_2$ analogue of the *trans*-1,4-diol

as shown by D labelling. In both cases the normalized rate for this reaction again reaches a maximum value at ion lifetimes of about 110 ps.

### Dyotropic hydrogen rearrangement in the molecular ions of 8,9-disubstituted-tricyclo[5.2.1.0$^{2,6}$] decenes

Upon EI the *exo* and *endo* isomers of the 8,9-disubstituted-tricyclo[5.2.1.0$^{2,6}$]decenes [1] and [2], respectively, fragment essentially differently. The ionized *endo* isomers [2] eliminate $C_5H_8$ via a retro-Diels–Alder reaction after migration of the hydrogen atoms from positions 8 and 9 to the double bond. This reaction is not observed for the *exo* isomers [1] where the distance between the hydrogen atoms of positions 8 and 9 and the double bond is too large for the hydrogen migration required to effect the retro-Diels–Alder reaction (Scheme 1).

From these observations, however, it cannot be concluded whether the double hydrogen migration in ionized [2] occurs in a concerted or stepwise fashion. If the hydrogen atoms would migrate in a concerted fashion, then it is an example of a dyotropic reaction, being defined as an uncatalysed process in which two σ-bonds simultaneously, but not necessarily via a fully symmetrical mode, migrate intramolecularly.




$R_1 = R_2 = OH$ or $OCH_3$
$R_1 = OAC$ ; $R_2 = H$

**Scheme 1** Double hydrogen transfer in the molecular ion of the *endo* isomer [2] of 8,9-disubsituted-tricyclo[5.2.1.0$^{2.6}$]decene (not possible in the molecular ion of the *exo* isomer [1] followed by elimination of $C_5H_8$ via a retro-Diels–Alder reaction. Reproduced with permission of the American Chemical Society from Kluft E, Nibbering NMM, Kühn H and Herzschuh R (1986) Dyotropic hydrogen rearrangement in the molecular ions of 8,9-disubstituted tricyclo [5.2.1.0$^{2,6}$] decenes as supported by field ionization kinetics. *Journal of the American Chemical Society* **108**: 7201–7203.

Such a suprafacial dyotropic transfer of hydrogen is a thermally 'allowed' $[_{\sigma}2_s + _{\sigma}2_s + _{\pi}2_s]$ pericyclic reaction.

Experimental support for either a concerted or a stepwise double hydrogen migration in ionized [2] before elimination of $C_5H_8$ via a retro-Diels–Alder reaction (Scheme 1) is obtained from application of FIK to both *exo* and *endo* isomers [1] and [2] respectively.

For example, the *exo* isomer [1a] with $R^1 = R^2 = OCH_3$ appears to eliminate, exclusively, $C_5H_6$ over all molecular ion lifetimes studied, while the corresponding *endo* isomer [2a] eliminates both $C_5H_6$ and $C_5H_8$, but not $C_5H_7^{\bullet}$, see Figures 4A and 4B, respectively.

As expected, the $C_5H_8$ loss is a delayed process with respect to the $C_5H_6$ loss because of the required double hydrogen rearrangement, i.e. this channel starts to compete with the $C_5H_6$ loss at molecular ion lifetimes of $\sim 10^{-10}$ s and becomes dominant beyond molecular ion lifetimes of $\sim 10^{-9.25}$ s. Important, however, is the absence of $C_5H_7^{\bullet}$ loss, which is expected to occur for a stepwise, but not for a concerted, double hydrogen rearrangement.

### Methyl loss from ionized pent-3-en-2-ol, $CH_3CH{=}CHCH(OH)CH_3$

A combined FIK and deuterium labelling study has revealed three pathways for the loss of methyl from the molecular ion of pent-3-en-2-ol, $CH_3CH{=}CHCH(OH)CH_3$. This is shown by the relative rate curves for loss of $CH_3^{\bullet}$ and $CD_3^{\bullet}$ from the molecular ion of pent-3-en-2-ol-1,1,1-$d_3$, $CH_3CH{=}CHCH(OH)CD_3$ in Figure 5.

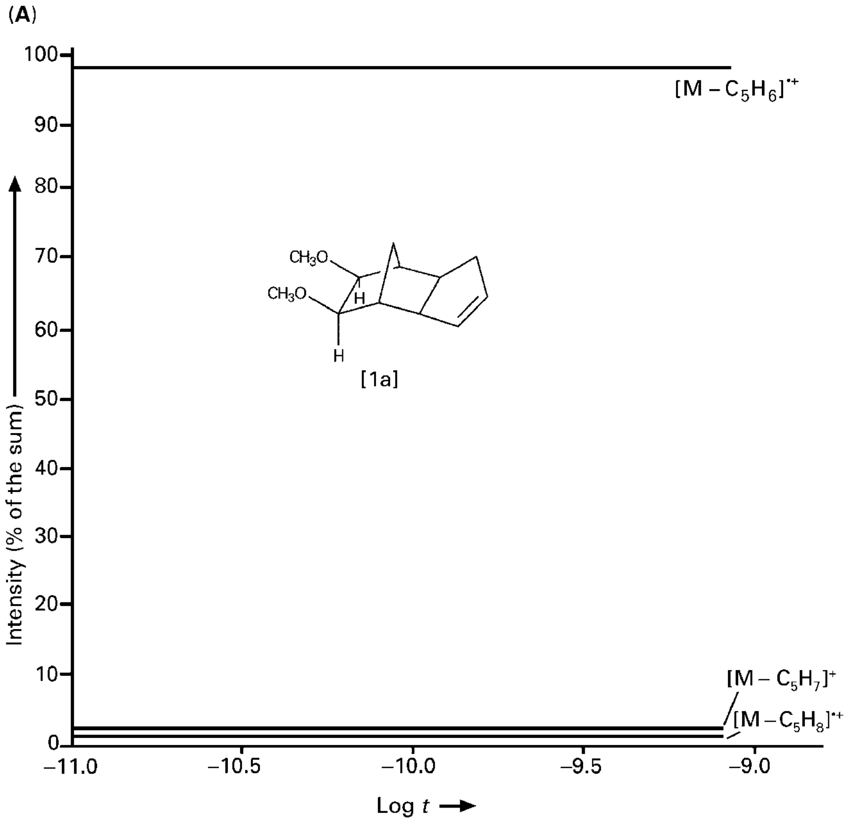
At short ion lifetimes, $CD_3^{\bullet}$ is the major loss, whereas at medium ion lifetimes $CH_3^{\bullet}$ elimination is the major channel. At longer ion lifetimes the $CD_3^{\bullet}$ loss again becomes increasingly important.

These observations have been rationalized mechanistically in Scheme 2.

Route [3] → [4] is operative at ion lifetimes $< 10^{-10.5}$ s and is suggested to lead to oxygen protonated crotonaldehyde, route [3] → [5] → [6] → [7] is dominant at ion lifetimes between $10^{-10.5}$ and $10^{-9}$ s and is suggested to lead to oxygen protonated methyl vinyl ketone, while route [3] → [5] → [8] → [9] increases at longer ion lifetimes and is suggested to produce the butyryl cation (Scheme 2).

The existence of different pathways is also indicated by the normalized rate curve for methyl loss from the molecular ion of unlabelled pent-3-en-2-ol, which contains two distinct maxima at ion lifetimes of $10^{-10.5}$ and $10^{-10.1}$ s after ionization (Figure 3).

In the normalized rate curve for $CD_3$ loss from the molecular ion of $CH_3CH{=}CHCH(OH)CD_3$ similar



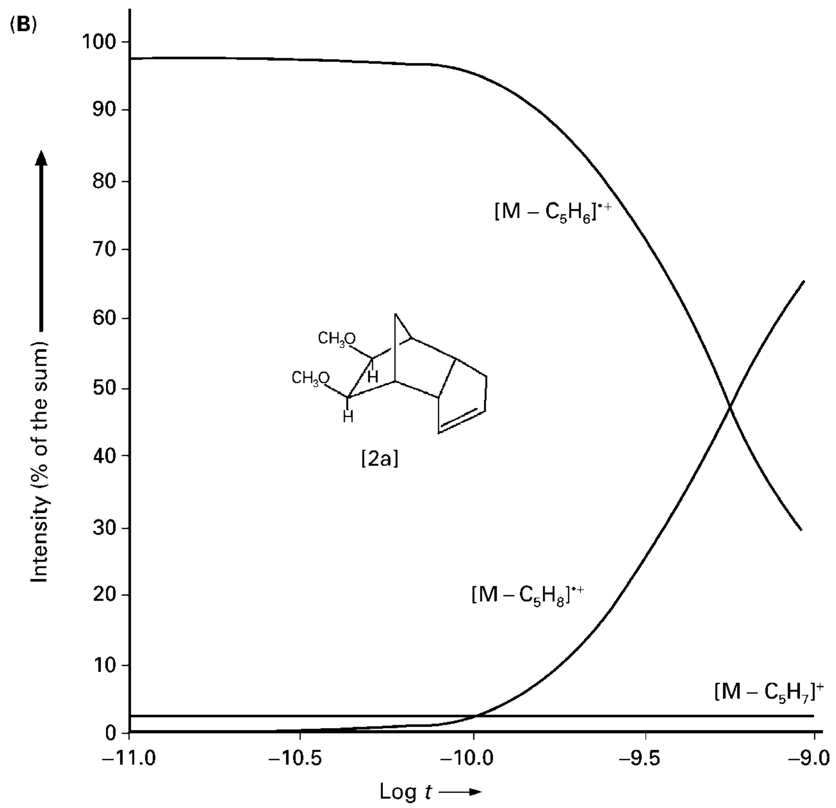
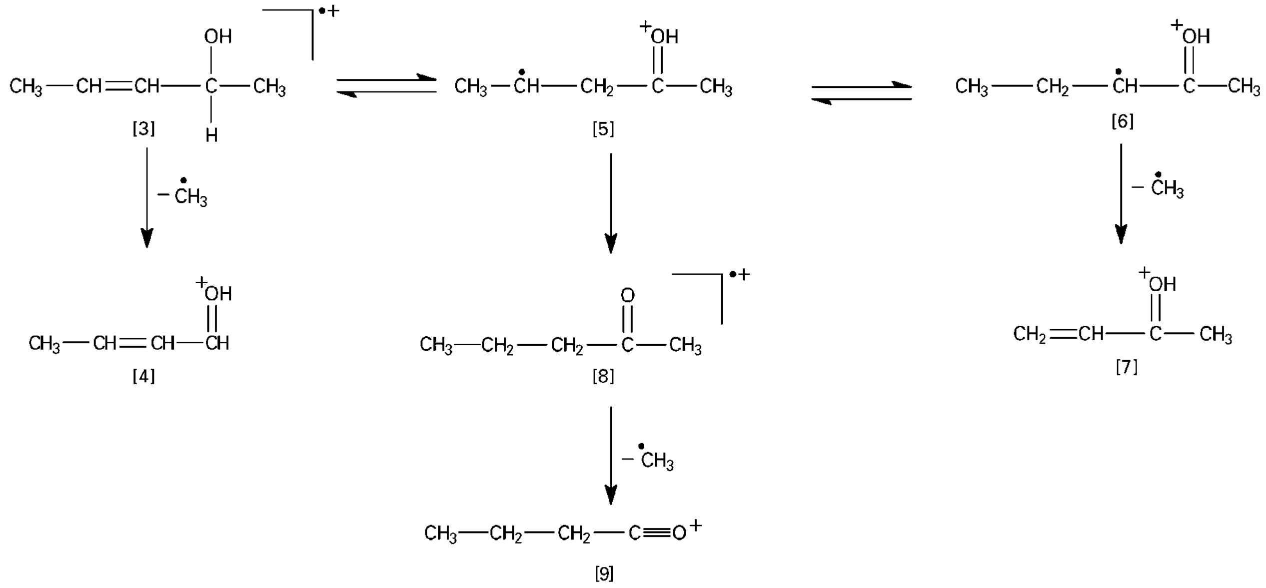




**Figure 4** Losses of $C_5H_6$, $C_5H_7^{\bullet}$ and $C_5H_8$, expressed in percentages of their sum, as a function of the lifetime of the molecular ions of (A) *exo-* and (B) *endo*-8,9-dimethoxytricyclo[5.2.1.0$^{2,6}$]decenes. Reproduced with permission of the American Chemical Society from Kluft E, Nibbering NMM, Kühn H and Herzschuh R (1986) Dyotropic hydrogen rearrangement in the molecular ions of 8,9-disubstituted tricyclo [5.2.1.0$^{2,6}$] decenes as supported by field ionization kinetics. *Journal of the American Chemical Society* **108**: 7201–7203.




**Scheme 2** Routes for methyl loss from ionized pent-3-en-2-ol. Reproduced with permission from Zappey HW, Ingemann S and Nibbering NMM (1991) Study of the rate of methyl loss and the structure of the product ion as a function of the lifetime of the molecular ion of penten-3-en-2-ol. *Organic Mass Spectrometry* **26**: 241–246.

maxima are observed, but in that for $CH_3$ loss only the maximum at $10^{-10.1}$ s is present. This means that the maximum at $10^{-10.5}$ s corresponds to the route [3] → [4] of Scheme 2, as expected for such a direct cleavage reaction, and that the maximum at $10^{-10.1}$ s is owing to the route [3] → [5] → [6] → [7] of Scheme 2. The ions [4 and 7] (and also [9]) formed

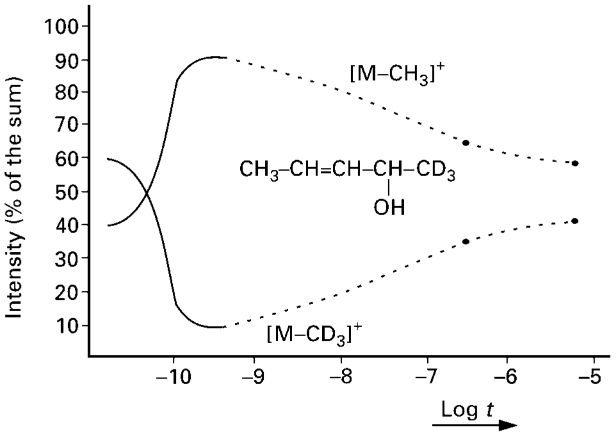



**Figure 5** Losses of $CH_3^{\bullet}$ and $CD_3^{\bullet}$ expressed in percentages of their sum, as a function of the lifetime of the molecular ion of pent-3-en-2-ol-1,1,1-$d_3$, $CH_3CH{=}CHCH(OH)CD_3$. Reprinted from the *International Journal of Mass Spectrometry and Ion Physics* **38**: Zwinselman JJ, Nibbering NMM, Middlemiss NE, Vajda JH and Harrison AG, A field ionization kinetics and metastable ion study of the fragmentation of some pentenols, 163–179, Copyright 1981, with permission from Elsevier Science.

by the different pathways are known to be non-interconverting species with distinct collision-induced dissociation mass spectra. Collision-induced dissociation experiments on the (M - methyl)$^+$ ions, formed at different molecular ion lifetimes, i.e. at times $< 10^{-10.6}$ and $10^{-10.0}$ s after ionization, have provided experimental support for their structures. The resulting time-resolved MS-MS spectra, in which the peaks are rather noisy owing to their low intensities, are presented in Figure 6.

The collision-induced dissociation MS-MS spectra of the independently made reference ions obtained by protonation of crotonaldehyde (ion [4] in Scheme 2) and methyl vinyl ketone (ion [7] in Scheme 2) are shown in Figure 7.

Figure 6A appears to match well Figure 7A which provides evidence that most of the (M- methyl)$^+$ ions formed at about $10^{-10.6}$ s have the protonated crotonaldehyde structure, i.e. ion [4] in Scheme 2.

Figure 6B is nearly identical to Figure 7B, showing that the (M- methyl)$^+$ ions at about $10^{-10.0}$ s have indeed, predominantly, the protonated methyl vinyl ketone structure, i.e. ion [7] in Scheme 2.

## List of symbols

$E$ = internal energy; $I_m$ = product ion current; $I_M$ = molecular ion current; $k(t)$ = normalized rate constant; $k(E)$ = time-independent rate constant; $t$ = time; $V$ = potential.

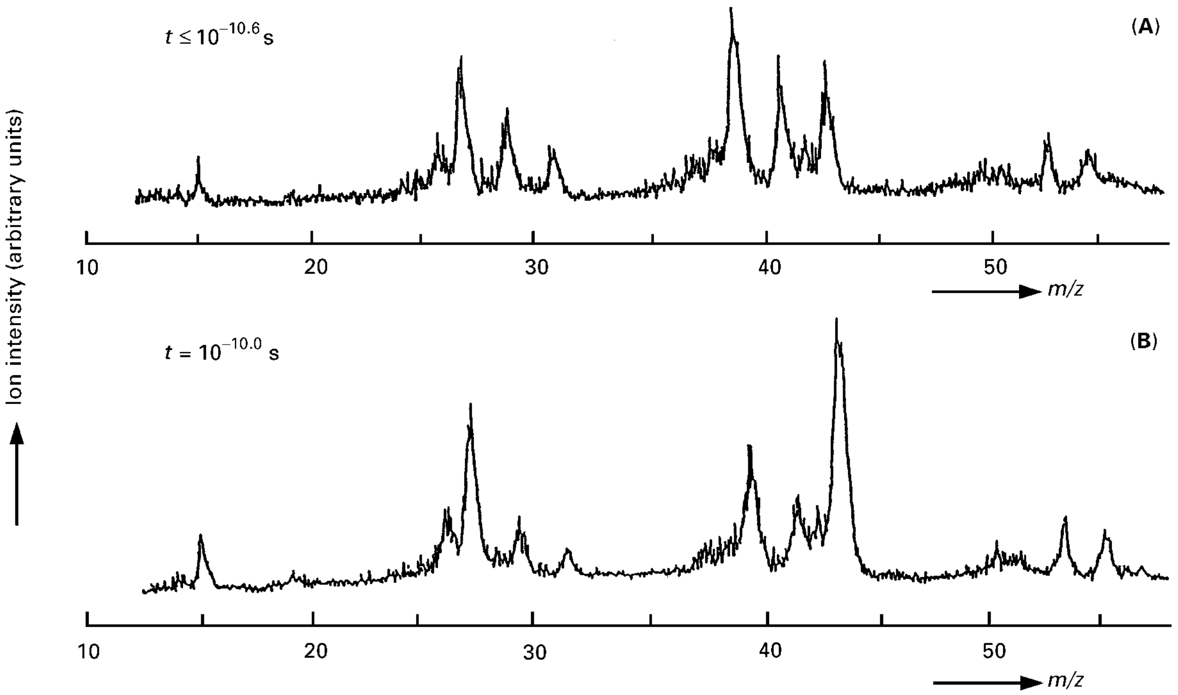



**Figure 6** Collision-induced dissociation MS-MS spectra of (M–methyl)$^+$ ions, generated upon FI from the molecular ion of pent-3-en-2-ol at molecular ion lifetimes of (A) ≤ $10^{-10.6}$ s and (B) $10^{-10.0}$ s. Reproduced with permission from Zappey HW, Ingemann S and Nibbering NMM (1991) Study of the rate of methyl loss and the structure of the product ion as a function of the lifetime of the molecular ion of penten-3-en-2-ol. *Organic Mass Spectrometry* **26**: 241–246.

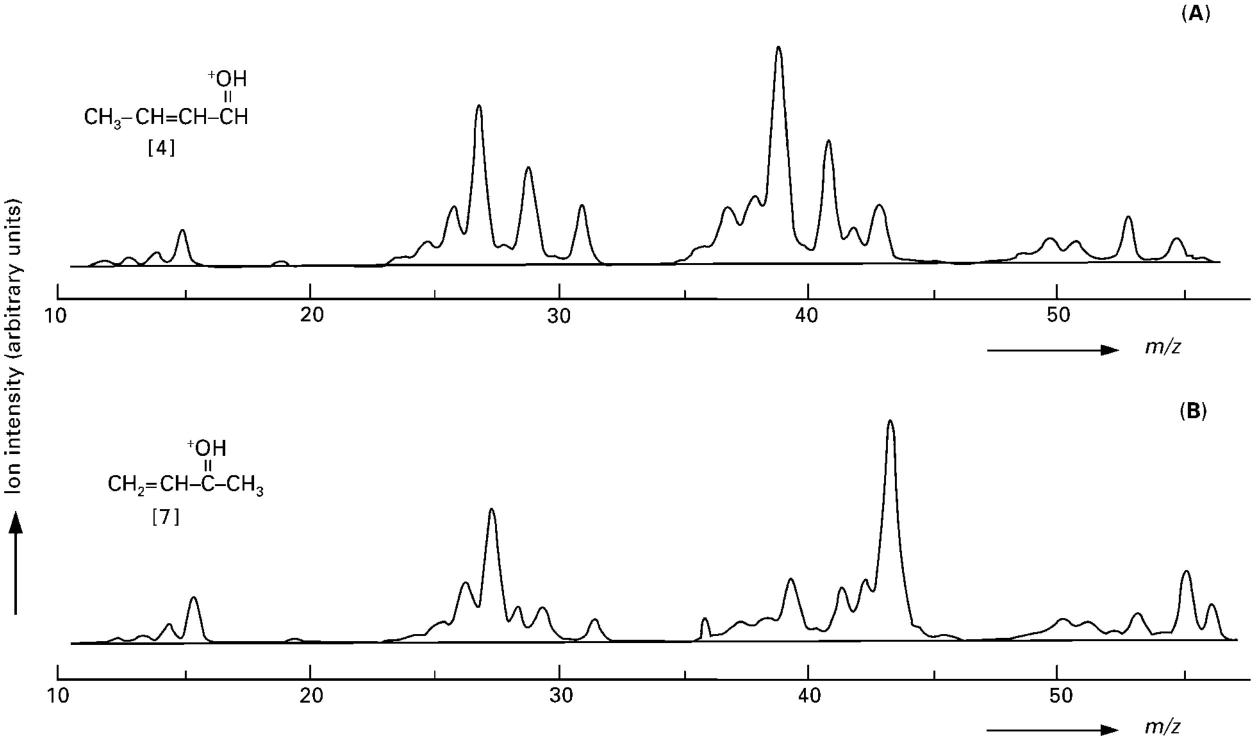



**Figure 7** Reference collision-induced dissociation MS-MS spectra of (A) protonated crotonaldehyde [4] and (B) protonated methyl vinyl ketone [7]. Reproduced with permission from Zappey HW, Ingemann S and Nibbering NMM (1991) Study of the rate of methyl loss and the structure of the product ion as a function of the lifetime of the molecular ion of penten-3-en-2-ol. *Organic Mass Spectrometry* **26**: 241–246.

*See also:* **Field Ionization Kinetics in Mass Spectrometry; Fragmentation in Mass Spectrometry; Ion Dissociation Kinetics, Mass Spectrometry; Ion Structures in Mass Spectrometry; Ionization Theory; Isotopic Labelling in Mass Spectrometry; Metastable Ions; MS–MS and MS*n*; Sector Mass Spectrometers; Statistical Theory of Mass Spectra; Stereochemistry Studied Using Mass Spectrometry.**

## Further reading

Beckey HD (1961) Measurement of extremely short decay times of organic ions with the field ionization mass spectrometer. *Zeitschrift für Naturforschung* **16a**: 505–510.

Beckey HD (1977) *Principles of Field Ionization and Field Desorption Mass Spectrometry*. Oxford: Pergamon.

Derrick PJ, Falick AM and Burlingame AL (1972) Mechanistic and kinetic studies at $10^{-11}$–$10^{-5}$ sec of the unimolecular gas-phase reactions induced by field ionization of cyclohexene and cyclohexene-3,3,6,6-$d_4$. *Journal of the American Chemical Society* **94**: 6794–6802.

Derrick PJ (1977) Field ionization and field desorption. In: Johnstone RAW (ed) *Specialist Periodical Reports on Mass Spectrometry*, Vol 4, pp 132–145. London: The Chemical Society.

Levsen K (1978) *Fundamental Aspects of Organic Mass Spectrometry*. Weinheim: Verlag Chemie.

Nibbering NMM (1984) Mechanistic studies by field ionization kinetics. *Mass Spectrometry Reviews* **3**: 445–477.

Nibbering NMM (1991) Time-resolved studies of unimolecular gas phase ion decompositions by field ionization kinetics. In: Jennings KR (ed) *Fundamentals of Gas Phase Ion Chemistry*, pp 333–349. Dordrecht: Kluwer Academic Publishers.

# Flame and Temperature Measurement Using Vibrational Spectroscopy

**Kevin L McNesby**, Citrus Heights, CA, USA

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES

**Applications**

## Introduction

The application of modern spectroscopic analysis to the study of flames began in the 1920s and 30s with breakthroughs in the understanding of atomic and molecular spectroscopy. The earliest spectroscopic investigations focused on understanding the line and band structure observed in the visible and ultraviolet regions of the spectrum when the light from the flame was dispersed by a grating or prism. One of the great breakthroughs of physics during this period was the understanding that the band structure observed in the emission spectra of flames originated from gas-phase molecular species.

For molecular species, understanding the appearance of flame spectra is simplified by assuming that the total internal energy, $E_T$, of a gas-phase molecule may be given to first order (Born Oppenheimer approximation) as:

$$E_T = E_{EL} + E_{VIB} + E_{ROT} \qquad [1]$$

Here, $E_{EL}$, $E_{VIB}$, and $E_{ROT}$ are respectively the quantized electronic, vibrational, and rotational energy of the molecular species. In the scientific literature, these energies are usually expressed in terms of wavenumbers (the reciprocal wavelength, expressed in centimetres, which is directly proportional to energy, and given the symbol $cm^{-1}$). All observed spectral features, in emission and absorption, are caused by changes in total energy, $\Delta E_T$, of the individual species present within the flame. Spectral features arise from the emission or absorption of a photon with energy corresponding to the difference between initial and final states of the transition. **Figure 1** illustrates the absorption of a photon corresponding to an electronic transition (**Figure 1A**) and to a vibrational transition (**Figure 1B**) for a diatomic molecule whose interatomic potential energy may be approximated by an anharmonic oscillator.

In general, changes in electronic, vibrational, and rotational energy correspond to emission or absorption of radiation in the visible (and ultraviolet), infrared, and microwave region of the electromagnetic spectrum, respectively. Typically, $\Delta E_{EL} > \Delta E_{VIB} > \Delta E_{ROT}$. For molecular species, changes in $E_{EL}$ generally are accompanied by changes in $E_{VIB}$ and

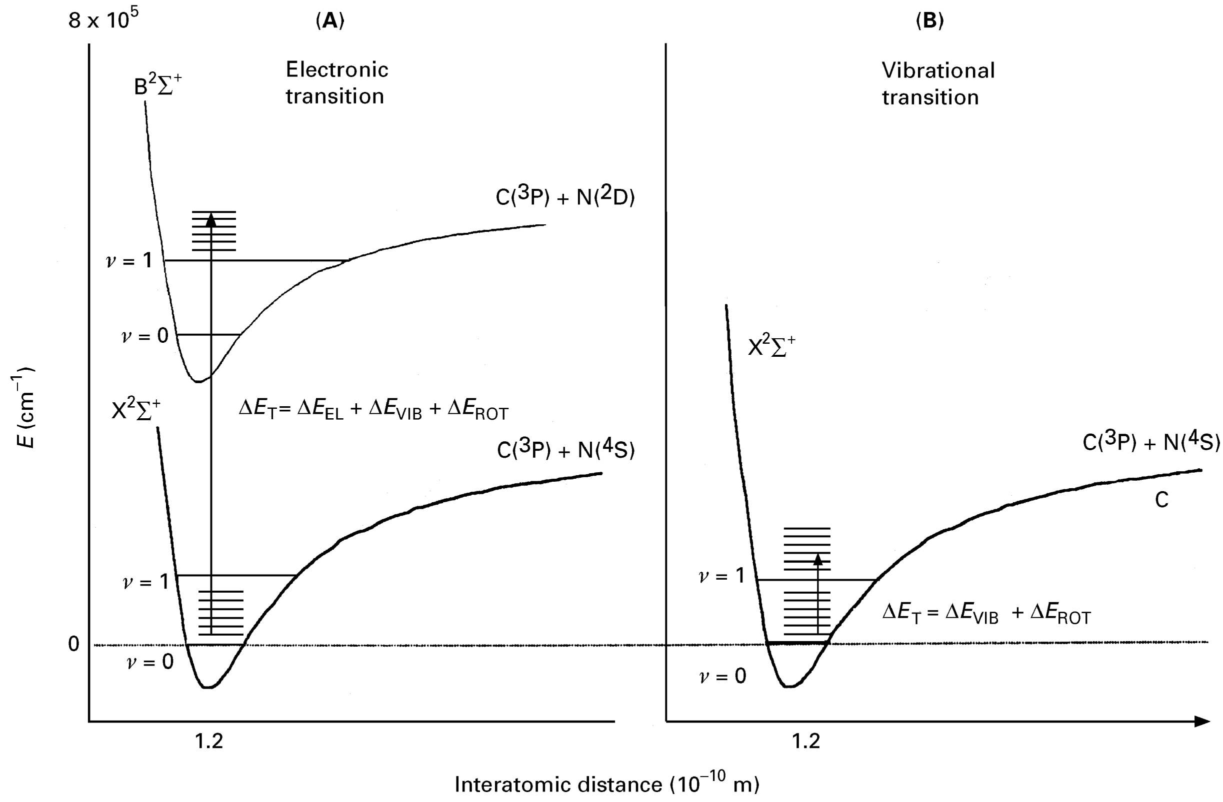



**Figure 1** Potential energy diagram for a diatomic anharmonic oscillator. Electronic (A) and vibrational (B) absorption transitions are illustrated. Values for interatomic distance, energy ($E$), and term symbols are for the radical flame species CN.

$E_{ROT}$, and changes in $E_{VIB}$ may be (and almost are) accompanied by changes in $E_{ROT}$. These general trends are illustrated in **Figure 1**. The band structure observed in the visible and ultraviolet spectra of molecular species in flames, in emission and absorption, is therefore understood to be the result of combined changes in electronic, vibrational, and rotational energies.

The visible and ultraviolet radiation from most flames usually accounts for less then 1% of the total emitted energy, with most of the energy emitted by a flame occurring in the infrared region of the spectrum. To see why this is not unexpected, it is useful to compare the radiation emitted by a flame with the radiation emitted by a blackbody. In a system at thermodynamic equilibrium (which, on a macroscopic scale, a flame is not), the distribution of radiation is given by Planck's blackbody radiation law:

$$I_\lambda = [2E_\lambda c_1 \lambda^{-5}/(\exp[c2/\lambda T] - 1)]A \, d\lambda \qquad [2]$$

Here, $I_\lambda$ is the wavelength-dependent radiant intensity normal to the surface of the radiator, $E_\lambda$ is the emissivity at wavelength $\lambda$, $c_1$ and $c_2$ are the first and second radiation constants and have the values $0.588 \times 10^{-8}$ Wm$^{-2}$ and $1.438 \times 10^{-2}$ m K$^{-1}$, respectively, $T$ is temperature in kelvin (K) and $A$ is the area of the surface in square metres (m$^2$).

**Figure 2** is a plot of Equation [2] at several temperatures. **Figure 2** shows that for a blackbody radiator, at temperatures up to 2300 K, the peak spectral radiancy always occurs in the infrared region of the spectrum. As the temperature increases above 2300 K, the peak of spectral radiancy moves to shorter wavelengths (towards the visible region of the spectrum). It is important to note that for a blackbody, the value of $E_\lambda$ is equal to 1 at all wavelengths. In flames, the value of $E_\lambda$ varies (and may approach 1) with wavelength, but is near zero for most wavelengths, indicating that a flame is not a blackbody, and that flame gases may not be in thermodynamic equilibrium. Nevertheless, in some cases where the emissivity of a flame species is known, measurements of spectral radiancy of flames may be used to calculate flame temperatures to within an accuracy of several kelvin.

When the regions of non-zero emissivity ($E_\lambda > 0$) in the flame emission spectrum are expanded along the wavelength scale, these regions exhibit detailed

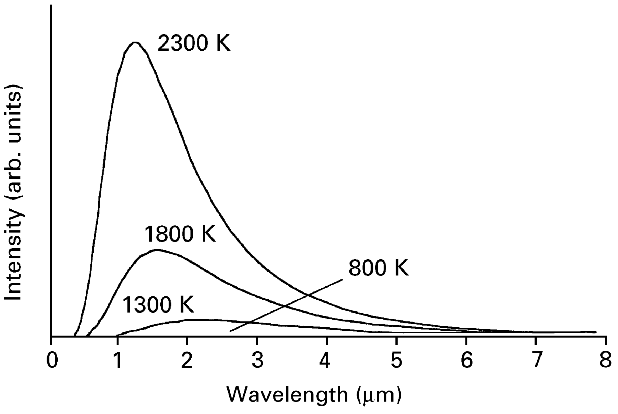



**Figure 2** Radiation from an ideal blackbody at several temperatures.

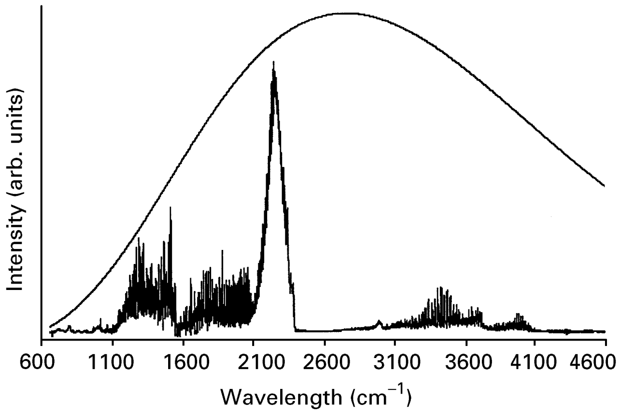



**Figure 3** The infrared emission spectrum measured using an FT-IR spectrometer from an 18 torr $CH_4/O_2$ flame and the blackbody radiation (1173 K) emitted over the same region. The infrared emission spectrum has not been corrected for the responsivity of the detector.

fine structure. In the visible and ultraviolet regions of the flame emission spectrum (in general, radiation with a wavelength between 1 μm and 200 nm), this fine structure represents changes in rotational and vibrational energy which accompany changes in electronic energy. In the more intense infrared region of the flame emission spectrum (in general, radiation with a wavelength between 1 and 30 μm), the observed fine structure is caused by transitions between rotational energy levels which occur with a change in vibrational energy, but with no change in electronic energy. The amount by which the vibrational and rotational energies may change during a transition is governed by selection rules, which are largely dependent upon the symmetry of the species involved in the transition.

**Figure 3** shows the raw emission spectrum from a premixed, reduced pressure (18 torr) stoichiometric $CH_4/O_2$ flame measured using a Fourier transform infrared (FT-IR) spectrometer and the calculated emission from a blackbody at 1173 K over the same spectral region. The peak temperature in this flame was measured (using a fine wire Pt-Pt/10%Rh thermocouple) to be near 2150 K. Comparison of flame emission spectra to calculated blackbody radiance must take into account the emissivity of the different species within the flame, reabsorption of emitted radiation by cold gases outside of the flame region, emission from species at different temperatures along the line of sight of the measurement, chemiluminescence, and the variation in sensitivity of the instrument detector with wavelength. **Figure 3** illustrates that care must be taken when estimating flame temperatures from measured spectra, since results from simple fits of a blackbody radiation curve to a measured spectrum may be inaccurate.

## Background

Emission and absorption spectra of flames may be continuous or banded. Continuous emission spectra may be modelled using the Planck blackbody equation (Eqn [2]). Modelling of banded spectra requires an understanding of the statistics that govern the way the population of a species is distributed among available energy levels. It is the dependence of population of molecular energy levels upon temperature, and the influence of this population upon band shape and spectral emissivity and absorption, which makes vibrational spectroscopy a useful tool for flame diagnostics.

The necessary mathematics for species and temperature measurements from infrared spectra of flames are outlined below. The example given is for a diatomic molecule, such as CO. It is assumed that the spectra are measured in absorption at low pressure, where the amount of light absorbed is less than approximately 5% of the incident intensity (often referred to as absorption in an optically thin medium). For extension to more complicated molecules (including more complicated diatomic molecules, such as NO), and to higher pressures, the reader should consult the 'Further reading' section for details.

### Population distribution among available energy levels

The classical Maxwell Boltzmann distribution law may be used to approximate the distribution of population among the quantized energy levels of a gas-phase diatomic molecule. For most diatomic molecules (such as CO, but not NO), each separate line shape observed in the infrared spectrum corresponds to a simultaneous change in vibrational

and rotational energy. For this reason, the spectral lines that make up the band structure observed in most infrared spectra are often called rovibrational transitions. For purposes here, it is convenient to describe the initial and final energy levels of the transition as rovibrational energy levels.

When describing two energy levels of a given species, the superscript $'$ denotes the state of higher energy, and the superscript $''$ denotes the state of lower energy. At thermodynamic equilibrium, the ratio of the number of molecules in the rovibrational energy level with vibrational quantum number $\nu'$ and rotational quantum number $J'$ to the number of molecules in the rovibrational energy level with vibrational quantum number $\nu''$ and rotational quantum number $J''$ is given by:

$$N_{J',\nu'}/N_{J'',\nu''} = [(2J'+1)/(2J''+1)] \times \exp(-(\Delta E_\nu + \Delta E_J)hc\,kT) \quad [3]$$

where $N_{J',\nu'}$ is the population of the higher rovibrational energy level, $N_{J'',\nu''}$ is the population of the lower rovibrational energy level, $\Delta E_\nu$ is the change in vibrational energy (units of $cm^{-1}$) and $\Delta E_J$ is the change in rotational energy (units of $cm^{-1}$) which occurs during the transition, $h$ is Planck's constant, $k$ is Boltzmann's constant, and $c$ is the speed of light. The quantity $[(2J'+1)/(2J''+1)]$ accounts for the degeneracy of rotational levels for a given value of $J$, providing a statistical weight to the levels with a given rotational energy.

For interpretation of measured spectra, it is useful to know the fraction of the total population ($N_T$) in the energy level from which the transition originates. For absorption (where the transition originates from the level of lower energy), this may be given by:

$$N_{J'',\nu''}/N_T = (N_{J'',\nu''}/N_{J=0,\nu=0})/(\sum N_{J,\nu}/N_{J=0,\nu=0}) \quad [4]$$

where the summation is over all values of $J''$ and $\nu''$ and $N_{J=0,\nu=0}$ is the population in the rovibrational energy level with rotational quantum number $J$ and vibrational quantum number $\nu$ both equal to 0. The denominator on the right side of Equation [4] may be written to a first approximation as:

$$\sum N_{J,\nu}/N_{J=0,\nu=0} = \sum (2J+1) \times \exp(-(E_\nu + BJ(J+1))hc/kT) \quad [5]$$

The first rotational constant, $B$, is inversely proportional to the moment of inertia of the rotating molecule, and determines the rotational energy level spacing. This means that, in general, the larger the mass of the rotating molecule, the closer is the spacing of the rotational energy levels. The right-hand side of Equation [5] may be factored into rotational and vibrational components:

$$\sum N_{J'',\nu''}/N_{J=0,\nu=0} = Q_R Q_V \quad [6]$$

where $Q_R$ and $Q_V$ are, respectively, the rotational partition function and the vibrational partition function:

$$Q_R = \sum (2J''+1)\exp(-(B''J''(J''+1))hc/kT) \quad [7]$$

$$Q_V = \sum \exp(-(E_{\nu''})hc/kT) \quad [8]$$

The summations in Equations [7] and [8] are over all $J$ and all $\nu$, respectively. Substituting Equations [3], [7], and [8] into Equation [4] gives:

$$N_{J'',\nu''}/N_T = (2J''+1) \times \exp(-(E_{\nu''} + B''J''(J''+1))hc/kT)/Q_R Q_V \quad [9]$$

Equation [9] shows how the fraction of total population in a given rovibrational energy level varies with temperature and rotational and vibrational quantum numbers. For most gas-phase diatomic molecules, if the population in a known rovibrational level is measured, Equation [9] allows the total population (and hence total pressure) of the gas to be calculated.

## Absorption of radiation

The Einstein transition probability of absorption, $B_{mn}$ (not to be confused with the first rotational constant $B$), predicts the energy removed ($I_R$) from an incident beam of radiation by an optically thin layer of absorbers for a transition from a lower state $m$ to an upper state $n$ as:

$$I_R = I_0 N_m B_{mn} \Delta x h\nu_{mn} \quad [10]$$

where $N_m$ is the number of molecules per unit volume in the energy level from which absorptions

occur, $I_0$ is the energy crossing unit area of absorbers per second, $\Delta x$ is the absorber thickness, and $\nu_{mn}$ is the energy (in $cm^{-1}$) of the monochromatic radiation exciting the transition. In Equation [10], $I_0 N_m B_{mn}$ is proportional to the number of transitions per second per unit volume produced by the radiation and $h\nu_{mn}$ is proportional to the energy removed from the incident beam per transition.

In an absorption experiment, the intensity of radiation exiting the absorbing medium, $I$, is described according to the Bouguer–Lambert Law (later restated by Beer for solutions):

$$I = I_0 \exp(-\sigma(\lambda) N_m x) \qquad [11]$$

where $x$ is the path length travelled by the light through the absorbing medium, and $\sigma(\lambda)$ is called the cross section for absorption. This cross section represents the 'effective area' that a molecule presents to the incident photons. When $\sigma(\lambda)N_m x$ is small (optically thin medium), Equation [11] may be rewritten:

$$I_0 - I = I_0(\sigma(\lambda) N_m x) \qquad [12]$$

Integrating Equation [12] over the wavelength range for which the absorption may occur gives:

$$\int (I_0(\lambda) - I(\lambda))\,\mathrm{d}\lambda = \int I_0(\lambda)(\sigma(\lambda) N_m x)\,\mathrm{d}\lambda \qquad [13]$$

The way in which the cross section for absorption ($\sigma(\lambda)$) varies with wavelength depends mainly upon the total pressure of the gas. For gas pressures above approximately 100 torr, the absorption is observed to occur with a Lorentzian line shape given by:

$$\sigma(\lambda) N_m x = 2.303 A_0 \gamma^2 / ((c/\lambda - c/\lambda_0)^2 + \gamma^2) \qquad [14]$$

where $A_0$ is the peak spectral absorbance ($-\log(I(\lambda)/I_0(\lambda))$) at the wavelength of maximum light attenuation by the gas, and $\gamma$ is the half width at half height (HWHH) of the spectral line.

Substituting Equation [14] into Equation [13], assuming that $I_0$ is invariant with wavelength over the absorption line width, and integrating over the full line width gives:

$$\begin{aligned} &\int (I_0(\lambda) - I(\lambda))\,\mathrm{d}\lambda \\ &= 2.303 I_0(\lambda) \int A_0 \gamma^2 / ((c/\lambda - c/\lambda_0)^2 + \gamma^2)\,\mathrm{d}\lambda \\ &= 2.303 I_0(\lambda) \pi A_0 \gamma \end{aligned} \qquad [15]$$

Substituting Equation [15] into Equation [10], and solving for the peak absorbance, $A_0$, gives

$$A_0 = N_m \Delta x \times B_{mn} h\nu_{mn} / 2.303\pi\gamma \qquad [16]$$

The number of molecules in the initial state, $N_m$, is related to the total number of molecules, $N_T$, through the vibrational ($Q_V$) and rotational ($Q_R$) partition functions:

$$\begin{aligned} N_m &= N_T(2J'' + 1) \\ &\quad \times \exp(-(E_{\nu''} + B''J''(J'' + 1))hc/kT)/Q_R Q_V \end{aligned} \qquad [17]$$

Substituting Equation [17] into Equation [16] gives:

$$\begin{aligned} A_0 &= [\Delta x B_{mn} h\nu_{mn} N_T (2J'' + 1)/2.303\pi\gamma Q_R Q_V] \\ &\quad \times \exp(-(E_{\nu''} + B''J''(J'' + 1))hc/kT) \end{aligned} \qquad [18]$$

Equation [18] is useful for extracting temperature and concentration information from measured values of peak absorbance ($A_0$) for individual transitions in a rovibrational band. By simultaneously fitting $T$ and $N_T$ to $A_0$ measured over a rovibrational band, temperature and gas pressure may be obtained. Equation [18] shows that it is necessary to know the half width at half height (HWHH) of each line in the spectrum used in the calculation, and the value of $B_{mn}$. This is not always trivial, since this value may be temperature and $J$-value dependent. Additionally, it is important to recognize peak absorbances must be corrected when measured with an instrument of moderate spectral resolution. The method for extracting true peak absorbance from peak absorbance measured at moderate resolution has been treated in detail by Anderson and Griffiths. Substituting Equation [18] into Equation [16], and expressing the result in the

form of the Bouguer–Lambert Law gives:

$$I(\lambda)/I_0(\lambda) = \exp[(-\Delta x B_{mn} h\nu_{mn}(\lambda) N_{\mathrm{T}}\gamma / \pi Q_{\mathrm{R}} Q_{\mathrm{V}}((c/\lambda - c/\lambda_0)^2 + \gamma^2)) \times (2J'' + 1)\exp(-(E_{\nu''} + B''J''(J''+1))hc/kT)] \quad [19]$$

Equation [19] allows direct comparison between high resolution measurements of transmittance ($I(\lambda)/I_0(\lambda)$) and transmittance calculated from spectral parameters, pressure of the absorbing gas, and the temperature. It should be noted that knowledge of the spectral line HWHH ($\gamma$) is required, as well as an instrument capable of measuring the transmittance at high resolution. In practice, values of $B_{mn}$ (usually converted to line strengths) for different rovibrational transitions are taken from the literature or estimated from measurements of total band strength. Also, the Voight line shape profile, which describes the convolution of a Gaussian line shape function (applicable to gases at low pressure) and a Lorentzian line shape function, is usually used to model spectral lines in flames at reduced pressure. As with fits to data based upon Equation [18], when spectra are measured on an instrument of moderate resolution (in general, an instrument resolution $>0.1\gamma$), the instrument line shape function is convoluted with the true line shape (Eqn [19]). When this is the case, this convolution must be included in the model.

Equation [19] describes the fully resolved band structure observed in high resolution infrared spectra of many gas-phase diatomic molecules, such as CO. **Figure 4** shows the transmittance ($I(\lambda)/I_0(\lambda)$) spectrum of CO, calculated using Equation [19], for a constant number of molecules at the temperatures shown in **Figure 2**. **Figure 4** shows that as the temperature increases, and more rotational energy levels becomes populated, the overall shape of the absorption band broadens to cover a wider spectral range. Additionally, because the total number of molecules is divided between a greater number of initial energy levels as temperature is increased, the intensity of individual rovibrational transitions changes with increasing temperature.

## Experimental methods narrative

By the 1950s, the mathematical and instrumental methods for determination of temperatures and species concentration from measurements of infrared spectra had been established, in large part spurred by electronics developed during the Second World War, by perfection of the method of commercial

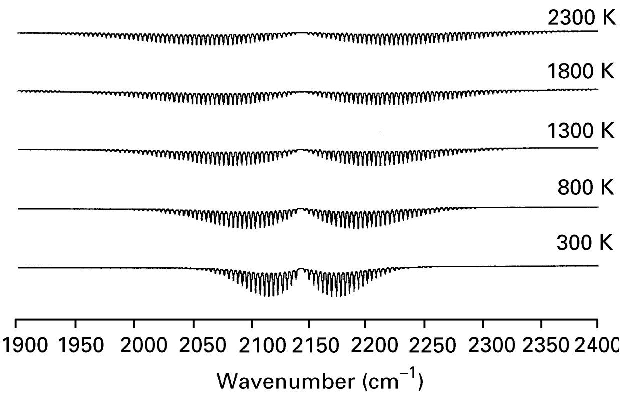



**Figure 4** Calculated transmittance of light through a fixed pressure of CO gas at different temperatures. Spectra are offset for clarity (each baseline corresponds to 100% transmission).

replication of diffraction gratings, and by publication in 1945 of Hertberg's book 'Infrared and Raman Spectra of Polyatomic Molecules'. During the 1950s, comparison of infrared emission spectra from high temperature sources in different laboratories was complicated by temperature gradients along the measurement line of sight, and by incomplete spectral parameters for gas absorption at high temperatures. To obtain spectral parameters for gases under study, several efforts were made to study gases in closed cells at controlled temperatures. Most of these efforts at studying gases at high temperature under static conditions employed absorption spectroscopy, in part to minimize self absorption along the line of sight, and also to enable using modulation of the source radiation (allowing discrimination of the high intensity, unmodulated background infrared emission). The success in fundamental studies of band structure and in predicting emissivities and changes in absorption at high temperatures and pressures led to an increase, by the early 1960s, in studies of radiation transfer for systems ranging from jet and rocket motors to high efficiency oil burners. A particular success which resulted from the study of gas-phase emissivities was the determination of thermal stress on NASA rocket motors from exhaust gas radiation.

With the development in the early 1970s of the Fourier transform spectrometer and the computer based fast Fourier transform (FFT), it became possible to achieve high resolution coupled with high energy throughput and phase-sensitive detection. However, the development of laser-based techniques (particularly laser-induced fluorescence (LIF) for measurement of flame radical species, such as OH, HCO, H, O, CH, and $C_2$) enabling direct measurement of species participating in flame propagation reactions caused a shift in focus of fundamental spectroscopic investigations of flame systems. This shift

led to a decrease in the late 1960s and early 1970s in the number of publications describing basic research that applied the techniques of dispersive infrared absorption and emission spectroscopy to flames. This decrease was offset by a considerable body of work on emission studies of hot gas sources, particularly smokestack and waste gas plumes from industrial sources.

By the late 1970s, laser-based techniques employing tunable infrared lasers began to be used for species measurements in flames. The majority of this work used tunable diode lasers (TDLs), semiconductor devices in which the output laser radiation wavelength is tuned by varying the temperature and diode injection current. Initial experiments used TDLs emitting narrow-line width (typically $< 10^{-4}$ $cm^{-1}$)-radiation in the mid-infrared spectral region. The narrow line width usually enables species specificity, even in congested spectral regions. Since the laser line width is typically several orders of magnitude less than the absorption line width, measurements of the fully-resolved absorption transition may be made, enabling the determination of line shape dependence upon pressure and temperature. An additional advantage to using tunable diode lasers, besides very narrows line widths, is the ability to tune the lasers, rapidly (kHz to MHz) over their output wavelength range. This enables phase-sensitive detection that minimizes the effect of the laser output noise, and also enables time-resolved measurements of dynamic systems.

Recently, tunable diode lasers operating in the near-infrared spectral region have begun to see application to the spectroscopy of flames and to flame gases. At this time, commercial availability of TDLs in the mid-infrared spectral range is greater than for the near-infrared spectral range. TDLs operating in the mid-infrared spectral range must be cooled to cryogenic temperatures. TDLs operating in the near-infrared spectral range operate near room temperature, and unlike mid-infrared radiation, the near-infrared radiation may be transmitted over long distances through optical fibres. However, since absorption in the near-infrared spectral region corresponds to a change in vibrational quantum number greater than unity (referred to as an overtone transition), the sensitivity to a given molecule is much less than for mid-infrared spectroscopy, which is usually used to measure rovibrational transitions in the fundamental vibrational band. Typically, mid-infrared TDLs are used when extreme sensitivity is required (ppb range). For many field based techniques which require transportability in rugged environments, near-infrared TDLs are more appropriate (ppm range).

## Experimental methods – dispersive and Fourier transform spectroscopy

### Emission measurements

Since the mid-1970s, most measurements of emission spectra of steady flames have used Fourier transform techniques. **Figure 5** shows the emission spectrum measured from a premixed, stoichiometric $CH_4/O_2$ flame (total pressure equal to 18 torr) to which 3% $CF_3Br$ has been added as a flame suppressant. When appropriate, reduced-pressure flames are often studied because at reduced pressure the flame region is expanded, allowing more detailed study. The emission spectrum shown in **Figure 5** was measured using a Fourier transform infrared (FT-IR) spectrometer at a resolution of 1 $cm^{-1}$.

In **Figure 5**, spectral features from several flame species are identified, but the spectrum is dominated by emission from $CO_2$ and gaseous $H_2O$. This is because higher temperatures allow emission over a wide range of rovibrational transitions (see **Figure 4**), from combinations of different vibrational states, and the geometry of the emitting species (specifically in this case $H_2O$) causes the spectrum to be complicated (relative to that for a diatomic molecule). The measured spectrum in **Figure 5** reports emission from species over a wide range of temperatures, because the line of sight from the flame centre to the detector includes gases at many different temperatures. Also, because many flame species are products of an ongoing chemical reaction (combustion), these species may be produced with vibrational and rotational population distributions which are not in equilibrium with translational

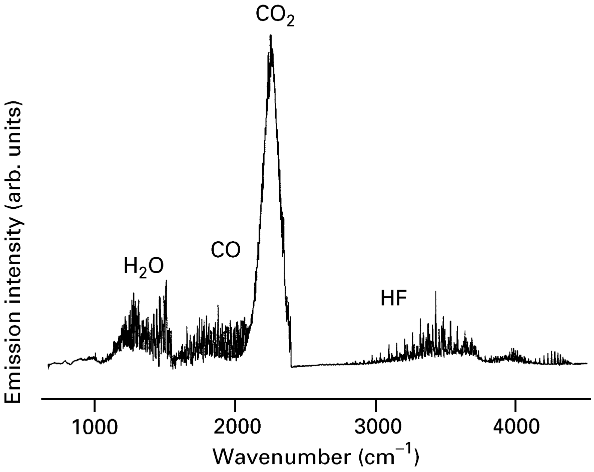



**Figure 5** The infrared emission spectrum measured from a premixed, stoichiometric $CH_4/O_2$ flame (total pressure 17 torr) to which 3% $CF_3Br$ has been added.

temperatures. Because of this, temperatures calculated from emission spectra using expressions similar to Equations [18] or [19] may give misleading results. When these factors are taken into account, temperature and partial pressures of gas species may be extracted from high resolution measurements of infrared emission spectra.

### Absorption measurements

**Figure 6** shows an absorption spectrum measured through an opposed-flow $CH_4$/air flame (total pressure 50 torr) using a FT-IR spectrometer at 1 $cm^{-1}$ resolution. It should be noted that this spectrum is less congested and has a slightly more regular appearance than that shown in **Figure 5**. Absorption features from several species are evident in the spectrum. For flame diagnostics using infrared absorption spectroscopy, CO is probably the most widely studied molecule. The fundamental rovibrational absorption band of CO ($v = 0$–1, centred at 2170 $cm^{-1}$) is well approximately by Equation [19], and has a spacing between rotational lines (2B) of approximately 3.6 $cm^{-1}$. Modelling of measured spectra is relatively straightforward and fully-resolved rovibrational lines may be measured at moderate instrument resolution (although the reported line shape may be just the instrument line shape function. In addition, since CO is a major product of hydrocarbon combustion, it is present in most flame systems. **Figure 7** shows the spectrum of the first overtone ($v = 0$–2) of CO, measured using a dispersive system, and the fit to the measured spectrum using an equation similar to Equation [19]. Since the resolution of the spectrometer used to measure the spectrum in **Figure 7** was insufficient to fully resolve the line shapes of the individual rovibrational transitions, it was necessary to convolute the instrument function with the true line shape function (Eqn [19]) to obtain an accurate fit of the measured to the calculated spectrum.

A major difficulty in obtaining quantitative absorption measurements of flame species is that absorption of radiation by cold gases in the probe beam line of sight may lead to errors. Currently, several methods of excluding contributions to absorptions by cold gases in the line of sight are used, with most methods employing some type of computer averaged tomography (CAT). Using these method (often referred to as inversion methods), multiple line of sight measurements at different relative orientations are made within a plane, or slice, through the flame. Using these multiple line of sight measurements, numerical methods are used to calculate species concentration at a given point within the plane. Applying this technique to multiple planes, or slices through the flame, can yield a three-dimensional reconstruction of species

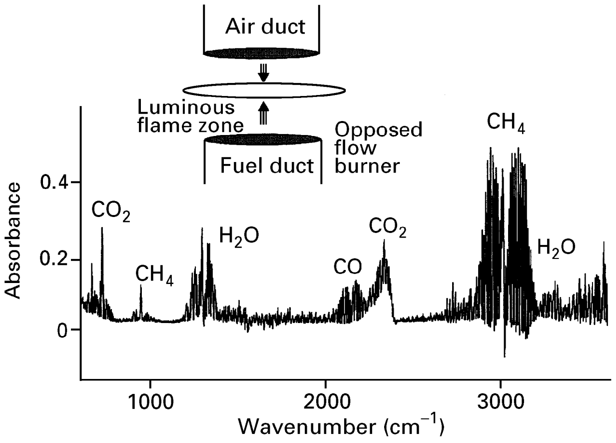



**Figure 6** The infrared absorption spectrum measured through a 50 torr opposed-flow $CH_4$/air flame. The inset shows the burner configuration used to produce the flame.

concentration within the flame. This technique has been used to measure CO in reduced- and atmospheric-pressure flames. A summation of inversion methods used to measure temperatures and species in flames has been given by Limbaugh.

It should be noted that when measuring flame absorption spectra using a FT-IR spectrometer, care must be exercised so that flame emission is not reflected through the interferometer and reported as an absorption signal. To minimize the emitted radiation which reaches the interferometer, the collimated probe infrared beam is brought to a focus at an iris prior to entering the flame region. After passing through the iris, the probe radiation is re-collimated prior to entering the flame. While useful for identification of species within the flame, FT-IR spectroscopy will not yield high spatial resolution unless the infrared probe beam is reduced in cross section (usually by being passed through a set of apertures) prior

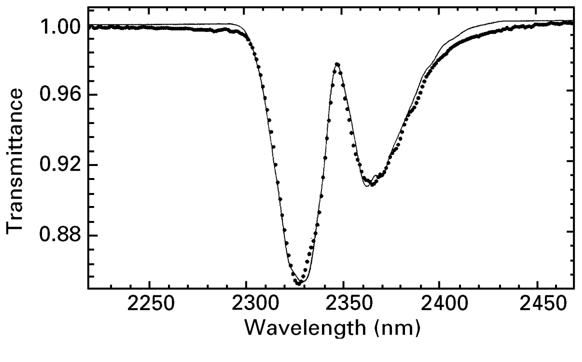



**Figure 7** The transmittance spectrum of the first overtone ($v = 0$–2) of CO. The gas was contained in a static cell at 1 atmosphere total pressure and a temperature of 423 K. The dotted line is the measured spectrum, the solid line represents the least-squares fit to the data.

to entering the flame region. This reduction in power of the infrared probe beam energy reduces the signal to noise ratio in the measured spectrum. A typical experimental setup for measuring infrared absorption spectra through a flame using an FT-IR spectrometer is shown in **Figure 8**.

Most commercial spectrometer systems employing broadband sources (typically FT-IR systems) are not capable of resolving the true line shape of an absorbing gas phase species because of limitations to instrument resolution. The instrument yield the convolution of the true absorbing gas line shape and the instrument response function. Computer-based programs used to retrieve species concentration and temperature must either take into account this convolution, or use for calculations the instrument response-corrected peak intensity of the rovibrational transition.

## Experimental methods – applications of tunable diode laser absorption spectroscopy (TDLAS)

Tunable diode lasers (TDLs) are semiconductor devices (typically GaAs) which are essentially light emitting diodes constructed with an optical resonator. Lasing is achieved by delivering a small current to the photodiode. Tuning is achieved by changing the temperature, and hence the Fermi level, of the device. The tuning range varies by type, but is usually between 2 $cm^{-1}$ and 30 $cm^{-1}$. The useful tuning range (that is, the range over which the device may be rapidly tuned without encountering changes in laser modes, which can mask absorption features) is usually less than 1 $cm^{-1}$. This restricted tuning range means that the devices are not used in a survey mode. In general, a TDL laser spectrometer system is designed to detect a single gas using a single, specially-fabricated, dedicated, laser diode. Occasionally, absorption transitions from different gases of interest may occur within the useful tuning range of a single laser, but such occurrences are usually fortuitous. Additionally, because of the limited tuning range, calculation of temperature and species concentration may be difficult when only one line is measured.

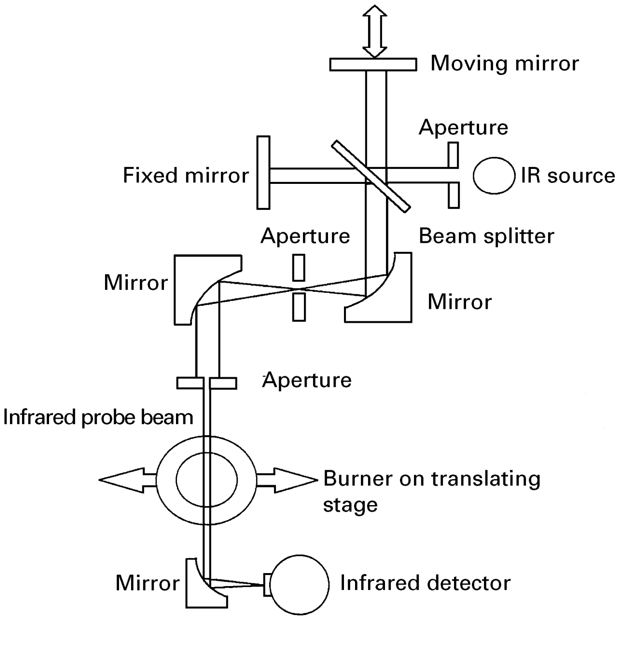



**Figure 8** Experimental apparatus for measurement of FT-IR absorption spectra through flames. The system is designed to minimize the amount of flame emission which reaches the interferometer. The burner is mounted on a translating stage (*X* and *Z* axes) to allow for tomographic analysis.

There are, however, several advantages to using TDLs for measurement of gas-phase flame species. These include high resolution (typically better than $1 \times 10^{-4}$ $cm^{-1}$), good spatial resolution (200 to 1 mm), reasonable output power (~1 mW), and the ability to scan over their spectral range on a millisecond or better timescale. Probably the most widely studied molecular flame species by tunable diode laser spectroscopy is CO. In addition to the reasons for study outlined above in the discussion of broadband source methods, CO possesses several fundamental ($\nu = 0$–$1$) and hot-band transitions ($\nu = 1$–$2$, $\nu = 2$–$3$) which occur within several line widths (approximately 0.05 $cm^{-1}$) of each other. At room temperature, populations of states from which hot-band transitions occur are very low. However, at flame temperatures, populations of vibrational states other than the $\nu = 0$ state may become appreciable. When temperatures (and also species concentrations) are calculated from simultaneous measurement of a fundamental and a hot-band transition, the technique is referred to as two-line thermometry.

**Figure 9** shows several spectra of CO measured in a low-pressure (20 torr), stoichiometric $CH_4/O_2$ flame using mid-infrared tunable diode laser spectroscopy. Because the wavelength of the diode laser radiation corresponds to single quantum number changes in CO (in this case a $\nu = 0$–$1$ and a $\nu = 1$–$2$ transition) simple absorption spectroscopy may be employed. Each spectrum is measured at a different height above the burner surface. For these experiments, the probe laser beam location was fixed, and the burner location was translated vertically. The infrared laser beam diameter through the flame region was 0.5 millimetres. As the diode laser radiation passes through different portions of the flame, the relative areas under the two peaks corresponding

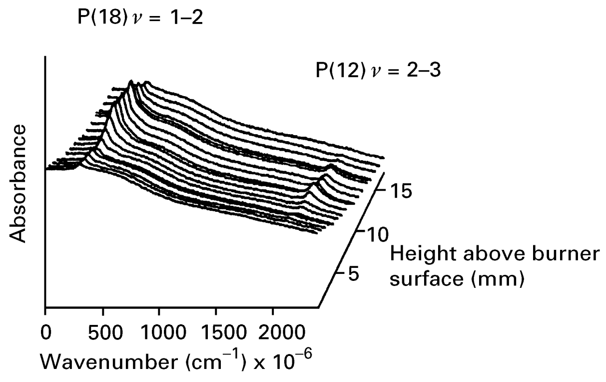



**Figure 9** Mid-infrared tunable diode laser spectra of CO, measured through a premixed $CH_4/O_2$ flame, at a total pressure of 20 torr, as a function of height above the burner surface. The maximum temperature calculated using two-line thermometry was 2150 K.

to absorption from different rovibrational levels changes. **Figure 10** shows the results of fitting an equation similar to Equation [19] to an absorption spectrum from **Figure 9**. Using this method, the CO vibrational temperature was calculated as a function of height above the burner surface ($T_{max}$ = 2150 K ±50 K). Additionally, tomographic analysis of the data showed that the error from cold gas absorption, for these experiments, was always less than 10% of calculated temperatures and partial pressures.

Tunable diode laser spectroscopy is also used for detection of many trace gas species in flames and in flame environments. The method used for measurements of small amounts of gases (often to the ppb range) relies on modulation spectroscopy using

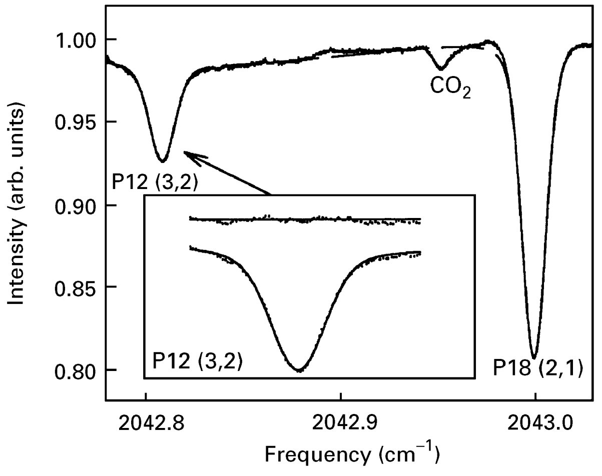



**Figure 10** A transmission spectrum measured through a low-pressure, premixed, $CH_4/O_2$ flame at 20 torr. The main absorption features are the two CO lines used to calculate temperature using the method of two-line thermometry. Calculated temperature was 2150 K.

phase-sensitive detection. The principal advantage to using modulation spectroscopy results from minimization of laser output noise by shifting detection to high frequencies. When phase-sensitive detection is employed using tunable diode lasers, the second derivative of the diode laser probe beam intensity with respect to wavelength is usually measured, since the wavelength at which the second derivative is a maximum coincides with the wavelength of maximum light absorption. The second-derivative signal peak height may be shown to be proportional to absorbance, $A$:

$$X''/V = kA \qquad [20]$$

Here, $X''$ is the peak height of the second-derivative signal (volts), $V$ is the DC voltage measured by the detector in the absence of any molecular absorbance, and $k$ is a constant which includes the measuring instrument and optics function. Letting $S$, denoted as the 2f signal peak to trough height (the distance between the maximum and minimum points on the 2f signal – see **Figure 11C**, equal $X''/V$, yields:

$$S = (k\alpha)\, LP \qquad [21]$$

The slope of a plot of $LP$ versus $S$ provides the value of $k\alpha$. A calibration of the system using known concentrations of the absorbing gas must be performed to determine the value of $k\alpha$. Once this value is known, gas pressure ($P$) may be obtained directly from Equation [21]. Care must be exercised so that changes in optical surfaces during measurement (as when measuring corrosive gases such as HF) do not affect the value of $k\alpha$, since this value is instrument-function dependent. For this reason, calibration should be performed at the beginning and end of each measurement set. Additionally, for gas concentrations which attenuate more than 5% of the incident light, the linear relationship between gas pressure and 2f signal peak height may no longer hold.

Modulation spectroscopy using tunable diode lasers (often referred to as 2f spectroscopy) has been extensively discussed in the literature. **Figure 11** gives a generic description of the signal processing employed in these experiments. Briefly, the laser output wavelength is scanned (10 Hz–10 kHz) through a spectral region where the gas being measured absorbs. The output at the detector during the laser scan (with the high-frequency modulation turned off) may be seen in the upper trace of

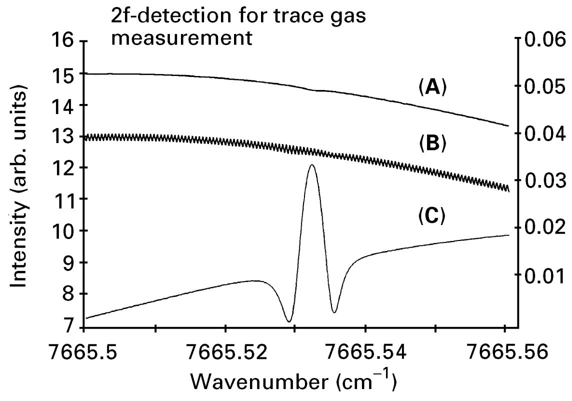



**Figure 11** Simulated signal versus wavelength of laser radiation after passing through a gas with an absorption feature near 7665.5 cm$^{-1}$. Graphs for (A) unmodulated laser radiation; (B) laser radiation with a small-amplitude, high frequency modulation; (C) demodulation at twice the high-frequency modulation in (B).

**Figure 11**. A high-frequency (10 KHz to MHz) wavelength modulation with an amplitude approximately equal to line width under investigation is superimposed on the laser drive current (middle trace in **Figure 11**). Demodulation at twice the frequency of the high-frequency laser drive modulation yields the second-derivative signal shown in the lower trace in **Figure 11**. It should be noted that the upper trace in **Figure 11** shows that the laser diode output power has a non-linear dependence on laser drive current. While this is non-ideal behaviour, this is a common trait of commercially available laser diodes. The non-linear power output dependence on laser drive current (exaggerated here for illustration) causes the sloping baseline for the second-derivative signal in **Figure 11**. For measurements at extremely low concentrations or for gases with small absorption cross sections, the non-linearity of laser diode output power with laser drive current may affect limits of detection.

Modulation spectroscopy enables discrimination against contributions to light attenuation by scattering from particulate matter using a single laser beam. Attenuation of laser radiation by a ro-vibrational transition of a small gas molecule is detected because the wavelength range of the scan is several times the width of the spectral absorption feature (typically on the order of 0.1 cm$^{-1}$ at atmospheric pressure). Because light scattering by particulate matter is nearly constant over the very small wavelength range of the laser scan, the change in detector signal intensity with the change in wavelength is effectively zero in the absence of any absorbing gas. However, because the 2f signal is also proportional to the background signal intensity, the second-derivative signal is divided by the DC signal to account for light scattering by particles.

While modulation techniques enable trace gas detection, special care must be exercised when making measurements in flames, especially when using near-infrared laser radiation. Significant errors in measured concentrations may arise because of beam steering as the laser radiation passes through the flame. In some cases, the beam steering may be severe enough to cause loss of signal at the detector. Secondly, great care must be exercised when measuring 2f spectra through regions of differing temperature and pressure, since the 2f signal is extremely sensitive to changes in line width of the absorbing species.

## Conclusion

Infrared spectroscopy continues to be a valuable tool for measuring species concentrations and temperatures in flames and combustion gases, especially in field-based studies. The advent of compact, fibre-coupled, tunable diode lasers operating at room temperatures is expanding the use of vibrational spectroscopy beyond the laboratory, and providing a useful complement to broadband methods currently in place.

## List of symbols

$A$ = area of radiator surface; $A_0$ = peak spectral absorbance; $B$ = first rotational constant; $B_{mn}$ = Einstein transition probability of absorption for states $m$, $n$; $c$ = speed of light; $c_1$, $c_2$ = radiation constants; $E$ = internal energy of gas-phase molecule; $E_\lambda$ = emissivity; $h$ = Planck's constant; HWHH = half width at half height; $I_0$ = energy crossing unit area of absorbers/second; $I_R$ = energy removed from incident beam by absorbers; $I_\lambda$ = radiant intensity at wavelength $\lambda$; $J$ = rotational quantum number; $k$ = constant; $k$ = Boltzmann's constant; $N$ = energy-level population; $Q_R$ = rotational partition function; $Q_V$ = vibrational partition function; $S$ = signal peak-trough height; $T$ = temperature (K); $V$ = DC voltage from detector for no absorbance; $x$ = path length through absorbing medium; $X''$ = peak height of second-derivative signal; $\Delta x$ = absorber thickness; $\gamma$ = half width at half height of spectral line; $v$ = vibrational quantum number; $\nu_{mn}$ = energy of exciting radiation; $\sigma(\lambda)$ = absorption cross-section.

*See also:* **Laser Spectroscopy Theory; Rotational, Spectroscopy Theory; Solid State NMR, Rotational Resonance; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading

Anderson RJ and Griffiths PR (1975) Errors in absorbance measurements in infrared Fourier transform spectrometry because of limited instrument resolution. *Analytical Chemistry* **47**: 2339–2347.

Anderson RJ and Griffiths PR (1977) Determination of rotational temperatures of diatomic molecules from absorption spectra measured at moderate resolution. *J. Quant. Spectrosc. Radiat. Trans.* **17**: 393–401.

Barnes AJ and Orville-Thomas WJ (eds) (1977) *Vibrational Spectroscopy – Modern Trends*. Amsterdam: Elsevier Scientific Publishing Company.

Bomse DS, Hovde DC, Oh DB, Silver JA and Stanton AC (1992) Diode laser spectroscopy for on-line chemical analysis. In: Bomse DS, Brittain H, Farquharson S, Lerner JM, Rein AJ, Sohl C, Todd TR and Weyer L (eds) *Optically Based Methods for Process Analysis*. Proc. SPIE **1681**: 138–148.

Daily JW (1997) Laser induced fluorescence spectroscopy in flames. *Prog. Energy Combust. Sci.* **23**: 133–199.

Daniel RG, McNesby KL and Miziolek AW (1996) The application of tunable diode laser diagnostics for temperature and species concentration profiles of inhibited low pressure flames. *Applied Optics* **35**.

Dasch (1994) *Applied Optics* **33.**

Fried A, Henry B, Wert B, Sewell S and Drummond JR (1998) Laboratory, ground-based and airborne tunable diode laser systems: performance characteristics and applications in atmospheric studies. *Applied Phys. B* **67**: 317–330.

Gaydon AG (1948) *Spectroscopy and Combustion Theory*, p 129. London: Chapman & Hall Ltd.

Gaydon AG (1974) *The Spectroscopy of Flames*, 2nd edn. London: Chapman & Hall Ltd.

Gaydon AG and Wolfhard HG (1953), *Flames: Their Structure, Radiation and Temperature*. London: Chapman & Hall Ltd.

Griffiths PR and de Haseth JA (1986) *Fourier Transform Infrared Spectroscopy*. New York: Wiley.

Hanson RK, Varghese PL, Schoenung SM, and Falcone PK (1980) Absorption spectroscopy of combustion gases using a tunable IR diode laser. *Laser Probes for Combustion Chemistry*. ACS Symposium Series 134. Washington DC: American Chemical Society.

Hertzberg G (1950) *Spectra of Diatomic Molecules*. New York: Van Nostrand & Co.

Hertzberg G (1950) *Electronic Spectra of polyatomic molecules*. New York: Van Nostrand & Co.

Limbaugh CC (1985) The infrared emission-absorption method for temperature and species partial pressure determination in flames and plumes. In: Wormhoudt Joda (ed) *Infrared Methods for Gaseous Measurements*. New York: Marcel Dekker Inc.

McNesby KL (1998) Informal survey of *Proceedings of the International Symposia (International) on Combustion* from 1948 to 1996. New York: Academic Press.

McNesby KL, Daniel RG and Miziolek AW (1995) Tomographic analysis of CO absorption in a low-pressure flame. *Applied Optics* **34**: 3318–3324.

McNesby KL, Daniel RG, Widder JM and Miziolek AW (1996) Spectroscopic investigation of atmospheric diffusion flames inhibited by halons and their alternatives. *Applied Spectroscopy* **50**: 126.

Miller JH, Elreedy S, Ahvazi B, Woldu F and Hassanzadeh P (1993) Tunable diode-laser measurement of carbon monoxide concentration and temperature in a laminar methane-air diffusion flame. *Applied Optics* **32**: 6082–6089.

Moore Walter J (1972) *Physical Chemistry*, 4th edn. New Jersey: Prentice-Hall Inc.

Oppenheim UP (1963) Spectral emissivity of the 4.3μ $CO_2$ band at 1200 K. *Ninth Symposium (International) on Combustion*, pp 96–101. New York: Academic Press.

Penner SS (1959) *Quantitative Molecular Spectroscopy and Gas Emissivities*. Reading, MA: Addison-Wesley Pub. Co.

Reid (1981) Second harmonic detection with tunable diode lasers – comparison of experiment and theory. *Applied Physics B* **26**: 203–210.

Silver JA (1992) Frequency modulation spectroscopy for trace species detection: theory and comparison among experimental methods. *Applied Optics* **31**: 707–717.

Steinfeld JI (1974) *Molecules and Radiation*. New York: Harper & Row.

Townes CH and Schawlow AL (1955) *Microwave Spectroscopy*. New York: McGraw-Hill.

Vanderhoff JA, Modiano SH, Homan BE and Teague MW (1997) Overtone absorption spectroscopy of solid propellant flames: CO and $N_2O$ concentrations. In Kuo KK (ed) *Challenges in Propellants and Combustion 100 Years after Nobel*, pp 876–884. New York: Begell House Inc.

Varghese PL and Hanson RK (1980) Tunable infrared diode laser measurements of line strengths and collision widths of $^{12}C^{16}O$ at room temperature. *Journal of Quantitative Spectroscopy & Radiative Transfer* **24**: 279.

Wilson EB, Jr., Decius JC and Cross PC (1955) *Molecular Vibrations*. New York: McGraw-Hill Co.

Wormhoudt JA and Conant JA (1985) High resolution infrared emission studies from gaseous sources. In: Wormhoudt Joda (ed) *Infrared Methods for Gaseous Measurements*, pp 1–45. New York: Marcel Dekker Inc.

# Fluorescence and Emission Spectroscopy, Theory

**James A Holcombe**, University of Texas at Austin, TX, USA

Copyright © 1999 Academic Press

ATOMIC SPECTROSCOPY
Theory

Like all atomic spectroscopic techniques, emission and fluorescence spectrometry require the production of free atoms (or ions) in the gaseous phase for detection. These two techniques require the population of a particular excited electronic state and determine concentration by monitoring the radiative relaxation process, i.e. by detection of the photon produced. In general, the intensity of the light emitted should be proportional to the gas-phase atom density and, ultimately, to the concentration of the analyte atoms in the sample being introduced to the atomization source. Likewise, it follows that increasing the excited-state population will enhance the sensitivity of the method.

The difference between the two spectroscopic approaches stems from the means by which the excited state is populated. In fluorescence, absorption of radiation at the proper frequency is used for excitation; in emission spectroscopy, collisional energy transfer with other atoms, molecules, ions or electrons produces the excited-state population.

## Atomic emission

### Excited-state population

Since excited-state populations are achieved by collisions with other charged or neutral particles within the source, the population of any given excited state can be calculated if local thermodynamic equilibrium (LTE) can be assumed to exist within the source. Under LTE, the fractional population of any given excited electronic state, $N_j$, is given by

$$\frac{N_j}{N_{\text{total}}} = \frac{g_j \exp\left(-E_j/k_{\text{B}}T\right)}{Z(T)} \qquad [1]$$

where $g_j$ and $E_j$ are the statistical weight and excitation energy of the $j$th state, respectively; $k_{\text{B}}$ is the Boltzmann constant and $T$ is the temperature (in kelvin). The partition function $Z(T)$ is the 'sum-over-states', or simply represents the relative population of all other states in the atom:

$$Z(T) = \sum_{i=0}^{\infty} g_i \exp\left(\frac{-E_i}{kT}\right) \qquad [2]$$

For relatively cool sources (e.g. $T < 5000$ K), $Z(T)$ can be approximated by the value for the statistical weight of the ground state, $g_0$, which implies that a majority of the atoms are present in the ground state ($E_0 = 0$). While an elevated population of the excited state is a *necessary condition* to realize a strong emission signal and good analytical sensitivity, it is not a *sufficient condition*. It is also necessary that there exists a high probability that, once populated, the excited state will radiatively relax to a lower energy state. In many instances the transition used terminates in the ground state of the atom or ion, and the line emitted is termed the resonance line. The probability that radiative relaxation will occur is indicated by the magnitude of the oscillator strength ($f$) or Einstein $A$ coefficient for spontaneous emission. Both of these parameters are indicators of the strength of the emission, with smaller values indicative of low probability or even 'forbidden' transitions. Examples where such transitions might have a low probability include transitions in which an apparent change in electron spin state is needed to accomplish the relaxation, i.e. a 'spin forbidden' transition. As a result of selection rules that dictate transitions with a high probability, not all possible transitions from an excited state to a lower-lying state occur with equal probabilities. In spite of this, the emission spectrum of elements with d or f electrons can exhibit hundreds to thousands of lines in the UV and visible region of the spectrum.

### Line widths

While radiative relaxation can produce relatively narrow spectral lines (e.g. 0.0001 nm or less, depending on the source and line), the lines are not truly monochromatic. At the very least each spectral line exhibits a 'natural line width' that is governed by the Heisenberg uncertainty principle

$$\Delta E\, \tau_{\text{r}} = \frac{h}{2\pi} \qquad [3]$$

where $\Delta E$ is the uncertainty in the position of the state, $\tau_r$ is the excited-state lifetime (typically 1–10 ns) and $h$ is the Plank constant. With rearrangement and

substitution of the Plank relationship ($E = h\nu$), the natural line width in hertz can be expressed as

$$\Delta\nu_N = \frac{1}{2\pi\tau_r} = \frac{A_{ji}}{2\pi} \quad [4]$$

where $A_{ji}$ is the Einstein $A$ coefficient for spontaneous emission from the $j$th to the $i$th state. With typical excited-state lifetimes of approximately 10 ns, the natural line widths are of the order of 0.000 01 nm.

Line widths in most sources are generally broader than the neutral line widths as a result of collisional and Doppler broadening. Collisional broadening is a result of collisions or near collisions between the excited-state atom or ion and other particles in the system that result in a perturbation of the relative location of the energy states of the atom at the time of emission. As a result, elevated pressures and temperatures in the source enhance the impact of collisional broadening on the line width. Doppler broadening is a consequence of phenomena similar to the familiar Doppler effect for sound, which results in change in pitch or frequency as a sound emitting source approaches and then passes an observer, for example for a passing train that is blowing its whistle. In a similar fashion, an emitting atom that is moving towards or away from the observer exhibits a frequency shift in the emitted radiation that is related to the relative direction and velocity of the emitter. Taken over an ensemble of atoms moving in random directions, this produces a broadened spectral line. Like collisional broadening, this is accentuated with increasing temperatures since the average molecular velocities increase with temperature. For sources at atmospheric pressure and temperatures of a few thousand degrees, these broadening events are similar in magnitude and produce spectral line widths in the range of 0.001–0.01 nm.

Other effects that can produce broadening or splitting of the spectral lines include Stark broadening and Zeeman splitting. Stark broadening results from local electric fields from neighbouring ions and electrons or from externally applied, strong electric fields. Zeeman splitting results from the interaction of a magnetic field with the particles or ions.

From Equation [1] it is apparent that elevated temperatures promote population of electronic excited states. However, competing with this enhanced excited state population is ionization:

$$\mathrm{M} \rightleftharpoons \mathrm{M}^+ + \mathrm{e}^- \quad [5]$$

This is an equilibrium condition and can be expressed by the Saha equation:

$$K_i = \frac{n_{\mathrm{M}^+} n_\mathrm{e}}{n_\mathrm{M}} \quad [6]$$

where $n$ represents the number density of the respective species. This equilibrium is quantitatively described by

$$\log K_i = 15.68 + \log \frac{Z_{\mathrm{M}^+}}{Z_\mathrm{M}} + 1.5 \log T - \frac{540 E_{\mathrm{ion}}}{T} \quad [7]$$

where $Z$ represents the partition functions for the respective species and $E_{\mathrm{ion}}$ is the ionization potential in eV. In brief, the equation predicts that reduced ionization potentials, low partial pressures of the neutrals in the source and elevated temperatures will enhance the fraction of the neutrals that are ionized. It is for this reason that in some high-temperature sources, such as inductively coupled plasmas, the resonance transitions of the *ions* produce more intense emission lines (and improved analytical detection limits) than neutral atom lines for the same element.

While the above discussion applies to many analytical atomic spectroscopic sources, much of the discussion and quantitative treatment requires the existence of local thermodynamic equilibrium, LTE. Some sources do not exist at LTE. Typically, sources that are at reduced pressure, which limits the rate of collisional distribution of energy, or those where energy is being input at a rapidly changing rate, may not exhibit LTE. As an example, glow discharge sources such as hollow-cathode lamps used in atomic absorption are operated in a low-pressure noble gas environment and exhibit extremely high ionization temperatures but very low kinetic gas temperatures; that is, the temperatures are not equal and, as a result, the system is not considered at LTE. This source, in spite of its kinetic gas temperature of only a few hundred degrees, exhibits an extremely high degree of ionization and population of some very high transition energy excited states.

## Calibration curves

In general, the intensity of light emitted from the excitation source should be proportional to the concentration of analyte atoms or ions within the source. This linear relationship should be valid as long as equilibrium processes within the source govern the

fraction of the total analyte population. An example where this may not be true can be seen in Equations [5]–[7], where ionization can be dependent on the partial pressure of the free atomic species. In the case of a flame source, low concentrations of the easily ionized elements (e.g. alkali metals) show measurably higher degrees of ionization. As a result, there can be curvature at the low concentration end of the calibration curve. Interestingly, the much hotter inductively coupled plasma does not exhibit the same nonlinearity at low concentrations. This is a result of the nearly constant electron density ($n_e$), which keeps the degree of ionization nearly independent of concentration.

Since resonance lines are often used for spectrochemical analysis, radiation emitted from the centre of an excitation source may pass through a substantial concentration of atoms in the ground state in the outer regions of the source. As a consequence, some of this light may be absorbed by the analyte atoms in this region. This is referred to as self-absorption. In the extreme, the absorbing atoms can reduce the intensity at the central wavelength of the emitting line to near zero (self-reversal). **Figure 1** shows intensity profiles with increasing analyte concentrations for an atomic emission line subject to self-absorption.

Consider light emitted from the hot centre region of a source (e.g. a flame) at an intensity of $I^0_{c,\lambda}$. The intensity is a function of both the analyte concentration and the wavelength, with maximum intensity emitted at the line centre. Since the intensity is proportional to the concentration $c$ of the analyte element in the sample,

$$I^0_{c,\lambda} = k'_\lambda c \qquad [8]$$

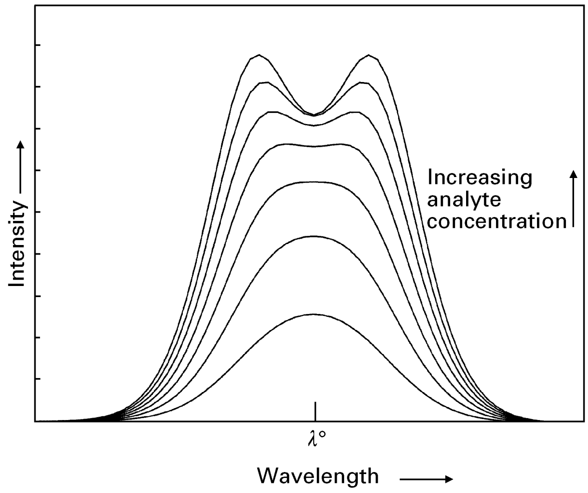



**Figure 1** Intensity profiles with increasing analyte concentration (bottom to top) for an emission line subject to self-absorption.

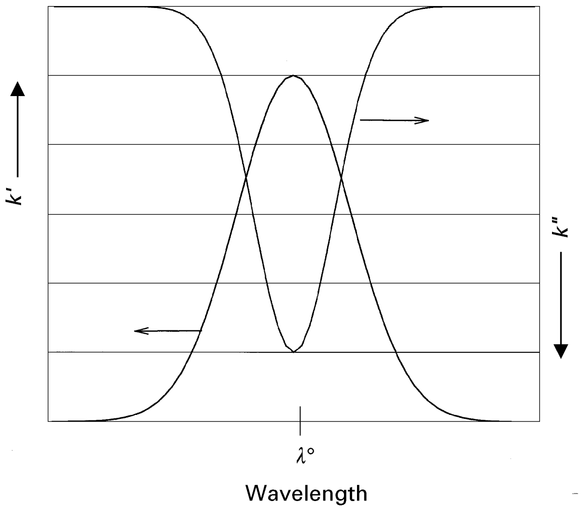



**Figure 2** Wavelength dependences of $k'$ and $k''$. Note that the vertical axis for $k''_\lambda$ has been inverted to aid presentation.

where $k'_\lambda$ is a constant that is wavelength dependent. The resonance emission also passes through the cooler outer regions of the source that also contain analyte atoms. However, they are emitting a negligible amount of light because of the cooler temperatures and the very strong temperature dependence of emission (Eqn [1]). Thus, these cooler atoms can absorb the resonance radiation without contributing significantly to the emission intensity. The absorption of this resonance radiation exhibits a Beer's law type of behaviour, where the intensity passing through this layer is related to the concentration $c$ and the incident intensity $I^0_{c,\lambda}$.

$$I_{c,\lambda} = I^0_{c,\lambda}\ \exp(k''_\lambda c) \qquad [9]$$

where $k''_\lambda$ is indicative of the strength of the absorption at any given λ. A graphical representation of the wavelength dependence of $k'$ and $k''$ is shown in **Figure 2**. (Note: The scale for $k''$ has been reversed for clearer presentation.) Since $I^0_{c,\lambda}$ is dependent on $c$ (Eqn [8]), the intensity emitted from the source after passing through this absorbing layer can be described as

$$I_c = \int k'_\lambda c\ \exp(k''_\lambda c)\ \partial\lambda \qquad [10]$$

From this expression two things become apparent: (a) the intensity detected ($I_c$) is not linearly dependent on $c$ unless $0 \le k''_\lambda << 1$ (i.e. there is a low probability of absorbing a photon) or $c$ is very small; and (b) as $c$ increases, the deviation from the ideal

calibration curve slope of $k'$ increases. This behaviour generates nonlinear calibration curves with greatest curvature at higher concentrations. Other instrumental factors (such as stray light in the spectrometer) can also affect the shape of the calibration curve but will not be dealt with here.

## Atomic fluorescence

### Excited-state population

Fluorescence, like emission, relies on radiative relaxation of an atom or ion from an excited electronic state to a lower-lying state. Unlike emission spectroscopy, population of the excited state occurs via absorption of a photon. Two of the possible fluorescence schemes are shown in **Figure 3**. If the excitation beam that is absorbed is the same as the detected fluorescent wavelength, it is termed resonance fluorescence. Nonresonance fluorescence denotes the case where the excitation and fluorescence wavelengths are different. The latter excitation/detection approach offers the advantage of minimizing scattered radiation from the excitation beam from reaching the detector. Additionally, by using a nonresonance line for analysis, self-absorption by the source of the fluorescent signal is minimized. (See also the discussion below on calibration curves.)

If the relaxation is a spontaneous emission process occurring after photon absorption has populated the excited state, then it follows that the intensity of the detected radiation will increase for transitions with larger Einstein $A$ coefficients, as in emission spectroscopy. However, population of the excited state is not governed by thermal or collisional processes, but instead by absorption of a photon of the proper frequency. Thus, it is logical that improved absorption should also yield improved sensitivity.

Absorption is a *stimulated* event, that is the number of atoms that are promoted to a higher energy level is dependent not only on the total number of atoms in the low-lying originating state that are capable of absorbing a photon, but also on the photon flux. For example, if one were to double the incident excitation beam intensity, one would expect twice as many atoms to absorb photons. (Note: This implies that the fraction of the total photons absorbed by the analyte is constant, which is consistent with the precepts of absorption where the ratio of $I^0/I$ is independent of $I^0$.) This dependence on the incident photon flux makes the absorption a stimulated process and the probability of absorbing a photon is governed by the magnitude of the Einstein $B$ coefficient for stimulated absorption. This $B$ coefficient is proportional to the Einstein $A$ coefficient and to $f$, the oscillator strength. Hence there is justification for the axiom that ‘a strong emitter is a strong absorber’ often cited in atomic spectroscopy.

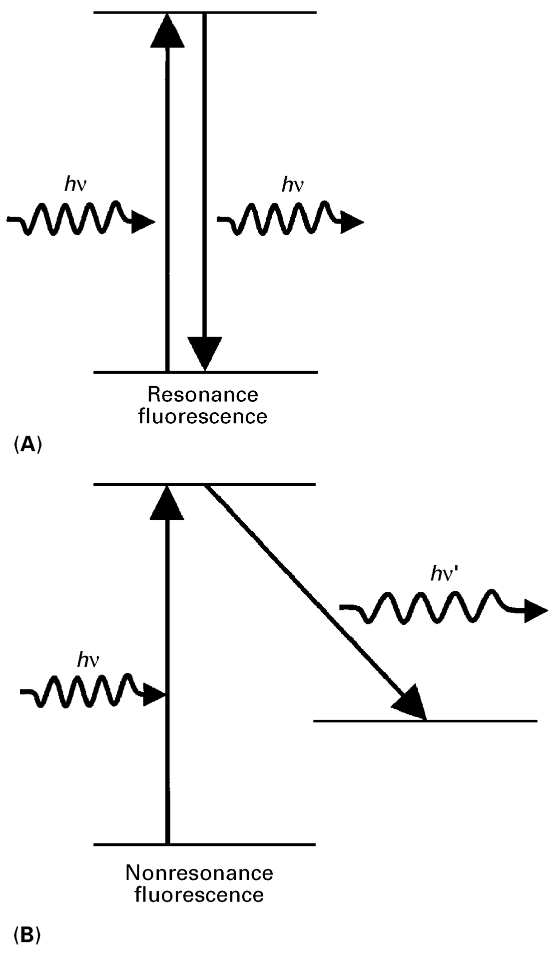



**Figure 3** Two fluorescence schemes. (A) Resonance fluorescence: excitation and detection at the same wavelength. (B) Nonresonance fluorescence: excitation and detection at different wavelengths, $\lambda_{det} > \lambda_{exc}$.

Since sensitivity is dependent on the excitation source intensity, it would seem that one should be able to increase the source intensity and thereby improve the sensitivity of the fluorescence technique for analysis. This is valid with low radiative fluxes. However, there is one more interaction between radiation and matter that must be introduced into this discussion to explain the limitation of this assumption: stimulated emission.

*Stimulated emission* is the process that occurs when an atom or ion (or molecule, for that matter) is in an excited state and a photon interacts with this atom to stimulate the relaxation to a lower energy state (where $\Delta E = h\nu$). The stimulated photon is emitted at the same wavelength as the incident photon, travels in the same direction as the incident

photon and is in phase with the incident photon. As a result of this process, the excited state has been diminished by this interaction with light. Since this is a stimulated process (like absorption), the number of particles that will participate in this interaction will depend on the incident flux and the number of atoms or ions located in the proper excited state. In short, as the excitation source intensity continues to increase, the population of the excited state increases from stimulated absorption, and the number of stimulated emission events also increases as the excited state population increases. Eventually a point is reached where the rate of the stimulated absorption process equals the rate of the stimulated emission and no net increase (or decrease) in population of the excited state occurs. The system is said to be 'saturated'. Additional increases in the incident radiation source intensity will have little impact on either the fraction of atoms or ions in the excited state or the spontaneous emission, i.e. fluorescence signal. As a result, small fluctuations in the source light intensity will not alter the detected signal for a given concentration of atoms in the atomization cell. Usually, source intensities comparable to that produced by lasers are needed to reach saturation.

Since the excitation process is dependent not only on the photon flux of the source but also on the number of atoms or ions in a state capable of absorbing the photon, most absorption/excitation steps involve transitions that originate in the ground electronic state where a majority of the analyte atoms or ions reside. Two-colour (or multiphoton) experiments can also be conducted in which multiple-step excitation is the objective. Similarly to nonresonance fluorescence discussed previously, multiphoton excitation also makes the fluorescent transition energy distinctively different from the excitation energy, and minimizes self-absorption if the terminating levels of the fluorescent wavelength is not the ground state. (See calibration curve discussion below.) Multicolour experiments are almost always restricted to high-intensity sources such as lasers.

Once a photon is absorbed, there is a finite possibility that it will relax radiatively. The fraction of the atoms or ions that fluoresce after they absorb radiation is referred to as the quantum yield, $\Phi$.

### Line widths

Parameters discussed previously for line broadening in emission spectrometry are also applicable for fluorescence spectrometry, especially since many of the same atom producers (flames, inductively coupled plasmas, etc.) are used for both techniques. In fluorescence, the width of the absorption profile can be as critical as that of the fluorescent line profile. This is particularly true when narrow-band excitation sources (e.g. lasers or low-pressure discharge lamps) are employed. At the atomic cloud becomes less transparent, the intensity of the excitation beam that reaches the 'analytical volume' in the source centre also decreases, since it is being attenuated by passing though absorbing atoms in the outer regions of the atomization source. The intensity distribution of this exciting beam is, of course, moderated by $k'_\lambda$. Narrow absorption line profiles and narrow excitation line widths (e.g. from laser and low-pressure discharge lamps) accentuate the impact that this absorption will have on the fluorescent intensity.

### Calibration curves

As in emission spectrometry, the intensity of the detected radiation should be proportional to concentration in the original sample, and this is generally true at low concentrations. As with emission, curvature at low concentrations can occur as a result of ionization interferences in some sources for the more easily ionized elements.

Also similarly to emission spectrometry, self-absorption (and self-reversal) of the fluorescent spectral lines can take place within the sample 'cell' if a resonance line is the line being monitored and detected and the atom density is sufficiently large. This alone can lead to a flattening of the calibration curve at elevated concentrations similar to that observed with emission spectroscopy. In addition, if the excitation beam is also a narrow line source used at a resonance wavelength of the analyte, the excitation beam intensity can be attenuated. This leads to lower populations of the excited state and decreased fluorescence signals when the beam attenuation becomes significant. In fact, with a line source (in contrast to use of a continuum excitation source) at elevated analyte concentrations, the calibration curve can actually roll over, i.e. exhibit a negative slope at higher concentrations.

## List of symbols

$A_{ij}$ = Einstein $A$ coefficient for spontaneous emission from $j$th to $i$th state; $c$ = analyte concentration; $E_j$ = excitation energy of the $j$th state; $f$ = oscillator strength; $g_j$ = statistical weight of the $j$th state; $h$ = Planck constant; $I$ = intensity of emitted/absorbed radiation; $k_B$ = Boltzmann constant; $k'_\lambda$ = proportionality constant ($I^0_{c,\lambda} = k'_\lambda c$); $k''_\lambda$ = constant expressing strength of absorption at $\lambda$; $K_i$ = equilibrium constant in Saha equation; $n_X$ = number density of species X; $T$ = temperature;

$Z(T)$ = partition function; $\Delta E$ = uncertainty in energy of a state; $\Delta\nu_N$ = natural line width; $\lambda$ = wavelength; $\nu$ = line/transition frequency; $\tau_r$ = lifetime of excited state; $\Phi$ = quantum yield.

*See also:* **Atomic Absorption, Theory; Atomic Absorption, Methods and Instrumentation; Atomic Emission, Methods and Instrumentation; Atomic Fluorescence, Methods and Instrumentation; UV-Visible Absorption and Fluorescence Spectrometers.**

## Further reading

Barnes RM (1976) *Emission Spectroscopy*. Stroudsburg PA: Dowden, Hutchinson and Ross.

Boumans PWJM (1966) *Theory of Spectrochemical Excitation*. Bristol: Adam Hilger.

Bower NW and Ingle JD Jr (1981) Experimental and theoretical comparison of the precision of atomic absorption, fluorescence, and emission measurements. *Applied Spectroscopy*, **35**: 317.

Ingle JD Jr and Crouch SR (1988) *Spectrochemical Analysis*. Upper Saddle River: Prentice Hall.

Kuhn HG (1962) *Atomic Spectra*. New York: Academic Press.

Mitchell ACG and Zemansky MW (1961) *Resonance Radiation and Excited Atoms*. New York: Cambridge University Press.

Vickers TJ and Winefordner JD (1972) Flame spectrometry. In: Grove EL (ed) *Analytical Emission Spectroscopy*, Vol 1, Part II. New York: Marcel Dekker.

Winefordner JD, Schulman SG and O'Haver TC (1972) *Luminescence Spectrometry in Analytical Chemistry*. New York: Wiley-Interscience.

# Fluorescence Microscopy, Applications

**Fred Rost**, Ashfield, NSW, Australia

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| **Applications** |

This article gives an overview of the current applications of the technique of fluorescence microscopy. The four sections describe, respectively, (1) the basic principles of fluorescence microscopy, (2) the types of information which can be obtained by fluorescence microscopy, (3) the technical ways in which fluorescence microscopy can be adapted to study various chemical species and (4) some examples of the range of biological, mineralogical and artificial specimens that can be studied.

## Principles of fluorescence microscopy

Fluorescence microscopy is a technique whereby fluorescent substances are examined in a microscope. It has a number of advantages over other forms of microscopy, offering high sensitivity and specificity.

In fluorescence microscopy, the specimen is illuminated (excited) with light of a relatively short wavelength, usually blue or ultraviolet (UV). The specimen is examined through a barrier filter that absorbs the short-wavelength light used for illumination and transmits the fluorescence, which is therefore seen as bright against a dark background (**Figure 1**). Because fluorescence is observed as luminosity on a dark background, fluorescent constituents of the specimen can be seen even in extremely small amounts. There are several different modes of fluorescence microscopy, of which the most important is confocal fluorescence microscopy.

In most modern fluorescence microscopes, epi-illumination is employed. This means that the light used for excitation is reflected onto the specimen through the objective, which acts as a condenser. Opaque or very thick objects can be examined using epi-illumination, even the skin of living people.

The position of fluorescence microscopy in relation to other techniques is summarized in **Table 1**. Conventional, transmitted-light, absorption microscopy is appropriate for coloured objects of resolvable size, and instrumentally is the simplest form of microscopy. Colourless, transparent objects can be studied only by retardation techniques (polarization, phase-contrast, interference); these techniques depend upon conversion of phase retardation into changes in intensity that can be seen by the eye. An exception is darkground illumination, which may reveal colourless transparent objects by reflection or refraction at interfaces of differing refractive indices. Darkground microscopy is otherwise suitable mainly for

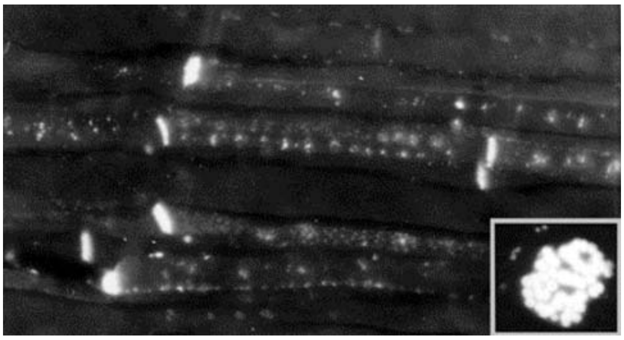


**Figure 1** Fluorescence photomicrographs of a section of a plant stem, cut longitudinally (main picture) and transversely (inset, at higher magnification). The tissue was stained with a fluorochrome, Aniline blue, to show the sugar-conducting tissue (phloem). Aniline blue stains specialized regions of the cell walls. Intense fluorescence is most obvious on the transverse (end) walls of elongated cells. The end walls have pores so that there is continuity from cell to cell, forming a continuous tube in which sugars may be moved down the plant from the leaves. A face view of one of the end walls is shown in the inset. There is a ring of fluorescence around each pore. There is also fluorescent staining, to a much lesser degree, on the longitudinal walls. This surrounds pores that allow sugar transport between adjacent tubes. Aniline blue stains 1→3β–glucans. Here it is staining callose, a glucan which is deposited to occlude the pores should the tubes become damaged, as when tissue is cut. This presumably blocks the cut part of the tubes, minimizing loss of sugars from the cut ends of the tubes, and may well also minimize entry of micro-organisms that would be attracted by leaking sugars. The tubes shown in the main picture are 70 μm in diameter. Field size of main picture approximately 500 × 1000 μm (0·5 × 1 mm), of inset approximately 100 × 100 μm. Pictures courtesy of Professor AE Ashford, University of New South Wales.

particulate matter, and (like fluorescence microscopy) can reveal the positions of particles too small to be resolved. Fluorescence microscopy is closely allied to transmission (absorption) microscopy in its range of application, but possesses particular advantages: great sensitivity for detection and quantification of small amounts of fluorescent substances or small particles, and the possibility of application to opaque objects. Since fluorescence involves two wavelength bands (excitation and emission), optical specificity can be substantially increased by careful selection of filter combinations to favour the excitation and emission of some particular fluorophore, and modern developments also permit time-resolution of the fluorescence lifetime.

Fluorescence microscopy, because of its complexity, gives more difficulty than usual in interpretation of the image. Factors which may affect the appearance of the image in a fluorescence microscope are related to the specimen, to the microscope optical system (particularly the filter combination) and to the observer's own optical and neurological characteristics. In particular, the use of a narrow-band barrier filter can be misleading, since it makes everything appear its specific colour, whereas a wide-band or long(wavelength)-pass filter allows differentiation of different colours. Even photography, apparently objective, may be misleading if not interpreted correctly.

The current definitive texts on fluorescence microscopy are those of Rost (see Further reading). There also exists several introductory works, such as that of Abramowicz, and a vast specialized literature. The major texts on confocal fluorescence microscopy are those of Pawley and of Wilson.

## Types of information to be obtained

Because many substances are fluorescent, or can be made so, fluorescence microscopy is widely applicable. It has a number of advantages over other forms of microscopy, offering high sensitivity and specificity. Because fluorescence is observed as luminosity on a dark background, specific constituents of the specimen can be seen even in minute amounts. Therefore, fluorescence microscopy is used mainly to detect and localize very small amounts of substances in specimens. The chemical substance being studied is either fluorescent of itself, made so by a chemical process or attached to a fluorescent label.

**Table 1** Applicability of fluorescence microscopy, compared with other techniques

| | | *Type of microscopy* | | | |
|---|---|---|---|---|---|
| *Specimen* | *Fluorescence* | *Absorption (transmission)* | *Retardation (DIC, Pol, etc.)* | *Reflection (incl. darkground)* |
| Coloured | Suitable | Suitable | Unsuitable | Suitable |
| Transparent | Impossible | Impossible | Suitable | Impossible |
| Opaque | Suitable | Impossible | Suitable | Suitable |
| Dynamic | Suitable | Impossible | Impossible | Impossible |
| Particles below limit of resolution | Suitable | Impossible | Unsuitable | Suitable |

Absorption microscopy is the conventional transmitted-light type. Retardation microscopy includes Nomarski interference-contrast (DIC), phase-contrast and polarization. Reflection microscopy includes darkground.

Fluorescence microscopy offers three significant capabilities. First, its high sensitivity allows very low concentrations of specific substances to be localized. Secondly, specific substances can be localized in structures smaller than the resolution of the microscope. Thirdly, fluorescence is particularly suitable for confocal microscopy, which offers optical sectioning, giving very clear imaging and the possibility of building up 3D reconstructions.

As in other forms of microscopy, there are three basic kinds of fluorescence microscopy: qualitative, quantitative and analytical. Qualitative fluorescence microscopy is concerned with morphology, or with whether something (e.g. an immunological reaction) is present. Quantitative fluorescence microscopy is concerned with finding out how much of a specific substance is present in a specified region of the specimen. Analytical fluorescence microscopy is the characterization of a fluorophore by measurement of excitation and emission spectra or other characteristics such as polarization or decay time. Kinetic studies essentially involve studying the fluorescence parameters described above over a period of time; examples include the study of fading rates, enzyme kinetics, time-resolved fluorescence and phosphorescence.

### Photomicrography

Images may be captured by photographic or digital means. Photography is best achieved by fast colour-positive (transparency) film in a small (35 mm) format. Film is certain to be largely replaced by solid-state electronic devices (CCDs) in the near future because these have a higher quantum efficiency, enabling the recording of fluorescence which is weaker and/or the use of a shorter exposure time. Confocal microscopy produces a digital image directly.

### Video intensification microscopy (analogue and digitized)

Attachment of a video camera to a fluorescence microscope facilitates viewing of the image, which is normally rather dim. A video camera, particularly one using a CCD, can readily be interfaced to a computer for image analysis. For details of video microscopy, see the book by Inoué and Spring and for an introduction to technique see the book by Herman, both listed under Further reading.

### Confocal fluorescence microscopy and 3D imaging

If both the illumination and detection systems are simultaneously focused onto a spot which is scanned over the specimen, this is called confocal fluorescence microscopy, and has several advantages. Most importantly, the confocal microscope has sharp depth discrimination, and therefore produces optical sections in the plane of focus. This is because the detector does not accept all the light from out-of-focus planes and so these are imaged less strongly than the in-focus plane. Background fluorescence is thereby sharply reduced. Three-dimensional data can be obtained by computer reconstruction of a series of optical sections. Confocal fluorescence microscopy offers greatly increased discrimination against background and some increase in resolution, and gives the possibility of optical sectioning and three-dimensional reconstruction. For detailed information, see the books by Gu, Matsumoto, Pawley, and Wilson listed under Further reading.

### Other modes

Several other modes of fluorescence microscopy have been developed, including automated fluorescence image cytometry (AFIC), fluorescence resonance energy transfer microscopy (FRETM), two-photon fluorescence microscopy, total internal reflection fluorescence microscopy, and standing-wave fluorescence microscopy.

### Combination with other techniques

Because the image seen in the microscope may consist of only a few small fluorescent areas in an otherwise black field, fluorescence microscopy is sometimes supplemented with other forms of microscopy to enable the specimen as a whole to be visualized and to show the position of fluorescent areas in relation to the rest of the specimen. Fluorescence microscopy is easily combined simultaneously or alternately with other methods for increasing object contrast, such as darkground illumination, interference-contrast (DIC) or phase-contrast microscopy. Combined electron microscopy and fluorescence microscopy of the same stain is possible with a small number of stains which contain heavy atoms.

### Quantification by microfluorometry

For quantification, microfluorometry has four advantages as compared with microdensitometric methods: distributional error does not occur, so that scanning or two-wavelength techniques are not required; optical specificity can be obtained by selection of appropriate wavelengths both for excitation and for measurement of emission; great sensitivity can be obtained; and thick or opaque objects can be examined. Disadvantages of microfluorometry include the necessity for artificial standards, and (usually) relatively rapid fading of the specimen. It is

important to ensure that the background fluorescence is kept to a minimum and that an adequate standard is available.

### Microspectrofluorometry

Excitation and/or emission spectra are measured from a specified area of the specimen using a microspectrofluorometer. Such instruments usually have to be assembled by addition of an appropriate light source and monochromators to a microfluorometer. This can be applied either for the identification of unknown fluorophores or to the investigation of the conditions of fluorescence of known fluorophores. The main application to date has been the identification of neurotransmitter amines in situ using formaldehyde-induced fluorescence.

## Kinetic studies: time-related and video fluorescence microscopy

Kinetic studies essentially involve studying the fluorescence parameters described above over a period of time; examples include the study of fading rates, enzyme kinetics, time-resolved fluorescence and phosphorescence.

Scanning of the image with a television camera, with the televised image shown on a video monitor, enables an otherwise dim image to be seen more conveniently. Time-lapse video recording can be used to study movements of fluorescent probes in living cells.

The specimen often photobleaches (fades) under irradiation. Photobleaching is put to good use in modern studies of molecular kinetics using video recording and subsequent measurements of fluorescence recovery after photobleaching (FRAP), which gives information about rates of diffusion of fluorescent species back into a bleached area.

### Fluorescence lifetime imaging microscopy (FLIM)

In this technique, the instrumentation provides time resolution of the fluorescence process, enabling spatial resolution of changes in fluorescence lifetime caused by environmental changes. This is useful for monitoring intracellular processes which cannot be studied by other means.

## Types of fluorescence

### Autofluorescence

Autofluorescence (primary fluorescence) is the fluorescence of naturally occurring substances, such as chlorophyll, collagen and fluorite. Most plant and animal tissues show some autofluorescence when excited with ultraviolet light (e.g. light of wavelength around 365 nm). Sometimes this autofluorescence is a nuisance as it may conceal or be confused with a specific fluorescence. The term autofluorescence is sometimes extended to cover the fluorescence of drugs and other exogenous substances which may be present in tissues, e.g. tetracycline antibiotics which accumulate in growing bone. In biology, autofluorescence is often regarded as a nuisance, masking the specific fluorescence of introduced substances, but it usually provides a guide to the morphology of the tissue, and is also worthy of study in its own right. In geology, fluorescence microscopy is widely used to study coal. In minerals, autofluorescence is most commonly due to the presence of trace amounts of activator substances, generally regarded as 'impurities'.

### Induced fluorescence

This is observed following a chemical reaction to convert some non-fluorescent substance, already present in tissue, into a fluorescent substance. The most important example is the demonstration of neurotransmitter amines such as noradrenaline, which can be converted into fluorescent substances by treatment with formaldehyde or glutaraldehyde.

### Fluorochromy

This is the process in which specimens are stained with fluorescent dyes, as compared with diachromy in which tissues are stained with coloured dyes. An example is shown in **Figure 1**, in which structures in a botanical specimen are stained with a fluorescent dye, Aniline blue. Many dyes which are used as coloured stains are also fluorescent, and can be used as either diachromes or fluorochromes. Hence, many conventionally stained preparations can be examined with either a conventional microscope or a fluorescence microscope. However, the use of fluorescence microscopy extends the range of dyes which can be used, because of the additional possibility of using dyes which are colourless or nearly so but fluoresce in the visible region after excitation with ultraviolet light. Fluorescence also gives increased sensitivity. Staining can also be carried out after chemical treatment, as in the Schiff reaction for carbohydrates and in the Feulgen reaction for nucleic acids. Some staining procedures commonly applied to biological material are listed in **Table 2**. For data and references on fluorochromes and fluorescent probes, see the Molecular Probes catalogue and the books by Kasten, Mason and Rost.

**Table 2** Some examples of fluorochromes (fluorescent stains)

| *Field* | *Application* | *Stain* |
|---|---|---|
| Bacteriology | Acid-fast bacteria | Auramine O |
| Botany | Plant cell walls | Calcofluor White |
| Cell biology | Chromosome banding | Quinacrine, Quinacrine mustard |
| Cell biology | Labels for affinity reactions | FITC, Texas Red isothiocyanate |
| Cell biology | Membrane lipids | ANS, SITS |
| Cell biology | Nucleic acids | Acridine Orange, bisbenzimide |
| Neurology | Nerve cell tracing | DiI |
| Pathology | Amyloid | Thioflavine T |

**Immunofluorescence** This is a special form of fluorochromy, in which the site of an antigen- antibody reaction is revealed by labelling one of the components of the reaction (the antigen, the antibody or complement) with a fluorescent dye. Such reactions allow specific proteins to be localized rapidly, reasonably accurately and with great sensitivity.

Similar techniques are used for the demonstration of other cytochemical affinity reactions, such as in nucleic acid hybridization and in the binding of lectins, hormones and other substances. I refer to these techniques as indirect fluorochromy, since the fluorochrome is bound indirectly (through at least one intermediate molecule) to the substrate (the substance intended to be stained). Immunofluorescence, nucleic acid hybridization and similar affinity reactions have made an enormous contribution to cell biology.

Double- and triple-staining techniques enable visualization of two or more affinity reactions in the same preparation, using differently coloured fluorochromes for each reaction. Special filter combinations are available to facilitate simultaneous visualization of the different colours, and wedge-free filters allow sequential photography of images in each colour without image shift.

### Enzymatically induced fluorescence

In general, enzyme activity is demonstrated by fluorescence microscopy as follows. A substrate is offered to the enzyme, which is allowed to act on the substrate to obtain a reaction product which is localized at the site of enzyme activity and is either fluorescent or easily rendered so. The technique is usually used for qualitative purposes only, to demonstrate the location of enzyme activity, but quantification of enzymatic activity and estimation of Michaelis-Menten constants ($K_m$) are possible. It is usually easier to demonstrate the presence of the enzyme protein using immunofluorescence, but of course this does not guarantee that the enzyme system is active.

## Applications

Fluorescence microscopy is widely used in biology and medicine, as well as in other fields. The literature is vast, and only a very brief outline of actual or potential applications can be presented here.

Fluorescence techniques can be applied to all kinds of material. Because fluorescence microscopy requires more complex and expensive instrumentation than conventional transmitted-light microscopy, fluorescence microscopy is usually reserved for those applications in which its high sensitivity is of importance: i.e., to examine substances present in low concentrations. Fluorescence microscopy can also be applied to detect particles below the resolution of a light microscope, and in histochemistry to visualize substances which cannot be seen by conventional microscopy - e.g. neurotransmitter amines. Biological material is commonly stained in some manner with a fluorescent stain.

### Biological applications

Probably the most important applications of fluorescence microscopy are to the study of living cells and tissues, to protein tracing by the Coons fluorescent antibody technique and to the study of nucleic acids by in situ hybridization.

Fluorescence microscopy has long been used in the study of living tissues (**Table 3**). This is possible because fluorescence microscopy reveals the position of very tiny amounts of fluorescent substances, which

**Table 3** Some examples of fluorescent probes used in living tissues

| *Application* | *Probes* |
|---|---|
| Bone growth | Tetracycline, chlortetracycline |
| Calcium ions | Fura-2, Indo-1, Calcium Green |
| Magnesium ions | Mag-Fura-2 |
| Membrane lipids | ANS, SITS, Nile Red |
| Mitochondria | Rhodamine 123, DASPI, DASPMI |
| Neoplasia | Haematoporphryin derivative |
| Neuronal tracing | Fast Blue, Lucifer Yellow |
| Nuclei | Acridine Orange, bisbenzimide, ethidium bromide |
| pH Indication | BCECF, carboxyfluorescein, dicarbocyanines |
| Streaming (plant cells) | Ethidium bromide |
| Viability | 6-Carboxyfluorescein diacetate |

can be introduced into living tissues or cells. Information about the environment of fluorescent probes in the tissues can be obtained by measuring changes in fluorescence spectrum or other characteristics. This technique is now widely applied to the intravital study of the concentration of various inorganic ions. The measurement of pH and of calcium ion concentration are particularly important.

The viability of cells is testable by using fluorescent esters, such as 6-carboxyfluorescein diacetate, which are non-fluorescent and which (being non-polar) can pass through the cell membrane. Inside the cell, the ester is hydrolysed to release the free fluorescent anion (6-carboxyfluorescein), which, if the membrane is intact, accumulates inside the cell. Fluorescence is also widely used to demonstrate the ramifications of nerve cells, by injection or takeup of fluorochromes into or by living cells.

### Mineralogical applications

Fluorescence microscopy is widely and routinely used for the study of coal.

### Technological applications

Fluorescence microscopy is ideal for studies of porosity in ceramics, using a fluorescent dye. It is also applicable to studies of semiconductors.

## List of symbols

$K_m$ = Michaelis- Menten constant.

*See also:* **NMR Microscopy; Rotational Spectroscopy, Theory; Scanning Probe Microscopy, Applications; Scanning Probe Microscopy, Theory.**

## Further reading


Abramowitz M (1993) Fluorescence Microscopy - the Essentials. New York: Olympus-America.

Bollinger A and Fagrell B (1990) Clinical Capillaroscopy. Toronto: Hogrefe & Huber.

Conn PM (ed) (1990) Methods in Neurosciences, Vol III, Quantitative and Qualitative Microscopy. San Diego: Academic Press.

Coulton GR and de Belleroche J (eds) (1992) In situ Hybridization: Medical Applications. Lancaster: Kluwer Academic.

Gu M (1996) Principles of Three-dimensional Imaging in Confocal Microscopes. Singapore: World Scientific.

Haugland RP (1996) In: Spence MTZ (ed) Handbook of Fluorescent Probes and Research Chemicals, 6th edn. Eugene, OR: Molecular Probes.

Herman B (1998) Fluorescence Microscopy, 2nd edn. Oxford: Bios Scientifica Publishers / The Royal Microscopical Society.

Inoué S and Spring KR (1997) Video Microscopy, 2nd edn. New York and London: Plenum Press.

Kasten F (1981) In: Clark, G. (ed) Staining Procedures, 4th edn, Chapter 3. Baltimore: Williams & Wilkins.

Kawamura A and Aoyama Y (eds) (1983) Immunofluorescence in Medical Science. Tokyo: University of Tokyo Press; Berlin/Heidelberg/New York: Springer Verlag.

Kotyk A and Slavick J (1989) Intracellular pH and its Measurement. Boca Raton, FL: CRC Press.

Mason WT (1993) Fluorescent and Luminescent Probes for Biological Activity. London: Academic Press.

Matsumoto B (ed) (1993) Cell Biological Applications of Confocal Microscopy. New York: Academic Press.

Pawley JB (ed) (1995) Handbook of Biological Confocal Microscopy, 2nd edn. New York: Plenum Press.

Polak JM and Van Noorden S (eds) (1983) Immunocytochemistry. Bristol: Wright.

Rost FWD (1991) Quantitative Fluorescence Microscopy. Cambridge: Cambridge University Press.

Rost FWD (1992) Fluorescence Microscopy, Vol I. Cambridge: Cambridge University Press. (Reprinted with update, 1996.)

Rost FWD (1995) Fluorescence Microscopy, Vol II. Cambridge: Cambridge University Press.

Rost FWD and Oldfield RJ (1999) Photography with a Microscope. Cambridge: Cambridge University Press.

Wilkinson DG (1994) In situ Hybridization. A Practical Approach. Oxford: IRL Press.

Wilson T (ed) (1990) Confocal Microscopy. San Diego: Academic Press.

# Fluorescence Polarization and Anisotropy

**GE Tranter**, Glaxo Wellcome Medicines Research Centre, Stevenage, Herts, UK

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| **Theory** |

Fluorescence polarization techniques have proved to be a very useful method of investigating the environment and motion of fluorescent molecules and fluorophores within larger macromolecules.

When molecules are illuminated by plane-polarized light, those with their transition electric dipole moment parallel with the plane of polarization will preferentially absorb the light (‘photoselection’). Subsequent emission of light from this excited electronic state, as the molecule returns to the original electronic ground state, will be similarly polarized parallel to the transition moment of each molecule. If the molecules are unable to rotate during the lifetime of the excited state, then the plane of polarization of the emitted light will be consistent and directly related to that of the exciting light (their electric vectors will be parallel). However, if the molecules are able to rotate within this time period, then the degree of polarization will be diminished. A similar description can be envisaged for fluorophoric groups within larger molecules; in this case the rotational relaxation involves the tumbling of the complete molecule together with the localized motion of the fluorophore with respect to the molecular environment.

Although rotational relaxation is the main mechanism for loss of polarization, other phenomena can also contribute to depolarization, not least radiationless transfer among fluorophores. Furthermore, if the emission occurs between a different pair of electronic states from that of the original absorption, with differently oriented transition moments with respect to the molecular axes, then a change in polarization may be observed.

## Steady-state anisotropy

In a typical fluorescence spectrometer, the excitation beam illuminating the sample and the optical path employed to detect the emission are orthogonal to each other and define a ‘horizontal’ plane. By employing linear polarizers before and after the sample, the plane of polarization of the excitation beam can be controlled and the polarization components of the emitted light determined. Specifically, by orientating the polarizers as either parallel or perpendicular to the horizontal plane, the plane of polarization of the excitation or detected emission can be described as either parallel to the plane (H, for horizontal) or perpendicular to it (V, for vertical). Four emission spectra can then be acquired:

$V_V$ = intensity of vertical emission from vertical excitation
$H_V$ = intensity of horizontal emission from vertical excitation
$V_H$ = intensity of vertical emission from horizontal excitation
$H_H$ = intensity of horizontal emission from horizontal excitation

In these circumstances, for an *ideal* instrument, the fluorescence polarization, $p(\lambda)$ for a sample at a given wavelength $\lambda$ is

$$p(\lambda) = \frac{(V_V - H_V)}{(V_V + H_V)} \qquad [1]$$

and the corresponding fluorescence anisotropy, $r(\lambda)$, as

$$r(\lambda) = \frac{(V_V - H_V)}{(V_V + 2H_V)} \qquad [2]$$

In principle, this would allow the fluorescence polarization or anisotropy to be determined with just two spectral acquisitions: $V_V$ and $H_V$. However, in practice, it is vital to correct for the differing efficiencies of the emission monochromator and detector towards vertically and horizontally polarized light, which may be substantial. Consequently, for isotropic samples, Equation [1] can be amended to

$$p(\lambda) = \frac{(V_V - \kappa(\lambda)H_V)}{(V_V + \kappa(\lambda)H_V)} \qquad [3]$$

where

$$\kappa(\lambda) = V_{\mathrm{H}}/H_{\mathrm{H}} \qquad [4]$$

and similarly for the fluorescence anisotropy,

$$r(\lambda) = \frac{(V_{\mathrm{V}} - \kappa(\lambda)H_{\mathrm{V}})}{(V_{\mathrm{V}} + 2\kappa(\lambda)H_{\mathrm{V}})} \qquad [5]$$

The basis for the correction ratio $\kappa$ is that, for horizontally polarized excitation of isotropic samples, the resulting vertically and horizontally polarized fluorescences detected should be of equal intensity; any difference may be attributed to the polarization dependence of the emission monochromator and detector.

The fluorescence polarization $p(\lambda)$ and anisotropy $r(\lambda)$ both take the value of unity for the limit of complete preservation of polarization during the excited state lifetime (e.g. if the molecules are immobile). Likewise, they both take the value of zero if there is a complete loss of polarization (i.e. a randomization of orientation). However, their values differ for states between these extremes. In principle $-1 < p < 1$ and $-0.5 < r < 1$, although the limits are rarely encountered; negative values may be indicative of processes beyond depolarization through rotational relaxation, such as transitions to other excited states.

If the only significant process for depolarization is rotational relaxation, then the fluorescence anisotropy, for a single molecule or fluorophore, may be given by a form of the Perrin equation:

$$r = \frac{r_0}{(1 + \tau/\phi)} \qquad [6]$$

where $\tau$ is the fluorescence lifetime, $\phi$ is the rotational correlation or relaxation time for the fluorophore's tumbling and $r_0$ is the intrinsic anisotropy in the absence of rotational relaxation. Small molecules (less than 1000 Da molecular mass) with fluorescence lifetimes of the order of $\tau = 10$ ns are able to tumble with relative ease in low-viscosity solvents (e.g. water), typically $\phi = 0.1$ ns and thus $\tau >> \phi$ and there is almost complete depolarization during the fluorescence lifetime ($r \rightarrow 0$). In contrast, a fixed fluorophore in a large protein may have the same fluorescence lifetime, but a rotational relaxation time of $\phi = 10$ ns, so that $\tau \sim \phi$ and there is significant preservation of the polarization.

As a convenient rule of thumb, the rotational relaxation time (in nanoseconds) for a spherical molecule of mass $M$ (g mol$^{-1}$) in a medium of viscosity $\eta$ (g cm$^{-1}$ s$^{-1}$) is

$$\phi \approx 0.03\,\eta M \quad \mathrm{ns} \qquad [7]$$

with, for example, $\eta = 0.01$ g cm$^{-1}$ s$^{-1}$ for water.

In the case of a small fluorescent molecule binding to a large macromolecule, the bound form should be distinguishable from the unbound form by having a substantially greater fluorescence anisotropy. Likewise, fluorophores rigidly held within a macromolecule would show significantly greater anisotropies than those exposed to the solvent and possessing some degree of motional freedom. Thus, one can employ fluorescence anisotropy to locate groups within a structure. This can prove particularly powerful if an environmental change (including perhaps the binding of a compound) induces a change in rigidity about observable fluorophores. Finally, if the variation of fluorescence anisotropy is plotted against wavelength, spectroscopic bands originating from distinct transitions or fluorophores may be distinguished by their differing anisotropies.

## Fluorescence lifetime anisotropy

By the use of fluorescence lifetime instrumentation, one can further determine the evolution of the fluorescence anisotropy with time during and beyond the lifetime of the excited state. In the simplest case of a fluorophore in solution, with a single fluorescence lifetime and depolarization through rotational relaxation alone, the fluorescence anisotropy will decay according to

$$r(t) = r_0 \exp(-t/\phi) \qquad [8]$$

whereas the total emission intensity, $I_{\mathrm{T}}(t)$ will decay as

$$I_{\mathrm{T}}(t) = I_{\mathrm{T}}(t_0)\exp(-t/\tau) \qquad [9]$$

Thus the rotational relaxation and fluorescence lifetime may be directly distinguished.

In more complex cases, perhaps involving binding kinetics or multiple environments, it is possible to deduce characteristic half-lives for the various processes. For example, one may identify subpopulations

of molecules that are bound in different modes and even characterize the residual motion of a molecule while it is bound to a site that is itself tumbling. Nonetheless, these complex applications all require a sufficiently robust hypothesis regarding the phenomena being modelled if the interpretations are to be valid.

## List of symbols

$H_H$ = intensity of horizontal emission from horizontal excitation; $H_V$ = intensity of horizontal emission from vertical excitation; $I_T$ = total emission intensity; $M$ = molar mass; $p(\lambda)$ = fluorescence polarization at wavelength $\lambda$; $r$, $r(\lambda)$ = fluorescence anisotropy (at wavelength $\lambda$); $r_0$ = intrinsic anisotropy; $t$ = time; $V_H$ = intensity of vertical emission from horizontal excitation; $V_V$ = intensity of vertical emission from vertical excitation; $\eta$ = viscosity; $\kappa(\lambda) = V_H/H_H$; $\tau$ = fluorescence lifetime; $\phi$ = rotational correlation time.

*See also:* **Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; NMR in Anisotropic Systems, Theory.**

## Further reading

Clark BJ, Frost T and Russell MA (eds) (1993) *Techniques in Visible and Ultraviolet Spectrometry*, Vol 4, *UV Spectroscopy*. London: Chapman and Hall.

Harris DA and Bashford CL (eds) (1987) *Spectrophotometry and Spectrofluorimetry, a Practical Approach*. Oxford: IRL Press.

Ingle JD and Crouch SR (1988) *Spectrochemical Analysis*. Prentice Hall International.

Miller JN (ed) (1981) *Techniques in Visible and Ultraviolet Spectrometry*, Vol 2, *Standards in Fluorescence Spectrometry*. London: Chapman and Hall.

# Fluorescence Spectroscopy in Biochemistry

*See* **Biochemical Applications of Fluorescence Spectroscopy.**

# Fluorescent Molecular Probes

**F Braut-Boucher** and **M Aubery**, Université, Paris V, France

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Methods & Instrumentation**

## Introduction

Fluorescent methods have been extensively developed for analytical purposes. Due to their specificity and sensitivity, fluorescent molecular probes (FMP) are presently largely used for the labelling of living cells. Numerous applications in biomedical research, including pharmacology and toxicology, are in progress. Initially, only flow cytometry was capable of quantifying fluorescent signals in cell biology and obtaining information such as cell type characterization, antigen expression, cell receptor localization, etc. More recently the most important advances have resulted from the combination of probe development for exploring specific cellular functions *in situ* and the use of performing instrumentation. Spectrofluorimetry associates the high sensitivity of fluorescence with the reproducibility of microplate reading. Moreover, new probes are expected in the near future and attempts are being made to develop automatic procedures for monitoring analysis and producing a variety of tests used in cell biology, microbiology, toxicology and oncology. Several reviews and monographs on molecular fluorescent probes are available in the bibliography and are listed in the Further reading section.

Micromethods are one of the most recent applications of fluorescence and efforts are devoted to the adaptation of apparatus capable of direct measurement on cell culture microplates that can be automatically scanned. In basic research, as well as in pharmaceutical drug development, this advanced technology represents a significant gain in productivity and cost control. This article deals first with probe labelling strategy, choice of suitable devices for microplate

reading and data analysis. Secondly, selected applications are discussed.

## Fluorescent molecular probes and practical considerations

Fluorescence theory and applications have been documented by Guilbault. According to this author, a fluorescent probe could be defined as 'a chromophore that undergoes changes in absorption characteristics as a function of environmental disturbance'. The molecule absorbs energy, rises to an excited state, and then returns to its normal state, concomitantly emitting energy as photons. This light emission is named fluorescence. The optimum difference between excitation and emission wavelengths (the Stokes shift) lies between 20 and 50 nm.

Several cellular functions are revealed by probes which react stoichiometrically and thus lead to a quantitative relationship between probe and target. In fact, fluorescent probes are versatile tools that can be used to reveal specific cell properties, such as information about cell structure, function, viability, proliferation or differentiation. Probes include fluorescent dyes, antibodies coupled with fluorochromes or fluorogenic substrates, and oligonucleotides. New fluorescent probes are being constantly developed, and additional fluorescent indicators can be found in catalogues or Web pages of a few specialized companies; among them are Molecular Probes and Fluka Biochemicals. A probe is generally selected because of its specific spectroscopic properties: fluorescence lifetime, fluorescence intensity, or functional characteristics: cellular incorporation, behaviour under pH variation. To minimize cellular probe disturbance, i.e. possible chemical interactions between probe and the tested xenobiotic, low probe concentrations (micro- to nanomolar) must be used. Moreover, to avoid photochemically induced reactive oxygen species production, all experiments have to be carried out with as little light present as possible. It is also necessary to consider quenching which results in fluorescence reduction by absorption of emitted light by a fluorophore or other substances present in the medium, such as phenol red.

Currently, there are a wide variety of highly fluorescent extrinsic dyes that may be grouped in two classes:

- fluorescent exclusion probes, which do not penetrate into the cell and so maintain its integrity;
- vital fluorescent probes, which are incorporated into living cells. The latter are more frequently used and can be divided into four sub-classes:
  - probes related to membrane integrity;
  - probes related to membrane potential;
  - probes related to DNA content;
  - probes related to ion content.

Fluorescent probes and their derivatives commonly used in microspectrofluorimetry are listed in **Table 1** with their specific spectra in UV, visible and near IR areas. The structures of some of the probes mentioned later are given in **Figure 1**.

## Concerning cells

Adherent or non-adherent cells could be obtained from either primary cultures or immortalized cell lines. While better reproducibility is generally obtained with cell lines, with careful monitoring very good results have also been obtained with primary cell cultures (see below). Concerning adherent cells, for better reproducibility, the cell monolayer should be used only when subconfluent. In such controlled condition, the cell outline is clearly visible (**Figure 2A**). Non-adherent cells in suspension must be carefully isolated before distribution into each well (**Figure 2B**). In both cases, monolayer variability or cell clusters should be avoided. Moreover, cellular autofluorescence has to be precisely measured before experiment, although there is a generally low fluorescence background when the labelling procedure is well adapted. The morphology and integrity of the

**Table 1** Fluorescent probes commonly used in microspectrofluorimetry

| *Fluorochrome* | *Excitation wavelength (nm)* | *Emission wavelength (nm)* |
|---|---|---|
| Acridine orange / RNA | 460 | 620 |
| Acridine orange / DNA | 485 | 530 |
| DAPI / RNA | 360 | 510 |
| DAPI / DNA | 360 | 460 |
| DIOC (3) / RNA | 482 | 610 |
| Ethidium bromide / DNA | 530 | 620 |
| Ethidium bromide / RNA | 370 | 620 |
| Hoechst 33258 | 360 | 460 |
| Hoechst 33342 | 365 | 502 |
| Propidium iodide | 530 | 620 |
| Alamar blue | 530 | 590 |
| Rhodamine 123 | 485 | 530 |
| BCECF | 485 | 530 |
| Calcein | 485 | 530 |
| Neutral red | 540 | 635 |
| Resorufin | 530 | 585 |
| Methylumbelliferone(4-MU) | 360 | 460 |

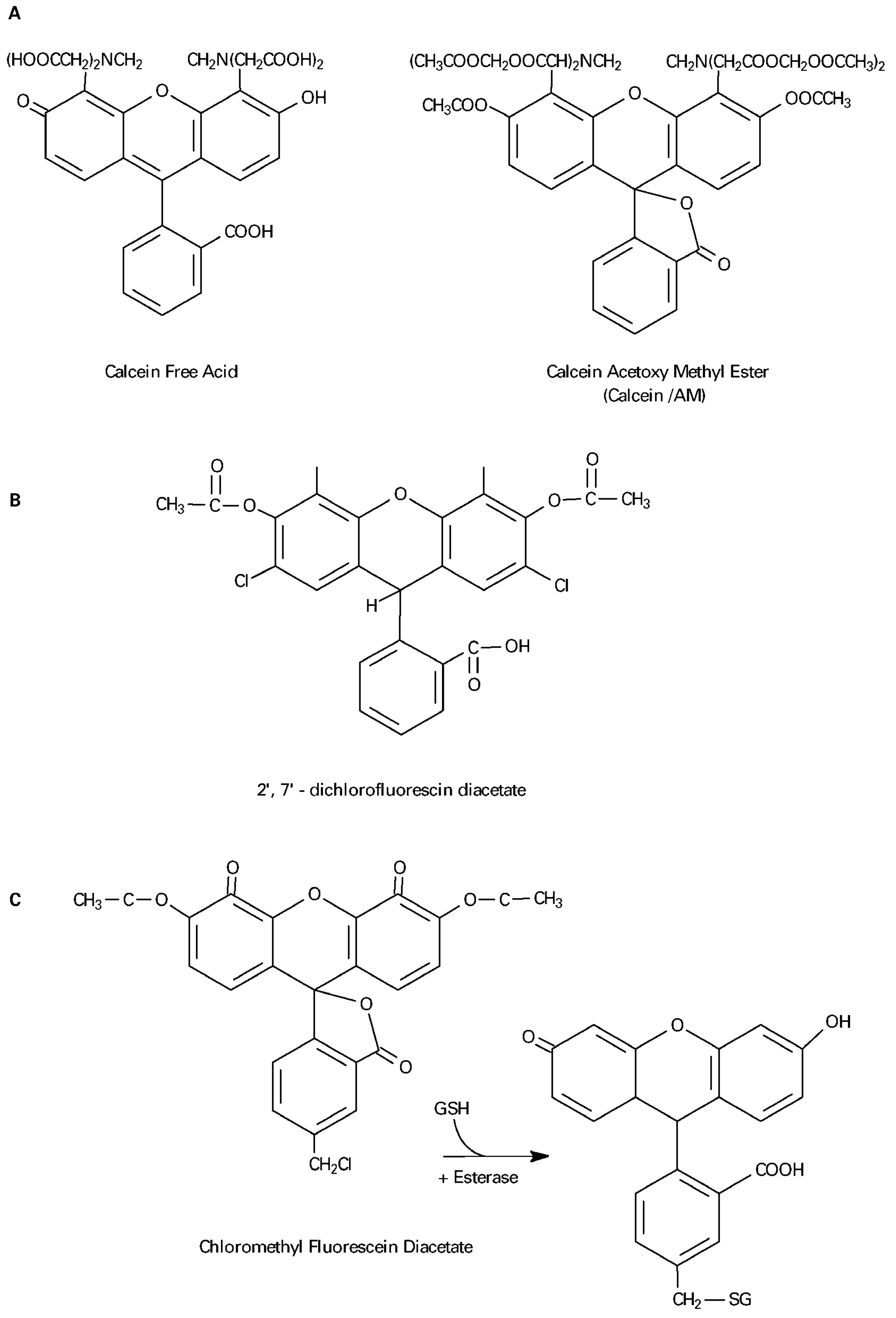



**Figure 1** Structure of some fluorescent probes used in microspectrofluorimetry.

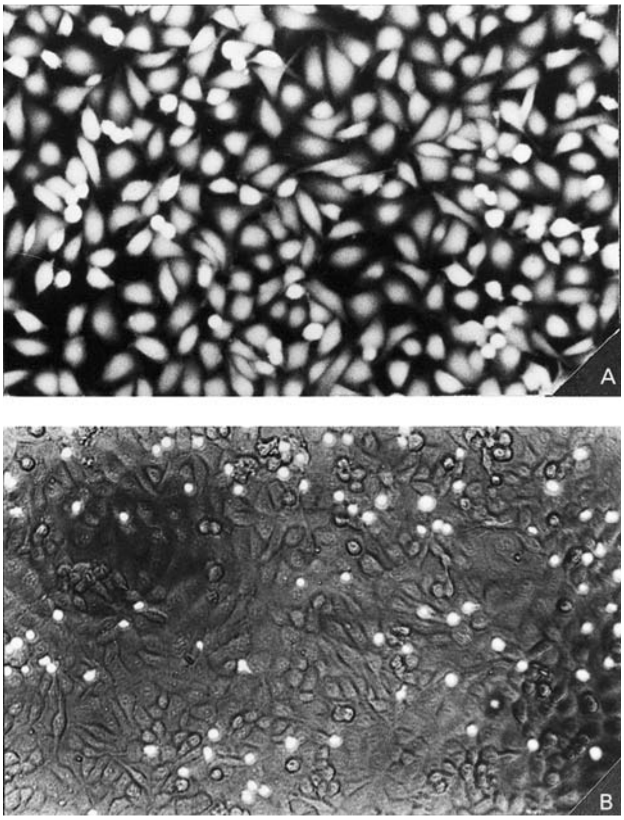



**Figure 2** (A) Endothelial cells (magnification × 200) in a portion of monolayer labelled with calcein-AM (5 μM for 30 min); (B) HL 60 (magnification × 200) in suspension labelled with calcein-AM (10 μM for 30 min) adhering onto an endothelial cell monolayer.

cells and their distribution on the bottom of the microplate have to be monitored to see if they are correctly reproduced from one set to another, paying particular attention to optimization of the culture support. The best optical design is represented by plates with a transparent flat bottom and black edges, but accepting some cross contamination between wells, and standard culture supports can be used in most apparatus. An optimal distribution of microplated cells is shown on **Figure 3**. The cell population could vary in terms of staining levels due, in part, to the state of the type of cells. In any case, the methodology has to be adapted to each cell type and controlled by fluorescence microscopy. Besides validating a new procedure, a comparison could also be made by using fluorescence-activated cell sorting (FACS).

### Choosing a fluorescent probe

The choice of a probe depends on the function or the major target to be explored. The first step of the labelling procedure is the determination of the concentration-response relationship between the probe and the fluorescent signal. Toxic effects of probes related to their concentration are easily controlled, as is non-release of the probe, when experiments are carried out in dilute samples. Higher concentration could induce leakage or quenching. The probe stability in the storage solvent and the culture medium has to be controlled. As deleterious effects and structural damage have been observed in loaded cells exposed to high light energy, potential phototoxic effects, i.e. peroxidative damages or necrosis, must be considered. Moreover, probes may be photolabile, and thus it is necessary to expose the sample in the wells for a short time, at the lowest possible light energy. In these experimental conditions, phototoxic interferences are minimized. In most apparatus, the microplate is scanned at a defined wavelength for 0.3 s, which results in limited fading.

## General labelling procedure

In all cases, whatever the probe used, concentration-response and kinetics curves should be established. These studies will indicate the optimal concentration and incubation time for the probe, depending on the cell type. As an example, calcein-AM is used at 20 μM for lymphocytes and 5 μM for keratinocytes, at 37°C, for a 30 min incubation period. Before loading, the cells are washed three times with PBS and then incubated with the appropriate concentration of the probe in PBS-BSA (0.2%) for 30 min. After elimination of the non-incorporated probe by centrifugation, labelled cells are distributed over the plate, in accordance with the scheme presented in **Figure 3**. The fluorescence

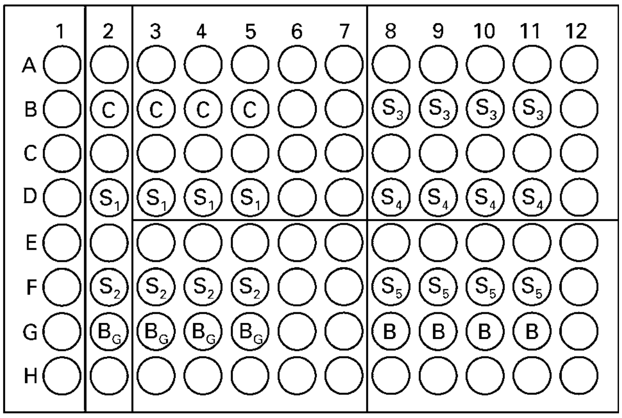



**Figure 3** Experimental setup with configuration of the 96-well microtitre plate. $B_G$, background fluorescence; B, blank (solvent and xenobiotic); C, control (cells loaded with probe without any treatment); S, samples (cells loaded with probe, treatment applied in quadruplicate; five concentrations of xenobiototic could be tested S1 to S5).

is then rapidly monitored using a cytofluorimeter. All manipulations concerning the loading of cells with fluorophores are carried out avoiding exposure to direct light. Generally, probes are aliquoted in a concentrated stock solution (0.1 to 1 mM), and diluted before use of the required final concentration. At low concentrations (below 50 μM), probes have little deleterious effect on cells. The efficiency of probe incorporation is temperature-dependent, with an optimum at 37°C. The excess probe is eliminated by gentle centrifugation and the cell pellet is resuspended in the medium and examined under an inverted microscope. Finally, the fluorescence is measured using a microplate reader.

### Selecting a microplate reader

The instrumentation for microspectrofluorimetry consists of two essential units: a source of radiation and a system for measuring the intensity of the fluorescence emission. The basic components of this apparatus are shown on **Figure 4**. The different

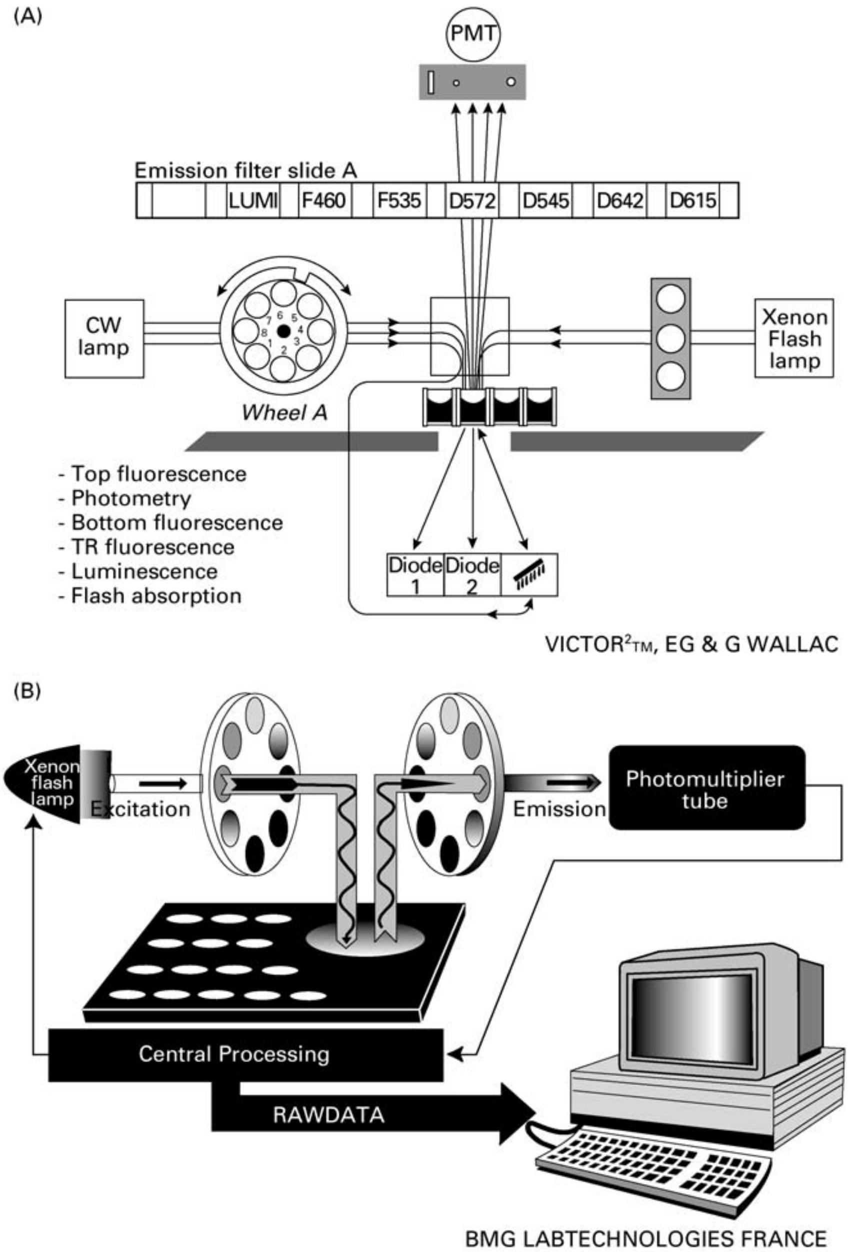



**Figure 4** (A) Diagram of an epiillumination pathway in a fluorimetric plate reader. (Reproduced by permission of EG & G WALLAC.) (B) Schematic overview of the main components of the Fluostar. (Reproduced by permission of BMG Labtechnologies, France.)

appliances available from various companies are compared in **Table 2**. As an energy source, the xenon lamp has a better continuum than that of the tungsten lamp in the ultraviolet spectrum. For this reason the xenon lamp is the most widely used source of radiation in spectrofluorimetry.

For quantitative measurements, excitation and emission monochromator can be set at the optimum wavelength for the probe. Difficulties are encountered with apparatus according to the chosen plate reading method above or below the plastic support of the microplate; to counter this, most manufacturers equip their instruments with two types of plate reader. To diminish interference with other biological molecules, longer wavelengths are preferred to shorter ones.

Environmental factors also have to be considered. The main inconvenience using fluorimetry comes from parameters such as (i) temperature: the decrease of fluorescence with temperature is estimated to be 1% per °C, consequently automated temperature control is of the upmost importance, (ii) light bleaching: it is necessary to operate out of direct light, (iii) pH. All appliances should serve as routine instruments and provide a wide range of excitation wavelengths. Most of them are provided with many accessories, such as automatic sample distributor, temperature regulator, shaking system, etc., as listed in **Table 2**.

### Collecting data (interpretation)

Each test must be repeated at least three times and each microplate has it own control with quadruple wells for each concentration. The intra-assay variability among quadruple concentrations must be less than 5%. The mean and standard deviation should be used in standard tests. Data could be exported into other programs and dealt with statistically. Plate-reader software, however, does possess graphical functions and data analysis capabilities. A wide range of graphical functions, such as linear regression and statistical tests, are included in most apparatus.

## Microspectrofluorimetry: biological applications

In the following section, selected applications of fluorescence are briefly discussed, and experimental results given. Our aim will be restricted to the presentation of some specific probes which target different mechanisms of cytotoxicity, oxidative stress, detoxification potential or cellular adhesive capacities. We demonstrate here that the selection of the most appropriate method for early detection of the cytotoxicity of xenobiotics is directly related to the mechanism of the toxic agent.

### Fluorescent probes for studying cell viability

To study cell viability, several probes are proposed, as indicated in **Table 3**. Two examples will be given in this section. They correspond to either membrane damage (calcein-AM) or disturbance of mitochondrial potential (Alamar blue). Cell viability is often estimated by the Neutral red assay which is widely used in classical spectrometry and, more recently, in fluorimetry because of its increased sensitivity. More often than not, the fluorogenic dye calcein

**Table 2** Some manufacturers of fluorescent microtitreplate readers: apparatus and specifications

| *Manufacturers and instruments* | *Fluorescence reading* | *Light source and power* | *Temperature control / plate shaking* | *Sensitivity (lowest detectable fluorescein concentration)* |
|---|---|---|---|---|
| BMG Fluostar II | 340–700 nm | Xenon 75 W | yes / yes | 10 fmol/well |
| EG & G Instruments | 340–850 nm Xenon flash | Xenon flash | yes / yes | 10 fmol/well |
| FISHER | 340–750 nm | Quartz–halogen 75 W | yes / yes | 10 fmol/well |
| LABSYSTEM Fluoroskan II | 320–670 nm | Quartz–halogen 30 W | yes / yes | 75 fmol/mL (4–MU) |
| TECAN | 320–670 nm | Xenon 75 W | yes / yes | 10 fmol/mL |
| PACKARD Fluorocount | 340–670 nm | Quartz–halogen 150 W | yes / yes | 5 fmol/well |
| MILLIPORE Cytofluor | 340–650 nm | Tungsten | yes / yes | 10 fmol/mL |
| DYNEX (Fluorolite 1000) | 300–880 nm Cold light fluorimetry | Tungsten/ halogen/xenon | no / yes | 10 fmol/mL (Hoechst 33258) |

**Table 3** Biological applications of fluorescent probes in microspectrofluorimetry

| *Tests* | *Fluorescent probes* | *Targets* | *References* |
|---|---|---|---|
| Cell proliferation | Hoechst 33258<br>Propidium iodide | DNA / apoptosis | Papadimitriou E and Lelkes P (1993) *Journal of Immunological Methods* **162**: 41–45 |
| Cellular viability | Calcein-AM | Cytoplasmic membrane | Braut-Boucher F *et al.* (1995) *Journal of Immunological Methods* **178**:41–51 |
| Cellular viability | Neutral red | Lysosomes | Noel-Hudson MS *et al.* (1997) *Toxicology in Vitro* **11**: 645–651 |
| Energetic metabolism | Alamar blue<br>Rhodamine 123 | Mitochondrial system | Shanan TA *et al.* (1994) *Journal of Immunological Methods* **175**: 181–187 |
| | | | Lilius H *et al.* (1996) *Toxicology in Vitro* **10**: 341–348. |
| Oxidative stress | H2DCFDA | Reactive oxygen species | Le Bel CP *et al.* (1992) *Chemical Research and Toxicology* **5**: 227<br>Braut-Boucher F *et al.* (1998) *Veterinary and Human Toxicology* **40**: 178 |
| Detoxification | Resorufin | Cytochrome P450 activity | Rat P *et al.* (1997) In van Zutphen LFM and Balls M, (eds), *Animal Alternatives*, *Welfare and Ethics*, pp 813–825. Amsterdam: Elsevier. |
| Enzymatic systems | CMFDA | Glutathione transferases | Lilius H *et al.* (1996) *Toxicology in Vitro* **10**: 341–348 |
| | Bromobimane | | Rat P *et al.* (1995) Cold light fluorimetry: a microtitration technology for cell culture to evaluate anethole dithiolethione and other thiols. *Methods in Enzymology* **252**: 331–341 |
| Adhesion | Calcein-AM | Cytosolic enzymes | Braut-Boucher F *et al.* (1995) A non-isotopic, highly sensitive, fluorimetric, cell-cell adhesion assay using calcein AM-labeled lymphocytes. *Journal of Immunological Methods* **178**: 41–51 |

acetoxymethylester (calcein-AM) can be used to provide a rapid and sensitive method for measuring cell viability as well as cell- cell and cell- substratum adhesion interactions. Calcein-AM is cleaved by intracellular esterases to produce highly fluorescent and well retained calcein, independently of pH.

The mitochondrial membrane potential can be monitored using several fluorescent probes. These probes are often cationic and accumulate in active mitochondria with a high membrane potential. A review on the use of rhodamine 123 has been published by Slavik. It is stated that the fluorescence intensity of rhodamine 123 in mitochondria of living cells could be quantified by various techniques such as fluorescence microscopy or FACS, and can also be used in microspectrofluorimetry. A new fluorimetric metabolic indicator for detecting mitochondrial activity (which reflects the viability of target cells) is Alamar blue. Alamar blue is a redox indicator that gives fluorescence and visible colour changes in response to cellular metabolic activity. It acts as an intermediate electron acceptor in the electron transport system between the final reduction of $O_2$ and cytochrome oxidase. Alamar blue spectra are shown in **Figure 5**.

**Performing calcein-AM assay in practice** To perform the assay, the probe is added at the concentration determined in preliminary experiments to each well before or after cell treatment with the xenobiotic, be it sodium dodecyl sulfate (SDS) (**Figure 6A**) dinitrophenol (DNP) (**Figure 6B**) or doxorubicin (**Figure 6C**). After incubation for 30 min, all wells are gently rinsed before scanning the microplate using an excitation wavelength of 490 nm and an emission wavelength of 530 nm. The calcein fluorescence in treated samples is compared to that of control cells and the relative fluorescence intensity is used to calculate the viability percentage.

**Performing Alamar blue assay in practice** After treatment with the studied xenobiotic, cells are carefully washed to eliminate proteins which decrease the rate of dye reduction. Alamar blue solution (10% final concentration in PBS) is added to each well and after two hours, supernatants are transferred to a separate 96-well culture plate and the fluorescence is monitored using an excitation wavelength of 560 nm and an emission wavelength of 590 nm.

The cytotoxicities of both DNP and SDS are presented in **Figure 6A** and **B**. For each compound, one can determine the cytotoxic concentration exerting 50% cell damage, i.e. 50% decrease of the fluorescent signal of control cells. Depending on the target, membrane or mitochondrial damage, the best test for SDS is the calcein-AM assay, whereas the Alamar blue assay is the most sensitive one for DNP.

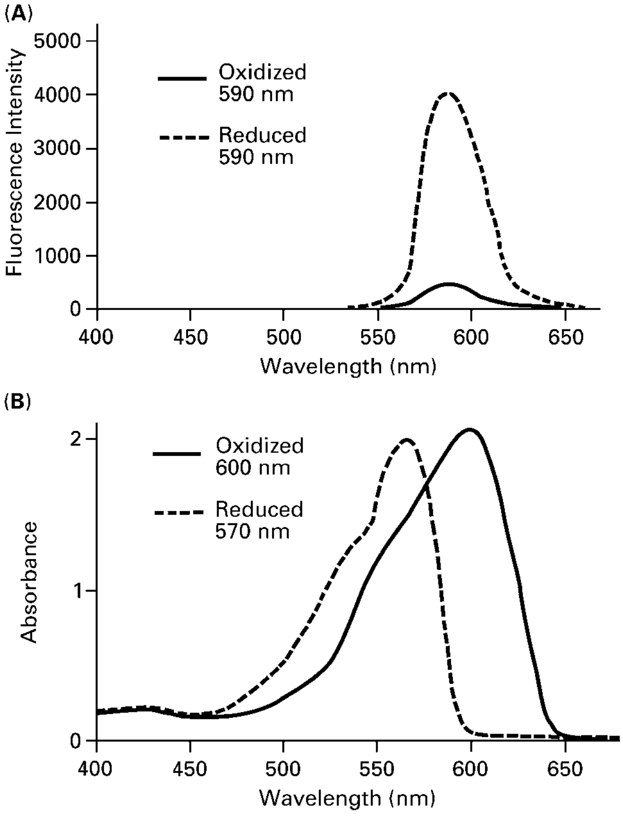



**Figure 5** Fluorescence and absorbance spectra of Alamar blue (oxidized or reduced) reprinted from Interchim, France.

## Probes for assessing oxidative stress

Intracellular peroxide levels were assessed using 2′,7′-dihydrodichlorofluorescein diacetate (H2DCFDA), an oxidation sensitive fluorescent probe. Once inside the cell, the probe is deacetylated by intracellular esterases forming 2′,7′-dichlorofluorescin (H2DCF), which is oxidized in the presence of a variety of peroxides, to a highly fluorescent compound, namely 2′,7′-dichlorofluorescein (DCF).

**Performing H2DCFDA assay in practice** Cells are loaded with 10 μM, H2DCFDA for microplate fluorescent analysis for 30 min. The fluorescence is quantified with the microtitreplate reader using an excitation wavelength of 480 nm and an emission wavelength of 538 nm, with three separate assays per treatment. This protocol was validated using well-known peroxide inducers, such as: H2O2, cumen hydroperoxide and the t-butylhydroperoxide (t-BHP). For each compound, a concentration-response relationship is observed.

The sensitivity of the assay allows the early detection of the formation of reactive oxygen species (ROS) induced by doxorubicin, when applied to a monolayer of endothelial cells. Cells treated with 20 μg/mL of doxorubicin for three hours revealed a 50% increase in ROS measured by the DCFDA assay (**Figure 6D**) which correlates to the cell viability estimated by the calcein-AM assay (**Figure 6C**), as aforementioned.

The study of the cytotoxicity of doxorubicin, a red coloured molecule for which cytotoxic effects involve at least three mechanisms (necrosis/apoptosis, mitochondrial activity and oxidative stress) connected in some way or another, is a prime example to demonstrate the relevance of fluorimetric methods.

## Probes for assessing intracellular sulfhydryl levels

An easy and sensitive method to evaluate the function of the cellular detoxification mechanisms, which could possibly be linked to stress conditions, is valuable. Glutathione (GSH), a tripeptide (Glu-Cys-Gly) which plays a key role in numerous metabolic processes, is a major intracellular reducing agent involved in detoxification of reactive species. A number of flow cytometric methods for measuring cellular GSH have been described using probes that form adducts with GSH. Cellular thiol levels could be analysed using either the monobromobimane (mBBr), which forms a fluorescent adduct with sulfhydryl groups, or 5-chloromethylfluorescein diacetate (CMFDA). CMFDA, a non-fluorescent molecule capable of permeating cells, becomes fluorescent after cleavage by cellular esterases. The glutathione transferase catalyses the reaction between CMF and the intracellular thiols retained inside the cell, whereas the unconjugated probe is released in the medium and eliminated. A decrease of fluorescence reveals a decrease of GSH content.

**Performing CMFDA assay in practice** In order to validate the CMFDA assay, intracellular sulfhydryl groups of the cells were depleted either by using N-ethylmaleimide (NEM), a thiol depleting agent, or by inhibiting GSH synthesis with buthionine sulfoximine (BSO). An overnight incubation with 2 mM BSO reduces the fluorescent signal intensity by 30%, compared to control cultures, whereas a 2 min incubation with 100 μM NEM induces a more acute depletion of 50%.

## Cell–cell adhesion assay

We have previously described a non-isotopic, highly sensitive, fluorimetric, cell-cell adhesion microplate assay using calcein-AM-labelled lymphocytes. To successfully perform the lymphocyte/keratinocyte adhesion assay, the following conditions are recommended:

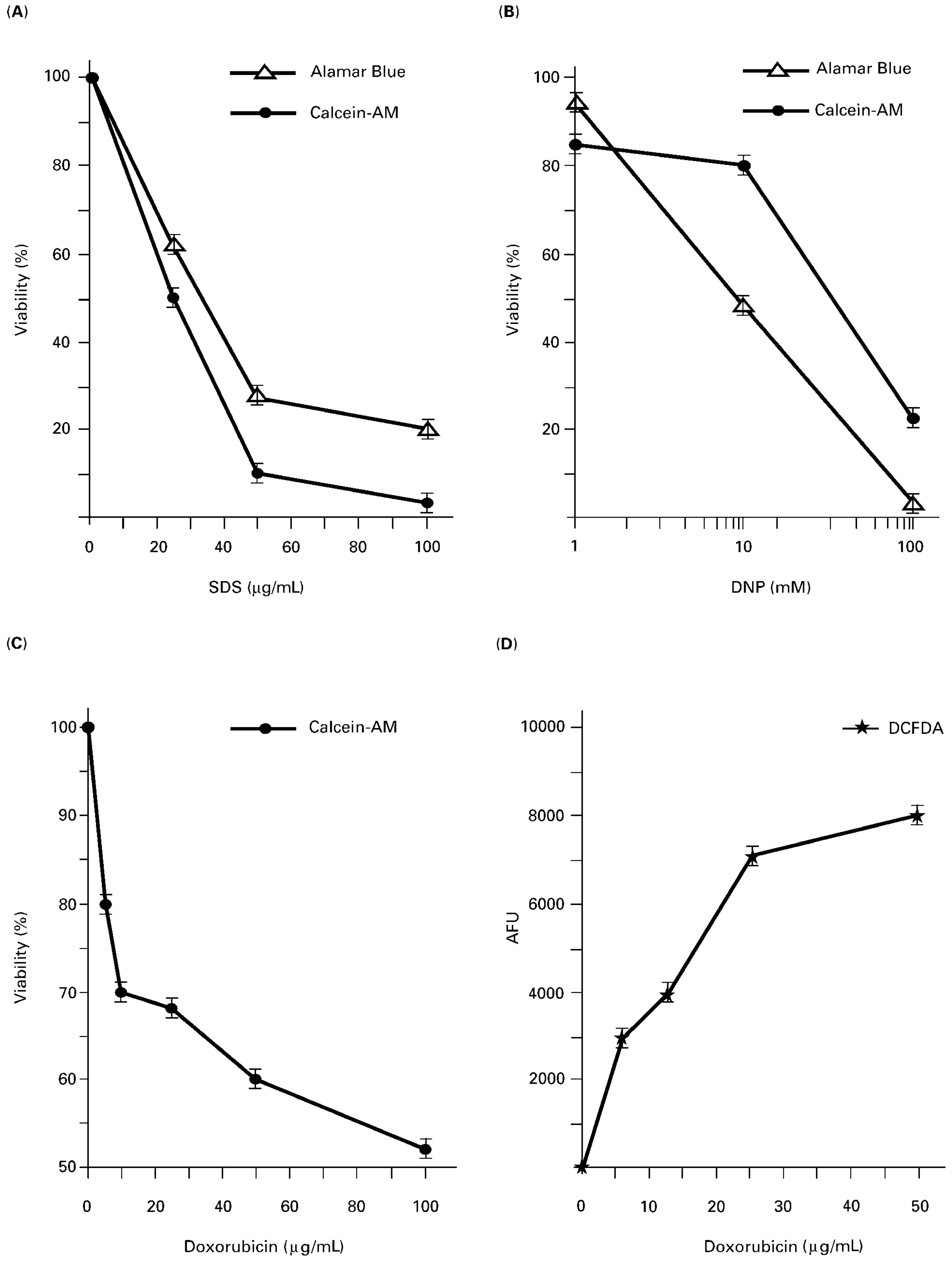



**Figure 6** (A) Cytotoxicity of SDS, Calcein-AM and Alamar blue assays; (B) Cytotoxicity of DNP, Calcein-AM and Alamar blue assays; (C) Cytotoxicity of doxorubicin, Calcein-AM assay; (D) Cytotoxicity of doxorubicin, H2DCFDA assay.

- the confluence of the keratinocyte monolayer should be checked in each well;
- the lymphocytes should be labelled with 20 μM calcein AM and incubated for 30 min;
- a defined lymphocyte density per well should be plated onto the monolayer;
- fluorescence should be monitored immediately using the fluorimetric microtitreplate reader to avoid probe release.

The percentage of lymphocytes adhering to the keratinocyte monolayer was calculated as follows:

$$\%\ \text{adhesion} = \frac{F_A}{F_T - F_R} \times 100$$

where $F_A$ is the fluorescence of adherent lymphocytes, $F_T$ the fluorescence of total lymphocytes deposited, and $F_R$ the released fluorescence. This equation takes into account the background fluorescence emitted by contol wells in the spectrum considered.

This method can be easily applied to any cells whose adhesive capacities have to be investigated and can also be used to manage cell- extracellular matrix adhesion.

## Concluding remarks

The aim of this article was to introduce microspectrofluorimetry, using fluorescence probes, as a new technology in analytical fluorescent applications and provide practical guidelines to the use of this powerful technique. When compared with conventional methods, its main advantages are:

- high sensitivity and specificity;
- simultaneous monitoring of different cellular functions;
- combining two or more probes, if the probes emit at different wavelengths;
- use of coloured molecules without interference;
- easy handling, low cost and possibility of batch studies;
- increasing number of available fluorescent probes and kits;
- avoiding the use of radiolabelled molecules.

Successful applications require the appropriate choice of model system and probes with the aim of defining the target cell functions. Under these conditions, microspectrofluorimetry offers innovative analytical tools for both research and development projects, as well as rapid and accurate assays for quality control. Recently, cytofluorimetric microtitration was recommended by the European Centre for Validation of Alternative Methods (see Further reading). These methods, based on the selection of appropriate test systems, could give strong evidence for possible molecular mechanisms involved in cellular functions.

*See also:* **Cells Studied By NMR; Colorimetry, Theory; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy**; **Fluorescence Microscopy, Applications.**

## Further reading

Brand L and Johnson M (1997) Methods in Enzymology. London: Academic Press.

Guibault GG (1996) Practical Fluorescence, 2nd edn. New York: Marcel Dekker.

Haugland RP (1996) In: Larrison KD (ed) Handbook of Fluorescent Probes and Research Chemicals, 6th edn. Molecular Probes, Eugene, Oregon.

Metezeau P, Ratinaud MH and Carayon P (1994) Cytometrie par fluorescence. Apports comparatifs de techniques flux, image et confocale. Paris: Editions INSERM.

Rat P, Chisten MO, Thevenin M, Warnet JM and Adolphe M (1995) Cold light fluorimetry: a microtitration technology for cell culture to evaluate anethole dithiolethione and other biothiols. Methods in Enzymology, **40**: 252- 331.

Rat P, Osseni R, Christen MO, Thevenin M, Warnet JM and Adolphe M (1997) Microtitration fluorimetric assays on living cells: new tools for screening in cell pharmacotoxicology. In: van Zutphen LFM and Balls M (eds) Animal Alternatives, Welfare and Ethics, pp 813-825. Amsterdam: Elsevier.

Ronot X, Paillasson S and Muirhead, KA (1996) Assessment of cell viability in mammalian cell lines. In: Al-Rubeai M and Emery AN (eds) Flow Cytometry Applications in Cell Culture, pp 177- 192. New York: Marcel Dekker.

Slavik J (1994) Fluorescent Probes in Cellular and Molecular Biology. Boca Raton, FL: CRC Press.

Wang YL and Taylor LD (1989) Methods in Cell Biology part A: Fluorescent analogs, labeling cells and basic microscopy. 29 and part B: quantitative fluorescence microscopy-imaging and spectroscopy 30 Fluorescence microscopy of living cells in culture, Part A and B, (Eds), Academic Press, Inc., San Diego, New York, Berkely, 498 pp.

White CE and Argauer RJ (1970) Fluorescence Analysis. A Practical Approach. New York: Marcel Dekker.

# Food and Dairy Products, Applications of Atomic Spectroscopy

**NJ Miller-Ihli** and **SA Baker**, U.S. Department of Agriculture, Beltsville, MD, USA



Food composition data are critical for consumers and health care professionals to make choices and recommendations based on the nutrient content of foods. Trace elements contribute significantly to human health with 25 elements identified as being of interest. A review of the nutrients for which recommended daily allowances (RDAs) exist shows that 7 of 19 are minerals (Ca, Fe, I, Mg, P, Se, and Zn). In addition, the National Research Council has established estimated safe and adequate daily dietary intakes (ESADDIs) rather than RDAs for an additional five elements (Cr, Cu, F, Mn, and Mo) and three electrolytes (Cl, K, and Na) have been identified as essential. Ten additional trace elements are considered essential but no human requirement level has been set for them: As, Ni, Si, B, Cd, Pb, Li, Sn, V, and Co. The status of fluorine as an essential nutrient has been debated and conflicting data exist but its valuable effects on dental health certainly make it of potential interest.

Recent reviews of the state of knowledge of food composition data shows that there is a significant lack of data for specific food groups and for several of the elements of interest, including the seven for which RDAs have been identified. This is owing, in large part, to the lack of suitable analytical methodology, which is why rugged, accurate methods are needed that use modern analytical methodology and instrumentation. Health professionals and the public are increasingly interested in the maintenance or promotion of good health and disease prevention, rather than the diagnosis and cure of diseases. Clearly good nutrition is an important part of good health so nutritionists need high quality food composition data so that they can assess the nutrient content of foods consumed. Recent reviews of food composition data still suggest the mineral composition of foods is not well characterized despite the extensive data found in U.S. Department of Agriculture's Nutrient Data Base for Standard Reference SR-12 (the former USDA Handbook 8). In addition, there is now interest in knowing the exact chemical form of the trace elements in foods since not all forms are biologically active.

Another issue which highlights the importance of continued research for methods for trace elements in foods and beverages is related to the Nutrition Labelling and Education Act of 1990 (NLEA); this is now law in the United States and provides nutritional information to the consumer. Fifteen mandatory nutrients must be listed on the labels, including three trace elements (Ca, Fe, and Na). Thirty-four additional 'voluntary' nutrients which affect human health may be required in the future and 12 of those are trace elements (K, P, Mg, Zn, I, Se, Cu, Mn, F, Cr, Mo and Cl).

Typically, essential elements have been at sufficiently high concentrations that, with the introduction of graphite furnace atomic absorption spectroscopy (GFAAS) more than two decades ago, detectability is not an issue. At present, researchers are interested in 'new' low-level trace elements with known or potential nutritional impact, which are present at lower concentrations and, because of potential interferences in GFAAS, it has been found that inductively coupled plasma mass spectrometry (ICP-MS) coupled with graphite furnace/electrothermal vaporization (ETV-ICP-MS) provides superior detectability. The most often utilized instrumentation for generation of food composition data includes: atomic absorption spectroscopy [flame (FAAS) and graphite furnace (GFAAS)]; flame atomic emission spectroscopy (Na, K); inductively coupled plasma atomic emission spectroscopy (ICP-AES) and ICP-MS.

In this article, the description of applications will highlight the selection of the instrumentation used as well as the sample preparation methods (often the most critical part of a complete analytical method). Calibration strategies will be discussed and any specialized interference effects will be highlighted. Strategies for validating and facilitating technology transfer of the methods will also be discussed and the need for commercial as well as in-house reference materials will be featured.

## Sampling (sample collection and processing)

One of the most overlooked aspects of analytical determinations relates to the suitability of the sample

provided for analyses. If food composition data are being generated it is very important to know if interest lies with the specific sample at hand or if that sample is supposed to be representative of a larger bulk sample (e.g. 1 jar from a lot totalling 100 000 jars). Work completed in the Food Composition Laboratory (USDA, Beltsville Human Nutrition Research Center, Beltsville, MD) is often in support of the Nutrient Data Laboratory which is responsible for the publication of the U.S. Department of Agriculture's Nutrient Data Base for Standard Reference SR-12 (formerly known as Handbook 8) which is the primary US database for food composition data. As such, the goal is to include data representative of the national US food supply. Because large-scale national sampling is too expensive, samples are often obtained nationwide (if possible) and brands are selected on the basis of popularity (consumption) to try to gain representative samples for analysis. Samples are typically brought into the laboratory and, depending on the form of the sample, a decision is made to either analyse multiple samples or to composite and homogenize the whole of the material (if possible) and to remove separate sample aliquots. Most often, a preliminary evaluation of the homogeneity of a selected number of analytes in the material, serves as a basis to select the sample size and number of replicate samples to be analysed.

Tools are available which facilitate sample homogenization without trace element contamination and equipment that is not available commercially can often be custom made. To dissect meat, titanium blade knives or high purity glass knives may be appropriate. Large-scale mixers (e.g. Robot-Coupe, Jackson, MS) equipped with a plastic-lined bowl and titanium blades to avoid possible contamination from the stainless steel apparatus have been used successfully to blend both common samples and a diet reference material (RM8431) sold by the National Institute of Standards and Technology (NIST, Gaithersburg, MD). Polythene trays and polypropylene storage containers are most often used without risk of contamination. An assessment of contamination risk is summarized in **Tables 1** and **2**. **Table 1** lists those elements prone to contamination while highlighting common sources of contamination. **Table 2** contains a list of materials (from most preferable to least) for use during sample preparation for trace element determinations. To further minimize the likelihood of contamination [particularly for low temperature (≤ 85–95 °C) wet ash preparations] glassware can be silanized; the coating will be retained on the surface of the glassware up to temperatures of approximately 350°C, providing excellent protection from contaminants leached from the glassware owing to matrix interaction with the glass. In addition, silanization prevents loss of analyte due to adsorption on the container walls.

**Table 1** Trace element determinations – risk assessment

| | |
|---|---|
| *Contamination potential* | |
| High potential | Al, Ca, Cr, Cu, Fe, Pb, Sn, Zn |
| Medium potential | Co, K, Mn, Mo, Na, Ni, V |
| Low potential | Mg, Se |
| *Sources of contamination* | |
| Glass | Al, Ca, Co, Cr, Cu, Fe, K, Mg, Na |
| Porcelain | Al, Cu, K, Na, V |
| Rubber | Co, Cu, Fe, Zn, Sn |
| Building materials | Al, Ca, Mg |
| Stainless steel | Co, Cr, Cu, Fe, Mn, Ni, Zn |
| Cosmetics | Fe, Pb, Zn, Cu |
| Skin | K, Na, Ni, Cu, Fe |
| Smoke | Fe, Pb, K, Ni |
| Haemolysis | Ca, Co, Cu, Fe, K, Mg, Mn, Zn |

**Table 2** Sample container materials

| *Material* | |
|---|---|
| Teflon | Most desirable (least trace element contamination) |
| PFA | |
| FEP | |
| Tefzel | |
| Polyethylene | ↓ |
| Polypropylene | ↓ |
| Quartz | ↓ |
| Platinum | ↓ |
| Borosilicate glass | Least desirable |

Contamination control, in a general sense, is not difficult and common sense can often provide protection from sample contamination, especially during the early sample handling stage. In an ideal world, samples are processed in clean rooms (e.g. class 100 clean room) designed to have minimal contamination because the air is filtered through hepa filters. Clean hoods and laminar air flow units can also prove useful, as well as wearing clothing that minimizes particulate generation, selecting powder-free gloves, taking care in cleaning glassware and plasticware with dilute, high purity acid, and covering gas cylinders which may be delivered coated with rust and peeling paint.

## Sample preparation (reagents and ashing methods)

One of the most important means of avoiding unnecessary risk for trace element contamination is

the judicious selection of reagents for sample preparation. Sub-boiling distilled or double-distilled acid is a good choice for sample treatment. As an example, Pb contaminant concentrations found in nitric acid (65–70% v/v) ranged from < 1 part per trillion (ppt) for sub-boiling distilled to 0.3 ppb for commercial 'high' purity acid. Although ppb levels may be suitable in many instances, sub-boiling distilled nitric acid (Seastar Chemicals, Sidney, British Columbia) provides even lower blanks with levels of < 50 ppt for Ca; < 20 ppt Fe and K; < 10 ppt Cr, Cu, Na, Mg, Ni, Sn, Zn; and < 1 ppt for Cd, Mn, and V, and is therefore highly recommended for ultratrace analyses.

Of course all reagents used throughout the course of any analyses provide potential sources of contamination and must be evaluated. Hydrogen peroxide comes in different levels of purity and some types prove to be particularly problematic for only a select number of elements. In the past, unacceptably high Cr levels have often been found. A peroxide called Perone (DuPont, Wilmington, DE), made for the semiconductor industry, has been evaluated in several laboratories and proved suitable for any number of elements (including Mn, Cu, Zn, Fe, Cr, and Co). Unfortunately this is not sold in small enough quantities to be suitable for the typical analytical laboratory. For analysts using matrix modifiers for GFAAS, the modifier can provide a significant level of contamination, indicating that very high purity materials should be used.

In all instances, environmental considerations are equally as important as reagent purity, when selecting a sample preparation method. In an ideal situation, samples can be handled under clean air conditions under laminar flow to prevent contamination. Potential sources of environmental contamination include the analyst (particles from skin or clothing), sample container material, air, gloves, heating block, muffle furnace, etc.

The two most common ashing procedures for food analyses are wet ashing and dry ashing procedures. Wet ashing may be done using a typical hot plate or heating blocks or it may be done using microwaves under increased pressure. Many modern microwave systems are now under temperature and pressure control and various methods are available in the literature. In a laboratory which processes many samples with one matrix type, microwave digestion can be quite useful. If, however, the matrices are variable, experiment design requires that samples of similar type be put in the microwave together. Otherwise, depending on the carbohydrate content, pressure will build up in the samples at different rates, requiring venting which can be problematic if the microwave system does not have temperature and pressure monitoring control for each vessel. Another great benefit is the use of microwave digestion as a reference to compare with wet ashing results. The microwave system provides the benefit of increased pressure, ensuring almost complete digestion (depending on the sample matrix). When used under optimum conditions, microwave digests provide a reliable dissolution of samples and minimize the risk of sample contamination while saving significant time. Microwave sample digestion is very much a growing field and the number of analysts using this sample preparation tool probably increased 10-fold between 1993 and 1998.

## Method selection and method validation

Several factors should be considered in selecting the appropriate sample preparation and instrumental detection method for a particular set of analyses. Factors include detection capability, ease of use, availability of instrumentation, time, cost, etc. One must consider the relative importance of the factors for their application but, most often, detection capability is the top priority. A good general rule is to use the simplest method possible that meets your analysis requirements. For example, if asked to determine Mn, Cu, and Zn in soft drinks, analysts should check to see if it could be carried out with a simple dilution directly using FAAS or simultaneous multielement ICP-AES (for greater speed). Once a method is selected it should be validated using one or more of the following strategies: (1) analysis of one or more standard reference materials; (2) comparison of results using an alternative method; (3) comparison of results with those from another laboratory; and (4) thorough evaluation of possible matrix interference effects (compare direct calibration with the method of additions). **Table 3** contains a list of commercial reference materials with matrices suitable for food and beverages analyses. One unfortunate aspect of solid commercial materials is that they are almost all freeze-dried and this is not a realistic representation of typical samples (e.g. meats, vegetables, fruits, etc.); analysts prefer to analyse 'as received'. Because these primary controls are so expensive, analysts should consider developing secondary, in-house controls that represent the food type/matrices most often analysed in their laboratory. These materials can be characterized using all the validation strategies listed above, with the assistance of colleagues and by using alternative methods.

**Table 3** Commercial reference materials

| *FCL-ID* | *Reference material (µg/g dry weight)* | *Number* | *Manufacturer* | *Ca* | *Cd* | *Co* |
|---|---|---|---|---|---|---|
| R - 0010 | Rice Flour SRM 1568 | 1568 | NIST | 140 ± 20 | 0.029 ± 0.004 | 0.02 ± 0.01 |
| R - 0011 | Rice Flour SRM 1568a | 1568a | NIST | 118 ± 6 | 0.022 ± 0.002 | (0.018) |
| R - 0020 | Wheat Flour SRM 1567 | 1567 | NIST | 190 ± 10 | 0.032 ± 0.007 | |
| R - 0021 | Wheat Flour SRM 1567a | 1567a | NIST | 191 ± 4 | 0.026 ± 0.002 | (0.006) |
| R - 0030 | Total Diet SRM 1548 | 1548 | NIST | 1740 ± 70 | 0.028 ± 0.004 | |
| R - 0040 | Carrot Powder | | ARC/CL | 1700 ± 45 | 0.0686 ± 0.0031 | |
| R - 0050 | Skim Milk Powder | | ARC/CL | 13040 ± 260 | (<0.005) | |
| R - 0060 | Pork | | ARC/CL | 150 ± 6 | 0.022 ± 0.0042 | |
| R - 0070 | Wheat Flour | | ARC/CL | 208 ± 4 | 0.039 ± 0.004 | |
| R - 0080 | Potato | | ARC/CL | 91 ± 6 | 0.035 ± 0.0016 | |
| R - 0090 | Total Diet | | HDP/CL | 2860 ± 124 | 0.021 ± 0.003 | |
| R - 0100 | Skim Milk (recommended values) | | ARC/CL | 13000 ± 291 | (<0.005) | |
| R - 0110 | Mixed Diet SRM 8431 (recommended values) | 8431 | NIST | 1940 ± 140 | 0.042 ± 0.011 | 0.038 ± 0.008 |
| R - 0120 | Non-Fat Milk Powder SRM 1549 | 1549 | NIST | 13000 ± 500 | 0.0005 ± 0.0002 | (0.004) |
| R - 0130 | Milk Powder IAEA A-11 | A-11 | IAEA | 12900 ± 800 | (0.526) | 0.005 ± 0.001 |
| R - 0140 | Oyster Tissue SRM 1566a | 1566a | NIST | 1960 ± 190 | 4.15 ± 0.38 | 0.57 ± 0.11 |
| R - 0150 | Bovine Liver SRM 1577a | 1577a | NIST | 120 ± 7 | 0.44 ± 0.06 | 0.21 ± 0.05 |
| R - 0151 | Bovine Liver SRM 1577b | 1577b | NIST | 116 ± 4 | 0.50 ± 0.03 | (0.25) |
| R - 0160 | Dogfish Muscle and Liver DORM1 | DORM1 | NRCC | | 0.086 ± 0.012 | 0.049 ± 0.014 |
| R - 0170 | Dogfish Muscle and Liver DOLT1 | DOLT1 | NRCC | | 4.18 ± 0.28 | 0.157 ± 0.037 |
| R - 0180 | Dogfish Muscle and Liver DORM2 | DORM2 | NRCC | | 0.043 ± 0.008 | 0.182 ± 0.031 |
| R - 0190 | Dogfish Muscle and Liver DOLT2 | DOLT2 | NRCC | | 20.8 ± 0.5 | 0.24 ± 0.05 |
| R - 0210 | Bovine Muscle Powder SRM 8414 | 8414 | NIST | 145 ± 20 | 0.013 ± 0.011 | 0.007 ± 0.003 |
| R - 0220 | Non-defatted Lobster Hepatopancreas LUTS-1 (as bottled) | LUTS-1 | NRCC | 203 ± 33 | 2.12 ± 0.15 | 0.051 ± 0.006 |
| R - 0230 | Non-defatted Lobster Hepatopancreas LUTS-1 (dry weight) | LUTS-1 | NRCC | 1360 ± 220 | 14.2 ± 1.0 | 0.34 ± 0.04 |
| R - 0240 | Lobster Hepatopancreas TORT-1 | TORT-1 | NRCC | 8950 ± 580 | 26.3 ± 2.1 | 0.42 ± 0.05 |
| R - 0250 | Spinach Leaves SRM 1570 | 1570 | NIST | 13500 ± 300 | (1.5) | (1.5) |
| R - 0251 | Spinach Leaves SRM 1570a | 1570a | NIST | 15270 ± 410 | 2.89 ± 0.07 | 0.39 ± 0.05 |
| R - 0260 | Apple Leaves SRM 1515 | 1515 | NIST | 15260 ± 150 | (0.014) | (0.09) |
| R - 0270 | Citrus Leaves SRM 1572 | 1572 | NIST | 31500 ± 1000 | 0.03 ± 0.01 | (0.02) |
| R - 0280 | Peach Leaves SRM 1547 | 1547 | NIST | 15600 ± 200 | 0.026 ± 0.003 | (0.07) |

**Table 3** *Continued*

| *FCL-ID* | *Cr* | *Cu* | *Fe* | *I* | *K* | *Mg* | *Mn* | *Na* |
|---|---|---|---|---|---|---|---|---|
| R - 0010 | | 2.2 ± 0.3 | 8.7 ± 0.6 | | 1120 ± 20 | | 20.1 ± 0.4 | 6.0 ± 1.5 |
| R - 0011 | | 2.4 ± 0.3 | 7.4 ± 0.9 | (0.01) | 1280 ± 8 | 560 ± 20 | 20.0 ± 1.6 | 6.6 ± 0.8 |
| R - 0020 | | 2.0 ± 0.3 | 18.3 ± 1.0 | | 1360 ± 40 | | 8.5 ± 0.5 | 8.0 ± 1.5 |
| R - 0021 | | 2.1 ± 0.2 | 14.1 ± 0.5 | (0.001) | 1330 ± 30 | 400 ± 20 | 9.4 ± 0.9 | 6.1 ± 0.8 |
| R - 0030 | | 2.6 ± 0.3 | 32.6 ± 3.6 | | 6060 ± 280 | 556 ± 27 | 5.2 ± 0.4 | 6250 ± 260 |
| R - 0040 | | | 12.6 ± 0.79 | | 10200 ± 350 | 350 ± 9 | 4.90 ± 0.17 | 485 ± 24 |
| R - 0050 | 0.0079 ± 0.0032 | 0.58 ± 0.06 | 4.53 ± 0.61 | 1.74 ± 0.07 | 17700 ± 590 | 1230 ± 24 | 0.45 ± 0.07 | 4870 ± 350 |
| R - 0060 | | 2.68 ± 0.283 | 52.84 ± 5.456 | | | (1010) | 0.30 ± 0.024 | |
| R - 0070 | 0.028 ± 0.005 | 2.35 ± 0.18 | 51 ± 3.79 | | 2200 ± 190 | 562 ± 10 | 12.87 ± 0.34 | |
| R - 0080 | 0.098 ± 0.018 | 3.87 ± 0.18 | 22.0 ± 2.0 | | 14100 ± 100 | 747 ± 16 | 8.1 ± 0.32 | 895 ± 42 |
| R - 0090 | | 3.18 ± 0.19 | 30.4 ± 0.9 | | 9420 ± 300 | 785 ± 25 | 12.9 ± 0.58 | 7870 ± 570 |
| R - 0100 | | 0.59 ± 0.148 | 4.54 ± 0.855 | | | 1230 ± 27 | 0.44 ± 0.104 | |
| R - 0110 | 0.102 ± 0.006 | 3.36 ± 0.33 | 37.0 ± 2.6 | | 7900 ± 420 | 650 ± 40 | 8.12 ± 0.31 | 3120 ± 160 |
| R - 0120 | 0.0026 ± 0.0007 | 0.7 ± 0.1 | 1.78 ± 0.10 | 3.38 ± 0.02 | 16900 ± 300 | 1200 ± 30 | 0.26 ± 0.06 | 4970 ± 100 |
| R - 0130 | (0.257) | 0.838 ± 0.165 | 3.65 ± 0.76 | (1.5) | 17200 ± 1000 | 1100 ± 80 | 0.373 ± 0.081 | 4420 ± 330 |
| R - 0140 | 1.43 ± 0.46 | 66.3 ± 4.3 | 539 ± 15 | 4.46 ± 0.42 | 7900 ± 470 | 1180 ± 170 | 12.3 ± 1.5 | 4170 ± 130 |
| R - 0150 | | 158 ± 7 | 194 ± 20 | | 9960 ± 70 | 600 ± 15 | 9.9 ± 0.8 | 2430 ± 130 |
| R - 0151 | | 160 ± 8 | 184 ± 15 | | 9940 ± 20 | 601 ± 28 | 10.5 ± 1.7 | 2420 ± 60 |
| R - 0160 | 3.60 ± 0.40 | 5.22 ± 0.33 | 63.6 ± 5.3 | | 15900 ± 1000 | 1210 ± 130 | 1.32 ± 0.26 | 8000 ± 600 |
| R - 0170 | 0.40 ± 0.07 | 20.8 ± 1.2 | 712 ± 48 | | 10100 ± 1000 | 1100 ± 150 | 8.72 ± 0.53 | 7260 ± 730 |
| R - 0180 | 34.7 ± 5.5 | 2.34 ± 0.16 | 142 ± 10 | | | | 3.66 ± 0.34 | |
| R - 0190 | 0.37 ± 0.08 | 25.8 ± 1.1 | 1103 ± 47 | | | | 6.88 ± 0.56 | |
| R - 0210 | 0.071 ± 0.038 | 2.84 ± 0.45 | 71.2 ± 9.2 | 0.035 ± 0.012 | 15170 ± 370 | 960 ± 95 | 0.37 ± 0.09 | 2100 ± 80 |
| R - 0220 | 0.079 ± 0.012 | 15.9 ± 1.2 | 11.6 ± 0.9 | | 948 ± 72 | 89.5 ± 4.1 | 1.20 ± 0.13 | |
| R - 0230 | 0.53 ± 0.08 | 107 ± 8 | 77.8 ± 6.0 | | 6360 ± 480 | 601 ± 28 | 8.02 ± 0.86 | |
| R - 0240 | 2.4 ± 0.6 | 439 ± 22 | 186 ± 11 | | 10410 ± 400 | 2550 ± 250 | 23.4 ± 1.0 | 36700 ± 2000 |
| R - 0250 | 4.6 ± 0.3 | 12 ± 2 | 550 ± 20 | | 35600 ± 300 | | 165 ± 6 | |
| R - 0251 | | 12.2 ± 0.6 | | | 29030 ± 520 | (8900) | 75.9 ± 1.9 | 18180 ± 430 |
| R - 0260 | –0.3 | 5.6 ± 0.24 | (80) | (0.3) | 16100 ± 200 | 2710 ± 80 | 54 ± 3 | 24.4 ± 1.2 |
| R - 0270 | 0.8 ± 0.2 | 16.5 ± 1.0 | 90 ± 10 | 1.84 ± 0.03 | 18200 ± 600 | 5800 ± 300 | 23 ± 2 | 160 ± 20 |
| R - 0280 | (1) | 3.7 ± 0.4 | 218 ± 14 | (0.3) | 24300 ± 300 | 4320 ± 80 | 98 ± 3 | 24 ± 2 |

**Table 3** *Continued*

| *FCL-ID* | *Ni* | *P* | *Pb* | *Se* | *Zn* |
|---|---|---|---|---|---|
| R - 0010 | (0.16) | | | 0.4 ± 0.1 | 19.4 ± 1.0 |
| R - 0011 | | 1530 ± 80 | (<0.010) | 0.38 ± 0.04 | 19.4 ± 0.5 |
| R - 0020 | (0.18) | | | 1.1 ± 0.2 | 10.6 ± 1.0 |
| R - 0021 | | 1340 ± 60 | (<0.020) | 1.1 ± 0.2 | 11.6 ± 0.4 |
| R - 0030 | (0.41) | 3240 ± 40 | (0.05) | 0.245 ± 0.005 | 30.8 ± 1.1 |
| R - 0040 | | | | | 6.9 ± 0.15 |
| R - 0050 | 0.059 ± 0.022 | | 0.016 ± 0.003 | 0.082 ± 0.0072 | 41.7 ± 1.00 |
| R - 0060 | | | 0.089 ± 0.013 | 0.0394 ± 0.031 | 103.80 ± 3.005 |
| R - 0070 | 0.153 ± 0.009 | 2090 ± 140 | 0.018 ± 0.007 | 0.057 ± 0.00545 | 14.6 ± 0.71 |
| R - 0080 | 0.193 ± 0.043 | 2370 ± 100 | 0.026 ± 0.0028 | | 9.0 ± 0.32 |
| R - 0090 | 0.271 ± 0.038 | | 0.043 ± 0.008 | 0.181 ± 0.017 | 28.9 ± 1.3 |
| R - 0100 | | | 0.016 ± 0.003 | 0.082 ± 0.0077 | 41.68 ± 1.056 |
| R - 0110 | 0.644 ± 0.151 | 3320 ± 310 | | 0.242 ± 0.030 | 17.0 ± 0.6 |
| R - 0120 | | 10600 ± 200 | 0.019 ± 0.003 | 0.11 ± 0.01 | 46.1 ± 2.2 |
| R - 0130 | (0.93) | 9100 ± 1020 | (0.27) | 0.0339 ± 0.0072 | 38.9 ± 2.3 |
| R - 0140 | 2.25 ± 0.44 | 6230 ± 180 | 0.371 ± 0.014 | 2.21 ± 0.24 | 830 ± 57 |
| R - 0150 | | 11100 ± 400 | 0.135 ± 0.015 | 0.71 ± 0.07 | 123 ± 8 |
| R - 0151 | | 11000 ± 300 | 0.129 ± 0.004 | 0.73 ± 0.06 | |
| R - 0160 | 1.20 ± 0.30 | | 0.40 ± 0.12 | 1.62 ± 0.12 | 21.3 ± 1.0 |
| R - 0170 | 0.26 ± 0.06 | | 1.36 ± 0.29 | 7.34 ± 0.42 | 92.5 ± 2.3 |
| R - 0180 | 19.4 ± 3.1 | | 0.065 ± 0.007 | 1.40 ± 0.09 | 25.6 ± 2.3 |
| R - 0190 | 0.20 ± 0.02 | | 0.22 ± 0.02 | 6.06 ± 0.49 | 85.8 ± 2.5 |
| R - 0210 | 0.05 ± 0.04 | 8360 ± 450 | 0.38 ± 0.24 | 0.076 ± 0.010 | 142 ± 14 |
| R - 0220 | 0.200 ± 0.034 | | 0.010 ± 0.002 | 0.641 ± 0.054 | 12.4 ± 0.8 |
| R - 0230 | 1.34 ± 0.23 | | 0.069 ± 0.011 | 4.30 ± 0.36 | 82.9 ± 5.4 |
| R - 0240 | 2.3 ± 0.3 | 8790 ± 210 | 10.4 ± 2.0 | 6.88 ± 0.47 | 177 ± 10 |
| R - 0250 | (6) | 5500 ± 200 | 1.2 ± 0.2 | | 50 ± 2 |
| R - 0251 | 2.14 ± 0.10 | 5180 ± 110 | (200) | 0.117 ± 0.0009 | 82 ± 3 |
| R - 0260 | 0.91 ± 0.12 | | 0.470 ± 0.024 | 0.050 ± 0.009 | 12.5 ± 0.3 |
| R - 0270 | 0.6 ± 0.3 | 1300 ± 200 | 13.3 ± 2.4 | (0.025) | 29 ± 2 |
| R - 0280 | 0.69 ± 0.09 | 1370 ± 70 | 0.87 ± 0.03 | 0.120 ± 0.009 | 17.9 ± 0.4 |

## Instrumentation and applications

The literature is the most informative source of insights on how to select the most appropriate sample preparation methods and instrumentation to make accurate and precise trace element determinations. Analysts worldwide are continuously publishing their successes (and often their failures), allowing their colleagues to benefit from their experiences. Since the bulk of trace element determinations for foods and biological samples are done by a handful of techniques, four methods represent 98% of all determinations: (1) flame atomic absorption spectroscopy (FAAS); (2) graphite furnace atomic absorption spectroscopy (GFAAS); (3) inductively coupled plasma atomic emission spectroscopy (ICP-AES); and (4) ICP-mass spectrometry (ICP-MS). Comments on both the sample preparation method and the instrumentation are available to the reader. It is often a good idea to combine the best portions of several different methods when developing a method, and it is always good to avoid mistakes documented by others.

### FAAS and GFAAS

An estimated 80% of all currently available trace element food composition data are the result of FAAS analyses after either wet ashing or dry ashing sample pretreatment. FAAS is a simple, robust, and easy to implement tool for the analysis of digests, and calibration can typically be accomplished using aqueous standards. Detection limits are in the sub-ppm range, making this method suitable for a wide range of elements (including Ca, Cu, Fe, Mg, Mn, Zn) in various sample matrices. Please note that Na and K are most often determined using flame emission spectroscopy rather than absorption on an AAS system.

GFAAS provides sub-ppb detection capability with μL-sized sample injections into a platform-containing graphite tube which is resistively heated to high (e.g. 2700°C) temperatures for sample atomization. Over the last decade, so-called STPF (stabilized temperature platform furnace) conditions established by Slavin have been almost universally adopted. STPF conditions call for the use of (1) platform atomization, (2) matrix modification, (3) rapid heating (1500°C $s^{-1}$ or more), (4) pyrolytically coated tubes; (5) fast digital electronics, (6) integrated absorbance measurements (peak area), (7) argon (stop-flow during atomization), and (8) Zeeman (or Smith–Hieftje) background correction. Methods developed with these criteria in mind will facilitate straightforward quantitation using aqueous standards to make external calibration curves, in most cases minimizing matrix interference effects and reducing the need for using the method of additions. Analytical figures of importance include sensitivity/characteristic mass, detection limit, accuracy, and precision.

The field of trace element analyses in nutrition is one of the most interesting areas. Trace elements serve as structural components of enzymes, vitamins, hormones, and protein-containing tissues. The trace element selenium helps in the defense mechanism against diseases and environmental risk. Selenium is the most promising trace element potentially involved in immune response. Chromium is a cofactor for several enzyme systems, and is required for insulin receptor interaction. Consequently, research was conducted using FAAS to assess the Se and Cr content of eight food categories (cereals, beans, vegetables, greens, fruit, condiments and spices, dried fruits, and edible flowers). Samples were wet ashed using a combination of three acids and samples were analysed using FAAS at 196 nm (Se) and 425 nm (Cr) using an air–acetylene flame. A total of 190 samples were analysed and from the study it was concluded that dried fruits have the highest Cr content (15–43.5 μg per 100 g) and that beans have the highest Se content (48.7–02.5 μg per 100 g). FAAS also recently proved useful in the monitoring of Pb in dinnerware where excess levels can provide increased risk to fetuses, children, and adults. A total of 0.9% of imported ceramic dinnerware and 2.5% of domestic ceramic dinnerware, evaluated over the course of 2 months in 1992, had levels in excess of the 3 ppm allowed limit for plates, saucers, and flatware. This work involved the use of Association of Official Analytical Chemists Official Method 973.32. FAAS was used by scientists at Behrend College (Erie, PA) to evaluate stainless steel cookware as a significant source of Ni, Cr, and Fe for ingestion. Nickel ingestion is potentially dangerous since Ni is implicated with health problems related to allergic dermatitis. Conversely, Cr and Fe are essential nutrients and stainless steel (typically 18% Cr, 8% Ni and 70–73% Fe) may provide a significant source. Sample preparation involved the addition of 5% acetic acid (Fisher), both cold and boiled, in each vessel for 5 min. Next the acetic acid was analysed using FAAS and the manufacturers' standard conditions. Measurable levels of all three elements were determined with only Ni being high enough (0.0–0.1 mg Ni per day) to pose a health threat, leading to a recommendation that Ni-sensitive patients avoid stainless steel cookware and that the industry switch to a non-Ni formulation.

As evidenced by the previous examples, FAAS is a powerful technique but it may not always provide

the necessary sensitivity for the determination of trace elements present at extremely low concentrations. This became apparent when the Swedish National Food Administration set out to develop a method for the determination of Pb, Cd, Zn, Cu, Fe, Cr, and Ni in dry foodstuffs after dry ashing at 450°C. Realising that contamination with elements such as Pb, Cr and Ni is probable, care was taken to acid-wash all plasticware associated with the analyses. FAAS was selected for Zn (213.9 nm air–acetylene, oxidizing), Cu (324.7 nm air–acetylene, oxidizing), and Fe (Fe 248.3 nm nitrous oxide–acetylene, oxidizing). Owing to the lower concentrations, GFAAS was used for Pb, Cd, Cr and Ni. Conditions were optimized based on the use of the appropriate resonance line but no one set of instrumental conditions proved acceptable for all four elements. Lead was analysed using platform atomization while Cd was measured in an uncoated tube and Cr and Ni were measured in a standard pyrolitic tube. Unfortunately, not using STPF conditions limited the capabilities of the method and the method of additions was required by most collaborating laboratories to get reasonably accurate results.

GFAAS serves as an excellent method for the direct determination of Pb in degassed cola beverages. Recent reports indicate, however, that GFAAS analyses of cola diluted with a solution of lanthanum to reduce chlorides and other matrix interferences required the use of the method of additions to obtain accurate results. If the authors had done in situ oxygen ashing, and used Pd or Pd combined with magnesium nitrate as a matrix modifier, all matrix interference effects could have been removed. It is likely that the biggest problem with the background was owing to the sugar (carbon) which could have been removed during the thermal pretreatment step using oxygen ashing – permitting the use of aqueous calibration standards rather than requiring the method of additions. Sugars and syrups have recently been analysed directly after diluting ~1 g of sugar per 10 mL 5% nitric acid and using oxygen ashing in the thermal pretreatment step. Instrumental GFAAS detection limits (DL) were 10 pg, corresponding to a method DL of 0.9 ng $g^{-1}$. Researchers often use ingenious approaches to improve GFAAS performance. This was the case when developing a strategy for the determination of La in food and water samples. The problem is that La has a strong affinity for the graphite, leading to carbide formation and memory effects. Although the graphite has been improved to reduce this, lining the tube with tantalum or tungsten foil can eliminate physical contact with the graphite and lead to increased sensitivity. The tungsten-lined tube provided a detection limit of 7.8 ng and a characteristic mass of 8.1 ng for La. Precisions were better than 10% RSD and the average accuracy was 90 ± 10%. GFAAS has been used by researchers at the US Food and Drug Administration to successfully determine Se in infant and enteral formulas. The method utilizes sample digestion on a hotplate after addition of magnesium nitrate–nitric acid. Following heating, digests were evaporated to dryness and placed in a 500°C muffle for 30 min to complete ashing. All Se is converted into $Se^{4+}$ by dissolving the ash in HCl (5:1) and holding the solution at 60°C for 20 min. The $Se^{4+}$ was subsequently reduced to $Se^{0}$ with ascorbic acid and collected on a membrane filter which was digested in nitric acid using microwave digestion. Following digestion and dilution, Se was determined using GFAAS. The recovery range for Se was 85–127% and analysed reference materials fell within the certified range for Se. The work was performed on a commercial system equipped with only a deuterium background correction. Peak height and peak area measurements both provided accurate results when using nickel nitrate for matrix modification. Finally, GFAAS has proved useful for the determination of Cr and Mo in medical foods. Both wet ashing and dry ashing proved acceptable and the detection limits were 0.24 ng $mL^{-1}$ for Cr and 0.67 ng $mL^{-1}$ for Mo. Both elements could be determined directly off the shelf and neither required the use of a matrix modifier. Optimum ashing temperatures of 1650 and 1600°C were found for Mo and Cr, respectively. The ideal atomization temperatures were 2400 and 2650°C, respectively, and all standard reference materials analysed provided results within the certified concentration range.

## ICP-AES and ICP-MS

To assist consumers in maintaining healthy dietary practices and to assure product safety, the Nutrition Labelling Education Act (NLEA) was published in 1990 and the compliance deadline was May, 1994. More than 17 000 food companies were affected by NLEA and labels for more than 250 000 products required modification. As such, Perkin-Elmer researchers developed a multielement ICP-AES method using an axially viewed plasma which they coupled with microwave sample digestion. ICP-AES measurements allowed both trace and macro elements to be determined simultaneously owing to the wide linear dynamic range. In contrast to many official methods, which are analyte and/or matrix specific, and which often require multiple dilutions, this ICP-AES method gives multielement data (seven elements; nine wavelengths), providing an elemental

fingerprint which proves useful in adulteration and/or contamination studies. Measurements were made on a Perkin-Elmer Optima [echelle polychromater with a segmented array charge coupled detector (SCD) equipped with a standard torch, a Scott spray chamber, and a cross-flow nebulizer]. Although only a single 10 ppm multielement standard was successfully used for calibration, the use of three or more standards is recommended to cover such a large range. The use of duplicate wavelengths for Ca and Mg, combined with spike recovery checks, assisted with method validation. Literature reports highlighted the benefit of the Multiwave (Perkin-Elmer, Norwalk, CT) microwave system which provides pressure monitoring in each vessel, allowing different sample types to be digested together. Simultaneous multielement ICP-AES provides significant time-savings and the establishment of compromise conditions does not typically lead to significant deterioration of performance as compared with single element determinations.

In the 1990s, ICP-MS determinations have started to play a significant role in generation of food composition data. This technique provides multielement isotopic data with sub-ppb detection limits with the same convenience as flame AAS or ICP-AES and it is much faster than GFAAS. As interest in low level trace element concentrations continues to grow, more data are generated by ICP-MS and often alternative methods are employed to validate the ICP-MS method and resultant data. Such was the case in the recent study of Pb in green vegetables where GFAAS (dry ashing) and isotope dilution–ICP-MS (ID-ICP-MS) (wet ashing) were both used. The end result was comparable detectability and accuracy for in-house and commercial control materials. The limit of detection for Pb was 1–3 $\mu g\ kg^{-1}$ and precisions ranged from 6% RSD for kale containing 500 $\mu g\ kg^{-1}$ Pb to 20% RSD for cabbage containing 3 $\mu g\ kg^{-1}$ Pb. These methods were used for a four-year study of Pb levels in green vegetables with good success. Another advantage of ICP-MS is that it provides excellent detection capability for non-metals of nutritional interest, including P and I. Methods for I have typically been dependent on radio-iodide measurements and the specialized equipment required is not widely available. In contrast, ICP-MS allows the direct determination of I in solution using highly sensitive, specific, and interference-free detection of monoisotopic iodine ($^{127}I$). Samples were wet ashed in closed steel bombs using a mixture of nitric and perchloric acids which converted any volatile I species into non-volatile species. The concentrations measured were in the range 0.15–4.59 $\mu g\ g^{-1}$ (dry mass) and the detection limit was 30 $ng\ g^{-1}$ (dry mass) for a 0.5 g (dry) sample diluted to a final volume of 20 mL. The importance of ICP-MS in food analysis is growing but probably won't be fully realized until the list of the mandatory nutrients for labelling has been expanded.

## Acknowledgements

Mention of trademark or proprietary products does not constitute a guarantee of warranty of the product by the U.S. Department of Agriculture and does not imply their approval to the exclusion of other products that may also be suitable. The author would like to acknowledge that this work is based on literature reports provided by many colleagues and regrets that the specified presentation format precluded referencing the many interesting applications individually. The references listed are of general interest and many of them are overviews which include a large number of pertinent literature citations.

*See also:* **Atomic Absorption, Methods and Instrumentation; Environmental and Agricultural Applications of Atomic Spectroscopy; Food Science, Applications of Mass Spectrometry; Food Science, Applications of NMR Spectroscopy; Inductively Coupled Plasma Mass Spectrometry, Methods; Microwave and Radiowave Spectroscopy, Applications; Microwave Spectrometers.**

## Further reading

Brown (1990) *Present Knowledge in Nutrition*. Washington, DC: International Life Sciences Institute, Nutrition Foundation.

Chang S, Rayas-Duarte P, Holm E and McDonald C (1993) Food. *Analytical Chemistry* **65**: 334R–363R.

Greenfield H and Southgate DAT (1992) *Food Composition Data, Production, Management and Use*. Barking: Elsevier Science Publishers.

Howard AG and Statham PJ (1997) *Inorganic Trace Analysis*: *Philosophy and Practice*. New York: Wiley.

Kingston HM and Jassie LB (1988) *Introduction to Microwave Sample Preparation*. Washington, DC: American Chemical Society Press.

National Research Council (1988) *Designing Foods, Animal Product Options in the Marketplace*. Washington, DC: National Academy Press.

Pauwel J, Stoeppler M and Wolf WR (1998) Proceedings 7th International symposium on biological and environmental reference materials, Antwerp, Belgium, April 21–25, 1997. *Fresenius Journal of Analytical Chemistry* **360**: 275–504.

Stewart K and Whitaker JR (1984) *Modern Methods of Food Analysis*. Westport CT: AVI Publishing Company.
Taylor A, Branch S, Crews H, Halls D, Owen L and White M (1997) Atomic spectrometry update – clinical and biological materials, food and beverages. *Journal of Analytical Atomic Spectrometry* **12**: 119R–221R.
Wolf W (1984) *Biological Reference Materials, Availability, Uses and Need for Validation of Nutrient Measurement*. New York: Wiley.

# Food Science, Applications of Mass Spectrometry

**John PG Wilkins**, Unilever Research, Sharnbrook, Bedfordshire, UK

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Applications**

The exquisite analytical sensitivity and specificity of MS have found numerous and diverse applications in the field of food science.

Trace analysis for undesirable chemical contaminants in foodstuffs is conducted in many laboratories around the world. This may be for quality control, regulatory or surveillance purposes. MS is frequently the analytical method of choice because of its ability to produce unequivocal data.

Organic compounds sought include naturally derived materials, such as mycotoxins and off-flavours (produced by rancidification or spoilage), and man-made/industrial chemicals, e.g. pesticides, veterinary drugs, environmental contaminants (such as polychlorodibenzo-*p*-dioxins, polychlorinated biphenyls, polynuclear aromatic hydrocarbons, etc.) and food tainting compounds (e.g. 2,4,6-trichloroanisole, the compound responsible for musty cork taint in wine, arising from the inappropriate use of wood preservatives). GC-MS and HPLC-API-MS are widely used for these types of analyses. Desirable food components present at trace levels, such as nutrients, are also determined using these techniques.

Trace levels of inorganic chemical species, e.g. lead, arsenic, cadmium, are also monitored in foodstuffs, often using ICP-MS. The advantage of MS over AAS is that several elements may be measured simultaneously and the concentrations of individual isotopes may be measured, facilitating metabolism/nutrient studies with stable isotope materials. Precise determination of isotope ratios (e.g. C, N and O) by IRMS is also important in agricultural and food authenticity studies. Accelerator mass spectrometry is used in tracer studies, for the determination of extremely low levels of carbon-14 (and other) isotopes.

The table below summarizes the ranges of analytical sensitivities required for different classes of food analysis.

Much work has been done on the characterization of food flavours and aromas by GC-MS. A significant recent advance has been the use of APCI and drift tube MS techniques for sensitive, on-line (‘real time’) monitoring of trace volatiles, e.g. in human breath during flavour release studies.

Characterization of unfractionated foodstuffs and related materials for screening or taxonomic purposes may be performed by pyrolysis MS and direct headspace MS. These techniques generate rather simple mass spectra that may be classified using pattern recognition/chemometric software.

MS is also used for the analysis of the more abundant (non-trace) components of food, e.g. oils and fats (triglycerides), proteins and carbohydrates. Analysis of these materials is often challenging, as they may comprise complex mixtures of isomeric compounds. Modern MS techniques (MALDI TOF and electrospray ionization) have become extremely important in protein and peptide studies.

**Table 1** Typical analytical sensitivities required for various classes of food contaminant/component

| *Class* | *Typical concentraion sought (by weight)* |
|---|---|
| Flavours | % - ppb |
| Food taints | % - ppq |
| Pesticide residues | ppm - ppb |
| Trace elements | ppm - ppb |
| Mycotoxins | ppb |
| Dioxins | ppt - ppq |

*See also:* **Biochemical Applications of Mass Spectrometry; Food and Dairy Products, Applications of Atomic Spectroscopy; Food Science, Applications of NMR Spectroscopy; Inorganic Chemistry, Applications of Mass Spectrometry; Peptides and Proteins Studied Using Mass Spectrometry; Time of Flight Mass Spectrometers.**

## Further reading

Belton PS, Mellon FA and Wilson RH (1993) Spectroscopy in Food Science, *Spectroscopy in Europe* **5**: 8–14.

Crews H (1993) A decade of ICP MS analysis. *International Laboratory* **23**: 38–42.

Gilbert J (ed.) (1987) *Applications of Mass Spectrometry in Food Science*. Amsterdam: Elsevier Applied Science.

Horman I (1979) Mass spectrometry in food science. *Mass Spectrometry* **5**: 211–233.

Self R, Mellon FA, McGaw BA, Calder AG, Lobley GE and Milne E (1996) The application of mass spectrometry to food and nutrition research. *NATO ASI Series, Series C* **475** (*Mass Spectrometry in Biomolecular Sciences*) 483–515.

# Food Science, Applications of NMR Spectroscopy

**Brian Hills**, Institute of Food Research, Norwich, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

The applications of NMR in food science have been the subject of several international conferences, reviews and books which are cited in the Further reading section. This intense interest in the NMR spectroscopy of food is driven not only by the commercial importance of food, but also by the intellectual challenge of unravelling the physicochemical properties of this exceedingly diverse and complex group of materials. In order to focus on food aspects, it will be assumed that the reader is familiar with the principles of NMR and magnetic resonance imaging (MRI), which are lucidly explained in other articles of this encyclopedia.

## Spatially resolved NMR applications

MRI is beginning to have a major impact on our understanding and control of the growth of food crops, on our assessment of fruit and vegetable quality and in developing the best methods of fruit and vegetable preservation. Moreover, by providing noninvasive, real-time images of moisture and lipid distributions, as well as spatial maps of temperature and food quality factors, MRI has the potential for making a major impact on food processing science. We therefore begin with a brief review of whole-plant functional imaging, then progress on to MRI applications in food processing and storage.

### Functional imaging of whole plants

The growth of agricultural crops is affected by many factors, including drought (osmotic stress), luminescence, disease and soil pollutants, such as heavy metals. By permitting noninvasive monitoring of a whole plant under realistic environmental conditions, MRI has become a powerful tool in the plant physiologists' armoury. The development of small seedlings can be observed directly inside adapted NMR tubes. Larger plants can be grown in controlled-environment boxes, which bathe the root system in nutrients, and control humidity, temperature and luminescence. The root system, stem or leaf areas can then be imaged noninvasively. Root imaging is now an established technique, though it is not applicable to all soil types because paramagnetic impurities in the soil can severely shorten the water relaxation time. Nevertheless, three-dimensional (3D) $T_1$-weighted imaging has been used to distinguish roots and soil in developing pine seedlings, to follow water depletion zones around the tap root, lateral roots and fine roots, and to follow the effects of symbiotic relationships, such as infection with mycorrhizal fungus, on water uptake. The increase in root network volume and surface area can also be measured from the 3D images.

A combination of flow imaging and chemical shift imaging has been used to monitor the effect of

environmental stresses on the flux of nutrients through the plant stem. For example, chemical shift imaging has been used to map the spatial distribution of sucrose, glucose, glutamate, lysine, arginine and valine in castor bean seedlings. If, at the same time, the velocity distribution through the vascular bundles in the stem is observed with phase-encoding MRI, the product of the concentration and flow rate gives the nutrient fluxes. The effects of varying osmotic stress and luminescence on the nutrient fluxes can then be monitored nondestructively. **Table 1** lists examples of this type of application.

**Table 1** Applications of functional imaging to the development of fruit and vegetables[a]

| *Plant* | *MRI technique* | *Phenomenon* |
|---|---|---|
| Castor bean seedlings | Flow imaging, chemical shift imaging | Effect of environmental stresses on nutrient fluxes |
| Maize and millet | Proton density mapping, $T_2$ mapping | Osmotic stress |
| Wheat grains | Water proton density, velocity and diffusion maps | Relationship between flow and grain development |
| Red raspberry | Gradient-echo signal intensity | Fungal infection |
| Strawberry | Proton density $T_1, T_2$ | Fungal infection |

### Imaging quality factors in intact fruit and vegetables

Understanding the development of the whole plant under realistic environmental conditions is only one aspect of food production. The quality of unprocessed, intact fruit, grains or vegetables as they are stored or transported is just as important. Here again MRI provides a useful noninvasive monitor of the effect of storage conditions on the ripening of fruit and vegetables, on the progress of detrimental changes such as disease and bruising and on the effects of preservation processes such as chilling and surface heat sterilization.

The ripening changes in tomato are of commercial importance because they are picked when they are green and transported, internationally, in boxes gassed with ethylene. Unfortunately, on arrival the tomatoes display a wide range of ripeness from green to the red-ripe fruit and require costly (and damaging) hand-sorting. If, when first picked and boxed, the degree of ripeness could be determined noninvasively by MRI, then ripening could be made more uniform and the second sorting stage rendered unnecessary. Motivated by these considerations, Saltveit and co-workers have shown that ripening in the early green stages is associated with increased water content, and hence increased MRI signal intensity, as the locular tissue 'liquefies'. Later stages of ripening are associated with increased image graininess as small air pockets develop in the pericarp tissue.

**Table 2** Representative MRI studies of quality factors in fruit and vegetables[a]

| *Plant* | *Quality factor/process* | *NMR parameter* |
|---|---|---|
| Red raspberry | Ripening | Spin-echo signal intensity |
| Strawberry | Ripening, bruising | Spin density, $T_2$ |
| Grape berries | Ripening, drying | Chemical shift |
| Red raspberry | Fungal infection | Gradient-echo signal intensity |
| Barley seeds | Ripening | Diffusion |
| Wheat grains | Ripening | Diffusion and flow |
| Tomato | Ripening | Spin density, $T_1$ |
| Mango | Heat sterilization | Relaxation-weighted signal intensity |
| Blueberry | Freezing–thawing | $T_1$ discrimination of sugar and water |
| Cherry | Ripening | $T_1$, $D$ and volume selective spectroscopy |
| Papaya | Heat injury | Spin density, $T_1, T_2$ |

[a] For references see Hills (1998) in Further reading section.

Strawberries have also been the subject of several MRI investigations. Bruising is associated with a lengthening of the apparent transverse relaxation time ($T_2$*) and an increase in apparent water proton spin density, consistent with the reduction of local magnetic susceptibility inhomogeneities as a result of the filling of the intercellular gas spaces with intracellular fluid released through membrane and cell wall damage. A similar increase in $T_2$* was found on infection with the fungus *Botrytis cinerea* and permitted the diseased and healthy tissue to be $T_2$*-contrasted and the rate of progress of the disease quantified as the rate of increase in the volume of infected tissue. Ripening of strawberries is associated with an increase in $T_2$ from about 16–25 ms in the immature green fruit to 25–100 ms in the red, ripe fruit. The ability to discriminate diseased, bruised and unripe regions from healthy ripe tissue is essential for the development of on-line MRI quality monitoring. **Table 2** lists a number of other MRI studies on the quality of fruit and vegetables, references to which can be found in the Further reading section.

### MRI studies of food processing

Many foods are processed before being packaged, distributed and consumed. The major processing

operations include drying, freezing, baking and extrusion, but novel processing methods, such as high-pressure treatment, are under development and can be used either alone or in combination with more conventional processing operations. By providing spatial maps of moisture, temperature and quality factors during processing, in either the off-line or the on-line mode, MRI has the potential of revolutionizing the food processing industry. The scope for this type of study is very wide and will be illustrated by a few key operations, beginning with drying.

**Drying** The output of MRI drying studies is a set of moisture profiles at known drying times. The usual strategy is then to fit the profiles with appropriate mass transport models, which are used subsequently to optimize the drying conditions for maximum product quality with minimum energy expenditure. In practice, complexities, such as shrinkage, case hardening and spatially varying diffusivities have meant that the MRI data do not always conform to simple drying models, and this has highlighted the need for more refined theoretical models. **Table 3** lists a number of MRI drying studies on foods such as pasta, potato and apple.

**Rehydration** The quality of extruded cereal products, such as pasta, depends on extrusion temperature and the variety of wheat used in the extrusion process. The effect of varying the ratio of hard and soft wheat on the rehydration kinetics of extruded spaghetti has been investigated by MRI. By measuring the moisture profiles at various times as the spaghetti is cooked in hot water it was shown that increasing the amount of soft wheat shifts the rehydration kinetics from a case II situation, where there is a fairly sharp water front ingressing into the pasta, towards rehydration by conventional Fickian diffusion. Other MRI studies have been reported for the rehydration of starch films and air-dried apple tissue. Besides simple Fickian diffusion, the Thomas Windle model provides a useful theoretical tool for analysing the rehydration kinetics, especially near the case II regime. **Table 3** lists a number of MRI rehydration studies.

**Table 3** Representative MRI processing studies[a]

| *Food* | *Process* | *MRI technique* |
|---|---|---|
| Model food gel | Drying | 2D spin density |
| Extruded pasta | Drying | Radial imaging |
| Apple, potato | Drying | 1D profiling |
| Pasta | Rehydration | $T_2$-weighted radial imaging |
| Starch gels | Rehydration | STRAFI[b] |
| Wheat grains | Boiling/steaming | 2D spin density |
| Dry beans | Rehydration | Spin warp imaging |
| Potato | Rehydration | 1D profiling |
| Pasta, potato | Freezing | Radial imaging |
| Meat | Freezing–thawing | Magnetization transfer imaging |
| Potato | Freeze-drying | Spin warp imaging |
| Cookies | Baking | Spin warp imaging |
| Potato | Frying | Chemical shift imaging |
| Model fluid | Extrusion | Flow imaging |
| Soya bean | Rehydration | Chemical shift imaging |

[a] For references see Hills (1998) in Further reading section.
[b] STAFI, stray field imaging.

**Freezing** Ice has a proton transverse relaxation time of just a few microseconds, so it does not contribute signal in conventional liquid-state MRI. Freezing is therefore associated with a progressive loss of signal and this provides a powerful tool for investigating freezing kinetics. The freezing of raw potatoz has been imaged at various resolutions. At low spatial resolution an ice front is observed to ingress into the food. But with smaller samples and higher spatial resolutions (~ 40 μm) there appears to be a gradation in signal intensity throughout the sample, indicating that different regions are associated with different amounts of unfrozen water. This observation is, perhaps, a consequence of the different freezing behaviour of starch granules, cell walls, vacuolar and cytoplasmic regions in the potato cell, though the possibility of ice 'channelling' in the extracellular spaces cannot be discounted. Like drying and rehydration, finding suitable theoretical models to quantify this type of behaviour presents a considerable theoretical challenge. **Table 3** includes a number of MRI food freezing studies. The potential of using MRI as an on-line monitor of food freezing has yet to be exploited but clearly has considerable potential.

**Freeze-drying** Despite the earlier statement that ice does not contribute signal in conventional MRI, it has, nevertheless, been possible to image the shrinkage of the frozen core in potato during freeze-drying. In this example there is sufficient unfrozen water in the 'frozen' potato to permit the core to be imaged. The absence of liquid water, which would saturate the signal, is another important factor, permitting high receiver gains during image acquisition. The time course of the sublimation of the frozen potato core was shown to follow a simplified freeze-drying model, but no other MRI freeze-drying studies have yet appeared.

**Temperature mapping and thermal transport** Most food processing operations involve not only mass

**Table 4** MRI temperature mapping in food systems[a]

| *Food* | *Temperature mapping technique* |
|---|---|
| Food gel | $T_1$ inversion recovery, snapshot FLASH [b] |
| Agar gel | Diffusion-weighting |
| Potato | Diffusion-weighting |
| Carrot | $T_1$ weighting |
| Potato pieces in 6% starch | Aseptic processing by chemical shift weighting |

[a] For references see Hills (1998) in Further reading section.
[b] FLASH, fast low-angle shot.

transport but also simultaneous heat transport. By providing real-time temperature maps, MRI can also be used to monitor heat transport. All that is required is an NMR parameter, such as a relaxation time, chemical shift or water diffusivity that depends uniquely on temperature over the desired range, so that, after calibration, a map of the parameter, or the parameter-weighted signal intensity, can be converted into a temperature map. **Table 4** lists a number of examples of MRI temperature mapping. Most recently the technique has been applied to the problem of aseptic food processing in which a fluid suspension of food particles is thermally processed in a continuous process. The problem here is knowing whether the central regions of the particles have attained sufficiently high temperature to ensure complete cooking and sterilization.

**Imaging quality factors in processed food** Processing involves not only mass and heat transport but also changes in quality factors associated with microstructure, phase changes, gelatinization, aggregation and chemical or enzymatic reactions, Some of these changing quality factors are not amenable to NMR measurements, colour and aroma being obvious examples. However, in many cases NMR parameters are sensitive to quality factors so that MRI can be used to image changes in food quality during processing and storage. Chocolate confectionery is a typical example of this type of study. The long-term storage stability of fat-containing foods, such as chocolate, can be affected by fat crystal growth and polymorphic transitions in the solid fat phase. The MRI signal from cocoa butter arises from regions of high lipid chain mobility and so depends both on the solid–liquid ratio and, because each polymorph is associated with different chain mobility, on the polymorphic state of the solid lipid components, For this reason transverse-relaxation time-weighted MRI has been used to distinguish regions of chocolate that have melted and resolidified from those that have remained solid throughout. The migration of liquid lipid components between two phases in multilayer confectionery products greatly affects the long-term quality of the product. For example, the transport of liquid triglycerides from a praline filling into the surrounding solid chocolate layer in a filled chocolate sweet can alter thermal stability and accelerate blooming on the chocolate surface. Accordingly, MRI has been used to image the transport of liquid lipids into chocolate in a model lamellar, 2-layer chocolate product, consisting of a layer of hazelnut oil and sugar layered over with dark chocolate. The lamellar structure permitted a 100 μm spatial resolution to be achieved and the time course of the lipid migration was followed over an 84 day period for two samples stored at 19 and 28°C. Intensity changes in the profiles were converted to liquid lipid concentrations by calibrating with homogeneous samples of known liquid-fat content. At 19°C, the liquid lipids were observed to accumulate at the interface between the hazelnut oil and the chocolate layers. In contrast, at 28°C, the lipid accumulated at the chocolate surface. Other examples where MRI has been used to image lipid transport and its effect on quality include the migration of fat into bread from peanut butter and the separation of cream from milk over a 9 h period.

The spatial distribution of fat in meat affects quality factors such as texture and has been imaged with chemical shift and $T_1$-weighted imaging. A number of feasibility studies have shown that MRI can detect bruises in apples, pears, onions and peaches, and detect the deterioration of apples in storage, watercore distribution in apples and core breakdown in pears. The quality of cereal products is also amenable to MRI analysis. The staling of baked biscuit ‘sweetrolls’ over several days storage has been imaged and shown to be associated with a migration of moisture from the inside to the outside of the sweetroll. **Table 5** lists a number of examples where MRI has been used to image quality factors.

## Flow imaging and food rheology

Knowledge of the rheological properties of food pastes, slurries and sauces, such as ketchup, mayonnaise and salad creams, is important both for quality assurance and for optimizing industrial flow and mixing processes. Unfortunately, many food slurries and pastes are opaque and do not lend themselves to flow studies with conventional techniques such as laser Doppler anemometry. Moreover, conventional rheological measurements are model-dependent in that it is necessary to fit the data by assuming a function relationship between the stress and strain (or strain rate) and to assume a set of boundary conditions (such as slip or stick) at the fluid–container

**Table 5** NMR studies of food quality factors[a]

| *Food* | *Quality factor* | *NMR technique* |
|---|---|---|
| Chocolate | Lipid migration | MRI |
| Model food emulsion | Crystallization kinetics | Localized spectroscopy |
| Creaming | Phase separation | $T_1$-weighted MRI |
| Egg-white and beer foam | Foam stability | MRI profiling |
| Meat | Curing/brining | $^{23}$Na NMR |
| Various fruits | Bruising | $T_2$- and $T_2$*-weighted MRI |
| Cookies (sweet rolls) | Retrogradation | $M_0$ maps |

[a] For references see Hills (1998) in Further reading section.

interface. This is satisfactory when it is known that the fluid has a simple rheological behaviour, such as that of a Newtonian or power-law fluid. However, many liquid foods do not conform to simple rheological models so their stress–strain relationship, usually called the flow curve, must be deduced in a model-independent way. MRI can do this by imaging the velocity profile at every point in a rheometer, including the boundary layer (at least within one voxel dimension). The rheometer can be of any of the standard types, including a cylindrical tube and rotational viscometers (or couette cell). **Table 6** lists a number of representative MRI rheological studies.

**Shear-induced nonrheological changes** Shear-induced aggregation in concentrated food emulsions such as dairy creams and shear-induced particle migration are just two examples of microstructural or compositional changes caused by shear. These changes can be studied with MRI methods by combining flow imaging with MRI spatial maps of NMR parameters, such as relaxation times or spin densities, which are sensitive to the microstructural changes. **Table 6** includes a number of examples of this approach.

**Flow and quality assurance** On-line quality control of food slurries and pastes often requires rapid, empirical, rheological tests based on flow times or flow lengths rather than quantitative measurements of well-defined rheological parameters such as shear viscosities. The Bostwick consistometer (which is used to determine the quality of puréed fruit and vegetable products), the Brabender Visco-amylograph (which is used to test the consistency of starches) and the Torque plasticorder (used for dough consistency), are examples of such empirical rheological testing. MRI studies of the flow within these various test rigs, such as the Bostwick consistometer, have helped provide a rheological rationale for these empirical measurements.

**Flow in extruders** Pioneering MRI studies on flow in an extruder have been reported by McCarthy and co-workers (see Further reading section) who designed a special nonmagnetic single-screw extruder for operation inside the magnet of an NMR spectrometer, and used carboxymethyl cellulose solution as a model extrudate. The flow patterns were compared with theoretical models of extruder flow for an incompressible Newtonian fluid. This type of study has considerable research potential because extrusion is used widely throughout the food industry for the manufacture of pasta, snacks, confectionery, pet foods and animal feeds. Despite this, extrusion is, perhaps, the least understood of all food processing operations, involving complex series of physical and chemical changes such as mixing, gelatinization, denaturation, evaporation, flavour production (and loss) and rheological changes. To date only the flow aspects of extrusion have been studied by MRI, but, by exploiting relaxation and diffusion weighted imaging techniques, the potential exists for exploring other associated physical and chemical changes.

## NMR diffusion studies of food microstructure

The use of pulsed gradients and fringe field gradients in NMR diffusion measurements has been widely documented in the research literature and is reviewed in other articles of this encyclopedia. Provided certain conditions are fulfilled, such as factorization of the effects of relaxation and diffusion and translational invariance, the pulsed-gradient techniques permit measurement of the average displacement

**Table 6** Flow imaging studies of food–related materials[a]

| *Fluid* | *Apparatus* | *Phenomenon* |
|---|---|---|
| Nutrient fluid | Hollow fibre bioreactor | Starling flow |
| 1.5% Carboxymethyl cellulose solution | Tube flow | Rheology |
| Tomato juice | Tube flow | Rheology |
| Fibrous suspensions | Tube flow | Rheology |
| Carboxymethyl cellulose solution | Extruder | Velocity field in an extruder |
| Particle suspensions | Tube/couette flow | Shear-induced migration |
| Xanthan gum | Tube flow | Shear thinning and apparent wall slip |
| Egg-white, cornflour/ water, tomato sauce | Couette, flow/ cone and plate flow | Shear-induced phase transitions |

[a] For references see Hills (1998) in Further reading section.

propagator, $P_{av}(R,\Delta)$, where $R$ is the distance diffused by a spin in the diffusion time, $\Delta$. Various theoretical models have been developed relating $P_{av}(R,\Delta)$ to microstructure in simple geometries, so that measurements of the average propagator can be used to gain information about food microstructure. In favourable circumstances, microstructural information on the submicrometre distance scale can be accessed by this approach, which greatly exceeds the spatial resolution in NMR microscopy. In the following we consider the application of diffusometry to various food systems, including gels, cellular tissue and emulsions.

### Diffusion studies of gels and starch-based systems

The microstructure of starch and gellan gum gels has been investigated by the stimulated-echo pulsed-gradient method as a function of increasing diffusion times, $\Delta$ (up to 1010 ms), and at various gradient amplitudes, $G$, at fixed gradient pulse duration $\delta$ and pulse separation, $t_1$. An effective water diffusion coefficient, $D(\Delta)$ was extracted by assuming the echo attenuation for unrestricted diffusion,

$$\ln[S(q,\Delta)/S(q=0)] = -D(\Delta)(4\pi q)^2(\Delta - \delta/3)$$

The gel microstructure caused the coefficient $D(\Delta)$ to decrease with increasing $\Delta$, and this decrease was fitted with the model of von Meerwall and Ferguson. This model represents the gel micropores as a series of parallel compartments of width, $a$, separated by barriers of permeability, $P$. Within each compartment the intrinsic water diffusivity was assigned the value, $D_0$. Pore sizes, $a$, of the order of 10–15 μm for starch gels and 5–15 μm for gellan gums were deduced and these varied slightly with gel concentration, and, for gellan gums, with ionic strength. These pore spacings appear to be rather large for these concentrated gel systems and, in the case of starch gels, may reflect the presence of residual granule structure resulting from incomplete gelatinization. The effects of retrogradation on starch gel microstructure were also followed in these studies, and were associated with changes in pore size and water diffusivity. Fringe field diffusion methods have been used to monitor water diffusion in gelatine gels of increasing concentration and in protein aerogels obtained by freeze-drying. The diffusion behaviour in the aerogel system was shown to be anomalous, and akin to diffusion in a fractal space. Echo attenuation of the pulsed-gradient stimulated echo of water in packed beds of potato starch granules appears to conform to diffusion in a lower dimensional space, lying between one and two dimensions, depending on the water content. This could be a consequence of the crystalline A- and B-type starch regions in the granule, channelling the water translational motion, but further work is needed to test this hypothesis.

### Diffusion studies of cellular plant tissue

The pulsed-gradient stimulated-echo sequence has been used to investigate the effects of electrical and conventional heating on intercell permeabilities in carrot tissue. The echo attenuation was interpreted with a lamellar cell model in which carrot cells were represented as lamellar regions of width, $W$, containing water of diffusivity, $D$, and separated by a barrier of effective permeability, $P$. Presumably this barrier arises from the combined effect of diffusion through the tonoplast, cytoplasm, plasmalemma and cell wall subcellular compartments. It was found that cooking increases the apparent diffusion coefficient and barrier permeability, which is consistent with the breakdown of cell structure, but there were also small, but statistically significant, differences between electrical and conventional cooking.

### Diffusion studies of food emulsions

The determination of droplet-size distributions in water-in-oil or oil-in-water emulsions by pulsed-gradient diffusion measurements is now a routine industrial method for monitoring the quality of food emulsions such as mayonnaise and salad creams. The method was first described in a classic paper by Packer and Rees (see Further reading section) and subsequently developed into a rapid analytical method. The theoretical model assumes spherical droplets and a log-normal size distribution, but other distributions have been considered. The MRI generalization of the method to spatially dependent droplet-size distributions in emulsions undergoing coalescence and/or phase separation has yet to be implemented, although a spatially dependent study of emulsion crystallization has been reported.

Cheese is, in essence, a distribution of fat droplets in a protein–water matrix. The fat droplet size distribution has been studied by the NMR diffusion method and showed a very broad Gaussian size distribution with significant differences between Swiss and cheddar cheeses. In this sample the water signal could be removed by using a long echo time, because the water transverse relaxation time was much shorter than that of the fat. The water diffusion was also shown to be consistent with 1D diffusion along the protein chains in the cross-linked protein matrix of the cheese. In this case the water signal was separated by Fourier transforming the echo obtained

at shorter echo times. This resulted in separate peaks for the fat and water.

## NMR relaxation studies of food

### Relaxation studies of food microstructure

NMR relaxation time measurements provide a unique window to view both food microstructure and the dynamic behaviour of molecules comprising the food matrix. The microstructural aspect arises because different aqueous microscopic compartments are usually associated with different intrinsic water-proton relaxation times. Provided the exchange of water magnetization between compartments is slow on the NMR measurement timescale, the distribution of water-proton relaxation times therefore reflects the number and relative volumes of the microscopic compartments. The proviso that the exchange is slow is important, because if the exchange is fast, then only a single average relaxation time will be observed and all compartmental information is lost. Like diffusion, theoretical models are required to relate the observed relaxation-time distribution to microstructure, and these models are usually based on numerical solution of the Bloch–Torrey equations. The Monte Carlo method is perhaps one of the most powerful solution tools for doing this. A cube is divided into $N^3$ volume elements and the morphology of the system is defined by assigning compartments (or regions) within the cube characterized by intrinsic relaxation times and diffusion coefficients. Magnetization density vectors are then randomly distributed throughout the cube, which then diffuse between voxels and relax with a rate corresponding to their location. In this way echo-decay amplitudes for the CPMG (Carr–Purcell pulse sequence, Meiboom–Gill modification) and pulsed-gradient diffusion sequences can be calculated and analysed with one of the standard deconvolution software packages.

Numerical solutions of the Bloch–Torrey equations have been used to interpret relaxation data for a variety of microscopically heterogeneous model systems, including randomly packed, water-saturated beds of Sephadex, silica and glass microspheres. Their application to plant cell tissue is perhaps more directly relevant to food. The three major compartments in a plant cell are the vacuole, cytoplasm and cell-wall region, though starch granules comprise an additional compartment in some foods, such as potato. The Bloch–Torrey equations describing relaxation and diffusive exchange between these intracellular compartments have been solved for an idealized three-compartment spherical cell morphology and used to predict the dependence of the relaxation-time distribution and pulsed gradient spin–echo (PGSE) diffusion behaviour of apple tissue, both fresh and during air drying. The latter showed that water is removed mainly from the vacuolar compartment during air-drying, and that this is associated with volume shrinkage of the tissue. The subcellular changes induced by the freezing and drying of potato and carrot have also been investigated in this way.

### Relaxation studies of aqueous solutions, gels and glasses

There are at least two proton pools in aqueous solutions of dissolved sugars, proteins and polysaccharides, namely the nonexchanging CH proton pool and the exchangeable pool of protons on water and biopolymer/solute hydroxyl and amino groups. Fast proton exchange between the water and exchangeable biopolymer protons means that any process, such as denaturation or gelation, that changes the rigidity of the biopolymer chains will result in an altered transverse relaxation time of both the exchangeable ‘water’ and the nonexchanging proton pools. The latter can be observed by replacing water with $D_2O$, or, at sufficiently high concentrations, as a faster relaxing component in the free induction decay (FID). The transverse relaxation rate of the exchangeable ‘water’ proton pool can be monitored as a slowly relaxing component in the FID or, more usually, with the CPMG pulse sequence, and can be used to follow the kinetics of gelation and denaturation. This approach has been used to follow the thermal denaturation of whey proteins and the gelation of pectins and starch.

Longitudinal and rotating-frame relaxation in these systems is complicated not only by proton exchange but also by the possibility of secular dipolar cross relaxation between all proton pools. This makes theoretical modelling difficult, although it does not prevent longitudinal relaxation being used empirically to monitor food processing changes.

At lower water contents, below ~ 50% by weight water, the signal from the nonexchanging solute proton pool appears clearly as a faster relaxing component in the FID and can be characterized with an effective $T_2$ by analysing the decay as either Gaussian or exponential. For example, in concentrated gelatine gels, the fast relaxing component has a $T_2$ of the order of a few tens of microseconds, and this increases with increasing water content as the gelatine chains are plasticized. Plasticization, which permits high-energy ‘strained’ interchain conformations to be relaxed, can result in increased local order and changes in texture. For this reason it has been investigated in a number of cereal proteins that are important in controlling the functional behaviour of dough and bread, such as

glutenin in wheat and chordein in barley. Retrogradation in starch-based foods is one of the most important examples of plasticization, and permits the partial recrystallization of amylopectin chains into A or B crystallites, and results in the staling of bread and cake products. The retrogradation kinetics can be followed from the concomitant decrease in the water-proton relaxation time, and the lowering of the glass transition temperature by plasticization can be modelled with the polymer science approach.

## NMR and food compositional analysis

### High-resolution NMR spectral studies

The application of high-resolution NMR methods to molecular structure elucidation of purified proteins, polysaccharides and lipids is too extensive to be treated in depth in this chapter. The use of $^1$H high-resolution solution-state NMR spectroscopy, in combination with enzymatic debranching and chromatographic purification, for elucidating the molecular structure of complex food polysaccharides, such as amylopectins, carrageenans, pectins and alginates, has been reviewed (see Gil and co-workers in the Further reading section). High-resolution liquid-state $^1$H and $^{13}$C NMR has also been widely used in the compositional analysis of lipids in vegetable and fish oils and **Table 7** indicates the extent of this type of study. Cross-polarization magic angle spinning (CP MAS) spectroscopy and other high-resolution solid-state NMR methods have been used to study structure–function relationships in the solid and gelled states of starches and other polysaccharides. Multinuclear NMR spectroscopy has been especially valuable for understanding cation–polysaccharide interactions in ionic food polysaccharides, and some examples are listed in **Table 8**.

The low-speed MAS technique is worthy of special mention because it overcomes spectral-line broadening artefacts from local field gradients created by differences in magnetic susceptibility across phase boundaries in microscopically compartmentalized foods. This technique has been used to obtain well-resolved spectra of sugars and metabolites in intact fruits, and to resolve water and oil peaks in multiple water-in-oil-in-water emulsions, found, for example, in low fat spreads. The detection of food adulteration is of increasing commercial importance and several high-resolution NMR methods, such as site-specific natural isotope fractionation (SNIF) NMR are being used routinely for food authentication, and have been extensively reviewed (see the Further reading section).

**Table 7** Recent NMR studies of food lipids[a]

| *Lipid* | *Nucleus* | *Analytical purpose* |
|---|---|---|
| Vegetable epoxidized oils | $^{13}$C | Epoxy acid content |
| Castor oil | $^{13}$C | Quantification of castor oil in edible oils |
| Minuta oil | $^{13}$C, $^1$H | Chemical changes during plant flowering and fruiting |
| Olive oil | $^{13}$C | Saturated fatty acid content. Grading of olive oils |
| Pomace oils | $^{13}$C | Composition of unsaponifiable matter in pomace oil |
| Fish oils | $^{13}$C, $^1$H | Quantification of oils in tuna and Atlantic salmon |
| Oxidative deterioration of oils | $^1$H | Ratio of olefinic to aliphatic protons as an index of deterioration |
| Canula (rapeseed) oils | $^{13}$C, $^1$H | Colour pigment formation during bleaching and deodorization |
| Liguloxide | $^{13}$C, $^1$H | Characterization of odour in tomato ketchup |
| Butter, baking and spreading fats | $^{13}$C | Compositional analysis |
| Milkfat | $^{13}$C, $^{31}$P | Solid fat content of milkfat fractions and blends |
| Tripalmitin | $^2$H | Lipid polymorphism and crystallization |
| Animal and vegetable fats | $^1$H | Solid fat content |

[a] For reference see Gil *et al* (1996) in Further reading section.

**Table 8** Multinuclear NMR studies of food biopolymers[a]

| *Nucleus* | *Biopolymer* | *Phenomenon* |
|---|---|---|
| $^{23}$Na, $^1$H, $^{13}$C | Carrageenans | Role of cation in gelation |
| $^{23}$Na, $^1$H | Pectins | Role of $^{23}$Na and degree of esterification on gelation |
| $^{31}$P | Starch | Number and location of phosphate ester groups and relationship to viscosity. Location of amylose in granule. |
| $^{87}$Rb, $^{133}$Cs, $^{127}$I | Carrageenans | Role of cation in gelation |
| $^{23}$Na, $^1$H | Xanthan gum | Sequestration of $Fe^{2+}$. Effect of $Na^+$ content on texture and stability |
| $^{25}$Mg, $^{43}$Ca, $^{31}$P | β-Casein | Ion binding sites |
| $^{111}$Cd, $^1$H | β-Lactoglobulin | Aggregation–gelation kinetics |

[a] For references see Gil *et al* (1996) in Further reading section.

### Low-resolution NMR analysis of food lipids

The NMR analysis of the oil content of seeds relies on the differences in proton transverse relaxation times of the constituents. Typical $T_2$ values of solid

macromolecules in seeds are just a few microseconds, and so they appear as the initial, rapidly decaying part of the FID. Water in seeds is strongly adsorbed to macromolecules and so has a short $T_2$ (of the order of a few milliseconds). In contrast, oil exists as liquid droplets in the seeds and therefore has a relatively long $T_2$ (of the order of hundreds of milliseconds). This means that the oil content can be determined directly from the FID as the ratio,

$$\text{Seed oil content } (\%, \text{ w/w}) = 100L/(L + fS)$$

where $L$ is the FID signal intensity measured after the solid signal has decayed and near the beginning of the liquid signal (typically after ~75 μs) and $S$ is the FID intensity of the slightly decayed solid signal. The correction factor, $f$, takes account of the decay during the instrument deadtime (typically the first 5–10 μs). The Hahn-echo pulse sequence is an alternative means of determining oil content. The water signal can be eliminated by choosing an echo time much longer than $5T_2$ of the seed water, the remaining oil signal then being calibrated with standard oil samples.

Alternatively, the echo amplitudes of the full CPMG echo-decay envelope acquired at short interpulse spacing can be analysed as a multiple exponential decay, whose relative weightings are proportional to the water and oil content of the seed. If this measurement is combined with the relative solid to liquid ratio determined from the FID, the complete solid–oil–moisture content of the seeds can be determined. Solid to liquid fat ratios in chocolate can be determined in a similar way. **Table 7** includes a number of such low-resolution studies of food lipids.

## List of symbols

$a$ = pore width; $D(\Delta)$ = water diffusion coefficient; $f$ = correction factor; $G$ = gradient amplitude; $L$ = FID liquid signal intensity; $M_0$ = initial magnetization; $P_{av}(R, \Delta)$ = average displacement propagator; $P$ = permeability; $q$ = diffusion wavevector defined as $(2\pi)^{-1}$ $\gamma G\delta$; $R$ = distance diffused by a spin; $S$ = amplitude of solid component in FID; $S$ = chemical shift; $S$ = FID solid signal intensity; $S(q, \Delta)$ = stimulated echo amplitude; $t_1$ = pulse separation time; $T_2$ = transverse relaxation time; $W$ = width of lamellar regions; $\Delta$ = diffusion time; $\gamma$ = gyromagnetic ratio.

*See also:* **Contrast Mechanisms in MRI; Diffusion Studied Using NMR Spectroscopy; Food and Dairy Products, Applications of Atomic Spectroscopy; Food Science, Applications of Mass Spectrometry; High Resolution Solid State NMR, $^{13}$C; Industrial Applications of IR and Raman Spectroscopy; Labelling Studies in Biochemistry Using NMR; MRI Applications, Biological; MRI Instrumentation; MRI Theory; MRI Using Stray Fields; NMR Data Processing; NMR Relaxation Rates; NMR of Solids.**

## Further reading

Gil AM, Belton PS, and Hills BP (1996) Applications of NMR to food science. *Annual Reports on NMR Spectroscopy* **32**: 1–49.

Hills BP (1998) *Magnetic Resonance Imaging in Food Science.* New York: Wiley.

Martin GJ (ed) (1996) Proceedings of the Third International Conference on applications of magnetic resonance in food science. *Journal of Magnetic Resonance Analysis* **2**: 1–264.

McCarthy KL, Kauten RJ and Agemura CK (1992) Application of NMR imaging to the study of velocity profiles during extrusion processing. *Trends in Food Science and Technology* **3**: 215–219.

McCarthy MJ (1994) *Magnetic Resonance Imaging in Foods.* New York: Chapman and Hall.

Packer KJ and Rees C (1972) Pulsed NMR studies of restricted diffusion: Droplet size distributions in emulsions. *Journal of Colloid Interface Science* **10**: 206–218.

# Forensic Science, Applications of Atomic Spectroscopy

**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

| ATOMIC SPECTROSCOPY |
| --- |
| **Applications** |

## Introduction

A number of studies have been published of the use of atomic spectroscopy in the furtherance of forensic science. This short article is intended to provide an outline of recent work that has been published in the open scientific literature. The representative Further reading at the end of this article should allow the reader an entry point into the scientific literature for details of methods and other studies.

## Reviews

The *Journal of Analytical Atomic Spectrometry* publishes periodical reviews of atomic spectroscopy, including its application to problems in metallurgy and industrial materials. This review also includes forensic studies and a recent example is given in the Further reading section.

## Analysis of glass

Fragments of glass can be transferred to a person during many criminal activities and recent work on the elemental analysis of glass has been reported using atomic emission spectroscopy. Results have been compared with those obtained using X-ray fluorescence.

## Gunshot residues

A review of many types of instrumental analysis used for the characterization of gunshot residues has appeared recently. This includes descriptions of the use of atomic absorption spectroscopy, neutron activation analysis and X-ray fluorescence as well as various mass spectrometry techniques after preconcentration of analytes using separation methods such as HPLC and capillary electrophoresis. In addition, atomic absorption spectroscopy has been used to characterize the brass used in different production lots of cartridge cases.

## Metal residues

Inductively coupled plasma atomic emission spectroscopy (ICP-AES) has been used to characterize the elemental composition of steel fragments found at crime scenes. In general the levels of chromium, nickel, cadmium, zinc and copper could be used as fingerprints to identify the source of the steel. The trace elemental composition of household aluminium foils can also, in some cases, be used to identify the brand name and production details.

*See also:* **Atomic Absorption, Methods and Instrumentation; Atomic Absorption, Theory; Atomic Emission, Methods and Instrumentation; Atomic Spectroscopy, Historical Perspective; Forensic Science, Applications of IR Spectroscopy; Forensic Science, Applications of Mass Spectrometry; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Crighton JS, Fairman B, Haines J, Hinds MW, Nelms SM and Penny DM (1997) Industrial analysis: metals, chemicals and advanced materials. *Journal of Analytical Atomic Spectrometry* **12**: R509–R542.

Curran JM, Triggs CM, Almirall JR, Buckleton JS and Walsh KAJ (1997) The interpretation of elemental composition measurements from forensic glass evidence. *Science and Justice* **37**: 241–244.

Heye CL and Thornton JI (1994) Firearm cartridge case comparison by graphite furnace atomic absorption spectrochemical determination of nickel, iron and lead. *Analytica Chimica Acta* **288**: 83–96.

Koons RD, Peters CA and Merrill RA (1993) Forensic comparison of household aluminium foils using elemental composition by inductively coupled plasma atomic emission spectrometry. *Journal of Forensic Sciences* **38**: 302–315.

Koons RD, Peters CA and Rebbert PS (1991) Comparison of refractive index, energy dispersive X-ray fluorescence and inductively coupled plasma atomic emission

spectrometry for forensic characterization of sheet glass fragments. *Journal of Analytical Atomic Spectrometry* **6**: 451–456.
Meng HH and Caddy B (1997) Gunshot residue analysis – a review. *Journal of Forensic Sciences* **42**: 553–570.
Poolman DG and Pistorius PC (1996) The possibility of using elemental analysis to identify debris from the cutting of mild steel. *Journal of Forensic Sciences* **41**: 998–1004.
Thornton JI (1994) The chemistry of death by gunshot. *Analytica Chimica Acta* **288**: 71–81.

# Forensic Science, Applications of IR Spectroscopy

**Núria Ferrer**, Universitat of Barcelona, Spain

Copyright © 1999 Academic Press

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
| --- |
| Applications |

## Introduction

Infrared spectroscopy has been applied to the characterization of molecules. Conventional dispersive infrared spectrometers have been replaced by Fourier-transform infrared equipment, which incorporates a Michelson interferometer and presents a series of advantages over dispersive systems, such as an improvement of energy and the simultaneous measurement of the whole spectral range.

Owing to these improvements, infrared spectroscopy has undergone a marked development in the last few years because of the possibility of adapting new accessories to the spectrometers and, therefore, of analysing samples whose infrared spectrum was impossible to obtain some years ago. Among these accessories are those which can measure specular and diffuse reflection, attenuated total reflectance and microscopes. One area in which the technique has improved dramatically is forensic analysis.

Many laboratories perform forensic analysis nowadays. Some of them seek to determine what is wrong with a process, while others are trying to copy a process from a competitor. The possibility of analysing and identifying microscopic particles, the accurate comparison of samples from different origins and the fact that the sample is not destroyed, have also converted the FT-IR technique into a valuable tool where criminal and social implications are present. Some of the forensic materials that can be identified by infrared spectroscopy are: paints, fibres, textiles, polymers, explosives, varnishes, adhesives, papers, gemstones, drugs, food additives, fur or textile treatments and physiological samples. An important feature is the possibility of analysing solid, liquid or gaseous samples. In criminal investigations it is important to attach the infrared spectrum to each sample, which is unique to every substance, like a fingerprint. In addition, in lawsuits where several parties are involved, infrared spectroscopy can contribute relevant information.

This article shows a number of applications of infrared spectroscopy to the forensic analysis of a range of samples using different accessories and methods. A number of examples are given, although applications are unlimited and any analysis might be of value in forensic research.

## Main accessories and methods used in forensic analysis

The two main restrictions of analysis by infrared spectroscopy are size and difficulties regarding handling. Three kinds of sample compartments are available depending on the size of the sample to be placed in the spectrometer:

- macro compartment: this employs a beam between 2 to 10 mm and any type of accessory might be fitted.
- Semimicro compartment: this requires the use of a beam condenser and the spot of the beam in the sample is about 1 or 2 mm.
- micro compartment: the size of the beam is less than 1 mm and a microscope needs to be used.

A beam condenser can be installed in the sample compartment of the conventional equipment, but the microscope needs the infrared beam to go out of the bench to an externally attached accessory. The microscope uses the source and beamsplitter of the infrared equipment to which it is coupled, but it has its own more sensitive detector. Obviously, large samples can also be located on the microscope, but in

such cases the measurement area is reduced to less than 100 µm. Therefore, if an analysis of the predominant material in the sample is required, its homogeneity must be assured and defects or pollutants must be avoided.

Depending on the manipulation that the sample requires, a range of accessories might be used. Many samples for forensic analysis must not be destroyed or manipulated either because they need to be kept for further analysis or because of their economic or historic value. KBr pellet measurement requires mixing the sample with the salt and, therefore, the possibility exists that it might be lost. Other methods such as diamond cell or reflection do not require any alteration or mixing of the sample with other substances. In the main, liquid and gaseous samples suffer no risk of destruction and it is only necessary to fit them in cells which are transparent in the desired range of measurement and which are also compatible with the material to be analysed. Solid samples can be mixed with salts in order to make an infrared transparent pellet, or can be placed on a suitable surface for further analysis.

### Diamond cell

Diamond cell is one of the most frequently used methods in forensic analysis. It is normally used when samples are very small, but can be separated from the matrix. The particle or fibre, which can be as small as 5 µm, is placed on a diamond window and put under pressure with the help of another diamond window. The pressure of the cell allows the sample to spread over the diamond increasing its surface area while decreasing in thickness. In this way, the infrared radiation is able to go through the sample and, therefore, its infrared spectrum can be obtained by transmission.

Diamond cells are about 4 $mm^2$ in aperture and depending on the area of the spot of the spread sample it is possible to utilize a beam condenser or a microscope. Despite the absorption of diamond in the mid infrared, the wavenumber reproducibility that FT-IR is able to achieve makes it possible to subtract the diamond spectrum and to obtain clean spectra of the sample. In forensic analysis this technique is applied to paints, varnishes, inks, fibres, crystals and particles in general, which can be put under pressure and suffer a deformation. To avoid excess of absorption, one of the cells can be removed. In the case of samples such as cork, tobacco and some polymers, which recover their original size after decompression of the cell, the alternative is to use two diamond windows with the sample in between, instead of removing one of them.

### Specular reflection

When a sample cannot be manipulated at all, is opaque and has a specular or polished surface, specular reflection is the technique to be applied in either the microscope or the macro compartment.

### Diffuse reflection

When the surface of the sample is not specular and there is a certain diffuse component of reflected light, or when analysing a cut gemstone, the analysis is appropriately conducted by using the diffuse reflection technique where all the light emitted by the sample is measured.

### Attenuated total reflectance (ATR)

Attenuated total reflectance is applied to samples where the composition of the surface needs to be measured. It is applied to soft samples such as paper, fur, polymers and textiles, which can achieve good contact with the crystal used. It is also used for aqueous samples and viscous or sticky samples such as pastes, soaps and detergents. The only consideration to be borne in mind is the need to obtain good optical contact between the surface of the sample and the crystal of the accessory, which can be made of diamond, ZnSe, Ge, Si, KRS-5, etc.

When the surfaces to be analysed are very small it is possible to use the attenuated total reflectance objective of the microscope. A unique internal reflection over the sample in touch with the crystal is enough to obtain the spectrum. This is the alternative when even a single particle cannot be removed from the matrix, which is the case for very ancient or valuable samples such as manuscripts, books or paintings.

## Separation techniques

Determination of complex mixtures is one of the most common problems in infrared analysis. Hyphenated techniques such as chromatography provides a useful way of separating mixtures into individual compounds. Gas chromatography (GC) has been by far the best application of combining infrared spectroscopy and a separation technique. One of the principal reasons has been the use of carrier gases which do not absorb in the infrared region.

## Specific analyses

### Analysis of paints

Infrared spectroscopy has been widely applied in the forensic analysis of paints. It is particularly useful in

car accidents where it is necessary to determine the origin of the paint adhering to a car. In some countries there are car paint libraries, and it is possible to find the model of the car and the year of manufacture by analysing the paint. In the absence of witnesses, or if a vehicle has left the scene of an accident, this analysis can be a crucial element in obtaining a verdict.

Another major application of the technique is in the analysis of ancient paints used in a particular period of history or a particular place. This is essential not only for the study of the pigments and materials used by different civilisations, but also for authenticating works of art, restoration work and in lawsuits involving damaged works of art.

**Figure 1** shows two spectra from two sections of a modern painting. The damage to the picture might have been caused by a heterogeneous application of oil or the poor conservation of the picture. A non-homogeneous distribution of oil over the surface was made evident when two sections of the picture were analysed. As it was possible to use part of the picture from under the frame, the amount of sample was sufficient to allow use of the attenuated total reflectance horizontal accessory.

A further application of this technique is in seeking to obtain high quality paints. Changes in pigment composition as well as in solvent composition can have a considerable impact in industry. Therefore, it is important to be able to determine quickly the composition of both solid and volatile fractions. In addition, increasingly stringent controls make it necessary to check for the presence of forbidden solvents in commercial paints.

**Figure 2** shows the spectrum of the volatile fraction of a commercial paint. The spectrum is obtained by placing a long strip of aluminium foil on the bottom of the gas cell with some paint on the foil. After a few minutes the number of molecules within the cell corresponding to the volatile fraction is sufficiently great to obtain the spectrum of the gas. In this particular case it was possible to identify xylol as a component of the gaseous fraction. Pigment analysis can be performed by covering a KBr pellet with a very thin layer of paint.

When paint samples are very small, for example in the case of a valuable painting, or an archaeological sample or the paint which remains attached to a lead bullet once it has passed through the bodywork of a car, the most widely used technique is to put small particles on a diamond cell and to obtain the infrared spectrum by transmission. If the sample cannot be removed from the matrix, microscopy with an attenuated total reflectance objective can be used.

### Fibre analysis

Forensic analysis of fibres constitutes another major application of infrared spectroscopy. The spherical structure of fibres makes it necessary to flatten the sample before measuring its spectrum, for which it is possible to use either a roller or a diamond cell. Once the fibre is prepared, the variable apertures of the microscope allow the masks to be adapted to the size of the fibre.

**Figure 3** shows the spectrum of a fibre found in a pharmaceutical product. Here, it was necessary to know whether the sample came from the filters in the industrial process or from the white coats of the workers. Although it was clear that the three fibres analysed were polyester, the technique was able to identify the kind of polyester and to determine its relation to fibres coming from the filters of the industrial process. In this case the sample consisted of a single fibre and, therefore, the use of a microscope was the only option. If the sample consists of a piece of textile or several fibres, other methods such as attenuated total reflectance, KBr pelleting following shattering after treatment in liquid nitrogen or diamond cell in the beam condenser can be used.

Other applications have been reported, in the field of medicine where suture threads have had harmful reactions in patients, and for analysing single fibres recovered from the scenes of crime.

The appearance of surface defects in animal furs is another reported application of attenuated total reflectance techniques. **Figure 4** shows an example of a valuable fur damaged by white spots. The manufacturer claimed that the client was responsible for damage to the fur because it had been stored in an inappropriate environment or because of treatments it had been exposed to. The client, on the other hand, claimed its original treatment during manufacture was to blame. Infrared analysis demonstrated that the white spots came from the migration of animal fat to the surface and bands were comparable to the fat extracted from the natural fur. Other substances, such as silicone, have been found in this kind of sample.

### Paper and ink analysis

Application of the infrared technique with microscope to inks in paper has been a useful tool. Examples include comparison of inks used on the same cheque, for example an additional zero to the total sum or in the signature, and documents which are signed and dated, where it is possible to analyse the inks from printers or hand typewriters to determine if they actually existed when the document was signed. Particles of toner from a photocopy can also

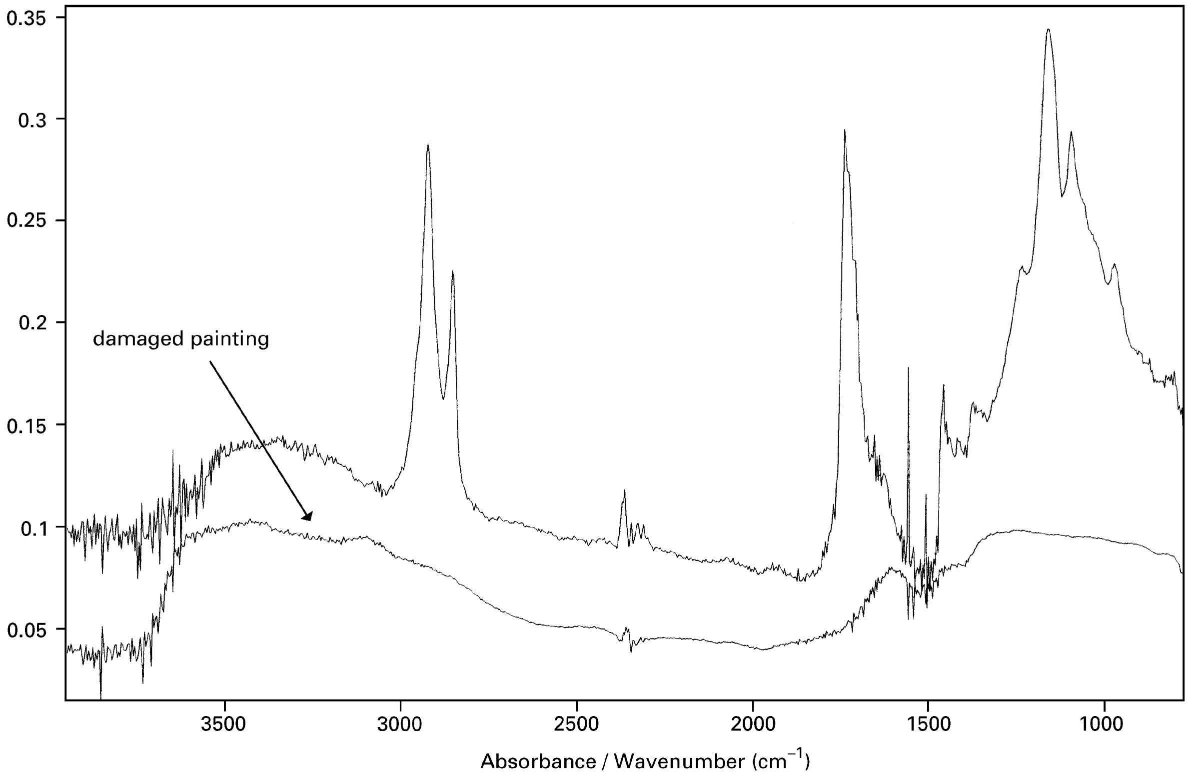



**Figure 1** Spectra of two pieces of painting: damaged painting does not show the bands corresponding to oil (ATR horizontal).

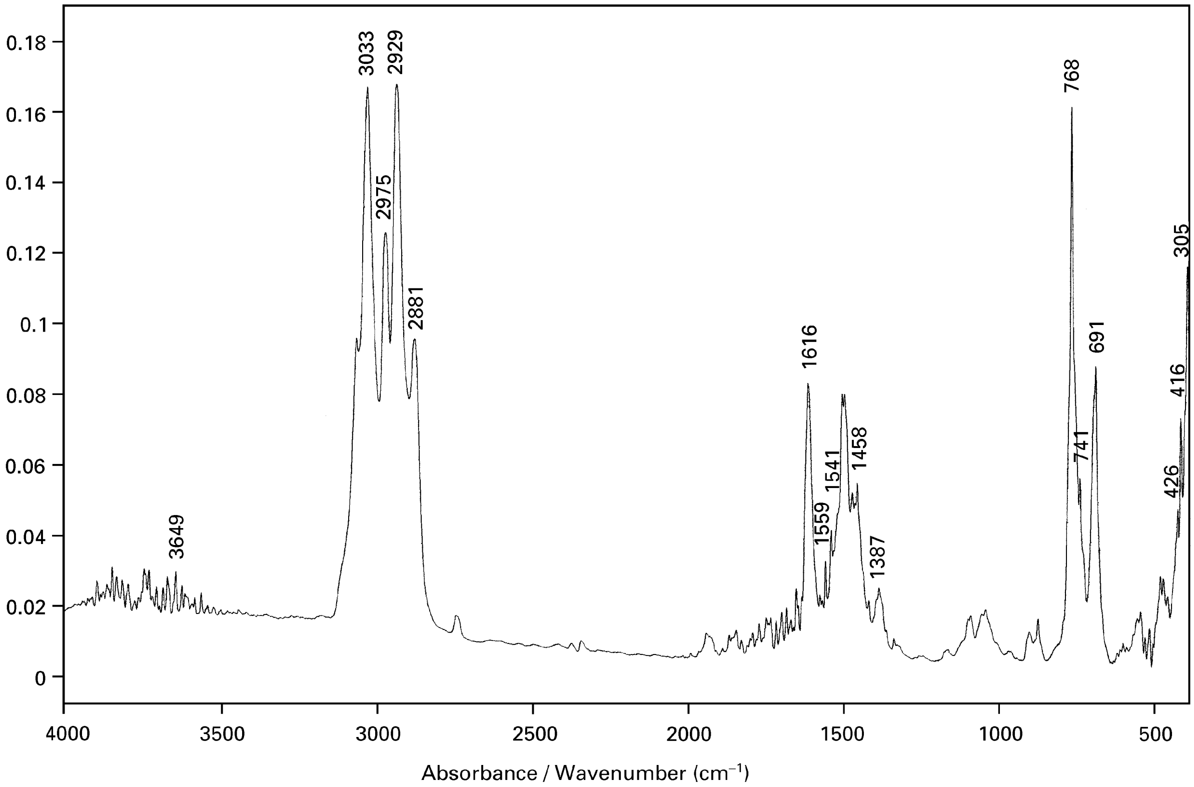



**Figure 2** Spectrum of volatile fraction (xylol) of an industrial painting (gas cell).

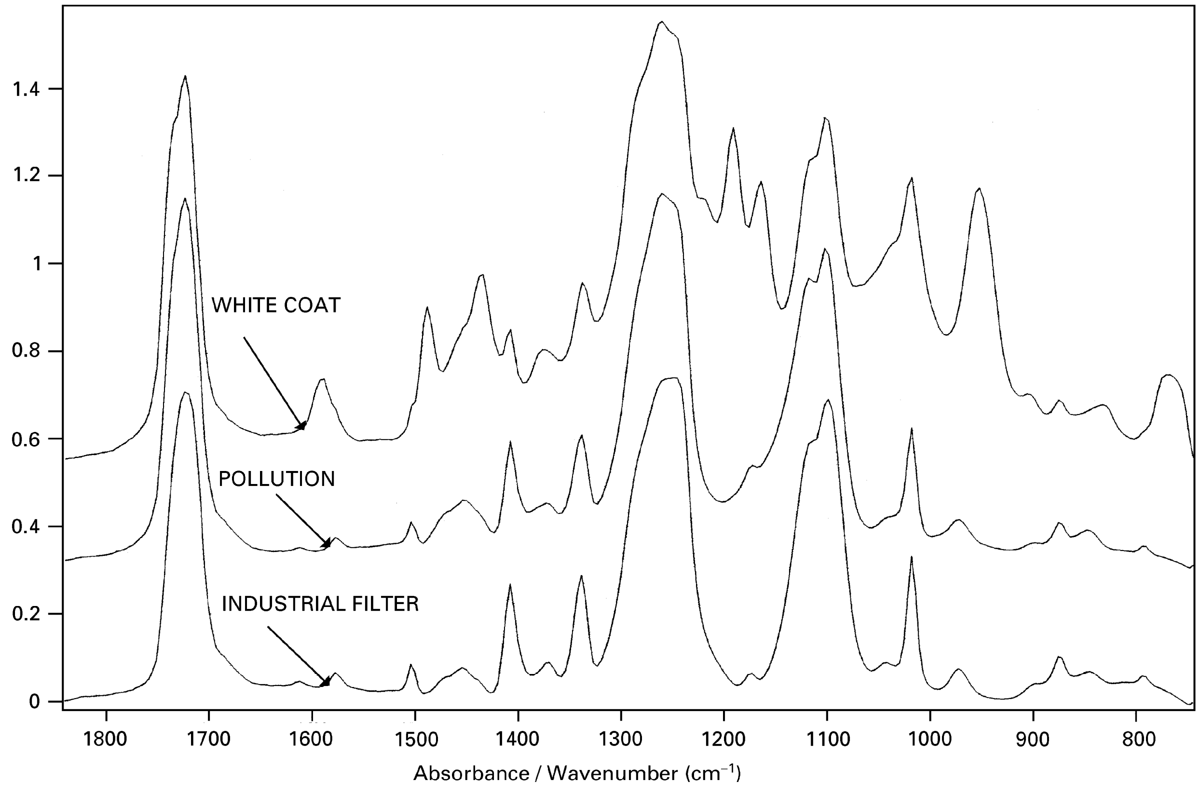



**Figure 3** Comparison of fibres from a polluted sample, white coats and an industrial filter (diamond cell in the microscope).

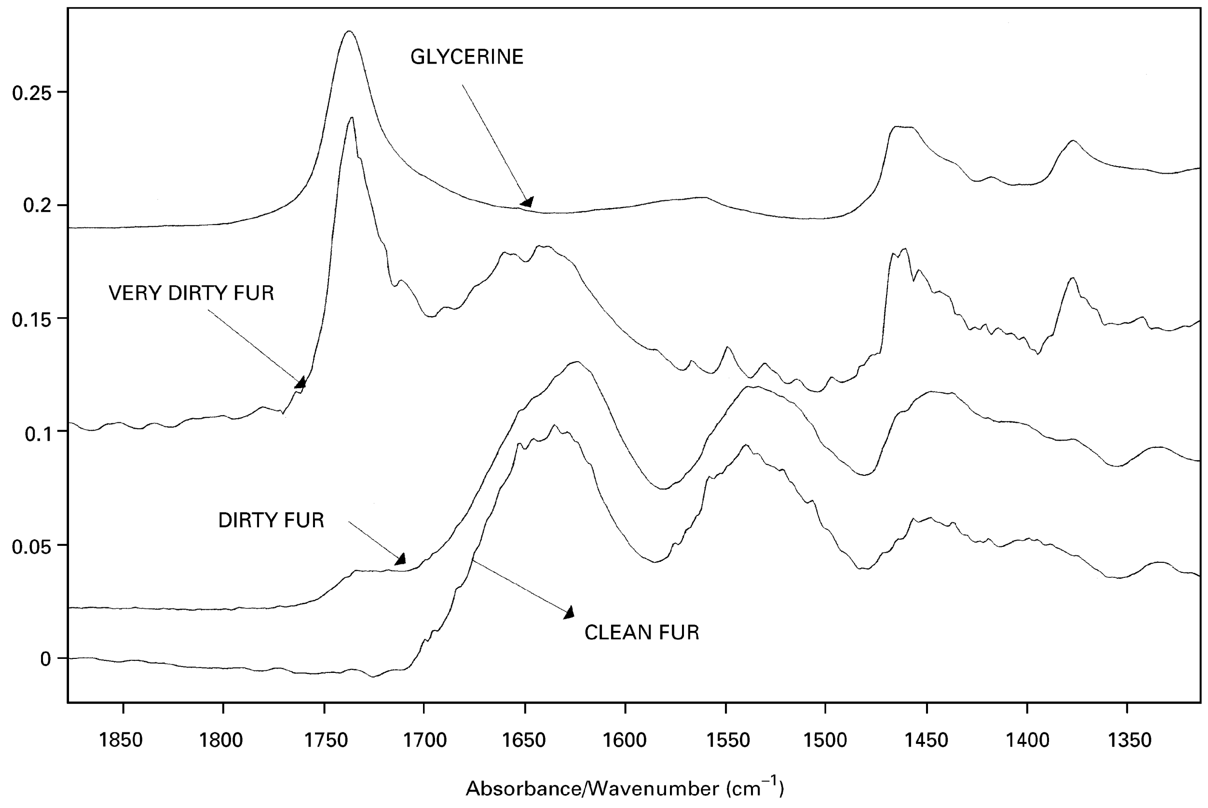



**Figure 4** Comparison of spectra of clean and dirty furs with glycerine spectrum (ATR horizontal).

help to solve cases by comparing the particles of the document with the toner from the photocopying machine used.

**Figure 5** shows the spectra obtained by transmission microscopy of two cotton fibres, one of which contains a blue ink spot. Comparison of the spectral differences between the clean fibre and that spotted with ink, shows that this difference can indeed be imputed to the ink.

Cellulose fibres also have a highly characteristic infrared spectrum. Acidity of paper due to the presence of acidic inks or biological degradation can be measured by analysing the infrared spectrum. **Figure 6** shows the spectral differences in certain ranges of the spectrum due to corrosion related to the acidity of ancient papers. Carbonate bands at 875 and 1797 $cm^{-1}$ decrease with decreasing pH, carbonyl groups increase with decreasing pH and hydroxyl groups from cellulose decrease with increasing pH. This kind of change arises as a result of poor conservation or the type of paper.

Spots on papers can also be measured. **Figure 7** shows the spectra of two samples from the same piece of paper, one of which is spotted. The attenuated total reflectance objective of the microscope allows us to identify the presence of paraffins. In such cases, a microscope is essential to obtain the infrared spectrum while the attenuated total reflectance objective allows the analysis to be carried out without destroying the document.

## Analysis of gemstones

There is an increasing number of gemstones, treated or synthetic, that need analysing and the technique that allows them to be distinguished visually is not particularly reliable. Infrared spectroscopy can provide objective evidence for the characterization of such samples and this is very important for both identification and certification purposes in commercial and legal practices.

Diamonds are one of the more interesting groups of gems to study. Only when diamonds present two parallel faces is it possible to use transmission to determine the infrared spectrum. Normally diamonds, especially those destined for jewellery, have a brilliant cut. This cut affects the light and makes it go out in all directions. In this case the only technique that allows us to obtain the infrared spectrum is diffuse refle.ction. Here the gem is placed horizontally over an aluminium mirror.

Natural diamond without impurities has a characteristic spectrum with bands that are common to both natural and synthetic diamonds, but most diamonds also present bands caused by defects in the crystalline lattice. We can distinguish three different ranges in the infrared spectrum. The first corresponds to the interval between 5000 and 2700 $cm^{-1}$ and is the range where the bands corresponding to hydrogen impurities appear (4499, 4169, 3237, 3107 and 2785 $cm^{-1}$) and is characteristic of the natural diamond. The second range corresponds to the common bands of the diamond structure but they do not reveal the difference between natural and synthetic diamond (2505, 2443, 2159, 2033 and 1970 $cm^{-1}$). Finally, the range between 1400 and 1000 $cm^{-1}$ identifies the impurities due to nitrogen as a specific defect in the crystalline lattice. According to the type of nitrogen substitution, there are different kinds of diamonds and it is in this range that it is possible to distinguish between the natural and the synthetic stones. **Figure 8** shows the spectra of a natural diamond with hydrogen and nitrogen impurities, and of a synthetic diamond.

The spectral differences of each diamond are in general great enough to distinguish between them. It is of particular value in identifying a gem with a high commercial value. If an infrared spectrum is attached to a gem, it can be used as proof in case of legal certification. The ratio of the hydrogen or nitrogen bands to the common bands of the second region of the diamond provides a value that is probably unique to the stone.

Another group of gems for which it has been possible to measure the differences between natural and synthetic specimens is emeralds. The water and carbon dioxide content in natural gems means they can be distinguished by measuring both mid and near infrared bands. **Figure 9** shows the infrared spectra of a natural and a fused phase synthetic emerald, where carbon dioxide and water are not present. Nowadays, techniques of manufacture of synthetic emeralds allow water to be incorporated in the lattice. In spite of this, infrared spectroscopy can distinguish between both kinds of water with the help of polarized radiation. **Figure 10** shows the near infrared spectra of a natural and a synthetic emerald using polarized light at 0° and 90°. In general, infrared spectroscopy enables us not only to see the differences between certain synthetic and natural specimens, but also to obtain the fingerprint of each gem with its impurities and, therefore, to use it in judicial and commercial certification.

## Polymer analysis

Polymers are a group of substances that can easily be investigated by infrared spectroscopy. Apart from analyses aimed at identifying small particles or the composition of packagings used for food or drinks,

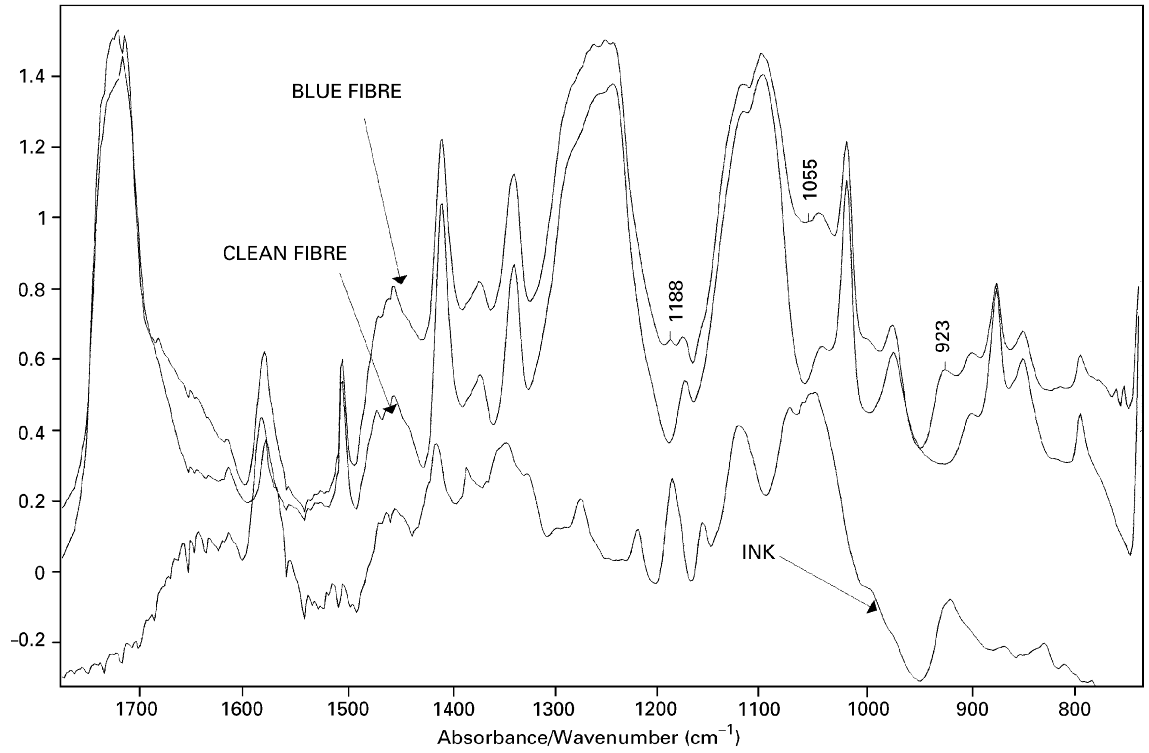



**Figure 5** Comparison of the difference between a clean and a blue cotton fibre with the suspected ink (diamond cell in the microscope).

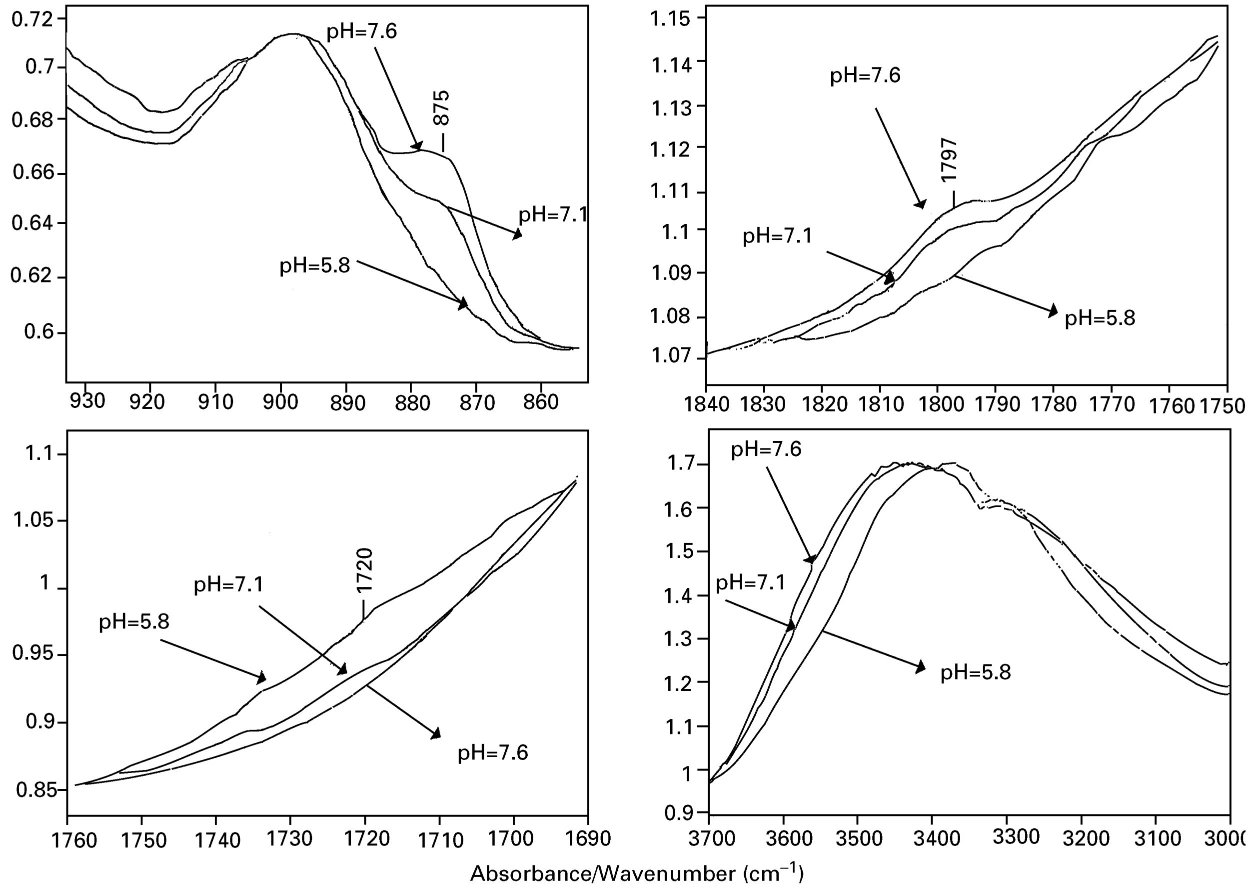



**Figure 6** Changes of four different regions of infrared spectrum with pH of paper (diffuse reflection accessory).

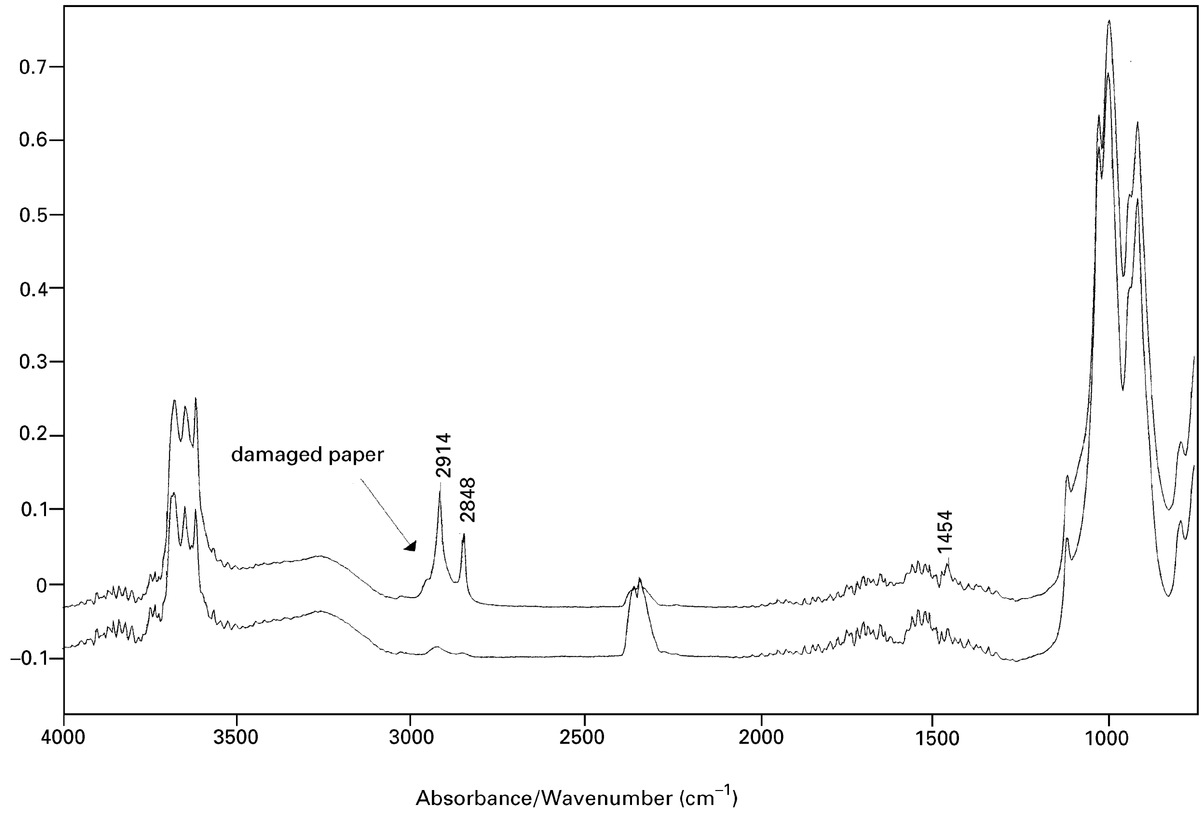



**Figure 7** Spectrum of a spot on the surface of a paper identified as paraffin (ATR microscope).

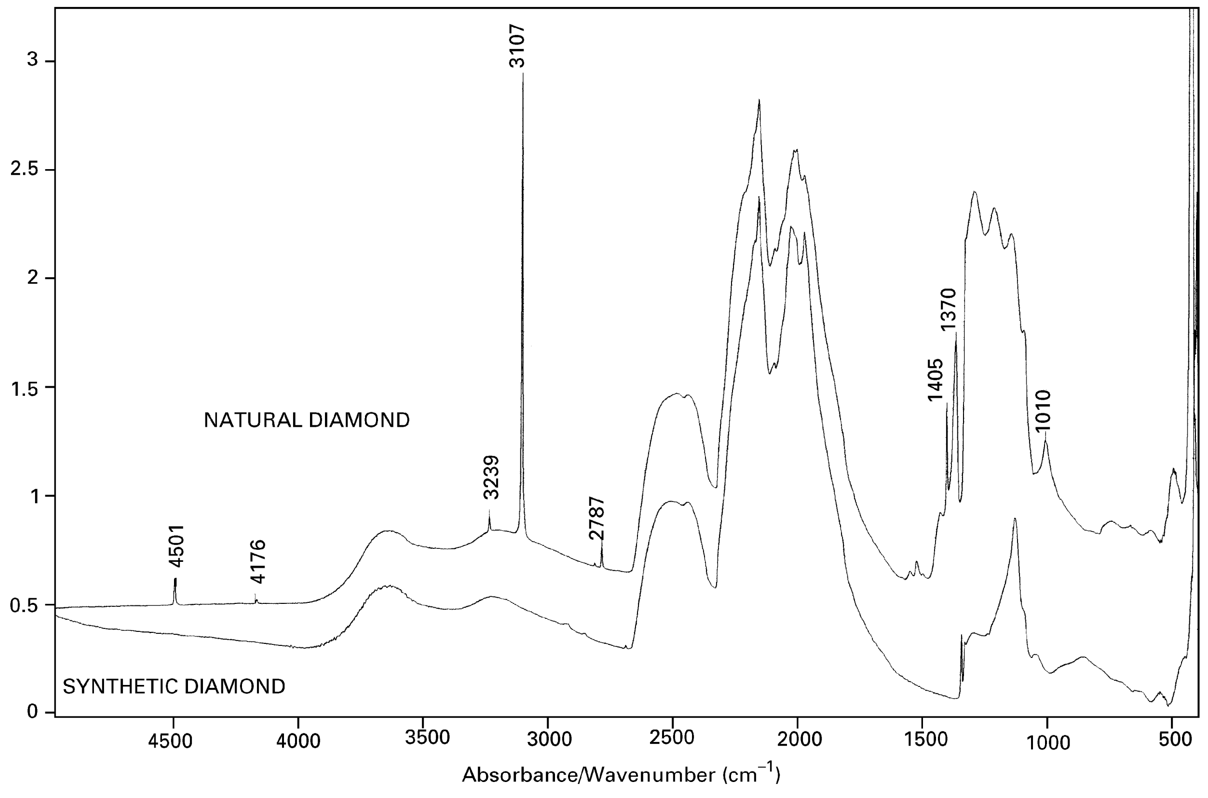



**Figure 8** Comparison of a natural and a synthetic diamond (diffuse reflection).

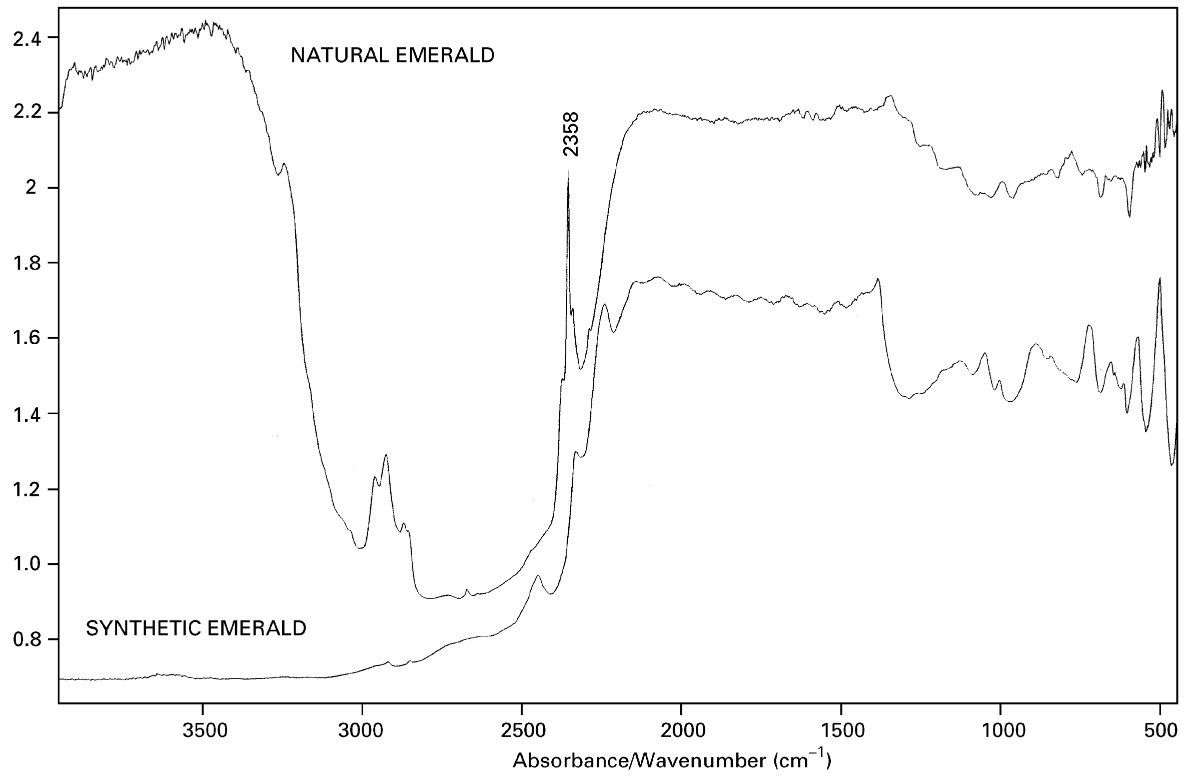



**Figure 9** Comparison of a natural and a synthetic emerald (diffuse reflection).

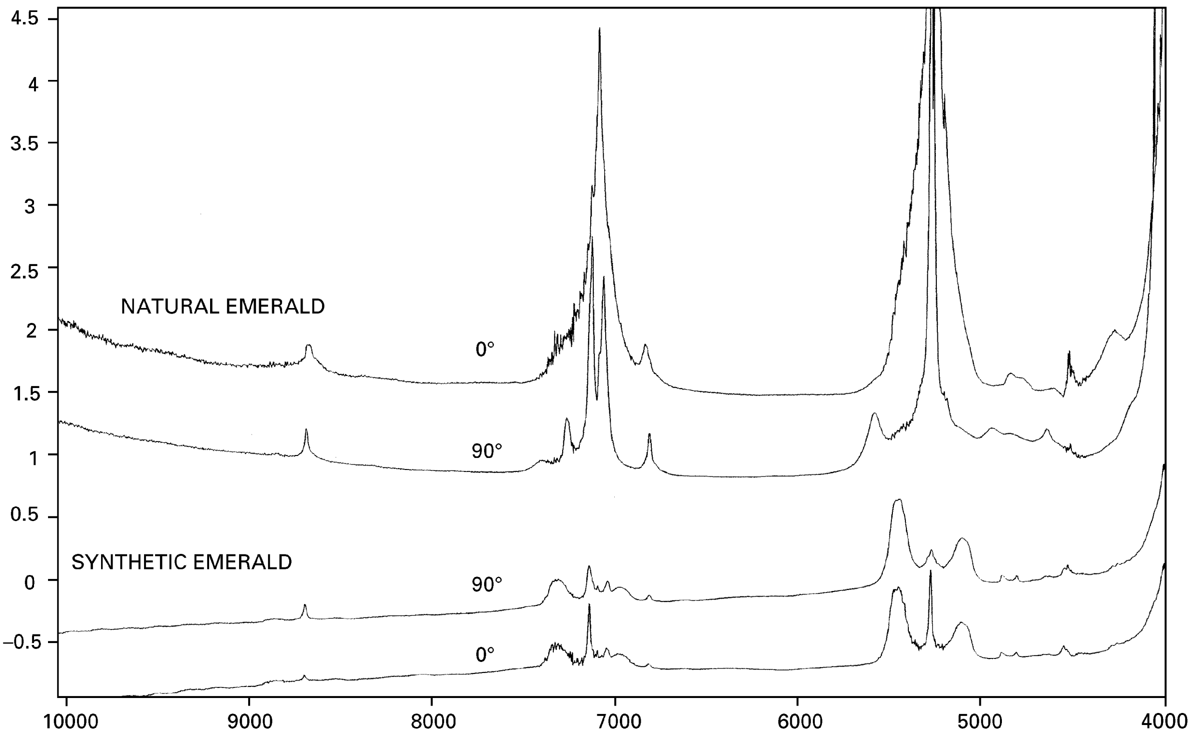



**Figure 10** Comparison of bands of water in the near infrared between a natural and a synthetic emerald using polarized light (reflection microscope).

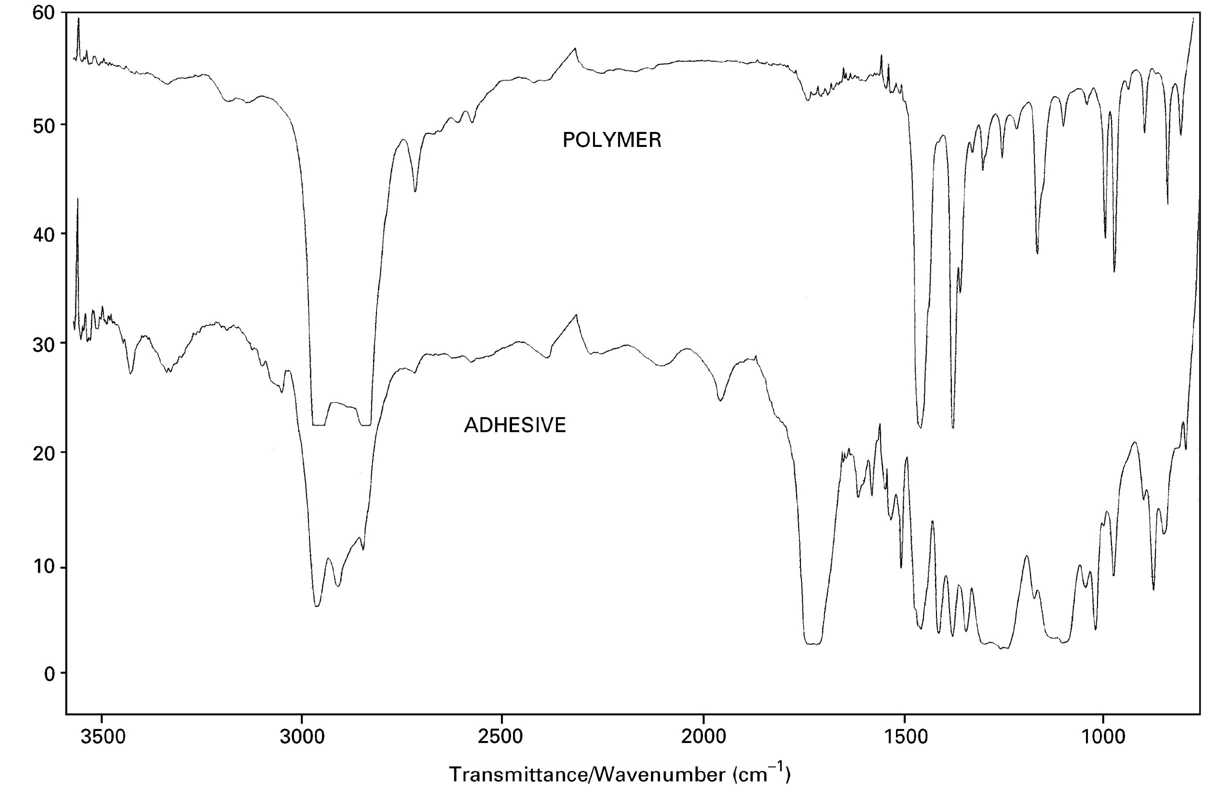



**Figure 11** Spectra of the polymer and the adhesive of a multilayer film (transmission microscope).

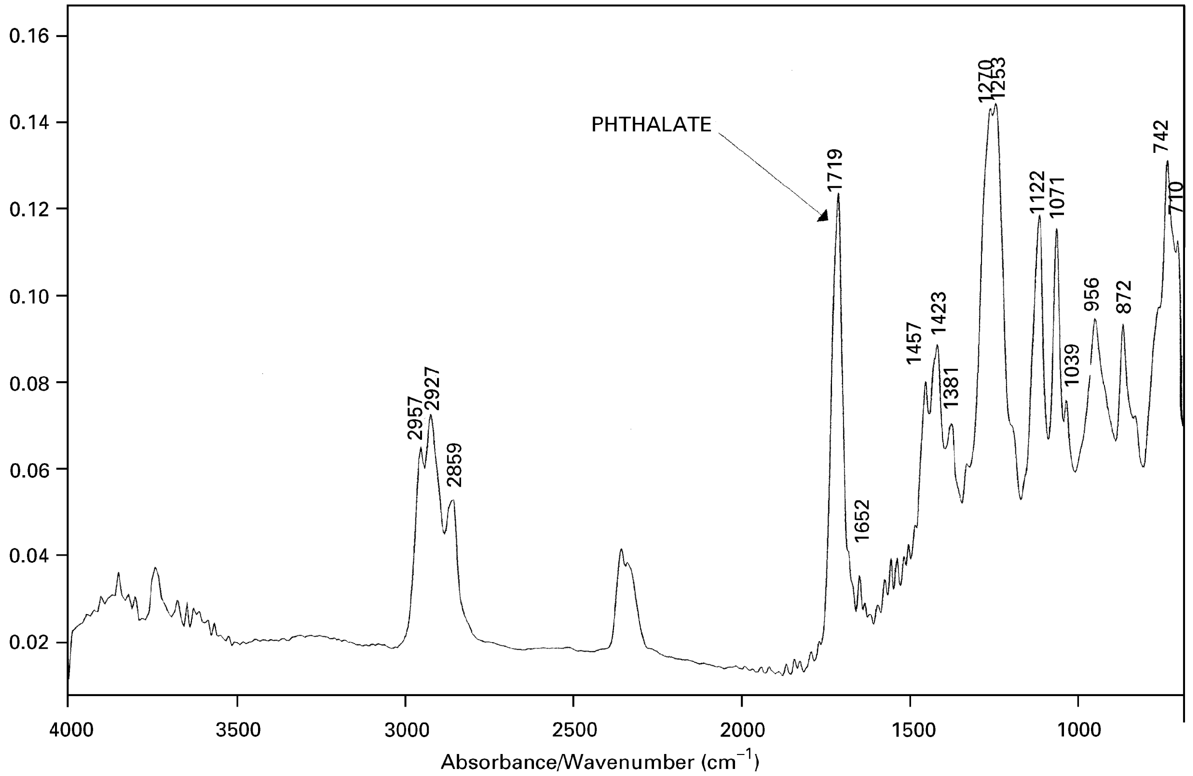



**Figure 12** Spectrum of a PVC polymer used in contact with drinking water, showing phthalate bands at 1719 cm$^{-1}$ (ATR horizontal).

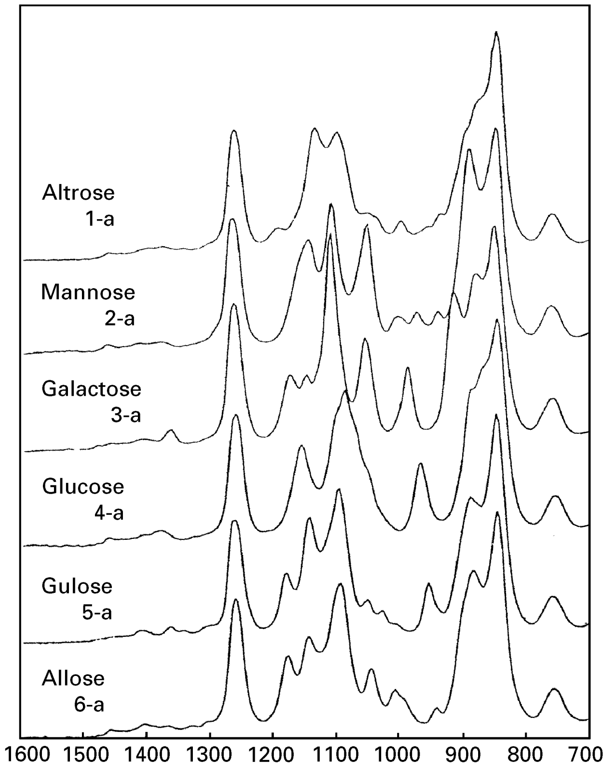



**Figure 13** GC-IR spectra of six aldohexoses.

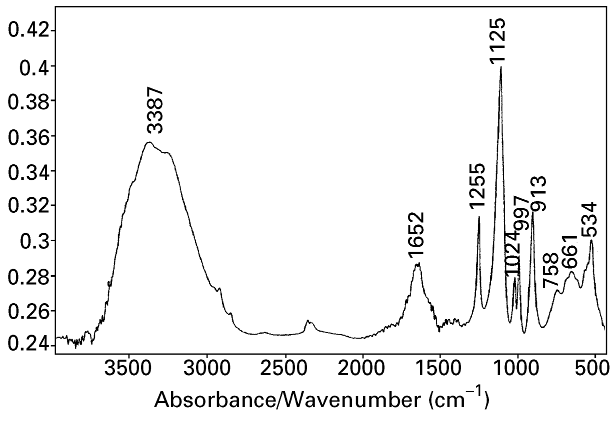



**Figure 14** Spectrum of particles on the surface of a frozen fish (diamond cell in beam condenser).

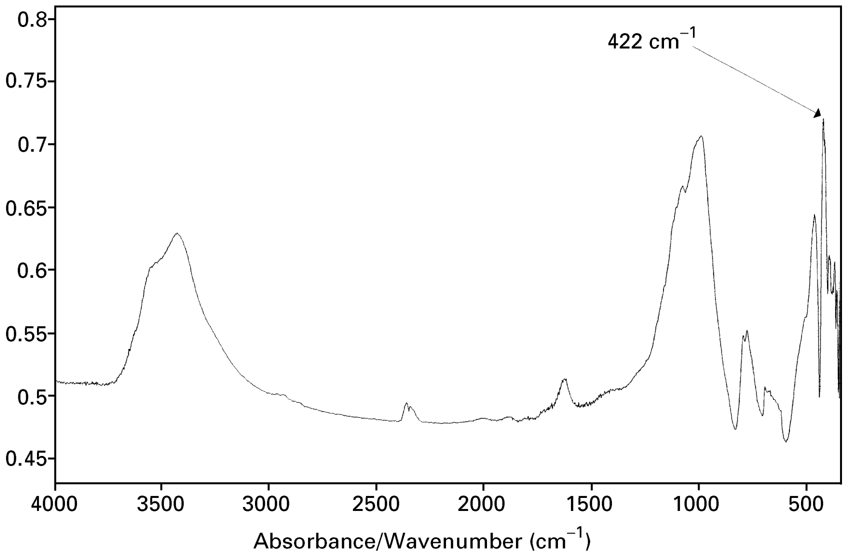



**Figure 15** Spectrum of pyrite from the spill of waste in Doñana, showing pyrite band at 422 cm$^{-1}$ (KBr pellet).

the use of infrared spectroscopy allows multilayer polymers and adhesives used with them to be determined. **Figure 11** shows the infrared spectra of a 50 μm thick packaging where the content was altered because of migration of components from the adhesive between the layers.

The use of phthalates as plasticizers in the polymers of food packagings can lead to the migration of these molecules to the food itself. Therefore, it is essential to check for their presence in such packagings. **Figure 12** shows the attenuated total reflectance spectra of a polymer in contact with drinking water where phthalates are detected.

### Analysis of physiological samples

Infrared spectroscopy has also been applied to the identification of particles found in organs such as the lung, kidney or liver. Some studies have been reported in which this technique has been applied to the analysis of malignant cells, by measuring the region of phosphodiesters of lymphocytes. This technique has been applied to the determination of samples of normal and malignant cervical cells. The presence of biomaterials such as polyesters, polyurethanes and silicones originating from implants can also be measured.

In the field of drugs, the analysis of human and animal hairs has been reported. Application of infrared microscopy in determining the presence of drugs in inner cuts of hair samples is a useful method for avoiding false positives from external contamination.

Some minerals, including carbonates and phosphates, from both inner and external parts of bones, have also been measured.

### Food analysis

A number of analyses show the possibility of measuring additives used as preservatives in foods. Adulteration of juices, vegetable oils and milk, as well as caffeine content, protein and lipid content or sugar composition are other examples that have been reported.

**Figure 13** shows the GC-IR spectra of six aldohexoses. None of the monosaccharides showed identical spectral patterns, so the technique was effective for their identification following separation by GC.

**Figure 14** shows the spectrum of particles removed from the surface of frozen fish. The particles were identified as pyrophosphate.

## Environmental samples

Analysis of solid samples which might contain particles from a contamination is another area for which infrared spectroscopy might be needed. In cases where it is necessary to prove that a particular waste has been introduced in a certain place, the technique allows the nature of the particles found to be analysed. Compensation is often available for those who have lost crops or cattle, but evidence of pollution must be proved.

Analysis of vapours from polluted soils or from stacks can also be carried out.

**Figure 15** shows the spectrum of particles from the spill near the natural park of Doñana (Spain) which caused an ecological catastrophe in 1998. The pool from which the spill originated was full of pyrite waste and a black wave covered crops over an area extending many kilometres. The band at 422 cm$^{-1}$

corresponds to pyrite ($FeS_2$). Infrared spectroscopy was able to identify this compound at a distance of many kilometres from the spill.

The identification of compounds that should not occur in a place, makes it possible to impute them to spills or pollution from other places.

## Conclusions

To sum up, infrared spectroscopy allows a rapid, initial determination of a substance sent for forensic analysis and can provide evidence that is of great, perhaps decisive, use.

It is often difficult to identify a substance by its infrared spectrum, especially one containing organic molecules, when there is no indication of what it might be. Inorganic molecules are easier to identify, at least those bands corresponding to the anionic part of the molecules: carbonates, silicates, nitrates, sulfates, phosphates, etc. Yet, frequently it is only necessary to compare samples or to follow up a suspicion of the presence of a substance.

Libraries play an important role in forensic analysis, but the best libraries are the ones built up by analysts themselves using suitable accessories and developing their own methods.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Chromatography-IR Methods and Instrumentation; Chromatography-IR Applications; Fibres and Films Studied Using X-Ray Diffraction; Forensic Science, Applications of Atomic Spectroscopy; Forensic Science, Applications of Mass Spectrometry; Geology and Mineralogy, Applications of Atomic Spectroscopy; Industrial Applications of IR and Raman Spectroscopy; Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; Medical Science Applications of IR; MRI of Oil/Water in Rocks; Polymer Applications of IR and Raman Spectroscopy.**

## Further reading

Brettel TA and Saferstein R (1997) Forensic science. *Analytical Chemistry 97 Application Reviews* 123R.

Coleman PB (ed) (1993) *Practical Sampling Techniques for Infrared Analysis*. Boca Raton, FL: CRC Press.

Ferraro JR and Krishnan K (eds) (1990) *Practical Fourier Transform Infrared Spectroscopy. Industrial and Laboratory Chemical Analysis*. San Diego: Academic Press.

McKelvy ML, Britt TR, Davis BL, GLillie JK, Lentz LA, Leugers A, Nyquist RA and Putzig CL (1996) Infrared spectroscopy. *Analytical Chemistry* **68**: 93R–160R.

Messerschmidt RR and Harthcock MA (eds) (1988) *Infrared Microspectroscopy Theory and Applications*. New York: Marcel Dekker.

Roush PB (ed) (1987) *The Design, Sample Handling and Applications of Infrared Microscopes*. Ann Arbor: ASTM.

# Forensic Science, Applications of Mass Spectrometry

**Rodger L Foltz**, **David M Andrenyak** and **Dennis J Crouch**, University of Utah, Salt Lake City, UT, USA

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Applications**

An important component of forensic toxicology is the determination of drugs in biological specimens when the analytical results have potential legal consequences. Mass spectrometry is generally the primary analytical tool for these investigations. This is because mass spectrometry, in combination with chromatographic methods of sample introduction, offers the best combination of high sensitivity and reliable identification. The areas of forensic toxicology where mass spectrometry has played a particularly important role include workplace drug testing, testing of motor vehicle operators for drug impairment, and postmortem toxicology.

Workplace drug testing typically involves the analysis of urine samples from job applicants and/or employees to determine whether they have recently used drugs of abuse such as heroin, marijuana, amphetamines or cocaine. Test samples are normally screened by a relatively inexpensive immunoassay, and all presumptively positive samples are submitted for mass spectral analysis to confirm the presence of the drug or drugs. Similarly, for determination of

impairment in motor vehicle operators, blood samples are often analysed by mass spectrometric methods to determine whether the impaired performance was associated with recent use of a psychoactive drug. In postmortem toxicology, additional sample matrices are often analysed by mass spectrometric methods in order to gain the greatest insight into whether drugs contributed to the cause of death. In each of these applications the biological specimens must be collected, transported, stored and analysed by procedures that are scientifically and legally supportable.

The application of mass spectrometry in forensic toxicology typically consists of extraction of the drug or drug metabolites from selected body fluids or tissues, chromatographic separation of the drug from other co-extractants, and mass spectrometric detection of the analyte. The chromatographic separation is most often performed by capillary column gas chromatography (GC), following chemical conversion of the analyte to a volatile derivative that has more favourable chromatographic and mass spectral characteristics. The combination of liquid chromatography and mass spectrometry (LC-MS) can also be used, but its adoption by forensic toxicology laboratories has been slow, primarily because of the high cost of LC-MS instrumentation. Nevertheless, LC-MS offers the important advantages of not requiring derivatization and of being amenable to analysis of high-molecular-mass compounds (such as peptides and proteins) and very polar compounds (such as glucuronide-conjugated metabolites).

GC-MS analyses can be performed using electron ionization (EI) or chemical ionization (CI), followed by either full-scan mass analysis or selected-ion monitoring. Electron ionization is favoured in forensic toxicology because it generates detailed, compound-specific mass spectra, and because more than 100 000 electron ionization reference mass spectra have been published. However, chemical ionization can often provide better sensitivity, depending on the chemical characteristics of the analyte. Most drugs contain basic functional groups that are easily protonated in the gas phase within a chemical ionization source, resulting in a simple mass spectrum dominated by a large protonated molecule ($MH^+$) ion peak. Drugs that contain functional groups with high electron affinity, such as halogens, can be ionized selectively and with very high efficiency by electron-capture negative-ion chemical ionization.

**Figure 1** compares the mass spectra resulting from analysis of a cocaine-containing sample using electron ionization and positive-ion chemical ionization, and illustrates the features typical of these two types of mass spectra. The electron ionization

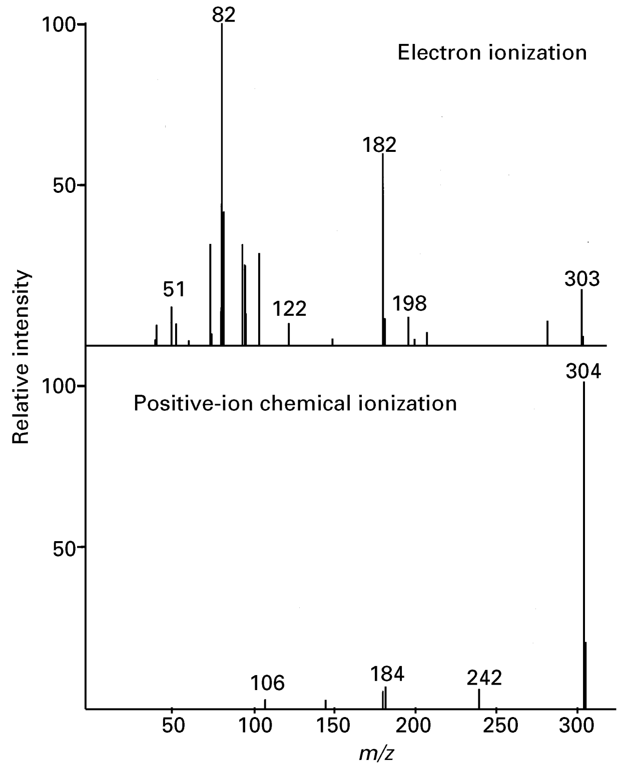



**Figure 1** Comparison of the electron ionization and positive-ion chemical ionization mass spectra of cocaine.

mass spectrum has a relatively low-intensity molecular ion ($M^+$) peak at *m*/*z* 303 and contains many fragment ion peaks. In contrast, the positive-ion chemical ionization mass spectrum shows a single intense peak (*m*/*z* 304) that corresponds to the protonated molecule ion ($MH^+$). The high sensitivity often afforded by chemical ionization is due to the way the analyte ions are concentrated at a single *m*/*z* value, rather than being distributed among many different *m*/*z* values.

Both full-scan recording and selected-ion monitoring are widely used in forensic toxicology. Each has advantages. A full-scan spectrum that closely matches a reference spectrum generally constitutes the most reliable identification of an analyte. However, monitoring just a few well-selected ions characteristic of an analyte is more sensitive than full-scan recording and can be a reliable identification if it is combined with a selective chromatographic separation.

## Workplace drug testing

The analysis of urine specimens from employees to determine whether they are inappropriately using drugs that can affect job performance falls under the

**Table 1** Initial test cutoffs

| *Drug class* | *Cutoff* (ng $mL^{-1}$) |
|---|---|
| Marijuana metabolites | 50 |
| Cocaine metabolites | 300 |
| Opiate metabolites | 300 |
| Phencyclidine | 25 |
| Amphetamine | 1000 |

**Table 2** Confirmation test cutoffs

| *Drug* | *Cutoff* (ng $mL^{-1}$) |
|---|---|
| 9-Carboxy-$\Delta^9$-THC[a] | 15 |
| Benzoylecgonine[b] | 150 |
| Morphine | 300 |
| Codeine | 300 |
| Phencyclidine | 25 |
| Amphetamine | 500 |
| Methamphetamine | 500 |

[a] The major urinary metabolite of $\Delta^9$-tetrahydrocannabinol, the psychoactive component of marijuana.
[b] The major urinary metabolite of cocaine.

broad category of forensic toxicology because of the legal consequences that can arise. In the United States, workplace drug testing is widely practised, and can be divided into 'regulated' and 'nonregulated' testing. 'Regulated drug testing' refers to programmes that fall under federal regulations, such as the 'Mandatory Guidelines for Federal Workplace Drug Testing Programs' issued in 1988. These 'mandatory guidelines' define how workplace drug testing must be performed, and include the requirement that the testing be performed by laboratories certified by the National Laboratory Certification Program. Another key component of the guidelines requires a two-tier testing process consisting of an initial series of immunoassays followed by confirmatory tests based on gas chromatography–mass spectrometry. For a sample to be reported as positive for a drug, the immunoassay screen must show the presence of a drug class at a concentration above the cutoff values shown in **Table 1** and the GC-MS confirmation assay must show a drug, or metabolite within the drug class, to have a concentration above the cutoff values listed in **Table 2**.

'Nonregulated' workplace drug testing refers to programmes that do not fall under federal regulations. These programmes may include testing for other psychoactive drugs such as benzodiazepines, barbiturates and lysergic acid diethylamide (LSD). Also, nonregulated drug testing programmes sometimes include analysis of other types of specimens, such as hair or sweat.

Federal regulations do not identify specific methods for performing GC-MS confirmations. However, certain mass spectrometric procedures have become widely adopted by laboratories that perform workplace drug testing. A laboratory must first demonstrate and document that each confirmation assay is specific for the target drug and that it is sufficiently sensitive to permit accurate measurement of the drug at its established confirmation cutoff.

Typically, an analytical batch of urine samples will include each of the following: (1) calibration standards with at least one standard at the analyte's cutoff concentration; (2) quality-control samples containing known concentrations of the analyte both above and below the analyte's cutoff; (3) a sample blank containing none of the analyte; and (4) the test samples submitted for confirmation. A specific amount of internal standard is added to each of these samples. Deuterium-labelled analogues, now commercially available for all common drugs of abuse, are generally the first choice as internal standards. After addition of the internal standard, each sample is extracted to separate the analyte from other components of the urine. This step may involve a simple liquid–liquid extraction with a water-immiscible organic solvent, extraction onto a solid adsorbent and elution with an organic solvent, or a combination of these procedures.

Most drugs and metabolites are chemically converted to more volatile and stable derivatives before GC-MS analysis. **Table 3** gives examples of derivatizations commonly used for GC-MS confirmation of drugs of abuse.

Before any test samples are analysed, the mass spectrometer's mass calibration is checked by

**Table 3** Examples of derivatization procedures for GC-MS analysis of common drugs of abuse

| *Drug group* | *Derivatization procedure* |
|---|---|
| Amphetamines | Acylation with acid anhydrides such as trifluoroacetic anhydride (TFAA), pentafluoropropionic anhydride (PFPA) or heptafluorobutyric anhydride (HFBA) |
| Opiates | Trimethylsilylation with bis(trimethylsilyl)trifluoroacetamide (BSTFA) or acylation with a fluorinated anhydride |
| 9-Carboxyl-$\Delta^9$-THC | Trimethylsilylation with BSTFA or alkylation with methyl iodide in dimethyl sulfoxide |
| Benzoylecgonine | Trimethylsilylation with BSTFA |
| Benzodiazepines | Trimethylsilylation with BSTFA |
| Barbiturates | Alkylation with trimethylanilinium hydroxide (TMAH) |

analysis of a reference material such as perfluorotributylamine (PFTBA), and the measured mass-to-charge (*m*/*z*) values for the major PFTBA ions are compared with the corresponding theoretical *m*/*z* values. A reference solution containing a known concentration of the analyte or analytes of interest is then injected into the GC-MS system to evaluate the instrument's performance. If the analyte's peak intensity, retention time and chromatographic peak shape are acceptable, analysis of a batch of samples can be initiated.

Fused-silica capillary gas chromatographic columns are universally used for analysis of drugs in forensic laboratories. Typically the columns are coated to a thickness of approximately 0.3 μm with relatively nonpolar methylsilicone or phenylmethylsilicone adsorptive phases and have internal diameters of 0.2–0.3 mm and are 10–30 meters in length. Extracted and derivatized samples injected into the GC-MS are most often detected by electron ionization and selected-ion monitoring (SIM) of three prominent analyte ions and two prominent internal standard ions. The analyte concentration is determined by reference to a calibration curve based on a plot of the ratio of the area of the analyte's quantitating ion to the area of the internal standard's quantitating ion versus analyte concentration. For a urine sample to be confirmed as positive for a drug, the analytical results must meet all of the following criteria:

1. The drug's measured concentration must be equal to, or greater than, the established cutoff concentration for the drug (see **Table 2**).
2. The ratios of peak areas for the monitored ions of the analyte and internal standard must fall within 20% of the corresponding peak area ratios for the calibration standard at the cutoff concentration.
3. The retention times for the analyte and the internal standard must be within 2% of those for the cutoff calibrator.
4. The chromatographic peaks for the analyte and the internal standard must be clearly separate from co-eluting compounds.

The analytical methods most widely practised in laboratories for workplace drug testing have been well validated and are highly reliable. However, some problems continue to be of concern. For example, ingestion of certain legitimate food products can result in a urine sample testing positive for a controlled drug. Poppy seeds, widely used in bakery products, contain low concentrations of morphine, and hemp oil contains low concentrations of $\Delta^9$-tetrahydrocannabinol, the psychoactive component of marijuana. Efforts to develop analytical methods for testing urine that would permit distinction between ingestion of these food products and nonmedical use of regulated drugs have so far been unsuccessful. However, methods are now available for determining whether urine found to be positive for methamphetamine reflects nonmedical use of methamphetamine or legitimate use of 'Vick's Inhaler', an over-the-counter medication which in the United States contains the *l*-enantiomer of methamphetamine. To make this distinction, the urine must be analysed by a method that can separate and identify each of the optical isomers of methamphetamine. This can be accomplished by derivatizing the extracted methamphetamine with a chiral derivatizing reagent such as *N*-trifluoroacetyl-L-prolyl chloride. This reaction forms separate diastereoisomers of *d*- and *l*-methamphetamine that are easily distinguished by GC-MS analysis.

## Testing of motor vehicle operators for drug impairment

Drug impairment of operators of motor vehicles (cars, buses, trains, airplanes, etc.) is a continuing threat to public safety. To counteract this threat there has been increased effort devoted to developing analytical methods for determining when an operator's impaired performance is due to use of psychoactive drugs. These analytical methods have also played an important role in investigations into how drugs of abuse, as well as some legitimate prescription drugs, can affect a person's ability to operate a motor vehicle safely. To better understand how drugs can affect human performance, controlled clinical laboratory experiments are required. These experiments are designed to record self-reported impairment and to detect physiological, perceptual and cognitive drug-induced changes that may impact the subject's performance. The most useful clinical-study designs incorporate collection of biological samples at predetermined times for measurement of the drug concentration. Here mass spectrometry is used extensively to quantify the studied drug and its metabolites. Drug concentrations can then be compared to the physiological and performance measures to determine whether a dose–response relationship exists for the drug being evaluated. These clinical studies provide invaluable data for assisting in interpreting cases with legal implications.

Mass spectrometry is also used to analyse biological samples collected from subjects suspected of operating a motor vehicle under the influence of a drug. In contrast to workplace drug testing, where

urine is the body fluid normally analysed, blood is the preferred body fluid for determining impaired performance. This is because drug concentrations in blood have been found to correlate better with the level of impairment at the time the blood was collected than drug concentrations in urine, which can vary widely depending on the subject's degree of hydration and other variables. For some drugs, collection and analysis of saliva can also be an effective method for determining impairment.

Alcohol is the drug most frequently associated with driving impairment, but its presence in the body is measured by a breathalyser or by analysis of blood using procedures that do not require the use of mass spectrometry. Analysis of blood or saliva for other drugs normally includes an initial screening procedure based on a series of immunoanalyses or by a chromatographic screening method, followed by confirmation and quantitative measurement of the drug by mass spectrometry in combination with gas or liquid chromatography.

The frequency with which specific drugs are detected in impaired motor vehicle operators varies from one country or region to another. In the United States, marijuana is the second most frequently detected drug in impaired drivers, with cocaine probably the third most frequently detected. Benzodiazepines have been shown to affect fundamental driving skills such as coordination and reaction time, and are reported to be frequently detected in drivers suspected of being drug impaired. Amphetamines and chemically related sympathomimetic drugs are sometimes used to reduce fatigue in drivers, but large doses of these drugs may lead to greater risk-taking and aggressive driving. Other drugs that have been detected in motor vehicle operators suspected of driving under the influence (DUI) include opiates, barbiturates, antihistamines and antidepressants.

The confirmation and quantitation of drugs identified in DUI cases is a challenge to the forensic toxicologist because drug concentrations are typically lower in blood than in urine or in tissue samples. The determination of marijuana use by DUI suspects is particularly challenging. Although $\Delta^9$-tetrahydrocannabinol (THC) and its major metabolites may be detected in blood following recent use, the concentrations of these analytes are usually in the low nanogram per milliliter range. THC and its metabolites have active hydroxyl or carboxyl groups that require derivatization prior to GC-MS analysis. Examples of derivatization procedures are shown in **Table 3**. GC-MS with electron ionization may be used for the analysis of THC and its metabolites, but this combination has limited use in forensic cases because it lacks sufficient sensitivity to routinely measure these analytes in blood samples. Better sensitivity can be achieved using GC-MS with negative-ion chemical ionization following conversion of the cannabinoids to derivatives containing electron-capturing substituents. Sub-nanogram- per -milliliter detection limits have also been achieved by trimethylsilylation of the extracted cannabinoids followed by gas chromatography coupled to a tandem mass spectrometer operated in the positive-ion chemical ionization mode.

Cocaine is rapidly metabolized, primarily to benzoylecgonine and ecgonine methyl ester, and this conversion can continue after the blood is drawn unless an esterase inhibitor such as potassium fluoride is added to the blood immediately. This step is important because the blood concentration of the parent cocaine is a better indicator of the degree of impairment than the concentrations of the inactive metabolites.

Mass spectral identification and quantification of benzodiazepines in blood is complicated by the large number of prescription benzodiazepine drugs available to the public. Treatment of blood extracts with strong acid has been used to convert benzodiazepines to benzophenones, which are more easily identified by GC-MS analysis. However, more recently developed methods using GC-MS with negative-ion chemical ionization or LC-MS with electrospray ionization take less time and offer better sensitivity.

## Postmortem toxicology

Postmortem forensic toxicology involves the analysis of drugs and poisons in body fluids and tissues collected from victims that have died under suspicious circumstances. Postmortem forensic toxicology is a key part of a death investigation in that it can provide information addressing the following questions and issues: (1) What drugs or toxins are present in the body? (2) Are the concentrations of the drugs or toxins in the lethal range? (3) Could drugs have caused impairment resulting in a fatal accident?

Currently in the United States there are no federally mandated regulations governing postmortem forensic toxicology. However, several forensic toxicology accreditation programmes have been established, and the Society of Forensic Toxicology and the American Academy of Forensic Sciences have published guidelines for forensic toxicology laboratory procedures and performance.

Unlike workplace drug testing, in which urine samples are analysed for a relatively small number of drugs and drug metabolites, postmortem forensic

toxicology often involves analysis for a wide variety of drugs and poisons in many different specimens, including gastric contents, vitreous humor, blood, liver, spleen, muscle and brain tissue. Analysis of blood is particularly important because drug concentrations in blood that are high enough to cause death are known for most drugs. However, drug concentrations in blood are generally lower than in urine or in tissues.

Before extraction and analysis, tissues must be homogenized. Water or buffer solutions such as 0.1 M sodium phosphate, pH 6, are added to the tissue sample prior to homogenization. It is important to record the weight of the tissue and the volume of the fluid in which the tissue is homogenized. This information will be used to express the amount of drug per gram of tissue.

Extraction efficiency and sample cleanliness can be a particular challenge in the analysis of postmortem samples because the samples are often collected from decomposed bodies. Products from decomposition can reduce the extraction efficiency and can produce interfering peaks. Also, tissue homogenates contain lipid material that must be separated from the drug analyte prior to GC-MS analysis. Basic drugs can be efficiently separated from lipid material by back-extraction. In this procedure the drug is extracted from the tissue homogenate into a water-immiscible organic solvent and then back-extracted into a dilute acid solution while the neutral lipid material remains in the organic solvent. The dilute acid solution is then made basic and the drug is re-extracted into an organic solvent.

Another cleanup procedure is to reconstitute the initial extract residue in a mixture of hexane and ethanol–water (80:20 v/v). Most drugs will partition into the ethanol–water layer and lipid compounds will partition into the hexane layer. Precipitation of proteins from the tissue homogenates prior to extraction is also helpful for improving cleanliness of the extract and increasing the extraction efficiency. Reagents that precipitate proteins include acetonitrile, methanol and trichloroacetic acid. Digestion of tissue samples with proteolytic enzymes such as subilisin A can also enhance the recovery of drugs in the extraction procedure.

Following extraction, the analyte is derivatized, if necessary, and analysed by GC-MS. As discussed previously, electron ionization continues to be favoured. However, use of chemical ionization is increasing because it often provides better sensitivity. Also, a chemical ionization mass spectrum will frequently provide valuable complementary information. For example, **Figure 2** shows electron ionization and chemical ionization mass spectra for two similar drugs, amitriptyline and cyclobenzaprine. The electron ionization mass spectra for the two drugs are virtually identical owing to the prominence of the common fragment ion at *m*/*z* 58. The positive-ion chemical ionization mass spectra show protonated molecule ion peaks that differ by two daltons, thereby permitting identification of the two drugs.

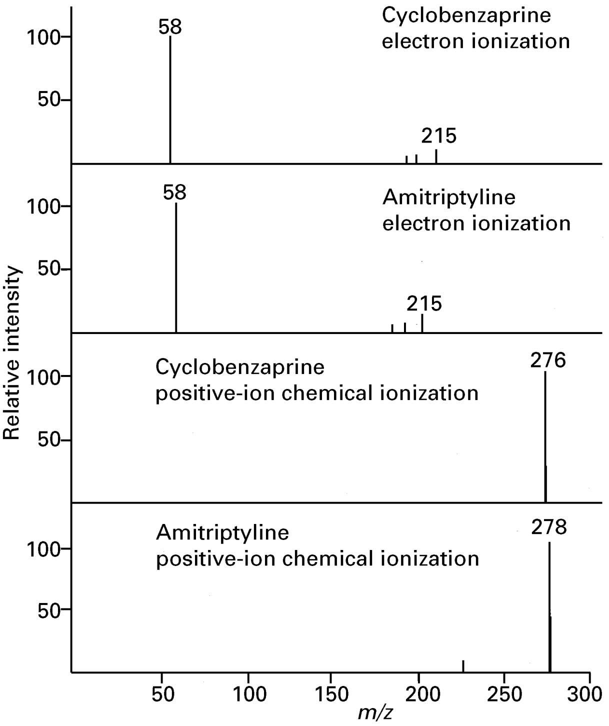



**Figure 2** Comparison of the electron ionization and positive-ion chemical ionization mass spectra of cyclobenzaprine and amitriptyline.

Of course, the relevant questions in a death investigation cannot be answered by mass spectrometric analysis and other analytical data alone. Final interpretation should always include careful consideration of all the information relating to the case.

## Future developments

The combination of gas chromatography and electron ionization mass spectrometry will continue to be an important analytical tool for identification and quantification of drugs and toxins in most forensic toxicology laboratories. However, the use of liquid chromatography with mass spectrometry is certain to increase, and is likely to become the most widely used method for identification and quantification of drugs and other toxins in forensic toxicology samples. Many variations of liquid

chromatography, such as capillary electrophoresis (CE) and capillary electrochromatography (CEC), have been successfully coupled to mass spectrometers and will provide the forensic toxicologist with an array of techniques that will permit many analytical tasks to be performed with improved sensitivity, accuracy and efficiency.

*See also:* **Biochemical Applications of Mass Spectrometry; Chemical Structure Information from Mass Spectrometry; Chemical Ionization in Mass Spectrometry; Chromatography-MS, Methods; Forensic Science, Applications of Atomic Spectroscopy; Forensic Science, Applications of IR Spectroscopy; Hyphenated Techniques, Applications of in Mass Spectrometry; Ion Trap Mass Spectrometers; Isotopic Labelling Mass Spectrometry; Medical Applications of Mass Spectrometry; MS-MS and MS$^n$; Negative Ion Mass Spectrometry, Methods; Quadrupoles, Use of in Mass Spectrometry.**

## Further reading

Baselt RC and Cravey RH (eds) (1995) *Introduction to Forensic Toxicology*, 4th edn. Foster City, CA: Chemical Toxicology Institute.

Baselt R (ed) (1989) *Advances in Analytical Toxicology*, Vol 2. Chicago, IL: Year Book Medical Publishers.

Curry AS and Thomas CC (1976) *Poison Detection in Human Organs*, 3rd edn. Springfield, IL: Charles C. Thomas.

Karch SB (ed) (1998) *Drugs Abuse Handbook*. Boca Raton, FL: CRC Press.

Moffat AC (ed) (1986) *Clarke's Isolation and Detection of Drugs*. London: The Pharmaceutical Press.

Swotinsky RB (1992) *The Medical Review Officer's Guide to Drug Testing*. New York: Van Nostrand Reinhold.

Yinon J (ed) (1995) *Forensic Applications of Mass Spectrometry*. Boca Raton, FL: CRC Press.

Yinon J (ed) (1987) *Forensic Mass Spectrometry*. Boca Raton, FL: CRC Press.

# Forestry and Wood Products, Applications of Atomic Spectroscopy

**Cathy Hayes**, Trinity College, Dublin, Ireland

Copyright © 1999 Academic Press

| ATOMIC SPECTROMETRY |
| --- |
| Applications |

## Introduction

Analytical methods of atomic spectroscopy have been used in forestry and wood product research since their earliest development. Nowadays, almost all of the spectroscopic techniques available are employed in the analysis of metals and trace elements in diverse samples of industrial and environmental origin. The techniques that find most regular application include flame atomic absorption spectroscopy (F-AAS), graphite furnace atomic absorption spectroscopy (GF-AAS), inductively coupled plasma atomic emission spectroscopy (ICP-AES) and, occasionally, also direct current plasma atomic emission spectroscopy (DCP-AES). In many applications F-AAS is a sufficiently sensitive and precise technique; however, in the analysis of some environmental samples for trace elements (forest soils, plant material and water) where concentrations may be very low (of the order of 100 ng mL$^{-1}$) the greater sensitivity of GF-AAS and ICP/DCP-AES is required. In considering the applications of atomic spectroscopy to forestry and wood products it is worthwhile remembering that the analytical method available for use is often determined by the volume and variety of analyses carried out in individual laboratories.

## Atomic spectroscopic techniques in use

F-AAS is probably the most widely used analytical method in forestry and wood product research for two reasons. Firstly, it is a long established technique, machines are relatively cheap and widely available and many of the analyses routinely performed by this method have now been optimized. Secondly, the concentrations of many nutrient elements that are important for forest growth, which are present in soil and water (e.g. Ca, Mg, Al, Fe, K, Si) and in wood treated with metal-containing preservatives (Cu, Cr, As, Zn), are well within the limits of detection (100 ppm) of the technique. Environmental samples of this kind are also usually available in large quantities. F-AAS offers a precise, rapid

method of analysis, and automation is possible for large numbers of samples.

GF-AAS, also known as electrothermal atomization atomic absorption spectroscopy (ET-AAS) or heated graphite atomizer atomic absorption spectroscopy (HGA-AAS), is widely available in research analytical laboratories where trace element work is carried out. GF-AAS is possibly the most sensitive method available in the atomic spectrometry range but machines are approximately twice as expensive as F-AAS set-ups. The detection limit for determinations is typically in the order of 10–100 ng $mL^{-1}$. Elements in soil/water samples most commonly determined by GF-AAS might include Ca and Mg, but also, in particular, toxic Cd, Cr, Cu, Ni and Pb. Details of the furnace program (the temperature, ramp, hold and purge gas flow characteristics of the drying, charring, atomization and cleaning steps) should be reported in published results more often than they are. The heightened sensitivity of this technique requires the careful use of analytical reagents to keep background contamination to a minimum. Matrix matching between samples and calibration solutions is critical in GF-AAS, and matrix modification (through the addition of ammonium nitrate, magnesium nitrate or palladium salts) may be necessary to prevent the loss of analytes before and in the atomization stage.

When DCP/ICP-AES becomes cheaper and therefore more widely available, the advantages offered by simultaneous multielement determinations combined with detection limits comparable to those of F-AAS (100 ng $mL^{-1}$), the technique will undoubtedly find greater application in the analysis of forestry and wood samples.

## Forestry

Forestry research embraces both the monitoring of environmental impacts on forest ecosystems as well as the impact of forestry on surrounding landscapes, in particular on water catchments. Studies carried out have analysed soil to monitor the effects of acid rain on forest nutrition in relation to observed disease susceptibility in trees from different geographical locations. It is postulated that immobilization of essential plant nutrients such as calcium and magnesium occurs in soil where toxic aluminium becomes increasingly mobile owing to the reduced soil pH attributed to acid rain. Routine monitoring of water catchments that surround forestry follows the level of movement of both nutrient elements (N and P) and other forest exudates (Mn, Cu, Zn, Sr and Ba) in run-off. Both F-AAS and GF-AAS are used in these analyses and offer acceptable sensitivity and precision.

## Wood and wood products

Timber harvested from commercial forests has several possible end users. Good quality timber (i.e. timber with good growth characteristics) is sawn to produce timber for the construction industry. Non-construction grade timber can end up being processed to make wood pulp and ultimately paper or used as sawn timber, joinery, fencing rails and posts, transmission poles, bulkheads and pilings. Recently, poorer quality timber has been used more and more to make value added commodities such as board materials like medium density fibre board (MDF), oriented strand board (OSB) and particle/chip board. Each of these areas generates the need for further monitoring for various reasons.

The awareness of the need to monitor paper and pulp mill effluent is not novel. The reduction, and even elimination, in the use of chlorine in the bleaching of pulp has focused environmental concern on the other elements present in wood pulp. Pulp and paper mill effluent may contribute to elevated metal concentrations in receiving waters and sediments. Consequently, the wood pulp industry, along with statutory monitoring bodies, make regular use of sensitive and quantitative techniques such as DCP-AES and GF-AAS in attempts to reduce metal levels.

The implication of wood dust and the metals from machined treated wood in the increased instance of carcinoma in people working with wood also requires the use of GF-AAS and F-AAS in the health and safety analysis of dusts from wood related work places (joinery manufacturers, sanding and sawing operations, particle board and plywood industries).

## Wood preservatives

The bulk of analytical work performed in the wood sector is associated with wood preservatives. Most temperate timbers are perishable and therefore require treatment with toxic combinations of elements to achieve protection from biodeterioration owing to attack by fungi, insects, termites and, in marine situations, wood-boring crustaceans and molluscs.

The most widely used preservatives are the class of inorganic waterborne salts that contain Cu, Cr, As, B and Zn. A list of commonly used preservatives of this type is presented in **Table 1**.

## Preservative penetration and retention

The necessity of monitoring the levels of preservatives (waterborne inorganic salts as well as organic solvents and creosote) in treated timber arises from the fact that such timber is supposed to be treated to specified levels or retentions (kg preservative salt $m^{-3}$ wood), depending on the biological hazard to

which it is expected to be exposed in service. There are five classes of biological hazard based on the expected exposure to wetting and these are shown in **Table 2**.

Technical standards exist which prescribe the levels of preservative salt necessary in timber exposed to a given hazard class. Treatment levels are defined in terms of required penetration depths and must be

**Table 1** Common inorganic waterborne salt type wood preservatives; active ingredients and nominal compositions as well as the recommended in-service situation are listed

| *Preservative* | *Active ingredient* | *Nominal % composition* | *Principal use/country of use* |
|---|---|---|---|
| Chromated zinc chloride (CZC) | Zinc as ZnO | 80 | Out of ground contact |
| | Chromium as $CrO_3$ | 20 | USA, Germany |
| Copper–chromium–arsenic (CCA) | Copper as $CuSO_4{\cdot}5H_2O$<br>as CuO | 33[a], 35[b]<br>18.1[c], 19.6[d], 18.5[e] | Ground contact, freshwater and marine situations |
| | Chromium<br>as $Na_2Cr_2O_7{\cdot}2H_2O$<br>as $CuO_3$ | <br>41[a], 45[b]<br>65.5[c], 35.3[d], 47.5[e] | Worldwide |
| | Arsenic as $As_2O_5{\cdot}2H_2O$<br>as $As_2O_5$ | 26[a], 20[b]<br>16.4[c], 45.1[d], 34[e] | |
| Copper–chromium–boron (CCB) | Copper as $CuSO_4{\cdot}5H_2O$ | 36 | Out of ground contact |
| | Chromium as $Na_2Cr_2O_7{\cdot}2H_2O$ | 40 | Continental Europe |
| | Boron as $H_3BO_3$ | 24 | |
| Ammoniacal copper arsenate (ACA) | Copper as CuO | 49.8 | Ground contact, freshwater and marine situations |
| | Arsenic as $As_2O_5$ | 50.2 | USA |
| Fluor chrome arsenate phenol (FCAP) | Fluoride as F | 22 | Out of ground contact |
| | Chromium as $CrO_3$ | 37 | |
| | Arsenic as $As_2O_5$ | 25 | Continental Europe, |
| | Dinitrophenol | 16 | USA |
| Ammoniacal copper, zinc arsenate (ACZA) | Copper as CuO | 50 | Out of ground contact |
| | Zinc as ZnO | 25 | |
| | Arsenic as $As_2O_5$ | 25 | Scandinavia, USA |

The country of use, is not intended as an exhaustive list.
[a] CCA type 1 (UK).
[b] CCA type 2 (UK).
[c] CCA type A (USA).
[d] CCA type B (USA).
[e] CCA type C (USA).

**Table 2** Biological hazard classes for timber exposed in service as defined by the European Standard EN 335-1:1992

| *Hazard class* | *General service situations* | *Description of wetting in service* | *Occurrence of biological agencies*[a] | | | |
|---|---|---|---|---|---|---|
| | | | *Fungi* | *Beetles*[b] | *Termites* | *Marine borers* |
| 1 | Above ground, covered (dry) | None | – | U | L | – |
| 2 | Above ground, covered (risk of wetting) | Occasionally | U | U | L | – |
| 3 | Above ground, not covered | Frequently | U | U | L | – |
| 4 | In contact with ground or fresh water | Permanently | U | U | L | – |
| 5 | In salt water | Permanently | U | U | L | U |

[a] U, Universally present within Europe. L, Locally present within Europe.
[b] The risk of attack can be insignificant, according to specific service situations.

matched to individual wood species' treatability. The more pertinent European and American standards covering these technical requirements are listed in **Table 3**.

Meeting the requirements of these standards generates an enormous amount of routine analysis in the preservative treatment plants that are responsible for correctly treating the timber, and also in research laboratories examining biodeterioration of timber and in industrial development laboratories. Inorganic waterborne elements are readily determined using F-AAS and this is the method recommended in the technical standards governing the analysis of inorganic waterborne preservatives, the European and American examples of which are included in **Table 3**.

## Naturally occurring levels of elements

The naturally occurring levels of the elements most frequently analysed in forestry soil and foliage samples and wood/wood product samples are presented in **Table 4**. The normal manner in which samples containing these elements are prepared so as to release the element as an analyte for trace element determination is indicated in the table. Where several analytical techniques are possible this is indicated in the footnotes. The most commonly recommended F-AAS and GF-AAS settings for the same elements are shown in **Table 5** along with potential interferences and their control. From these two tables it is readily appreciated that, for most determinations involving forestry related or wood samples, contamination is not usually problematic and background levels can frequently be ignored.

## General instrument settings

Sample instrument settings for F-AAS, GF-AAS and ICP-AES analyses are shown in **Table 6**. However, instrument manufacturers, recommendations should always be followed.

## Analytical procedures

### General comments

In the use of AAS or AES, quality control can be maintained by the use of standard reference materials in analyses. Although expensive, the practice of including reference samples provides extra control to the data generated. The heterogeneity of elemental distributions in ecological samples requires adequate field sampling to generate a statistically significant number of samples on which accurate determinations can be made. In many instances the pooling of samples after initial processing (grinding, drying, sieving etc.) is possible (and has been reported for

**Table 3** National standards referring to preservative treatment of wood and also to the laboratory methods used to analyse the preservative content of wood and treating solutions (including F-AAS)

| *Standard No* | *Title* |
|---|---|
| EN 335-1: 1992 | *Durability of wood and wood-based products.* Hazard classes of wood and wood-based products against biological attack. Part 1. Classification of hazard classes |
| EN 335-1: 1992 | *Durability of wood and wood-based products.* Hazard classes of wood and wood-based products against biological attack. Part 2. Guide to the application of hazard classes to solid wood |
| EN 335-3: 1996 | *Durability of wood and wood-based products.* Definition of hazard classes of biological attack. Part 3. Application to wood based panels |
| EN 350-2: 1994 | *Durability of wood and wood-based products.* Natural durability of solid wood. Part 2. Guide to natural durability and treatability of selected wood species of importance in Europe |
| EN 351-1: 1996 | *Durability of wood and wood-based products.* Preservative-treated solid wood. Part 1. Classification of preservative penetration and retention |
| EN 599-1: 1997 | *Durability of wood and wood-based products.* Performance of preventive wood preservatives as determined by biological tests. Part 1. Specification according to hazard class |
| BS 5666-3: 1991 | Methods of analysis of wood preservatives and treated timber. Part 3. Quantitative analysis of preservatives and treated timber containing copper/chromium/arsenic formulations |
| BS 4072: 1978 | Specification for wood preservation by means of copper/chrome/arsenic compositions |
| AWPA Standard P5-98 (1998) | Standards for waterborne preservatives |
| AWPA Standard C1-98 (1998) | All timber products – preservative treatment by pressure processes |
| AWPA Standard A11-93 (1993) | Standard method for analysis of treated wood and treating solutions by atomic absorption spectroscopy |

EN = European standard; BS = British Standard; AWPA = American Wood Preservers' Association Standard.

soils and herbaceous plant matter) but the decision must be made whether to sacrifice the spatial pattern for a mean concentration. The practice is not recommended for wood. Matrix matching of calibration and test solutions is critical, especially where the more sensitive GF-AAS is concerned. If acid digests have been carried out on samples the same acids must be used in calibration standards. Analytical grade chemicals should be used in all determinations and, where specified, distilled and deionized water may be preferred over distilled water.

## Applications of atomic spectroscopy in forestry and wood products

### Fine estuarine sediments

Particulate organic matter (POM) is largely associated with fine sediment fractions, i.e. silts (2–20/50 µm) and clays (< 2 µm). Inorganic copper–chromium–arsenic salts (CCA) leached from submerged or floating treated wood in coastal situations also tend to be associated with the organic matter in this fraction. The settlement of fine particulates (POM, silts, clays) where Cu, Cr and As may be present is dependent on the flushing characteristics of the water body as well as the age of the structure from which preservative components are leaching. In slow currents where recently treated timber is present, Cu, Cr and As can be detected several metres from treated wooden emplacements such as pilings, floating jetties and bulkheads. It is important that experimental design takes these facts into account when deciding on a sampling regime.

Fine sediments should be allowed to settle out in volumetric cylinders before subsampling so as to reduce the risk of contamination from metal sieves. Sediment is then dried (150°C) and a subsample wet ashed in a 3:1 mix of nitric acid and perchloric acid ($HNO_3$–$HClO_4$). *Caution*: perchloric acid is highly explosive when it comes into contact with organic-matter. Therefore raw organic samples should always be pre-mixed in warmed nitric acid *before* the addition of perchloric acid. Once perchloric acid has been added, the digest mix should not be allowed to boil dry. Only experienced personnel should handle and store this reagent.

Determinations of Cu, Cr and As can be made on the filtered liquid fraction, using the recommended settings in the manual accompanying the atomic spectrometer used. Elemental levels determined by F-AAS methods from fine coastal sediments have been found in the range 25–2000 µg Cu per g sediment, 40–180 µg Cr per g sediment and 10–100 µg As per g sediment. The possibility of carrying out hydride generation before the determination of arsenic is an option on many machines.

### Algae and animal tissues

Plant and animal material (e.g. algae, crustaceans, molluscs and bivalves) collected to determine the transfer of metals from the environment to biota within habitats can be collected at field sites and frozen until laboratory analysis is possible. Defrosted or fresh material can then be oven dried (150°C) before a subsample is selected and wet ashed in a 3:1 mix of nitric and perchloric acid ($HNO_3$–$HClO_4$). Note the cautions expressed above for the handling of this mixture.

The filtered leachate is analysed in F-AAS using the settings recommended by the manufacturer of the machine for each element under determination. In cases where the elemental concentrations are likely to be very low (e.g. control samples from 'clean' sites) it may be necessary to carry out analysis using GF-AAS (or ICP-AES if available) with a simpler acid digestion. Typical levels of CCA encountered in coastal biota where preserved wood is submerged are in the range 5–10 µg $g^{-1}$ (As) 5–15 µg $g^{-1}$ (Cr) and 60–150 µg $g^{-1}$ (Cu).

### Trees and herbaceous plant material

Foliage samples are collected fresh and then stored cold or frozen until analysis. Dried/defrosted samples are oven-dried at 60–80°C and then ground to a fine powder in a suitable mill. Beware: steel mills may lead to contamination by Fe, Mn, Mo and Co while other mills can lead to contamination by Cu and Zn. Ceramic ball mills may lead to contamination by Al, B, Ca, K or Na.

Ground samples are dissolved in a wet digestion procedure on a heating block using a 3:1:2 mix of $HNO_3$–$H_2SO_4$–$H_2O_2$ (typically 1–5 g in 1.5:0.5:1.0 mL of acid mix). Lanthanum (as lanthanum chloride, 1000 µg $mL^{-1}$) is added as a releasing agent to prevent phosphate interference in the analysis of Ca and Mg by F-AAS. Aluminium levels may need to be determined by the more sensitive GF-AAS, in which case the chances of contamination need to be minimized by using smaller volumes of acid in the digestion mix (1 mL $HNO_3$, 0.2 mL $H_2SO_4$, 0.2 mL $H_2O_2$).

### Forest soils

Soils are collected in the field using stainless steel soil probes which sample to depths of 40–50 cm, maintaining the stratigraphic integrity of the sample.

**Table 4** List of common soil minerals and methods for sample preparation and analysis; Naturally occurring levels (in mineral soils, freshwater, etc.) are also included. With the permission of Blackwell Scientific Publications data based on Allen SE (1989) In: Allen SE (ed) *Chemical Analysis of Ecological Materials*, 2nd edn. Oxford: Blackwell Scientific Publications.

| *Element* | *Fusion type* | *Acid digestion[a] (as an alternative to fusion)* | *Sample size(g) (to give 100 mL solution)* | *Mineral soil ($\mu g\ g^{-1}$)* | *Organic soil ($\mu g\ g^{-1}$)* | *Soil extracts ($\mu g\ g^{-1}$)* | *Plant materials ($\mu g\ g^{-1}$)* | *Animal tissues ($\mu g\ g^{-1}$)* | *Rain water ($\mu g\ L^{-1}$)* | *Fresh water ($\mu g\ L^{-1}$)* |
|---|---|---|---|---|---|---|---|---|---|---|
| Al | $Na_2CO_3$ or NaOH | | 0.1 | 1–12% | 0.05–0.5% | 10–200 (mg per 100 g) | 0.01–0.1% | 0.001–0.02% | 2–100 | 100–2000 |
| B | $Na_2CO_3$ + 30 mL water, then 20% v/v $H_2SO_4$ until melt dissolves | | 0.25 | 5–50 | 2–20 | 0.2–5 | 10–80 | 0.4–2 | 100 | 10–500 |
| Ca | Any fusion method | HF–$HClO_4$ | 0.1 (less for calcareous soils) | 0.5–2% (excl. calcareous soils) | 0.1–0.5% | 10–200 (mg per 100 g) (excl. calcareous soils) | 0.3–2.5% (high in calcareous species) | 0.03–0.3% (excl. bones) | 0.1–3 mg $L^{-1}$ | 0.1–100 mg $L^{-1}$ |
| Cl | $Na_2CO_3$ + $NaNO_3$ to improve flux Dissolve melt in water Adjust pH to 8 with 0.1 M $H_2SO_4$ | | 1.0 | 0.004–0.08% | 0.03–0.2% | 1–40 (excl. saline soils) | 0.04–0.4% | 0.2–0.8% | 1–25 mg $L^{-1}$ | 2–100 mg $L^{-1}$ |
| Co | Any fusion method | HF–$HClO_4$ | 0.5 | 1–50 | 0.2–1 | 0.05–4 | 0.1–0.6 | 0.02–0.1 | 0.05–0.5 | 0.5–2.5 |
| Cu | $Na_2CO_3$ or $KHSO_4$ | HF–$HClO_4$ | 0.25 | 5–80 | 6–40 | 0.1–3 | 2.5–25 | 10–100 | 0.2–2 | 2–50 |
| Fe | $Na_2CO_3$ | HF–$HClO_4$ | 0.1 | 0.5–10% | 0.02–0.5% | 50–1000 | 40–500 | 100–400 | 5–150 | 50–1000 |
| Mg | Any fusion method | HF–$HClO_4$ | 0.1 | 0.2–2% | 0.05–0.3% | 4–50 mg per 100 g | 0.1–0.5% | 0.05–0.2% | 0.1–2.0 mg $L^{-1}$ | 0.5–20 mg $L^{-1}$ |
| Mn | Any fusion method If solution is pink–brown add a few drops of $H_2SO_3$(5%w/v $SO_2$) and boil before dilution to volume | HF–$HClO_4$ | 0.1 | 200–2000 | 50–500 | 5–500 | 50–1000 | 5–50 | 0.4–3 | 1–80 |
| Mo[b] | | | 0.6 or > | 0.5–10 | 0.05–0.5 | 0.01–0.2 | 0.1–0.8 | 0.03–0.3 | 0.05–0.3 | 0.1–2 |
| N[c] | | | 0.5–1.0 | 0.1–0.5% | 0.5–1.5% | $NH_4^+$-N 0.2–3 mg per 100 g $NO_3$-N 0.1–2 mg per 100 g | 1–3% | 4–10% | $NH_4^+$-N 0.1–0.8 mg $L^{-1}$ $NO_3$-N 0.05–0.4 mg $L^{-1}$ | $NH_4^+$-N 0.1–2 mg $L^{-1}$ $NO_3$-N 0.05–3 mg $L^{-1}$ |
| P | Fusion followed by boiling with $H_2SO_4$ to convert into orthophosphate | HF–$HClO_4$ provided all HF removed or<br>digest with $HNO_3$ (2 mL) + $HClO_4$(1 mL) + $H_2SO_4$(0.5 mL) or<br>digest with $H_2SO_4$–$H_2O_2$ for 2 h | 0.1<br>0.05<br>0.2 | 0.02–0.15% | 0.01–0.2% | 0.3–8 mg per 100 g | 0.05–0.3% | 0.3–4% (excl.bone) | 2–50 g $L^{-1}$ | 5–500 g $L^{-1}$ |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| K | Any fusion method | HF–$HClO_4$ | 0.1 | 0.3–2% | 0.02–0.2% | 5–50 mg per 100 g | 0.5–5% | 0.3–1.5% | 0.1–1 mg $L^{-1}$ | |
| | Fusion reagents may give high blanks | | | | | | | | | |
| Si | NaOH | | 0.1 | 20–60% | 0.2–2% | | 0.05–0.5% | 0.05–0.3% | 0.01–0.1 | 0.5–8 mg $L^{-1}$ |
| Na | | HF–$HClO_4$ | 0.1 | 0.1–2% | 0.02–0.1% | 2–20 mg per 100 g | 0.02–0.3% | 0.2–1.0% | 0.5–15 mg $L^{-1}$ | 2–100 mg $L^{-1}$ |
| S | Mix soil with 2.5 g $Na_2CO_3$, preheat to 450°C for 30 min before fusion | | 0.5 | 0.03–0.3% | 0.03–0.4% | 100–500 | 0.08–0.5 | 0.2–0.8% | 0.4–4 mg $L^{-1}$ | 2–150 mg $L^{-1}$ |
| | Addition of 0.2 g $NaNO_3$ or 0.1 g $Na_2O_2$ improves the flux | | | | | | | | as $SO_4^-$-S | as $SO_4^-$-S |
| Ti | $Na_2CO_3$ | HF–$HClO_4$ | 0.5 | 0.2–1.5% | 0.01–0.08% | 0.1–2 | 0.4–8 | 0.01–0.1 | 0.03–0.2 | 0.4–4 |
| | Evaporate solution obtained and add 30 ml 50% v/v $H_2SO_4$. Evaporate to white fume stage. Add 10 ml 50% v/v $H_2SO_4$ and filter | Not in presence of high soil P | | | | | | | | |
| Zn | Any fusion method | HF–$HClO_4$ | 0.1 | 20–300 | 10–50 | 1–40 | 15–100 | 100–300 | 1–15 | 5–50 |
| Cr[d] | | | | | | 10–200 | 0.05–0.5 | 0.01–0.3 | | 0.1–0.5 |
| As[d] | | | | | | 0.5–30 | 0.1–1 | 0.1–0.5 (accumu-lates in certain organisms) | | 0.2–1 |
| Pb[d] | | | | | | 2–20 | 0.05–3 | 0.1–3 | | 1–20 |
| Hg[e] | | | | | | 0.1–1 | 0.005–0.1 | 0.03–0.3 | | 0.3–3 |

Note: for the naturally occurring levels all measurements are on a dry weight (µg $g^{-1}$) basis unless otherwise stated. Dissolution methods (fusion and acid digestion) for the determination of total element concentrations in soil are shown.

[a] Perchloric acid is explosive if handled incorrectly. Consult safety manuals for information on correct handling.

[b] This element presents difficulty in determination. Consult specialist texts for methodology.

[c] Various chemical methods for the determination of $NH_3$-N, $NO_3$-N, $NO_2$-N and tot. N exist. Consult specialist texts.

[d] Various methods exists. Consult specialist tests for methodology.

[e] Fusion/dry ashing would cause volatilization of Hg, other methods preferred. Consult specialist texts for methodology.

**Table 5** Summary of flame techniques applicable to the analysis of ecological materials. With the permission of Blackwell Scientific Publications, data based on Allen SE (1989) In: Allen SE (ed) *Chemical Analysis of Ecological Materials*. 2nd edn. Oxford: Blackwell Scientific Publications.

| *Element* | *Wave-length (nm)* | *Mode* | *Flame* | *Suitable ranges (ppm)* | *Potential interferences* | *Control of interference* | *Comments* |
|---|---|---|---|---|---|---|---|
| Al | 309.3 | A | $N_2O–C_2H_2$ | 0–20 | Fe, Ti, Ca, HOAc | Include Fe in standards | Very hot flame essential EA possible |
| Ca | 422.7 | E | $Air–C_2H_2$ $O_2–C_2H_2$ | 0–40 | Al, P, Si, S | Add releasing agent (Sr or La salt) | A more sensitive than E |
| | | A | $Air–C_2H_2$ | 0–10 0–40 | As above | As above | |
| | | A | $N_2O–C_2H_2$ | 0–40 | Na, K | Include ionization buffer | |
| Co | 240.7 | A | $Air–C_2H_2$ | 0–50 | Fe if high | Extract | Prior concn normally reqd. |
| | | EA | | 0–1 | | | EA preferable |
| Cr | 357.9 | EA | – | 0–0.1 | Ca, Fe | Standard compensation | EA usually essential |
| Cu | 324.8 | A | $Air–C_2H_2$ | 0–5 | None significant | | Prior concn desirable EA preferable |
| Fe | 248.3 | A | $Air–C_2H_2$ | 0–20 | Si | Add $CaCl_2$ | |
| Hg | 253.7 | A | Flameless | 0–1 | None | | Hg vapour is produced which is passed into an abs cell in the light path |
| Mg | 285.2 | A | $Air–C_2H_2$ | 0–3 | Ca, Al, P, Si | Add releasing agent (Sr or La salt) | |
| Mn | 279.5 | A | $Air–C_2H_2$ | 0–2 | None significant | | Samples and standards to be in same valency state |
| | | | | 0–20 | | | |
| Mo | 313.3 | A | $Air–C_2H_2$ | 0–3 | | | Prior concn with organic extraction |
| K | 766.5 | | $Air–C_2H_2$ } Air–propane | 0–10 0–100 | Na, Ca, $H_2SO_4$, HCl | Include in standards Dilute sample | A has no marked advantage over E |
| Na | 589.0 | | $Air–C_2H_2$ } Air–propane | 0–5 0–20 | Ca, $H_2SO_4$, HCl | Include in standards Dilute sample | E preferable |
| Ni | 232.0 | EA | – | 0–0.2 | None serious | | EA essential |
| Pb | 217.0 | EA | – | 0–0.5 | None serious, Cl possible | | EA desirable |
| | | A | $Air–C_2H_2$ | 0–10 | | | Prior concn necessary |
| Zn | 213.9 | A | $Air–C_2H_2$ | 0–1 | None significant | | |

E=emission, A=atomic absorption, EA=electrothermal atomization.

Before analysis soil may be split into different levels, most metals are associated with the organic fraction in the active root zone 10–15 cm down. Soil should be air-dried for several days (~7) and then ground in a mortar and pestle and passed through stainless steel (2 mm) and 1 mm mesh sieves. After thorough mixing, soils can be stored in polythene bags until analysed.

Available nutrient elements are extracted from ~2 g subsamples which are treated with ammonium acetate (100 mL) and placed on a mechanical shaker for 1 h and then the resulting solution is gravity filftered. Concentrated HCl can be added (2–3 drops) to decrease the bacterial activity in the sample. Lanthanum (as lanthanum chloride at 1000 μg $mL^{-1}$) is added to the filtrate so as to prevent the interference of phosphate in the determination of Ca and Mg. Soil element concentrations are usually determined by F-AAS. Certain micronutrients such as Mo occur at levels near the limits of detection of F-AAS and are best determined by GF-AAS. Naturally occurring levels of Cr, Pb and Ni are also best determined by the more sensitive GF-AAS. Only chromium absorbance experiences potential

**Table 6** Sample instrumental setting for analyses in F-AAS (A) and GF-AAS (B)

(**A**) **F-AAS** Settings for the measurement of Cu, Cr and As in CCA treated wood preservatives (BS 4072:1974)

| *Parameter* | *Copper* | *Chromium* | *Arsenic* |
|---|---|---|---|
| Wavelength (nm) | 324.8 | 357.9 or 429.0 | 193.7 or 197.2 |
| Slit width (mm) | 0.08 | 0.06 to 0.10 | 0.30 |
| Lamp current (mA) | 4 | 8 | 7 |
| Scale expansion | up to × 10 | up to × 10 | up to × 10 |
| Burner | 10 cm acetylene | 10 cm acetylene | 10 cm propane |
| Burner height (cm) | 1.0 | 0.5 | 1.4 |
| Acetylene flow rate (P of 0.7 kg $cm^{-2}$)/$cm^{-3}$ $min^{-1}$ | 1000 | 1800 | – |
| Air flow rate (P of 2.1 kg $cm^{-2}$)/L $min^{-1}$ | 5 | 5 | – |
| Hydrogen flow rate (P of 0.7 kg $cm^{-2}$)/$cm^{-3}$ $min^{-1}$ | – | – | 1800 |
| Argon flow rate (P of 2.1 kg $cm^{-2}$)/L $min^{-1}$ | – | – | 5 |

(**B**) **GF-AAS** Settings for the determination of aluminium in conifer foliage. Reproduced from Lee CE, Cox JM, Foster DM, Humphrey HL, Woosley RS and Butcher D (1997) Determination of aluminium, calcium, and magnesium in Fraser fir (*Abies fraseri*) foliage from five native sites by atomic absorption spectrometry: the effect of elevation upon nutritional status. *Microchemical Journal* **56**: 236–246.

Analytical wavelength: 394.4 nm
Chemical modifier: magnesium nitrate (2 µg as Mg)
Bandpass: 0.40 nm
Background correction: continuum source

*Furnace program*

| | *Dry step* | *Char step* | *Atomization step* | *Clean step* |
|---|---|---|---|---|
| Temperature (°C) | 110 | 1100 | 2300 | 2500 |
| Ramp (s) | 20 | 20 | 0[a] | 1 |
| Hold(s) | 15 | 30 | 4 | 4 |
| Purge gas flow | Low | Medium | None | High |

These conditions represent improved optimization compared with conditions recommended in manufacturers manual for Al determination

[a] Maximum power heating.

interference, from Ca and Fe, which is controlled by several means, including the addition of 2% ammonium chloride ($NH_4Cl$) to samples and standards.

## Solid wood

The analysis of solid wood for elements (both naturally occurring and as preservative components) is a relatively straightforward determination using spectroscopic techniques. Some values for naturally occurring elements in hardwoods and softwoods are presented in **Table 7**, the levels of toxic metals reported for spruce are noteworthy. The methodology detailed here for the analysis of wood preserving elements is widely accepted.

Only Ca, Mg and phosphate are reported to occur in appreciable amounts in wood and experiments carried out in the presence of excess Ca, Mg and phosphate showed no interference in typical digest mixes [0.5 M sulfuric acid and 0.3% (w/v) sodium sulfate] in the determination of Cu, Cr, As and Zn. Phosphate does not interfere with As determination, indicating that an insignificant amount of the phosphate present is ionic. Good extractives are not reported to cause interference in the determination of Cr and Cu and they can enhance the As signal unless the burner height is maintained between 1.3 and 1.5 cm (1.4 cm is frequently used). There is no interelement interference in the determination of Cu, Cr and As solutions. Preservative compounds are extracted from treated wood prepared as thin sections (0.2 mm thick), sawdust (produced by sawing a suitably representative part of the treated wood) or wood flour (prepared by hammer milling wood chips).

The decomposition step is carried out in a mixture of 0.5 M sulfuric acid and 100 volume hydrogen peroxide. The copper(II) sulfate, potassium/sodium dichromate and arsenic pentoxide in CCA preservatives undergo fixation to form potassium/sodium sulfate by-products. Any potential interference in the chromium and arsenic signals from these compounds in the leachate is overcome by an excess of sodium

**Table 7** Naturally occurring metal ion concentrations in wood of various kinds; values on a mg per kg dry weight basis

| *Source* | *Na* | *Al* | *Fe* | *Mn* | *Zn* | *Cu* | *Pb* | *Cd* | *Cr* | *B* |
|---|---|---|---|---|---|---|---|---|---|---|
| Spruce | 76–133 | 3.2–8.3 | 1.6–3 | 32–271 | 5.7–22.9 | 0.7–1.3 | 0.3–1.5 | 0.2–0.5 | | |
| Birch | 96 | 11 | 21 | 76 | | | | | | |
| Pine | 35 | 16 | 24 | 76 | | | | | | |
| Hardwoods | | | | | | 0.2–2.1 | | | 0.2–4.6 | 1.5–7.2 |
| Softwoods | | | | | | 0–3.0 | | | 0.2–4.8 | 0.8–6.8 |

sulfate (at 0.3% w/v) added to samples and calibration standards. Wood that has been submerged in water for a prolonged period can experience leaching of preservative salts and so the concentration of potassium/sodium salts in treated wood can be very variable.

The analytical sensitivity for arsenic has been improved from 0.60 μg $mL^{-1}$ to 0.10 μg $mL^{-1}$ by the use of an argon–hydrogen flame (instead of air–acetylene) and a 10 cm propane burner (instead of a 10 cm acetylene burner) analysed at a wavelength of 197.2 nm, and this flame is recommended in arsenic analysis. The argon–hydrogen combination is highly explosive and should only be used by trained personnel.

Standard curves can be prepared in the range of 5–35 μg $mL^{-1}$ Cu and 10–70 μg $mL^{-1}$ Cr and As for preservative concentrations in the region of 3% w/v. The acid decomposition described is suitable for samples in the range of 3–8 g.

The comminution of wood probably represents the most time-consuming part of the sample preparation procedure. The production of wood flour requires the tedious reduction of wooden samples to a small enough size (*ca.* matchstick) before feeding into a hammer mill. The hammer mill itself can be a cause of contamination in the form of Cr or Cu from the hammers or the chamber though this is probably insignificant compared with the levels of Cu and Cr in the treated wood. Though wood flour represents the maximum surface area to volume ratio for the action of the acid decomposition, the use of thin wood sections (0.2 mm thick) or sawdust is acceptable providing they are obtained from areas of treated wood deemed representative of the whole, i.e. not from the ends or surfaces, which can be over-treated with respect to the rest of the timber.

## Water

The requirements of the analysis of water from various sources is one area of environmental research that is driving efforts to improve the limit of detection of spectroscopic analysis. Elements are frequently in very dilute concentrations and the presence of dissolved salts (chiefly sodium chloride and magnesium salts in marine samples) can offer some interference in sensitive analyses.

The pulping of hardwood and softwood during the manufacture of paper generates waste water that contains elements such as Na, Al, Mn, P, Fe, Mn, Zn, and metals such as Cd, Cu, Cr, Ni, Hg and Pb. These elements and metals are released in mill effluent to receiving waters and elevated concentrations can build up in fine sediments at some distance (up to 12 km) from the effluent source.

Water samples can be handled in several ways but an important precaution is to prevent the sorption of metals to sample bottles. This is achieved by the addition of 0.02 vol.% suprapure nitric acid. Both plasma AES and GF-AAS methods are suitable for the determination of the elements mentioned. Wavelengths (nm) considered more sensitive for elements in organic matrices analysed by DCP-AES include 568.82 (Na), 308.215 (Al), 259.94 (Fe), 403.076 (Mn) and 202.548 (Zn).

The determination of Cr, Ni and Pb require the addition of $NH_4H_2PO_4$ to samples for matrix modification. Most elements in pulp effluent occur in μg $L^{-1}$ concentrations (As, Ba, Co, Li, Sr, Sn) while some, such as Na, Ca, Mg, Si and K (naturally present in wood or added to pulp during processing in various chemicals), occur in mg $L^{-1}$ concentrations. Mass spectrometry may be employed for certain elements that are present in very low concentrations < 1 μg $L^{-1}$ (Ag, Mo, Sb, Tl) and the combination of elements determined by this method offers a characterization of pulp and paper mill effluent in water. Mercury is determined by cold vapour AAS where the detection limit is ~0.1 μg-Hg $L^{-1}$.

## Air

The concentration of metals in the air of wood-processing factories has recently become a more prominent health and safety issue and the analysis of metal-containing dust samples is carried out by AAS and AES methods.

For the analysis of metal elements in the air of factories that process wood which has been treated with

wood preservatives (such as Cu, Cr and B) the finest dust particles are collected by suction onto glass fibre filter paper. The concentration of dust (in mg $m^{-3}$ air) is calculated by dividing the mass of the dust (mass of filter paper after filtering minus the mass of filter paper before filtering) by the volume of air filtered.

Elements are determined in dust by first drying the dust sample over silica gel. Fine dust samples are very mobile and difficulties may be experienced transferring samples. A dust sample is leached in 20 mL $HNO_3$ (1 M) mixed with 1 mL $H_2O_2$ (30%) in an Erlenmeyer flask and shaken on a mechanical shaker for 20 h. After filtration, Cu and Cr determinations may be made in F-AAS as described above.

Hexavalent chromium (the more toxic form of Cr) can be determined separately by extraction from dried dust with diphenylcarbazide followed by photometric analysis. Boron is extracted from silica gel dried dust by leaching in the $HNO_3$ (1 M) mix described above. After 20 h in the solution on a mechanical shaker the sample is filtered, extracted with 2-ethylhexa-1,3-diol and curcumin and analysed by photometry.

To date, research has shown that in less than half of the workplaces examined the concentration of dust in workplace air was above the recommended threshold (2 mg dust per $m^{-3}$ of air), and in fewer than a third of cases the levels were five times the recommended threshold value. It has also been shown that the concentration of metals in the dust did not have its origin exclusively in preservative treated wood but that there was a significant load of metals from external air. The concentration of hexavalent chromium in dust was in line with natural levels in air and well below the safe threshold levels of 50 000 ng $m^{-3}$.

## Future developments

Analytical improvements in atomic spectroscopy are most likely to affect the analysis of water and air samples related to environmental impacts of wood-products and the health and safety aspects of working with wood and wood preservatives. Atomic absorption spectrometry is now such a thoroughly tried and tested analytical technique that regular major advances are unlikely.

In spite of the advent of ICP based techniques, F-AAS and GF-AAS continue to be used routinely in many laboratories for the analysis of waters, soils and plant material. As might be expected, most of the significant developments in the AES field have been associated with ICP-AES, and especially with aspects of sample introduction to the ICP, though reports of novel developments in wood related research are not plentiful.

*See also:* **Atomic Spectroscopy, Historical Perspective; Environmental and Agricultural Applications of Atomic Spectroscopy; Inductively Coupled Plasma Mass Spectrometry, Methods; Quantitative Analysis; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Allen SE (1989) In: Allen SE (ed) *Chemical Analysis of Ecological Materials*, 2nd edn. Oxford: Blackwell Scientific Publications.

Holmbom B, Harju L, Lindholm J and Gröning A-L (1994) Effect of a pulp and paper mill on metal concentrations in the receiving lake system. *Aqua Fennica* **24**: 93–100.

Lambert M (1969) The determination of copper, chromium and arsenic in preservative treated timber by the method of atomic absorption spectrophotometry. *Journal of the Institute of Wood Science* **4**: 27–36.

Lee CE, Cox JM, Foster DM, Humphrey HL, Woosley RS and Butcher D (1997) Determination of aluminium, calcium, and magnesium in Fraser fir (*Abies fraseri*) foliage from five native sites by atomic absorption spectrometry: the effect of elevation upon nutritional status. *Microchemical Journal* **56**: 236–246.

Williams AI (1972) The use of atomic absorption spectrophotometry for the determination of copper, chromium and arsenic in preserved wood. *Analyst* **97**: 104–110.

# Fourier Transformation and Sampling Theory

**Raúl Curbelo**, Bio-Rad, Cambridge, MA, USA

Copyright © 1999 Academic Press

| FUNDAMENTALS OF SPECTROSCOPY |
| --- |
| Theory |

In Fourier transform spectroscopy (FTS), the output of the detector can be represented as a value that is a function of the independent variable, usually time. For a given finite measurement time $T$, it can be considered a transient signal that lasts $T$ seconds. This transient is the interferogram in FT-IR or the free induction decay in FT-NMR. In either case, this transient function contains all the spectral information being measured, but the spectral amplitude for each frequency is encoded in the detector signal and in general is not discernible directly. The Fourier transform is a mathematical operation that can be used to change this function of time into an amplitude that is a function of frequency, which is a representation of the desired spectrum.

Signal processing has always been a critical aspect in spectroscopy and especially in FTS. The generalized use of computers as components in spectrometers to implement the Fourier transform and/or other digital signal processing (DSP) tasks requires, as a first step, that the signals used be discrete amplitude, discrete time-sampled representations of continuous amplitude, continuous time (analogue) signals, such as the output of a detector. The sampling process must meet certain minimum requirements for the resulting sample sequence to contain all the information of the original signal. The output of the sampling process is a continuous amplitude, discrete time representation of the original signal. To obtain a discrete amplitude sequence, the amplitude of each sample is quantized with finite precision by a digital to analogue converter (ADC) and is then represented by a number. The resulting sequence of numbers is the input to the DSP.

## The Fourier series

In the artificial case that the spectrometer measures the same sample multiple times, the successive transients will be identical, and the entire sequence will be a periodic signal of period $T$.

The Fourier series provides a representation of the signal as a sum of sinusoidal components of different frequencies. For a periodic signal of period $T$ it is easy to visualize that the component frequencies will be $f_k = k/T$, where $k = 0, 1, 2, 3 \ldots$, because these sinusoids are the only ones periodic in $T$. These components are called harmonics of the signal. The sinusoid of frequency $f_k$ is the $k$th harmonic.

Sinusoids are used to study and represent signals in linear time invariant systems because an input sinusoid generates an output from a linear time invariant system that is a sinusoid of the same frequency. (See Mason and Zimmerman in the Further reading section). The Fourier series representation of a periodic signal $v(t)$ with period $T$ is

$$v(t) = A_0 + \sum_{k=1}^{\infty} \{A_k \cos(2\pi k/T)t + B_k \sin(2\pi k/T)t\} \quad [1]$$

for $k = 1 \ldots \infty$, and $-\infty < t < \infty$; $2\pi k/T$ is the radian or angular frequency of the $k$th harmonic and $A_k \cos(2\pi k/T)t$ and $B_k \sin(2\pi k/T)t$ are the harmonic components of $v(t)$. $A_k$ is the amplitude of the cosine component of frequency $k/T$. $A_0$ is the average of $v(t)$ over a period $T$, and $A_k$ and $B_k$ are the averages of $2v(t)\cos(2\pi k/T)t$ and $2v(t)\sin(2\pi k/T)t$ respectively, also over a period $T$.

For an even function of $t$, that is $v_e(t) = v_e(-t)$ the above simplifies to

$$v_e(t) = A_0 + \sum_{k=1}^{\infty} A_k \cos(2\pi k/T)t \quad [2]$$

For an odd function of $t$, that is $v_o(t) = -v_o(-t)$ we have

$$v_o(t) = \sum_{k=1}^{\infty} B_k \sin(2\pi k/T)t \quad [3]$$

The Fourier series can be written in a more compact form with the use of the complex exponential function $\exp(i\omega t) = \cos(\omega t) + i\ \sin(\omega t)$, where $\omega$ is the radian frequency $2\pi f$, and $\omega_0 = 2\pi/T$. Then

$$v(t) = \sum_{k=-\infty}^{\infty} V_k \exp(ik\omega_0 t) \quad [4]$$

where $V_k$ is the average of $v(t)\exp(-ik\omega_0 t)$ over the period $T$:

$$V_k = 1/T \int_{-T/2}^{T/2} v(t)\exp(-ik\omega_0 t)\,dt \qquad [5]$$

$V_k$ has the dimension of voltage, and in the general case will be complex. This expression leads to the Fourier transform, also called the Fourier integral. If we assume $v(t)$ is periodic with period $T$, and different than zero for $|t| < T/2$, we can define

$$V(k\omega_0) = \int_{-T/2}^{T/2} v(t)\exp(-ik\omega_0 t)\,dt \qquad [6]$$

$V(k\omega_0)$ has dimensions of voltage per unit frequency. It is independent of $T$, and does not change when $T$ increases, while the Fourier coefficients $V_k$ become smaller as $T$ increases.

$$V_k = (1/T)V(k\omega_0) = (\omega_0/2\pi)V(k\omega_0) \qquad [7]$$

$$\begin{aligned} v(t) &= (1/T)\sum_{k=-\infty}^{\infty} V(k\omega_0)\exp(ik\omega_0 t) \\ &= \sum_{k=-\infty}^{\infty} V(k\omega_0)\exp(ik\omega_0 t)(\omega_0/2\pi) \end{aligned} \qquad [8]$$

For $T$ approaching infinity, $\omega_0$ tends to zero, and if we define $\omega = k\omega_0$, in the limit

$$v(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} V(\omega)\exp(i\omega t)\,d\omega \qquad [9]$$

and

$$V(\omega) = \int_{-\infty}^{\infty} v(t)\exp(-i\omega t)\,dt \qquad [10]$$

where $V(\omega)$ is the Fourier transform of $v(t)$. The term $V(\omega)$ is also called the frequency spectrum of $v(t)$ or simply the spectrum. This is the Fourier representation of a single pulse, because we defined it by selecting one pulse and moving all others in the sequence to infinity.

In the case of the periodic signal, the coefficients of the Fourier series have the dimension of voltage. The Fourier transform of a pulse $v(t)$, $V(\omega)$, has the dimensions of voltage multiplied by time, and we can think of it as a voltage density spectrum with a dimension of voltage per unit frequency. Therefore, the area under $V(\omega)$ has the dimension of voltage.

**Table 1** lists some of the properties of the Fourier transform. The functions $v(t)$ and $u(t)$ are defined for $-\infty < t < \infty$, and their respective Fourier transforms $V(\omega)$ and $U(\omega)$ are defined for $-\infty < \omega < \infty$. In property 6 of **Table 1** the convolution is defined for $-\infty < \xi < \infty$, where $\xi$ is the angular frequency. It is used here to distinguish it from the independent variable $\omega$ of the result.

## Discrete time signals and sampling

In the development of the Fourier transform above we have assumed that the signal $v(t)$ was a continuous time signal defined for $-\infty < t < \infty$. These signals are also commonly called analogue signals.

Discrete time signals may be generated by a discrete time process or may be the result of a sampling process on a continuous time signal.

The common method to obtain a discrete representation of a continuous time signal is to take periodic samples of the signal, and create a sequence $v[n]$ from the continuous time signal $v_c(t)$ defined by the equality

$$v[n] = v_c(nT_s) \qquad \text{for} \quad -\infty < n < \infty \qquad [11]$$

**Table 1** Some properties of the Fourier transform

| | Function | Fourier transform |
|---|---|---|
| | $v(t)$ $\quad -\infty < t < \infty$ | $V(\omega)$ $\quad -\infty < \omega < \infty$ |
| | $u(t)$ | $U(\omega)$ |
| 1 | $A\,u(t) + B\,v(t)$ | $A\,U(\omega) + B\,V(\omega)$ |
| 2 | $v(-t)$ | $V(-\omega)$ |
| 3 | $v(t/a)$ | $a\,V(a\omega)$ for $a$ real positive |
| 4 | $v(t-t_0)$ | $V(\omega)\exp(-i\omega t_0)$ |
| 5 | $dv(t)/dt$ | $i\omega V(\omega)$ |
| 5 | $\int_{-\infty}^{\infty} v(t)u(\tau - t)\,dt$ | $V(\omega)U(\omega)$ |
| 6 | $v(t)u(t)$ | $\frac{1}{2\pi}\int_{-\infty}^{\infty} V(\xi)U(\omega-\xi)\,d\xi = W(\omega)$ |
| 7 | $\int_{-\infty}^{\infty} \|v(t)\|^2\,dt$ | $\int_{-\infty}^{\infty} \|V(\omega)\|^2\,d\omega$ |

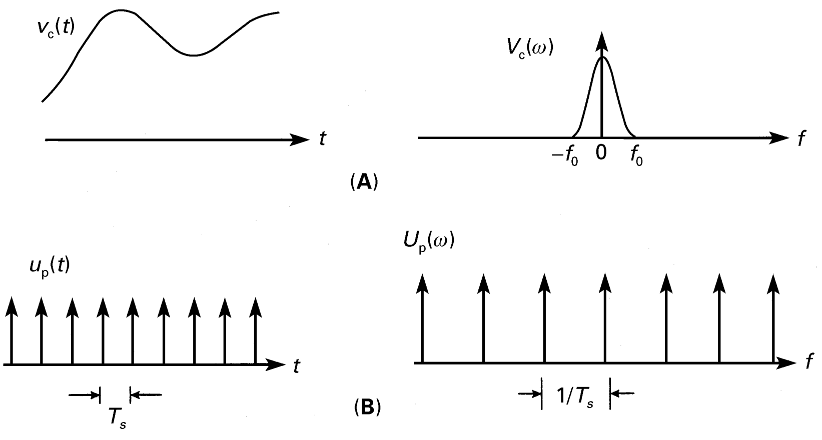



**Figure 1** Sampling the continuous signal with a periodic impulse train.

where $T_s$ is the sampling period, and its inverse, $f_s = 1/T_s$, is the sampling frequency, in samples per second.

The sampling operation is not invertible in general, that is given $v[n]$, it is not possible in general to reconstruct $v_c(t)$ because many continuous signals can be represented by $v[n]$. The sampling theorem due to Nyquist removes this uncertainty by imposing a restriction on the signal $v_c(t)$.

**Figure 1A** shows a portion of the continuous time signal $v_c(t)$ defined for $-\infty < t < \infty$, and the Fourier transform of $v_c(t)$, the spectrum $V_c(\omega)$. We have chosen $v_c(t)$ such that $V_c(\omega) = 0$ for $|f| > f_0$. ($|\omega| > 2\pi f_0$). **Figure 1B** shows a portion of the periodic train of unit impulses $u_p(t)$, defined for $-\infty < t < \infty$, and a portion of the Fourier transform of $u_p(t)$, the spectrum $U_p(\omega)$, $-\infty < f < \infty$.

It is convenient to represent the sampling process in two steps, as shown in **Figure 2**. First the multiplication of $v_c(t)$ with a periodic train of unit impulses $u_p(t)$, which results in the train of impulses $v_s(t)$ for $-\infty < t < \infty$, and second, the conversion of the train of impulses to a sequence of samples $v[n]$.

The product of the continuous signal $v_c(t)$ with the periodic train of impulses generates the train of impulses $v_s(t)$ that is zero except at integer multiples of $T_s$.

**Figure 2A** shows a portion of the train of impulses $v_s(t)$ that has a spectrum $V_s(\omega)$ for $-\infty < \omega < \infty$, where $\omega$ is the radian frequency $2\pi f$. This spectrum can be obtained using property 6 of the Fourier transform in **Table 1**, which shows that the transform of the product of two functions is the convolution of the transforms. The convolution of $V_c(\omega)$ with the periodic train of impulses $U_p(\omega)$ of **Figure 1** results in periodically repeated copies of $V_c(\omega)$, separated by $1/T_s$, a portion of which is shown in **Figure 2A**.

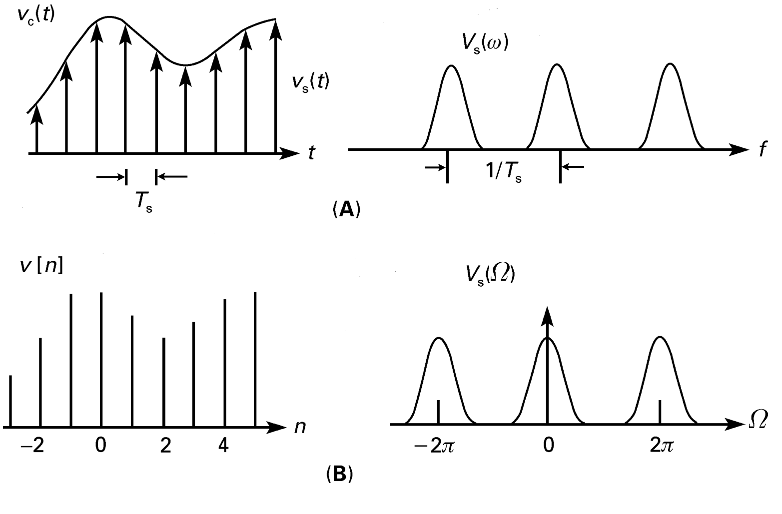



**Figure 2** Impulse train converted into a discrete time sequence.

## Sampling theorem

The sampling theorem due to Nyquist (see Mason and Zimmermann in the Further reading section) can be visualized from **Figure 2A**. The sampling frequency, $f_s = 1/T_s$, must be at least twice the highest frequency $f_0$ contained in the signal, otherwise the spectrum of $V_s(\omega)$ will be corrupted by the overlap of adjacent replicas of $V_c(\omega)$ and it will not permit the recovery of the original signal. When the above condition is met, the spectrum of $V_s(\omega)$ contains all the information of the original spectrum $V_c(\omega)$, therefore the original signal is recoverable from $V_s(\omega)$, for example by passing the sampled signal through a low pass filter that rejects frequencies above $f_0$ but passes all the frequency components below $f_0$.

Sampling at a rate higher than $1/T_s$ has the undesirable effect of increasing the computational cost of any digital signal process that follows. One exception is the case of the recovery of the original signal with a simple filter. The trade off between sampling rate and filter complexity is more common in the case of narrow band signals and analogue filters.

The discrete sequence $v[n]$ for $-\infty < n < \infty$ in **Figure 2B** is indexed on the integer variable $n$, which does not have information about the sampling rate and represents the samples of $v_c(t)$ with finite numbers instead of the areas of impulses as in $v_s(t)$. The spectrum of $v[n]$ is $V_s(\Omega)$ for $-\infty < \Omega < \infty$, the discrete Fourier transform of $v[n]$, where $\Omega$ is in radians.

### Discrete time signals and the Fourier transform

A sequence that represents a sinusoid has the general form for $-\infty < n < \infty$:

$$v[n] = A\exp[\mathrm{i}(\Omega_0 n + \phi)] \qquad [12]$$

By analogy with the continuous time case, $\Omega_0$ is called the frequency of the complex sinusoid, and $\phi$ is called the phase. Here $n$ is a dimensionless integer number. Therefore, the dimension of $\Omega_0$ is radians. One can specify the units of $\Omega_0$ in radians per sample, for a better analogy with the continuous time case.

A more important difference with the continuous time case becomes apparent when we consider the frequency $(\Omega_0 + 2\pi)$. In this case

$$\begin{aligned} v[n] &= A\ \exp[\mathrm{i}(\Omega_0 + 2\pi)n] \\ &= A\exp(\mathrm{i}\Omega_0 n)\ \exp(\mathrm{i}2\pi n) \\ &= A\ \exp(\mathrm{i}\Omega_0 n) \end{aligned} \qquad [13]$$

It is easy to see that sinusoids and complex exponential sequences that differ in frequency by a multiple of $2\pi$ are identical. For these sequences, we only need to consider frequencies in an interval of length $2\pi$, such as $-\pi$ to $\pi$, or 0 to $2\pi$.

Discrete time periodic sequences are also different from continuous periodic signals in the behaviour of the frequency of the harmonics. A periodic discrete time sequence can be defined as

$$v[n] = v[n + N] \qquad [14]$$

for $-\infty < n < \infty$, where the period $N$ has to be an integer. A periodic discrete time sinusoid sequence,

$$A\cos(\Omega_0\, n + \phi) = A\cos(\Omega_0\, n + \Omega_0\, N + \phi) \qquad [15]$$

requires that $\Omega_0 N$ must be a multiple of $2\pi$. Therefore

$$\Omega_0\, N = 2\pi\, m \qquad [16]$$

where $m$ is an integer. The same result applies to complex exponential sequences.

A further conclusion from Equations [13] and [16] is that, for a periodic sequence of period $N$, the only frequencies that can be different are

$$\Omega_m = 2\pi\, m/N,\ \text{for}\ m = 0, 1, \cdots, N-1 \qquad [17]$$

Furthermore, low and high frequencies appear differently than in continuous signals. In the discrete time sinusoid $v[n] = A\ \cos(\Omega_0 n + \phi)$ as $\Omega_0$ increases from 0 to $\pi$, $v[n]$ cycles faster, but as $\Omega_0$ increases from $\pi$ to $2\pi$, $v[n]$ cycles slower and slower, and the sequence for $\Omega_0 = 2\pi$ is the same as the one for $\Omega_0 = 0$.

In general we can define the discrete Fourier transform of the sequence $v[n]$ as

$$V(\Omega) = \sum_{n=-\infty}^{\infty} v[n]\ \exp(-\mathrm{i}\Omega n) \qquad [18]$$

that is valid whenever $v[n]$ meets

$$\sum_{n=-\infty}^{\infty} |v[n]| < \infty \qquad [19]$$

**Table 2** Some properties of the discrete Fourier transform

| | *Function* | *Discrete Fourier transform* |
|---|---|---|
| | $v(n)$ $-\infty < n < \infty$ | $V(\Omega)$ $-\pi < \Omega < \pi$ |
| | $u(n)$ | $U(\Omega)$ |
| 1 | $a\,v[n] + b\,u[n]$ | $a\,V(\Omega) + b\,U(\Omega)$ |
| 2 | $v[n-m]$ ($m$ an integer) | $\exp(-i\Omega m)V(\Omega)$ |
| 3 | $\exp(i\Omega_0 n)$ | $V(\Omega-\Omega_0)$ |
| 4 | $nv[n]$ | $i\,dV(\Omega)/d\Omega$ |
| 5 | $v[n].u[n]$ | $\frac{1}{2\pi}\int_{-\pi}^{\pi} V(\theta)U(\Omega-\theta)\,d\theta$ |
| 6 | $\sum_{n=-\infty}^{\infty} \lvert v[n]\rvert^2$ | $\frac{1}{2\pi}\int_{-\pi}^{\pi} \lvert V(\Omega)\rvert^2\,d(\Omega)$ |

Using similar reasoning to that used for the continuous time case, we can write

$$v[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} V(\Omega)\,\exp(i\Omega n)\,d\Omega \qquad [20]$$

Equations [18] and [20] form a discrete Fourier transform pair that relate the discrete sample space with the frequency space.

**Table 2** lists some of the properties of the discrete Fourier transform. The functions $v[n]$ and $u[n]$ are defined for $-\infty < n < \infty$, and their respective discrete Fourier transforms $V(\Omega)$ and $U(\Omega)$ are defined for $-\pi < \Omega < \pi$. In property 5 of **Table 2** the convolution is defined for $-\pi < \theta < \pi$ where $\theta$ is the angular frequency. It is used here to distinguish it from the independent variable $\Omega$ of the result.

## List of symbols

$A_0$ = average of $v(t)$ over $T$; $A_k$ = amplitude of cosine components of frequency $k/T$; $f_k$ = sinusoid frequency ($k$th harmonic); $T$ = signal time period; $v_c(t)$ = continuous time signal; $v_c(\pm)$ = continuous time signal; $v[n]$ = discrete sequence; $v(t)$ = periodic signal; $V(\omega)$ = Fourier transform of $v(t)$; $\theta$, $\xi$ = angular frequency; $\phi$ = phase of complex sinusoid; $\omega$ = the radian frequency $2\pi f$; $\Omega_0$ = frequency of a discrete sinusoid.

*See also:* **FT-Raman Spectroscopy, Applications; Raman and IR Microspectroscopy.**

## Further reading

Bracewell RN (1986) *The Fourier Transform and its Applications*, 2nd edn. New York: McGraw-Hill.

Carslow HS (1952) *Introduction to the Theory of Fourier's Series and Integrals*, 3rd edn. New York: Dover Publications.

Dym H and McKean HP (1972) *Fourier Series and Integrals*. New York: Academic Press.

Jerry AJ (1997) The Shannon sampling theorem – its various extensions and applications: a tutorial review. *Proceedings of the Institute of Electrical and Electronic Engineers* **65**: 1565–1595.

Mason S and Zimmermann HJ (1960) *Electronic Circuits, Signals and Systems*. New York: John Wiley and Sons.

Nyquist H (1928) Certain topics in telegraph transmission theory. *American Institute of Electrical Engineers Transactions* 617–644.

Oppenenheim AV and Schafer RW (1975) *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

Oppenenheim AV and Schafer RW (1989) *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

Papoulis A (1962) *The Fourier Integral and its Applications*. New York: McGraw-Hill.

Shannon CE (1949) Communications in the presence of noise. *Proceedings of the Institute of Radio Engineers* **37**: 10–21.

# Fragmentation in Mass Spectrometry

**Hans-Friedrich Grützmacher**, Universität Bielefeld, Germany

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
| --- |
| Theory |

## Fragmentation reactions of radical cations

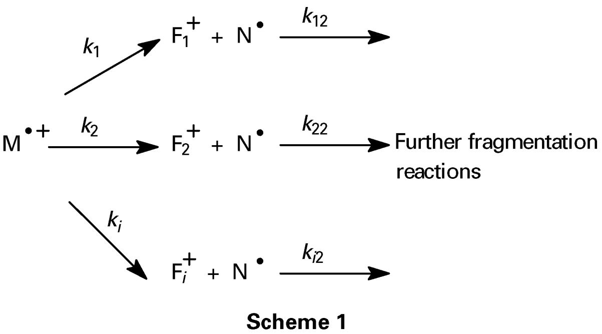
Organic molecules are usually even-electron species with spin paired valence electrons. The one-electron oxidation of these molecules, which is performed by oxidation with certain transition metal salts, by electrochemical oxidation, by photoinduced single-electron transfer (PET), by ionization using high energy radiation or electron impact, results in odd-electron radical cations. Sometimes these organic radical cations form stable salts which can be isolated; a well-known example is 'Wurster's red', the bromide of the *N*,*N*-dimethyl-1,4-diaminobenzene radical cation. Mostly, however, organic radical cations are highly reactive, and the study of radical cations as reactive intermediates of organic reactions is an important and active field of research both with respect to reaction mechanism and to synthetic application. In the context of mass spectrometry, the radical cations are generated by electron impact ionization (EI) or by photoionization (PI) in the ion source of the mass spectrometer as energetically excited molecular ions with a rather broad range of excess energy. Because of the low pressure (usually < $10^{-6}$ mbar) and the short residence time of the ions within the ion source deactivation of the excited molecular ions by collision does not occur and radiative deactivation is too slow as described by a review by Dunbar (see Further reading section). Consequently the excess energy of the radical ions created by EI or PI is used for unimolecular chemical reactions, which eventually results in the formation of fragment ions. The surviving molecular ions and the fragment ions formed constitute the mass spectrum of the compound analysed. This illustrates that 'obtaining a mass spectrum' is principally a kinetic experiment, and kinetics laws postulate that in the case of parallel reactions of the same reaction order, which is inherently first order in the case of mass spectrometric fragmentations, the relative concentrations (intensities $I_i$) of the individual products match the ratio of the corresponding rate constants $k_i$ (**Scheme 1**).

Therefore, a detailed interpretation of a mass spectrum both for understanding the reactivity of organic radical cations and for a structure analysis of the compound studied requires the knowledge of the rate constant $k_i$ of the individual fragmentation steps and the dependence of $k_i$ on the structure of the reactants and their excess energy.

The unimolecular fragmentation of excited radical cations in a mass spectrometer refers to the general problem of the reaction kinetics of the decomposition of energized species isolated from their environment. The Rice, Ramsperger, Kassel, Marcus (RRKM) theory, the quasi-equilibrium theory (QET) of Rosenstock and Wahrhaftig, and related 'statistical' theories have been developed to deal with this problem. These theories are discussed in detail in the other articles. Using certain prerequisites they result in Equation [1], showing the dependence of the rate constant $k_i$ on the excess energy $E$, the number $s$ of vibrational degrees of freedom of the reactant ion, the activation energy $E_{i0}$ and the so-called frequency factor $\nu_i$ of fragmentation $i$.

$$k_i(E) = \nu_i \times \left(\frac{E - E_{i0}}{E}\right)^{s-1} \qquad [1]$$

While the oversimplifications used in deriving Equation [1] prevent any genuine calculation of $k_i(E)$ it is useful to demonstrate the factors that influence $k_i$, and its relation to reaction kinetics in a condensed phase at thermal equilibrium. The activation energy $E_{i0}$ has the same meaning as in condensed phase kinetics, and in the case of cleavage of only one bond of the reactant radical cation during the fragmentation process, $E_{i0}$ corresponds to the enthalpy of reaction $\Delta H_r$. This approximation is valid



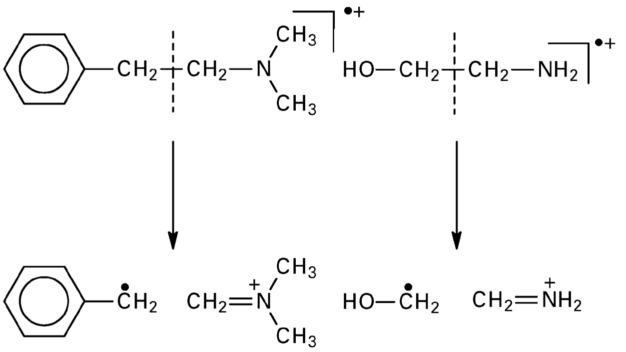

Scheme 1

because the reverse process of the fragmentation, i.e. bond formation between the ionized fragment (a cation) and the neutral fragment (a radical), requires no or only a very small activation energy (the so-called 'reverse' activation energy). If the excess energy $E$ of the reactant ion (the 'parent ion') is very large ($E >> E_{i0}$) the rate constant $k_i$ approaches the limiting value of $\nu$. The frequency factor $\nu$ reflects the frequency of that vibration of the reactant ion which is transformed into translation along the reaction coordinate of the fragmentation. The frequency factor $\nu$ is large if the transition corresponds simply to the elongation of the bond to be broken. The situation is more complicated in the case of a fragmentation by a rearrangement of the reactant radical cation which requires the making of new chemical bonds in a particular conformation as well as bond cleavage. As a consequence of the required conformation of the transition state the frequency factors $\nu$ of fragmentations by rearrangement are smaller than those of direct bond cleavages, and a rearrangement can compete successfully for the decomposition of the excited reactant ion only if the activation energy $E_0$ is small. Generally, a small activation energy is expected for an energetically favourable fragmentation that yields stable ionized and neutral fragments, if any extra reverse activation energy can be neglected.

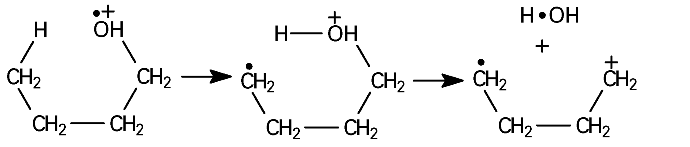
Applying these general considerations to the fragmentation of organic molecules in a mass spectrometer one expects three types of fragmentation reactions. Fragmentations of type I correspond to the direct cleavage of a weak bond of the reactant ion. Since this reaction has a large frequency factor and needs only a small activation energy, the rate constant will be large for any energy regime of the parent ion and the parent ion will always decompose preferably by this pathway. The result is a 'one-peak mass spectrum' that exhibits only the intense signal of the respective fragment ion in addition to a weak peak of the parent ion. Examples of type I fragmentations are found in the mass spectra of many organo-element compounds, which decompose by cleavage of the weak element–carbon bond, and compounds such as *N*,*N*-dimethyl-β-phenethyl amine (**Scheme 2**), which fragments into a stable ion $(CH_3)_2N^+{=}CH_2$ and a stable benzyl radical $C_6H_5CH_2^{\bullet}$ on electron impact ionization, or β-amino ethanol, which fragments into a stable ion $H_2N^+{=}CH_2$ and a stable $^{\bullet}CH_2OH$ radical.

A type II fragmentation corresponds to decomposition of the parent ion via a rearrangement into a stable fragment ion and a stable neutral fragment. Because of the stable products the reaction enthalpy of this process is usually small, as in the case of the type I fragmentation. However, the small frequency factor causes a slow rise of the rate constant with the excitation energy of the parent ion. Thus, this rearrangement process is the favoured decomposition pathway for parent ions of low excitation energy. A typical type II fragmentation process is the elimination of $H_2O$ from the molecular ions of *n*-butanol and its higher homologues. Specific studies have shown that this process is mainly a 1,4-elimination and most likely proceeds as a two step process (**Scheme 3**) with a slow initial H-atom transfer to the hydroxy group.

A type III fragmentation corresponds to the cleavage of covalent bond compounds, e.g. a C–H or a C–C bond. Since these bonds are rather strong the activation energy of this process (comparable to the reaction enthalpy) is relatively large, but the cleavage is entropically favoured by a large frequency factor. Hence, reactant ions with plenty of excess energy follow this fragmentation pathway. In principle there could be a fourth type of fragmentation, an energetically unfavourable rearrangement process that exhibits a large activation barrier and a small frequency factor. Such a fragmentation will be slow for all energy regimes of the reactant ions and cannot compete with the other types of fragmentation. If a process of this type appears in the mass spectrum, there is either no other fragmentation path or a 'hidden' rearrangement of the reactant ion to another isomer precedes the fragmentation. Examples of the first situation are the mass spectra of benzene and polycyclic arenes, which show a distinct signal for the loss of a methyl radical from the parent ion. The only 'simple' fragmentation of these radical ions is the loss of an H atom, which, however, requires about 3–4 eV activation energy because of the strong $C_{arene}$–H bond. Therefore, the deep-seated rearrangement necessary for the elimination of a methyl radical is possible but most molecular ions of the arene do not fragment at all during their residence time within a



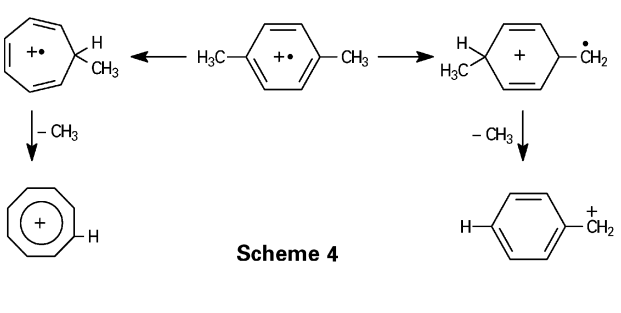

**Scheme 2**




Scheme 3

mass spectrometer. An example of the second situation is the loss of a methyl substituent from the molecular ions of xylene which would require cleavage of a strong $C_{arene}$–C bond and which should not compete effectively with loss of H to generate a stable methyl-benzyl cation (**Scheme 4**). An investigation of the fragmentation mechanism by Grotemeyer and Grützmacher showed (see the Further reading section) that a direct loss of the methyl substituent occurs only for highly excited molecular ions whereas most of the molecular ions rearrange either by H-migration or by a skeletal rearrangement into a methyl cycloheptatriene radical cation before elimination of the methyl group.

The peak pattern of the EI mass spectrum of a standard organic compound arises from competition between type II and type III fragmentations. This will be illustrated below by a few examples to familiarize the reader with the concepts used for the interpretation of the fragmentations of radical ions in a mass spectrometer. However, it has to be remembered that the peak pattern in the mass spectrum of a compound depends very much on the method used for the ionization and on the type of mass spectrometer and scan mode used to acquire the spectrum. The former determines the energy content of the parent ions created in the ion source of the instrument and the latter influences the reaction time allowed for the ions to decompose. If the ionization method chosen supplies only a relatively small amount of excess energy to the molecular ion, the rates of most fragmentations (apart from those of type I) are small and the mass spectrum will consist mainly of the peak for intact parent ions. At long reaction times ($> 10$ μs) peaks of the product ions of type II fragmentation will appear because of the small activation energies of such processes. 'Soft' ionization




Scheme 4

methods that produce this category of spectra are field ionization (FI), charge exchange (CE) ionization with a primary ion of low ionization energy ($C_6F_6$, $CS_2$), and multiphoton ionization (MUPI) as described by Grotemeyer and Schlag. In particular, MUPI can be used to steer the ionization into a 'soft' or a 'hard' mode by adjusting the number and energy of the photons absorbed by the molecules of the sample. Ionization of molecules is most often achieved by electron impact (EI) with electrons of 70 eV kinetic energy and is a 'hard' ionization process that produces parent ions having a broad distribution of excess energy up to several eV.

The type of instrument, ion source and scan mode used to obtain the mass spectrum define the reaction time during which the fragment ions that appear in the mass spectrum are generated. A routine sector field mass spectrometer with several kV ion accelerating voltage has a residence time for the ions within the ion source of $< 1$ μs; this time is extended slightly for a quadrupole mass spectrometer using lower accelerating voltages. All ions appearing in the mass spectrum recorded under standard conditions have to be formed within the ion source, and a short residence time is favourable for the observation of fragmentations with a large frequency factor, i.e. type I and type III fragmentations which originate from highly excited parent ions. However, with *tandem-mass spectrometry* the reaction time can be extended by more than an order of magnitude. At a reaction time of 10 μs or longer the highly excited parent ions have decomposed already and only ions with small amounts of excess energy remain. Since the only fragmentation pathways open to these gently excited ions are those with small activation energies, the decomposition of these so called *metastable ions* occurs predominantly by type I and type II processes. Sometimes this is not very useful for analytical applications of mass spectrometry because the rearrangement and isomerization may obscure the original structure of the molecules analysed, but these low energy processes of the radical cations within a mass spectrometer are more closely related to chemical reactions under thermal conditions and are of eminent importance for an understanding of radical ion chemistry.

## Fragmentation mechanisms and fragmentation/structure relationship

The fragmentation reactions observed in an EI mass spectrometer are usually discussed with respect to their reaction mechanisms and the effect of structural parameters of the compound studied on these mechanisms. The types of arguments used in this discussion follow closely the practice used in mechanistic organic chemistry, considering in particular the stability of individual fragmentation products, intermediates and transition states. Some of the more general fragmentation mechanisms observed in the EI mass spectra of organic compounds are discussed in the following examples.

### Example 1: type I fragmentation and charge distribution on the fragments

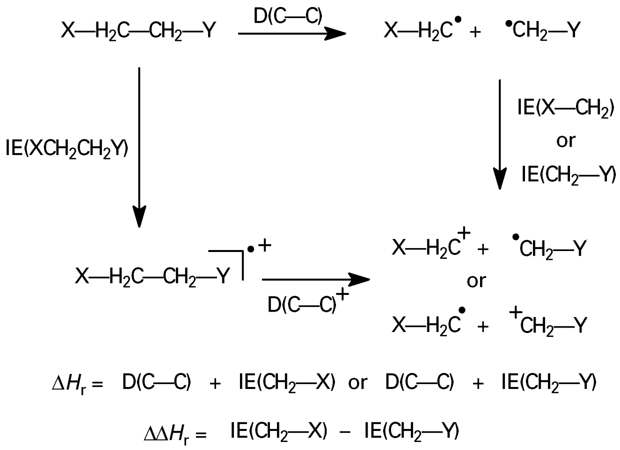
For a compound of structure X–$CH_2$–$CH_2$–Y where X and Y correspond to functional groups, one expects an easy cleavage of the central C–C bond in the radical cation, generating the fragments X–$CH_2$ and Y–$CH_2$. Only one of these fragments will carry the positive charge while the other one remains as a neutral radical. Stephenson's rule states that in a fragmentation the positive charge appears predominantly with the fragment of lowest ionization energy (IE). This rule is explained easily by the (relative) reaction enthalpy $\Delta H_r$ of the competing fragmentations (**Scheme 5**). Starting from the neutral molecule the $\Delta H_r$ for the fragmentation routes a and b are given by Equations [a] and [b], and since the dissociation energy of the central bond is the same for both processes, that with the lowest IE will exhibit the lowest $\Delta H_r$ and will be preferred. Examples of this effect are given in **Table 1** for neopentyl derivatives $(CH_3)_3$ C–$CR_2$–Y and β-phenethyl derivatives for the competition between a *t*-butyl or benzyl group and the fragment $CH_2$–Y.

**Table 1** Ionization energy (IE) and relative intensity (rel. int.) in the 70 eV mass spectra of some neopentyl derivatives $(CH_3)_3C$–$CR_2$–Y and β-phenethyl derivatives $C_6H_5$–$CH_2$–$CH_2$–Y

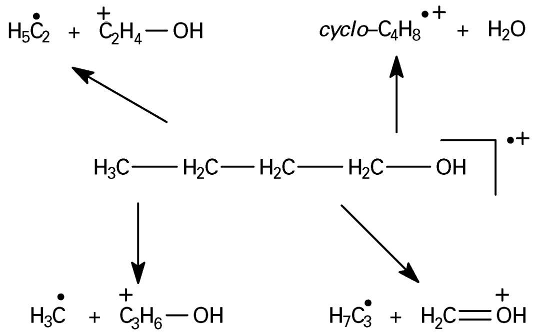
| *Compound* | *Fragment (CH₃)₃C* | | *Fragment CH₂–Y* | |
|---|---|---|---|---|
| *(CH₃)₃CCH₂Y* | *IE*(eV) | *rel. int.* (%) | *IE* (eV) | *rel. int.* (%) |
| $(CH_3)_3CCH_2NH_2$ | 6.7 | 2 | 6.1 | 100 |
| $(CH_3)_3CCH_2OH$ | 6.7 | 100 | 7.56 | 9 |
| $(CH_3)_3CCH(CH_3)OH$ | 6.7 | 100 | 6.7 | 93 |
| $(CH_3)_3CC(CH_3)_2OH$ | 6.7 | 15 | < 6.7 | 100 |
| | *Fragment C₆H₅CH₂* | | *Fragment CH₂–Y* | |
| $C_6H_5CH_2CH_2NH_2$ | 7.2 | 11 | 6.1 | 100 |
| $C_6H_5CH_2CH_2OH$ | 7.2 | 100 | 7.56 | 3 |

### Example 2: competing bond cleavages and fragmentation by rearrangement

A similar example of the effect of the relative reaction enthalpy $\Delta H_r$ on the relative abundance of competing fragmentations is obtained by an estimation of $\Delta H_r$ for the (hypothetical) C–C and C–O bond cleavages within the molecular ion of *n*-butanol and for the loss of a $H_2O$ molecule. From the data given in **Scheme 6** it is seen that C–C bond cleavage next to the α-C atom carrying the OH group ('α-cleavage') is the energetically most favoured but that a fragmentation of the molecular ion by loss of $H_2O$ is even more favoured. The EI mass spectrum of *n*-butanol indeed exhibits large peaks due to α-cleavage by formation of the $^+CH_2OH$ (100% rel. int.) and to loss of $H_2O$ (~50% rel. int.) from the molecular ion as the only abundant primary fragmentation process. However, loss of $H_2O$ requires a rearrangement by hydrogen migration and is hindered by a small frequency factor (type II fragmentation). By investigat-




**Scheme 5**




**Scheme 6**

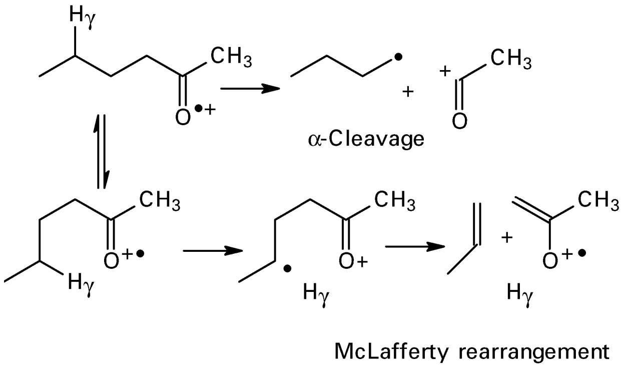
ing specifically isotopically labelled *n*-butanols it was shown that loss of $H_2O$ occurs by a 1,4-elimination and is most likely a two-step process, in which the transfer of the hydrogen atom to the OH group in the first step has an additional activation barrier. Hence α-cleavage (a type III fragmentation) is predominant. However, the long-lived metastable molecular ions of *n*-butanol, which have more time for reactions and which are less excited, decompose only by loss of $H_2O$. If a similar estimation is made for *n*-butylamine it turns out that elimination of $NH_3$ is not energetically favoured over the α-cleavage of the molecular ion, resulting in the very stable ion $^+CH_2NH_2$ (*m/z* 30). This latter process corresponds to a type I fragmentation, and the 70 eV mass spectrum of *n*-butylamine (as most other *n*-alkyl amines) is a 'one-peak' mass spectrum, exhibiting only one intense signal at *m/z* 30.

### Example 3: the McLafferty rearrangement

An example related to Example 2 is the fragmentation of the molecular ions of hexan-2-one by α-cleavage at the carbonyl group and by a McLafferty rearrangement, as shown in **Scheme 7**. α-cleavage at a carbonyl group bond gives rise to a stable acetyl cation $CH_3CO^+$, *m/z* 43, and a $^{\bullet}C_3H_7$ radical while transfer of a H atom from the γ–$CH_2$ group to the C=O group accompanied by cleavage of the C(α)–C(β) bond produces the radical cation of the acetone enol and a propene molecule. The latter process exhibits a smaller reaction enthalpy and probably also a smaller activation energy than the α-cleavage, but because of the larger frequency factor $\nu$ of this simple bond cleavage the acetyl cation is normally the base peak in the mass spectrum at *m/z* 43.

The fragmentation of the radical cation of a carbonyl compound by a rearrangement corresponding to H atom transfer from a γ position to the carbonyl group and cleavage of the bond between the atoms at the $\alpha$ and $\beta$ positions (**Scheme 7**) is known as the *McLafferty* rearrangement and is a general reaction of excited radical cations of aldehydes, ketones, esters and amides. Because of this the McLafferty rearrangement is one of the most widespread and reliable fragmentations for mass spectrometric structure analysis. Its mechanism is related to the well-known Norrish type II photoreaction of carbonyl compounds and has been studied in much detail. Normally the McLafferty rearrangement is a two-step process composed of a 1,4-hydrogen migration to the carbonyl-O atom followed by bond cleavage in the intermediate. This intermediate corresponds to a unique class of radical cations named *distonic ions*, which are characterized by the positive charge and the radical cation residing in separate molecular orbitals, in contrast to 'normal' radical ions (or *molecular ions*) in which the positive charge (the 'electron hole') and the radical electron are located in the same molecular orbital. Distonic ions and their characteristics have been reviewed by Hammerum and by Kenttämaa (see the Further reading section). The properties of distonic ions will be discussed in more detail in Example 7 below. In the case of the McLafferty rearrangement the distonic intermediate is composed of an oxonium ion (the protonated carbonyl group) and an alkyl radical. The rate of its formation is favoured sterically by the six-membered transition state and depends on the bond energy of the H–X bond (X = C,O,N,S) at the γ position which is cleaved during the process. Any structural feature which lowers the dissociation energy of this bond (for example tertiary C–H bond versus primary C–H bond) enhances the intensity of the product ion of the McLafferty rearrangement in the mass spectrum. The charge may remain with either of the two fragments of the McLafferty rearrangement according to Stevenson's rule; usually this is the oxygen-containing species. Another interesting detail of the mechanism is the possibility that the two fragments stay bound to each other in the gas phase, forming an intermediate *ion–neutral complex*. (The properties of these intermediate ion–neutral complexes will be discussed in Example 7 and have been reviewed by Morton and others; see the Further reading section). During the lifetime of this complex the components may undergo intermolecular reactions, usually H atom transfers that give rise to the 'McLafferty +1' fragment ion seen in the mass spectra of many esters of long chain alcohols.

Fragmentations related to the McLafferty rearrangement, and often denoted also as a McLafferty rearrangement, are observed in the mass spectra of other organic compounds with a γ-H atom available to a suitable H acceptor group. An important

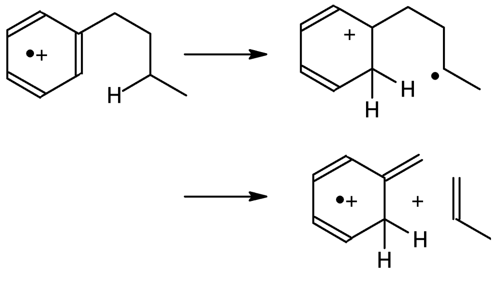



**Scheme 7**

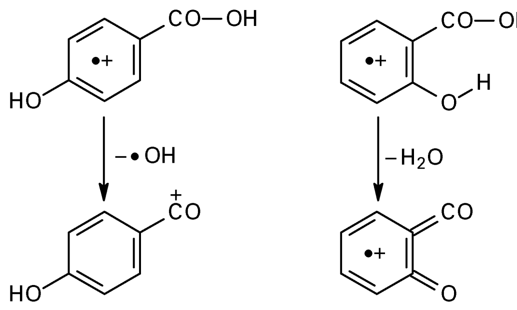
example of this type of fragmentation is radical cations of long-chain alkyl arenes, which contain the ionized arene systems as the H acceptor group (**Scheme 8**). The 1,4-migration of the H atom from the $\gamma$ position of the sidechain creates a distonic ion composed of a protonated arene (an *arenium ion*) and an alkyl radical. Subsequent bond cleavage within the sidechain gives rise to an alkene fragment and the tautomer of a methyl arene ('isotoluene'), which usually carries the positive charge. An interesting feature of this fragmentation is proton migration within the arenium ion by successive 1,2-shifts (a 'ring walk') which may initiate further reactions.

### Example 4: the *ortho* effect and steric effects on fragmentations

While the EI mass spectra of the positional isomers of many di- and polysubstituted aromatic compounds are almost identical, the mass spectra of isomers carrying certain substituents in the *ortho* positions are clearly different from the *meta* and *para* isomers. A typical example is the mass spectra of *para*-, *meta*-, and *ortho*-hydroxybenzoic acid. The mass spectrum of the *para* isomer is easily analysed; the main fragment ions arise from a series of simple bond cleavages starting with the loss of an $^{\bullet}$OH radical by $\alpha$-cleavage within the carboxy group (**Scheme 9**). Analogous reactions are observed in the mass spectrum of the *meta* isomer. In contrast the mass spectrum of the *ortho* isomer (salicylic acid) is dominated by a signal of the fragment ions $[M-H_2O]^{\bullet+}$. An investigation by McLafferty in 1959, of the mechanism of this *ortho* effect showed that the $H_2O$ molecule lost from the molecular ion is generated by a specific transfer of the H atom of the hydroxy group to the neighbouring carboxy group. The loss of $H_2O$ is energetically very favourable, and the frequency factor for the H rearrangement is large because the H donor (the hydroxy group) and the H acceptor (the carboxy group) are fixed in a sterically favourable arrangement of the *ortho* disubstituted benzene ring. Similar intense




Scheme 8



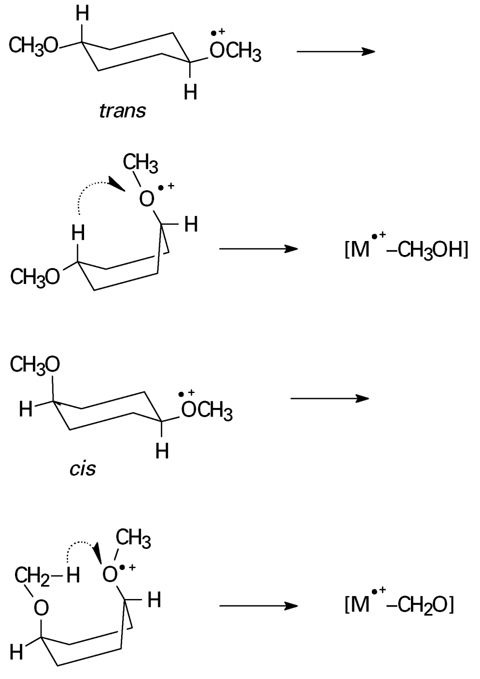

Scheme 9

fragmentations by this *ortho* effect are observed in the mass spectra of many *ortho* substituted arenes, the prerequisites being the presence of a H donor group H–Y– (H–$CR_2$–, H–O–, H–NR–, H–S–) and the possibility to eliminate a thermodynamically stable molecule H–X (H–Cl, H–OH, H–OR, H–$NH_2$, etc.) after H transfer to the acceptor group in the *ortho* position. Thus, fragmentation by the *ortho* effect is a very reliable and useful reaction for mass spectrometric structure analysis.

A special aspect of the *ortho* effect, in addition to the 'driving force' of the reaction by elimination of a stable molecule, is the fixed arrangement of the two groups, allowing their intramolecular interaction. This proximity effect or neighbouring group effect, which guarantees a favourable frequency factor $\nu$ of the fragmentation by rearrangement, is present in many other semi-rigid organic molecules and can be used to differentiate mass spectrometrically between stereoisomers. An early example of this type of effect is the mass spectra of *cis*- and *trans*-cyclohexanediols and their di-O-methyl ethers, which can be easily distinguished by their EI mass spectra because of the stereospecific fragmentations shown in **Scheme 10**. However, it should be noted that this type of effect requires that the molecular ions of the stereoisomers do not interconvert before fragmentation.

### Example 5: fragmentations by intramolecular substitution

The generation of the principal fragment ions in the EI mass spectrum of benzalacetone is depicted in **Scheme 11**. These fragment ions can be pictured as arising, at least formally, by a series of simple bond cleavages, with the exception of the formation of the ions $[M-H]^+$. The mass spectra of deuterated derivatives demonstrate that the H atom is lost from the aromatic ring. This is explained by an intramolecular attack of the (nucleophilic) carbonyl group on the ionized benzene ring, eventually giving rise to

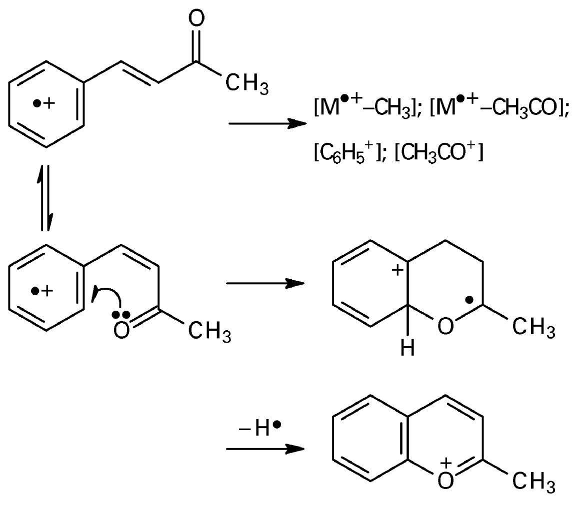



**Scheme 10**

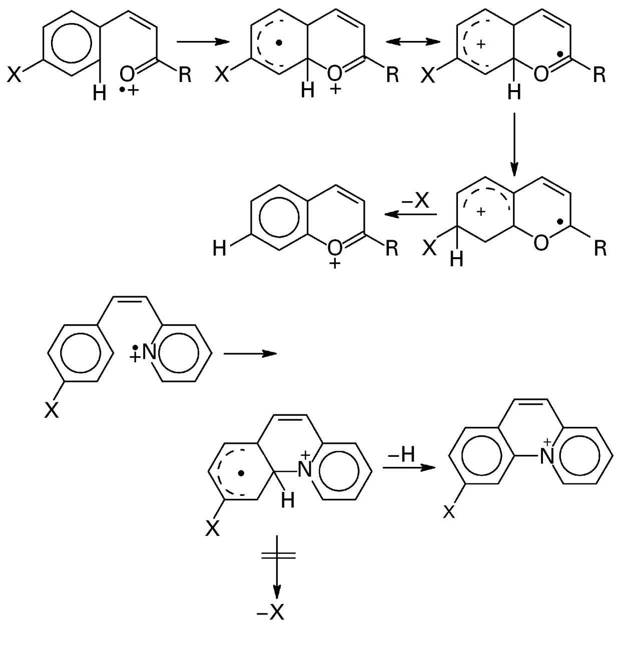
a very stable benzopyrylium ion via an intermediate related to the $\sigma$-complex of aromatic substitution.

This type of fragmentation by an intramolecular substitution has been studied in detail by the author and is a rather common reaction of unsaturated or aromatic radical cations of a suitable structure. The substitution occurs by attack of a nucleophilic group on an ionized C=C double bond or an ionized aromatic ring and is driven by the generation of a stable cyclic fragment ion. By this process a substituent at the attacked position is lost, and the process is particularly abundant in the mass spectra if a weak bond is cleaved, for instance during the loss of a Cl or Br substituent, but methyl or alkyl groups are also easily eliminated.

Interestingly, two definitively different intermediates are generated by the intramolecular substitution of an aromatic ring which exhibit different reactivities depending on whether the positive charge is fixed at the attacking group. In the case of the intermediate created from cinnamic aldehyde, however, both the radical electron and the positive charge are delocalized over the total $\pi$-electron system (**Scheme 12**). As a consequence the positive charge may also appear at the (formerly) aromatic ring, giving it the properties of a true $\sigma$-complex. In these $\sigma$-complexes a proton migrates by successive 1,2-shifts around the ring (‘*ring walk*’) before further fragmentation. Consequently, substituents which are good ‘mass spectrometric leaving groups’ because of a small bond energy (such as Cl or Br) are lost not only from the attacked position but also from all other positions after H migration by protolytic bond cleavage (a well-known process of substituted arenes in acidic solutions). The other type of intermediate is represented by the intramolecular substitution within the molecular ion of 2-stilbazole or other related heteroaromatic analogues of stilbene. In this case the positive charge of the intermediate is localized outside of the attacked aromatic ring in a quaternary ammonium ion and only the radical electron is delocalized in the $\pi$-electron system. This type of intermediate is therefore a distonic radical cation. The migration of H in radicals by a 1,2-shift is forbidden by the Woodward–Hoffman rules. Hence, from this




**Scheme 11**




**Scheme 12**

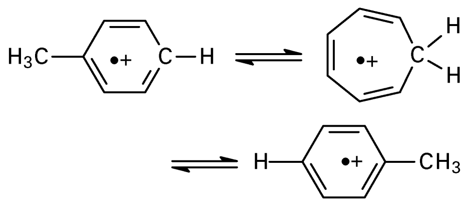
type of intermediate, substituents (or a H atom) are lost specifically from the attacked position of the aromatic ring.

### Example 6: the toluene–cycloheptatriene radical cation rearrangement

The examples of fragmentation reactions of radical cations discussed so far correspond either to direct bond cleavages or to bond cleavages preceded by H migration or intramolecular bond formation. These rearrangements do not alter the connectivity of C and other atoms of the original molecules. However, fragmentations of radical cations by 'skeletal rearrangements' are also known. The most familiar example is the interconversion of the radical cations of toluene and cycloheptatriene, which interchanges C atoms within the C skeleton and which gives rise to stable tropylium ions (**Scheme 13**) after H loss.

The rearrangement of $C_7H_8$ radical cations has been studied in detail both by experimental methods and by theoretical calculations and the mechanism is well understood. The crucial feature is the initial generation of a distonic ion by a 1,2-H shift from the methyl substituent to the aromatic ring. This is followed by cyclization to a bicyclic norcarane-like structure which eventually ring opens to the cycloheptatriene radical cation. This last step exhibits the highest activation barrier of the rearrangement and makes it a high energy process. However, a fragmentation of the toluene radical cation by a direct C–H bond cleavage also needs a large activation energy. Hence, the rearrangement competes successfully with this fragmentation, and a $C_7H_8$ molecular ion may undergo several reversible toluene–cycloheptatriene rearrangements before the final fragmentation.

It often is assumed that this type of toluene–cycloheptatriene rearrangement applies to other polysubstituted alkylbenzenes, alkylarenes and even alkyl heteroarenes. Indeed, this rearrangement explains the abundant loss of a whole substituent from an arene molecular ion in spite of the normally strong bond to a C(sp$^2$) atom of the arene, which has to be cleaved to detach the substituent. Rearrangement into a substituted cycloheptatriene radical cation and proton migrations within the ionized seven-membered ring can place any substituent at the C(sp$^3$) atom of the cycloheptatriene, facilitating the loss of the substituent X from a >CH–X group. Nonetheless, one must realize that the toluene–cycloheptatriene rearrangement is a high energy process of excited ions, and other, less energy demanding, fragmentation mechanisms of polysubstituted arenes into substituted derivatives of the $C_7H_7^+$ ion may override this rearrangement. Thus, the loss of a methyl radical from *t*-butylbenzene radical cations occurs by a direct 'benzylic' cleavage, producing the stable cumyl cations, and the loss of a methyl radical from the radical cations of *p*-xylene is preceded by the toluene–cycloheptatriene rearrangement only for highly excited ions that fragment in the ion source of a mass spectrometer, while the only gently excited metastable *p*-xylene radical cations decompose by H transfer to the aromatic ring, a proton ring walk to the position of the second methyl group and protolytic bond cleavage (**Scheme 4**).

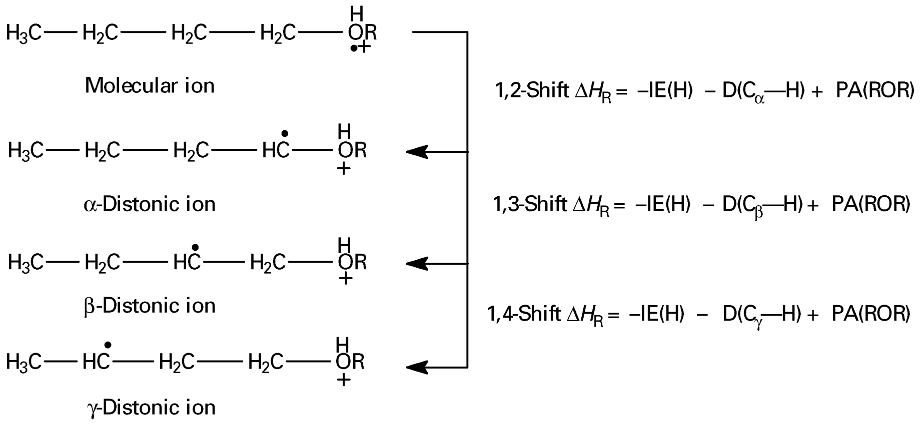



**Scheme 13**

### Example 7: distonic radical cations and intermediate ion–neutral complexes

The examples of mass spectrometric fragmentations of radical cations presented above show that in many cases a process which formally appears to be a simple bond cleavage corresponds to a multistep mechanism and includes several intermediates. The reason for this detour in the fragmentation route is usually the lower activation energy of the multistep process, which converts the original radical cation into an isomer more suitable for fragmentation into thermodynamically stable products. This is especially true if decompositions of metastable ions in a mass spectrometer are investigated, since these ions have long lifetimes and hence do not rely only on fragmentations with a large frequency factor. However, it is always essential that the intermediate ions of the multistep mechanism correspond to energetically favourable structures compared with the original radical cation. Two types of such intermediate ions have gained importance in explaining mass spectrometric fragmentations and will be discussed in this last example: distonic ions and intermediate ion–neutral complexes.

As mentioned before, distonic ions are distinguished from normal molecular ions by the occurrence of the positive charge and the radical cation in separate molecular orbitals. Several examples of intermediate distonic ions have been discussed above, and these distonic ions arise usually by hydrogen migration, for example during the initial step of the McLafferty rearrangement. Other modes for the

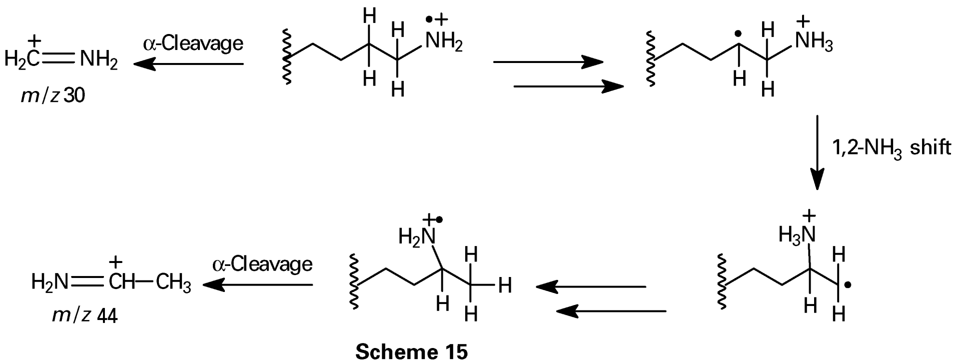



**Scheme 14**

generation of distonic ions are the ring cleavage of cyclic molecular ions or the addition of a radical cation to a C=C double bond. Depending on the minimum number of atoms between the positive charge and the radical site, α-, β-, γ- or even more distant distonic ions are known. A hydrogen shift within an aliphatic molecular ion to a heteroatom generates distonic ions composed of an onium structure and a radical. The stability of these distonic ions relative to the original molecular ion is important for their role as reaction intermediates, and this can be estimated from the thermochemical cycle (**Scheme 14**) generated from the energy needed for homolytic cleavage of the C–H bond, the ionization energy of hydrogen, and of the proton affinity at the heteroatom. For a series of isomeric distonic ions with different distances between the positive charge and the radical site the proton affinity of the heteroatom and ionization energy of the H atom are identical. Hence, the essential feature for the relative stability within this series is the dissociation energy of the C–H bond. It can be shown that this is more or less constant, with exception of the smaller C–H bond energy adjacent to the heteroatom. It follows that in this series of isomeric distonic ions the α-distonic ion should always be the most stable. However, the 1,2-H shift necessary to convert the conventional molecular ion into the α-distonic isomer is usually associated with a high activation energy and a direct isomerization does not take place. Instead, other distonic ions are formed by less energy demanding 1,4-, 1,5-, 1,6- or even more distant hydrogen shifts, and these distonic ions are eventually transformed into the stable α-distonic isomer by further H shifts.

It is of interest to note that distonic ions also isomerize by shifts of the protonated group that contains the heteroatom. A typical example is the 1,2-$NH_3$ shift in the α-distonic ion of an *n*-alkyl amine. This shift converts the molecular ion of the *n*-alkyl amine into the molecular ion of an *s*-alkylamine and is responsible for the small but distinct signals at *m/z* 44 that are found in addition to the expected base peak at *m/z* 30 in the mass spectra of these amines (**Scheme 15**). Originally it was supposed that the ions were generated by 'β-cleavage' and formation of protonated ethylene imine.

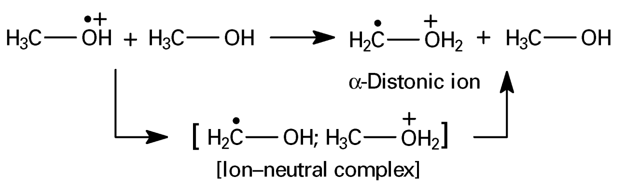
The main factors that determine the relative stability of a distonic ion with respect to molecular ions containing different heteroatoms are, on the one hand, the ionization energy of the molecule and, on the other hand, the proton affinity of the heteroatom at the basic centre of the molecule. Considering these quantities for different heteroatoms it can shown that distonic ions containing an oxonium structure




**Scheme 15**

should be especially stable because alcohols and ethers exhibit rather high ionization energies and also large proton affinities. In *n*-alkylamines the amino group also exhibits a large proton affinity, but the ionization energy of amines is also quite low, favouring the stability of the conventional molecular ion. These expectations are confirmed by high level *ab initio* calculations by Radom (see the Further reading section). Of particular stability is the distonic isomer of the molecular ion of methanol, and it has been shown that the methanol molecular ions may isomerize (catalytically!) into the α-distonic isomer during a biomolecular collision with a neutral methanol molecule (**Scheme 16**). It is possible that a similar isomerization of molecular ions into distonic isomers by collision with a neutral molecule may occur in different systems.

Because parent ions, fragment ions and the neutral fragments travel through the mass spectrometer at the very high speed resulting from the acceleration by the high electric field of the ion source, the relative velocities of a fragment ion and the corresponding neutral fragment after a fragmentation are usually quite small and the fragmentation products drift only slowly apart. Since ions and neutrals attract each other by ion–dipole and ion–induced-dipole forces, the fragment ion and neutral fragment may be held by electrostatic attraction during the first stages of the dissociation in an ion–neutral complex. This gives the components of the intermediate ion–neutral complex of a mass spectrometric fragmentation a chance to undergo an exothermic bimolecular reaction within the complex. After the dissociation, the result of this reaction mimics the fragmentation mechanism of a rearrangement process. However, since the components within the intermediate ion–neutral complex are more or less free to rotate with respect to each other, this fragmentation route by intermediate ion–neutral complexes avoids the small frequency factor of a rearrangement with a sterically strained transition state. Since the bimolecular reaction within the complex has to be fast to compete with the direct dissociation of the ion–neutral complex, only rather simple reactions (without a large additional activation energy) are observed, i.e. exothermic hydrogen or proton transfer or nucleophilic addition of carbenium ions to electron lone pairs or π-electrons.



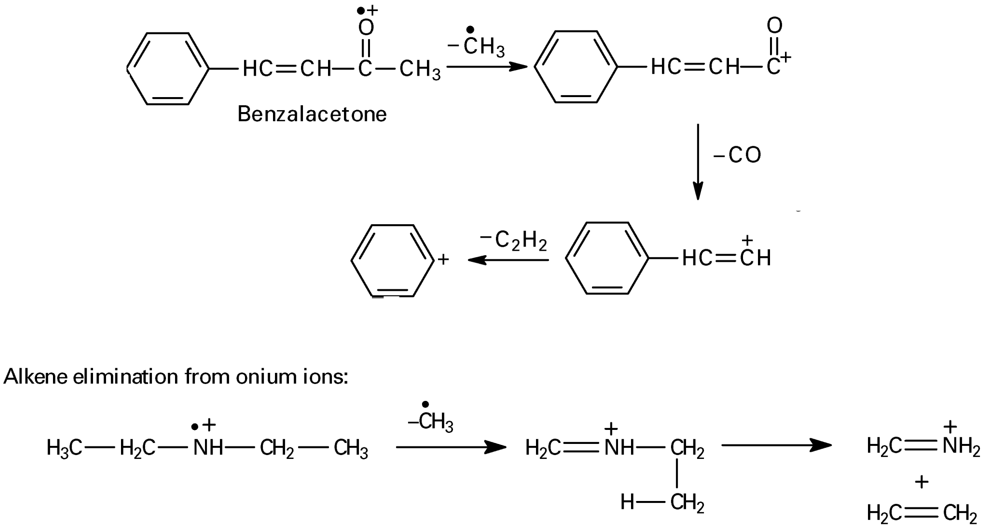

**Scheme 16**

The first fragmentation of this type was described by Longevialle and Botter (see the Further reading section) in the mass spectra of certain amino steroids, in which a fragment ion generated at the right hand side of the steroid molecule transported a proton over a very large intramolecular distance to a functional group at the left hand side of the steroid. Probably the most ubiquitous fragmentation via intermediate ion–neutral complexes is the loss of alkane molecules that originate from H abstractions by an alkyl radical in an ion–neutral complex generated by a simple C–C bond cleavage of an alkyl chain. Traeger and coworkers (see the Further reading section) have shown that this fragmentation is especially abundant in the mass spectra of cycloalkanes carrying alkyl sidechains. Indeed, the EI mass spectra of these compounds are distinguished by a series of peaks at even *m*/*z* values which arise by this mechanism. Other interesting examples for this fragmentation mechanism are the migration of onium ions (and other stable carbenium ions) over long molecular distances within ion–neutral complexes before the terminating ion–molecule reaction. This corresponds to a reactive interaction between two functional groups which are far apart in the original molecule. Thus, this type of fragmentation provides an explanation for a formal rearrangement by migration of functional groups within the original molecular ion.

## Secondary fragmentation

As discussed in the introduction and shown in **Scheme 1**, the fragmentations of organic molecular ions, which are induced by electron impact in a mass spectrometer and which give rise to the peak pattern of the EI mass spectrum, correspond to a kinetic scheme of parallel and consecutive unimolecular reactions. Any fragment ion that still contains sufficient energy for further fragmentation will do so, and these fragmentations correspond to the reactions of an isolated and energetically excited species described by QET or RRKM theory. However, the radical cations (molecular ions) generated by EI from the neutral molecules may decompose either by elimination of small neutral molecules (loss of $H_2O$, see Example 2; loss of alkene, see Example 3) or by loss of radicals. In the former case the newly generated secondary fragment ions correspond to radical cations, while in the latter case the fragmentations give rise to secondary even-electron cations. The different electronic nature of the fragment ions has an effect on the preferred pathways for further fragmentations.

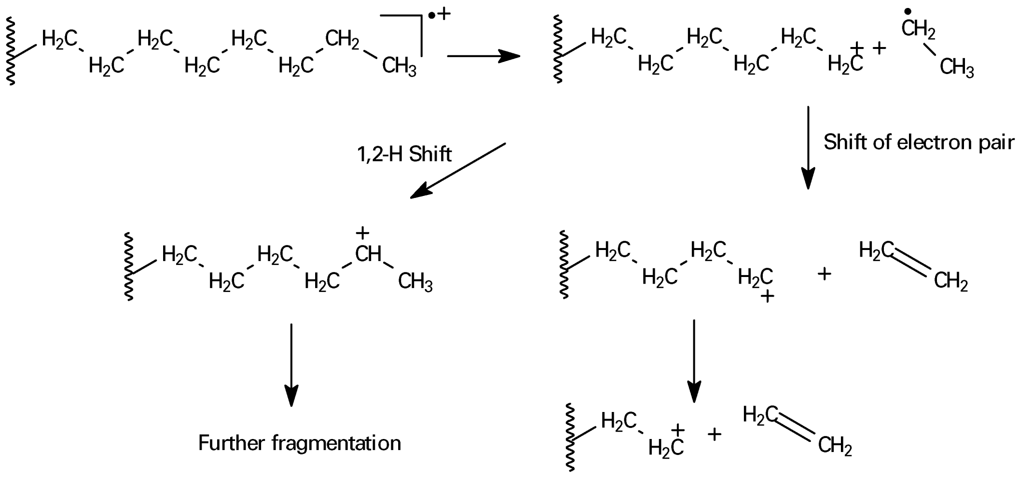
The types of fragmentation reactions observed for secondary radical ions are more or less those




Scheme 17

discussed in the previous section. Thus, the fragmentations of the $[M^{\bullet+}-H_2O]$ fragment ions in the mass spectra of aliphatic alcohols concur with those of the radical cations of the corresponding alkenes (or alkylcyclobutanes), and because of these further fragmentations of secondary fragment ions the mass spectra of long chain alcohols and alkenes are very similar. In comparing the fragmentations of secondary radical cations with radical cations produced directly by EI from stable neutral molecules one has to account for the different excess energy present in radical cations originating from EI and from fragmentation, and, more importantly, for the possibility that fragment ions may correspond to distonic isomers or tautomers of the molecular ions. Thus, the secondary $C_3H_6O^{\bullet+}$ radical cations created by the McLafferty rearrangement of the molecular ions of hexan-2-one (**Scheme 7**) correspond to the ionized enols of acetone and not acetone radical cations, and further fragmentation may require 're-ketonization'. Similarly, elimination of an alkene fragment from the molecular ions of alkylbenzenes gives rise primarily to an isotoluene radical cation (**Scheme 8**) which differs in reactivity from the toluene molecular ion. In most cases these differences alter the details of the fragmentation mechanisms, but not the principal routes of a further fragmentation.

Secondary fragment ions which are even-electron cations prefer the elimination of neutral molecules in




Scheme 18

their further decompositions. This has inspired the formulation of the so-called 'odd–even-electron' rule which states that odd-electron radical cations may yield either odd-electron or even-electron secondary ions on decomposition while even-electron cations produce only smaller even-electron cations as products. However, many exceptions to this rule are known, in particular for the mass spectra of large polyalkylated arenes and of organo-element compounds. In individual cases it is clearly the stability of the product ions which directs further fragmentations. Nevertheless, the rule can still be used as a first guide for the interpretation of mass spectra of organic compounds. The elimination of a small molecule from an even-electron cation is achieved by a shift of an electron pair only or by hydrogen migration. A typical example for the former mechanism is the fragmentation cascade observed for the $[M–CH_3]^+$ ion in the mass spectrum of benzalacetone while the latter fragmentation is typical of alkene eliminations from the onium ions produced by α-cleavage (**Scheme 17**).

Even-electron cations are known to rearrange easily by 'electron-sextet rearrangements' of the Wagner–Meerwein type. In particular, alkyl cations and other hydrocarbon cations rearrange easily by this mechanism. Thus, 1,2-hydrogen shifts and 1,2-alkyl shifts in an alkyl cation compete successfully with a direct fragmentation by alkene loss as depicted in **Scheme 18**. The formation of an unstable primary alkyl cation by direct bond cleavage within a molecular ion is very likely associated with a concerted 1,2-H shift to create a more stable secondary fragment ion. As a consequence of these rearrangements the abundant $C_3H_7^+$ and $C_4H_9^+$ ions observed in the mass spectra of *n*-alkanes at *m*/*z* 43 and *m*/*z* 57 correspond to stable isopropyl cations and *t*-butyl cations.

## List of symbols

$E$ = excess energy; $E_{i0}$ = activation energy; $k_i$ = rate constant; $\nu_i$ = frequency factor; $\Delta H_r$ = reaction enthalpy.

*See also:* **Ion Energetics in Mass Spectrometry; Ion Imaging Using Mass Spectrometry; Ion Structures in Mass Spectrometry; Isotopic Labelling in Mass Spectrometry; Metastable Ions; Photoionization and Photodissociation Methods in Mass Spectrometry; Stereochemistry Studied Using Mass Spectrometry**.

## Further reading

Barkow A, Pilotek S and Grützmacher H-F (1995) *European Mass Spectrometry* **1**: 525–537.

Bowen RD (1991) *Accounts of Chemical Research* **24**: 364–371.

Budzikiewicz H, Djerassi C and Williams DH (1967) *Mass Spectrometry of Organic Compounds*. San Francisco: Holden-Day.

Dunbar RC (1992) *Mass Spectrometry Reviews* **11**: 309–339.

Fax MA and Channon M (1988) *Photoinduced Electron Transfer*. Amsterdam: Elsevier.

Gauld JW and Radom L (1997) *Journal of the American Chemical Society* **119**: 9831–9839.

Grotemeyer J and Grützmacher H-F (1982) In: Maccoll A (ed) *Current Topics in Mass Spectrometry and Chemical Kinetics*. London: Heyden.

Grotemeyer J and Grützmacher H-F (1982) *Organic Mass Spectrometry* **17**: 353–359.

Grotemeyer J and Schlag EW (1988) *Angewandt Chemie, International Edition English* **27**: 447–460.

Grützmacher H-F (1993) *Organic Mass Spectrometry* **28**: 1375–1387.

Hammerum S (1988) *Mass Spectrometry Reviews* **7**: 123–202.

Levsen K (1968) *Fundamental Aspects of Organic Mass Spectrometry*. Weinheim: Verlag Chemie.

Lifshitz C (1997) *International Review of Physical Chemistry* **16**: 113–139.

Longevialle P and Botter R (1983) *Organic Mass Spectrometry* **18**: 1–8.

Marcus RA (1952) *Journal of Chemical Physics* **20**: 359–364.

Mattay J (ed) (1992) In: *Photoinduced Electron Transfer V*, Topics in Current Chemistry, Vol 168.

McAdoo DJ and Morton TH (1993) *Accounts of Chemical Research* **26**: 295–302.

McLafferty FW (1966) *Chemical Communications* **78**: 80.

McLafferty FW and Gohlke RS (1959) *Analytical Chemistry* **31**: 2076–2082.

Morton TH (1983) *Tetrahedron* **38**: 3195–3243.

Rosenstock HB and Krauss M (1968) *Adv. Mass Spectrometry* **4**: 523–545.

Schwarz H (1981) *Topics in Current Chemistry* **76**: 1–31.

Splitter JS and Turecek F (eds) (1994) *Applications of Mass Spectrometry to Organic Stereochemistry*. Weinheim: VCH.

Stirk KM, Kiminkinen LKM and Kenttämaa HI (1992) *Chemical Reviews* **92**: 1649–1665.

Traeger JC, McAdoo DJ, Hudson CE and Giam CS (1998) *Journal of American Society Mass Spectrometry* **9**: 21–28.

Wurster C and Sendtner R (1879) *Berichte der deutschen chemischen Gesellschaft* **12**: 1803–1813.

Yoshida K (1984) *Electrooxidation in Organic Chemistry*. New York: Wiley.

# FT-Raman Spectroscopy, Applications

**RH Brody**, **EA Carter**, **HGM Edwards** and
**AM Pollard**, University of Bradford, UK

Copyright © 1999 Academic Press



## Introduction

A major advantage proposed for the application of Raman spectroscopy to many problems is the minimal sample preparation required for presentation of the specimen to the spectrometer. Unlike most other analytical techniques, no chemical or mechanical pretreatment is necessary; of particular relevance to biological and biomedical studies is the ability to obtain Raman spectra from specimens which are in their state of natural hydration without further desiccation. The weak Raman scattering of hydroxyl groups and silica means that water and glass will not strongly affect the observation of Raman spectra. In other cases, specimens which have been prepared for optical microscopy and protected using standard glass cuvettes or slide cover-slips can also be studied using a Raman microscope without any changes being effected.

The swamping of weaker Raman spectra by strong fluorescence emission in the visible region of the electromagnetic spectrum is an occupational hazard of Raman spectroscopy using laser excitation, particularly in the blue and green regions (400–520 nm). This arises from the high latent energy of excitation at these wavelengths. In some cases sample fluorescence arises from impurities in the specimen, and in these cases sample treatment may be indicated. This may be a trivial or difficult task, depending on the particular specimen being studied. For example, the swabbing of biological tissue, which is common practice to remove surface contaminants, is neither effective nor admissible when archaeological biomaterials are being studied, which may have involved the absorption and concentration of fluorescent materials over a long period of time.

It has often been reported that prolonged exposure of a specimen to a laser beam will produce a Raman spectrum with enhanced signal-to-noise (the little-understood physical phenomenon of 'fluorescence burnout' which is illustrated in **Figure 1**); by this means, good quality Raman spectra can be recorded in the presence of fluorescent material. However, the removal of the fluorescence emission is only temporary and the method is not always appropriate. Some biomaterials, such as skin and wool, exhibit strong fluorescence in the blue region of the spectrum and their Raman spectra generally cannot be recorded using laser radiation around 4880 nm.

The choice of exciting line wavelength is therefore of critical importance for the observation of Raman spectra of biomaterials without attendant fluorescence emission. A particular advantage of the move towards red excitation (600–800 nm) and particularly the near-infrared (1064 nm) is the minimization of fluorescence from biological tissues and organs, such as skin, because of lower excitation energies. Other methods of combatting fluorescence have involved the use of pulsed laser excitation and d.c.-chopping or synchronization of the scattered radiation to electronically filter out the background emission.

When coupled to a microscope a Raman instrument with a high spatial resolution (approximately 5 µm) is then available which can produce spectra that are largely fluorescent-free. The Raman microscope exhibits a high signal-to-noise ratio and is ideal for the characterization of weak scattering samples and analysis of specific regions within a sample.

Another attachment that can be effected to a Raman spectrometer is a remote probe. This consists of a fibre optic cable that passes the laser beam to a sample outside the conventional sampling chamber. This can be of use in many different applications, the most obvious of which is where the sample is too large or complex to fit into the instrument. In vivo biological studies utilize a fibre optic probe for the investigation of human tissue. Industrial process monitoring uses a Raman spectroscopic probe for online quality control during manufacture.

The linear dependence between intensity of scattering and species concentration means Raman spectroscopy is not considered to be a sensitive technique for trace quantities of material, unless these are localized, can be identified microscopically and have a large molecular scattering cross-section factor. However, the linearity of dependence between concentration and observed band intensities can be used to good effect in quantitative Raman spectrometry, especially for solution equilibria, for which important physicochemical data can be obtained, such as rate constants, equilibrium constants and enthalpies of chemical bond formation or breakage. Hence,

**Figure 1** Dissipation of fluorescence in wool on extended exposure to laser irradiation. (A) 0, (B) 30 and (C) 210 s. (633 nm, 5 s acquisition time).

whereas in absorption spectrometry weak bands can be enhanced using larger path lengths, this is not practicable in Raman spectroscopy.

The effect of the input of high energy visible laser radiation into samples with consequent possible increase fluorescence emission may also be marked in local sample heating. This is particularly important for valuable samples that must not suffer damage during analysis; also, coloured materials may absorb the laser radiation particularly effectively and literally burn in the process. For this reason, high laser powers focused onto stable mineral or geological materials are not acceptable for sensitive biomaterials and an experienced Raman spectroscopist would always attempt to use low laser powers initially on samples whose reaction to laser radiation is not known. Even when sample degradation or damage does not occur through laser absorption, for this reason, special devices based on a spinning sample illumination system exist where the species is permitted to have only low residence times in the laser focus, so reducing the localized heating.

With 1064 nm infrared excitation, generally used in FT-Raman studies, the effect of sample heating is noted first in emission at higher wavenumber shifts (> 2000 $cm^{-1}$), but the problem is exacerbated when the sample itself is being thermally investigated. For this reason, the 'background' emission of infrared radiation at higher wavenumbers with 1064 nm excitation varies with sample temperature until near 200 °C it becomes effectively unworkable, e.g. the Raman spectrum of $S_8$ at elevated temperatures.

## Biological materials

Fluorescence and the inherent weakness of the Raman effect seriously hampered the use of Raman spectroscopy for the routine analysis and investigation of biological samples. Until the commercial development of the FT-Raman technique in 1986 some researchers went to quite extreme lengths to overcome these problems. This is well demonstrated by workers who found it necessary to 'burn out'

fluorescence in wool for up to 20 hours before being able to acquire an acceptable Raman spectrum.

The keratins are a specialized group of structural fibrous proteins that are characterized by their high cystine content and can be classified according to the amount of sulfur present in the protein. Structures such as wool, hair, hooves, horns, claws, beaks and feathers are classified as 'hard' keratins because the sulfur concentration in these proteins is greater than 3%. Keratin proteins containing less than 3% sulfur, such as the stratum corneum (the outermost layer of skin), are classified as 'soft' keratins. Although these keratin proteins have a similar molecular composition major spectral differences have been observed in: the intensities of the C–S and S–S stretching vibrations; the conformation of the disulfide bond; and the position, line shape and bandwidth of vibrations associated with the proteins secondary structure.

**Figure 2** shows the FT-Raman spectra of wool (**Figure 2A**) and hydrated stratum corneum (**Figure 2B**) over the wavenumber regions 2600–3600 $cm^{-1}$ and 400–1800 $cm^{-1}$. The presence of water near the $\nu$(CH) stretching region can be clearly seen; the characteristic amide I, $\nu$(CONH), $\delta(NH_2)$ and $\delta(CH_2)$ modes are seen in the region 1200–1700 $cm^{-1}$. Lipid modes are near 1000–1200 $cm^{-1}$ and the prominent $\nu$(SS) band in wool near 500 $cm^{-1}$ is also noteworthy.

To date, Raman spectroscopy has been used for the characterization of stratum corneum, epidermal membrane and dermal tissue. Of current interest is the administration of therapeutic agents across the skin barrier (transdermal drug delivery); however, there are a number of difficulties which involve the supply, storage and use of biohazardous human material.

## Medical/pharmaceutical

The use of Raman spectroscopy in medical diagnostics is an ever-expanding field of application. Again, features such as the nondestructive and noninvasive nature of the technique mean that the technique is well suited to this type of application. Raman spectra have often been masked by fluorescence from some of the naturally occurring constituents of animal tissues. Raman spectra of cells and tissues are

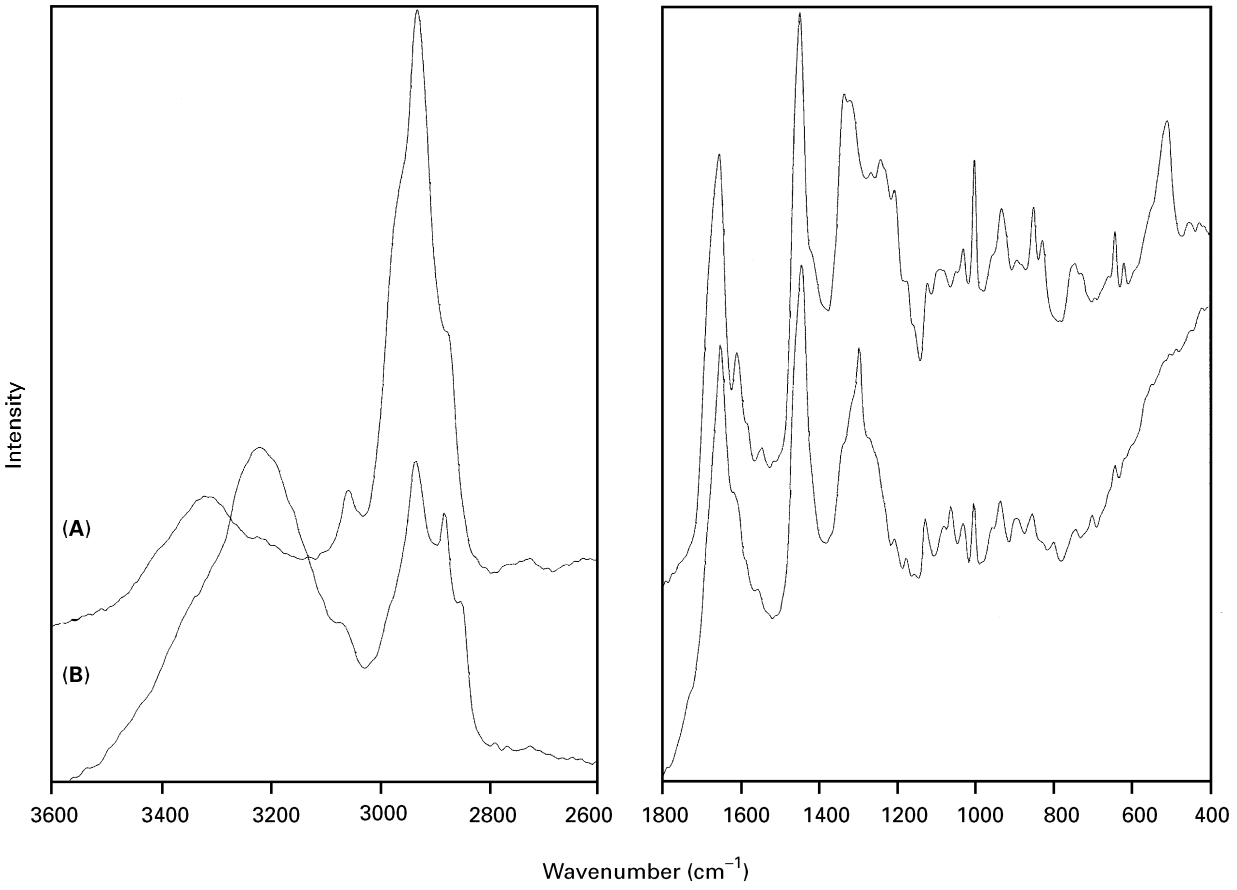



**Figure 2** FT-Raman spectra of (A) untreated Merino wool, and (B) hydrated stratum corneum.

usually excited in the near infrared with radiation at 780, 830 or 1064 nm as the probability of exciting fluorescence is reduced at these longer wavelengths.

Raman spectroscopy has been used for the compositional analysis of gallstones and kidney stones; the distribution of the mineral and organic components of a human tooth have been mapped by FT-Raman microspectroscopy; the characterization of normal and pathological tissues including breast, liver, colon and brain. Furthermore, analysis can be performed *in vitro*, on biopsy samples, or *in vivo* whereby spectra are recorded using a fibre optic cable. Examination of complex Raman spectra that contain a large number of bands as in biological tissues can be somewhat subjective, and it is often difficult to detect subtle changes or differentiate between similar spectra. To overcome this problem it is necessary to automate the interpretation of spectra by developing a non-subjective technique for analysis, such as a neural network. To date, neural networking has been successfully used to differentiate between Raman spectra of basal cell carcinoma and healthy skin.

Interactions between hydroxyapatite crystallites, inorganic ions, cellular components and the organic matrix in living tissues form part of the complex process known as biomineralization. An example of biomineralization is the production of calcified tissues such as bone and teeth. The Raman spectrum of bone is dominated by a very intense vibration at 960 $cm^{-1}$, which is attributed to the phosphate stretch of the hydroxyapatite component.

Bone implants are commonly made of metal coated in a similar material to bone, i.e. synthetic hydroxyapatite, to improve the biocompatility and to aid bonding between the natural bone and the implant. Raman studies of bone implants utilize the spectral difference between bone and that of synthetic hydroxyapatite. The synthetic material lacks many of the characteristic Raman vibrations of bone and those of phosphate have significantly reduced bandwidths. Thus, it is possible to study the interface between the implant and the recipient's bone. Biomineralization of implants coated with materials other than hydroxyapatite can be followed by monitoring the intensity of the phosphate vibration related to those produced by the coating material.

## Art and archaeology

The identification of materials that make up an object is important to art historians and archaeologists for a variety of reasons. Information can be provided on trade routes that existed at the time, the development of artistic styles and techniques and the understanding of chemical technology. This information may help to authenticate and date an object, and aid in the process of restoration and conservation. Identification of the products of degradation will help in devising a regime for treatment to arrest or reverse the deterioration, thus aiding the conservator in the preservation of the object. In many cases previous conservation treatments are a problem as the materials used breakdown and damage the object. Often treatment records have not been kept, or are lost, and analysis is required to identify the materials previously employed so they can be successfully removed.

Some problems have been found in analysing archaeological materials from certain burial environments where absorption of coloured and/or fluorescent impurities has taken place. Special care has also to be taken when handling rare or very fragile samples. Often very low laser powers are necessary in order to prevent thermal damage from occurring. These problems are being overcome and Raman spectroscopy is now being applied to the identification of a diverse range of material types including pigments, resins, waxes, gums, ivory and ancient skin.

### Pigments

Pigments can be found on a wide range of objects such as paintings, frescoes, manuscripts and china. Often only a very small section of the pigment will remain perhaps only several grains. It is important to be able to identify the pigment (often *in situ*) without harming the object in anyway. Raman microscopy has been found to be one of the best techniques for identifying pigments. It is sufficiently sensitive to analyse pigment grains, generally does not suffer from interference (from surrounding media such as binders) and most importantly it is nondestructive. It is also possible to identify components in a pigment mixture, as the spatial resolution of the technique is about $\leq$ 2–10 μm. A wide range of pigments have been identified using Raman spectroscopy; for example, it is possible to differentiate between the red pigments: red lead, red ochre and vermilion used in antiquity and to determine the mixture composition quantitatively and nondestructively (see **Figures 3B**, **3C** and **3E**).

Pigments used in the wall paintings in Winchester Cathedral and Sherborne Abbey have been analysed using FT-Raman spectroscopy. Wall paintings from circa 1175–1185 are found in the Holy Sepulchre Chapel, North Transept of Winchester Cathedral. They illustrate several scenes around the Crucifixion and are amongst the most important paintings of

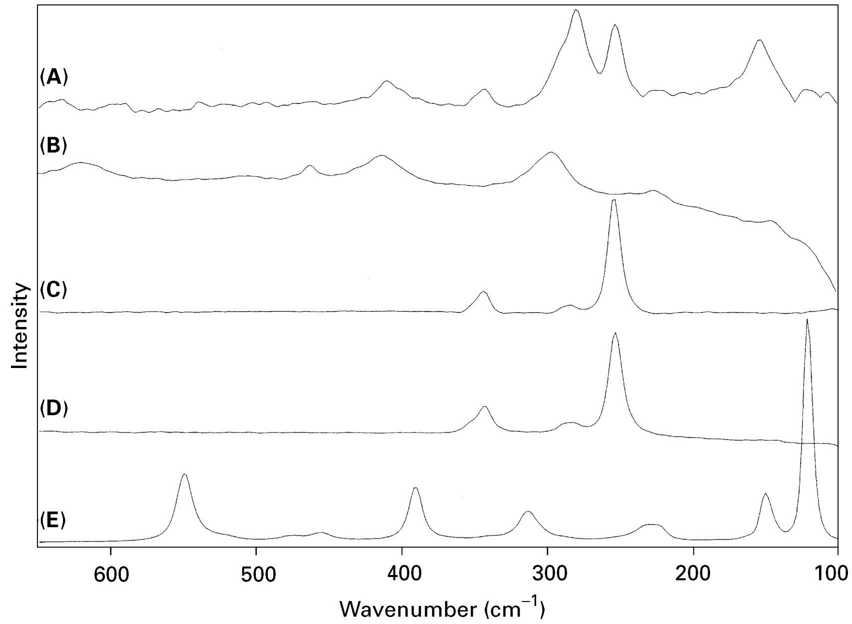



**Figure 3** FT-Raman spectra of (A) pigment from wall painting in Winchester Cathedral (baseline corrected) (B) red ochre, (C) vermilion, (D) pigment from wall painting in Sherborne Abbey, and (E) red lead pigment.

their period in the UK. Those found at the Chapter House at Sherborne Abbey are much more fragmentary and date to the late 12th/early 13th Century. The Raman spectra shown in **Figure 3** demonstrate that the pigment from Sherborne Abbey (**Figure 3D**) is pure vermilion (**Figure 3C**) whereas that from Winchester Cathedral (**Figure 3A**) is a mixture of red ochre (**Figure 3B**) and vermilion (**Figure 3C**). Adulteration of pigments was practiced for economic reasons even though it was recognized that the resulting pigment mixture stability could be adversely affected. Also, variation in pigment composition over a painting could indicate to a conservator that some more recent renovation has been undertaken.

Analysis of frescoes at the palazzo Farnese, a mansion at Caprarola, Italy, helped to characterize chemically the biodeterioration that had occurred. A circular courtyard at the mansion has inner walls at ground and first floor levels bearing frescoes painted by Zuccari in the 1560s. In recent years serious biodeterioration through lichen invasion has affected the frescoes. *D. massiliensis* forma *sorediata*, which produces oxalate encrustations which are almost 2 mm thick in some places is responsible for the disfigurement and the chemicals involved in this process have been identified by Raman spectroscopy.

### Archaeological resins

Raman spectroscopy can be used to identify different natural resins. By comparing spectra from ancient resins with a database of modern materials it is sometimes possible to identify the source of the ancient sample. An example of this use is the identification of the different forms of the resin 'dragon's blood'. Since ancient times this resin has been used for a diversity of medical and artistic purposes. This rich red resin has been used to make varnishes for a variety of objects including reputedly 16–18th century Italian violins such as those made by Stradivarius. Many cultures have at least one indigenous natural resin which they call 'dragon's blood' but the botanical sources are not necessarily the same. Today, the primary commercial supply of dragon's blood is the resin from the fruit of the Southeast Asian rattan, or cane-palm, *Daemonorops draco* (*Palmae*). In Western antiquity, in the Roman Empire, one of the main sources was *Dracaena cinnabari* L., which was once prevalent around the Mediterranean coast, but now endemically localized to Socotra Island, off the Horn of Africa. **Figure 4** presents the FT-Raman spectra of these two resins. Clearly the spectra are very different and provide a means of identification of a botanical source of 'dragon's blood' resin. Elsewhere, other trees have been tapped for their 'dragon's blood' such as *Eucalyptus resinifera* (actually a gum) in Australia. The FT-Raman technique is capable of distinguishing between fresh resins from these three sources mentioned. Identification of the resin can play a deciding role in the cleaning or stabilization of an object coated in resin and with the provenancing of the artefact.

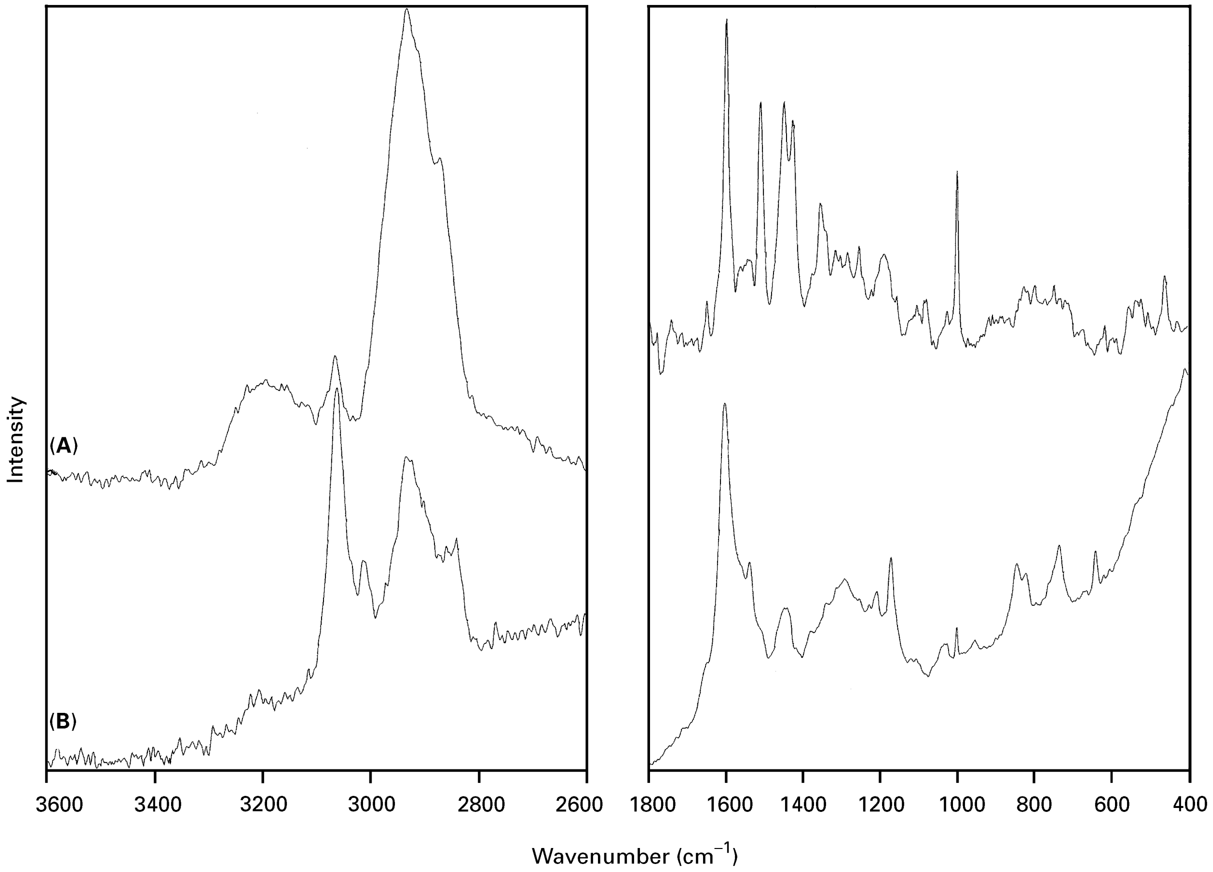



**Figure 4** FT-Raman spectra of (A) *Daemonorops draco (Palmae)*, and (B) *Dracaena cinnabari*.

## Forensic science

The field of forensic science, scientific analysis used within the legal system, is very wide-ranging; narcotics and explosives are amongst the areas of analysis covered. Many different techniques are used within the field to identify samples found at a crime scene or seized by police. Techniques used must fulfil certain requirements: sensitive enough to deal with often very small or trace samples, provide clear, characteristic results which easily distinguish different materials and minimal sample size and preparation. The latter point is one which Raman spectroscopy can easily fulfil, as the technique is both nondestructive and noninvasive. No sampling is required, which means that further analysis is possible should it be necessary, for example in the case of an appeal. Raman spectroscopy fulfils the other requirements as it is both sensitive and provides characteristic spectra for different substances.

Raman spectroscopy has been applied to a range of different materials of interest to the forensic scientist. A wide range of narcotics can be identified including heroin, cocaine and amphetamines. The wavelength of excitation has to be varied, as some samples are highly fluorescent under certain conditions. For example, heroin fluoresces when excited at 633 nm but a clear, characteristic spectrum can be obtained with resonance Raman using 244 nm excitation. Fluorescence can also be a problem if a narcotic has been ‘cut’ (mixed) with a fluorescent material. The spectra of these ‘street’ drugs require baseline correction in order to be successfully compared with a database of ‘pure’ standards or again a lower wavenumber excitation can be employed. Narcotics often enter the laboratory in a crime scene bag and may also still be in their original packaging – often wrapped in plastic. It has been found that it is possible to take a Raman spectrum of the sample through the packaging (including black plastic), subtracting the spectrum of the packaging to get a spectrum of the sample alone. This is very useful in forensic science as this minimizes any contamination that could occur to the sample during preparation for analysis. Along with the lack of any sample preparation required and with provision of a suitable database, a computerized analytical instrument means that relatively unskilled users

could potentially use Raman spectroscopy to analyse any seizure more quickly.

Raman spectroscopy has also been applied to the identification of different explosives. Most explosives, both nitro-containing (i.e. 2,4,6-trinitrotoluene (TNT)) and non-nitro-containing (i.e. TATP), produce high quality, low fluorescence Raman spectra. Some plastic explosives (i.e. Semtex) have some fluorescence originating from the binder materials but this can be overcome by use of anti-Stokes bands and Boltzmann correction of the data.

Raman microscopy has been used for the analysis of gunshot residues. Chemical analysis of the materials that leave the barrel of a gun on firing can give information about the gun, the ammunition, the shooting distance and direction. All of this information is important for investigators of any shooting incident. The residues that leave a gun on firing are deposited on the victim, user, weapon and within the immediate vicinity. Gunshot residues are composed of the metallic components of the bullet, partly burnt propellant particles and particles of elements originating from the primer, the bullet, the propellant and the propellant additives. Raman microspectroscopy has illustrated that different metal anions contained in gunshot residue such as carbon, barium carbonate, lead sulfate, lead oxide and mixtures of these compounds are easily identified.

### Real or fake?

An area that combines the interests of historians and forensic scientists is whether an object has been constructed from the material from which it is purported to have been made. Rare or valued materials are often replaced by cheaper, more easily available imitations, some of which are of such good quality that they deceive even expert examination. Examples of these materials are ivory and amber, both having been simulated using a wide range of materials, including modern polymers. In some instances great effort is given by the forger to copy the weight and texture of the original material. Ivory is much heavier than many modern plastics so fillers such as calcite (calcium carbonate) may be added so that the weight then more closely matches that of ivory. On highly carved objects it can be very difficult to identify the material as genuine ivory, for example, or an imitation. The 'hot pin' test has been used to identify modern plastics, over ivory. Plastic materials will tend to melt or burn when a hot pin is applied but it will have no effect on genuine ivory. This rather crude method has one quite obvious drawback in that a plastic object will be marked. Even though the item is not made of ivory it may still be of some value, which would be affected by any marking. FT-Raman spectroscopy provides a good alternative; a result can be obtained almost as quickly and the object is not damaged. The spectra of ivory and amber are clearly different to polymeric imitations, so a definitive identification can be given where previously it has proved difficult, or impossible, to do so visually. This is illustrated in **Figure 5** which presents the FT-Raman spectra of African elephant tusk, micarta (phenolic thermoset resin), an imitation ivory and calcite. It can be seen that the spectrum of the imitation ivory contains vibrational features which are characteristic of polymer mixtures, in this case polystyrene and polymethylacrylate. The spectrum of micarta indicates that it probably contains either polyethylene or polybutylene terephthalate.

## Gemmology

Analytical techniques commonly used in gemmology include X-ray and neutron diffraction, scanning electron microscopy and, more recently, FT-Raman microspectroscopy. Traditional identification is based on the gems' unique physical, chemical and optical properties. These include specific gravity, cleavage, hardness, toughness, fracture, refraction, transparency, lustre and sheen. However, some gemstones have similar properties for example taaffeite ($BeMgAl_4O_8$) and musgravite ($Mg_2Al_6BeO_{12}$) have essentially identical chemical composition and gemmological properties. The advantage of utilizing Raman spectroscopy for analysis lies in the non-destructive nature of the technique, rapid identification and the ability to analyse both carved and rough surfaces.

Other applications include the detection of synthetics and imitations, the detection of composite or assembled stones and the investigation of inclusions to assist in the identification of the origin of the gemstone. In order to hide surface cracks, improve colour or provide protection for soft stones, gemstones may undergo certain enhancement treatments. For example, they may be treated with oil, artificial resins or waxes to fill any fissures or fractures thus improving their clarity. These foreign substances produce distinctive Raman spectra from which their presence may be identified.

## Polymers and plastics

Characterization of polymers and their reaction processes can be carried out using Raman spectroscopy. Both qualitative and quantitative information can be provided about stereoregularity, chemical nature,

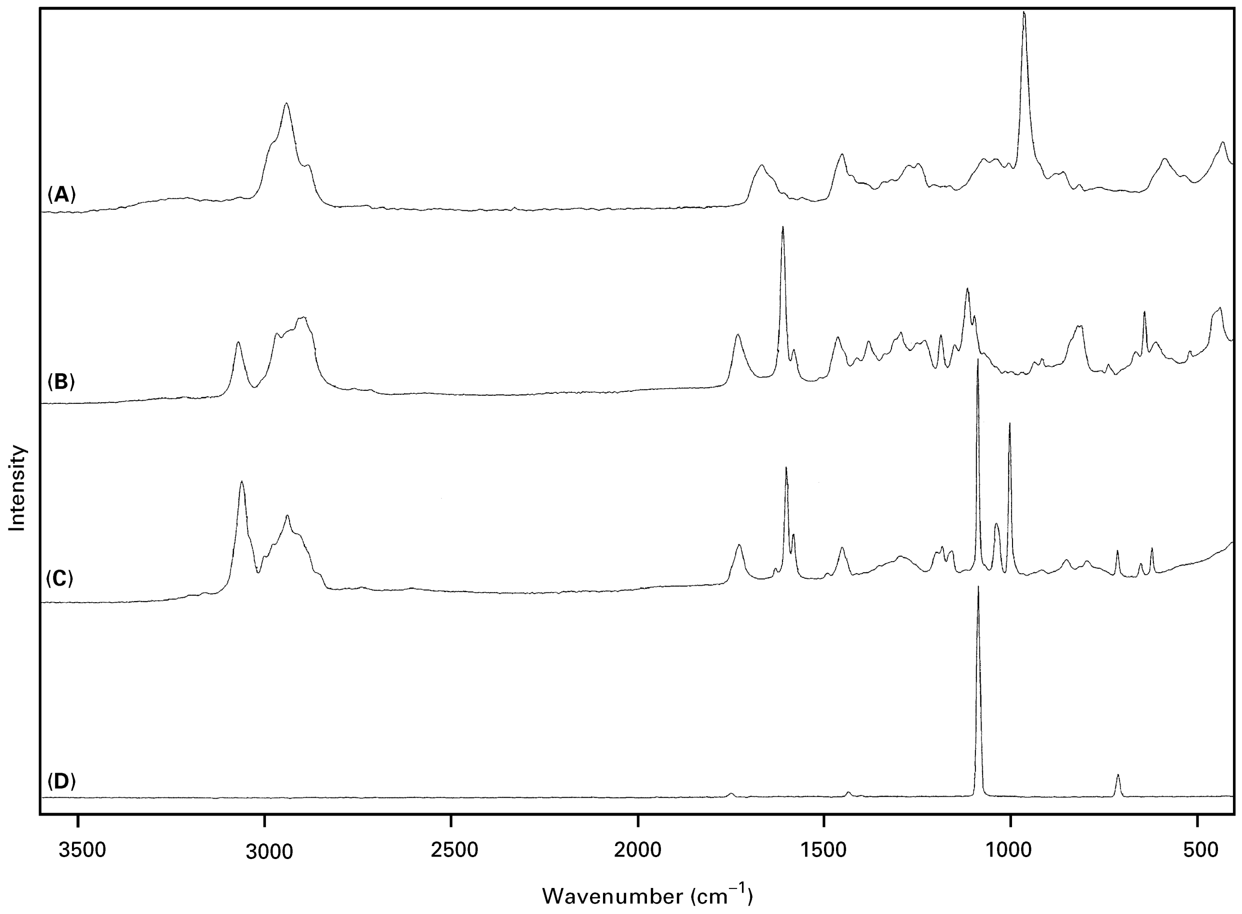



**Figure 5** FT-Raman spectra of (A) African elephant ivory, (B) micarta, (C) an imitation ivory and (D) calcite.

orientation, conformation and three-dimensional state of order in a polymer. The polymerization process may also be monitored, allowing the effect of different parameters on the course of the reaction to be analysed. Many different polymers and polymer reactions have been studied using Raman spectroscopy in recent years. For example, epoxysilanes that are used to obtain hybrid polymers via a sol-gel process.

Process control and remote Raman sensing are now possible through the application of fibre optics. Remote measurements can now be taken directly in the reaction vessel. The real-time monitoring of chemical reactions can aid in the streamlining of the process and thus produce cost savings. The minimal fluorescence produced using near-infrared excitation means that FT-Raman spectroscopy is becoming widely accepted as an industrial tool. An example is the analysis of waterborne polymer emulsions, i.e. styrene/butadiene copolymers, which are used as binders in coated paper and carpet manufacture. Raman spectroscopy is useful in this case due to little interference being caused by the aqueous phase or the scattering effects from the emulsion particles, which can be a problem with other spectroscopic techniques.

Assessment of the mechanical properties and engineering applications of polymers is made by consideration of orientation, conformation and three-dimensional state of order of the polymer during the deformation processes. FT-Raman spectroscopy is also applied to the analysis of transient structural changes induced by polymer deformation. For a clear understanding of a deformation mechanism, characterization of structural changes is required online. Studies have been performed on a wide variety of polymers during elongation and recovery, to monitor phase transitions and alterations in crystallinity and anisotropy. Particularly useful for this purpose has been the simultaneous analysis of mechanical and spectroscopic properties, called rheooptical measurements.

## Inorganic/organometallic materials

A typical FT-Raman spectrum covers the wavenumber range 100–3500 $cm^{-1}$, which encompasses the

skeletal vibrational modes of organic molecules normally occurring in the functionality region (1000–2000 $cm^{-1}$) and the bonds between heavy metals and carbon, oxygen and sulfur which all occur below 500 $cm^{-1}$. This ability to record the spectra of ‘inorganic’ and ‘organic’ components within a sample is a valuable attribute of Raman spectroscopy. For example, the Raman spectrum of mercury (II) sulfide (commonly known as vermillion), shown in **Figure 3C**, exhibits a strong Raman band due to $\nu$(HgS) at 257 $cm^{-1}$ which is not normally seen in infrared spectrometry unless special instrumentation is used.

The ability of Raman spectroscopy to detect coordinate bands in inorganic and organometallic compounds lends itself particularly well to the observation of metal–metal, metal–halogen and metal–phosphorus clusters such as $Ag_4I_4(PMe_3)_4$, metals–oxygen bonds in corrosion studies involving zirconium(IV) oxide and iron(III) oxide in nuclear fuel claddings and steels, and metal–carbon bonds in organometallic compounds used in catalysis.

Studies of metal ion coordination with anion ligands in mixed organic/aqueous solvent systems have provided novel information about the metal ion coordination numbers and mechanism of crystallization processes from Raman spectroscopy. Also, thermodynamic parameters such as enthalpy of band formation or breakage have been evaluated from Raman spectra of solution studied over a range of temperatures.

## Geological materials/solid states

Access to the low-wavenumber regions of the vibrational Raman spectra of minerals and solids of geological relevance have yielded information on crystal lattice symmetries and on the effect of ionic impurities. Databases of geological materials now rival those of organic compounds and affiliations to solid-state physics, crystal engineering and to substrate characterization for thin-film devices are numerous.

Some of the most recent affiliations in this area include the analysis of extraterrestrial material in the form of Martian meteorites (Shergotty–Nakhla–Chassigny classification) and in the characterization of simulated soils for planetary exploration.

*See also:* **Art Works Studied Using IR and Raman Spectroscopy; Astronomy, Applications of Spectroscopy; Biochemical Applications of Raman Spectroscopy; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy; Forensic Science, Applications of Atomic Spectroscopy; Forensic Science, Applications of IR Spectroscopy; Forensic Science, Applications of Mass Spectrometry; Forestry and Wood Products, Applications of Atomic Spectroscopy; Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy; Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; IR Spectral Group Frequencies of Organic Compounds; Matrix Isolation Studies By IR and Raman Spectroscopies; MRI of Oil/Water in Rocks; Nonlinear Raman Spectroscopy, Applications; Nonlinear Raman Spectroscopy, Instruments; Nonlinear Raman Spectroscopy, Theory; Polymer Applications of IR and Raman Spectroscopy; Rayleigh Scattering and Raman Spectroscopy, Theory; Surface-Enhanced Raman Scattering (SERS), Applications; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading

Carmona P, Navarro R and Hernanz A (eds) (1997) *Spectroscopy of Biological Molecules*: Modern Trends. Netherlands: Kluwer Academic Publishers.

Chase DB and Rebelled JF (eds) (1994) *Fourier Transform Spectroscopy from Concept to Experiment*. Cambridge: Academic Press.

Edwards HGM (1996) Raman Spectroscopy Instrumentation. In: *Encyclopedia of Applied Physics*, Vol 16, pp 1–43. New York: American Institute of Physics, VCH Publishers.

Hendra PJ (ed) (1997) Special issue – applications of Fourier transform Raman spectroscopy VI. *Spectrochimica Acta* **53A**: 1–128.

Hendra PJ (ed) (1997) Special issue – applications of Fourier transform Raman spectroscopy VII. *Spectrochimica Acta* **53A**: 2245–2422.

Hendra PJ and Agbenyega JK (1993) *Raman Spectra of Polymers*. Chichester: Wiley.

Hendra PJ, Jones C and Warnes G (1990) *Fourier Transform Raman Spectroscopy: Instrumentation and Chemical Applications*. New York: Ellis Horwood.

Heyns AM (ed) (1998) *Proceedings of the 16th International Conference on Raman Spectroscopy*. Chichester: John Wiley & Sons.

Lawson EE, Barry BW, Williams AC and Edwards HGM (1997) Biomedical applications of Raman spectroscopy. *Journal of Raman Spectroscopy* **28**: 111–117.

Special Issue – Raman Spectroscopy in Art, Medicine and Archaeology. *Journal of Raman Spectroscopy* **28**: 77–197.

Special Issue – Raman Spectroscopy (Commemorative Issue). *Journal of Raman Spectroscopy* **26**: 587–927.

Turrell G and Corset J (eds) (1996) *Raman Microscopy: Developments and Applications*. London: Academic Press Limited.

# Gas Phase Applications of NMR Spectroscopy

**Nancy S True**, University of California, Davis, CA, USA

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

NMR spectroscopy has been used to study many aspects of the physics and chemistry of molecules in the gas phase for more than 50 years. The third report of the observation of the magnetic resonance phenomenon, which was published in 1946, described a study of $H_2$ gas. The signal/noise ratio obtained from a 10 atm $H_2$ sample (line width 600 ppm at 0.68 T) was ~15/1 in that study. Nuclei in larger molecules have longer relaxation times and in many cases produce NMR spectra with natural line widths of less than 1 Hz. Currently, it is possible to obtain $^1$H and $^{19}$F spectra with good signal/noise ratios from gases present at partial pressures of less than 1 torr. With this sensitivity, NMR spectroscopy can be used to address many questions about molecular interactions and dynamics in the gas phase. Experimentally measured chemical shifts and spin–spin coupling constants of gas-phase molecules are a benchmark for theoretical calculations of NMR parameters. Temperature- and pressure-dependent relaxation rate constants, $T_1$, $T_2$, and $T_{1\rho}$, have been measured for several different nuclei in diatomic and small polyatomic molecules in the gas phase. These data provide information about relaxation mechanisms, rotationally inelastic collision cross sections, coupling constants and intermolecular forces. NMR spectroscopy is an important tool for studying hydrogen bonding in the gas phase. Thermodynamic parameters for keto–enol tautomeric equilibria and conformational equilibria of gases have been determined from measurements of temperature-dependent relative intensity ratios. Temperature- and pressure-dependent rate constants for low-energy intramolecular chemical exchange processes of gases including conformational interconversion, ring inversion and Berry pseudorotation have been measured using NMR spectroscopy. The motion of xenon atoms diffusing in zeolite matrices and the motion of gas-phase molecules in protein cavities can be studied using NMR spectroscopy. Gas-phase NMR spectroscopy can be used to study chemical reactions involving volatile species. Density-dependent chemical shielding and gas-phase spin–lattice relaxation were reviewed a few years ago. The application of gas-phase NMR spectroscopy to chemical exchange processes was reviewed more recently.

## Principles

### Gas-phase chemical shifts and coupling constants

The general features of high-resolution NMR spectra of gases resemble those of liquids, but the actual resonance frequencies in the gas phase and in solution are different owing to differences in the local environment. A gas sample has a smaller bulk magnetic susceptibility and weaker intermolecular interactions than a liquid. The bulk magnetic susceptibility difference causes the entire gas-phase NMR spectrum to be offset a few ppm from that of the corresponding liquid. The direction and magnitude of the offset depend on the magnetic field strength, the shape of the sample, its orientation in the magnetic field and its chemical composition. The magnetic field experienced by the sample, $B_0(s)$, is $B_0(s) = B_0[1+(4\pi/3 - \alpha)\kappa_s]$, where $B_0$ is the permanent magnetic field, $\kappa_s$ is the volume susceptibility, a negative quantity which is much smaller for a gas than for a liquid owing to density differences, and $\alpha$ is a shape factor that is zero for a cylinder aligned along the magnetic field axis and $2\pi$ for a cylinder aligned perpendicular to the field axis. For the parallel arrangement of a cylindrical sample in a 7.2 T superconducting magnet, the gas-phase spectrum is shifted a few ppm to higher frequency from the spectrum of the corresponding liquid.

In the gas phase, intramolecular factors primarily determine the relative resonance frequencies of the nuclei in a molecule. In a liquid, intermolecular interactions are also an important factor and the relative resonance frequencies of nuclei located at different sites within a molecule can be phase dependent. The largest phase-dependent shifts in relative resonance frequencies occur for nuclei in positions near the molecule's exterior and for nuclei capable of strong intermolecular interactions such as hydrogen bonding. **Figure 1** shows gas-phase NMR spectra of a series of alcohols obtained at 421 K. The major difference between these spectra and those of the corresponding liquid alcohols is the position of the hydroxyl proton resonance which, in each case, is near 0 ppm relative to gaseous TMS. The hydroxyl proton resonances of the corresponding alcohols in

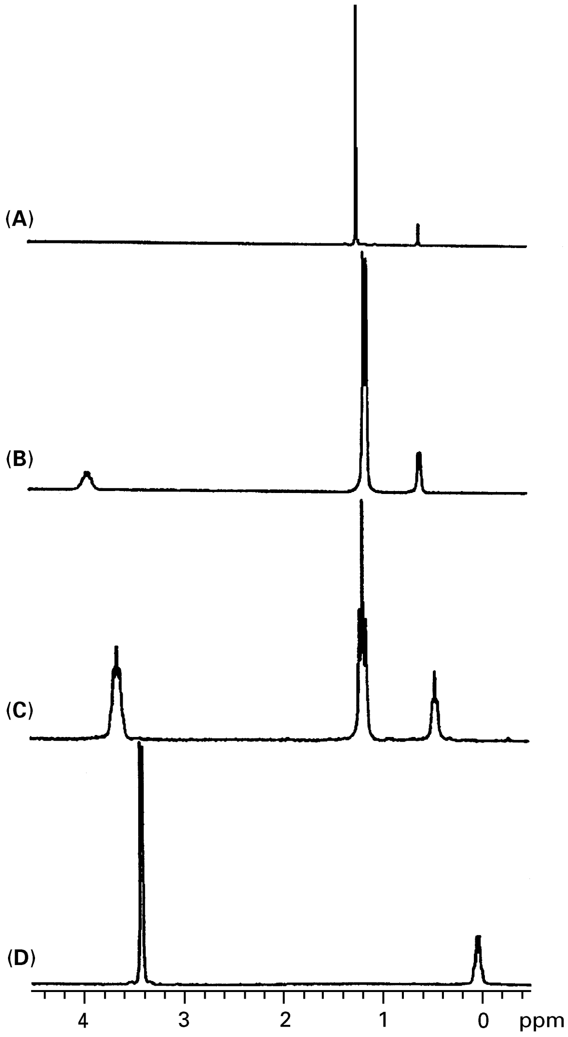



**Figure 1** $^1H$ NMR spectra of gaseous *t*-butanol (A), isopropanol (B), ethanol (C), and methanol (D) at their respective vapour pressures at 421.8 K. Reprinted from Chauvel JP Jr and True NS (1985) Gas-phase NMR studies of alcohols. Intrinsic acidities. *Chemical Physics* **95**: 435–441, © 1985, with permission from Elsevier Science.

the liquid phase are several ppm lower than that of liquid TMS owing to participation in intermolecular hydrogen bonding. In contrast, the position of the resonances of the protons attached to the carbon atoms relative to TMS is similar for both gas-phase and liquid alcohols. For cases where intermolecular interactions are weaker, gas and solution spectra are more similar, but small relative changes are often observed. For example, the limiting $^1H$ methyl chemical shift differences for dimethylacetamide and dimethylformamide are slightly smaller in the gas phase than in solution.

The temperature and density dependences of the chemical shifts of nuclei of gas-phase molecules have received significant theoretical and experimental attention. The chemical shifts of most nuclei at a constant temperature exhibit a linear density dependence for moderate densities up to ~50 atm. Intermolecular effects are deshielding for almost all gas-phase molecules, except for molecules with $\pi^*$ electronically excited states. The temperature dependence of the chemical shift of nuclei of gas-phase molecules at constant gas density can be used to probe the effects of ro-vibrational shielding. Deshielding occurs with increasing temperature as the population of rotationally and vibrationally excited states increases and the average internuclear distances in the molecule increase.

The phase dependence of spin–spin coupling constants is usually small since spin–spin coupling mechanisms are primarily intramolecular in nature. However, changes in conformational equilibria and molecular geometry can result from intermolecular interactions and can affect the magnitude of spin–spin coupling constants. For example, the magnitude of the spin–spin coupling between two nuclei separated by three bonds is dependent on the effective torsional angle and will be phase dependent if the conformational equilibrium established in the gas phase differs from that in solution. For substituted ethanes such as dimethoxyethane, the average $^3J_{HH}$ coupling constants in the gas phase have been measured as a function of temperature and used to determine conformational equilibria and barriers to internal rotation using models based on the Karplus equation. Solvent effects are determined by comparison with results of similar experiments performed on condensed-phase samples. Small phase-dependent differences in the $^1J_{NH}$ and $^2J_{NH}$ coupling constants of formamide have been observed and can be attributed to changes in bond lengths that occur upon solvation. For molecules that do not exhibit conformational isomerism or form strong intermolecular interactions, gas-phase and solution-phase spin–spin coupling constants are very similar.

## Relaxation

Nuclear spin–lattice relaxation of gas-phase molecules occurs primarily via the spin–rotation (SR) mechanism. The magnitude of the magnetic field generated by the rotational motion of the molecule changes at a rate that is dependent on the rotationally inelastic collision frequency. Scalar coupling of the nuclear spin angular momentum to this time-dependent field provides an efficient relaxation pathway that is generally not available in condensed phases where rotational motion is hindered. For medium-sized molecules, $^1H$ and $^{13}C$ $T_1$ values are typically on the order of a few hundred milliseconds

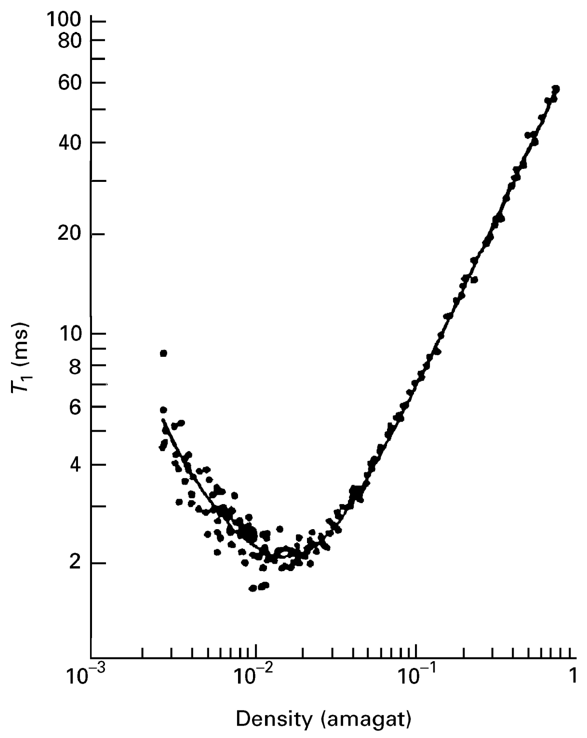



**Figure 2** The density dependence of the proton $T_1$ of ammonia gas at 300 K. The density of an ideal gas at STP is defined as 1 amagat. The solid line is the best fit of a theoretical model that allowed for multiple relaxation times to the data. Reproduced with permission of the American Physical Society from Lemaire C and Armstrong RL (1984) Proton spin longitudinal relaxation time in gaseous ammonia and hydrogen chloride. *Journal of Chemical Physics* **81**: 1626–1631.

and $^{19}F$ $T_1$ values are typically on the order of 10 to 100 ms at a gas pressure of 1 atm.

**Figure 2** shows the characteristic nonlinear density dependence of $T_1$ in the gas phase. Most NMR relaxation studies of gases report gas densities in units of amagats where 1 amagat is the density of an ideal gas at 1 atm and 273 K, i.e. $2.687 \times 10^{19}$ molecules $cm^{-3}$, or 44.612 moles $m^{-3}$. At the $T_1$ minimum, the collision frequency is approximately equal to the Larmor precession frequency of the nucleus. The $T_1$ minimum can be used to determine the rotationally inelastic cross section, $\sigma^{NMR}$, and the average spin–rotation (SR) coupling constant. The density-dependent $T_1$ data for the $^1H$ resonance of $NH_3$, shown in **Figure 2**, is consistent with a $\sigma^{NMR}$ of 105(8) Å$^2$. This value is almost twice as large as the geometric cross section of the molecule, 60 Å$^2$, indicating that long-range interactions are important in the relaxation process. The effective $^1H$ SR coupling constant of $NH_3$ is 20.3(0.1) kHz, considerably larger than the $^1H$–$^1H$ dipole–dipole coupling constant in this molecule.

Relaxation via dipole–dipole interactions and quadrupolar interactions has also been observed for nuclear spins of molecular gases. Dipole–dipole relaxation is competitive with spin–rotation relaxation of $H_2$ in the gas phase. For $H_2$, the SR coupling constant is ~110 kHz and the dipole–dipole coupling constant is 143 kHz. This mechanism has not been found to contribute significantly to spin–lattice relaxation of larger molecules, where dipole–dipole distances are greater. The lack of observable nuclear Overhauser enhancement (NOE) in gaseous benzene indicated that relaxation via dipole–dipole interactions is not significant at densities between 7 and 81 mol $m^{-3}$ at 381 K. For $^{14}N$ in nitrogen gas, the dominant relaxation mechanism is through intramolecular quadrupolar interactions. Experiments on ClF established that there is a significant contribution to $T_{1\rho}{}^F$ from modulation through quadrupolar relaxation of $^{35}Cl$ and $^{37}Cl$ of $J(^{35}Cl–F)$ and $J(^{37}Cl–F)$. The dependence of $T_{1\rho}{}^F$ on the spin-lock field strength allowed values of these two spin–spin couplings to be obtained.

Dipole–dipole interactions occurring during collisions provide a relaxation mechanism for atoms in the gas phase. Xenon-129 atoms in the gas phase have spin–lattice relaxation times on the order of minutes which are dependent on gas density, temperature and the amount of other gases in the sample. For xenon-129 adsorbed in cavities inside zeolite matrices relaxation occurs via dipole–dipole interactions with hydrogen atoms that are bound to the sides of the cavities.

## Methods

### Sample preparation

For many applications, gas-phase samples can be prepared using a routine vacuum line consisting of mechanical and diffusion pumps, a cold trap, thermocouple and capacitance manometer gauges and a manifold with attachments for sample tubes and gas inlets. NMR sample tubes are either sealed with a glass torch or constructed with self-sealing top assemblies. Preparation of samples at elevated pressures requires quantitative transfer operations. With few exceptions, gas-phase NMR samples used to study conformational processes contained a volatile bath gas in addition to the sample molecule, which usually has a low vapour pressure. Molecules with low vapour pressures tend to behave nonideally even at very low pressures and this factor must be taken into consideration when converting measured gas pressures into densities.

Sample diffusion and convection are significant problems in temperature-dependent studies and in relaxation experiments using gases. In most NMR

probes the heater is positioned below the bottom of the sample and the top of the sample protrudes above the spinner. With this arrangement a significant temperature gradient occurs along the length of a standard NMR tube when working at elevated temperatures. This causes convection in the sample and makes it impossible to establish, maintain and measure the temperature. Diffusion can cause significant line broadening at low pressures and can result in large systematic errors in relaxation measurements. To minimize these problems, short NMR tubes are frequently used for gas-phase studies. They provide better temperature control and minimize diffusion of magnetized molecules from the active volume region of the receiver during monitoring of the free induction decay. For relaxation studies of gases, the sample must be confined to the active volume of the receiver for accurate measurements of relaxation times and the tube length cannot exceed the receiver coil dimensions. Short tubes can be placed directly inside a standard size coaxial outer tubing to facilitate sample spinning or they can be attached to a spinner via a stem.

For samples containing total gas pressures of 3 atm and below, sample cells constructed from standard thickness 12, 10, or 5 mm o.d. NMR-quality glass tubing have been employed. Sample cells constructed from heavy-wall NMR-quality glass tubing, single-crystal sapphire and vespel can be used for studies of gases at pressures above 1 atm. Special precautions are necessary at all stages of assembling, testing and using any kind of high-pressure NMR apparatus because of the hazards involved. Samples in the 5–50 atm range can be prepared in small-diameter tubes with 2 mm thick walls that can withstand pressures of 50 atm. The maximum pressure is limited by the ratio of the outer diameter to the inner diameter of the tube. Pyrex and borosilicate glasses have been used. Xenon gas has been studied at pressures up to 200 atm in borosilicate glass (3.9 mm o.d./1.2 mm i.d.) Methane and ethane have been studied at pressures up to 300 atm in Pyrex tubes (5 mm o.d./0.5 mm i.d.) Sample cells constructed from single-crystal sapphire epoxied to titanium flanges that attach to vacuum lines have been used to prepare samples containing gases at pressures up to about 130 atm. A newer design allows use with 5 or 10 mm tubes, is light enough to allow spinning in most spectrometers and allows attainment of proton line widths of 0.5 Hz in 300 MHz spectrometers. Heavy-walled Vespel tubes have been used to study reactive systems at elevated pressures. This polyimide material is resistant to chemical attack but heavy walls are required and after several pressurizations the tubes tend to deform.

Specially designed NMR probes capable of withstanding high internal pressures have been described. Improved sensitivity and resolution have been obtained with toroidal probes. They consist of a toroidal-shaped RF detector inside a pressure vessel with RF and gas connectors mounted on the top and bottom, respectively, and are useful for studying pressurized reactor effluent.

### Spectral acquisition

Gas-phase NMR spectra can be acquired using procedures similar to those employed for liquid samples. Most gas-phase NMR spectra are obtained without use of a frequency lock. The field drift of most superconducting magnets is only a few hertz per day and does not contribute significantly to observed line widths of spectra acquired in a few hours or less. The absence of a lock necessitates that shimming must be done by observing the magnitude and shape of the free induction decay. Inclusion of several torr of TMS gas ensures a free induction decay of adequate magnitude for shimming. For very high-resolution work or long runs, a deuterated bath gas such as neopentane-$d_{12}$ or isobutane-$d_{10}$ can be used, but the low sensitivity in the lock channel on most spectrometers requires the presence of a partial pressure of at least 0.5 atm of these molecules to achieve a stable lock.

Acquisition of gas-phase NMR spectra is facilitated by rapid spin–lattice relaxation. This is especially true for $^{13}C$ studies of gases where $T_1$ values can be hundreds of times shorter than in condensed phases and transients can be acquired much faster. However, relaxation times and mechanisms in the gas phase preclude many experiments involving intermolecular and intramolecular polarization transfer that are routinely performed on solution samples.

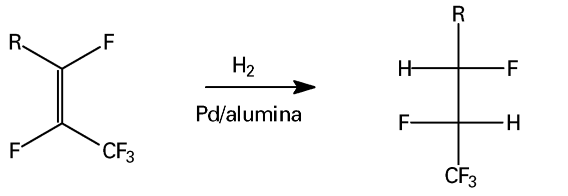
## Applications

Specific applications of gas-phase NMR spectroscopy to research in heterogeneous catalysis, conformational dynamics and absorption processes in zeolites are described below. The application of gas-phase NMR spectroscopy to many other research areas is described in the Further reading section at the end of this article.

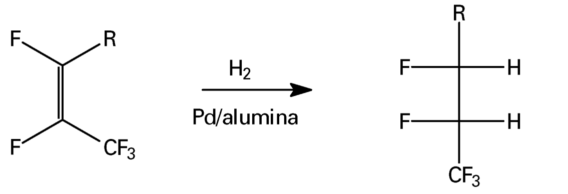
### Heterogeneous catalysis

Gas-phase NMR can be used to study the kinetics of reactions that consume or generate volatile species. In a recent study, gas-phase NMR spectroscopy was used to follow the kinetics of the hydrogenation of *cis*/*trans* mixtures of perfluoro-2-butenes ($R{=}CF_3$) and perfluoro-2-pentenes ($R{=}C_2F_5$) over palladium

supported on alumina (Eqns [1] and [2]).




[1]




[2]

Samples were prepared in 12 mm NMR tubes and contained a total gas pressure of ~3.1 atm at 333 or 338 K. The time dependence of the composition of the sample was determined from $^{19}$F NMR spectra. Since the $T_1$ values of all the $^{19}$F nuclei present are short, it was possible to obtain spectra, the result of 16 transients, with a signal/noise ratio of 1000 in 2.9 s. The data obtained in these experiments would be very labour intensive to obtain by conventional chemical techniques. **Figure 3** shows typical data obtained in this experiment. The *erythro*/*threo* product ratio was the same as the starting *cis*/*trans* reactant ratio, indicating that the olefins did not isomerize during the reaction and hydrogenation of the *cis* olefins produces the corresponding *erythro* hydrofluorocarbons while hydrogenation of the *trans* olefins produces the corresponding *threo* hydrofluorocarbons. The rate of hydrogenation was faster for the *trans* isomer of both perfluoro-2-butene and perfluoro-2-pentene, with the rate constant ratio, $k_{trans}/k_{cis}$, ranging from 2 to 3.4. It was also found that the *cis* isomers are preferentially absorbed on the catalyst, a fact that may account for their lower relative reactivity.

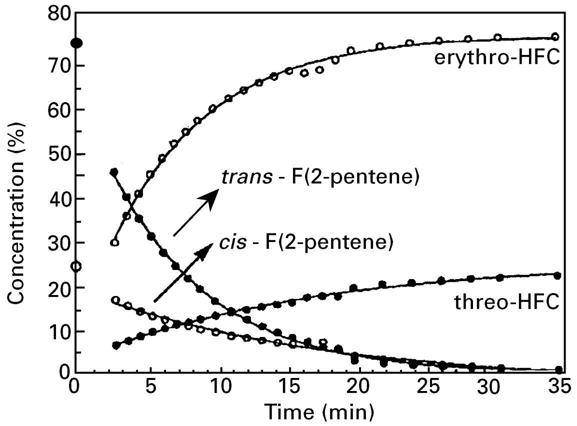



**Figure 3** Time dependence of the concentrations of reactants and products of the Pd/$Al_2O_3$-catalysed hydrogenation of a 24.8/75.2% *cis*/*trans* mixture of perfluoro-2-pentene at 65°C obtained by $^{19}$F gas-phase NMR. Reprinted with permission from Kating PM, Krusic PJ, Roe DC and Smart BE (1996) Hydrogenation of fluoroolefins studied by gas phase NMR: a new technique for heterogeneous catalysis. *Journal of the American Chemical Society* **118**: 10 000–10 001. © 1996 American Chemical Society.

## Conformational kinetics

For many gas-phase molecular processes with activation energies in the 5–20 kcal mol$^{-1}$ range, temperature- and pressure-dependent rate constants can be obtained from analysis of exchange-broadened NMR line shapes. Frequently, rate constants can be obtained with accuracy comparable to that obtained in the best liquid-phase studies. This technique can be applied successfully to gases that have at least 1 torr of vapour pressure at temperatures at which slow or intermediate exchange can be observed. Temperature-dependent gas-phase rate constants and kinetic parameters can be compared with similarly obtained parameters in condensed phases to determine the direction and magnitude of solvent effects on these processes. Gas-phase results also provide a critical test for theoretical calculations. The pressure dependence of rate constants for chemical exchange processes in the gas phase provides information on the microscopic dynamics of these processes in isolated systems.

Rate constants for the chemical exchange processes that occur in alkyl nitrites, cyclohexane, substituted cyclohexanes, sulfur tetrafluoride and formamide are pressure dependent. The mechanism for these thermally initiated, unimolecular gas-phase processes, reported by Lindemann in 1922, involves competition between the reaction and collisional deactivation of the critically energized molecule, A* (Eqn [3]).

$$
\begin{aligned}
\mathrm{A} + \mathrm{M} &\xrightarrow{k_a} \mathrm{A}^* + \mathrm{M} \\
\mathrm{A}^* + \mathrm{M} &\xrightarrow{k_d} \mathrm{A} \\
\mathrm{A}^* &\xrightarrow{k(E)} \mathrm{P}
\end{aligned}
\qquad [3]
$$

where $k_a$, $k_d$ and $k(E)$ are the rate constants for activation, deactivation and reaction. The rate constant for reaction, $k(E)$, is energy dependent. The solution of this mechanism yields the macroscopic

pseudounimolecular rate constant, $k_{uni}$ (Eqn [4]).

$$k_{uni} = \frac{1}{[A]}\frac{d[A]}{dt} = \int_{E=0}^{E} \frac{k(E)\cdot k_a/k_d}{1 + k(E)/k_d[M]} dE \quad [4]$$

At large [M] the integrand reduces to $k(E)k_a/k_d$ and $k_{uni}$ is a constant. At small [M] the integrand reduces to $k_a$[M] and $k_{uni}$ is a linear function of the gas density. The zero of energy is taken at the energy at threshold. The details of the fall-off curve, i.e. the rate constant $k_{uni}$ as a function of gas density, provides information about intramolecular energy transfer in the critically energized molecule. If this is ergodic and rapid on a timescale that is short compared to the reaction rate constant, statistical theories such as RRKM theory can be used to calculate $k(E)$. The molecular conditions that favour statistical behaviour are high densities of states in the reactant and efficient coupling mechanisms. Reactions of large molecules at high internal energies can generally be modelled adequately with statistical theories. The processes that can be studied by gas-phase NMR spectroscopy occur at much lower energies where state densities are sparse and coupling mechanisms are inefficient. The shape and location of the experimental fall-off curves of chemical exchange processes provide a test of the applicability of statistical kinetic theories to these low-energy processes.

**Figure 4** shows exchange-broadened gas-phase $^1$H NMR spectra of *n*-propyl nitrite obtained at 240.6 K as a function of $CO_2$ bath gas pressures. Each sample contained 2 torr of *n*-propyl nitrite. Rate constants for the *syn–anti* conformational exchange process, obtained from the spectral simulations, range from 315 $s^{-1}$ at 720 torr to 112 $s^{-1}$ at 11.8 torr. The pressure dependence of these rate constants was compared with predictions using RRKM theory. The agreement obtained indicated that vibrational redistribution in *n*-propyl nitrite is statistical or nearly so at the average internal energy required for the *syn–anti* conformational interconversion process, which is ~12 kcal mol$^{-1}$. The observed fall-off curve is dependent on the bath gas used in the study and relative collision efficiencies for activation of the conformational process can be obtained from studies that vary the bath gas. Smaller nitrites, $SF_4$ and formamide have lower state densities and are currently the subjects of detailed analysis.

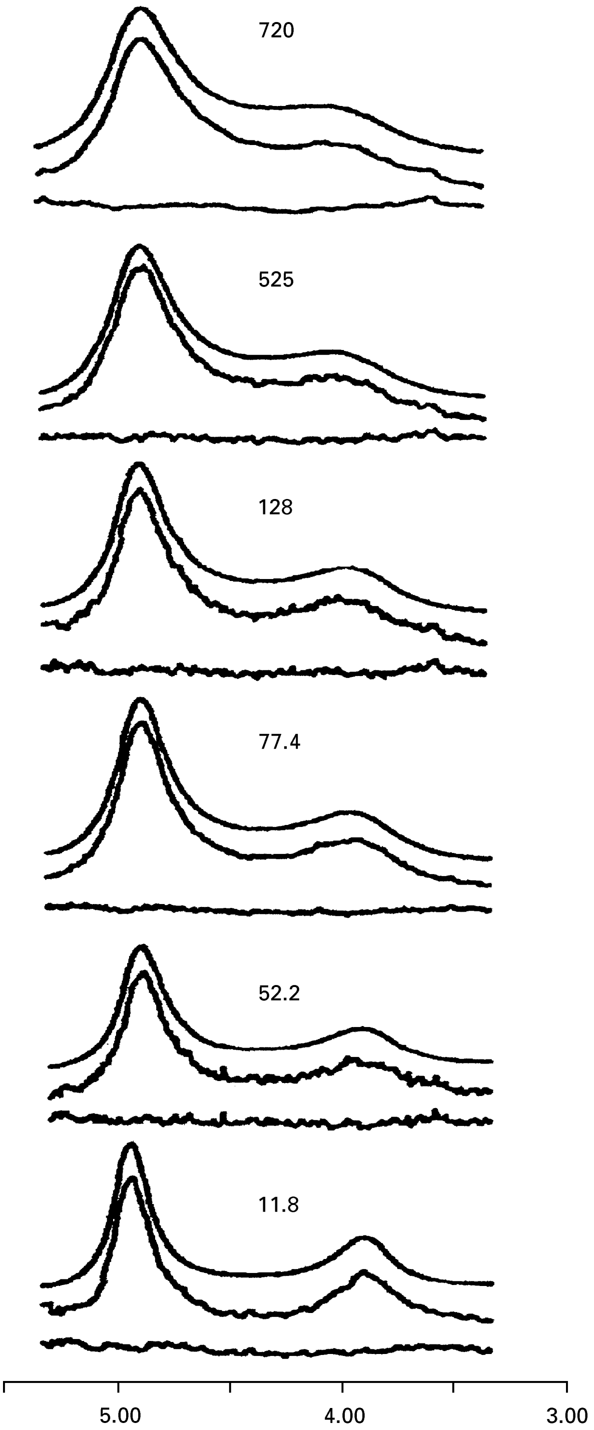



**Figure 4** Gas-phase $^1$H NMR spectra of *n*-propyl nitrite at 240.6 K. The spectral region shown corresponds to the methylene protons attached to the carbon bonded to the oxygen. Samples contained 2 torr of *n*-propyl nitrite and various pressures of $CO_2$. Spectral labels refer to the total sample pressure in torr. In each plot the top trace is the simulated spectrum, the middle trace is the experimental spectrum and the lower trace is the difference between the two. Reprinted with permission from Moreno PO, True NS and LeMaster CB (1990) Pressure-dependent gas-phase $^1$H NMR studies of conformational kinetics in a homologous series of alkyl nitrites. *Journal of Physical Chemistry* **94**: 8780–8787. © 1990 American Chemical Society.

## Absorption processes in zeolites

Several factors make xenon a very useful probe nucleus for studying absorption processes in

microporous materials such as zeolites. It is chemically inert, monatomic and of a convenient size, and its chemical shift is very sensitive to environmental factors and is amenable to theoretical modelling. Two of the stable isotopic species of xenon, $^{131}Xe$ and $^{129}Xe$ have magnetic moments. $^{129}Xe$, which has a spin of $\frac{1}{2}$, a natural abundance of 26.44%, a receptivity 31.8 times greater than $^{13}C$ and $T_1$ values on the order of seconds, is the most extensively used. The effects of environmental interactions on the chemical shift of xenon have been modelled using grand canonical Monte Carlo simulations, allowing the determination of the cluster size of xenon in cavities in a zeolite from chemical shift measurements. A recent study of xenon adsorbed in AgA zeolite, which is a system of silicoaluminate cages containing charged Ag clusters separated by 8-membered Si–O–Si rings, demonstrated that up to 10 different cluster sizes could be distinguished by their differences in chemical shifts. **Figure 5** shows static NMR spectra of $^{129}Xe$ in AgA zeolite. Spectral labels refer to the zeolite composition, dehydration temperature, and average number of xenon atoms per cage. Chemical shifts are referenced to xenon gas present in spaces between the crystallites. Exchange between cages is slow and the $^{129}Xe$ NMR spectrum consists of a series of lines corresponding to cages with different numbers of xenon atoms. The $^{129}Xe$ resonances from high field to low have chemical shifts that range over 200 ppm and are assigned to cages containing from 2 to 8 xenon atoms. Xenon intercage exchange dynamics can be studied using 2D-EXSY experiments. If chemical exchange occurs on the order of the mixing time, off-diagonal cross peaks appear in the 2D spectrum between the

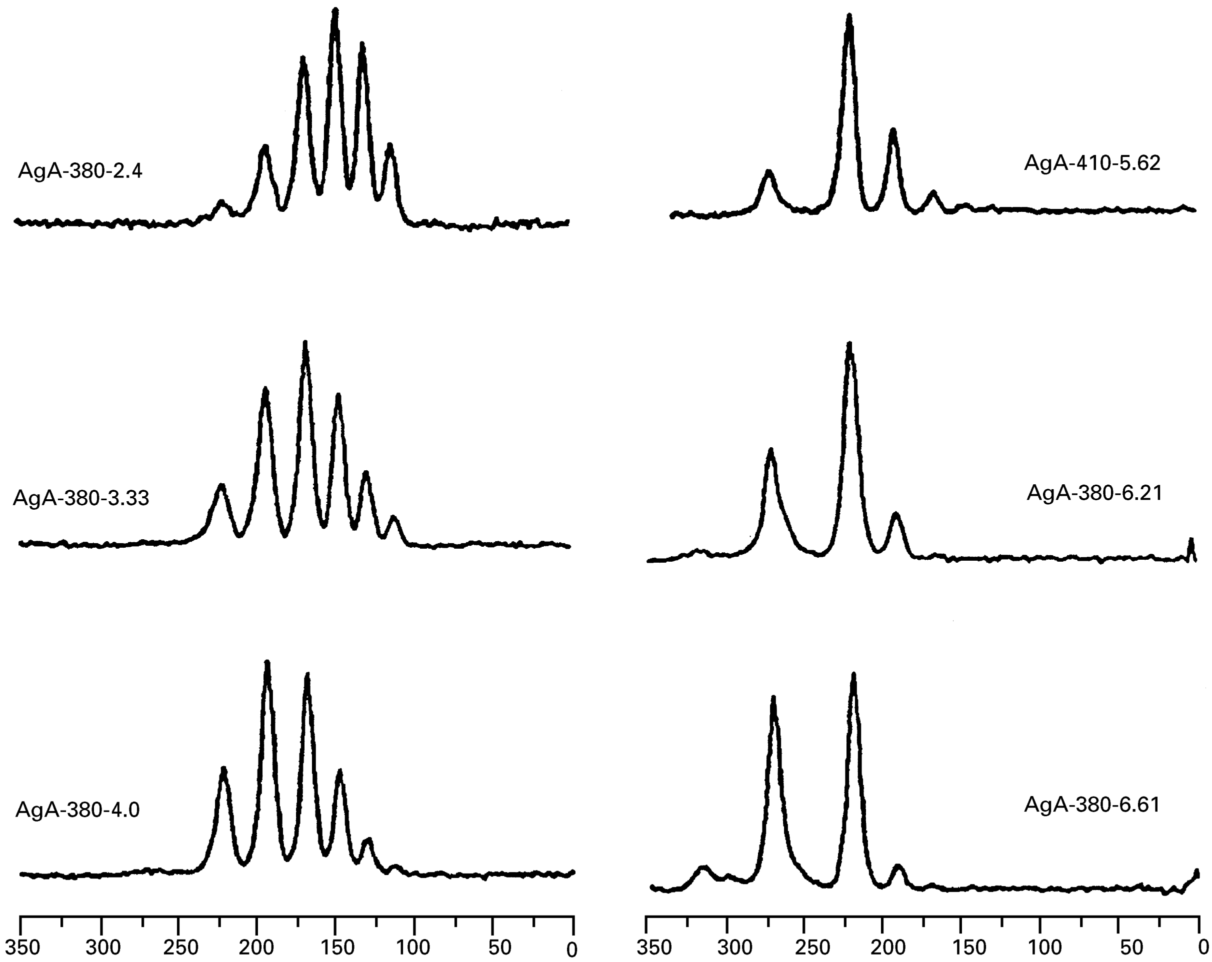



**Figure 5** NMR spectra of xenon in AgA zeolite. The axis in each spectrum is labelled in ppm with 0.00 ppm corresponding to the chemical shift of xenon gas in the space between the crystallites in the sample. Spectral labels refer to the zeolite composition, sample preparation temperature, and the average number of xenon atoms per α cage. The progression of lines to lower field corresponds to resonances from 1 xenon atom per α cage to 8 xenon atoms per α cage, respectively. Reprinted with permission from Moudrakovski IL, Ratcliffe CI and Ripmeester JA (1998) $^{129}Xe$ NMR study of adsorption and dynamics of xenon in AgA zeolite. *Journal of the American Chemical Society* **120**: 3123–3132. © 1998 American Chemical Society.

sites involved in the exchange. The cross peaks provide information about the exchange pathways and their intensities are a function of the exchange rate constant. Rate constants for exchange between cages indicate that sorption energies decrease at higher loading.

## List of symbols

$B_0$(s) = magnetic field experienced by sample; $k_a$, $k_d$, $k(E)$ = rate constants for activation, deactivation, reaction; $k_{uni}$ = macroscopic pseudo-unimolecular rate constant; $\alpha$ = shape factor; $\kappa_s$ = (sample) volume susceptibility; $\sigma^{NMR}$ = rotationally inelastic cross section.

*See also:* **High Resolution IR Spectroscopy (Gas Phase) Applications; High Resolution IR Spectroscopy (Gas Phase) Instrumentation; NMR Relaxation Rates; Parameters in NMR Spectroscopy, Theory of; Xenon NMR Spectroscopy.**

## Further reading

Armstrong RL (1987) Nuclear magnetic relaxation effects in polyatomic gases. *Magnetic Resonance Review* **12:** 91–135.

Dybowski C and Bansal N (1991) NMR spectroscopy of xenon in confined spaces: clathrates, intercalates, and zeolites. *Annual Reviews of Physical Chemistry* **42:** 433–464.

Jameson CJ (1991) Gas-phase NMR spectroscopy. *Chemical Reviews* **91:** 1375–1395.

True NS (1996) Gas phase studies of chemical exchange processes. In: Grant DM and Harris RK (eds) *Encyclopedia of NMR*, pp 2173–2178. New York: Wiley.

True NS and Suarez C (1995) Gas-phase NMR studies of conformational processes. In: Hargittai M and Hargittai I (eds) *Advances in Molecular Structure Research*, Vol 1, pp 115–155. Stanford, CT: JAI Press.

# GC–IR Applications

*See* **Chromatography-IR, Applications.**

# GC–MS Methods

*See* **Chromatography-MS, Methods; Hyphenated Techniques, Applications of in Mass Spectrometry.**

# Geology and Mineralogy, Applications of Atomic Spectroscopy

**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

| ATOMIC SPECTROMETRY |
|---|
| Applications |

## Introduction

With their high level of sensitivity and specificity, atomic absorption and atomic emission spectroscopy methods have been applied to the analysis of elemental metal content in samples of geological, mineral and metal ore relevance. This article highlights some recent studies in the scientific literature and should serve to direct the reader to appropriate articles for further investigation.

## Review

A recent review by Balaram gives an account of the various instrumental techniques used in geological analysis. These include atomic absorption spectroscopy and inductively coupled plasma atomic emission spectroscopy as well as X-ray fluorescence, isotope dilution mass spectrometry and neutron activation analysis in addition to recent developments in inductively coupled plasma mass spectrometry especially in rare earth element analysis.

## Atomic absorption spectroscopy

A background to this technique is given elsewhere in this Encyclopedia. Tungsten has been determined in ores as its thiocyanate complex after separation using an Amberlite resin. Arsenic has been measured by pre-reduction of As(V) to As(III) using L-cysteine. Much study has been devoted to the quantitation of gold in rocks and ores. This has been achieved using a chelate-forming resin for preconcentration or a two-stage solvent extraction process or by treatment with $MnO_2$ or $KMnO_4$ and HCl. Additionally, bismuth, indium and lead have been quantified using electrothermal atomic absorption spectroscopy.

## Atomic emission spectroscopy

The theory of, and the instrumental methods used in, atomic emission spectroscopy are given elsewhere in this Encyclopedia. The simultaneous determination of hafnium, scandium and yttrium in rare earth element geological materials has been described using separation and concentration stages.

*See also:* **Atomic Absorption, Methods and Instrumentation; Atomic Absorption, Theory; Atomic Spectroscopy, Historical Perspective; Fluorescence and Emission Spectroscopy, Theory; Inductively Coupled Plasma Mass Spectrometry, Methods; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Acar O, Turker AR and Kilic Z (1997) Determination of bismuth and lead in geological samples by electrothermal AAS. *Fresenius Journal of Analytical Chemistry* **357**: 656–660, **360**: 645–649.

Balaram V (1996) Recent trends in the instrumental analysis of rare earth elements in geological and industrial materials. *Trends in Analytical Chemistry* **25**: 475–486.

Celkova A, Kubova J and Stresko V (1996) Determination of arsenic in geological samples by Hg AAS. *Fresenius Journal of Analytical Chemistry* **355**: 150–153.

Lihareva N and Delaloye M (1997) Determination of Hf, Sc and Y in geological samples together with the rare earth elements. *Fresenius Journal of Analytical Chemistry* **357**: 314–316.

Singh AK (1996) Rapid procedure for the determination of gold at sub ppm levels in geological samples by atomic absorption spectrometry. *Talanta* **43**: 1843–1846.

Soylak M, Elci L and Dogan M (1995) Spectrophotometric determination of trace amounts of tungsten in geological samples after preconcentration on an Amberlite XAD-1180. *Talanta* **42**: 1513–1517.

Yokoyama T, Yokota T, Hayashi S and Izawa E (1996) Determination of trace gold in rock samples by a combination of 2-stage solvent extraction and graphite furnace atomic absorption spectrometry—the problem of iron interference and its solution. *Geochemical Journal* **30**: 175–181.

# Germanium NMR, Applications

*See* **Heteronuclear NMR Applications (Ge, Sn, Pb).**

# Glow Discharge Mass Spectrometry, Methods

**Annemie Bogaerts**, University of Antwerp, Belgium

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

## Introduction (synopsis)

This article describes the basic characteristics and methodology of glow discharge mass spectrometry (GDMS). First, the working principles of the glow discharge source will be explained, and its use for mass spectrometry will be clarified, followed by a short historical background of GDMS. Further, the various glow discharge source configurations and mass spectrometers used in GDMS will be outlined. GDMS can be operated in three different electrical operation modes, with either a direct current, radiofrequency powered or pulsed-glow discharge system. The glow discharge can also be 'boosted', by combining it either with a laser, a graphite furnace, a microwave discharge, magnetic fields or external gas jets, in order to improve the analytical results. The three electrical operation modes and the boosting methods will be briefly discussed. Moreover, the quantification methods in GDMS will be described, as well as possible solutions to overcome spectral interferences. Finally, although GDMS is particularly suitable for the analysis of conducting materials, considerable effort has been undertaken also to analyse nonconductors, and three commonly used methods to achieve this will be briefly discussed.

## Principle of the glow discharge and its use for mass spectrometry

A glow discharge is a kind of plasma, i.e. a partially ionized gas, consisting of positive ions and electrons, and a large number of neutral atoms. It is formed when a cell, consisting of an anode and a cathode, is filled with a gas at low pressure (e.g. 1 torr; 1 torr = 133.3 Pa). In glow discharges used for mass spectrometry, argon is most frequently used as the filling gas. A potential difference (of the order of 1 kV) is applied between the two electrodes, and creates gas breakdown (i.e. the splitting of the gas into positive ions and electrons). The ions are accelerated towards the cathode and cause the emission of electrons upon bombardment at the cathode. The electrons arrive in the plasma, and give rise to excitation and ionization collisions with the argon gas atoms. The excitation collisions (and the subsequent decay, with emission of light) are responsible for the characteristic name of the 'glow' discharge. The ionization collisions create new ion-electron pairs. The ions are again accelerated towards the cathode, giving rise to new electrons. The electrons can again produce ionization collisions, creating new electron-ion pairs. Hence, the latter processes make the glow discharge a self-sustaining plasma.

The use of the glow discharge as an ion source for mass spectrometry is based on the phenomenon of sputtering. The material to be analysed serves as the cathode of the glow discharge. The argon ions from the plasma (and also fast argon atoms) that bombard the cathode can also release atoms of the cathode material, which is called sputtering. The sputtered atoms arrive in the plasma where they can be ionized. Thus formed ions of the material to be analysed can be detected with a mass spectrometer, giving rise to GDMS. **Figure 1** illustrates the basic principles of the glow discharge and its coupling to mass spectrometry.

Typical discharge conditions used for GDMS are about 1 kV discharge voltage, an argon gas pressure in the order of 1 torr, and a d.c. discharge current in the mA range. The detection limits of this technique are generally in the ppb range.

## Short history of GDMS

The glow discharge has been known as an ion source for mass spectrometry for more than 60 years. Gas discharges were indeed already used in the 1920s and 1930s as ion sources in the first mass spectrographs of

Aston and Thompson. However, the early popularity was followed by a decline into relative obscurity during the next 30 years, due to the development of the simple electron impact ion source. There was indeed more interest at that time for the analysis of organic samples with relatively high vapour pressure; hence simple ionization in the gas phase by an electron beam was largely sufficient. When the interest also shifted to the analysis of inorganic materials with lower vapour pressure, there was again need for other sources, with sufficient energy for atomization and ionization. Since spark and arc discharges were already widely used as excitation sources for atomic emission spectrometry, it was obvious that these sources could also be applied for mass spectrometry. These sources exhibit a high sensitivity, large applicability and only a few spectral interferences, but they do not yield a stable ion population. This led to the reexamination of the glow discharge as a stable, low energy ion source. The first come-back of GDMS was due to Coburn and coworkers; but later, Harrison and co-workers in particular were pioneers in the development of modern GDMS.

## Glow discharge source configurations

As mentioned above, the glow discharge is formed in a cell consisting of two electrodes. These two electrodes can be mounted in five different geometries (see **Figure 2**).

The *coaxial cathode* (**Figure 2A**) is the most widely used source configuration in GDMS applications. Samples can be made in pin-form (with a few millimetres exposed to the discharge) or in disk-form (partly shielded so that only the top part is exposed to the discharge). The sample acts as cathode whereas the anode is formed by the cell body itself.

The *planar diode* (**Figure 2B**) is the simplest analytical source. It is used for analysing samples in disk-form. The cathode (sample) and anode are in parallel configuration and are placed inside a tube.

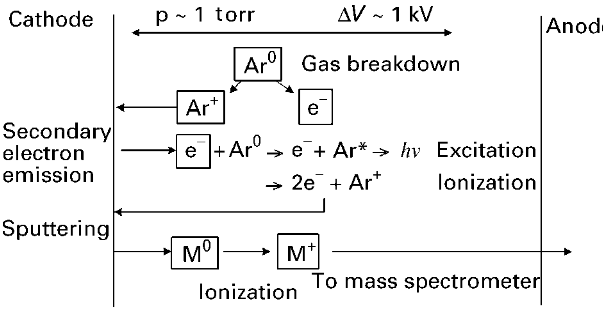



**Figure 1** Schematic of the basic processes in a glow discharge.

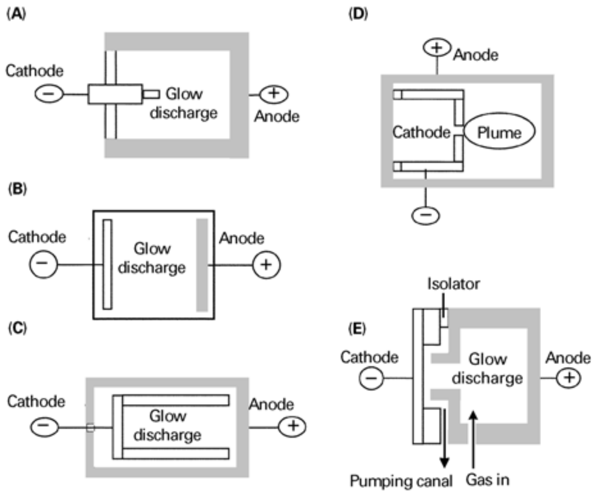



**Figure 2** Different glow discharge source configurations: (A) coaxial cathode; (B) planar diode; (C) hollow cathode lamp; (D) hollow cathode plume; (E) Grimm source. Reprinted from Bogaerts A and Gijbels R (1998) Fundamental aspects and applications of glow discharge spectrometric techniques (Review). *Spectrochimica Acta Part B* **53**: 1–42, with permission from Elsevier Science.

In the *hollow cathode lamp* (**Figure 2C**) the cathode forms a cavity rather than a pin or disk. It can be considered as three planar cathodes placed so close to each other that their negative glow regions coalesce into a single negative glow. This results in increased sputtering and ionization/excitation, yielding much better analytical sensitivity. A disadvantage of this source is the extensive machining required to make hollow cathodes from metal samples. Because most of the sputtering occurs at the cathode base, studies have been performed using a disk sample as the base of the cathode.

In the *hollow cathode plume* (**Figure 2D**) the sample is mounted in the base of the hollow cathode, in which an orifice is also made. A highly energetic flamelike plume, where excitation and ionization processes occur, is ejected through this hole. Due to the high atom population, this geometry is also characterized by a high sensitivity. Nevertheless, it is rarely used for practical analyses.

In the *Grimm configuration* (**Figure 2E**) the cell body (anode) approaches the cathode very closely (at a distance smaller than the thickness of the cathode dark space), so that the discharge is constricted to a well-defined part of the sample surface. It is therefore called an 'obstructed discharge'. A similar concept is also used in the standard cells for analysing flat samples in commercial mass spectrometers (e.g. the VG9000 instrument, see below). Moreover, the Grimm source possesses an additional pumping canal close to the cathode, which reduces the pressure near

the cathode, thereby minimizing redeposition. This geometry can, however, only be used for flat samples. Although this source is extensively used in glow discharge optical emission spectrometry (GD-OES; particularly for in-depth analysis), and forms the basic design for all commercial optical emission instruments, it is not so frequently used in GDMS.

## Mass spectrometers coupled to the glow discharge

To date, all common mass analysers have been explored for use in GDMS. The first commercial GDMS instruments used a double-focusing mass analyser, permitting the acquisition of high-resolution spectra with high sensitivity. Figure 3 shows a schematic of the 'VG9000' GDMS instrument, with first a magnetic and then an electrostatic sector (VG Elemental, Thermo Group). This instrument was made commercially available in 1985. About 60 instruments of this type are currently being used worldwide. Beside this VG9000 instrument, two other double-focusing GDMS instruments are used nowadays: the Kratos 'Concept', from which only two instruments have been manufactured, and the 'Element' from Finnigan MAT, which is essentially an inductively coupled plasma mass spectrometer (ICP-MS) that also became recently available with a glow discharge source. Figure 4 shows a typical GDMS mass spectrum obtained with the VG9000 glow discharge mass spectrometer.

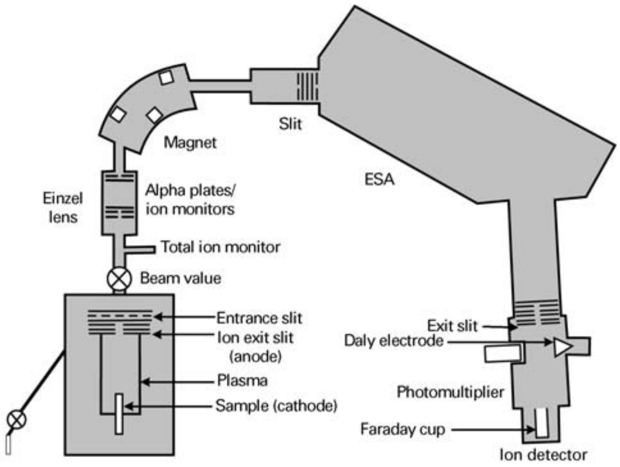



**Figure 3** Schematic of the VG9000 glow discharge mass spectrometer (VG Elemental, Thermo Group).

The wide expansion of modern GDMS began, however, with the quadrupole-based mass analysers, which are mainly being employed for fundamental and development research of GDMS. Nevertheless, this research resulted also in the commercial availability of quadrupole GDMS systems, e.g. the 'VG Gloquad' (VG Elemental, Thermo Group).

Moreover, glow discharges have also been coupled to ion trap mass spectrometers, double and triple

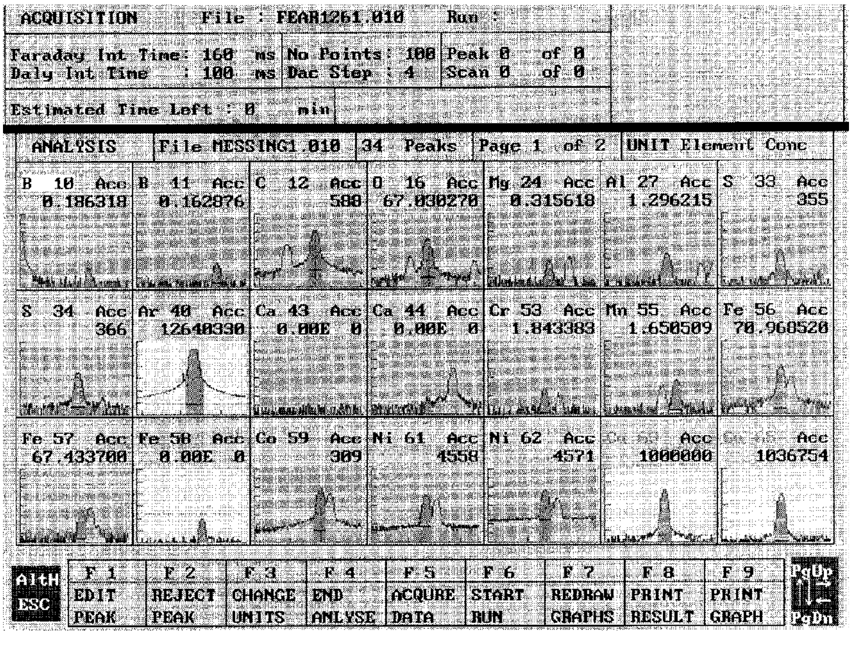



**Figure 4** Typical GDMS mass spectrum obtained with the VG9000 glow discharge mass spectrometer (VG Elemental, Thermo), for specific elements in a copper sample, using an argon glow discharge. The VG9000 software provides a separate window for each of the isotopes to be measured, instead of showing the complete mass spectrum. The values shown at the top of each window give the isotopic-corrected concentrations of the elements, in ppm.

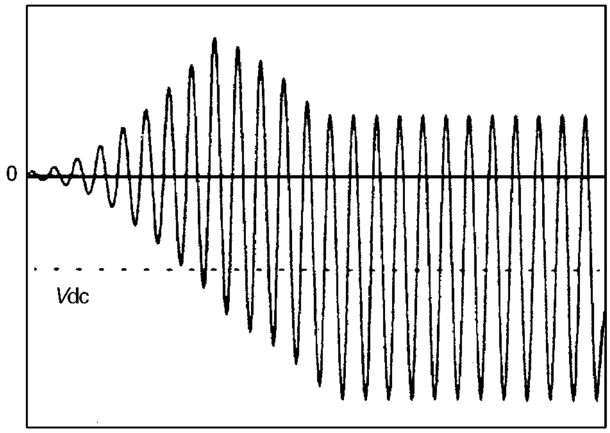



**Figure 5** Formation of a d.c. bias voltage in an RF-powered glow discharge with a nonconducting sample.

quadrupole instruments, time-of-flight mass spectrometers, and Fourier-transform mass spectrometers, but these investigations can rather be considered as research topics. Indeed, the commercial GDMS systems available at present employ only double-focusing and quadrupole-based mass spectrometers.

## Glow discharge electrical operation modes

The simplest and cheapest operation mode is the *direct current* (*d.c.*) mode. Voltages are typically 500- 1500 V, yielding electrical currents in the order of mA. This type of discharge mode is the oldest and the most widely used in glow discharge applications. However, it has the serious drawback of not being able to analyse nonconducting samples directly. Indeed, since in a glow discharge the sample to be analysed acts as the cathode, which is sputter bombarded by positive ions, it must be conducting. If not, the surface will be charged, preventing the positive ions from bombarding further. Due to this drawback of the d.c. mode, attention has been drawn during the 1990s to the *radiofrequency* (*RF*) operation mode.

The *radiofrequency* mode is indeed able to analyse nonconductors directly, since the positive charge accumulated during one half-cycle will be neutralized by negative charge accumulation during the next half-cycle, so that no charging occurs. Operation with RF-power of a glow discharge using a nonconducting sample yields a negative d.c. bias voltage on the sample surface (see Figure 5). Indeed, during the half-cycles in which the nonconducting electrode is positive, surface charging will occur much faster than in the half-cycles in which the electrode is negative, due to the much higher mobility of the electrons compared to the positive ions. The self-bias phenomenon permits the establishing of a time-averaged cathode and anode in the glow discharge, so that sputter-bombardment of positive ions on the cathode is still possible. Since the electrons try to follow the RF electric field, they oscillate between the two electrodes and spend more time in the plasma before they are lost, which results in a higher ionization efficiency. This leads to the second advantage of RF discharges, i.e. they can be operated at much lower pressures for the same current than d.c. discharges, which is interesting for reducing redeposition and spectral interferences. The capability of RF-powered GDMS for direct analysis of nonconductors was demonstrated already in the 1970s by Coburn and Kay and by Harrison and co-workers. However, it took until the late 1980s before RF-GDMS was revisited by Marcus and collaborators. Since then, extensive work has been done in this field. RF-discharges used for GDMS have been combined with quadrupole mass spectrometers, a Fourier-transform mass spectrometer, an ion trap mass spectrometer, a time-of-flight system and two sector-based mass spectrometers, but up to now there is no commercial RF-GDMS instrument available.

The third mode of operation of a glow discharge is the *pulsed mode*, which can be employed in combination with a conventional d.c. or with an RF glow discharge. Voltage and current are applied only during short periods of time (generally the millisecond range). Hence, compared to a normal d.c. discharge, higher peak voltages and peak currents can be obtained for the same average power. Therefore, more highly energetic gaseous ions can be produced, yielding more sputtering, a higher concentration of analyte atoms in the plasma and hence better analytical sensitivity. In addition to the better sensitivity, the pulsed mode has a second advantage for mass spectrometry, i.e. the analytically important ions and the interfering ions are formed during a different time in the pulse. By coupling this ‘time-resolved’ production of ions to a time-resolved detection (time-of-flight mass spectrometer), spectral interferences in the mass spectrum can be reduced. Moreover, the construction of a pulsed dual discharge system allows for simultaneous analyses with two electrodes, rendering possible the *in situ* calibration of an unknown sample against a reference standard. Recently, Harrison and co-workers introduced the microsecond-pulsed glow discharge. Due to the still higher peak currents and voltages that can be obtained during the short pulses, this source exhibits an even better analytical sensitivity. As shown in Figure 6, the microsecond-pulsed copper signals are one order of magnitude

higher than those in the d.c. glow discharge, and they can still be further enhanced, when the pulse power is increased.

## Boosting of the simple GDMS systems

Beside these three electrical operation modes, which make use of the glow discharge in its simple form as a spectroscopic source, the glow discharge can also be employed in a hybrid construction, in combination with lasers, graphite furnaces, microwave discharges or magnetic fields. The common purpose of these constructions is to increase the sputtering (atomization) and/or ionization (or excitation), and hence to improve the analytical sensitivity of the glow discharge.

### Laser-based methods

The development of cost-effective laser systems has generated a variety of laser techniques that can be coupled to a glow discharge. The atomization and the ionization/excitation steps occur independently of each other in the glow discharge, and a laser can be employed to enhance either of these two steps.

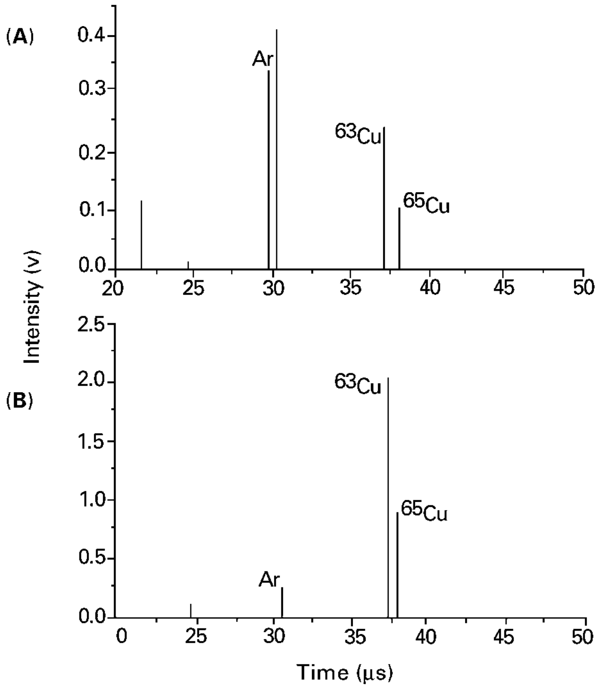



**Figure 6** Mass spectra of a d.c. (A) and a microsecond-pulsed (B) glow discharge, at 1 torr gas pressure. Conditions for d.c.: 800 V, 3 mA; conditions for pulsed: 2 kV, 15 μs. Reprinted from Hang W, Baker C, Smith BW, Winefordner JD and Harrison WW (1997) Microsecond pulsed glow discharge time-of-flight mass spectrometry: analytical advantages. *Journal of Analytical Atomic Spectrometry* **12**: 143–149, with permission of the Royal Society of Chemistry.

First, the laser can be utilized to ablate material from the sample cathode, enhancing the atomization step. The laser can also ablate material from a secondary sample (not the cathode) into the glow discharge. In the latter case, the sample must not be conductive, expanding the analytical applications of glow discharges to the analysis of nonconductors without the need to apply RF powers (see above). Beside the enhanced atomization, the possibility of performing spatially resolved measurements is an additional advantage of this method.

Second, the laser can also be used to enhance the excitation/ionization processes in the discharge. The usefulness of this laser enhancement has been demonstrated in a variety of application areas, such as optogalvanic effect spectroscopy (i.e. the laser results in alternations in the ionization rate of the discharge, which are electrically detected by the resulting changes in voltage or current), laser enhanced ionization (i.e. directly measuring the electrons released when the ionization in the discharge is enhanced due to laser photons) and resonance ionization mass spectrometry (i.e. the laser is used for selectively ionization enhancement of sputtered species in the discharge, increasing both the sensitivity and selectivity in GDMS).

### Furnace atomization nonthermal excitation spectrometry (FANES)

By analogy to the laser ablation glow discharge technique, FANES also makes use of an external atomization source (i.e. thermal vaporization from a graphite furnace) whereas the atoms created are excited and/or ionized in the glow discharge plasma. This source type is markedly different from a classical glow discharge source. Indeed, sample volatilization is accomplished thermally and the rate of volatilization can be three orders of magnitude higher than typical sputtering rates in a glow discharge. Moreover, this technique is mostly operated at much higher pressures than a conventional glow discharge, which results in a longer residence time for the analyte atoms in the plasma, and hence higher sensitivity.

### Microwave boosted glow discharge

Coupling of auxiliary microwave power to a glow discharge, in order to enhance the ionization and excitation, can be performed in two different ways. An antenna can be inserted in the plasma to match the microwave power, or an additional microwave induced plasma can be applied. The latter technique has, in particular, been applied to GD-OES, where considerable enhancement of analytical optical

emission lines has been attained. However, this technique is not commonly used for GDMS.

### Magnetron discharges

Another method to improve the performance of a conventional glow discharge, is based on magnetic enhancement. One simple device among magnetically enhanced glow discharges is the magnetron glow discharge. In this device, permanent magnets are used to form a magnetic field of a few hundred gauss in the plasma. Electrons in the plasma are forced to move in closed-loop trajectories parallel to the cathode surface. Hence, the electron path length is increased and the ionization of the discharge gas is significantly enhanced. Therefore, the magnetron discharge can operate at much lower pressures than conventional discharges. Lower pressure operation provides higher ion and electron kinetic energies, leading to higher atomization and excitation/ionization efficiencies, and hence a better analytical sensitivity.

### Gas-assisted sputtering glow discharges

Another way to increase the analytical sensitivity of a glow discharge is to use a gas-jet discharge. Due to the gas jet impinging on the sample surface, the sputter ablation is improved by both reduced redeposition and increased sample transport. This results in a higher sputtered atom population in the plasma and hence better analytical sensitivity. The gas-jet glow discharge was first developed for glow discharge atomic absorption spectrometry (GD-AAS), but a few applications in GDMS have also been reported.

## Quantification in GDMS

To obtain quantitative concentrations in GDMS, two main approaches can be used. The easiest approach is the ion beam ratio method. The assumption is made that the ratio of the ion current for any one isotope with respect to the total ion current (except the signal arising from the discharge gas ions) is representative of the concentration of that isotope in the sample. Since the matrix ion signal is generally large compared to the individual trace species, especially for a high purity metal or semiconductor, the matrix ion current is, to a good approximation, equal to the total ion current and the matrix atoms can be assumed to have a concentration of 100%. Since this method can not correct for the variation in analytical sensitivity among different elements (e.g. due to variations in sputtering and ionization of the elements), it provides only semiquantitative results, i.e. accuracies of a factor of 2–3.

Real quantitative results require the differences in elemental sensitivities to be characterized using standards similar to the material under study. This characterization generates relative sensitivity factors (RSFs) that can be employed to correct the measured ion beam ratios. Since RSFs vary only slightly between matrices of the same general composition, exact matrix matching is not required to yield quantitative results with accuracies of 15–20%. The RSF method of quantification is the most widely employed in GDMS. Generally, the RSFs of different elements in GDMS lie within one order of magnitude, which makes GDMS a technique with rather uniform sensitivity for most elements. Experimental RSFs have been reported in the literature for different kinds of matrices, and some empirical models based on fitting parameters have been developed to predict RSFs. Moreover, it was found that the theoretically calculated RSF values correlated better with the experiment when 1% $H_2$ was added to the argon discharge gas. This suggests that RSFs could be more accurately predicted theoretically in a gas mixture of Ar + 1% $H_2$, leading to satisfactory quantitative results without the need to analyse a standard. This might be very interesting, because solid reference materials with known concentrations at the (ultra-) trace level are not commonly commercially available. However, this method is not yet routinely used in practice.

## Spectral interferences in GDMS

A common problem in GDMS, as in most mass spectrometric methods, is the occurrence of spectral interferences in the mass spectrum, e.g. by impurity gas ions, various types of cluster ions or multiply charged ions. This problem can partly be overcome by using a high-resolution mass spectrometer, such as a double-focusing instrument or a Fourier-transform mass spectrometer, or by using high purity gases and special purification systems to suppress these interferences. However, some interference problems cannot be overcome in this way. For example, $^{103}$Rh (rhodium is monoisotopic) in a copper matrix ($^{63}$Cu has 69% abundance) is severely limited by the $^{40}$Ar $^{63}$Cu cluster. Even if one could separate both peaks with a high-resolution mass spectrometer (a resolution of $M/\Delta M = 7620$ is required), the tailing of the huge cluster peak would prevent one from reaching low limits of detection for rhodium. In this case, the problem can be solved using an alternative plasma gas, e.g. neon instead of argon. It is demonstrated that both discharge gases exhibit similar analytical performance if a correspondingly higher pressure is

used for neon; and hence, specific interference problems can be overcome. To illustrate this, Figure 7 presents a part of the mass spectrum (*m*/*z* 90- 100) of a pure iron sample, both in argon and in neon. It appears that most peak intensities in this mass range are significantly decreased in the neon discharge. This clearly demonstrates that these peaks are interfering peaks due to $ArFe^+$ clusters, which can be avoided by using neon as the discharge gas.

Moreover, another method has recently been proposed in the literature, which tries to suppress the cluster interferences by sampling from a reversed hollow cathode ion source. Indeed, it was found that the analyte ions are characterized by a peak at high energy, whereas argon ions and cluster ions possess a peak at low energy. By sampling only high energy ions, the argon ion and cluster ion interferences can be suppressed. Finally, the clusters can also be used for quantification. Indeed, since argides, dimers and doubly charged analyte ions may be less disturbed by spectral interferences, they can therefore sometimes be better employed for quantification than singly charged analyte ions, as has also been recently demonstrated.

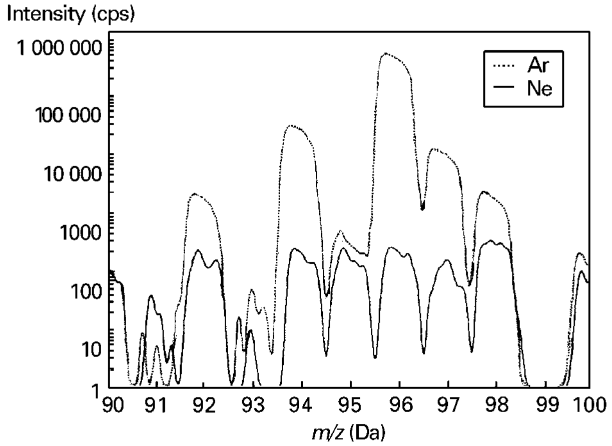



**Figure 7** Mass spectra of a pure Fe sample in argon and in neon (*m*/*z* 90–100 range). The difference between argon and neon is caused by $ArFe^+$ interferences. Reprinted from Jakubowski N and Stüwer D (1989) Comparison of Ar and Ne as working gases in analytical glow discharge mass spectrometry. *Fresenius' Journal of Analytical Chemistry* **355**: 680–686, with permission of Springer-Verlag, Berlin.

## The analysis of nonconductors by GDMS

Since the sample in the glow discharge acts as the cathode bombarded by positive ions, the concept seems to restrict the applications of GDMS to the analysis of electrically conducting materials, because nonconductive materials will be charged. The analysis of conducting materials (e.g. high-purity metals or alloys) forms, indeed, the most important field of application of GDMS. Nevertheless, much effort has also been applied to analyse nonconducting materials with GDMS, to greatly widen the application field. As mentioned above, RF-discharges are often used to analyse nonconductors directly. Nevertheless, there is no commercial RF-GDMS instrument available up to now, which is a serious drawback for routine analysis.

In a d.c. discharge, nonconductors can, however, also be analysed when applying certain modifications. Two methods are reported in the literature. The first consists of mixing the nonconducting sample as a powder with a conductive binder (Cu, Ag, Ga) and pressing it into an electrode. This method is generally well established, as follows from the large number of papers in the literature. However, in addition to the increase in sample preparation time compared to direct analysis of conducting solids, the mixing with the conductive matrix can introduce contamination. Other problems arise from the trapping of water vapour and atmospheric gases in the sample during the compaction process. Also, the time to reach steady state conditions with a composite cathode can be prohibitively long.

The second approach in d.c.-GDMS is the use of a metallic secondary cathode diaphragm in front of the flat nonconducting sample surface. Due to the redeposition of a part of the sputtered metal atoms from the secondary cathode, a very thin conductive layer is formed on the nonconductive material. The sampling depth is large enough (~5 Å, or 0.5 nm) to allow atomization of the nonconducting sample as well. The principle of this method is explained schematically in Figure 8. This method is rather new, but has already been successfully applied to the analysis of glass, ceramics, marble, polymers and even atmospheric particulate matter (aerosols).

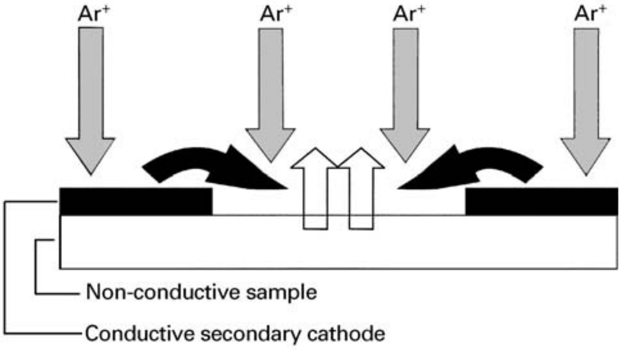



**Figure 8** Schematic representation of the sputter-redeposited film formation on the nonconducting sample surface.

*See also:* **Atomic Absorption, Theory; Inorganic Chemistry, Applications of Mass Spectrometry; Laser Applications in Electronic Spectroscopy; Laser Microprobe Mass Spectrometers; Laser Spectroscopy Theory; Mass Spectrometry, Historical Perspective; Microwave Spectrometers.**

## Further reading

Bogaerts A and Gijbels R (1997) Modeling of glow discharges: what can we learn from it? *Analytical Chemistry* 68: 719A-727A.

Bogaerts A and Gijbels R (1998) Fundamental aspects and applications of glow discharge spectrometric techniques. *Spectrochimica Acta Part B* 53: 1-42.

Harrison WW (1988) Glow discharge mass spectrometry. In: Adams F, Gijbels R and Van Grieken R (eds) *Inorganic Mass Spectrometry*, Chapter 3. New York: Wiley.

Harrison WW (1988) Glow discharge mass spectrometry: a current assessment. *Journal of Analytical Atomic Spectrometry* 3: 867-872.

Harrison WW (1992) Glow discharge: considerations of a versatile analytical source. Plenary lecture. *Journal of Analytical Atomic Spectrometry* 7: 75-79.

Harrison WW and Bentz BL (1988) Glow discharge mass spectrometry. *Progress in Analytical Spectroscopy* 11: 53-110.

Harrison WW, Hess KR, Marcus RK and King FL (1986) Glow discharge mass spectrometry. *Analytical Chemistry* 58: 341A-356A.

Harrison WW, Barshick CM, Klingler JA, Ratliff PH and Mei Y (1990) Glow discharge techniques in analytical chemistry. *Analytical Chemistry* 62: 943A-949A.

King FL and Harrison WW (1990) Glow discharge mass spectrometry: an introduction to the technique and its utility. *Mass Spectrometry Reviews* 9: 285-317.

King FL, Teng J and Steiner RE (1995) Glow discharge mass spectrometry: trace element determinations in solid samples. *Journal of Mass Spectrometry* 30: 1061-1075.

Marcus RK (1996) Radio-frequency powered glow discharges: opportunities and challenges. *Journal of Analytical Atomic Spectrometry* 11: 821-828.

Marcus RK, Harville TR, Mei Y, Shick CR (1994) Radio-frequency powered glow discharges: elemental analysis across the solids spectrum. *Analytical Chemistry* 66: 902A-911A.

# Gold NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# H

## Hafnium NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# Halogen NMR Spectroscopy (Excluding $^{19}$F)

**Frank G Riddell**, The University of St Andrews, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

The higher halogens, chlorine, bromine and iodine all possess NMR active nuclei, all of which are quadrupolar. Chlorine has two NMR active isotopes $^{35}$Cl (75%) and $^{37}$Cl (25%) of which $^{35}$Cl is the isotope of choice owing to its higher magnetogyric ratio and natural abundance. Bromine also presents two NMR active isotopes $^{79}$Br (50.5%) and $^{81}$Br (49.5%). Despite its marginally lower natural abundance, $^{81}$Br is the isotope of choice because of its slightly higher magnetogyric ratio. Iodine has only one NMR active nucleus $^{127}$I (100%). More NMR studies have been reported on chlorine than on the other two halogens for two reasons: (i) the chloride ion is the most abundant intracellular anion in normal mammalian cells, giving a biological incentive for NMR studies and (ii) both isotopes of chlorine have relatively low quadrupole moments, reducing unfavourable quadrupolar interactions and consequently enhancing visibility relative to the other halogens. This chapter reviews the use of NMR to study halides in biological systems, in solution chemistry and in solids. In particular it concentrates on the use of quadrupolar interactions from the isotopes available. There have been relatively few reviews of halogen NMR (excluding $^{19}$F) and as a consequence several leading papers from the primary literature are included in the Further reading section list at the end of this article.

## Nuclear properties

The nuclear properties of the five NMR active isotopes of the halogens are presented in **Table 1**.

### Quadrupolar relaxation and visibility

The NMR spectra of the halogens are dominated by the fact that all the isotopes are quadrupolar. Indeed, the majority of the reported studies of halogen NMR make use of the resulting quadrupolar interactions. Many of the problems that arise and solutions adopted are similar to those involved with the alkali metals. Quadrupolar nuclei have an asymmetric distribution of charge which gives rise to an electric quadrupole moment. Apart from when the nucleus is in an environment with cubic or higher symmetry, the quadrupole moment interacts with the electric field gradient (EFG) experienced by the nucleus, giving rise to, among other things, quadrupolar relaxation. The strength of the quadrupolar interaction between the quadrupole moment ($eQ$) and the electric field gradient ($eq$) is given by the quadrupolar coupling $e^2qQ/h$. This can take values from very small to hundreds of MHz depending on the magnitudes of $Q$ and $q$.

In solution, modulation of the EFG at the quadrupolar nucleus by isotropic and sufficiently rapid

**Table 1** Nuclear properties of the halogens other than $^{19}$F

| Isotope | Spin | *Natural abundance (%)* | *$10^{-7}$ Magnetogyric ratio $\gamma$ (rad $T^{-1}$ $s^{-1}$)* | *$10^{28}$ Quadrupole moment Q ($m^2$)* | *NMR frequency* Ξ (MHz) | *Relative receptivity $D^C$* |
|---|---|---|---|---|---|---|
| $^{35}$Cl | 3/2 | 75.53 | 2.624 | $-8.2 \times 10^{-2}$ | 9.809 | 20.2 |
| $^{37}$Cl | 3/2 | 24.47 | 2.184 | $-6.5 \times 10^{-2}$ | 8.165 | 3.78 |
| $^{79}$Br | 3/2 | 50.54 | 6.726 | 0.33 | 25.140 | $2.28 \times 10^2$ |
| $^{81}$Br | 3/2 | 49.46 | 7.250 | 0.27 | 27.100 | $2.79 \times 10^2$ |
| $^{127}$I | 5/2 | 100 | 5.390 | –0.79 | 20.146 | $5.41 \times 10^2$ |

Ξ is the observing frequency in a magnetic field in which $^1$H is at 100 MHz.
$D^C$ is the receptivity relative to $^{13}$C.
Quadrupole moments *Q* are the least well determined parameters in this table.
Data taken from Mason J (1987). In: Mason J (ed) *Multinuclear NMR*, p 623. New York: Plenum Press.

molecular motions (where $\omega_0\tau_c << 1$) leads to relaxation according to the expression:

$$\frac{1}{T_1}=\left(\frac{3}{40}\right)[(2I+3)/I^2(2I-1)](1+\eta^2/3)e^2qQ/\hbar)^2\tau_c$$

where $\eta$ is the asymmetry parameter associated with the EFG.

For all of the nuclei in this chapter the quadrupolar couplings for covalently bound halogens are typically many tens for MHz, depending upon the particular case (~70 MHz for chloroalkanes). This can be thought of as arising from the asymmetric distribution of electrons in the covalent bond giving rise to a large EFG at the nucleus. For molecules in solution that tumble with typical correlation times ($\tau_c$) of $10^{-12}$ s, this gives rise to $T_1$ and $T_2$ values that are typically in the submillisecond range. As a consequence, line widths range from the very broad to the invisibly broad with the latter case being the norm. With very broad line widths due to short $T_2$ values, much if not all of the halide resonance can be lost in the dead time between the pulse and the start of signal acquisition. Thus, quadrupolar relaxation has a devastating effect upon the signals of the covalently bonded halogen nuclei in solution, rendering them invisible in all but the most favourable circumstances. In passing, it is worth noting that the $^{35}$Cl line width in a typical small molecule like $CCl_4$ is about 2000 ppm at an observing frequency of 10 MHz. This has to be compared with a chemical shift range of around 1000 ppm from aqueous $Cl^-$, which is amongst the most shielded cases, to $ClO_4^-$, which is one of the most deshielded cases.

Quadrupolar couplings are smaller in cases where chlorine is covalently bound to the second and third row elements such as silicon, phosphorus, sulfur and arsenic (Si–Cl ~35 MHz), making observation easier but the lines are still very broad.

The anions themselves in solution present a different picture since they are subject to much lower quadrupolar interactions. In aqueous solution the anions are solvated by hydrogen bonding. At any one instant the pattern of hydrogen bonded water molecules around the anion does not have spherical symmetry but is always close to it. Thus the quadrupolar couplings are much lower than for the covalently bound halogens but are not zero. Typically, in aqueous solution and in the absence of extraneous influences, both isotopes of chlorine show $Cl^-$ line widths of 10–15 Hz, both isotopes of bromine show $Br^-$ line widths of ~ 400 Hz, $^{127}I^-$ has a line width of over 1000 Hz.

Covalent compounds with the halogen nucleus at a site of tetrahedral or octahedral symmetry, as in $ClO_4^-$ or $IF_6^+$, present a low EFG at the halogen nucleus and frequently have relatively sharp resonances. Indeed the spectra of the tetrahedral perhalate ions in acetonitrile solution show exceptionally sharp lines that enable secondary isotope effects on shielding to be observed. The shielding difference between $(Cl^{16}O_4)^-$ and $(Cl^{18}O^{16}O_3)^-$ is 0.090 ppm.

In cases where molecular motion is restricted ($\omega_0\tau_c$ is not <<1) the situation is more complex. The quadrupolar interaction with the nucleus shifts the energies of the Zeeman levels according to the square of the quantum number to a first approximation. Thus, the energy level splittings for a nucleus with $I=\frac{3}{2}$ (e.g. $^{35/37}$Cl and $^{79/81}$Br) become as illustrated in **Figure 1**. With rapid isotropic motion, as described above, the multiple line pattern will collapse into a single line. In the absence of rapid isotropic motion the relaxation rate of the outer transitions, which combine to have 60% of the total intensity, is different from that of the inner transition which has 40% of the total intensity. The frequencies of the outer transitions also shift from the inner transition (dynamic frequency shift). There are three principal consequences of these changes for cases

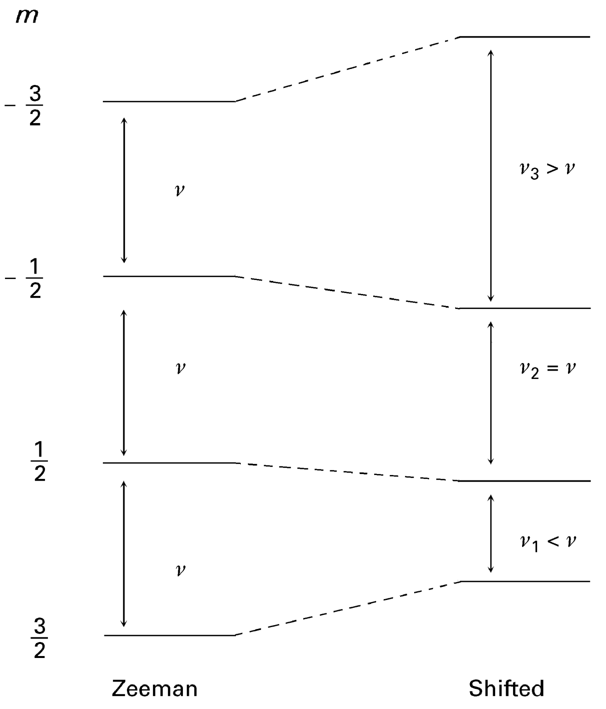



**Figure 1** Changes in the energy levels for a spin $\frac{3}{2}$ nucleus subject to a quadrupolar interaction.

where the motion of halogen is restricted. Firstly, line shapes become a double and not a single Lorentzian, secondly two relaxation times are apparent, and finally, where the more rapid relaxation time becomes very short, partial or total invisibility of the signal from the outer transitions may occur. For example, studies on dog erythrocytes showed a relatively sharp $^{35}$Cl line (line width ~40 Hz) but only 40% visibility and in renal tubules the $^{35}$Cl signal was reported to be invisible. An excellent review of quadrupolar relaxation effects has been given by Springer (listed in the Further reading).

In solids the situation is even more complex with, for example, the possibility of multiple quantum relaxation pathways. A discussion of this is outwith the scope of this article. In general, however, relaxation times for halogens in solids are several orders of magnitude longer than for those in solution and this, in turn, in favourable cases, does allow effects due to the halogen nuclei to be observed.

## Biological chloride

Chlorine is the eighth most abundant element in the human body. Although it is present in large amounts as chloride ion in extracellular spaces the intracellular concentration of chloride is high and it is the most abundant intracellular anion in normal mammalian cells. For example, its concentration in the human erythrocyte is ~150 mM. Chloride is involved in many important biochemical pathways. The most used membrane transport system in mammals is band III protein whose role is to exchange $HCO_3^-$ inside the cell for $Cl^-$ outside, thus facilitating the removal of $CO_2$. Abnormal chloride transport is implicated in several serious medical conditions such as cystic fibrosis and Duchenne muscular dystrophy. Many good reasons exist, therefore, to develop NMR methods to study biological chloride.

One of the main problems in using NMR to study biological chloride is that the chemical shift of aqueous $^{35}Cl^-$ is essentially independent of the ion's surroundings, making differentiation of intra- and extracellular chloride difficult. Further difficulties arise because of the binding of chloride to proteins, some of which may contain paramagnetic ions, which can greatly increase the line width. If the exchange rate between the free and the bound chloride is slow this leads to a loss of signal.

The first problem can be met by using a contrast reagent (either a shift or a relaxation agent) in one of the compartments, normally extracellular. Aqueous shift reagents that have been employed include Co(II) either as $Co^{2+}$ cations or as the triglycinate complex $[Co(Gly)_3]^-$. Manganese (II) has been shown to be an extremely efficient relaxation agent for $^{35}Cl^-$ in which rapid exchange of chloride for water in the inner ligation sphere brings the chloride ions transiently under the paramagnetic influence of the five unpaired electrons on the manganese.

It has been argued that the relaxation agent $Mn^{2+}$ is superior to Co(II)-based shift reagents. First, the concentrations of Co(II) reagents needed to visualize intracellular or intravesicular chloride are considerably higher than those of $Mn^{2+}$. Secondly, all contrast reagents induce paramagnetic line broadening of the extracellular $^{35}$Cl resonance that is proportional to the square of the magnetic field whilst the induced shift is only linearly dependent on the field. Therefore, on moving to the high fields required to attain satisfactory signal-to-noise, $Mn^{2+}$ has the advantage of requiring even lower amounts of reagent to achieve the same line widths, whilst Co(II)-based shift reagents require even higher concentrations to achieve the same relative separation of resonances.

Cobalt(II) as a shift reagent has been employed to visualize the $^{35}Cl^-$ resonance from chloride ions inside vesicles and erythrocyte ghosts. The latter observation allowed measurements of the kinetics of chloride/sulfate exchange mediated by the band III protein. The same group also used Co(II) to observe $Cl^-$ inside the cells of an alga, *Chlorella pyrenoidosa* and inside the cells of a suspension of cultured plant cells from *Nicotiana tobbaccum*. In both of these

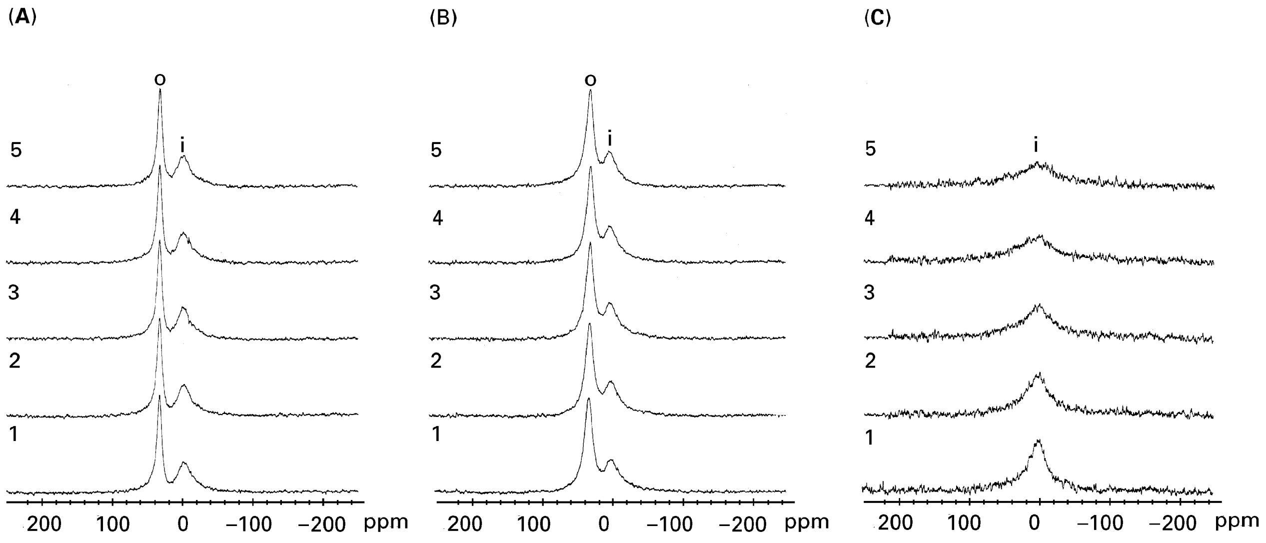



**Figure 2** The $^{35}$Cl NMR spectra at the mid-points of data accumulation (1) 20, (2) 60, (3) 100, (4) 140 and (5) 180 minutes of human erythrocyte suspensions at 70% haematocrit containing (A) 60 mM $[Co(Gly)_3]^-$, 40 mM $CoCl_2$ (B) and 20 mM $MnCl_2$ (C)$^-$. The symbols i and o denote the resonances originating from the intra- and extracellular halide, respectively. From Lin W and Mota de Freitas D (1996) *Magnetic Resonance in Chemistry* **34**: 768–772. Copyright John Wiley & Sons Limited. Reproduced with permission.

cases relatively sharp intracellular resonances that were suitably separated from the extracellular resonance were observed.

For many years it was believed that the $^{35}Cl^-$ resonance from chloride ions inside human erythrocytes was invisible owing to interaction of the ions with the paramagnetic protein haemoglobin. However, the shift reagent $[Co(Gly)_3]^-$ and the relaxation agent $Mn^{2+}$ have been used in separate studies that demonstrate the visibility of the intracellular chloride as a line that is ~700 Hz wide (11.75 T magnet). The intracellular $^{35}Cl^-$ resonance shows a double Lorentzian line shape and two $T_2$ values of ~0.09 and 1.2 ms, in keeping with a halide ion undergoing restricted motion. **Figure 2** shows typical spectra using both shift and relaxation agents. In the report using the shift reagent the exchange of intracellular $Cl^-$ with extracellular hypophosphite $(H_2PO_2)^-$ as a typical monovalent anion was followed by $^{35}$Cl NMR. The exchange followed first order kinetics with a rate constant of $2.4 \pm 0.8 \times 10^{-3}$ min$^{-1}$. This work shows the potential of $^{35}$Cl NMR to study chloride transport kinetics in cellular systems such as erythrocytes.

Manganese(II) has been used in conjunction with model vesicle systems to examine membrane transport induced by natural and synthetic anion carriers.

There are problems, however, with the use of $Co^{2+}$ and $Mn^{2+}$ as contrast reagents because it has been shown that biological membranes are very slowly permeable to these ions. It is, however, claimed that the shift reagent $[Co(Gly)_3]^-$ does not show membrane permeability and is, therefore, more suitable for studies on model systems.

The other way that $^{35}$Cl NMR can be used in biological systems is in the study of chloride bound to proteins or other macromolecules. Sometimes this can be done by direct visualization of the resonance or from its free induction decay. More commonly it is performed by relaxation time measurements provided that there is rapid exchange of bound and free anions, because bound chloride relaxes very much more rapidly than free chloride. Frequently, double quantum filtered (DQF) spectra are employed which give useful information on chlorine bound to sites on the macromolecule.

The most studied system has undoubtedly been the integral membrane band III protein which mediates the one-for-one exchange of hydrogencarbonate and chloride ions across the membrane, thus playing an integral part in respiration. NMR proved useful in verifying a ping-pong mechanism for the mode of action of this protein. Binding sites for chloride at both surfaces have been identified although the cytoplasmic face has the larger number. Chloride affinities for the binding sites determined by NMR, DQF spectra and $^{35}Cl^-$ and $^{37}Cl^-$ relaxation rates at a variety of field strengths, both in absence and presence of various inhibitors, have also given large amounts of information about the operation of this important protein.

In a typical example, DQF $^{35}$Cl NMR was used to study the binding of $Cl^-$ to external sites on the surface of red blood cells. A DQF $^{35}$Cl signal was

observed in cell suspensions containing 150 mM KCl, but the DQF signal was totally eliminated by adding 500 μM 4,4′-dinitrostilbene-2,2′-disulfonate (DNDS), an inhibitor that interferes with $Cl^-$ binding to the band III protein transport site. This result shows that only the binding of $Cl^-$ to transport sites of band III can give rise to a $^{35}Cl$ DQF signal from red blood cell suspensions. In accordance with this concept, analysis of the single quantum FID revealed that signals from buffer and DNDS-treated cells could be fitted with a single exponential function, whereas the FID signals of untreated control cells were biexponential. The band III-dependent DQF signal is thus caused, at least in part, by non-isotropic motions of $Cl^-$ in the transport site, resulting in incompletely averaged quadrupolar couplings.

Other biological systems that have been investigated by such techniques include sarcoplasmic reticulum membranes, superoxide dismutase, dromedary haemoglobin, the oxygen-evolving complex of spinach photosystem II, the zinc binding sites of human α-2-macroglobulin, alkaline phosphatase and halorhodopsin chromoprotein.

## Spin–spin coupling

The rapid quadrupolar relaxation of halogens, in general, effectively masks fine structure arising from scalar coupling between the halogen and other covalently bound magnetically active nuclei. Exceptions occur when the halogen experiences a low EFG as, for example, when it is bound in sites of high symmetry. Thus, for the octahedral $IF_6^+$ ion, the $^{19}F$ resonance appears as a sextet due to coupling to the $^{127}I$ nucleus ($I = 5/2$). Analogously, the $^{19}F$ spectrum of $BrF_6^+$ appears as two quartets due to coupling to the two almost equally populated Br isotopes (for both of which $I = 3/2$). In the perhalate ions $XO_4^-$, when labelled with $^{17}O$, it is possible to measure coupling between the halogen and the oxygen. **Figure 3** shows the $^{127}I$ spectrum of singly $^{17}O$-labelled $IO_4^-$. The six lines arising from coupling between the $^{127}I$ and the $^{17}O$ are visible. For non-symmetric sites where direct determination of coupling constants is difficult it is sometimes possible to make use of the scalar contribution to the relaxation of a dipolar nucleus ($I = 1/2$) coupled to the quadrupolar halogen. Typical values of one–bond $J$ couplings to Cl, Br and I are given in **Table 2**.

**Table 2** Illustrative values of one-bond coupling constants to Cl, Br and I

| *Compound* | *Spin–spin coupling constant (Hz)* |
|---|---|
| HCl | $J(^{35}Cl-^{1}H) = 41 \pm 2$ |
| $PCl_3$ | $J(^{35}Cl-^{31}P) = 120$ |
| $SnCl_4$ | $J(^{35}Cl-^{119}Sn) = 375$ |
| $PbCl_4$ | $J(^{35}Cl-^{207}Pb) = 705$ |
| $ClF_3$ | $J(^{35}Cl-^{19}F) = 260$ |
| $ClF_6^+$ | $J(^{35}Cl-^{19}F) = 337$ |
| HBr | $J(^{79}Br-^{1}H) = 57 \pm 3$ |
| $SnBr_4$ | $J(^{81}Br-^{119}Sn) = 920$ |
| $BrF_6^+$ | $J(^{79}Br-^{19}F) = 1575$ |
| $SnI_4$ | $J(^{127}I-^{119}Sn) = 1097$ |
| $IF_6^+$ | $J(^{127}I-^{19}F) = 2730 \pm 15$ |
| $ClO_4^-$ | $J(^{35}Cl-^{17}O) = 83 \pm 3$ |
| $BrO_4^-$ | $J(^{79}Br-^{17}O) = 408 \pm 6$ |
| $BrO_4^-$ | $J(^{81}Br-^{17}O) = 440 \pm 6$ |
| $IO_4^-$ | $J(^{127}I-^{17}O) = 489 \pm 6$ |

## Residual dipolar coupling in solids

In the solid state the relaxation times of quadrupolar nuclei are typically several orders of magnitude greater than in solution. One important consequence of these longer relaxation times is the ability, sometimes, to determine indirect spin–spin coupling constants, $J$, involving quadrupolar nuclei. These experiments are performed typically by high resolution cross-polarization magic-angle spinning experiments on spin $\frac{1}{2}$ nuclei that are coupled to the quadrupolar nucleus. The longer quadrupolar relaxation times in solids mean that direct dipolar interactions involving quadrupolar nuclei are not completely averaged by the MAS. This is because the magnetic moments of the quadrupolar nuclei are quantized, not only by the magnetic field, $B_0$ but also by interaction with the electric field gradient tensor. Consequently the magic angle (described by the familiar $3\cos^2\theta - 1 = 0$) is no longer magic and fails to reduce the dipolar interaction to 0. Only since 1993 have examples of residual dipolar coupling of chlorine nuclei to $^{13}C$ been reported but many examples are now known. The residual dipolar couplings of chlorine to directly bonded $^{13}C$ are of the order of several hundred Hz and are field dependent, the coupling decreasing with increasing magnetic field. The observed $^{13}C$ doublets contain overlapping components from coupling of both $^{35}Cl$ and $^{37}Cl$ to the $^{13}C$ nucleus. Analogous residual dipolar coupling between $^{79/81}Br$ and $^{13}C$ has been observed in 1,4-dibromobenzene and 4-bromo-3,5-dimethylpyrazole.

## Shielding

From a consideration of the section above on quadrupolar relaxation of covalent compounds it

will be realized that, in all but exceptional cases, the large line widths of the resonances give rise to a low precision in the determination of chemical shifts for Cl, whilst shielding studies for Br and I are virtually impossible. Only in molecules in which the halogen is present in a site of high symmetry can chemical shifts (relative to the free halide ion in aqueous solution as standard) be determined with precision. In the case of halogens present in non-ligated anions the chemical shift generally has a low dependence on the accompanying cation. Ligation can, however, have a substantial effect upon the chemical shift (see below). The largest deshieldings, relative to $Cl^-$, are to be found in species such as $ClO_3F$ (978.1 ppm)

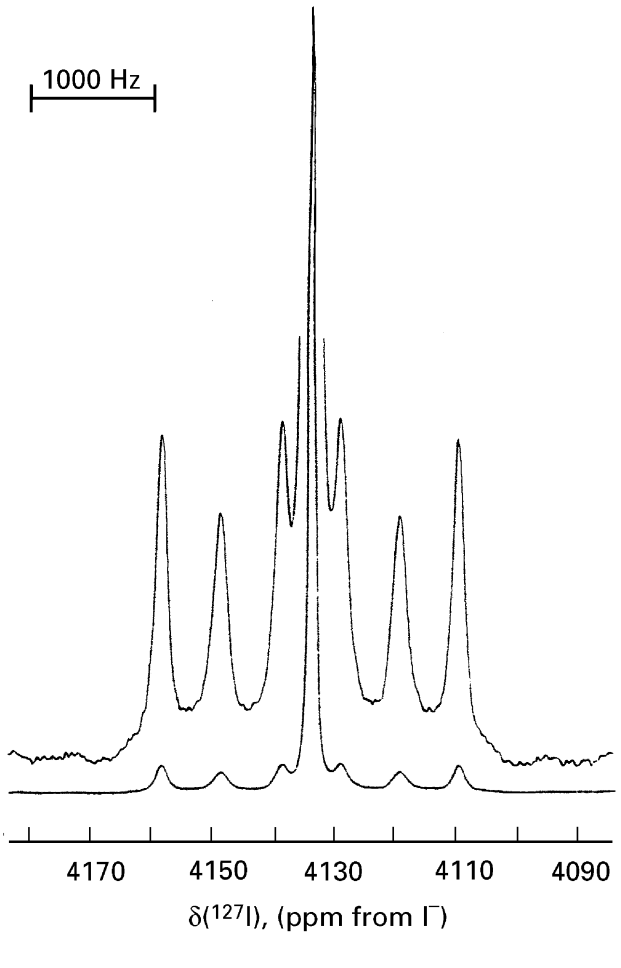



**Figure 3** $^{127}I$ spectrum (50.05 MHz) of 0.1 M $^{17,18}O$-enriched $Et_4NIO_4^-$ in $CH_3CN$ at 24°C. The central line arises from $I^{16}O_4^-$ (46%) and $I^{18}O^{16}O_3^-$ (17%); the sextet arises from $I^{17}O^{16}O_3^-$ (29%) and $I^{17}O^{18}O^{16}O_2^-$ (6%). The vertical expansion reveals the slight variation in line widths of the sextet. From Dove MFA, Sanders JCP and Appelman EH (1995) Comparative multinuclear ($^{35}Cl$, $^{79}Br$, $^{81}Br$, $^{127}I$ and $^{17}O$) magnetic resonance study of the perhalate anions $XO_4^-$ (X = Cl, Br or I). *Magnetic Resonance in Chemistry* **33**: 44–58. Copyright John Wiley & Sons Limited. Reproduced with permission.

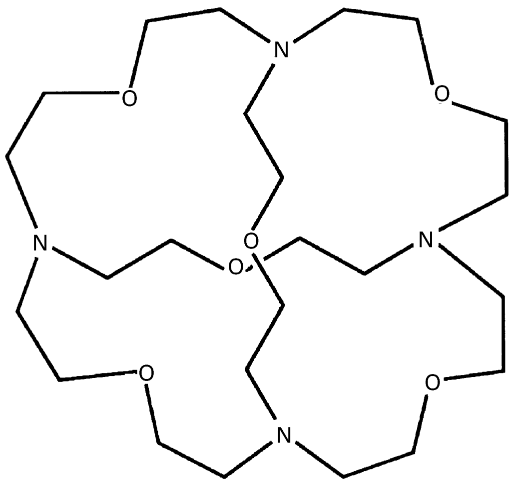
and $ClO_4^-$ (1002.5 ppm). The corresponding aqueous shifts for perbromate and periodate are 2478 and 4089 ppm, respectively. The normal chemical shift reference for halogen NMR is the frequency of the aqueous halide ion. Differences in the chemical environment of halogens are generally far more readily studied by using relaxation methods than by looking for chemical shift differences.

## Halide–ligand interactions

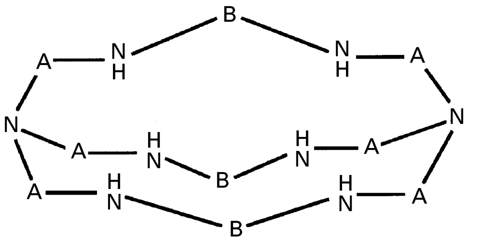
Halide ions may be chelated by ligands that contain suitable arrays of hydrogen-bonding N–H groups. The binding is stronger if the NH is present as part of a positively charged ammonium ion. The ligands (shown as [1] and [2]) form tetra- and hexaammonium ions, respectively, and bind a halide strongly inside the cavity with hydrogen bonds pointing inwards. The chelation of halide ions by ligands of this type is demonstrated by the generation of very substantial chemical shifts for the halide ions either as the ligand is added to an aqueous solution, or as the pH of the solution is lowered to generate the ammonium ion centres. Line widths of the encapsulated $^{35}Cl^-$ ions range from tens to thousands of Hz, depending upon the particular case. The larger line widths demonstrate that there are appreciable electric field gradients present at the chloride nucleus due to transient deformations of the coordination sphere away from ideal symmetry.

**Figure 4** illustrates the use of $^{35}Cl$ NMR to study the binding of chloride ions to the ligand [1] (as the $4H^+$ species) as the ratio of free to bound $Cl^-$ is increased. The chloride ion enters the centre of the ligand and is held in position by four inward facing hydrogen bonds from the $N^+$–H groups. The signal near 50 ppm is from the coordinated $^{35}Cl^-$ and is relatively sharp owing to the tetrahedral site symmetry. The signal at 0 ppm is from free $^{35}Cl^-$, which at relatively small amounts of free $Cl^-$ is 300 Hz wide due principally to dynamic line broadening that arises from exchange with the bound $Cl^-$. At an equal concentration of bound and free $Cl^-$ the contributions to the line width of the free $^{35}Cl^-$ of 80 Hz are 36 Hz from exchange and 44 Hz from quadrupolar interactions. At higher relative concentrations of free $Cl^-$ the exchange contribution to the line width diminishes still further but the free $^{35}Cl^-$ line remains relatively broad probably owing to attraction of the free $Cl^-$ ions to the exterior of the triply positively charged complex.

Halide ions show a weak association with cyclodextrins which can be visualized by the quadrupolar line broadening of the halide resonances. Competition between bromide ions and other ions for the




[1]




[2] A = $-CH_2-CH_2-$

B = $-(CH_2)_2-O-(CH_2)_2-$

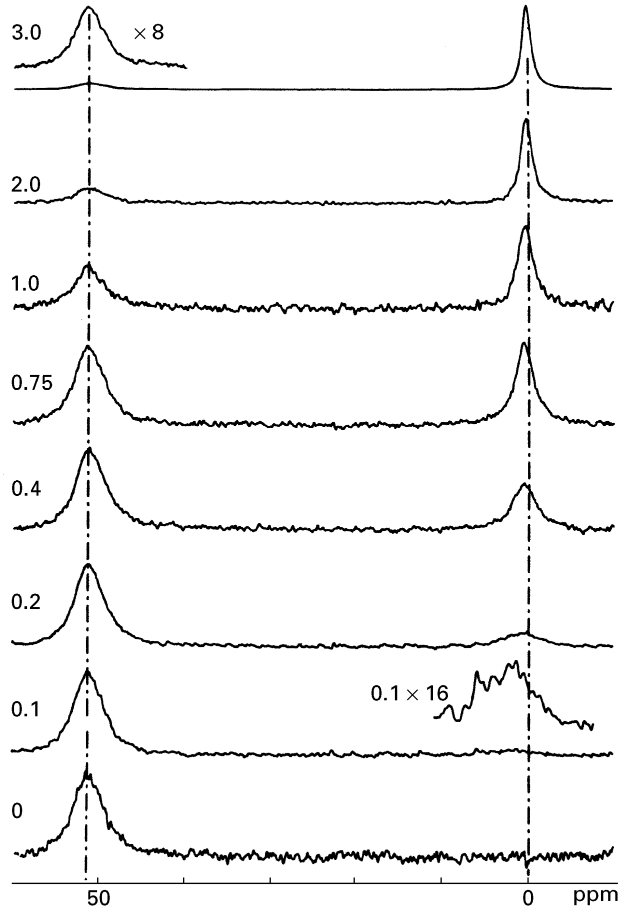



**Figure 4** $^{35}Cl$ NMR spectra (at 39.2 MHz) of aqueous solutions of the $Cl^-$ complex of ligand [**1**]-4H$^+$ (20 mM) and uncomplexed $Cl^-$ as a function of the ratio of free $Cl^-$ to $Cl^-$ complex. The shifts are given in ppm relative to external aqueous NaCl (0.1 M); pH 1.5 and 20°C. Reprinted with permission from Kintzinger JP, Lehn J-M, Kauffmann E, Dye JL and Popov AI (1983) Anion co-ordination chemistry – $^{35}Cl$ NMR studies of chloride anion cryptates. *Journal of the American Chemical Society* **105**: 7549–7553. Copyright 1983 American Chemical Society.

binding site in cyclodextrins can be studied by the reduction in the $^{81}Br^-$ line width as other electrolytes are added to an aqueous solution containing cyclodextrin and KBr.

## Halide–cation interactions

Interaction between ions of opposite charge is an important aspect of the chemistry of electrolyte solutions. The line widths of quadrupolar nuclei are an important probe of such interactions since the closer two ions of opposite charge approach the greater the EFG at their nuclei. This in turn can lead to a stronger quadrupole interaction, more rapid quadrupolar relaxation and larger line widths. Thus studies of the concentration dependence of quadrupolar relaxation are probes for ionic interactions. Such concentration dependence varies markedly with the cation present. Most monovalent ions show a low dependence of quadrupolar relaxation with concentration but this is not the case with alkyl substituted ammonium ions. In one typical such study $^{35}Cl$, $^{81}Br$ and $^{127}I$ NMR was used to study the interaction of halide ions with alkylammonium ions. The intrinsic line widths (line width at $\frac{1}{2}$ height corrected for temperature and viscosity) of the halide ions, plotted against molality ($m$) of the salt, show distinct plateau in the region 1–4 $m$ with the effect being greatest for iodide. These findings suggest the presence of contact ion pairs in these concentration ranges.

Similarly, the $T_2$ values of $^{35}Cl^-$ ions show a pH dependence in the presence of polybasic molecules such as polyamines, or amino acids such as histidine or arginine. The relaxation rates increase as the pH is lowered and show steepest changes at pH values equivalent to the p$K_a$ values for the various titratable groups. This is consistent with increasing ionic interactions between the anion as the number of positive charges on the organic substrate is increased.

## Halogens in conducting and semiconducting lattices

The ionic mobility of halogen-containing ions in solids, polymeric matrices and solutions may be studied by halogen NMR techniques. The self-diffusion of perchlorate ions in aqueous solutions of $NaClO_4$, $LiClO_4$ and $Mg(ClO_4)_2$ has been studied using pulsed field gradient techniques.

Both $^{81}$Br and $^{35}$Cl MAS NMR have been used in the investigation of sodium, silver and halo-sodalite semiconductor supralattices. Useful information is available on the environments of the encapsulated clusters within the sodalite lattice. The $Na_4X$ (X = Cl, Br) tetrahedra provide a symmetric environment around the halide and give rise to narrow resonances for specific locations. The spectra are sensitive to the distribution of anion empty cavities or to $I^-$-containing cavities.

$^{35}$Cl and $^{79\&81}$Br $T_1$ measurements between 77 and 700 K shed light on ionic transport in (solid) silver halides. A hopping process due to thermally activated diffusion of defects accounts for the data. The chloride ion conductor $CH_3NH_3GeCl_3$ has a variety of solid state phases and shows electrical conductivity above 398 K. The mobile ion in the high temperature phase was deduced to be $Cl^-$ because a disordered perovskite structure at the Cl site was found by Rietveld analysis of X-ray powder data and confirmed by $^{35}$Cl NMR. At temperatures above 364 K a $^{35}$Cl resonance becomes visible. This suggests that the rate of isotropic motion at the $Cl^-$ site is becoming comparable with or greater than the quadrupolar coupling, thus reducing the importance of the quadrupolar interactions.

## Two-dimensional spectroscopy

Two-dimensional multiple-quantum (2D2Q) $^{35}$Cl NMR spectra of systems exhibiting electric quadrupolar effects have been obtained. In such spectra the single quantum spectrum is displayed along one axis against the double quantum spectrum along the other. Such spectra give information about the quadrupolar interactions present in ordered systems.

## List of symbols

$I$ = nuclear spin quantum number; $Q$ = quadrupole moment; $\eta$ = asymmetry parameter; $\tau_c$ = correlation time for molecular rotation; $\omega_0$ = Larmor frequency.

*See also:* **Contrast Mechanisms in MRI; Membranes Studied By NMR Spectroscopy; Quadrupoles, Use of in Mass Spectrometry; Relaxometers; Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy.**

## Further reading

Dove MF, Sanders JCP, and Appelman E (1995) Comparative multinuclear ($^{35}$Cl, $^{79}$Br, $^{81}$Br, $^{127}$I and $^{17}$O) magnetic resonance study of the perhalate anions $XO_4^-$ (X = Cl, Br or I). *Magnetic Resonance in Chemistry* **33**: 44–58.

Drakenberg T, and Forsén S (1983). In: Lambert JB and Riddell FG (eds) *The Multinuclear Approach to NMR Spectroscopy*, NATO ASI Series C, p 405. Boston: Reidel.

Kintzinger JP, Lehn J-M, Kauffmann E, Dye JL and Popov, AI (1983) Anion co-ordination chemistry – $^{35}$Cl NMR studies of chloride anion cryptates, *Journal of the American Chemical Society* **105**: 7549–7553.

Lindman B and Forsén S (1976) *Chlorine Bromine and Iodine NMR. Physico-Chemical and Biological Applications.* Heidelberg: Springer Verlag.

Lindman B and Forsén S (1978) The halogens – chlorine, bromine and iodine. In: Harris RK and Mann BE (eds) *NMR and the Periodic Table.* London: Academic Press.

Liu DS, Kennedy SD and Knauf PA (1996) Source of transport site asymmetry in the band-3 anion exchange protein determined by NMR measurements of external $Cl^-$ affinity. *Biochemistry* **35**: 15228–15235.

Liu DS, Knauf PA and Kennedy SD (1996) Detection of $Cl^-$ binding to band-3 by double quantum filtered $^{35}$Cl nuclear magnetic resonance. *Biophysics Journal* **70**: 715–722.

Price WS, Kuchel PW and Cornell BA (1991) A $^{35}$Cl and $^{37}$Cl NMR study of chloride binding to the erythrocyte anion transport protein. *Biophysical Chemistry* **40**: 329–337.

Riddell FG and Zhou Z (1995) NMR visibility of $^{35}Cl^-$ in human erythrocytes. *Magnetic Resonance in Chemistry*, **33**: 66–69.

Roberts JE and Schnitker J (1993) Ionic quadrupolar relaxation in aqueous solution – dynamics of the hydration sphere. *Journal of Physical Chemistry* **97**: 5410–5417.

Shachar-Hill Y and Shulman RG (1992) $Co^{2+}$ as a shift reagent for $^{35}$Cl NMR of chloride with vesicles and cells. *Biochemistry* **31**: 6272–6278.

Springer CS (1996) Biological systems, spin-3/2 nuclei. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, p 940. Chichester: Wiley.

Xu Y, Barbara TM, Rooney WD and Springer CS (1989) Two-dimensional multiple-quantum NMR spectroscopy of isolated half-integer spin systems. 2. $^{35}$Cl examples. *Journal of Magnetic Resonance* **83**: 279–298.

# Heteronuclear NMR Applications (As, Sb, Bi)

**Claudio Pettinari**, **Fabio Marchetti** and
**Giovanni Rafaiani**, Università di Camerino, Italy

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

All the elements of group 15 of the periodic table have at least one isotope suitable for study by NMR spectroscopy. The magnetically active isotopes of arsenic, antimony and bismuth ($^{75}$As, $I = \frac{3}{2}$; $^{121}$Sb, $I = \frac{5}{2}$; $^{123}$Sb, $I = \frac{7}{2}$; $^{209}$Bi, $I = \frac{9}{2}$) have high abundance and relatively high sensitivity; nevertheless they have been little employed in multinuclear NMR spectroscopy owing to their high spin quantum numbers, which produce large quadrupolar moments. In fact, an efficient quadrupolar mechanism dominates their relaxation behaviour and gives rise to extremely broad resonances; this has limited the application of $^{75}$As, $^{121,123}$Sb and $^{209}$Bi NMR spectroscopy to species having a high degree of symmetry. Only a few NMR studies have been reported in the last three decades, and only a small amount of data, often useful in the determination of the molecular structure, is currently available.

## Properties of the nuclei

$^{75}$As is 100% abundant with a receptivity of 2.5 × $10^{-2}$ relative to the proton; $^{121}$Sb is 57% abundant with a receptivity of $9.2 \times 10^{-2}$; and $^{209}$Bi is 100% abundant with a receptivity of $0.14 \times 10^{-2}$; by contrast, $^{123}$Sb is about 43% abundant with a receptivity of $1.9 \times 10^{-2}$. In the case of antimony, the nuclide generally chosen for NMR determinations is $^{121}$Sb, presumably owing to the lower abundance of $^{123}$Sb.

The group 15 nuclides can be studied by NMR spectroscopy only when the E-containing compounds (E = As, Sb, Bi) are highly symmetric, with E in the oxidation state +5. These nuclides are spectroscopically silent in nonsymmetric compounds or in the +3 oxidation state. Narrow line widths arise from quadrupolar nuclei residing at the centre of a highly symmetric ligand environment. In addition, in oxidation state +5, no lone pairs remain in the valence shell and tetrahedral or octahedral symmetry can be achieved through bonding of four or six donor ligands, as in the case of $(SbCl_4)^+$ and $(SbCl_6)^-$. By contrast, when the group 15 elements are in oxidation state +3, there is a lone pair in the valence shell and the structures are generally characterized by substantial electric field gradients at the central atom, which prevents observation of NMR resonance.

The main nuclear properties of group 15 nuclides are shown in **Table 1**.

## Experimental aspects

$^{75}$As, $^{121}$Sb, $^{123}$Sb and $^{209}$Bi NMR spectra can be obtained efficiently using a broad-band probe. The observation frequencies are 85.637 MHz ($^{75}$As), 119.696 MHz ($^{121}$Sb), 64.819 MHz ($^{123}$Sb) and 80.381 MHz ($^{209}$Bi) when $^1$H NMR is at 500 MHz. The ideal spectral width setting is 20 kHz, to yield a data point resolution of 4.9 Hz per data point and an acquisition time of 0.205 s. The number of transients that must be accumulated varies with the nuclide under observation and with the concentration of the sample; however, typical values are 14 000 ($^{75}$As), 7000 ($^{121}$Sb), 36 000 ($^{123}$Sb) and 200 000–400 000 ($^{209}$Bi). A suitable line-broadening

**Table 1** Selected nuclear properties of magnetically active isotopes of As, Sb and Bi

| *Isotope* | *$^{75}$As* | *$^{121}$Sb* | *$^{123}$Sb* | *$^{209}$Bi* |
|---|---|---|---|---|
| Nuclear spin quantum number, *I* | $\frac{3}{2}$ | $\frac{5}{2}$ | $\frac{7}{2}$ | $\frac{9}{2}$ |
| Natural abundance (%) | 100 | 57.25 | 42.75 | 100 |
| Magnetogyric ratio ($10^{-7}$ rad $T^{-1}$ $s^{-1}$) | 4.5804 | 6.4016 | 3.4668 | 4.2986 |
| Quadrupole moment | 0.3 | −0.53 | −0.68 | −0.4 |
| Resonance frequency / MHz referred to $^1$H TMS resonance at 100 MHz | 17.122 | 23.930 | 12.958 | 16.069 |
| Receptivity relative to $^{13}$C | 143 | 520 | 111 | 777 |
| Receptivity relative to $^1$H | $2.5 \times 10^{-2}$ | $9.2 \times 10^{-2}$ | $1.9 \times 10^{-2}$ | 0.14 |

parameter can be applied by the exponential multiplication of the free induction decay prior to Fourier transformation. A line broadening factor equal to the original line width will produce the optimum signal-to-noise ratio.

The chemical shift standards generally used (at room temperature) are 0.1 M $[(Et_4N)^+(AsF_6)^-]$ in $CH_3CN$ for $^{75}As$, 0.3 M $[(Et_4N)^+(SbCl_6)^-]$ in $CH_3CN$ for $^{121,123}Sb$, and a saturated solution of $[(Me_4N)^+(BiF_6)^-]$ in $CH_3CN$ for $^{209}Bi$. In some cases, 0.1 M $[(Et_4N)^+(SbF_6)^-]$ in $CH_3CN$ as standard (0.0 ppm) for antimony has also been used. All the antimony data reported here are referred to $[(Et_4N)^+(SbCl_6)^-]$ at 0 ppm.

## Chemical shifts

$^{75}As$, $^{121,123}Sb$ and $^{209}Bi$ chemical shift ranges are very large (**Tables 2** and **3**). A positive sign indicates a chemical shift to higher frequency (lower shielding) with respect to the reference compound.

In the case of arsenic(+5) derivatives, the chemical shift range spans approximately 700 ppm, with the tetrahedral oxoanion $(AsO_4)^{3-}$ and the cation $(AsH_4)^+$ occupying the deshielded and the shielded ends of the range, respectively.

Although there is limited chemical shift data in the case of $^{75}As$ and $^{121}Sb$ it is possible to observe some general patterns analogous to those found for

**Table 2** $^{75}As$ NMR data for selected arsenic-containing compounds

| *Compound* | *Symmetry* | *Solvent* | $\delta$ (ppm) | *Multiplicity* | $^nJ(^{75}As–M)$(Hz) |
|---|---|---|---|---|---|
| $[AsF_6]^-$ | $O_h$ | $CH_3CN$ | 0.0 | st | 932 $^1J(^{75}As–^{19}F)$ |
| $[AsClF_5]^-$ | $O_h$ | $CH_3CN$ | 12.2 | q of d | ax.897 $^1J(^{75}As–^{19}F)$ eq.1009 $^1J(^{75}As–^{19}F)$ |
| [*cis*-$AsF_4Cl_2]^-$ | $O_h$ | $CH_3CN$ | –0.3 | q | 1013 $^1J(^{75}As–^{19}F)$ |
| [*trans*-$AsF_4Cl_2]^-$ | $O_h$ | $CH_3CN$ | –42.4 | qu | 1259 $^1J(^{75}As–^{19}F)$ |
| [*fac*-$AsF_3Cl_3]^-$ | | | | | |
| [*mer*-$AsF_3Cl_3]^-$ | | | | | |
| [*cis*-$AsF_2Cl_4]^-$ | $O_h$ | $CH_3CN$ | –102.2 | t | 985 $^1J(^{75}As–^{19}F)$ |
| [*trans*-$AsF_2Cl_4]^-$ | | | | | 925 $^1J(^{75}As–^{19}F)$ |
| $[AsFCl_5]^-$ | $O_h$ | $CH_3CN$ | –212.4 | d | 1005 $^1J(^{75}As–^{19}F)$ |
| $[AsCl_6]^-$ | $O_h$ | $CH_3CN$ | –391.8 | s | 555 $^1J(^{75}As–^1H)$ |
| $[AsF_5Br]^-$ | $O_h$ | $CH_3CN$ | 102.5 | st | |
| $[AsO_4]^{3-}$ | $T_d$ | $CH_3CN$ | 369 | s | |
| $[AsH_4]^+$ | $T_d$ | HF | –291 | q | |
| $[AsMe_4]^+$ | $T_d$ | $H_2O$ | 206 | s | |
| $[AsEt_4]^+$ | $T_d$ | $H_2O$ | 249 | s | |
| $[AsPr^n_4]^+$ | $T_d$ | $H_2O$ | 230 | s | |
| $[AsBu^n_4]^+$ | $T_d$ | $H_2O$ | 234 | s | |
| $[AsEt_3Me]^+$ | $T_d$ | $H_2O$ | 242 | s | |
| $[AsEt_3Pr^n]^+$ | $T_d$ | $H_2O$ | 243 | s | |
| $[AsEt_3Bu^n]^+$ | $T_d$ | $H_2O$ | 245 | s | |
| $[AsPr^n_3Me]^+$ | $T_d$ | $H_2O$ | 225 | s | |
| $[AsPr^n_3Et]^+$ | $T_d$ | $H_2O$ | 234 | s | |
| $[AsPr^n_3Bu^n]^+$ | $T_d$ | $H_2O$ | 231 | s | |
| $[AsPr^i_3Et]^+$ | $T_d$ | $H_2O$ | 258 | | |
| $[AsPh_4]^+$ | $T_d$ | $H_2O$ | 217 | s | |
| $Cs[As(OTeF_5)_6]$ | $O_h$ | $CH_3CN$ | –28.9 | st | 420 $^2J(^{75}As–^{125}Te)$ |
| $NMe_4[As(OTeF_5)_6]$ | $O_h$ | $CH_3CN$ | –29.1 | st | 430 $^2J(^{75}As–^{125}Te)$ |
| $Te(OTeF_5)_3[As(OTeF_5)_6]$ | $O_h$ | $SO_2$ | –28.2 | st | 432 $^2J(^{75}As–^{125}Te)$ |
| $H_3AsW_{12}O_{40}$ | $T_d$ | $H_2O$/dioxane | 298 | s | |
| $H_3AsMo_{12}O_{40}$ | $T_d$ | $H_2O$/dioxane | 337 | s | |

$Pr^n$ = *n*-propyl; $Bu^n$ = *n*-butyl; $Pr^i$ = isopropyl; st = septet, q = quartet; qu = quintet; t = triplet; d = doublet; s = singlet.

**Table 3** $^{121}$Sb NMR data for selected antimony-containing compounds

| *Compound* | *Symmetry* | *Solvent* | *δ* (ppm) | *Multiplicity* | $^{n}J(^{121}Sb–M)$ (Hz) |
|---|---|---|---|---|---|
| $[SbF_6]^-$ | $O_h$ | $CH_3CN$ | 86.6 | st | 1938 $^1J(^{121}Sb–^{19}F)$ |
| $[SbF_5Cl]^-$ | $O_h$ | $CH_3CN$ | 149.3 | q of d | ax. 1859 $^1J(^{121}Sb–^{19}F)$<br>eq. 2088 $^1J(^{121}Sb–^{19}F)$ |
| [*cis*-$SbF_4Cl_2]^-$ | $O_h$ | $CH_3CN$ | 183.9 | q | 2083 $^1J(^{121}Sb–^{19}F)$ |
| [*trans*-$SbF_4Cl_2]^-$ | $O_h$ | $CH_3CN$ | 187.5 | qu | 1980 $^1J(^{121}Sb–^{19}F)$ |
| [*fac*-$SbF_3Cl_3]^-$ | | | | | |
| [*mer*-$SbF_3Cl_3]^-$ | | | | | |
| [*cis*-$SbF_2Cl_4]^-$ | $O_h$ | $CH_3CN$ | 174.4 | t | 1981 $^1J(^{121}Sb–^{19}F)$ |
| [*trans*-$SbF_2Cl_4]^-$ | | | | | |
| $[SbCl_5F]^-$ | $O_h$ | $CH_3CN$ | 114.9 | d | 2079 $^1J(^{121}Sb–^{19}F)$ |
| $[SbCl_6]^-$ | $O_h$ | $CH_3CN$ | 0.0 | s | |
| $[SbCl_4]^-$ | $T_d$ | $SO_2ClF$ | 847.0 | s | |
| $[Sb(OTeF_5)_6]^-$ | $O_h$ | $SO_2ClF$ | 72.9 | s | |
| $[SbBr_4]^+$ | $T_d$ | $SO_2ClF$ | 95.3 | s | |
| $[Sb(OTeF_5)_6]^+$ | $O_h$ | $SO_2ClF$ | 73.1 | s | |
| $NMe_4[Sb(OTeF_5)_6]$ | $O_h$ | $CH_3CN$ | 72.7 | s | |
| $[SbCl_5Br]^-$ | $O_h$ | $CH_3CN$ | −380 | s | |
| $[SbCl_4Br_2]^-$ | $O_h$ | $CH_3CN$ | −780 | s | |
| $[SbCl_3Br_3]^-$ | $O_h$ | $CH_3CN$ | −1180 | s | |
| $[SbCl_2Br_4]^-$ | $O_h$ | $CH_3CN$ | −1590 | s | |
| $[SbClBr_5]^-$ | $O_h$ | $CH_3CN$ | −2005 | s | |
| $[SbBr_6]^-$ | $O_h$ | $CH_3CN$ | −2430 | s | |
| $[SbS_4]^{3-}$ | $T_d$ | $H_2O$ | 1032 | s | |
| $[SbMe_4]^+$ | $T_d$ | $H_2O$ | 780 | s | |
| $[Sb(OH)_6]^-$ | $O_h$ | $H_2O$ | 296 | s | |

st = septet, q = quartet; d = doublet; qu = quintet; t = triplet; s = singlet.

other nuclides, such as group 14 elements. For example, even if there seems to be no regularity among the shifts of the $AsR_4^+$ compounds, the ordering in terms of alkyl substituent is consistent with that found for group 14 and group 15 alkyls: labelling the atoms of the structural unit $[As(alkyl)_4]^+$ as $As_\alpha$-$C_\beta$-$C_\gamma$-$C_\delta$, methyl substitution for one hydrogen at the β position (Δβ) deshields by 11 ppm, Δγ shields by 5 ppm, and Δδ shields by 1 ppm. This trend in shielding effect as one proceeds down the chain is characteristic of methyl substitution.

In $(AsF_{6-n}Cl_n)^-$ ($n = 0$–6) derivatives, the trend observed for F/Cl exchange in $^{75}$As chemical shift also suggests an increasing shielding with increasing $n$, owing to lower electronegativity of chloride with respect to fluoride donor ligands; a further example is the difference in chemical shift for $(AsF_5Cl)^-$ (12.2 ppm) and $(AsF_5Br)^-$ (102.5 ppm).

In the case of the two heteropoly anions $(AsW_{12}O_{40})^{3-}$ and $(AsMo_{12}O_{40})^{3-}$ ('Keggin' anions), the chemical shifts reflect the influence of electron density surrounding the arsenic atom, which is dependent on the nature of peripheral metal atoms: as expected, the As atom will be shielded to a greater degree in the former anion that in the latter, the same trend being observed for phosphate anions. Moreover, both arsenic-containing heteropoly anions display chemical shifts that are upfield from the relatively deshielded arsenate ion $(AsO_4)^{3-}$.

The antimony(+5) chemical shift range spans approximately 3500 ppm. Rather more information is available for $^{121}$Sb than for $^{75}$As NMR chemical shifts. To date, $(SbS_4)^{3-}$ and $(SbBr_6)^-$ occupy the deshielded and shielded ends of the range, respectively.

In $(SbF_6)^-$ and $(SbCl_6)^-$ the small difference between the antimony shifts (~87 ppm) compared with those for analogous $^{31}$P and $^{75}$As species (152 and 392 ppm, respectively) is probably due to the electronic excitation energy that in both antimony and, more likely, arsenic is responsible for the anomalous paramagnetic shielding term.

$^{121}Sb$ chemical shifts fall into the normal halogen dependence category and, in the case of the mixed halide compounds $(SbX_nY_{6-n})^-$, it is possible to assign the *cis* and *trans* isomers by the pairwise additivity parameter, calculated using Equation [1]:

$$\delta_{\mathrm{calc}} = \frac{n}{6}\,\delta_{\mathrm{SbCl}_6^-} + \frac{6-n}{6}\,\delta_{\mathrm{SbBr}_6^-} \qquad [1]$$

The dominant contribution to the shielding, $\sigma_p$, can be expressed in the form

$$\sigma_{\mathrm{p}} = -\left(\frac{2e^2\hbar^2}{3\,\Delta m\,c^2}\right)\left(\left\langle\frac{1}{r^3}\right\rangle_{\mathrm{p}} P_{\mathrm{u}} + \left\langle\frac{1}{r^3}\right\rangle_{\mathrm{d}} D_{\mathrm{u}}\right) \qquad [2]$$

The large variation in Sb shielding arises from the dominant and negative paramagnetic shielding term, $\sigma_p$. The magnitude of $\sigma_p$ is directly proportional to changes in the inverse mean excitation energy ($\Delta E^{-1}$), the mean inverse cubes of the p and d electron–nucleus distances ($\langle r^{-3}\rangle_{np}$ and $\langle r^{-3}\rangle_{nd}$), and the valence imbalance in the p and d orbitals centred on the Sb atom ($P_u$ and $D_u$).

The relative chemical shifts of the $SbX_4^+$ and $SbX_6^-$ ions follow the valence imbalance terms and are dependent upon the electron density and relative electronegativities of the halogens. The electronegativity difference of chlorine and bromine accounts quantitatively for the high-frequency shifts of $(SbCl_4)^+$ (847.0 ppm) and $(SbCl_6)^-$ (0.0 ppm) relative to those of $(SbBr_4)^+$ (95.3 ppm) and $(SbBr_6)^-$ (–2430 ppm), and the formal positive charge, which is expected to reside mainly on the Sb atom, accounts for the high-frequency shift of the $(SbX_4)^+$ cations compared to the $(SbX_6)^-$ anions, which have their formal negative charge located on the halogens.

Only one value is reported for antimony compounds having a symmetry other than octahedral or tetrahedral: $SbCl_5$ resonates at +509 ppm, but the line width is about 8000 Hz. This very broad absorption band illustrates the effect on the rate of quadrupolar relaxation as the geometry around the antimony nucleus is reduced from $O_h$ or $T_d$ to trigonal bipyramidal ($C_{3v}$).

In the case of the heaviest bismuth nuclide, only a few $^{209}Bi$ NMR data are available yet. For example, taking $(BiF_6)^-$ (0.0 ppm, 2700 Hz $^1J(^{209}Bi–^{19}F)$) as reference, the anion $[Bi(OTeF_5)_6]^-$ is shifted to +126.7 ppm (2269 Hz $^1J(^{209}Bi–^{125}Te)$) (**Figure 1**), contrary to what is observed in analogous $[As(OTeF_5)_6]^-$ and $[Sb(OTeF_5)_6]^-$ compounds, which exhibit low frequencies with respect to $(AsF_6)^-$ (**Figure 2**) and $(SbF_6)^-$, respectively. This anomalous behaviour could be caused by relativistic effects of the bismuth atom.

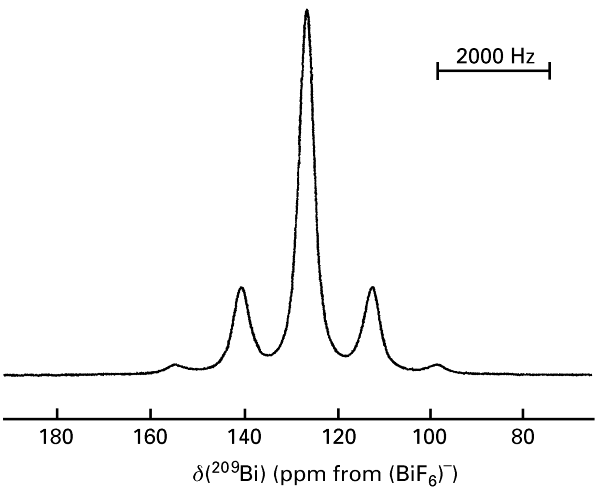



**Figure 1** $^{209}Bi$ NMR spectrum of $[Bi(OTeF_5)_6]^-$. Reproduced with permission from Mercier HPA, Sanders JCP and Schrobilgen GJ (1994) Hexakis(pentafluorooxotellurato)pnictate(V) anions, $M(OTeF_5)_6$ (M = As, Sb, Bi): a series of very weakly coordinating anions. *Journal of the American Chemical Society* **116**: 2921–2937. © 1994 American Chemical Society.

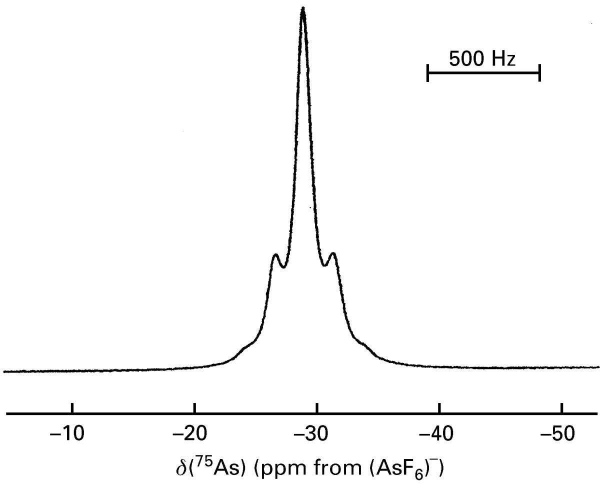



**Figure 2** $^{75}As$ NMR spectrum of $[As(OTeF_5)_6]^-$. Reproduced with permission from Mercier HPA, Sanders JCP and Schrobilgen GJ (1994) Hexakis(pentafluorooxotellurato)pnictate(V) anions, $M(OTeF_5)_6$ (M = As, Sb, Bi): a series of very weakly coordinating anions. *Journal of the American Chemical Society* **116**: 2921–2937. © 1994 American Chemical Society.

## Coupling constants

In the case of $^{75}As$, only one $J(^{75}As–^1H)$ (555 Hz) value, related to the $(AsH_4)^+$ cation, has been reported, whereas several $^1J(^{75}As–^{19}F)$ values have been obtained from the spectra of $(AsF_nCl_{6-n})^-$ derivatives,

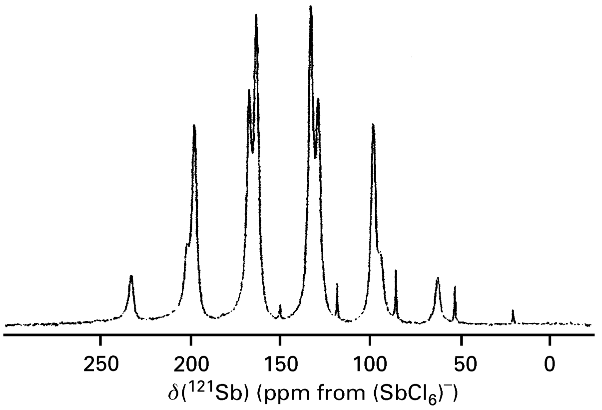



**Figure 3** $^{121}Sb$ NMR spectrum of $[SbF_nCl_{6-n}]^-$. Reproduced with permission from Dove MFA and Sanders JCP (1992) Synthesis and characterization of chlorofluoroantimonates(V). *Journal of the Chemical Society, Dalton Transactions* 3311–3316.

and these range from 897 to 1259 Hz depending on the position of the fluorine atom in the octahedron around arsenic. More recently, several $J(^{75}As–^{125}Te)$ values have been measured (420–430 Hz) in $[As(OTeF_5)_6]^-$ compounds.

Also in $[SbF_nCl_{6-n}]^-$ compounds, some $^1J(^{121}Sb–^{19}F)$ values have been observed (**Figure 3**): the typical range is 1859–2088 Hz and, as with arsenic, the values depend on the fluorine position around the antimony atom.

## Relaxation behaviour

For a quadrupolar nucleus, the relaxation under conditions of extreme narrowing can be described by

$$\Delta\nu_{\frac{1}{2}} = \frac{1}{\pi T_2} = \frac{1}{\pi T_1} = \frac{3\pi}{10}\left(\frac{2I+3}{I^2(2I-1)}\right) \times \left(\frac{e^2qQ}{h}\right)^2\left(1+\frac{\eta^2}{3}\right)\tau_c \quad [3]$$

where $\Delta\nu_{1/2}$ is the line width at half-height, $T_2$ is the spin–spin relaxation time, $T_1$ is the spin–lattice relaxation time, $I$ is the nuclear spin quantum number, $e$ is the charge on the electron, $Q$ is the nuclear quadrupole moment, $q$ is the electric field gradient (EFG) along the principal $z$ axis, $\eta$ is the symmetry parameter for the EFG, and finally $\tau_c$ is the rotational correlation time.

A careful inspection of several factors in Equation [3] reveals that line widths are dramatically reduced for nuclides having a high value of $I$ or a small value of $Q$, or both. Moreover, $\tau_c$ serves to increase the $\Delta\nu_{1/2}$ of the quadrupolar anion with increasing ionic/molecular radius, increased ion pairing and increasing solvent viscosity (decreasing temperature).

The resonance line widths for $(MF_{6-n}X_n)^-$ increase as $n$ increases from 0 to 5 (with the exception of $(SbF)^-Cl_5$), and are also greater for X = Br than for X = Cl. This behaviour reflects the increase in the quadrupolar relaxation rate of the central nucleus as its octahedral environment is distorted by the replacement of F by the larger halogens. An analogous behaviour has been noted in the bromochloroantimonates(+5).

Although the larger radius of the $(SbBr_4)^+$ cation should result in a longer $\tau_c$ and larger $\Delta\nu_{1/2}$ values than for $(SbCl_4)^+$, the opposite effect is found; this is attributed to the greater Lewis acidity of the latter cation, which is expected to form with the very weakly basic $SO_2ClF$ solvent a stronger, albeit weak, donor–acceptor bond than $(SbBr_4)^+$. An equilibrium concentration of the adduct leads to an increase in the EFG at the antimony nucleus and increased $\tau_c$, both of which serve to decrease the relaxation time and increase $\Delta\nu_{1/2}$.

A more complete study of the factors influencing relaxation behaviour has been made on arsonium salts by altering the alkyl group, the temperature, the solvent, the concentration and the counterion. It has been shown that the line width of the resonances increases with increasing size of the alkyl group, i.e. with increasing molecular mass of the species under investigation; in fact, the isotropic motion of a solute in a mobile liquid can be correlated to the correlation time $\tau_c$ by

$$\tau_c = \frac{4\pi v a^3}{3k_B T} \quad [4]$$

where $v$ is the viscosity of the solution and $a$ is the radius of the 'spherical' molecule. Moreover in the case of these small molecules, the hypothesis of extreme narrowing conditions ($\tau_c^2\omega^2 \ll 1$, i.e. fast molecular motions with respect to the resonance frequency) seems very likely to be valid.

Combining Equations [3] and [4], a linear plot of $\Delta\nu_{1/2}$ vs the experimentally measurable value $v/T$ is observed. The $\Delta\nu_{1/2}$ values, unlike $^{75}As$ chemical shifts, are generally strongly dependent on the type of solvent employed, partly surely owing to ion pairing; in fact, the line widths for $(AsEt_4)^+Br^-$ are reported to be 168, 455, 670 and 700 Hz in water, acetonitrile, dimethylformamide and dimethyl sulfoxide, respectively. Finally, the concentration of the sample influences the arsenic line width, whereas the counterion seems not to play a significant role in temperature variation.

## Applications

NMR spectroscopy of $^{75}$As, $^{121,123}$Sb and $^{209}$Bi is arousing increasing interest. $^{75}$As, $^{121,123}$Sb and $^{209}$Bi NMR studies have been found to be very useful for structural characterization in solutions of very weakly coordinating anions such as $[M(OTeF_5)_6]^-$ (M = As, Sb, Bi). These are able to stabilize Meerwein salts (tertiary carbonium salts) and cations such as $Cs^+$, $Rb^+$, $K^+$ and $Na^+$ can help in the development of the field of stable carbocation chemistry and of protosolvated superelectrophiles.

$^{75}$As NMR spectroscopy has been applied in solution in elucidating the structures of heteropoly oxometalates $(AsM_{12}O_{40})^{3-}$ (M = W or Mo), currently known as 'Keggin' anions, which are important in both fundamental and applied studies.

In some cases, $^{75}$As relaxation of the $(AsF_6)^-$ ion has been employed for probing anion sites in proteins by analysing the band shape of the $^{19}$F multiplets in terms of the transition probabilities for $^{75}$As relaxation. Outside the motional narrowing limit, the transition probabilities $P_{m,m+1}$ and $P_{m,m+2}$ for $^{75}$As relaxation become unequal. However, the actual transition probabilities dominating the line shape are weighted averages of those in the free and bound states, since binding usually occurs under fast exchange limit conditions. In this way, it has been possible to reproduce the experimental $^{19}$F line shape for the binding of $(AsF_6)^-$ to human serum albumin.

Finally, a $^{75}$As solid-state NMR study has been done on the measurement of isotropic indirect coupling to gain a better understanding of the electronic nature of semiconductor solids of the binary 13–15 type compounds, such as GaAs. $^{75}$As NMR has also been used to study the microscopic magnetic behaviour of an incipient heavy-electron compound UAs. It has been reported on the basis of the temperature dependence of the energy scale of the spin fluctuation that the relative importance of the exchange interaction with respect to the single-site dynamics in the moment relaxation generally increases on going from UP to UAs.

## List of symbols

$a$ = radius of a spherical molecule; $c$ = speed of light; $D_u$ = d orbital valence imbalance; $e$ = electron charge; $\hbar$ = Planck constant/2π; $I$ = nuclear spin quantum number; $J$ = coupling constant; $k_B$ = Boltzmann constant; $m$ = mass; $P_u$ = p orbital valence imbalance; $P_{m,n}$ = transition probabilities; $q$ = electric field gradient; $Q$ = nuclear quadrupole moment; $r$ = electron–nucleus distance; $T$ = temperature; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $\delta$ = chemical shift; $\Delta E$ = mean excitation energy; $\Delta\nu_{1/2}$ = line width at half-height; $\eta$ = electric field gradient asymmetry parameter; $\sigma$ = viscosity; $\sigma$ = shielding parameter; $\tau_c$ = rotational correlation time; $\omega$ = angular frequency.

*See also:* **Nitrogen NMR; NMR Principles; NQR Applications; NQR, Theory; Parameters in NMR Spectroscopy, Theory of; Solid State NMR, Methods; Solid State NMR Using Quadrupolar Nuclei; Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules.**

## Further reading

Casteel WJ Jr, Kolb P, Leblond N, Mercier HPA and Schrobilgen GJ (1996) *Inorganic Chemistry* **35**: 929–942.

Collins MJ and Schrobilgen GJ (1985) *Inorganic Chemistry* **24**: 2608–2614.

Dixon KR (1987) Phosphorus to bismuth. In: Mason J (ed) *Multinuclear NMR*, pp 369–402. New York: Plenum Press.

Dove MFA and Sanders JCP (1992) *Journal of the Chemical Society, Dalton Transactions* 3311–3316.

Drakenberg T (1986) Nuclear magnetic resonance of less common quadrupolar nuclei. In: Webb GA (ed) *Annual Reports on NMR Spectroscopy*, Vol 17, 231–283. London: Academic Press.

Harris RK (1978) In: Harris RK and Mann BE (eds) *NMR and the Periodic Table*, Vol 11, 379–383. London: Academic Press.

Kidd RG (1983) Group V atom NMR spectroscopy other than nitrogen. In: Lambert JB and Riddell FG (eds) *The Multinuclear Approach to NMR Spectroscopy*, NATO ASI Series C, 379–382. Boston: Reidel.

Mercier HPA, Sanders JCP and Schrobilgen GJ (1994) *Journal of the American Chemical Society* **116**: 2921–2937.

Morgan K, Sayer BG and Schrobilgen GJ (1983) *Journal of Magnetic Resonance* **52**: 139.

Wehrli FW (1979) Nuclear magnetic resonance of the less common quadrupolar nuclei. In: Webb GA (ed) *Annual Reports on NMR Spectroscopy*, Vol 9, p 177. London: Academic Press.

# Heteronuclear NMR Applications (B, Al, Ga, In, Tl)

**Janusz Lewiński**, Warsaw University of Technology, Poland

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
**Applications**

## Introduction

The naturally occurring nuclei of the Group 13 elements, $^{10}B$ and $^{11}B$, $^{27}Al$, $^{69}Ga$ and $^{71}Ga$, $^{113}In$ and $^{115}In$, $^{203}Tl$ and $^{205}Tl$, are NMR-accessible magnetically. The receptivities of these nuclei are quite high and, except for thallium, the resonance frequencies are convenient for most multinuclear spectrometers. The relevant NMR properties of these nuclei are summarized in **Table 1**. With the exception of thallium all the nuclei possess a quadrupolar moment. Thallium nuclei with spin-1/2 have a relaxation time much longer than that characteristic of the lighter nuclei. Due to the quadrupolar nature of those nuclei, resonance signals are broadened in all environments, except those where the electric field gradient approaches zero, while the quadrupolar line width effect increases down the group. In the case of $^{115}In$, the quadrupole moment is significantly large and only highly symmetric compounds are observable using $^{115}In$ NMR, even these producing extremely broad lines. From the two naturally occurring boron isotopes, $^{11}B$ is more advantageous for NMR investigation and only this nucleus is usually studied. This is due to its higher degree of abundance and its lower nuclear spin. For gallium the $^{69}Ga$ nucleus, due to its significantly lower receptivity and despite its greater abundance, is considerably less favourable than the $^{71}Ga$ nucleus. $^{115}In$ and $^{205}Tl$ have superior receptivities and higher natural abundances than the other natural isotopes of these elements, and are preferred for indium and thallium NMR. Later in this article, the NMR spectroscopy of the preferred nuclei will be discussed.

Nuclei of the Group 13 atoms have been used as NMR probes, especially in view of being able to identify molecular structures, to monitor reaction courses, detect intermediate species, and indicate chemical equilibria in exchange reactions. In general, the relative importance of the heteronuclear NMR technique is dictated by the chemistry of certain elements and because of this $^{11}B$ NMR and $^{27}Al$ NMR have been used most extensively. Structural evidence of molecules is mainly deduced from the chemical shift values. Correlations between chemical shifts and chemical parameters, such as coordination environment, different types of π-interaction, electronegativity, charges, etc. may be observed. Very often no one factor controls the chemical shift, but a combination of several interactions may be balanced to control the chemical shifts. Therefore, conclusions with respect to the structure of compounds in solution can be drawn usually from the chemical shifts when a sufficient amount of experimental data is available for comparison. In some cases, efficient techniques for calculating NMR chemical shifts may provide valuable data, strongly supporting the interpretation of experimental chemical shifts (particularly the individual gauge for the localized orbital method, IGLO, and the gauge-including atomic orbital method, GIAO).

Since a chemical shift is the most informative characteristic of heteronuclear NMR, it is, therefore, of interest to indicate briefly the factors affecting the chemical shift of a nucleus. Generally, the chemical shifts are affected by local diamagnetic ($\sigma_{dia}$) and paramagnetic ($\sigma_{para}$) contributions to the nuclear screening constant ($\sigma$): $\sigma \cong \sigma_{para} + \sigma_{dia}$. The diamagnetic contribution in the Group 13 nuclei is generally assumed to be relatively small compared with the observed chemical shift range, and remains roughly constant for each nucleus despite dramatic changes of the nucleus chemical environment, since in chemical bonding the core electrons are relatively unperturbed. Hence, the paramagnetic shielding term is considered to be responsible for the variation observed in the chemical shifts and it results from the local anisotropic motion of the valence electrons surrounding the nucleus in question. The local paramagnetic term $\sigma_{para}$ depends to a first approximation on the average electronic excitation energy $\Delta E$, the average radius of the valence orbitals and the symmetry of the distribution of bonding electrons around the nucleus in question. A convenient form of description of the local paramagnetic shielding term is the expression based on Ramsey's nonrelativistic approach:

$$\sigma_{\mathrm{para}} \sim \Delta E^{-1} \langle r^{-3} \rangle_{2\mathrm{p}} P_{\mathrm{u.}}$$

The value of $\sigma_{para}$ is larger the lower the energy of the excited states involving a rotation of charge ($\Delta E$ is small), the more asymmetric the distribution of p electrons (the $P_u$ term is large) and the closer they are to

**Table 1** NMR properties of nuclei of the Group 13 atoms. Reproduced with permission of Chapman & Hall from *Chemistry of Aluminium, Gallium, Indium and Thallium* (1993) Downs AJ (ed), Chapter 1. London: Chapman & Hall

| *Property* | $^{10}B$ | $^{11}B$ | $^{27}Al$ | $^{69}Ga$ | $^{71}Ga$ | $^{113}In$ | $^{115}In$ | $^{203}Tl$ | $^{205}Tl$ |
|---|---|---|---|---|---|---|---|---|---|
| Natural abundance (%) | 19.82 | 80.18 | 100 | 60.108 | 39.892 | 4.33 | 95.67 | 29.524 | 70.476 |
| Nuclear spin ($h/2\pi$) | +3 | −3/2 | +5/2 | −3/2 | −3/2 | +9/2 | +9/2 | +1/2 | +1/2 |
| Magnetic moment ($\mu/\mu_N$) | +1.8006 | +2.56227 | +5.529 | +5.541 | +1.622258 | +1.638215 | +2.6886 | +3.64151 | +2.01659 |
| Relative sensitivity ($^1H$ = 1.00) | $1.99 \times 10^{-2}$ | 0.17 | 0.21 | $6.91 \times 10^{-2}$ | 0.14 | 0.34 | 0.34 | 0.18 | 0.19 |
| Receptivity ($^{13}C$ = 1.00) | 22.1 | 754 | 1170 | 237 | 319 | 83.8 | 1890 | 289 | 769 |
| Magnetogyric ratio ($\times 10^7$ rad $T^{-1}s^{-1}$) | 2.8740 | 8.5794 | 6.9704 | 6.420 | 8.158 | 5.8493 | 5.8618 | 15.3078 | 15.4584 |
| Quadrupole moment ($\times 10^{-3}$ $m^2$) | 8.5 | 4.1 | 15 | 17.8 | 11.2 | 80 | 81 | – | – |
| Frequency (MHz) ($^1H$ = 100 MHz; 2.3488 T) | 10.746 | 32.084 | 26.057 | 24.003 | 30.495 | 21.866 | 21.914 | 57.149 | 57.708 |
| Width factor (Al = 1)[a] | 0.20 | 0.31 | 1.00 | 5.85 | 2.32 | 6.6 | 6.7 | – | – |
| Normal reference | | $Et_2O{\cdot}BF_3$ | $[Al(OH_2)_6]^{3+}$ | | $[Ga(OH_2)_6]^{3+}$ | | $[In(OH_2)_6]^{3+}$ | | $TlNO_3$(aq) |
| Approximate range of chemical shifts (ppm) | | 250 | 350 | | 1400 | | 1100 | | >5500 |

[a] Width factor = $Q^2(2I + 3)\, I^2\, (2I - 1)$ and is the nuclear contribution to quadrupole relaxation, normalized to Al = 1.

the origin (the radial term is large). An electronegative neighbouring atom increases the effective nuclear charge on the atom in question, thereby increasing the $\langle r^{-3}\rangle_p$ term and increasing $\sigma_{para}$. But even for closely analogous compounds, the electronegativity argument must be used with care since other effects may alter the radius and energy terms in $\sigma_{para}$ in a complex fashion. Therefore, for Group 13 compounds the substituent effect, for instance, may affect the shielding, often in a way not yet completely understood. According to the above consideration concerning the factors determining the chemical shift, a greater range of chemical shifts may be expected for the three-coordinate than for the four- and higher coordinate Group 13 compounds. In the former there is a p orbital available for π-interaction with the heteroatom lone pair. Due to this interaction a significant variation in orbital angular momentum will be present resulting from the wide variation of strength and number of π-donor ligands. This type of interaction is not available for the four-coordinate complexes since all the valence orbitals are involved in the σ bond formation.

Much less attention has been paid to magnetic relaxation measurement of the Group 13 element nuclei which may provide useful information in the study of molecular motion and interactions, and the symmetry of the nuclear environment. Nuclear magnetic relaxation is usually described both in terms of the methods of coupling of the local magnetic field with the spin system (the spin–lattice and spin–spin relaxation) and the different mechanisms. There are a number of different nuclear magnetic relaxation mechanisms: the quadrupolar mechanism which governs the relaxation behaviour of most quadrupolar nuclei, the spin–spin rotation and the chemical shift anisotropy mechanisms (important for spin-1/2 heavy metal ions), the magnetic dipole–dipole interaction of two nuclei (which is the predominant relaxation mechanism for instance for $^{13}C$ nuclei with directly bonded hydrogen), the scalar coupling mechanism, and relaxation induced by the presence of unpaired electrons. The contribution of the individual relaxation mechanisms to the overall relaxation rate is additive. Nuclei with spin $I \geq 1$ have unfavourable characteristics arising from the quadrupole moment which interacts with the electric field gradient created by the surrounding electrons. This leads to a decrease of relaxation times and, accordingly, an increase of the line widths.

## $^{11}$B NMR spectroscopy

Boron was among the first elements to attract the attention of NMR spectroscopists in the 1950s. The $^{11}B$ nucleus has a relatively small quadrupole

moment and the NMR line widths are relatively sharp, ranging between 2 and 100 Hz, usually in the 30–60 Hz range. Occasionally, such as for $(R_2N)_2B^+$ borinium cations, the line widths increase greatly to 400–800 Hz owing to the more rapid relaxation associated with an increase in the nuclear field gradient. Boron NMR spectroscopy has found widespread application in boron chemistry. The $^{11}$B chemical shifts, intensity of $^{11}$B resonances and coupling constants between boron and other nuclei are very useful for the elucidation of molecular structures and for control of reactions involving boron compounds. The peak area of a resonance is proportional to the number of boron atoms in the respective chemical environment. The chemical shift is affected by the chemical environment of the boron nucleus under investigation, i.e. the coordination number and the nature of substituents directly attached to the boron atom, and the chemical shift can be readily correlated with structure. The $^{11}$B chemical shifts of boron compounds cover a range of about 250 ppm. The solvent effect on $\delta$ $^{11}$B is of small magnitude and can usually be neglected. Noticeable solvent shifts (up to 4 ppm) have been observed only for some polyhedral boranes and related compounds. $BF_3 \cdot OEt_2$ is a reference compound, usually employed as an external standard. Several correlations of $\delta$ $^{11}$B with structural parameters in molecules have been established. The number of reported $^{11}$B NMR data is enormous. Almost any new boron-containing compound has been subjected to structural analysis by $^{11}$B NMR. In general, $^{11}$B NMR measurements are applied in two areas: organoboranes and miscellaneous compounds, and boron polyhedral borane derivatives. The first area encompasses research in organic synthesis chemistry and on the structure and dynamics of boron compounds in solution. $^{11}$B NMR provides important information on the electronic environment of a boron atom and can be useful in predicting some of the chemical behaviour of boron compounds. The second area is largely a domain of elucidation of the structure and bonding in polyhedral boranes, carboranes, azaboranes, metallaboranes and related compounds.

### Organylboranes and miscellaneous compounds

Two-coordinate boron compounds are relatively rare. The $^{11}$B chemical shifts of $(R_2N)_2B^+$ borinium cations cover a narrow range of 35–38 ppm. However, the $\delta$ $^{11}$B values of iminoboranes, R–B≡N–R′, fall in the range of 3 to 22 ppm and are sensitive to the nature of the substituents at the site of the imino nitrogen atom. The trend in the $^{11}$B shifts corresponds closely to that observed for the $^{13}$C chemical shifts of similarly substituted alkynes.

$^{11}$B chemical shifts have proved to be an extremely valuable tool for distinguishing between three- and four-coordinate boron compounds. Chemical-shift data of representative compounds are given in **Table 2**. Furthermore, a schematic structure representation of compounds [1] – [5] with the corresponding chemical shift values, provides a good illustration of the power of $^{11}$B NMR for intimate understanding of the nature of π-interactions in boron compounds. In trialkylboranes of a trigonal planar structure [1] the p orbital remains unoccupied (π-bonding is not an obvious possibility), resulting in a great imbalance of bonding electrons around the boron atom and leading to a large shift of the $^{11}$B NMR signal to low field ($\delta \sim 86$ ppm). By contrast, another triorganylborane compound, 7-borabicyclo[2.2.1]heptadiene ([2], $\delta$ –5 ppm), exhibits a significant upfield shift, and the magnitude of this shift is rationalized in terms of the interaction of the empty p orbital on boron with the π-clouds of two C=C bonds. $^{11}$B NMR spectroscopy has also been commonly employed to provide information about the contribution of π-bonding between boron and carbon centres in many other cyclic and open-chain unsaturated organylboron compounds, i.e. vinylboanes, alkynylboranes, allylboranes, borinenes, boroles and boron derivatives of six-membered rings.

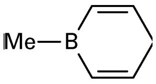
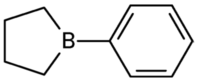
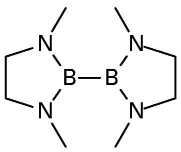
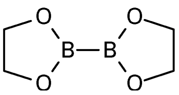
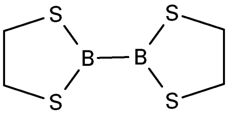
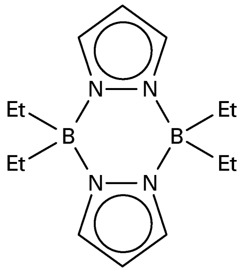
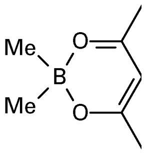
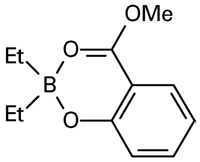
Referring to the $R_3B$ compounds, replacement of the R substituent by a strong electronegative atom or ligand X with a lone pair available for π-interaction results in a high-field shift. For example, the $\delta$ $^{11}$B values for three-coordinate $R_2BNR'R''$ and $R_2BOR'$ compounds are found approximately in the range 44–50 ppm and 50–55 ppm, respectively. Hence, the $^{11}$B signal of $Et_2B$(benzoinato) [4] at 55 ppm is indicative of a three-coordinate environment at boron with the uncoordinated carbonyl group of the bifunctional benzoinato ligand. In this case $^{11}$B NMR gives an accurate picture of the π-interaction and is additionally a good indicator of the relative strength of a boron–oxygen π-bond vs. boron–carbonyl dative bond. It is worth noting that the solution structure of compound [4] is consistent with that found in the solid state. From the differences in $^{11}$B chemical shifts of three-coordinate diorganylboranes $R_2BX$ the following order of increasing π-interaction for various X results: $PR_2 < SR < Cl < OR < NR_2 < F$. This is only a qualitative interpretation and a quantitative correlation should not be expected from chemical shift values since the magnitude of the shift may be influenced by other electronic and steric effects, e.g. various functionalities of the ligand X have a remarkable influence on the $^{11}$B nuclear shielding.

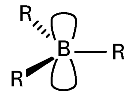
Further successive replacement of an alkyl group by the ligand X leads to a progressive upfield shift of $^{11}B$ resonances and the magnitude of shift depends on the nature of the substituent X. Minor changes in the chemical shift values are observed for sulfide and then for chloride derivatives, and more significant changes are for X = OR, $NR_2$ and F. A great variety of noncyclic $RBX_2$ and cyclic RB(X,X) compounds have been studied by $^{11}B$ NMR spectroscopy. It is interesting to note that for $BX_3$ compounds the order of increasing substituent effect is somewhat different from that for $R_2BX$ diorganylboranes: SR < Cl < $NR_2$ < OR < F. $^{11}B$ NMR is extremely useful in the investigation of the chemistry of three-coordinate diborane compounds, X′X″B–BX′X″. The presence of the B–B bond is reflected in the $^{11}B$ resonance low-field shift compared with that in the corresponding organylboranes RBX′X″.

$^{11}B$ NMR is very useful in the investigation of the rich chemistry of transition metal complexes with boron-containing heterocycles like borole and borabenzene derivatives or various boron–nitrogens and boron–oxygen cyclic compounds. In general, the metal–boron bonding interaction is reflected by a $^{11}B$ high-field shift compared with that of the corresponding heterocycles. In the triple-decker complexes, the boron of the bridging ligand (to which two metals are bonded) is more shielded than that in the end or capping boron-containing ring to which only one metal atom is bonded.

$^{11}B$ NMR provides valuable information concerning the interaction between three-coordinate boranes

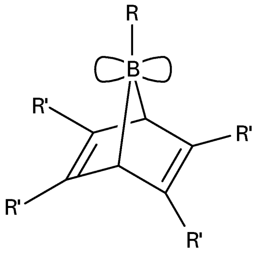
**Table 2** $^{11}B$ chemical shifts of some representative three-coordinate and four-coordinate organylboranes and miscellaneous compounds

| *Compound* | *δ $^{11}B$* | *Compound* | *δ $^{11}B$* | *Compound* | *δ $^{11}B$* |
|---|---|---|---|---|---|
| **CN = 3** | | | | | |
| $Me_3B$ | 86.6 | $Ph_3B$ | 68.0 | $[MeB(NH)]_3$ | 34.5 |
| $(CH_2{=}CH_2)_3B$ | 56.4 | $Me_2BPh$ | 77.6 | $(MeBO)_3$ | 33.2 |
|  | 52.8 |  | 84.5 | $(^nBuBS)_3$ | 68.4 |
| $Me_2BF$ | 60.1 | $MeBF_2$ | 8.2 | $BF_3$ | 10.0 |
| $Et_2BCl$ | 78.0 | $EtBCl_2$ | 63.4 | $BCl_3$ | 46.5 |
| $Et_2BBr$ | 81.9 | $EtBBr_2$ | 65.6 | $BBr_3$ | 38.7 |
| $Et_2BI$ | 84.4 | $EtBI_2$ | 55.9 | $BI_3$ | −7.9 |
| $Me_2BNMe_2$ | 44.6 | $MeB(NMe_2)_2$ | 33.5 | $B(NMe_2)_3$ | 27.6 |
| $Me_2BOMe$ | 53.0 | $MeB(OMe)_2$ | 29.5 | $B(OMe)_3$ | 18.3 |
| $Me_2BSMe$ | 73.6 | $MeB(SMe)_2$ | 66.3 | $B(SMe)_3$ | 61.0 |
|  | 33.7 |  | 31.5 |  | 68.3 |
| **CN = 4** | | | | | |
| $Me_3B{\cdot}NMe_3$ | 0.1 | $H_3B{\cdot}NMe_3$ | −8.3 | $BMe_4^-$ | −20.2 |
| $(CH_2{=}CH_2)_3B{\cdot}NMe_3$ | −3.0 | $H_3B{\cdot}SMe_2$ | −20.1 | $B(OMe)_4^-$ | 2.7 |
| $Me_3B{\cdot}PMe_3$ | −12.3 | $F_3B{\cdot}NMe_3$ | 0.6 | $BF_4^-$ | −1.6 |
| $(C_6F_5)_3B{\cdot}O{=}C(N^iPr_2)Ph$ | −0.1 | $Cl_3B{\cdot}NMe_3$ | 10.2 | $BCl_4^-$ | 6.7 |
| $(C_6F_5)_3B{\cdot}O{=}C(H)Ph$ | 5.0 | $Br_3B{\cdot}NMe_3$ | −3.3 | $BBr_4^-$ | −23.8 |
| $(C_6F_5)_3B{\cdot}O{=}C(OEt)Ph$ | 19.2 | $I_3B{\cdot}NMe_3$ | −54.4 | $BI_4^-$ | −127.5 |
|  | 2.2 |  | 13.0 |  | 17.5 |



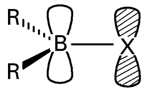

[1]

($\delta \approx 86$ ppm)



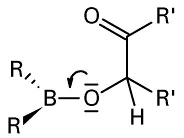

[2]

(R = Me, R' = Ph; $\delta$ −5 ppm)



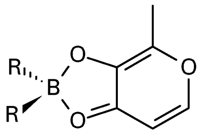

[3]

($\delta \approx$ 40–55 ppm)




[4]

(R = Et, R' = Ph; $\delta$ 55 ppm)




[5]

(R = Et; $\delta$ 22 ppm)

and donor ligands and enables prediction of some of the chemical behaviour. The formation of the four-coordinate Lewis acid–base adducts is evident from the $^{11}$B NMR spectrum. When the coordination number of boron increases from three to four, an upfield shift of the $^{11}$B resonance is generally observed. The interaction between a borane and a donor ligand often results in a dynamic equilibrium and $^{11}$B NMR measurements at variable temperatures are particularly helpful for investigations of this type of equilibrium. The influence of various ratios of donor and acceptor molecules upon the equilibrium can be determined by $^{11}$B NMR spectroscopy. However, the utility of $^{11}$B NMR to determine the Lewis acid strength of three-coordinate boranes is strongly limited, since the magnitude of chemical shift change upon coordination reflects the $\pi$-bonding effect in a parent borane as well as the sensitivity to steric factors. Similarly, an observed order of ligand basicity usually corresponds to certain boron-system affinities and is strongly affected by steric effects. The structures of dialkylboron derivatives of saturated and unsaturated hydroxy carbonyl compounds, $Et_2B$ (benzoinato) [4] and $Et_2B$ (maltolato) [5], are a good illustration of the importance of $\pi$-bonding between boron and oxygen. The $^{11}$B signal of [5] at 22 ppm is significantly shifted to high field when compared with that for complex [4] and the chemical shift value falls in the region associated with neutral, four-coordinate boron nuclei. This result is consistent with the chelation of the carbonyl group to the boron atom and the four-coordinate chelate structure [5] in solution. Thus, these data, based on $^{11}$B NMR and X-ray crystallographic studies, provide strong evidence that for organylborane derivatives of unsaturated hydroxy carbonyl compounds the $\pi$-interaction of the alkoxide oxygen lone pairs with the chelate-ligand-extended $\pi$ system may dominate and effectively obstruct the boron–oxygen $\pi$-interaction.

Many coupling constants between directly bonded nuclei can be obtained from $^{11}$B NMR spectra, especially for four-coordinate boron, where in many cases the quadrupolar relaxation rate is sufficiently slow. The accuracy is somewhat limited because of the line width of the resonance signals. Therefore, the use of a high magnetic field strength is advantageous. One-bond $^{11}$B–X coupling constants from 1–5 Hz in $BF_4^-$ up to more than 1000 Hz for X = $^{119}$Sn or $^{207}$Pb have been observed. For terminal boron–hydrogen bonds, the $^1J(^{11}B–^1H)$ values are in the range 100–190 Hz, but when a hydrogen bridge is involved $^1J(^{11}B–^1H)$ is less than 80 Hz. Thus, it is possible to use these coupling constants to determine the coordination at each boron in a boron hydride when resolution permits. Despite the wide use of $^{11}$B NMR for the characterization of boron compounds, relatively little attention has been paid to $^{11}$B relaxation. Typical values for relaxation times of $^{11}$B nuclei are in the range of 10–200 ms, commonly 10–50 ms.

## Polyhedral boranes

The development of boron cluster chemistry is intimately coupled with $^{11}$B NMR which has been the most efficient method of elucidating the structure

and bonding in polyhedral boranes and related clusters. At present, based on a great amount of $^{11}B$ NMR experimental data, meaningful conclusions with respect to the structure of boron skeleton compounds in solution can be drawn from the chemical shifts. Polyhedral boranes are electron-deficient systems involving multicentre bonding, i.e. B–H–B, B–B–B, B–C–B, and with highly delocalized skeletal electrons. The majority of boron vertices in borane clusters are formally $sp^3$ hybridized; deviation from the ideal tetrahedral geometry, i.e. the imbalance of bonding electrons, is the most important factor affecting the chemical shift of individual skeletal atoms. In borane cluster compounds the influence of electron density of individual skeletal atoms on chemical shift is significant, but is not the primary factor. $^{11}B$ chemical shift values of polyhedral boranes and related clusters span approximately 75 to −60 ppm. Many empirical rules for predicting $\delta$ $^{11}B$ values have been established which are based on various effects including those related to the coordination number, bridging hydrogens, *endo*- and *exo*-cluster substituents, cluster shape and antipodal atoms. For example, the replacement of a boron nucleus in a polyhedral borane by a heteroatom results in a downfield $^{11}B$ chemical shift for the boron across the cage from the heteroatom, usually the so called antipodal effect. The largest effect was observed for the antipodal boron B(10) in the 10-vertex series *closo*-1-$EB_9H_9$ heteroboranes, e.g. $\delta$ $^{11}B$ values of B(10) are −2.0, 28.4 and 74.5 ppm for E = $BH^{2-}$, $CH^-$ and S, respectively. At present, novel two-dimensional (2D) NMR techniques, i.e. $^{11}B$–$^{11}B$ 2D COSY and $^1H$–$^{11}B$ 2D spectra, usually allow unambiguous $^1H$ and $^{11}B$ signal assignments for most polyhedral boranes. In applying the *ab initio*–IGLO–NMR technique, the various competitive structures are subject to *ab initio* structural optimization, following which IGLO calculations are used to predict the sets of $^{11}B$ chemical shift values to be expected of each *ab initio* optimized structure.

## $^{27}$Al NMR spectroscopy

The first $^{27}Al$ NMR measurements were performed in the early 1960s. However, the 1980s and 1990s have brought a considerable increase in investigations of the nature of aluminium compounds using $^{27}Al$ NMR spectroscopy. The usual reference for chemical shifts is $[Al(H_2O)_6]^{3+}$ (an acidic solution of Al(III) salt, line width several Hz) or $Al(acac)_3$ in benzene solution (line width ~100 Hz; acac = acetylacetonate). The $^{27}Al$ NMR chemical shifts of the most often studied organometallic and inorganic aluminium (III) compounds cover a range of approximately 350 ppm, from 290 to −60 ppm. However, there are some rare examples such as the $[Cp_2Al]^+$ cation ($\delta$ −126.4 ppm) (Cp = $C_5H_5$) or the salts with transition metals, i.e. $Co(AlCl_4)_2$ ($\delta$ −225 ppm). The line width of $^{27}Al$ resonances is highly sensitive to the symmetry of the aluminium coordination sphere and ranges from a few Hz for highly symmetric species to a few kHz. However, in some cases the resonance may even be undetectably broad and a given spectrum does not display all the response of the aluminium nuclei present. In the standard $^{27}Al$ NMR spectrum the intensity ratio of resonances cannot usually be interpreted quantitatively with respect to the number of aluminium atoms in the respective chemical environment, due to this strong effect on line widths.

Currently, almost any new organoaluminium compound is routinely characterized by $^{27}Al$ NMR. Earlier investigations suggested that the observed $^{27}Al$ chemical shift is diagnostic of the compound coordination state. However, an increasing amount of experimental data has shown that the chemical shift ranges corresponding to the different coordina-

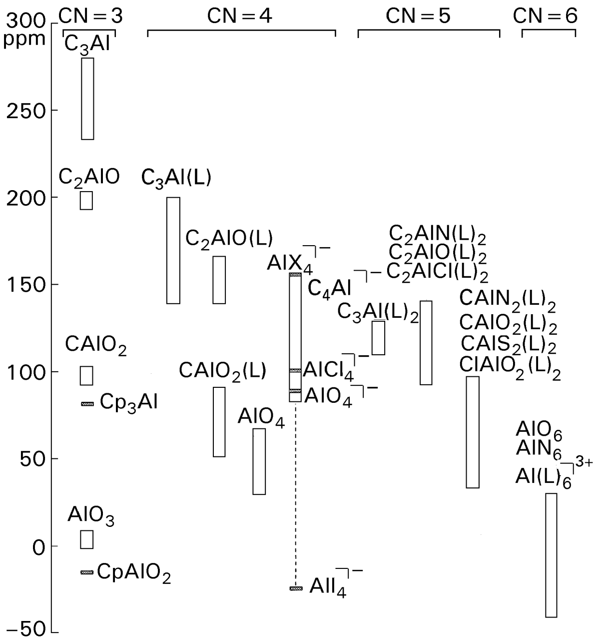



**Figure 1** $^{27}Al$ chemical shift ranges of aluminium compounds, demonstrating the dependence of aluminium shielding on the nature of the coordination environment and the coordination number (CN) of the aluminium atom. L is a monodentate ligand (e.g. $NR_3$, pyridine, $OR_2$, $R_2C{=}O$, $H_2O$, $PR_3$, $SR_2$) or may denote a neutral chelating group from a monoanionic bidentate ligand (e.g. the carbonyl group in chelate complexes [6] and [9], hence the coordination environment of these complexes is described as $C_2AlO(L)$ and $C_2AlO(L)_2$, respectively) as well as one of the bridging alkoxide oxygen atoms in associate compounds (e.g. an alkoxide oxygen atom in dimer [8]). X is a ligand carrying one negative charge (e.g. H, R, OR, OH, halogen).

tion states largely overlap and a certain value may not be simply interpreted according to the number of ligands. Shielding data subdivided with respect to the aluminium coordination number, presented graphically in **Figure 1** and for a series of compounds collated in **Table 3**, show that the nature of substituent functions significantly governs the $^{27}$Al NMR chemical shifts. Hence, although characterization of organoaluminium compounds by $^{27}$Al NMR has recently become routine, the $^{27}$Al NMR spectra still pose significant interpretational difficulties. Additionally, the important difficulty in resonance identification is also due to a background signal that in some cases (especially for dilute solutions) may dominate in a measured spectrum (**Figure 2**). This relatively broad 'probe head signal' falls at about 60 ppm and when this region of the spectrum is of interest then it is advisable to use relatively highly concentrated solutions and to identify possible extraneous signals with blank samples.

Because of the relatively fast relaxation of the $^{27}$Al nucleus, spin–spin coupling is not generally observed in $^{27}$Al NMR spectra. However, a number of coupling constants have been reported. The $^1J_{AlH}$ and $^2J_{AlH}$ coupling constants were found to be 174 Hz and 6.3 Hz, respectively. The observed $^1J_{AlP}$ and $^2J_{AlP}$ coupling constants are in the range of 90–290 Hz and 6–30 Hz, respectively. The $^1J_{AlF}$ is ~19 Hz, $^1J_{AlB}$ is ~9 Hz, $^1J_{AlC}$ is ~73 Hz and $^1J_{AlN}$ is ~22 Hz. Only a few investigations have been made of the spin–lattice and spin–spin relaxation times of the $^{27}$Al nucleus. Relaxation of the $^{27}$Al nucleus is dominated by the quadrupolar interaction, as expected from its relatively high quadrupolar moment. For highly symmetrical species the relaxation rates are a few Hz and other relaxation mechanisms may be significant. For example, the relaxation of the $^{27}$Al nucleus in the solvate complex $[Al(CH_3CN)_6]^{3+}$ and in anions $[AlCl_nBr_{4-n}]^-$ proceeds by means of the quadrupole and scalar mechanisms.

### Organoaluminium (III) and miscellaneous compounds

Various factors that influence the aluminium nuclear shielding may be employed for specifying different

**Table 3** $^{27}$Al chemical shifts of some representative three-, four-, five- and six-coordinate aluminium compounds

| *Compound* | *δ $^{27}$Al* | *Compound* | *δ $^{27}$Al* | *Compound* | *δ $^{27}$Al* |
|---|---|---|---|---|---|
| **CN = 3** | | | | | |
| $^i$Bu$_3$Al | 276 | $^t$Bu$_3$Al | 255 | (mesityl)$_3$Al | 260 |
| $^i$Bu$_2$Al(bht) | 196 | $^{ii}$BuAl(bht)$_2$ | 109 | Al(bht)$_3$ | 3 |
| Cp$_3$Al | 81.7 | Cp$_2$MeAl | 72.7 | Cp$_2$Al(bht) | –15 |
| **CN = 4** | | | | | |
| [Me$_3$Al]$_2$ | 153 | Me$_3$Al(thf) | 182 | Et$_4$Al$^-$ | 154 |
| [Et$_3$Al]$_2$ | 152 | Et$_3$Al(py) | 167 | AlH$_4$$^-$ | 103 |
| [Me$_2$AlH]$_3$ | 159 | Me$_2$AlH(thf) | 179 | AlCl$_4$$^-$ | 102 |
| [Me$_2$AlCl]$_2$ | 180 | Me$_2$AlCl(pyrazole) | 157 | AlBr$_4$$^-$ | 80 |
| [Me$_2$AlOMe]$_3$ | 152 | Me$_2$Al(bht)(py) | 134 | AlI$_4$$^-$ | –27 |
| [Et$_2$AlOEt]$_2$ | 151 | MeAl(bht)$_2$(py) | 72 | AlBr$_2$I$_2$$^-$ | 37 |
| [Et$_2$AlNEt$_2$]$_2$ | 160 | Al(Omesityl)$_3$(py) | 50 | Al(OH)$_4$$^-$ | 80 |
| Me$_2$Al(dpa) | 151 | [Cl$_2$AlOEt]$_3$ | 98 | [($^t$BuO)$_3$Al]$_2$ | 48 |
| Et$_2$Al(acac) | 140 | Cp*$_3$Al(thf) | 151 | [AlCl$_3$]$_2$ | 91 |
| Cl$_2$Al(acac) | 88 | Me(Cp*)AlCl]$_2$ | 8.5 | [AlBr$_3$]$_2$ | 75 |
| (Ph$_3$SiO)$_2$Al(acac) | 53 | [Cp*AlCl$_2$]$_2$ | –53 | [AlI$_3$]$_2$ | –11 |
| **CN = 5** | | | | | |
| Me$_2$Al(hacet)(py-Me) | 118 | [Me$_2$Al(lak)]$_2$ | 114 | MeAl(hacet)$_2$ | 68 |
| Me$_2$Al(amket)(py-Me) | 120 | [Et$_2$AlO(CH$_2$)$_2$NEt$_2$]$_2$ | 112 | MeAl(amket)$_2$ | 70 |
| MeAl(hacet) | 64 | [Et(EtO)Al(hacet)]$_2$ | 60 | [Cl(EtO)Al(hacet)]$_2$ | 60 |
| **CN = 6** | | | | | |
| Al(acac)$_3$ | 0.0 | [Al(acac)$_2$(thf)$_2$]$^+$ | 4 | [Al(OH)$_6$$^{3-}$] | 0.0 |
| Al(dpa)$_3$ | 26 | [Al(H$_3$ppma)$_2$]$^{3+}$ | –12.9 | [AlF$_6$]$^{3-}$ | –1.2 |
| Al(maltolate)$_3$ | 38 | [Al(MeCN)$_6$]$^{3+}$ | –33 | [Al(NCS)$_6$]$^{3-}$ | –33.7 |

bht = deprotonated 2,6-di-*tert*-butyl-4-phenol; Cp = $C_5H_5$; Cp* = $C_5Me_5$; thf = tetrahydrofuran; py = pyridine; py-Me = γ-picoline; acac = acetylacetonate; hacet = deprotonated 2′-hydroxyacetophenone; lak = deprotonated ethyl lactate; amket = deprotonated 2′-aminoacetophenone; H$_3$ppma = tris[4-(phenylphosphinato)-3-methyl-3-azabutyl]amine; dpa = deprotonated 2,2-dipyridylamine.

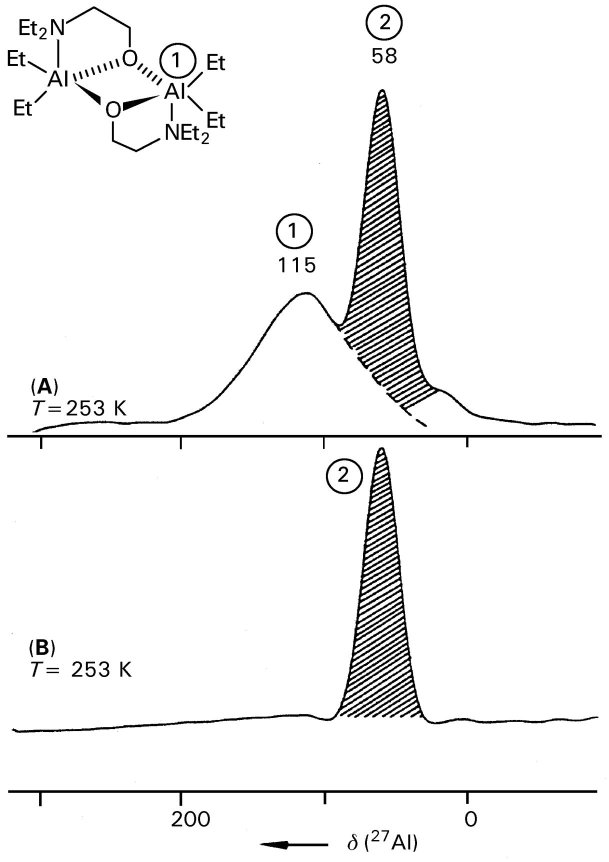
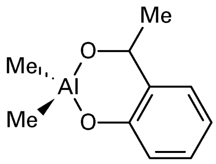



**Figure 2** Background signal (2) in the $^{27}$Al NMR spectrum of a multinuclear probe head. (A) $^{27}$Al NMR spectrum of a 0.1 M solution of $[Et_2AlO(CH_2)_2NEt_2]_2$. (B) Spectrum obtained with the test-tube filled with solvent only. For resolution enhancement, the same Lorentz–Gauss transformation was used in both cases before Fourier transformation. Reprinted with permission of Wiley-VCH from Benn R and Rufinska A (1986) *Angewandte Chemie, International Edition in English* **25**: 861–881.

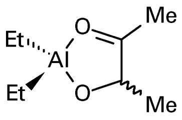
organoaluminium species in solution. Thus, the chemical shifts for the three-coordinate aluminium compounds range from 260 to 0 ppm, and the electronic nature of the ligand determines the extent of metal shielding. The majority of large chemical shift resonances of large chemical shift are generally exhibited by the coordinatively unsaturated three-coordinate $R_3Al$ compounds (~260 ppm). Successive replacement of the alkyl substituent by a sterically hindered aryloxide ligand leads to a progressive decrease in chemical shift of the $^{27}$Al resonance in the series of three-coordinate monoaryloxide (~190 ppm), bisaryloxide (~100 ppm) and trisaryloxide (~0 ppm) compounds. These phenomena can be explained in terms of a π-interaction between the aluminium empty p orbital and the aryloxide oxygen lone pairs. The increasing number of strong electronegative aryloxide ligands attached to aluminium probably enhances the π-acceptor strength of aluminium and strengthens the Al–O π-bonding, which leads to decreasing chemical shift. In the corresponding four-coordinate Lewis acid–base adducts, $R_3Al(L)$, the aluminium nucleus is significantly shielded due to the new dative bond formed and $^{27}$Al chemical shifts are substantially smaller than that of the parent monomeric three-coordinate compound. A similar behaviour results when the alkyl groups in $R_3Al(L)$ are successively replaced by aryloxide substituents, although in this case the observed decrease in chemical shift is not as great as for three-coordinate compounds. For the four-coordinate $AlX_4^-$ anionic species, where X = H, R, halogen, OH or OR, an appreciable increase in chemical shift is observed for the tetraalkylaluminium anions and a decrease is found for the $AlI_4^-$ anion. The $\delta$ $^{27}$Al values of $AlX_4^-$ generally follow the electronegativity of X, and some discrepancies may be attributed to the anisotropic contribution of various ligands because no π-bonding is likely to take place in these anions. These highly symmetrical species can be easily identified by their chemical shift and by the very sharp resonance.

Recently, extensive X-ray work has revealed a large variety of solid-state structures of cyclopentadienylaluminium compounds. In combination with the X-ray data, $^{27}$Al NMR spectroscopy in solution provides information about the contribution of π-bonding between aluminium and carbon centres. For example, inspection of selected data for the three- and four-coordinate cyclopentadienyl–aluminium complexes collated in **Table 3** shows that the shielding effect is most pronounced for situations where the aluminium centre, irrespective of its coordination number, bears electronegative substituents. This is consistent with the expectation that the electron-withdrawing group should enhance π-bonding in that Al–Cp interaction.

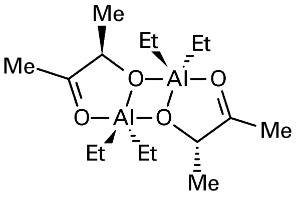
The $\delta$ $^{27}$Al values for an extensive series of the four-coordinate $R_2Al(\mu\text{–}X)_2AlR_2$ compounds with alkoxide or amide bridging ligands fall in a very narrow range of approximately 160–140 ppm. In the same range are observed chemical shifts of a substantially different class of compounds, the monomeric four-coordinate dialkylaluminium chelate complexes $R_2Al(X,Y)$, where X,Y = O,O′–, O,N′– or N,N′–unsaturated bifunctional ligands with both unsaturated bifunctional atoms involved in a π conjugated bond system. However, a pronounced change in the shielding of aluminium resulting from the replacement of both R substituents in the $R_2Al(O,O')$-type complexes by chloride or alkoxide ligands is reflected by increased shielding of the $^{27}$Al resonance of approximately 60 ppm or 110 ppm, respectively. A



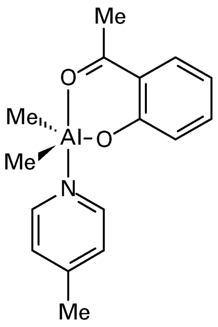
[6]
($\delta$ 144 ppm)



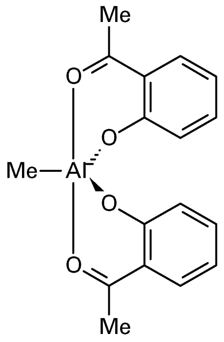
[7]
($\delta$ 122 ppm)



[8]
($\delta$ 131 ppm)



[9]
($\delta$ 128 ppm)



[10]
($\delta$ 64 ppm)

unique shielding situation arises for the dialkylaluminium chelate derivatives of saturated ketone-functionalized alcohols. This group of compounds exists in solution as an equilibrium mixture of the four-coordinate monomer $R_2Al(O,O')$ and the five-coordinate dimer $[R_2Al(O,O')]_2$. The monomeric chelate complexes have the $^{27}Al$ shielding greater than those of related dialkylaluminium complexes derived from the unsaturated hydroxy carbonyl compound. This is illustrated by complexes $Me_2Al$(hacet) [6] and $Et_2Al$(acet) [7] (where hacet = deprotonated 2-hydroxyacetophenone, acet = deprotonated 3-hydroxy-2-butanone) and the corresponding chemical shift values. Furthermore, there is observed a decreased shielding of the $^{27}Al$ resonance with the association of monomeric species to the five-coordinate adduct which is opposite to the expected coordination number effects, e.g. [7] ($\delta$ 122 ppm) and $[Et_2Al(acet)]_2$ [8] ($\delta$ 131 ppm). The observed relatively high shielding of aluminium in the four-coordinate dialkylaluminium derivatives of saturated ketone-functionalized alcohols was reasonably explained by $\pi$-donation from the alkoxide-oxygen lone-pair orbitals to the $\sigma^*$ antibonding orbital of the dative bond between the aluminium atom and the oxygen atom of the chelating carbonyl group. This type of interaction is rather unlikely for the related aluminium compounds with unsaturated chelating ligands due to the involvement of oxygen lone pairs in a chelate conjugated bond system. Thus $^{27}Al$ NMR spectroscopy may be diagnostic of another type of $\pi$-interaction, i.e. $\pi$-bonding between the alkoxide oxygen lone pairs and Al–X $\sigma^*$ antibonding orbitals in four-coordinate aluminium alkoxides.

Simple monomeric five-coordinate aluminium complexes are usually relatively unstable and/or highly reactive (when not stabilized by multidentate ligands) and $^{27}Al$ NMR spectroscopy becomes a powerful tool for the characterization of these species in solution. The available $\delta$ $^{27}Al$ data show that the aluminium chemical shifts in the series of the five-coordinate organoaluminium compounds varies with the number of alkyl groups attached to aluminium. The $\delta$ $^{27}Al$ values of the five-coordinate dialkylaluminium complexes span a relatively narrow range of approximately 135–100 ppm, irrespective of whether they occur as monomeric or dimeric moieties, e.g. $R_2Al(O,O')\cdot L$ or $[R_2Al(O,O')]_2$. The nature of chelate ligands bonded to aluminium does not affect significantly the $^{27}Al$ chemical shift. Furthermore, for the dialkylaluminium chelate derivatives of unsaturated hydroxy carbonyl compounds, $R_2Al(O,O')$, the observed shielding effect on the five-coordinate adduct formation, $R_2Al(O,O')\cdot L$, is

approximately in the range of 40–15 ppm, e.g. the shielding effect for the chelate complex [6] and its adduct with γ-picoline [9] is 16 ppm. For the monoalkylaluminium five-coordinate compounds, an increase in shielding with respect to the dialkylaluminium derivatives is observed and their $\delta$ $^{27}$Al fall in the range of 70–50 ppm (see for example $MeAl(hacet)_2$ [10]).

### $Al^{3+}$ species in solution

For years, $^{27}$Al NMR spectroscopy has been extensively used to investigate the physical and chemical properties of aluminium salts in polar solvents with regard to solvation, ion pairing, hydrolysis and complexation. These studies involve mainly the octahedrally coordinated aluminium species since the solvates of $Al^{3+}$ are octahedral, except where steric restrictions favour lower coordinated complexes. A correlation between the chemical shifts of aluminium hexasolvated complexes and the donor number (DN) (empirical scale of Lewis basicity) of the ligand has been observed, as might be expected for a metal ion with a large charge density whose chemical shift is dominated by the $\sigma_{para}$ term. Al(III) solvation and ion pair $[Al(H_2O)_nL_{6-n}]^{3+}$ complexes can be readily studied. Quadrupole broadening usually precludes the observation of signals for geometrical isomers of these complexes. The $\delta$ $^{27}$Al for certain $Al^{3+}$ species is essentially independent of solution concentration and solvent composition. This indicates that the metal-ion electronic environment and subsequent chemical shift depend only on the composition of the principal solvation shell, and are not influenced by long-range outer-shell effects. For the octahedrally coordinated aluminium species $[Al(L^1)_{6-n}(L^2)_n]^{3+}$, when a $L^1$ ligand is gradually replaced by a $L^2$ ligand, the induced shift is usually additive. Very often the ligand exchange rate is slow enough to permit the observation of several solvated species and the kinetic and thermodynamic parameters may be determined.

$^{27}$Al NMR investigations play an important role in the control of the degree of hydrolysis and further of the polymerization of the aqueous Al(III) solution. Various aggregates give distinctive resonances of different $AlO_4$, $AlO_5$ and $AlO_6$ moieties, e.g. species such as the dimer $[Al_2(OH)_2(H_2O)_8]^{4+}$ and the polyoxocation $[AlO_4Al_{12}(OH)_{24}(H_2O)_{12}]^{7+}$ were identified. $^{27}$Al NMR spectroscopy can be used to study the characterization of $Al^{3+}$ species in aqueous solution over a wide range of concentration and pH. For example, a conceptual basic drawback in the interpretation of aluminium toxicological data lies to a great extent in the ill-defined nature of the aluminium species administered at physiological pH and in the absence of strongly complexing agents. Hence, there are numerous applications of $^{27}$Al NMR measurements to probe biological processes, e.g. $^{27}$Al NMR is a useful probe for the Al(III)–phosphate system and in this respect studies of aqueous solutions of Al(III) salts with biologically relevant phosphates were carried out.

### Aluminium(I) compounds

Monovalent aluminium compounds have been synthesized and characterized during the last decade. The calculated $\delta$ $^{27}$Al for the Al(I) compounds cover a relatively large range, over 1000 ppm, and indicate that the shift depends strongly on the nature of the bonded ligand. There are only a few experimental data. For example, the chemical shift for AlCl in toluene–ether solution is 35 ppm and for the tetramer $[Cp^*Al]_4$ and the monomer Cp*Al ($Cp^* = C_5Me_5$ in benzene the corresponding values are –80 ppm and –150 ppm, respectively. The observed line widths for the Al(I) compounds range from 100 to 1800 Hz.

### Solid-state $^{27}$Al MAS NMR

The increasing popularity of $^{27}$Al NMR is strongly connected with high-resolution magic angle spinning (MAS) solid-state NMR, which allows investigation of polycrystalline and amorphous materials, which are of great practical importance. Understanding the structure of these materials is fundamental to many problems in materials science. High-resolution solid-state NMR is particularly valuable in quantifying the extent and nature of disorder. Another reason for the interest in solid-state studies is that the fluxional behaviour is often frozen out and consequently the NMR results are more amenable to correlation with X-ray studies, and these studies can act as a bridge between the crystal and solution structures. For example, $^{27}$Al MAS NMR spectra have been extensively used to inspect the structural character of the aluminium species in zeolites. One of the most important applications yields the state of aluminium, shows the transformation of the zeolite symmetry, and yields access to the chemical status of the extra-framework aluminium species after stream treatment of zeolites. $^{27}$Al MAS NMR is one of the main techniques for investigating cement, glasses, sol-gel precursors of the zeolites and other ceramic materials. Simple MAS spectra are usually limited by the low resolution generally caused by second-order broadening, and special techniques such as double rotation or two-dimensional MAS may be used to enhance the resolution.

## $^{71}$Ga NMR spectroscopy

$^{71}$Ga NMR spectroscopic studies are much less common due to the characteristically broad resonances, associated with the large quadrupole moment, and low receptivity, and partly because Ga chemistry has attracted much less attention than that of B or Al. The chemical shifts are referred to $[Ga(H_2O)_6]^{3+}$. $^{71}$Ga NMR data are rather scant and at present no meaningful conclusion with respect to the structure of gallium compounds in solution could be drawn from an interpretation of the chemical shift values. However, in view of the significant structural similarity of analogous gallium and aluminium compounds it is reasonable to expect that the general observation concerning the correlation between $^{27}$Al shifts and the structure of aluminium species may be transferred to gallium. **Table 4** presents the $^{71}$Ga chemical shifts for some gallium compounds.

$^{71}$Ga NMR spectroscopic studies have been used in the considerable research of the chemical nature of gallium(III) hydrolysis species. For example, it was found that the base hydrolysis of gallium(III) salt solutions led to the $[GaO_4Ga_{12}(OH)_{24}(H_2O)_{12}]^{7+}$ polyoxocation. However, much less of this species is formed than in the case of aluminium. Furthermore, $^{71}$Ga NMR spectroscopy can be used to study the structure and stability of gallium complexes. The clinical use of $^{67}$Ga as a radioactive tracer in tumour scanning and the inhibition of cellular incorporation *in vivo* by citrate, phosphate and other chelating agents served as the impetus for these studies. For example, in order to provide complexes with potentially greater *in vivo* stability with respect to acid-catalysed dissociation, model complexation studies with amino phosphonic acids have been carried out. It was found that the tripodal amino phosphinato ligand, tris[4-(phenylphosphinato)-3-methyl-3-azabutyl]amine ($H_3$ppma), forms stable complexes of octahedral geometry at extremely low pH, a regime in which potentiometry is of no use for studying the solution chemistry. Other methods were sought to determine the stability constants, with NMR spectroscopy proving to be the only suitable tool. The highly symmetric environment around the metal centre in $[Ga(H_3ppma)_2]^{3+}$ produces a low electric field gradient at the gallium nucleus, leading to an unusually narrow line (50 Hz) in the $^{71}$Ga NMR spectrum. The narrow resonances made the complex amenable to stability constant studies via a combination of $^{71}$Ga and $^{31}$P NMR spectroscopies as illustrated in **Figure 3**. The chemical shifts observed for the gallium complexes are indicative of octahedral geometries. The gradual disappearance of the resonance for $[Ga(H_2O)_6]^{3+}$ (Ga) with the concomitant growth of the resonance for $[Ga(H_3ppma)_2]^{3+}$ ($GaL_2$) is evident as $R$ is increased (concentration ratio $R = [H_3ppma]/[Ga^{3+}]$). The $[Ga(H_3ppma)(H_2O)_3]^{3+}$ species (GaL) is barely observable as an insignificant broad resonance and is only seen when $R$ is small. The relative intensities of the gallium resonances correlate with their respective $^{31}$P NMR spectra. The latter consist of three major resonances corresponding to the two diasteroisomers of $GaL_2$ and to the free ligand L, respectively. The formation constant of the $GaL_2$ complex was determined from the integration of the $^{71}$Ga and $^{31}$P NMR spectra.

The +1 oxidation state of gallium represents a relatively unexplored area. There are only a few available data. For example, the $^{71}$Ga chemical shift for benzene solutions of $GaAlX_4$ compounds shows a considerable dependence on the type of $AlX_4^-$ anion and a decreased shielding of about 130 ppm was observed: from −680 ppm ($AlCl_4^-$) and −635 ppm ($AlBr_4^-$) to −552 ppm ($AlI_4^-$). The $^{71}$Ga chemical shifts of $GaAlX_4$ were found to be solvent- and temperature-dependent. The observed line widths for the Ga(I) species are in the range of 80–1200 Hz.

## $^{115}$In NMR spectroscopy

The first NMR measurements of In(III) salts in aqueous solution were reported in the 1950s. However, $^{115}$In has a relatively large quadrupole moment which reduces its observability. $^{115}$In NMR studies of indium complexes in solids or melts of indium metals, alloys and salts have been carried out. Relatively few experimental results are available in

**Table 4** $^{71}$Ga chemical shifts of selected gallium compounds

| *Compound* | $\delta$ $^{71}$Ga | *Compound* | $\delta$ $^{71}$Ga | *Compound* | $\delta$ $^{71}$Ga |
|---|---|---|---|---|---|
| $[Me_3Ga]_2$ | 720 | $Me_3Ga(OEt_2)$ | 560 | $GaH_4^-$ | 682 |
| $[Me_2GaCl]_2$ | 355 | $Me_3Ga(NMe_3)$ | 430 | $Ga(OH)_4^-$ | 192 |
| $[MeGaCl_2]_2$ | 295 | $Me_3Ga(SMe_2)$ | 365 | $GaCl_4^-$ | 250 |
| $[GaCl_3]_2$ | 221 | $[Me_2GaOMe]_2$ | 340 | $GaBr_4^-$ | 69 |
| $[GaBr_3]_2$ | 40 | $GaBr_3(OEt_2)$ | 125 | $GaI_4^-$ | –27 |
| $[Ga(MeCN)_6]^{3+}$ | −76 | $[Ga(dmf)_6]^{3+})$ | −25 | $[Ga(H_2O)_6]^{3+}$ | 0 |

dmf = dimethylformamide.

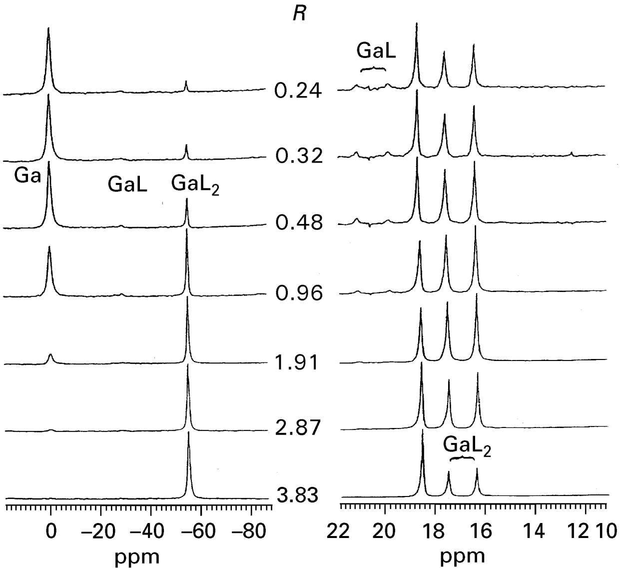



**Figure 3** $^{71}Ga$ (65.7 MHz) (left) and $^{31}P$ (121.0 MHz) (right) NMR spectra for the stability constant study of the Ga(III)–$H_3$ppma system ($R$ = [$H_3$ppma]/[$Ga^{3+}$]). Reprinted with permission of the American Chemical Society from Lowe MP, Rettig SJ and Orvig C (1996) *Journal of the American Chemical Society* **118**: 10446–10456.

solution. The adopted reference of chemical shift is $[In(H_2O)_6]^{3+}$ and the reported line widths even for this symmetrical species are in the range of 1500–2000 Hz. For example, $^{115}$In NMR studies of indium halide complex formation in an aqueous solvent mixture have been carried out. The chemical shifts were used along with the line width to elucidate the chemical equilibria and the exchange processes of indium halide complexes in HCl, HBr and HI aqueous solution and in water–acetone mixtures. The $^{115}$In chemical shifts in the series of four-coordinate $InX_4^-$ ions are 420 ppm ($InCl_4^-$), 180 ppm ($InBr_4^-$) and −440 ppm ($InI_4^-$) and follow a trend similar to that found for the analogous Al and Ga species. Furthermore, $\delta$ $^{115}$In for $[In(H_3ppma)_2]^{3+}$ in $CD_3OD$ solution is −14.6 ppm. The relaxation rates of the $^{115}$In nucleus were obtained for $In(ClO_4)_3$ in aqueous $HClO_4$, i.e. spin–lattice and spin–spin relaxation times of 0.33 ms and 0.24 ms, respectively.

## $^{205}$Tl NMR spectroscopy

$^{205}$Tl is among the most sensitive NMR nuclides and was investigated quite early on. The first liquid- and solid-state spectra for thallium were measured in 1953. The resonance frequency of $^{205}$Tl demanded a single-frequency tuned probe, except for new probe designs at lower fields. Thallium provides significant contrast from the lighter Group 13 elements both in terms of nuclear properties (see **Table 1**) and chemical behaviour. For thallium the +1 oxidation state is normally the more stable, an exception among the Group 13 elements, where the +3 oxidation state dominates. $^{205}$Tl is the fourth most receptive spin-$\frac{1}{2}$ nuclide. The sensitivity of the NMR parameters, such as chemical shift, the spin–spin coupling constant and the relaxation time, to the changes in the chemical environment of the thallium nucleus is more pronounced than is that of the lighter Group 13 nuclei; this is quite reasonable in view of the greater polarizability of the sixth-row elements. Therefore, $^{205}$Tl NMR spectroscopy is quite useful for investigating such phenomena as ion-pairing, solvent–solute interactions, phase changes, long-range substituent effects and binding constants.

The chemical shift ranges are approximately 3400 ppm for Tl(I) and 5000 ppm for Tl(III) compounds. The lower oxidation state $^{205}$Tl is appreciably shielded. The $^{205}$Tl chemical shift is highly sensitive to the anion, solvent, concentration, phase change, substituents and temperature effects. For example, the chemical shift of the $^{205}$Tl(I) ion increased by about 2000 ppm as the solvent changed from water to liquid ammonia, and that of TlCl increased by 925 ppm from the melt to the solid. The anion- and concentration-dependent shifts are in the range of 10–100 ppm. The coupling constants involving the $^{205}$Tl nucleus are usually large and vary significantly with changes in the physical and chemical environments. For example, $^1J_{TlC}$ is nearly 6000 Hz for $CH_3TlX$ salts and $^2J_{TlH}$ ranges from −187 Hz to −1120 Hz for different organothallium(III) compounds. Only a few coupling constants have been observed for Tl(I) compounds due to the highly ionic nature of these complexes.

### Tl(I) species in solution

$^{205}$Tl NMR has been used to determine ion-pair formation constants and the thermodynamic parameters of ion-pairing processes for Tl(I) salts in a number of solvents. The values of ion-pair formation constants have been found to be dependent on the dielectric constant of the solvent. A correlation between the $^{205}$Tl chemical shift and the DN of the solvent has been indicated: the more basic the solvent, the lower the shielding of the $Tl^+$ ion. There have also been numerous studies of the preferential solvation of the $Tl^+$ ion in mixed solvent systems. The relative solvating abilities were analysed in terms of solvent donor–acceptor properties and the solvent DN was found to be most important in the preferential solvation.

An important recent application is the use of $^{205}Tl$ NMR measurements to probe biological processes in which $Tl^+$ replaces alkali cations; the $Tl^+$ ion has an ionic radius of 1.40 Å and in terms of size is close to $K^+$. $^{205}Tl$ NMR has been useful in the study of a wide range of the Tl(I) complexes with synthetic and biological ionophores. The mode of complexation may be indicated by the solvent effect. For example, the chemical shifts of the cryptate complexes of the $Tl^+$ ion are usually independent of solvent, while those of the various Tl(I)–crown complexes are very solvent dependent. The latter reflect the open axial position of the Tl(I)–crown, where the solvent and anion interaction with the complexed ion occurs. A series of the Tl(I)–antibiotics complexes has been characterized by the $^{205}Tl$ chemical shift and the spin–lattice relaxation. The trend in chemical shifts reflected the diversity of the ligand used to bind the $Tl^+$ ion. On the other hand, the magnitude of the chemical shift anisotropy was diagnostic of the symmetry of the $Tl^+$ binding site, and the magnitude of the dipolar contribution to the spin–lattice relaxation indicated the proximity of hydrogen atoms. Hence, $^{205}Tl$ NMR shows great promise as a probe of the nature of metal binding sites.

Much effort has been expended in studying the relaxation behaviour of the $^{205}Tl^+$ ion in solution and has found application in the study of enzyme structure. The spin–lattice relaxation time of the Tl(I) was found to be very sensitive to environmental effects. For example, the $^{205}Tl^+$ spin–lattice longitudinal relaxation times $T_1$ were found to vary by over one order of magnitude for the solvent series, ranging from 1.8 s in water to 0.08 s in *n*-butylamine. The relaxation time of the Tl(I) ion aqueous solution was independent of concentration, and spin–rotation was found to be the dominant relaxation mechanism. In nonaqueous solution ion pairing becomes important and the longitudinal relaxation time of the $Tl^+$ ion is dominated by the chemical shift anisotropy mechanism. Furthermore, in the case of $Tl^+$ complexes with antibiotics, the longitudinal relaxation time is determined by the contribution from spin rotation, dipolar, and chemical anisotropy shift mechanisms. A strong influence of dissolved dioxygen on the $^{205}Tl^+$ relaxation rate has also been found. Therefore, it is advisable to use a degassed solution before attempting relaxation measurements.

### Organothallium(III) and miscellaneous compounds

The $^{205}Tl$ chemical shifts of organothallium(III) compounds cover a range of over 2000 ppm and they display solvent dependence, although not as striking as that for Tl(I). Solvent basicity and ion pairing contribute to the solvent dependence. The dependence of the $^{205}Tl$ chemical shift of $Me_2TlX$ salts upon solvent, concentration, X anion and temperature has been studied. The latter factor was found to affect the chemical shifts the most. Substituent changes produce large chemical shifts, for example the difference in chemical shift between $Me_3Tl$ and $Et_3Tl$ is 130 ppm. The relaxation behaviour of $^{205}Tl(III)$ has been investigated for $Me_2Tl^+$ in solution as a function of temperature and magnetic field strength and the dominant relaxation mechanism was found to be chemical shift anisotropy.

## List of symbols

$I$ = spin; $J$ = coupling constant; $P_u$ = elements of the bond-order charge-density matrix (expresses population of the *np* orbitals and a *p*-electron 'imbalance' about the nucleus in question); $r$ = radius of the valence *p* orbitals ($r^{-3}$ is usually called the orbital expansion term); $\delta$ = chemical shift; $\Delta E$ = average electronic excitation energy; $\sigma$ = nuclear screening constant; $\sigma_{para}$ and $\sigma_{dia}$ = local paramagnetic and diamagnetic, respectively, contributions to the nuclear screening constant.

*See also:* **NMR Relaxation Rates; NMR Spectrometers; Parameters in NMR Spectroscopy, Theory of; Relaxometers; Solid State NMR Using Quadrupolar Nuclei; Solid State NMR, Methods.**

## Further reading

Akitt JW (1989) Multinuclear studies of aluminium compounds. *Progress in NMR spectroscopy* **21**: 1–149.

Dechter JJ (1982) NMR of metal nuclides. Part I. The main group metals. *Progress in Inorganic Chemistry* **29**: 351–369.

Harris RK and Mann BE (1978) Group III – aluminium, gallium, indium and thallium. In: *NMR and the Periodic Table*, Chapter 9A, pp 279–308. New York: Academic Press.

Hermanek S (1992) $^{11}B$ NMR spectra of boranes, main-group heteroboranes, and substituted derivatives. Factors influencing chemical shifts of skeletal atoms. *Chemical Reviews* **92**: 325–362.

Hinton JF, Metz KR and Briggs RW (1988) Thallium NMR spectroscopy. *Progress in NMR Spectroscopy* **20**: 497–513.

Nöth H and Wrackmeyer B (1978) Nuclear magnetic resonance spectroscopy of boron compounds. *NMR Basic Principles and Progress* **14**: 1–461.

Siedle AR (1988) $^{11}B$ NMR spectroscopy. *Annual Reports on NMR Spectroscopy* **20**: 205–314.

Wrackmeyer B (1988) Nuclear magnetic resonance spectroscopy of boron compounds containing two-, three- and four-coordinate boron. *Annual Reports on NMR Spectroscopy* **20**: 61–203.

# Heteronuclear NMR Applications (Ge, Sn, Pb)

**Claudio Pettinari**, Università di Camerino, Italy

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

Group 14 of the periodic table is surely one of the most interesting for the NMR spectroscopist. With the exception of germanium, all the other members have favourable NMR properties: tin, silicon and lead have at least one spin $\frac{1}{2}$ isotope of good sensitivity.

Silicon and tin have been studied extensively, and several applications have been found in chemistry. Only few reports are available on germanium, which has only one magnetic isotope, $^{73}$Ge, characterized by a relatively large quadrupole moment; this has limited the number of compounds studied, and only a little more data is available on lead. There is a sizeable literature on tin, beginning with the first investigation in 1961. For example, there has been extensive compilation of solution and solid-state $^{119}$Sn chemical shifts and coupling constants, mainly in organometallic chemistry. These data are often useful in the determination of molecular structures, not only of small molecules but also of polymeric and cluster compounds. In particular, solution data are the main tools for gaining information on the progress of a selected reaction, on the existence of donor–acceptors and/or solute–solvent interactions, of dissociation or autoassociation, and of intramolecular coordination.

Moreover in the last few years, one-dimensional (1D) and two-dimensional (2D) proton detected heteronuclear correlation NMR techniques involving $^{119}$Sn and $^{207}$Pb nuclei have been used for applications in organometallic NMR. In particular, the development of organometallic chemistry has been affected markedly by the ever-increasing capabilities of NMR instruments. Prominent NMR parameters such as chemical shifts, coupling constants and relaxation times have also been exploited to gain information on organo-E (E = Ge, Sn, Pb) compounds not readily available by other methods.

## Properties of the nuclei

The nuclear properties are summarized in **Table 1**.

Germanium possesses only one magnetic isotope, $^{73}$Ge, with spin $\frac{9}{2}$. This nucleus shows a combination of properties unfavourable to magnetic resonance studies, so that only a limited range of molecules can be studied: the natural abundance is 7.76%, the gyromagnetic ratio is very small ($\gamma = -0.9332 \times 10^7$ rad $T^{-1}$ $s^{-1}$), and the nuclear quadrupole is moderately large. The sensitivity with respect to $^1$H at natural abundance and constant field is $1.08 \times 10^{-4}$. Since quadrupolar relaxation dominates, dipole–dipole relaxation and nuclear Overhauser effects (NOE) are not important and unsymmetric germanium compounds often give very broad $^{73}$Ge NMR signals.

Tin possesses three magnetically active tin isotopes with spin $I = \frac{1}{2}$ ($^{115}$Sn, $^{117}$Sn, $^{119}$Sn). Two of them, $^{117}$Sn (7.51%) and $^{119}$Sn (8.58%), have an appreciable natural abundance and receptivities ~20 and 25 times greater, respectively, than that of $^{13}$C. The low natural abundance makes the $^{115}$Sn resonance unfavourable. The basic resonance frequencies for $^{115}$Sn, $^{117}$Sn, and $^{119}$Sn are 98.16, 106.90, and 111.87 MHz on a spectrometer that operates at a field strength of 7.05 T with $^1$H NMR at 300 MHz.

In tin NMR experiments, the negative gyromagnetic ratio can lead to disadvantageous NOE behaviour, which can be suppressed by inverse gated decoupling or by using relaxation rate enhancement additives. This problem is not present in the NMR of lead compounds: in fact $^{207}$Pb, the only lead isotope

**Table 1** NMR properties[a] of magnetic isotopes of germanium, tin and lead

| *Nucleus* | *Spin* | *NA*(%) | $R^H$ | $R^C$ | $\gamma$ ($10^7$ rad $T^{-1}$ $s^{-1}$) | $\Xi$ (MHz) | *Q* ($10^{-28}$ $m^2$) |
|---|---|---|---|---|---|---|---|
| $^{73}$Ge | $\frac{9}{2}$ | 7.76 | $1.08 \times 10^{-4}$ | 0.622 | –0.9332 | 3.488 315 | –0.17 |
| $^{115}$Sn | $\frac{1}{2}$ | 0.35 | $1.24 \times 10^{-4}$ | 0.705 | –8.792 | 32.718 780 | – |
| $^{117}$Sn | $\frac{1}{2}$ | 7.51 | $3.49 \times 10^{-3}$ | 19.8 | –9.578 | 35.632 295 | – |
| $^{119}$Sn | $\frac{1}{2}$ | 8.58 | $4.51 \times 10^{-3}$ | 25.6 | –10.021 | 37.290 662 | – |
| $^{207}$Pb | $\frac{1}{2}$ | 22.6[b] | $2.1 \times 10^{-3}$ | 11.9 | 5.6264 | 20.920 597 | – |

[a] NA is the natural abundance (%); $R^H$ and $R^C$ are the receptivities relative to that of $^1$H and $^{13}$C, respectively; $\gamma$ is the gyromagnetic ratio; $\Xi$ is the resonance frequency for the reference compound $(Me)_4$E (E = Ge, Sn or Pb) in a magnetic field for which $(Me)_4$Si has a proton resonance of 100 MHz; *Q* is the electric quadrupole moment.

[b] Varies with sources.

with spin $\frac{1}{2}$ (the natural abundance is ~22% and the sensitivity to detection is 0.09% of that of the proton), has a positive gyromagnetic ratio. The maximum $^{207}$Pb-{$^1$H} nuclear Overhauser effect is 350%.

## Experimental aspects

Samples for solution studies are prepared in the same manner as for $^1$H NMR spectroscopy. In most of the experiments, the standard reference compound is the tetramethyl-E (E = Ge, Sn, Pb) derivative, although sometimes $GeCl_4$ and $SnCl_4$ are used as references in $^{73}$Ge and $^{119}$Sn NMR spectra, respectively. The tetramethyl-E compound can be used externally; however, on modern NMR spectrometers it is generally preferable to use a known absolute frequency.

There are great difficulties in determining the chemical shifts of $^{73}$Ge nuclei by direct observation owing to its low natural abundance and receptivity and large quadrupole moment. Coupling to $^{73}$Ge is rarely reported; however, if appreciable $^1$H–$^{73}$Ge coupling is present, the observation of $^{73}$Ge resonances can be achieved by heteronuclear double-resonance experiments. Special pulse sequences such as INEPT (insensitive nuclei enhancement by polarization transfer) which effect nuclear spin polarization transfer from nuclei of high gyromagnetic ratio ($\gamma$), usually $^1$H, to nuclei of small $\gamma$, and allow dramatic increases in signal-to-noise ratio and corresponding decreases of spectral accumulation time, have been used in the study of simple symmetrical molecules.

The $^{119}$Sn nucleus is generally preferred in tin NMR experiments because of its greater natural receptivity. $^{119}$Sn chemical information is obtained from 1D $^{119}$Sn NMR and $^nJ(^{119}$Sn, $^1$H), $^nJ(^{119}$Sn, $^{13}$C) and $^nJ$ ($^{119}$Sn, $^{117}$/$^{119}$Sn) coupling, or by using the $^1$H-{$^{119}$Sn} heteronuclear double magnetic resonance (HDMR) method, polarization transfer techniques such as INEPT and DEPT (distortionless enhancement polarization transfer), or proton-detected heteronuclear $^1$H-{$^{119}$Sn} correlation techniques such as HQMC (heteronuclear multiple-quantum coherence). HQMC NMR spectroscopy can be easily applied to the structure elucidation of organotin compounds, while in those containing at least a single Sn–N bond it is possible to measure very efficiently the $J(^{119}$Sn, $^{15}$N) coupling constant from $^{15}$N satellites in $^{119}$Sn NMR spectra by application of the Hahn-echo extended (HEED) INEPT experiment.

The magnitude and sign of $^nJ(^{119}$Sn,X) and $^nJ(^{207}$Pb,X) coupling constants are fundamental requisites for interpreting the bonding and structure in tin and lead compounds. Relative signs of the coupling constants can be obtained efficiently from 2D heteronuclear shift correlation (HETCOR) experiments. This method is helpful also in determining small long-range coupling constants (e.g. $^4J(^{117/119}$Sn,$^1$H)), which are normally not resolved in 1D $^1$H NMR spectra.

In order to obtain highly resolved, selective solid-state information, in recent years much work has been done on solid-state $^{119}$Sn and $^{207}$Pb NMR using magic angle spinning (MAS) and cross-polarization (CP) methods.

## Chemical shifts

In agreement with Ramsey's theory, the variation of the 'paramagnetic term' $\sigma_p$ is the dominating factor in the nuclear screening (Eqn [1]) for all three elements, i.e. germanium, tin and lead.

$$\sigma = \sigma_p + \sigma_d + \sigma_n \qquad [1]$$

In Equation [1] $\sigma_d$ is the diamagnetic contribution to shielding; $\sigma_n$ includes all diamagnetic and paramagnetic nonlocal contributions; and $\sigma_p$ is a function of at least three factors: the effective nuclear charge, the p and d electron imbalance term, and the average excitation energies; $\sigma_p$ and $\sigma_d$ may be calculated in principle for a given system, but the theory has not yet been developed, so that this model may be applied with success only when calculating the difference in shielding in closely related compounds, whereas for heavy-atom substituents $\sigma_p$ does not reflect the influence of nearest neighbours.

Germanium, tin and lead chemical shifts cover a very large range if inorganic and organometallic compounds are examined: more than 1200 ppm for $^{73}$Ge, 5000 ppm for $^{119}$Sn, and 12 000 ppm for $^{207}$Pb. **Tables 2**, **3** and **4** exhibit selected $^{73}$Ge, $^{119}$Sn and $^{207}$Pb data, all referred to $(Me)_4E$ (E = Ge, Sn or Pb), a negative sign indicating a higher shielding or a shift to lower frequencies.

### Isotope effects

In the case of tin, the primary isotope effect (dependence of the screening upon the type of nuclide of tin investigated) should be considered. The chemical shift difference [$\delta(^{119}$Sn) – $\delta(^{117}$Sn)] has been measured for a number of organotin compounds in noncoordinating solvents, and in most cases the primary isotope effect is negligibly low. Secondary isotope effects are more important: for example, $^{119}$Sn chemical shifts have been measured for $Me_2SnH_2$, $Me_2SnHD$ and $Me_2SnD_2$ and a shift of ~1 ppm to lower frequencies has been observed when one H is

**Table 2** Selected $^{73}$Ge NMR chemical shifts in ppm relative to $Me_4Ge$

| *Compound* | $\delta$(ppm) | *Compound* | $\delta$(ppm) |
|---|---|---|---|
| $Me_4Ge$ | 0 | $Ge(CH{=}CH_2)_4$ | −58.7 |
| $Et_4Ge$ | 17.3 | $Me_3Ge(CH{=}CH_2)$ | −15.6 |
| $Ph_4Ge$ | −31.6 | $Me_2Ge(CH{=}CH_2)_2$ | −30.6 |
| $GeH_4$ | −283.7 | $MeGe(CH{=}CH_2)_3$ | −44.9 |
| $EtGeH_3$ | −186.4 | $Ge(C{\equiv}CH)_4$ | −173 |
| $Et_2GeH_2$ | −88.4 | $MeGe(C{\equiv}CH)_3$ | −118 |
| $Et_3GeH$ | −15.7 | $Me_2Ge(C{\equiv}CH)_2$ | −77 |
| $GeCl_4$ | 30.9 | $Me_3Ge(C{\equiv}CH)$ | −34 |
| $GeBr_4$ | −311.3 | $Ge(OMe)_4$ | −37.8 |
| $GeI_4$ | −1081.8 | $Ge(OEt)_4$ | −43.9 |
| $GeCl_3Br$ | −47.8 | $Ge(OPr)_4$ | −49.7 |
| $GeCl_2Br_2$ | −131.3 | $Ge(OCH_2CF_3)_4$ | −48.1 |
| $GeClBr_3$ | −219.4 | Germacyclohexane | −131.2 |
| $GeCl_4$(Bipy) | −313.7 | 1-Me-Germacyclohexane | −65.3 |
| $GeCl_4$(phen) | −319.4 | 1-Ph-Germacyclohexane | −69.4 |
| $[Ge(NCS)_6]^{2-}$ | −442.5 | $Ge(NCS)_4$(Bipy) | −351.8 |

replaced by one D. The largest isotopic effect has been observed with the $[SnH_3]^-$ anion, where the $^{119}$Sn chemical shift moves by 9.84 ppm on perdeuteration. The isotopic effects in the case of germanium and lead compounds have not been widely studied: however, it has been found that they are generally greater for $^{207}$Pb and less for $^{73}$Ge with respect to $^{119}$Sn.

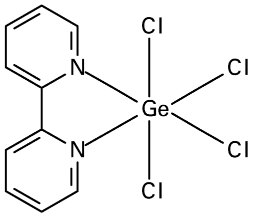
### Solvent and temperature dependence of tin, germanium and lead chemical shifts

Tin, germanium and lead chemical shifts are not very sensitive to temperature in the absence of chemical changes (e.g. autoassociation, dissociation, change in coordination number of the atom of the nucleus examined). Sometimes variations in conformational populations and electrostatic solvent–solute interactions can produce appreciable effects, and it has been found, for example, that shielding of the $^{119}$Sn nucleus generally increases roughly linearly with increasing temperature. Analogously, change in solvent, in the absence of solvent–solute chemical interactions, autoassociation or dissociation affects the tin, germanium and lead chemical shifts only slightly with respect to their chemical shift ranges.

### Coordination number dependence of tin, germanium and lead chemical shifts

Taking into account that relevant excitation energies are fairly constant, the paramagnetic term $\sigma_p$, which represents the imbalance of charge, could be related to local symmetry. It has been established that an increase in coordination number produces a significant increase in the shielding of Sn(IV) and Pb(IV).

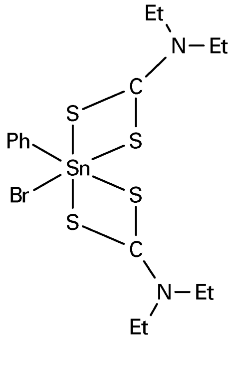
Germanium(IV) compounds have been less studied in this regard; however, they seem to show the same behaviour. For example the $GeCl_4$-2,2′-bipyridine adduct [1] containing a hexacoordinate germanium has $\delta(^{73}Ge)$ ~340 ppm upfield shifted with respect to the four-coordinate compound $GeCl_4$.



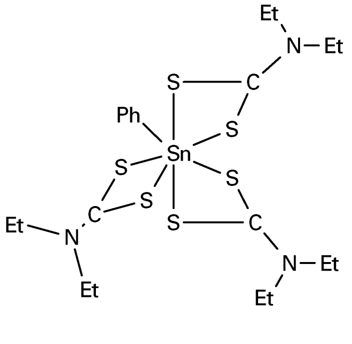

[1] $\delta$ ($^{73}$Ge) : −313.7 ppm

The knowledge of $\delta(^{119}Sn)$ chemical shifts is very important for predicting the geometry of coordination polyhedra of tin compounds in solution. In the case of tetravalent tin compounds, an increase in tin coordination number from four to five corresponds to an increase in tin shielding by ~150–200 ppm. Similar increases are generally associated with six and seven coordination. The $\delta(^{119}Sn)$ for $PhSn(edtc)_2Br$ [2], a hexacoordinate pseudooctahedral tin complex, is −704.1 ppm, whereas $\delta(^{119}Sn)$ for $PhSn(edtc)_3$ [3] in which the central Sn has coordination number seven and the shape is distorted pentagonal bipyramidal is −806.5 ppm.

The $^{119}$Sn chemical shift also provides effective means of studying covalent solute–solvent interaction, autoassociation, dissociation, and intramolecular coordination. For example, the $^{119}$Sn chemical shift is nearly concentration independent when tin(IV) and organotin compounds such as $Me_2SnCl_2$



[2] $\delta$ ($^{119}$Sn) : −704.1 ppm



[3] $\delta$ ($^{119}$Sn) : −806.5 ppm

**Table 3** Selected $^{119}Sn$ NMR chemical shifts in ppm relative to $Me_4Sn$

| *Compound* | $\delta$(ppm) | *Notes* | *Compound* | $\delta$(ppm) | *Notes* |
|---|---|---|---|---|---|
| $Me_4Sn$ | 0 | Neat liquid | $Me_3SnCl$ | +168.9 | 5% in $CH_2Cl_2$ |
| $(CD_3)_4Sn$ | +2.6 | Neat liquid | $Me_3SnBr$ | +128.0 | Benzene |
| $Et_4Sn$ | +1.4 | 20% in $CCl_4$ | $Me_3SnI$ | + 38.6 | 3–20% toluene |
| $Pr^i_4Sn$ | −43.9 | Neat liquid | $Ph_3SnCl$ | − 44.7 | $CDCl_3$ |
| $Ph_4Sn$ | –137 ± 2 | 30% in $C_2H_3Cl_3$ | $Ph_2SnCl_2$ | −32 | $CH_2Cl_2$ |
| $Me_3SnEt$ | +3.0 | $CCl_4$ | $Me_2SnCl_2$ | +137 | 30% in $CH_2Cl_2$ |
| $(Me_3Sn)_4C$ | +49.8 | $C_6D_6$ | $(Me_3Sn)_3P$ | +36.3 | 65% in benzene |
| $Me_3SnPh$ | −28.6 | 20% $CH_2Cl_2$ | $[(Cl_3Sn)_5Pt]_3^-$ | −142.0 | Acetone-$d_6$ |
| $Me_3Sn(C{\equiv}CH)$ | −68.1 | $CH_2Cl_2$ | $Ph_3SnOH$ | −86 | $CH_2Cl_2$ |
| $Me_3SnC{\equiv}CH$-Ph | −69.0 | $CDCl_3$ | $Me_3SnOH$ | +11.8 | Saturated $CH_2Cl_2$ |
| $Et_3SnCH{=}CH_2$ | −42.0 | 50% $CCl_4$ | $MeSnH_3$ | −346 | 60% in toluene |
| $Me_3SnPh$ | −30 | Neat liquid | $Me_2SnH_2$ | −229 | Neat liquid |
| $Me_2SnPh_2$ | −60 | Neat liquid | $Me_3SnH$ | −104.5 | Benzene |
| $MeSnPh_3$ | −98 | $CH_2Cl_2$ | $Ph_3SnH$ | −148 | Neat liquid |
| $SnCl_4$ | −150 | Neat liquid | $MeSnCl_3 \cdot 2DMSO$ | −457 | Saturated DMSO |
| $SnBr_4$ | −638 | Neat liquid | $Me_2SnCl_2 \cdot 2DMSO$ | −246 | Saturated DMSO |
| $SnI_4$ | −1701 | $CS_2$ | $Me_3SnCl \cdot DMSO$ | −86 | Saturated $CH_2Cl_2$ |
| $SnCl_3Br$ | −265 | 1:1 M $SnBr_4/SnCl_4$ | $SnCl_2$ | −388 | In 12 M HCl |
| $SnCl_2Br_2$ | −387 | 1:1 M $SnBr_4/SnCl_4$ | $Cp_2Sn$ | −2199 ± 10 | $C_6H_{12}$ |
| $SnClBr_3$ | −509 | 1:1 M $SnBr_4/SnCl_4$ | $Me_3SnOMe$ | +120.9 | Benzene |
| $MeSnCl_3$ | +18.7 | Saturated benzene | $Me_2Sn(OMe)_2$ | −126.3 | Benzene |
| $PhSnCl_3$ | −64 | $CDCl_3$ | $MeSn(OBu^i)_3$ | −452.0 | 50% in mesitylene |
| $(Me_3Sn)_2NMe$ | +81 | 20% $C_6D_6$ | $Me_3SnNMe_2$ | +75.5 | 10% in benzene |

**Table 4** Selected $^{207}Pb$ NMR chemical shifts in ppm relative to $Me_4Pb$

| *Compound* | $\delta$(ppm) | *Notes* | *Compound* | $\delta$(ppm) | *Notes* |
|---|---|---|---|---|---|
| $Me_4Pb$ | 0 | Containing 15% toluene | $Me_3PbCl$ | +432 | Saturated in $CH_2Cl_2$ |
| $Et_4Pb$ | +73.3 | Neat liquid | $Me_3PbBr$ | +367 | Saturated in $CH_2Cl_2$ |
| $Vi_4Pb$ | −251 | Neat liquid | $Me_3PbI$ | +203.6 | 10% in $CH_2Cl_2$ |
| $Me_3Pb(NMePh)$ | +226 | 10% in $C_6D_6$ | $Me_3ViPb$ | −65.3 | Neat liquid |
| $Me_3Pb(SnMe_3)$ | −324 | Saturated in $C_6D_6$ | $Me_3(SMe)Pb$ | +214 | Neat liquid |
| $Ph_3PbGePh_3$ | −271.5 | $CDCl_3$ | $Ph_3PbSnPh_3$ | −256.5 | $CDCl_3$ |
| $Me_3(Pr^i)Pb$ | 62 | Saturated $CDCl_3$ | $Me_2PbCl_2$ | −222 | DMSO |
| $Et_3Pb(O_2CMe)$ | 317 | $CDCl_3$ | $(CH_2CMe_3)_3Pb$ | +336 | Saturated $CH_2Cl_2$ |
| $Et_2Pb(O_2CMe)_2$ | −441 | $CDCl_3$ | $[Pb(CH_2CMe_3)_3]_2$ | −237 | Saturated $CH_2Cl_2$ |
| $Pb(O_2CMe)_4$ | −1869 | – | $Ph_3Pb(O_2CMe)$ | −93 | – |
| $PbCl_2$ | −4750 | Solid | $Pb(NO_3)_2$ | −2961.2 | 1.0 M in $H_2O$ |
| $PbO_2$ | +4550 ± 30 | Powder, 98% pure | $Pb(ClO_4)_2$ | −2950 | In $H_2O$ |
| $Pb_9^{4-}$ | −4152 | Naked anion cluster | $Me_3PbPbMe_3$ | −281 | Saturated in $C_6H_6$ |
| $Pb_2Ph_6$ | −79.8 | Saturated $CDCl_3$ | $Pb(2\text{-naphthyl})_6$ | −74.6 | $CDCl_3$ |

Vi = vinyl.

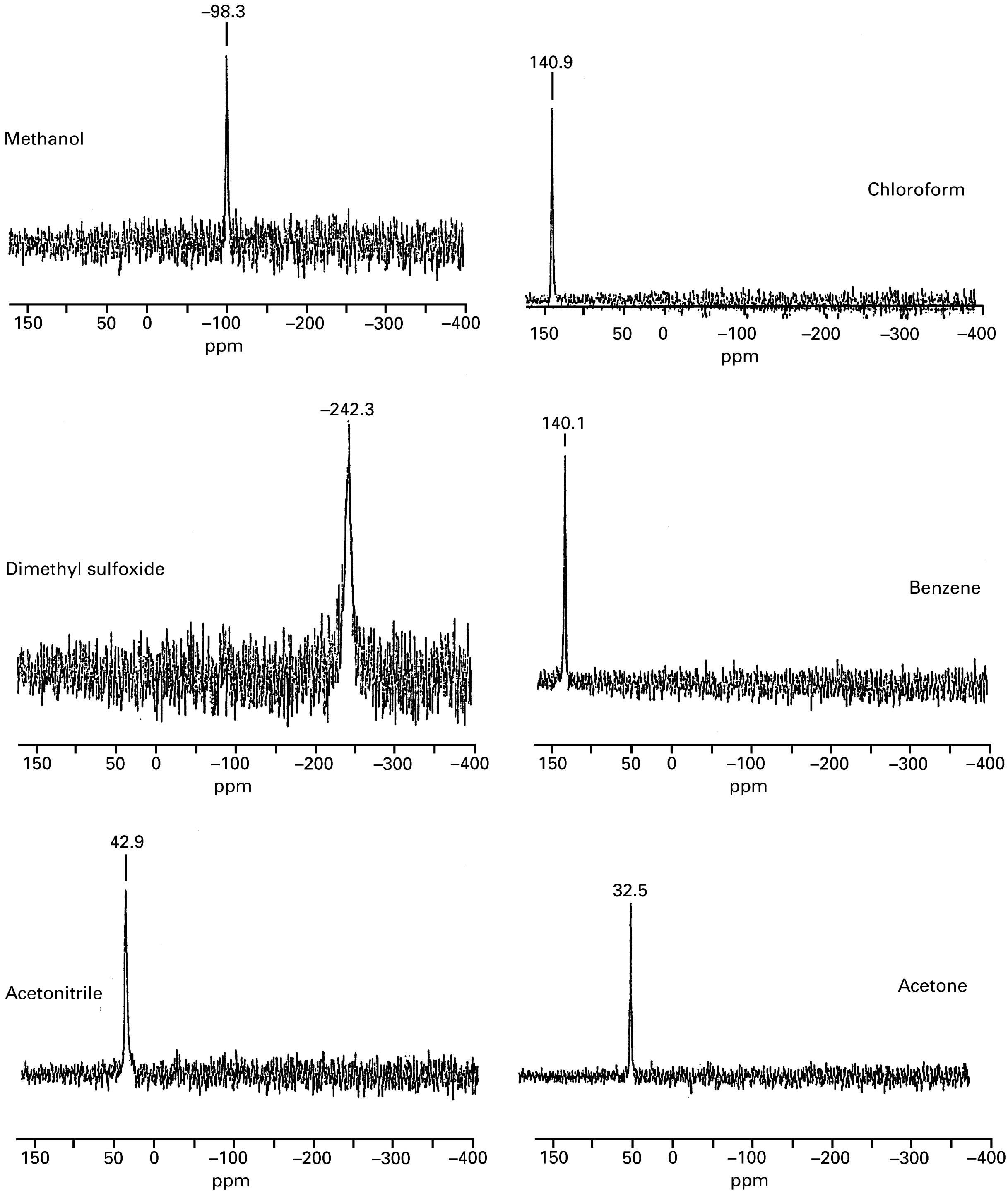



**Figure 1** $^{119}$Sn NMR spectra of $Me_2SnCl_2$ in chloroform, benzene, acetone, methanol, dimethyl sulfoxide and acetonitrile (concentration ~ 40 mg $mL^{-1}$).

are dissolved in noncoordinating solvents such as chloroform or benzene (**Figure 1**). On the other hand, when $Me_2SnCl_2$ is dissolved in various donor solvents such as acetonitrile, pyridine and dimethyl sulfoxide, the chemical shift shows dependence on the concentration.

The magnitude of the chemical shift can permit determination of association constants between the tin compound and the solvent. For example, formation of a 1:1 complex of trimethyltin chloride in dimethyl sulfoxide may be written as in Equation [2].

$$\mathrm{Bu^n_3SnCl + DMSO \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Bu^n_3Sn(DMSO)Cl} \quad [2]$$

The equilibrium constant can be calculated from the concentration dependence found for $^{119}Sn$ chemical shifts (Eqn [3]):

$$K = \frac{k_1}{k_2} = \frac{\delta_o(\delta_c - C\,\delta_o)}{\delta_o{}^2(1-C) + C\,\delta_o{}^2 - \delta_o\delta_c} \quad [3]$$

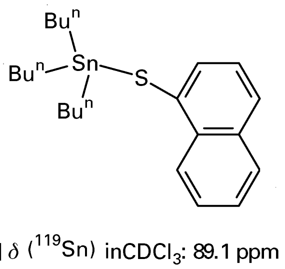
where $\delta_o$ is the $^{119}Sn$ chemical shift for $Bu^n{}_3SnCl$, $\delta_c$ is the $^{119}Sn$ chemical shift for the $Bu^n{}_3Sn(DMSO)Cl$, and $C$ is the molar fraction of $Bu^n{}_3SnCl$.

This type of behaviour is also displayed by lead and organolead compounds; the changes in $\delta(^{207}Pb)$ are generally larger than changes in $\delta(^{119}Sn)$: for example, the $^{207}Pb$ chemical shift for derivative $Pb(Ph)_2(O_2CMe)_2$ is –688 ppm in $CHCl_3$ and –859 ppm in pyridine.

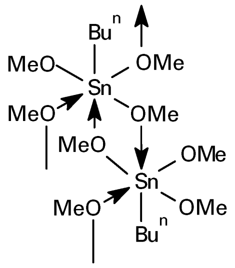
Organotin(IV) thiolates show a low tendency to form complexes with molecules of coordinating solvent. However, it has been observed that although the $\delta(^{119}Sn)$ for the organotin(IV) thiolate compound $Bu^n{}_3Sn[S(1\text{-}C_{10}H_7)]$ [4] is concentration independent in noncoordinating solvent, the variation of $\delta(^{119}Sn)$ with temperature in pyridine(Py) is at first a sigmoidal curve which is typical of the equilibrium (Eqn [4]).

$$\mathrm{Bu^n_3Sn[S(1\text{-}C_{10}H_7)] + Py \rightleftharpoons Bu^n_3\,Sn[S(1\text{-}C_{10}H_7)](Py)} \quad [4]$$

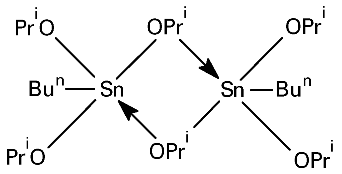
In several organometal derivatives, changes in the coordination number can arise from self-association, as in the case of organotin alkoxides. This self-association depends on steric and electronic effects and it can be studied with $^{119}Sn$ NMR. Several $R_{4-n}Sn(OR')_n$ compounds exhibit values of chemical shifts suggesting self-association as neat liquid or in noncoordinating solvents. For example, butyltin(IV) trialkoxides with an unbranched alkoxy group (such as $BuSn(OMe)_3$, [5], are six-coordinate at room temperature, the association presumably involving octahedral structures with both bridging and nonbridging alkoxy oxygen atom. Isopropoxy groups on tin, for example in [6], increase the steric effect and inhibit six-coordination. At room temperature, the chemical shift is typical of an essentially five-coordinate situation.



[4] $\delta$ ($^{119}Sn$) in$CDCl_3$: 89.1 ppm



[5]



[6]

## Effective nuclear charge and presence of low-lying excited states

Electronegative substituents generally lead to higher-frequency germanium, tin and lead chemical shifts. The effect of ligand electronegativity can be illustrated by the changes of $^{119}Sn$ chemical shifts along the series $R_nSnCl_{4-n}$ (R = alkyl or aryl) as the value of the $n$ changes from 2 to 0 (**Figure 2A**). However, the plot of the $\delta(^{119}Sn)$ for the compounds $R_nSnCl_{4-n}$ (R = alkyl or aryl; $n$ = 0, 1, 2, 3 or 4) in **Figure 2B**, indicates clearly that there are no obvious general relationships between $\delta(^{119}Sn)$ and the ligand electronegativity. The U-shape trend is usually attributed to change in the paramagnetic term due to p electron imbalance, generally maximum at $n = 2$.

The $^{73}Ge$ chemical shift of $GeCl_4$, $GeBr_4$, $GeI_4$ and the mixed tetrahalogermanes is not a linear function of halogen coordination but is consistent with the pattern of nuclear shielding established for halogens bonded to main-group elements: as expected, the nuclear shielding increases in the order Cl < Br < I.

The $^{119}Sn$ and $^{207}Pb$ chemical shifts follow the same trend shown by $^{73}Ge$, but the larger lead and tin shift range allows a clearer distinction between closely related species (**Figure 3**).

In a series of methylvinylgermanes, the $^{73}Ge$ chemical shifts were compared with those of the corresponding methylvinylsilanes, and the correlation $\delta(^{73}Ge) = 1.96\ \delta(^{29}Si) + 2.37$ was found. It has been observed that the magnitude of the chemical shift cannot be simply attributed to the electron densities, but relative contributions of the diamagnetic and paramagnetic terms to the $^{73}Ge$ chemical shift should be considered.

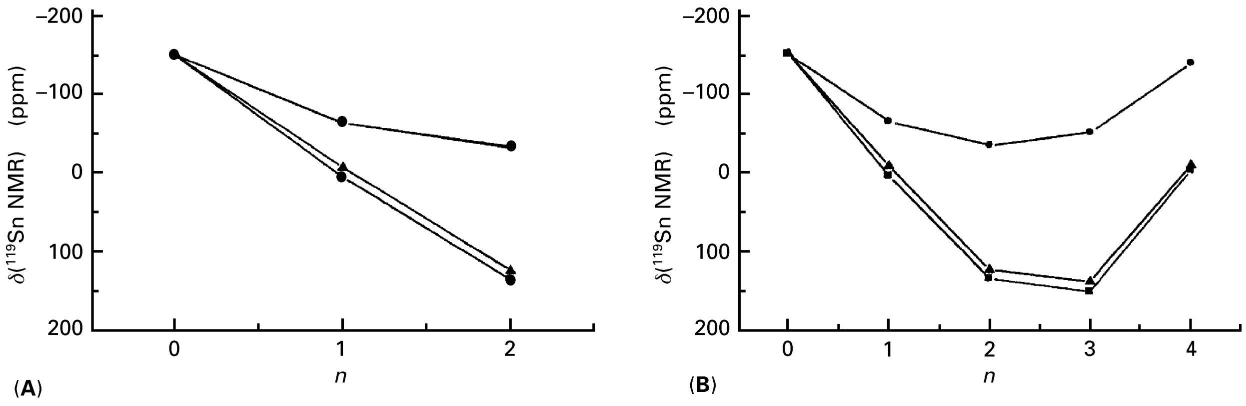



**Figure 2** Plot of δ($^{119}$Sn) against *n* for a series of monomeric tin(IV) compounds $R_nSnCl_{4-n}$. ■, Me; ●, Ph; ▲, Bu.

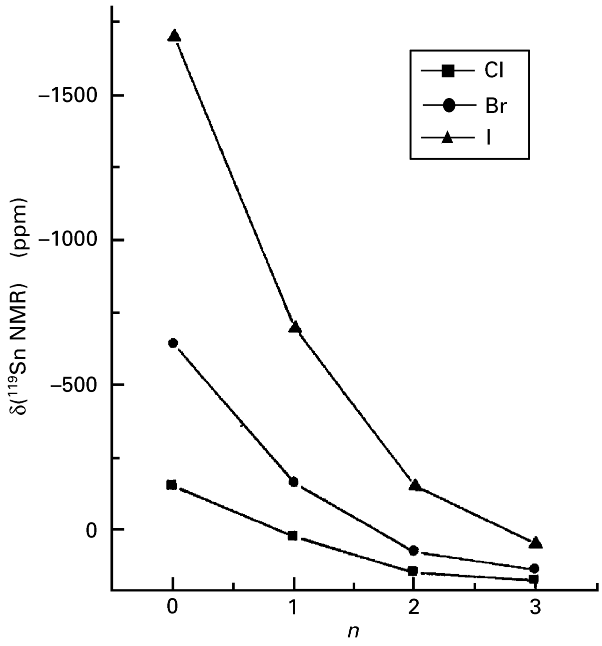
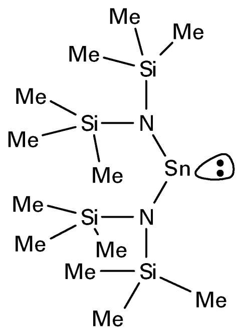
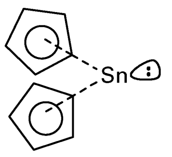



**Figure 3** Plot of δ($^{119}$Sn) against *n* for a series of monomeric methyltin(IV) halides. $Me_nSnX_{4-n}$. ■, Cl; ●, Br; — ▲, I.




[7] δ ($^{119}$Sn) : +776 ppm




[8] δ ($^{119}$Sn) : −2199 ppm

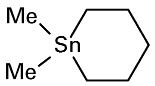
From a comparison between ethylvinylgermanes and ethynylstannanes, a similar linear correlation $\delta(^{73}Ge) = 0.481\ \delta(^{119}Sn) - 0.892$ was obtained.

Several studies have been made to evaluate the effects of the paramagnetic circulation of charge: in the monomeric bis[*N*,*N*-bis(trimethylsilyl)amino]-tin(II) compound [7], it has been found that circulation of charge from the trigonal plane into the underoccupied tin $p_\pi$ orbitals deshields the tin atoms and produces a strong high-frequency shift (+776 ppm in $C_6D_6$) of the $^{119}$Sn resonance. On the other hand, MO calculations indicate in monomeric tin(II) compounds such as stannocene, $(C_5H_5)_2Sn$ [8], a highly shielded tin atom (−2199 ppm in $C_6H_{12}$) as a consequence of an inefficient paramagnetic charge circulation due to the high LUMO energy of this molecule.

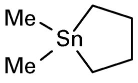
When tin is bonded to transition metals, large $^{119}$Sn chemical shift ranges have been found that cannot be explained in terms of simple inductive and coordination number effects. The shielding could be interpreted in terms of $d_\pi$–$d_\pi$ overlap and the effect on the average excitation energy.

### Effects of interbond angles at the metal atoms

Changes in the bond angles in molecules in which the metal atom is a part of a ring system generally effect anomalous shifts of the resonances with respect to those found in analogous acyclic molecules. This is shown, for example, with tin–sulfur and lead–sulfur and with tin–carbon and lead–carbon bonds. A reduction of the cyclic interbond angles at the metal in five-membered rings generally decreases shielding. $^{119}$Sn NMR data can be used to distinguish between six-membered and five-membered cyclic compounds: the δ($^{119}$Sn) for compound [9] is −42.5 ppm, whereas that for the five-membered derivative [10] it is +53.5 ppm.




[9] $\delta$ ($^{119}$Sn) : –42.5 ppm




[10] $\delta$ ($^{119}$Sn) : +53.5 ppm

### Other substituent effects

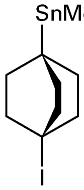
In the unassociated tin compounds $R_{4-n}SnX_n$ ($n$ = 0–4), as the electron–releasing power of the organic group R increases, the tin atom is generally more shielded and the $\delta(^{119}Sn)$ moves to a higher field. This high-field shift, which cannot be explained in terms of $p_\pi$–$d_\pi$ bonding between the $\pi$ electron of the electron negative groups and the unfilled 5d orbital of the tin atom, is probably due to ring current effects by the $\pi$ electron and/or a finite contribution to the shift from the diamagnetic term.

The values of the $\delta(^{119}Sn)$ chemical shift of diphenyltin(IV), dibenzyltin(IV) and di-*n*-butyltin(IV) compounds exhibit a good mutual correlation which can be expressed as in Equation [5].

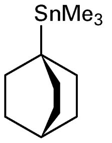
$$\delta(^{119}\mathrm{SN})[\mathrm{R}_2^* \ \mathrm{Sn}] = a\delta(^{119}\mathrm{Sn})[\mathrm{R}_2^\# \ \mathrm{Sn}] + b \qquad [5]$$

where $[R_2^*Sn]$ and $[R_2^\#Sn]$ represent pairs of diorganotin(IV) compounds with different organic substituents, and *a* and *b* represent the slope and the intercepts of the linear correlation, respectively.

In some cases it has also been observed that a remote substituent in molecules in which direct $\pi$ transmission mechanisms are prohibited can have a very strong influence on the nuclear shielding: for example, the $^{119}$Sn substituent chemical shift (SCS) for derivative [11] (relative to parent molecule [12], which is due to both 'polar-field' and 'residual contribution' where the latter is probably due to 'through-bond' and/or 'through space electron delocalization', is very sensitive to polar substituent influences, like that derived from iodide groups.




[11] $\delta$ ($^{119}$Sn) : +14.54 ppm




[12] $\delta$ ($^{119}$Sn) :+ 0.46 ppm

## Coupling constants

The valence electron-mediated nuclear spin–spin couplings, $J(^{73}Ge,X)$, $J(^{117}Sn,X)$, $J(^{119}Sn,X)$ and $J(^{207}Pb,X)$, are dominated by the 'Fermi contact' term arising from the interaction of the nuclear spin magnetic dipole with the electron spin, which has a finite density at the nucleus. If the contact term is expressed with Pople and Santry's MO treatment and it is permissible to use the mean electronic excitation energy approximation, the coupling constants can be related to the hybridization (i.e. 's' electron density) and effective nuclear charge of the atoms participating in the bond. This implies that the parameters $J$(E,X) (E = Ge, Sn or Pb) are useful in the discussion of the structure and bonding.

The reduced couplings $K$(Ge,X) and $K$(Sn,X) are of opposite sign with respect to $J$(Ge,X) and $J$(Sn,X), respectively, owing to the negative magnetogyric ratios of $^{73}$Ge and $^{115/117/119}$Sn.

In the case of X = $^1$H, $^{19}$F, $^{31}$P, highly abundant spin-$\frac{1}{2}$ nuclei, the $^nJ$(E,X) data are obtained conveniently from the X atom NMR spectra, whereas in the case of couplings with rare spin-$\frac{1}{2}$ nuclei (X = $^{13}$C, $^{29}$Si, $^{15}$N), the $^nJ$(E,X) can be obtained also from E NMR spectra by using Hahn-echo extended pulse sequences. In tin derivatives, in the case of couplings to quadrupolar X nuclei (X = $^{11}$B, $^{14}$N), it frequently happens that splitting due to $^{119}$Sn–X coupling cannot be detected as a consequence of fast quadrupolar relaxation of X. The magnitude of $J$ can be obtained from Equation [6].

$$J(\mathrm{Sn,X}) = \frac{\pi}{2}\left(\frac{3}{T_2{}^{\mathrm{SC}}(^{119}\mathrm{Sn})\ S_\mathrm{x}(S_\mathrm{x}+1)\ T^{\mathrm{Q}}_{(\mathrm{x})}}\right)^{1/2} \qquad [6]$$

where $S_x$ is the spin $I$ of the quadrupolar nucleus, $T_2{}^{SC}(^{119}Sn)$ is the transverse scalar relaxation time, and $T_X^Q$ is the quadrupolar relaxation time.

The most correct way to obtain the coupling sign of $J$(E,X), which is helpful in the discussion of the nature of chemical bonding, is based on the analysis of an energy level diagram that describes the interaction in a spin system. In a system containing at least three different magnetically active nuclei (E, A and X), two of them are defined as active spins, whereas the other is the passive one. If the relative signs of the coupling $J$(E,A) and $J$(E,X) and the absolute sign of one of these two are known, it is possible to determine the absolute sign of the other.

Spin couplings involving $^{73}$Ge have been generally reported only in highly symmetrical compounds; **Table 5** gives a selection of the values the of coupling constants in germanium compounds of tetrahedral symmetry. The signs of these coupling constants are indicated only by comparison with those

**Table 5** Some representative spin coupling involving $^{73}Ge$

| *Compound* | *Coupling constant* | (Hz) |
|---|---|---|
| $GeH_4$ | $^1J(^{73}Ge,^1H)$: | –97.6 |
| $Me_4Ge$ | $^1J(^{73}Ge,^{13}C)$: | –18.7 |
| $GeF_4$ | $^1J(^{73}Ge,^{19}F)$: | –97.6 |
| $Me_4Ge$ | $^2J(^{73}Ge,^1H)$: | +3.0 |
| $(MeO)_4Ge$ | $^3J(^{73}Ge,^1H)$: | –1.9 |
| $(MeS)_4Ge$ | $^3J(^{73}Ge,^1H)$: | –2.5 |

of similar group 14 element compounds and not by double resonance experiments.

## One-bond couplings

In the case of couplings with electropositive elements, the one-bond reduced coupling constants show the trend $^1K(\text{Ge,X}) < {}^1K(\text{Sn,X}) < {}^1K(\text{Pb,X})$. In fact, the increase of $^1K(\text{E,X})$ with increasing size of E corresponds to the change of valence s electron densities $\Psi_M(O)^2$. This relationship is not valid when X is a less electropositive or more polarizable atom. **Table 6** gives a selection of values of a one-bond coupling constant involving $^{119}Sn$ and $^{207}Pb$.

### $^1J(E,^1H)$

Very few one-bond coupling constants between $^{119}Sn$ or $^{207}Pb$ and H have been measured. This is in part due to the instability and the difficulty of preparing suitable species. Limited data are available for $^{119}Sn$ and indicate that the magnitude of the $^1J(^{119}Sn–H)$ in $R_{4-n}SnH_n$ increases with $n$ when R = alkyl and decreases with $n$ when R = phenyl, owing to greater electronegativity of phenyl with respect to alkyl groups. In $SnH_3^-$ the presence of an electron lone pair at tin leads to a very low s character for hybrid orbitals in Sn–H bonds, and hence to a very small magnitude of coupling.

### $^1J(E,^{13}C)$

Owing to the large number of organolead and organotin species, there is a sizeable literature of data on $^1J(E,^{13}C)$ coupling constants (E = $^{119}Sn$ or $^{207}Pb$).

The magnitude of $^1J(^{119}Sn,^{13}C)$ organotin compounds has been related quantitatively to the C–Sn–C bond angles. A simple linear relationship between $^1J(^{119}Sn,^{13}C)$ coupling constants of a group of dimethyltin(IV) compounds and their C–Sn–C angle ($\theta$) has been found on the basis of the analysis of a large set of experimental data (Eqn [7]).

$$|^1J(^{119}\text{Sn},^{13}\text{C})| = (10.7 \pm 0.5)\theta - (778 \pm 64) \quad [7]$$

**Table 6** Some representative one-bond spin $^1J(^{119}Sn,X)$ and $^1J(^{207}Pb,X)$

| *Compound* | *X* | *E* | *¹J(E,X) values* (Hz) |
|---|---|---|---|
| $SnH_4$ | $^1H$ | $^{119}Sn$ | –1 930 |
| $MeSnH_3$ | $^1H$ | $^{119}Sn$ | –1 850 |
| $Me_2SnH_2$ | $^1H$ | $^{119}Sn$ | –1 797 |
| $Me_3SnH$ | $^1H$ | $^{119}Sn$ | –1 744 |
| $Ph_3SnH$ | $^1H$ | $^{119}Sn$ | –1 935.8 |
| $MeSnH_2Cl$ | $^1H$ | $^{119}Sn$ | –2 228 |
| $Me_4Sn$ | $^{13}C$ | $^{119}Sn$ | –340 |
| $Me_3SnCl$ | $^{13}C$ | $^{119}Sn$ | –380 |
| $Me_2SnCl_2$ | $^{13}C$ | $^{119}Sn$ | –566 |
| $Ph_2SnCl_2$ | $^{13}C$ | $^{119}Sn$ | –785 |
| $Me_2Sn(acac)_2$ | $^{13}C$ | $^{119}Sn$ | –966 |
| $Me_3SnLi·3THF$ | $^{13}C$ | $^{119}Sn$ | +155 |
| $(Me_3Sn)_2$ | $^{13}C$ | $^{119}Sn$ | –240 |
| $[SnF^6]^{2-}$ | $^{19}F$ | $^{119}Sn$ | 1 550 |
| $[Sn(OH)F_5]^{2-}$ | $^{19}F$ | $^{119}Sn$ | 1 776,127.8 |
| *cis*-$[Sn(OH)_2F_4]^{2-}$ | $^{19}F$ | $^{119}Sn$ | 1 820,1 518 |
| *trans*-$[Sn(OH)_2F_4]^{2-}$ | $^{19}F$ | $^{119}Sn$ | 1 956 |
| $(Me_3Sn)_3P$ | $^{31}P$ | $^{119}Sn$ | +832.5 |
| $SnCl_4(PEt_3)_2$ | $^{31}P$ | $^{119}Sn$ | 2 383.1 |
| $Me_3SnPH_2$ | $^{31}P$ | $^{119}Sn$ | +463 |
| $Me_3SnSnMe_3$ | $^{119}Sn$ | $^{119}Sn$ | +4240 |
| *trans*-$Pt(SnCl_3)_2(PEt_3)_2$ | $^{195}Pt$ | $^{119}Sn$ | 20 585 |
| $[Rh(SnCl_3)Cl_5]^{3-}$ | $^{103}Rh$ | $^{119}Sn$ | 864 |
| $Me_2PbH_2$ | $^1H$ | $^{207}Pb$ | +2 454.6 |
| $Me_3PbH$ | $^1H$ | $^{207}Pb$ | +2 295 |
| $Ph_4Pb$ | $^{13}C$ | $^{207}Pb$ | +481 |
| $Me_4Pb$ | $^{13}C$ | $^{207}Pb$ | +320 |
| $Me_3PbBr$ | $^{13}C$ | $^{207}Pb$ | +246 |
| $Me_3PbPbMe_3$ | $^{13}C$ | $^{207}Pb$ | +28 |
| $Ph_2Pb(O_2CMe)_2$ | $^{13}C$ | $^{207}Pb$ | 1 203 |
| $PhPb(O_2CMe)_3$ | $^{13}C$ | $^{207}Pb$ | 2 100 |
| $Me_3PbSeMe$ | $^{13}C$ | $^{207}Pb$ | –1170 |
| $Me_3PbCH{=}CHMe$ | $^{13}C$ | $^{207}Pb$ | 268 |
| $Me_3PbPbMe_3$ | $^{207}Pb$ | $^{207}Pb$ | +290 |
| $Me_3PbPbMe_3$ | $^{119}Sn$ | $^{207}Pb$ | –3 570 |
| $[Pt(PR_3)_2(PbPh_3)_2]$ | $^{195}Pt$ | $^{207}Pb$ | 14 500–18 500 |
| $Pb_9Pt$ | $^{195}Pt$ | $^{207}Pb$ | 4 122 |

A similar relation has been described for di-*n*-butyltin(IV) compounds (Eqn [8]).

$$|^1J(^{119}\text{Sn},^{13}\text{C})| = (9.99 \pm 0.73)\theta - (746 \pm 100) \quad [8]$$

and for tetra-, tri- and diphenyltin(IV) compounds (Eqn [9])

$$|^1J(^{119}\text{Sn},^{13}\text{C})| = (15.56 \pm 0.84)\theta - (1160 \pm 101) \quad [9]$$

Equation [7] is not valid for cyclic methyltin(IV) compounds.

The sign of $^{1}J(^{119}Sn,^{13}C)$ is generally negative for all tetraorganotin(IV) compounds, whereas $^{1}J(^{207}Pb,^{13}C)$ is positive for tetraorganolead(IV) derivatives.

In $R_4Sn$ (R = alkyl) compounds, $^{1}J(^{119}Sn,^{13}C)$ decreases in magnitude with increasing chain length and branching. It has also been found that in the tetraorganotin derivatives the magnitude of $^{1}J(^{119}Sn,^{13}C)$ increases with change of the hybridization of the carbon atoms from $sp^3$ to sp.

The replacement of alkyl groups with electronegative substituents causes an increase in the magnitude of $^{1}J(^{119}Sn,^{13}C)$ as the result of a greater s character in the Sn–C hybrid orbital.

Very small values of $^{1}J(^{119}Sn,^{13}C)$ are found in cyclopentadienyltin(IV) compounds owing to the low carbon s character in the pentahapto link to tin, whereas the small values in allyl- and benzyltin(IV) compounds may be ascribed to $\sigma$–$\pi$ conjugative connection of the polarized $\sigma$(Sn–C) bond with the adjacent $\pi$ electron system.

As in the case of the $\delta(^{119}Sn)$ chemical shift, moderate linear correlations have been found between the values of triphenyltin(IV) compounds and those of analogous triorganotin compounds $R_3SnX$ (R = $Bu^n$, Bz, Vi) which can be expressed as Equation [10]

$$^{1}J(^{119}Sn,^{13}C)(Ph_3Sn) = a\,^{1}J(^{119}Sn,^{13}C)(R_3Sn) + b \qquad [10]$$

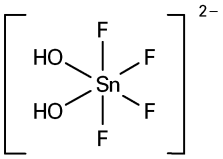
*a* and *b* being slope and intercepts, respectively.

### $^{1}J$(E,X) (X = $^{15}$N, $^{31}$P, $^{19}$F, $^{119}$Sn, $^{207}$Pb)

Several studies have been carried out on the bis(amino)stannylenes, where the coupling constants $^{1}J(^{119}Sn,^{15}N)$ are generally large and negative in accordance with a Ψ-pseudotrigonal-bipyramidal structure, whereas in stannylphosphane compounds $^{1}J(^{119}Sn,^{31}P)$ are generally positive and strongly dependent not only on changes in hybridization and substituent electronegativity, but also on s-overlap integral.

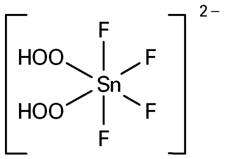
It has been found that the values of $^{1}K(^{119}Sn,^{31}P)$ for $(R_3Sn)_3P$ derivatives become less negative if the phosphorus lone electron pair is used in metal complexation.

$^{1}J(^{119}Sn,^{19}F)$ coupling constants are reported for hydroxo- and peroxotin(IV) fluoro complexes [13] and [14] and for $[SnF_6]^{2-}$ species. Their magnitude is strongly dependent on the stereochemistry.



[13]

$^{1}J(^{119}Sn,^{19}F)$:
1820 Hz (*trans*-OH)
1518 Hz (*trans*-F)



[14]

$^{1}J(^{119}Sn,^{19}F)$:
2064 Hz (*trans*-OOH)
1692 Hz (*trans*-F)

Several negative terms contribute to the extent of $^{1}K(E,^{19}F)$ so that no values are known for lead–fluorine and only a few tin–fluorine couplings have been reported.

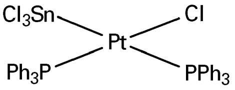
$^{1}J(^{119}Sn,^{119}Sn)$ values are generally positive and show clear dependence upon tin hybridization. When very bulky substituents are on tin atoms, the tin–tin s-overlap integral is small, whereas the mutual polarizability term changes sign, so that very small or negative $^{1}J(^{119}Sn,^{119}Sn)$ values are found. Large values of $^{1}J(^{119}Sn,^{119}Sn)$ have been detected in tin–tin complexes containing electronegative substituents or characterized by greater coordination number.

$^{1}J(^{119}Sn,^{207}Pb)$ values are normally negative and often of very small magnitude owing to reduced lead–tin s-overlap.

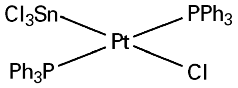
$^{1}J(^{207}Pb,^{207}Pb)$ are positive, and values are also very small for the same reason: lead–lead compounds are in fact easily dissociated in solution owing small lead–lead overlap.

### $^{1}J$(E, transition metal)

Many tin–transition metal coupling constants are known owing to the ease of obtaining [$(SnCl_3)$–transition metal] derivatives. The $^{1}J(^{119}Sn$, transition metal) values are generally large and are used to characterize complexes having tin–metal bonds also in solution. It has been observed that in square planar Pt(II) complexes [15] and [16], the magnitude of the $^{1}J(^{195}Pt,^{119}Sn)$ coupling constants depends on the magnitude of the *trans* effect of the respective substituent.



[15] $^{1}J(^{195}Pt,^{119}Sn)$: 16 321 Hz



[16] $^{1}J(^{195}Pt,^{119}Sn)$: 28 052 Hz

## Geminal coupling constants, $^2J$(E,Y)

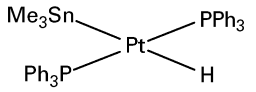
The geminal coupling constants $^2J$(E–X–Y) (**Table 7**) are sensitive not only to the factors reported for one-bond couplings, but also to the nature of the intervening atom X, to the metal–X–Y bond angle, to the stereochemistry of the molecule and finally to the nature of other substituents.

### $^2J$(E–X–$^1$H)

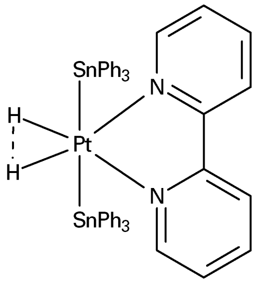
Several data have been reported, in particular for X = C. There is a linear relationship between $^2J(^{119}Sn,^1H)$ and $^2J(^{207}Pb,^1H)$ and also between $^2J(^{119}Sn,^1H)$ and $^1J(^{119}Sn,^{13}C)$ in related methyltin(IV) and methyllead(IV) derivatives. In several dimethyltin(IV) complexes the values of $^2J(^{119}Sn,^1H)$ can be used to calculate the bond angles C–Sn–C in solution (Eqn [11]).

$$(\text{angle C}-\text{Sn}-\text{C}) = 0.0161[\,^2J(^{119}\text{Sn},^1\text{H})_{\text{Me}})]^2 + 133.4 \qquad [11]$$

The sign of $^2J(^{119}Sn,^1H)$ is generally positive. Exceptionally the coupling can be negative, as in $Me_3Sn^-$ and diorganostannylenes, whereas the $^2J(^{207}Pb,^1H)$ are negative and normally greater than $^2J(^{119}Sn,^1H)$ in related compounds. The magnitude of $^2J$(E–C–$^1$H) increases with increasing coordination number on the metal. $^2J(^{207}Pb,^1H)$ values are generally less sensitive than $^2J(^{119}Sn,^1H)$ to structural geometrical differences. In several octahedral and square planar complexes, the magnitude of the $^2J(^{119}$Sn–X–$^1$H) across a transition metal X can be used to distinguish the *trans* (compound [17]) and *cis* positions (compound [18]) of Sn and H.




[17] $^2J(^{119}Sn,^1H)$: 1740 Hz



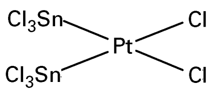

[18] $^2J(^{119}Sn,^1H)$: ± 27 Hz

### $^2J$(E–X–$^{13}$C)

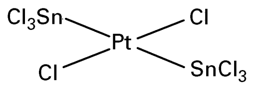
The $^2J(^{119}$Sn–X–$^{13}$C) couplings are generally small and positive in aliphatic and tetracoordinate tin derivatives, whereas they are large and negative in compounds containing an Sn–$C_{olefinic}$–C bond and close to 0 in compounds containing an Sn–C=C bond. In some cases the values are exceptionally large, as in Sn(C≡CBu)$_4$.

### Other geminal couplings

The couplings best studied are those involving $^{119}$Sn, $^{117}$Sn, $^{19}$F and $^{31}$P. For example, in bis(triorganotin)oxides the geminal couplings $^2J(^{119}Sn,^{119}Sn)$ indicate a linear structure in solution. The $^2J(^{119}$Sn–X–$^{119}$Sn), where the X intervening atom is a transition metal, show a strong dependence on the geometry and on the nature of other ligands on the metal or on tin atoms: for example, the $^2J(^{119}$Sn–Pt–$^{119}$Sn) is 2601 Hz in *cis*-[PtCl$_2$(Cl$_3$Sn)$_2$] compound [19] and 36 286 Hz in *trans*-[PtCl$_2$(Cl$_3$Sn)$_2$] (compound [20]).




[19] $^2J(^{119}Sn,^{119}Sn)$: 2601 Hz




[20] $^2J(^{119}Sn,^{119}Sn)$: 36286 Hz

A similar geometrical dependence has been shown by $^2J(^{119}$Sn–Pt–$^{31}$P). Two-bond couplings involving $^{19}$F are detected in $(CF_3)_4Ge$ (26.3 Hz), in $(CF_3)_4Sn$ (531 Hz) and also in the organolead derivatives $(Me)_3Pb(CF_3)$ (240 Hz) and $(Me)_2Pb(CF_3)_2$ (379 Hz), which show a strong dependence on the effective nuclear charge.

**Table 7** Selected germinal coupling constants $^2J$ (E,Y)

| *Compound* | *X* | *E* | *$^2J$(E,Y) values* (Hz) |
|---|---|---|---|
| $MeSnH_3$ | $^1$H | $^{119}$Sn | +63.6 |
| $Me_2SnH_2$ | $^1$H | $^{119}$Sn | +58 |
| $Me_2SnH$ | $^1$H | $^{119}$Sn | +56.5 |
| $Me_4Sn$ | $^1$H | $^{119}$Sn | +53.9 |
| $Me_2SnCl$ | $^1$H | $^{119}$Sn | +58.2 |
| $Me_2SnCl_2$ | $^1$H | $^{119}$Sn | +68.9 |
| $MeSnCl_3$ | $^1$H | $^{119}$Sn | +96.9 |
| $Me_2SnPH_2$ | $^1$H | $^{119}$Sn | 62.5 |
| $(Bu_2ClSn)_2O$ | $^{119}$Sn | $^{119}$Sn | 74 |
| $(Me)_3SnCF_2H$ | $^{19}$F | $^{119}$Sn | 265.6 |
| $Ph_2SnCl_2$ | $^{13}$C | $^{119}$Sn | 63 |
| $Me_2SnPh$ | $^{13}$C | $^{119}$Sn | 36.6 |
| $Et_4Sn$ | $^{13}$C | $^{119}$Sn | +23.5 |
| $Me_4Pb$ | $^1$H | $^{207}$Pb | −60.5 |
| $Me_2PbCl_2$ | $^1$H | $^{207}$Pb | −70 |
| $Me_2PbH_2$ | $^1$H | $^{207}$Pb | −76 |
| $Vl_4Pb$ | $^1$H | $^{207}$Pb | 213.3 |
| $Ph_4Pb$ | $^{13}$C | $^{207}$Pb | −68 |
| $Me_2PbPbMe_3$ | $^{13}$C | $^{207}$Pb | +92 |

**Table 8** Selected vicinal and long-range coupling constants $^nJ$(E,Y)

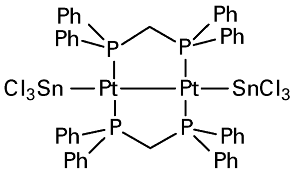
| *Compound* | *Y* | *E* | *n* | $^nJ$*(E,Y) values* (Hz) |
|---|---|---|---|---|
| $Et_4Sn$ | $^1$H | $^{119}$Sn | 3 | −71.2 |
| $Vi_4Sn$ | $^1$H | $^{119}$Sn | 3 | 183.1 |
| $EtSnCl_3$ | $^1$H | $^{119}$Sn | 3 | −242 |
| $Et_2SnCl_2$ | $^1$H | $^{119}$Sn | 3 | −130.8 |
| $PhSnCl_3$ | $^1$H | $^{119}$Sn | 3 | 122 |
| $(MeS)_4Sn$ | $^1$H | $^{119}$Sn | 3 | −66 |
| $(Me_2N)_4Sn$ | $^1$H | $^{119}$Sn | 3 | −51.0 |
| $Me_3SnSMe$ | $^1$H | $^{119}$Sn | 3 | −37.5 |
| $(Me_3Sn)_2C{=}CH_2$ | $^1$H | $^{119}$Sn | 3 | 124 (*cis*); 208 (*trans*) |
| $(CH_3)_2B–N(SnMe_3)_2$ | $^1$H | $^{119}$Sn | 3 | 33.0 (*cis*); 48.4 (*trans*) |
| $Me_2Sn(CF{=}CF_2)_2$ | $^{19}$F | $^{119}$Sn | 3 | 25 (*cis*); 29 (*trans*) |
| $Et_3Sn(CH_2)_2SnEt_3$ | $^{119}$Sn | $^{119}$Sn | 3 | 920 |
| $Bu_4Sn$ | $^{13}$C | $^{119}$Sn | 3 | 52.0 |
| $BuSnCl_3$ | $^{13}$C | $^{119}$Sn | 3 | 120.0 |
| $Ph_2SnCl_2$ | $^{13}$C | $^{119}$Sn | 3 | 90 |
| $Ph_2SnCl_2$ | $^{13}$C | $^{119}$Sn | 4 | 16 |
| $Me_3SnPh$ | $^{13}$C | $^{119}$Sn | 4 | 10.8 |
| $Et_4Pb$ | $^1$H | $^{207}$Pb | 3 | +128.6 |
| $Ph_4Pb$ | $^{13}$C | $^{207}$Pb | 3 | 80 |
| $Ph_4Pb$ | $^{13}$C | $^{207}$Pb | 4 | 20 |
| $(4\text{-}MeC_6H_4)_4Pb$ | $^1$H | $^{207}$Pb | 6 | 5.4 |
| $Me_3PbPbMe_3$ | $^1$H | $^{207}$Pb | 3 | +22.9 |
| $(Me_3CCH_2)_4Pb$ | $^1$H | $^{207}$Pb | 4 | 5.3 |

## Vicinal and long-range coupling constants, $^nJ$(E,Y) ($n \geq 3$)

Many vicinal coupling constants (**Table 8**) are known, with different combinations of intervening atoms. In most cases the $^3J(^{119}Sn,X)$ data are dependent on the respective dihedral angles and follow a Karplus-type relationship. The sign of $^3J(^{119}Sn,^1H)$ and $^3J(^{119}Sn,^{13}C)$ is negative and the magnitudes of $^3J(^{119}Sn,^1H)$ and $^3J(^{119}Sn,^{13}C)$ are usually greater than that of $^2J(^{119}Sn,^1H)$ and $^2J(^{119}Sn,^{13}C)$, respectively.

In several olefinic compounds it has been observed that $^3J(^{119}Sn,^{13}C)_{trans}$ is greater than $^3J(^{119}Sn,^{13}C)_{cis}$. The magnitude of $^3J(^{119}Sn,^1H)$ and $^3J(^{119}Sn,^{13}C)$ is strongly dependent on the nature of the substituents, an increase being observed in the presence of electropositive substituents on the C=C double bond.

$^3J(^{119}Sn,^{119}Sn)$ often behaves in a similar manner to $^3J(^{119}Sn,^1H)$ and $^3J(^{119}Sn,^{13}C)$. For example, $^3J(^{119}Sn,^{119}Sn)$ is 491 Hz in *cis*- and 1012 Hz in *trans*-$Me_3SnCH{=}CHSnMe_3$, respectively. An exceptionally large value of $^3J(^{119}Sn,^{119}Sn)$ is found for the derivative [21] owing to an additive coupling pathway across the Pt–Pt bond, whereas very small $^3J(^{119}Sn,^{119}Sn)$ are observed across the C≡C bond.




[21] $^3J(^{119}Sn,^{119}Sn)$: 25430 Hz

Several long-range coupling constants (**Table 8**) are reported for many different X nuclei, but they have not been systematically studied. Analogously to $^nJ(^1H,^1H)$ ($n \geq 3$), the $^nJ$(E,Y) couplings are best transferred by π system or by σ–π interaction.

## Relaxation behaviour

The dominant relaxation mechanism for $^{73}$Ge is quadrupolar, and it is possible to obtain $T_1$ directly from the line width.

The tin and lead compounds so far described are generally small, highly symmetrical molecules in which the spin-rotation (SR) mechanism is important.

In the tetrahalogermanes $GeX_4$ (X = Cl, Br or I), $T_1$ is almost exclusively dominated by the quadrupole relaxation mechanism, while $T_2$ is dominated by the combination of the scalar coupling and the quadrupole relaxation mechanism in the high-temperature region. The contribution of scalar coupling to the relaxation mechanism decreases in the order

$$GeCl_4 > GeBr_4 > GeI_4$$

Relaxation studies on liquid tetrahalotin(IV) compounds, carried out to determine temperature and field dependences of relaxation, indicated that $T_1$ for $SnCl_4$, $SnMe_4$ and $SnMeCl_3$ is dominated by a spin-rotation mechanism, while both scalar (SC) and spin-rotation (SR) mechanisms contribute to $T_1$ for $SnBr_4$ and $SnI_4$. On the other hand, $T_2$ for $SnCl_4$, $SnBr_4$ and $SnI_4$ is scalar dominated. When the molecules have hydrogens directly linked to tin, $^{119}$Sn/$^1$H dipolar relaxation is significant.

In the derivative $Bu_3SnSnBu_3$, the relaxation is about 75% $T_1^{DD}$.

In a 3 M solution of $Pb(ClO_4)_2$, the dominant $T_1$ relaxation is spin rotation. At higher concentration, CSA (chemical shift anisotropy) and dipolar contributions are also present and important. SC and SR mechanisms contribute to $T_1$ for neat $PbCl_4$. CSA is significant in $PbMe_3Cl$ in DMSO, which probably coordinates the lead atom.

## Applications

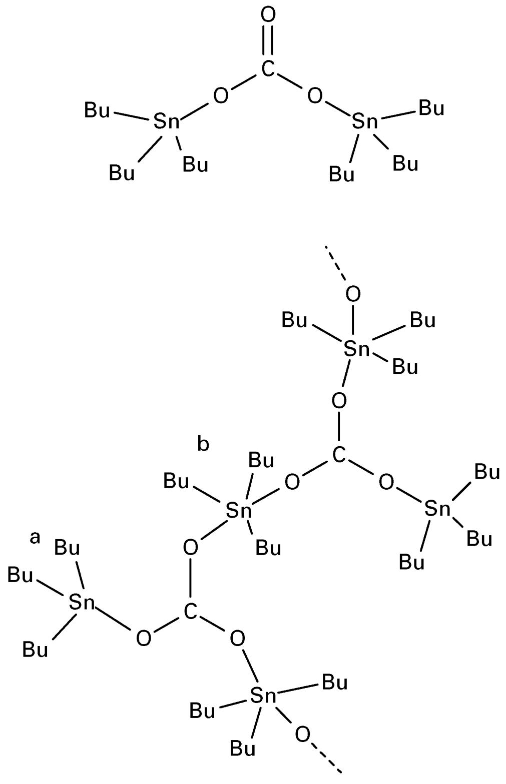
### Solid-state NMR

Solution and solid-state structures of metal and organometal derivatives are often not identical. High-resolution solid-state NMR experiments on metallic and organometallic compounds are important tools for bridging the information gap between X-ray diffraction results and solution-state NMR data. Until now, most of the high-resolution work in tin(IV), organotin(IV), lead and organolead(IV) chemistry has dealt with the combined use of cross-polarization and magic angle spinning (CP/MAS) techniques. The experimental conditions for $^{119}Sn$ and $^{207}Pb$ spectra are different from those usually employed for $^{13}C$ CP/MAS spectra of organic and organometallic compounds owing to smaller gyromagnetic ratios and larger shielding anisotropies. The usual CP/MAS situation for $^{119}Sn$ and $^{207}Pb$ is the presence of only a few resonances with large shielding anisotropies, so that it is very easy to determine shielding tensor components from one-dimensional CP/MAS spectra.

An example of application of CP/MAS is the $(Bu^{i}_3Sn)_2CO_3$ derivative [22]: the $^{119}Sn$ solution spectrum gives only one resonance at +101.7 ppm, in accordance with a monomeric tetracoordinated tin atom. The crystal structure of this compound clearly indicates a two-Sn environment: a polymeric chain of trigonal-bipyramidal *trans*-$R_3SnO_2$ units linked to each other via bridging carbonate anions and tetrahedral $R_3SnO$ units bonded to the third oxygen atom of the $CO_3^{2-}$ bridging group. In the $^{119}Sn$ CP/MAS spectrum, the presence of both tin fragments, a and b, is well resolved.

CP/MAS NMR spectra are also used to clarify the controversial view of the structures of $R_2SnX_2$ and $R_3SnX$ species in the solid and solution state: for example, the tricyclohexyltin hydroxide, which is a linear polymer in the solid state ($\delta = -217$ ppm) is monomeric in solution ($\delta = +11.6$ ppm). The diorganotin derivatives, which are monomeric in both phases, exhibit very similar shifts both in $^{119}Sn$ CP/MAS and $^{119}Sn$ solution NMR spectra. Small chemical shifts are observed in diorganotin and triorganotin derivatives weakly associated in the solid state and not associated in solution.

The CP/MAS NMR technique is also useful for investigating dynamic processes in solids such as reorientational processes and structural phase transitions: an example of solid-state phase transition monitored by variable temperature $^{119}Sn$ CP/MAS is given by $Me_3Sn–C{\equiv}C–C{\equiv}C–SnMe_3$, for which a structural phase change occurs in the temperature range 232–248K.




[22]

### Other applications

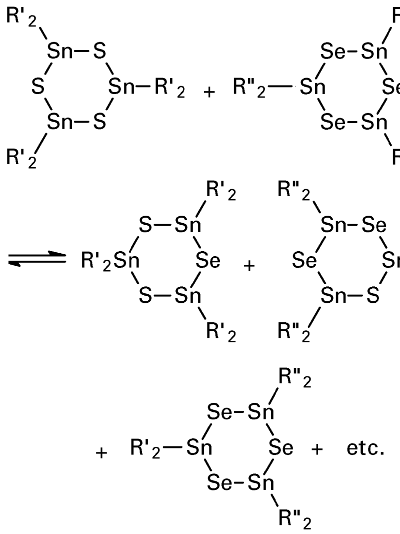
The $^{73}Ge$ NMR chemical shifts of several methyl- and phenyl-substituted germacyclohexanes have been determined, and it has been shown that the chemical shifts are very sensitive to steric effects and can provide a useful tool for the conformational analysis of organogermanium compounds.

The progress of reactions involving tin and lead compounds can readily be monitored by $^{119}Sn$ and $^{207}Pb$ NMR. For example, the various products derived from the 1,1-organoboration of 1-alkynyltin compounds can be easily identified by $^{119}Sn$ NMR, compounds [23, 24, 25].

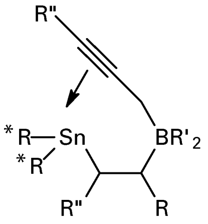
The content of complex reaction mixtures can be also analysed: for example, several species and intermediates are detectable in exchange reactions by $^{119}Sn$ NMR techniques, e.g. compounds in Equation [12].

1D and 2D $^{1}H$–$^{119}Sn$ HQMC experiments are useful in the investigation of weak coordination in functionalized triphenylvinyltin(IV) compounds of the type $Ph_3Sn–CH{=}CH–C(OH)RR'$.

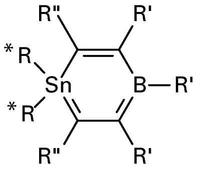
The HQMC technique is widely employed in the study of the condensation product of salicylaldoxime (*o*-HO–N=CH–$C_6H_4$–OH) with di-*n*-butyl oxide, which in solution exists as a mixture of several coordination complexes such as [26]. In this case, two-dimensional $^{1}H$–$^{119}Sn$ HQMC spectroscopy yields essential data for the structure elucidation of



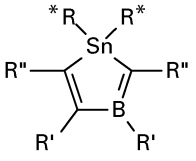
[12]



[23]

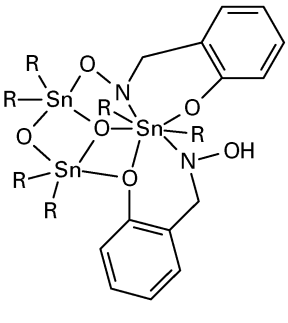


[24]



[25]

most of these complexes and also of the transient species, and allows several correlations between geometric parameters and $^{n}J(^{119}Sn,^{1}H)$ to be established.



[26]

## List of symbols

$T_1$ = spin–lattice relaxation times; $T_2$ = spin–spin relaxation time; $T^{Q}$ = quadrupolar relaxation time; $T^{SC}$ = transverse scalar relaxation time; $\delta$ = chemical shift; $\gamma$ = gyromagnetic ratio; $\sigma$ = shielding constant ($\sigma_d$ = diamagnetic, $\sigma_n$ = nonlocal contribution, $\sigma_p$ = paramagnetic).

*See also:* **NMR Principles; NMR Pulse Sequences; Parameters in NMR Spectroscopy, Theory of; $^{29}$Si NMR; Solid State NMR, Methods; Solid State NMR Using Quadrupolar Nuclei; Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules; Two-Dimensional NMR, Methods.**

## Further reading


Dechter JJ (1982) NMR of metal nuclides. Part I. The main group metals. In: Lippard SJ (ed). *Progress in Inorganic Chemistry*, pp 285–385. New York: Wiley.

Hani R and Geanangel RA (1982) $^{119}$Sn NMR in coordination chemistry. *Coordination Chemistry Reviews* **44**: 229–246.

Harris RK, Kennedy JD and McFarlane W (1978) Group IV–silicon, germanium, tin and lead. In: Harris RK and Mann BE (eds) *NMR and Periodic Table*, pp 340–371. New York: Academic Press.

Harrison PG (1989) Investigating tin compounds using spectroscopy. In: Harrison PG (ed) *Chemistry of Tin*, pp 60–117. Glasgow: Blackie.

Hunter BK and Reeves LW (1967) Chemical shift for compounds of the group IV elements silicon and tin. *Canadian Journal Chemistry* **46**: 1399–1414.

Kayser F, Biesemans M, Gielen M and Willem R (1996) Two dimensional $^{1}$H–$^{119}$Sn proton detected correlation spectroscopy in coordination chemistry of hypervalent organotin compounds. In: Gielen M, Willem R and Wrackmeyer B (eds), *Advanced Applications of NMR to Organometallic Chemistry*, pp 45–86. Chichester: Wiley.

Kennedy JD and McFarlane W (1987) Silicon, germanium, tin and lead. In: Mason J (ed) *Multinuclear NMR*, pp 305–333. New York: Plenum Press.

Petrosyan VS (1977) NMR spectra and structures of organotin compounds. *Progress in NMR Spectroscopy* **11**: 115–148.

Sebald A (1996) Solid state NMR applications in organotin and organolead chemistry. In: Gielen M, Willem R and Wrackmeyer B (eds), *Advanced Applications of NMR to Organometallic Chemistry*, pp 123–157. Chichester: Wiley.

Smith PJ and Smith L (1973) Application of $^{119}$Sn chemical shifts to structural tin chemistry. *Inorganica Chimica Acta Reviews* **7**: 11–33.

Smith PJ and Tupciauskas AP (1978) Chemical shifts of $^{119}$Sn nuclei in organotin compounds. *Annual Reports on NMR Spectroscopy* **8**: 291–370.

Wrackmeyer B (1985) $^{119}$Sn NMR parameters. *Annual Reports on NMR Spectroscopy* **16**: 73–185.

Wrackmeyer B (1996) Indirect nuclear $^{119}$Sn–X spin–spin coupling. In: Gielen M, Willem R and Wrackmeyer B (eds), *Advanced Applications of NMR To Organometallic Chemistry*, pp 87–122. Chichester: Wiley.

# Heteronuclear NMR Applications (La–Hg)

**Trevor G Appleton**, The University of Queensland, Brisbane, Australia

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
**Applications**

## Introduction

NMR spectroscopy for this group of metals is dominated by the three nuclei: $^{183}$W, $^{195}$Pt and $^{199}$Hg. Each of these will be dealt with in turn, followed by brief mention of other nuclei with non-zero values of the nuclear spin quantum number, *I*.

## Tungsten, $^{183}$W

The only tungsten isotope with a non-zero value of *I* is $^{183}$W, $I = \frac{1}{2}$, (14.4% natural abundance). Its receptivity relative to that of $^{1}$H is low ($1.06 \times 10^{-5}$). Its resonance frequency in a magnetic field of 2.35 T, where protons resonate at 100 MHz, is 4.17 MHz. This nucleus is radioactive, an $\alpha$-particle emitter with a long half-life (> $1 \times 10^{17}$ y). 'Satellite' peaks from coupling with $^{183}$W are readily observed in the $^{31}$P NMR spectra of tertiary phosphine complexes, the $^{1}$H NMR spectra of hydrides, the $^{19}$F NMR spectra of fluoro complexes and the $^{13}$C NMR spectra of carbonyl complexes. The earliest measurements of tungsten chemical shifts were indirect measurements through INDOR (internuclear double resonance) experiments. Direct measurement is now commonplace, but large sensitivity enhancements are possible with two-dimensional indirect spectroscopy, using the highly receptive 100% abundant nuclei $^{31}$P, $^{1}$H or $^{19}$F.

A solution of $Na_2[WO_4]$ (1 M in $D_2O$, pD 9) is now becoming accepted as the standard reference for $^{183}$W NMR in preference to $WF_6$. Relaxation times for $^{183}$W are frequently long. Since a major relaxation mechanism is through chemical shift anisotropy (CSA), shorter, more favourable relaxation times are obtained when the environment about the tungsten nucleus is less symmetric. Addition of a paramagnetic substance, such as $[Cr(acac)_3]$ can be used to shorten relaxation times which are inconveniently long. The tungsten chemical shift is very sensitive to the environment about the nucleus. The overall range is large (~6000 ppm). Shifts for some representative compounds are given in **Table 1**.

Variations in chemical shifts for heavy atom nuclei are dominated by changes in $\sigma_p$, the paramagnetic contribution to total nuclear screening. For a transition metal complex, a simplified expression for $\sigma_p$ is

$$\sigma_p \propto -\langle r^{-3}\rangle \sum_{j}^{\text{occ}} \sum_{k}^{\text{unocc}} (E_k - E_j)^{-1} C \qquad [1]$$

where *r* is an average distance of valence d-electrons from the nucleus, $E_k$ and $E_j$ are energies of unoccupied and occupied molecular orbitals, respectively, and *C* is a sum containing the coefficients of the metal s, p and d orbitals used in the summation to develop the various molecular orbitals. It will be noted that the tungsten nucleus in the series $[W(CO)_3(Cp)X]$ becomes progressively more shielded as the halide is changed from chloride to bromide to iodide, which is the 'normal halide dependence' order for transition metal complexes. This is opposite to the order expected from the effect of $X^-$ on excitation energies, which must therefore be

**Table 1** $^{183}$W chemical shifts for some representative compounds

| *Compound* | $\delta_W$ *(ppm)* |
|---|---|
| W(0) | |
| $W(CO)_6$ | −3505 |
| $W(CO)_5(PMePh_2)$ | −3324 |
| W(I) | |
| $[\{W(CO)_3(Cp)\}_2]$ | −4040 |
| W(II) | |
| $[W(CO)_3(Cp)Cl]$ | −2406 |
| $[W(CO)_3(Cp)Br]$ | −2584 |
| $[W(CO)_3(Cp)I]$ | −2996 |
| $[W(CO)_3(Cp)Me]$ | −3549 |
| $[W(CO)_3(Cp)H]$ | −4017 |
| W(IV) | |
| $[W(Cp)_2H_2]$ | −4671 |
| W(VI) | |
| $WF_6$ | −1120 |
| $[W_{10}O_{22}]^{4-}$ | −43 |
| $[WO_4]^{2-}$ | 0 |
| $[W_6O_{19}]^{2-}$ | +59 |
| $[WCl_6]$ | +2181 |

Cp = $\eta^5$-cyclopentadienyl.

outweighed by the effects of metal–ligand covalency on $r$ and the coefficients $C$. For the W(VI) complexes $WF_6$ and $WCl_6$, however, the order expected from excitation energies is observed.

One area where $^{183}W$ NMR has made a major contribution is in the study of iso- and heteropolytungstate ions. Since the tungsten atoms are usually in highly distorted octahedral environments, the relaxation times are short enough to allow convenient direct observation. The sensitivity of $\delta_w$ to the tungsten environment, and the observation of W–O–W coupling patterns, facilitate structural assignments. Paramagnetic heteropolytungstate ions have been included in these studies.

Where similar series of tungsten and molybdenum compounds have been studied by $^{183}W$ and $^{95}Mo$ NMR, respectively, there are close parallels between the two series. $^{95}Mo$ ($I = \frac{5}{2}$) is quadrupolar, and lines are relatively broad in unsymmetrical environments. Magic-angle spinning (MAS) solid-state $^{183}W$ NMR spectra may be obtained, but long spectrometer times are usually required because of long spin–lattice relaxation times.

## Platinum, $^{195}Pt$

The only platinum nucleus with magnetic properties is $^{195}Pt$, $I = \frac{1}{2}$, (33.7% abundance). The resonance frequency in a magnetic field of 2.35 T is approximately 21.4 MHz. 'Satellite' peaks from coupling with $^{195}Pt$ were observed in $^1H$ and $^{31}P$ NMR spectra in the 1960s, and much of the early work on $^{193}Pt$ detection used INDOR methods. Direct one-dimensional observation of $^{195}Pt$ NMR spectra is now routine. Because $^{195}Pt$ relaxation times are short for most compounds, there need be only a very short delay between pulses, allowing rapid accumulation. Two-dimensional inverse detection methods are also being increasingly used.

The most commonly used reference for $^{195}Pt$ NMR is an aqueous solution of $Na_2[PtCl_6]$. As a number of authors have pointed out, there are some disadvantages in using this substance. At high magnetic fields, peaks due to different combinations of chlorine isotopes may be resolved, and the resonance frequency is significantly dependent on temperature. This leads to some uncertainty (up to ± 5 ppm) in reported chemical shifts, but this is not very significant in the context of the overall chemical shift range for platinum (~15 000 ppm). However, this reference does have the great advantage of convenience, and is therefore unlikely to be supplanted. Other suggestions have been $[Pt(CN)_6]^{2-}$ (– 3863 ppm) and $\Xi_{Pt}$, a frequency of exactly 21.4 MHz in a magnetic field in which the resonance frequency of tetramethylsilane (TMS) is exactly 100 MHz (–4533 ppm). An aqueous solution of $K_2[PtCl_4]$ ($\delta_{Pt}$ –1624 ppm) is sometimes used as a secondary reference. Of greater concern than the small uncertainty in the reference frequency is the possibility, with a very wide chemical shift range for $^{195}Pt$, of observing 'folded' peaks. If peaks are not carefully 'checked for folding', reported shifts may be in error by hundreds of ppm!

Chemical shifts for some representative platinum compounds are given in **Table 2**. The order of ligands in their effect on $\delta_{Pt}$ is similar to that on other well-studied transition metals (e.g. $^{59}Co$, $^{103}Rh$). In particular, ligands with heavy donor atoms tend to cause increased shielding of the metal nucleus so that, for example, the shielding increases from $Cl^-$ to $Br^-$ to $I^-$. As discussed above for W(II) compounds, this may be understood in terms of the heavier donor atoms having a greater effect on $r$ and $C$ in Equation [1].

While $\delta_{Pt}$ depends primarily on the donor atom set, the nature of the whole ligand does have an important effect (compare, for example, $[Pt(NH_3)_4]^{2+}$ and $[Pt(NO_2)_4]^{2-}$). Geometric isomers will also have different shifts, which are greatest when the different ligands bound to the metal are very different in their *trans* influence (compare the *cis* and *trans* pairs $[PtCl_2(NH_3)_2]$ and $[Pt(NH_3)_2(H_2O)_2]^{2+}$). There are also 'ring-size' effects, evident in the difference between the shifts for $[\{Pt(NH_3)_2(\mu\text{-}OH)\}_n]^{n-}$ with $n = 2$ (four-membered ring) and $n = 3$ (six-membered ring), and for *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ and $[Pt(en)(H_2O)_2]^{2+}$ (five-membered chelate ring).

Chemical-shift anisotropy relaxation is important for platinum complexes, especially for square planar complexes, and at high magnetic field strengths. It is also enhanced when there is significant hydrogen-bonding between the solvent and solute (as with diammineplatinum compounds in water), and when platinum is coordinated to bulky ligands (both of which will decrease the rate of tumbling in solution). Fast CSA relaxation leads to broadening of peaks and loss of resolved splittings from couplings. It can also effectively decouple platinum from another nucleus which is being observed. Coupling constants involving $^{195}Pt$ [e.g. $^1J(Pt–^{31}P)$, $^1J(Pt–^{15}N)$, $^1J(Pt–^1H)$, $^1J(Pt–^{13}C)$, $^2J(Pt–CH_3)$, $^2J(Pt–CF_3)$] can provide useful structural information (e.g. through the *trans* influence of the *trans* ligand on these couplings) which may therefore be lost at high magnetic fields. In compounds containing platinum–platinum bonds, the magnitude of $J$(Pt–Pt) does not appear to be related simply to bond length or bond strength, although $J$(Pt–Pt) is affected in the same way as other coupling constants by the *trans* influence of the ligands *trans* to the metal–metal bond.

**Table 2** $^{195}$Pt chemical shifts for some representative compounds

| *Compound* | $\delta_{Pt}$ |
|---|---|
| Pt (0) | |
| $[Pt(PPh_3)_3]$ | −4583 |
| [Pt(1,5-cyclooctadiene)$_2$] | −4636 |
| $[Pt(PMe_2Ph)_4]$ | −4728 |
| $[Pt(PCy_3)_2]$ | −6501 |
| Pt(I) (Pt—Pt bond) | |
| $[\{PtCl_2(CO)\}_2]^{2-}$ | −4162 |
| Pt(II) | |
| $[Pt(H_2O)_4]^{2+}$ | +31 |
| $[\{Pt(NH_3)_2(\mu\text{-}OH)\}_2]^{2-}$ | −1153 |
| $[PtCl_3(H_2O)]^-$ | −1180 |
| *trans*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ | −1374 |
| $[\{Pt(NH_3)_2(\mu\text{-}OH)\}_3]^{3-}$ | −1505 |
| *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ | −1584 |
| $[PtCl_4]^{2-}$ | −1620 |
| $[PtCl_3(NMe_3)]^-$ | −1715 |
| $[Pt(en)(H_2O)_2]^{2+}$ | −1914 |
| *trans*-$[PtCl_2(NH_3)_2]$ (indmf) | −2101 |
| *cis*-$[PtCl_2(NH_3)_2]$ (indmf) | −2104 |
| $[Pt(NO_2)_4]^{2-}$ | −2166 |
| $[Pt(NH_3)_4]^{2+}$ | −2580 |
| $[PtBr_4]^{2-}$ | −2690 |
| $[PtCl_3(SMe_2)]^-$ | −2757 |
| $[PtCl_3(PMe_3)]^-$ | −3500 |
| $[Pt(CN)_4]^{2-}$ | −4746 |
| Pt(III) (Pt—Pt bond) | |
| $[\{(H_2O)Pt(\mu\text{-}SO_4)_2\}_2]^{2-}$ | +1753 |
| $[\{ClPt(\mu\text{-}P(O)_2O(O)_2P)_2\}_2]^{4-}$ | −4236 |
| Pt(IV) | |
| $[PtF_6]^{2-}$ | +7314 |
| $[Pt(OH)_6]^{2-}$ | +3277 |
| $[PtCl_6]^{2-}$ | 0 |
| *fac*-$[PtMe_3(H_2O)_3]^+$ | −1794 |
| $[PtBr_6]^{2-}$ | −1860 |
| $[Pt(CN)_6]^{2-}$ | −3866 |

When platinum is bound by ligands with donor atoms having quadrupolar nuclei (e.g. $^{14}$N, $^{75}$As), NMR peaks are frequently broad, because the rate of quadrupole-induced relaxation is such that $^{195}$Pt is only partially decoupled from the quadrupolar nucleus.

Solid state $^{195}$Pt NMR spectra, including CP-MAS spectra, have been obtained. Much higher spinning rates are required for platinum(II) compounds than for most platinum(lV) complexes, because of their much greater shielding anisotropy.

## Mercury, $^{199}$Hg

Mercury has two isotopes with $I > 0$, $^{199}$Hg ($I = \frac{1}{2}$; 16.8% abundance) and $^{201}$Hg ($I = \frac{3}{2}$; 13.2% abundance). Because of the large quadrupole moment of $^{201}$Hg, this nucleus is not readily observed, and only $^{199}$Hg is of significance for NMR spectroscopy. Its resonance frequency in a magnetic field with strength 2.35 T is 17.9 MHz. There is general agreement on the use of neat dimethylmercury as reference, despite the high toxicity of this substance. An aqueous solution of $Hg(ClO_4)_2$ (–2253 relative to $Me_2Hg$) has been suggested as an alternative but its shift is dependent on concentration and temperature.

Common coordination behaviour for Hg(II) involves strong linear coordination by two ligands, often with additional weaker interactions from solvent molecules, counter-ions, etc. In some cases, (e.g. halides $X^-$) higher coordination numbers are possible (e.g. tetrahedral $[HgX_4]^{2-}$). There is considerable ligand lability, so that averaged signals are often observed. There is therefore often some uncertainty about the precise nature of the species being observed.

Chemical shift anisotropies are also large (except for highly symmetrical $[HgX_4]^{2-}$), leading to efficient relaxation, and some line broadening at high magnetic fields. Two-dimensional NMR spectra may be readily obtained.

Some $^{199}$Hg chemical shift data are given in **Table 3** for some representative compounds whose structures appear to be fairly well defined in solution. As with typical transition metals, such as $^{195}$Pt, and $^{113}$Cd, the nuclear shielding increases in halide complexes as the halide ion becomes heavier (‘normal halide dependence’)

Coupling constants between $^{199}$Hg and other nuclei [e.g $^{13}$C, $^{31}$P and $^1$H($^2J$, $^3J$)] have been extensively studied. In a series of complexes (e.g. $CH_3HgX$) the coupling constants [$^2J$(Hg–$CH_3$) in this instance] may be used to establish a *trans* influence series which is similar to that obtained from coupling constants involving $^{195}$Pt.

**Table 3** $^{199}Hg$ shifts for some representative compounds

| *Compound* | $\delta_{Hg}$ *(ppm)* |
|---|---|
| Hg(I) | |
| $[Hg_2]^{2+}(ClO_4)_2/H_2O$(satd.) | −1614 |
| Hg(II) | |
| $Me_2Hg$ | 0 |
| MeHgCl | −813 |
| MeHgBr | −915 |
| $[MeHg(NH_3)]BF_4$ | −943 |
| MeHgI | −1097 |
| $[MeHg(OH_2)]^+$ | −1150 |
| $[HgCl_4]^{2-}$ | −1200 |
| $[HgBr_4]^{2-}$ | −1810 |
| $[Hg(H_2O)_6]ClO_4)_2$ | −2253 |
| $[HgI_4]^{2-}$ | −3510 |

## Other elements, La–Hg

### Lanthanum

Nuclei with $I > 0$ are $^{139}La$ ($I = \frac{7}{2}$; 99.91% abundance) and $^{138}La$ ($I = 5$; 0.09% abundance). Because of its low natural abundance, $^{138}La$ is of no practical importance for NMR spectroscopy. The common oxidation state, La(III), is diamagnetic but co-ordination complexes are usually labile, and quadrupole-induced relaxation causes broad line widths, except when the environment is highly symmetric, as in $[LaX_6]^{3-}$ complexes, which are presumably octahedral. Both the chemical shift and line width can provide information about the average environment of the $La^{3+}$ ion. As expected for a $d^0$ ion, the metal nucleus is less shielded in $[LaBr_6]^{3-}$ (+1090 ppm from aqueous $La^{3+}$) than in $[LaCl6]^{3-}$ (+851 ppm).

### Other lanthanides

Diamagnetic oxidation states tend to be unstable in aqueous solution, except for Yb(II) and Lu(III). Since $^{171}Yb$ (14.3% abundance) has $I = \frac{1}{2}$, there have been some solution studies on Yb(II) complexes. $^{169}Tm$ (100% abundance) also has $I = \frac{1}{2}$, but diamagnetic compounds are not stable in solution. $^{175}Lu$ ($I = \frac{7}{2}$, 97.4% abundance) has a large quadrupole moment, which limits its use. Solid state spectra have been obtained on paramagnetic compounds using nuclei with relatively low quadrupole moments: $^{141}Pr$ ($I = \frac{5}{2}$; 100% abundance), $^{151}Eu$ ($I = \frac{5}{2}$; 47.8% abundance), $^{159}Tb$ ($I = \frac{3}{2}$; 100% abundance), $^{165}Ho$ ($I = \frac{7}{2}$; 100% abundance) and with $^{169}Tm$ ($I = \frac{1}{2}$; 100% abundance).

### Hafnium

$^{171}Hf$ ($I = \frac{7}{2}$; 18.5% abundance) and $^{179}Hf$ ($I = \frac{9}{2}$; 13.8% abundance) have large quadrupole moments, causing severe line broadening.

### Tantalum

$^{181}Ta$ ($I = \frac{7}{2}$; 99.99% abundance) also possesses a large quadrupole moment so that, even in octahedral species such as $[TaCl_6]^-$, lines are broad.

### Rhenium

$^{185}Re$ ($I = \frac{5}{2}$; 37.1% abundance) and $^{187}Re$ ($I = \frac{5}{2}$; 62.9% abundance) also have large quadrupole moments, so that the line widths are large even for the symmetric species $[ReO_4]^-$.

### Osmium

$^{187}Os$ has $I = \frac{1}{2}$, but its low natural abundance (1.64%) and its very low receptivity make it difficult to observe. $^{189}Os$ ($I = \frac{3}{2}$; 16.1% abundance) has a large quadrupole moment, leading to very broad lines, except in highly symmetric environments. In some cases, $^{187}Os$ spectra may be obtained by indirect detection, using more sensitive nuclei such as $^{31}P$. Coupling constants between $^{187}Os$ and other nuclei (e.g. $^{31}P$, $^{13}C$) may be easily observed, and provide structural information.

### Iridium

$^{191}Ir$ ($I = \frac{3}{2}$; 37.3% abundance) and $^{193}Ir$ ($I = \frac{3}{2}$; 62.7%) have large quadrupole moments, low receptivity and low resonance frequencies, which make observation difficult.

### Gold

$^{197}Au$ ($I = \frac{3}{2}$; 100% abundance) has a large quadrupole moment. With the preferred linear [Au(I)] and square planar [Au(III)] geometries, quadrupole-induced relaxation is very fast, and signals have not been observed in solution.

## List of symbols

$I$ = nuclear spin quantum number; $J$ = coupling constant; $\sigma_p$ = paramagnetic contribution to total nuclear screening.

*See also:* **Fourier Transformation and Sampling Theory; Heteronuclear NMR Applications (Sc–Zn); Heteronuclear NMR Applications (Y–Cd); NMR Relaxation Rates; NMR Spectrometers.**

## Further reading

Appleton TG, Clark HC and Manzer LE (1973) The trans influence: its measurement and significance. *Coordination Chemistry Reviews* **10**: 335–422.
Brevard C, Pregosin PS and Thouvenot R (1991) Group 6 chromium to tungsten. In: Pregosin PS (ed) *Transition Metal Nuclear Magnetic Resonance*, pp 59–89. Amsterdam: Elsevier.
Goodfellow RJ (1987) Group VIII transition metals. In: Mason J (ed) *Multinuclear NMR*, pp 521–561. New York: Plenum.
Goodfellow RJ (1987) Post-transition metals, copper to mercury. In: Mason J (ed) *Multinuclear NMR*, pp 563–589. New York: Plenum.
Granger P (1991) Groups 11 and 12 copper to mercury. In Pregosin PS (ed) *Transition Metal Nuclear Magnetic Resonance*, pp 265–346. Amsterdam: Elsevier.
Pregosin PS (1982) Platinum-195 nuclear magnetic resonance. *Coordination Chemistry Reviews* **44**: 247–291.
Pregosin PS (1986) Platinum NMR spectroscopy. *Annual Reports on NMR Spectroscopy* **17**: 285–349.
Pregosin PS (1991) Group 10 (nickel to platinum). In: Pregosin PS (ed) *Transition Metal Nuclear Magnetic Resonance*, pp 216–263. Amsterdam: Elsevier.
Rehder D (1987) Early transition metals, lanthanides and actinides. In: Mason J (ed) *Multinuclear NMR*, pp 479–519. New York: Plenum.
Rehder D (1991) Groups 3–5 scandium to tantalum. In: Pregosin PS (ed) *Transition Metal Nuclear Magnetic Resonance*, pp 1–58. Amsterdam: Elsevier.
Wrackmeyer B and Contreras R (1992) $^{199}$Hg NMR parameters. *Annual Reports on NMR Spectroscopy* **24**: 267–329.

# Heteronuclear NMR Applications (O, S, Se and Te)

**Ioannis P Gerothanassis**, University of Ioannina, Greece

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
**Applications**

## NMR spectroscopic properties and techniques for $^{17}$O, $^{33}$S, $^{77}$Se and $^{123,125}$Te

### Nuclear properties

Both $^{17}$O and $^{33}$S are quadrupolar nuclei with very low natural abundance (**Table 1**). The NMR active isotopes of selenium and tellurium are spin $\frac{1}{2}$ nuclei. The receptivity of $^{77}$Se is about three times larger than that of $^{13}$C. The element tellurium has two active isotopes, $^{123,125}$Te. $^{125}$Te has been used in the majority of tellurium NMR investigations owing its significantly better receptivity.

### Experimental techniques

The stringent requirements in $^{17}$O and $^{33}$S NMR studies of compounds at natural abundance are the high concentrations and extensive signal averaging needed. Recording of spectra can be greatly facilitated by the use of enriched samples. Synthesis with oxygen isotopes involves rather straightforward organic reactions. The quadrupolar moments of $^{17}$O

**Table 1** Nuclear properties of $^{17}$O, $^{33}$S, $^{77}$Se, $^{123}$Te and $^{125}$Te

| *Isotope* | *Spin* | *Natural abundance (%)* | *NMR frequency[a] (MHz)* | *Receptivity* | |
|---|---|---|---|---|---|
| | | | | *Relative[b]* | *Absolute[c]* |
| $^{17}$O | $\frac{5}{2}$ | 0.037 | 54.227 | $2.91 \times 10^{-2}$ | $1.08 \times 10^{-5}$ |
| $^{33}$S | $\frac{3}{2}$ | 0.76 | 30.678 | $2.26 \times 10^{-3}$ | $1.71 \times 10^{-5}$ |
| $^{77}$Se | $\frac{1}{2}$ | 7.58 | 76.270 | $6.93 \times 10^{-3}$ | $5.26 \times 10^{-4}$ |
| $^{123}$Te | $\frac{1}{2}$ | 0.87 | 104.831 | $1.80 \times 10^{-2}$ | $1.57 \times 10^{-4}$ |
| $^{125}$Te | $\frac{1}{2}$ | 6.99 | 126.387 | $3.15 \times 10^{-2}$ | $2.91 \times 10^{-3}$ |

[a] At 9.395 T ($^{1}$H frequency: 400 MHz).
[b] Relative to proton, at constant field, for equal number of nuclei.
[c] Product of relative receptivity and natural abundance.

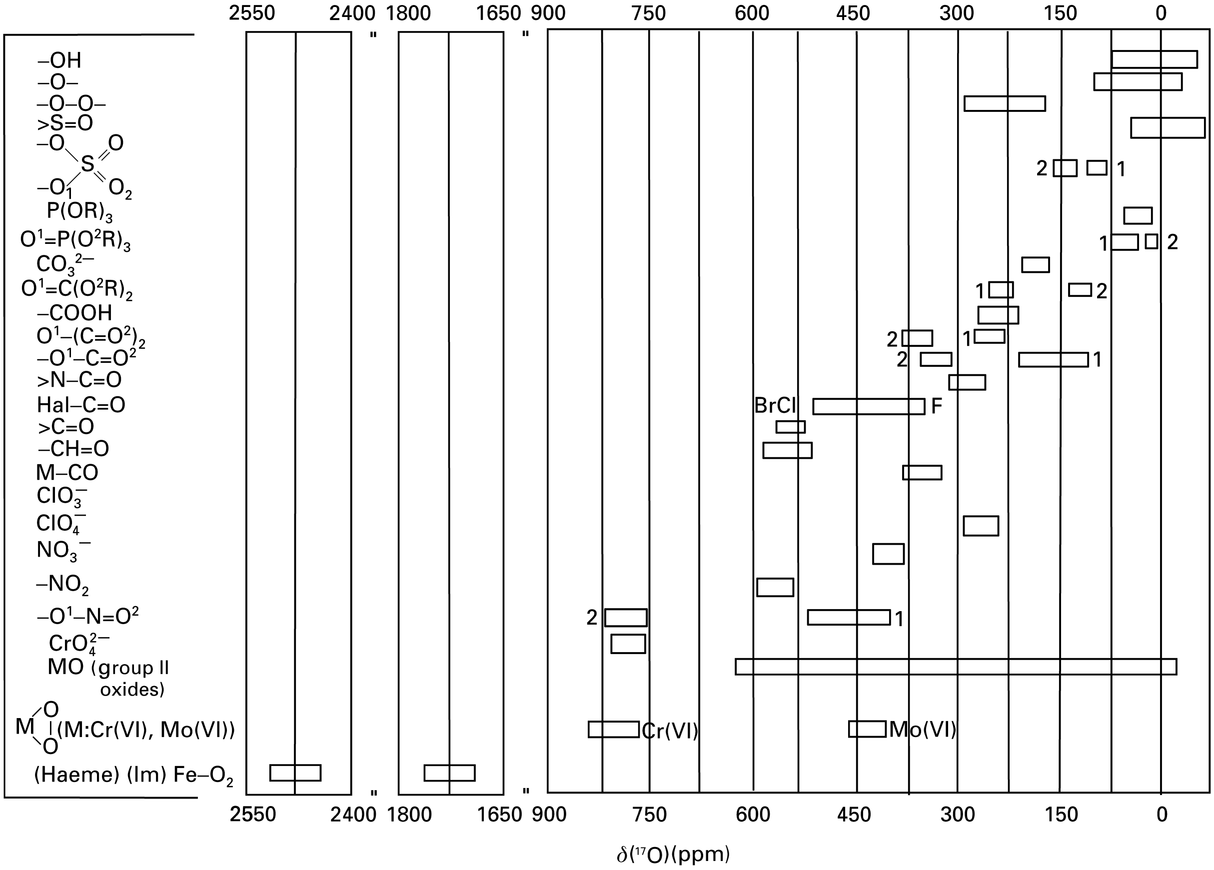



**Figure 1** $^{17}O$ chemical shift ranges of selected diamagnetic oxygen functional groups.

and $^{33}S$ result in very effective relaxation and, thus, extensive line broadening with $T_1$ and $T_2$ relaxation times often in the range below 1 ms. A possible way to attenuate extensive line-broadening effects is to use higher temperatures and/or solvents of low viscosity. Observation of very broad (>1 kHz) resonances, however, is a potential problem due to baseline distortions. The most severe origin of baseline artifacts is the transient response of the NMR probe, often referred to as 'acoustic ringing', caused by the generation of ultrasonic waves from the action of the RF pulse. Several methods have been proposed to overcome this problem, with particular emphasis on multipulse sequences.

The longitudinal relaxation of $^{77}Se$ and $^{125}Te$ at high fields is governed by the spin–rotation (SR) mechanism with typical $T_1$ values in the range of 1–30 s. In the solid state the relaxation times are very long; consequently the use of MAS NMR with cross polarization from $^1H$ is necessary. The $^{77}Se$ and $^{125}Te$ shieldings indicate a strong temperature dependence (several ppm/K), and so extreme care is necessary to maintain constant temperature, particularly when broadband decoupling is employed.

As is common with many heteronuclei, a variety of reference compounds and referencing procedures (both internal and external) have been suggested. The most commonly used standards ($\delta = 0$, as external reference) are $H_2O$ (a temperature dependence of the $^{17}O$ shielding has been reported) or 1,4-dioxane at 30°C for $^{17}O$; $(NH_4)_2SO_4$ or $Cs_2SO_4$ in water for $^{33}S$; $Me_2Se$ in $CDCl_3$ for $^{77}Se$ and $Me_2Te$ in $CDCl_3$ for $^{125}Te$.

## $^{17}O$ NMR parameters and applications

### $^{17}O$ shieldings

The total range of $^{17}O$ shieldings of diamagnetic compounds is about 2600 ppm (**Figure 1**). For compounds containing oxygen bonded only to carbon and/or hydrogen there is a clear correlation between C–O $\pi$ bond order and high frequency shift. This type of correlation arises from the multiple bond term in the Karplus–Pople equation. Bridging oxygens indicate a strong deshielding for $\pi$-accepting acyl substituents which introduce $\pi$ character into the C–O bond of the bridging oxygen.

In nitrogen compounds, chemical shift values are also determined largely by π bonding effects. Bridging oxygen resonances are observed at higher frequencies in nitrites (R–*O*–N=O), where π bonding is appreciable. A lesser but nonetheless significant effect is observed for bridging oxygens in nitrates (R–*O*–$NO_2$). The most striking feature of the $^{17}O$ shifts observed for P–O bonds is the small chemical shift range (20–260 ppm) which cannot be related to P–O bond order in a straightforward fashion. However, $^{17}O$ resonances for non-equivalent oxygens in phosphates and polyphosphates could be resolved and could be utilized in stereochemical and binding site studies. For S–O bonding, similar trends to those of P–O bonds were observed. Modification of π-bond order in a X=O group due to substituents on X results in a decrease (or increase) in the net charge on X as the charge on O increases (or decreases). The chemical shift variations experienced by the nuclei X and O are then opposite and roughly proportional to the ratio $\langle r^{-3}\rangle 2P_X/\langle r^{-3}\rangle 2P_O$.

For a series of oxyanions and a series of C=O and N=O compounds a linear correlation between $\delta(^{17}O)$ and the inverse of the lowest energy $n \rightarrow \pi^*$ transition was found. A linear correlation exists between acetone chemical shifts and the maximum of the $n \rightarrow \pi^*$ transition for acetone–water solutions or for acetone in different solvents. Similar trends have also been observed in a variety of substituted acetophenones; however, it was suggested that such trends may be fortuitous. In a series of aliphatic ethers a correlation between $\delta(^{17}O)$ and the first ionization potential was suggested.

Compounds that contain oxygen–oxygen bonds and oxygen–fluorine bonds display large high frequency shifts. The ability, therefore, of $^{17}O$ NMR to unambiguously distinguish between non-equivalent oxygens has found application in several structural and mechanistic studies. Many stereochemical studies have dealt with polynuclear, anionic oxo complexes (polyoxoanions) of the early transition elements in their highest oxidation states. Despite their structural complexity, these polyoxoanions often yield well-resolved $^{17}O$ NMR spectra since large shielding ranges are observed and samples may be studied at elevated temperatures in nonviscous solvents. Furthermore, $^{17}O$ enrichment can easily be obtained.

### Indirect spin coupling constants

Because of the breadth of $^{17}O$ resonances, only relatively large coupling constants can be measured and hence the majority of those recorded refer to couplings from directly bonded nuclei. $^1J$(O,H) in water and in alcohols has a value close to 80 Hz. The $^1J$(P,O) couplings involving P=O groups cover a fairly wide range (145–210 Hz) and for a series of compounds of type $OPX_3$ (X=F, O in OMe, N in $NMe_2$ and C in Me) there is a consistent drop in coupling with decreasing electronegativity of X. The oxygens in P–OMe groups within phosphates give rise to a consistently smaller coupling constant, compared with P=O, of between 90 and 100 Hz. The reduced coupling constants decrease with increasing atomic number of the central atoms along the isoelectronic sequence $VO_4^{3-}$, $CrO_4^{2-}$ and $MnO_4^-$. An application of $^{17}O$ spin–spin couplings has been made to investigate $H_3O^+$ species.

### Nuclear quadrupole coupling constants (χ)

Experimental χ values cover a range from –8 to 17 MHz. The experimental χ values for a series of compounds in the solid state with C=O, –O–, PO, NO, and SO bonds show the following trends:

$$\begin{array}{ccccccccccc} \mathrm{O-O} & > & \mathrm{N-O} & > & \mathrm{NO_2} & > & \mathrm{C=O} & > & \mathrm{P-OR} \sim \mathrm{S{=}O} & > \\ 16-17 & & 16 & & 13 & & 11-9 & & 9.5-9 & \end{array}$$

$$\begin{array}{ccccccc} \mathrm{ROH} & > & \mathrm{SO_2} & > & \mathrm{P=O} \sim \mathrm{C=O} & > & \mathrm{XO}_n^{m-} \\ 9-8 & & 6.7 & & 4 & & 1 \end{array}$$

Strong hydrogen bonding reduces the χ values by up to 2.6 MHz. The χ values of the terminal CO groups in metallocarbonyls were found to be between 1–3 MHz although values below 0.1 MHz have also been observed. It was suggested that increased dπ–pπ back bonding from the metal will increase electron density perpendicular to the C–O axis and so reduce χ values.

### Dynamic and kinetic studies

Numerous dynamic and kinetic studies have been performed using $^{17}O$ NMR. Line shape analysis is not straightforward because the line width itself is temperature dependent. Nevertheless, it has been shown that dynamic $^{17}O$ NMR is a convenient alternative to $^{13}C$ NMR with, for example, methyl carbonyl compounds. A substantial number of publications have been concerned primarily with studies of the hydration (solvation) of ions in solution and the determination of rates and activation parameters for ligand substitution (a rapid development of high-pressure NMR has occurred in the last twenty years). There is an extensive literature of $^{17}O$ NMR spectroscopy used for the study of hyperfine interactions between the unpaired electrons in paramagnetic molecules and ions and the $^{17}O$ nucleus. The

magnitudes and temperature dependences of line widths (or relaxation time measurements) of the 'bound' and 'free' resonances, and the chemical shift separations between those resonances, can in principle be used to derive information about the hyperfine splitting constant, the rate of the solvent interchange process and the solvation number. However, because the shifted resonance of the solvated water molecules is often very broad, it is usually not possible to derive the hydration number of the paramagnetic cation from straightforward area measurements. In these cases, the hydration numbers may be obtained by the use of theory developed by Swift and Connick.

#### Torsion angle and hydrogen bonding effects

$^{17}$O NMR spectroscopy is a powerful method for the detection of steric effects in molecules in which steric interactions are characterized by rotation of functional groups around single bonds to relieve van der Waals interactions or on rigid systems in which steric interactions are partially accommodated by bond angle and bond length distortions. Applications include aromatic and heteroaromatic nitro compounds, aryl ketones, aldehydes, 1,2-diketones, aromatic carboxylic acids, esters and amides.

$^{17}$O NMR appears to be especially promising for use in hydrogen bonding interaction studies because of the large chemical shift range of the oxygen nucleus. The dominance of intramolecular hydrogen bonding effects over substituent effects was clearly demonstrated in acetophenones, benzaldehydes and hydroxynaphthoquinones. Detailed studies of the influence of intermolecular solute–solvent hydrogen bonding in amides and peptides demonstrate that both long range dipole–dipole interactions and hydrogen bonds at the amide oxygen induce significant and specific shielding of the $^{17}$O nucleus.

#### Water orientation in lyotropic phases and bilayers – hydration of proteins and DNA

$^{17}$O NMR studies have several advantages compared with $^{1}$H and $^{2}$D NMR for investigating water orientation in lyotropic phases and bilayers and the hydration of proteins and DNA since the relaxation effect is large and the intramolecular origin of the electric field gradient makes the quadrupolar interaction virtually independent of the molecular environment. Recently, it has been suggested that the relaxation dispersion in aqueous protein and DNA solutions in the MHz frequency range is not owing to a long-lived hydration layer at the biomolecular surface, but to a small number of crystallographically well-defined water molecules.

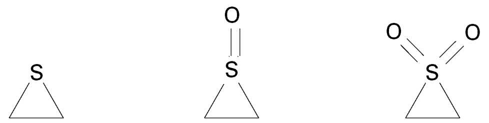
#### The solid state

Under conditions of magic-angle spinning (MAS) the powder band shape of the ($\frac{1}{2}$, $-\frac{1}{2}$) transition of quadrupolar nuclei with non-integral spin does not depend on first-order quadrupolar interactions but only upon second-order quadrupolar effects which do not eliminate the possibility of obtaining meaningful NMR spectra. In this case operation at very high magnetic fields will generally be most advantageous since the maximum second-order quadrupolar line width of the ($\frac{1}{2}$, $-\frac{1}{2}$) transition decreases linearly with the magnetic field. Typical results are those obtained for a range of oxides, oxyanions and high-temperature oxide superconductors.

## $^{33}$S NMR parameters and applications

#### $^{33}$S shieldings

$^{33}$S shieldings cover a range of about 1000 ppm. Four regions (with several exceptions) can be identified: (a) singly bonded sulfur, (b) multiple bonded sulfur, (c) sulfur in delocalized π systems and (d) S=O bonds. The sulfur nucleus in thiols and thioethers is strongly shielded ($\delta$ = –300 to –510 ppm). Oxidation to sulfoxides results in 300 –400 ppm deshielding ($\delta$ = +40 to –100 ppm). Further oxidation to sulfones does not significantly change $\delta(^{33}S)$ values ($\delta$ = +40 to –30 ppm). Sulfur in three- and four-membered ring compounds ([1]–[3]) is more shielded than in the corresponding acyclic and cyclic analogues of larger ring sizes (similar trends have been observed for $^{1}$H, $^{13}$C and $^{17}$O nuclei).



[1] $\delta$ = –574 ppm [2] $\delta$ = –214 ppm [3] $\delta$ = –89 ppm

The resonance range of sulfones is about 60 ppm and it has been used to follow the analysis of sulfur-containing oils after oxidation of the sulfur species. The values of $\delta(^{33}S)$ in sulfonamides are quite similar to those of sulfones. Sulfoximines [R–(O=)S(=NR″)–R′] resonate at lower frequencies than the corresponding sulfones [R–S(=O)$_2$–R′], which may be attributed to the lower electronegativity of the doubly bonded NR group in the sulfoximine replacing the oxygen in sulfones. The opposite is seen when sulfimines [R–S(=NR″)–R′] are compared with the corresponding sulfoxides [R–S(=O)–R′].

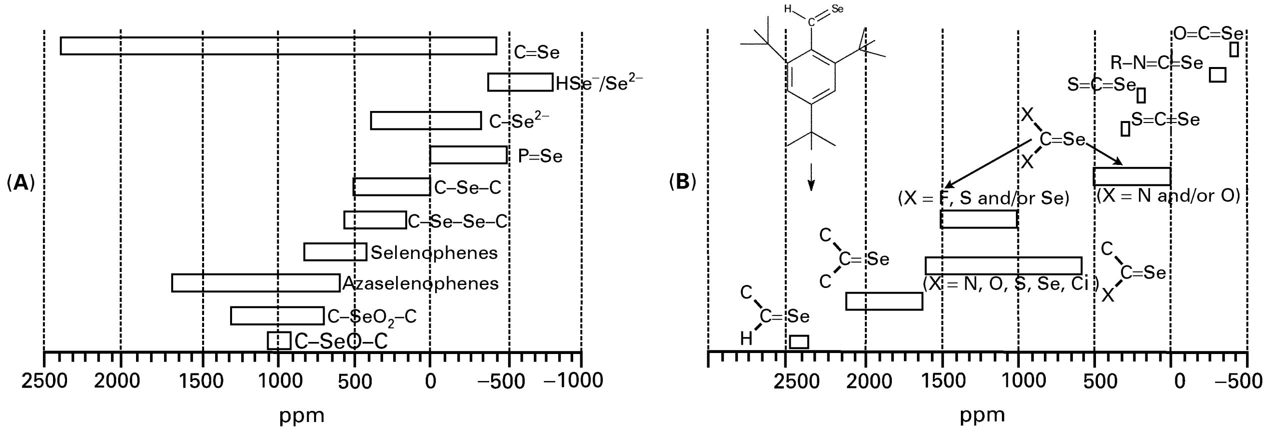
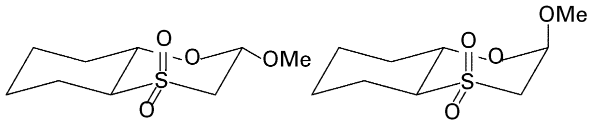



**Figure 2** (A) $^{77}$Se chemical shift ranges of selected selenium functionalities (excluding metal complexes). (B) $^{77}$Se chemical shift ranges of compounds containing C=Se groups (excluding metal complexes). Reproduced by permission of John Wiley and Sons from Duddeck H (1996) Sulfur, selenium and tellurium NMR. In Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, pp 4623–4635. Copyright 1996. John Wiley & Sons.

Sulfuric acid and sulfates resonate around $\delta = 0$. Significant deshieldings have been observed when oxygen is replaced gradually by sulfur in thiomolybdate anions. Thiotungstates show a similar trend. Sulfides cover a wide range, over 600 ppm (e.g. –680 and –42 ppm for $Li_2S$ and BaS, respectively). This can be attributed to different bond characters (ionicities) in the crystal.

Substituent effects in diorganyl sulfones, 3- and *trans*-3,4-substituted thiolane 1,1-dioxides and 4-substituted 2-thiolene 1,1-dioxides have been investigated and correlations with $^{13}$C and $^{17}$O in analogous compounds as well as with Taft (Hammett) electronic parameters have been reported. Stereochemical effects on $^{33}$S shielding have been studied such as the stereomeric sulfones [4] and [5].




[4] $\delta = -10$ ppm [5] $\delta = -20$ ppm

### $^{33}$S Coupling constants

There is a limited number of $^{33}$S coupling constants in the literature due to severely broadened signals. $^2J(^{33}S–^1H)$ coupling constants of 3 and 4.5 Hz were measured for $Me_2SO_2$ and sulfolane and a $^3J(^{33}S–^1H)$ coupling constant of 6 Hz for 2,5-dihydrothiophene 1,1-dioxide. One-bond $^{33}S–^{19}F$ coupling of ~252 Hz has been found for $SF_6$.

### $^{33}$S Relaxation times

The line widths of $^{33}$S cover a range of 10 to 300 Hz for $[MoS_nO_{4-n}]^{2-}$ and $[WS_nO_{4-n}]^{2-}$ ($n$ = 1–4), >4 kHz for sulfoxides and >9 kHz for sulfonium salts. The narrow $^{33}$S signal in solid ZnS can be attributed to cubic symmetry and magnetic dilution of this compound.

## $^{77}$Se NMR parameters and their applications

### $^{77}$Se shieldings

The total range of $^{77}$Se shieldings is about 3300 ppm. **Figure 2** illustrates the resonance ranges of various selenium functional groups. Electropositive substituents such as $H_3Si$ and $H_3Ge$ attached to selenium result in very strong shielding ($\delta = -666$ and –612 ppm, respectively). Selenide ions are even more shielded, e.g. $K_2Se$ ($\delta = -670$ ppm) and $H_3Si–Se^-Li+$ ($\delta = -736$ ppm). Increasing electronegativity of substituents directly attached to selenium in R–Se–X (R-organyl) results in deshielding. In selenium halides ($SeX_2$), strong shielding is observed with decreasing electronegativity of X in the series X=Cl ($\delta = 1758$ ppm), Br ($\delta = 1474$ ppm) and I ($\delta =$ 814 ppm). Deshielding, but with an opposite trend to the halogen effect, is observed for the oxygenated species: $SeOF_2$ ($\delta = 1378$ ppm), $SeOCl_2$ ($\delta =$ 1479 ppm) and $SeOBr_2$ ($\delta = 1559$ ppm).

Alkyl substituents in dialkyl and arylalkyl selenides have a significant deshielding effect on $\delta(^{77}Se)$, with β-effects being larger than α-effects. Analogous trends have been observed for phenylselenenylalkanes (PhSe–R), alkylselenols (RSeH), alkylselenolates ($RSe–Na^+$) and in dialkyl diselenides (RSeSeR). The γ-effects are substantially shielding. The large magnitudes of α-, β- and γ-effects have been discussed in terms of intramolecular polarizability effects. The

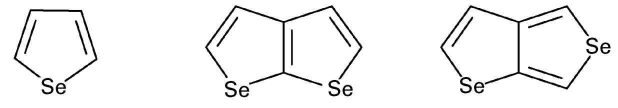
$\delta(^{77}Se)$ of selenoanisoles (Ph–Se–Me) substituted at the aromatic ring have been expressed in terms of dual substituent parameters (DSP).

Oxidation of Se(II) and Se(IV) generally results in significant deshielding: $Me_2Se$ ($\delta$ = 0 ppm), $Me_2SeBr_2$ ($\delta$ = 389 ppm) and $Me_2SeCl_2$ ($\delta$ = 448 ppm). The magnitudes of substituent effects of alkyl groups are approximately four times smaller in the selenoxides than in the corresponding selenides. Halogenated species show the opposite tendency.

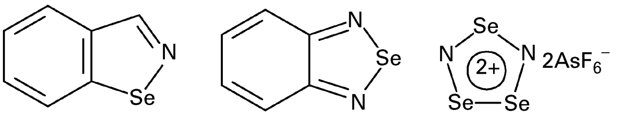
Se(VI) species are shielded compared with the structurally related Se(IV) analogues: $H_2SeO_4$ ($\delta$ = 1001 ppm) and $H_2SeO_3$ ($\delta$ = 1300 ppm). The $\delta(^{77}Se)$ of hypervalent tetraarylselenium compounds are not significantly different compared with those of divalent diaryl selenides.

The C=Se functional group shows shieldings which cover the whole of the known range of $\delta(^{77}Se)$: selenoaldehyde ($\delta$ = 2398 ppm) and O=C=Se ($\delta$ = –447 ppm). Shieldings strongly depend on the nature of the atoms and substituents attached to the carbon of the C=Se group. Furthermore, correlations with $n \rightarrow \pi^*$ transitions and $\delta(^{17}O)$ in C=O analogues have been found.

Cyclic compounds with selenium in the ring(s) have been comprehensively investigated. Annulation of conjugated ring systems to selenophene causes considerable shieldings. Replacement of carbon by an electronegative element such as nitrogen atom results in a deshielding by up to 900 ppm. Even minor strain effects within the heterocycle can cause significant shielding. As expected, positively charged unsaturated ring systems cause extreme deshielding effects.




$\delta$ = 605 ppm   $\delta$ = 579 ppm   $\delta$ = 429 ppm



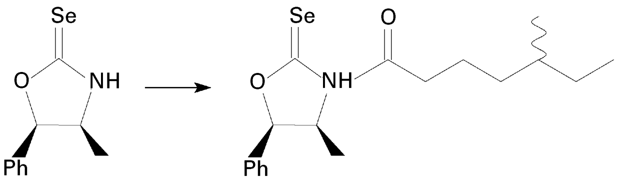

$\delta$ = 1013 ppm   $\delta$ =1511 ppm   $\delta$ = 2434 ppm

P=Se bonds indicate negative $\delta(^{77}Se)$ values which was attributed to the dipolar mesomeric form II: –P=Se (I) $\leftrightarrow$ –$P^+$–$Se^-$ (II). Shieldings depend strongly on the nature of the sustituent at phosphorus and the branching of carbon substituents.

There is a large number of $\delta(^{77}Se)$ of selenium-metal compounds which, according to Duddeck, can be classified into three types: (a) selenium with covalent bonds to main group metals. Deshielding is usually observed compared with the corresponding selenide anions. (b) Selenium σ-, π-, μ- or η-coordinated to transition metal atoms. The $\delta(^{77}Se)$ of these complexes cover the whole shift range. In most cases deshielding is observed compared with the free selenium-containing ligand. Strong shieldings have been observed in the case of the $\eta^4$-cobalt complexes compared with the free heterocycles. (c) Selenium ligands with other atoms (usually phosphorus or nitrogen) directly coordinated to the metal atom. The $\delta(^{77}Se)$ of this type of metal complexes do not differ significantly from those of the free ligands.

Variable temperature NMR has been utilized in the conformational analysis of selenium-containing compounds such as the ring inversion process of phenylselenenylcyclohexane, 2,11-diselena[3.3] metacyclophane, [3.3]diselena- and [4.4]tetraselenacyclophanes and 1,4,5,7-tetrahydro-3*H*-2,6-benzodiselenines.

$^{77}Se$ NMR has been utilized as a tool for chiral recognition which occurs across several bonds such as in the derivatizing agent (4*S*,5*S*)-(–)-4-methyl-5-phenyloxazolidine-2-selone.

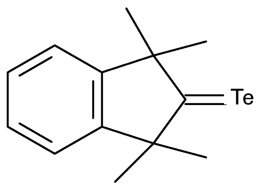



CP MAS $^{77}Se$ NMR in the solid has been utilized to investigate conformational properties in cyclic compounds and comparison with X-ray structures. $\delta(^{77}Se)$ were found to be strongly dependent on the crystallographic non-equivalences in the unit cell.

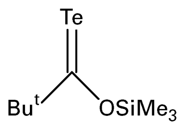
### $^{77}Se$ coupling constants

$^{77}Se$ coupling constants are abundant in the literature since they can be investigated in a straightforward way from the spectra of the coupling partner such as $^1H$, $^{13}C$ and $^{31}P$ nuclei. The *J* values have been reviewed extensively so in this article only a synopsis will be given.

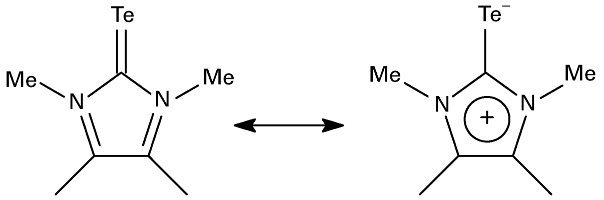
**$^nJ(^{77}Se, ^1H)$** One-bond coupling constants are, presumably, positive with values between 65.4 Hz ($H_2Se$) to 5.3 Hz for Cp–W$(CO_3)$SeH. Geminal couplings are mostly positive and vary between 4 and 16 Hz for hydrogen atoms in aliphatic fragments, but can increase to 50 Hz if the fragment has further geminal heteratoms or if it is in a heteroatom ring. Two- and three-bond (vicinal) coupling constants show some stereochemical dependence.



[6] $\delta$ = 2858 ppm



[7] $\delta$ = 1418 ppm



[8] $\delta$ = –168 ppm

**$^nJ$($^{77}$Se, X where X = $^{13}$C, $^{15}$N, $^{19}$F, $^{29}$Si, $^{31}$P, $^{77}$Se, $^{125}$Te and $^{129}$Xe)** One-bond couplings $^1J$($^{77}$Se, $^{13}$C) are negative and strongly depend on the hybridization of the carbon atom. Typical values are –45 to –100 Hz for Csp$^3$ — Se. Geminal couplings (< 15 Hz for aliphatic carbons) and vicinal couplings (3–12 Hz) are very little investigated and it is not clear whether a Karplus-type dependence exists.

There are very few reports of $^{77}$Se–$^{15}$N couplings (60.1 Hz for $CF_3Se–NH_2$). One-bond $^{77}$Se–$^{19}$F coupling constants considerably exceed 1000 Hz and strongly depend on the stereochemical position of the fluorine atoms (for $SeF_4$ the coupling to the axial fluorine atom is 284 Hz and that to the equatorial one 1206 Hz). Geminal couplings can be substantial (~120 Hz in fluorinated 1,3-diselenetanes) and four-bond coupling constants of ~12 Hz have been observed.

Few data of one bond $^{77}$Se–$^{29}$Si have been published and are generally in the region of 110–49 Hz. One-bond $^{77}$Se–$^{31}$P couplings are negative and show a strong dependence on both the bond order between phosphorus and selenium (–800 to –1200 Hz for P=Se double bonds and –200 to –620 Hz for single bonds) and the stereochemistry of the compounds investigated.

One-bond $^{77}$Se–$^{77}$Se couplings have been observed in a number of diselenides (+22 to –67 Hz) and polyselenide anions with charged selenide atoms (250 ± 20 Hz). Two-bond couplings can have large values, particularly in cyclic sulfur–selenium compounds (95–114 Hz). Three-bond couplings (3–16 Hz) have been observed and four-bond couplings up to 16 Hz have been observed in sulfur–selenium cyclic compounds.

**Couplings to metal nuclei** Typical one-bond couplings to metal nuclei are $^{119}$Sn, 500–1700 Hz; $^{207}$Pb, 780–150 Hz; $^{103}$Rh, 20–44 Hz; $^{113}$Cd, 126–195 Hz; $^{183}$W, 109–14 Hz; $^{195}$Pt, 630–74 Hz (a strong dependence on the nature of the substituents and the stereochemistry has been emphasized) and $^{199}$Hg, –1270 to –751 Hz.

## $^{125}$Te NMR spectral parameters and their applications

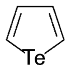
### $^{125}$Te shieldings

The total range of $^{125}$Te shieldings is nearly 5000 ppm. The $\delta$($^{125}$Te) of phosphine tellurides, $R_3PTe$, lie between –513 (R = Me) and –1000 ppm (R = Pr$^i$) and have been interpreted in terms of the character of the P–Te bond.

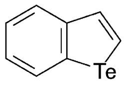
The effects of replacing α-hydrogen atoms of dimethyl telluride and phenylmethyl tellurides or ditellurides are in close analogy to those in the corresponding selenides. The same holds for diorganyltellurium dichlorides and difluorides. Similarly, $\delta$($^{125}$Te) in *ortho*-substituted tellurophenetols can be correlated with analogous selenophenetols.

Telluroketones [6] indicate strong deshielding ($\delta$ = 2858 ppm). If a heteroatom is attached to the tellurocarbonyl group, such as in [7], then a shielding is observed. The unusually high shielding ($\delta$ = –168 ppm) for [8] has been interpreted in terms of the dipolar canonical form.

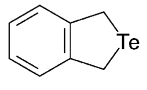
The ($^{125}$Te) of cyclic tellurium compounds indicate similar trends to those with the selenium analogues. Halogenation of the tellurium atom indicates strong deshielding. Furthermore, positively charged saturated and unsaturated ring systems cause extreme deshielding effects.



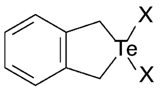
$\delta$ = 782 ppm



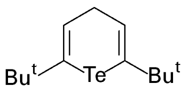
$\delta$ = 727 ppm



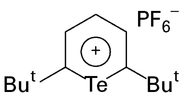
$\delta$ = 269 ppm



X = Br, $\delta$ = 940 ppm

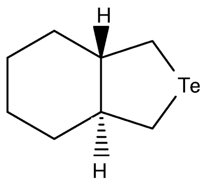


$\delta$ = 257 ppm

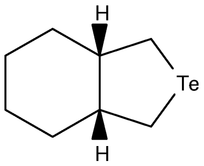


$\delta$ = 1304 ppm

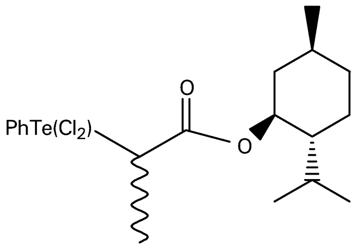
Variable temperature $^{125}$Te NMR has been utilized to investigate ring inversion in phenyltellurenyl-substituted cyclohexane derivatives. $^{125}$Te NMR was shown to be an excellent tool for chiral recognition. Thus, the $^{125}$Te chemical shifts of the stereoisomeric bicyclic compounds [9] and [10] differ by 34 ppm. Interestingly long-range effects, through several bonds, of the menthyl residue in [11] on tellurium shieldings have been observed.




[9] $\delta = 175$ ppm




[10] $\delta = 209$ ppm




[11] $\delta = 989/983$ ppm

Very few CP MAS $^{125}$Te NMR studies in the solid have been reported; however, they clearly demonstrate the great sensitivity of the method in investigating even minor distortions from octahedral crystallographic symmetry.

### $^{125}$Te coupling constants

$^{125}$Te coupling constants are less abundant in the literature than those involving $^{77}$Se and are of the opposite sign due to negative magnetogyric ratio. The ratio $J(^{125}\text{Te,X})/J(^{77}\text{Se,X})$ is 2–3 (in absolute value) and very probably reflects differences in the Fermi contact interaction. One-bond coupling constants $^1J(^{125}\text{Te},^1\text{H})$ are negative (–59 Hz for $H_2Te$), geminal couplings in dialkyltellurides are ~20 Hz and −90 Hz in tellurophene. Vicinal couplings are smaller ($Et_2Te$, –22.7 Hz).

One-bond $^1J(^{125}\text{Te},^{13}\text{C})$ couplings show a strong dependence on the hybridization state of the carbon atom ($Me_2Te$, 162 Hz; $(H_2C{=}CH)_2Te$, 285 Hz; Me–Te–C≡C–Bu, 531 Hz). One-bond $^1J(^{125}\text{Te},^{19}\text{F})$ values are very large ($TeF_6$, 3688 Hz; $TeF_7^-$, 2876 Hz; $Ph_2TeF_2$, 530 Hz). The $^1J(^{125}\text{Te},^{31}\text{P})$ couplings strongly depend on the bond order ($R_3P{=}Te$, 1548–1743 Hz; $(Bu^t)_2P$–Te–$P(Bu^t)_2$, 451 Hz).

$^1J(^{183}\text{W},^{125}\text{Te})$ and $^1J(^{195}\text{Pt},^{125}\text{Te})$ couplings have also been reported.

## List of symbols

$J$ = coupling constant; $T_1$, $T_2$ = relaxation constants; $\chi$ = nuclear quadrupole coupling constant.

*See also:* **High Pressure Studies Using NMR Spectroscopy; NMR in Anisotropic Systems, Theory; NMR of Solids; NMR Relaxation Rates; Solid State NMR, Methods; Solid State NMR Using Quadrupolar Nuclei.**

## Further reading

Barbarella G (1995) Sulfur-33 NMR. *Progress in NMR Spectroscopy* **25**: 1.

Boykin DW (ed) (1991) *$^{17}$O NMR Spectroscopy in Organic Chemistry*. Boston: CRC Press.

Duddeck H (1995) Selenium-77 nuclear magnetic resonance spectroscopy. *Progress in NMR Spectroscopy* **27**: 1–323.

Duddeck H (1996) Sulfur, selenium & tellurium NMR. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, pp 4623–4635. Chichester: Wiley.

Gerothanassis IP (1997) Methods of avoiding the effects of acoustic ringing in pulsed Fourier transform nuclear magnetic resonance spectroscopy. *Progress in NMR Spectroscopy* **19**: 267–329.

Gerothanassis IP (1996) Oxygen-17 NMR. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, pp 3430–3440. Chichester: Wiley.

Kintzinger JP (1981) Oxygen NMR, characteristic parameters and applications. In: Diehl P, Fluck E and Kosfeld R (eds) *NMR, Basic Principles and Progress*, Vol 17, pp 1–64. New York: Springer.

Pearson JG and Oldfield E (1996) Oxygen-17 NMR: applications in biochemistry. In: Grant DM and Harris RK (eds). *Encyclopedia of Nuclear Magnetic Resonance*, pp 3440–3443. Chichester: Wiley.

**Table 1** NMR properties of 3d transition metal nuclei[a]

| | $\nu^b$ (MHz) | Spin | $Q^c$ | $\gamma^d$ | Natural abundance (%) | $r^e$ | Reference compound |
|---|---|---|---|---|---|---|---|
| $^{45}$Sc | 24.330 | 7/2 | −0.22 | +6.5088 | 100 | 0.30 | $[Sc(H_2O)_6]^{3+}$ [h] |
| $^{47}$Ti[f] | 5.643 | 5/2 | +0.29 | −1.5106 | 7.28 | $1.5\times10^{-4}$ | $TiCl_4$[i] |
| $^{49}$Ti[f] | 5.644 | 7/2 | +0.24 | −1.5110 | 5.51 | $1.5\times10^{-4}$ | $TiCl_4$[i] |
| $^{50}$V | 9.988 | 6 | +0.21 | +2.6721 | 0.24 | $1.3\times10^{-4}$ | $VOCl_3$[j] |
| $^{51}$V* | 26.350 | 7/2 | −0.052 | +7.0492 | 99.76 | 0.38 | $VOCl_3$j |
| $^{53}$Cr | 5.636 | 3/2 | −0.15[g] | −1.5007 | 9.55 | $8.6\times10^{-3}$ | $[CrO_4]^{2-}$ [k] |
| $^{55}$Mn | 24.840 | 5/2 | +0.55 | +6.6453 | 100 | 0.175 | $[MnO_4]^-$ [l] |
| $^{57}$Fe | 3.238 | 1/2 | – | +0.8687 | 2.19 | $7.4\times10^{-7}$ | $Fe(CO)_5$[m] |
| $^{59}$Co | 23.727 | 7/2 | +0.42 | +6.3015 | 100 | 0.28 | $[Co(CN)_6]^{3-}$ [n] |
| $^{61}$Ni | 8.936 | 3/2 | +0.16 | −2.3948 | 1.19 | $4.2\times10^{-5}$ | $Ni(CO)_4$ |
| $^{63}$Cu* | 26.515 | 3/2 | −0.21 | +7.1088 | 69.09 | 0.064 | $[Cu\{P(OMe)_3\}_4]^+$ [o] |
| $^{65}$Cu | 28.404 | 3/2 | −0.195 | +7.6104 | 30.91 | 0.035 | $[Cu\{P(OMe)_3\}_4]^+$ [o] |
| $^{67}$Zn | 6.257 | 5/2 | +0.15 | +1.6778 | 4.11 | $1.2\times10^{-4}$ | $[Zn(H_2O)_6]^{2+}$ [p] |

[a] The nucleus usually employed in NMR experiments is denoted by an asterisk.
[b] Measuring frequency at 2.35T (where $^1$H of TMS resonates at 100 MHz).
[c] Electric nuclear quadrupole moment, in $10^{-28}$ m$^2$.
[d] Magnetogyric ratio, in $10^7$ rad s$^{-1}$ T$^{-1}$.
[e] Receptivity relative to $^1$H = 1.
[f] $^{47/49}$Ti NMR signals appear 'paired', separated by 266 ppm.
[g] The quadrupole moment for $^{53}$Cr is still under debate. Values for $|Q|$ as low as 0.04 have also been reported.
[h] Dilute, aqueous solution of $Sc(ClO_4)_3$.
[i] $[TiF_6]^{2-}$ is also used as a standard ($\delta$ = –1163 ppm).
[j] An aqueous, alkaline solution of vanadate ($[VO_4]^{3-}$, $\delta$ = –576 ppm) is sometimes employed as a standard in solution NMR, $Na_3[VO_4]$ in solid state NMR.
[k] Dilute, aqueous solution of $K_2[CrO_4]$.
[l] Saturated aqueous solution of $K[MnO_4]$.
[m] The following standards are frequently employed: $[Fe(CN)_6]^{4-}$ ($\delta$ = +2497ppm), $(\eta^5\text{-}C_5H_5)_2Fe$ dissolved in $CDCl_3$ ($\delta$ = −954.2 ppm).
[n] Alternative standards are: $Co(acetylacetonate)_3$ ($\delta$ = +12 500 ppm), $[Co(NH_3)_6]^{3+}$ (+8150 ppm), $[Co(ethylenediamine)_3]^{3+}$ (+7120 ppm).
[o] Me = $CH_3$. Alternatively (but not recommended) $[Cu(NCMe)_4][ClO_4/BF_4]$ in MeCN ($\delta$ = – 83 ppm) may be used.
[p] The standard commonly used is an aqueous solution of $Zn(ClO_4)_2$.

one has to take care that the signals appear in the correct window (i.e. are not folded). For the quadrupolar nuclei, it is advisable to set the pulse width to 60–90° and to apply a line broadening factor. Relaxation delays can usually be avoided. Except in special cases, $^1$H broad-band decoupling is unnecessary. For very large molecules (metalloproteins), high magnetic field strengths are recommended to cope with second-order shift contributions to the $+\frac{1}{2}\rightarrow-\frac{1}{2}$ quadrupole component. See Individual nuclei and applications, vanadium, below).

Commercially available equipment is used for solid-state static, MAS, OFF-MAS and double rotation techniques, and precautions have to be taken to discriminate between resonance signals and rotation side-bands by measuring the spectra at different spinning frequencies. The information that can be extracted from spectra under the static versus non-static conditions (where quadrupole perturbations, dipole interaction and chemical shift anisotropy effects are suppressed) will be exemplified for solid state Sc NMR spectra below. Note again that the position of the central $+\frac{1}{2}\rightarrow-\frac{1}{2}$ transition which, owing to large quadrupole coupling anisotropy often is the only accessible one, is not coincidental with the isotropic shift $\delta_{iso}$.

**Table 1** also contains chemical shift reference compounds (used as the zero point of the shift scales) and, in the table footnote, standards employed in practice, including their chemical shifts with respect to the reference if not identical. Chemical shifts depend on concentration, isotopic composition (in the first but also the second coordination sphere), nature of the solvent, ionic strength, counterion and other influence of the medium, and these dependences are proportional to the intrinsic shielding sensitivity of a nucleus. Where the shielding sensitivity is exceedingly large ($^{59}$Co), these effects are significant and have to be taken into account when preparing sample and standard solutions. Coupling constants

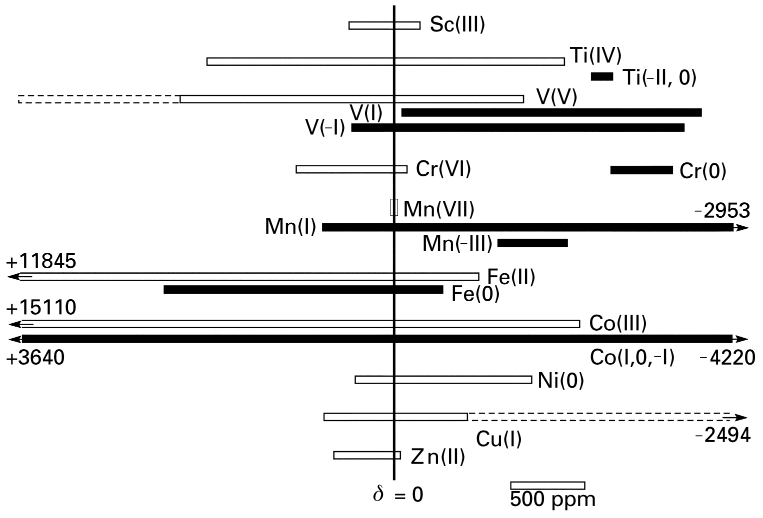



**Figure 1** Chemical shift ranges for the 3d series metals. Open bars correspond to high, solid bars to low oxidation states. Chemical shift values $\delta$ have negative signs to the right of the vertical $\delta = 0$ line, which connects the references given in Table 1. Where the shift range extends beyond the scale of the figure, the limiting $\delta$ values are indicated.

are less dependent on medium effects, but relaxation times can also be very dependent on solution conditions. Other factors influencing shielding (and relaxation times) are temperature and pressure (see below). Chemical shifts should also be corrected with respect to field-lock shifts if different deuterium locks are used for the (external) standard/reference and the sample, and with respect to influences of the bulk susceptibility, if different sample shapes are compared (e.g. cylindrical versus spherical).

## Chemical shifts (nuclear shielding)

**Figure 1** is a summary of chemical shift ranges for the 3d series nuclei in their more common oxidation states in diamagnetic compounds, representing the present state of knowledge but also, to a certain extent, the intrinsic shielding sensitivities. Shift ranges for high and low oxidation states are represented by open and solid bars, respectively. Metals that form closed shell complexes ($d^0$: $Sc^{III}$; $d^{10}$: $Cu^{I}$, $Zn^{II}$) have relatively narrow shift ranges, while open shell systems ($d^4$: $V^{I}$; $d^6$: $V^{-I}$, $Mn^{I}$, $Fe^{II}$, $Co^{III}$; $d^8$: $Fe^0$, $Co^{I}$) give rise to extended shift ranges. Particularly noteworthy is the overall shift range of ~20 000 ppm for $^{59}Co$.

There is a tendency for deshielding of the nucleus (downfield shift with respect to the magnetic field, high-field shift with respect to the RF field) in compounds with high-valency metal nuclei, and for shielding in compounds containing the metal nuclei in a low oxidation state. However, as demonstrated by the large overlapping areas of open and solid bars for the same metal in **Figure 1**, no clear correlation between shielding and oxidation state exists, indicating that the nature of the coordination environment of the metal centre has to be considered. The factors influencing shielding are complex but can be described in a simplified manner by the expression $\sigma = \sigma(\text{dia}) + \text{const.}(\Delta E^{-1})_{\text{av}}\ (C^2 r^{-3})_{\text{av}}$, where $\sigma$ is the overall shielding and $\sigma$(dia) the (essentially constant, since dominated by the core electrons) diamagnetic contribution. The second term, the paramagnetic contribution $\sigma$(para), the main factor influencing variations in shielding (and hence chemical shift), contains the mean HOMO-LUMO splitting $\Delta E$, the LCAO coefficient $C$ of relevant orbitals ($C = 1$ in the crystal field approximation), and a measure for the expansion of the orbitals, $\langle r^{-3}\rangle$. The main contributions to this term arise from the valence d orbitals, but in closed shell systems valence p contributions have to be taken into account and may become predominant. Based on this relationship, the following shielding trends can be predicted and have been verified by experiment.

(1) A strong ligand field decreases $\sigma$(para) via an increase of $\Delta E$ and thus gives rise to high overall shielding. In compounds with a low-valency metal centre, strong ligands are $\pi$ acceptors such as CO, $PR_3$ and CNR; in compounds with a high-valency metal centre, 'hard' ligands such as oxo and nitrogen functional groups induce a large $\Delta E$ splitting. Consequently, the $^{51}V$ nucleus is shielded in $[VO_4]^{3-}$ ($\delta = -540$) as well as in $[V(CO)_6]^-$ (–1952), and deshielded in $[VS_4]^{3-}$ (+1395) as well as in $[V(NO)_2(thf)_4]^+$ (+269 ppm).
(2) Mainly in open shell compounds, shielding is additionally influenced through the $(C^2 r^{-3})_{3d}$ term. This factor becomes small for ligands such

as $H^-$, $R^-$, $I^-$, $SR_2$ or, more generally, for ligands with a high polarizability/low electronegativity. In a series such as $[Co(NO)_2X]_2$, shielding increases in the sequence Cl ($\delta = +3640$), Br (+3390), I (+2790 ppm), known as the ‘normal halogen dependence of chemical shielding’. Corresponding trends (‘normal polarizability dependence’) are commonly found in the series amine, phosphine, arsine and stibine, as well as in the series oxo, thio, seleno and telluro functional ligands. The inverse halogen (polarizability) dependence, i.e. a decrease of shielding (mainly as a consequence of a decrease in $\Delta E$) usually prevails in closed shell $d^0$ systems; cf. $Cp_2TiF_2$ ($\delta = -1052$), $Cp_2TiCl_2$ (–772), $Cp_2TiBr_2$ (–668 ppm).

(3) Bulky ligands and strained ring structures (e.g. 3- and 4-membered chelate rings) induce deshielding in open shell and increase shielding in closed shell systems with respect to sterically normal conditions. This sensitivity of shielding to steric effects often allows for the distinction of conformers and diastereomeric pairs of enantiomers.

(4) Non-innocent ligands (typically catecholates, cat) can give rise to substantial deshielding of the metal nucleus, e.g. in VO(sb)cat (sb is a tridentate Schiff base dianion) $\delta = +480$ ppm instead of an expected $\delta \approx -540$ ppm for a vanadium compound with an ‘innocent’ $N_xO_y$ donor set.

(5) Since the number of relevant transitions between highest occupied and lowest unoccupied molecular orbitals increases with a decrease in local symmetry, structural isomers (e.g. *cis*/*trans*, *fac*/*mer*) exhibit different shielding situations.

### Secondary factors influencing shielding

Apart from effects arising directly from the coordinating functions and the arrangement of the ligands, various second-sphere and outer-sphere effects have to be considered, some of which will be noted here. (i) Ligand substituents, e.g. Z in phosphines $PZ_3$, very effectively influence the steric and electronic nature (donor/acceptor properties) of the ligand. $P(OR)_3$, for example, induces stronger shielding than P(alkyl)$_3$ and $PMe_3$ stronger shielding than $P(Bu^t)_3$. (ii) The shielding variations observed as the counterion or the solvent varies can indicate contact-ion pairing (e.g. in the series $[Co(en)_3]X_3$, X = Cl, Br, I: $\Delta\delta = 32$ ppm) and solvation, respectively. (iii) In the solid state, distinct crystallographic sites are represented by distinct resonance signals. The $\delta_{iso}$(solid) are not usually equal to $\delta_{iso}$(solution).

Isotopic substitution leads to increased shielding for the heavier isotope (e.g. $^{12}CO$ versus $^{13}CO$; $C^{14}N$– versus $C^{15}N$–). This isotope effect can be traced back to the same origin as the temperature and pressure dependence of shielding, i.e. they can be related to the vibrational fine structure of the highest electronic levels participating in electronic transitions: as the temperature decreases, vibronically excited levels become decreasingly populated; $\Delta E$ increases, and overall shielding decreases. An equivalent effect is responsible for increased shielding with increased pressure and an increase of shielding for a heavier with respect to a lighter isotopomer. **Table 2** provides selected examples and **Figure 2** illustrates isotope effects for $C_5H_5V(CO)_4$.

### Anisotropy of chemical shifts

Under anisotropic conditions (solid state, mesophases), the chemical shift consists principally of three components (a parallel ($\delta_\parallel$) and two perpendicular ones ($\delta_1$ and $\delta_2$, where $\delta_\perp = \frac{1}{2}(\delta_1 + \delta_2)$), the separation of which depends on the molecular anisotropy which, for quadrupolar nuclei, correlates with the extent of quadrupole perturbation and hence with the size of the nuclear quadrupole coupling constant $e^2qQ/h$. The extent of anisotropy is expressed by the molecular asymmetry parameter $\eta$, defined by $\eta = (\delta_2 - \delta_1)/\delta_\parallel - \delta_{iso}$), and the chemical shift anisot--ropy $\Delta\delta = |\delta_\parallel - \delta_\perp|$. Selected values for $\delta_{iso}$, $\Delta\delta$ and $\eta$ are listed in **Table 3** together with $e^2qQ/h$ values. The chemical shift anisotropy

**Table 2** Isotope, temperature and pressure effects on shielding

| Isotope | Compound | Detected by | Isotope shift[a] | Temperature gradient[b] | Pressure gradient[c] |
|---|---|---|---|---|---|
| $^{1/2}H$ | $[Sc(H_2O)_6]^{3+}$ | $^{45}Sc$ | –0.52 | | |
| | $C_5H_5V(CO)_4$ | $^{51}V$ | –0.715 | 0.61 | |
| | $[(CpV(CO)_3H]^-$ | $^{51}V$ | –4.7 | | |
| | $[Co(NH_3)_6]^{3+}$ | $^{59}Co$ | –5.2 | 1.55 | |
| $^{12/13}C$ | $[V(CO)_6]^-$ | $^{51}V$ | –0.27 | 0.30 | –0.032 |
| | $[Co(CN)_6]^{3-}$ | $^{59}Co$ | –0.85 | 1.38 | –0.018 |
| $^{14/15}N$ | $[Co(CN)_6]^{3-}$ | $^{59}Co$ | –0.197 | | |
| $^{16/18}O$ | $[V(Co)_6]^-$ | $^{51}V$ | –0.10 | | |
| | $[MnO_4]^-$ | $^{55}Mn$ | –0.60 | | |

[a] The one- and two-bond isotope shifts are indicated per isotopic substitution.
[b] The temperature gradients (tg) are given in ppm per degree, averaged over a temperature ($T$) range of ~230–320 K ($^{51}V$) and 290–350 K ($^{59}Co$); the (tg)/$T$ relation is not strictly linear.
[c] kg cm$^{-2}$, for a pressure range of 1000 to 9000 kg cm$^{-2}$ and at 291 K.

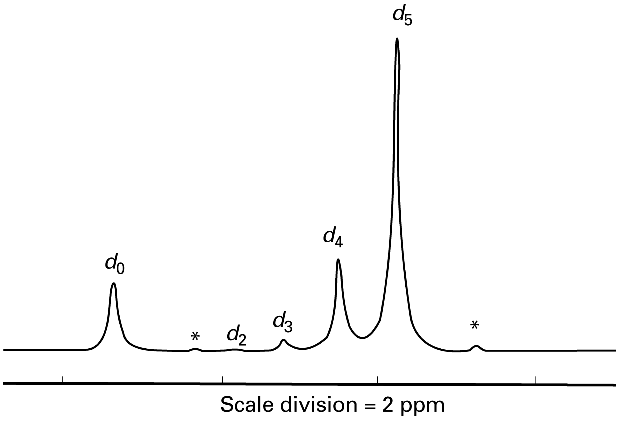



**Figure 2** Two-bond deuterium isotope shifts on the $^{51}V$ chemical shifts in $\eta^5$-$\{C_5{}^1H_{5-n}{}^2H_n\}$ $V(CO)_4$ (assigned $d_n$). The asterisks indicate the high-field components of the natural-abundance $^{13}C$ satellite doublets for $d_0$ and $d_5$. Reproduced with permission from Rehder D, Hoch M Jameson CJ (1990) Deuterium isotope effects and bonding in carbonylvanadium complexes. *Magnetic Resonance Chemistry* **28**: 138–144. © John Wiley & Sons Ltd.

produces an important relaxation mechanism in nonquadrupolar nuclei ($^{57}Fe$; see below).

## Scalar coupling constants

Information on electron-mediated (scalar) spin–spin coupling $J$ may be obtained from the NMR spectrum of the metal nucleus $N_M$ or the ligand nucleus $N_L$. When $N_M$ is the NMR probe, two limiting cases have to be considered. One case is where coupling is fully effective and thus gives rise to a resolved first-order multiplet for the M resonance. This is the case if M is a spin-$\frac{1}{2}$ nucleus ($^{57}Fe$), or if the quadrupole moment of M is small ($^{51}V$) and L is a spin-$\frac{1}{2}$ nucleus or again a quadrupole nucleus with a small quadrupole moment ($^{14}N$). Examples are given in **Figure 3A** and **Figure 3B(ii)**. For the medium quadrupole nuclei, resolved $J$ coupling is restricted to compounds where M is in a local symmetry providing small electrical field gradients at the nucleus and hence less effective quadrupole relaxation. The second case is where coupling is completely quenched. This is the case if fast relaxation leads to relaxation decoupling of $N_M$ and $N_L$, or in the case of intermolecular exchange if it is accompanied by bond rupture. Fast relaxation is encountered with quadrupolar nuclei in all but highly symmetric environments, and with long molecular correlation times $\tau_c$. Bulky and/or large molecules, and a high viscosity of the medium (polar solvents, low temperature, high concentration) will give rise to slow molecular tumbling and hence long $\tau_c$.

If $N_L$ is the NMR probe, we again have to deal with two limiting cases. One is when all of the coupling information is available. This is the case if M belongs to the low quadrupole category nuclei and/or M is at a highly symmetric site. The spectrum appears as a resolved nonbinomial multiplet of

**Table 3** Solid-state chemical shifts $\delta_{iso}$, chemical shift anisotropies $|\Delta\delta|$, asymmetry parameters $\eta$, and nuclear quadrupole coupling constants $e^2qQ/h$

| *Compound* | *Nucleus* | $\delta_{iso}{}^a$ | $\lvert\Delta\delta\rvert$ | $\eta$ | $e^2qQ/h$ *(MHz)* |
|---|---|---|---|---|---|
| $Sc(O_2CMe)_3$ | $^{45}Sc$ | –40 | – | 0 | 5 |
| $\alpha$-$NaVO_3$ | $^{51}V$ | –578 | 365 | 0.70 | 2.95 |
| $VOCl_3$ | $^{51}V$ | ~0 | Not reported | 0.09 | 5.7 |
| $[NH_4]_2[VO_2nta]^b$ | $^{51}V$ | –507 | 800 | 0.04 | 8.80 |
| $CpV(CO)_4$ | $^{51}V$ | −1534 | Not reported | 0.11 | 2.79 |
| $KMnO_4$ | $^{55}Mn$ | ~0 | Not reported | 0.121 | 1.57 |
| $Mn(CO)_5Cl$ | $^{55}Mn$ | ~ −1000 | 1350 | 0.1 | 13.9 |
| $Co(tpp)(im)_2{}^c$ | $^{59}Co$ | +8940 | 2497 | 0.17 | 5 |
| $Co(acac)_3$ | $^{59}Co$ | +12505 | 1029 | 0.38 | 5.6 |
| $FeCo_3(CO)_{12}H$ | $^{59}Co$ | −2833 | 1140 | 0.23 | 12.19 |
| Myoglobin.CNEt$^d$ | $^{57}Fe$ | +9223 | 1288 | – | – |
| CuI$^e$ | $^{63}Cu$ | ~ −410 | Not reported | Not reported | 1.7 |
| $Zn(O_2CMe)_2.2H_2O$ | $^{67}Zn$ | ~ 0 | Not reported | 0.87 | 5.3 |

[a] Referenced against the standards given in **Table 1**.
[b] nta = nitrilotriacetate (3−).
[c] tppp = tetraphenylporphyrin, im = imidazole.
[d] From solution studies.
[e] In $CuBr_xI_{1-x}$ crystals.

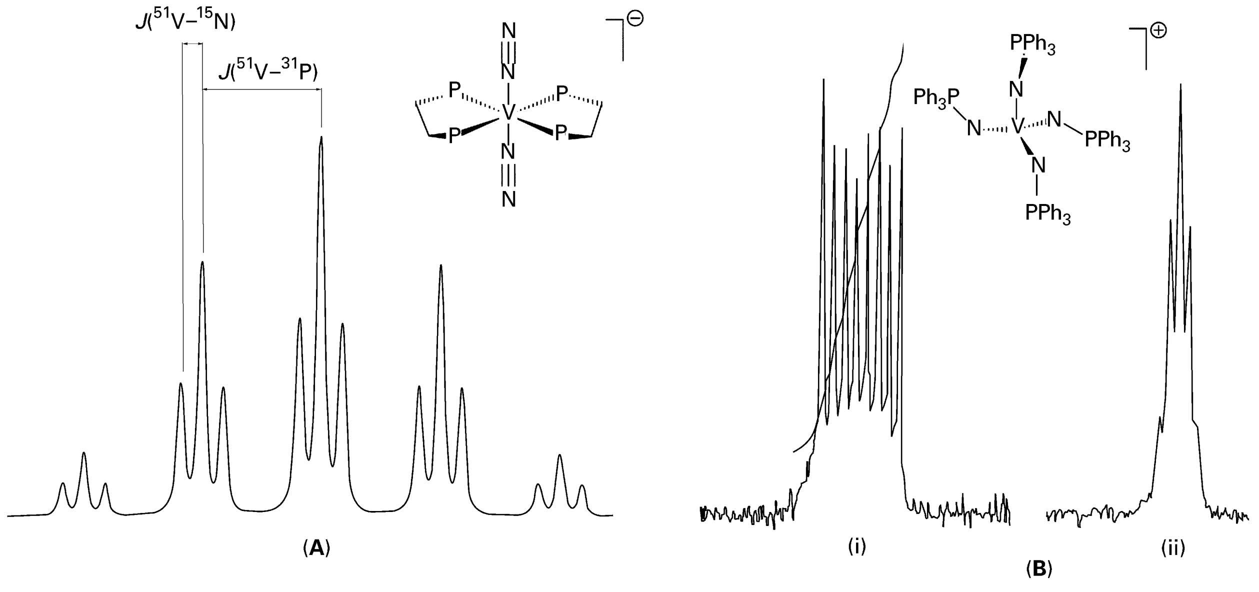



Figure 3 (A) One-bond scalar couplings in the $^{51}V$ NMR spectrum of *trans*-$[V(^{15}N)_2(Me_2PCH_2CH_2PMe_2)_2]^-$: $^1J(^{51}V–^{31}P)$ = 314 (quintet), $^1J(^{51}V–^{15}N)$ = 57 Hz (triplet splitting). Reproduced with permission of the Royal Society of Chemistry from Rehder D, Woitha C, Priebsch W and Gailus H (1992) Structural characterization of a dinitrogenvanadium complex, a functional model for vanadiumnitrogenase. *Chemical Communications* 364–365. (B) Two-bond coupling $^2J(^{51}V–^{31}P)$ = 120 Hz in the $^{31}P$ (i) and $^{51}V$ (ii) NMR spectra of $[V(NPPh_3)_4]^+$. Reprinted with permission from Aistars A, Doedens RJ and Dohesty NM (1994) Synthesis and characterization of vanadium(V) mono-, bis-, tris- and tetrakis(phosphoraniminato) complexes. *Inorganic Chemistry* **33**: 4360–4365. © 1994 American Chemical Society.

multiplicity $2I + 1$ ($I$ is the nuclear spin of M; **Figure 3B(i)**) or as the envelope of such a multiplet, a typical plateau-like signal where the width at half-height divided by $(2I + 1)$ provides a good estimate for $J$. In the solid state $J(N_MN_L)$ multiplets can be distorted owing to the anisotropy of $J$ and the quadrupole coupling interaction. An example is $(Ph_3P)_2CuCl$, where the $^{31}P$ quartet spacings are 1.43, 1.27 and 10.94 kHz. The second limiting case is again complete relaxation decoupling, resulting in a comparatively sharp resonance signal for $N_L$. More or less broad signals with a Lorentzian–Gaussian line shape are typical for intermediate situations.

Coupling constants $J(N_MN_L)$, in most cases one-bond couplings $^1J$, are available for all of the 3d series nuclei except for $^{67}Zn$. They are dominated by the Fermi contact term and the main factors in this term contributing to the size of $J$ are the s electron densities $|S(0)|^2$ at the M and L nuclei, which are susceptible to the metal oxidation state and donor/acceptor properties of the ligands. For a direct comparison of coupling constants of different nuclei, it is advisable to employ the reduced coupling constant $K \propto J/\gamma_M\gamma_L$; $\gamma$ is the magnetogyric ratio, see **Table 1** for the $\gamma_M$ values. **Figure 4** summarizes $^1J/\gamma_M\gamma_L$ ranges for various M–L pairs in relation to the $|S(0)|^2$ terms.

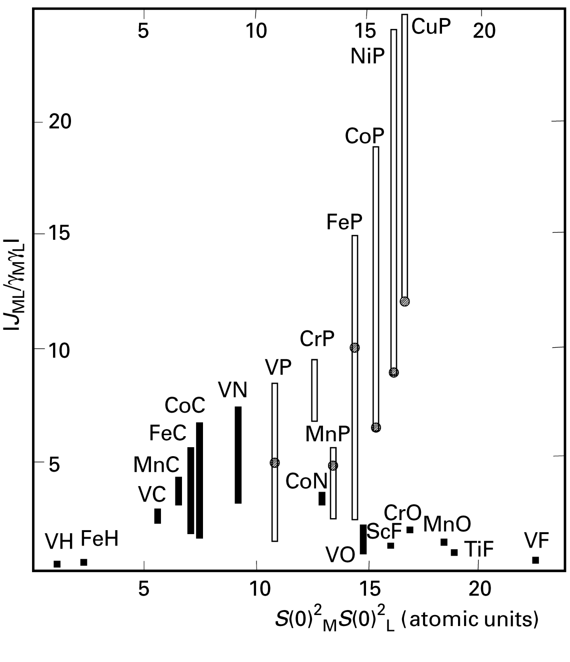



Figure 4 Reduced one-bond M–L coupling constants versus the product of the s electron densities at the respective nuclei (assuming effective zero charge). Open bars represent M–P (phosphine) coupling; the shaded circle represent phosphites $[P(OR)_3]$. $S(0)^2$ values have been taken from Pregosin PS and Kunz RW (1979) *NMR Basic Principles and Progress* **16**: 80–85.

## Relaxation and line widths

The width of a resonance line, usually measured at half-height and quoted as $W_{1/2}$ (or $\Delta\nu_{1/2}$) in units of Hertz, is mainly determined by the spin–spin relaxation time $T_2$ ($W_{1/2} = (\pi T_2)^{-1}$) which, in nonviscous solutions and other than highly symmetric complexes, usually has about the same magnitude as the spin–lattice relaxation time $T_1$. In the solid state, $T_1$ and $T_2$ can deviate from each other considerably. The knowledge of $T_1$ values is a necessary precondition for optimizing NMR detection, especially for nuclei with long $T_1$, viz. $^{57}$Fe. In the absence of quadrupolar ligand nuclei, the main relaxation mechanism in $^{57}$Fe compounds with a nonisotropic coordination environment is the chemical shift anisotropy $\Delta\delta$: $T_1$ is inversely proportional to $(\Delta\delta)^2$, the square of the applied magnetic field $B_0$ and is a function of the correlation time $\tau_c$. $\tau_c$ gains importance for large and/or bulky molecules; compare $T_1$ ($B_0 \approx 9$ T) = 4.2 s for ferrocene, and 0.02 s for myoglobin·CO.

Chemical shift anisotropy has also been noted for cobalt complexes as a $B_0$-dependent component to relaxation which, for quadrupolar nuclei, is otherwise dominated by the quadrupolar relaxation mechanism: $T_2^{-1} \propto f(I)(e^2qQ/h)^2\ (1 + \eta^2/3)\tau_c$, where $f(I)$ is a function of the nuclear spin, and $e^2qQ/h$ and $\eta$ (= 0 under axial symmetry) are as defined above (cf. also **Table 2**). $\tau_c$ describes the intimacy of interaction between solvent and solute molecules and hence the ease by which a solute molecule tumbles in a solution. As noted previously, $T_1$ and $T_2$ are very similar. For example, in aqueous $ZnSO_4$, $^{67}$Zn $T_1$ and $T_2$ are 9.3 and 9.8 ms; in aqueous Co(acac)$_3$ the $^{59}$Co $T_1$ and $T_2$ are both 1.7 ms, and in neat $VOCl_3$ its $^{51}$V $T_1$ and $T_2$ are both 17 ms while the $^{50}$V $T_1$ and $T_2$ are both 5 ms. The latter example ($VOCl_3$) clearly demonstrates the importance of the size of the nuclear quadrupole moment $Q$ (cf. **Table 1**), since all of the other parameters are identical. On the other hand, in a series of similar compounds containing the same metal nucleus (i.e. the same $f(I)$ and $Q$), $T_2$ reflects variations of $\tau_c$ and the electric field gradient $q$ at the nucleus. For solutions, long $T_2$ (narrow resonance lines) can hence be expected for nuclei with a reasonably small $Q$ (e.g. $^{51}$V), for small, nonbulky molecules in nonviscous media (influence via $\tau_c$), for molecules of cubic point symmetry and for a few other cases where $q$ becomes small, e.g. octahedral complexes of $C_{3v}$ symmetry, half-sandwich complexes having local $C_{3v}$ or $C_{4v}$ symmetry, or trigonal-pyramidal complexes of $C_{3v}$ symmetry. The field gradient $q$ also reflects the donor and acceptor power and related properties of the ligands.

Finally, it should be remembered that relaxation is also influenced by chemical exchange.

## Individual nuclei and applications

### Scandium and titanium

Scandium-45 is a relatively easily accessible nucleus owing to its 100% natural abundance and high magnetogyric ratio, but the comparatively large quadrupole $Q$ has restricted its application. In the solid state, the large $Q$ can provide valuable information on the quadrupolar interaction as exemplified for $ScCl_3 \cdot 6H_2O$ and $Sc(O_2CCH_3)_3$ in **Figure 5**. Relevant data can be extracted from static spectra, which depending on the strength of the applied magnetic field, reveal first-order or second-order quadrupole patterns (**Figure 5B**), while under MAS conditions (**Figure 5C**) the main spectral information is the exact isotropic chemical shift.

The concentration, counterion and pH dependences of $\delta(^{45}Sc)$ in aqueous solutions of scandium salts have revealed equilibria between outer-sphere versus inner-sphere complexes (contact ion pairs versus direct coordination of the anion), in addition to the equilibria between $[Sc(H_2O)_6]^{3+}$, which is present at low pH, its deprotonation ($[Sc(H_2O)_5(OH)]^{2+}$) and oligomerization products ($[Sc_2(H_2O)_x(OH)_2]^{4+}$, $[Sc_3(H_2O)_y(OH)_4]^{4+}$), present at pH values up to 4, and $[Sc(OH)_4]^-$, the species that is formed at high pH.

Owing to the very similar magnetogyric ratios, $^{47/49}$Ti, NMR spectra are twinned, the two resonances separated by only 266 ppm. The slightly smaller $Q$ and larger spin of the $^{49}$Ti nucleus make the low-field (high-frequency) twin the better-resolved one. As in the case of $^{45}$Sc, the sizeable quadrupole moments restrict the observability of $^{47/49}$Ti resonances to rather symmetric molecules. $[Ti(CO)_6]^{2-}$ marks the high-field, $TiI_4$ the low-field margin of the Ti chemical shift scale. Mono- and bis(cyclopentadienyl) complexes are in between. In the context of their application in Ziegler–Natta polymerization, a dependence of chemical shift upon the electronic nature of the substituents on cyclopentadienyl (Cp) is noteworthy, an effect which is also observed for $\delta(^{51}V)$ and $\delta(^{57}Fe)$ in Cp-vanadium and Cp-iron complexes.

### Vanadium, chromium and manganese

The nucleus $^{51}$V is the only one among the quadrupolar transition metals with a relatively low quadrupole moment *and* a high receptivity, and is hence a particularly well-suited NMR probe even in

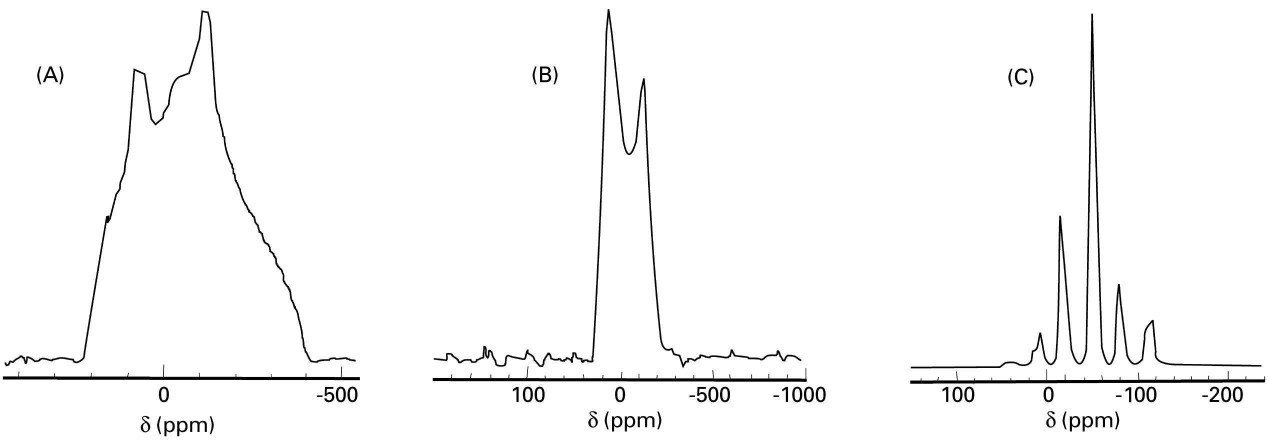



**Figure 5** Solid state $^{45}Sc$ NMR spectra A: Static of $ScCl_3.6H_2O$ at 8.45 T; B: static of $Sc(O_2CCH_3)_3$ at 3.52 T; C: MAS of $Sc(O_2CCH_3)_3$ at 8.45 T, showing the central line and spinning side-bands. Reproduced with permission of the Royal Society of Chemistry from Oldfield E (1987) Solid-state scandium-45, yttrium-89 and lanthanum-139 nuclear magnetic resonance spectroscopy. *Chemical Communications*, 27–29.

compounds with no actual symmetry. Well-resolved coupling patterns, however, require compounds of medium to high local symmetry, as shown in **Figures 2** and **3** for complexes of idealized $C_{4v}$, $D_{4h}$ and $T_d$ symmetries. A large body of $J$ and $\delta$ data has been accumulated (see Further reading for compilations), many of them in the context of investigations into biological/biomimetic vanadium systems, vanadium catalyst systems, correlations between NMR parameters and ligand properties, and vanadium speciation analysis.

More than a dozen species have been detected in aqueous vanadate solutions, and their interconversion and protonation/deprotonation equilibria have been studied by quantitative 1D-NMR, 2D-EXSY, and a combination of $^{51}V$ NMR with potentiometric measurements. The latter procedure has turned out to be a powerful tool for investigating the speciation in ternary systems containing, along with vanadates and protons, a biogenic ligand as a third component. **Figure 6** is an illustrative example for such a system. While the extreme narrowing conditions ($\omega\tau_c << 1$) are fulfilled in these systems, this is no longer so for large $V^V$–protein molecules with a tightly bound vanadium, placing the vanadium outside the extreme narrowing (though still within the motional narrowing) regime. Here, acquisition of spectra is restricted to the quadrupolar central $+\frac{1}{2}\rightarrow-\frac{1}{2}$ transition, which makes up only about 20% of the overall intensity (and is subject to a second-order shift contribution) but, contrasting the other three quadrupole components, does not suffer from severe relaxation broadening. The two slightly different metal ion binding sites in the C- and N-terminal lobes of transferrin have thus been characterized by $^{51}V$ NMR.

The main oxidation states investigated by $^{51}V$ NMR are the diamagnetic –I (carbonylvanadates), +I (half-sandwich complexes) and +V states (e.g. vanadates and other oxovanadium(V) compounds), but other oxidation states such as –III, low-spin +III and binuclear, and strongly antiferromagnetically coupled +IV have also been studied. Irrespective of the oxidation state, the concepts of electronic and steric ligand influences upon $\delta(M)$ outlined above are valid throughout. The steric effect has been employed to distinguish between diastereomers based on pairs of enantiomers in $V^I$ and $V^V$ complexes. Depending on the separation of the chiral centres, the diastereomer splittings amount to ~2 – 10 ppm.

Chemically there are many similarities between V, Cr and Mn in their low-valency diamagnetic states ($d^6$: $V^{-I}$, $Cr^0$, $Mn^{+I}$) and in their highest, closed shell ($d^0$) oxidation states. Where data are available (mainly for Cr-$d^6$ and Mn-$d^6$), dependences of chemical shifts $\delta(^{53}Cr)$ and $\delta(^{55}Mn)$ upon the donor/acceptor properties and sensitivities towards steric effects have been documented that parallel those reported for $\delta(^{51}V)$ in V-$d^6$ and V-$d^4$ complexes.

The relative receptivity of the $^{55}Mn$ nucleus compares with that of $^{51}V$; $Q(^{55}Mn)$ is, however, an order of magnitude larger than $Q(^{51}V)$, limiting the accessibility of $^{55}Mn$ as an NMR probe. The manganese shift range is flanked by $[Mn(CO)_3(NCMe)_3]^+$ at the low-field margin, and the formally $Mn^{-III}$ complexes $Mn(NO)_3L$ at the high-field margin. The only $Mn^{VII}$ compound studied so far by NMR is $KMnO_4$, which gives a very sharp resonance line in solution in accordance with its $T_d$ symmetry but a very broad feature in the solid state as a consequence of effective

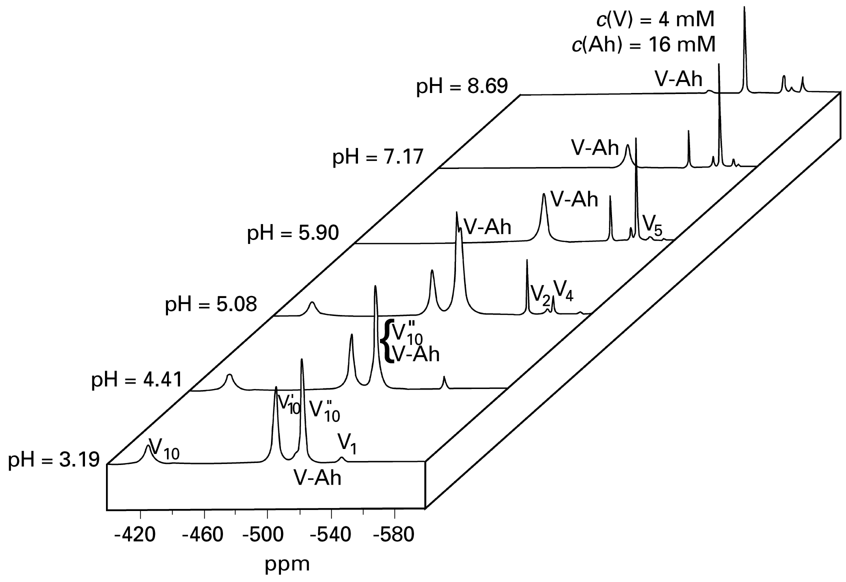



**Figure 6** pH-dependent 131.5 MHz $^{51}$V NMR spectra of the system vanadate/H+/alanylhistidine (Ah). The resonances correspond to decavanadate ($V_{10}$), vanadate-Ah complexes (V-Ah), monovanadate ($V_1$), divanadate ($V_2$), tetravanadate ($V_4$) and pentavanadate ($V_5$) in chemical equilibrium with each other. Reproduced with permission from Elvingsn K, Fritzsche M, Rehder D and Pettersson L (1994) Potentiometric and $^{51}$V NMR study of aqueous equilibria in the $H^+$-vanadate(V)-alanylhistidine system. *Acta Chimica Scandinavica* **48**: 881.

interaction between $Q$ and a crystallographically imposed field gradient $q$. For $|Q(^{53}Cr)|$, reported values range between 0.04 and $0.15 \times 10^{-28}$ m$^2$. NMR experience with this nucleus places it in the medium quadrupole category which, along with the low receptivity, makes $^{53}$Cr a difficult nucleus to detect. The $\delta(^{53}Cr)$ shift range is flanked by $[CrO_3Br]^-$ (+670) and *cis*-$[Cr(CO)_2(PF_3)_4]$ (–1898 ppm).

## Iron, cobalt and nickel

Iron-57 ($I = \frac{1}{2}$) is among the least NMR-sensitive nuclei of the periodic table. Isotopic enrichment or steady-state techniques (in the case of long $T_1$), high magnetic field strengths and large sample volumes may be applied in the case of direct detection, but 1D double resonance and 2D indirect techniques are nowadays much more commonly employed in spectrum acquisition wherever coupling to a sensitive ligand nucleus in the coordination sphere (such as $^1$H, $^{13}$C, $^{15}$N, $^{31}$P) is observable. The chemical shift range spanned by $^{57}$Fe, ~12 000 ppm, is surmounted only by the $^{59}$Co shift range of ~20 000 ppm. In the case of $^{59}$Co, this extended shift range, reflecting a high intrinsic shielding sensitivity, somewhat neutralizes the disadvantage of the relatively large $Q(^{59}Co)$, leading to line widths commonly of several kHz. Nonetheless, abundant data on hexa- and tetracoordinated cobalt complexes in the oxidation states +III, +I and –I are in fact available. In contrast, $^{61}$Ni, again a medium quadrupole category nucleus, but with a very low receptivity, causes severe detection problems. Only recently, through the use of special devices such as solenoidal coils, has a comprehensive series of NMR parameters been obtained on mixed phosphine/carbonylnickel(0) ($d^8$) complexes, which exhibit chemical shift dependences on the nature of the substituents Z in the phosphine $PZ_3$ comparable to what is known from open shell complexes of other transition metals.

Recent applications of $^{57}$Fe and $^{59}$Co NMR have been directed towards correlations between chemical shift and kinetic parameters on the one hand and the investigation of iron–porphyrin and cobalt–porphyrin systems on the other. In the half-sandwich complexes $CpFe(CO)_2R$ ([1] in **Figure 7**) shielding decreases with increasing bulk of R ($\delta(^{57}Fe)$ values cover the range of +684 for R = Me to +805 for R = Bu$^s$), and $\delta(^{57}Fe)$ in turn correlates linearly with the rate constant $k$ for the insertion of CO in the reaction

$$CpFe(CO)_2R + PPh_3 \rightarrow CpFeCO(PPh_3)COR$$

(increase in $k$ as shielding decreases). The sensitivity of shielding to steric effects is also evident in the discrimination of conformers of the complexes (diene) $Fe(CO)_2L$ [2]. The catalytic activity of the complexes Cp′Co(diene) ([3] in **Figure 7**) in the formation of substituted pyridines [4] by co-trimerization of

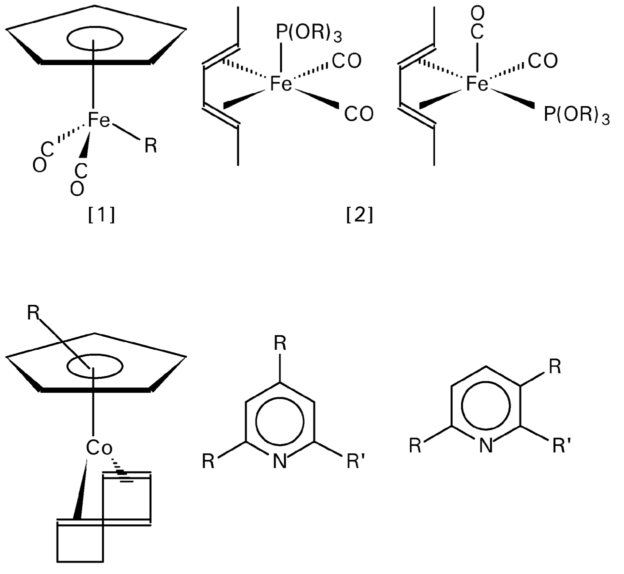



**Figure 7** Applications of $^{57}$Fe and $^{59}$Co NMR. For compounds [1] $\delta(^{57}\text{Fe})$ values correlate with the size of R and the rate of insertion of R into the Fe–CO bond. Compounds [2] are different conformers discernible by their $\delta(^{57}\text{Fe})$ values. Compounds [3] catalyse the formation of [4]; the ratio of [4a] and [4b] depends on the $\delta(^{59}\text{Co})$ values, which are related to the nature of R.

alkynes and isonitriles is closely related to $\delta(^{59}\text{Co})$, which in turn is related to the nature of Cp′: an electron-rich Cp′, such as $C_5Me_5$ gives rise to high shielding of the $^{59}$Co nucleus and induces the preferred formation of the symmetrically substituted pyridine [4a].

The presence in porphyrin and related systems of cobalt (cobalamines) and iron (heme proteins) has initiated metal NMR studies into native (myoglobin, cytochrome *c*) and model systems (substituted porphyrins) with the objective of elucidating the influence of the chemical nature of axial ligands (CO, isonitriles, N-donors) upon parameters such as $\delta$ and $T_1$ (and $e^2qQ/h$ in the case of $^{59}$Co), and hence of the functionally interesting metal centre. The differing binding strengths of isonitriles and carbon monoxide in myoglobins produce a change of the chemical shift anisotropy (CSA), viz. $\Delta\delta = 3600$ ppm for Mb·CO versus 1250 ppm for Mb·CNR, and thus clearly different CSA-induced $T_1$ values (17 ms for Mb·CO, 140 ms for Mb·CNR).

### Copper and zinc

$Cu^+$ and $Zn^{2+}$ are closed shell $d^{10}$ systems and exhibit relatively low intrinsic shielding sensitivities. This fact, together with the unfavourable size of the quadrupole moment, the low receptivity in the case of $^{67}$Zn, and the ease of chemical exchange in the case of Cu complexes has restricted NMR access to these biologically important metal nuclei. The $\delta(^{63/65}\text{Cu})$ range spans ~1000 ppm, with an extension from the high-field margin (–517 ppm, occupied by $[CuI_4]^{3-}$) to –2494 ppm for the cluster compound $[Cu_3Fe_3(CO)_{12}]^{3-}$. $^{67}$Zn chemical shifts encompass about 450 ppm. $Cu^I$ complexes with N-donor ligands such a nitriles and pyridines undergo exchange of the kind $[CuL_4]^+ \rightleftharpoons [CuL_3]^+ + L$, where the lower symmetry component considerably broadens the resonance line. Despite this complication, some insight has been obtained by $^{63}$Cu NMR into the distribution of $Cu^+$ in mixed-solvent systems comprising nitriles and phosphites. A few series of tetrahedral Cu and Zn complexes, viz. $[Cu(\text{phosphite})_4]^+$, $[Cu(ER_2)_4]^+$ (E = S, Se, Te) and $[Zn(ER_2)_4]^{2+}$ (E = S, Se) have been investigated. The latter are of interest in the context of biogenic $\{ZnCys_4\}$ centres present in alcohol dehydrogenase and thioneins.

The $^{63/65}$Cu probe is also increasingly used to elucidate solid-state features of copper-containing superconducting materials.

## List of symbols

$J$ = scalar coupling constant; $q$ = electric field gradient; $Q$ = nuclear quadrupole moment; $r$ = receptivity; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $W_{1/2}$, $\Delta\nu_{1/2}$ = resonance line width [at half-height]; $\gamma$ = magnetogyric ratio; $\delta_{\text{iso}}$ = isotropic chemical shift; $\eta$ = molecular asymmetry parameter; $\sigma$ = shielding; $\sigma$(dia) = diamagnetic contribution to shielding; $\sigma$(para) = paramagnetic contribution shielding; $\tau_c$ = molecular correlation time.

*See also:* **Chemical Exchange Effects in NMR; High Pressure Studies Using NMR Spectroscopy; NMR in Anisotropic Systems, Theory; NMR Relaxation Rates; Parameters in NMR Spectroscopy, Theory of; Solid State NMR Using Quadrupolar Nuclei; Solid State NMR, Methods.**

## Further reading

Brevard C and Granger P (1981) *Handbook of High Resolution Multinuclear NMR*. New York: Wiley.

Gielen M, Willem R and Wrackmeyer B (1996) *Advanced Applications of NMR to Organometallic Chemistry*. Chichester: Wiley.

Howarth OW (1990) Vanadium-51 NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* **22**: 453–485.

Juranic N (1989) Ligand field interpretation of metal NMR chemical shifts in octahedral $d^6$ transition metal complexes. *Coordination Chemistry Reviews* **96**: 253–290.

Laszlo P (1983) *NMR of Newly Accessible Nuclei*. New York: Academic Press.
Mason J (1987) *Multinuclear NMR*. New York: Plenum Press.
Pregosin PS (1991) *Transition Metal Nuclear Magnetic Resonance*. Amsterdam: Elsevier.
Rehder D (1984) NMR of the first transition series nuclei (Sc to Zn). *Magnetic Resonance Reviews* **9**: 125–237.
Rehder D (1986) Applications of transition metal NMR spectroscopy in coordination chemistry. *Chimia* **6**: 186–199.
Rehder D (1991) Metal NMR of organometallic (d-block) systems. *Coordination Chemistry Reviews* **110**: 161–210.

# Heteronuclear NMR Applications (Y–Cd)

**Erkki Kolehmainen**, Department of Chemistry, University of Jyväskylä, Jyväskylä, Finland

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

The elements yttrium (Y), zirconium (Zr), niobium (Nb), molybdenum (Mo), technetium (Tc), ruthenium (Ru), rhodium (Rh), palladium (Pd), silver (Ag) and cadmium (Cd) form a series of transition metals characterized by the electronic configurations $Kr4d^n5s^2$ where $n = 1–10$. This series of d-block metals is located in the 5th row of the periodic table, between the series from Sc to Zn (4th row) and from La to Hg (6th row). The NMR spectroscopy of transition metals is a rapidly increasing area of magnetic resonance with new experimental methods and new applications for both the solution and, especially, for solid state. There is an extensive literature on multinuclear NMR studies applied to the structures of polyoxometalates of many transition metal nuclei, especially Nb and Mo. The NMR data available for the series from Y to Cd depends strongly on the element. A literature search using the *Chemical Abstracts* data base reveals that in case of Y, Zr, Tc, Ru and Pd the number of reports is limited while for $^{113}$Nb, $^{95}$Mo, $^{103}$Rh and $^{113}$Cd there exists much experimental data. Cadmium-113 alone makes up about 20% of the total NMR papers dealing with these d-block metals. However, this situation might change significantly in the future. For example, with the advent of high temperature superconductors one might expect that $^{89}$Y NMR could become a very common method in characterizing yttrium-containing superconducting materials.

Generally, in this article the main aim is to cover novel NMR techniques and promising new results, supplemented with characteristic reference data available in the excellent reviews and articles catalogued in the Further reading section at the end of this article. In selecting and arranging the reference material it has been the goal to demonstrate obvious trends in NMR data (mainly collected in tables of characteristic chemical shifts) that depend on the structural characteristics and measuring conditions: solution vs. solid state, the ligand in a metal complex and its substitution, coordination number, solvent, temperature, etc. From these tables the reader can get a hint as to where to hunt for a ‘hidden’ NMR signal. The text itself, dedicated to each of the NMR-nuclei concerned, serves as a general guide to select an appropriate measuring technique and to estimate if it is worth examining a given NMR-nucleus when compared with the other possible nuclei and techniques available.

## Properties of the nuclei

Some important properties of the NMR active isotopes of Y, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag and Cd are collected in **Table 1**. From the practical point of view the receptivity ($R$) relative to that of $^{13}$C (=1) is one of the most important terms in estimating how easily a nucleus can be observed by NMR. In the case of nuclei with, $I = \frac{1}{2}$ such as $^{89}$Y, $^{103}$Rh, $^{107}$Ag, $^{109}$Ag, $^{111}$Cd and $^{113}$Cd, which are characterized by sharp NMR spectral lines, this parameter can be used as a direct basis in this estimation. However, a nuclear quadrupole moment can produce serious broadening of the spectral lines and hence decrease the signal-to-noise ratio in the NMR spectra. In the case of transition metal nuclei this signal broadening can be the most serious limiting factor from a practical NMR point of view. For example, $^{105}$Pd should be a promising NMR nucleus based on its receptivity

**Table 1** Isotope, spin (*I*), Natural abundance (%NA), frequency ($B_0$) relative to $^1$H (at 100 MHz), receptivity (*R*) relative to $^{13}$C (= 1), magnetogyric ratio (γ), magnetic moment (μ) and nuclear quadrupole moment ($Q_m{}^2$)

| *Isotope* | *I* | *NA (%)* | $B_0$ | *R* | $10^7$ ($\gamma$ rad $T^{-1}s^{-1}$) | *μ* | $10^{28}Q$ ($m^2$) |
|---|---|---|---|---|---|---|---|
| $^{89}$Y | $-\frac{1}{2}$ | 100 | 4.92 | 0.67 | –1.31 | –0.137 | – |
| $^{91}$Zr | $\frac{5}{2}$ | 11.2 | 9.37 | 6.05 | –2.50 | –1.304 | –0.21 |
| $^{93}$Nb | $\frac{9}{2}$ | 100 | 24.55 | 2740 | 6.54 | 6.17 | –0.32 |
| $^{95}$Mo | $\frac{5}{2}$ | 15.9 | 6.52 | 2.88 | 1.74 | –0.914 | –0.015 |
| $^{97}$Mo | $\frac{5}{2}$ | 9.55 | 6.65 | 1.84 | –1.78 | –0.934 | 0.17 |
| $^{99}$Tc | $\frac{9}{2}$ | 0 | 22.5 | – | 6.05 | 5.66 | –0.13 |
| $^{99}$Ru | $\frac{5}{2}$ | 12.7 | 3.39 | 0.82 | –1.23 | –6.41 | 0.076 |
| $^{101}$Ru | $\frac{5}{2}$ | 17.0 | 4.94 | 1.54 | –1.38 | –0.719 | 0.44 |
| $^{103}$Rh | $-\frac{1}{2}$ | 100 | 3.17 | 0.18 | –0.85 | –0.088 | – |
| $^{105}$Pd | $\frac{5}{2}$ | 22.3 | 4.58 | 1.43 | –0.76 | –0.642 | 0.8 |
| $^{107}$Ag | $-\frac{1}{2}$ | 51.8 | 4.05 | 0.20 | –1.08 | –0.113 | – |
| $^{109}$Ag | $-\frac{1}{2}$ | 48.2 | 4.65 | 0.28 | –1.24 | –0.131 | – |
| $^{111}$Cd | $\frac{1}{2}$ | 12.8 | 21.2 | 7.01 | –5.70 | –0.595 | – |
| $^{113}$Cd | $\frac{1}{2}$ | 12.2 | 22.2 | 7.59 | –5.93 | –0.622 | – |

and natural abundance but it suffers from a huge nuclear quadrupole moment such that in, for example, $K_2PdCl_6$ the line width at half-height is 25 kHz. Other limiting factors often met in the early days of NMR, viz. the limited spectral range of the spectrometer hardware, probe and amplifiers, are nowadays not usually problematic. A modern NMR spectrometer can be tuned and matched routinely to observe nuclei from $^{103}$Rh to $^{31}$P, $^{19}$F and $^1$H. This flexibility and the increased sensitivity of NMR spectrometers, e.g. inverse 2D detection techniques, are pushing the detection limits of transition metal nuclei to small concentrations and allowing complex structural characterizations.

## Experimental methods

The variety of measuring techniques and applications in heteronuclear NMR is enormous. In the case of $I = \frac{1}{2}$ nuclei the same basic rules as for $^1$H, $^{19}$F and $^{13}$C are true. However, one should remember that negative gyromagnetic ratios as in case of silver-107/109 and cadmium-111/113 isotopes prevent the signal enhancement by NOE effects. This is different to the case for $^{13}$C NMR. For $I \geq 1$ nuclei the nuclear quadrupole moment opens an efficient relaxation channel and magnetization (free induction) decays often very rapidly. Starting the acquisition as soon as possible after the excitation pulse to avoid the loss of magnetization causes some distortions on the spectral base line (rolling) owing to the acoustic ringing of the probehead. This distortion can be eliminated by ‘an antiringing’ pulse sequence (at the expense of sensitivity) or by using base line correction routines that are now available as a standard in the spectrometer software.

Two-dimensional experimental techniques such as homo- and heteronuclear chemical shift correlations are frequently used in heteronuclear NMR applications. COSY (correlation spectroscopy) and INADEQUATE (incredible natural abundance double quantum transfer experiment) are suitable methods if the natural abundance is not very low, as in the case of $^{113}$Cd. The most promising 2D techniques are, however, inverse detected heteronuclear chemical shift correlations such as $^1$H,X HSQC (heteronuclear single quantum coherence), $^1$H,X HMQC (heteronuclear multiple quantum coherence) and $^1$H,X HMBC (heteronuclear multiple bond correlation) and, especially, $^{31}$P,X HMQC and $^{31}$P,X HMBC. If the scalar coupling constants between a proton (phosphorus or any nucleus of high natural abundance and high gyromagnetic ratio used for free induction decay (FID) detection) and the heteronucleus is small (<2 Hz) the delay in the evolution of these couplings increases to such a length that a significant loss of magnetization occurs and inverse heteronuclear multiple quantum (or bond) correlation becomes impractical. Using the proton for detection of heteronuclear chemical shift correlation can be done easily by a standard two-channel spectrometer equipped with a direct or inverse dual (proton and X) probehead. To observe heteronuclear chemical shift correlation by $^{31}$P or

some other heteronucleus the spectrometer should have two broad band channels and a triple resonance probehead to be able to perform proton decoupling during an experiment if necessary.

Low natural abundances, low gyromagnetic ratios and high quadrupole moments of several NMR-active isotopes of the series from yttrium to cadmium can cause difficulties in solution state NMR. However, these formidable obstacles in the solution state can be turned to advantages in solid state NMR. Low natural abundance (i.e. dilute spins) can help in solid state measurements because the problem raised by dipolar interactions between like spins is eliminated. Sensitivity problems found with low gyromagnetic ratios in the solution state can be overcome in the solid state by the cross-polarization (CP) technique by which the magnetization of the observed nucleus (S) can be increased by a factor, $\gamma_H/\gamma_S$. Further, a nuclear quadrupole moment can result in considerable sensitivity to local nuclear site symmetry, thus giving useful structural information.

The major technique to eliminate dipolar interactions in the solid state NMR is MAS (magic-angle spinning). Together with CP and high-power proton decoupling it has become the most often applied method in solid-state NMR. However, in this connection it is worth mentioning some other techniques such as double-rotation (DOR) and dynamic-angle spinning (DAS) which could remarkably increase the potential of the NMR of quadrupolar nuclei in the future. In DOR, MAS and DAS techniques the sample must be rotated at high speed, which can cause some technical problems. These are avoided by manipulation of nuclear spins by pulse sequences such as two-pulse free induction decay (TPF) and nutation spectroscopy. 2D NMR nutation spectroscopy of powder samples has been applied, for example, in $^{93}$Nb NMR.

## Yttrium-89

Although 89Y is characterized by $I=\frac{1}{2}$, its long $T_1$ relaxation time, low magnetic moment and low measuring frequency have caused problems in $^{89}$Y NMR. On the other hand, its 100% natural abundance and the increasing interest in organoyttrium and solid state chemistry (high-temperature superconducting materials, e.g. $YBa_2Cu_3O_{7-x}$) have created strong interest in this NMR nucleus. Especially promising in this respect is the CP-MAS technique. At first, CP-MAS was applied in the case of the ‘common’ NMR nuclei such as $^{13}$C, $^{27}$Al, $^{29}$Si and $^{31}$P but has now expanded into ‘more difficult’ low $\gamma$-nuclei such as $^{89}$Y. The first solid state $^{89}$Y CP-MAS NMR spectra of $Y(NO)_3\cdot 6H_2O$, $Y_2(SO_4)_3\cdot 8H_2O$, $YCl_3\cdot 6H_2O$, $Y(OAc)_3\cdot 4H_2O$ and $Y(acac)_3\cdot 3H_2O$ were reported by Merwin and Sebald in 1990. Later, mono-, di- and polymetallic samples were characterized by $^{89}$Y CP-MAS NMR, revealing a direct correlation between the number of $^{89}$Y NMR signals and the unique solid state environments of the $^{89}$Y nuclei.

In **Table 2** are collected some characteristic $^{89}$Y NMR chemical shifts. Nowadays, a common reference compound for $^{89}$Y NMR is aqueous $Y(NO_3)_3$ although aqueous $YCl_3$ also has been used. In fact, both salts form, in water, the same hexahydrate cation, $[Y(H_2O)_6]^{3+}$. Generally, the chemical shifts of the yttrium salts show similar concentration trends to those observed for scandium and lanthanum.

## Zirconium-91

The number of papers dealing with $^{91}$Zr NMR is still small. However, some $^{91}$Zr NMR chemical shifts of organometallic compounds (mainly zirconocene derivatives) and inorganic ions are available and are collected in **Table 3**. The line widths, $LW_{1/2}$ of these compounds vary between 19–3100 Hz. The most often used standard is $Cp_2ZrBr_2$ because it shows a sharper resonance line than any other zirconium complex. Von Philipsborn and co-workers (Organisch-Chemisches Institut, Universität Zürich), have used zirconium-91 NMR chemical shifts and line widths as indicators of coordination geometry distortions in zirconocene complexes. Further, the experimental trends in $^{91}$Zr NMR chemical shifts of these zirconocene complexes are well reproduced computationally by molecular orbital calculations using the

**Table 2** Some characteristic $^{89}$Y NMR chemical shifts from $[Y(H_2O)_6]^{3+}$ ($\delta = 0$ ppm)

| *Compound* | *Solvent/conditions* | *$\delta$ ($^{89}$Y)/(ppm)* |
|---|---|---|
| $Y(NO_3)_3$ | 0.77 M in $H_2O-OCMe_2$ at 298 K | − 4.9 |
| $Y(NO_3)_3\cdot 6H_2O$ | Solid | − 53.2 |
| $YCl_3\cdot 6H_2O$ | Solid | +58 |
| $Y(OAc)_3.4H_2O$ | Solid | +46.7 |
| $Y(acac_3)_3.3H_2O$ | Solid | +21.5 and +27.9 |
| $Y_2(SO_4)_3.8H_2O$ | Solid | − 40 and − 46.2 |
| $Y_2O_3$ | Solid | +270 and +315 |
| $Y_2Sn_2O_7$ | Solid | +150 |
| $Cp'_3Y(THF)^a$ | THF | − 371 |
| $Cp'_2YCl(THF)$ | THF | − 103 |
| $[YCl_6]^{3-}$ | Nitrobenzene | +240 |
| $[Y(EDTA)]^-$ | $H_2O$ | +126 |

$^a$Cp′ = $C_5H_4Me$, THF = tetrahydrofuran.

**Table 3** Some characteristic $^{91}Zr$ NMR chemical shifts from $Cp_2ZrBr_2$ ($\delta$ = 0 ppm)

| *Compound* | *Solvent* | *$\delta$ ($^{91}Zr$) (ppm)* |
|---|---|---|
| $Zr(NEt_2)_4$ | – | +875 |
| $Zr(BH_4)_4$ | Benzene-$d_6$ | +41 |
| $H_2ZrCl_6$ | Conc HCl | +601 |
| $[ZrF_6]^{2-}$ | $D_2O$ | –191 |
| $Cp_2ZrCl_2$[a] | – | –122 |
| $Cp_2ZrBr_2$ | THF[b] | 0 |
| $Cp_2ZrI_2$ | – | +126 |
| $Cp^*_2ZrCl_2$[c] | THF[b] | +82 |
| $Cp_2ZrCl$(vinyl) | THF[b] | +16 |
| CpZr($\eta^3$-allyl)(ip)[d] | – | +132 |
| $Cp_2Zr$ (ip) | – | –325 |

[a] Cp = $C_5H_5$.
[b] THF = tetrahydrofuran.
[c] Cp* = $C_5Me_5$.
[d] ip = $\eta^4$-isoprene.

**Table 4** Some characteristic $^{93}Nb$ NMR chemical shifts from $[NbCl_6]^-$ ($\delta$ = 0 ppm)

| *Compound* | *Solvent/temperature (K)* | *$\delta$($^{93}Nb$)/ppm)* |
|---|---|---|
| $Cu_3[NbS_4]$ | Solid | +380 |
| $Cu_3[NbSe_4]$ | Solid | +1110 |
| $Cu_3[NbTe_4]$ | Solid | +2590 |
| $[NbW_5O_{19}]^{3-}$ | – | –888 |
| $NbOCl_3$ | MeCN | –450 |
| $[NbOCl_4]^-$ | MeCN | –480 |
| $[NbOF_4]^-$[a] | MeCN | –1228 |
| $[NbOBr_4]^-$ | MeCN | –210 |
| $[NbSCl_4]^-$ | MeCN | +500 |
| $[NbSeCl_4]^-$ | MeCN | +970 |
| $[NbF_6]^-$ | MeCN | –1550 |
| $[NbCl_6]^-$ | MeCN | 0 |
| $[NbBr_6]^-$ | MeCN | +735 |
| $[HNb(CO)_5]^{2-}$ | $NH_3$/223 | –2122 |
| $[Nb(CO)_6]^-$ | $NH_3$/223 | –2136 |
| $[Nb(CO)_6]^-$ | THF/298 | –2121 |
| $[Nb(CO)_5NH_3]^-$ | $NH_3$/223 | –1880 |

[a] $J$(F, Nb) = 410 Hz.

IGLO (individual gauge for localized orbitals) or GIAO (gauge including AOs) SCF methods employing large basis sets. In the solid state, Zr-metal, $ZrH_2$, ZrC, ZrCo, $ZrO_2$ and $ZrO_2$ doped with $Y_2O_3$ and MgO have been characterized by $^{91}Zr$ NMR. Solid state $^{91}Zr$ NMR line shapes have been found to be related to the Zr (or Nb) content in $Zr_{1-x}Nb_x$ alloys. These applications suggest an increasing interest in $^{91}Zr$ NMR in metallurgy and material sciences.

## Niobium-93

There exist several reports on $^{93}Nb$ NMR, as expected based on a natural abundance of 100% and a receptivity of 2740 (**Table 1**). In fact, the receptivity of $^{93}Nb$ is one of the most promising among all the NMR-active nuclei. However, the influence of broad resonance lines again effectively diminishes its usefulness although its other NMR-properties are favourable. In **Table 4** are collected some characteristic $^{93}Nb$ NMR chemical shifts. In general the chemical shift trends of $^{93}Nb$ are similar to those observed for $^{51}V$ NMR. Recommended reference compounds for niobium-93 are $NbOCl_3$ in acetonitrile or the $[NbF_6]^-$ anion in concentrated aqueous HF, although $[NbCl_6]^-$ in acetonitrile has also been used. In addition to inorganic salts and their anions, and various carbonyl niobium complexes collected in **Table 4** there exist several cyclopentadienyl niobium complexes. The $^{93}Nb$ NMR chemical shifts of these orgnaometallic complexes are clearly shielded, varying between 1290 and 2258 ppm upfield from the reference. Their $LW_{1/2}$ vary from 20 Hz for $CpNb(CO)_4$ in THF–$CH_2Cl_2$ to 10 300 Hz in $[CpNb(AuPPh_3)(CO)_3]^-$.

In addition to the previously mentioned 2D nutation NMR spectroscopy, an interesting new method in solids, first-satellite spectroscopy, applied to $^{93}Nb$ NMR of $LiNbO_3$ (and $^{27}Al$ NMR in $Al_2O_3$) has been reported by McDowell and his co-workers. This technique is experimentally easier and more sensitive than the other methods for studying wide, quadrupolar broadened spectra. Further, the experimental data are considerably simpler and easier to interpret than obtained by the other methods. Examples of the significance of $^{93}Nb$ NMR in materials science include novel solid state studies on ion ordering and ion shifts in relaxor ferroelectrics, structural distortions and intrinsic defects in $LiNbO_3$ and studies on superconducting Nb–Cu multilayers by the field cycling method.

## Molybdenum-95 and -97

Molybdenum has two naturally abundant NMR-active isotopes (**Table 1**). The nuclear properties of the isotope $^{95}Mo$ are more favourable than those of $^{97}Mo$. Malito has reviewed more than 150 original papers on $^{95}Mo$ NMR. The vast majority of those reports are solution state studies whilst very few of them deal with solid state $^{95/97}Mo$ NMR. This

scarcity of solid state $^{95/97}$Mo NMR data is owing to the relatively low NMR receptivity, long probe dead-times and acoustic ringing associated with the pulse-NMR observation of these low-frequency isotopes. Recently, however, it has been demonstrated that solid state MAS $^{95}$Mo NMR studies of organometallic complexes of molybdenum are feasible, particularly if the compounds are known to give relatively narrow $^{95}$Mo NMR spectral lines in solution.

As mentioned before (poly)oxometalates and related structures are traditional topics for multinuclear NMR studies. In the case of molybdenum, there exist systematic experimental NMR and *ab initio* molecular orbital studies on $[MoO_{4-n}X_n]^{2-}$ anions (X = S or Se and $n$ = 0–4). It has been shown that the $^{95}$Mo NMR chemical shifts of these anions can be reliably estimated by such calculations. The paramagnetic shielding term mainly determines the chemical shift variation, ligand induced changes in molybdenum valence d-orbitals correlate with the changes in the metal chemical shift and the excitation energy term, $\Delta E$, in the Karplus–Pople equation is related to the character of the ligand.

Some characteristic $^{95}$Mo NMR chemical shifts are collected in **Table 5**. A common reference for $^{95}$Mo NMR is aqueous (0.2 M) $Na_2MoO_4$ ($\delta$ = 0 ppm). The chemical shift variation of $^{95}$Mo is from +4000 to −2000 ppm and clearly depends on the oxidation state of the metal. $^{95}$Mo NMR line widths can be very broad (>5 kHz). Generally, slower molecular motion (as a consequence of increased relative molecular mass) results in broader resonance lines. This fact can be a limiting factor in studies of molybdenum-containing metalloenzymes and other biopolymers by $^{95}$Mo NMR. A very recent area of biochemical interest is the Mo(VI) complexes of carbohydrates such as alditols and corresponding acids. For example, it has been shown by $^{1}$H, $^{13}$C, $^{17}$O and $^{95}$Mo NMR that D-galacturonic and D-glucuronic acids in their aqueous solutions form complexes with molybdate anions. Although the acids predominately exist in the pyranose forms, their complexes involving the less stable α- and β-furanose anomers as well as the α-pyranose form was detected also. The 2:1 complexes with the α-pyranose forms, insofar as they involve metal binding to the ring oxygen atom, are considered to play an important role in the oxidation of the acids by Mo(VI).

## Technetium-99

Technetium is an artificially produced element used for medical imaging. Reports dealing with $^{99}$Tc NMR are very rare. In addition to unfavourable NMR properties (**Table 1**) a more complex situation regarding practical measurements comes from the fact that $^{99}$Tc is radioactive. Nevertheless, there are some $^{99}$Tc NMR chemical shifts available, as collected in **Table 6**, referenced to the signal of $TcO_4^-$ ($\delta$ = 0 ppm). $^{99}$Tc NMR line widths can vary strongly from < 100 Hz to > 3 kHz.

## Ruthenium-99 and -101

Ruthenium has seven stable isotopes, of which $^{99}$Ru and $^{101}$Ru are NMR active. Although their natural abundances and resonance frequencies are quite similar the smaller quadrupole moment of ruthenium-99 makes the former more favourable from a NMR spectroscopic point of view. The $^{99}$Ru NMR chemical shift range is enormous (> 18 000 ppm). The shifts of ruthenium compounds and complexes follow roughly the same trends as those of iron and osmium. The reference compound in $^{99}$Ru NMR is the $[Ru(CN)_6]^{4-}$ anion ($\delta$ = 0 ppm). Some characteristic $^{99}$Ru NMR chemical shifts are collected in **Table 7**.

Ruthenium complexes with nitrogen-containing ligands show very interesting photochemical properties which can be monitored by $^{99}$Ru NMR. In materials science, solid state $^{99}$Ru and $^{99}$Ru NMR has been recently applied in characterizing the packing of ruthenium metal. This result represents the first detection of Ru NMR in a paramagnetic solid.

## Rhodium-103

Traditionally $^{103}$Rh NMR spectroscopy has been regarded as problematic owing to the very low gyromagnetic ratio and long relaxation time of rhodium-103 although its natural abundance is 100% and $I = \frac{1}{2}$ (**Table 1**). However, by using modern broad band spectrometers rhodium-103 is an easy nucleus to observe. Taking into account the catalytic significance of rhodium compounds, $^{103}$Rh NMR spectroscopy has turned out to be a very powerful method and there exist plenty of experimental data. Some characteristic $^{103}$Rh NMR chemical shifts are collected in **Table 8**. In the case of $^{103}$Rh NMR, several chemical shift references have been used but nowadays *mer*-$[RhCl_3((SCH_3)_2)_3]$ is generally accepted. Owing to the low gyromagnetic ratio of rhodium-103, its detection is now based on polarization transfer techniques such as INEPT (insensitive nuclei enhanced by polarization transfer), DEPT (distortionless enhancement by polarization transfer) or HMQC although earlier INDOR (internuclear double resonance) was also utilized. If a rhodium-bound proton is used for the detection of rhodium-103 an enhancement of $(\gamma_H/\gamma_{Rh})^{5/2}$ = 5635 can be achieved in $^{1}$H, $^{103}$Rh HMQC. If the proton is not directly

**Table 5** Some characteristic $^{95}$Mo NMR chemical shifts from $[MoO_4]^{2-}$ ($\delta$ = 0 ppm)

| *Compound* | *Oxidation state* | *$\delta$ ($^{95}$Mo)/(ppm)* | *Compound* | *Oxidation state* | *$\delta$ ($^{95}$Mo)/(ppm)* |
|---|---|---|---|---|---|
| $Mo(CO)_6$ (in DMF) | 0 | –1850 | $K_4[Mo_2Cl_8]$ | 2 | +3816 |
| $Mo(CO)_6$ (in $CD_3CN$) | 0 | –1855 | $Mo_2[OCH(CH_3)_2]_6$ | 3 | +2444 |
| $Mo(CO)_6$ (in $(CH_3)_2CO$) | 0 | –1856 | $Mo_2[OC(CH_3)_3]_6$ | 3 | +2645 |
| $Mo(CO)_6$ (in $CDCl_3$) | 0 | –1857 | $Mo_2[N(CH_3)_2]_6$ (in toluene) | 3 | +2430 |
| $Mo(CO)_6$ (in $C_6H_6$) | 0 | –1858 | $Mo_2[N(CH_3)_2]_6$ (in hexane) | 3 | +2420 |
| $Mo(CO)_6$ (in $CH_2Cl_2$) | 0 | –1855 to –1858 | $(Cp)_2Mo_2[\mu\text{-}S_2C_2(CF_3)_2]_2$ | 3 | +2301 |
| $Mo(CO)_5CH_3CN$ | 0 | –1440 | $K_4[Mo(CN)_8]$ | 4 | –1309 |
| $Mo(CO)_5Cl^-$ | 0 | –1513 | $[(MoO_3)_2(EDTA)_2]^{4-}$ | 4 | +63 |
| $Mo(CO)_5Br^-$ | 0 | –1540 | $[MoH_2(Cp)_2]$ | 4 | –2507 |
| $Mo(CO)_5I^-$ | 0 | –1660 | $(Cp^*)_2Mo_2(\mu\text{-}S)_2(\mu\text{-}S_2)$ | 4 | +440 |
| $Mo(CO)_4(CH_3CN)_2$ | 0 | –1307 | $(Cp^*)_2Mo_2(\mu\text{-}S)_2(\mu\text{-}Se_2)$ | 4 | +770 |
| $Mo(CO)_3(CH_3CN)_3$ | 0 | –1114 | $[MoOCl_2P(CH_3)_3]$ | 4 | +1890 |
| $Mo_2(CO)_6(Cp)_2$ | 1 | –1856 | $[MoOCl_2(PPh_2CH_3)_3]$ | 4 | +2180 |
| $Mo_2(CO)_6(Cp^*)_2$ | 1 | –1701 | $[Mo_2O_4(EDTA)\cdot H_2O]^{2-}$ | 5 | +609 |
| $Mo_2(CO)_4(Cp)_2$ | 1 | +182 | $[Mo_2O_4(EDTA)\cdot 2H_2O]^{2-}$ | 5 | +982 |
| $Mo_2(CO)_4(Cp^*)_2$ | 1 | +133 | $(Cp^*)_2Mo_2O_2(\mu\text{-}S)_2$ | 5 | –93 |
| $Mo_2(CH_3CO_2)_4$ | 2 | +3702 | $(Cp^*)_2Mo_2S_2(\mu\text{-}S)_2$ | 5 | +478 |
| $Mo_2(CH_3CH_2CH_2CO_2)_4$ | 2 | +3682 | $(Cp^*)_2Mo_2O_2(\mu\text{-}Se)_2$ | 5 | +131 |
| $Mo_2(CH_3CH_2CH_2CH_2CO_2)_4$ | 2 | +3661 | $Na_2MoO_4$ (pH 7) | 6 | –11.71 |
| $Mo(\eta^6\text{-}C_6H_6)_2$ | 2 | –1362 | $Na_2MoO_4$ (pH 9) | 6 | –10.33 |
| $Mo(\eta^6\text{-}C_6H_6CH_3)_2$ | 2 | –1270 | $MoO_2(NCS)_4^{2-}$ | 6 | –155 |
| $Mo(\eta^5\text{-}C_7H_9)(\eta^7\text{-}C_7H_7)$ | 2 | –469 | $MoNCl_3$ | 6 | +952 |
| $Mo(\eta^6\text{-}C_7H_8)_2$ | 2 | +358 | $Mo_2O_5^{2-}$ | 6 | +120 |
| $Mo(Cp)(CO)_3Cl$ | 2 | –836 | Mo + D-glutaric acid | 6 | +45 to +115 |
| $Mo(Cp)(CO)_3Br$ | 2 | –956 | Mo + *threo*-alditol | 6 | +22 |
| $Mo(Cp)(CO)_3I$ | 2 | –1248 | Mo + *erythro*-alditol | 6 | +30 to +34 |

bound to rhodium and $^1$H, $^{103}$Rh HMBC must be used instead of HMQC, the delay required for polarization transfer [generally $J(^1H, ^{103}Rh \leq 2\ Hz)$] is so long that an extensive loss of magnetization occurs. The magnitude of $^1J(^{31}P, ^{103}Rh)$ is generally not less than 80 Hz and therefore $^{31}$P is an excellent nucleus for indirect detection of rhodium NMR resonances although the enhancement by differences in gyromagnetic ratios is not as favourable as in the case of proton. **Figure 1** shows the $^{31}$P, $^{103}$Rh HMQC contour map obtained from a mixture of $[Rh(H)(PPh_3)_4]$ (0.02 M), CySH (0.2 M) and $Ph_3P$ (0.2 M) in 10% pyridine–toluene at –25°C. If the rhodium-103 nucleus is in an unsymmetric environment it can also be detected easily by direct observation but in a symmetric environment this can be rather difficult.

There exist many papers on $^{103}$Rh NMR. For example, small rhodium particles have been characterized by the $^{103}$Rh NMR line broadening and chemical shift changes which have been interpreted in terms of both the electron-orbital (chemical) and electron-spin (Knight) shifts. Further, alkylrhodoximes, $[Rh(Hdmg)_2RL]$, where Hdmg is the monoanion of dimethylglyoxime, L is $H_2O$, pyridine or $PPh_3$ and R is alkyl or an halomethyl, have been extensively studied by $^{103}$Rh NMR spectroscopy. The $^{103}$Rh shielding decreases in the order R = Et > Me

**Table 6** Some characteristic $^{99}$Tc NMR chemical shifts from $[TcO_4]^-$ ($\delta$ = 0 ppm)

| *Compound* | *$\delta$ ($^{99}$Tc)/(ppm)* |
|---|---|
| $Tc(CO)_3ClPPh_3$(*trans*) | –1481 |
| $Tc(CO)_3BrPPh_3$(*trans*) | –1555 |
| $Tc(CO)_3BrPPh_3$(*cis*) | –1460 |
| $Tc(CO)_3(CH_3CN)_3$ | –2853 |
| $Tc(CO)_3(CH_3CN)(PPh_3)_2$ | –3213 |
| $Tc(DMPE)_3^{+a}$ | –13 |

[a] 1,2-Bis(dimethylphosphino)ethane.

**Table 7** Some characteristic $^{99}$Ru NMR chemical shifts from $[Ru(CN)_6]^{4-}$ ($\delta$ = 0 ppm)

| *Compound* | *δ ($^{99}$Ru)(ppm)* |
|---|---|
| Ru($\eta^5$-$C_5H_5$) | −1270 |
| $Ru_3(CO)_{12}$ | −1208 |
| $[Ru(CO)_3Cl_2]_2$ | +1204 |
| $[Ru(CO)_3Cl_3]^-$ | +1090 |
| $[Ru(CO)_2Cl_4]^{2-}$ | +2523 |
| $RuO_4$ | ∼+2000 |
| $[Ru(NH_3)_6]^{2+}$ | ∼+7750 |
| $[Ru(H_2O)_6]^{2+}$ | +16 050 |

**Table 8** Some characteristic $^{103}$Rh NMR chemical shifts to high frequency from Ξ ($^{103}$Rh) = 3.16 MHz

| *Compound* | *δ ($^{103}$Rh)/(ppm)* |
|---|---|
| [($\eta^5$-$C_5H_5$)Rh($\eta^4$-$C_4H_4$)] | −2057 |
| [($\eta^5$-$C_5H_5$)Rh($\eta^4$-$C_8H_8$)] | −348 |
| [($\eta^5$-$C_5H_5$)Rh$(CO)_2$] | −1322 |
| [($\eta^5$-$C_5Me_5$)RhMo$O_4$$]_4$ | +4079 |
| $[Rh_6(CO)_{16}]$ | −426 |
| $[Rh_6(CO)_{15}C]^{2-}$ | −313 |
| $[Rh(CO)_2Cl_2]^-$ | +84 |
| [Rh(acac)$(CO)_2$] | +309 |
| [Rh(cod)Cl$(PPh)_2$] | +409 |
| $[Rh(cod)Cl]_2$ | +1112 |
| [Rh(cod)(acac)] | +1306 |
| [Rh(acac)(*p*-quinone)] | +2065 |
| $[RhCl_6]^{3-}$ | +8075 |
| $[Rh(H_2O)_6]^{3-}$ | +9931 |

> Bu > Pr$^i$ > Bu$^i$ > *neo*-pent > Bu$^t$. In the case of [$(P_2)$Rh(hfa-cac)], where P = bidentate chelating phosphane and hfacac = hexafluoroacetylacetonate, correlations between solid state structures, $^{103}$Rh NMR chemical shifts and catalytic activities ($CO_2$ hydrogenation) have been investigated. $^{103}$Rh NMR chemical shift correlations have also been clarified for rhodium(III) complexes with cyanide and sulfur-donor ligands. A very interesting report describes a correlation between the rhodium-103 shielding in rhodium enamide complexes and the stereoselectivity of dihydrogen ($H_2$) addition to diastereomeric olefin complexes.

## Palladium-105

Palladium belongs to the same group of elements as nickel and platinum. The only report dealing with

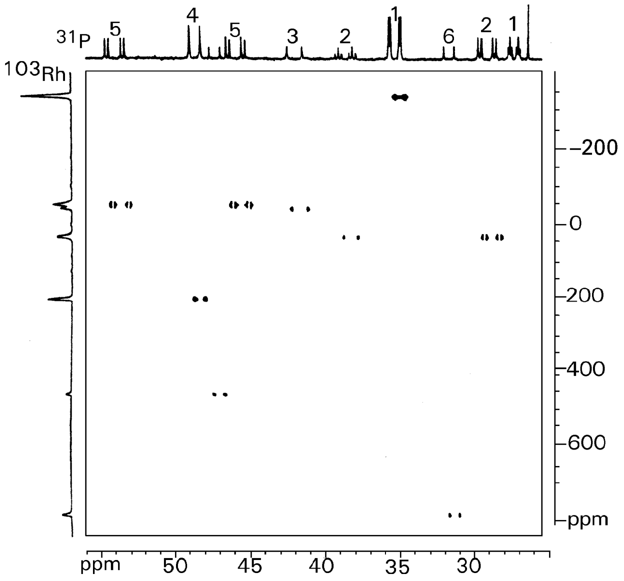



**Figure 1** $^{31}$P, $^{103}$Rh HMQC spectrum obtained from a mixture of [Rh(H)$(PPh_3)_4$] (0.02 M), CySH (0.2 M) and $Ph_3P$ (0.2 M) in 10% pyridine–toluene at −25°C. Reproduced with permission of John Wiley and Sons from Carlton L (1997) Rhodium-103 NMR of carboxylate and thiolate complexes by indirect detection using phosphorus. *Magnetic Resonance in Chemistry* **35**: 153–158. Copyright 1997, John Wiley & Sons.

$^{105}$Pd NMR is on $K_2[PdCl_6]$ for which the $LW_{1/2}$ is 25 kHz.

## Silver-107 and -109

From the NMR spectroscopic point of view, silver suffers from poor sensitivity owing to the low gyromagnetic ratio of both NMR-active isotopes. Further, long relaxation times make it necessary to incorporate long delays within the pulse sequence to avoid saturation effects during acquistion. Sensitivity enhancement by NOE effects is not possible owing to the negative $\gamma$ value of both NMR-active silver nuclei. In spite of these unfavourable nuclear properties, there exist several papers on $^{109}$Ag NMR. Although the natural abundance of the isotope-109 is less than that of 107, the larger $\gamma$ value of 109 gives a better receptivity. As in the case of rhodium-103, direct observation of silver-109 is not useful and various polarization transfer techniques such as INEPT and HMQC are recommended. An example has been published by Berners-Price and co-workers where a retro-INEPT 2D $^{31}$P-{$^{109}$Ag} pulse sequence (published by Bodenhausen and Ruben for $^{15}$N-{$^1$H}) was utilized.

Remarkable progress in solid state $^{109}$Ag NMR has taken place during the last few years. The first $^{109}$Ag CP-MAS NMR spectrum was published by Merwin and Sebald in 1992. Some isotropic solid state $^{109}$Ag

NMR chemical shifts are included in **Table 9** as well as a collection of characteristic solution-state data of silver-109.

## Cadmium-111 and -113

Among the NMR-active isotopes included in this article the most exhaustively studied is certainly $^{113}$Cd. The reasons for this are quite obvious: $I = \frac{1}{2}$ and a relatively good receptivity. Cadmium has also another NMR-active isotope ($^{111}$Cd) characterized with $I = \frac{1}{2}$ but with a somewhat smaller receptivity than $^{113}$Cd. This is why $^{113}$Cd is preferred over $^{111}$Cd. According to a review by Granger, the number of publications dealing with Cd NMR from 1980 to 1987 is ~160. Since then the number of papers on Cd NMR has continuously increased. Therefore, including all the data available here is not possible. Nevertheless, owing to the significance of $^{113}$Cd NMR in biological, inorganic, organic and organometallic chemistry extra attention is given to this NMR-active nucleus. At first sight it looks strange that $^{113}$Cd NMR could possess any biological interest because cadmium is known to be a very toxic element for living organisms. However, the $Cd^{2+}$ cation resembles closely the divalent cations of calcium and zinc which are involved in very many biological complexes. From an NMR spectroscopic point of view these two cations are very impractical. Replacing calcium and zinc by cadmium in these complexes is often easy and opens new versatile NMR spectroscopic possibilities in this area of research. $^{113}$Cd NMR studies in biochemistry and bio-inorganic chemistry have been reported by Sadler and Viles on the binding sites of $Cd^{2+}$ and $Zn^{2+}$ in serum albumin and by Chung and co-workers on the binding of $Cd^{2+}$ in soil fulvic acid.

The direct observation of $^{113}$Cd is generally easy, although in some cases long relaxation times may cause prolonged acquisition times. This can be avoided by using a paramagnetic additive such as $Gd^{3+}$ in the sample. If there exists a coupling between proton and cadmium, polarization transfer techniques such as HMQC are useful, as in the case of rhodium and silver. Owing to the relatively high natural abundance, homonuclear COSY can also be useful alternative in the case of $^{113}$Cd. **Figure 2** shows the DQF (double quantum filtered) $^{113}$Cd–$^{113}$Cd COSY spectrum of cadmium metallothionein 2 (MT-2).

Although the observation of $^{113}$Cd NMR spectral lines is easy and obtainable by a plethora of various measuring techniques, there is an obvious danger of oversimplification of the chemical shift data. This is

**Table 9** Some characteristic $^{109}$Ag NMR chemical shifts from aqueous $AgNO_3$ ($\delta$ = 0 ppm)

| *Compound* | *Solvent/temperature (K)* | *δ ($^{109}$Ag)/ (ppm)* |
|---|---|---|
| $Ag(pyridine)_2^+Ac^-$ | $C_5H_5N$ | +350 |
| $Ag(pyridine)_2^+NO_3^-$ | $C_5H_5N$ | +259 |
| $Ag(C_5H_5)^+NO_3^-$ | $H_2O$/300 | +128 |
| $Ag(CH_3CN)^+NO_3^-$ | $H_2O$/300 | +108 |
| $Ag(DMSO)^+NO_3^-$ | $H_2O$/300 | +94 |
| $Ag(C_6H_6)^+ClO_4^-$ | – | +350 |
| $Ag(NH_3)_2^+NO_3^-$ | Conc. aq. $NH_3$/300 | +593 |
| $[(C_2H_5O)_3P]_4Ag^+SCN^-$ | $CH_2Cl_2$ | +1371 |
| $[Ag(dppe)_2]^+NO_3^-$ | $CDCl_3$/300 | +1338 |
| $Ag[CH_3CHC(OH)CO_2]$ | Solid | +345.9,+320.2 |
| | | +219.7,+210.7 |
| $Ag(CH_3CO_2)$ | Solid | +401.2, +382.7 |
| Ag(*p*-$CH_3C_6H_4SO_3$) | Solid | +44.1 |
| Ag(acac) | Solid | +471.6 |
| $AgN(SO_2Me)_2$ | Solid | +210.3 |
| $AgN(SO_2Me)_2 \cdot \frac{1}{4}H_2O$ | Solid | +32.5 |

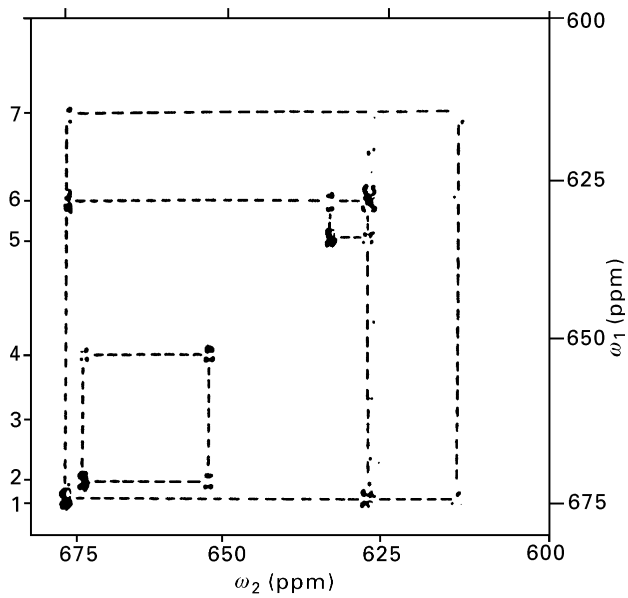



**Figure 2** DQF $^{113}$Cd–$^{113}$Cd COSY spectrum of MT-2. Reproduced with permission of the American Chemical Society from Frey MH, Wagner G, Vašak M, Sørensen OW, Neuhaus D, Wörgötter E, Kägi JHR, Ernst RR and Wüthrich K (1985) Polypeptide-metal cluster connectivities in metallothionein 2 by novel $^1$H–$^{113}$Cd heteronuclear two-dimensional NMR experiments. *Journal of the American Chemical Society* **107**: 6847–6851. Copyright 1985, American Chemical Society.

because chemical exchange often occurs between different complexation or binding sites (including ligand and solvent) for cadmium and the observed Cd NMR line is a time average of the different chemical shifts from these different sites. Therefore it is recommended that $^{113}$Cd NMR studies are carried out at low temperature to 'freeze out' these different species. **Figure 3** shows the $^{113}$Cd NMR spectra of five different Cd(II) complexes of multiple piperazine–pyridine ligands measured in ethanol at −50°C.

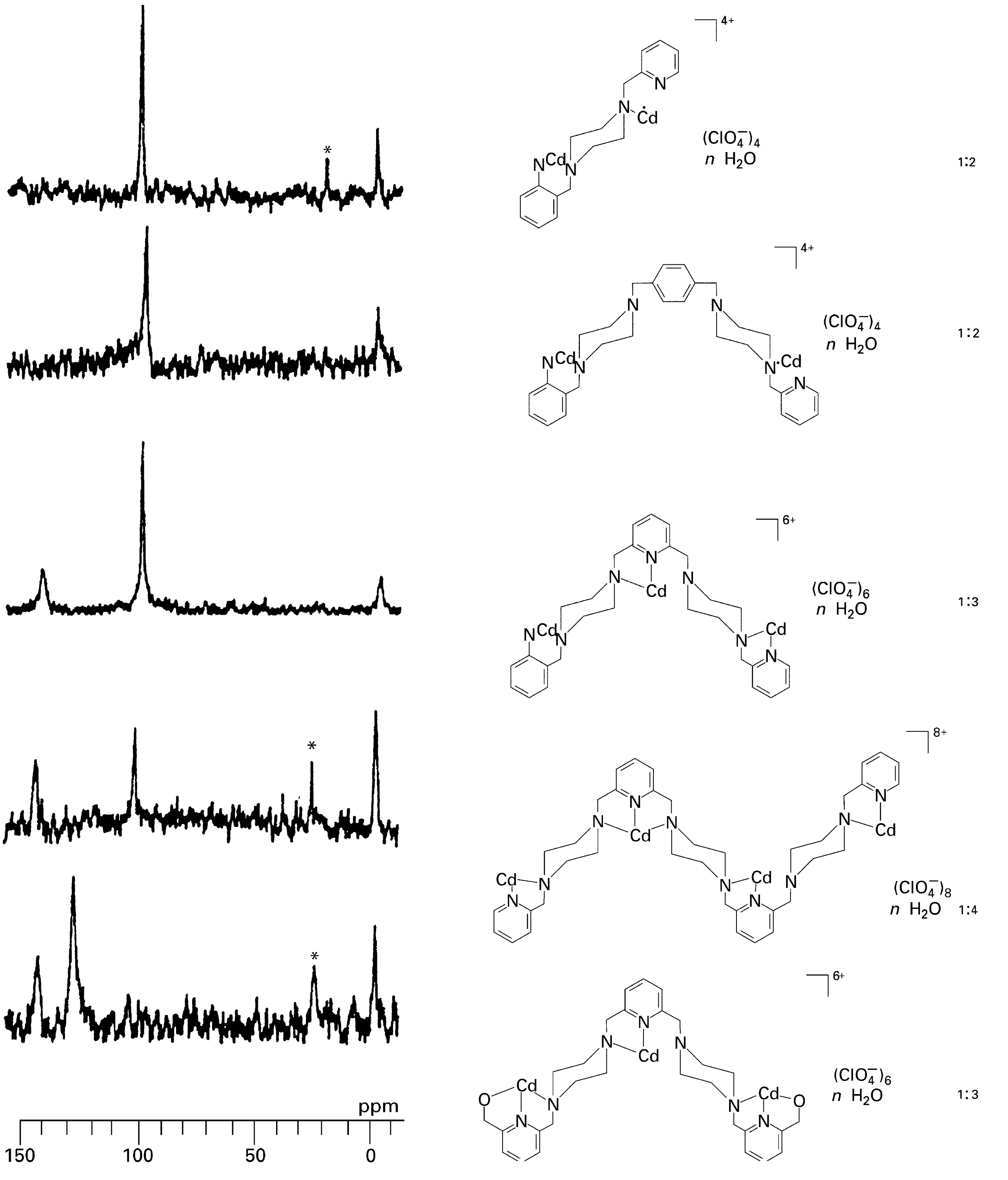



**Figure 3** $^{113}$Cd NMR spectra of the Cd(II) complexes of five different piperazine–pyridine ligands measured in ethanol at −50°C. The signal of the uncomplexed $Cd^{2+}$ cation is marked by an asterisk. Reproduced with permission of VCH Verlagsgesellschaft from Ratilainen J, Airola K, Kolemainen E and Rissanen K (1997) Regioselective complexation of new multiple piperazine/pyridine ligands: differentiation by $^{113}$Cd-NMR spectroscopy. *Chemische Berichte/Recueil* **130**: 1353–1359.

$Cd^{2+}$ cations can form salts and complexes with numerous anions and molecules that bear free electron pairs. For example, in supercooled water at –80°C five different $CdI_n^{2-n}$ ($n = 0$–4) species are observed with the chemical shifts –86, +20, +43, +122 and +101 ppm, respectively. Generally, $^{113}$Cd NMR chemical shifts are referenced to the signal of 0.1 M aqueous $Cd(ClO_4)_2$. For low-temperature measurements this reference is not suitable and 0.1 M $Cd(ClO_4)_2$ in ethanol can be used. In **Table 10** are collected some characteristic $^{113}$Cd NMR chemical shifts. In addition to the monometal compounds and complexes, cadmium can form clusters with several cadmium and/or zinc atoms which also are characterized by $^{113}$Cd NMR.

As noted before, the isotropic $^{113}$Cd NMR chemical shift determined in solution is influenced by exchange processes, temperature, concentration and solvent effects which may limit the interpretation of the structural data. In the solid state, complete information concerning the shielding tensor can be obtained without the difficulties mentioned above. From single-crystal NMR experiments the orientation of the principal elements of the shielding tensor with respect of the coordinate system based on the molecule are obtained. Further, this knowledge provides information on the coordination of cadmium, for example in the active site of metalloproteins such as parvalbumin and concanavalin A. There exist many solid-state NMR studies on cadmium complexes with ligands that contain nitrogen, oxygen and sulfur donor groups. A novel application of solid state $^{113}$Cd NMR is an investigation on host–guest systems. In these studies $^{113}$Cd NMR has been used to differentiate the possible coordination environments.

**Table 10** Some characteristic $^{113}$Cd NMR chemical shifts from 0.1 M aqueous $Cd(ClO_4)_2$ ($\delta = 0$ ppm)

| *Compound* | $\delta$ (*$^{113}$Cd) (ppm)* |
|---|---|
| $[Cd(CH_3S)_4]^{2-}$ | +663 |
| $[Cd(C_2H_5S)_4]^{2-}$ | +648 |
| $[Cd(n\text{–}C_3H_7S)_4]^{2-}$ | +647 |
| $[Cd(n\text{–}C_4H_9S)_4]^{2-}$ | +644 |
| $[Cd(C_6H_5S)_4]^{2-}$ | +590 |
| $[Cd(C_6H_5CH_2S)_4]^{2-}$ | +646 |
| $[Cd(SCH_2CH_2S)_2]^{2-}$ | +829 |
| $[Cd(C_6H_5Se)_4]^{2-}$ | +541 |
| $[Cd(H_3O)_6]^{2+}$ | 0.0 |
| $[Cd(NH_3S)_6]^{2+}$ | +287.4 |
| $[Cd(pyridine)]^{2+}$ | +10 |
| $[Cd(pyridine)_2]^{2+}$ | +40 |
| $[Cd(pyridine)_3]^{2+}$ | +70 |
| $[Cd(pyridine)_4]^{2+}$ | +95 |
| $Cd(CF_3SO_3)_2[P(n\text{-}C_4H_9)_3]$ | +167 |
| $Cd(CF_3SO_3)_2[P(n\text{-}C_4H_9)_3]_2$ | +379 |
| $Cd(CF_3SO_3)_2[P(n\text{-}C_4H_9)_3]_3$ | +505 |
| $Cd(CF_3CO_2)_2[P(n\text{-}C_4H_9)_3]$ | +194 |
| $Cd(CF_3CO_2)_2[P(n\text{-}C_4H_9)_3]_2$ | +431 |
| $Cd(CF_3CO_2)_2[P(n\text{-}C_4H_9)_3]_3$ | +485 |
| $Cd(ClO_4)_2[P(n\text{-}C_4H_9)_3]$ | +163 |
| $Cd(ClO_4)_2[P(n\text{-}C_4H_9)_3]_2$ | +354 |
| $Cd(ClO_4)_2[P(n\text{-}C_4H_9)_3]_3$ | +506 |
| $CdF_2$(solid) | –233 |
| $CdCl_2$(solid) | +211 |
| $CdBr_2$(solid) | +41 |
| $CdI_2$(solid) | –672 |
| $[Cd(S\text{-}2,4,6\text{-}Pr^i_3C_2H_2)_2(bpy)_2]$ | |
| Single crystal: $\sigma_{11}$ | +814 |
| $\sigma_{22}$ | +630 |
| $\sigma_{33}$ | +32 |
| $[Cd(S\text{-}2,4,6\text{-}Pr^i_3C_2H_2)_2(phen)]$ | |
| Solid | +371.5 |
| Solution | +450 |

## Summary

In an article with a limited length and depth of coverage like this, dealing with 10 elements and 14 NMR-active nuclei, it is impossible to discuss all the important NMR parameters. Therefore, most attention has been paid to NMR chemical shifts while spin–spin coupling constants and relaxation times are only mentioned in the contexts where they should be taken into account to obtain useful NMR results. However, as well as direct NMR detection, one should remember, for example, the usefulness of the (inverse) two-dimensional pulse sequences and the MAS technique to utilize homo- and heteronuclear spin–spin coupling constants for polarization transfer between nuclei to enhance the sensitivity of the measurement. These approaches have increased the number of heteronuclear NMR applications dramatically and should play an important role in future studies.

## List of symbols

$I$ = nuclear spin quantum number; $J$ = coupling constant; $R$ = receptivity; $T_1$ = relaxation time; $\gamma$ = gyromagnetic ratio; $\delta$ = chemical shift.

*See also:* **$^{13}$C NMR, Methods; Heteronuclear NMR Applications (B, Al, Ga, In, Tl); Heteronuclear NMR Applications (La–Hg); Heteronuclear NMR Applications (Sc–Zn); Magnetic Resonance, Historical Perspective; MRI Theory; NMR Pulse Sequences; NMR of Solids; NMR Spectrometers; Parameters in NMR Spectroscopy, Theory of; Solid State NMR, Methods.**

## Further reading


Asaro F, Costa G, Dreos R, Pellitzer G and von Philipsborn W (1996) Steric effects in organometallic compounds. A $^{103}$Rh NMR study of alkylrhodoximes. *Journal of Organometallic Chemistry* **513**: 193–200.

Berners-Price SJ, Sadler PJ and Brevard C (1990) Tetrahedral, chelated, silver(I) diphosphine complexes. Rapid measurements of chemical shifts and couplings by two-dimensional $^{31}$P-{$^{109}$Ag} NMR spectroscopy. *Magnetic Resonance in Chemistry* **28**: 145–148.

Bühl M, Hopp G, von Philipsborn W, Beck S, Prosenc M-H, Rief U and Brintzinger H-H (1996) Zirconium-91 chemical shifts and line widths as indicators of coordination geometry distortions in zirconocene complexes. *Organometallics* **15**: 778–785.

Burgstaller A, Ebert H and Voitländer J (1993) NMR properties of hcp Ru metal-first detection of the $^{99}$Ru and $^{101}$Ru NMR in a paramagnetic solid. *Hyperfine Interactions* **80**: 1015–1018.

Carlton L (1997) Rhodium-103 NMR of carboxylate and thiolate complexes by indirect detection using phosphorus. *Magnetic Resonance in Chemistry* **35**: 153–158.

Chung KH, Rhee SW, Shin HS and Moon CH (1996) Probe of cadmium(II) binding on soil fulvic acid investigated by $^{113}$Cd NMR spectroscopy. *Canadian Journal of Chemistry* **74**: 1360–1365.

Davies JA and Dutremez S (1992) Solid state NMR studies of *d*-block and *p*-block metal nuclei: applications to organometallic and coordination chemistry. *Coordination Chemistry Reviews* **114**: 201–247.

Malito J (1997) Molybdenum-95 NMR spectroscopy. *Annual Reports on NMR Spectroscopy* **33**: 151–206.

McDowell A, Conradi MS and Haase J (1996) First-satellite spectroscopy, a new method for quadrupolar spins. *Journal of Magnetic Resonance, Series A* **119**: 211–218.

Merwin LH and Sebald A (1990) The first $^{89}$Y CP-MAS spectra. *Journal of Magnetic Resonance* **88**: 167–171.

Pregosin PS (ed) (1991) *Transition Metal Nuclear Magnetic Resonance, Studies in Inorganic Chemistry*, Vol 13. New York: Elsevier.

Ratilainen J, Airola K, Kolehmainen E and Rissanen K (1997) Regioselective complexation of new multiple piperazine/pyridine ligands: differentiation by $^{113}$Cd NMR spectroscopy. *Chemische Berischte/Recueil* **130**: 1353–1359.

Sadler PJ and Viles JH (1996) $^{1}$H and $^{113}$Cd NMR investigations of $Cd^{2+}$ and $Zn^{2+}$ binding sites on serum albumin: competition with $Ca^{2+}$, $Ni^{2+}$, $Cu^{2+}$ and $Zn^{2+}$. *Inorganic Chemistry* **35**: 4490–4496.


# High Energy Ion Beam Analysis

**Geoff W Grime**, University of Oxford, UK

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| **Methods & Instrumentation** |

The interactions between medium energy (1–5 MeV) atomic ions and the atoms of a sample produce a wide variety of reaction products; electromagnetic radiation from the electron shells or gamma rays from the nucleus, scattered particles from the primary beam which may be modified in direction and energy, or particles of a different species emitted from the nucleus. Any of these may be analysed spectroscopically using suitable detectors to provide information about the elemental and in some cases isotopic composition of the sample.

## Analysis using MeV ions

### Interactions of MeV ions in matter

MeV ions interact with matter through collisions with the electrons and nuclei of the atoms of the material. Ion–electron collisions give rise to reactions associated with the electron shells of atoms (photon emission, Auger electrons, conduction electrons or holes, secondary electron emission) and nuclear collisions give rise to reactions involving the nucleus

(nuclear scattering, atomic displacement, sputtering, nuclear reactions resulting in gamma ray or particle emission). The products of these reactions may be detected and used for analytical purposes.

In nuclear scattering theory, the relative probability of a reaction taking place is expressed as a cross-section, $\sigma$. Each atom or nucleus is represented by a disk of area $\sigma$, and if a beam particle passes through this disk a reaction takes place resulting in the detection of a particle or photon in a detector with 1 steradian (sr) solid angle. Thus high yield reactions have a high cross-section. Values of cross-section are of the order of $10^{-24}$ cm$^2$, and this quantity is called 1 barn. Reaction cross-sections are usually quoted as barn sr$^{-1}$. The yield of a particular reaction may be calculated using the expression

$$Y = \Omega \sigma N n_{\mathrm{T}} \qquad [1]$$

where $Y$ is the number of reaction products (photons, particles) detected in a detector of solid angle $\Omega$ sr when a sample containing $n_{\mathrm{T}}$ target atoms per square centimetre is bombarded with $N$ projectile ions. Here, the cross-section, $\sigma$, is assumed to be in square centimetres.

Cross-sections for ion–atom reactions vary over a wide range depending on the specific reaction, the ion and atom (or nuclear) species, the ion energy and the angle of detection. Clearly for analytical purposes, high cross-section reactions are preferable, but this also depends on the relative intensity of any competing reactions which produce a background signal. In many cases, ion beam analysis (IBA) reactions are free from background so that even low cross-section reactions may be used for analysis.

### PIXE (proton induced X-ray emission)

In PIXE, the basic interaction between the beam and the sample which gives rise to the emission of characteristic X-rays is analogous to the processes involved in electron induced X-ray emission (EDX, or electron probe microanalysis, EPMA). In each case, the charged particle in the beam creates inner electron shell vacancies in the sample atoms, leaving them in an excited state which decays by electron transitions from higher shells, each of which causes the emission of an X-ray quantum whose energy is determined by the atomic number of the atom. This can be detected either using a wavelength dispersive crystal spectrometer (WDX) or an energy dispersive lithium drifted silicon detector (Si-Li) (EDX). The use of WDX detectors is not common in PIXE.

Because of the use of Si-Li detectors, EDX and PIXE share many characteristics, especially the advantage of rapid multielemental analysis, but the use of MeV protons as the incident particle gives one crucial advantage over the technologically simpler use of keV electrons in an electron microscope, and that is in reducing the background of broad spectrum non-characteristic X-rays. In EDX, the major source of background is created by the scattering and deceleration of the electrons of the primary beam as they collide with the electrons in the sample. Each collision gives rise to a large change in momentum of the primary electron and causes the emission of broad spectrum ‘Bremsstrahlung’ radiation which has a maximum energy equal to the beam energy. Bremsstrahlung is the major source of background in EDX and is a fundamental limitation to the minimum detectable limit (MDL).

In contrast, MeV ions lose negligible momentum in each collision with an electron, which means that the primary Bremsstrahlung is effectively absent, permitting the detection of weak characteristic X-ray lines and giving a reduction in MDL of 100 to 1000 times over EDX. PIXE is not totally free from background; the electrons created during the initial ionization generate a background at low energies (secondary electron Bremsstrahlung), which limits the MDL for elements in the range Na to Fe (depending on the primary beam energy), but a substantial improvement (two to three orders of magnitude) over EDX is achieved. **Figure 1** presents a comparison of EDX and PIXE on the same sample of brain tissue showing the improvement in sensitivity resulting from the lack of Bremsstrahlung.

Like EDX, the lightest element that may be detected is determined by the response of the detector to low energy X-rays, and with conventional beryllium window detectors, the lightest element that can be routinely detected is sodium. PIXE has a high cross-section (typically 10–1000 barn sr$^{-1}$) which makes it very well suited for use with microprobe systems.

Each element has a characteristic spectrum of X-ray emission lines with energies depending on the atomic number. These lines are categorized into groups depending on which electron shell was ionized in the collision. For example, the K series consists of lines denoted as $K_{\alpha 1}$, $K_{\alpha 2}$, … $K_{\beta 1}$, $K_{\beta 2}$ … resulting from transitions between the L subshells and the K subshells. The $K_\alpha$ and the $K_\beta$ sublines have energies which are close in value, so that with a detector of moderate energy resolution such as Si-Li, only two peaks are observed. The L series, resulting from transitions between the M subshells and the L shells is more complex with three main groups ($L_\alpha$, $L_\beta$, $L_\gamma$) and several less intense peaks. In heavy elements (above Ba), the M series (N shell to M shell)

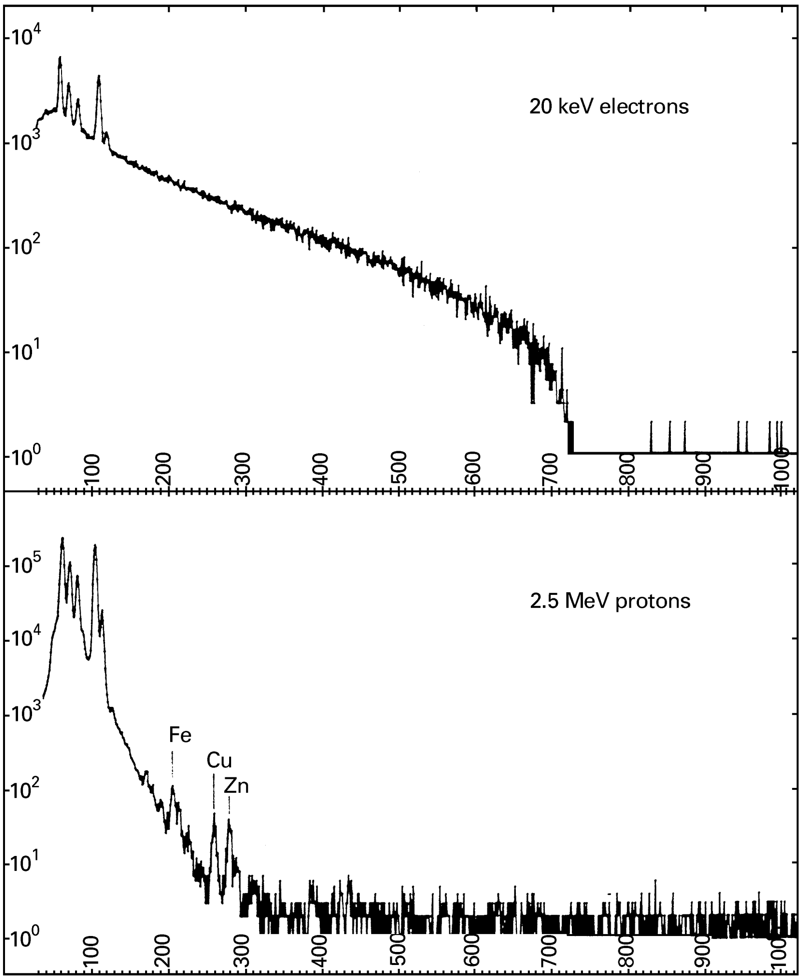



**Figure 1** Energy dispersed X-ray spectra from the same sample excited by 20 keV electrons (top) and 2.5 MeV protons (bottom). The enhancement of the detection limit for the trace elements caused by the absence of primary Bremsstrahlung in the PIXE spectrum can be seen clearly. Reproduced with permission of Wiley from Johansson SAE, Campbell JL and Malmqvist KG (1995) *Particle-Induced X-ray Emission Spectrometry*. New York: Wiley.

may also be observed, but the M lines are not often used for quantification. **Table 1** shows the energy of the most intense K, L and M lines from selected elements, with those detectable using a typical Si-Li detector. The table shows that by using both the K series and the L series all elements heavier than Na may be detected using PIXE. It also highlights the possibility of interference between the series of different elements (e.g. As $K_\alpha$ has almost the same energy as Pb $L_\alpha$). When all the sublines of each series of each element in the sample are taken into consideration, the probability of overlaps is very high, and spectrum processing software must be capable of handling this.

**Figure 2** shows the PIXE spectrum from a single ambient aerosol particle, showing the major and trace elements detected.

**Absorbers** One characteristic of PIXE analysis is the wide dynamic range of the significant peaks. Intense peaks from major elements in the sample may contain millions of counts, while in the low background region of the spectrum, significant peaks containing less than 10 counts may be observed. For this reason, PIXE spectra are normally presented with a logarithmic vertical scale. Another consequence of this is that spurious responses from intense peaks in the detector can have a drastic effect in swamping the small peaks from trace elements. For this reason, the careful use of X-ray absorbers (filters) is crucial to obtaining the best performance with PIXE. For example, when analysing bone, an absorber (typically of Al or a polymer foil) may be fitted to reduce the intensity of the Ca peak to zero or a few parts per thousand to minimize the

**Table 1** Energies (in keV) of the $K_\alpha$, $L_\alpha$, and $M_\alpha$ lines of selected elements. Energies in italic face may be detected using a typical Si-Li detector.

| *Element* | *$K_\alpha$ energy* | *$L_\alpha$ energy* | *$M_\alpha$ energy* |
|---|---|---|---|
| Li | 0.052 | | |
| O | 0.523 | | |
| Na | *1.041* | | |
| Si | *1.740* | | |
| Ca | *3.691* | 0.341 | |
| Fe | *6.403* | 0.704 | |
| As | *10.534* | *1.282* | |
| Zr | *15.774* | *2.042* | 0.350 |
| Ag | *22.162* | *2.984* | 0.568 |
| Ba | 32.191 | *4.467* | 0.972 |
| W | 59.310 | *8.396* | *1.774* |
| Pb | 74.957 | *10.549* | 2.342 |
| Pu | 103.653 | *14.297* | *3.350* |

background and spurious peaks due to the intense flux of Ca X-rays. This is crucial in improving the detection limits for the metals. Fitting an absorber means that X-ray lines of lower energy than the line being removed are lost completely, and it is becoming more common now to use two detectors simultaneously, one unfiltered for major elements and one with a suitable filter for the trace elements.

The art of PIXE analysis lies in the selection of the optimum filter.

**Quantification of PIXE spectra** **Figure 2** shows that a PIXE spectrum consists of peaks which are approximately Gaussian sitting on top of a continuous background. The background-subtracted area of each peak is proportional to the concentration of the associated element, and a number of software packages are now available to perform the processing required to extract the areas and convert these to true concentrations.

For a sample consisting of a thin film (so that proton energy loss and X-ray absorption can be neglected), the yield of characteristic X-ray photons of energy $E_x$ from element $Z$ induced by particles of energy $E$ is given by

$$N = C_z Q \Omega \varepsilon(E_x) Y(Z, E) \qquad [2]$$

where $N$ is the total number of photons detected using a detector of solid angle $\Omega$, $Q$ is the total beam charge, $C_z$ is the concentration of the element and $\varepsilon(E_x)$ is the dependence of detector efficiency on X-ray energy. $Y(Z, E)$ is the yield of the interaction expressed as counts per unit of concentration per unit of charge per unit of solid angle. This, in turn, is derived from the ionization cross-section (the probability of creating a vacancy) and the fluorescence yield (the number of photons emitted in each X-ray line for each vacancy).

In more realistic thick samples, the calculation must also take account of the energy loss of the particle as it penetrates the sample and also the absorption of the X-rays as they emerge from the sample. To do this, a knowledge of the bulk composition of the sample is required (together with any variation with depth) and both the stopping power of the ion in the matrix and the X-ray attenuation coefficient must be known. The calculation then involves a numerical integration of the total X-ray yield from each sublayer of the sample.

All of the physical parameters involved in this calculation (ionization cross-section, fluorescence yield, stopping power, X-ray attenuation and detector efficiency) have been measured and parameterized with sufficient accuracy that quantitative PIXE analysis may be carried out without routine reference to standards once the fixed system parameters have been determined. A small number of computer

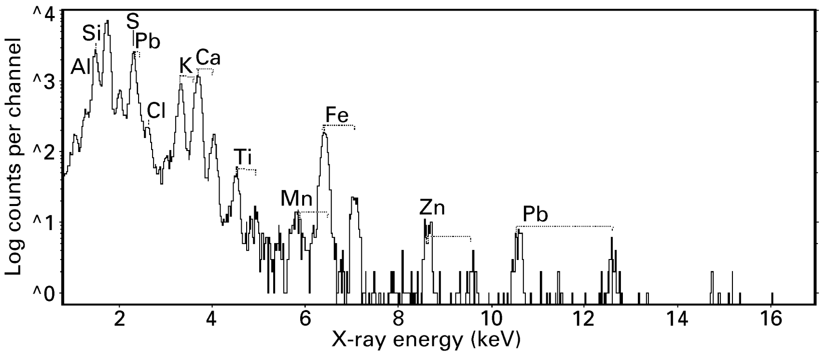



**Figure 2** Proton-induced X-ray spectrum from a single ambient aerosol particle of 2 μm diameter. The spectrum was collected in 5 min using a beam of 3 MeV protons focused to a diameter of 1 μm at a current of 100 pA.

programs have been developed to carry out PIXE spectrum processing, and special mention must be made of GUPIX, developed at the University of Guelph, Ontario, which embodies the most detailed physical model of the PIXE process and also offers the capability to iterate the sample matrix composition (including user-defined 'invisible' elements such as oxygen) until the matrix composition matches the values obtained from the peak areas.

**Accuracy and detection limit of PIXE** The major sources of error in a PIXE measurement and the above calculation are as follows:

(1) *Inaccuracies in the basic assumptions about the target*. The calculation assumes that the sample is a flat homogenous slab of the given composition. Any departure from this model (surface roughness, compositional variation with depth, etc.) will introduce an error in the calculated relative intensity of the X-ray lines.
(2) *Inaccuracies in the physics database*. These may contribute errors to the relative X-ray intensities, but it is believed that the database is of sufficient accuracy that these are relatively insignificant compared with other sources of error.
(3) *Numerical and statistical errors in processing the spectrum*. The quality of the final fit depends on the fitting procedure and the amount of detail in the mathematical model. The precision of each peak determination also depends on the quality of the statistics in the experimental spectrum. Peaks with few counts will not be fitted so well as intense peaks.
(4) *Limitations in the model used for the detector response function and absorber transmission*. The response of the detector varies with X-ray energy (and also depends on any absorbers used). This can be simulated using a simple physical model of the detector and parameterized expressions for X-ray absorption coefficient. However, for the highest accuracy it is necessary to analyse standard samples and derive a correction factor for each X-ray line of interest.
(5) *Inaccuracies in determining the charge and detector solid angle*. These affect the absolute accuracy but not the relative concentrations. For this reason it is more accurate to use PIXE to determine element ratios rather than absolute values. The charge is difficult to measure accurately since the beam currents involved are generally quite small and also large numbers of secondary electrons are created in the beam impact on the sample, which, if they are not returned to the sample, will give an error in the measurement. Beam current normalization is often carried out by using PIXE simultaneously with another technique (e.g. Rutherford backscattering (RBS), which also offers the possibility of determining the sample matrix independently).

When all these sources of error have been addressed, it should be possible to achieve a relative accuracy of the order of 5–10% and an absolute accuracy of 10–20%.

The detection limit depends on the height of the peak relative to the background (which now can mean both the continuum background and also any overlapping X-ray lines or detector artefacts). The detection limit is conventionally defined as 3 times the square root of the area of the background within one standard deviation of the peak centre (converted to concentration using the same factor as for the peak). This depends on all the parameters of the experiment, and **Figure 3** shows the theoretical detection limits for trace elements in a carbon matrix. This indicates that values of the order of 0.1–1 ppm may be expected under optimum conditions.

## RBS (Rutherford backscattering)

Particle backscattering spectrometry, also known as Rutherford backscattering or RBS involves measuring the recoil energy of backscattered particles from the primary beam to obtain analytical information. Ions recoiling from direct elastic collisions with the nuclei of atoms in the sample lose energy according to the mass of the target atom; recoils from heavy nuclei have a higher energy than recoils from light nuclei. If the target atom is not at the surface, energy is also lost during the passage through the sample to and from the reaction site (**Figure 4**), so measuring the energy distribution of recoiling atoms gives

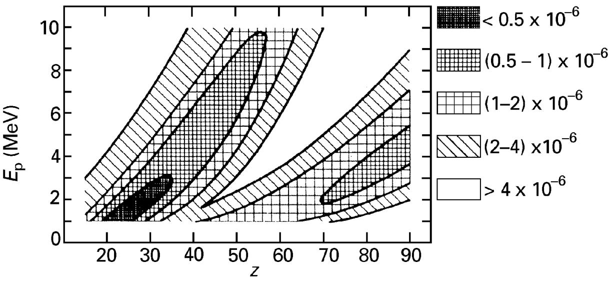



**Figure 3** Minimum detectable limit of trace elements in a carbon matrix as a function of atomic number and proton energy. The calculation assumes a beam charge of 1 μC and curves are presented for K and L lines. Reproduced with permission of Wiley from Johansson SAE, Campbell JL and Malmqvist KG (1995) *Particle-Induced X-ray Emission Spectrometry*. New York: Wiley.

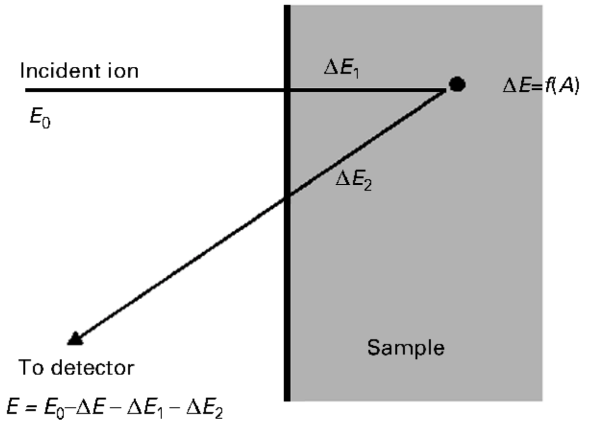



**Figure 4** Typical geometry for Rutherford backscattering. The incident particle loses energy $\Delta E_1$ going through the sample to the scattering site, loses energy $\Delta E$ (determined by the Rutherford formula) in the collision and loses energy $\Delta E_2$ on the recoil path to the detector.

information on the major element composition and depth distribution in the sample. This technique is loosely complementary to PIXE in that the greatest mass resolution occurs for the light elements which are invisible using PIXE. The cross-sections are lower than for PIXE (typically 0.1–10 barn $sr^{-1}$) therefore the minimum detectable limit is not so low as for PIXE (typically 0.1%), but RBS can be used for determining major element composition and depth profiles. In certain favourable cases (e.g. heavy inclusions embedded in a light matrix) RBS can be used to carry out nondestructive three-dimensional mapping. Used simultaneously with PIXE, RBS provides, in principle, analysis of all elements above He in the periodic table.

**Interpreting RBS spectra** The classical treatment of the scattering of two charged particles yields the following expressions for the recoil energy and for the cross-section:

$$E = E_0 \frac{m_1^2}{(m_1 + m_2)^2} \left[ \cos\theta + \left( \frac{m_2^2}{m_1^2} - \sin^2\theta \right)^{\frac{1}{2}} \right]^2 \quad [3]$$

$$\frac{d\sigma}{d\Omega} \approx 1.296 \left( \frac{zZ}{E_0} \right)^2 \left[ \sin^{-4}\left( \frac{\theta}{2} \right) - 2 \left( \frac{m_1}{m_2} \right) + \ldots \right] \quad [4]$$

where $E$ is the recoil energy of a particle of mass $m_1$, charge z and initial energy $E_0$ scattering through a total angle $\theta$ from a stationary nucleus of mass $m_2$ and charge $Z$. The form of these expressions is shown in **Figure 5** for 3 MeV protons and alpha particles recoiling off nuclei up to mass 100. These curves show that the best mass resolution (change in recoil energy with mass) occurs for low masses. At high masses, the difference in recoil energy is too small to identify the mass of the scattering nucleus uniquely. The curves also show that the cross-section increases with mass. Alpha particles give both a higher cross-section and a better mass resolution, and these are normally used for RBS applications.

Because of the interplay of mass and depth information in RBS spectra, interpreting the data is not so straightforward as, for example, in PIXE spectroscopy where each peak is uniquely associated with a particular element. At present, RBS spectrum interpretation relies on simulating the spectrum from a proposed model of the sample using a suitable computer program, for example the commonly used program, RUMP. The parameters of the model can then be refined, either manually or using fitting routines, until a good match is obtained to the experimental data. Thus, some insight into the physics of the situation is required in order to decide on a suitable starting point for the modelling procedure. Some of the points to be considered are presented in **Figures 6A–E** which show simulated RBS spectra. In **Figure 6A** recoils from a very thin $SiO_2$ film give two narrow peaks, one at higher energy from Si and one at lower energy from O. The relative areas of the peaks are proportional to the ratio of Si to O. In **Figure 6B** a thicker film of $SiO_2$ gives two broader peaks, as recoils from deeper inside the sample occur at lower energies. The energy width of the peaks increases with increasing sample thickness. In the case of an infinitely thick sample, **Figure 6C**, the responses from Si and O merge and instead of seeing peaks, the presence of the two elements is seen as steps or edges

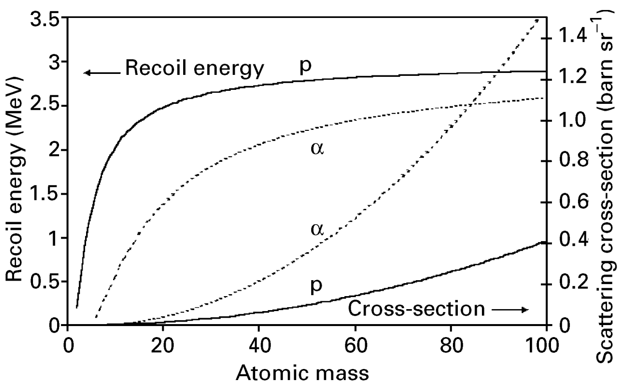



**Figure 5** Recoil energy (left axis) and scattering cross-section (right axis) for 3 MeV protons (solid line) and alpha particles (dashed line) backscattering from nuclei of different masses. Scattering angle, $\theta$, 150° (i.e. beam to detector angle 30°).

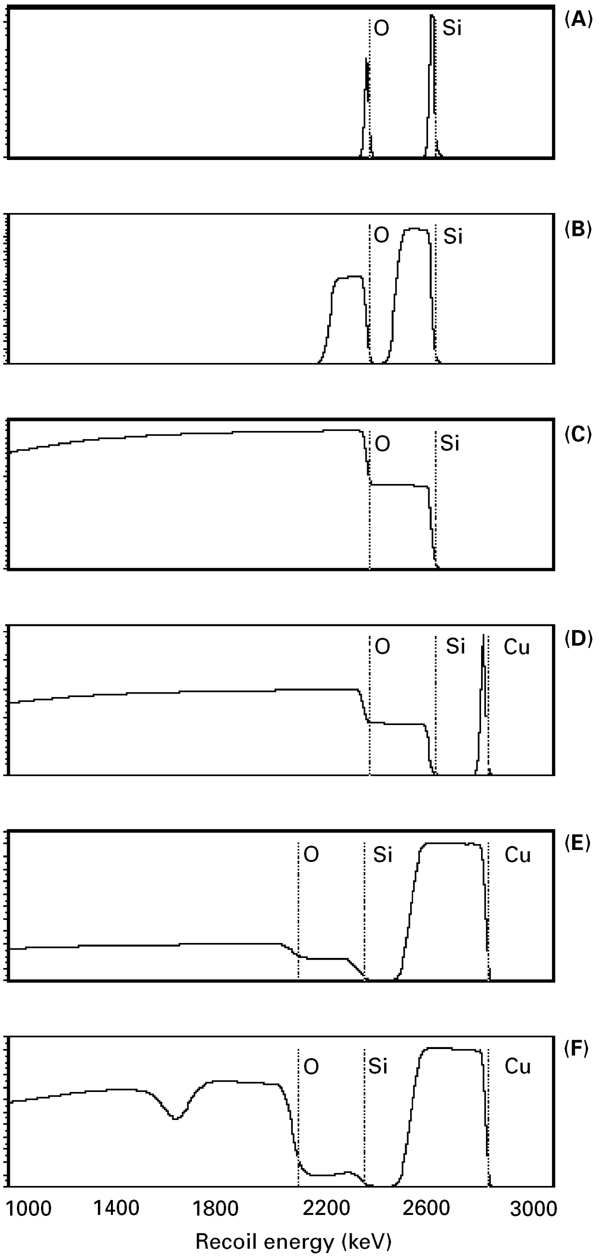



**Figure 6** Simulated RBS spectra using 3 MeV protons scattered at an angle of 150° for the following samples: (A) 0.1 μm film of $SiO_2$, (B) 5 μm film of $SiO_2$, (C) Thick sample of $SiO_2$, (D) 0.01 μm Cu film on $SiO_2$, (E) 2.5 μm Cu film on $SiO_2$. Simulations (A)–(E) carried out using the Rutherford cross-section (Eqn [3]). (F) Simulation of 2.5 μm Cu film on $SiO_2$ (case (E)) using true light element cross-sections for Si and O.

corresponding to the recoil energy of that element at the surface. When a thin Cu film is added to the surface of the $SiO_2$, this is seen as a sharp peak at higher energy (**Figure 6D**), and if the Cu layer is made thicker (**Figure 6E**), the Cu peak becomes broader, but also the Si and O edges are shifted to lower energy because of the energy loss in the Cu film.

The use of RBS with protons on light nuclei is complicated by nuclear reactions (see below) and the cross-section may be much higher (or lower) depending on the energy and the scattering angle. This is illustrated in **Figure 6F**, which shows the same simulation as in **Figure 6E**, but now using true cross-sections rather than the theoretical Rutherford formula. It can be seen that the yield for O is significantly higher than the Rutherford case due to nuclear reactions, and this can be exploited to increase the sensitivity to O. The dip in the spectrum observed at about 1600 keV is due to a nuclear reaction in $^{16}O$ which reduces the cross-section over a limited energy range.

In spite of the relative complexity of interpreting the spectra, RBS with alpha particles is widely used for materials analysis for characterizing thin film structures or depth profiles of impurities or dopants. RBS with protons can be used simultaneously with PIXE to help to determine the bulk matrix composition and also the incident beam charge.

## Nuclear reaction analysis (NRA)

If the incident ion penetrates the repulsive electrostatic field surrounding the nucleus of a target atom, it is possible for nuclear reactions to take place involving the weak and strong nucleonic forces. This can result in the formation of an excited unstable compound nucleus which may decay by the emission of particles or gamma rays to give a final nucleus which may be different from the original. Nuclear reactions in general have a much lower cross-section than atomic reactions such as PIXE (typically 10–100 mb $sr^{-1}$), but they often have resonant behaviour, so that provided the conditions are carefully chosen, NRA may give a useful yield and unique information. The cross-section and the types of reactions observed are strongly dependent on the isotope of the target atom and the beam energy and do not vary in a systematic way with atomic number, as do the yields and X-ray energies in PIXE. Nuclear reactions are often described using the notation A(a,b)B, where A and B are the initial and final states of the target nucleus, a is the projectile ion and b is the emerging reaction product (particle or gamma rays).

Nuclear reactions tend to be most important for light nuclei, as the Coulomb repulsion is less for nuclei with low $Z$. There are three type of nuclear reaction which are of interest for analytical purposes and these are described below. In general, because of the rapid and irregular variation of cross-section with energy, NRA techniques are difficult to quantify and in the majority of cases, quantitative data is obtained by comparison with known standards.

**Resonant elastic scattering** At certain energies, the yield of RBS analysis may be strongly enhanced (or reduced) by interactions with the target nucleus. One example of this is resonant scattering of protons from $^{12}C$, i.e. $^{12}C(p,p)^{12}C$. At an energy of 1.76 ± 0.1 MeV, the cross-section for backscattered protons from $^{12}C$ increases by a factor of 60 relative to the classical Rutherford formula and this can be exploited to improve the sensitivity for carbon. This is shown in the simulations of **Figure 7**, for a 10 nm thick layer of carbon on a copper substrate.

**Nuclear reactions resulting in gamma ray emission** Most nuclear reactions result in the emission of gamma rays of a well defined characteristic energy and these can be energy-analysed using a suitable detector and used for analysis. This technique is sometimes referred to as particle-induced gamma emission or PIGE. PIGE is useful for fluorine determination, for example, since the reaction $^{19}F(p, \alpha\gamma)^{16}O$ has a particularly high yield, though many other light elements up to Al can also be detected with a reasonable detection limit. **Figure 8** shows the gamma ray spectrum of a tourmaline, a mineral rich in light elements such as F, Li and B, bombarded with 3 MeV protons showing the characteristic lines from the light elements in the sample.

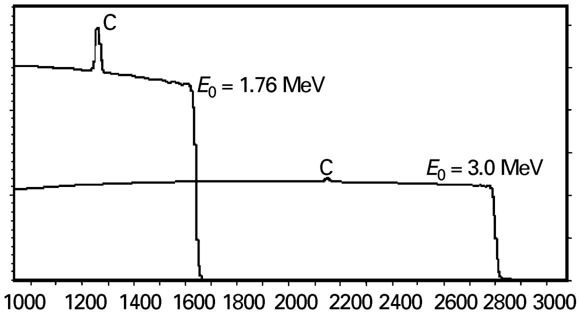



**Figure 7** Simulated RBS spectra of a 10 nm layer of C on a Cu substrate using protons of 3 MeV and 1.76 MeV, which correspond to the (p,p) resonance in $^{12}C$.

**Nuclear reactions resulting in particle emission** Some nuclear reactions result in the emission of a particle (proton or alpha) with an energy higher than the primary beam energy so that they can be detected unambiguously in the detector used for RBS analysis. Relatively few reactions are useful for analytical purposes. One example is $^{7}Li\ (p,\alpha)\ ^{4}He$, which occurs when Li is bombarded with 2–3 MeV protons. Two alpha particles are emitted which have energies of 7–8 MeV and can be detected with no interference from the spectrum of backscattered protons.

## Equipment for ion beam analysis

### Accelerators

The major component of any nuclear microbeam facility is the particle accelerator. This must generate ions of the species and energy required to carry out the analysis (protons and alpha particles with energies of a few MeV), but the requirement of operating with a focusing system may also impose additional requirements on the energy spread and brightness of the beam.

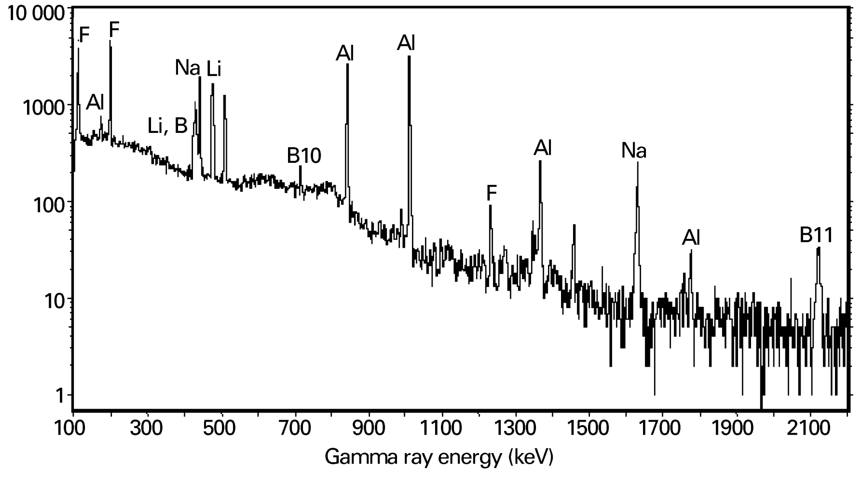



**Figure 8** Gamma ray spectrum from a tourmaline bombarded with 3 MeV protons. Gamma ray lines corresponding to F, Li, B, Na and Al are marked. The unmarked lines are created by background sources within the fabric of the laboratory.

The accelerators most suitable for IBA are the electrostatic type, in which a static voltage equal to the beam energy required is generated, and the particles are accelerated in a single step (or two steps if the 'tandem' principle is used). The voltage may be generated either by an electrostatic system such as the popular 'van de Graaff' system, or by a voltage multiplier stack, in which an alternating voltage is rectified in such a way as to generate a DC voltage many times greater than the amplitude of the applied voltage. **Figure 9** shows schematically the principle of electrostatic accelerators.

Historically a nuclear accelerator formed part of a major facility and many IBA installations still compete for beamtime at a nuclear physics accelerator institute. In contrast, modern small accelerators represent a relatively modest investment in capital and support costs and increasingly now, new small accelerators are being purchased specifically for analytical applications. It is notable, for example, that the Louvre Museum in Paris now houses a modern accelerator facility (Accelerateur Grand Louvre pour Analyses Elémentaire – AGLAE) dedicated to the analysis of art objects and archaeological samples.

### Detectors

The majority of detectors used in IBA are based on the reverse biased semiconductor junction. Ionizing radiation creates electron–hole pairs in the depletion region of the junction which drift under the applied electric field to the electrodes where they create a voltage pulse which is amplified. The material and dimensions of the detector depend on the radiation to be detected, and must be chosen so that the incident photon or particle has a high probability of interacting with an atom in the depletion region (**Table 2**). In order to reduce the noise due to thermal electrons, photon (X-ray and gamma) detectors must be operated with the detector crystal and preamplifier at a low temperature (usually with liquid nitrogen cooling, but newer semiconductor materials such as cadmium zinc telluride can give good performance with Peltier junction cooling devices). For particle detectors, other sources of noise dominate and these can be operated at room temperature.

## Applications of nonspatially resolved IBA

PIXE and RBS are the major spectroscopic techniques used for materials analysis because of their high yield and relative ease of quantification. In many cases, however, the use of nonspatially resolved, or 'broad beam' PIXE has been displaced to some extent by inductively coupled plasma mass spectrometry (ICPMS), which offers the same advantages of rapid multielemental analysis with the additional advantage of very low detection limits and greater accessibility. However, PIXE still has advantages of speed and ease of sample preparation and is used routinely in a number of applications, especially ambient aerosol filter analysis. Broad beam RBS still has a wide application in characterizing thin films and diffusion profiles of impurities in solids, for example in measuring thickness of metallization layers on silicon or depth profiles following implantation of dopants into semiconductors. The technique may also be combined with channelling, in which the variation of yield as the beam is aligned

**Table 2** Properties of semiconductor detectors used for spectroscopic analysis of charged particles, X-rays and gamma rays

| *Name ('common name')* | *Silicon surface barrier detector (surface barrier, SSB)* | *Lithium drifted silicon (Si-Li)* | *High purity germanium (HpGe)* |
|---|---|---|---|
| Radiation | MeV charged particles | X-rays | Gamma rays |
| Physical form | Thin disk of Si typically 10–30 mm diameter with depletion layer formed beneath an insulating 'barrier' on the surface | Single crystal Si disk typically 10 mm diameter, 4 mm thick. Li diffused into Si to compensate defects and impurities to ensure efficient charge collection | High purity single crystal Ge cylinder typically 20 mm diameter, 80 mm long |
| Must be cooled? | No | Yes | Yes |
| Depletion layer thickness | Typically 100 µm | ~4 mm | ~80 mm |
| Entrance window | None | Varies: typically 10 µm Be. Also ultra-thin polymer windows to extend lower energy range | Thin (~1 mm) Al |
| Energy range | ~10 keV to MeV. Upper limit depends on depletion depth | ~1 keV–50 keV. Lower limit depends on window; upper limit depends on crystal thickness | ~30 keV–10 MeV. Lower limit depends on window and electronic noise; upper limit depends on crystal dimensions |

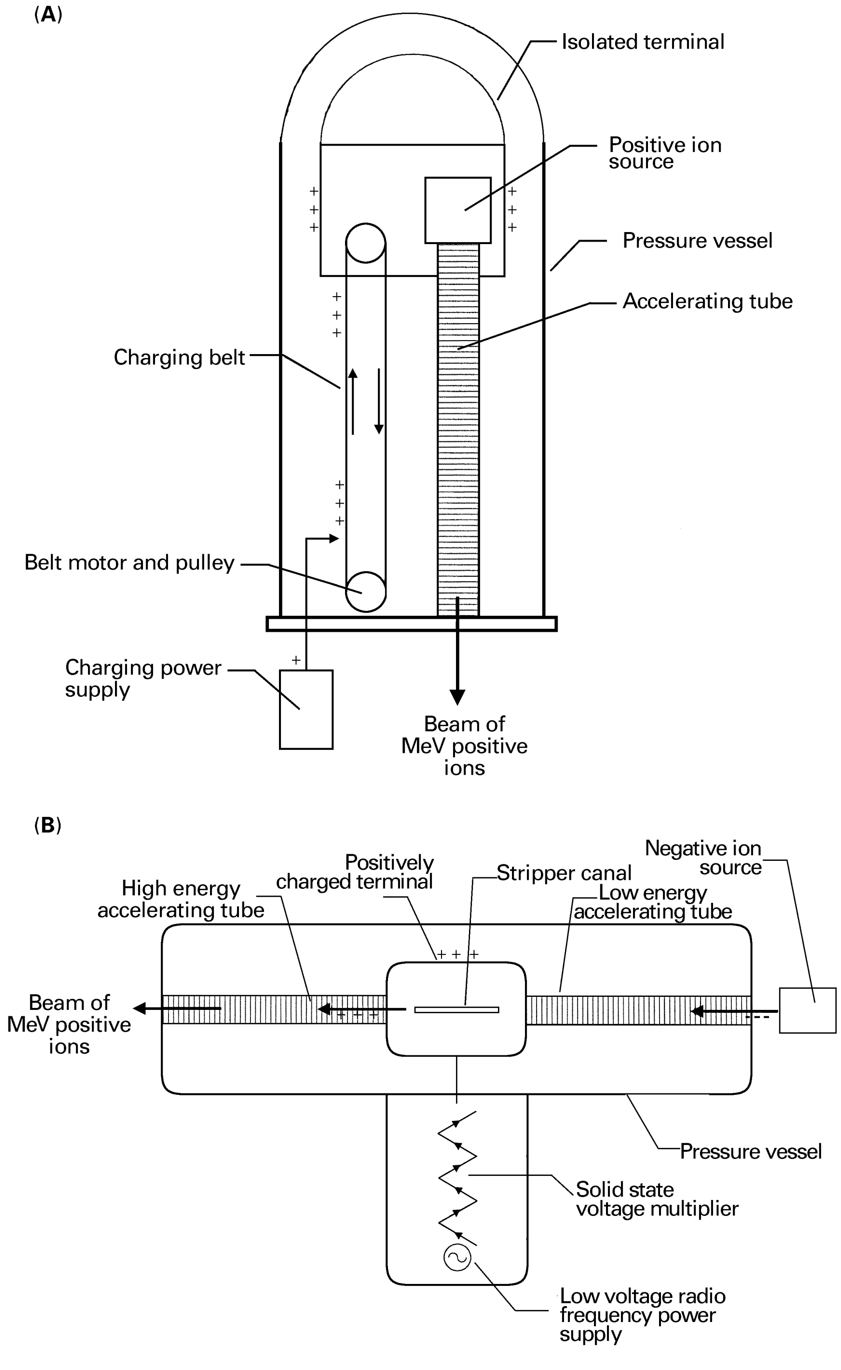



**Figure 9** Schematic diagram of (A) single ended accelerator and (B) tandem accelerator. The terminal voltage may be generated either by a moving belt or chain (the van de Graaff or Pelletron principle) or by rectifying a high frequency alternating voltage (the Tandetron). In the single-ended accelerator a positive ion source is mounted in the terminal and the beam energy is equal to the terminal voltage. In the tandem accelerator an external *negative* ion source injects ions into the accelerator where they are accelerated towards the positive terminal. In the terminal they pass through a region of low pressure gas (the stripper) where collisions with the gas molecules remove some, or all, of the electrons. The resulting positive ion is further accelerated. The final energy is given by $V(1+q)$ where $V$ is the terminal voltage and $q$ is the charge state after stripping. The accelerator assembly is normally enclosed in a pressure vessel containing several atmospheres of an insulating gas such as $SF_6$ to avoid the possibility of electrical discharges.

with one of the major axes of a single crystal sample can be used to obtain information on lattice quality and the localization within the lattice of impurity atoms. A full discussion of channelling is beyond the scope of this article.

PIXE (and to a certain extent RBS) are now more commonly used in the spatially-resolved mode. This is because their high yield contributes to the attainment of high spatial resolution.

## List of symbols

$C_z$ = concentration of element; $E$ = energy; $m$ = particle mass; $n_T$ = number of target atoms; $N$ = number of projectile atoms, number of photons; $q$ = terminal charge; $Q$ = total beam charge; $V$ = voltage; $Y$ = number of reaction products; $Y(Z,E)$ = yield of interaction; $z$ = charge; $\sigma$ = cross-section; $\varepsilon$ = detector efficiency; $\Omega$ = solid angle.

*See also:* **Atomic Absorption, Methods and Instrumentation; Atomic Absorption, Theory; Atomic Emission, Methods and Instrumentation; Atomic Fluorescence, Methods and Instrumentation; Fluorescence and Emission Spectroscopy, Theory; Geology and Mineralogy, Applications of Atomic Spectroscopy; Inductively Coupled Plasma Mass Spectrometry, Methods; Proton Microprobe (Method and Background); X-Ray Emission Spectroscopy, Applications; X-Ray Emission Spectroscopy, Methods; X-Ray Fluorescence Spectrometers; X-Ray Spectroscopy, Theory.**

## Further reading

Breese MBH, Jamieson DN and King PJC (1996) *Materials Analysis Using a Nuclear Microprobe*. New York: Wiley.

Johansson SAE and Campbell JL (1988) *PIXE, A Novel Technique for Elemental Analysis*. Chichester: Wiley.

Johansson SAE, Campbell JL and Malmqvist KG (1995) *Particle-Induced X-ray Emission Spectrometry*. New York: Wiley.

Watt F and Grime GW (1987) *Principles and Applications of High Energy Ion Microbeams*. Bristol: Hilger.

# High Pressure Studies Using NMR Spectroscopy

**Jiri Jonas**, University of Illinois at Urbana-Champaign, IL, USA

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

The use of pressure as an experimental variable in NMR studies of chemical and biochemical systems leads to added complexity in instrumentation but the unique information gained from the combination of high pressure and NMR techniques justifies fully its use. There are several fundamental reasons for carrying out NMR experiments at high pressures.

First, to separate the effects of density and temperature on various dynamic processes, one has to perform the measurements as a function of pressure. Second, for liquids the use of pressure enables one to extend the measurement range well above the normal boiling point and thus to study supercritical fluids. Third, as noncovalent interactions play a primary role in stabilization of biochemical systems, the use of pressure allows one to change, in a controlled way, the intermolecular interactions without the major perturbations produced by changes in temperature and/or chemical composition. Fourth, pressure affects chemical equilibria and reaction rates. The following standard equations define the reaction volume, $\Delta V$ and the activation volume, $\Delta V^{\ddagger}$:

$$\Delta V = -\left[\frac{RT\,\partial \ln K}{\partial P}\right]_T, \quad \Delta V^{\ddagger} = -\left[\frac{RT\,\partial \ln k}{\partial P}\right]_T$$

where $K$ is the equilibrium constant, and $k$ is the reaction rate. With the knowledge of $\Delta V^{\ddagger}$ values one can draw conclusions about the nature of reaction and its mechanism. Fifth, the phase behaviour and dynamics of molecular solids and model membranes can be explored more completely by carrying out high pressure experiments. Sixth, the combination of advanced 2D, and 3D NMR techniques with high pressure capability represents a powerful new experimental tool in studies of protein folding.

The usual range of pressures used to investigate chemical and biochemical systems is from 0.1 MPa to 1 GPa (0.1 MPa = 1 bar; 1 GPa = 10 kbar); such pressures only change intermolecular distances and affect conformations but do not change covalent bond distances or bond angles.

## Pressure effects on NMR spectra

This heading is related to the fact that the effects of pressure and/or density on chemical shifts and spin–spin-coupling in molecular systems are relatively small and difficult to interpret. In contrast, most of the high pressure NMR studies of chemical or biochemical systems deal with pressure or density effects on dynamical behaviour of the readily compressible gases, molecular liquids and aqueous solution of biomolecules. Increasing pressure changes both the density and shear viscosity of the liquids which in turn changes the relaxation times and chemical exchange. The majority of NMR studies deal with liquids of low viscosity for which $T_1 = T_2$ and the condition of extreme motional narrowing applies. In the case of rotational motions of a spherical top molecule the following expression can be used to calculate the rotational correlation time, $\tau_R$, from proton relaxation times

$$\frac{1}{T_1} = \frac{3}{2}\hbar^2\gamma^4\sum_{i,j} r_{ij}^{-6}\tau_R$$

where $\gamma$ is the gyromagnetic ratio for hydrogen and $r_{ij}$ is the distance between protons. However, for proton-containing molecules one has to separate the intramolecular and intermolecular contributions to the observed proton relaxation times. In the case of nuclei with spin $I > \frac{1}{2}$ the relaxation studies provide direct information about rotational motions because the electric quadrupole interactions provide the dominant relaxation mechanisms. For nuclei with spin $I = 1$ (e.g. $^2$H, $^{14}$N) the relaxation rate due to the quadrupolar interaction is given by the well-known expression

$$\frac{1}{T_1} = \frac{3}{8}(1 + \chi^2/3)(e^2qQ/\hbar)^2\tau_R$$

where $\chi$ is the asymmetry parameter, $e^2qQ/\hbar$ is $2\pi$ times the quadrupole coupling constant in hertz, and $\tau_R$ is the rotational correlation time.

It is well established that the rotational correlation time, $\tau_R$, for the isotopic rotation of a spherical top molecule can be expressed in terms of the macroscopic viscosity, $\eta$, using the Debye equation

$$\tau_R = \frac{4\pi a^2\eta}{3k_B T}$$

where $\eta$ is the viscosity, $a$ is the radius of the molecule and the other symbols have their usual meaning.

In the case of the intermolecular contribution to relaxation times of proton-containing molecules and for the effect of pressure on self-diffusion one can use the Stokes–Einstein expression, which relates the shear viscosity, $\eta$, to the self-diffusion coefficient $D$:

$$D = \frac{k_B T}{6\pi a\eta}$$

where all symbols have their usual meaning.

For non-hydrogen bonded liquids the increase of pressure slows down both rotational and translational motions with the result that for low viscosity liquids one shortens the spin–lattice relaxation times and decreases the diffusion.

The slowing down of exchange owing to increased pressure in chemically exchanging systems also leads to changes in the NMR spectra. As discussed elsewhere in this encyclopedia, high resolution NMR spectroscopy is widely used to study chemical exchange processes. As an example of the effect of pressure on chemical exchange **Figure 1** shows the pressure dependence of hindered rotation about the C–N amide bond in *N*,*N*-dimethyltrichloracetamide (DMTCA). From the observed line shapes it is straightforward to calculate the observed reaction rate and from its pressure dependence one can calculate the volume of activation $\Delta V^\ddagger$. In the case of DMTCA the experimental rotation rate, $k$, decreases with increasing pressure and the correlation of the rates with measured viscosity (shear viscosity, $\eta$) shows that the hindered rotation in DMTCA falls in the strongly coupled diffusive regime.

There is yet another important origin of observed pressure changes in the NMR spectra of biomolecules in aqueous solutions. NMR studies of protein denaturation using temperature or chemical means are well established but pressure also leads to reversible denaturation as discussed in more detail in the section on applications. In the case of model membranes the use of pressure leads to very rich barotropic behaviour and high pressure NMR techniques can establish pressure–temperature phase diagrams for the membrane system studied.

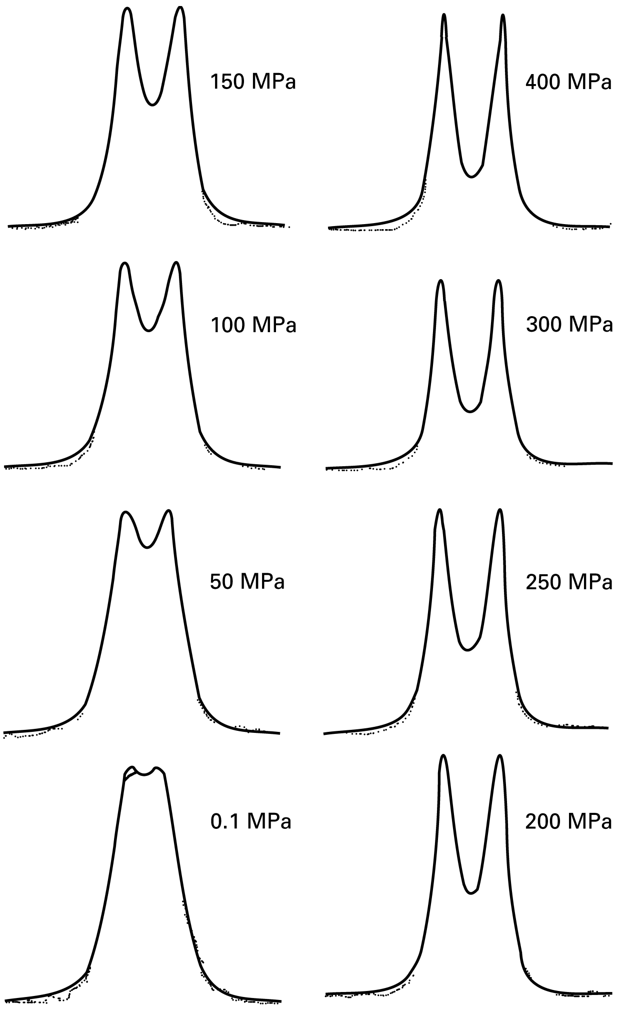



**Figure 1** Pressure effects on the line shape of the N-methyl proton NMR spectrum of DMTCA in pentane at 282.3 K. The dotted lines denote the experimental line shapes, and the full lines denote the calculated line shapes.

In addition to gases and liquids the NMR spectra of molecular solids exhibit sensitivity to pressure. The relaxation behaviour is changed and, in addition, pressure also leads to changes in phase-transition temperatures.

The effect of pressure on the phase-transition temperature can be evaluated using the Clausius–Clapeyron equation:

$$\frac{\mathrm{d}T}{\mathrm{d}P} = T\left(\frac{\Delta V}{\Delta H}\right)$$

where $T$ is the transition temperature, $P$ is the pressure, $\Delta V$ is the molar volume change and $\Delta H$ is the change in molar enthalpy.

## High pressure NMR instrumentation

Since high resolution NMR spectroscopy represents a spectroscopic technique of major importance for chemistry and biochemistry the overview of high pressure NMR instrumentation focuses on high resolution, high pressure NMR techniques.

Thanks to advances in superconducting magnet technology one can achieve a high homogeneity of the magnetic field over the sample volume even without sample spinning and, therefore, it is possible to construct high resolution NMR probes for work at high pressures. The high resolution, high pressure NMR equipment consists of three main parts: (a) pressure generating and pressure measuring systems; (b) nonmagnetic high pressure vessel; and (c) a NMR probe that includes a sample cell. All components necessary for building a high pressure setup, such as hand pumps, high pressure tubing, and valve intensifiers are currently available from commercial sources. **Figure 2** shows a schematic drawing of a high pressure generating system capable of producing a hydrostatic pressure up to 1 GPa. In addition to high pressure NMR instrumentation based on a high pressure vessel made from non-magnetic metallic alloy, one can also obtain high resolution NMR spectra using glass capillaries or special sample cells made from sapphire. **Table 1**, which gives characteristic performance features of high pressure NMR probes, also includes information on the diamond anvil cell (DAV) which is suitable for broad-line work up to pressures of 8 GPa. However, the extremely small sample size makes DAV probes difficult to operate and relatively insensitive.

A schematic drawing of a high pressure NMR vessel, and a sample cell used for high resolution NMR experiments at pressures up to 950 MPa, is shown in **Figure 3**. It is important to note that even at pressures of 1 GPa and sample diameter of 8 mm one can achieve a resolution of $3 \times 10^{-9}$. The natural abundance $^{13}$C NMR spectrum of 2-ethylhexylcyclohexanecarboxylate recorded at 80°C and 500 MPa as shown in **Figure 4** serves as an illustrative example of the high quality NMR spectra which can be routinely obtained at high pressures.

## Applications of high pressure NMR spectroscopy

### Dynamic structure of liquids

Theoretical and experimental evidence indicate that many physical properties of liquids may be

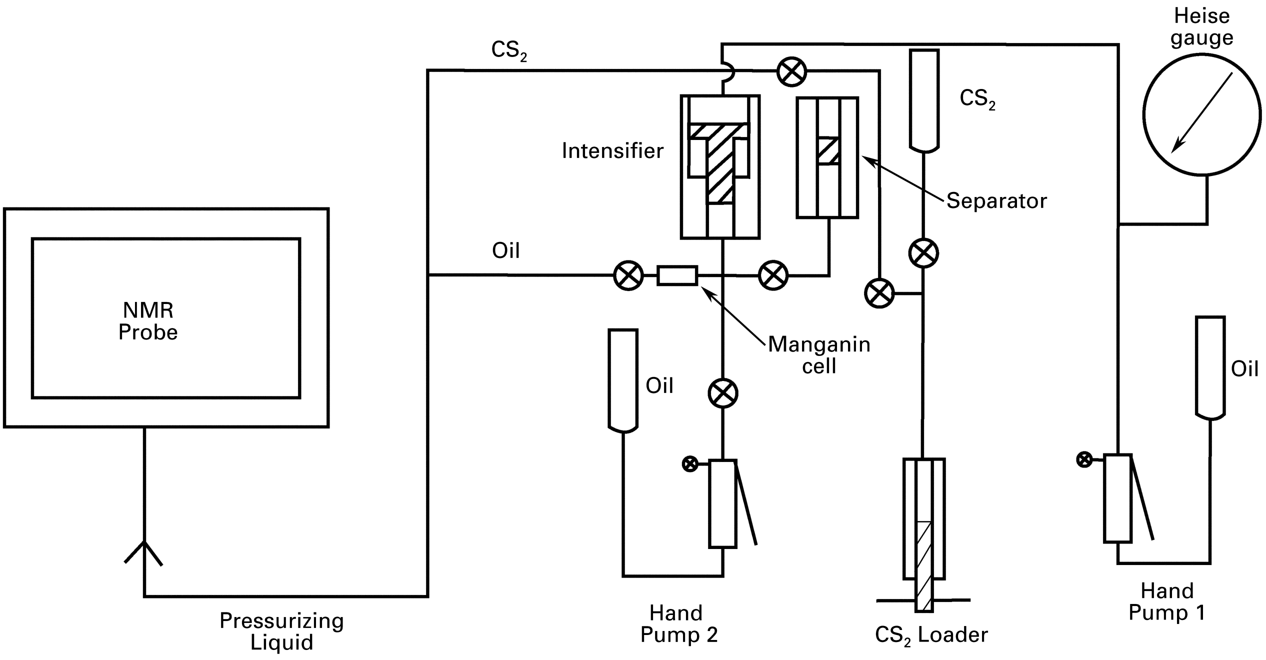



**Figure 2** Schematic diagram of the high pressure generating equipment.

**Table 1** Characteristic performance features of high pressure NMR probes

| Probe type | *Pressure range (MPa)* | *Temperature range (K)* | *Sample diameter (mm)* | *Resolution* ($\times 10^{-9}$) | $^{1}$*H Frequency (MHz)* |
|---|---|---|---|---|---|
| Capillary cell | 0.1–250 | *a* | 1–1.5 | 3 | *a* |
| Sapphire cell | 0.1–100 | *a* | 3.4 | 1.5 | *a* |
| HP vessel | 0.1–950 | 253–373 | 8 | 3 | 60–500 |
| DAV[b] | 0.1–8000 | 200–300 | <<1 | *c* | 1–100 |

[a] Temperature and frequency range is determined by the commercial probe used.
[b] Diamond anvil cell.
[c] Broad line.

determined, to a large extent, by the size and shape of the constituent atoms or molecules. Owing to the close packing of molecules in liquids, even a small change in density can produce a considerable change in the molecular dynamics of the liquid; therefore, to test rigorously a theoretical model of a liquid, or a model of a specific dynamic process in a liquid, one must perform isochoric, isothermal, and isobaric experiments. **Figure 5**, which shows the temperature dependence of the self-diffusion coefficient, $D$, in liquid tetramethylsilane at constant density and constant pressure, illustrates the importance of separating the effects of density and temperature on molecular motions. Four main research directions can be identified in the high pressure NMR studies of liquids: (a) studies of self-diffusion; (b) investigations of reorientational motions; (c) studies of angular momentum behaviour; and (d) tests of applicability of hydrodynamic equations at the molecular level.

The effect of pressure on the anisotropic reorientation of acetonitrile-$d_3$ in the liquid state can serve as in illustrative example of high pressure NMR studies of molecular reorientation in liquids. The deuteron and nitrogen spin–lattice relaxation times of acetonitrile-$d_3$ have been measured as a function of pressure up to 200 MPa at 23°C, using the NMR $T_1$ relaxation technique. Since the quadrupole coupling constants for nitrogen and deuterons in acetonitrile-$d_3$ are known, the experimental $T_1$ data are interpretable in terms of the rotational diffusion constants for motion perpendicular $D_\perp$ and parallel $D_\parallel$ to the symmetry axis.

From **Figure 6**, which plots the pressure dependence of $D_\perp$ and $D_\parallel$, it is readily apparent that $\Delta V^{\ddagger}(D_\perp) >> \Delta V^{\ddagger}(D_\parallel)$. This experimental finding

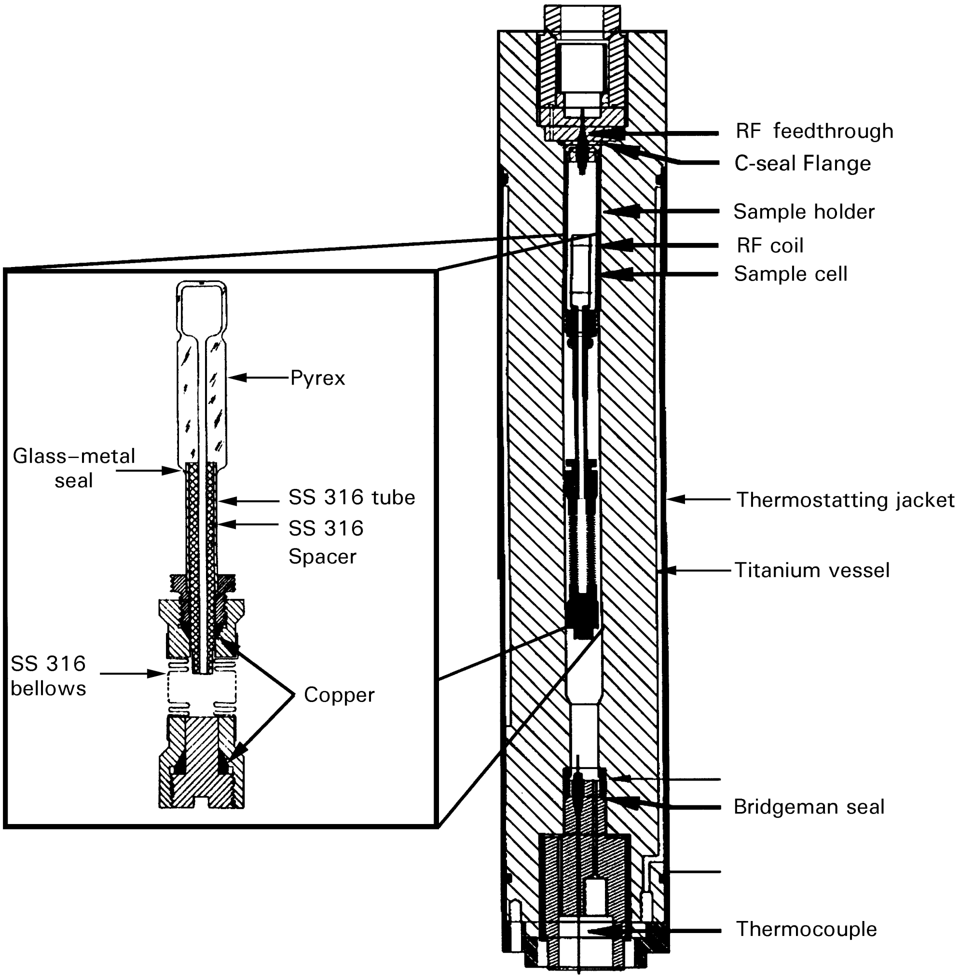



**Figure 3** Schematic drawing of a high pressure NMR probe and sample cell.

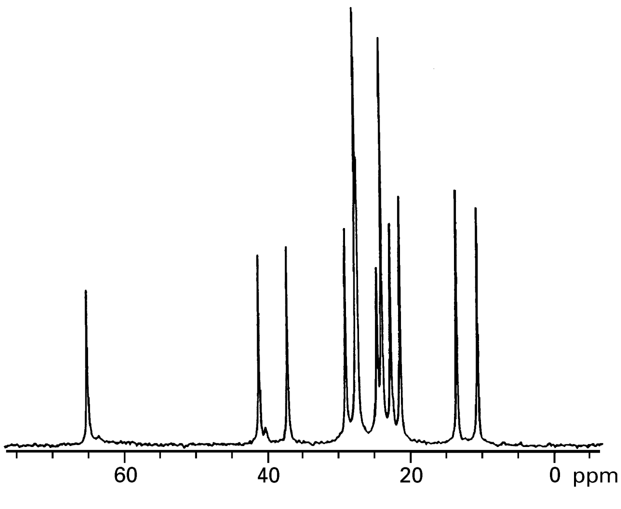



**Figure 4** High resolution natural abundance $^{13}C$ NMR spectrum of liquid 2-ethylhexylcyclohexanecarboxylate at 80°C and 500 MPa.

reflects the difference in frictional torques connected with the reorientation about the individual molecular axis, or more generally stated, reflects the symmetry of intermolecular potentials. The low $\Delta V^{\ddagger}(D_{\parallel})$ value indicates that the intermolecular potential energy is largely independent of the angle of orientation about the main symmetry axis and thus the rotational frictional coefficient is very small.

## Supercritical fluids

NMR techniques can be used to study supercritical fluids as the employed pressures range from 0.1 to 80 MPa. The reason for renewed interest in the properties of supercritical fluids can be traced to the great promise of supercritical fluid extraction techniques. High pressure NMR spectroscopy can not only be used to investigate transport and intermolecular interactions in compressed supercritical

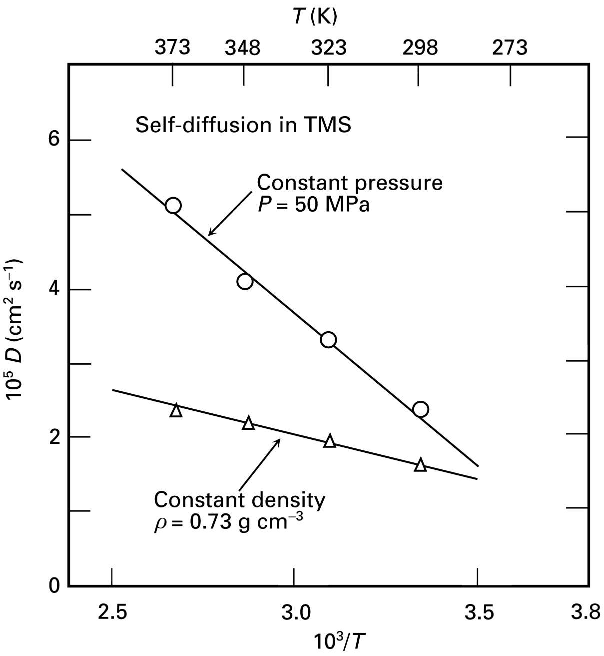



**Figure 5** Temperature dependence of self-diffusion in liquid tetramethylsilane (TMS) at (○) constant pressure and (Δ) constant density.

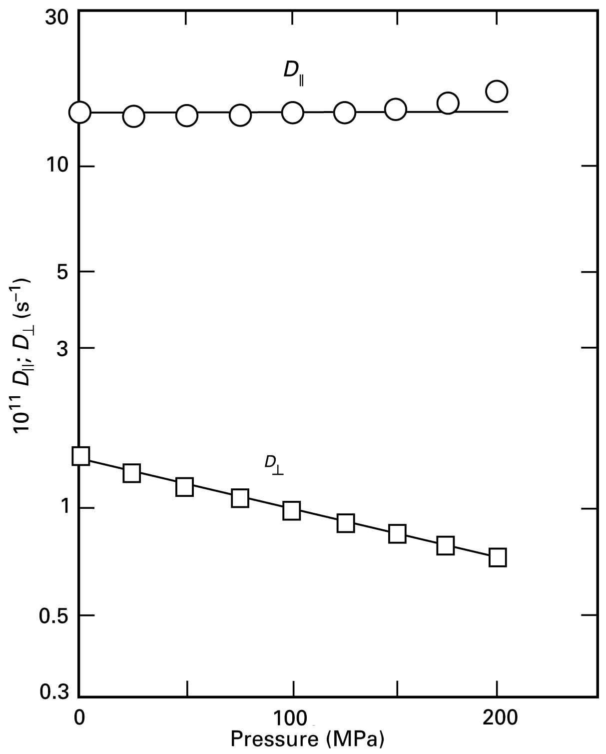



**Figure 6** Pressure efffects on the rotational diffusion constants $D_\perp$ and $D_\parallel$ of acetonitrile-$d_3$ at 23°C.

fluids but an in situ NMR technique has also been developed for the determination of the solubility of solids in supercritical fluids.

The basic idea behind the determination of the solubility of solids in supercritical fluids is based on the radically different spin–spin relaxation rate between dissolved and solid material ($T_2$, solid $\ll T_2$, dissolved). Using the 90°–$t$–180° spin echo sequence with a pulse separation of $t$ << 1 ms ensures that no contribution to the NMR echo signal can result from the quickly relaxing protons in the solid material. **Figure 7** gives the experimental solubilities for solid naphthalene in supercritical carbon dioxide as determined by the spin echo NMR technique. The three isotherms measured were selected near the upper supercritical end point (UCEP) for naphthalene in $CO_2$. In **Figure 7** the 58.5°C isotherm shows a large increase in solubility at about 23.5 Mpa; the slope of the isotherm is near zero. This indicates that one operates in close proximity to UCEP. This NMR technique can yield both solubility data and phase information when studying equilibria in supercritical fluid mixtures.

### Phase transitions in molecular solids

Because molecular solids exhibit significant compressibility even in the pressure range 0.1 to 900 MPa one can use high pressure NMR relaxation techniques to investigate phase transitions. The high degree of molecular rational freedom in plastic crystals (orientationally disordered crystals)

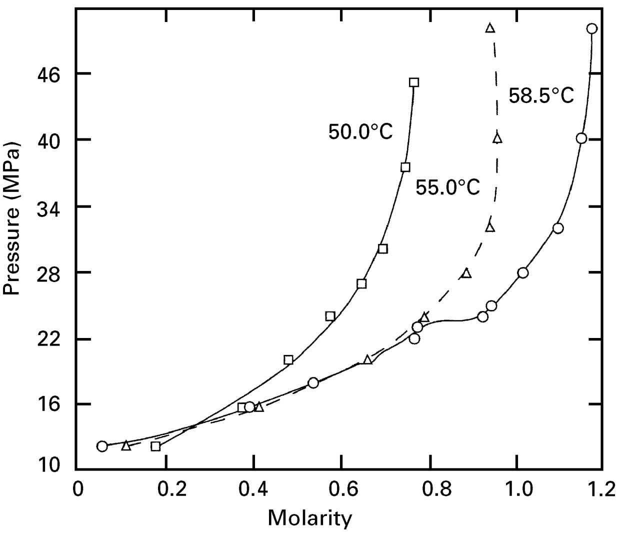



**Figure 7** Experimental solubilities for solid naphthalene in supercritical $CO_2$ expressed in moles of napthalene dissolved per litre of solution.

provides a good illustrative example. Adamantane, $C_{10}H_{16}$, is a saturated hydrocarbon with an fcc crystal structure in its high temperature plastic phase (α-phase) whereas at lower temperatures it undergoes a phase transition to a tetragonal structure (β-phase) in which the rotation is quenched. Taking advantage of the abrupt change in $T_1$ or $T_2$ at the phase transition provides the means of detecting the precise phase-transition pressure. At the phase-transition pressure, the $T_1$ of adamantane decreases abruptly by a factor of 40 in passing from the orientationally disordered plastic α-phase to the β-phase. **Figure 8** shows the pressure dependence of the proton spin–lattice relaxation times, $T_1$, in solid adamantane in the α- and β-phases at different temperatures.

The relaxation time can be expressed in terms of the rotational correlation time $\tau$, and the

$$\frac{1}{T_1} = C\left(\frac{\omega_0\tau}{1+\omega_0^2\tau^2} + \frac{4\omega_0\tau}{1+4\omega_0^2\tau^2}\right) \qquad [1]$$

second moments ($M^2$) where $\omega_0$ is the Larmor frequency (Eqn [1]). The constant $C$ is proportional to the proton second moment expressed in angular frequency units. For the measured $T_1$ one can assume $\omega_0\tau >> 1$ for the β-phase rotation. Then the expressions for the relaxation times are as follows:

$$(1/T_1)_\alpha = 5\omega_0\tau C_\alpha \qquad \text{for } \omega_0\tau \ll 1$$
$$(1/T_1)_\beta = 2C_\beta/\omega_0\tau \qquad \text{for } \omega_0\tau \gg 1 \qquad [2]$$

where $C_\alpha\omega_0 = 8.91 \times 10^9$ radian$^2$ s$^{-2}$ and $2C_\beta/\omega_0 = 2.07 \times 10^{-7}$.

Assuming that rotational motion is a thermally activated process and that the second moments are independent of temperature and pressure within our experimental range, one can obtain the activation volume $\Delta V^\ddagger(T)$ at constant temperature from the pressure dependence of $\ln(T_1)$:

$$T_1(T,P) = A(T)\exp\{\pm[\Delta V^\ddagger(T)/RT]P\} \qquad [3]$$

and the activation enthaply, $\Delta H^\ddagger(P)$, at constant pressure:

$$T_1(T,P) = B(P)\exp[\pm\Delta H^\ddagger(P)/RT] \qquad [4]$$

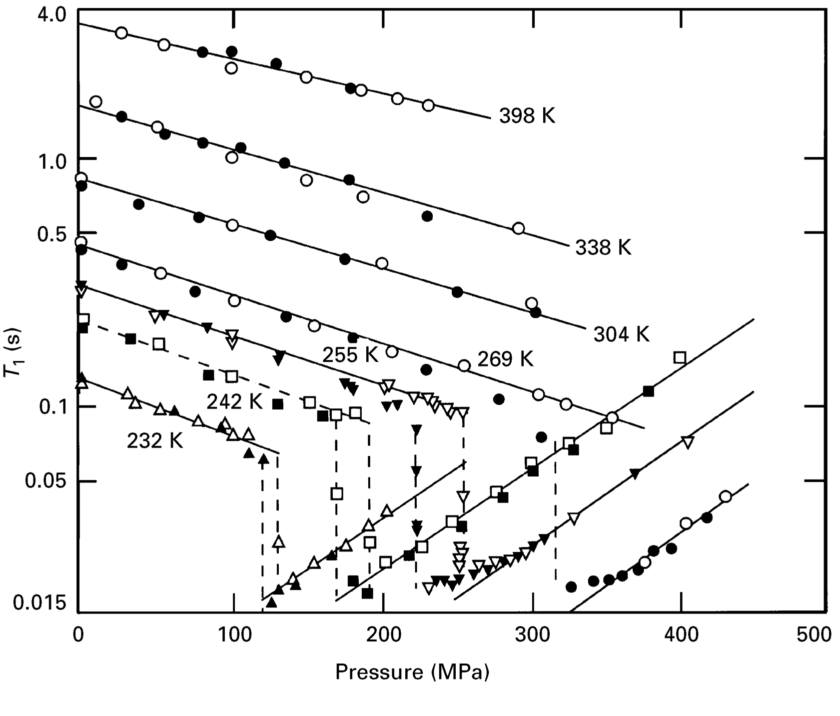



**Figure 8** Pressure dependence of the proton spin–lattice relaxation times, $T_1$, of solid adamantane in its two phases at different temperatures. The open symbols denote values obtained in the compression run whereas the filled symbols give values for the decompression run.

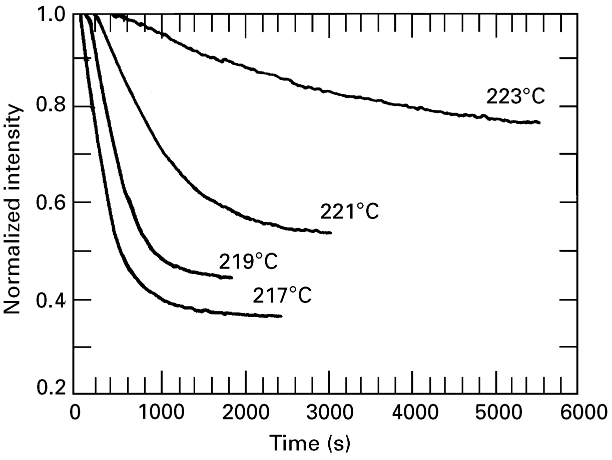



**Figure 9** Normalized spin-echo amplitude as a function of time in polyethylene at 400 MPa and temperatures of 217 to 223°C.

where Equations [3] and [4] have a plus sign when $\omega_0\tau >> 1$, and a minus sign when $\omega_0\tau << 1$.

The interesting feature of the experimental results given in **Figure 8** is the hysteresis of the observed transition pressure for the compression $\alpha \rightarrow \beta$ run and the decompression $\beta \rightarrow \alpha$ run. Analysis of this hysteresis can provide valuable information about the nature of order–disorder transitions.

### Pressure effects on crystallization kinetics of polymers

NMR relaxation or line width measurements are sensitive to the degree of crystallinity in solid polymers, because the protons in the crystalline lattice experience strong dipole–dipole interactions which cause fast spin–spin relaxation and line broadening. As a result, line width and relaxation studies have been used to measure crystallinity in a broad range of polymers. The use of high pressure as an experimental variable in NMR studies of polymer crystallinity offers new details about polymer crystallization. The study of pressure and temperature effects on the kinetics of crystallization of linear polyethylene can be used as an illustrative example of this specific application of high pressure NMR techniques. One can determine the amount of amorphous polyethylene present from the initial maximum by spin-echo measurements 90°–$\tau$–180°–$\tau$–echo with $\tau$ equal to 300 μs. **Figure 9**, which plots the amplitude of the echo signal against time, shows the crystallization data at 400 MPa for polyethylene. Before each measurement the sample was equilibrated at the desired temperature and a pressure 75 MPa below the crystallization pressure. The very fast relaxation of the crystalline material prevented it from contributing to the echo intensity. The isotherms as depicted in **Figure 9** show the characteristic crystallization behaviour: an induction time before crystallization, a primary crystallization, and a gradual decrease in crystallization rate as the final crystallinity is approached and secondary crystallization and perfection pressures occur. The quantitative analysis of the high pressure crystallization kinetics yields information about the formation of extended crystals, the value of the Avrami coefficient, the mechanism of extended chain crystallization, and the effect of pressure on the surface energies of the crystal nuclei. High pressure NMR experiments are interesting because of their direct sensitivity to polymer chain dynamics on the microscopic scale. In this way NMR complements techniques such as dilatometry which are sensitive to bulk properties.

### Mechanistic studies of chemical reactions in solution

High pressure NMR studies of various chemical reactions provide information about the reaction mechanism. Almost all reactions exhibit a characteristic pressure dependence of the reaction rate which can be used to calculate the volume of activation, $\Delta V^{\ddagger}$. The $\Delta V^{\ddagger}$, extrapolated to ambient pressure, together with the partial molar volume for the reactants and products, or with the reaction volume

**Table 2** High pressure kinetic studies of solvent exchange in inorganic systems

| *System* | *Solvent* | *Pressure (bar)* | *Temperature (K)* | *Nuclei* |
|---|---|---|---|---|
| $[Cu(DMF)_6]^{2+}$, [Cu(H2O)6]$^{2+}$ | DMF, $H_2O$ | 2000 | 217–300 | $^{17}O$ |
| $[Ln(PDTA)(H_2O)_2]^-$ (Ln = Tb, Dy, Er, Tm, Yb) $[Er(EDTA)(H_2O)_2]^-$ | H2O | 2000 | 272–292 | $^{17}O$ |
| $[Ln(DTPA\text{-}BMA)(H_2O)]$ (Ln = Nd, Eu, Tb, Dy, Ho) | $H_2O$ | 2000 | na | $^{17}O$ |
| $[M(en)_3](CF_3SO_3)_2$ (M = Co, Fe, Mn) | en$^a$ | 2000 | 278–332 | $^{14}N$ |
| $[Cp^*M(H_2O)_3]^{2+}$ (M = Rh, Ir) | $H_2O$ | 2000 | 288–333 | $^{17}O$ |

$^a$ Ethylenediamine.

of the overall reaction, allows one to create a volume profile that describes the volume changes along the reaction coordinate. The location of the transition state with respect to reactants and products determines the mechanism of the reaction investigated. Both organic and inorganic reactions can be studied. Specifically, solvent and ligand exchange, and ligand substitution reactions in inorganic systems have extensively been studied by high pressure NMR techniques. **Table 2** gives the activation parameters of solvent exchange studied in inorganic systems.

## Biochemical applications

### Model membranes

Phospholipid bilayers and monolayers have important roles in nature as components of cell membranes

**Table 3** Examples of NMR studies of the pressure effects on model membranes[a]

| *System* | *Experiment* | *Result* |
|---|---|---|
| DPPC[b] | Natural abundance $^{13}$C, $T_1$, $T_2$ | Phase transitions |
| DPPC-$d_{62}$ DPPC-$d_{62}$-TTC[c] | $^2$H line shapes | Phase diagram; order parameter; pressure reversal of the anaesthetic effect of tetracaine |
| DPPC-TTC[e] | $^{31}$P line shapes, $T_1$ | Structure and dynamics of the head group; phase diagram |
| DPPC-$d_2$ (2,2); (9,9); (13,13) | $^2$H line shapes, $T_1$, $T_2$ | Order parameters; chain motions |
| DPPC-$d_{62}$-cholesterol | $^2$H line shapes | Phase diagram |
| DPPC<br>POPC[d] | $^1$H $T_1$ and $T_{1\rho}$ | Lateral diffusion |

[a] Pressure range from 0.1 to 500 MPa.
[b] DPPC = dipalmitoylphosphatidylcholine.
[c] TTC = tetracaine.
[d] POPC = palmitoyloleylphosphatidylcholine.

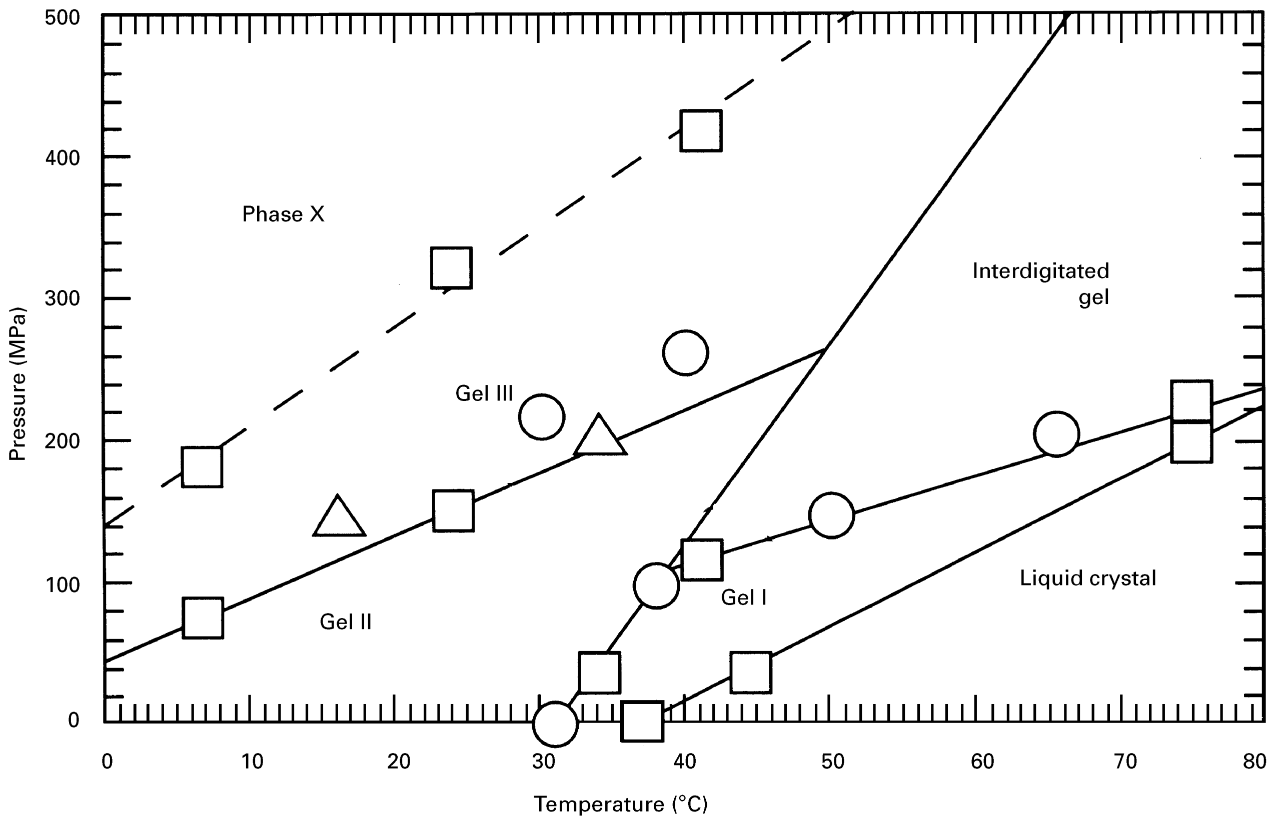



**Figure 10** Pressure–temperature phase diagram of DPPC-$d_{62}$.

and lipoproteins. In cell membranes, phospholipid bilayers constitute the permeability barriers between aqueous compartments. Aside from their structural interfacial roles, aggregated phospholipids modulate the functions of associated proteins and enzymes by direct binding and by physical effects. Therefore, synthetic phospholipid bilayers, in multilamellar or small bilayer vesicle forms, have become models for the study of the structural and dynamic properties of natural phospholipid aggregates.

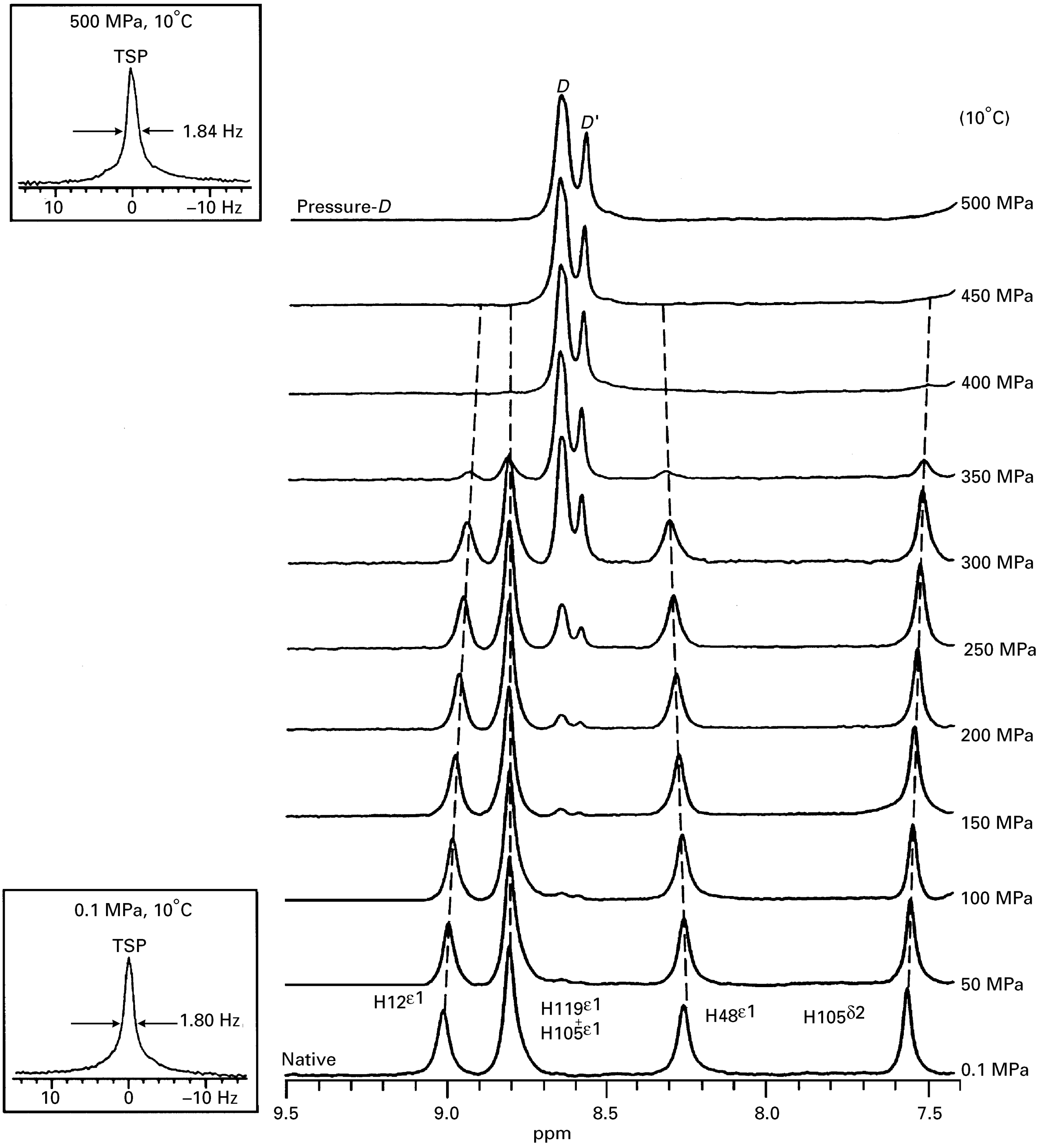



**Figure 11** Histidine region of the $^1$H NMR spectra of RNase A in $D_2O$ at various pressures. The standard in the insets is sodium 3-(trimethylsilyl) tetradeuteriopropionate (TSP).

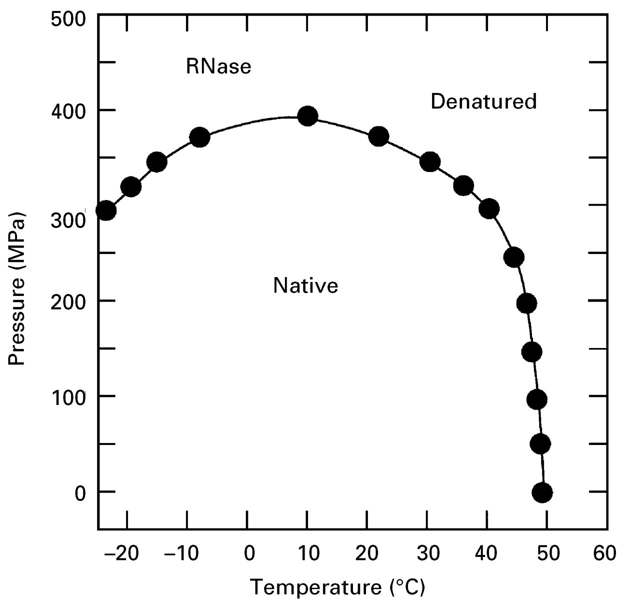



Figure 12 Phase diagram of RNase A.

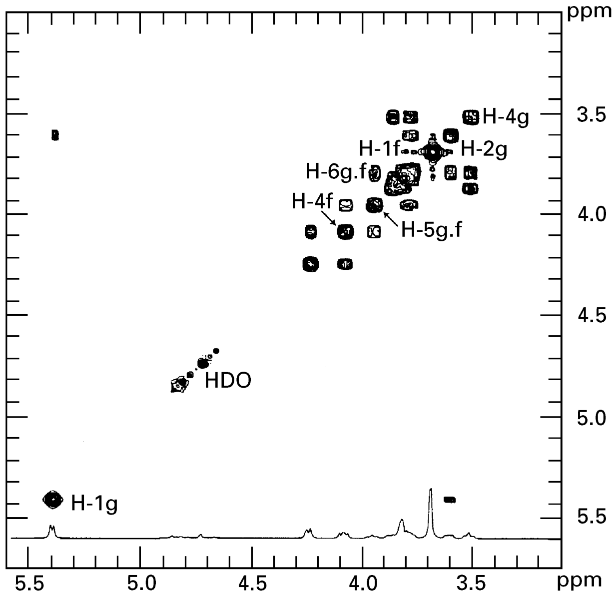



Figure 13 2D $^1$H Correlation spectroscopy spectrum with presaturation pulse for solvent suppression of 2 mM sucrose at 25°C and 475 MPa pressure recorded at 500 MHz.

Since the 1970s, NMR methods, in particular $^2$H NMR, have been applied very effectively to investigate the biophysical properties of phospholipid bilayers. NMR experiments with high pressure and variable temperature measurements allow one to access diverse gel phases in phospholipid bilayers and to study the order and dynamics of phospholipids in bilayers as a function of volume changes.

**Table 3** gives illustrative examples of high pressure NMR studies of model membranes and indicates the type of information that can be obtained from these studies. **Figure 10** shows the pressure–temperature phase diagram of deuterated dipalmitoylphosphatidylcholine, DPPC-$d_{62}$, as determined by high pressure NMR techniques using direct measurement of the deuteron quadrupole splitting of the methyl group and first moment analysis.

## Pressure denaturation of proteins

The relationship between the amino acid sequence and the structure of the native conformation of proteins represents the key unsolved problem of biochemistry and biology. Temperature and chemical perturbation represent the usual approach used in studies of protein stability and the folding problem. However, it is advantageous to use pressure to study protein solutions as pressure perturbs the protein environment in a controlled way by changing only intermolecular interactions. Moreover, by taking advantage of the phase behaviour of water, high pressure can lower the freezing point of an aqueous solution and allow the study of cold denaturation of proteins. The reversible pressure denaturation of bovine ribonuclease A (RNase A) can be used as an illustrative example for pressure denaturation. In this specific protein the $\varepsilon 1$ hydrogen signals of the four histidine residues are well resolved from other proton peaks in the 1D $^1$H NMR spectrum of the native RNase A in $D_2O$. These peaks have been assigned and used to investigate the folding and unfolding of the protein. **Figure 11** shows the pressure effects on the proton NMR spectrum in the histidine region of RNase A at pH 2.0 and 10°C. As expected with increasing pressure, the intensity of the native histidine peaks decreases and at about 400 MPa their disappearance indicates that the protein is denatured. The appearance of peaks *D* and *D′* with increasing pressure signals the denaturation.

It is important to point out that during the pressure-assisted cold denaturation many proteins show significant residual secondary structure which parallels the structure of folding intermediates. It appears that pressure denaturation and pressure-assisted cold denaturation of proteins studied by high resolution NMR techniques provide novel information about the folding of proteins. In addition, these experiments allow the determination of phase diagrams for proteins. **Figure 12** shows the

pressure–temperature phase diagram for RNase. Above the curve the protein is in the denaturated state.

The most promising direction of the high pressure studies of proteins is the application of advanced 2D and 3D NMR techniques. **Figure 13** shows the high quality of 2D NMR spectra which can be routinely obtained at high pressures. Very often the 2D spectra obtained at high pressure offer new insights into the pressure-induced unfolding and pressure-induced dissociation of proteins when compared with results obtained by other high pressure spectroscopic techniques such as fluorescence methods.

## List of symbols

$a$ = molecular radius; D = Self-diffusion coefficient; $I$ = nuclear spin quantum number; $k$ = rate constant; $K$ = equilibrium constant; $P$ = pressure; $r_{ij}$ = distance between protons; $R$ = gas constant; $t$ = pulse separation time; $T$ = absolute temperature; $T_1$, $T_2$ = relaxation times; $\Delta V$ = reaction volume; $\Delta V^{\ddagger}$ = activation volume; $\gamma$ = gyromagnetic ratio; $\eta$ = shear viscosity; $\chi$ = asymmetry parameter; $\omega_0$ = Larmor frequency.

*See also:* **Biofluids Studied By NMR; Diffusion Studied Using NMR Spectroscopy; High Pressure Studies Using NMR Spectroscopy; Labelling Studies in Biochemistry Using NMR; Membranes Studied By NMR Spectroscopy; NMR Relaxation Rates; Proteins Studied Using NMR Spectroscopy; Relaxometers; Two-Dimensional NMR, Methods.**

## Further reading

Benedek GB (1963) *Magnetic Resonance at High Pressure*. New York: Interscience Publishers.

Jonas J (1973) NMR studies in liquids at high pressure. *Advances in Magnetic Resonance* **6**: 73–139.

Jonas J (1975) Nuclear magnetic resonance at high pressure. *Annual Reviews in Physical Chemistry* **26**: 167–190.

Jonas J (1978) Magnetic resonance spectroscopy at high pressure. In: Kelm H (ed) *Proceedings NATO ASI, Series C, High Pressure Chemistry*, pp 65–110. Dordrecht: Reidel.

Jonas J (1982) Nuclear magnetic resonance at high pressure. *Science* **216**: 1179–1184.

Jonas J (1991) High pressure NMR studies of the dynamics in liquids and complex systems In: Jonas J (ed) *High Pressure NMR*, pp 85–128. Vol 24 in Monographs NMR Basic Principles and Progress. Heidelberg: Springer-Verlag.

Jonas J (1993) High pressure NMR studies of chemical and biochemical systems. *Proceedings NATO ASI, Series C, High Pressure Chemistry, Biochemistry and Materials Science*, **401**: 393–441. Dordrecht: Reidel.

Jonas J and Jonas A (1994) High pressure NMR spectroscopy of proteins and membranes. *Annual Reviews in Biophysical and Biomolecular Structure* **23**: 287–318.

Jonas J and Lamb DM (1987) Transport and intermolecular interactions in compressed supercritical fluids. In: Squires TE and Paulaitis ME (eds) *Supercritical Fluids Chemical and Engineering Principles and Applications*, pp 15–28. Vol 329 in ACS Symposium Series. Washington DC: American Chemical Society.

van Eldik R (1987) High pressure studies of inorganic reaction mechanisms. In: *Proceedings NATO ASI* **197**: 333–356.

# High Resolution Electron Energy Loss Spectroscopy, Applications

**Horst Conrad**, Fritz Haber Institute of the Max Planck Gesellschaft, Berlin, Germany
**Martin E Kordesch**, Ohio University, Athens, OH, USA

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
Applications

## Introduction

High resolution electron energy loss spectroscopy, known as 'HREELS', is a method for measuring the vibrational spectra of molecules adsorbed on surfaces, and some properties of clean surfaces in ultrahigh vacuum. The technical aspects of the technique, theoretical basis and some examples of HREELS will be discussed.

HREELS is an 'electrons in–electrons out' spectroscopy that uses electrons typically in the 2–20 eV energy range as a probe, and detects electrons that have undergone both elastic and inelastic scattering from a surface or adsorbed layer. In practice, electrons are detected that have lost up to about 500 meV (≈4000 $cm^{-1}$, 8.06 $cm^{-1}$ = 1 meV) relative to the primary beam energy in the scattering event. This range includes the most common vibrational modes of small molecules, and their overtones and combinations.

A major advantage of the HREELS method is the benefit derived from electron detection over optical detectors and methods, and a variety of scattering mechanisms. The major technical complexity of the HREELS method is the 'resolution' of 'the spectrometer', in this case the production of an intense, well-collimated beam of electrons with an energy width of less than 10 meV at the chosen primary energy (usually about 5 eV).

Detailed and comprehensive analysis of vibrational spectroscopy performed with electrons can be found in the classic work by Ibach and Mills. The details of the modern HREELS spectrometer are explored by the acknowledged master of the technique Harald Ibach in his monograph. A guide to surfaces and adsorbates studied with HREELS can be found in the proceedings of the 'Vibrations at Surfaces' Conferences, several of which are listed in the 'Further reading' section. A general introduction is often included in surface science or surface chemistry textbooks, for example the books by Lüth, Prutton and Somorjai. The World Wide Web is a resource for manufacturers of HREELS apparatus, e.g. Omicron Vakuumphysik GmbH (omicron-instruments.com) and LK Technologies, Inc. (lktech.com).

The ideal surfaces for study are smooth, metal single crystals that are clean and well-ordered on an atomic level. Such a specimen is 'highly reflecting' and maintains the signal in the specularly reflected electron beam. Practical surfaces range from polished single crystals, cleaved surfaces and semiconductor wafers to polycrystalline metals. The adsorbate can be a monoatomic or molecular species, such as atomic hydrogen, or large macromolecules and polymers. Insulators can be measured in special circumstances.

Many aspects of surface science and surface spectroscopy are concerned with the geometrical structure of surfaces, the composition of the surface and the identification of adatoms that may be present. Vibrational spectroscopy is a method for direct measurement of specific *chemical bonds* of adsorbed atoms and molecules, both between the adsorbate and the surface and the adatoms themselves. In the early days of HREELS, the 1970s, an added attraction for this type of spectroscopy was the ability to observe adsorbate–surface bonding modes (often < 125 meV = 1000 $cm^{-1}$), because the infrared spectrometers of the day used 'grating' spectrometers, and IR detectors that were useful only above 1600 $cm^{-1}$. The low cost and versatile Fourier transform infrared spectrometer (FTIR) and improved detector technology have eclipsed HREELS for routine surface chemical bond analysis. There are, however, some surface processes that can only be observed with electrons. Some diagnostic benefits that are related to the scattering mechanisms operative in HREELS continue to be useful for surface science. It should be noted that HREELS is usually performed on a 'known' adsorbate, with a focus on the details of a specific adsorption system. HREELS is seldom used for the identification of 'unknown' adsorbate species.

## Basic concepts and requirements

### Method

In HREELS experiments, electrons produced by a thermionic emission source are passed through an electrostatic monochromator, so that an electron beam of narrow angular width and narrow energy spread is produced that is then scattered from a surface. Electrons which scatter from the surface are analysed with an electrostatic analyser (similar or identical to the monochromator) and detected with an electron multiplier. The HREELS data is acquired as 'count rate' (= current exiting the analyser) *vs* the energy loss range scanned, from $E_p$ to ($E_p$- $E_{scan}$), for example from 5 to 4.5 eV, that includes $E_{loss}$. In a typical HREELS spectrum, a 'no-loss' or 'elastic' peak will be observed in the specular scattering direction with energy $E_p$, and a FWHM of $\Delta E_p$, and loss peaks will be observed in accordance with certain selection rules, at energies $E_{loss} = E_p - \hbar\omega$, where $\omega$ is the frequency of oscillation of the vibrational mode, and an energy width of $\Delta E_{loss} \geq \Delta E_p$. A general schematic is shown in **Figure 1**.

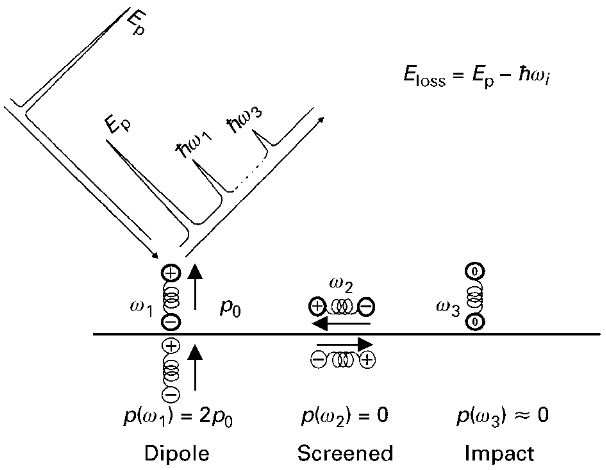



**Figure 1** An overview of high resolution electron energy loss spectroscopy. Top left: the incident electron beam is shown as a narrow, intense peak on the intensity *vs* energy loss axes. The specularly reflected beam is shown with loss peaks due to adsorbed molecules, with modes $\hbar\omega$. Centre: The scattering mechanism is illustrated with the three diatomic molecules 'adsorbed' on the surface with perpendicular and parallel orientation relative to the surface. Mode $\omega_1$ has a dynamic dipole moment $p_0$ which is perpendicular to the surface, and induces a second image dipole in the same direction, so that the electron scatters from a combined dipole moment of $2p_0$. This is the dipole scattering process. The mode $\omega_2$ is parallel to the surface, and the induced image dipole cancels the molecular dynamic dipole moment. The mode is 'screened' and is not present in the spectrum if there is no impact contribution to the scattering. Mode $\omega_3$ is shown with the dynamic dipole moment equal to zero (the orientation is not relevant). The mode will be observed as an 'impact' mode.

The 127° cylindrical capacitor analyser is currently the most advanced device for HREELS. The functional components of the HREELS spectrometer are shown as a block diagram in **Figure 2**; a representative schematic of a 127° cylinder spectrometer is also shown in **Figure 3**. There are various other HREELS instrument designs with individual merits.

HREELS operation requires specialized precautions and procedures. Many are determined by the trade-off between the spectrum intensity and resolution, the production of the low-energy electron beam, and mechanical restrictions. Practical size limits for HREELS analysers are about 100 mm radius. The thermionic electron source has an energy width of the order of 1 eV. Only a small fraction of the emitted current is passed by the monochromator. Electrons with 5 eV kinetic energy are easily deflected by stray electric and magnetic fields. All of the spectrometer parts are coated with graphite to create a uniform work function on parts in the electron path. The spectrometer itself is usually baked after the UHV chamber itself is baked, so that non-uniformities in the analyser due to unwanted adsorbed contaminants are removed. No magnetic parts are allowed in the vicinity of the spectrometer. Most HREELS chambers are doubly shielded against magnetic fields (< 1 mG $cm^{-1}$ for the residual homogeneous DC field), and all of the electric potentials applied to the spectrometer are shielded and regulated to within a few percent of 1 mV. The specimen

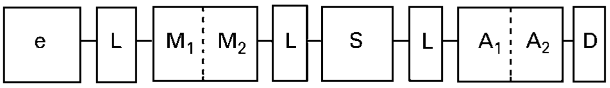



**Figure 2** The necessary part of an HREELS spectrometer: e, electron source, L, lens to focus electrons onto the entrance slit of the monochromator, M, which can have one or two stages. Another lens, L, focuses the beam onto the sample, S. A lens, L, focuses electrons scattered from the surface onto the entrance slit of the analyser, A, which can also have two stages. The electrons are counted at the detector, D.

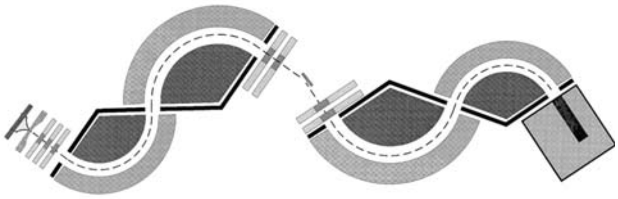


**Figure 3** A typical 127° two-stage cylindrical capacitor analyser. The individual parts correspond to the components described in **Figure 2**. Usually, one of the arms of the spectrometer is made to rotate in-plane about the specimen axis to allow off-specular measurement.

must be introduced between the HREELS monochromator and analyser with no vibration, no magnetic parts, and no 'stray' potentials such as those that may arise from charging of ceramic insulators, etc. It is often beneficial to be able to cool the specimen to 100 K or lower, and sputter and heat the surface for cleaning.

The 'mechanics' of HREELS is now manageable using personal computer based spectrometer control. Typical spectral resolution of modern instruments is 1- 5 meV. At the lower end of this range physical effects due to natural vibrational line widths are observable. In the early days of HREELS (which extend to the present, because many of the older spectrometers are still in use), single-pass cylindrical (127°) capacitor spectrometers, both fixed and with a movable analyser or monochromator, double-pass cylindrical (127°) capacitor analysers, spherical capacitor spectrometers and even some dual-use photoelectron spectrometers with enhanced resolution ranges were used, with spectral resolution in the 7- 12 meV range.

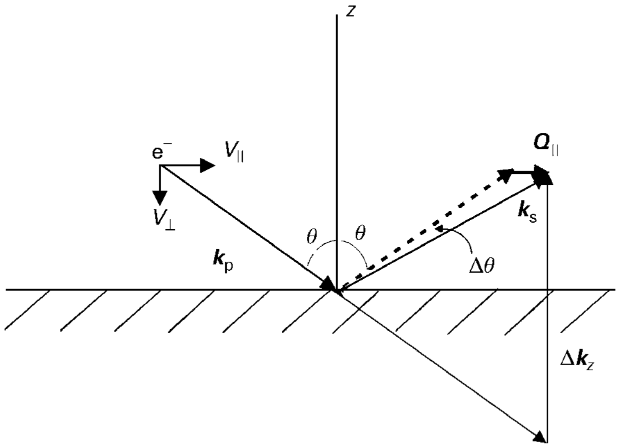



**Figure 4** A schematic representation of specular reflection ($\theta$) and momentum conservation in HREELS scattering events. The momentum parallel to the surface is conserved. The quantity $\boldsymbol{Q}_{\|}$ is the parallel momentum transfer, and determines the angular deflection $\Delta\theta$. The incident electron is characterized by velocity and momentum components $\boldsymbol{v}$ and $\boldsymbol{k}_p$, $\boldsymbol{k}_s$ is the scattered electron momentum, and $\Delta\boldsymbol{k}_z$ the change in momentum in the $z$ direction.

## Scattering mechanism

A non-mathematical description of the HREELS scattering mechanisms will be given briefly. In 'dipole scattering' an incident electron scatters from the long-range oscillating dipole field that arises when the adatom oscillates on the surface. In this mode, collecting only electrons scattered into the near-specular direction, HREELS information is essentially identical to data obtained from surface infrared vibrational spectroscopy. Electrons scatter from the long-range dipole fields of elementary excitations of the solid or adsorbate. A change in the dipole moment during the vibration is detected; consequently, there must be a non-zero dipole derivative with respect to time or a 'dynamic' dipole moment. **Figure 4** illustrates the scattering geometry for HREELS.

Energy and momentum conservation are simply expressed as: $E_{\text{loss}} = E_p \pm \hbar\omega_s$, where the momentum of the incident ('primary') electron is $k_p$, and $k_s$ for the scattered electron, so that

$$E_{\text{loss}} = \frac{\hbar^2 k_s^2}{2m} = \frac{\hbar^2 k_p^2}{2m} \pm \hbar\omega_s(Q_{\|})$$

where $Q_{\|}$ is the momentum transfer parallel to the surface, and the vector

$$\boldsymbol{k}_{s\|} = \boldsymbol{k}_{p\|} \pm \boldsymbol{Q}_{\|} + \boldsymbol{G}_{\|}$$

where $G$ is a reciprocal lattice vector (usually, $G_{\|} = (0,0)$, $G_{\|} \neq (0,0)$ is diffraction). The dipole excitations have a long-range potential, which ensures that the interaction time is large and that the probability for scattering is also high. Both losses and gains are possible.

In order to pass through the analyser entrance slit, the angular deflection (momentum transfer) caused by the collision must be small enough to keep the electron within a few degrees of the specular direction. An estimate for this deflection in dipole scattering is given by $\delta\theta = \hbar\omega_s/2E_p$, so that for 200 meV loss energy and 5 eV primary energy, the deflection is of the order 1°.

It is possible to detect electrons that are not scattered into the specular direction, either by rotation of the sample or moving the HREELS analyser or monochromator. The possibility of data collection 'off specular' sets HREELS apart from its optical counterparts. The most important 'off-specular' scattering mechanism is known as 'impact' scattering. This name is intended to denote a *short-range* inelastic scattering mechanism unrelated to the long-range dipole scattering mechanism. (The elastic 'impact' process is electron diffraction.) As a functional, but not entirely correct, definition, any intensity that remains in the 'off-specular' direction spectrum is termed 'impact' scattering. The intensity of 'impact modes' for a mode $u$ is $I_{\text{impact}} = \frac{1}{2}(Q{\cdot}u)^2 I_0$, where $u$ is the displacement of the atom from its equilibrium position, $Q$ is the change in wavevector of the electron, and $I_0$ is the elastic scattering intensity. The

wave vector transfer $Q$ can be large, so that impact scattering is not restricted to the specular direction. Experimentally, the intensity of the 'dipole' loss peaks exhibits the same dependence on off-specular angle as the elastic peak intensity. As a consequence, impact modes can be much more intense relative to the dipolar contribution at high off-specular angles. Because most analysers are constructed for about 50° incidence from the sample surface normal, 25-30° off specular are the largest practical values.

A detailed examination of the electron-adatom-surface interaction must be divided into several parts: either direct scattering from the adatom in the specular direction, specular reflection at the surface followed by (forward) scattering from the adatom, or adatom scattering followed by specular reflection. Multiple combinations of these individual events can also result in an electron detected at the analyser. For each interaction, there are elastic and inelastic versions. A schematic representation is shown in Figure 5. To date, HREELS theory has dealt only with the large divisions of scattering events into elastic/inelastic and dipole/impact processes. Other phenomena have been observed, and can be regarded as significant scattering types observed in HREELS spectra, even if they are not entirely distinct from the dipole and impact mechanisms (dipole or impact scattering may still be *part* of the process). Molecules with large static dipole moments have a large scattering cross section, so that 'dipole' scattering from a partially covered surface may cause diffuse scattering of electrons, much like a rough surface causes diffuse optical reflection. There are many forms of resonance scattering: an electron can be scattered into a surface state or resonance, causing an increased residence time at the surface, which can influence spectral features. The incident electron can induce surface resonances, such as image potential states, which also capture the electron at the surface for increased residence time and possible non-specular scattering. Molecules can capture electrons in 'negative ion resonances', which also have consequences for peak intensities and multiple or combination scattering peaks. Ionic crystals and semiconductors can have surface phonon modes in the energy range which is accessible to HREELS. Some illustrative examples are given below.

## Selection rules and symmetry

Selection rules for dipole scattering from surface-adatom symmetry configurations follow those for IR spectroscopy. Relevant discussion can be found in Ibach and Mills, and general references on symmetry, normal modes and selection rules in

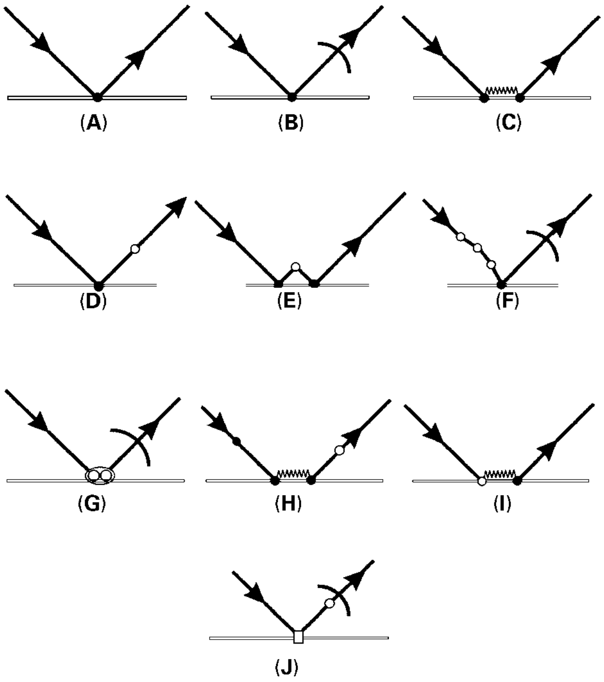



**Figure 5** A schematic representation of scattering phenomena that can be observed in HREELS spectra. The diagram notation is loosely based on McRae and Ibach and Mills. The incident electron comes from the left. In each diagram, a time-reversed event is often possible, and these are not shown. Events that involve small-angle forward scattering preserve the specularly scattered intensity, and contribute to HREELS; large-angle scattering contributes less, and some of these diagrams are omitted. In (A) the no-loss, specular scattering from a surface is shown (a filled, black circle at the surface). The no-loss or elastic peak is a result of process (A). Process (B) represents dipolar elastic diffuse scattering (denoted with an arc on the exit beam), such as observed for no-loss peak intensity decreases upon adsorption (see **Figure 14**, **Figure 15B**). Diagram (C) shows a resonant process, where an electron is scattered elastically into and out of a surface resonance (wavy line) in the specular direction. An example is a clean-surface reflectivity measurement (see **Figure 17**) with a surface resonance and/or image potential states. In these processes, in order to scatter into a weakly bound surface resonance, energy conservation by an Umklapp process may occur. Diagram (D) shows the elastic, specular reflection at the surface followed by a loss (open circle) and small angle deflection. This is 'dipole scattering'; the reverse, loss then reflection, is equally important, but not shown. Most HREELS spectral features are due to processes (A) and (D). Diagram (E) shows an unlikely, high-angle scattering, inelastic process that is detected in the specular direction. The process (F) represents multiple, small-angle inelastic scattering followed by specular reflection, but at angles unrelated to the incident electron, which results in a 'diffuse' reflected beam, even though each inelastic scattering event was nearly specular (see **Figure 14**, HCN 'ice'). Process (G) represents a molecular resonance. The electron scatters out of this resoanance isotropically. The diagrams (H) and (I) show combinations of inelastic scattering and surface resonances with specular reflection. Process (J) represents diffuse (open square box) impact scattering followed by an inelastic loss event. Many more combinations of these mechanisms are observed in special cases.

spectroscopy. Briefly, for dipole scattering, those modes with a non-zero dynamic dipole moment are allowed. Because the electric field induced by the incident electron is perpendicular to the surface, only dipoles with a component in this direction can interact with the incident electron. This implies that the normal component ($z$) of the dynamic dipole moment must be totally symmetric with respect to all of the symmetry operations of a point group. Often, the rule is stated that fundamental modes observed in dipole scattering must belong to the $A_1$, $A'$ and A totally symmetric representations. This is true for any surface that can effectively screen the parallel dipole moment of the adsorbate motion (most often, metals). The possibility remains that a mode exists that is normal to the surface, but has no dynamic dipole moment, or that a parallel mode that belongs to the correct representation (i.e. dipole allowed) could be observed due to charge transfer to orbitals that create a dynamic dipole moment normal to the surface.

The impact scattering selection rule is more complex, and can be found in Ibach and Mills.

Group frequencies for adsorbates are usually similar to those determined from IR spectroscopy for coordination compounds and from bulk measurements. Several reference works of this type are available.

## Applications

### Clean surfaces

In the 2-20 eV range, surface excitations related to the dielectric response function and free carrier density of the surface are observed. Clean-surface loss spectra can include features due to surface phonons, surface plasmons, interband transitions and surface optical phonons in ionic insulators. The probe depth for these phenomena in HREELS is about 10 nm. Transitions between surface states can also be observed in the loss spectrum. Some examples are given in the section related to hydrogen adsorption and surface states below.

Clean surfaces can be probed with electrons to determine vibrational, electronic and structural properties. Diffraction with electrons is well known; practical low-energy electron diffraction is performed with display analysers in the 50-500 eV range. Bulk electronic excitations such as plasmons are usually observed in the 10-50 eV range. While some HREELS spectrometers are capable of 200 eV primary beams for this kind of measurement, plasmons are more often investigated with analyser types designed for higher energy, such as those used for Auger electron spectroscopy or X-ray and ultraviolet photoelectron spectroscopy.

### Adsorbates on surfaces

Chemisorbed molecules Literally hundreds of adsorbate-surface systems have been investigated in the circa 20 years over which HREELS measurements have been made. A selection of work will be presented that represents the major principles of HREELS and what can be learned from this technique. It is not comprehensive, and an electronic literature search program is recommended for up-to-the-minute references to the 'state of the art'. Modes are often denoted by Greek letters, the most common are ν for stretching modes, δ for bending modes and ρ for wag, scissor or twisting modes. Aromatic molecules have several different labelling conventions.

In Figure 6 the relative importance of charge transfer upon adsorption, bond order, the dipole moment perpendicular to the surface and screening is shown in a comparison of three isoelectronic adsorbates on Pd and Cu surfaces. In the earlier discussion (Figure 1), the adsorbate-surface modes were neglected for clarity. In Figure 6, NO, CO and CN HREELS spectra are shown; these adsorbates are present on the surface as $NO^+$, CO and $CN^-$. Each of these diatomic molecules will have a

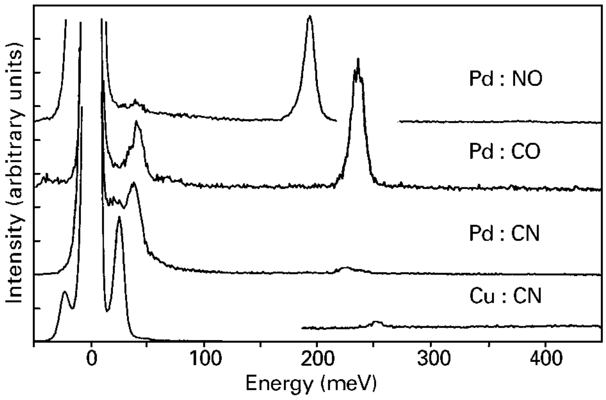



**Figure 6** The HREELS spectra of saturation coverage of adsorbed NO, CO and CN on Pd(111) and Cu(111). In each case, a metal–adsorbate (M–R, R= NO,CO,CN) mode and a molecular stretching mode (ν(A–B)) is observed. For NO and CO, the relative bond order is observed in the position of the ν(N–O) and ν(C–O) stretching modes; the relative strength of the metal–adsorbate and molecular stretching dynamic dipole moments is seen in the ratio of the M–NO to ν(N–O) modes, 1:12; M–CO to ν(C–O), 1:2. For CN on Pd(111) and Cu(111), the ν(C–N) mode is parallel to the surface. The difference in charge transfer and bond strength is seen in the ratio M–CN to ν(C–N) which is 20:1 for Pd and 190:1 for Cu. The spectra are not normalized to each other, and are not intended for absolute comparison.

stretching mode and a molecule-surface vibrational mode known as a 'frustrated translation' perpendicular to the surface. For NO and CO, the axis passing through the molecule along its length is perpendicular to the surface; for CN on Pd, the C-N axis and the CN stretching motion are parallel to the surface and partially screened, on Cu the CN motion is again parallel to the surface and hence very weak, but the Cu-$CN^-$ mode (near the elastic peak) is very intense due to the motion of the *entire* ion perpendicular to the surface. A gain is also observed. Adsorbed NO is also an ion, but the difference in charge transfer compared to adsorbed CN is significant, as is revealed by the difference in the adsorbate-metal modes.

Using a modern spectrometer with better resolution, Figure 7 shows a spectrum of an ordered OH overlayer on Ag(110) generated by the reaction of water with pre-adsorbed oxygen (the oxygen spectrum is the lower curve in Figure 7). The OH radical is stabilized by the chemisorption bond to the silver surface and forms an ordered $(1 \times 2)$ layer. Note that the δ(O-H) mode is broadened relative to the elastic peak, and that phonons due to the silver surface are observed as both losses and gains (45, 83 $cm^{-1}$).

In Figure 8, the use of dipole selection rules to determine molecular bond orientation is shown. On Cu(111), for adsorbed CN, the CN axis is parallel to the surface. The C-N stretching mode is weak (screened), and measured to lie at 254 meV (2045 $cm^{-1}$); the Cu-CN mode is very strong. When CN is adsorbed on an oxygen pre-covered surface, the C-N stretching mode is observed at a higher frequency, 266 meV (2140 $cm^{-1}$), and several times more intense. No new modes are observed, so that

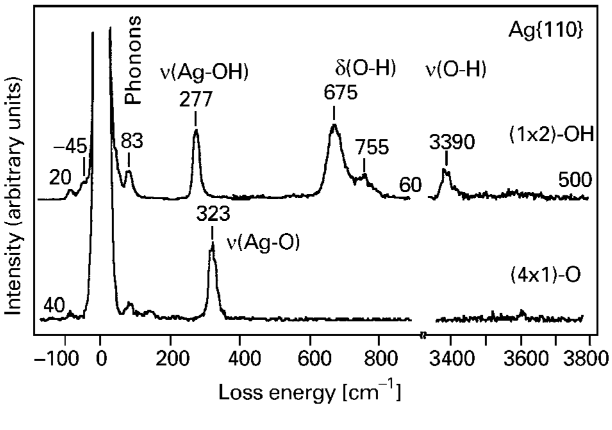



**Figure 7** HREELS spectrum of OH and Ag(110) [top], and O [bottom]. A modern HREELS spectrometer was used with 2 meV elastic peak width. Note phonon modes and the width of the δ(O–H) relative to the ν(Ag–OH) mode.

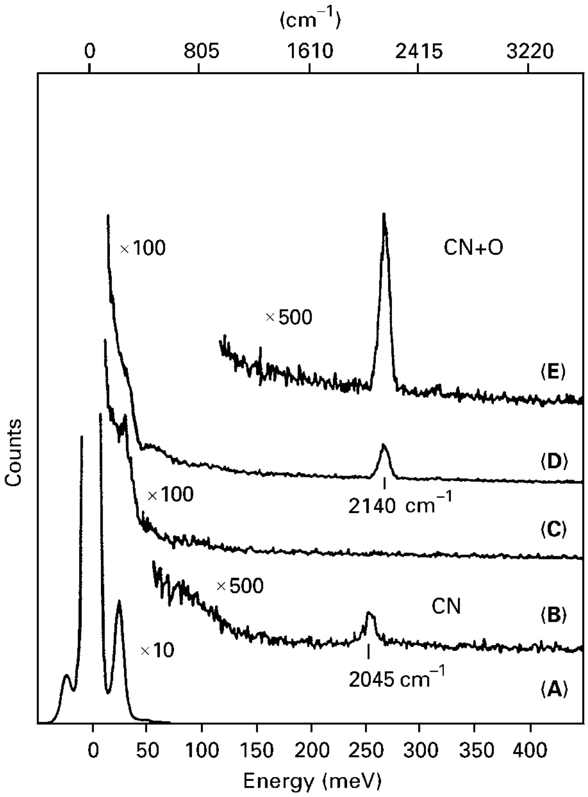



**Figure 8** The reorientation of absorbed CN on oxygen pre-covered Cu(111). The bottom spectra (A) and (B) show 100 L ($L = 1 \times 10^{-6}$ torr $s^{-1}$) $C_2N_2$ dissociatively (as CN) absorbed on Cu(111) at 300 K. The loss and gain of the CN–Cu absorbate–metal frustrated translation is clearly evident at ±24 meV. Spectrum (C) shows the Cu(111) surface after 60 L oxygen exposure at 300 K. Spectra (D) and (E) show the surface after exposure to 20 L $C_2N_2$ after the oxygen exposure. The CN stretching mode is more intense, and shifts from 254 to 266 meV, and shows a 'dipolar' angular dependence. Reprinted from *Journal of Electron Spectroscopy and Related Phenomena*, **44**, M.E. Kordesch, HREELS of the adsorption and reaction of CN on clean and oxygen covered Cu (111), 154, 1987, with permission from Elsevier Science.

an O-CN bond is not formed, and the Cu-CN mode is reduced in intensity. These spectra are interpreted to indicate that the CN molecule initially bonds to the surface as an ion, with the CN axis parallel to the surface. In the presence of pre-adsorbed oxygen, the vertical bonding geometry, with the CN axis perpendicular to the surface, is observed. The shift in frequency is consistent with this observation, and the fact that oxygen adsorbs on Cu as $O^-$ and thereby alters the CN bond geometry. Similar orientation effects are observed for molecules such as pyridine and benzene, thiophene, furan and pyrrole. A molecule may be observed to reorient as a function of surface coverage (i.e. a second layer), or as a function of the packing density of the chemisorbed layer (reorientation in the first layer).

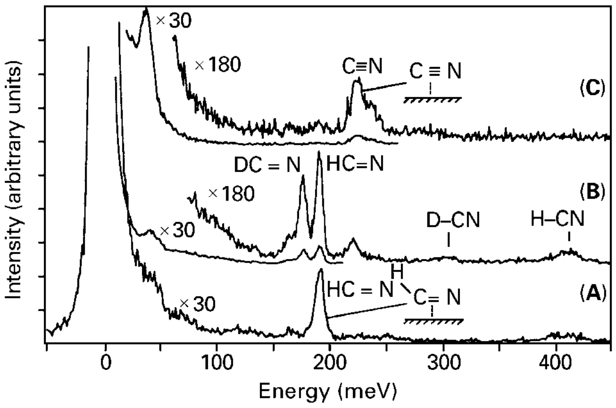



**Figure 9** Deuteration and hydrogenation of CN to form DCN and HCN in situ during HREELS observation. (A) HREELS of 10 L HCN at 300 K. (B) 10 L NCCN at 300 K, after 30 min exposure to $D_2$ at $1 \times 10^{-8}$ torr and hydrogen form the residual gas. (C) 10 L NCCN at 300 K. The absorbed CN comibnes to form DCN and HCN. Reprinted from *Surface Science*, **175**, M.E. Kordesch, The hydrogenation of CN on Pd(111) and Pd(100), L687–L692, 1986, with permission from Elsevier Science.

Another useful diagnostic method in HREELS (and vibrational spectroscopy in general) is the use of isotopic substitution to identify a particular vibrational mode by substitution of a different mass isotope and determining the shift of the mode frequency upon substitution. In Figure 9, the HCN molecule adsorbed on Pd(111) at 300 K is exposed to deuterium while the HREELS spectrum is monitored. New modes are observed; the C–H mode is replaced with a C–D mode at a new frequency related to the change in the mass of D relative to H. The altered mass of DCN relative to HCN also shifts the bending and molecule–substrate modes.

New HREELS spectrometers are able to distinguish shifts due to oxygen, carbon and nitrogen isotopes, among others, with no alteration in the adsorption geometry, simplifying the interpretation of the spectra. The shifts are not as simple as the H/D substitution, and must be calculated from first principles. In Figure 10, the reaction of isotopically labelled $CO_2$ with preadsorbed O to form adsorbed carbonate on Ag(111) is shown, with the relative

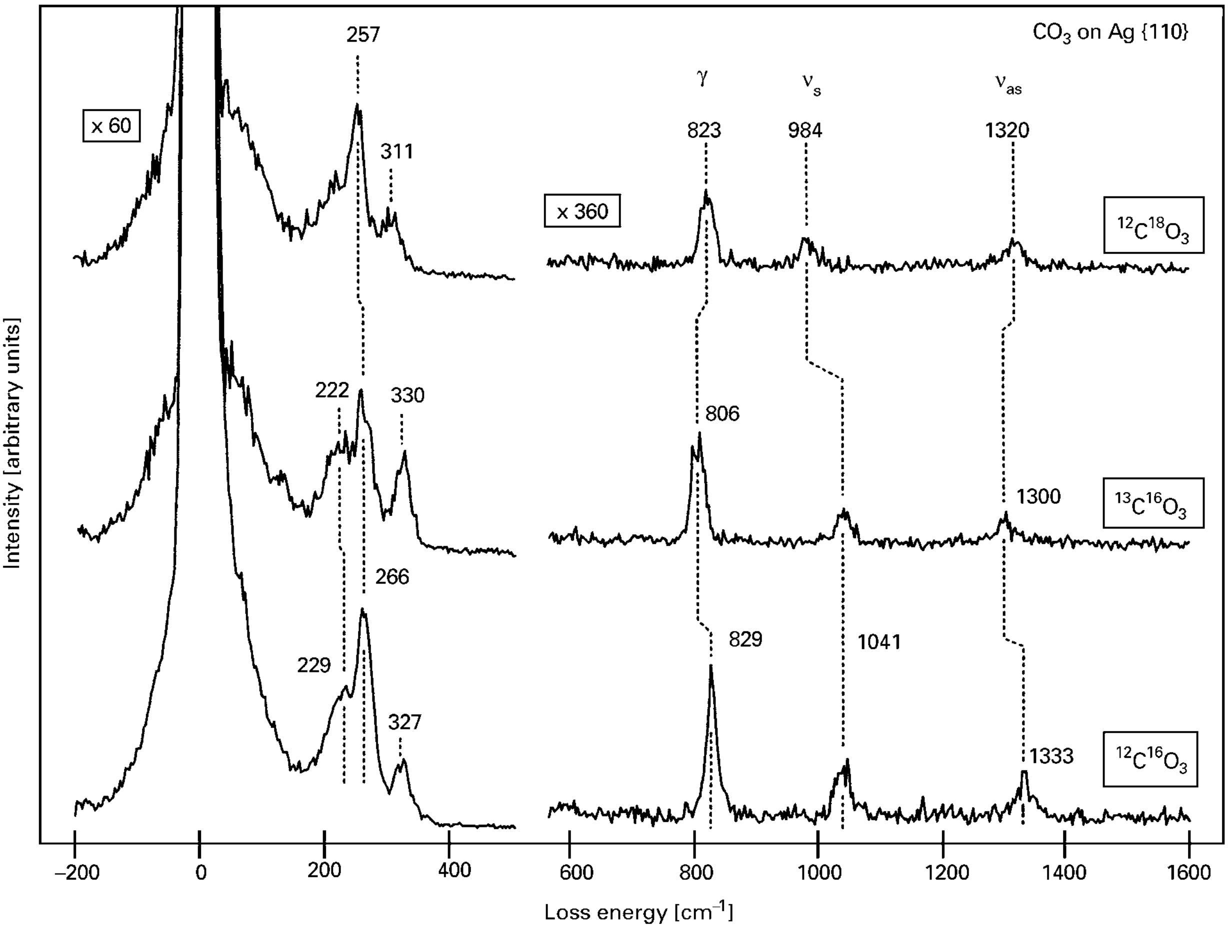



**Figure 10** Carbonate formation on Ag(110). The new HREELS spectrometers can distinguish isotopically labelled molecular vibrations. The shifts are complex and must be calculated from first principles.

shifts of the peak positions due to the differences in isotope masses. The assignment of the carbonate modes is included in Figure 11. The frequencies are for $(CO_3)^-$ ions in solution. The delta-mode is not seen in specular HREELS. The extrinsic modes of $CO_3$ are at 229 and 266 cm$^{-1}$ for the normal carbonate; the higher mode is assigned to the substrate stretch, the lower mode is a libration mode. The losses around 330 to 311 cm$^{-1}$ are due to residual oxygen not bonded in a carbonate, shifted in accordance with the isotopic substitution, because the layers are prepared by pre-adsorption of $^{18}O$ or $^{16}O$ at 300 K and forming in the (4×1) reconstructed surface. During the reaction (exposure to $CO_2$, $^{13}C^{16}O_2$ or $^{12}C^{18}O_2$) only part of the oxygen reacts, the rest is compressed into the higher coverage reconstruction, the (2 × 1) surface, and survives. The LEED pattern is a superposition of (2 × 1)O and (1 × 2) $CO_3$.

Reactions can be observed in HREELS, most commonly the dissociation of large molecules into smaller fragments after heating. A correlation with temperature-programmed desorption is very useful. The reaction of small molecules to form larger compounds is also of interest. In Figure 12, the reaction of benzonitrile with oxygen on Cu(111) to form an 'N-bonded oxide' (or fulminate) is shown. The two reactants are adsorbed at low temperature, where no reaction occurs. The substrate is heated to successively higher temperatures and then observed in HREELS. The benzonitrile is adsorbed with the plane of the benzene ring parallel to the surface, and the benzene- CN bond is observed to remain intact in the spectrum during the reaction. At 170 K the new

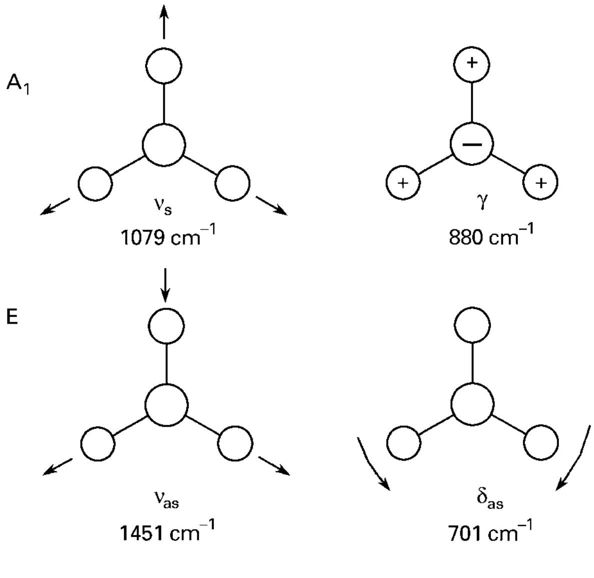



**Figure 11** Normal modes for the carbonate ion.

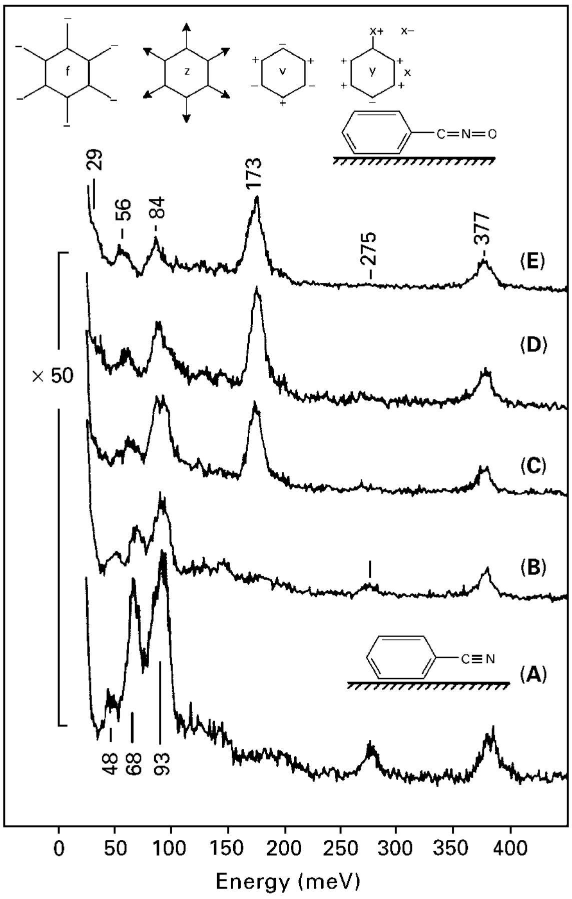



**Figure 12** The reaction of benzonitrile and oxygen on Cu(111) to form an 'N-bonded oxide' or CN–O compound. Top left: Whiffen mode symbols for the monosubstituted benzenes: f: in-phase H motion relative to the ring plane (93 meV), Z: C–H stretching motion (377 meV), v: ring torsion (84–86 meV), y: a substituent-sensitive mode where the substituent (small x at the top of the benzene ring) moves out-of phase with the ring C atom opposite to the point of attachment (86–56 meV), x: the in-phase version of y (unresolved). (A) 100 L oxygen + 10 L Phenyl-CN (benzonitrile) at 100 K. (B), as (A) heated to 210 K, (C) heated to 250 K, (D) heated to 300 K. (E) 100 L oxygen + 10 L PhCN exposure at 300 K. The CN–O bond is observed at 173 meV. Reprinted from *Surface Science*, **211/212**, M.E. Kordesch, An HREELS characterization of a surface CN-oxygen compound formed on Cu(111), 1047, 1989, with permission from Elsevier Science.

CN- O stretching band is observed at 171 meV (1370 cm$^{-1}$). This example illustrates how HREELS is applied to specific chemical bonds.

The deuterium substitution reaction mentioned above raises the question of time-resolved HREELS. Some attempts have been made to use parallel detection methods in HREELS, so that time-resolved spectra can be acquired. The experimental realization of this method is complex; Wilson Ho has pioneered the implementation of time-resolved HREELS, and some results are given in his review.

FTIR instruments may prove more efficient in this area.

Large molecules pose a difficult interpretation task for HREELS. A molecule with many, closely spaced modes will produce a spectrum with unresolved vibrational modes. A large molecule may not have a single adsorption geometry, making the application of selection rules difficult. Some success has been achieved with polymers. An example is shown in Figure 13; the observation of Fomblin Y, a perfluoropolyether oil, on polycrystalline molybdenum. The polymer has a molecular mass of over 10 000, and contains several different monomer constituents. The fluorine groups, and some pendant group modes, can be observed. The individual fluorine groups, CF, $CF_2$, $CF_3$ are not usually resolved (157 meV, 1256 $cm^{-1}$), even in IR spectra, but several overtones of the C-F modes are observed. Useful information can still be obtained, even from very large molecules or polymers.

As the surface coverage increases, as in a condensed multilayer 'ice', the dipole selection rules are relaxed, because the screening of parallel modes is less effective in the outermost layers of the 'ice', because these layers are further from the surfaces. In Figure 14, a HREELS spectrum of HCN multilayer 'ice' condensed on Pd at 100 K is shown. The HCN molecule has three normal modes, ν(C-H) and ν(C-N) stretching modes along the molecule axis, and a doubly degenerate δ(H-CN) bending mode. There are at least 20 modes evident in this spectrum. At low temperature, HCN is thought to form hydrogen-bonded chains (a 'catamer'). The HCN molecule in the chain has a strong static dipole moment, and the motion of the chains parallel to the surface (a libration, or frustrated rotation) serves to induce a strong dynamic dipole moment in the direction perpendicular to the chains. The libration is observed along with the usual HCN modes.

The HCN 'ice' spectrum illustrates several aspects of HREELS. The strong dynamic dipole moment due to the libration mode causes multiple scattering from the molecules, so that there are several replica 'primary' peaks almost as intense as the elastic peak. Consequently, each of the usual HCN loss peaks generates a spectrum with the same spectral intensity relationships as the 'no-loss' peak. The multiple

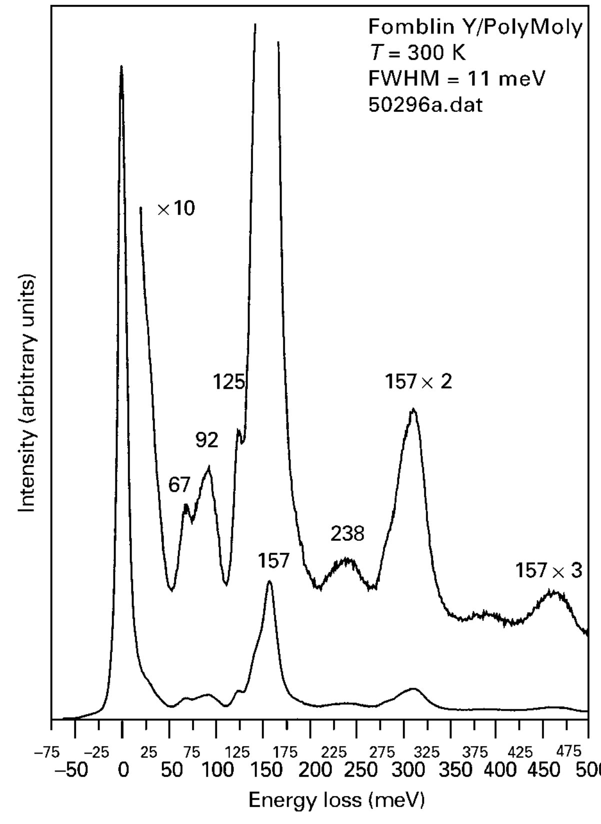



**Figure 13** HREELS spectrum of Fomblin Y on polycrystalline molybdenum. The CF, $CF_2$, $CF_3$ modes are unresolved, at 157 meV. Overtones at 314 and 471 meV are observed. Other modes: δ(OCO) = 67 meV, ν(OCO) or ν(CC) = 92 meV, the mode at 125 meV is characteristic of Fomblin Y, a stretching mode in a monomer segment, $\nu(CF_3\text{–}C\text{–}CF_2)$.

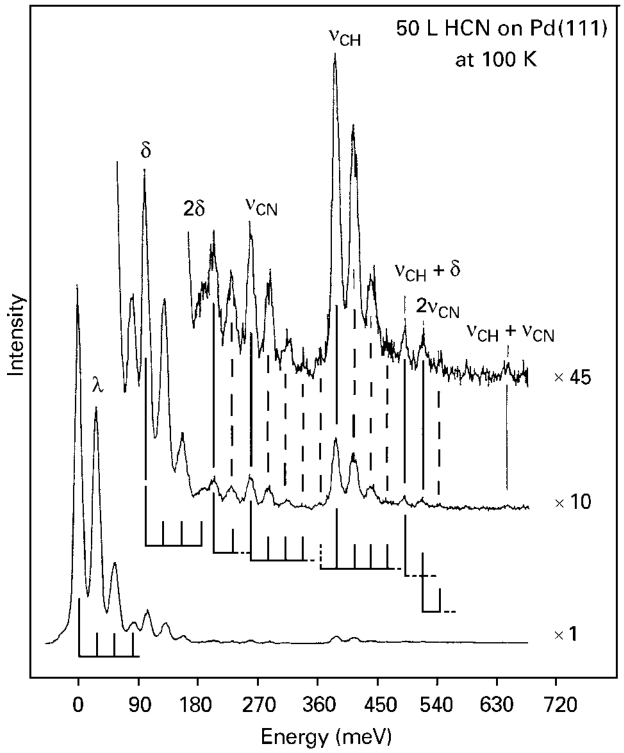



**Figure 14** HREELS spectrum of 50 L HCN exposure at 100 K on Pd(111). The libration mode λ, the HCN bending mode δ, and the CN and CH stretching modes $\nu_{CH}$ and $\nu_{CN}$ are identified with some of their combinations. The vertical ladders show some of the Poisson series for the libration mode.

scattering intensities follow the Poisson relationship for multiple random events. The chain formation distorts the molecular potential, so that overtones and combinations are possible due to anharmonic coupling of the normal modes.

A common phenomenon observed upon adsorption of molecules is the decrease in the elastic peak intensity as a function of coverage. Figures 15A, B show the intensity and angular width of the elastic peak as a function of HCN coverage on Pd. The intensity correlates well with layer formation on the surface, much like RHEED oscillations. Even though the scattering mechanism is decidedly 'dipolar', the surface is made 'rough' by a partial coverage of HCN. As the surface layer is completed, the specular beam sharpens, but the phenomenon is repeated as the second layer adsorbs.

In dielectric layers, where a surface phonon mode may occur, or in ionic crystals, multiple scattering from the surface phonon mode can result in Poisson 'replicas' of the no-loss peak. These modes are referred to as 'Fuchs-Kliewer' modes; they are a general feature of HREELS spectra of ionic and polar materials, and metal oxides. Ordered overlays on surfaces can also exhibit collective modes, but at submonolayer coverages the HREELS loss peaks are due almost exclusively to single oscillations of the fundamentals. Substrate (silver) phonon modes are shown at 10 meV (83 $cm^{-1}$) in Figure 7.

Single crystals are understandably the most interesting substrates for surface science, because other spectral data and calculations are made easier by an ordered surface. For technology, however, such as lubrication studies, or adhesion, a polycrystalline surface is more relevant. Some success has been achieved with polycrystalline metal surfaces, but a case-by-case test for suitability is necessary. The spectrum in Figure 13 (Fomblin Y perfluoropolyether) was nearly identical to that obtained on Mo(100), with the exception of a slightly broader elastic peak.

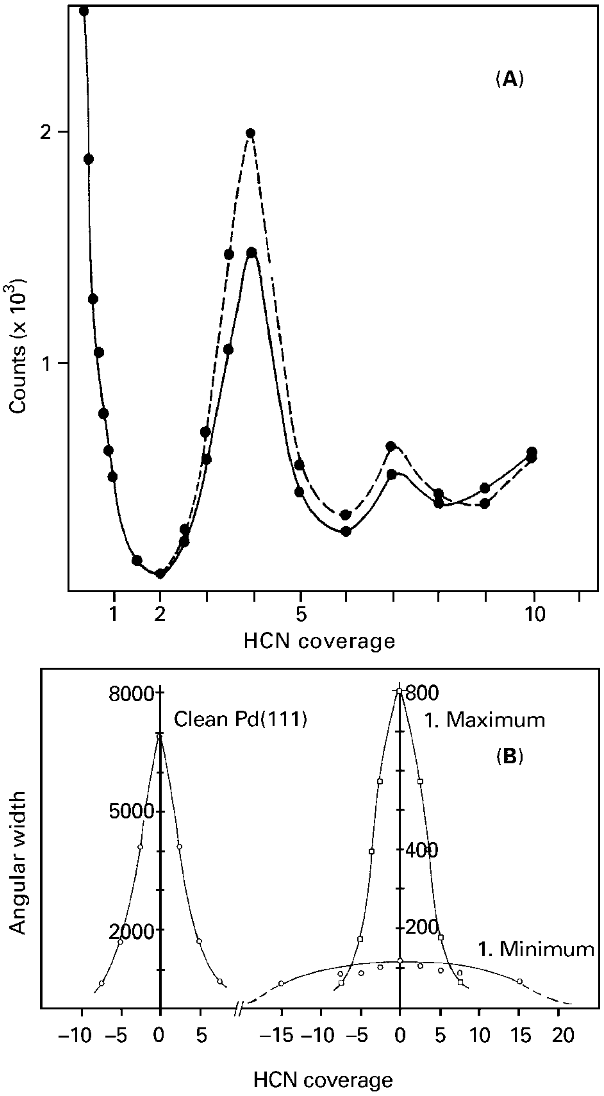
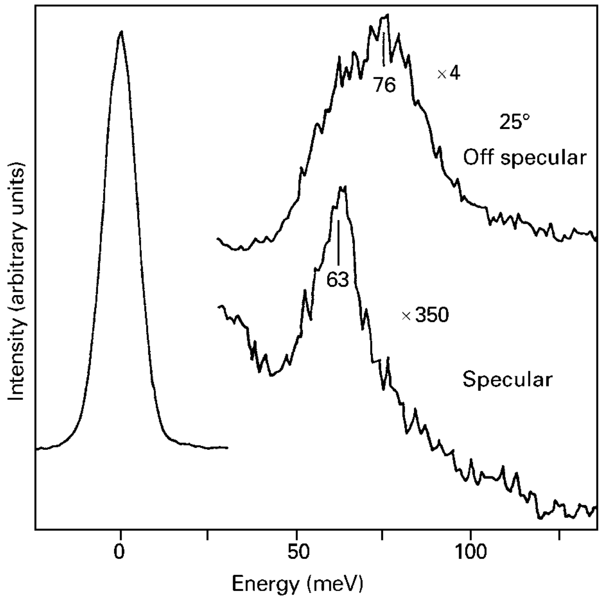



**Figure 15** (A) The intensity of the no-loss peak as a function of HCN exposure in Langmuir at 100 K. The dotted line is with the spectrometer optimized to compensate for work function changes, the solid line is 'as measured'. (B) The angular width of the no-loss peak, measured for clean Pd(111), at the first minimum, 2 L, and at the first maximum in **Figure 9A**, about 4 L.

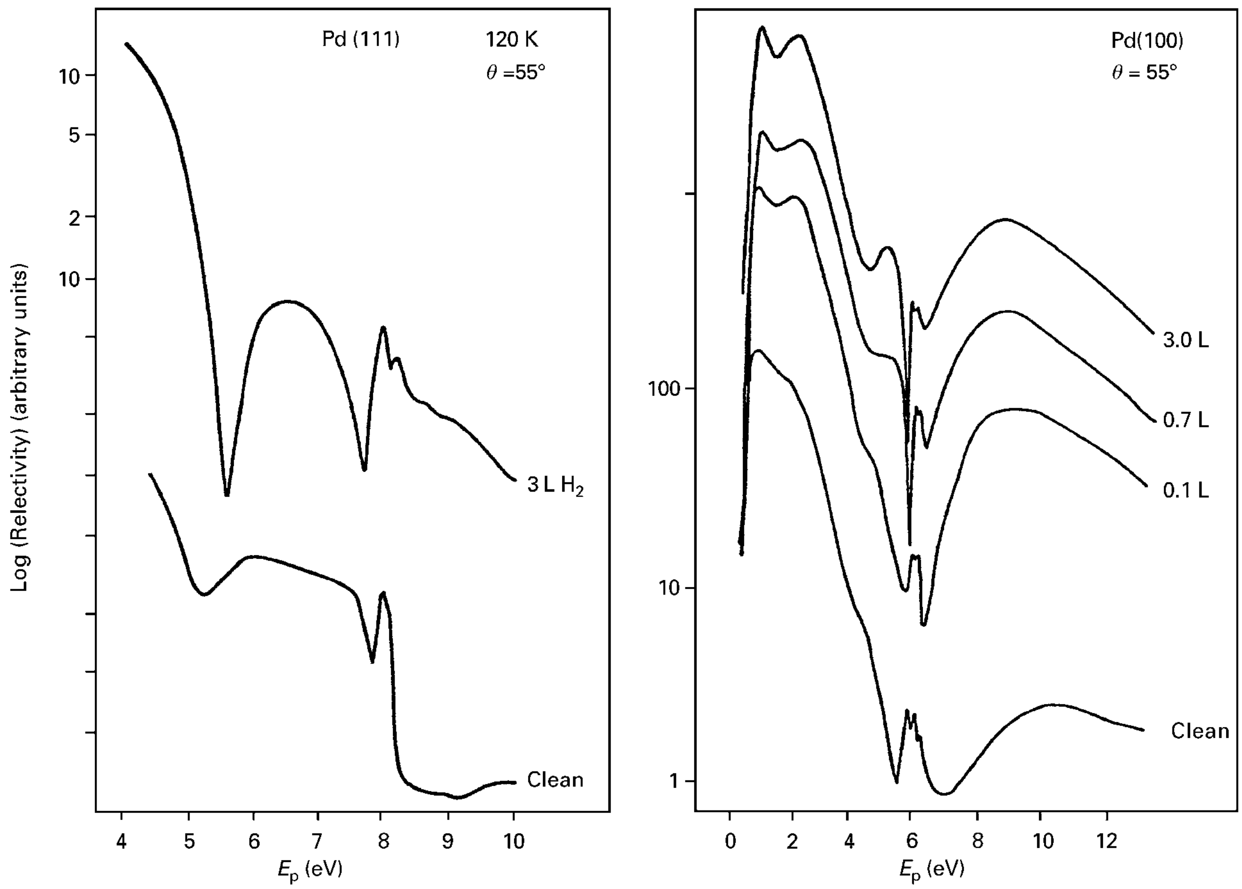



**Figures 16** Specular and off-specular HREELS spectra of H adsorbed on Pd(100). The perpendicular mode is at 63 meV. The parallel mode is at 76 meV. Reprinted from *Surface Science*, **178**, M.E. Kordesch, Surface resonances in vibrational spectroscopy of hydrogen on transition metal surfaces: Pd(100) and Pd(111), 578–588, 1986, with permission from Elsevier Science.

Hydrogen Hydrogen is difficult to detect with many surface analytical methods. In HREELS, the adsorption of hydrogen and deuterium is relatively easy, both experimentally, and in 'first instance' interpretation of the spectral features. In the case of hydrocarbons and other molecules, the CH, NH or OH modes are often studied. The CH modes in aromatic molecules are observed to be partially 'impact' modes. Figure 7 shows several OH modes; in particular it is possible to observe natural line width broadening for the δ(O- H) mode on Ag(110) with the new high resolution spectrometers.

On many transition metals, hydrogen adsorbs dissociatively, with atomic hydrogen adsorbed in high-symmetry sites. The modes perpendicular to the surface are usually dipole modes; parallel, frustrated translation, modes have some impact scattering character and can be observed in off-specular spectra. Figure 16 is an example of specular and off-specular HREELS spectra of H adsorbed on Pd(111).

Atomic hydrogen is small and adsorbs near to the surface. H modes can directly couple to surface states and resonances, such as image potential states, that may overlap in the near-surface region. An empirical verification of this phenomenon, observed on several metals (Pd, Pt, Rh, Ru) is the enhancement of surface resonances by adsorbed H, and the enhancement of H vibrational modes at primary energies which correspond to the population of the surface resonance with electrons from the incident HREELS beam. In Figure 17, the reflectivities of the Pd(111) and (100) surface with and without adsorbed hydrogen are shown. The relative intensities of the H frustrated translation and rotation (perpendicular and parallel) modes are shown in Figure 18.




**Figures 17** Surface reflectivity of Pd(111) and Pd(100) with and without adsorbed H. The reflectivity is on log scale. The addition of hydrogen shifts and intensifies the lowest energy surface resonance on Pd(111) (5.5 eV). The sharp drop in reflectivity at 8 eV corresponds to the emergence of a surface diffraction beam, and opens a new channel for electron interaction with the surface. The image potential states are just below this emergence threshold. On Pd(100) the curves are similar, but the energy scale is reduced due to the different crystal structure of the surface and different-sized surface Brillouin Zone. Reprinted from *Surface Science*, **178**, M.E. Kordesch, Surface resonances in vibrational spectroscopy of hydrogen on transition metal surfaces: Pd(100) and Pd(111), 578–588, 1986, with permission from Elsevier Science.

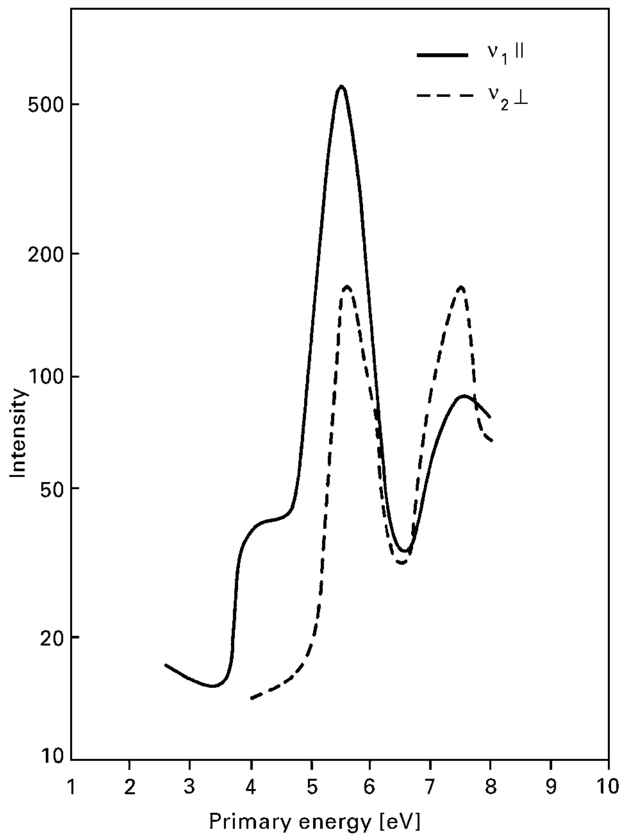



**Figure 18** The intensity of the parallel and perpendicular H Vibrational modes at different primary energies on Pd(111). The maxima in the vibrational enhancement correspond to minima in the electron reflectivity curves. Reprinted from *Journal of Electron Spectroscopy and Related Phenomena*, **38**, M.E. Kordesch, Surface resonances on Pd(111)/H observed with HREELS, 289–298, 1986, with permission from Elsevier Science.

## List of symbols

A = symmetry group designation; $G$ = reciprocal lattice vector; $I$ = intensity; $k$ = crystal momentum; $m$ = electron mass; $p$ = dipole moment; $\Delta E$ = energy spread.

*See also:* **Hydrogen Bonding and Other Physicochemical Interactions Studied By IR and Raman Spectroscopy; IR Spectral Group Frequencies of Organic Compounds; IR Spectroscopy, Theory; Photoelectron Spectrometers; Polymer Applications of IR and Raman Spectroscopy; Surface Studies By IR Spectroscopy; Symmetry in Spectroscopy, Effects of.**

## Further reading


Harris DC and Bertolucci MD (1989) *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy.* New York: Dover Publications, Inc.

Herzberg G (1989) *Molecular Spectra and Molecular Structure: I. Spectra of Diatomic Molecules, II. Infrared and Raman Spectra of Polyatomic Molecules.* Malabar, Florida: Robert E. Krieger Publishing Co.

Ho W (1993) High resolution electron energy loss spectroscopy. In: Rossiter BW and Baetzold RC (eds) *Investigations of Surfaces and Interfaces – Part A.* Physical Methods of Chemistry Series, 2nd edn, Vol IXA, Ch 4, pp 209-320. New York: Wiley.

Ibach H and Mills DL (1982) *Electron Energy Loss Spectroscopy and Surface Vibrations.* New York: Academic Press.

Ibach H (1990) *Electron Energy Loss Spectrometers, Springer Series in Optical Sciences,* Vol 63. Berlin: Springer Verlag.

Kesmodel L (1995) High resolution electron energy loss spectroscopy. In: Hubbard AT (ed) *The CRC Handbook of Surface Imaging and Visualization,* pp 223-237. Boca Raton: CRC Press, Inc.

Lüth H (1993) *Surfaces and Interfaces of Solids,* 2nd edn. Berlin: Springer Verlag.

McRae EG (1979) Electronic surface resonances of crystals. *Reviews of Modern Physics* 51: 541-568.

Nakamoto K (1986) *Infrared and Raman Spectra of Inorganic and Coordination Compounds,* 4th edn. New York: Wiley.

Prutton M (1994) *Introduction to Surface Physics.* Oxford: Clarendon Press.

Somorjai GA (1993) *Introduction to Surface Chemistry and Catalysis.* New York: John Wiley and Sons.

Weidhem J, Mueller U and Dehnike K (1982) *Schwingungsspektroskopie.* New York: Thieme.

Yates JT Jr (1998) *Experimental Innovations in Surface Science.* New York: Springer Verlag.


### Vibrations at Surfaces Conferences:


1st: Kernforschungsanlage Jülich Reports, Jülich, Germany, Conference 26 (1978).

2nd: Caudano R, Gilles J-M and Lucas AA (1982) *Vibrations at Surfaces.* New York: Plenum.

3rd: Brundle CR and Morawitz H (1982) *Vibrations at Surfaces 1982.* Amsterdam: Elsevier.

4th: King DA, Richardson NV and Holloway S (eds) (1986) *Vibrations at Surfaces 1985.* Amsterdam: Elsevier.

5th: Bradshaw AM and Conrad H (eds) (1987) *Vibrations at Surfaces 1987.* Amsterdam: Elsevier.

6th: Chabal YT, Hoffmann FM and Williams GP (eds) (1990) *Vibrations at Surfaces 1990.* Amsterdam: Elsevier.

7th: Proceedings in *Journal of Electron Spectroscopy and Related Phenomena* 64/65 (1993).

8th: Proceedings in *Surface Science* 368 (1996).

# High Resolution IR Spectroscopy (Gas Phase) Instrumentation

**Jyrki K Kauppinen** and **Jari O Partanen**,
University of Turku, Finland

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
Methods & Instrumentation

The definition of high resolution in IR spectroscopy has been changing with time. In this article, the limit of high resolution is defined as 0.1 $cm^{-1}$. This is the full width at half height (FWHH) of the line. In scientific language, the higher the resolution, the smaller the FWHH. In history, grating and even prism spectrometers, which have a lower resolution than 0.1 $cm^{-1}$, have also been regarded as high-resolution instruments, but the development of spectroscopic instrumentation has pushed the limit of high resolution towards smaller FWHH. Nowadays, high-resolution IR spectra are generally recorded by Fourier-transform IR (FT-IR) spectrometers, or by laser spectrometers.

Michelson, in the 1880s, invented his interferometer, which was used to study the speed of light and to fix the standard metre with the wavelength of a spectral line. The use of the Michelson interferometer in spectroscopic applications was started almost 60 years later.

The first FT-IR spectrum was computed in 1949 by Peter Fellgett, who used an interferometer to measure light from celestial bodies. Transforming an interferogram into a spectrum was, however, a laborious task, and FT-IR spectroscopy for years remained the field of only a few advanced research groups, who had large computers for time-consuming Fourier transforms. The first commercial FT-IR spectrometers appeared at the end of the 1960s, when microcomputers became available. The development of the Cooley–Tukey algorithm in 1965 for a quick performance of a Fourier transform (the fast Fourier transform, FFT) was an important step for FT-IR spectroscopy. The FT-IR spectrometers of the 1960s worked with a resolution of a few wavenumbers, but they were the high-resolution technology of their own time.

There are a few commercial high-resolution spectrometer models, and research laboratories have built their own instruments, which can achieve a resolution of 0.001 $cm^{-1}$, or even better.

The history of laser spectrometers is short, starting effectively at the beginning of the 1980s.

## Basic definitions

IR radiation consists of electromagnetic waves, which oscillate with a frequency of $3 \times 10^{11}$ to $4 \times 10^{14}$ Hz. The corresponding wavelength range is $10^3$ to 0.78 μm. The infrared waves travel with the group velocity of light $c_n = c_0/n$, where $c_0$ is the velocity of light in a vacuum ($2.99792458 \times 10^8$ m $s^{-1}$) and $n$ is the refractive index of the medium. IR radiation has shorter wavelength than microwaves but longer wavelength than visible light. The IR region may be subdivided into three: far-IR ($10^3$ to 30 μm), mid-IR (30–3 μm) and near-IR (3–0.78 μm).

Generally, in spectroscopy, the interaction between the matter and the electromagnetic waves takes place via absorption or emission. The matter can absorb the electromagnetic waves with wavelength $\lambda$ in the transition $E'' \rightarrow E'$ if

$$hc/\lambda = E' - E'' \qquad [1]$$

where $h$ is Planck's constant ($6.626076 \times 10^{-34}$ J s) and $E''$ and $E'$ are lower and higher energy levels, respectively, of the atoms or the molecules. Equation [1] is a basic result of quantum physics. The emission is the inverse process, where the energy drop of the atoms or the molecules is emitted as an electromagnetic quantum $hc/\lambda$. In IR spectroscopy, generally the absorption is examined. When IR radiation is penetrating through the sample material (solid, liquid or gas) the absorption is described by Beer's law

$$I(\nu) = I_0(\nu)\exp[-\beta(\nu)Cd] \qquad [2]$$

where $I_0(\nu)$ and $I(\nu)$ are incident and transmitted intensities, respectively, $\nu = 1/\lambda$ is the wavenumber of radiation, $\beta(\nu)$ is the absorptivity, $C$ is the concentration and $d$ the thickness of the sample. For a solid sample $\beta(\nu)C = \alpha(\nu)$ is the absorption coefficient. The

fraction $[I_0(\nu) - I(\nu)]/I_0(\nu)$ is called absorption, $I(\nu)/I_0(\nu)$ transmittance, and $-\ln[I(\nu)/I_0(\nu)] = \beta(\nu)Cd$ absorbance. Typical curves $I_0(\nu)$, $I(\nu)$, $I(\nu)/I_0(\nu)$ and $-\ln[I(\nu)/I_0(\nu)]$ are shown in **Figure 1**. The absorbance as a function of the wavenumber $\nu$ is called simply the (IR) absorption spectrum.

## Fourier transform interferometers

### Michelson interferometer

**Figure 2** illustrates a basic optical layout of the original Michelson interferometer and **Figure 3** shows schematically the ray travelling at the angle $\theta$ with the optical axis of the interferometer. If the beamsplitter divides the incident intensity $I_0$ into two equal parts, the output intensity is simply $I = 2I_0(1 + \cos\phi)$, with $\phi = (2\pi/\lambda)2nd\cos\theta = 2\pi\nu x$, where $x$ is the optical path difference. If $\theta = 0$, $n = 1$, and $S_0$ is a monochromatic point source, then $x = 2d$, and

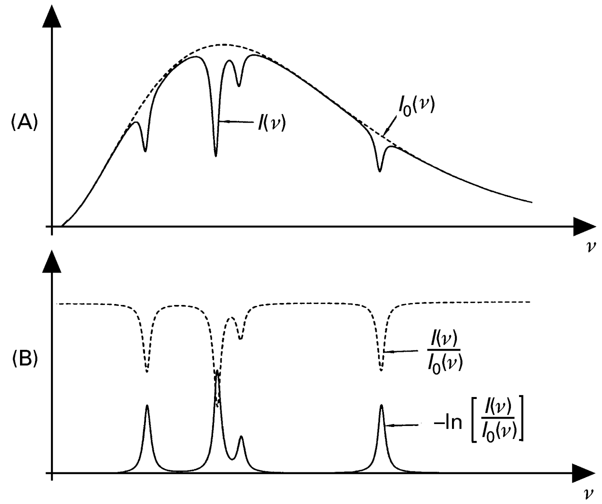



**Figure 1** (A) Incident (or background) spectrum $I_0(\nu)$ and transmitted spectrum $I(\nu)$. (B) Transmittance spectrum $I(\nu)/I_0(\nu)$ and absorbance spectrum $-\ln[I(\nu)/I_0(\nu)]$.

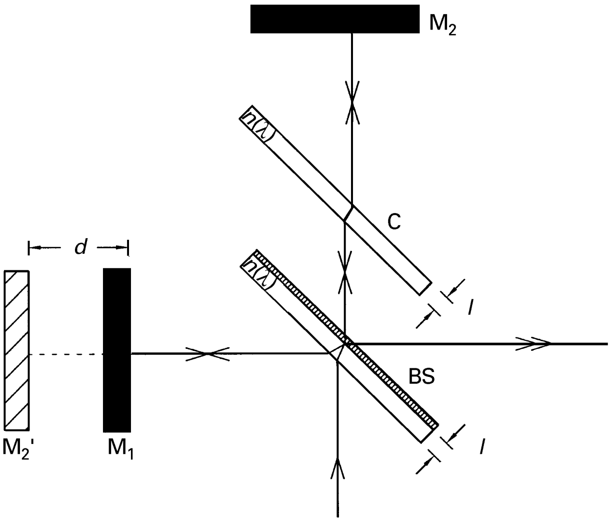



**Figure 2** Optical arrangement of the Michelson interferometer. $M_1$ is a fixed and $M_2$ a moving mirror. BS is a beamsplitter and C a compensating plate. $M_2'$ is the image of $M_2$ formed by BS.

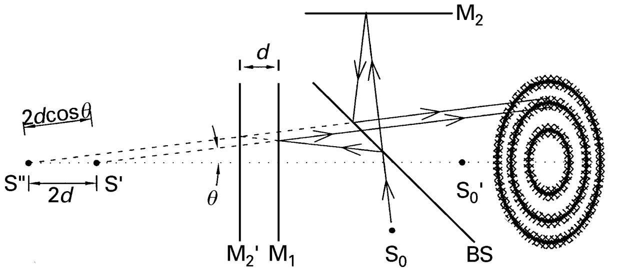



**Figure 3** Michelson interferometer produces two coherent images S′ and S″ of the real source $S_0$. These coherent sources generate the interference fringe pattern.

$$I = 2I_0[1 + \cos(2\pi\nu 2d)] \quad [3]$$

Let us assume that $S_0$ has a wide-band continuous spectrum $E(\nu)$ as shown in **Figure 4**, and it is still a point source. Now the interference signal from the region between $\nu$ and $\nu + d\nu$ is given by $dF(x,\nu) = 2E(\nu)[1 + \cos(2\pi\nu x)]\,d\nu$, according to Equation [3]. At a given optical path difference $x$, the total signal is

$$F(x) = 2\int_0^\infty E(\nu)[1 + \cos(2\pi\nu x)]\mathrm{d}\nu \quad [4]$$

where $F(x)$ is called an interference record. By subtracting from $F(x)$ the constant term

$$\frac{1}{2}F(0) = 2\int_0^\infty E(\nu)\mathrm{d}\nu \quad [5]$$

an interferogram is obtained:

$$I(x) = F(x) - \frac{1}{2}F(0) = 2\int_0^\infty E(\nu)\cos(2\pi\nu x)\mathrm{d}\nu \quad [6]$$

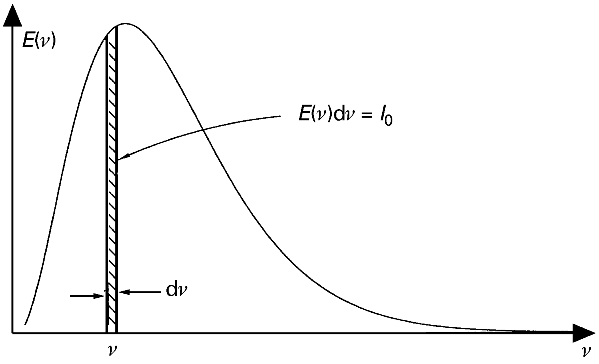



**Figure 4** Wide band continuous spectrum $E(\nu)$ can be expressed as the sum of monochromatic infinitely narrow ($d\nu$) lines.

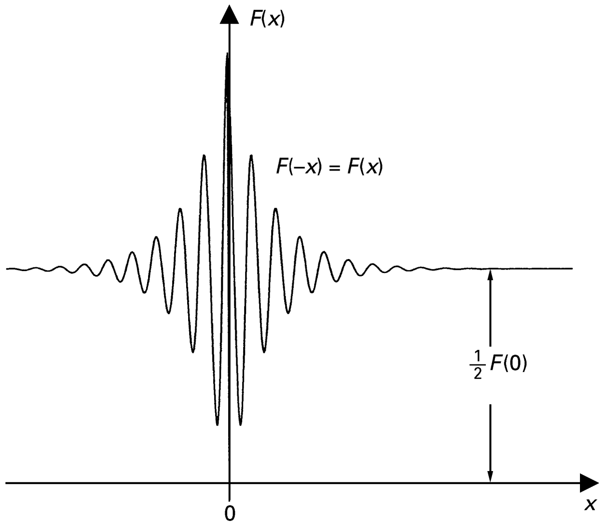



**Figure 5** Typical interference record $F(x)$.

If it is assumed that $E(-\nu) = E(\nu)$, the computations become easier, because

$$I(x) = \int_{-\infty}^{\infty} E(\nu)\cos(2\pi\nu x)\mathrm{d}\nu = \int_{-\infty}^{\infty} E(\nu)\exp(\mathrm{i}2\pi\nu x)\mathrm{d}\nu = \mathcal{F}\{E(\nu)\} \quad [7]$$

Hence, $I_x$ and $E(\nu)$ form a Fourier transform pair, and they can be written as follows

$$I(x) = \int_{-\infty}^{\infty} E(\nu)\exp(\mathrm{i}2\pi\nu x)\mathrm{d}\nu = \mathcal{F}\{E(\nu)\}$$
$$E(\nu) = \int_{-\infty}^{\infty} I(x)\exp(-\mathrm{i}2\pi\nu x)\mathrm{d}x = \mathcal{F}^{-1}\{I(x)\} \quad [8]$$

where $F\{\ \}$ and $F^{-1}\{\ \}$ are the Fourier transform and the inverse Fourier transform, respectively. Furthermore, one can notice that $F\{F^{-1}\{I(x)\}\} = I(x)$. A typical interference record is shown in **Figure 5**.

Because in practice it is impossible to make the inverse Fourier transform $F^{-1}\{I(x)\}$ of the continuous $I(x)$, one has to use the discrete Fourier transforms. Thus the interferogram $I(x)$ is sampled at the discrete points $x_j = jh$, let us say, between $x = -L$ and $x = L$. The spectrum is now given by

$$E^h(\nu) = h\sum_{j=-N}^{N-1} I_j \exp(-\mathrm{i}2\pi\nu jh) \quad [9]$$

In practice, the discrete Fourier transform is computed by the FFT in order to minimize the needed computation time. It can be shown that

$$E^h(\nu) = W^h(\nu) * E(\nu) \quad [10]$$

where * means convolution, and the instrumental profile due to the truncation of the interferogram at $\pm L$ and the discrete sampling is given by

$$W^h(\nu) = \sum_{k=-\infty}^{\infty} 2\,\mathrm{sinc}\left[2\pi\left(\nu + \frac{k}{h}\right)\right] \quad [11]$$

And now the instrumental resolution (FWHH of the instrumental function) is given by $\Delta\nu = 1.207/(2L)$, where $L$ is the maximum optical path difference (please note, a point source is assumed). Furthermore, the optimum sampling interval is given by $h = 1/(2\nu_{max})$ if $E(\nu) = 0$ with $|\nu| > \nu_{max}$. The optimum sampling interval guarantees there is no aliasing in the spectrum.

## Collimated beam and extended source in Michelson interferometer

A practical Michelson interferometer (**Figure 6**) has an extended source, as shown in **Figure 7**, and a collimator mirror. The parabolic collimator mirror may be off-axis, as in **Figure 6**, or on-axis, as in **Figure 7**. When $a$ denotes the area of the source, $A$ the area of the collimator mirror, $f$ the focal length

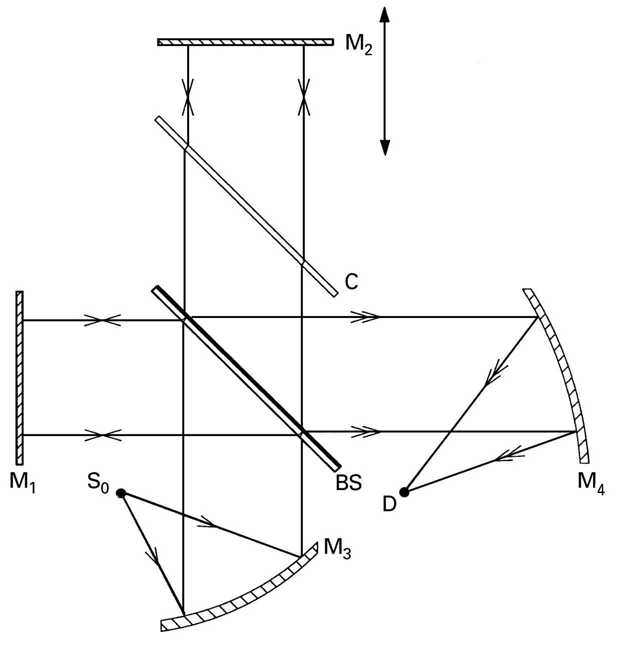



**Figure 6** Optical layout of a practical Michelson interferometer. $S_0$ is a black-body source and D an IR detector; $M_3$ and $M_4$ are parabolic off-axis mirrors. $M_3$ and $M_4$ may also be normal parabolic mirrors as in **Figure 7**.

of the collimator, $\Omega_0 = A/f^2$ the solid angle of the collimator, and $\Omega = a/f^2$ the solid angle of the source, then the signal is proportional to

$$aA/f^2 = \Omega A = a\Omega_0 \qquad [12]$$

On the other hand, the optical path difference is now $x \cos\alpha \approx x[1 - \Omega'/(2\pi)]$, where $x = 2nd$ and the angle $\alpha$ between the optical axis and the beam varies from 0 to $\theta$, which is the maximum of $\alpha$, while $\Omega'$ is changing from 0 to $\Omega$. Thus the total interferogram is

$$I_\Omega(x) = \int_0^\Omega I\left[x\left(1 - \frac{\Omega'}{2\pi}\right)\right]\mathrm{d}\Omega' =$$
$$\Omega \int_{-\infty}^{\infty} E(\nu)\mathrm{sinc}\left(\frac{\nu x \Omega}{2}\right)\exp\left[\mathrm{i}2\pi\nu x\left(1 - \frac{\Omega}{4\pi}\right)\right]\mathrm{d}\nu \qquad [13]$$

In the case of a point source,

$$\lim_{\Omega \to 0} I_\Omega(x) = \int_{-\infty}^{\infty} E(\nu)\exp(\mathrm{i}2\pi\nu x)\,\mathrm{d}\nu$$

and thus the total instrumental function can be

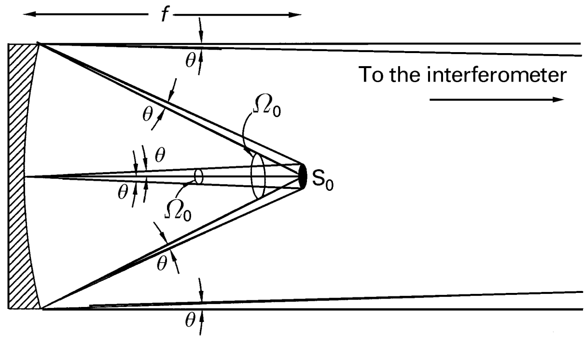



**Figure 7** Extended IR radiation source $S_0$ at the focal point of the parabolic collimator mirror M.

given by

$$W^h_{L,\Omega}(\nu) = W_\Omega(\mp\nu_0, \nu) *$$
$$\left\{\sum_{k=-\infty}^{\infty} 2 \quad \mathrm{sinc}\left[2\pi\left(\nu + \frac{k}{h}\right)L\right]\right\} \qquad [14]$$

This has to be used in Equation [10] instead of $W^b_L(\nu)$. As one can see, the instrumental line shape function [i.e. the observed line shape if $E(\nu) = \delta(\nu + \nu_0) + \delta(\nu - \nu_0)$] is given by

$$W_{L,\Omega}(\nu) = W_\Omega(\nu_0, \nu) * 2L\,\mathrm{sinc}(2\pi\nu L) \qquad [15]$$

This situation is illustrated in **Figure 8** with three different $\Omega$ values. The optimum situation is the case where the aperture broadening $W_\Omega(\nu_0,\nu)$ and the truncation broadening $2L\mathrm{sinc}(2\pi\nu L)$ have approximately equal widths, i.e.

$$\frac{\nu_0 \Omega}{2\pi} \approx \frac{1.21}{2L} \qquad [16]$$

In **Figure 8** the optimum situation is roughly in the second row. Furthermore, it can be seen that the position of the line at $\nu_0$ is given by $\nu_{\mathrm{obs}} = [1 - \Omega/4\pi]\nu_0$ = constant $\times\, \nu_0$. Now, according to Equation [12], the signal is proportional to $aA/f^2 = \Omega A$, and for a given resolution, when $\Omega$ is fixed, the only parameter which has an effect on the signal, is the area $A$ of the collimator mirror. The area $a$ of the radiation source and the focal length $f$ of the collimator mirror can be selected freely provided that $a/f^2 = \Omega$. A similar effect can occur due to the area of the detector. As mentioned earlier, the instrumental profile plays an important role in describing the behaviour of the

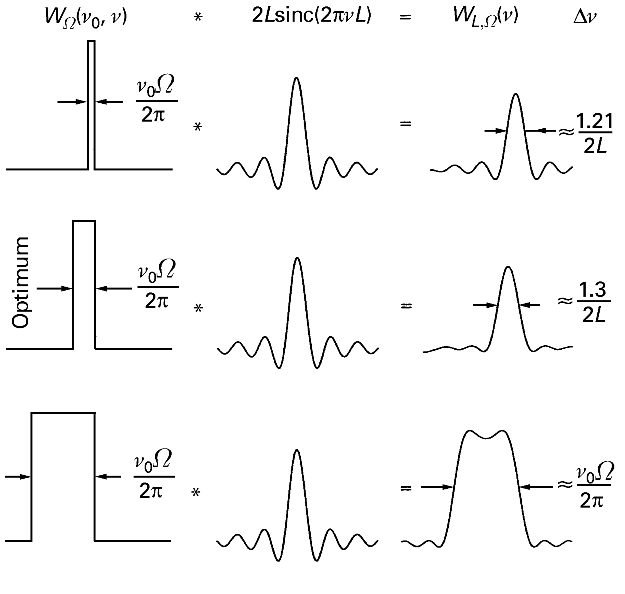



**Figure 8** Total instrumental line shape function $W_{L,\Omega}(\nu)$ with three different $\Omega$-values.

spectrometer. Especially in high-resolution gas-phase spectroscopy, where the instrumental profile $W_{L,\Omega}$ and the real (Lorentzian) line shape have approximately equal widths, the convolution by instrumental profile changes substantially the real line shape.

### Other FT interferometers

The most difficult disadvantage of the Michelson interferometer is tilting of the moving plane mirror during scan. It is well known that the tilting angle of the moving mirror should be less than about $\lambda/(8D)$, where $\lambda$ is the wavelength of the light under study and $D$ is the diameter of the moving mirror. Thus the driving of the mirror is very difficult with high-resolution FT-IR spectrometers. One way to solve this problem is by using a dynamic alignment system, which measures the tilting angle and tries to keep it at zero during the run. However, as the resolution increases, the probability of malfunction of the dynamic alignment system during a very long run increases.

Another solution is to use cube-corner mirrors, which consist of three mutually orthogonal plane mirrors. The mirror system reflects any incident ray back in the opposite direction. The first very high resolution cube-corner interferometer that really worked was the Oulu interferometer in Finland. The Oulu interferometer is basically a Michelson interferometer, where the moving and the fixed mirror are cube-corners. If the corners are perfect, the tilt problem completely disappears. The only disadvantage of this type of interferometer is a shearing problem, i.e. the lateral shift of the moving cube-corner. This is the case where the moving cube-corner does not follow the optical axis of the interferometer. The lower the resolution the more serious this disadvantage is.

A cube-corner interferometer is improved if two cube-corners are set to move back to back. Two cube-corners fixed back to back as a single moving part completely eliminates the tilting and the lateral shifts, if the cube-corners are perfect (90° × 90° × 90°), and the corners coincide exactly. Even further improvements are emerging at the University of Turku, Finland, namely a super high-resolution interferometer with a resolution of $4 \times 10^{-4}$ cm$^{-1}$ which has the moving cube-corner pair back-to-back and the beam focus at the beamsplitter (see **Figure 9**). This type of interferometer is almost ideal for very high resolution IR spectrometry.

An alternative interferometer to the cube-corner is the cat's-eye interferometer, where the cube-corners are replaced by cat's-eye retroreflectors. These are components that consist of one parabolic and one spherical mirror, and reflect an incident ray back in the opposite direction. Cat's-eye interferometers are also used in high-resolution spectroscopy.

## Instrumental details of FT-IR spectrometers

### Radiation source

Black-body radiators are used as broad band radiation sources in IR spectrometers. The only adjustable parameter of the source is the temperature; the higher the temperature, the higher the intensity. Typical temperatures are 1000–1800 K. The most commonly used sources in mid-IR are Globars (silicon carbide), Nernst glowers and nichrome coils. In the far-IR below 100 cm$^{-1}$ a high-pressure mercury lamp is quite good. Nowadays, ceramic sources have become more popular. The radiation sources may need a cooling system, like water cooling.

### Beamsplitter

The beamsplitter is one of the most expensive and sensitive components of an interferometer, and must be chosen carefully.

A pellicle beamsplitter is a high tensile strength elastic membrane, which is stretched like a drumhead over a flat frame. The membrane has a high refractive index and its absorption is negligible. The most common membrane material is Mylar. An ideal beamsplitter divides the incoming intensity into two equal parts. This is, however, achieved only in a definite wavenumber region, due to the interference

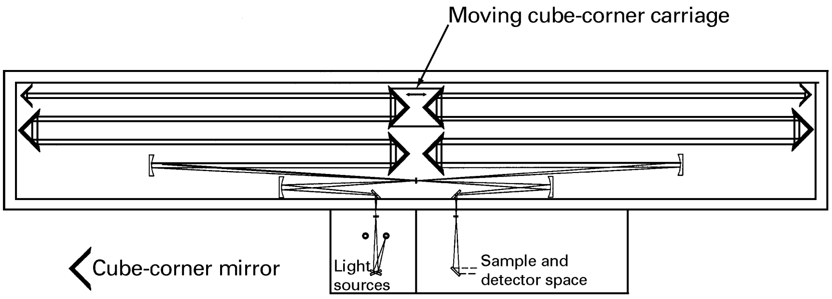



**Figure 9** Layout of the Turku high-resolution interferometer.

phenomenon in a thin film. A Mylar beamsplitter works with wavenumbers less than 1000 $cm^{-1}$.

In the near-IR region a beam-splitting film should be very thin. In this case it is necessary to use a low-absorption dielectric coating, which is deposited on a suitable substrate plate. A thin Ge layer deposited on a KBr substrate, with a KBr compensating plate, is a good beamsplitter in the wavenumber region 1000–4000 $cm^{-1}$.

## Optics

With the help of optical components the IR radiation is carried from the source through the whole spectrometer and finally to the detector. Mostly the needed accuracy is so-called geometrical accuracy. This means that the beam has to be guided to the appropriate places to within 0.1–1 mm. However, higher accuracy is needed in order to measure the wavelength of radiation in the interferometer. It is said that there is a need for interferometric accuracy, i.e., everything should be correct within a small fraction of wavelength, say 0.01–0.1 μm.

In IR spectroscopy there are only a few materials transparent enough; this means that lens optics is not usually used. Instead, the use of mirror optics is preferred and the transmission of any material is avoided. However, in certain circumstances, for convenience sake, attempts are made to use some materials, e.g. polyethylene and KBr lenses. Thus the weakest points in IR optics are, besides the beamsplitter, only the windows of the gas cell. Mirrors are often manufactured from Pyrex glass coated with metal. Gold coating has a good reflection coefficient especially in far-IR and aluminium is also quite good over the whole IR region. In IR spectrometers, plane mirrors, spherical mirrors, paraboloidal mirrors, off-axis paraboloidal mirrors, ellipsoidal mirrors, toroidal and off-axis ellipsoidal mirrors are used. In planning IR instruments one must bear in mind that a perfect imaging takes place only in plane mirrors. All others have some kind of aberration, like spherical aberration. These aberrations limit the performance of instruments. However, there are some special imaging situations where perfect imaging is possible, for example, a spherical mirror images perfectly only on the optical axis from the centre of curvature back to the centre of curvature. A paraboloidal mirror images from infinity to the focal point on the optical axis and the off-axis paraboloid images from infinity to the focal point, which is out of the optical axis. An ellipsoidal mirror images from one focus to another. Hence these mirrors should preferably be used, so that conditions are as near perfect as possible. This also minimizes aberrations. In IR spectroscopy it is very important that as much radiation power as possible is transmitted from the source through the whole optics to the detector.

## Gas cell

The most common sample compartment in high-resolution FT-IR spectrometers is the White cell, where the beam passes a path of 1–1000 m. In a long-path gas cell, however, a small f/number (focal length of the mirror/diameter of the pupil) and a large size of the circular image may lead to a loss of effectivity. In the Oulu interferometer this problem was solved by an alternative multiple-path gas cell, where the images at the focal planes occur on a circle around the principal axis of the cell. The aberrations remain small in spite of a small f/number, because the successive images are relatively near to the principal axis of the optics.

## Detectors

The detector of the FT-IR spectrometer is very important, because it determines the maximum possible signal-to-noise ratio ($S/N$; where $S$ = signal and $N$ = noise) and the minimum recording time of the

spectrum. IR detectors can be divided into two categories: thermal detectors and quantum detectors. Thermal detectors sense the change of temperature of a detector material due to absorption of the IR radiation. The output signal is generated in the form of a thermal electromotive force (e.g. thermocouples), the change in resistance of a conductor (bolometers) or semiconductor (thermistor bolometers), or the change in pressure of a noble gas (pneumatic detectors). Thermal detectors respond to radiation over a wide range of wavenumbers, but they are usually slow with a time constant typically from 0.001 to 0.1 s. In pyroelectric bolometers, a heat-sensing ferroelectric material changes the degree of polarization when the temperature changes due to absorption. The voltage responsivity of a pyroelectric detector at frequency $f$ is roughly proportional to $1/f$. The most commonly used material for the pyroelectric detector is deuterated triglycine sulfate.

In quantum detectors, IR radiation causes electrons to be excited to a higher energy level. In an n-type semiconductor, electrons in the valence band are unable to increase the conductivity. If IR radiation excites the electrons to the conduction band, they can act as current carriers. For good sensitivity, it is necessary to cool the detector. For certain detectors, such as PbS and PbSe, it may be sufficient to use a thermoelectric cooler to maintain their temperature just below ambient. Mercury cadmium telluride detectors must usually be maintained at liquid-nitrogen temperature (77 K), whereas others may require cooling to the temperature of liquid helium (4.2 K).

In the far-IR, where photon energies are very low, quantum detectors cannot be used at all. Two types of thermal detectors, each operating at liquid-helium temperatures, have been used. The first one is the germanium bolometer, often doped with a low level of copper, gallium or antimony. For increased responsivity, these detectors are cooled down to 1.5 K by pumping the detector cryostat. The second type is the InSb hot-electron detector.

The sensitivity of IR detectors is expressed by the noise equivalent power (NEP) of the detector. This NEP is the ratio of the root mean square (rms) noise voltage $V_{rms}$ in the detector to the voltage responsivity, i.e. NEP = $V_{rms}/R$. If $V_{rms}$ is in units of V $Hz^{-1/2}$ and $R$ in units of V/W, then NEP is in units W $Hz^{-1/2}$. Usually the manufacturer gives a specific detectivity $D^*$ as a function of wavelength or wavenumber. It is defined as $D^* = \sqrt{A_D}/\text{NEP}$, where $A_D$ is the area of the detector. Some typical $D^*$-curves are shown in **Figure 10**.

## Electronics

Nowadays, IR spectrometers are usually computer controlled. Thus the main component of the electronics is a microprocessor, which controls all the operations of the spectrometer, collects the data, and computes, saves, displays and plots final spectra. The computer is very important to perform any kind of data treatment needed in IR spectroscopy. The other electronic parts of FT-IR spectrometers are a preamplifier and a main amplifier of the IR detector, an

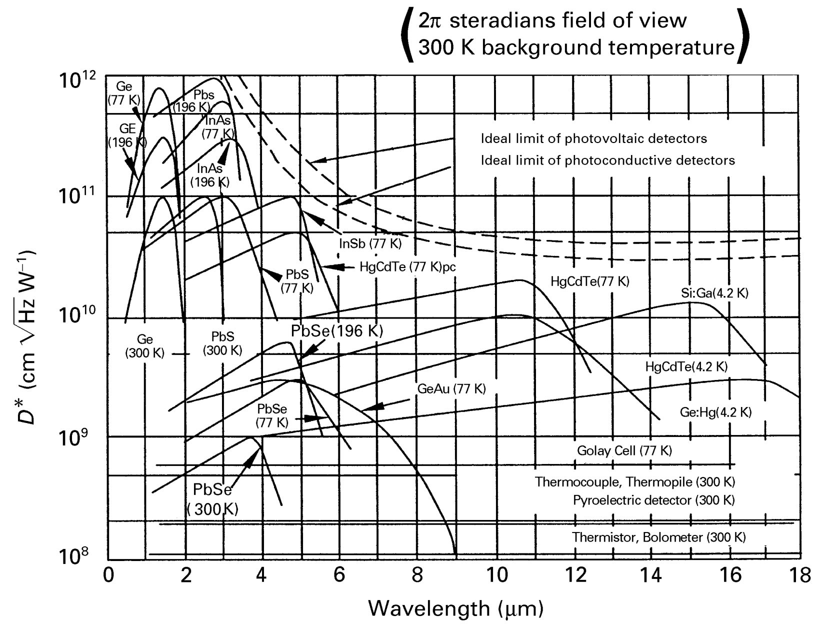



**Figure 10** $D^*$ curves for the most used IR detectors (Hamamatsu Photonics K. K., Solid State Division, 1990).

analogue-to-digital converter (AD converter), a set of bandpass electronic filters and power supplies. It is important that the numbers of bits and data rate of the AD converter are high enough.

## Advantages of FT-IR spectrometers

### Jacquinot advantage (throughput advantage)

In an interferometer the total radiation power on the detector is much higher than that in a grating instrument. In practice, the FT-IR spectrometer has an approximately 200 times larger $S/N$ than the grating spectrometer with the same resolution and recording time. The reason is that at the optimum resolution the area of the circular aperture (the Jacquinot stop) in the FT-IR spectrometer is 200 times larger than the slit in the grating spectrometer.

### Fellgett advantage (multiplex advantage)

In FT-IR spectrometers all the wavenumbers $\nu$ are recorded simultaneously. Let us assume that the wavenumber region under study consists of $M$ resolution elements, i.e. $\nu_2 - \nu_1 = M\Delta\nu$. If the noise of a detector and an amplifier is greater than the noise of a radiation source, then the $S/N$ is proportional to the square root of the total observation time $T$, i.e. $S/N \propto \sqrt{T}$. In fact $S \propto T$ and $N \propto \sqrt{T}$. In FT-IR spectrometers one observes all the times $T$, all the $M$ spectral elements, and thus $(S/N)_{FT} = a\sqrt{T}$; in grating spectrometers the observation time of each spectral element is only $T/M$, and $(S/N)_G = a\sqrt{T/M}$, where in both equations $a$ denotes the same constant. And finally $(S/N)_{FT}/(S/N)_G = \sqrt{M}$, which is the Fellgett multiplex advantage. The Jacquinot advantage and the Fellgett advantage together give $(S/N)_{FT}/(S/N)_G \approx 200\sqrt{M}$ in IR.

### Other advantages

It is useful to define a $K$-factor as $K = \Delta\nu_{true}/\Delta\nu$, where $\Delta\nu$ is the FWHH of the instrument function $W_{\Omega,L}(\nu_0, \nu)$ and $\Delta\nu_{true}$ is the FWHH of the true line shape. The optimum instrumental resolution $\Delta\nu_0 = \Delta\nu_{true}/K_0 \approx 1.21/(2L_0)$ is the case where the maximum distortion of observed lines is roughly equal to the noise of the spectrum and the optimum $K$-value is given by $K_0 = 2\ln[(S/N)_0]/(1.21\pi) \approx \log_{10}[(S/N)_0]$. The greatest difference between FT and grating instruments is the instrumental profiles $2L\mathrm{sinc}(2\pi\nu L)$ (in FT) and $L'\mathrm{sinc}^2(\pi\nu L')$ (in grating). **Table 1** shows the optimum $K$-values, i.e., $K_0$, as a function of $(S/N)$, when the instrumental profile is sinc and sinc$^2$ shape. **Table 1** shows that, for

**Table 1** Optimum $K$-values as a function of the signal-to-noise ratio ($S/N$) are listed for sinc- and sinc$^2$-type instrumental functions

| *S/N* | *Sinc* | *Sinc*$^2$ |
|---|---|---|
| 10 | 1.21 | 3.59 |
| 20 | 1.58 | 7.19 |
| 40 | 1.95 | 14.4 |
| 60 | 2.16 | 21.6 |
| 80 | 2.31 | 28.7 |
| 100 | 2.43 | 35.9 |
| 200 | 2.79 | 71.9 |
| 400 | 3.16 | 144 |
| 600 | 3.37 | 216 |
| 800 | 3.53 | 287 |
| 1 000 | 3.64 | 359 |
| 2 000 | 4.01 | 719 |
| 4 000 | 4.37 | 1 437 |
| 6 000 | 4.59 | 2 156 |
| 8 000 | 4.74 | 2 874 |
| 10 000 | 4.86 | 3 593 |

example, with $(S/N) = 1000$, the optimum value $K_0$ for grating spectrometers is 100 times higher than that for FT-IR spectrometers. In other words, a 100 times higher resolution is needed in the case of the grating spectrometer. This is usually impossible in practice. Now, it can be concluded that the optimum instrumental resolution of the FT-IR spectrometer is the lowest possible one or the FT-IR spectrometer measures a spectrum with the smallest possible distortions of observed lines. This is the resolution advantage of FT-IR spectrometers.

The fact that $\nu_{obs} = \text{constant} \times \nu_{true}$ is the linearity advantage of FT-IR spectrometers. It is roughly valid even though the interferometer does not work correctly, for example, in the case of phase errors in the interferogram. Furthermore, no ghost lines exist in the spectrum. However, in the grating spectra ghost lines are possible in the case of malfunction of the instrument.

## Use of FT-IR spectrometers

### Performance

The instrumental resolution of high-resolution FT-IR spectrometers is, in practice, of the order of 0.001 $cm^{-1}$. In addition to high resolution, a good absolute accuracy of the line positions in the spectrum is maintained. A wavenumber accuracy of $10^{-9}$, or even better, has been achieved.

The main field of use of high-resolution FT-IR spectrometers is molecular spectroscopy. The

spectrometers are used to study the geometric structures of molecules and their behaviour in rotations and vibrations. IR spectra are effective in revealing quantum mechanical phenomena and interactions in molecules.

An example of a typical rotation–vibration band, measured with a high-resolution FT-IR spectrometer, is shown in **Figure 11**. It is the bending band $\nu_2$ of $CO_2$ near 667 $cm^{-1}$. This spectrum was recorded by the Oulu spectrometer with an instrumental resolution of 0.002 $cm^{-1}$. The relative errors of the wavenumber compared to model equations are of the order of $10^{-8}$. The vibration energy levels of different normal modes are sometimes very close to each other. In this case these normal modes can be coupled so that the energy levels are shifted apart. Now the observed wavenumber $\nu_{obs}$ cannot be expressed in a closed form and the rotational lines of the spectrum are shifted. With the proper quantum mechanical calculations it is possible to calculate the shifts. The high resolution and accuracy in wavenumber scale make an FT-IR spectrometer a very effective instrument to study these kinds of interactions in molecular spectra.

## Calibration

In principle, the calibration of FT-IR spectrometers is performed with regard to intensity and wavenumber. However, the intensity calibration is not usually necessary because the absorbance scale is used, $-\ln(I/I_0)$, where the ratio $I/I_0$ automatically accomplishes the intensity calibration. The most important task in calibration is the wavenumber calibration. Usually the calibration takes place by simultaneous measurement of the known IR spectrum. For example, the user simply adds reference gas (known spectrum) into the sample gas under measurement. Using the known

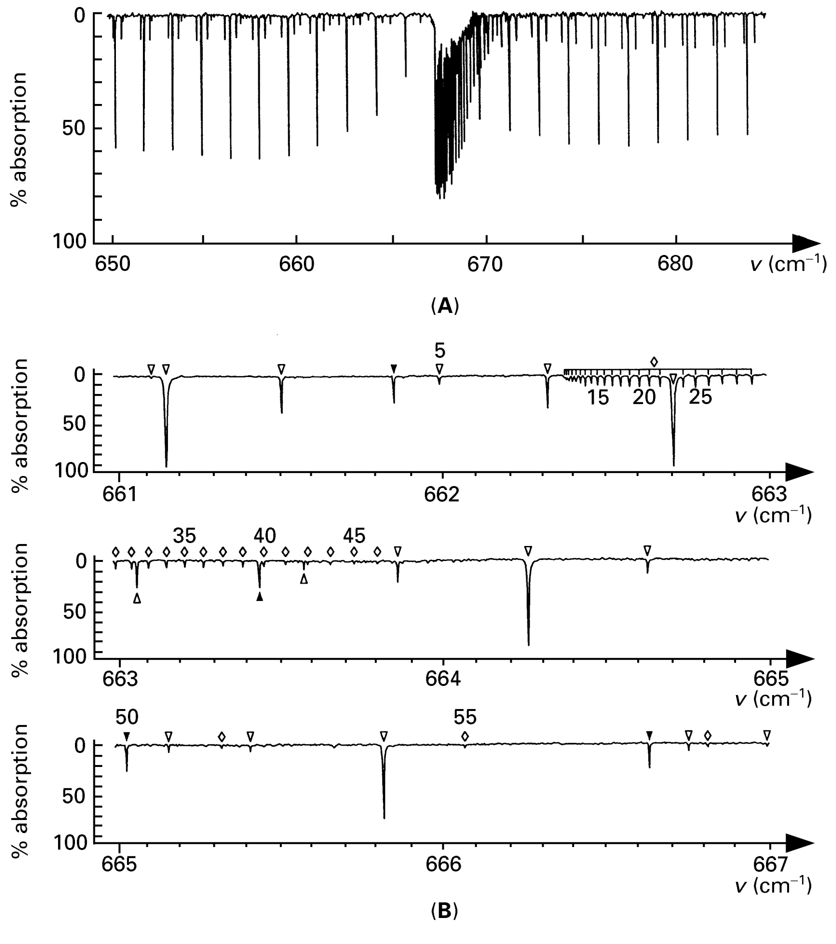



**Figure 11** A part of the $\nu_2$ bending band of $CO_2$ near 667 $cm^{-1}$ measured by the Oulu FT-IR spectrometer with a resolution of 0.0045 $cm^{-1}$ (A) and 0.002 $cm^{-1}$ (B). In (B), ∇ indicates $^{12}C^{16}O_2$, ▼ indicates $^{13}C^{16}O_2$ and ◇ indicates $^{16}O^{12}C^{18}O$.

spectral lines of the reference gas it is possible to fix the wavenumber scale with the uncertainty, which is roughly the uncertainty of the reference spectral lines. The linearity advantage of FT-IR spectrometers makes it possible to reach a perfectly linear wavenumber scale. This is why only one reference line is needed, i.e. the reference laser line, to measure the optical path difference of the interferometer. The situation is completely different in grating, prism and laser spectrometers, where the wavenumber scale is more or less nonlinear with respect to the scanning parameter, like the rotation of grating or prism. Therefore in these spectrometers it is important to use many reference lines measured simultaneously with the sample spectrum in order to achieve a good wavenumber scale.

### Data handling and computation

Nowadays, all modern IR spectrometers are equipped with a (micro) computer. Of course, the computer calculates absorbance $-\ln(I/I_0)$, displays the spectrum on the screen or plots the spectrum on paper, saves the spectra on disk, and so on. Modern FT-IR spectrometers are able to record very accurate spectra, where $S/N$ is high and $|\mathrm{d}\nu/\nu|$ is very small, typically $S/N \approx 10^5$ and $|\mathrm{d}\nu/\nu| \approx 10^{-9}$. Also absorbance accuracy is quite high in FT-IR spectrometers. The spectra with high accuracy and high $S/N$ include plenty of spectral information. Often spectral information is not in a proper form in the original recorded spectra. In order to derive all the possible information from high-quality IR spectra, good data-treatment algorithms are needed. If the spectrum consists of overlapping lines, resolution enhancement methods, like derivation, Fourier self-deconvolution, linear prediction (or LOMEP, line-shape optimized maximum entropy linear prediction) or band fitting and factor analysis, are very useful. Other possible data treatment methods, which the users may need, are, for example, smoothing, background elimination, interference fringe pattern elimination, peak-finding algorithm, maximum-entropy method, deconvolution and multi-component analysis methods. In modern IR spectrometers good data treatment algorithms greatly increase the usefulness of IR spectroscopy in many applications.

## Laser spectrometers

Laser spectrometers are not very useful in IR; they have more applications in the visible region, for example, LIDAR (light detection and ranging). The reason for this is the lack of good, cheap commercial lasers for IR. Wavenumber regions, due to tuning methods, are quite narrow, and the wavenumber calibration is a problem because of the lack of good calibration lines. In some cases it is also difficult to maintain a single-mode operation of the laser. The advantages of the laser spectrometer are a very high resolution, even sub-Doppler resolution, and a good $S/N$ of the spectra.

There are two main techniques when using lasers as IR spectrometers. The first method is simply to tune the laser wavelength over a narrow wavelength region and simultaneously detect the laser signal $I(\nu)$ transmitted through the gas cell. The second possibility is to tune the wavelength of the absorption lines by applying the magnetic (Zeeman effect) or electric (Stark effect) field to the sample gas. When the wavelength of the tuned spectral lines matches the laser wavelength, absorption takes place. Thus absorption spectra are recorded as a function of the magnetic or electric field magnitude (strength).

A diode laser spectrometer is described as a typical example of this kind of spectrometer. A simple p–n junction laser is run by applying the forward bias voltage across the junction. The parallel polished crystal faces perpendicular to the junction form the mirrors acting as a laser resonator. The coherent radiation emits at the wavenumber $\nu_k = k/(2nL)$, where $k$ is an integer, $n$ is the refractive index and $L$ is the length of the resonator. The wavenumbers $\nu_k$ are called the longitudinal modes of the laser. The number of modes is typically from three to ten. Operation in IR needs the cooling of the diode laser. Tuning can be accomplished by changing the term $nL$, which depends on the temperature. Temperature can be fine-controlled by changing current through the junction. It is possible to tune typically 200 cm$^{-1}$ by changing the temperature and about 2 cm$^{-1}$ by changing the current. Thus a spectrum has to be run in parts of 1–2 cm$^{-1}$ by current tuning at a constant temperature. The temperatures of the parts are selected so that the parts together form a continuous spectrum over a desired region. Running the diode laser is very troublesome, because current tuning is needed at very many different temperatures, which makes calibration, for example, very difficult. In addition, a single-mode operation of the diode laser is needed, in other words, the integer $k$ must be a constant. This can be performed by scanning a grating monochromator simultaneously with the laser tuning. The grating spectrometer follows the wavenumber $\nu_k = k/(2nL)$ of the mode. The change in wavenumber during tuning is measured by recording the interference fringes of a long Fabry–Perot etalon. The absolute value of the wavenumber in each part has to be fixed by calibration lines. In

many applications the diode laser spectrometer solely is not sufficient. However, it is a very good instrument for looking at details over a wide region spectrum recorded by a FT-IR spectrometer. Thus the laser spectrometer is a supplementary instrument to the FT-IR spectrometer.

## List of symbols

$a$ = source area; $A$ = collimator mirror area; $A_D$ = detector area; $c_0$ = velocity of light in a vacuum; $C$ = concentration; $d$ = distance; $D$ = mirror diameter; $D^*$ = specific detectivity; $E''$ and $E'$ = lower and higher energy levels; $E$ = spectrum; $f$ = focal length, or frequency; $F$ = interference record; $F\{\ \}$ = Fourier transform; $F^{-1}\{\ \}$ = inverse Fourier transform; $h$ = Planck's constant; $i = \sqrt{-1}$; $I$ = radiation intensity; $k$ = an integer; $K$ = resolution enhancement factor; $L$ = maximum optical path difference; $M$ = number of resolution elements; $n$ = refractive index; $N$ = noise; $S$ = signal; $(S/N)_{FT}$ = signal-to-noise-ratio of a FT-IR spectrometer; $(S/N)_G$ = signal-to-noise-ratio of a grating spectrometer; $S_0$ = radiation source; $T$ = observation time; $V_{rms}$ = root-mean-square voltage; $W$ = instrumental function; $\alpha$ = angle between optical axis and beam; $\alpha(\nu)$ = absorption coefficient; $\beta(\nu)$ = absorptivity; $\delta$ = Dirac's delta function; $\lambda$ = wavelength; $\nu$ = radiation wavenumber; $\Omega$ = solid angle.

*See also:* **Fourier Transformation and Sampling Theory; FT-Raman Spectroscopy, Applications; Gas Phase Applications of NMR Spectroscopy; High Resolution IR Spectroscopy (Gas Phase), Applications; Hydrogen Bonding and Other Physicochemical Interactions Studied By IR and Raman Spectroscopy; Laboratory Information Management Systems (LIMS); Laser Spectroscopy Theory; Light Sources and Optics; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading

Bell RJ (1972) *Introductory Fourier Transform Spectroscopy*. New York: Academic Press.

Bracewell R (1965) *The Fourier Transform and its Applications*. New York: McGraw-Hill.

Brigham EO (1974) *The Fast Fourier Transform*. Englewood Cliffs: Prentice-Hall.

Chamberlain J (1979) *The Principles of Interferometric Spectroscopy*. Chichester, UK: John Wiley.

Griffiths PR and de Haseth JA (1986) *Fourier Transform Infrared Spectrometry*. New York: John Wiley.

Herzberg G (1945) *Infrared and Raman Spectra of Polyatomic Molecules*. New York: Van Nostrand.

Mackenzie MW (ed) (1998) *Advances in Applied Fourier Transform Infrared Spectroscopy*. New York: John Wiley.

Marshall AG and Verdun FR (1990) *Fourier Transforms in NMR, Optical, and Mass Spectrometry, A Users Handbook*. New York: Elsevier.

# High Resolution IR Spectroscopy (Gas Phase), Applications

**E Canè** and **A Trombetti**, Università di Bologna, Italy

Copyright © 1999 Academic Press

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
|---|
| Applications |

The IR region of the electromagnetic spectrum extends from the visible ($\lambda \cong 800$ nm, $\tilde{\nu} \cong 12\,500$ cm$^{-1}$) to the sub-mm waves ($\lambda < 1$ mm; $\tilde{\nu} > 10$ cm$^{-1}$) and the IR photon carries an energy in the $2.48 \times 10^{-21}$ – $1.99 \times 10^{-24}$ J range. This is also the energy exchanged by an atom or a molecule when IR radiation is emitted or absorbed. The interaction of the IR radiation with particles corresponds to transitions between rotation–vibration energy levels in molecules or between Ryberg levels in atoms.

High-resolution IR spectroscopy is confined to the study of atoms and molecules in the gas phase at low pressures, where the interaction energy between particles is orders of magnitude lower than in the condensed phases. The intermolecular interactions broaden the energy levels, and, consequently, the spectral lines, preventing the observation of finer details in the spectra. In the IR region the width of the spectral lines in the gas phase is caused mainly by the Doppler effect and by the broadening due to

collisions. The concept of high resolution in IR spectroscopy has changed with the introduction of new experimental techniques such as FT-spectroscopy and laser techniques. Presently, an instrumental resolution of $2 \times 10^{-3}$ cm$^{-1}$ (measured as FWHM) or higher is achievable with commercial FT interferometers in the 10–4000 cm$^{-1}$ range: this is comparable to the Doppler line width for a molecule with a mass of 50 u at room temperature, calculated using the equation:

$$\Delta\tilde{\nu} = 2\tilde{\nu}\left(\frac{2Tk_B \ln 2}{mc^2}\right)^{1/2} \quad [1]$$

where $k_B$ is the Boltzmann constant, $m$ is the particle mass, $T$ is the absolute temperature and $c$ is the speed of light. Thus the Doppler width can be assumed as the limit of high-resolution FTIR spectrometers that use a broadband radiation source of low spectral brightness. With the introduction of lasers in the infrared, sub-Doppler resolution can be obtained using the saturation lamb-dip or other specialized techniques. Sub-Doppler resolution can also be performed with molecular beams or in supersonic jet expansions, where molecules are rotationally and vibrationally cooled to a few K. Moreover, the enhancement of sensitivity of the new experimental techniques has made possible the detection and analysis of many unstable or transient species such as van der Waals complexes, free radicals and ions.

The main applications covered in this study are the accurate determination of rotation and rotation–vibration molecular energies; the determination of the molecular geometry of simple molecules; the evaluation of force field and of the vibration– and rotation–vibration interactions; the measurement of pressure broadening and pressure shift of the spectral lines; the determination of electric dipole moments via laser-Stark spectroscopy; the studies of intramolecular dynamics; the calculation of rate constants, equilibrium constants and other thermodynamic data; the evaluation of relaxation times.

## Rotational and rovibrational energies

The results of the analysis of a high-resolution IR spectrum of a molecule is usually a large and accurate set of molecular constants. With these constants, rotational and rovibrational energy levels can be calculated within the range of the quantum numbers of the transitions from which they have been obtained. This is the primary application of high-resolution gas phase spectra in the IR region. The results of the analysis of the $\tilde{\nu}_1$, $2\tilde{\nu}_2$ and $2\tilde{\nu}_3 + \tilde{\nu}_6$ IR bands of $COF_2$ are presented as an example in **Table 1**. The wavenumbers of about 3200 spectral lines have been measured and assigned to specific rovibrational transitions, identified through the rotational quantum numbers of the lower and upper states of the transition. After the assignments a least-squares analysis was performed, where the parameters of **Table 1** and their standard deviations were determined.

The most important molecular constants are $\tilde{\nu}_0$, the energy of the vibrational levels above the rotationless ground state and the rotational constants $B_x$, $B_y$ and $B_z$. The centrifugal distortion constants, $D$ and $d$, $H$ and $h$, are orders of magnitude smaller than $B$ and account for the deviations from the rigid rotor energy levels. The constants $W$ are necessary to reproduce the measured spectrum since the $\tilde{\nu}_1$ and $2\tilde{\nu}_2$ states are in Fermi resonance. The constants $\xi$ take into account the Coriolis interaction between the $\tilde{\nu}_1$ and $2\tilde{\nu}_3+\tilde{\nu}_6$. For the latter state no transitions have been observed and its constants have been obtained from the perturbation induced on the levels of $\tilde{\nu}_1$. The RMS in **Table 1** is the standard deviation normalized to a single rovibrational transition of the calculated and measured spectrum. This quantity should be very close to the precision of the wavenumber measurement if the theoretical model used for the analysis is satisfactory.

## Determination of molecular geometry

Accurate molecular geometries can be obtained from the analysis of the high-resolution IR spectra. The analysis of rovibrational spectra provides, among other important data, the rotational constants of the molecules, which in turn allow the determination of the principal moments of inertia and, eventually, of the molecular geometry. Pure rotational spectroscopy provides the same information, usually for the ground state, except when the molecule under study has no permanent electric dipole moment. The principal moments of inertia are related to the bond lengths and bond angles of the molecule by equations that vary according to the geometry. Equilibrium geometry corresponds to the minimum of the potential energy surface of the molecule. Owing to the zero-point energy this minimum does not correspond to a stationary state of the molecule. The rovibrational analysis of fundamentals and overtones gives the rotational constants of the excited vibrational state, and of the lowest molecular state, which is above the minimum of the potential energy curve by the zero-point energy. The bond lengths obtained from the ground state rotational constants are approximately the mean values of the lengths in the ground state and

**Table 1** Parameters for the $2\nu_2$, $\nu_1$, and $2\nu_3$+$\nu_6$ states of $COF_2$. Reprinted with permission of Academic Press from D' Cunha *et al* (1997) *Journal of Molecular Spectroscopy* **186**: 363–373

| | $2\nu_2$ state | | $\nu_1$ state | | $2\nu_3$+$\nu_6$ state |
|---|---|---|---|---|---|
| $\nu_0$ | 1 922.516 220(49) | | 1 935.940 117(47) | | 1 935.145(30) |
| $B_x$ | 0.390 996 08(127) | | 0.393 161 46(93) | | 0.394 881 1[a] |
| $B_y$ | 0.391 157 49(124) | | 0.390 005 56(107) | | 0.392 324 0(985) |
| $B_z$ | 0.194 993 62(21) | | 0.195 528 25(26) | | 0.196 974 7(632) |
| $D_J\times10^6$ | 0.446 129(451) | | 0.437 270(163) | | 0.200 4(209) |
| $D_{JK}\times10^6$ | −0.766 714(945) | | −0.742 981(653) | | 0.104 2(243) |
| $D_K\times10^6$ | 0.353 319(598) | | 0.339 445(640) | | 4.007 2(161) |
| $d_1\times10^6$ | −0.041 021(348) | | −0.041 513(211) | | −0.042 227[b] |
| $d_2\times10^6$ | −0.008 716(227) | | −0.008 832(183) | | −0.010 031[b] |
| $H_J\times10^{11}$ | 0.123 47(2150) | | 0.025 39[a] | | 0.089 36[b] |
| $H_{JK}\times10^{11}$ | −0.462 29(6740) | | −0.155 58(2910) | | −0.375 49[b] |
| $H_{KJ}\times10^{11}$ | 0.556 43(7710) | | 0.217 96(3480) | | 0.480 50[b] |
| $H_K\times10^{11}$ | −0.238 54(2940) | | −0.032 93[a] | | −0.194 57[b] |
| $h_1\times10^{11}$ | −0.009 40[c] | | 0.009 40[b] | | 0.009 40[b] |
| $h_2\times10^{11}$ | 0.003 66[b] | | 0.003 66[b] | | 0.003 66[b] |
| $h_3\times10^{11}$ | 0.008 25[c] | | −0.008 25[b] | | −0.008 25[b] |
| $W_0$ | | 13.844 5[a] | | | |
| $W_J\times10^4$ | | −0.074 8[a] | | | |
| $W_K\times10^4$ | | 0.042 0[a] | | | |
| $W_{XY}\times10^4$ | | 1.167 9(66) | | | |
| $\xi$ | | | | −0.006 711(95) | |
| $\xi_J\times10^5$ | | | | −0.126 49(615) | |
| $\xi_K\times10^5$ | | | | −0.127 92(544) | |
| RMS dev | 0.000 39 | | 0.000 36 | | |
| Lines | 1 529 | | 1 675 | | |

All values are in $cm^{-1}$.
[a] Constrained to the value obtained with a damping factor of 0.0001.
[b] Constrained to the ground state value.
[c] Sign changed to conform with $III^l$ representation (see text).

are usually slightly larger than the equilibrium ones, and the geometry is called the $r_0$-structure. The equilibrium geometry, $r_e$-structure, can be determined from $B_e$ if the rovibration interaction constants $\alpha_i$, defined by the equation

$$B_{\nu_i} = B_e - \sum \alpha_i \left(\nu_i + \frac{1}{2}\right) \quad [2]$$

for all the normal vibrations of the molecule are known. Typical values of a $r_0$-structure are compared with a $r_e$-structure in **Table 2** for $C_2H_2$. Since for an asymmetric top molecule there are three principal moments of inertia a maximum of three molecular parameters, bond length and angles, can be obtained. Moreover, in planar molecules only two moments of inertia are independent and in linear molecules such as acetylene the two moments of inertia are equal. Thus, determination of geometry for polyatomic molecules seems to be impossible. Isotopic substitution is the way to increase the number of independent moments of inertia so that the structure parameters of polyatomic molecules can also be inferred. In fact, molecules differing by one or more isotopes have the same $r_e$-structure, while the $r_0$-structure is slightly different since the zero-point energy is isotope dependent. For large polyatomic molecules the determination of the $r_e$-structure is nearly impossible. In benzene, $C_6H_6$, owing to the hexagonal symmetry, only the $r_{C-H}$ and $r_{C-C}$ interatomic distances are needed to determine the geometry. From the analysis of the high-resolution IR or UV spectrum only one independent rotational constant is obtained and an isotopically substituted benzene such as $C_6D_6$ has to be studied to obtain the $r_0$-structure. Since each isotopomer has 20 normal vibrations, and thus 20 $\alpha_i$ values, the analysis of 40 high-resolution bands is needed for the $r_e$-structure. To date, about 15 fundamental bands of $C_6H_6$ and $C_6D_6$ have been rotationally analysed.

## Force field evaluation

Accurate *ab initio* theoretical calculations of harmonic and anharmonic force fields of molecules as large as benzene have recently been performed. By comparing the experimentally determined molecular constants with the calculated ones, the reliability of theoretical calculations can be assessed. All the molecular constants of $COF_2$ in **Table 1** can be evaluated from the theoretical anharmonic force field. For molecules such as ammonia the computed fundamental wavenumbers are within 3 cm$^{-1}$ or better of the experiment. For anharmonic constants and centrifugal distortion constants, which are much smaller than the fundamental frequencies, the deviations are on average around 15%. The experimental determination of anharmonic constants $x_{ij}$ that appear in the equation for the vibrational energy of a polyatomic molecule with the harmonic frequencies $\omega_i^0$ and without degenerate vibrations:

$$G_0(v_1, v_2, v_3 \ldots) = \sum_i \omega_i^0 v_i + \sum_i \sum_{j \geq i} X_{ij}^0 v_i v_j \quad [3]$$

relies on the analysis of overtone and combination bands. In **Table 1** the data under the headings $\tilde{\nu}_1$, $2\tilde{\nu}_2$ and $2\tilde{\nu}_3 + \tilde{\nu}_6$ refer to a fundamental, an overtone and a combination band, respectively. The relations between $\tilde{\nu}_0$ of **Table 1** and $\omega_i^0$, $x_{ij}^0$ are given by

$$\text{for } \tilde{\nu}_1 \qquad \tilde{\nu}_0 = \omega_1^0 + x_{1,1}^0 \quad [4]$$

$$\text{for } 2\tilde{\nu}_2 \qquad \tilde{\nu}_0 = 2\omega_2^0 + 4x_{2,2}^0 \quad [5]$$

$$\text{for } 2\tilde{\nu}_3 + \tilde{\nu}_6 \qquad \tilde{\nu}_0 = 2\omega_3^0 + \omega_6^0 + 4x_{3,3}^0 + 2x_{3,6}^0 + x_{6,6}^0 \quad [6]$$

The assessment of the reliability of theoretical calculations is particularly important when the experiment is unable to determine some of the molecular constants, which in turn can be calculated by *ab initio* methods. In benzene, for example, there are about 300 anharmonic constants and there is little hope that all of these could be obtained experimentally. The equilibrium structure $r_e$ is also a very important result of the theoretical calculations of molecular force field, and is to be compared with the experimental structure (see **Table 2**).

**Table 2** $r_0$- and $r_e$-structures of acetylene

| | $r_0$ [a] (pm) | $r_e$ (pm) | $r_e$ [b] (pm) |
|---|---|---|---|
| $r_{C\equiv C}$ | 120.85 | 120.241 (9)[c] | 120.33 |
| $r_{C-H}$ | 105.72 | 106.25 (1)[c] | 106.05 |

[a] From $B_0$ of $C_2H_2$ and $C_2D_2$.
[b] Theoretical equilibrium structure.
[c] Standard errors in parentheses refer to the last significant digits.

## Laser Stark IR spectroscopy

The laser Stark technique applied to IR spectroscopy is based on bringing in resonance, with a fixed frequency IR laser (such as $CO_2$, CO, $N_2O$ lasers), the molecular transitions modulated via an electric field (Stark effect). From an analysis of the Stark spectra very accurate values of the electric dipole moments of molecules in the ground and excited vibrational states are obtained. **Table 3** presents some of the dipole moments that have been determined.

## Pressure broadening and pressure shift

The pressure shift and pressure broadening of the spectral lines are primarily due to the collisions of the molecules. The study of these effects provides information on the mechanism of the molecular interactions which are mainly of dipole–dipole type. In addition, the experimental determination of the line broadening and line shift parameters is instrumental in the spectroscopic study of the atmosphere, where the absorption or emission of IR radiation takes place from regions where the pressure may change by several orders of magnitude. These effects are relatively small; for example, for the P and R branches of the $\nu_3$ band of $^{12}C^{16}O_2$ at 2350 cm$^{-1}$, broadened by $N_2$ at 296 K, broadening coefficients in the 0.066–0.089 cm$^{-1}$ atm$^{-1}$ range and shift coefficients of –0.0016–0.0037 cm$^{-1}$ atm$^{-1}$ have been reported by Devi and co-workers.

**Table 3** Electric dipole moments from laser Stark spectroscopy

| *Molecule* | *Vibrational state* | $\mu$ (D)[a,b] |
|---|---|---|
| $PH_3$ | $v_2 = 1$ | 0.574 20(27) |
| | $v_4 = 1$ | 0.579 04(32) |
| $HNO_3$ | $v_2 = 1$ | 2.12(4) |
| HCOOH | Ground state | 1.395 7(9) |
| | $v_3 = 1$ | 1.421 6(10) |
| $CH_3F$ | Ground state | 1.857 |
| | $v_3 = 1$ | 1.905 |
| $ND_3$ | Ground state | 1.492(8) |
| | $v_2 = 1$ | 1.352(3) |
| $AsH_3$ | Ground state | 0.213(3) |
| | $v_2 = 1$ | 0.218(3) |

[a] 1 D = 3.335 64 × $10^{-30}$ cm
[b] Standard errors in parentheses refer to the last significant digits.

## High-resolution IR spectroscopy of weakly bound complexes

High-resolution IR spectroscopy has been used to elucidate the potential energy surfaces that govern the weak intermolecular forces between molecules, since it provides information on the vibrationally excited region of the potential energy surface of the ground electronic state. This region controls the long- and short-range collisional dynamics of the monomers. This application was delayed until the advent of adequately sensitive detection methods and of an appropriate light source. The experimental methods are based on either direct absorption techniques of light obtained from a laser source (from a tunable diode laser or from difference frequency mixing of a dye and $Ar^+$ lasers) by complexes in a cooled multiple pass gas cell or on FTIR techniques. A rotationally resolved spectrum can be recorded even though the line widths are Doppler limited or broadened by pressure or predissociation effects. From the analysis of these spectra thermodynamic properties, vibrational term values, rotational constants and bond dissociation energies may be extracted. Alternatively, supersonic molecular beam methods can be used with direct absorption laser spectroscopy or bolometric techniques. These experiments supply additional information on samples at rotational–vibrational temperatures ranging from 1–2K for expansion from an ultracold pinhole, to 10–30K for slit expansion. The spectroscopic techniques developed in the FAR-IR allow the investigation of the low-frequency intermolecular modes in complexes. Additional spectroscopic data on the large amplitude motions associated with these low frequency vibrations have been obtained from the analysis of the hot and/or combination bands observed in the mid/near-IR. Some of the complexes analysed so far are $(HF)_2$, $(DF)_2$, $(HCl)_2$, $(CO_2)_2$, $(HCN)_2$, $(HCCH)_2$, $(NH_3)_2$, $H_2$–inert gas, $H_2$–HF, HF–inert gas, HCl–inert gas, HF–$N_2$, HF–OCO, OC–HF, $C_2H_2$–HF, HCN–HF, HF–DF, $(DF)_3$. The HF and DF van der Waals complexes have attracted much attention because of their fundamental importance in the chemistry of hydrogen bonded species. Most of the spectra recorded in the near-IR are structured since the upper states are metastable with respect to vibrational (or rotational) predissociation. The main channel leading to the breakdown of the complex into its constituent subunits is the coupling of the intramolecular mode with the dissociating, low-frequency intermolecular modes. This coupling causes a redistribution of vibrational energy in the complex that may excite the intermolecular bond up to dissociation. The excited states have to last at least for the time needed to complete an end-over-end rotation in the complex to observe a rotationally resolved structure in the spectrum. From the experimental values of FWHM line width it is possible to calculate the lifetime of the metastable state, according to

$$\tau = 1/2\pi\Delta\nu_{\mathrm{FWHM}} \qquad [7]$$

for a Lorentzian absorption line shape and a single exponential decay of the upper state population. However, another possible contribution to the homogeneous line width comes from intramolecular vibrational relaxation (IVR). Consequently, the collision-free line width depends in principle on the combined 'predissociation/relaxation' pathway. From the analysis of high-resolution spectra a very large range of dissociation/relaxation behaviour results [$\tau > 3 \times 10^{-4}$ s for $\nu_1$ of Ar–HF (Huang and co-workers) to $\tau \geq 10$ ps for $\nu_2$ of the linear $(HCN)_3$ (Jucks and Miller)] indicating strong sensitivity to small differences in the potential energy surfaces. In many dimers a strong dependence of the line widths from specific vibrational couplings has been observed as in $(HF)_2$ [$\Delta\nu_1 = 13.4$ MHz, $\Delta\nu_2 = 408$ MHz (Pine and co-workers and Huang and co-workers)], HCN–HF [$\Delta\nu_1 = 2722$ MHz, $\Delta\nu_2 = 11.8$ MHz, $\Delta\nu_3 = 558$ MHz (Wofford and co-workers)]. The analysis of spectroscopic data may give insight into the coupling between intermolecular bending and stretching modes, as indicated by the results of the Coriolis analysis in the Ar–HF dimer between the $\nu_2$ perpendicular band and the $3\nu_3$ overtone of the intermolecular stretching vibration (Lovejoy and co-workers). Additional probes of intermolecular interactions are the intensity changes and the amount of the shift of the frequency of the monomer fundamentals in complexes from their unperturbed position. From IR experiments it is possible to infer the correct equilibrium structure of a complex, i.e. that corresponding to the absolute min-imum of the potential surface. Information about large amplitude motions from hot bands and combination bands analysis are also very useful to this end. Results indicate that for many complexes, i.e. $CO_2$, $N_2O$ dimers, and HCN trimers, several structures, corresponding to a different minima in the potential surface, are possible and that the isomerization involves motion along intermolecular coordinates. In the case of $Ar_nHF$ clusters a satisfactory agreement is observed between the isomer equilibrium structures predicted on the basis of pairwise additive potentials and the experimental observations (see **Figure 1**).

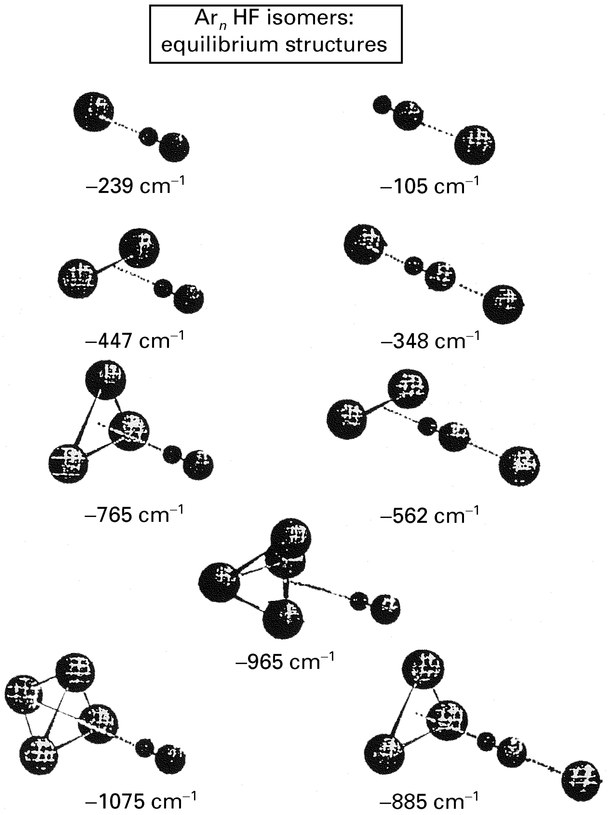



**Figure 1** Equilibrium structures and energies for $Ar_nHF$ ($n = 1$–4) predicted from pairwise additive potentials. In all cases, the lowest predicted energy structures (left) are in excellent agreement with experimental observation. Reprinted with permission of Annual Reviews from Nesbitt DJ (1994) High-resolution, direct infrared laser absorption spectroscopy in slit supersonic jet: intermolecular forces and unimolecular vibrational dynamics in clusters. *Annual Review of Physical Chemistry* **45**: 369–399.

## High-resolution IR spectroscopy of free radicals and ions

IR laser spectroscopy has been employed to study free radicals and ions because of its sensitivity in the detection of compounds present in small concentrations. The spectroscopic characterization of radicals and ions is valuable in the fields of chemical kinetics, of astronomy and of plasma diagnosis. The experimental techniques used to produce unstable molecules and to record the spectra are here briefly summarized. Free radicals can be generated using the discharge-flow method, while ionic species can be obtained using a hollow-cathode discharge cell that, combined with magnetic-field modulation or amplitude-discharge modulation, allows the detection of the selected ionic species among neutral stable molecules. Alternatively, velocity modulation IR laser spectroscopy can be used to selectively detect ionic species among neutral ones, exploiting the characteristic Doppler shift of charged molecules. $HD^+$, $HeH^+$, $NeH^+$, $ArH^+$, $HCO^+$ and $HNN^+$ are some of the ions studied with this technique in the infrared and whose fundamental frequencies and molecular constants have been determined. Both radicals and ions may be generated by photolysis or photoionization of a stable precursor seeded in a rare-gas free-jet expansion. Another powerful technique with which to study ions employs a fast ion beam extracted from a plasma ion source and velocity tuned. This is merged with $CO_2$ laser light to induce transitions from a vibrational bound level near the dissociation limit to a repulsive electronic state or to a predissociating bound level above the dissociation limit. Most of the systems studied are diatomic, such as $HD^+$, $HeH^+$, $He_2^+$, $H_2^+$ and $D_2^+$, and the information obtained concerns the description and spectroscopy of weakly bound states, to a infer deeper understanding of ion–neutral reaction dynamics and of the charge–dipole induced interaction. In addition, polyatomic unstable species have been the subject of detailed study with high-resolution IR spectroscopy, among those examined are CCH, $C_3$, $C_4$, $C_9$, $C_{13}$, $NO_3$, $HO_2$, HCCO and $C_3H_3$ radicals, and $H_3^+$, $SiH_3^+$, $CH_3CNH^+$, $CH_2^+$, $NH_3^+$ and $H_2O^+$ ions. The results of such experiments may be valuable in the field of chemical kinetics since they allow the determination of rate constants and an understanding of some aspects of the mechanism of the reaction. For example, the observation, in a hollow cathode discharge, of the $\nu_1$ band of $HOC^+$ at 3268 cm$^{-1}$ allowed the measurement of the abundance ratio [$HCO^+$]:[$HOC^+$] in the laboratory from the relative intensity of the IR high-resolution lines of the two species, using the transition dipole moments for the band of $HCO^+$ (0.168 D) and $HOC^+$ (0.350 D), and assuming that the rotational temperature of both ions was equal to the cell temperature (223 K) (Amano). This information was particularly valuable in studying the formation and depletion of $HOC^+$ in the interstellar medium. A rate equation analysis was performed to rationalize the observed dependence of the relative intensities of $HOC^+$ and $HCO^+$ lines as a function of the pressures of $H_2$ and CO. The rate constants of some of the reactions considered in the rate equation analysis were then determined. The contribution that the analysis of the high-resolution IR spectra of such kind of molecules brings to astronomy is the determination with great accuracy of the molecular constants that allow one to search for the rotational lines in laboratory microwave spectra. These results may lead to the identification of unassigned interstellar lines. The nature and the amount of a species present in plasmas may

be determined by high-resolution IR spectroscopy, as in the case of $SiH_3^+$, a transient molecular ion observed and analysed in a silane discharge plasma. The $\nu_2$ and $\nu_4$ bands, i.e. the out-of-plane bending and the in-plane degenerate bending vibration respectively, have been recorded and analysed (Davies and Smith), taking into account the Coriolis interaction between the two vibrationally excited states.

## Molecular dynamics

Frequency domain spectroscopic techniques can be widely used to determine the homogeneous intramolecular vibrational redistribution (IVR) of energy, a phenomenon observed when molecules have enough energy to break bonds. An advantage of the frequency domain techniques is the ability to measure the rotational state-dependence of the IVR rates. Briefly, the background theory is that a single rotational state in an excited anharmonic normal-mode vibrational state (the bright state) carries all the oscillator strength and is coupled to near-resonant vibrational dark states through perturbation terms of the Hamiltonian. As a result of this coupling the bright state oscillator strength is redistributed among the dark states. The width of the intensity distribution is a measure of the time scale of energy localization in the bright state. Thus, one goal of IVR studies is the measurement of the relaxation rates that span a wide range (50 fs to 10 ns) and are dependent on molecular structure. The two major experimental methods used for IVR studies of isolated molecules in their ground electronic state can be divided into single photon and double resonance techniques. The general requirements of these experimental methods are both high resolution and sensitivity. Quantitative IVR rates of many molecules have been determined: they span from $10^8$–$10^{13}$ $s^{-1}$, with $3 \times 10^9$ $s^{-1}$ being a typical value. These rates may be sufficiently slow to permit mode-selective chemistry in some systems. From a comparative analysis of the IVR rate values for the acetylenic C–H stretching for many molecules of the type $(CY_3)_3XCCH$ (see **Table 4**), it appears that the lifetime and density of states are not correlated, even if the state density must be large enough to provide a bath for energy redistribution. The IVR rate is controlled by the coupling of a few dark states that interact with the bright one through low order terms of the potential energy surface, which in turn must be efficiently coupled to other non-resonant states to allow fast energy flow. Other structural aspects of IVR such as the role of large amplitude motion in IVR dynamics, the change of IVR rates by structural modification, and the degree of dependence of the IVR rates on modes are currently being studied.

**Table 4** Experimental IVR lifetimes and states densities for the acetylenic C–H stretch in molecules of the form $(CY_3)_3XCCH$. Adapted with permission of Annual Reviews from Lehmann KK, Scoles G and Pate BH (1994) Intramolecular dynamics from eigenstate-resolved infrared spectra. *Annual Review of Physical Chemistry* **45**: 241–274.

| *Molecule* | *Level of excitation* | *Density of states*[a] (cm) | *Lifetime* (ps) |
|---|---|---|---|
| $(CH_3)_3CCCH$ | $v = 1$ | $4.9 \times 10^2$ | 200 |
| | $v = 2$ | $6.2 \times 10^5$ | 110 |
| $(CD_3)_3CCCH$ | $v = 1$ | $2.8 \times 10^3$ | 40 |
| | $v = 2$ | $7.6 \times 10^6$ | < 40[b] |
| $(CH_3)_3SiCCH$ | $v = 1$ | $1.0 \times 10^4$ | 2 000 |
| | $v = 2$ | $2.9 \times 10^7$ | 4 000 |
| $(CD_3)_3SiCCH$ | $v = 1$ | $1.0 \times 10^5$ | 830 |
| | $v = 2$ | $6.0 \times 10^8$ | 140 |
| $(CH_3)_3SnCCH$ | $v = 1$ | $1.0 \times 10^6$ | 6 000 |
| | $v = 2$ | $1.0 \times 10^9$ | > 1 000[c] |
| $(CF_3)_3CCCH$ | $v = 1$ | $4.2 \times 10^6$ | 60 |
| | $v = 2$ | $1.0 \times 10^{11}$ | < 40[b] |

[a] The state densities (states $cm^{-1}$) are for vibrational bath states of $A_1$ symmetry, which can couple via anharmonic interactions.
[b] Upper limit of the lifetime. The overtone absorption was unobservable, despite observation of the fundamental.
[c] Lower limit of the lifetime. The Q branch structure was too sharp to obtain an estimate of the Lorentzian wing. Other inhomogeneities (either isotopes, torsional, or K structure) were also present.

## Thermodynamic and kinetic data from high-resolution spectroscopic IR studies

An example of this application is the FTIR and IR laser diode spectroscopy study of the reversible gas-phase reaction

$$HONO + NH_3 \rightleftarrows H_3N - HONO$$

A kinetic study has been carried out (Pagsberg *et al.*) by monitoring selected absorption signals of the rovibration transitions of the $\nu_3$ and $\nu_5$ IR bands of HONO as a function of time for different partial pressures of ammonia at a total pressure of 20 mbar. At 298 K the rate constant of the forward reaction and the equilibrium constant are $k_f = 2.2 \pm 0.2 \times 10^3$ $M^{-1}$ $s^{-1}$ and $K_c = 1.5 \pm 0.2 \times 10^5$ $M^{-1}$, respectively. From the observed temperature dependence of the equilibrium constant expressed in terms of the

van't Hoff equation, the values of $\Delta_r H^\circ_{298} = -49.4 \pm 3.3$ kJ mol$^{-1}$ and $\Delta_r S^\circ_{298} = -111.7 \pm 4.2$ J mol$^{-1}$ K$^{-1}$ have been derived.

## List of symbols

$B_e$ = equilibrium rotational constant; $B_{v_i}$ = rotational constant in the $v_i$ vibrational mode; $B_x$ = rotational constant; $D$, $d$ = quartic centrifugal distortion constants; $G_0$ ($v_i$) = vibrational term value when the energy levels are referred to the ground state as zero; $H$, $h$ = sextic distorsion constants; $k_B$ = Boltzmann constant; $k_f$ = rate constant of the forward reaction; $K_c$ = equilibrium constant; $r_{C-H}$ = interatomic distance between C and H atoms; $T$ = (thermodynamic) temperature; $v_i$ = vibrational quantum number; $W$ = Fermi resonance parameter; $x^0_{ij}$ = vibrational anharmonic constants when the energy levels are referred to the ground state as zero; $\alpha_i$ = rovibration interaction constant; $\Delta_r H^\circ_{298}$ = standard reaction enthalpy; $\Delta_r S^\circ_{298}$ = standard reaction entropy; $\Delta\nu$ = full line width at half intensity maximum; $\mu$ = dipole moment; $\tilde{\nu}(v_i)$ = transition wavenumber in vacuum; $\xi$ = Coriolis constant; $\tilde{\nu}$ = wavenumber in vacuum; $v_i\tilde{\nu}_0$ = wavenumber of a rovibrational band origin; $\tau$ = relaxation time; $\omega^0_i$ = harmonic vibration wavenumber when the energy levels are referred to the ground state as zero.

*See also:* **High Resolution IR Spectroscopy (Gas Phase) Instrumentation; Interstellar Molecules, Spectroscopy of; IR Spectroscopy, Theory; Laser Magnetic Resonance; Microwave and Radiowave Spectroscopy, Applications; Rotational Spectroscopy, Theory.**

## Further reading


Amano T (1988) High resolution infrared spectroscopy of molecular ions. *Philosophical Transactions of the Royal Society of London, Section A* **324**: 163–178.

Barrow RF and Crozet P (1997) Gas-phase molecular spectroscopy. *Annual Reports on the Progress of Chemistry, Section C, Physical Chemistry* **93**: 187–256.

Brian JH and Brown JM (1992) High resolution infrared spectroscopy. In: Andrews DL (ed) *Applied Laser Spectroscopy*, pp 185–225. New York: VCH.

Davies PB and Smith DM (1994) Diode laser spectroscopy and coupled analysis of the $\nu_2$ and $\nu_4$ fundamental bands of $SiH_3^+$. *Journal of Chemical Physics* **100**: 6166–6174.

Devi VM, Benner DC, Rinsland CP and Smith MAH (1992) Measurements of pressure broadening and pressure shifting by nitrogen in the 4.3 μm band of $^{12}C^{16}O_2$. *Journal of Quantum Spectroscopy and Radiation Transfer* **48**: 581–589.

Gudemann CS and Saykally RJ (1984) Velocity modulation infrared laser spectroscopy of molecular ions. *Annual Review of Physical Chemistry* **35**: 387–418.

Hirota E (1994) Rotational and vibrational spectra of free radicals and molecular ions. *Annual Reports on the1 Progress of Chemistry, Section C, Physical Chemistry* **91**: 3–36.

Hollas JM (1982) *High Resolution Spectroscopy*. London: Butterworth.

Huang ZS, Jucks KW and Miller RE (1986) The argon-hydrogen fluoride binary complex: an example of a long lived metastable system. *Journal of Chemical Physics* **85**: 6905–6909.

Huang ZS, Jucks KW and Miller RE (1986) The vibrational predissociation lifetime of the HF dimer upon exciting the "free-H" stretching vibration. *Journal of Chemical Physics* **85**: 3338–3341.

Jacox ME (1998) Vibrational and electronic energy levels of polyatomic transient molecules. *Journal of Physical and Chemical Reference Data* **27**: 115–412.

Jucks KW and Miller RE (1988) Near infrared spectroscopic observation of the linear and cyclic isomers of the hydrogen cyanide trimer. *Journal of Chemical Physics* **88**: 2196–2204.

Lehmann KK, Scoles G and Pate BH (1994) Intramolecular dynamics from eigenstate-resolved infrared spectra. *Annual Review of Physical Chemistry* **45**: 241–274.

Lovejoy CM, Schuder MD and Nesbitt DJ (1986) High resolution IR laser spectroscopy of the Van der Waals complexes in slit supersonic jets. *Journal of Chemical Physics* **85**: 4890–4902.

Nesbitt DJ (1994) High-resolution, direct infrared laser absorption spectroscopy in slit supersonic jet: intermolecular forces and unimolecular vibrational dynamics in clusters. *Annual Review of Physical Chemistry* **45**: 367–399.

Pagsberg P, Ratajczak E, Sillesen A and Lotaika Z (1994) Kinetics and thermochemistry of the reversible gas phase reaction HONO + $NH_3 \rightarrow H_3N$-HONO studied by infrared diode laser spectroscopy. *Chemical Physics Letters* **227**: 6–12.

Pine AS, Lafferty WJ and Howard BJ (1984) Vibrational predissociation, tunneling, and rotational saturation in the HF and DF dimers. *Journal of Chemical Physics* **81**: 2939–2950.

Rao KN and Weber A (eds) (1992) *Spectroscopy of the Earth's Atmosphere and Interstellar Medium*. New York: Academic Press.

Weber WH, Tanaka K and Tanaka T (eds) (1987) Stark and Zeeman techniques in laser spectroscopy. *Journal of the Optical Society of America, Part B* **4**: 1141–1172.

Wofford BA, Bevan JW, Olson WB and Lafferty WT (1985) Rovibrational analysis of $\nu_3$ HCN–HF using Fourier transform infrared spectroscopy. *Journal of Chemical Physics* **83**: 6188–6192.

# High Resolution Solid State NMR, $^{13}C$

**Etsuko Katoh**, National Institute of Agrobiological Resources, Tsukuba, Japan
**Isao Ando**, Tokyo Institute of Technology, Japan

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

## Introduction

After the first NMR experiment for obtaining high-resolution spectra of solids was carried out with the high-speed magic-angle spinning (MAS) method in 1958, the cross polarization (CP) and multipulse methods were developed as important for high-resolution solid-state NMR experiments. On the basis of these methods a CP MAS technique that combines MAS and CP has been conventionally used to obtain high-resolution solid-state NMR spectra.

In the solution state, the observable NMR chemical shift is often the average value of some preferred conformations of a molecule because of its rapid motion, but in the solid state the chemical shift is often characteristic of a specified conformation because of the highly restricted molecular motion. Thus, the chemical shifts can be used as well as the spin–lattice relaxation time ($T_1$), the spin–spin relaxation time ($T_2$) and $T_1$ in the rotating frame ($T_{1\rho}$) to provide information on the dynamics. At present, $^{13}C$ CP MAS has been widely employed to characterize the structure and dynamics of solid samples such as polymers, biopolymers, catalysts etc. Here, we will introduce some applications of $^{13}C$ CP MAS NMR for the characterization of the structure and dynamics of polymers, including polypeptides and proteins and also briefly describe the applications of solid-state $^{13}C$ NMR spectroscopy to food materials and coals. There is also a large body of literature on high-resolution solid-state $^{13}C$ NMR spectroscopy of small molecules, although the methods used are largely the same as those given here.

## Synthetic polymers

Polymer materials are almost always used as solids. It is well known that the properties of polymers depend on their structure and dynamics. Therefore, to design new polymer materials, it is necessary to understand more deeply the property–structure relationships. The X-ray diffraction method has provided the structure of polymers which have high crystallinity. However, most polymers have low crystallinity and so structural information about the non-crystalline region, which is the major component, cannot be obtained by X-ray studies. Therefore, X-ray diffraction is of limited use in the structural analysis for such systems. Chain segments in the non-crystalline region are sometimes in a mobile state and so the X-ray diffraction method provides no structural or dynamic information. On the other hand, solid-state NMR provides knowledge of the structure and dynamics of a sample irrespective of whether the region studied is crystalline or non-crystalline.

### Polyethylene

Polyethylene (PE) is a typical plastic. The polymer chain exhibits a *trans* and two kinds of *gauche* conformations, between which the energy difference is small and so the polymer takes various configurations under different conditions. PE has two kinds of crystal forms, as determined by several spectroscopic methods such as NMR, X-ray diffraction, electron diffraction, IR and neutron diffraction methods. One form is the orthorhombic form which occurs under normal conditions and another is the monoclinic form which occurs under high pressure or drawn conditions. In both crystal forms the conformation is the all-*trans* zigzag conformation, but the chain arrangements are different from each other. The all-*trans* zigzag planes in the orthorhombic form are perpendicular to each other, and in the monoclinic form they are parallel. Further, there also exist the non-crystalline and interfacial phases. Their existence strongly affects the physical properties of the polymer. Therefore, it is very important to analyse the detailed structure of PE in the solid state to understand its physical properties.

**Figure 1** shows $^{13}C$ CP MAS NMR spectra of single crystal PE (SC-PE), melt-quenched PE (MQ-PE) and drawn PE (DR-PE). Only one sharp peak (33.0 ppm) appears in the SC-PE spectrum. However, MQ-PE and DR-PE have two and three peaks, respectively. This means that MQ-PE and DR-PE have at least two or three kinds of magnetically inequivalent carbon atoms. These peaks have been assigned by $T_1$ measurements; peak a (31 ppm) comes from the non-crystalline region, which is a

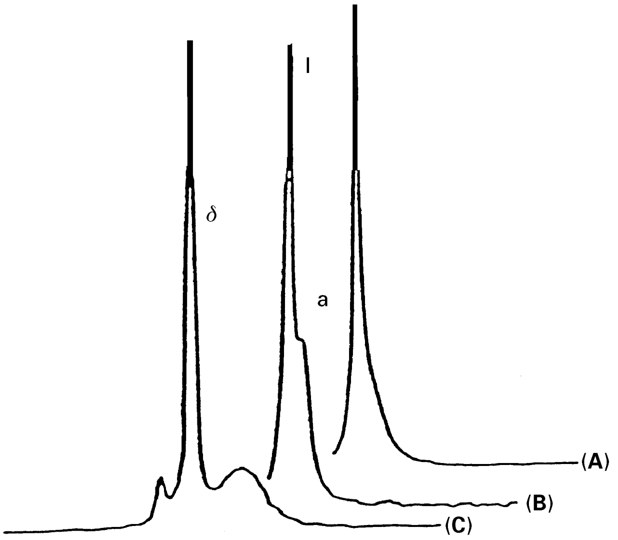



**Figure 1** $^{13}$C CP MAS NMR spectra of polyethylene. (A) SC-, (B) MQ- and (C) DR-PE. Parts (A) and (B) reproduced with permission of Elsevier Science from Ando I, Sorita T, Yamanobe T, Komoto T, Sato H, Deguchi K and Imanari M (1985) *Polymer* **26**: 11 864. Part (C) reproduced with permission of Elsevier Science from Van der Hart DL and Khoury F (1984) *Polymer* **25**: 1589.

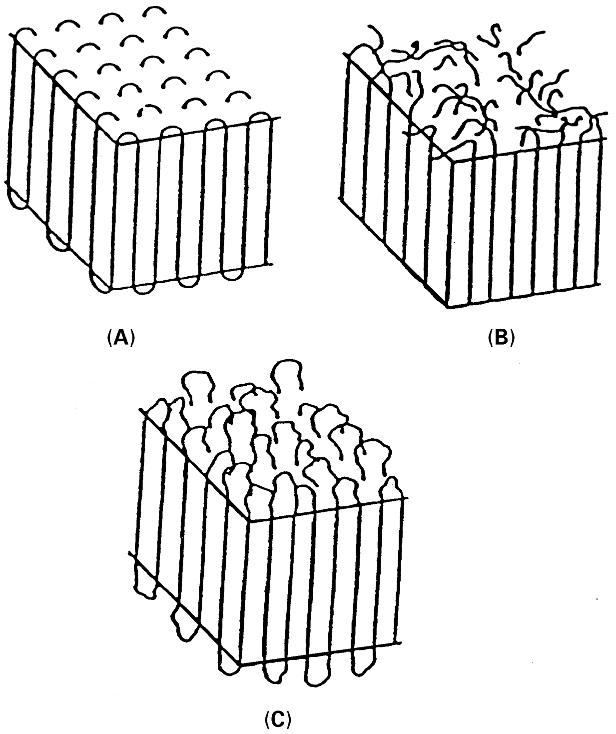



**Figure 2** Schematic drawing of the conformation models of a polyethylene single crystal. (A) sharp-fold, (B) switch-board and (C) loose-loop model. Reproduced with permission of Elsevier Science from Ando I, Sorita T, Yamanobe T *et al* (1985) *Polymer* **26**: 1864.

mobile state, and peak I (33 ppm) from the crystalline region, which is in an immobile state. Further, the high-frequency peak of DR-PE (35.0 ppm) is assigned to the methylene carbons in the monoclinic crystal region. These assignments were justified by quantum chemical shielding calculations with the tight-binding (TB) sum-over-states (SOS) method. A peak from the interfacial region between the crystal and non-crystalline regions has also been reported.

The crystal structure of PE has been proposed from models such as sharp-fold, switch-board and loose-loops models as shown in **Figure 2A–C**. To obtain detailed information on the fold structure, SC- and MQ-PE have been studied by high-resolution solid-state $^{13}$C NMR. This study reveals that SC- and MQ-PE take the adjacent re-entry type of macroconformation in addition to loose and long loops in the fold surface, as shown in **Figure 3**. The number of carbon atoms in the *trans* zigzag chain from one fold to the next is estimated to be ~100, which leads to an estimate of the stem length of about 125 Å. Its magnitude is consistent with the crystal thickness (120–150 Å) measured directly for PE single crystals by electron microscopy.

The physical properties of polymers in the solid state are strongly affected by temperature. Variable-temperature (VT) NMR techniques should provide useful information on the structural and dynamic aspects of polymers in the solid state. The advantage of VT solid-state NMR techniques is that the temperature dependence of the conformation and molecular motion can be studied by observing the chemical shift. The structure of ultrahigh relative molecular weight PE (UHRMW-PE) has been investigated in the solid state by VT $^{13}$C CP MAS NMR. The $^{13}$C NMR spectrum of UHRMW-PE at room temperature has two peaks (the low and high frequency peaks are designated by A and I, respectively) as shown in **Figure 4**. Their chemical shifts agree with those for the melt-crystallized PE (MC-PE), and so peaks I and A have been assigned to the crystalline region with the *trans* zigzag conformation and to the non-crystalline region, respectively. When the temperature is decreased, peak A shifts to low frequency but peak I shifts to high frequency. The $^{13}$C chemical shift behaviour for peak A of UHRMW-PE is similar to that of MQ-PE. At –108°C, the molecular motion is frozen and the chemical shift for peak A is about 32 ppm. Therefore, peak A is due to the methylene carbons in the *trans* conformation in the non-crystalline region since in the frozen state a methylene carbon in the *gauche* conformation should appear at about 27 ppm because of the γ-*gauche* effect. On the other hand, the $^{13}$C chemical

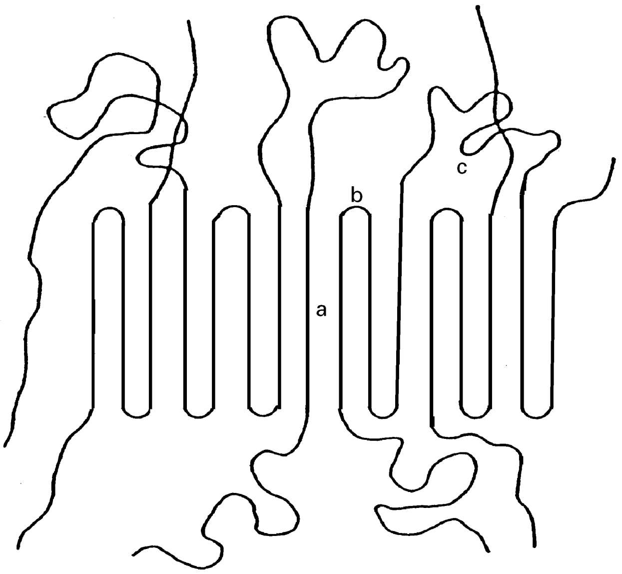



**Figure 3** Schematic illustration of the folded chain conformation form of melt-quenched polyethylene. (a) *trans* zigzag, (b) sharp-fold and (c) loose and long loop. Reproduced with permission of Elsevier Science from Ando I, Sorita T, Yamanobe T, *et al.* (1985) *Polymer* **26**: 1864.

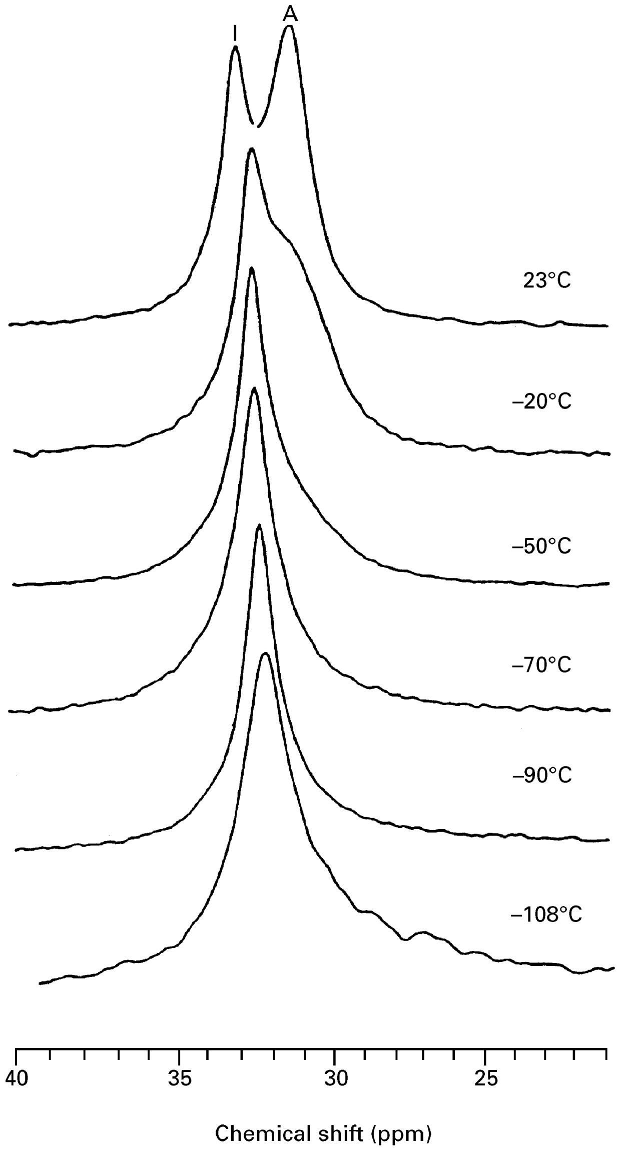



**Figure 4** $^{13}C$ CP MAS NMR spectra of ultrahigh relative molecular mass polyethylene as a function of temperature. Reproduced with permission of John Wiley & Sons from Akiyama (1990) *Journal of Polymer Science, Part B: Polymer Physics* **28**: 587.

shift of peak I shifts to low frequency, which is opposite to the case of MQ-PE, with decreasing temperature. At –108°C, the chemical shift is about 32 ppm, and peaks A and I coalesce. It is difficult to conclude that the cause for the low frequency shift of peak I is a change of structure in going from the orthorhombic form to a different crystal structure. It is known that the $^{13}C$ shifts of methylene carbons in the orthorhombic, triclinic and monoclinic forms are about 33, 34 and 35 ppm, respectively. From these results, the $^{13}C$ shift of 32 ppm at –108°C indicates that the crystalline structure is different from other crystal structures found at room temperature. The other possibility for the low frequency shift is a distortion of the orthorhombic form.

The relaxation parameters, $T_1$, dipolar-dephasing relaxation times ($T_{DD}$), etc. can provide useful information on the dynamics of polymer motion in the solid state and to monitor solid-state polymer phase transitions.

The non-crystalline region of two kinds of PE samples, using single $^{13}C$-labelled solution-crystallized PE (PE-SL) and single $^{13}C$-labelled melt quenched PE (MQ-PE-SL), has been studied by VT $^{13}C$ CP MAS NMR. The dynamics of the non-crystalline region have been discussed by measuring $^{13}C$ $T_1$ and $T_{DD}$ from –120 to 44°C. Each of these spectra consists of three peaks, corresponding to an orthorhombic crystalline peak at 33.0 ppm, a monoclinic crystalline peak at 34.4 ppm (a small shoulder on the left-hand side of the orthorhombic peak) and a non-crystalline peak which appears at 30.8–31.3 ppm. The $^{13}C$ $T_1$ data of these samples over a wide range of temperature, obtained using the inversion–recovery method with the pulse saturation transfer (PST) pulse sequence, are given. The PST pulse sequence enhances the intensity of mobile methylene carbons. The $T_1$ values for samples PE-SL and MQ-PE-SL are plotted against the inverse of the absolute temperature ($1/T$) in **Figure 5A**. It has been suggested previously that local molecular motion in the non-

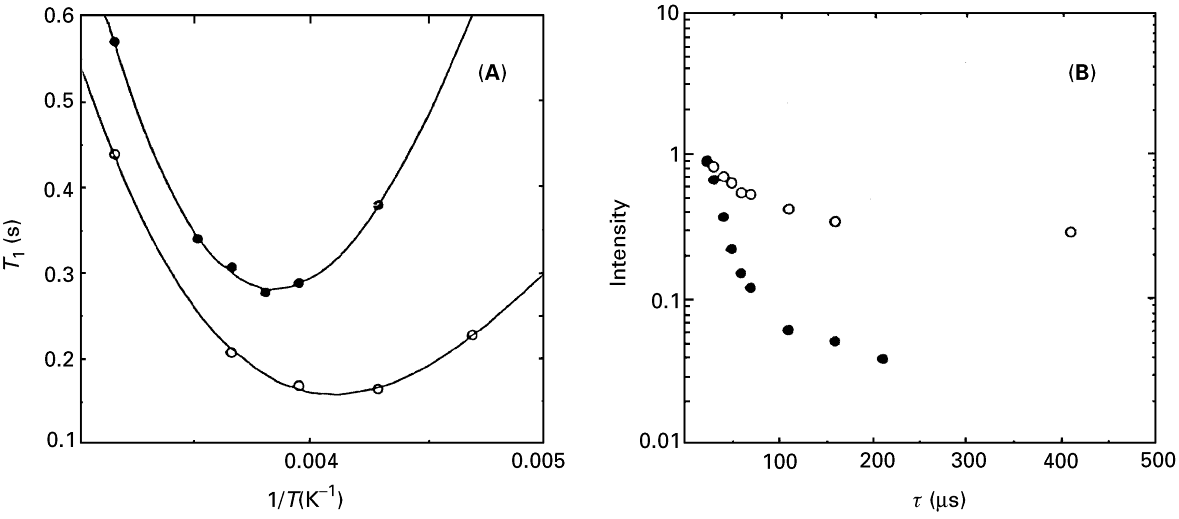



**Figure 5** (A) Plots of non-crystalline $^{13}C$ spin–lattice relaxation times: $T_1$ of PE-SL samples (●) and MQ-PE-SL (O) versus the reciprocal absolute temperature. (B) Intensity of non-crystalline peaks of a sample of PE-SL and MQ-PE-SL versus delay time $\tau$. The peak intensity was obtained from computer simulation of the $^{13}C$ partially relaxed dipolar-dephasing NMR spectra. Reproduced with permission of John Wiley & Sons from Chen Q, Yamada T, Kurosu H, Ando I, Shioino T and Doi Y (1992) *Journal of Polymer Science, Part B: Polymer Physics* **30**: 591.

crystalline region of PE is independent of the degree of crystallinity, higher-order structures or morphologies. However, these suggestions are not supported by the experimental results of PE-SL and MQ-PE-SL for two reasons: one is the difference in $T_1$ values and the other is the large difference in the temperature of the $T_1$ minimum of the non-crystalline region between PE-SL and MQ-PE-SL, as shown in **Figure 5A**. These facts show that local molecular motions in the non-crystalline regions of the PE-SL sample are more constrained than that of MQ-PE-SL. To study whether the dynamic behaviour of the two kinds of non-crystalline region is also different on the $T_2$ time-scale, $T_{DD}$ values of samples of PE-SL and MQ-PE-SL were measured over a wide temperature range. The relative intensity of the non-crystalline peak obtained from computer simulation was plotted against the delay time τ in **Figure 5B**. It can be clearly seen that the non-crystalline peak of the sample of PE-SL relaxes more quickly than that of the MQ-PE-SL sample. The value of $T_{DD}$ depends on molecular motion, carbon–proton dipolar interactions, MAS rate and spin diffusion. Fundamentally, it can be said that the dipolar dephasing time in the non-crystalline region becomes a measure of molecular motion because of the high mobility. Therefore, the longer $T_{DD}$ value of MQ-PE-SL, compared with that of PE-SL obviously suggests that the carbon–proton dipolar interaction is partially averaged by molecular motion on the $T_2$ time-scale.

### Poly(vinyl alcohol)

Poly(vinyl alcohol) (PVA) is used as a high-performance fibre and as a barrier material against oxygen. Detailed information about the structure and molecular motion is needed in order to develop and design this material. The $^{13}C$ CPMAS NMR spectra of PVA have been measured at room temperature. The methine carbon resonance split into three peaks in the CP MAS NMR spectrum. In the case of solution state NMR, the three methine carbon resonances and relative intensities correspond to the triad tacticity. However, the chemical shifts of the three methine carbons in the solid state are about 77, 71 and 65 ppm **(Figure 6)** and the relative intensities are not consistent with the triad tacticity in solution. The chemical shifts of two high-frequency peaks move significantly to higher frequencies in the solid state. These peaks, on the basis of the formation of intramolecular hydrogen bonds, have been assigned as follows. The highest frequency peak of the methine carbon is assigned to the *mm* triad with two intramolecular hydrogen bonds, the second highest one is assigned to the *mm* and *mr* triad with one intramolecular hydrogen bond and the other peak is assigned to the *mm*, *mr* and *rr* triads with no intramolecular hydrogen bonds.

Most recently, the $^{13}C$ CP MAS NMR spectra of PVA gels were measured to clarify the structure of their immobile component. In the $^{13}C$ CP MAS NMR

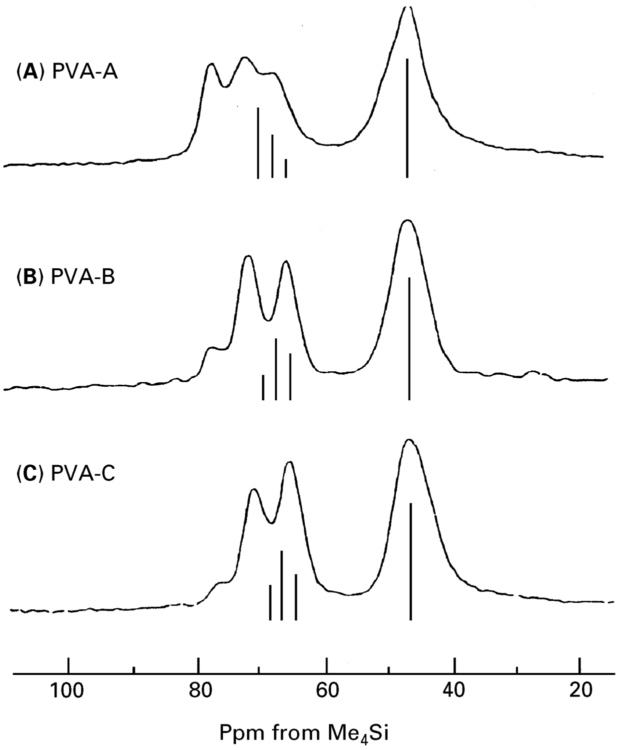



**Figure 6** $^{13}C$ CP-MAS spectra of poly(vinyl alcohol) at room temperature. (A) syndiotactic, (B) isotactic and (C) atactic. Line spectra represent the $^{13}C$ spectra in DMSO-$d_6$ solution at 353 K. Reproduced with permission of American Chemical Society from Terao T, Maeda S and Saika A (1983) *Macromolecules* **16**: 1535.

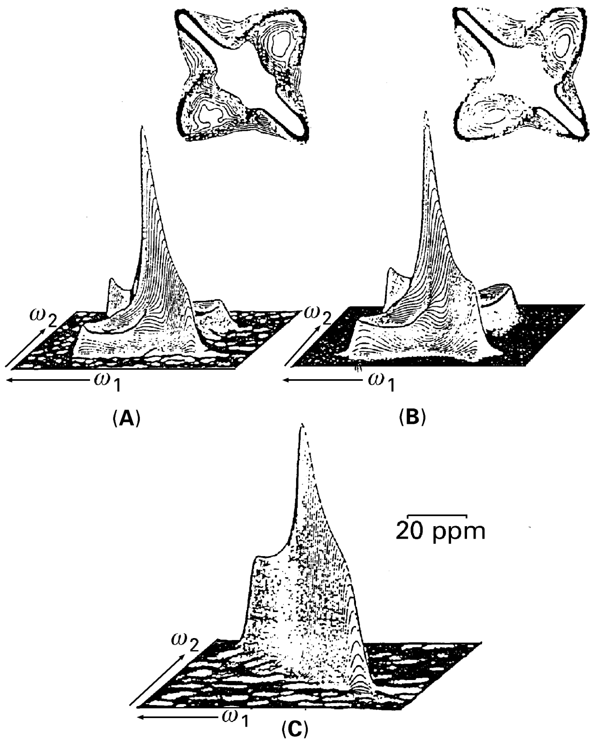



**Figure 7** (A) Pure absorption-mode 2D exchange spectra of isotactic POM at $T = 360$ K and mixing time $t_m = 2$ s, (B) theoretical spectrum of (A) with $x = 2\omega$, $t_m = 2.5$ s and (C) $T = 252$ K and $t_m = 1$ s. Reproduced with permission of Academic Press from Hagemeyer A, Schmidt-Rohr K and Spiess HW (1989) *Advances in Magnetic Resonance* **13**: 85.

spectra, the three carbon peaks were clearly observed, and originate from the formation of strong intermolecular or intramolecular hydrogen bonds between OH groups such as PVA in the solid state as mentioned above. Further, the molecular motion of immobile and mobile regions of PVA gel through the observation of $^{13}C$ $T_1$ has been studied.

## Polyoxymethylene

The molecular motion of polyoxymethylene (POM), which is the first member of the polyether series, has been investigated using the two-dimensional (2D) exchange experiment. The observed and calculated 2D exchange spectra of POM are shown in **Figures 7A** and **7B**, respectively. The spectrum observed at 252 K shows only diagonal signals and thus no molecular motion occurred during the mixing time. However, if molecular motion occurred during the mixing time, the spectrum would show steric effects as shown in **Figure 7C**. The calculations were performed by using the model of helical jump motions, assuming a one-dimensional random walk in continuous time and elementary processes of 200° jumps. Thus, at $T = 360$ K the subspectra can be identified for the discrete steps a given $CH_2$ group of POM experiences, as the helix rotates.

## Fluoropolymers

Fluoropolymers have excellent properties and have been widely used for various purposes under severe conditions. Although high-resolution solid-state NMR is a powerful tool with which to elucidate structures and dynamics of polymers in the solid state as mentioned above, the $^{13}C$ CP MAS experiment often fails to provide detailed information on the structures of fluoropolymers because their spectra gives rise to broad peaks. This is mainly due to two dominant factors, namely large dipolar coupling and the shielding anisotropy of fluorine nuclei. Although the high-speed magic-angle spinning (HS MAS) method leads to a significant reduction in such broadening for $^{19}F$ NMR in the solid state and has been used to characterize fluoropolymers, only a few $^{13}C$ solid-state NMR experiments have been reported. The simultaneous application of proton and fluorine decoupling drastically reduces line broadening caused by $^1H$–$^{13}C$ and $^{19}F$–$^{13}C$ dipolar interactions. The two kinds of poly(vinylidene fluoride), non-polar α phase

and polar β phase, have been characterized by means of solid-state triple-resonance spectroscopy. There was a small difference (~ 4 ppm) in the chemical shift positions for these distinct crystalline modifications.

### Blend polymers

Multicomponent organic polymer materials such as polymer blends are important in industry. Their macroscopic properties are determined not only by the molecular properties of the individual polymers, but also by the miscibility, morphology and structure of the interfaces between the polymers. Solid-state NMR is a useful tool to study the microscopic structure of heterogeneous polymer materials.

The miscibility of PMMA and poly(vinylidene fluoride) (PVDF) has been established by various methods. Both $^{19}$F–$^{13}$C and $^{19}$F–$^{1}$H CP and molecular miscibility in blends of PVDF and isotactic, syndiotactic and atactic PMMAs were investigated by triple resonance $^{1}$H, $^{19}$F, $^{13}$C solid state CP MAS NMR. The fraction of nonmixed PMMA was determined for these blends, and it was found that this fraction was smaller for isotactic than for atactic and syndiotactic PMMAs.

A study of the polymer blends of poly(vinyl phenol) and poly(methyl acrylate) using 2D $^{13}$C–$^{1}$H correlation NMR have been reported. From the spectrum, a direct interaction between the hydroxyl hydrogen of poly(vinyl phenol) and the carbonyl carbon of poly(methyl acrylate) can be deduced. The miscibility in annealed 50:50 blends of a random, copolyester Vectra-A, containing 73% *p*-hydroxybenzoic acid and 27% hydroxynaphthoic acid, and poly(ethylene terephthalate) (PET) has been discussed by using a standard 2D exchange experiment. The small cross peaks between the peak of aliphatic PET and that of quaternary Vectra-A carbons diagonal lines indicated intimate mixing.

## Biopolymers

In the solution state, the NMR chemical shifts of biopolymers with internal rotations are often averaged values for all internal rotations because of rapid interconversion by rotation about peptide bonds. In the solid state, however, chemical shifts are often characteristic of specific conformations because of highly restricted rotation about the peptide bonds. The NMR chemical shift is affected by a change in the electronic state arising from the conformational change. NMR chemical shifts in the solid state, therefore, provide useful information about the electronic state and conformation of a biopolymer with fixed structure, especially membrane-bound polymers. Several studies of typical synthetic polypeptides and membrane proteins using high-resolution $^{13}$C NMR in the solid state are given here.

### Synthetic polypeptides

Synthetic polypeptides consist of repeating sequences of certain amino acids and their structures are not as complicated as those of proteins. For this reason, synthetic polypeptides are sometimes used as model compounds for proteins. Their preferred conformations are classified as α helix, β sheet, ω helix and so on. High-resolution solid-state $^{13}$C NMR spectroscopy has proved to be a very powerful tool for determining the structure of polypeptides in the crystalline state. For example, $^{13}$C CP MAS NMR spectra of solid poly(L-alanine) ($[Ala]_n$) show the Cα, Cβ and C=O carbon signals to be well resolved between the α helix and β sheet forms. The chemical shifts of the Cα and C=O carbons of the α helix are displaced significantly to high frequency by 4.2 and 4.6 ppm, respectively, relative to those of the β sheet form, while the shift of the Cβ carbon of the α helix is displaced to low frequency by about 5 ppm with respect to that of the β sheet. For this reason, the value of the $^{13}$C shift can be used to describe the local conformation. In addition, the $^{13}$C shifts of randomly coiled $[Ala]_n$ in trifluoroacetic acid solution have values between those of the α helix and β sheet forms. The absolute $^{13}$C shifts of the Cα and Cβ carbons are affected by the chemical structure of the individual amino acid residues and can be used effectively for conformational studies of particular amino acid residues in polypeptides and proteins. On the other hand, the C=O shifts do not seem to be affected by residue structure and can be used for diagnosing the main chain conformation.

Recently, the relation between the amide proton shift and the conformation of synthetic polypeptides has been studied in the solid state using $^{1}$H combined rotation and multipulse spectroscopy (CRAMPS). It has been determined that the amide chemical shifts of a synthetic polypeptide are 8.0–8.1 ppm for the α helix and 8.6–9.1 ppm for the β sheet. It was shown that high-resolution solid-state $^{1}$H NMR such as CRAMPS is also a useful tool for studying the molecular structure of polypeptides in addition to high-resolution solid-state $^{13}$C NMR.

Poly(β-benzyl-L-aspartate) (PBLA) in the solid state adopts various conformations such as a right-handed $\alpha(\alpha_R)$ helix, a left-handed $\alpha(\alpha_L)$ helix, a left-handed $\omega(\omega_L)$ helix and a β sheet form, depending on the temperature. VT $^{13}$C CP MAS NMR spectra of PBLA are shown in **Figure 8**. From this, it is clear that the conformational changes such as $\alpha_R$ helix → $\omega_L$ helix → β sheet occur with increasing temperature.

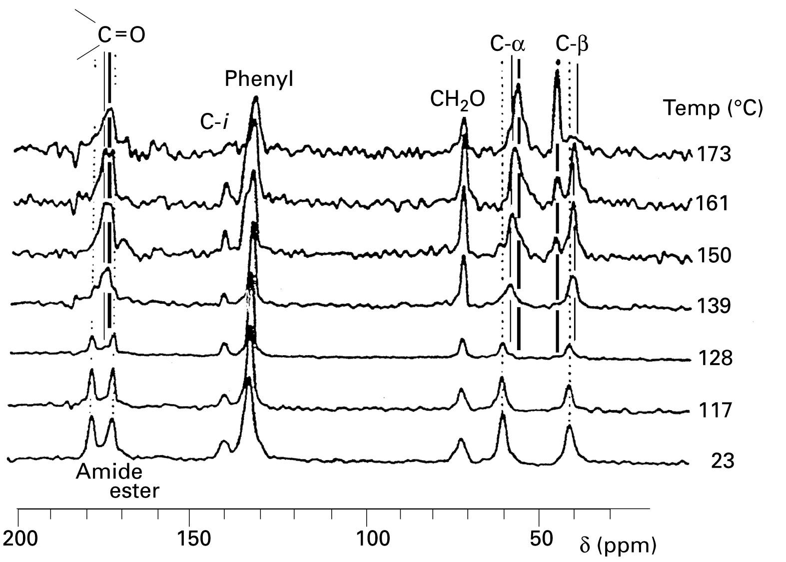



**Figure 8** VT $^{13}C$ CP MAS NMR spectrum with spinning sideband suppression of PBLA in the solid state at various temperatures: dashed line, $\alpha_R$ helix form, light solid line, $\omega_L$ helix form and bold solid line, β sheet form. Reproduced with permission of American Chemical Society from Akieda T, Mimura H, Kuroki S, Kurosu H and Ando I (1992) *Macromolecules* **25**: 5794.

Poly(γ-*n*-octadecyl-L-glutamate) (POLG) forms a crystalline phase composed of paraffin-like crystallites together with the α helical main chain, packing into a characteristic layer structure. The polymer forms thermotropic cholesteric liquid crystals by the melting of the side-chain crystallites. The conformational changes of POLG have been studied extensively by high-resolution solid-state NMR. **Figure 9** shows the $^{13}C$ CP MAS NMR spectrum of POLG at room temperature together with the assignment of peaks. At room temperature, $^{13}C$ shift values of the C=O and Cα carbons are 176.0 and 57.6 ppm, respectively. These values show that the main chain takes an α helix conformation. On the other hand, the interior $CH_2$ signal splits into two peaks. From reference data of *n*-alkanes and polyethylene, it is seen that the peak at 33.4 ppm is the carbon signal in the all-*trans* zig-zag conformation in a crystalline state and the peak at 30.6 ppm is the carbon signal in the non-crystalline phase or the liquid phase.

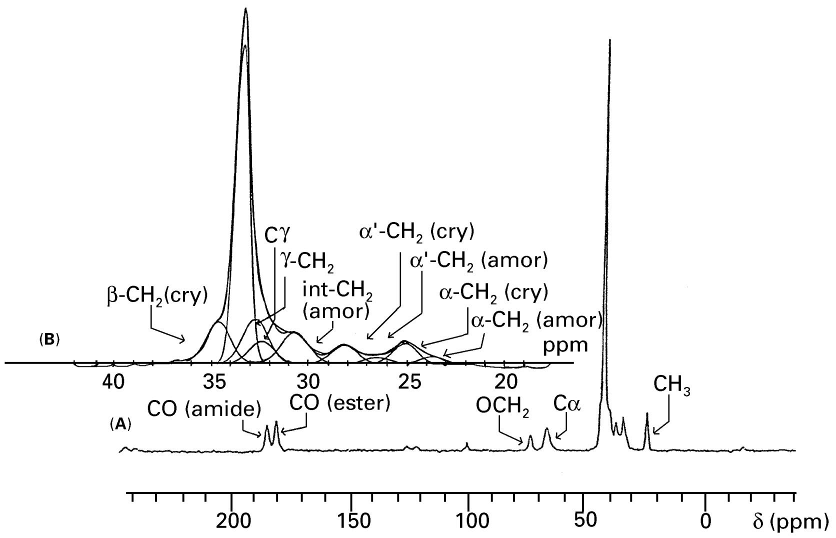



**Figure 9** (A) $^{13}C$ CP MAS NMR spectrum of POLG in the solid state at room temperature. (B) The methylene region is expanded and deconvoluted by computer-fitting with a Gaussian function. (cry and amor indicate crystalline and non-crystalline state, respectively.) Reproduced with permission of Elsevier Science from Katoh E, Kurosu H and Ando I (1994) *Journal of Molecular Structure* **318**: 123.

**Figure 10** shows the VT $^{13}C$ CP MAS NMR spectra of POLG over the temperature range −40 to 230°C. The *n*-alkyl peaks change noticeably as the temperature is increased. The peak of the carbon signal in the non-crystalline state disappears above 50°C and the intensity of the interior $CH_2$ carbon in the non-crystalline state increases noticeably. This is due to the melting of side-chain crystallites. On the other hand, $^{13}C$ shift values of the C=O and Cα carbons are independent of temperature. It is shown that the main chain takes on a right-handed α helix conformation within the temperature range −40 to 230°C.

For samples which show a variation in molecular motion with temperature it is useful to measure the $T_1$ value using both the inversion–recovery method with the PST pulse sequence and Torchia's pulse sequence, depending on the type and rate of molecular motions. The molecular motions of POLG have been investigated within the temperature range −40 to 230°C, using both methods. **Figure 11** shows the temperature dependence of $T_1$ for some typical carbons. According to the BPP theory, $T_1$ decreases, passes through a minimum (at a correlation time: $\tau_c$) and increases again when molecular motion increases, i.e., temperature is increased. The $^{13}C$ $T_1$ values of POLG are very short in spite of the fact that, as seen from the chemical shifts, the *n*-alkyl side-chains of POLG are in the crystalline state in the temperature range −40 to 40°C. If the *n*-alkyl side-chains are in the crystalline state, the $^{13}C$ $T_1$ values must be very long by analogy with the *n*-alkane data. However, the $^{13}C$ $T_1$ values for individual carbons of the *n*-alkyl side-chains are very short and very close to the $T_1$ values of *n*-alkanes in the rotator phase; in addition, the $^{13}C$ $T_1$ values for individual carbons of *n*-alkyl side-chains are almost the same as for *n*-alkanes in the rotator phase. From these results, it can be said that the *n*-alkyl side-chains of POLG are in a phase similar to the rotator phase, in which they undergo fast rotation such as libration, along the *trans* zigzag chain axis. The effect of the transitions at 45°, 60° and between 90° and 110°C on the individual carbon of the side-chains has also been investigated. The transition at 45°C comes from the melting of the side-chains, while other changes come from changes of the liquid crystalline phase, which has also been proved by other means.

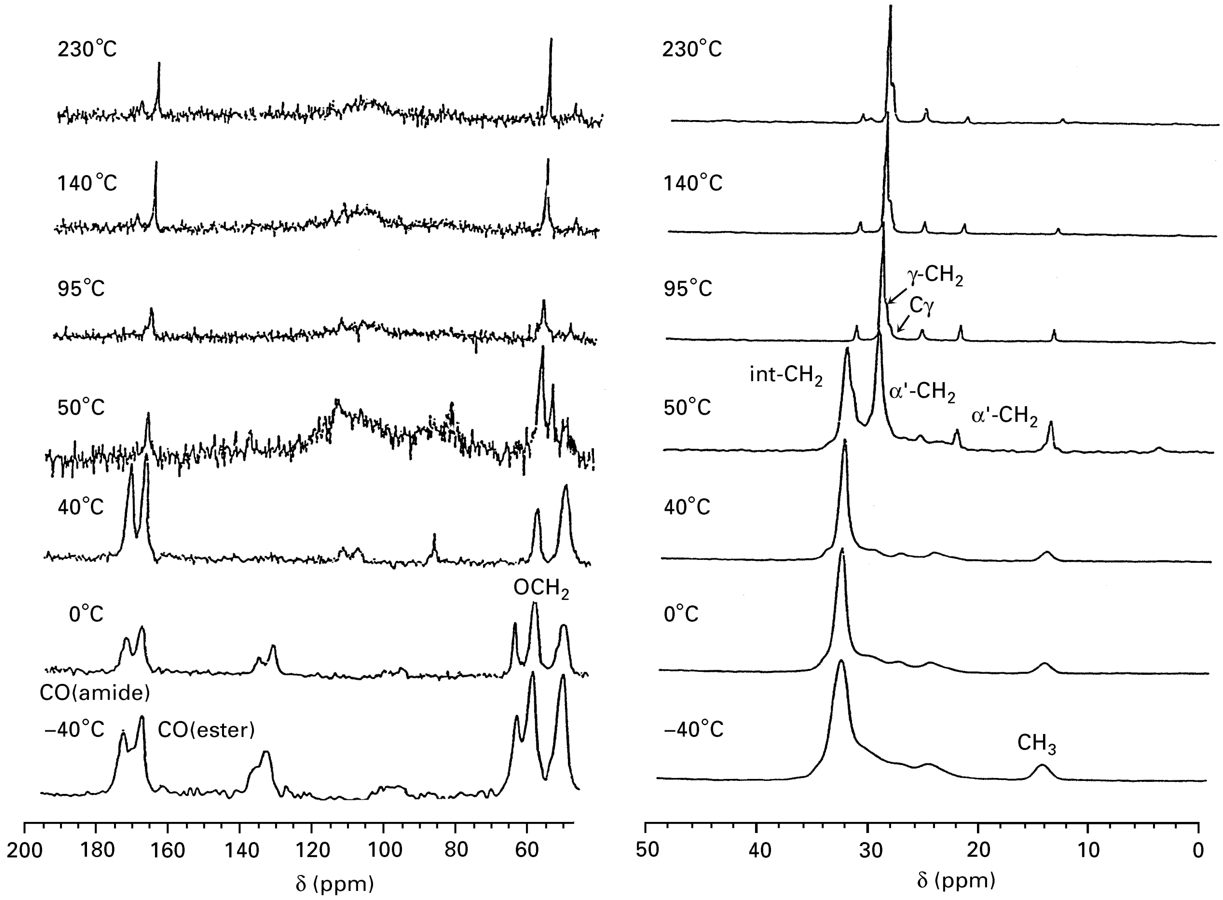



**Figure 10** VT $^{13}C$ CP MAS NMR spectra of POLG as a function of temperature. Reproduced with permission of Elsevier Science from Katoh E, Kurosu H and Ando I (1994) *Journal of Molecular Structure* **318**: 123.

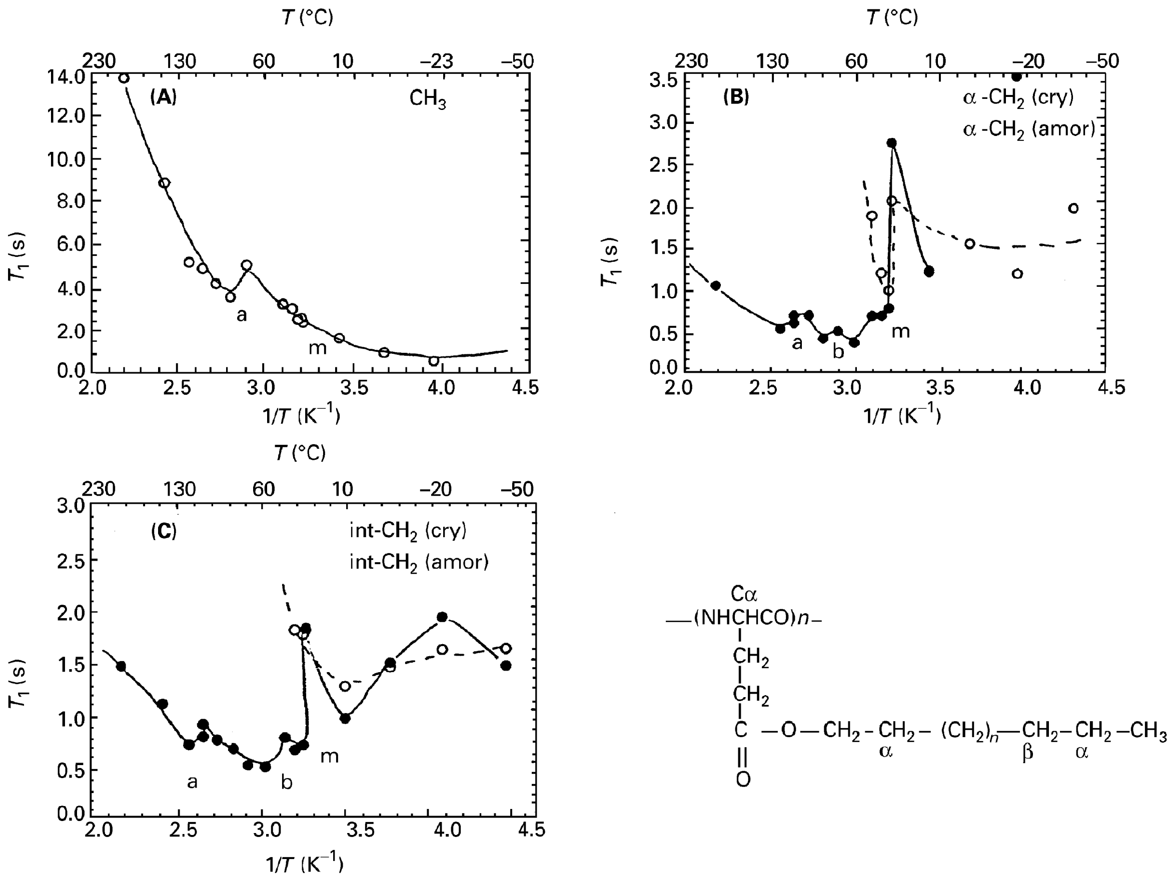



**Figure 11** Temperature dependence of $^{13}C$ $T_1$ for individual carbons of POLG. (A) $CH_3$, (B) α-$CH_2$ and (C) int-$CH_2$ (a and b indicate transition temperatures. m indicates melting point.) For $T_1$ measurement in the temperature range −40 to 50°C Torchia's pulse sequence was used, and in the temperature range 50 to 230°C, the inversion–recovery pulse sequence was used. Reproduced with permission of Elsevier Science from Katoh E, Kurosu H, Kuroki S and Ando I (1994) *Journal of Molecular Structure* **326**: 145.

It has been demonstrated that the introduction of fluorine atoms into polymers induces new physical properties by changes in intra- and intermolecular interactions. Poly(γ-glutamate)s with *n*-fluoroalkyl side-chains have been studied using high-resolution solid-state NMR. From the experimental results, it was found that the conformation of the main chain depends on the carbon number of the *n*-fluoroalkyl side-chains. The main chain of poly(γ-alkyl-L-glutamate) with short *n*-fluoroalkyl side-chains takes the helix form and the poly(γ-alkyl-L-glutamate) with long *n*-fluoroalkyl side-chains takes the β sheet form. These conformational behaviours in the solid state are different from that of poly(γ-alkyl-L-glutamate) without *n*-fluoroalkyl side-chains.

## Fibrous proteins

Since fibrous proteins generally have periodical amino acid sequences and higher-order structure, the clarification of their fine structure in the solid state becomes very important not only when discussing the physical and chemical properties, but when obtaining information about the molecular design of synthetic polypeptides. The conformation-dependent $^{13}C$ CP MAS NMR chemical shifts are particularly useful for the determination of the conformational features of fibrous proteins such as silk fibroin, collagen and wool keratin. For example, the conformational transition of *S*-carboxymethyl keratin that has low-sulfur fractions (SCMK low-sulfur) has been estimated. For SCMK low-sulfur heated at 200°C for 3 hr under vacuum, the $^{13}C$ MAS NMR spectrum shows that each signal becomes broader than those of other treated specimens. This indicates the existence of various conformations and/or different microenvironments in the heated SCMK low-sulfur. Thus, it can be said that the random coil form appears by heating. On the other hand, from the X-ray diffraction, the α helix form completely vanishes in SCMK low-sulfur under the same conditions. The difference between the results from X-ray diffraction and NMR spectroscopy suggests that only the packing of the ordered structure (α helix form) in SCMK

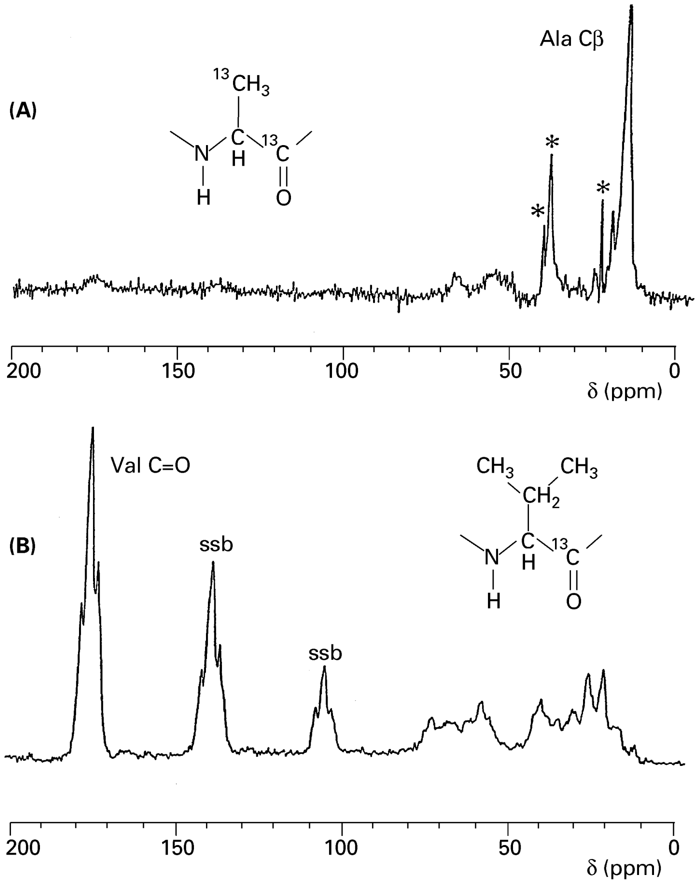



**Figure 12** $^{13}C$ CP MAS NMR spectra of (A) [3-$^{13}C$] Ala bR and (B) [1-$^{13}C$]Val bR in the purple membranes. (*) signals of $^{13}C$-labelled lipids. Reproduced with permission of Elsevier Science from Tuzi S, Naito A and Saito H (1994) *Biochemistry* **33**: 15 046.

low-sulfur is disrupted by heating, while the secondary structure is retained.

## Membrane proteins

Membranes provide a vital interface between a biological organism and its environment. Furthermore, they provide a means for the formation of compartments within a cell. In every sense, they are vital to living systems. Indeed, since membranes and membrane proteins cannot easily be crystallized, NMR spectroscopy is a particularly powerful method for probing their structure and function. [1-$^{13}C$]Ala, [3-$^{13}C$]Ala and [1-$^{13}C$]Val selected labelling of bacteriorhodopsin (bR) is a very convenient probe for conformation and dynamics studies. **Figure 12** shows the $^{13}C$ CP MAS NMR spectra of [3-$^{13}C$]Ala and [1-$^{13}C$]Val bR. The peaks of [3-$^{13}C$]Ala bR have been assigned to the $\alpha_{II}$ helix form at 16.3 ppm (60%), $\alpha_I$ helix form at 14.9 ppm (20%) and the loops at 17.2 ppm (20%), taking a variety of turn structures. Further, from the VT CP MAS $^{13}C$ NMR spectrum of [3-$^{13}C$]Ala labelled bR, well-resolved signals were observed at both ambient temperature and −20°C but were broadened considerably at temperatures below −40°C. This situation was interpreted as showing that interconversion between several slightly different conformations is taking place. It was found that the exchange process was strongly influenced by the manner of organization of the lipid bilayers, and depended upon the presence or absence of cations responsible for electric shielding of the negative charge at the polar head groups. The secondary structures of [3-$^{13}C$]Ala-labelled bR were not always identical at temperatures between ambient and lower temperatures, since the $^{13}C$ shifts and relative peak intensities from purple membrane preparations containing these salts changed with temperature in the range −110 to 23°C. In particular, some residues involving Ala residues at the $\alpha_{II}$ helix and loop region were converted at temperatures below −60°C into a conformation of bR involving a $\alpha_I$ helix.

The rotational resonance (RR) method is one of several solid-state NMR approaches that have been developed recently for determining weak dipolar

couplings under conditions of MAS. To measure heteronuclear dipolar interactions, such as between $^{13}C$ and $^{15}N$, rotational echo double resonance (REDOR) and transferred-echo double resonance (TEDOR) NMR have been developed. The measurement of weak dipolar couplings in MAS experiments using any of these methods greatly enhances solid state NMR studies of biological systems since the dipolar couplings are directly related to internuclear distances which in turn provide constraints for the determination of molecular structure.

The interatomic distances in crystalline specimen of $^{13}C$, $^{15}N$ doubly labelled simple peptides, N-acetyl-Pro-Gly-Phe, evaluated from REDOR data were compared with those from X-ray diffraction studies and found to justify such a novel approach. The REDOR-derived conformation of this peptide was β turn type I, consistent with the X-ray diffraction study. The maximum deviations of the distances determined by NMR and X-ray diffraction is 0.08 Å despite the complete neglect of the dipolar interactions with the labelled nuclei of neighbouring molecules and natural abundance nuclei. The precision and accuracy given by $^{13}C$ REDOR experiments are of the order of 0.05 Å. Distinction between the two types of β turn forms, including the β turn type II, found in the monoclinic crystal of this peptide (whose interatomic distances are different by about 0.57 Å) is made possible only by very accurate REDOR measurement.

More recently, it is the ability of solid-state NMR to give completely resolved spectra of immobile proteins that has enabled the structures of larger membrane proteins to be determined in the definitive environment of lipid bilayers. Further, high-resolution solid-state $^{13}C$ NMR spectra of the [3-$^{13}C$]Ala-labelled fragment of bR incorporated into dimyristoylphosphatidylcholine bilayers have been recorded. This approach has proved useful for the assignment of peaks as well as dynamic features of the transmembrane helixes in lipid bilayers.

## Coals and coal products

The complete molecular structure of coal is one of the unsolved secrets of nature. Of the various analytical methods available, NMR spectroscopy has proved to be of special importance in coal research. Solid-state NMR techniques allow the structure of coal and coal products to be investigated in a direct and nondestructive way.

The accuracy of aromaticity measurements on coals by CP MAS $^{13}C$ NMR, together with additional problems posed by high-field measurements and spectral editing, and some emerging techniques, have been discussed. It is suggested that a combination of low field, single-pulse excitation with long relaxation delays and the use of a suitable reagent to quench paramagnetic centres is the most satisfactory, albeit time consuming, recipe for obtaining reasonably reliable results on unknown samples.

One particular coal, Illinois No. 6, has been measured by 2D $^{13}C$–$^{1}H$ chemical shift correlation spectroscopy. A variant of dipolar shift correlation spectroscopy has been described and this is very convenient for determining the relative amounts of methylene, methine and methyl or non-protonated carbons in a complex mixture such as coal.

## Foods

It has long been recognized that NMR can be of use to the food scientist and food processor. However, it is only within the last 15 years that there has been a consistent and widespread growth of the use of NMR. One reason for this is the development of new solid-state NMR methods.

Work on starch provides an excellent example of the information obtainable from solid-state NMR. Native starch granules have crystalline and non-crystalline regions. The crystalline regions are made up of an ordered arrangement of polymer chains in the form of double helices, whereas the amorphous regions contain single chains. The two polymorphs of starch, A (from cereals) and B (from tubers), have different packing arrangements of the double helices. The CP MAS NMR spectra of various native starches clearly demonstrate distinguishable crystalline and amorphous regions.

$^{13}C$ CP MAS and single-pulse (SP) MAS spectra have been obtained for galactomannans and glucomannans as powders, hydrates or gels. Comparison of CP and SP experiments on galactomannan gels did not reveal any chemical shift changes, suggesting that the conformations of mobile and rigid segments are similar. CP MAS experiments have been performed on purified carrageenans and agaroses and on the seaweeds which are the source of these materials. Although the lines were broad, characterization of the intact algae was possible and various substituents were identified.

## List of symbols

$T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $T_{1\rho}$ = spin–lattice relaxation time in the rotating frame; $T_{DD}$ = dipolar-dephasing relaxation time.

*See also:* **Food Science, Applications of NMR Spectroscopy; High Resolution Solid State NMR, $^1$H, $^{19}$F; Membranes Studied By NMR Spectroscopy; Neutron Diffraction, Theory; NMR of Solids; NMR Pulse Sequences; NMR Spectrometers; Powder X-Ray Diffraction, Applications; Proteins Studied Using NMR Spectroscopy; Solid State NMR, Methods; Solid State NMR, Rotational Resonance.**

## Further reading


Ando I and Asakura T (1998) *Solid State NMR of Polymers*. Amsterdam: Elsevier Science.

Ando I, Yamanobe T and Asakura T (1990) *Progress in NMR Spectroscopy* **22**: 210.

Belton PS, Colquhoun IJ and Hills BP (1993) *Annual Reports in NMR Spectroscopy* **26**: 1.

Dec SF, Wind RA and Maciel GE (1987) *Macromolecules* **20**: 349.

Evans JS (ed) (1995) *Biomolecular NMR Spectroscopy*. New York: Oxford University Press.

Katoh E, Sugisawa H, Oshima A, Tabata Y, Seguchi T and Yamazaki T (1998) *Radiation Physics and Chemistry* (in the press).

Katoh E, Sugimoto H, Kita Y and Ando I (1995) *Journal of Molecular Structure* **355**: 21.

Komoroski RA (ed) (1986) *High Resolution NMR of Synthetic Polymers in Bulks*. Florida: VCH.

Kurosu H, Ando S, Yoshimizu H and Ando I (1994) *Annual Reports in NMR Spectroscopy* **28**: 189.

Harris RK and Jackson P (1991) *Chemical Reviews* **91**: 1427.

Holstein P, Scheler U and Harris RK (1997) *Magnetic Resonance in Chemistry* **35**: 647.

Meiler W and Meusinger R (1991) *Annual Reports in NMR Spectroscopy* **23**: 375.

Saito H and Ando I (1989) *Annual Reports in NMR Spectroscopy* **21**: 210.

Schmidt-Rohr K and Spiess HW (1994) *Multidimensional Solid State NMR and Polymers*. London: Academic Press.


# High Resolution Solid State NMR, $^1$H, $^{19}$F

**Anne S Ulrich**, Friedrich-Schiller-University of Jena, Germany

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| Applications |

## Introduction

Hydrogen ($^1$H) and fluorine ($^{19}$F) possess the highest magnetic moments amongst all isotopes in the periodic table (besides the radioactive tritium, $^3$H), which gives them the greatest NMR sensitivity and strongest dipolar couplings. Their spin of $I = \frac{1}{2}$ makes them well suited for structural investigations, and $^1$H is indeed the most widely used nucleus in solution-state NMR. In the solid state, however, the very advantage of strong dipolar interactions, conveying long-range distance information, turns to a disadvantage when rapid spin diffusion occurs between nuclei in an extensively coupled network. The high abundance of $^1$H in organic compounds thus leads to substantial line-broadening, and the same happens to $^{19}$F in a perfluorinated or protonated environment. Several different line-narrowing approaches have been developed to achieve high resolution, namely magic-angle spinning (MAS), multiple-pulse sequences, a combination of both (CRAMPS), or the use of uniaxially oriented samples. Some important experiments will be illustrated here that are applicable to both $^1$H and $^{19}$F for determining chemical shift values, segmental mobilities, and through-space correlations. Depending on the material, both nuclei have been fruitfully applied in the analysis of inorganic crystals organic solids, glasses, polymers, liquid crystals and biological systems. The extensive literature on broadline NMR studies of these systems will not be considered here.

The high-frequency channel of a standard solid-state NMR spectrometer tends to be optimized for $^1$H, especially when $^1$H-decoupling is required for observing $^{13}$C and other low-band nuclei. For $^{19}$F NMR experiments it is necessary to re-tune the transmitters, amplifiers and probes, and some extra hardware may need to be implemented when $^{19}$F and $^1$H are to be used simultaneously. The areas of $^1$H and $^{19}$F NMR have thus been pursued by different groups and with different objectives. If any general trends are to be identified, solid-state $^1$H NMR appears to be moving towards higher resolution by working at high magnetic field strengths (currently up to 20 T), by applying high MAS speeds (beyond 35 kHz), by making use of multidimensional experiments (HETCOR, multiquantum), and with an awareness of the physical limitations to the inherent spectral

line widths. Concerning the solid state $^{19}$F NMR analysis of (per-)fluorinated materials, an observation of $^{19}$F is often preferred over the otherwise very popular $^{13}$C, because in these systems the $^{13}$C resonances are strongly broadened and not so readily decoupled from the $^{19}$F spins. With yet another motivation to study dilute $^{19}$F-labels in a protonated environment, namely in organic and biological systems, many $^{19}$F NMR experiments can be realized under efficient $^1$H-decoupling that are analogous to $^{13}$C NMR.

## Properties of the nuclei

The nuclear spin properties of $^1$H and $^{19}$F are summarized and compared in **Table 1**. The high gyromagnetic ratios $\gamma$ provide both isotopes with a much higher sensitivity than $^2$H (1% of $^1$H), $^{13}$C (1.6%) or $^{15}$N (0.1%), to name a few, and these factors are further enhanced by the respective natural abundances. The homo- and heteronuclear dipolar Hamiltonian also depends on $\gamma$, and it is given for a dilute pair of spins ($I_1$,$I_2$) by

$$\mathscr{H}_{\mathrm{D}} = -\left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_1\gamma_2\hbar^2}{r_{12}^3}\left(\frac{3\cos^2\theta - 1}{2}\right)\times(3I_{1z}I_{2z} - \mathbf{I}_1\mathbf{I}_2) \quad [1]$$

The coupling within a pair (or within a small group) of nuclei is well defined by their relative positions (distance $r$ and angle $\theta$), which makes their interaction inhomogeneous. The line shapes or dipolar splittings of dilute labels can thus be directly analysed to measure local structural parameters or motional averaging. In an extended network of abundant spins, on the other hand, the dipolar interactions are more complex than in Equation [1] and they become highly nonlinear. This leads to homogeneous line-broadening by spin diffusion, which is much harder to suppress experimentally.

The main difference between $^1$H and $^{19}$F lies in the ranges of their isotropic and anisotropic chemical shifts (anisotropic = dependent on the orientation of the molecule in space, i.e. on its alignment with respect to the magnetic field direction). Unlike $^1$H, the $^{19}$F atom is rather polarizable, which makes the chemical shift very sensitive to local charge densities and hydrogen-bonding, and which also leads to significant solvent and temperature effects. **Table 2** provides a comprehensive list of isotropic $^{19}$F chemical shifts for inorganic and organic solid compounds, as compiled by Miller. Referencing in the solid state usually relies on comparison with a separate sample of TMS for $^1$H, or $CFCl_3$ for $^{19}$F, rather than using internal or external chemical shift standards. An extensive list of chemical shielding anisotropies for $^1$H and $^{19}$F has been published by Duncan.

**Table 1** Nuclear spin properties of $^1$H and $^{19}$F

| *Property* | *$^1$H* | *$^{19}$F* |
|---|---|---|
| Natural abundance | 99.98% | 100% |
| Gyromagnetic ratio $\gamma$ ($10^7$ rad s$^{-1}$ T$^{-1}$) | 26.7522 | 25.1815 |
| Relative sensitivity (at constant field) | 1.00 | 0.83 |
| Resonance frequency at 11.74 T | 500 MHz | 470.4 MHz |
| Chemical shift range (total) | −20–30 ppm | −400–800 ppm |
| (organic compounds) | 0–12 ppm | −200–0 ppm |
| Shielding anisotropy (CSA) | 30 ppm | 200 ppm |
| Dominant relaxation mechanism | Dipolar | Dipolar, CSA |

## Methods for studying abundant and dilute spins

In the following overview of methods and applications in high-resolution solid-state $^1$H and $^{19}$F NMR it is instructive to consider different kinds of situations that the experiment may need to address. Depending on the abundance of spins, different line broadening mechanisms will be encountered, and different techniques are required to achieve high resolution. After a discussion of four such categories, some representative applications to various materials will be illustrated.

### $^1$H NMR of abundant spins

In organic compounds, in polymers and in biological materials the $^1$H-density is high and spin diffusion leads to rapid dephasing of the magnetization. The resulting homogenous line broadening can only be fully suppressed by MAS when the spinning frequency $\nu_r$ is much higher than the static line width (which is typically around 50 kHz in rigid solids). Since modern commercial NMR probes can reach speeds of up to 35 kHz with small rotors, this is not usually sufficient to achieve complete homonuclear dipolar decoupling in rigid solids. Only samples with a high intrinsic mobility, such as liquid crystals and rubber-like materials, will give narrow lines under high-speed MAS, with well resolved isotropic chemical shifts. Note that the chemical shielding anisotropy does not usually represent a problem for $^1$H NMR, because it represents an inhomogenous interaction. It is efficiently narrowed by MAS already at low spinning speeds, resulting in a narrow isotropic line plus a set of side-bands that are separated by multiples of $\nu_r$.

**Table 2** Chemical shifts of fluorine in inorganic and organic solie samples. Reproduced with permission of Elsevier Science from Miller JM (1996) *Progress in NMR Spectroscopy*, Vol 28, pp 255–281. Amsterdam: Elsevier.

| *Sample* | *Chemical group* | *Chemical shift (relative to $CFCl_3$ in ppm)* |
|---|---|---|
| *Inorganic samples* | | |
| LiF | $F^-$ | –204 |
| LiF | $F^-$ | –130[a] |
| NaF | $F^-$ | –221 |
| NaF | $F^-$ | –221 |
| NaF | $F^-$ | –126[a] |
| NaF | $F^-$ | –224 |
| KF | $F^-$ | –130 |
| KF | $F^-$ | –130.2 |
| KF | $F^-$ | –123 |
| $KF \cdot 2H_2O$ | $F^-$ | –133 |
| KF–$CaF_2$ | $F^-$ | –121 |
| KF–alumina | $F^-$ | –159 |
| KF–alumina | $F^-$ | –115 |
| | $AlF_6^{3-}$ | –156 |
| KF–silica | $SiF_6^{2-}$ | –129 |
| | $SiF_4^-$ | –116 |
| | $SiF_4$/HF | –161 |
| $KAlF_4$ | $AlF_4^-$ | –155 |
| $K_3AlF_6$ | $AlF_6^{3-}$ | –155 |
| $K_2SiF_6$ | $SiF_6^{2-}$ | –92 |
| RbF | $F^-$ | –90 |
| RbF | $F^-$ | –88 |
| $RbF \cdot H_2O$ | $F^-$ | –113 |
| RbF–alumina | $F^-$ | –109 |
| | $AlF_6^{3-}$ | –139 |
| CsF | $F^-$ | –8 |
| CsF | $F^-$ | –79[a] |
| CsF–$CaF_2$ | $F^-$ | –79 |
| $CsF \cdot 2H_2O$ | $F^-$ | –97 |
| CsF–alumina | $F^-$ | –116 |
| CsF–alumina | $F^-$ | –88[a] |
| | $AlF_6^{3-}$ | –111[a] |
| $Cs_2SiF_6$ | $SiF_6^{2-}$ | –92 |
| $MgF_2$ | $F^-$ | –194.8 |
| $CaF_2$ | $F^-$ | –104.8 |
| $CaF_2$ | $F^-$ | –107.7 |
| $CaF_2$ | $F^-$ | –106.4, –107.0 |
| $SrF_2$ | $F^-$ | –84.1 |
| $CdF_2$ | $F^-$ | –192.1 |
| $Hg_2F_2$ | $F^-$ | –95.8 |
| $HgF_2$ | $F^-$ | –196.4 |
| $SnF_2$ | $F^-$ | –110.4 |
| $SnF_4$ | $F^-$ | –146.9 |
| $(C_4H_9)_3SnF$ | $F^-$ | –145 |
| $AlF_3$ | $F^-$ | –174 |
| $\alpha$-$PbF_2$ | F(1) | –20.5 |
| | F(2) | –57.7 |
| | Mobile $F^-$ | –39.0 |
| $Na_3AlF_6$ | $F^-$ | –189 |

**Table 2** *Continued*

| *Sample* | *Chemical group* | *Chemical shift (relative to $CFCl_3$ in ppm)* |
|---|---|---|
| $K_3AlF_6$ | $F^-$ | –190 |
| $Al_2(F,OH)_2SiO_4$ (topaz) | $F^-$ | –140 |
| $Na_2SiF_6$ | $F^-$ | –152 |
| $Na_2SiF_6$ | $F^-$ | –135 |
| $Na_2SiF_6$ | $F^-$ | –151[a] |
| $K_2SiF_6$ | $F^-$ | –135 |
| $NH_4F-$silica | $SiF_6^{2-}$ | –128 |
| $NH_4F-$alumina | $F^-$ | –178 |
| | Al–F | –166 |
| $Et_4NF-$alumina | $F^-$ | –123 |
| $Et_4NF{\bullet}2H_2O$ | $F^-$ | –130 |
| $Et_4NF-$cytosine | $F^-$ | –148 |
| $Bu_4NF-$alumina | $F^-$ | –106 |
| $Bu_4NF{\bullet}3H_2O$ | $F^-$ | –109 |
| $Bu_4NF-$4-cyanophenol | $F^-$ | –144 |
| $Bu_4NF-$uracil | $F^-$ | –149 |
| $Bu_4NF-HOCPh_3$ | $F^-$ | –147 |
| HAP/$SnF_2$ dentifrice | FHAP | –103 to –99 |
| | $F^-$ | –119 |
| FAP | $Ca_5(PO_4)_3F$ | –98.9 |
| FAP | $Ca_5(PO_4)_3F$ | –101.0 |
| FAP | $Ca_5(PO_4)_3F$ | –99.0 |
| Sb(III) subs., FAP | $1F^-/Sb^{3+}$ | –97.5 |
| | $2F^-/Sb^{3+}$ | –95.4 |
| | $1F^-/Sb^{3+}$ | –89.9 |
| Synthetic FAP | $Ca_5(PO_4)_3F$ | –98.9 |
| Heated FAP | $Ca_5(PO_4)_3F$ | –100 |
| Mineral | $Ca_5(PO_4)_3F$ | –100.4 |
| Synthetic FHAP | $Ca_5(PO_4)_3F_x(OH)_{1-x}$ | –102.3 |
| FHAP | $Ca_5(PO_4)_3F_x(OH)_{1-x}$ | –99 to –103 |
| | $F^-$(NSA) | –119 |
| $CaF_2$–FAP | $F^-(CaF_2)$ | –105 |
| | FAP | –99 |
| Gallophosphate cloverite | $F^-$ in DR cages 4Ga | –67.7 |
| | 3 Ga, 2Al | –72.8 |
| | 2 Ga, 2Al | –81.0 |
| | 1 Ga, 1Al | –89.2 |
| | 4Al | –95.5 |
| | HO–Al–F | –145.6 |
| | F in sodalite type cages | –178.2 |
| Flurophosphate glass | $F^-$ coord. to Al/Ba | –99 to –90 |
| Fluorinated $AlPO_4$ | Al–F | –142 |
| | $F^-\cdot\cdot$ CHAH$^+$ | –123 |
| Fluorinated $AlPO_4$ | $F^-$ ion pair | –119 to –123 |
| | Al$-$F | –135 to –139[a] |
| Hexamethylenediamine | $F^-$ occluded in double 4-membered rings | –67.8 |
| subs. $Ga_2O_3-P_2O_5$ | Ga...$F^-$... Ga | –92.2, –113 |
| Fluorinated $SiO_2$ | $F^-$ | –63 to –75 |
| Templated with | Si$-$F | –155[a] |
| $R_4NF$ | | |
| | $F^-$ Interaction with Ge/Ti/Al/Ga | –140 |
| | $F^-...NH_4^+$ | –128 |

**Table 2** *Continued*

| *Sample* | *Chemical group* | *Chemical shift (relative to $CFCl_3$ in ppm)* |
|---|---|---|
| Octadecasil | $F^-$ in 4-membered ring cages | –38 |
| Fluoride subs, silicate | $F^-$ | –64 |
| F subs. LEV type zeolite | $F^-$ in framework | –82 |
| | Si–F | –120 |
| | Al–F | –150, –175 |
| | $SiF_6^{2-}$ | –129 |
| Zn,Ti substituted layered silicate | $F^-$ | –146.6, –169.1, –186.0 |
| Ga layered silicate | $F^-$ in dioctahedral sheets | –108 |
| F–hectorite | $F^-$ in trioctahedral Mg/Li-containing sheets | –182.8[a] |
| F–montmorillonite | $F^-$ in trioctahedral Mg-containing sheets | –176.2[a] |
| Barasym SMM 100 | $F^-$ in dioctahderal Mg-containing sheets | –152.0[a] |
| Synthetic dioctahedral | $F^-$ in dioctahedral Al-containing sheets | –131.9[a] |
| 2:1 layered aluminosilicate | $F^-$ | –133.2[a] |
| Tremolite | $F^-$ | –171.7 |
| Fluorscandium pargasite | $F^-$ | –169.6 |
| NaF–zeolite A(LTA) | $F^-$ in sodalite structure | –179 |
| KF–montmorillonite K10 | $SiF_6^{2-}$ | –129 |
| $NH_4$ –montmorillonite K10 | $SiF_6^{2-}$ | –136 |
| RbF–montmorillonite K10 | $SiF_6^{2-}$ | –123 |
| | $AlF_4^-$ | –141 |
| CsF–montmorillonite K10 | $AlF_6^{3-}$ | –113 |
| RbF–silica | $SiF_6^{2-}$ | –122 |
| Cd/Zn/$CuF_2^-$ montmorillonite K10 | $F^-$ | –151 |
| $CuF_2$–silica | $SiF_6^{2-}$ | –149 |
| $ZnF_2$–silica | $F^-/F^-H_3O^+$ | –124 |
| | $SiF_6^{2-}$ | –149 |
| $CdF_2$–silica | $F^-/F^-H_3O^+$ | –124 |
| $NaF_4$–silica | $F^-$ | –128 |
| $NaBF_4$–silica | $F^-$ | –128 |
| [(PS)$H^+$] [$OTeF_5^-$] | $F^-$ triclinic, axial | –12.8 |
| (PS)$H^+$ = protonated | $F^-$ triclinic, equatorial | –30.7 |
| 1.8-bis(dimethylamino)- | $F^-$ orthorhombic, axial | –17.2 |
| naphthalene | $F^-$ orthorhombic, equatorial | –33.1 |
| [N(*n*-$Bu_4^+$][$OTeF_5^-$)] | $F^-$ axial | –14.3 |
| | $F^-$ equatorial | –33.6 |
| [N(*n*-Bu)$_4^+$] | $F^-$ axial | –27.7, –33.1 |
| [H($OTeF_5$)$_2^-$] | $F^-$ equatorial | –42.4, –43.3 |
| *Organic samples* | | |
| $(CF_2CFCl)_n$ | CFCl | –130.2 |
| | $CF_2$ | –106.2, –100.2 |
| $CH_2CF_2$/$CF_3CFCF_2$/$CF_2CFCl$ co- and ter-polymers | $–CH_2CF_2CF(CF_3)CF_2CH_2–$ | –71.9, –72.1, –72.7 –72.9 |
| | $–CH_2CF_2CF(CF_3)CH_2CF_2–$ | –76.4, –76.6, –77.6 |
| | | –77.9 |
| | $–CH_2CF_2CH_2–$ | –82.1 |
| | $–CF_2CH_2CF_2CH_2CF_2–$ | –91.4, –91.8, –90.8 |
| | | –91.5, –89.7 |
| | $–CFClCH_2CF_2CH_2CF_2–$ | –90.9 |
| | $–CF_2CH_2CF_2CF(CF_3)CF_2–$ | –104.4, –104.7 |
| | $–CF_2CH_2CF_2CF_2CFCl–$ | –108.8, –109.1 |
| | $–CF(CF)_3CH_2CF_2CF_2CF(CF_3)–$ | –112.8 |
| | $–CF_2CH_2CF_2CF_2CF(CF_3)–$ | –111.3, –111.6 |

**Table 2** *Continued*

| *Sample* | *Chemical group* | *Chemical shift (relative to $CFCl_3$ in ppm)* |
|---|---|---|
| | $–CF_2CH_2CF_2CF_2CH_2–$ | –115.2, –116.7 |
| | | –116.4 |
| | $–CH_2CF_2CF_2CH_2CH_2–$ | –116.8, –116.4 |
| | $–CH_2CF_2CF_2CF(CF)_3CH_2–$ | –119.3, –119.6 |
| | | –120.0, –120.2 |
| | $–CH_2CF_2CF_2CFClCH_2–$ | –119.9, –120.2 |
| | $–(CF_2)_n–$ | –123.2, –123.6 |
| | | –123.7, –123.8 |
| | $–CF_2CF_2CFClCH_2CF_2–$ | –122.3 |
| | $–CF_2CFClCH_2–$ | –127.7, –129.8 |
| | | –130.5 |
| | $–CH_2CF_2CF_2CF_2CH_2–$ | –127.3, –126.7 |
| | $–CF_2CF(CF_3)CF_2CF_2CF(CF_3)–$ | –128.4, –130.5 |
| | | –131.3 |
| | $–CH_2CF(CF_3)CF_2CF_2CF(CF_3)–$ | –135.7 |
| | $–CF_2CF_2CF_2(CF_3)CH_2CF_2–$ | –184.9, –185.0 |
| | | –185.4, –185.8 |
| $(CF_2CF_2)_n$ | $–CF_2CF_2CF_2–$ | –123.2 |
| $(–C(CF_3){=}C(CF_3)–)_n$ | $CF_3–$ | –50.2 |
| | $CF_3–$ | –54.1 |
| | $CF_3–$ | –82.4 |
| | $CF_3CF_2–$ | –127.9 |
| | $–CF_2CF_2CF_2–$ | –123.7 |
| Poly(vinyl difluoride) | Imperfections | –113 |
| | C–F | –89 |
| $C_8F_{17}SO_3Na$ | $CF_3–$ | –83.0 |
| | $CF_3CF_2–$ | –129.8 |
| | $–CF_2CF_2CF_2–$ | –124.7 |
| | $–CF_2SO_3–$ | –118.2 |
| Perfluoronaphthalene ($C_{10}F_8$) | $F_1$ | –146.9 |
| | $F_2$ | –148.4 |
| | $F_3$ | –151.7 |
| | $F_4$ | –152.6 |
| $C_6F_5NH_2$ | $F_{2.6}$ | –165.2 |
| | $F_{3.5}$ | –167.7 |
| | $F_4$ | –177.8 |
| AgTFA | $CF_3CO_2^-$ | –71.5 |
| $CFCl_3$/PU foam | $CFCl_3$ | –2.0 |
| $CFCl_2CH_3$/PU foam | $CFCl_2CH_3$ | –45.6 |
| $CHCl_2CF_3$/ PU foam | $CHCl_2CF_3$ | –79.6 |
| $C_6H_4F$–co-polymer | *p*-F | –108.5 |
| $CF_4$/plasma treated diamond | CF | –148 |
| | $CF_2$ | –106, –123 |
| | $CF_3$ | –78 |
| $C_6F_6$ sorbed on polystyrene | Dual mode sorbed | –117, –161 |
| | Mobile $C_6F_6$ | –169 |
| Bis(trifluoromethyl) aniline/polystyrene | Dual mode sorbed | –55.7 |
| $C_6F_6$ sorbed on butyl rubber | Dual mode sorbed | –153.4 |
| $CF_2CFHCF_3$ subs. adamantane | $CF_3$ | –75 |
| | $CF_2$ | –122, –130 |
| | CFH | –209 |

**Table 2** *Continued*

| *Sample* | *Chemical group* | *Chemical shift (relative to $CFCl_3$ in ppm)* |
|---|---|---|
| Fluorinated steroids | CF | –165.6, –165.2 |
| | | –171.4 |
| | CHF | –187.9, –180.1 |
| | $CH_2F$ | –192.3 |

All shifts have been recalculated with respect to $CFCl_3$ as the reference.
[a] Possibly an inconsistent assignment. Use with caution.

As an alternative to averaging in real space, dipolar decoupling can also be achieved in spin space by using multiple-pulse sequences such as WAHUHA-4, MREV-8 or BR-24. The efficiencies and spectral sweep-widths of these sequences are limited by the cycle time, e.g. to average a dipolar width of 50 kHz the cycle would have to be short compared to 20 μs. This places considerable technical demands on the probe design, namely short ring-down times (<<10 μs) and the capacity for short high-power pulses under high duty cycle. Homonuclear decoupling efficiency is further improved by Combined Rotation And Multiple-Pulse Spectroscopy (CRAMPS). Even though this kind of experiment was not considered in the past to be easily implemented, modern spectrometers can be more routinely adjusted with respect to accurate phase cycling and timing. CRAMPS measurements are usually carried out under quasi-static conditions (typically 2–3 kHz MAS), but recent experiments have also been designed for fast MAS (10–20 kHz) in combination with short windowless multiple-pulse sequences. To assign simple compounds, the line widths in a CRAMPS spectrum are often sufficient, even though they are still broader than in solution state $^{1}H$ NMR.

**Figure 1** illustrates some sources of line-broadening that cannot be prevented even on a perfectly well adjusted spectrometer. The resolution of the glassy polymer polystyrene is about an order of magnitude worse than that of crystalline adipic acid, which is partially attributed to the dispersion of chemical shifts in the amorphous material. Additionally, it has recently been recognized that the anisotropic molecular magnetic susceptibility provides a dipolar broadening mechanism for aromatic groups, which limits their peak resolution to about 0.3 ppm. Further difficulties in a CRAMPS experiment will be encountered when a group undergoes motions with a correlation time comparable to the pulse width, to the multiple-pulse cycle time, or to the sample spinning speed, as this can lead to severe losses in signal intensity.

Multidimensional spectroscopy plays an essential role in solution state $^{1}H$ NMR, by providing correlations through bonds (COSY, TOCSY), through space (NOESY), or to heteroatoms (HMQC, HSQC). Homonuclear 2D spectra of solids are only feasible when the sample possesses an intrinsically high mobility to give a sufficient resolution, which is the case for liquid crystals. Heteronuclear correlation (HETCOR) experiments, on the other hand, are a powerful tool to map even the very broad $^{1}H$ signals of a rigid molecule onto a second $^{13}C$ dimension via the dipolar coupling through space. Since $^{13}C$ possesses a chemical shift range of about 200 ppm, compared with 12 ppm for $^{1}H$, the increase in resolution makes it possible to assign resonances in relatively large organic compounds and to study hydrogen-

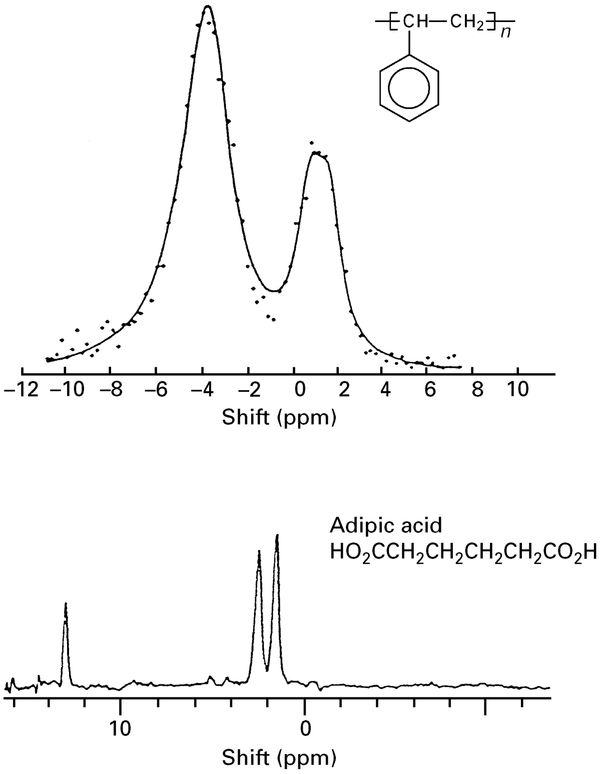



**Figure 1** $^{1}H$ NMR CRAMPS spectra of amorphous polystyrene (top) and crystalline adipic acid (bottom), showing different degrees of resolution. Reproduced with permission of Elsevier Science from Gerstein BC (1998) $^{1}H$ NMR. In: Ando I and Asakura T (eds) *Solid State NMR of Polymers*, Vol 84, pp 166–189. Amsterdam: Elsevier Science.

bonding in biological solids. During the evolution and mixing periods of a HETCOR experiment, the suppression of $^{1}$H–$^{1}$H homonuclear and $^{1}$H–$^{13}$C heteronuclear couplings can be achieved by multiple-pulse sequences and/or MAS. In the related wide-line separation experiment (WISE), where $^{1}$H-line-widths are used as a measure of the respective segmental mobilities, the homonuclear interaction is allowed to be active during the evolution period. Whereas HETCOR is applicable to sites carrying any number of protons, the dipolar chemical shift correlation experiment (DIPSHIFT) has been designed to determine the accurate bond distance in a particular heteronuclear spin-pair, e.g. of $^{1}$H–$^{15}$N in polypeptides.

Another sophisticated way of correlating spins with one another is possible by multiquantum spectroscopy, which can in principle resolve the connectivities of atoms that are present as an isolated cluster. Various multiple-pulse experiments have been designed to recouple isolated spins under fast MAS via multiquantum (or zero-quantum) coherences. When faced with an abundant network of $^{1}$H, however, it is usually only feasible to monitor build-up rates qualitatively as a measure of the dipolar coupling strength. For example, when molecular motion reduces this interaction in polymers, the signal intensities after quantum production can be interpreted in terms of the respective local dynamics.

A different kind of $^{1}$H NMR application in polymers is the use of long-range spin diffusion to study the heterogeneity of the sample. In this experiment a gradient of magnetization is set up between different regions that can be differentiated either by their mobilities (i.e. amorphous and crystalline) or by their chemical shifts. During a subsequent storage period the magnetization will diffuse over distances of up to 30 nm per second, such that domain sizes can be estimated by choosing suitable delay times. In these systems a CRAMPS experiment may be applied to resolve groups of signals, and multiple-pulse sequences have even been used as a chemical shift filter to ‘catch’ a resonance and keep it aligned with the static magnetic field direction. Strictly, however, the experiments mentioned in this paragraph are derived from broadline experiments and do not contribute themselves to an improvement in resolution.

### $^{19}$F NMR of abundant spins

The situation with abundant $^{19}$F spins, e.g. in fluoropolymers and inorganic fluorides, is similar to what has been outlined for $^{1}$H. The same basic experiments apply, although the much broader range of $^{19}$F isotropic chemical shifts (see **Table 2**) affords an intrinsically better resolution. $^{19}$F MAS spectra of rigid solids will be accompanied by a higher number of spinning side-bands due to the increased chemical shielding anisotropy (CSA). Technically it becomes more demanding to excite uniformly the large spectral width of $^{19}$F (90° pulse length of typically 1–2 μs), especially in multiple-pulse experiments which are quite sensitive to offset. To cover a broad spectral width in a CRAMPS experiment, it is essential to use the shortest possible cycle times. To that aim, the ring-down time of a probe can be reduced by Q-spoiling, which involves the addition of resistance to the RF circuit but has the disadvantage of requiring higher power levels for a given 90° pulse length. Alternatively, the signal can be built up in a point-wise fashion at the end of an incremented number of cycles. Since this single-point approach dispenses with the need for sampling during the pulse train, windowless or semi-windowless sequences can be used, although this method is associated with an obvious penalty in measurement time. In terms of hardware, a fast digitizer is required for $^{19}$F NMR, and fluorine-containing materials should be avoided in the probe components. The relaxation mechanisms of $^{19}$F are somewhat more complicated than those of $^{1}$H, as they are affected not only by dipolar couplings but also by the chemical shielding anisotropy and hence magnetic field strength.

### $^{19}$F NMR of dilute spins in the presence of protons

Dilute $^{19}$F spins in a protonated environment do not suffer from direct homonuclear line broadening, but a new problem arises from the possibility of cross-relaxation to $^{1}$H. In principle these heteronuclear interactions are inhomogenous, and MAS should be effective for narrowing. However, the situation becomes complicated when strong $^{1}$H–$^{1}$H coupling leads to spin diffusion within the proton network, because this leads to a homogeneous dephasing of the $^{19}$F magnetization. Therefore, to obtain well resolved $^{19}$F NMR spectra strong $^{1}$H-decoupling is essential (ideally up to 250 kHz). A double-resonance $^{19}$F{$^{1}$H} probe should thus be capable of handling high power levels on both channels, and efficient filters are required to separate the two frequencies which are only 6% apart. With increasing magnetic field strengths the filtering becomes easier. Probe design may nevertheless require new approaches, such as transmission lines or cavity resonators, especially when a third channel is to be included for a low-frequency nucleus. As a consequence of high-power $^{1}$H-decoupling, the $^{19}$F resonances will be affected by the Bloch–Siegert shift,

which is given by the ratio

$$\Delta = (\gamma_F\ B_{1H})^2/(\omega_F{}^2 - \omega_H{}^2) \qquad [2]$$

where $\gamma_F\ B_{1H}$ is the magnitude of the $^1$H-decoupling field at the $^{19}$F resonance frequency. On a 4.7 T spectrometer (200 MHz for $^1$H) the shift is typically a few ppm towards lower frequency, but it becomes less significant at higher magnetic field strength.

**Figure 2** illustrates the effects of $^1$H-decoupling, of MAS, and of intrinsic molecular mobility on the $^{19}$F NMR spectrum of the fluorinated sterol diflucortolon-21-valerate (DFC). While the static $^{19}$F line shape of the crystalline material shows no recognizable features (**Figure 2A**), $^1$H-decoupling reveals the two underlying CSA tensors (**Figure 2B**). With additional MAS at moderate speed (**Figure 2C**), the isotropic chemical shifts are resolved (as marked by an asterisk), and the principal values of the tensors can be calculated from the spinning sideband intensities using the Herzfeld–Berger approach. To demonstrate the molecular mobility in a liquid crystalline environment, in the next example (**Figure 2D**) the hydrophobic DFC was incorporated at a total concentration of 10% (w/w) into lipid bilayers composed of dimyristoylphosphatidylcholine (DMPC). Two narrow lines appear in this $^{19}$F {$^1$H} MAS spectrum (**Figure 2D**), corresponding to the two $^{19}$F-substituents on those sterol molecules that are partitioned into the bilayer. Since they undergo rapid long-axial rotation and restricted anisotropic tumbling, their total line width is less than the MAS spinning speed (5 kHz) and no sideband patterns are observed. Additionally, a residual contribution of crystalline material is visible in the same spectrum (**Figure 2D**), whose integrated intensity shows that DFC can only be dissolved up to 3.5% (w/w) in hydrated DMPC bilayers. Notably, the isotropic chemical shifts of DFC in the hydrophobic bilayer interior differ significantly from those in the crystalline environment, whereas they are rather similar (within 1 ppm) to the values obtained in $CHCl_3$ solution (**Figure 2E**).

An elegant way to distinguish $^{19}$F-substituents based on their mobility or their proximity to protons is provided by cross-polarization (CP) from $^1$H to $^{19}$F, allowing, for example, the suppression of PTFE background signals. Having been originally developed under static conditions, this experiment can be used under MAS with some slight modifications. By choosing a short CP contact time, the $^{19}$F-signals from immobile groups are selected since they are engaged in a rapid transfer of dipolar magnetization. Alternatively, the dipolar dephasing experiment can be used

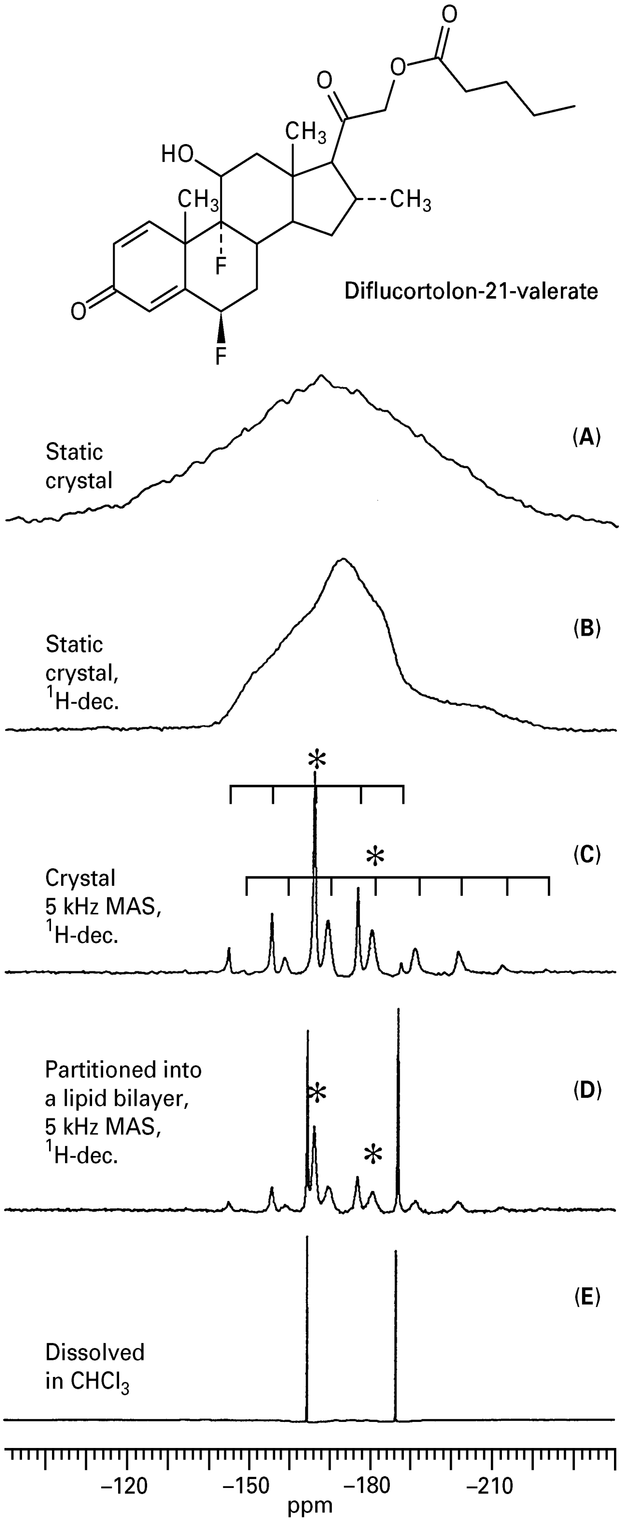



**Figure 2** $^{19}$F NMR spectra of diflucortolon-21-valerate (DFC) under different experimental conditions and in different environments, illustrating the effects of $^1$H-decoupling, MAS and intrinsic molecular mobility.

to select resonances of $^{19}$F-spins that are in comparatively mobile domains or remote from protons. CP MAS can also be combined with inversion recovery

measurements, since the $^{19}F$ relaxation tends to be affected by cross-relaxation to $^{1}H$. The $T_1$-values of $^{19}F$ are obtained either by direct polarization or by a suitable post-CP variable delay, whereas $T_1$-relaxation of the protons is revealed by inserting a pre-CP delay on $^{1}H$ during the $^{19}F$ {$^{1}H$}-experiment.

Homonuclear $^{19}F$–$^{19}F$ correlations provide information about spin diffusion through space, as, for example, in a 2D spin-exchange experiment. This approach requires sufficiently fast MAS and $^{1}H$-decoupling to achieve complete resolution, since any magnetization exchange due to spectral overlap of the $^{19}F$ resonances has to be avoided. On the other hand, it may be desirable to stimulate a controlled magnetization exchange by introducing a certain spectral overlap under moderate MAS speeds. That is, when the spinning side-bands of one signal are allowed to overlap with the side-band pattern of a second signal, these two spins are recoupled by Rotational Resonance (RR). The RR experiment and related broadband recoupling schemes were originally developed to measure weak pairwise couplings between selective $^{13}C$-labels and thus to determine their accurate distances (up to 7 Å). When applied to a fluorinated sterol, the RR effect was observed as a significant line-broadening rather than a dipolar fine-structure, which is consistent with rapid spin diffusion. Despite these difficulties, fluorinated substituents are desirable for RR and related experiments, because $^{19}F$–$^{19}F$ distances could theoretically be extended up to 16 Å. In fact, long-range heteronuclear distances ($^{19}F$ to $^{13}C$ or $^{31}P$) have been successfully determined by Rotational Echo Double Resonance (REDOR) and Transferred Echo Double Resonance (TEDOR) experiments. Owing to the fast spin diffusion of the $^{19}F$ magnetization it was usually necessary to observe the REDOR effect on the low-frequency channel under $^{19}F$-dephasing. Overall, many of the experiments known for $^{13}C$ can be adapted for dilute $^{19}F$ spins provided that efficient $^{1}H$-decoupling is available, and there is much potential for further development.

### High-resolution $^{1}H$ and $^{19}F$ NMR of oriented samples

Instead of averaging away the anisotropic nuclear spin interactions by MAS or multiple-pulse sequences, it is possible to take advantage of the anisotropy of these interactions, provided that macroscopically oriented samples are available. This kind of static solid-state NMR approach is entirely different from the experiments discussed above. It can lead to highly resolved $^{1}H$ and $^{19}F$ spectra with narrow lines, whose position carries information about the alignment of individual molecular segments. Oriented samples are typically prepared by stretching polymers or suitable proteins into fibres, or by aligning polymer films, liquid crystals or biological membranes on planar supports. For the most basic experimental set-up, the symmetry axis of the uniaxially oriented sample is aligned parallel to the static magnetic field direction. That way all groups of any one kind in the molecule will experience the same nuclear spin interaction. This interaction is characterized by a single value of the relevant anisotropy tensor, for example by the CSA tensor or by the inhomogenous dipolar coupling to another dilute spin (see Eqn [1]). The resulting $^{1}H$ or $^{19}F$ NMR signal will thus consist of a narrow line, whose resonance position reveals the geometric relationship between the respective tensor and the macroscopic axis of the oriented sample. In contrast, a non-oriented static sample gives rise to a broad powder line shape, and an MAS spectrum is simply reduced to the isotropic resonance. In practice, the effective line width of an oriented sample is limited by the quality of alignment, plus any residual broadening.

## Applications of $^{1}H$ and $^{19}F$ NMR

### Inorganic solids

The classic inorganic compound investigated by solid state $^{1}H$ NMR is $CaSO_4 \cdot xH_2O$. As a powder it yields the well-known Pake pattern, and oriented single crystals have been used to analyse the angle-dependence of the homonuclear dipolar coupling. Today, most $^{1}H$ NMR applications employ fast MAS, preferentially at high field. Studies of hydroxyl groups in zeolites, in glasses, and on the surface of catalysts yield information about their local structure and acidity. Using $^{19}F$ MAS or CRAMPS experiments, the isotropic chemical shifts of many inorganic fluorides have been determined (see **Table 2**). Here, $CaF_2$ and its fluorohydroxyapatite derivatives $Ca_5(PO_4)_3\,F_x(OH)_{1-x}$ are of particular interest due to their relevance in human dental enamel, in bone implants, as well as in fluorescent lights. In addition, the fluoride species formed upon adsorption onto alumina, silica, clays, zeolites or related materials have attracted attention in view of their role as catalysts or molecular sieves.

### Organic molecules

Solid-state $^{1}H$ NMR constitutes a powerful approach to investigate the hydrogen-bonding and ionization states of small organic compounds, with similar applications in structural biology. High-

resolution $^{1}H$ MAS and CRAMPS experiments have provided isotropic chemical shifts and CSA tensors for many simple molecules. In several instances a direct correlation with hydrogen-bonding lengths could be demonstrated, e.g. for amino acid carboxyl groups. Well resolved $^{1}H$–$^{13}C$ HETCOR spectra have been obtained for various drugs (e.g. ibuprofen) and oligopeptides. With regard to $^{19}F$ NMR, apart from the isotropic chemical shift values the $^{19}F$ CSA tensors have been characterized for various aliphatic (e.g. silver trifluoroacetate, $C_8F_{17}SO_3Na$) and aromatic (e.g. pentafluoroaniline, perfluoronaphthalene) compounds by analysing the spinning sideband intensities. Pharamaceutically active sterols with one, two, and three $^{19}F$-substituents (cf. **Figure 2C**) were used to establish more advanced $^{19}F\{^{1}H\}$ and triple-resonance $^{13}C\{^{19}F\}\{^{1}H\}$ NMR experiments. Since the 1D spectra were found to be rather sensitive to sample polymorphism, they can be used to detect different crystalline forms that may exhibit different pharmacological activities.

### Polymers and rubbers

High-resolution $^{1}H$ NMR spectra of polymers are most readily obtained when the material possesses an intrinsically mobile backbone or side-chains (e.g. an elastomer), when measured at high temperature, or when a deuterated solvent has been imbibed [e.g. $CDCl_3$ in PMMA, or a swollen functionalized resin carrying a bound product]. For more rigid polymers, sufficiently narrow line widths have been achieved with the aid of fast MAS or CRAMPS to resolve aliphatic and aromatic resonances. Proton line widths and relaxation times have thus been used to describe segmental mobilities and crystallinity, to assess phase separation in co-polymers (e.g. PS–PVME), or to examine hydrogen-bonding and acidity (e.g in PVA or polysiloxanes). Unlike the case with protonated polymers, where $^{13}C\{^{1}H\}$ NMR tends to be the method of choice, fluorinated polymers give strongly broadened $^{13}C$ NMR spectra, which creates the need for direct $^{19}F$-observation. Owing to the wide chemical shift range of $^{19}F$ NMR, simple broadline/relaxation NMR studies were quite popular in the past, but many high-resolution experiments have been demonstrated since. Studies of commercially important materials include PTFE, $(CF_2CF_2)_n$; PTrFE, $(CF_2CFH)_n$; PCTFE, $(CF_2CFCl)_n$; and PVDF, $(CF_2H_2)_n$; as well as co-polymers of the latter material with HFP, $(CF_2CFCF_3)$; TFE, $(CF_2CF_2)$; and CTFE, $(CF_2CFCl)$. Some advanced triple-resonance REDOR experiments have been performed with $^{19}F$-labels in dendrimers and in nanospheres, to visualize their packing density and the distribution of sequestered ligands, respectively.

### Liquid crystals and lipid bilayers

The fast long-axial rotation ($\tau_c \approx 10^{-10}$ s) of liquid crystalline molecules leads to a partial averaging of the anisotropic interactions, thus improving the basic spectral resolution. For example, fast MAS of hydrated phospholipid dispersions yields well-resolved $^{1}H$-NOESY spectra with distinct NOE cross peaks. Another important property of liquid crystals lies in the local alignment of the molecules about a director axis, which makes it useful to prepare macroscopically oriented samples. In a mechanically oriented sample, quasi-isotropic $^{1}H$ NMR spectra should be accessible by aligning the static sample axis along the magic angle. That way the intrinsic rotational diffusion of the molecules will average all anisotropic interactions to zero, although in practice any mis-alignment and slow collective motions of the molecules will reintroduce some line broadening. As an example, **Figure 3** illustrates the $^{1}H$ NMR spectra of an un-oriented dispersion of DMPC (**Figure 3A**) versus an oriented sample on glass plates (**Figure 3B**) aligned at the magic angle. Resolution can now be further enhanced by slowly spinning the oriented sample around the magic angle, which is called Magic-Angle Oriented Sample Spinning (MAOSS). The spectrum in **Figure 3A** shows narrow lipid signals ($^{1}H$ line widths of 10–30 Hz) at a spinning speed of only 200 Hz, which compares favourably with the resolution achieved for lipid dispersions under high-speed MAS (14 kHz) or at a much higher magnetic field strength (17.6 T). It appears hopeful that even the signals of membrane-bound polypeptides (e.g. gramicidin, M13 coat protein) may be resolved using either of these MAS approaches.

Selective $^{19}F$ labelling of lipids has contributed to a better understanding of the structural and motional properties of artificial model bilayers as well as biological membranes (e.g. of the bacterium *Acholeplasma laidlawii* when fed with $^{19}F$-labelled fatty acids). On the basis of early broadline NMR investigations, resolution has been improved by using oriented samples and multiple-pulse experiments to select certain anisotropic spin interactions. For example, the local order parameter of a particular acyl-chain segment can be calculated from the dipolar splitting of the corresponding $CF_2$-labelled group. **Figure 4** illustrates how the $^{19}F$–$^{19}F$ dipolar coupling of 4,4-DMPC-$F_2$ can be extracted, which is not resolved in the normal $^{19}F$ NMR spectrum of an un-oriented lipid dispersion (**Figure 4A**), unless

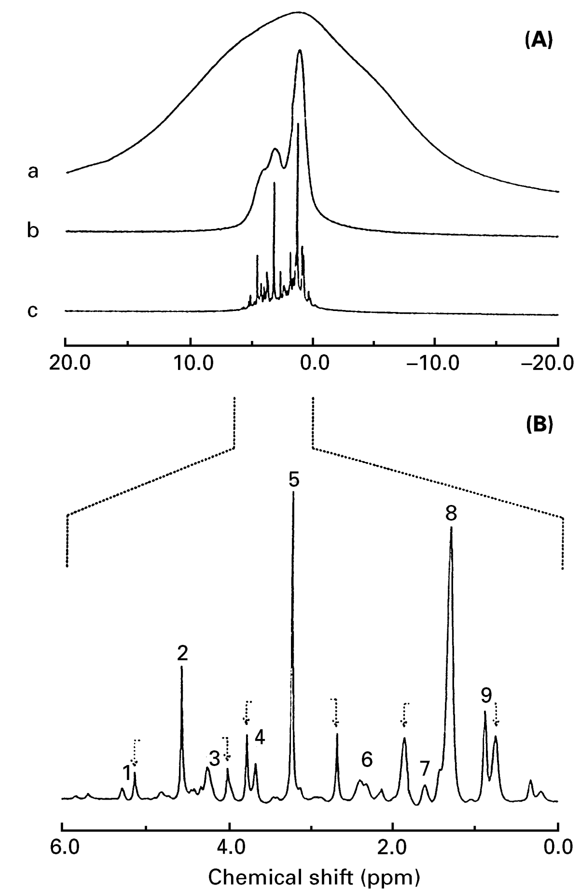



**Figure 3** $^{1}H$ NMR spectra of DMPC bilayers as a non-oriented lipid dispersion (A), and as a macroscopically oriented sample that is aligned along the magic angle (B). The resolution is further improved by MAOSS at 200 Hz (see text). Reproduced with permission of Academic Press from Glaubitz C and Watts A (1998) *Journal of Magnetic Resonance* **130**: 305–316.

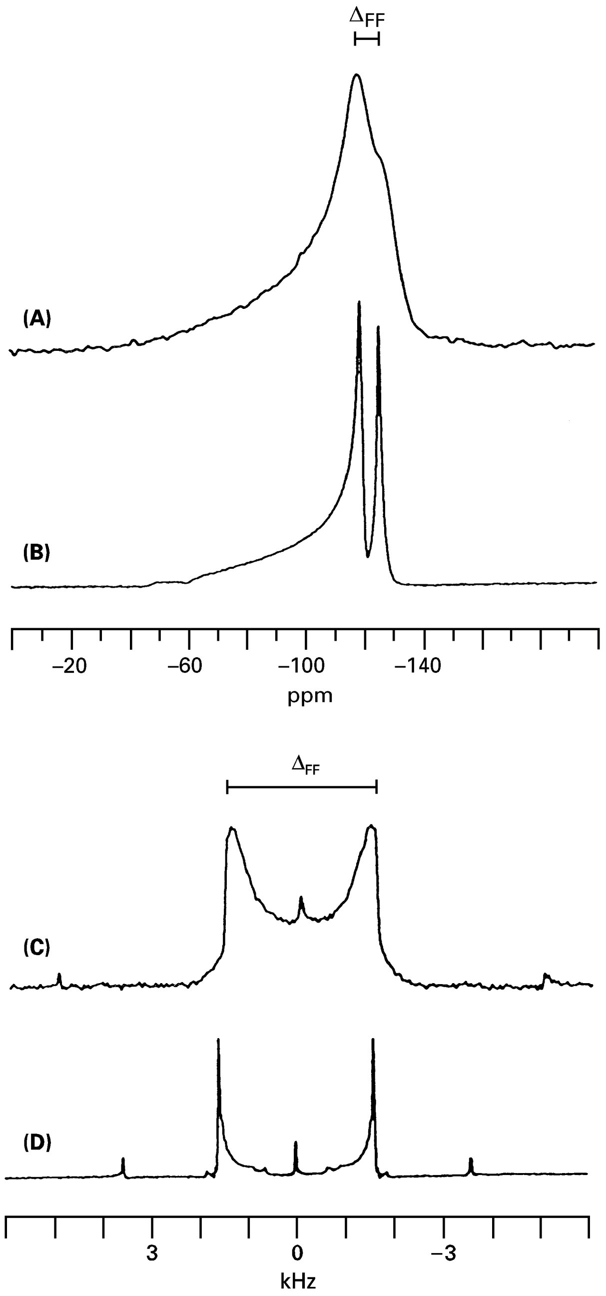



**Figure 4** $^{19}F$ NMR spectra of 4,4-DMPC-$F_2$ as a non-oriented lipid dispersion without $^{1}H$-decoupling (A), and with $^{1}H$-decoupling (B). The dipolar splitting $\Delta_{FF}$ is directly resolved without $^{1}H$-decoupling in the CPMG experiment (C), and the lines are sharpened further when an oriented sample is prepared and aligned perpendicular to the magnetic field (D). Adapted with permission of Academic Press from Grage SL and Ulrich AS (1999) *Journal of Magnetic Resonance* **138**: 98–106.

additional $^{1}H$-decoupling is applied (**Figure 4B**). Alternatively, and without any need for $^{1}H$-decoupling, the Carr–Purcell–Meiboom–Gill (CPMG) multiple-pulse sequence is able to suppress all unwanted line-broadening interactions such that only the desired homonuclear $^{19}F$ coupling remains in the shape of a Pake pattern (**Figure 4C**). By extending the CPMG experiment to an oriented sample that is aligned with its symmetry axis perpendicular to the magnetic field (**Figure 4D**), an even better resolution (< 50 Hz line width) is achieved.

## Peptides and proteins

Solid-state NMR lends itself to the structural analysis of proteins that do not crystallize, or when details about their functional mechanism or hydrogen-bonding pattern are to be investigated. For example, long-range distances have been measured by heteronuclear REDOR experiments on enzymes prepared with specifically $^{19}F$-labelled side-chains in the frozen or lyophilized state. In this way it could be demonstrated that the substrate binding cleft of 5-enolpyruvylshikimate-3-phosphate synthase closes upon ligand binding, whereas the binding pocket of tryptophan synthase remains largely unperturbed.

Another, different kind of NMR approach has been developed to investigate the structure of membrane-associated peptides and proteins. This method relies on the use of oriented membrane samples, and it measures the $^{1}H$–heteronuclear dipolar coupling and $^{15}N$-CSA to determine the orientation of a labelled $^{1}H$–$^{15}N$ bond in the polypeptide backbone. To build a full molecular structure from such orientational constraints, usually a series of selectively labelled peptides have to be prepared and the alignment of each single backbone segment is then measured step-by-step. Recently, a large number (>100) of signals from a uniformly $^{15}N$-labelled membrane protein were successfully resolved, using a 2D PISEMA experiment with homonuclear Lee–Goldburg decoupling. Once an assignment strategy for this novel approach is developed, solid-state NMR will provide a set of tools for a comprehensive structural analysis of membrane proteins.

## List of symbols

$\hbar$ = Planck's constant/2π; $I$ = nuclear spin quantum number; $r$ = distance between spins; $T_1$ = relaxation time; $\Delta$ = Bloch–Siegert shift; $\gamma$ = gyromagnetic ratio; $\mu_0$ = permittivity of vaccum; $\nu_r$ = spinning frequency; $\tau_c$ = correlation time; $\theta$ = angle between internuclear vector and magnetic field direction; $\omega$ = Larmor frequency.

*See also:* **$^{19}F$ NMR Applications, Solution State; High Resolution Solid State NMR, $^{13}C$; Liquid Crystals and Liquid Crystal Solutions Studied By NMR; Membranes Studied By NMR Spectroscopy; NMR of Solids; NMR in Anisotropic Systems, Theory; Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules; Solid State NMR, Rotational Resonance; Solid State NMR, Methods; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Peptides; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals.**

## Further reading


Duncan TM (1990) *A Compilation of Chemical Shift Anisotropies*. Chicago: Farragut Press.

Gerstein BC (1998) $^{1}H$ NMR. In: Ando I and Asakura T (eds) *Solid State NMR of Polymers,* Vol 84, pp 166–189. Amsterdam: Elsevier Science.

Glaubitz C and Watts A (1998) Magic angle-oriented sample spinning (MAOSS): a new approach to biomembrane studies. *Journal of Magnetic Resonance* **130**: 305–316.

Grage SL and Ulrich AS (1999) Structural parameters from $^{19}F$ homonuclear dipolar couplings, obtained by multipulse solid state NMR on static and oriented systems. *Journal of Magnetic Resonance* **138**: 98–106.

Grant DM and Harris RK (eds) (1996) *Encyclopedia of Nuclear Magnetic Resonance*, Chapters on Bilayer Membranes: $^{1}H$ & $^{19}F$ NMR; CRAMPS; HETCOR in Organic Solids; Proton Chemical Shift Measurements in Solids. Chichester: Wiley.

Hafner S and Spiess HW (1998) Advanced solid-state NMR spectroscopy of strongly dipolar coupled spins under fast magic angle spinning. *Concepts in Magnetic Resonance* **10**: 99–128.

Harris RK, Monti EA and Holstein P (1998) Fluoropolymers. In: Ando I and Asakura T (eds) *Solid State NMR of Polymers*, Vol 84, pp 667–712. Amsterdam: Elsevier Science.

Harris RK, Monti GA and Holstein P (1998) $^{19}F$ NMR. In: Ando I and Asakura T (eds) *Solid State NMR of Polymers*, Vol 84, pp 253–266. Amsterdam: Elsevier Science.

Harris RK and Jackson P (1991) High-resolution fluorine-19 magnetic resonance of solids. *Chemical Reviews* **91**: 1427–1440.

Miller JM (1996) Fluorine-19 magic angle spinning NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* **28**: 255–281.

# Hole Burning Spectroscopy, Methods

**Josef Friedrich**, Technische Universität München, Germany

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| **Methods & Instrumentation** |

## Introduction

Much of our information on the microscopic features of matter in various aggregate states has been obtained from spectroscopy. How detailed such information can be depends on the resolution of the experiment. During the evolution of modern physics it often happened that theory predicted features that could not be verified experimentally because the resolution of the associated spectroscopic techniques was too low. As an example, consider the famous experiment by Pound and Rebka in 1960 by which the gravitational red shift of an electromagnetic wave could be verified. This red shift was predicted by Einstein in 1911 on the basis of general relativity, but the verification had to wait until there was a technique available with the necessary resolution. The required resolution is incredibly high, of the order of $10^{15}$, but, nevertheless, could be achieved by making use of the Mössbauer effect. A resolution of $10^{15}$ means that the frequency of an electromagnetic wave of $10^{15}$ Hz can be determined with an accuracy of 1 Hz. The highest resolution that can be achieved in the spectroscopy of a particular system (e.g. of atoms, molecules, nuclei, etc.) is determined by the natural width of the spectral lines in these systems. According to the Heisenberg principle, this width is determined by the lifetime of the excited state. However, in almost all real systems this ultimate limit cannot be achieved in a straightforward way because the environment of the system broadens the respective levels. The Mössbauer effect is one technique where the natural line width can be exploited for high-resolution spectroscopy. Hole burning techniques also work at this ultimate limit of resolution, but the physical principle is totally different. Nevertheless, because of the ability to do spectroscopy at the limit determined by the natural line width, the hole burning technique is sometimes called the 'optical Mössbauer effect'. The following surveys how the hole burning technique evolved during the last few decades of the twentieth century, its physical basis, and the fields of spectroscopy and technology in which it is used.

## Hole burning: a survey

Hole burning stands as a generalized synonym for all kinds of saturation spectroscopic techniques in inhomogeneously broadened bands. At the frequency where saturation is performed, a dip appears in the spectrum, the so-called hole. Generally speaking, the hole burning technique aims to unravel the features of the homogeneous line shape of the transition involved, which is obscured by the presence of strong inhomogeneous broadening. Hole burning was demonstrated for the first time in 1948 in NMR spectroscopy. Since then, it has been used in almost all fields of spectroscopy: NMR, ESR, dielectric, IR and optical spectroscopy. Hole burning experiments have been applied to almost all aggregate states: gases, liquids, crystals, glasses and polymers. More recently biological materials have also attracted increasing interest.

The power of the technique rests with a special variation of hole burning, so-called persistent spectral hole burning. Persistent spectral hole burning was discovered in 1974 by Gorokhovskii, Kharlamow and their respective co-workers. It is based on population storage in a long-lived intermediate, which is very often a photochemical state. In this case the method is called photochemical hole burning. It is the persistence of the holes that has made the hole burning method an important and attractive technique in a very broad field of spectroscopy techniques and applications. The application to optical data storage in the frequency domain, with the possibility of increasing the bit density beyond the diffraction limit, has attracted much interest. Information storage in the time domain is also a promising field of application. In addition, hole burning is a technique highly suited to storing multiple holograms in narrow frequency domains. Spectral holes can be exploited for frequency stabilization of laser radiation and for shaping of laser pulses as well as for ultra-narrow optical filters. In spectroscopy, the method has gained most attention in optical spectroscopy of the solid state. The following survey focuses on this field of study.

## The homogeneous line shape function: the perfect crystal case

Consider an optical transition of a dye probe molecule doped a low concentration into a perfect crystal lattice (**Figure 1A**). Since all the probes have the same local environment, their absorption frequencies coincide and the line shape of the transition considered is representative of the line shape of a single probe molecule. Excitation of the probe molecule is accompanied by a charge redistribution in the excited state. This leads to a different equilibrium configuration of the lattice molecules in the excited state. As a consequence, there is a certain probability that the optical excitation will be accompanied by excitation of lattice motions that give rise to so-called phonon side bands. The intensity distribution is determined by the Franck–Condon principle. The relative intensity of the transition with no lattice phonon excitation, the so-called zero-phonon line, is given by the Debye–Waller factor $\alpha$:

$$\alpha = \frac{\text{Intensity in zero-phonon line}}{\text{Total intensity in the band}}$$

As a rule, the zero-phonon line is much narrower than the accompanying phonon transitions. Consequently, its peak intensity is much higher than the peak intensity of the phonon wing (**Figure 1B**). The reason for the large line width of the transitions into phonon states is the short lifetime of these states, which is largely determined by vibrational relaxation.

In comparision with the phonon states, the lifetime $T_1$ of the purely electronic state can be orders of magnitudes longer. According to the Heisenberg law, the width $\Delta\omega_0$ of the zero-phonon line is given by

$$\Delta\omega_0 = \frac{1}{T_1} \quad [1]$$

$\Delta\omega_0$ is called the natural width of the transition. As a rule, the transition attains this width only at very low temperatures, i.e. at around or below a few kelvin. At higher temperatures, the line width is determined by the thermal fluctuations of the lattice, which lead via some coupling mechanism to thermal fluctuations of the energy levels involved, i.e. of the electronic ground and excited states. Since these fluctuations disrupt the phase of the polarization, they are called dephasing processes. The decay time $T_2^*$ of the phase coherence is called the pure dephasing time. These dephasing processes provide an additional contribution to the line width, so that the total line width $\Delta\omega_h$ is determined by two contributions:

$$\Delta\omega_h = \frac{1}{T_1} + \frac{2}{T_2^*} \quad [2]$$

$\Delta\omega_h$ is called the homogenous line width. As a rule, the associated line shape is Lorentzian. As the pure dephasing contribution $2/T_2^*$ dies out rapidly as the temperature approaches absolute zero, $\Delta\omega_h$ approaches its ultimate limit of $\Delta\omega_0$ at very low temperature. As an example, consider an organic dye molecule. $T_1$ is typically of the order of some tens of nanoseconds, so that $\Delta\omega_0$ is of the order of some tens

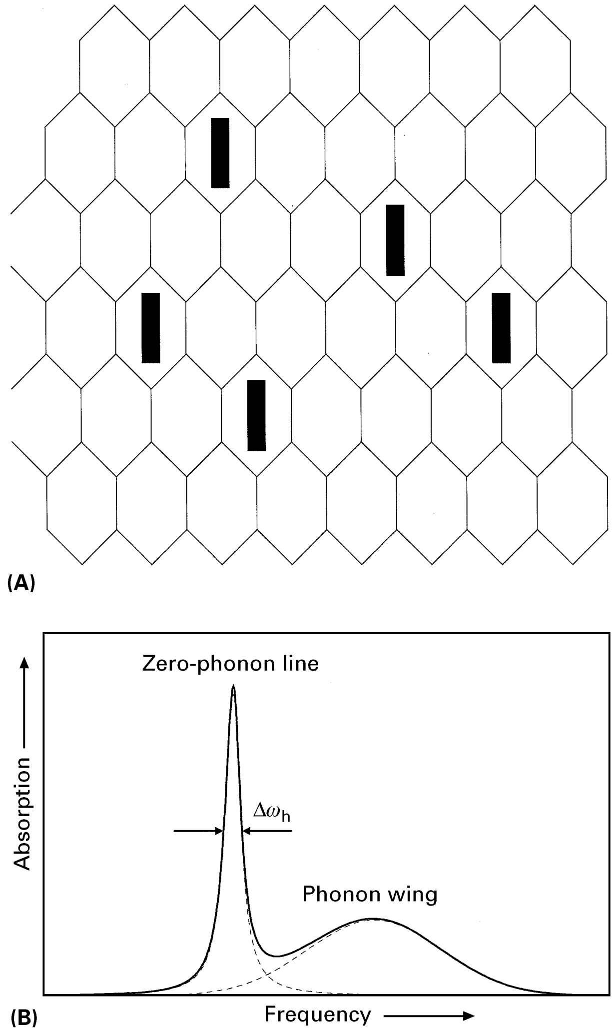



**Figure 1** (A) Sketch of a perfect host lattice (honeycomb) doped with probe molecules (black bars). (B) Line shape of an ensemble of identical probe molecules in a perfect host lattice.

of MHz. In this case the ultimate limit of spectroscopic resolution is around $10^7$.

## Inhomogeneous broadening

A basic requirement for spectral hole burning to be observed is inhomogeneous line broadening. In NMR, inhomogeneous broadening originates mainly from field inhomogeneities. In the gas phase, inhomogeneous broadening comes from the velocity distribution of the molecules, which leads to a frequency distribution via the Doppler effect. In the solid state, inhomogeneous line broadening occurs because of structural disorder. For example, in a crystal there are lattice defects that cause statistically varying strain fields. These strain fields lead to statistically varying frequency shifts of the zero-phonon transitions of the probe molecules. As a consequence, the zero-phonon frequencies are spread out and the line experiences a broadening $\Delta\omega_i$. Whereas the homogenous line broadening $\Delta\omega_h$ is dynamic in nature, inhomogeneous broadening is, at sufficiently low temperature, essentially static in nature.

In a more general way this situation is sketched in **Figure 2A**. Each of the probe molecules sees a different environment and hence their absorption frequencies are different.

In organic crystals, the inhomogenous broadening is of the order of a wavenumber; in glasses it is of the order of several hundred wavenumbers. In this latter case, even the phonon side bands are largely buried beneath the inhomogeneous envelope (**Figure 2B**).

In contrast to the homogeneous line shape, which is Lorentzian, the inhomogeneous distribution is governed by the statistics of large numbers and hence is Gaussian. At sufficiently low temperatures, the ratio of the inhomogenous to the homogenous width may cover 5–6 orders of magnitude.

## Hole burning

Suppose the dye probe molecule is photoreactive. Then, to understand the essential features of hole burning, it suffices to consider three states only: the ground state $S_0$, the first excited singlet state $S_1$, and the photoproduct state $P_0$ (**Figure 3**). For simplicity we assume that $P_0$ has an infinite lifetime. The absorption spectrum of $P_0$ is shifted as compared to that of $S_0$ because $P_0$ represents a structurally different molecule. Consider **Figure 4**: narrow-bandwidth laser light is tuned to some position within the inhomogeneous band. We denote the respective frequency by $\nu_L$. Excitation leads to photochemistry; hence, population is transferred to $P_0$ and, concomitantly, $S_0$ is depleted. However, depletion in $S_0$ can only occur in a spectral range around $\nu_L$ roughly determined by the homogenous line width $\Delta\omega_h$. At temperatures of a few kelvin, $\Delta\omega_h$ is close to $\Delta\omega_0$ and hence is extremely narrow. As a consequence, the depletion dip is extremely narrow. The depletion dip is called the 'spectral hole'. If $P_0$ has an infinite lifetime, the lifetime of the hole is, of course, also infinite. In this case the hole is persistent. However, in the usual terminology holes are called persistent if their lifetime is much longer than the typical hole burning timescale determined by burning and reading the hole. A typical order of magnitude for this timescale is a few minutes. For vanishingly small depletion and sufficiently low power of the burning laser, the shape of the depletion hole is Lorentzian with a width close to $\Delta\omega_h$. This depletion hole is, for instance, detected by scanning the

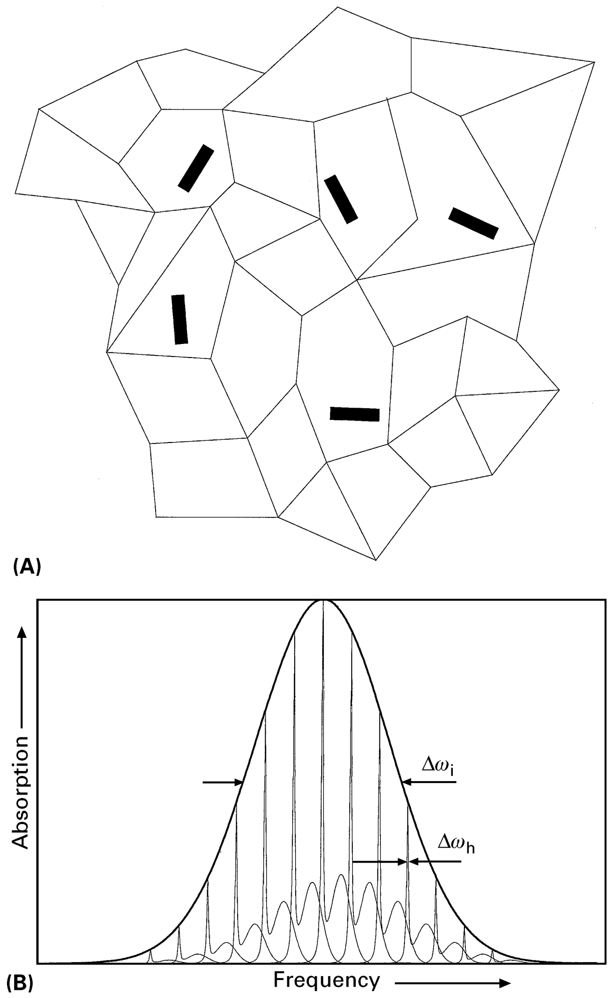



**Figure 2** (A) Sketch of a disordered lattice doped with probe molecules (black bars). Note that each probe molecule has a different environment. (B) Inhomogeneous line broadening: the inhomogeneous line of width $\Delta\omega_i$ is built from an ensemble of homogeneous lines of width $\Delta\omega_h$.

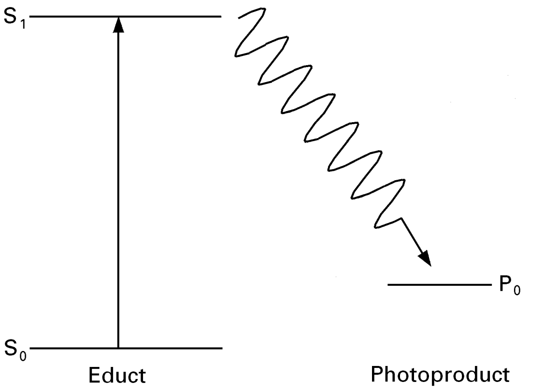



**Figure 3** Three levels are necessary for persistent hole burning. $P_0$ has a long lifetime and acts as a storage state.

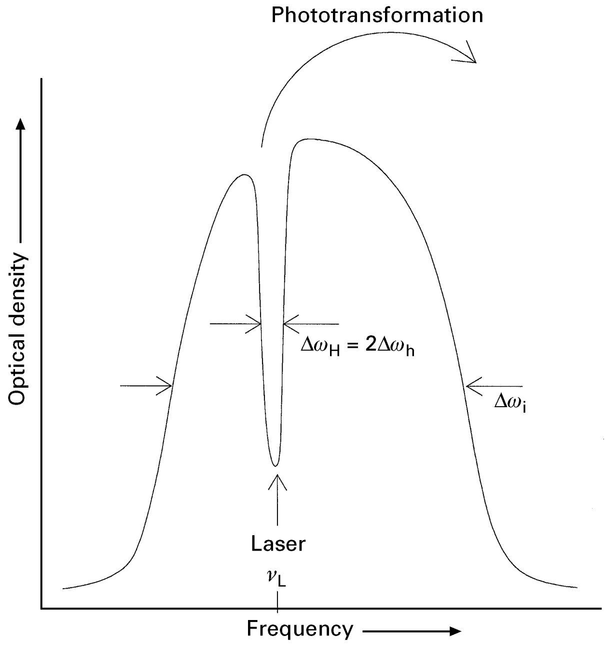



**Figure 4** Hole burning: laser radiation with frequency $\nu_L$ creates a population dip and, concomitantly, a hole in the inhomogenous band.

laser over the depletion range. The resulting line shape of the recorded hole is a convolution of the depletion hole with the homogeneous line shape, i.e. a convolution of a Lorentzian with a Lorentzian that is itself a Lorentzian of twice the width:

$$\Delta\omega_{\mathrm{H}} \Rightarrow 2\Delta\omega_{\mathrm{h}} \qquad [3]$$

where $\Delta\omega_H$ is the hole width.

**Figure 5** shows a hole in the inhomogenous absorption band of protoporphyrin IX-substituted myoglobin. The insert shows the hole on an enlarged scale.

## Saturation broadening

Saturation spectroscopy of two-level systems results in a power-dependent line width of the transition concerned, a phenomenon termed power broadening. In a similar fashion, line broadening due to strong saturation can also occur in a three-level system with a long-lived storage state $P_0$. However, in this case, it is not the power that determines the broadening but rather the number of irradiating photons, which is proportional to the irradiated energy dose or to the laser fluence. Hence, we call this broadening 'fluence broadening'. It is easy to see why the irradiated energy and not the power is the important factor: the power needed to burn a hole with a certain depth can be made arbitrarily small by increasing the irradiation time accordingly. This has no effect on the hole area so long as $P_0$ has an infinite lifetime. However, if the irradiated energy increases, photochemistry in the centre of the line slows down because the number of absorbers decreases, but it is still going on in the wings. As a consequence, the centre of the line is more strongly depleted and the line flattens and becomes broader. Because of this saturation broadening, the homogeneous line width must be determined from a plot of the hole width as a function of laser fluence in the limit of vanishing fluence. Likewise, one can plot the hole width as a function of the hole depth (**Figure 6**): as the depth approaches zero, the hole width converges to $2\Delta\omega_h$.

## Hole burning techniques

The deeper the holes, the better is the signal-to-noise ratio. However, deep holes are artificially broadened

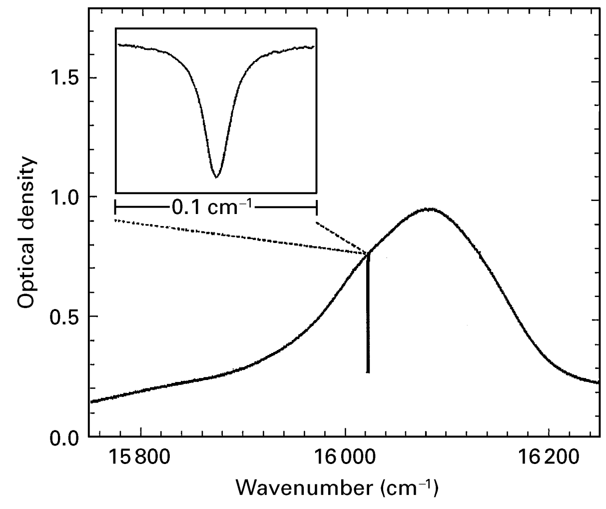



**Figure 5** A spectral hole in the absorption spectrum of myoglobin complexed with protoporphyrin IX at $T = 1.5$ K.

owing to saturation. Hence, techniques have been developed for reading shallow spectral holes with a high signal-to-noise ratio.

The most straightforward way to read a hole is to scan the laser over the frequency range where burning has been performed. In the range of the hole, the number of absorbers is diminished and so there is an increased transmission. The transmission technique requires a baseline correction since the hole is detected on top of an inhomogeneous background absorption. For an optical density around 1, relative transmission changes of a few per cent can easily be detected in this way.

If the optical density at the burning frequency is well below 1, the fluorescence excitation mode is more appropriate for reading the hole. In this mode it is not the transmission but rather the fluorescence that is measured. If the laser is scanned across the spectral range of the hole, the fluorescence decreases. The relative change of the fluorescence as a function of frequency directly follows the shape of the hole. A baseline correction is required in this case, too.

There are some elaborate techniques that are extremely sensitive because they are based on zero-background signals. The most important of these are the dichroitic and the holographic techniques.

In the dichroitic technique, the hole is burnt with a polarized laser field, say in the $z$ direction (**Figure 7**). The polarized laser radiation leads to an anisotropic distribution of the vector $N$, denoting the number density of absorbers oriented with their transition dipole moment at an angle $\theta$ to the polarization of the burning field. Reading is performed with the sample between two crossed polarizers $P_1$ and $P_2$, inclined at say $-45°$ and $+45°$ to the $z$ direction. If the sample is isotropic, it does not alter the polarization direction, so the reading laser field $E_1$ will be $P_1$-polarized and will be blocked by $P_2$. On the other hand, if at frequency $\nu - \nu_L$ the sample has a preferred transmission for $z$-polarized light, then the reading laser field $E_1$ will be rotated towards the $z$ axis. As a consequence, the respective field vector $E_1'$ will acquire a component $E_2'$ along the $P_2$ direction and light will be transmitted. In this way the hole is measured against zero background intensity. Accordingly, the signal-to-noise ratio can be orders of magnitude better than with the transmission or fluorescence detection modes and tiny changes in optical density can be measured. However, two things have to be accounted for: the degree of anisotropy is not constant over the frequency range of the hole, and it also depends on the depth of the hole. This so-called polarization bleaching is very similar to saturation broadening. For strong polarization bleaching, the shape of the hole may be flattened or may even show a double peak structure. Apart from this effect, the hole shape may also suffer from interference with light which is not totally blocked by $P_1$ and $P_2$ as a result of small degrees of misalignment.

In the holographic technique two beams, 1 and 2, interfere at an angle of $2\theta$ in the sample and write a spatial grating of the absorption coefficient and of the index of refraction (**Figure 8**). The modulation of this grating is frequency dependent via the line shape function. If one of the writing beams, say 2, is blocked, the other beam can serve as the reading beam. If it is tuned across the frequency range of the hole, the grating deflects the reading beam (1) into

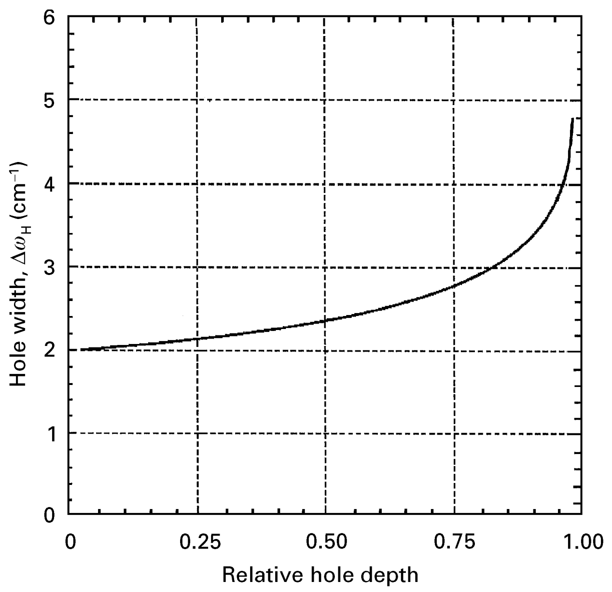



**Figure 6** Saturation broadening due to laser fluence. Reproduced with permission from Dick B (1989) Habilitationsschrift, University of Göttingen.

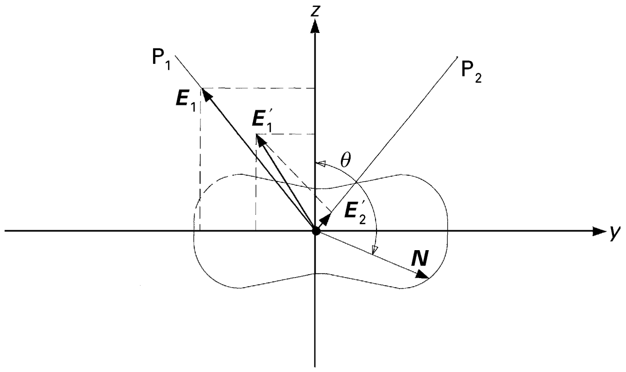



**Figure 7** Exploiting the dichroism of holes for a low-noise reading process. ***N*** denotes the number density of absorbers with their dipole moments in the direction $\theta$. $P_1$ is the polarizer, $P_2$ the analyser.

the direction of beam 2. The intensity of the diffracted light is proportional to the modulation depths of the absorption coefficient and of the index of refraction, and hence is strongest at the centre of the hole and vanishes far in the wings. Although the holographic technique reproduces the true line shape of the hole as long as the saturation of the hole is weak, it has characteristic properties that have to be taken into account. First, strong bleaching results in a strong deviation from a sinusoidal grating, which leads to higher diffraction orders. Second, strong bleaching also leads to a strong distortion of the Lorentzian line shape, which may eventually show a double peak structure.

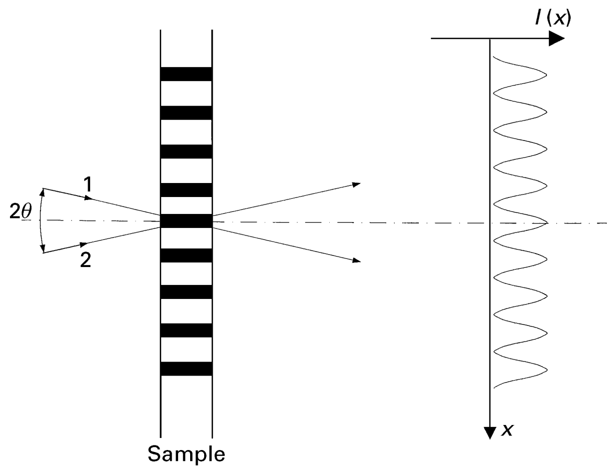



**Figure 8** Holographic hole burning: burning is performed via two beams, 1 and 2, which interfere in the sample and create a spatial grating $I(x)$ of the absorption coefficient and the index of refraction.

## Applications of hole burning

### Spectroscopy

Hole burning has been used in a very broad field of scientific problems ranging from solid-state physics to biological physics. The most straightforward application is the measurement of the homogeneous line width. Apart from measuring vibrational relaxation in the excited electronic state in this way, most effort has been put into measuring the homogeneous line width of the purely electronic $S_1 \leftarrow S_0$ absorption of dopant molecules in glasses. The homogeneous line width in a glassy host material has quite specific properties compared to that in crystalline materials (**Figures 9A** and **B**): in the low-temperature regime, $T < 4$ K, the homogeneous width in a glass is an order of magnitude larger than in a crystal (**Figure 9A**). In a crystal, it is already close to its lifetime-limited value and hence is temperature independent (**Figure 9B**). In a glass, it depends on temperature down to the lowest temperatures measured in these types of experiments (~50 mK); however, its temperature dependence is weak. It is usually well described by a temperature power law $\sim T^{\alpha}$ with $\alpha$ between 1 and 2. In a crystal, the temperature dependence goes as $T^7$ or is very often activated. The reason for the specific features of the optical line width in glasses is the presence of specific degrees of freedom, so-called TLS (two-level system) modes, which reflect the disorder in the structure and which cannot be frozen out, not even at the lowest temperatures reached so far in solid-state materials.

The TLS degrees of freedom reflect the simplest approach to characterizing the complex energy landscape of disordered materials such as glasses and polymers. These materials are never in true thermodynamic equilibrium, and, hence, there are always some structural dynamics covering an extremely broad range of relaxation rates of up to 15 orders of magnitude, from nanoseconds to months. These structural dynamics create locally fluctuating strain

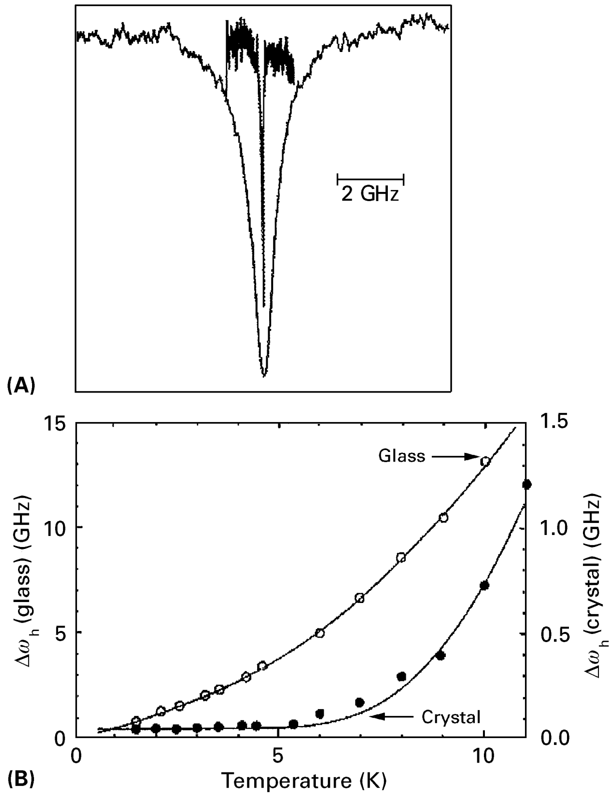



**Figure 9** (A) Hole widths of chlorine-doped benzophenone glass (broad hole) and crystal (narrow hole). (B) The homogeneous line widths of chlorine-doped benzophenone glass and benzophenone crystal as a function of temperature. Note the different scales for glass and crystal.

and/or electric fields that, in turn, lead to fluctuations of the electronic energy levels of the dye probe. These fluctuations are reflected in a time-dependent line broadening, called spectral diffusion (**Figure 10**). Measurement of the spectral diffusion broadening gives insight into the structural dynamics and the associated features of the energy landscape. This is especially interesting for biological materials.

Another major field in which the hole burning technique can be exploited concerns measuring the influence of external fields, such as electric, magnetic or pressure fields, on the electronic energy levels. For instance, measuring the Stark effect (electric field) even in good-quality crystals usually requires special modulation techniques since the spectral changes are rather small. In disordered materials, owing to the large inhomogeneous broadening, it is difficult to obtain good-quality data unless unusually large changes of the dipole moments are involved. In hole burning, moderate field strengths (10 kV $cm^{-1}$) and pressure levels (1 MPa) yield easily detectable line shifts and field-induced changes in the line widths. Moreover, the high resolution and the associated selectivity of these experiments enables one to investigate field effects as a function of frequency within the inhomogeneous band. The solvent shift can be varied simply by tuning the laser frequency and hence becomes a parameter of the experiment. From hole burning Stark effect experiments (**Figure 11**), information on the molecular dipole moments involved, on the respective polarizabilities, on matrix fields and on associated symmetries can be obtained. High-quality quantum chemical calculations are usually desirable to support the interpretation of the experiments because there are always several parameters that are not very well known. From pressure experiments (**Figure 12**), information on local compressibilities, on molecule–matrix interactions and on the vacuum absorption frequency of the probe molecule can be obtained.

Hole burning has been used widely to investigate relaxation processes of coherent tunnelling in the solid state. An outstanding example in this context is the methyl group. Consider a dye probe molecule

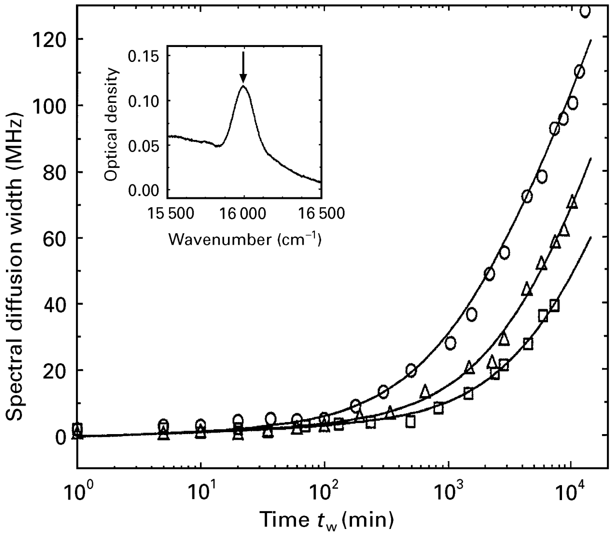



**Figure 10** Spectral diffusion broadening as a function of waiting time $t_w$ for different ageing times. The sample is protoporphyrin IX in a dimethylformamide/glycerol glass; temperature is 100 mK. Insert: Broad-band absorption spectrum. The arrow marks the wavenumber where hole burning was performed.

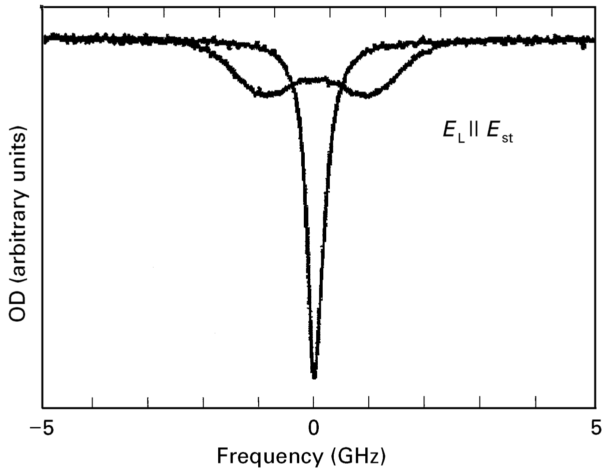



**Figure 11** Deformation of a spectral hole in an external field. The external field $E_{st}$ (about 12 kV $cm^{-1}$) is parallel to the laser field. The sample is protoporphyrin IX-substituted myoglobin in a glycerol/water glass. The double peak structure is indicative of a chromophore with a well-defined dipole moment.

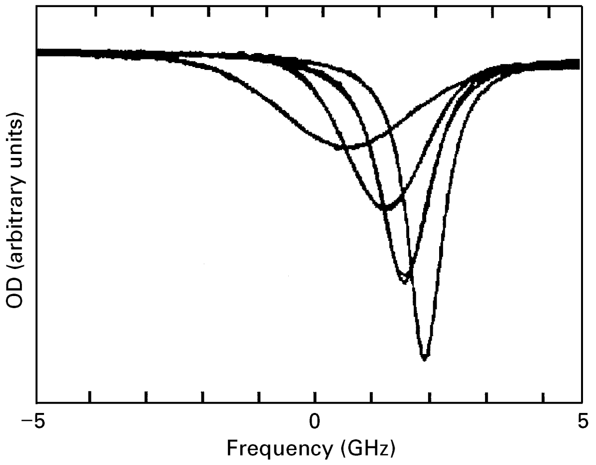



**Figure 12** Shift and broadening of a spectral hole as a function of pressure. Maximum pressure is 0.6 MPa. The sample is resorufin in glycerol glass at liquid-helium temperature.

with one or several methyl groups attached, for instance dimethyl-s-tetrazine (**Figure 13**). Irrespective of how disordered the host matrix is, the methyl group always has at least a 3-fold rotational symmetry axis (along the C–C bond), which results in a 3-fold periodic potential $V = V_3 \cos 3\varphi$, where $\varphi$ is the angle of rotation around the C–C bond and $V_3$ is the barrier height. The lowest states in this potential interact via the tunnelling interaction. As a result, the three states form a split pair, with A and E symmetry, where the E state is still doubly degenerate. Since the $C_3$ group is isomorphic to the group of even permutations of the protons, the wavefunctions of the methyl rotor have to be totally symmetric according to the Pauli principle. Relaxation between the tunnelling states requires an interaction, which breaks the permutation symmetry. These interactions necessarily contain the nuclear spin, and so are very small. Accordingly, the lifetimes of these states can be extremely long (e.g. months at 1 K) and thus population can be stored in these states. This is where hole burning comes in. Suppose laser light excites a molecule in its A-rotor (E-rotor) state from $S_0$ to $S_1$ (**Figure 14**). From $S_1$, intersystem crossing to the triplet state $T_1$ occurs. In the $T_1$ state the electron spin turns around the nuclear spin via the hyperfine interaction, the rotor converts from the A to the E state or from the E to the A state. Light irradiation thus creates a population redistribution among the rotational tunnelling states. The point is that this population redistribution is accompanied by a frequency transformation because the tunnelling splitting $\delta^*$ in $S_1$ may be different from that in $S_0$. Hence, a hole appears at the laser frequency, which is accompanied by two antiholes shifted symmetrically by the difference $|\delta - \delta^*|$ between the tunnelling splitting in $S_0$ and $S_1$ (**Figure 13**). The hole offers an easy way to measure the rotational tunnelling relaxation. All one has to measure is how the hole recovers or how the antiholes vanish. **Figure 15** shows the tunnelling relaxation rate as a function of temperature for the protonated and the deuterated rotor. The rate increases with $T^7$, indicating a Raman-type relaxation process. Deuteration speeds up the relaxation instead of slowing it down because of the nuclear quadrupole interaction associated with the deuterons and the fact that the methyl rotor in the example considered is almost free.

Summarizing the application of hole burning in spectroscopy, it should be noted that there is also a large field in which it is used in elucidating photochemical reactions. It is also increasingly used in studies of photosynthesis, to shed light on the energy and electron transfer processes, on the nature of the states involved, on electron phonon coupling and

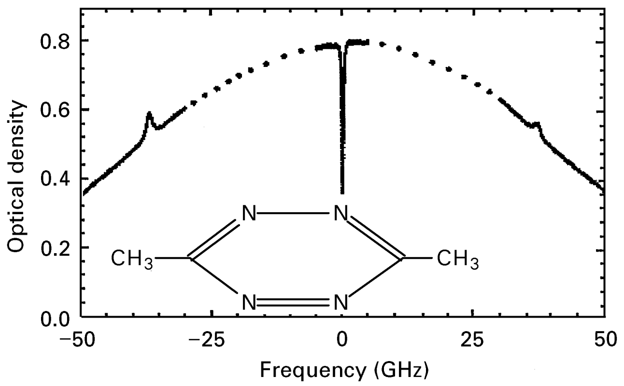



**Figure 13** Hole burning via nuclear spin conversion in rotational tunnelling states of methyl groups. The sample is dimethyl-s-tetrazine in *n*-octane at 1.4 K. The central hole at the laser frequency is accompanied by two antiholes symmetrically shifted by 37 GHz.

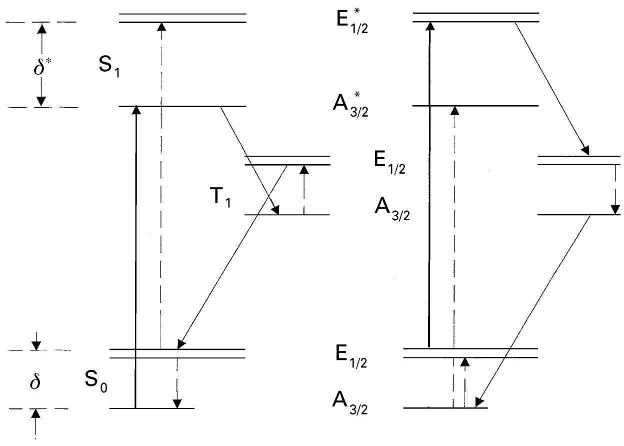



**Figure 14** The level scheme of a methyl rotor attached to a chromophore. $\delta$ and $\delta^*$ are tunnelling splitting in $S_0$ and $S_1$, respectively. $T_1$ is the lowest triplet state. E and A label the symmetry species of the rotor states, the subscripts characterize the total nuclear spin of the methyl protons.

exciton coherence. Further, the hole burning technique offers great power to investigate the features of the energy landscape of proteins and the associated dynamics in conformation space.

## Technology

In 1976 the use of hole burning was suggested for frequency-domain optical data storage. **Figure 16** sketches the principle. The storage medium could, for instance, be a dye-doped polymer film. In each spatial spot (**Figure 16A**) a number of about $\Delta\omega_i/\Delta\omega_H$ bits can be stored in the frequency domain (**Figure 16B**), with $\Delta\omega_i$ being the inhomogeneous width and $\Delta\omega_H$ the hole width. $\Delta\omega_i/\Delta\omega_H$ could be as large as $10^4$–$10^5$. As a consequence, the capacity of

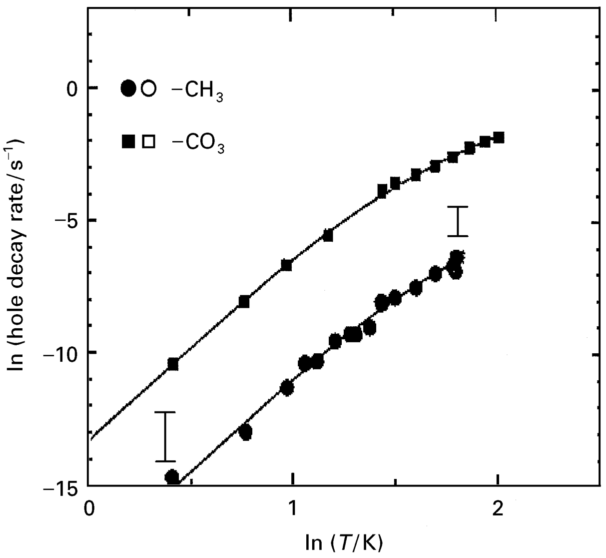



**Figure 15** Rotational tunnelling relaxation rate as a function of temperature in a log–log representation. Fitted curves are based on Raman-type relaxation processes. Note that the tunnelling relaxation in the deuterated rotor is faster by two orders of magnitude. The sample is dimethyl-s-tetrazine in *n*-octane.

an optical storage device could be increased by several orders of magnitude beyond the diffraction limit, i.e. to about $10^{12}$ bits cm$^{-2}$. So far such a memory has not been realized because too many requirements have to be met simultaneously: the high bit density is achieved only at liquid helium temperatures; at higher temperatures the information is, at least partly, lost; appropriate dye molecules should absorb in the frequency range where common laser diodes work; the respective Debye–Waller factors as well as the photochemical yields should be high so that information can be written and read in the nanosecond time regime; and repeated reading may deteriorate the stored information.

Solutions could be found for many of these requirements. For instance, the use of photon-gated hole burning ensures nondestructive read out. Many materials have been discovered in which this concept could be successfully demonstrated; however, attempts to unify all the requirements in one system have failed so far.

Instead of storing data bits in the frequency domain, they can be stored in the time domain, for instance as a pulse sequence $E_S$ as shown in **Figure 17A**. Writing of the information is performed by interference of the signal $E_S$ with a short reference pulse $E_R$ in the photochemical storage material. If the spectral width of the reference pulse $E_R$ is much larger than the width of the whole signal train $E_S$, which in turn is much larger than the homogeneous width of the respective optical transitions of the storage material, then the whole Fourier spectrum of the signal train can be stored as a grating in frequency space, i.e. as many holes in the inhomogeneous band whose width, of course, has to be even larger than the width of the reference pulse. **Figure 17B** demonstrates how

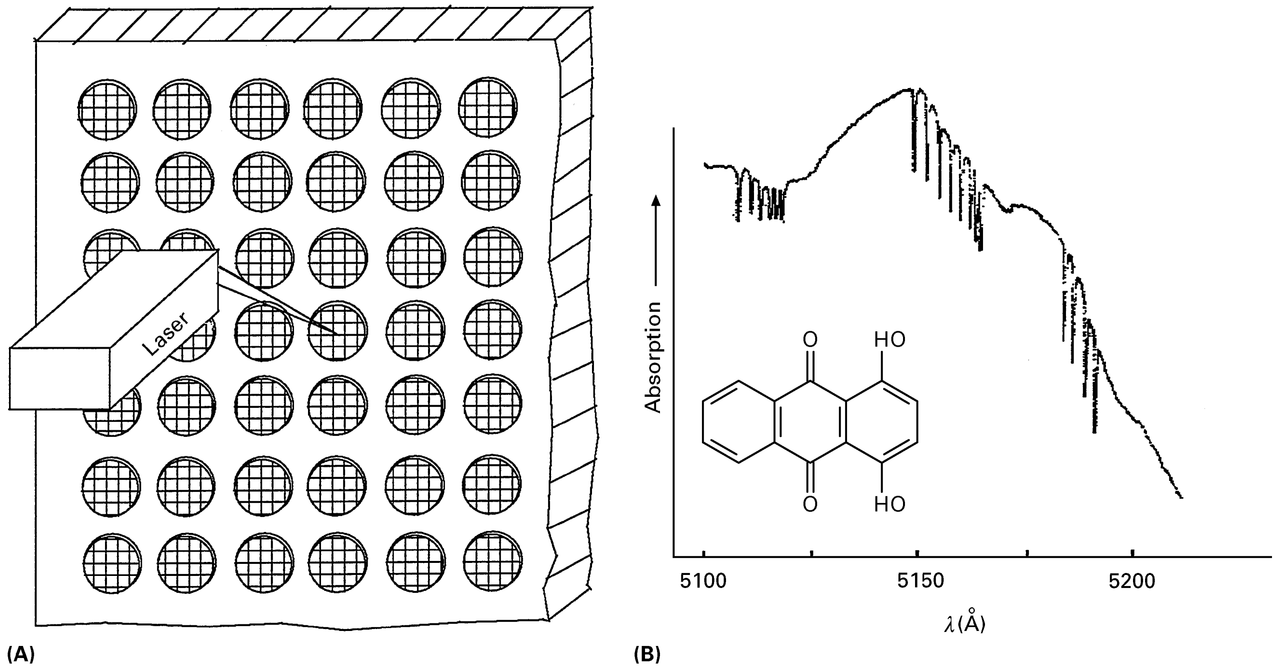



**Figure 16** (A) Hole burning and data storage. (B) In each spatial spot many bits can be stored in the frequency or wavelength domain. The dye probe, in this case, is quimizarin whose structure is shown in the insert.

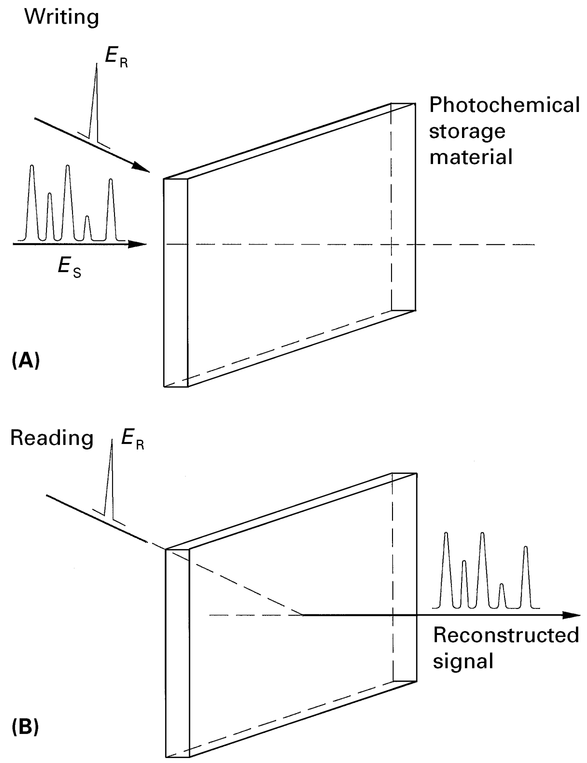



**Figure 17** (A) Data storage in the time domain using the 'PASPE' technique. (B) Reading the information. The echo is stimulated by a single pulse.

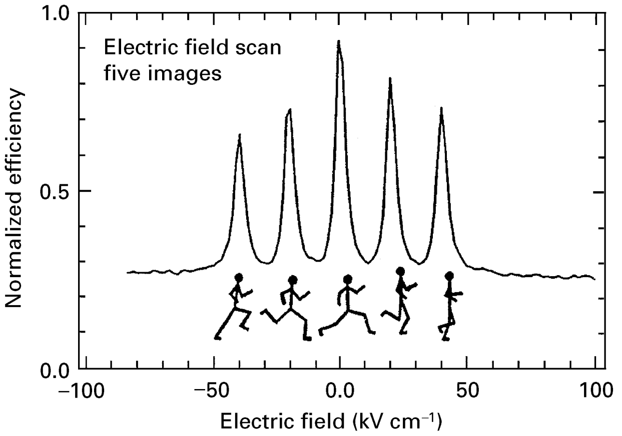



**Figure 18** Five holograms stored as holes in the electric field domain. By scanning the field, the five holograms can be read out serially, resulting in a movie showing a running jogger. Reproduced by permission of the Society of Photo-Optical Instrumentation Engineers (SPIE) from Wild UP and Renn A (1988) *Proceedings of the SPIE*. **910**: 61.

the stored information is recovered. If a single pulse $E_R$ hits the sample some time (hours, days, weeks) later, the sample regenerates the whole pulse sequence. Depending on how the reference pulse is applied in the writing and reading procedure, the time-reversed signal train can be generated as well. This is nothing other than a special form of the stimulated photon echo where the contrast of the stored information can be enhanced by accumulating the information over many writing cycles in the long-lived photochemical state. Accordingly, the technique has been called PASPE – photochemically accumulated stimulated photon echo.

As well as storing information bitwise, it is also possible to store information in an analogue fashion as a hologram of a picture. The concept is very similar to that shown in **Figures 17A** and **B**, but with $E_R$ and $E_S$ being the reference and the object waves, respectively. The reference wave can, for instance, be a plane wave. The object wave carries all the information of the picture. The two waves interfere for some time in the storage medium and create a hole at the laser frequency. In the frequency domain this hole is sharp, but in space it is distributed over a larger part of the storage material. Again, one can store many holograms in the frequency range covered by the inhomogeneous band. The holograms can be read by scanning the reading laser beam over the frequency interval where the holograms are stored. If the reading laser beam is directed along the direction of the reference beam, the burnt-in hologram regenerates the original object wave. Likewise, reading can be performed in the electric field domain by making use of the Stark effect. In this case the holograms are stored with an electric field applied in a way that each hologram has its own specific Stark field. The frequency of the burning (and reading) laser is fixed. In reading the holograms, the Stark field is tuned. Whenever the field reaches one of the specific values used during the storage process, a holographic interference pattern with a spatially well-resolved modulation of the absorption coefficient and index of refraction appears that scatters the light from the reading laser such that the object wave is regenerated. **Figure 18** shows an example: five pictures are stored as holograms in the five holes shown. Since each hole has its specific Stark field, reading is performed by scanning the electric field. In this way, a movie is generated, showing a jogger, as indicated in the figure.

## List of symbols

$E$ = field vector ($E = |E|$); $E_R$ = reference pulse; $E_S$ = signal pulse sequence; $N$ = vector of number density of absorbers oriented at $\theta$ to the polarization of the burning field; $T_1$ = lifetime of purely

electronic state; $T_2^*$ = pure dephasing time; $\alpha$ = Debye–Waller factor, temperature-law exponent; $\delta$ = tunnelling splitting; $\theta$ = angle between transition dipole moment and burning field polarization, angle between holographic beams; $\nu$ = frequency; $\Delta\omega_0$ = natural width of zero-phonon line; $\Delta\omega_h$ = homogeneous width of zero-phonon line; $\Delta\omega_H$ = hole width; $\Delta\omega_i$ = inhomogeneous line width.

## Further reading

Bloembergen N, Purcell EM and Pound RV (1948) Relaxation effects in nuclear magnetic resonance absorption. *Physical Review* **73**:679–712.

Borczyskovski C, Oppenländer A, Trommsdorff HP and Vial J-C (1990) Optical measurements of methyl-group tunneling and nuclear-spin conversion. *Physical Review Letters* **65**: 3277–3280.

De Caro C, Renn A and Wild UP (1991) Hole burning, Stark effect, and data storage: 2: holographic recording and detection of spectral holes. *Applied Optics* **30**: 2890–2898.

Dick B (1989) Polarisationsspektroskopie permanenter und transienter spektraler Löcher, Habilitationsschrift, Universit of Göttingen.

Friedrich J and Haarer D (1984) Photochemical hole burning: A spectroscopic study of relaxation processes in polymers and glasses. *Angewandt Chemie, International Edition in English* **23**: 113–140.

Gorokhovskii AA, Kaarli RK and Rebane LA (1974) Hole burning in the contour of a pure electronic line in a Shpol'skii system. *JETP Letters* **20**: 216–218.

Kharlamow B, Personov RI and Bykovskaya LA (1974) Stable ‘gap’ in absorption spectra of solid solutions of organic molecules by laser irradiation. *Optics Communications* **12**: 191–193.

Levenson MD, Macfarlane RM and Shelby RM (1980) Polarization-spectroscopy measurement of the homogeneous linewidth of an inhomogeneously broadened color-center band. *Physical Review B* **22**: 4915–4920.

Meixner A, Renn A and Wild UP (1989) Spectral hole-burning and holography. I. Transmission and holographic detection of spectral holes. *Journal of Chemical Physics* **91**: 6728–6736.

Moerner WE (ed) (1998) *Persistent Spectral Hole Burning: Science and Applications*. Berlin: Springer-Verlag.

Moerner WE (1995) Molecular electronics for frequency domain optical storage: persistent spectral hole-burning – a review. *Journal of Molecular Electronics* **1**: 55–71.

Pinsker M and Friedrich J (1996) Hole burning spectroscopy and quantum phenomena in methyl groups. *Molecular Crystals and Liquid Crystals* **291**: 97–102.

Rebane A (1998) Compression and recovery of temporal profiles of picosecond light signals by persistent spectral hole-burning holograms. *Optics Communications* **67**: 301–304.

Thorn-Leeson D, Wiersma DA, Fritsch K and Friedrich J (1997) the energy landscape of myoglobin: An optical study. *Journal of Physical Chemistry* **101**: 6331–6340.

Wegener H (1995) *Der Mössbauer-Effekt und seine Anwendung in Physik and Chemie*. Mannheim: Hochschultaschenbücher-Verlag.

Wild UP and Renn A (1988) Spectral hole burning and hologram storage. *Proceedings of the SPIE* **910**: 61–65.

# HPLC-IR, Applications

*See* **Chromatography-IR, Applications.**

# HPLC-MS, Methods

*See* **Chromatography-MS, Methods; Hyphenated Techniques, Applications of in Mass Spectrometry.**

# HPLC-NMR, Applications

*See* **Chromatography-NMR, Applications.**

# HREELS, Applications

*See* **High Resolution Electron Energy Loss Spectroscopy, Applications.**

# Hydrogen Bonding and Other Physicochemical Interactions Studied By IR and Raman Spectroscopy

**AS Gilbert**, Beckenham, Kent, UK

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

Vibrational spectra are sensitive to intermolecular interactions. This is most clearly evidenced by the differences seen in spectra of condensed phases compared to gases at low pressure which show not just the loss of rotational sidebands but also changes in profile, intensity and position. Hydrogen bonding (H-bonding) provides the most dramatic effects but even the weak van der Waals forces can significantly modify spectra. In liquids, collisions can distort molecular configuration leading to local breakdown of symmetry.

## Hydrogen bonds

### Introduction

The hydrogen bond is a weak, fairly directional, interaction represented here as a broken line in the scheme A–H---B. Atom A is electronegative, the A–H bond being therefore slightly ionic in character, and atom B possesses an area of basicity such as lone pairs on nitrogen, oxygen and halogen or $\pi$ electron rings (here B is a group of atoms) in certain aromatic systems. The energies of interaction are of an order of magnitude less than covalent bonds but greater than the very simple non-directional non-covalent forces (van der Waals or dispersion forces) that occur between all molecules. The unit A is generally oxygen, nitrogen or halogen but S–H and in certain circumstances even C–H can engage in H-bonding.

Theoretical and experimental studies suggest that the hydrogen bond is not purely electrostatic but has some covalent character. It is observed from crystallographic and other investigations that the length of the A–H bond increases and the distance between A and B decreases with increasing strength of interaction. In the limit of the weakest H-bonds, A and B are separated by about the sum of their collision radii. This is slightly greater, by about 0.5 Å than the sum of the van der Waals radii.

**Figure 1** shows IR spectra of isopropanol in solution ($CCl_4$) at various concentrations, the pathlength being adjusted to compensate as nearly as possible for the changes in molarity. It illustrates the dramatic effects of H-bonding (between the hydroxyl groups) on the O–H stretching mode. The free bond yields a relatively narrow band above 3600 cm$^{-1}$ but the singly and doubly H-bonded groups give broad and intense absorptions near 3500 cm$^{-1}$ and 3350 cm$^{-1}$ respectively.

This behaviour of the stretching band $\nu$ (A–H), of shift to lower frequency, increase in overall intensity and band broadening is highly characteristic and indicative of H-bonding. In extreme cases shifts of more than 2000 cm$^{-1}$ are observed while breadth and intensity increases can be two orders of magnitude or so.

The band shift and intensity changes can be crudely rationalized on the basis of a decrease in force constant due to weakening (lengthening) of A–H and increased polarization on interaction with B.

Raman spectra show similar behaviour to IR in terms of shift and band breadth but the intensity increase is not observed. **Figure 2** illustrates part of the spectrum of a mixture of deuterium chloride and dimethyl ether in the gas phase. The stretching vibration of unassociated DCl can be seen at 2085 cm$^{-1}$ along with rotational side bands, while H-bonded DCl yields a broad absorption underlying the sidebands centred near 1885 cm$^{-1}$.

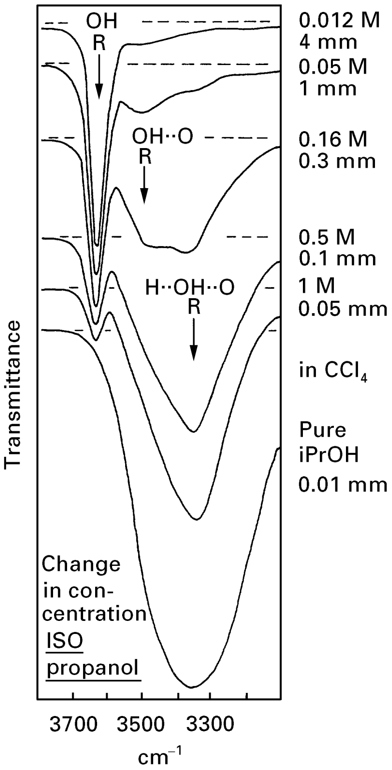



**Figure 1** Effect of H-bonding on the O–H stretching vibration of isopropanol. Reproduced with permission from Colthup NB, Daly LH and Wiberley SE (1975) *Introduction to Infrared and Raman Spectroscopy*. New York, San Francisco and London: Academic Press.

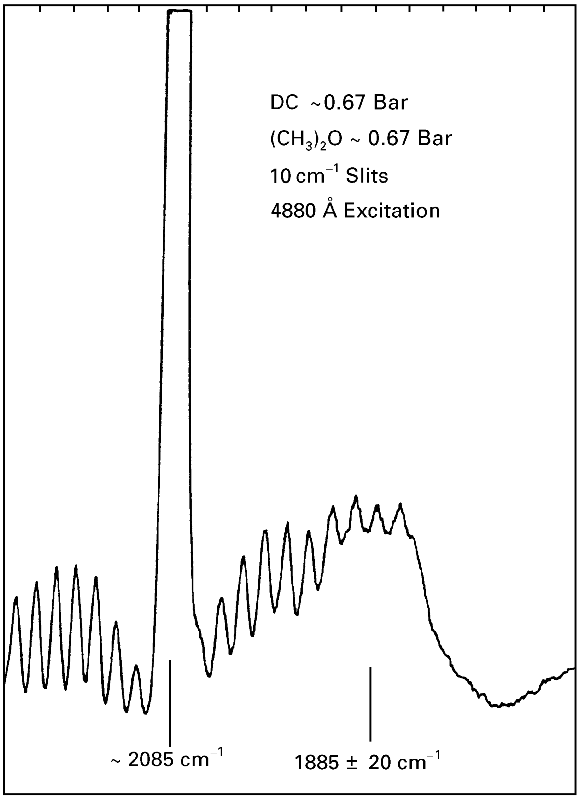



**Figure 2** Part of the Raman spectrum of a gaseous mixture of $(CH_3)_2O$ and DCl. Reproduced with permission from Gilbert AS and Bernstein HJ (1974) Measurement of rotational temperatures by Raman spectroscopy: application of Raman spectroscopy to the acquisition of thermodynamic values in a chemical system. In: Lapp M and Penney CM (eds) *Laser Raman Gas Diagnostics*, pp 161–169. New York: Plenum Press.

The bending vibrations of A–H are affected much less than the stretching vibration. Some broadening and intensification takes place but the small shift is to higher wavenumbers.

The new vibration, $\nu$(A–B), the overall stretching of the system and one which directly reflects H-bond strength, is found at low frequencies usually no higher than about 200 $cm^{-1}$. Probably the strongest H-bonded system in existence, the symmetrical bifluoride ion $(F–H–F)^{-1}$, has $\nu$(F–F) near 600 $cm^{-1}$.

Finally, the vibrational modes of bonds directly attached to B are significantly affected. Carbonyl groups which have been extensively studied as H-bond acceptors show downward shifts in $\nu$(C=O) of as much as 40 $cm^{-1}$.

## Dependence on temperature and concentration

In gas and liquid phases the extent of H-bonding decreases with rise in temperature and quantitative band-intensity measurements can give values for various thermodynamic quantities.

In solution similar measurements on concentration dependencies can also provide information on energetics but in addition can be used to allow

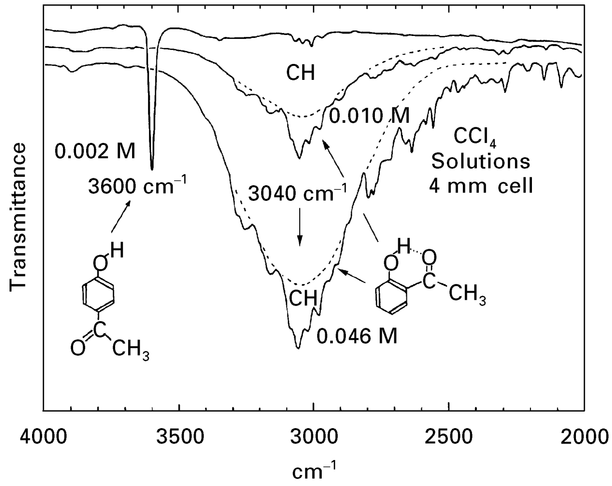



**Figure 3** The IR spectra of hydroxyacetophenones in $CCl_4$ solution. Reproduced with permission from Colthup NB, Daly LH and Wiberley SE (1975) *Introduction to Infrared and Raman Spectroscopy*. New York, San Francisco and London: Academic Press.

qualitative conclusions about structure to be made. For instance, **Figure 3** shows the IR spectra of *ortho* and *para* hydroxyacetophenones in solution. The *para* isomer in dilute solution is almost entirely free but the *ortho* isomer by contrast exhibits complete internal H-bonding which, as would be expected, is unaffected by change in concentration.

### Correlations with bond length

In crystals the stretching vibrating is observed to possess a sort of half-parabolic relationship with overall H-bond distance (i.e. length A–B). **Figure 4** shows data from a number of O–H---O systems. Note that data from other families (e.g. N–H---O)

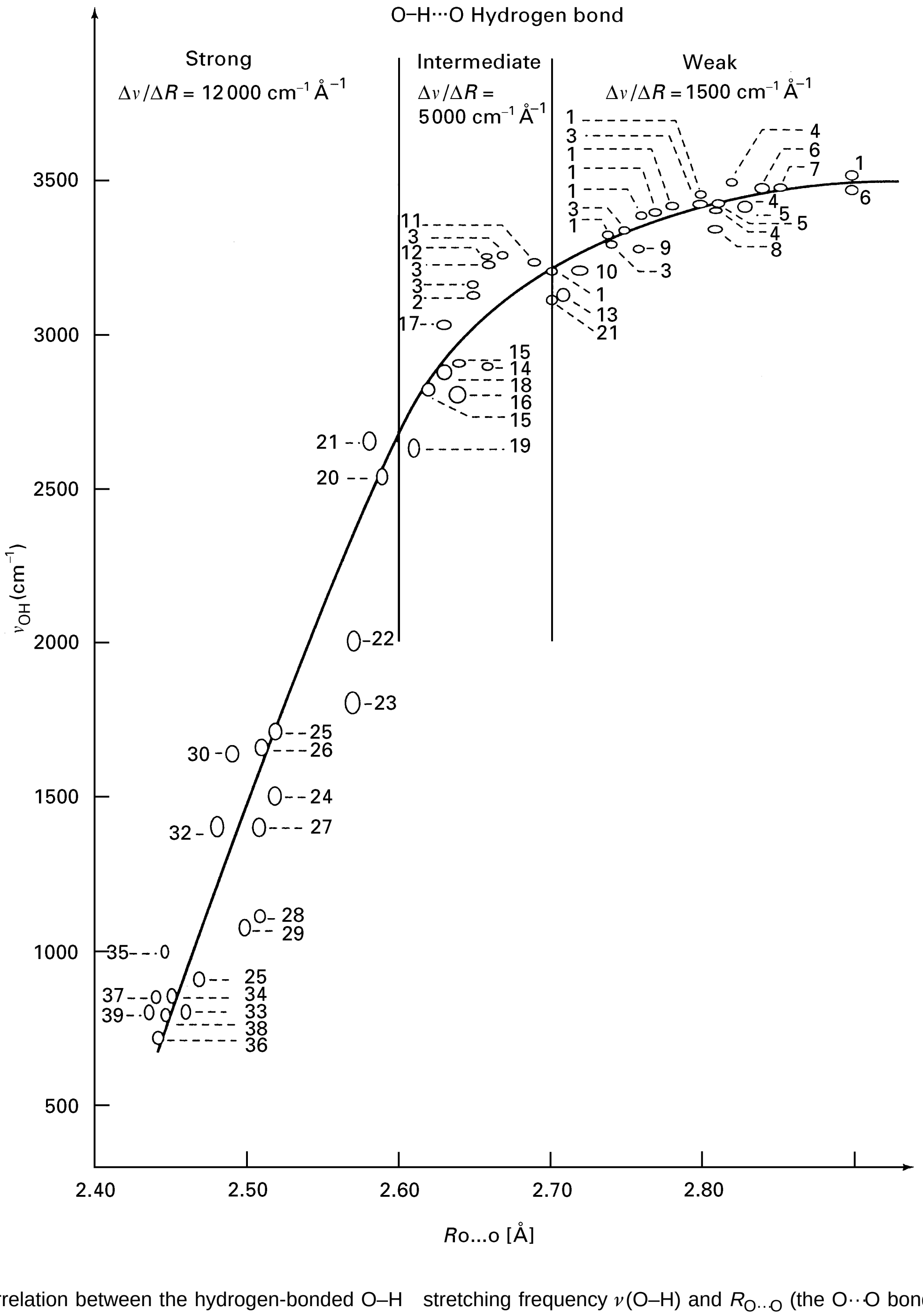



**Figure 4** Correlation between the hydrogen-bonded O–H stretching frequency $\nu$(O–H) and $R_{O\cdots O}$ (the O···O bond distance) for O–H---O H-bonds in crystals. The weakly bonded examples are mostly salt hydrates, the strongest examples (those with frequencies below 1000 cm$^{-1}$) are in the main organic singly-ionized diacid anions (oxygen analogues of bifluoride ion). Reproduced with permission of Springer-Verlag from Novak A (1974) Hydrogen bonding in solids: correlation of spectroscopic and crystallographic data. *Structure and Bonding* **18**: 177–216.

fall along separate curves. The sum of the van der Waals radii for two oxygens is about 2.8 Å.

## Correlations with acidity/basicity and energy of interaction

The strength of H-bonding increases with greater proton-donating ability of the A–H moiety and increasing proton-accepting power of the base B. **Figure 5** shows the effect of differing basicity on isopropanol at constant concentration. Here the solvent itself plays the role of the base. In benzene (inert) the hydroxyl groups are mostly free (the weak sideband indicates a small amount of OH to OH bonding) but are almost entirely H-bonded to either tetrahydrofuran, pyridine or triethylamine. The shift $\Delta\nu$(A–H) is proportional to solvent p*K*a, the values being −2.1, 5.2 and 11 respectively.

The Badger–Bauer rule states that there is a linear relationship (in solution) between the shift $\Delta\nu$(A–H) and $\Delta H$, the enthalpy of H-bond formation, but it was formulated originally from early and somewhat limited data. More comprehensive work has since shown that there are deviations from linearity, most obviously in the region of weaker bonding. Each individual system again has a separate relationship, each system here being a one-proton donor (e.g. phenol) with a range of organic bases or vice-versa in a particular solvent.

The solvents utilized must of course be inert; $CCl_4$ or benzene are commonly used. Detailed work suggests that family types must be very exclusive to have a smooth relationship, for example a range of nitrogen bases must be all of the same electronic configuration and must be free of the complications of steric interference.

When a single carbonylic base is used, $\Delta\nu$(C=O) is found also to yield largely linear relationships with $\Delta H$. Other studies find quite reasonable linear relations between band shift and $\Delta G$ (the free energy difference derived from association constants) within specific families of donor and acceptor. While there is some theoretical justification for a linear relationship between shift and enthalpy (or more specifically with $\Delta E$, the internal energy change, in the gas phase) the linear relation with $\Delta G$ is only empirical as the latter cannot be directly related with $\Delta H$ (or $\Delta E$).

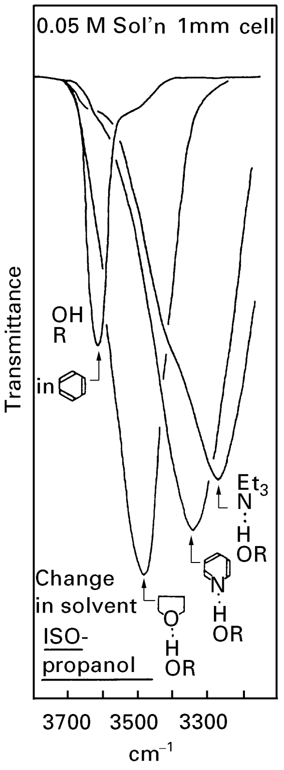



**Figure 5** Effect of changing the solvent on the O–H stretching vibration of isopropanol. Reproduced with permission from Colthup NB, Daly LH and Wiberley SE (1975) *Introduction to Infrared and Raman Spectroscopy.* New York, San Francisco and London: Academic Press.

## Estimation of thermodynamic quantities

IR absorbance spectroscopy obeys the Beer–Lambert law whereby band absorbance is linearly related to concentration of the molecule in solution. This allows the estimation of the association constant, and thereby ΔG, by measuring the decrease in the intensity of the free A–H band on formation of an H-bond along with knowledge of the total amounts of each substance present. The degree of association (whether dimer, trimer etc.) can be found by examining the way H-bond formation varies with change in concentration. The enthalpy can be estimated by measuring the association constant over a range of temperatures and employing the van't Hoff equation, subject to the usual caveats when using this relationship.

## Strong H-bonds

On the basis of the simple electrostatic picture a much stronger H-bond should result if either A–H or B are ionized. This is indeed found from crystallographic data, the bifluoride ion being a well known example. Many very short and strongly bonded O–H–O systems exist; **Figure 6A, B** shows the IR spectra of sodium hydrogen diacetate, which contains $[CH_3C\text{–}(=O)O\text{--}H\text{--}O(O=)CCH_3]^-$, and its singly deuterated equivalent. The O–O distance is slightly less than 2.45 Å and the proton is located equidistant (or very nearly so) from the two oxygens. It should be noted that the spectra are displayed here

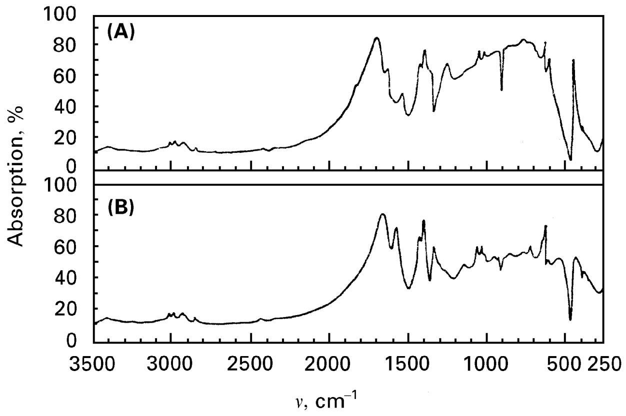



**Figure 6** The IR spectra of (A) $NaH(CH_3CO_2)_2$ and (B) $NaD(CH_3CO_2)_2$. Reproduced with permission from Hadzi D (1965) Infrared spectra of strongly hydrogen bonded systems. *Pure and Applied Chem*istry **11:** 435–453.

in absorbance so that the bands go upwards. Both $\nu$(OH) and $\nu$(OD) yield very broad bands starting at around 2500 cm$^{-1}$ and extending to below 500 cm$^{-1}$. The band maximum is at about 750 cm$^{-1}$ in the first case and somewhere around 600 cm$^{-1}$ in the other. Both spectra display a sharp nick (window) at about 900 cm$^{-1}$ and such features often abound in spectra of these types of compounds to give very bizarre appearances.

Isotope shifts (H$\rightarrow$D) in such compounds are often much lower than normal (which is slightly less than $\sqrt{2}$) though the jagged profiles of the bands can make estimation difficult. All the phenomena suggest that the potential surfaces for vibration are quite anharmonic.

Many medium-strength H-bonds also yield strange features in their stretching bands, carboxylic acids being a case in point.

### Explanations for bandwidth increase

Surprisingly there is as yet no accepted explanation for the phenomenon of band broadening in H-bonding. It may be that a number of different causes are responsible. A major factor is undoubtedly the abnormality of the vibrational potential functions which appear both from experiment and theory to be often considerably more anharmonic than usual.

Various suggestions for the likely effects of anharmonicity include very short lifetime of the excited vibrational state (uncertainty broadening) and breakdown of the Born–Oppenheimer approximation (i.e. motion of electrons and nuclei are no longer independent, which might be viewed as leading to a multiplicity of potential functions). Special factors may operate in the liquid state; in water the combination of distribution of bond lengths and high polarizability may be connected not only with broad bands but also with the presence of extensive continua that underly the spectrum. In solids, breakdown of selection rules for acoustic modes could contribute to breadth. Widespread interactions with other vibrational modes, including Fermi resonance, are responsible for the strange and broken profiles and can give the appearance of band plurality.

In some instances strong reflection from solids may artificially broaden IR absorbance spectra. This is certainly the case for the bifluoride ion.

## Non-specific forces in the liquid state

Compared to the vapour, band shifts similar in sign though smaller in general magnitude to H-bonding, even in its absence, are observed in liquids and solutions. Thus stretching vibrations are slightly shifted to lower wavenumber; for instance $\nu$(C=O) from $PhC(=O)CH_3$ is 1709 cm$^{-1}$ in the vapour but 1697 cm$^{-1}$ in hexane and 1692 cm$^{-1}$ in $CCl_4$. As a number of different effects are likely to influence the electronic structure and bond dynamics, quantitative predictions are difficult. Attempts to relate shifts to bulk properties only, such as dielectric constant, while ignoring nearest-neighbour interactions have therefore met with little success.

While many of these interactions are essentially very weak they nevertheless can affect vibrational spectra quite significantly. Solute–solvent interactions though only causing minor changes in band profile or position can severely interfere with quantitative analysis. This is because the spectrum of the pure solvent will differ from when it is involved in solution, preventing accurate subtraction if solvent and solute bands overlie each other.

Solute–solvent interactions can occasionally be put to good use, as in the spectral resolution of enantiomeric pairs. By themselves monostereoisomers have identical spectra but when dissolved in an optically active solvent such as D- (or L-)2-octanol (available commercially) yield slightly different spectra. This is because of the asymmetry of the respective interactions which effectively create diastereoisometric pairs.

### Collisional interactions

Collisions between molecules can cause distortion and alter the symmetry. Carbon disulfide is linear with a centre of symmetry (the two C=S bonds are equal in length) and thus the symmetric-stretching vibration at about 650 cm$^{-1}$ is forbidden in the IR. It is however observed as a weak band in IR spectra of liquid $CS_2$. Normally forbidden bands can also make

appearances when charge-transfer complexes are formed, iodine–pyridine being an example.

Collisions are also responsible for inducing transient polarizations in non-polar molecules which result in a characteristic broad absorbance in the liquid in the range 40–100 cm$^{-1}$. Both geometric distortion and induced local dipole moments are involved.

## Matrix isolation

It is possible to control and distinguish many interactions by use of the matrix-isolation technique. This involves highly diluting a species of interest with an 'inert' and IR-transparent gas, such as argon or nitrogen, and condensing the mixture at very low temperature (liquid helium) onto an IR-transparent window (or a Raman cell). By varying the dilution it is possible to isolate single molecules or controlled associations of molecules in 'cages' of the matrix gas. Reactive species can be trapped and studied at leisure. The very low temperature solid phase has the advantage of eliminating translations and collisions yet also yielding generally very sharp bands which can aid in distinguishing mixed species.

**Figure 7** shows IR spectra of methanol in various phases including an argon matrix. The bands observed are all from O–H stretching and demonstrate the advantage of the matrix technique. The vapour-phase spectra are dominated by rotational side bands; only H-bonded multimer can be seen in pure liquid and solid, though solution shows dimer as well. Spectra from the matrix at two different dilutions can however distinguish dimer, trimer, tetramer and multimer.

Simple molecular aggregates are amenable to detailed *ab initio* molecular orbital calculations of vibrational modes, which can be compared to experiment to resolve structural dilemmas. For instance theory predicts that the linear-asymmetric $H_2O$ dimer is more stable than the cyclic-symmetric alternative. Consistent with theory, the low-temperature matrix spectrum can only be fitted to the linear structure.

## Inelastic neutron scattering

This technique complements IR and Raman spectroscopy in the study of H-bonding in the solid state.

## List of symbols

$R$ = O···O bond distance; $\Delta E$ = internal-energy change; $\Delta G$ = free-energy difference; $\Delta H$ = enthalpy of bond formation; $\Delta \nu$ = O–H stretching frequency.

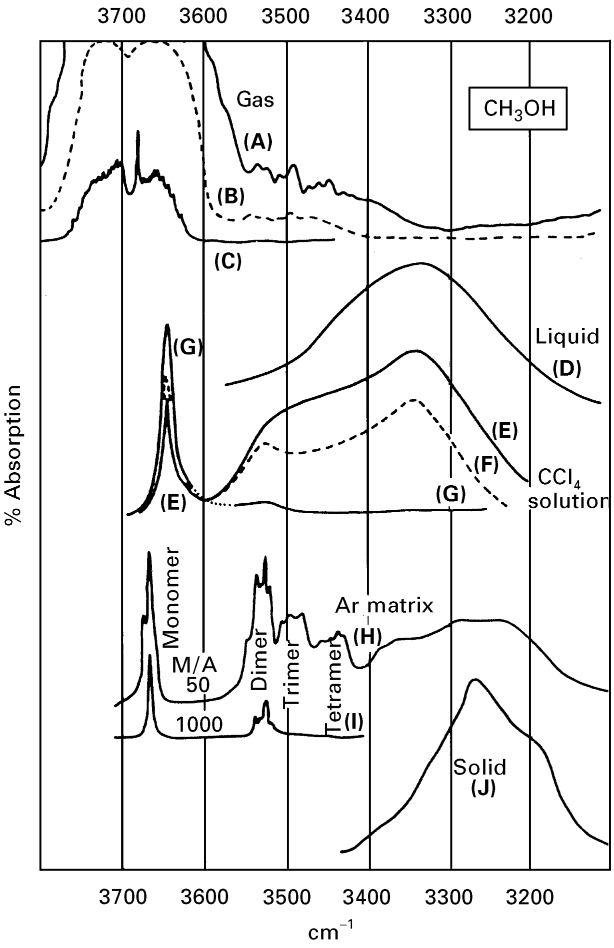



**Figure 7** IR spectra of methanol in various phases. (A)–(C) gas; (D) pure liquid; (E)–(G) decreasing concentration in $CCl_4$ solution; (H)–(I) decreasing concentration in Ar matrix; (J) pure solid. Reproduced with permission from Hallam HE (1973) *Vibrational Spectroscopy of Trapped Species*. Wiley.

*See also:* **IR Spectrometers; IR Spectroscopy Sample Preparation Methods; IR Spectroscopy, Theory.**

## Further reading

Abraham MH, Duce PP, Prior DV, Barratt DG, Morris JJ and Taylor PJ (1989) Hydrogen bonding: solute proton donor and proton acceptor scales for use in drug design. *Journal of the Chemical Society, Perkin Transactions II*, 1355–1375.

Bellamy LJ (1968) *Advances in Infrared Group Frequencies*. London: Methuen.

Clark RGH and Hester RE (eds) (1989) Spectroscopy of matrix isolated species. In: *Advances in Spectroscopy 17*. London and New York: Wiley.

Frohlich H (1993) Calculating the degree of association, equilibrium constant and bond energy for hydrogen bonding in benzyl alcohol and phenol. *Journal of Chemical Education* **70**: A3–A6.

Hallam HE (1973) *Vibrational Spectroscopy of Trapped Species*. London and New York: Wiley.
Hadzi D (ed) (1997) *Theoretical Treatments of Hydrogen Bonding*. London and New York: Wiley.
Jeffrey GA (1997) *An Introduction to Hydrogen Bonding*. Oxford University Press.
Kamlet MJ, Solomonovici A and Taft RW (1979) Linear solvation energy relationships, 5: correlations between infrared $\Delta\nu$ values and the $\beta$ scale of hydrogen-bond acceptor basicities. *Journal of the American Chemical Society* **101**: 3734–3739.
Pimentel GC and McClellan AL (1960) *The Hydrogen Bond*. San Francisco: Freeman.
Rao CNR, Dwivedi PC, Ratajczak H and Orville-Thomas WJ (1975) Relation between O–H stretching frequency and hydrogen bond energy: re-examination of the Badger–Bauer rule. *Journal of the Chemical Society, Faraday Transactions II* **71**: 955–966.
Scheiner S (1997) *Hydrogen Bonding: A Theoretical Perspective*. Oxford University Press.
Tomkinson J (1992) The vibrations of hydrogen bonds. *Spectrochimica Acta* **48A**: 329–348.

# Hyphenated Techniques, Applications of in Mass Spectrometry

**WMA Niessen**, hyphen MassSpec Consultancy, Leiden, The Netherlands

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| **Applications** |

## Introduction

One of the most fascinating fields of instrumental development in mass spectrometry (MS) is hyphenation: the on-line coupling of various techniques to MS. Apart from the obvious combinations of straightforward coupling of gas chromatography (GC) or liquid chromatography (LC) to MS, a wide variety of other combinations has been described. This contribution pays attention to the rationale of hyphenated techniques and briefly indicates a number of typical applications.

The key idea behind hyphenation is the significant gain in signal-to-noise ratio, and thus improvement in detection limit, that can be achieved by multidimensional methods. This concept was nicely pictured by Yost and co-workers, in the early 1980s (**Figure 1**). While in hyphenation the response or signal achieved decreases with the increasing number of couplings or dimensions, the (chemical) noise decreases even faster due to the increased selectivity, thus resulting in an improved signal-to-noise ratio. Another important issue is the ability to avoid sample losses and sample contamination in on-line rather than off-line combinations.

In principle, one can discriminate between various approaches to hyphenation, depending on the primary objective:

(a) On-line separation technique coupled to MS, e.g. GC-MS and LC-MS.
(b) On-line sample pretreatment in combination with a separation technique coupled to MS.
(c) On-line multidimensional separation techniques coupled to MS.
(d) On-line coupling of a separation technique to multiple detection strategies, one of which is MS.

Obviously, a combination of these strategies is also possible. Examples of these approaches are given below.

## Tandem mass spectrometry

The first hyphenated approach to be considered is the on-line combination of MS and MS, i.e. tandem mass spectrometry (MS-MS). A variety of combinations of different mass analysers have been described, including quadrupole and magnetic-sector analysers as $MS_1$, and quadrupole, magnetic-sector, ion-trap and time-of-flight analysers as $MS_2$. Instruments like triple-quadrupoles are widely used for MS-MS, either as stand-alone systems with sample introduction via a solids insertion probe or flow-injection analysis, or in on-line combination with GC or LC. The work of Yost and co-workers and of Hunt and colleagues exemplify these methods.

In the most common mode, i.e. the product-ion mode, $MS_1$, is used to select a precursor ion with a particular $m/z$ from the variety of ions generated in the ion source. The mass selected ions are dissociated via collisions with an inert gas in a collision cell, and

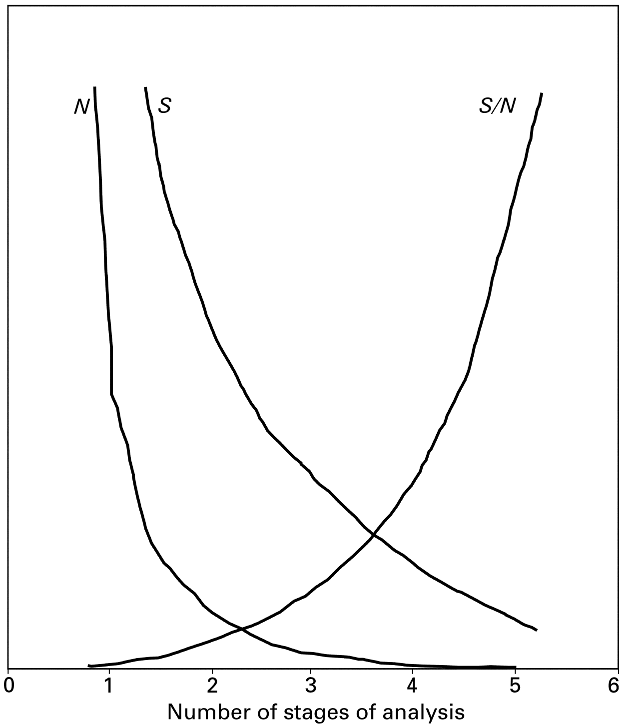



**Figure 1** Hyphenation of analytical techniques. Dependence of the signal *S*, the noise *N*, and the signal-to-noise (*S*/*N*) ratio on the number of analytical stages.

the product-ions are subsequently mass-analysed by $MS_2$. In this setup, $MS_1$ can be considered as a separation technique, while $MS_2$ is operated as a conventional mass analyser and detector. Analogies between MS-MS and GC-MS have frequently been drawn (**Figure 2**). It has been argued whether a chromatographic separation is actually still required in a combination with MS-MS, given the excellent selectivity that can be achieved. However, it is currently generally agreed that some separation is required, at least in the analysis of samples of biological or environmental origin, in order to avoid rapid contamination of the ion source and in order to reduce and/or avoid analyte ion suppression effects, i.e. in electrospray ionization.

MS-MS is currently very widely used in combination with chromatographic separation methods, especially LC. The obvious reason for this is the frequent use of soft ionization techniques in LC-MS interfacing, i.e. electrospray and atmospheric-pressure chemical ionization. MS-MS allows additional structural information as well as the molecular mass information to be obtained. On-line LC-MS-MS is currently the method-of-choice in quantitative bioanalysis in (pre-)clinical pharmacological studies during drug development in pharmaceutical industries. In these studies, the instrument is operated in selective-reaction monitoring (SRM) mode, selecting a particular precursor ion in $MS_1$ and selecting and detecting one or more product ions in $MS_2$. The excellent selectivity, the high level of confidence of identity, and the ability to use stable isotopically labelled internal standards are important arguments for the use of SRM in this type of routine applications.

## On-line chromatography–mass spectrometry

The on-line combination of a chromatographic separation technique (GC, LC, but also thin-layer chromatography (TLC), capillary zone electrophoresis (CZE) and supercritical fluid chromatography (SFC)) with MS enables the mass spectrometric characterization of components in complex mixtures after separation with minimal or no sample loss. It is especially useful in the identification of minor or trace components that are difficult to collect by fractionation of the column effluent or would be easily lost. Furthermore, as already indicated above, on-line GC-MS and LC-MS-MS are important tools in quantitative bioanalysis as well. Obviously, fractionation in large series of samples for routine quantitative applications would be extremely time-consuming and ineffective.

GC-MS plays an important role in many application areas, including the characterization of components in petroleum and derived products and in essential oils, the identification and quantitation of compounds of environmental interest such as polychlorodibenzodioxins and related compounds, polycyclic aromatic hydrocarbons, pesticides and herbicides, and a variety of other microcontaminants. In addition, GC-MS is important in the analysis of compounds of pharmacological, forensic and/or toxicological interest, including drugs, anaesthetics, steroids, growth hormones and drugs of abuse.

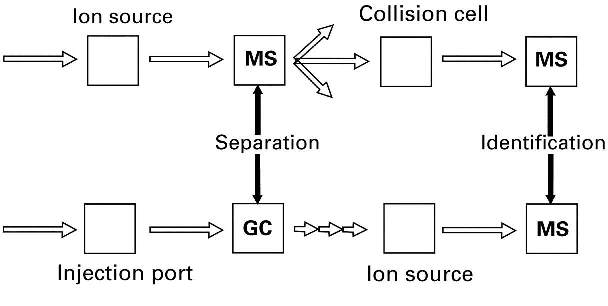



**Figure 2** Schematic comparison of MS-MS and GC-MS.

LC-MS is applied in complementary fields, where the analytes are not amenable to GC-MS. LC-MS is for instance applied in the identification and quantitation of pesticides, herbicides, surfactants and (sulfonated) azo dyes in environmental samples, in the identification of drugs during drug development, their degradation products and metabolites, in the quantitative bioanalysis of drugs and related compounds in biological tissues and fluids, in the characterization of natural products, such as alkaloids, taxoids, toxins, as well as endogenous compounds like acylcarnitines, prostaglandins, bile acids. Furthermore, LC-MS plays an important role in biochemical and biotechnological applications of MS, via the electrospray MS and electrospray LC-MS analysis of biomacromolecules like peptides, proteins and DNA fragments.

Methods and some applications of GC-MS and LC-MS are discussed in a separate contribution.

The on-line combination of TLC and MS via FAB, liquid SIMS, or MALDI has also been frequently described. The TLC-MS combination, recently reviewed by Somsen and co-workers, is applied for a wide variety of compounds, including drugs and their metabolites, antibiotics, steroids, alkaloids, lipids, bile acids, porphyrins, dyes and peptides.

## On-line sample pretreatment

With the advent of LC-MS technologies, the on-line combination with various sample pretreatment strategies received considerable attention. The rationale for on-line sample pretreatment is to avoid sample losses and sample contamination during the off-line transfer from one step of the analytical procedure to another. In addition, an on-line procedure greatly facilitates automation of the complete procedure, thereby speeding up the analysis. The most successful and most widely applied approach is on-line solid-phase extraction (SPE) in combination with LC-MS. A number of instruments have been developed and commercialized for this combination, e.g. the Varian AASP, the Gilson ASPEC, the Spark Holland Prospekt, and the Merck OSP-2. In all these automated systems, the sample constituents within a certain polarity range are trapped onto a short cartridge column (or Empore disk) containing reversed-phase LC packing material. The cartridge or disk may be washed with water to remove hydrophilic sample constituents. Subsequently, the analytes of interest are desorbed and transferred to an LC column for separation. These on-line SPE strategies serve for both sample pretreatment and analyte preconcentration. While on-line SPE-LC-MS was initially mainly developed for quantitative bioanalysis, the most important current application is in environmental analysis, i.e. in the selective sample pretreatment and analyte preconcentration of pesticides, their degradation products as well as other microcontaminants from surface water as an on-line part of the SPE-LC-MS or SPE-LC-MS-MS analysis. The results of the group of Brinkman show that analyte preconcentration by factors up to $10^3$ or $10^4$ are feasible in pesticide analysis from surface water, enabling concentration detection limits well below 0.1 $\mu g L^{-1}$.

Consecutively, the same group demonstrated that SPE can also be applied in an on-line combination with GC-MS. An example is the SPE-GC-MS analysis of 10 mL of river Rhine water spiked at the 0.5 $\mu g L^{-1}$ level with 80 microcontaminants, such as chlorobenzenes, aromatic compounds, anilines, phenols and organonitrogen and organophosphorus pesticides.

In addition to SPE, other sample pretreatment methods have been combined with GC-MS (see Goosens and co-workers), LC-MS or to MS(-MS) directly, e.g. on-line membrane sample introduction, solid-phase micro-extraction (SPME), supercritical fluid extraction and membrane dialysis.

A variety of on-line sample pretreatment procedures have been described for the coupling to on-line capillary zone electrophoresis (CZE)-MS, including capillary isotachophoresis, electrodialysis, liquid–liquid electroextraction and SPE on Empore disks. The latter system was applied to the analysis of the neuroleptic drug haloperidol in patients' urine by on-line Empore-disk SPE-CZE-MS.

## Multidimensional separation techniques

Multidimensional separation techniques have been developed with a number of objectives such as increased peak capacity and/or improved resolution for the separation of highly complex samples, shortened analysis time via heartcut and partial analysis of fractions from complex samples, and enhanced detection of trace components. Both GC-GC, LC-GC, and LC-LC methods as well as various other multidimensional combinations involving CZE or SFC have been described and coupled to mass spectrometry. In most cases, the multidimensional approach comprises a combination of two chromatographic columns, either containing two different stationary phases (both GC and LC), or developed with two different mobile phases (LC only). The sample is injected onto the first column. Part of the chromatogram is heart-

cut, either sampled onto a retention gap or short trapping column or transferred directly, and subsequently analysed on the second column, which is then interfaced to an MS. A typical setup for LC-LC-MS is shown in **Figure 3**.

Several examples of on-line multidimensional GC-MS were reviewed by Ragunathan and co-workers including the characterization of essential oils in order to determine the composition, to identify or quantify enantiomers, or to study the chemotaxonomy of the oil, and the determination of coplanar polychlorobiphenyls congeners.

The on-line combination of LC-GC-MS is applied in a variety of analytical problems, such as the analysis of (trace amounts of) aromatic compounds in complex fossil fuel fractions or in vegetable oil, the identification of microcontaminants in surface water, and the identification of impurities in pharmaceutical products.

The on-line combination of LC-LC-MS has been investigated for a number of reasons. In addition to the general prospects of multidimensional separation techniques, especially enhanced selectivity, there was special interest in the ability to perform LC-LC with two different mobile-phase compositions. In this way, it should be possible to avoid problems with mobile-phase incompatibility due to the use of non-volatile mobile-phase constituents. A good example of this approach is the determination of enantiomers of β-blockers in plasma samples, described by Edholm and co-workers. Racemic mixtures of a β-blocker like metoprolol can be separated on a $\alpha_1$-acid glycoprotein column. However, the chromatography requires the use of a 20 mM phosphate buffer (pH 7) in the mobile phase, which is not compatible with on-line LC-MS. Therefore, the chiral column was coupled via a set of two trapping columns to a common reversed-phase LC column. After separation, the two enantiomers were separately trapped onto two short columns and subsequently transferred to the second LC column for analysis with a mobile phase containing ammonium acetate, which is well compatible with the thermospray LC-MS interfacing applied in this study. Subsequently, it was demonstrated that by the use of MS-MS the second separation step could be avoided: the effluent from the trapping column can be directly introduced into the MS-MS system, operated in SRM mode. This phase-system switching approach is well suited for solving this type of mobile-phase incompatibility problem.

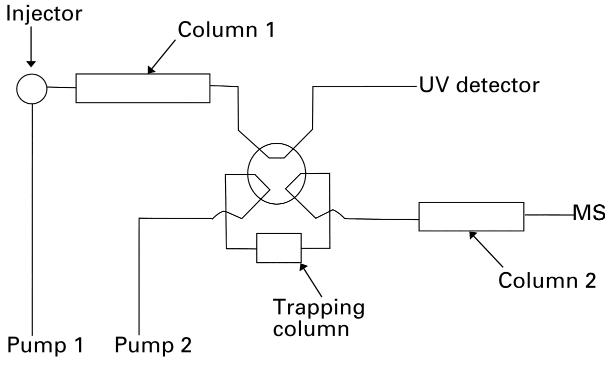



**Figure 3** Schematic diagram for a coupled-column LC-LC-MS system.

## Multiple detection strategies

The increasingly more difficult analytical problems to be solved is an important impetus in all developments in hyphenated techniques, but especially in multidimensional detection strategies. Development of adequate separation techniques for the analysis of complex samples is often a difficult task, independent of the question whether one is interested in only one compound present at trace level, or in characterization and structure elucidation of most sample components. Multidimensional separation techniques were developed in order to achieve the separation of the components of very complex samples. MS is an extremely powerful technique for the identification and structure elucidation of unknowns. The sensitivity of MS as a fairly universal technique in this area is unsurpassed. However, in many instances the information obtained from MS is not sufficient to enable unambiguous identification of the unknown compounds. Furthermore, in certain application areas it is obligatory to identify the components with such a high level of confidence that independent results from a second technique in addition to MS are required, as for instance for impurity screening of pharmaceutical products. Improved levels of confidence in identification and structure elucidation can be achieved by the use of multiple detection strategies. Although detection with a variety of detectors can be achieved via multiple analysis of the sample on systems equipped with different detectors, there has been a growing interest and need for systems which enable the use of multiple detectors coupled to a single analytical separation system. This is especially important in two areas, i.e. in the analysis of very complex samples where multiple detection strategies coupled to one single data-acquisition and processing system can greatly facilitate the interpretation of the data, and in the high throughput analysis where minimization of the analysis time is of utmost importance.

Several dual or multiple detection systems involving MS have been described. In this respect, one should not forget that the mass spectrometer is a destructive detector. Therefore, the mass spectrometer should be the final detector in a series. Combination with another destructive detector, e.g. a flame ionization detector, is possible only after splitting the column effluent, while in-line coupling with a nondestructive detector, e.g. a UV absorbance detector, is only possible when this detector does not result in too much chromatographic peak broadening. In several cases, the second nondestructive detector is used in parallel rather than in series, e.g. because the sensitivities of both detectors are widely different, or in order to avoid the loss of chromatographic resolution due to excessive peak broadening by the first detector in the series. The most widely applied dual detector strategies are those where MS is combined with another spectrometric technique, in order to obtain complementary structural information.

Two types of spectrometric detectors have been widely used in an in-line combination with GC-MS, i.e. Fourier-transform infrared (FT-IR) and atomic emission spectrometry (AED). Both FT-IR and AED are used in parallel with the MS, i.e. after a split.

GC-FT-IR-MS systems are commercially available. When the measured IR spectrum of an unknown is not present in the IR spectral library, the FT-IR part of the system can help in the identification of an unknown by providing information on functional groups and the structural backbone. While homologues cannot readily be identified from IR spectra but easily discriminated by MS, the FT-IR part can be used as an important tool in discriminating between structural isomers where MS generally fails. Typical applications of GC-FT-IR-MS can be found in the analysis of essential oils, flavours and fragrances, and in the identification of isomeric reaction products or other types of isomeric compounds. The potential of a dual detection GC-FT-IR-MS system was also explored for the identification of illicit drugs, e.g. in tracing banned stimulants used by athletes.

The combination of GC-AED-MS to our knowledge is not (yet) commercially available. The element specificity of the AED can be of great help in identifying particular classes of compounds in complex samples, e.g. in environmental screening. The AED provides information on the presence of certain elements, e.g. halogens, nitrogen, or phosphorus, in the chromatographic peaks detected. This facilitates the search for relevant peaks as well as the interpretation of the mass spectra of particular compounds in a complex chromatogram. In addition, the AED can provide a fairly accurate determination of the elemental composition, thereby often excluding the need for accurate mass determination by expensive magnetic sector instruments. Furthermore, GC-AED-MS can be applied in the analysis of organometallic compounds, e.g. organomercury and organotin compounds.

The availability of nondestructive flow-through detectors for LC, e.g. UV absorbance and fluorescence detectors, readily leads to the application of dual detection systems in LC-MS. While the additional information of single-wavelength UV detection is generally limited, the availability of efficient and sufficiently sensitive UV-photodiode array detectors enables the on-line acquisition of UV spectra as well as mass spectra (**Figure 4**). Although the structural information available from a UV spectrum is generally limited, especially in a relatively undefined solvent mixture during gradient elution, the on-line combination can successfully be used in facilitating identification when appropriate UV spectral libraries are available. The combination of retention time in a well defined LC system, the match of a UV spectrum, and the molecular mass (and possible structural information) from MS are especially useful for confirmation of identity, e.g. in multiresidue screening for pesticides in environmental samples or antibiotics in food products. Furthermore, an excellent feature of UV-PDA systems in combination with appropriate software is peak purity assessment, based on changes of the UV spectra over the chromatographic peak. This feature may be helpful in impurity profiling by LC-UV-PDA-MS.

Although the use of FT-IR in combination with LC is of growing importance, the on-line or in-line combination of FT-IR and MS is not yet frequently applied. However, the on-line combination of LC-MS and NMR is currently introduced as an important tool in identification of impurities, degradation products, and/or metabolites of pharmaceutical products. A recent example of on-line LC-NMR-MS is the characterization of ibuprofen metabolites in human urine. Simultaneous acquisition of electrospray MS and NMR spectra of synthetic drug products in an ‘open-access’ setting has also been reported.

## Final considerations and perspectives

Hyphenation in relation to MS is an important research topic because in many instances the optimum tuning of the two or more parts of the setup requires

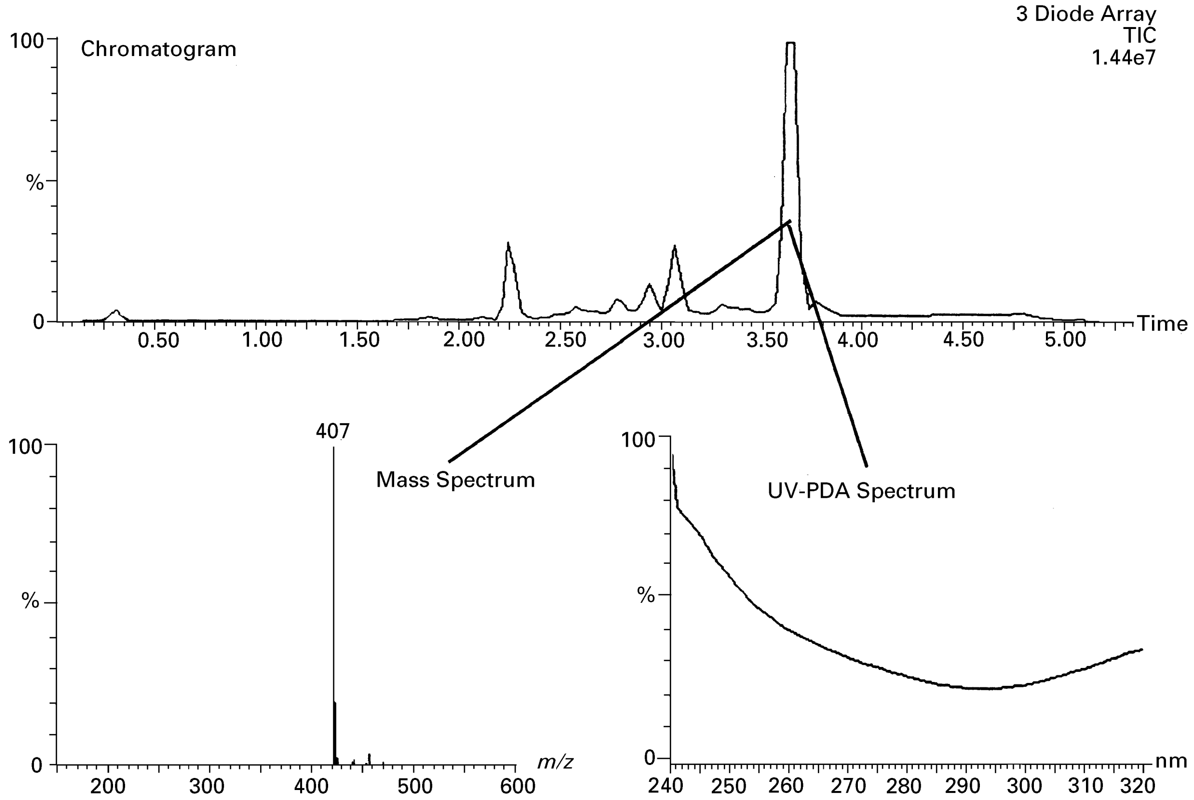



**Figure 4** An example of typical results from an LC-UV-PDA-MS setup for dual detection of separated compounds.

considerable attention. Good knowledge of both parts is required in order to avoid the often inevitable compatibility problems and to obtain best possible results. However, hyphenated MS techniques have already widely proved their value in solving real analytical problems.

In principle, the use of advanced data-processing software could be considered as a hyphenated approach as well, but in general is not. However, one can question whether the use of principal component analysis and similar data-processing techniques does not have a similar impact on the analytical procedure and its result as the hardware hyphenation discussed above.

Software for both controlling the hyphenated combination of techniques and efficient processing of the data is of utmost importance to the breakthrough and general use of a hyphenated technique. In this respect, the development of automatic tuning and calibration software for GC-MS and more recently LC-MS, of computer spectral library searching after electron ionization, and of efficient quantitation software packages after both GC-MS and LC-MS can be considered as important steps in the development of hyphenated techniques. Control of the complete (often multivendor) instrumentation, including the GC or LC chromatograph, the sample processor or autosampler, on-line or off-line second (scanning) detectors from within a single software platform, often the MS control and acquisition software, is also essential and obligatory for the success of the technique. Some instruments enable high levels of automation and control via the use of macro scripts or instrument control languages. Such approaches may be used for optimization and/or control of the various steps in the hyphenated technique, but also for artificial-intelligence type of applications where the software makes decisions concerning the analytical strategy to follow or the type of experiments to perform on the basis of the data acquired. The latter is, for instance, applied in data-dependent product-ion scanning during GC-MS-MS and LC-MS-MS.

Progress is also made in software for mass spectral interpretation, especially for the so-called deconvolution of (mixed) ion envelopes of multiply charged protein ions generated by electrospray ionization, and for the peptide sequencing by MS-MS techniques. In that respect, the development of tools to search extensive protein and DNA databases using peptide maps, peptide molecular masses, LC retention time data, and sequence tags is an enormous

step forward in the routine application of hyphenated MS-MS.

Excellent examples of the hybrid of hyphenated hardware, automated sample processing via computer control of the instrumentation, and automated data processing can be found in approaches recently developed in relation to combinatorial chemistry. In order to screen a particular combinatorial library in a 96-well plate, high throughput electrospray LC-MS is applied. The data are postprocessed by means of a browser, which compares the molecular mass measured with expected values, taking into account the various cationized and anionized species that may be generated during electrospray ionization. By means of green and red colours, the software indicates on a screen representation of the 96-well plate which samples were found to be present and which were not. In addition to this, and using similar software strategies, the use of automated fraction collection after preparative LC under control of the data acquired by MS has been demonstrated. This hyphenated approach allows the purification of particular products from extensive combinatorial libraries based on the results of a first biological screening and prior to a more advanced biological screening, using the purified products.

*See also:* **Atmospheric Pressure ionization in Mass Spectrometry; Biochemical Applications of Mass Spectrometry; Chemical Structure Information from Mass Spectrometry; Chromatography-MS, Methods; Isotopic Labelling in Mass Spectrometry; Medical Applications of Mass Spectrometry; MS-MS and MS$^n$.**

## Further reading

Brotherton HO and Yost RA (1984) Rapid screening and confirmation for drugs and metabolites in racing animals by tandem mass spectrometry. *American Journal of Veterinary Research* **45**: 2436.

Edholm L-E, Lindberg C, Paulson J and Walhagen A (1998) Determination of drug enantiomers in biological samples by coupled column liquid chromatography and liquid chromatography–mass spectrometry. *Journal of Chromatography* **424**: 61.

Hankemeier Th, Van Leeuwen SPJ, Vreuls RJJ and Brinkman UATh (1998) Use of a presolvent to include volatile organic analytes in the application range of on-line solid-phase extraction–gas chromatography–mass spectrometry. *Journal of Chromatography A* **811**: 117.

Hogenboom AC, Speksnijder P, Vreeken RJ, Niessen WMA and Brinkman UATh (1997) Rapid target analysis of microcontaminants in water by on-line single-short-column liquid chromatography combined with atmospheric-pressure chemical ionization tandem mass spectrometry. *Journal of Chromatography A* **777**: 81.

Hunt DF, Shabanowitz J, Harvey TM and Coates ML (1983) Analysis of organics in the environment by functional group using a triple quadrupole mass spectrometer. *Journal of Chromatography* **271**: 93.

Kitson FG, Larsen BS and McEwen CN (1996) *Gas Chromatography and Mass Spectrometry. A Practical Guide*. London: Academic Press.

Niessen WMA (1998) *Liquid Chromatography–Mass Spectrometry*, 2nd edn. New York: Marcel Dekker.

Ragunathan N, Krock KA, Klawun C, Sasaki TA and Wilkins CL (1995) Multispectral detection for gas chromatography. *Journal of Chromatography A* **703**: 335.

Slobodnik J, Hogenboom AC, Vreuls JJ *et al* (1996) Trace-level determination of pesticide residues using on-line solid-phase extraction–column liquid chromatography with atmospheric pressure ionization mass spectrometric and tandem mass spectrometric detection. *Journal of Chromatography A* **741**: 59.

Somsen GW, Morden W and Wilson ID (1995) Planar chromatography coupled with spectroscopic techniques. *Journal of Chromatography A* **703**: 613.

Vreuls JJ, Bulterman A-J, Ghijsen RT and Brinkman UATh (1992) On-line preconcentration of aqueous samples for gas chromatographic–mass spectrometric analysis. *Analyst* **117**: 1701.

Walhagen A, Edholm L-E, Heeremans CEM *et al* (1989) Coupled-column chromatography–mass spectrometry. Thermospray liquid chromatographic–mass spectrometric and liquid chromatographic–tandem mass spectrometric analysis of metoprolol enantiomers in plasma using phase-system switching. *Journal of Chromatography* **474**: 257.

# I

# *In Vivo* NMR, Applications, $^{31}$P

**Ruth M Dixon** and **Peter Styles**, University of Oxford, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

## Introduction

The phosphorus nucleus has occupied centre stage in the development of *in vivo* NMR spectroscopy for two important reasons. First, the NMR properties of the nucleus are well suited to *in vivo* NMR (100% natural abundance, reasonable sensitivity and adequate but not excessive chemical shift). Second, key metabolites, particularly those involved in the production of energy, contain phosphorus and are present in substantial concentrations in living tissue. It is for these reasons that the first NMR spectra of living cells (a suspension of erythrocytes), and later of excised frog muscle, were obtained using $^{31}$P NMR. In both instances, resonances were observed from adenosine triphosphate and inorganic phosphate, the position of the latter being an indicator of intracellular pH. Additional peaks were seen from 2,3-diphosphoglycerate (2,3-DPG) in red blood cells and creatine phosphate in skeletal muscle. Despite the comparative simplicity of $^{31}$P NMR tissue spectra, the technique offers unique insight into the regulation of critical energetic pathways. The early studies of *ex vivo* samples were soon extended to *in vivo* examination of intact animals and later to humans, experiments which were made possible by the development of large superconducting magnets, the use of local surface coil probes and the invention of a variety of techniques for localizing the source of the NMR signal.

## Techniques for $^{31}$P spectroscopy *in vivo*

### Signal acquisition and spatial localization

Perhaps the most critical step in the development of *in vivo* spectroscopy was the realization that it was not necessary to surround the 'sample' with a conventional NMR probe, but that a small local coil (a surface coil) could be placed against an animal to preferentially receive signals from the tissues adjacent to this coil. The surface coil has two advantages, namely excellent sensitivity close to the coil and inherent localizing properties owing to the rapid fall-off of the magnetic flux away from the coil. However, when deep tissues are to be investigated, additional localization is required. In choosing an appropriate technique for obtaining $^{31}$P NMR spectra from a live animal or man, careful consideration must be given to the spatial resolution and signal-to-noise ratio which can reasonably be expected. As the phosphorus nucleus has only 6% of the sensitivity of the proton, it follows that these expectations will be somewhat more modest than for $^{1}$H spectroscopy. The usual approach is to adopt the least demanding methodology appropriate for any particular application. As a consequence, in many instances simple surface coil localization is used to give adequate localization and provide the best signal-to-noise ratio in short acquisition times. When this simple approach is inadequate, a range of single and multiple voxel methods can be considered, but care must be taken to choose methods which work well at short echo times to avoid excessive $T_2$ relaxation and minimize problems due to the $J$ modulation of peaks such as adenosine triphosphate (ATP).

Dealing first with the single voxel approach, image selected *in vivo* spectroscopy (ISIS) is frequently the method of choice because the spins are kept predominantly along the $z$ axis, thus minimizing $T_2$ and $J$ coupling effects. Other localizing schemes (e.g. stimulated echo acquisition mode spectroscopy (STEAM) and point resolved spectroscopy (PRESS)) which are commonly used for proton spectroscopy can give reasonable results, but are more sensitive to the aforementioned problems.

Where multivoxel localization is advantageous, chemical shift imaging techniques can be applied in one, two or three dimensions, but relatively poor spatial resolution is obtainable and long acquisition times are usually required.

Finally, it will often be advantageous to implement hybrid protocols in which, for example, two dimensions might be defined by an ISIS sequence, and phase encoded chemical shift imaging (CSI) implemented along the third axis, thus producing a set of spectra that represent slices through a well-defined column of tissue.

### Proton decoupling

The quality of *in vivo* $^{31}$P spectra can be improved by the use of proton decoupling. Two benefits are available. Some of the phosphorus spectral lines are significantly broadened by scalar coupling to hydrogen. Irradiation of the protons during acquisition removes this interaction and sharpens the spectra. The effect is particularly advantageous in the phosphomonoester (PME) and phosphodiester (PDE) regions of the spectrum where conventional acquisition will fail to resolve, for example, glycerophosphocholine and glycerophosphoethanolamine. Together with the increased resolution, proton decoupling also increases the signal-to-noise ratio due to peak narrowing. A further increase in signal is available by low power decoupling during the interpulse relaxation period. Here, the nuclear Overhauser effect (NOE) can result in improved sensitivity of up to about 50% owing to interactions between the phosphorus nuclei and neighbouring protons.

### Magnetisation transfer

Under certain conditions, rates of chemical exchange can be measured using saturation- or inversion-transfer methods. These work by magnetically labelling one chemical species and observing the transfer of this magnetic signature to a second compound which is coupled by chemical exchange. For the technique to be viable, there needs to be significant exchange during a time scale which is of the order of the longitudinal relaxation time $T_1$. Such reactions include the synthesis and hydrolysis of ATP (by measuring exchange between ATP and inorganic phosphate) and the fast exchange catalysed by the creatine kinase enzyme.

## The *in vivo* $^{31}$P NMR spectrum

**Figure 1** shows spectra from human skeletal muscle, heart, liver and brain. Each contains the three peaks from ATP and one from inorganic phosphate. ATP shows three signals due to the different chemical environments of its three phosphorus atoms. These signals are further split by $^{31}$P–$^{31}$P *J*-coupling, but this splitting is not always detectable *in vivo* owing to the relatively broad lines. These peaks are normally designated ‘ATP’ but also contain minor contributions from other nucleotides, such as guanosine triphosphate, NADH, NADP, UDP-sugars and ADP. The signal which contains the least contamination from other compounds is the one from the central (beta) phosphate, which appears at the upfield end of the spectrum, and so this peak is generally used for the measurement of ATP levels. In addition, spectra from muscle, heart and brain contain a prominent peak from creatine phosphate. This compound is absent from liver. Liver and brain spectra also contain substantial peaks designated as phosphomonoester (PME) and phosphodiester (PDE). These regions of the spectra can contain metabolites involved in membrane metabolism and also phosphorylated sugars involved in glycolysis. An important feature of the $^{31}$P spectrum is that the position of the inorganic phosphate ($P_i$) peak is pH dependent, thus providing a sensitive marker of intracellular acidity/alkalinity.

### Biological relevance of the peaks

**Adenosine triphosphate (ATP)** ATP is the universal energy carrier in all cells. When energy is required for cell functions such as preserving ionic gradients, muscular contraction, etc., the terminal phosphate is cleaved by ATPase enzymes to form adenosine diphosphate (ADP) and inorganic phosphate, an

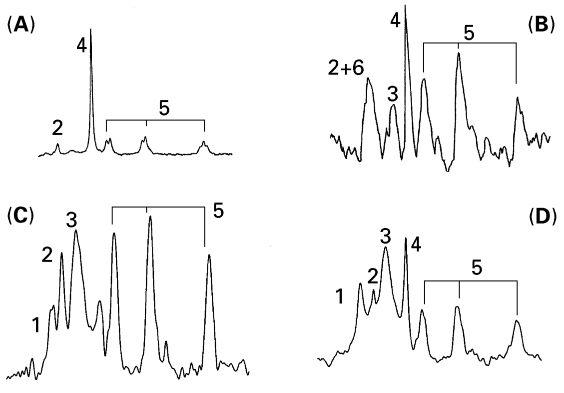



**Figure 1** $^{31}$P spectra of human organs (A) calf muscle, (B) heart, (C) liver, (D) brain. Peak assignments: 1. PME, 2. Pi, 3. PDE, 4. PCr, 5. ATP, 6. 2,3-DPG. Reproduced with permission of Academic Press from Fu, Schroeder, Jones, Santini and Gorenstein (1998) *Journal of Magnetic Resonance* **77**: 577–582.

energy producing reaction:

$$\mathrm{ATP + H_2O \rightarrow ADP + P_i}$$

ATP is resynthesized by one of several pathways, most usually by ATP synthesis using energy obtained from oxidative or anaerobic metabolism.

**Inorganic phosphate ($P_i$)** This compound may be considered as the waste product of ATP hydrolysis or conversely as a substrate for ATP synthesis. It may also play a role in regulating the supply of energy via oxidative metabolism in the cells' mitochondria. At physiological pH, $P_i$ exists in two ionic forms in fast equilibrium via the reaction:

$$\mathrm{H_2PO_4^- \rightleftharpoons HPO_4^{2-} + H^+}$$

This reaction has a p$K_a$ of 6.75, and so the relative concentrations of $P_i^-$ and $P_i^{2-}$ vary with the pH. Owing to the fast exchange between these two species, the position of the $P_i$ peak in the spectrum reflects the weighted average of the two chemical shifts, and is thus a direct measure of pH.

**Creatine phosphate (PCr)** This molecule is present in abundance in muscle, heart and brain cells. Through the fast equilibrium reaction catalysed by creatine kinase, PCr buffers the ATP concentration by phosphorylating ADP, the reaction being:

$$\mathrm{PCr + ADP + H^+ \rightleftharpoons ATP + Cr}$$

PCr is depleted during sudden energetic demands such as vigorous muscle contraction.

**Adenosine diphosphate (ADP)** ADP is a substrate for ATP production, and a product of its hydrolysis. ADP stimulates respiration (oxidative metabolism) – breakdown of ATP increases ADP which in turn activates the processes which provide energy for ATP synthesis. In tissues such as muscle, much of the ADP is bound to macromolecules and therefore invisible to NMR. However, the free (and therefore metabolically active) ADP usually present in only micromolar concentrations, may be deduced using the creatine kinase equilibrium (see above).

**Phosphomonoesters (PME)** Certain intermediaries in the glycolytic pathway such as glucose and fructose phosphates resonate in this region of the spectrum and are sometimes detected. More usually, the PME peak comprises the lipid precursors phosphocholine and phosphoethanolamine.

**Phosphodiester (PDE)** This region also contains compounds involved in membrane metabolism, specifically glycerophosphocholine (GPC) and glycerophosphoethanolamine (GPE).

**Broad resonances** In addition to the sharp peaks that have just been mentioned, there will often be underlying broad 'humps' in the phosphorus spectrum, arising from phosphorus moieties that are immobile or tumbling very slowly. Phospholipid head groups in cell membranes produce a broad and asymmetric signal centred under the PDE part of the spectrum. Bone also produces a broad resonance, several kilohertz wide, which is easily visible in surface coil spectra of the head, but is not usually present in localized spectra due to its very short $T_2$.

## Applications in skeletal muscle

Skeletal muscle is a tissue that is particularly suited to study by $^{31}$P NMR. From a technical standpoint, muscle is an ideal target because it exists close to the body surface and is not obscured by any other pool of phosphorus-containing metabolites. From a biochemical standpoint, muscle is especially interesting because its energetic demands vary by at least two orders of magnitude (from rest to exercise) and so it provides an example of exquisite metabolic control which can be studied non-invasively. Finally, muscle can be stressed readily by exercise within the NMR instrument, thus facilitating the real-time investigation of the dynamics of energetic regulation.

Animal studies of skeletal muscle often employ direct stimulation of the sciatic nerve to induce contraction and associated metabolic response. Basic metabolic control can be investigated using such a model, as can a wide range of disease-related processes. Genetically abnormal animals provide models of diseases such as muscular dystrophy and $^{31}$P NMR spectroscopy provides a tool for investigating metabolism, pH regulation and possible therapeutic interventions.

In a typical human examination, a subject will perform an exercise regime within the magnet, perhaps squeezing a handgrip or pressing down a loaded pedal. The NMR signals are usually collected using a surface coil placed over the exercising muscle and a series of spectra are collected to follow the metabolic state before, during and after the exercise. A typical time course is shown in **Figure 2**. The several metabolite concentrations and pH are then used to quantitatively assess the metabolic response to the stress, and, by using models of metabolic control, unique insight can be gained into the complementary

processes of mitochondrial respiration and anaerobic metabolism.

### Muscle metabolism

**Figure 3** is a block diagram of the main energy-producing pathways in skeletal muscle (and other tissues, except that the creatine kinase pathways may not be present). In resting muscle, ATP is synthesized only by oxidative metabolism, but during exercise glycolysis (anaerobic metabolism) and PCr breakdown via the creatine kinase reaction provide additional ATP production. Although only three compounds are directly detected by $^{31}$P NMR, others may be deduced by making assumptions based on data derived from biopsy analysis. In particular, an estimate of free [ADP] is obtained using assumed values of ATP and total creatine concentrations, and applying the known value for the creatine kinase equilibrium constant.

In using these data to investigate metabolic control *in vivo*, two periods within an exercise regime are of particular importance. At the start of exercise, energy is produced primarily by glycolysis and a breakdown of creatine phosphate. The latter is measured directly while the former produces lactate which is reflected in the pH. At the start of recovery, glycolysis stops, the energy demands fall dramatically and oxidative metabolism restores the PCr. Using models for proton efflux and metabolic control by ADP, it is possible to assess the relative contributions of the various energy pathways in both normal and diseased muscle.

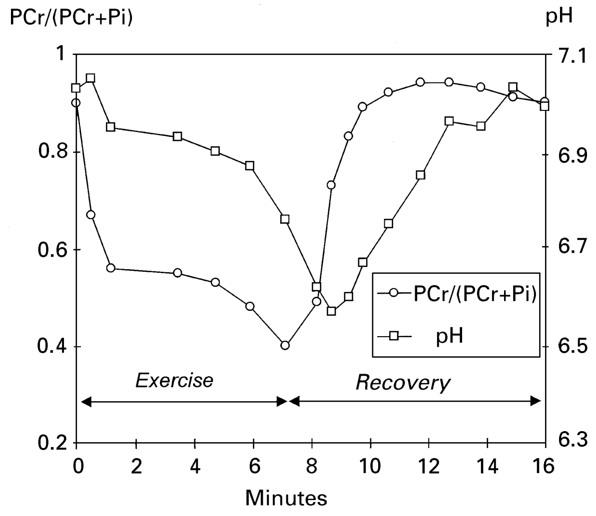



**Figure 2** A typical time course for changes in PCr and pH during exercise and recovery in the human forearm muscle.

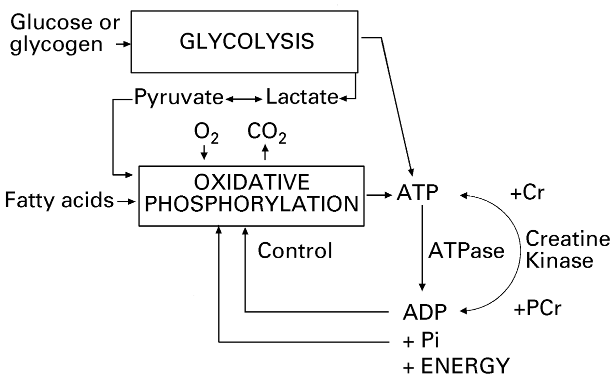



**Figure 3** A schematic diagram of energy metabolism in skeletal muscle, heart and brain. The creatine kinase pathway is absent in liver.

### Diseases of skeletal muscle

In man, the understanding and quantification of muscle energetics has facilitated the study and, sometimes, diagnosis of a wide range of muscle diseases. Often these stem from inherited genetic abnormalities, including the dystrophies, mitochondrial myopathies and so-called energy storage diseases. Muscle can also be affected by arterial disease and systemic abnormalities secondary to, for example, renal and cardiac failure. Many studies have looked at these phenomena and monitored the response to training or therapeutic intervention.

## Applications in the heart

Basic NMR research of the heart has concentrated on studies of the organ *ex vivo* (the excised heart, perfused with physiological buffer, can be maintained in a viable and beating condition for many hours). *In vivo* studies in both animals and man are more difficult – the heart is a moving target, is filled with blood and overlaid with skeletal muscle which gives a similar $^{31}$P NMR signature. The approaches that have been applied in animal systems often involve exposing the organ and suturing a surface coil to the heart wall, thus facilitating the measurement of localized metabolic events. In man, single voxel ISIS is sometimes used, or chemical shift imaging, usually as a hybrid experiment (see above) where spectra are obtained from parallel slabs of tissue. The biochemistry of the heart is in many ways similar to skeletal muscle. However, the dynamic range of energetic demand is less extreme than in muscle and in normal controls it is difficult to induce significant changes in the $^{31}$P spectrum from the heart. In the diseased heart, the main findings have been that cardiac energetics in the resting subject are

substantially stable until the onset of heart failure, when PCr falls significantly. In patients with coronary artery disease, metabolic levels can be modulated by exercise. As an alternative to exercise, the positive inotrope atropine-dobutamine has been administered to normal subjects, and at high cardiac output a reduction in the PCr:ATP ratio has been observed.

Intracellular pH is difficult to measure because the $P_i$ peaks can be obscured by 2,3-DPG in the blood. However, it is possible to overcome this problem using proton decoupling to sharpen the resonances, or saturation transfer to identify the intracellular $P_i$ unambiguously.

## Applications in brain

Using animal models, many disorders of the brain have been studied, including the effects of ischaemia (e.g. models of stroke and birth asphyxia), systemic disturbances (e.g. hepatic encephalopathy) and genetic abnormalities (e.g. models of Duchenne dystrophy). Spatial localization is usually kept to a minimum – often the scalp is removed to avoid contamination by signals from the cranial musculature, and a single surface coil is used to obtain the $^{31}$P NMR signal either from the whole brain or, by appropriate placing of the coil, predominantly from a single hemisphere. Almost invariably it is the energy metabolism of the brain which is of interest.

There have been many studies using $^{31}$P NMR to assess disease processes in man. Perhaps the most interesting have been the investigation of birth asphyxia in neonates where energy metabolite levels and intracellular pH give direct measures of brain damage and provide a good prognostic index. Other disorders which have been investigated include stroke, multiple sclerosis, epilepsy, hydrocephalus, tumours, dementias and systemic encephalopathies. In all cases there have been reports of altered energy metabolism (reduced PCr or increased $P_i$) and, in multiple sclerosis, changes have been observed in the broad phospholipid peak. The pH data are somewhat more variable – acidic values have been reported in acute stroke and hydrocephalus, but alkaline values have been observed in chronic stroke and brain tumours. One of the main limitations in performing $^{31}$P NMR spectroscopy in the human brain is that the spatial resolution is inadequate for interrogating focal lesions of moderate size. Nevertheless the global pictures that have emerged give insight into the way that brain energetics respond in disease.

## Applications in liver

The liver is relatively simple to study *in vivo* by $^{31}$P spectroscopy. On a macroscopic level it is homogeneous, and localization is generally straightforward. Since hepatocytes do not express creatine kinase, the liver contains no detectable phosphocreatine and so the degree of contamination of the liver spectra by signals from overlying muscle can be readily assessed (and in some cases corrected for).

Localization strategies have ranged from simple surface coil detection (with a hard 180° inversion pulse at the surface to suppress the muscle signal) through Fourier series window approaches and ISIS to one-, two- or three-dimensional chemical shift imaging. Proton decoupling and NOE irradiation have occasionally been employed, but the technical requirements and power deposition considerations have limited these to a few cases. In animal studies, localization methods are often avoided by exposing the liver and using a surface coil placed directly on the organ.

### Liver metabolism

The liver's role is to maintain chemical homeostasis and it can, in certain circumstances, be stressed by means of a metabolic load. One example has been to administer an oral fructose load which causes minimal spectral changes in normal subjects but, at very low dosage, results in an increase in the fructose-1-phosphate peak in patients with fructose intolerance. Other metabolic stresses have been administered to investigate galactose intolerance, glycogen storage diseases and cirrhosis. The rat liver (in situ or perfused) has been studied to elucidate details of carbohydrate and fatty acid metabolism, sometimes by a combination of $^{13}$C and $^{31}$P NMR. The control of pH in the liver has also been studied.

### Liver diseases

Some common diseases that involve the liver have also been studied by $^{31}$P NMR. These include alcoholic liver disease, hepatitis, cirrhosis of various etiologies and liver tumours. The findings have often allowed the diseased liver to be distinguished from the healthy one, but the changes have not usually been specific to a particular disease and the interpretation of the changes is often speculative. Overall, these studies have shown that the major long-term changes in liver spectra involve the PME and the PDE signals, which probably signify changes in phospholipid metabolism (see below) while the energetics of the liver

(assessed from the $P_i$:ATP ratio and the pH) remains normal until a very late stage of disease. Another area that shows promise is in assessing the transplanted liver, either before or after transplantation.

### Spectral interpretation

The PME and the PDE signals in the liver spectrum are each composed of several components, which cannot generally be resolved, except with the use of proton decoupling. Aqueous extracts show that the PME signal contains primarily phosphoethanolamine and phosphocholine, with contributions from 2,3-DPG from blood and other minor components such as AMP, glucose-6-phosphate, *sn*-glycerol-1-phosphate and 3-phosphoglycerate. The PDE signal (at the low magnetic field strengths used for whole-body studies) is mostly composed of signals from phospholipid headgroups, broadened by chemical shift anisotropy and proton–phosphorus dipolar coupling, and also contains signals from GPE and GPC. These can be resolved by proton decoupling as can the PME signal, but this has rarely been undertaken.

## Applications in cancer

From the first $^{31}$P NMR studies of tumours *in vivo*, the differences between tumour spectra and those from other tissues were obvious. The tumour spectra were dominated by the PME signal, and the pH was found to be higher than in most other tissues. While these findings had potential use in distinguishing cancerous from normal tissue, it was hoped that subtle differences in the spectra would also help in determining the type or grade of the tumour, but these hopes have to a large extent not been realized. Indeed some non-cancerous tissue, such as the regenerating liver or the neonatal brain, also show these characteristics.

The PME signal in most tumour types is predominantly composed of phosphoethanolamine, with variable contributions from phosphocholine. These molecules are precursors (and breakdown products) of membrane phospholipids, and it has been suggested that high levels may indicate a rapid rate of membrane turnover, although this has not been shown experimentally. Where proton decoupling has been applied, the relative amounts of phosphoethanolamine, phosphocholine, GPE and GPC can be assessed separately, and this may lead to a better biochemical explanation of the changes.

### pH of tumour cells

Because tumour cells tend to be glycolytic, their pH was expected to be low. It was therefore a surprise to find that the NMR-determined pH was high (about 7.4, similar to that in blood plasma). These findings have been explained by the intracellular pH being kept high by the extrusion of protons into the extracellular space, which has a low pH, as shown by microelectrode studies. This 'reverse pH gradient' (opposite to that found in most cells) has implications for the uptake of drugs, and drug design.

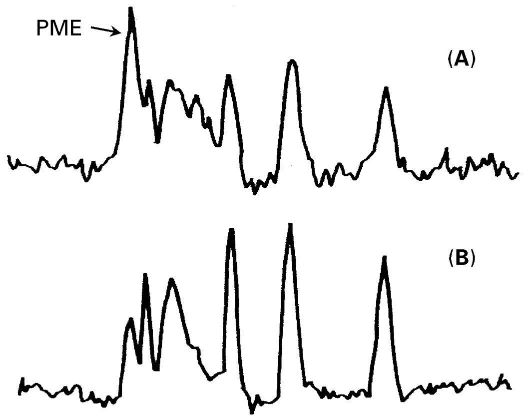



**Figure 4** An example of response to chemotherapy in the lymphomatous liver. Note the dramatic change in the PME peak: (A) before chemotherapy, (B) following chemotherapy.

### Response to therapy

Various groups have studied the response of tumours to therapy, both in patients and in experimental models. The results of such studies are often affected by partial volume effects, i.e. apparent metabolic changes in the tumour may simply be the result of the tumour shrinking, and the spectra containing more signal from the surrounding tissue. The most consistent findings are that tumours that respond to chemotherapy show a greater decrease in the PME:ATP ratio than do non-responders (**Figure 4**). In some cases, these changes take place before any measurable change in tumour size. Changes in pH and Pi:ATP ratios have also been reported. Experimental manipulation of tumour pH and blood flow, together with radiotherapy or chemotherapy, in implanted tumours in animals have led to a greater understanding of the way in which tumours respond to therapy.

## List of symbols

$J$ = coupling constant; $T_1$ = longitudinal relaxation time; $T_2$ = transverse relaxation time.

*See also:* **Biofluids Studied By NMR; Cells Studied By NMR; *In Vivo* NMR, Methods; *In Vivo* NMR,**

**Applications – Other Nuclei; Isotopic Labelling in Biochemistry, NMR Studies; Medical Applications of Mass Spectrometry; Nuclear Overhauser Effect; $^{31}$P NMR; Perfused Organs Studied Using NMR Spectroscopy.**

## Further reading

Bottomley PA (1994) MR spectroscopy of the human heart: the status and the challenges. *Radiology* **191**: 593–612.

Gadian DG (1995) *NMR and its Applications to Living Systems*. Oxford: Oxford University Press.

Gillies RJ (ed) (1994) *NMR in Physiology and Biomedicine*. San Diego: Academic Press.

Kemp GJ and Radda GK (1994) Quantitative interpretation of bioenergetic data from $^{31}$P and $^{1}$H magnetic resonance spectroscopy studies of skeletal muscle: an analytic review. *Magnetic Resonance Quarterly* **10**: 43–63.

Negendank W (1992) Studies of human tumours by MRS: a review. *NMR in Biomedicine* **5**: 303–324.

# *In Vivo* NMR, Applications, Other Nuclei

**Jimmy D Bell**, **E Louise Thomas** and **K Kumar Changani**, Hammersmith Hospital, London, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

## Introduction

*In vivo* nuclear magnetic resonance spectroscopy (or *in vivo* MRS as it has become known) is increasingly being used to provide direct, localized biochemical information on animal and human metabolism in health and disease. Dedicated systems for animal studies use magnetic fields ranging from 2.0 to 7.5 T. Most human MRS systems range from 1.5 to 2.0 T, although 4 T systems are available. Clearly these systems offer spectra with a quality and sensitivity no better than that obtained with high-resolution NMR systems in the mid-1960s. However, even within these constraints, *in vivo* MRS has contributed greatly to our understanding of the metabolic processes associated with a wide range of clinical and metabolic disorders by providing researchers with information that could not be obtained previously even with tissue biopsy. The application of *in vivo* MRS has expanded rapidly in recent years and its value as a research tool depends on the organ under investigation as well as the nuclei being utilized. This article reviews the work carried out to date in *in vivo* MRS using $^{13}$C, $^{19}$F, $^{15}$N, $^{23}$Na and other less common nuclei.

A list of nuclei currently being used for *in vivo* biomedical research is shown in **Table 1**. Although many of these nuclei are routinely used in high-resolution *in vitro* NMR spectrometers, their implementation in *in vivo* MRS systems required substantial technical development and many are available only to a handful of research groups around the world. This is partly due to the considerable cost associated with developing multinuclear capabilities into commercial systems. Accordingly, most commercial MR systems have very restricted spectroscopic capabilities, with $^{1}$H and $^{31}$P being the nuclei generally available. This has, of course, restricted the application of less common nuclei in biomedical research. A summary of the main areas of research and application of some of these nuclei is shown in **Table 2**. Results of many of these studies have made considerable contributions to important areas of clinical and biochemical research, while other studies have concentrated on demonstrating technical feasibility, although they suggest the prospect of major areas of future research.

## *In vivo* $^{13}$C MRS

Whole-body $^{13}$C MRS is potentially the most powerful tool available for the study of human metabolism *in vivo*. The nuclide's low natural abundance (1.1%) allows the study of important biological molecules such as glycogen and lipids, while at the same time allowing the use of $^{13}$C-enriched compounds in dynamic measurements of intermediate metabolism. However, several technical issues need to be addressed before its implementation *in vivo*, including a second transmit channel, optimum localization and proton decoupling. The latter has been one of the major stumbling blocks for the application of *in vivo*

**Table 1** NMR characteristics of some nuclei of biomedical interest

| *Nucleus* | *Spin* | *Relative sensitivity* | *Natural abundance (%)* | *Observation frequency for $^1H$ observation at 100 MHz* |
|---|---|---|---|---|
| $^2H$ | 1 | $9.65\times10^{-3}$ | $1.5\times10^{-2}$ | 15.35 |
| $^7Li$ | $\frac{3}{2}$ | $2.93\times10^{-1}$ | 92.58 | 38.86 |
| $^{11}B$ | $\frac{3}{2}$ | $1.65\times10^{-1}$ | 80.42 | 32.08 |
| $^{13}C$ | $\frac{1}{2}$ | $1.59\times10^{-2}$ | 1.11 | 25.14 |
| $^{15}N$ | $\frac{1}{2}$ | $1.03\times10^{-3}$ | 0.37 | 10.13 |
| $^{17}O$ | $\frac{5}{2}$ | $2.91\times10^{-2}$ | 0.04 | 13.56 |
| $^{19}F$ | $\frac{1}{2}$ | $8.3\times10^{-1}$ | 100 | 94.08 |
| $^{23}Na$ | $\frac{3}{2}$ | $9.25\times10^{-1}$ | 100 | 26.45 |
| $^{27}Al$ | $\frac{5}{2}$ | $2.06\times10^{-1}$ | 100 | 26.06 |
| $^{29}Si$ | $\frac{1}{2}$ | $7.84\times10^{-3}$ | 4.70 | 19.86 |
| $^{87}Rb$ | $\frac{3}{2}$ | $1.7\times10^{-1}$ | 27.85 | 32.72 |
| $^{129}Xe$ | $\frac{1}{2}$ | $2.12\times10^{-2}$ | 26.44 | 27.66 |

$^{13}C$ MRS, because of the potential problem of power deposition during proton decoupling. This problem has now been addressed by the use of the newer pulse train decoupling sequences, which allow full decoupling of signal with low power deposition. The need for a second transmit channel with frequency synthesizer, phase modulator and power amplifier, together with the need for specifically designed $^1H/^{13}C$ coil systems, means that the actual cost of a standard *in vivo* scanner is substantially increased, hence limiting its availability. Thus, although many of the technical problems associated with the use of $^{13}C$ MRS *in vivo* have been overcome, it is still not routinely available to the great majority of research groups.

**Table 2** *In vivo* applications of NMR spectroscopy

| *Nucleus* | *Application* | *Organ/tissue* |
|---|---|---|
| $^{13}C$ | Glycogen | Liver, muscle |
| | Amino acids, glucose | Brain |
| | Lipids | Adipose tissue, liver |
| $^{19}F$ | Drug metabolism | Liver, brain, muscle |
| | Anaesthetics | Brain |
| | Phospholipids | Tumour |
| | Intracellular $Ca^{2+}$ | Brain, kidney, spleen |
| | Glucose metabolism | Brain |
| $^{15}N$ | Glutamine metabolism | Brain |
| $^{23}Na$ | Electrolyte balance | Muscle, brain, liver, heart |
| $^2H$ (D) | $D_2O$, methionine, glucose | Brain, adipose tissue, liver |
| | Choline | Kidney |
| $^{27}Al$ | Gastric emptying | Gut |
| $^{129}Xe$ | Lung volume | Lungs |
| $^{17}O$ | Blood flow | Brain, liver |
| $^{87}Ru$ | $Na^+/K^+$-ATPase | Muscle |
| $^7Li$ | Psychiatric disorders | Brain, muscle |
| $^{11}B$ | Therapy agents | Liver, brain, tumour |
| $^{29}Si$ | Breast implants | Liver |

## Application of *in vivo* $^{13}C$ MRS to carbohydrate metabolism

Carbohydrate metabolism is one area of research that has significantly benefited from *in vivo* $^{13}C$ MRS since signals from sugar carbons occur in a region of the spectrum that is relatively unpopulated by other resonances. Using specific carbon labelling on different substrates, various metabolic pathways have been studied. Further information has been gained by labelling two adjacent carbon atoms within the same molecule; this produces a characteristic $^{13}C$–$^{13}C$ spin–spin coupling, helping to elucidate important information about particular enzyme pathways.

**Liver** The first *in vivo* experiments, performed in 1981, explored the potential application of $^{13}C$ MRS to glucose/glycogen metabolism. Following infusion of $^{13}C$-enriched glucose in rats, well-resolved resonances from both glucose and glycogen could be observed, enabling the time course of the relative changes in concentration to be established simultaneously. One area of debate at the time was the degree of glycogen visibility in liver. It was initially thought that glycogen, being such a large macromolecule, would be only partly NMR 'visible'. However, the most recent consensus is that glycogen is in fact 100% visible and that $^{13}C$ MRS can therefore provide reliable noninvasive measurements of hepatic glycogen concentrations over physiological ranges.

Glycogen repletion studies have been carried out in animals and humans. Infusions of [1-$^{13}C$]glucose and [3-$^{13}C$]alanine to fasted rats suggested that 30% of the total glycogen formed originated from the direct conversion of glucose in the short term. However, over the longer term, active glycogen synthesis from [3-$^{13}C$]alanine occurred for a longer period of time, with a rate constant 3-fold higher than from [1-$^{13}C$]glucose, implying greater efficiency of glycogen synthesis from the indirect route. These studies have also highlighted potential problems associated with substrate cycling, since following entry into the Krebs cycle the label may be assimilated into other pathways and then re-introduced into the gluconeogenic pathway. Definitive mechanisms of glycogen repletion therefore become difficult to assess *in vivo*, unless measurements are performed immediately following administration.

Human studies have concentrated on direct observation of glycogen repletion using natural abundance $^{13}C$ or after administration of [1-$^{13}C$]glucose. Early work compared pre- and post-meal hepatic glycogen repletion rates in individuals starved for 30 hours, showing repletion rates 3.8 times greater in the post-meal states compared with the fasted basal rate. Hepatic gluconeogenesis in healthy subjects and those with non-insulin-dependent diabetes mellitus (NIDDM) was also investigated by $^{13}C$ MRS and conventional methods. Concentration of glycogen was lower in NIDDM patients compared to controls (**Figure 1**). Furthermore, following infusion of [6-$^{3}H$]glucose, gluconeogenesis in the NIDDM subjects was found to be 60% greater than in control subjects and accounted for 88% of the total glucose production compared with 70% in the normal subjects. This suggests that increased gluconeogenesis accounts for the increase in whole-body glucose production in overnight-fasted NIDDM subjects.

$^{13}C$ MRS has also been applied to determine the metabolic alterations caused by glycogen storage diseases. Patients with glycogen type IIIA storage disease were shown to have increased levels of hepatic glycogen compared with normal volunteers. This disease manifests owing to the inactivation of amylo-1,6-glucosidase, which is required for glycogen hydrolysis. It was concluded that *in vivo* $^{13}C$ MR spectroscopy could play an important role in the assessment of other glycogen storage diseases where current clinical practice requires needle biopsies to be performed.

$^{13}C$ MRS has also been used to assess enzyme capacities and kinetics in animals. Following infusion of [1-$^{13}C$]butyrate (90% enriched), rates of ketogenesis could be monitored by measuring the formation of acetoacetate and β-hydroxybutyrate within the carbonyl region of the $^{13}C$ MR spectrum. Further studies of $^{13}C$-labelled fatty acids, such as [1,3-$^{13}C$]butyrate, assessed the metabolic pathways of ketone body production in fasted and diabetic rats. Increased β-hydroxybutyrate production was seen in diabetic rats; it was thought that a large fraction of butyrate evaded β-oxidation to acetyl-CoA and was converted directly to acetoacetyl-CoA. Different physiological states can therefore be monitored by *in vivo* $^{13}C$ MRS since labelling patterns following $^{13}C$-enriched metabolite administration would be different.

*In vivo* $^{13}C$ MR spectroscopy has also been applied to monitor the pharmokinetics of $^{13}C$-labelled xenobiotic products in hepatic metabolism. One such study used a phenacyl imidaolium compound that effects glycogen synthesis. It was shown that glycogen synthesis increases 2-fold in drug-treated rats and it was suggested that augmentation of the indirect pathway is the probable mechanism. Further studies have shown that long-term monitoring of drug concentrations is possible using $^{13}C$-labelled compounds, which may be important for monitoring of hepatic tumour treatments.

**Muscle** Most human $^{13}C$ MRS studies in muscle have concentrated on the study of glycogen synthesis and its role in NIDDM, glycogen storage diseases and muscle myopathies. Studies of insulin action on muscle glycogen synthesis have shown that there appears to be an impairment of muscle's ability to either take up glucose or assimilate it as glycogen in NIDDM.

Recent studies have used *in vivo* $^{13}C$ MRS to study muscles following intense exercise. This, together with interleaved $^{31}P$ MRS, has provided important information on the phosphorylation states of

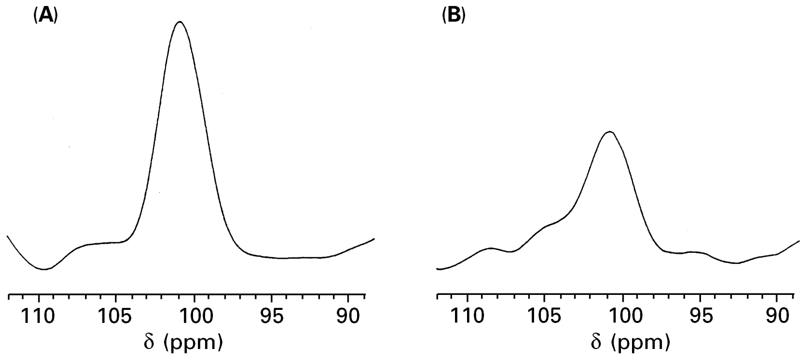



**Figure 1** Typical *in vivo* $^{1}H$-decoupled $^{13}C$ MRS spectra at 22.5 MHz of the C-1 position of liver glycogen from a control subject (A) and from a type II diabetic patient (B) 4 hours after a liquid meal. Reproduced with permission from Magnusson I, Rothman DL, Katz LD, Shulman RG and Shulman GI (1992) Increased rate of gluconeogenesis in type II diabetes mellitus. *Journal of Clinical Investigation* **90**: 1323–1327.

glycogen synthase that could be correlated with post-exercise glycogen synthesis rates. Studies of this kind could provide important information on enzyme pathways that may be defective in particular muscle myopathies. Acetate metabolism has also been assessed in rabbit muscles to determine rates of anaplerotic flux (i.e. determination of label of glutamate and glutamine) compared with basal flux of the Krebs cycle. Different muscle types (muscle fibre/oxidative or nonoxidative) have dissimilar anaplerotic flux rates, which may again be relevant for particular disease types.

**Heart** *In vivo* $^{13}$C MR spectroscopy of the heart is difficult owing to continuous contraction of the heart. For this reason, together with the relative insensitivity of the $^{13}$C nucleus, which requires a large number of averages, *in vivo* MRS studies are few. Consideration of the metabolites within blood must also be made so that the spectral acquisitions contain information purely reflecting heart metabolism. The bulk of studies on the heart have therefore been conducted using isolated, perfused organs. However, there have been a few *in vivo* studies that have looked at glycogen synthesis rates following infusions of acetate and glucose. These studies can provide information about physiological conditions that may effect carbohydrate utilization in favour of fatty acid metabolism, with implications for coronary heart disease.

**Brain** *In vivo* $^{13}$C MR spectroscopy of the brain has been used to determine *in vivo* levels of glutamate, glutamine, $\gamma$-aminobutyrate (GABA), and lactate following intravenous infusions of enriched precursors. Brain tumours in rats have been studied to investigate carbohydrate metabolism following administration of [1-$^{13}$C]glucose. Rat gliomas produced lactate and glutamate/glutamine as well as glycogen. Further studies used combinations of labelled substrates such as glucose and acetate to determine the relative flux through the tricarboxylic acid cycle together with their relative conversion rates into the cerebral amino acids, important for neurotransmission. Both these studies suggested that *in vivo* $^{13}$C MRS is ideal for studying the individual carbons of different metabolites in the brain. Furthermore, cerebral compartmentation of glial and neuronal cells may also be studied because of their different metabolism. Two separate glutamate pools and two separate tricarboxylic acid cycles, one preferentially metabolizing acetate and the other mainly using unlabelled acetyl-CoA and a predominant production of GABA in the glutamate pool lacking glutamine synthase, have been identified, accounted for by either the glial or neuronal cells.

A further application of *in vivo* $^{13}$C MRS has been to study the metabolite changes associated with mild or severe hypoxia as well as electroshock. Changes associated with lactate and the cerebral amino acids were correlated with both electroshock and degree of hypoxia. This work may be useful in characterizing the metabolic state following epilepsy, stroke, trauma, tumours and other pathological conditions in humans. In humans, total brain glutamate concentration has been assessed following [1-$^{13}$C]glucose infusions, while in stroke it has helped to determine metabolic turnover. Some of these studies will benefit from the recent introduction of proton-observed carbon-edited (POCE) techniques, which can provide better relative sensitivity and resolution. This technique has been very useful, for example, in determining the effects on brain metabolism of drugs such as inhibitors of succinate dehydrogenase and the GABA-transaminase inhibitor vigabatrin.

### Application of *in vivo* $^{13}$C MRS to lipids

The main application of *in vivo* $^{13}$C MRS to the field of lipids has been the study of adipose tissue composition. Although the natural abundance of the $^{13}$C isotope is low, the high content of lipids in adipose tissue has enabled acquisition of $^{13}$C MR spectra *in vivo* in a matter of minutes, with or without proton decoupling. This has allowed the use of this technique on subjects (e.g. neonates and in pregnancy) where power deposition from decoupling, though small, it is still unacceptable.

**Adipose tissue** Adipose tissue is principally composed of triglycerides and it has been shown to play an important role in major diseases, including NIDDM and cardiovascular disease. Preliminary studies in the early 1980s showed that *in vivo* $^{13}$C MR spectroscopy could be used to detect differences in the lipid composition of adipose tissue and liver in rats fed a diet high in polyunsaturated fatty acids. The first human studies showed that linoleic acid (18:2*n*–6), usually the principal polyunsaturated fatty acid present in human adipose tissue, dominated the polyunsaturated signal observed *in vivo*, allowing estimation of this stored essential fatty acid. Further work showed that fatty acid composition of adipose tissue closely correlated with dietary fat content, with subjects on long-term vegan diet showing significantly higher levels of polyunsaturated fatty acids and reduced saturated fatty acids compared to omnivores (general public) and vegetarians. Indeed, there was no difference in the adipose tissue composition of vegetarians compared to omnivores, suggesting that the former have substituted the animal fat (mainly saturated) with saturated fat in dairy products.

*In vivo* $^{13}C$ MR spectroscopy has also been used to study adipose tissue composition in babies. The work showed that the influence of maternal diet on infant adipose tissue composition and changes due to gestational age and during postnatal development could all be detected noninvasively by *in vivo* $^{13}C$ MR spectroscopy. Preterm infants had significantly lower levels of unsaturated adipose tissue fatty acids than the infants born at full term. Furthermore, the level of unsaturated fatty acids continued to increase as the infants matured (from birth to 6 weeks, to 6 months). Interestingly, the authors also found that infants breast-fed by vegan mothers had 70% more polyunsaturated fatty acids than infants breast-fed by omnivore mothers, the long-term consequences of which (beneficial or harmful) are unknown. The same authors showed that *in vivo* $^{13}C$ MRS could also measure the effect of intensive exercise on adipose tissue composition of adult human volunteers. A decrease in polyunsaturated fatty acids was observed, suggesting that exercise is an independent factor for adipose tissue composition.

In addition, *in vivo* $^{13}C$ MR spectroscopy has been applied to the study of adipose tissue composition in disease. Children with cystic fibrosis were shown to have lower levels of polyunsaturated adipose tissue fatty acids than healthy children, possibly owing to a disorder in essential fatty acid metabolism that may be partly responsible for the development of the disease. Further studies with *in vivo* $^{13}C$ MRS in disease have shown a significant increase in saturated adipose tissue fatty acids following transplantation and subsequent weight gain in malnourished patients with liver cirrhosis. It was suggested that this increase in saturated fatty acids may be secondary to a general repletion of membrane polyunsaturated fatty acids or the use of essential fatty acids for biosynthesis of eicosanoids in the postoperative period.

**Liver** $^{13}C$ MRS has also been used for noninvasive quantification of hepatic triglyceride content. Diagnosis of hepatic steatosis (fatty liver) is important because this is a reversible condition and early detection can help to prevent irreversible liver damage. The condition is normally diagnosed by a liver biopsy, but the invasive nature of this procedure and the problems associated with biopsies make it difficult to monitor patients at risk of developing hepatic steatosis. In a study of 15 patients with varying degrees of fatty infiltration it was shown that by quantifying the intensity of the $CH_2$ resonances from the *in vivo* $^{13}C$ MR spectrum of the liver it was possible to determine hepatic lipid content. Excellent correlation with conventional liver biopsy measurements was observed. As the authors suggested, this technique will in future become a valuable tool in the diagnosis and follow-up in patients with hepatic steatosis.

**Brain** The application of *in vivo* $^{13}C$ MR spectroscopy to the study of brain lipids has been rather limited, partly because most of the lipids in the brain are present in the form of membranes. This gives rise to broad and uninformative signals. An interesting possibility for improving the discriminatory power of MRS, as applied to lipids, is the use of $^{13}C$-enriched fatty acids. However, at present such studies would be prohibitively expensive as the subjects would be required to consume gram quantities of $^{13}C$-enriched fatty acid.

## *In vivo* $^{19}F$ MRS

Fluorine-19 is a highly NMR-sensitive and naturally abundant nucleus. An important advantage of $^{19}F$ MRS is the fact that it allows direct observation of fluorinated compounds and their metabolites in the human body without background signal from tissue. *In vivo* $^{19}F$ MR spectroscopy has been used to detect noninvasively anticancer and psychoactive drugs and other fluorinated compounds and to study their metabolism. It has also been used to monitor the effects of anaesthetics and to measure intracellular pH and calcium levels.

One of the major applications of *in vivo* $^{19}F$ MR spectroscopy has been in the field on oncology, detecting and monitoring the metabolism of fluorine-containing anticancer drugs. An important chemotherapeutic compound is the fluorinated drug 5-fluorouracil (5-FU). Much of the work with $^{19}F$ MRS has concentrated on determining optimal drug administration schedules and studying the effects of this drug on both tumour and healthy tissue. $^{19}F$ MRS has allowed detection and serial measurement of 5-FU and its metabolites in patients undergoing chemotherapy, increasing our understanding of the variation in efficacy of chemotherapy in different subjects (**Figure 2**). For example, levels of 5-FU ‘trapping’ by tumours following intravenous administration of the drug have been measured. The results reveal that patients whose tumours trapped the drug had a significantly better response to chemotherapy than those patients whose tumours did not trap 5-FU. This differential in response to therapy could not be ascertained from the standard plasma concentration measurements. From this study it was concluded that $^{19}F$ MRS can be used to identify patients who are likely to respond to chemotherapy with 5-fluorouracil and can therefore aid selection of the optimal

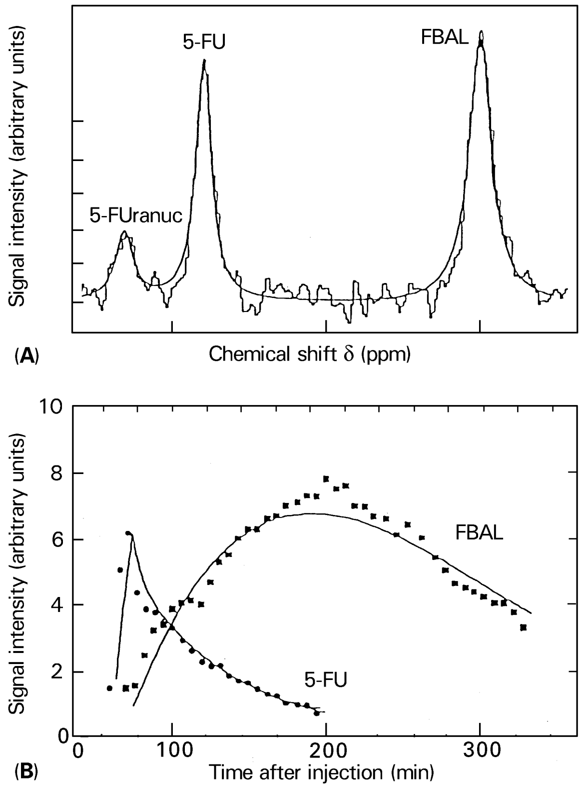



**Figure 2** (A) Liver *in vivo* $^{19}$F MR spectrum at 60 MHz following administration of 5-fluorouracil (5-FU). (B) Time course of 5-FU and its major catabolite α-fluoro-β-alanine (FBAL). (FUranuc = 5-fluorouracil nucleoside + nucleotides). Reproduced with permission from Semmler W, Bacher-Baumann P and Guckel F (1990) Real time follow-up of 5-fluorouracil metabolism in the liver of tumor patients by means of F-19 MR spectroscopy. *Radiology* **174**: 141–145.

treatment for individual patients. Studies of 5-FU metabolism in patients with protein calorie malnutrition (a common condition in patients with cancer) have shown increased toxicity from this drug. Therefore, identifying alterations in 5-FU metabolism may be an important tool in reducing toxicity during chemotherapy. Interestingly, like $^{31}$P MRS, *in vivo* $^{19}$F MRS may also be used to measure pH noninvasively. This may be useful for $^{19}$F MRS studies of tumours, as the uptake and retention of 5-fluorouracil is dependent on tumour pH.

An important area of research using *in vivo* $^{19}$F MRS has been the study of the pharmacokinetics of fluvoxamine and fluoxetine, drugs used to treat obsessive-compulsive disorder and depression, respectively. Several groups have shown that it is possible to detect these compounds noninvasively in the brain. Brain fluvoxamine levels were shown to reach a steady state 30 days after consistent daily dosing, substantially more rapidly than reported for fluoxetine, which was shown to plateau after 6–8 months of treatment. Brain concentrations of both drugs were shown to correlate with plasma levels, though the absolute concentrations were substantially higher in the brain. Comparison of the *in vivo* measurement of fluoxetine with *in vitro* samples taken at autopsy suggest that fluorinated drugs may not be 100% visible *in vivo*. These are based on the *in vitro* findings from just one brain that had been fixed in formaldehyde, so further studies would be required to confirm these findings.

Some of the other applications of *in vivo* $^{19}$F MRS to the study of fluorinated drugs have included the study of fluorinated antibiotics and antifungal agents in animals and humans. Orally administered fleroxacin (a fluoroquinolone antibiotic agent) could be detected in the liver and calf muscle by *in vivo* $^{19}$F MRS. Furthermore, its washout and metabolism were studied over a 24-hour period. There are important implications of this work, and it will in future be possible to ensure that drugs reach their target site in appropriate concentrations.

*In vivo* $^{19}$F MRS has also been used to study the pharmacokinetics and metabolism of anaesthetic agents such as halothane, isoflurane and desflurane in experimental animals. It has also been shown that anaesthetic agents can be observed noninvasively in the brains of patients immediately following surgery. The principal aim of these studies was to elucidate the mechanism of action of these compounds and eventually help in the development of more effective and selective anaesthetics. Unfortunately, factors such as sensitivity and localization of the region of interest (which are particularly relevant to the study of anaesthetics) have greatly limited the impact of this research.

Animal studies have shown that *in vivo* $^{19}$F MRS can be used in metabolic studies. It has been shown that 3-fluoro-3-deoxy-D-glucose in the brain is metabolized primarily in the aldose reductase sorbitol pathway, suggesting that $^{19}$F MRS may be a useful tool for the further elucidation of this pathway. Further studies have shown that it is also possible to study the metabolism of fluorinated galactose, tryptophan and protein *in vivo* and determine intracellular calcium levels. The concentration of this cation was shown to be the same in the kidney, spleen and brain of rats, at approximately 200 nM. These measurements were obtained by infusing the calcium indicator 5F-BAPTA (=5,5-difluoro-1,2-bis(*o*-aminophenoxy)ethane-$N,N,N^1,N^1$-tetraacetic acid) intravenously or intraventricularly into the animal, and so the potential application to human subjects appears limited.

It has also been possible to follow the time course of recovery from nerve injury in skeletal muscle

using *in vivo* $^{19}$F MRS. Muscle injury leads to significant decrease in the energy state and local circulation dynamics, which appear to reverse during recovery. While $^{31}$P MRS was used to determine the energy status of the cells, $^{19}$F MRS was utilized to determine local circulation dynamics by measuring blood volume. The results indicate that the recovery processes of circulation precede those of energy state.

Many of the above studies have shown that in the future $^{19}$F MRS is likely to become a valuable tool for monitoring drug metabolism at the site of action, though at the moment this is limited by the availability of *in vivo* MR scanners with $^{19}$F capabilities.

## Less common nuclei

In the last decade there has been a significant increase in the number of 'less common' nuclei that are being utilized for *in vivo* MRS (see **Table 1**). The use of some of these nuclei was initially limited by technical constraints and/or inherent NMR characteristics. However, recent studies have shown that these nuclei can be used to increase our understanding of some important physiological processes.

A number of studies have utilized *in vivo* MRS to assess the extent of delivery of pharmacological substances to target tissue. Nitrogen-15 MRS in combination with intravenous $^{15}NH_4^+$ infusion into animal models of hepatic encephalopathy (**Figure 3**) enabled the determination of glutamine synthetase activity *in vivo*. The results show that *in situ* the activity of this enzyme is kinetically limited by the levels of some substrates and cofactors, and that the activity of phosphate-activated glutaminase was maintained at a low level by a suboptimal concentration of $P_i$ and the strong inhibitory effect of glutamine. The introduction of the heteronuclear multiple quantum coherence (HMQC) transfer technique is likely to further these studies by probing the kinetics of glutamine synthesis at physiological concentrations.

Similarly, lithium-7 MRS has been used to examine the pharmacokinetics of lithium, a drug that is widely used in the treatment of a number of psychiatric disorders. Initial studies showed that there was a slow accumulation of lithium in the brain, which may be responsible for the delay in therapeutic response that is often seen following the initiation of therapy. Subsequent studies showed that lithium levels in the brain and muscle were lower than the average serum concentration. From the ratio of these *in vivo* measurements, the authors hypothesized that the minimal effective concentration of brain lithium level for maintenance treatment of bipolar disorder could be determined. However, in a similar study, researchers found that in a subgroup of patients there was only a weak correlation between serum and brain lithium levels, which may account for the failure of lithium therapy in some patients with serum lithium levels within the therapeutic range. This suggests that measurement of *in vivo* brain lithium concentration by MRS may be more clinically relevant than simply assessing serum concentration.

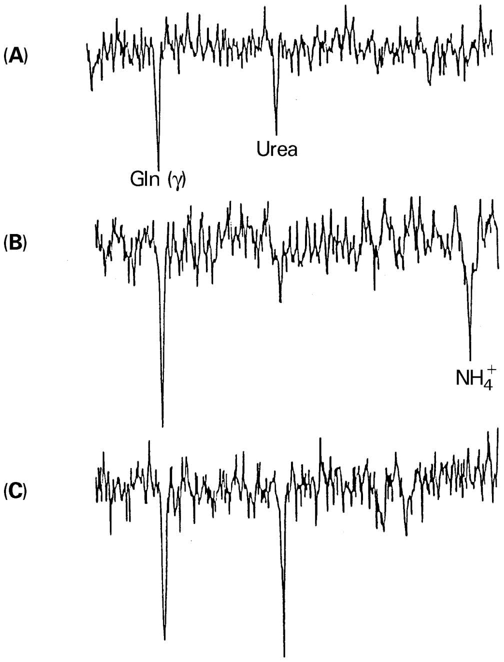



**Figure 3** *In vivo* $^{15}$N spectra at 24.3 MHz from cerebral region of (A) a control rat given low-rate $^{15}NH_4^+$ infusion, showing cerebral [γ-$^{15}$N]glutamine at a near physiological concentration; (B) a portacaval-shunt rat showing cerebral and blood $^{15}NH_4^+$, in addition to [γ-$^{15}$N]glutamine; (C) a control rat showing cerebral and blood [$^{15}$N]urea and lower [γ-$^{15}$N]glutamine. Peaks are downgoing following NMR convention for $^{15}$N spectroscopy. Reproduced with permission from Kanamori K, Parivar F and Ross BD (1993) A $^{15}$N NMR study of *in vivo* cerebral glutamine synthesis in hyperammonemic rats. *NMR in Biomedicine* **6**: 21–26.

Boron-11 MRS has been used to assess the delivery of boron neutron capture therapy (BNCT) agents to a target lesion. Uptake and retention of these compounds was determined *in vivo* in humans and animals and the results show that there is differential retention between normal and abnormal tissue. Furthermore, $^{1}$H-observed $^{10}$B-edited MRS has been shown to provide much higher sensitivity and could be utilized to monitor the distribution and excretion of boron agents noninvasively in patients about to undergo BNCT.

Sodium-23 magnetic resonance spectroscopy has been used to yield information on electrolyte balance,

which appears to reflect physiological changes associated with cell function. In myotonic dystrophy, an inherited multisystem disease that leads to muscle malfunction, $^{23}$Na MRS was shown to be a sensitive method for quantification of disease progression. Quantification of intracellular and extracellular sodium by MRS has also generated important information on organ function following transplantation or injury. In heart transplantation experiments in rats, $^{23}$Na MRS showed that there was a steady increase in intracellular sodium 3 days prior to rejection, followed by a sharp increase after rejection. In the liver, intracellular sodium levels were shown to be a sensitive indicator of hepatic dysfunction after major systemic injury. In the kidney, sodium compartmentation (intracellular, vascular and intraluminal) was determined using the shift reagent TmDOTP5, while in the brain this shift reagent was shown to be a useful tool in determining compartmental sodium concentration and measuring blood flow kinetics.

The activity of ion transport systems, such as sodium- and potassium-activated adenosine triphosphatase, has been studied *in vivo* by the use of rubidium-87 MRS. This has been possible because rubidium has been shown to substitute for potassium in a number of transmembrane transport systems, accumulating in the intracellular space. Standard *in vitro* methods of determining $Na^+/K^+$-ATPase activity, which also use rubidium, give highly variable measurements and in some cases contradictory results. Many of these problems appear to have been overcome by the use of *in vivo* $^{87}$Rb MRS, especially when sequential measurements are required. In a longitudinal study of spontaneously hypertensive rats, $^{87}$Rb MRS showed that skeletal muscle rubidium rose at a faster rate in hypertensive rats than in control animals, which is consistent with a marked increase in $Na^+/K^+$-ATPase activity. This type of experiment emphasizes the value of *in vivo* MRS since rubidium kinetics can be determined sequentially on the same animal, minimizing inter/intra-subject variability as well as the number of animals required in the study.

*In vivo* deuterium ($^2$H) MRS has been used to characterize amino acid metabolism, body iron content, brain and kidney metabolism and body fat utilization rates in rodents. These studies rely on the use of deuterium labelling or the existence of natural-abundance deuterium in water or lipids. For example, deuterium-labelled methionine was used to confirm the dominant contribution of the glycine/sarcosine shuttle to the metabolism of excess methionine, while deuterium-labelled glucose was used to show that systemic glucose level influences brain metabolic recovery following partial ischaemia. The renal distribution and metabolism of methyl groups was assessed by the use of deuterium-labelled choline. The results show that the concentration of trimethylamines, metabolic products of choline, was higher in the cortex and inner medulla than in the outer medulla, but that the inner medullary fraction was more liable to diuresis.

By combining $D_2O$ dilution techniques with $^2$H MRS it has been possible to study fat turnover *in vivo*. The rate of loss of deuterium from body fat, after a short period of $D_2O$ administration, has been shown to reflect utilization of body fat. This technique was used successfully to demonstrate that induction of diabetes in mice did not affect the utilization of fat as a metabolic fuel. Given that $D_2O$ administration is regularly used with human subjects, this opens up the possibility, for the first time, of determining fat turnover noninvasively in human subjects.

A number of studies have used nuclei such as oxygen-17, aluminium-27 and xenon-129 in conjunction with *in vivo* magnetic resonance spectroscopy for structural or dynamic measurements in different organs. $^{129}$Xe signals were detected in the lung and head from human volunteers following inhalation of hyperpolarized $^{129}$Xe gas, allowing the determination of the temporal evolution of the gas-phase and dissolved-phase components. Most of this work, however, was used to demonstrate the feasibility of obtaining xenon images of the lung. $^{129}$Xe imaging now allows the possibility of imaging the lung at resolution equivalent to that of conventional imaging of normal tissue. $^{27}$Al MRS has been applied to the visualization of gastric emptying and measurement of gastric phospholipids, while $^{17}$O MRS has been used to explore cerebral blood flow. Although these studies have again concentrated mainly on their technical feasibility, they do open up important areas of clinical and biochemical research.

## Conclusions

The application of *in vivo* multinuclear MRS to biological sciences is an ever-expanding area of research. Present-day clinical NMR systems routinely allow the use of $^1$H and $^{31}$P MR spectroscopy in clinical and scientific research; however, the use of other nuclei is still restricted to a relatively small number of research groups around the world. This is principally for commercial reasons rather than scientific or technical considerations. Demands for an increase in the multinuclear capability of commercial systems is of course increasing as the usefulness of nuclei such

as $^{13}C$, $^{19}F$ and $^{15}N$ continues to be demonstrated both in animal models and in human subjects. It is conceivable that in the near future these technical capabilities will be routinely available in most clinical systems as they are today available in most *in vivo* animal systems. This will no doubt greatly help to further elucidate metabolic mechanisms in health and disease and in the development and monitoring of the efficacy of novel clinical treatments, including gene therapy.

*See also:* **$^{13}C$ NMR, Methods; Drug Metabolism Studied Using NMR Spectroscopy; $^{19}F$ NMR, Applications, Solution State; Heteronuclear NMR Applications (B, Al, Ga, In, Tl); Heteronuclear NMR Applications (Ge, Sn, Pb); *In vivo* NMR, Applications – $^{31}P$; *In vivo* NMR, Methods; Labelling Studies in Biochemistry Using NMR; Nitrogen NMR; NMR Spectrometers; Nuclear Overhauser Effect; Perfused Organs Studied Using NMR Spectroscopy; Xenon NMR Spectroscopy.**

## Further reading

Arai T, Mori K, Nakao S *et al* (1991) *In vivo* oxygen-17 nuclear magnetic resonance for the estimation of cerebral blood flow and oxygen consumption. *Biochemical and Biophysical Research Communications* **179**: 954–961.

Bartels M and Albert K (1995) Detection of psychoactive drugs using $^{19}F$ MR spectroscopy. *Journal of Neural Transmission, General Section* **99**: 1–6.

Cox IJ (1996) Development and applications of *in vivo* clinical magnetic resonance spectroscopy. *Prog Biophys Mol Biol* **65**: 45–81.

Eng J, Berkowitz BA and Balaban RS (1990) Renal distribution and metabolism of [$^{2}H_9$] choline. A $^{2}H$ NMR and MRI study. *NMR in Biomedicine* **3**: 173–177 (and references therein).

Gadian DG (1995) *NMR and its Applications to Living Systems*. Oxford: Oxford Science Publications, Oxford University Press.

Kabalka GW, Davis M and Bendel P (1988) Boron-11 MRI and MRS of intact animals infused with a boron neutron capture agent. *Magnetic Resonance in Medicine* **8**: 231–237.

Kanamori K, Parivar F and Ross BD (1993) A $^{15}N$ NMR study of *in vivo* cerebral glutamine synthesis in hyperammonemic rats. *NMR in Biomedicine* **6**: 21–26.

Kaspar A, Bilecen D, Scheffler K and Seelig J (1996) Aluminium-27 nuclear magnetic resonance spectroscopy and imaging of the human gastric lumen. *Magnetic Resonance in Medicine* **36**: 177–192.

Kushnir T, Knubovets T, Itzchak Y *et al* (1997) *In vivo* $^{23}Na$ NMR studies of myotonic dystrophy. *Magnetic Resonance in Medicine* **37**: 192–196 (and references therein).

Leach MO (1994) Magnetic resonance spectroscopy applied to clinical oncology. *Technology and Health Care* **2**: 235–246.

Magnusson I, Rothman DL, Katz LD, Shulman RG and Shulman GI (1992) Increased rate of gluconeogenesis in type II diabetes mellitus. *Journal of Clinical Investigation* **90**: 1323–1327.

Mugler JP, Driehuys B, Brookman JR *et al* (1997) MR imaging and spectroscopy using hyperpolarized $^{129}Xe$ gas: preliminary human results. *Magnetic Resonance in Medicine* **37**: 809–905.

Semmler W, Bacher-Baumann P and Guckel F (1990) Real time follow-up of 5-fluorouracil metabolism in the liver of tumor patients by means of F-19 MR spectroscopy. *Radiology* **174**: 141–145.

Soares JC, Krishnan KR and Keshavan MS (1996) Nuclear magnetic resonance spectroscopy: new insights into the pathophysiology of mood disorders. *Depression* **4**: 14–30 (and references therein).

Syme PD, Dixon RM, Allis JL, Aronson JK, Grahame-Smith DG and Radda GK (1990) A non-invasive method of measuring concentrations of rubidium in rat skeletal muscle *in vivo* by $^{87}Rb$ nuclear magnetic resonance spectroscopy: implications for the measurement of cation transport activity *in vivo*. *Clinical Science* **78**: 303–309 (and references therein).

Yoshida SH, Swan S, Teuber SS and Gershwin ME (1995) Silicone breast implants: immunotoxic and epidemiologic issues. *Life Science* **56**: 1299–1310.

Young IR and Charles HC (1996) *MR Spectroscopy: Clinical Applications and Techniques*. London: Martin Dunitz.

# *In Vivo* NMR, Methods

**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
**Methods & Instrumentation**

## Introduction

NMR spectroscopy is a technique that has found widespread applicability in many areas of science. Although originally a physics method for measuring nuclear magnetic moments, it soon became invaluable in studies of identification and structuring of organic and inorganic chemicals with the discovery of the chemical shift and the resolution of spin–spin coupling splittings. Later the application of NMR spectroscopy was amplified by the realization that pieces of tissue could yield reasonably resolved NMR spectra (initially of the $^{31}$P nucleus) and this sparked a major research effort in the use of NMR spectroscopy to study biochemical systems such as intact cell suspensions and perfused tissues.

A logical next step was the investigation of whole living animals and humans and nowadays there is much research activity in these fields. Many of the *in vivo* NMR studies have been in the area of clinical medicine and by analogy with magnetic resonance imaging (MRI), where the 'nuclear' part has been dropped, *in vivo* NMR spectroscopy is often referred to as simply magnetic resonance spectroscopy (MRS), really a misnomer as there are other types of magnetic resonance spectroscopy such as electron paramagnetic resonance (EPR).

## Methods

In order to obtain high resolution NMR spectra from a subject *in vivo*, it is necessary to localize the NMR excitation and detection to a particular region of interest. There are two principal ways of achieving this. Of course, the NMR magnet has to be large enough to accommodate the animal or human subject, or at least that part which is being studied (e.g. leg or arm muscles) and therefore *in vivo* NMR experiments are usually conducted using horizontal bore magnets designed for MRI investigations. Some small animals and species such as invertebrates can be studied using conventional high resolution NMR spectrometers with specially developed MRI probes.

The simplest approach to NMR detection uses a RF transmitter and receiver coil placed on the surface of the subject inside the NMR magnet. The size of the coil and its siting determines which tissue will be examined, and the depth of penetration of the detection region can be effected by the strength of the RF pulse used. This method has been used extensively for $^{31}$P NMR spectroscopy. Surface coil methods when combined with field gradients can also give greater spatial discrimination. One example is the DRESS technique where a field gradient is applied perpendicular to the plane of the surface coil and a frequency selective RF pulse is used to excite only the region of the tissue where the Larmor frequency is achieved.

The second approach makes use of pulsed field gradients and techniques adapted from MRI. In MRI, a three-dimensional region of tissue (usually a slice of square cross-section) is excited selectively by a combination of RF pulses and field gradients and, according to the method used, an image is produced based upon parameters such as the proton density, relaxation times or diffusion coefficients of the water in that region. For *in vivo* spectroscopy, a different pulse sequence is used which gives the high resolution NMR spectrum from that active region. One of the first methods was termed ISIS (image selected *in vivo* spectroscopy). This could be the $^{31}$P or $^{1}$H NMR spectrum. If a $^{1}$H NMR spectrum is required then it is also necessary to suppress the strong resonance from water. In addition, if $^{1}$H NMR spectra are obtained by this means then the pulse sequence is usually based on some form of the spin–echo experiment in order to suppress the peaks from fast relaxing protons which occur in the less mobile components of the tissue such as proteins or membrane lipids. One of the popular methods is called PRESS (point resolved spectroscopy). More complicated methods are also used in which information on spectroscopy and imaging is obtained simultaneously, techniques such as chemical shift imaging (CSI) where up to 4-dimensional data sets are obtained in which there are up to three spatial axes and one spectroscopic axis. From such data it is possible to obtain spectra at any voxel in the image or alternatively to determine the spatial distribution of any given metabolite detected spectroscopically.

## Applications

### $^{31}P$ NMR spectroscopy

The first *in vivo* experiments were carried out using $^{31}P$ NMR spectroscopy. This nuclide is particularly favourable because it has 100% natural abundance, high intrinsic sensitivity, the number of organic phosphorus compounds is relatively few and such molecules are important in cellular processes particularly for energy balance.

Most studies have concentrated on muscle with much emphasis on the heart. However, a typical $^{31}P$ NMR spectrum obtained from a human forearm is shown in **Figure 1** using a surface coil. This shows the main features which comprise the three signals of nucleotide phosphates, in this case mainly ATP, and phosphocreatine. Saturation transfer experiments have been carried out on these resonances to determine the rate of phosphate exchange between the phosphocreatine and ADP which yields creatine and ATP. This reaction is catalysed by the important enzyme creatine kinase. In addition, a major area of research is the study of the effects of brain ischaemia, such as occurs after a stroke, on the $^{31}P$ NMR spectrum. In the brain, unlike muscle, peaks are also seen from a wider variety of phosphorus-containing species such as inorganic phosphate, phosphomonoesters such as AMP or glucose phosphate and phosphodiesters such as ADP or glycerophosphorylethanolamine.

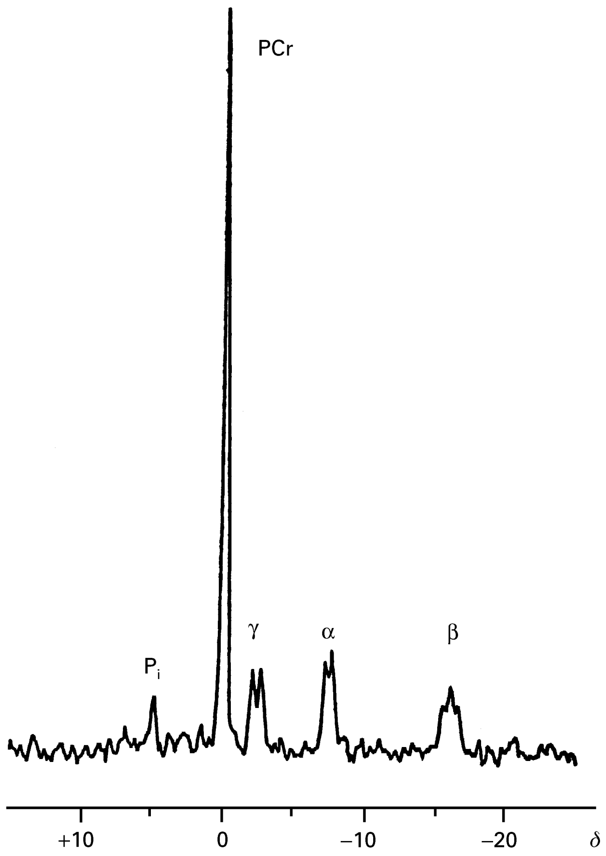



**Figure 1** 40.5 MHz $^{31}P$ NMR spectrum from a human forearm measured using a 4 cm diameter surface coil. $P_i$ – inorganic phosphate, PCr – creatine phosphate, α, β, and γ – phosphate groups at the α, β and γ positions in adenosine triphosphate. $^{31}P$–$^{31}P$ spin–spin coupling causes the splittings on the α, β and γ peaks. Chemical shifts are references to PCr at 0 ppm. Reproduced with permission from Friebolin H (1998) *Basic One- and Two-Dimensional NMR Spectroscopy*, 3rd edn. Weinheim: Wiley-VCH.

The chemical shift of inorganic phosphate is pH dependent and is thus a marker of intracellular pH. Typically, this is measured using an equation of the form

$$\mathrm{pH} = \mathrm{p}K_\mathrm{a} + \log\left(\frac{\delta - \delta_\mathrm{a}}{\delta_\mathrm{b} - \delta}\right)$$

where $pK_a$ is $-\log K_a$, the proton dissociation constant for the monoanionic species of phosphate, $\delta$ is the chemical shift difference between inorganic phosphate and phosphocreatine, and $\delta_a$ and $\delta_b$ are the chemical shifts of the diprotonated and monoprotonated forms of phosphate respectively. Usually, $pK_a$ is taken as 6.77 with $\delta_a$ and $\delta_b$ being 3.27 ppm and 5.78 ppm respectively, although different research groups use slightly different values.

### $^{1}H$ NMR spectroscopy

Nowadays, the bulk of publications on *in vivo* NMR spectroscopy are concerned with $^{1}H$ NMR spectroscopy of human brain. Many small molecule metabolites have been detected and much effort has gone into attempts to quantify these components. Many studies are concerned with altered levels of metabolites in various brain diseases. A typical $^{1}H$ NMR spectrum of human brain acquired *in vivo* is shown in **Figure 2**.

### $^{19}F$ NMR spectroscopy

*In vivo* $^{19}F$ NMR spectroscopy can be a very useful technique for monitoring the distribution of fluorinated drugs and their metabolites. These include the PET agent 2-fluoro-2-deoxyglucose, the study of its 3-fluoro-isomer as a probe of aldose reductase activity in brain, the elimination from the brain of fluorinated anaesthetics, metabolism of halothane in the liver, the distribution and catabolism of the anticancer drug 5-fluorouracil and the uptake of trifluoromethylthymidine into mouse tumours.

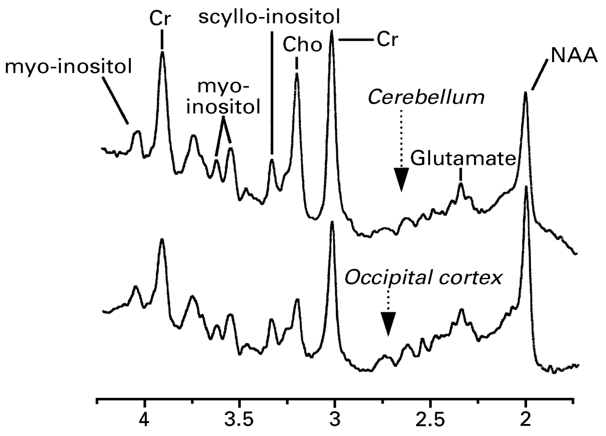



**Figure 2** $^1H$ NMR spectra from the brain of a healthy 43 year old male human volunteer. The upper spectrum is from the cerebellum and the lower spectrum is from the occipital cortex. The spectrum was measured at 170 MHz using a 4T magnet with a 15 cm diameter semi-volume coil. Localization was achieved using a pulse sequence based on the STEAM method with a nominal volume of 27 mL. Water peak suppression used selective Gaussian pulses. Cho – choline groups, Cr – creatine, NAA – *N*-acetylaspartate. Reproduced with permission from Seaquist ER and Gruetter R (1998) Identification of a high concentration of scyllo-inositol in the brain of a healthy human subject using $^1H$ and $^{13}C$ NMR. *Magnetic Resonance in Medicine* **39**: 313–316.

A number of fluorinated molecules which have pH-sensitive chemical shifts can be taken up by cells and used to monitor intracellular pH.

### $^{13}C$ NMR spectroscopy

A number of reports of the use of *in vivo* $^{13}C$ NMR spectroscopy are in the literature. These include studies of lipid deposits, glycogen and its breakdown in the liver and the monitoring of the metabolism of $^{13}C$-labelled substances such as $^{13}C$-glucose in various tissues such as the brain.

*See also:* **Biofluids Studied By NMR; $^{13}C$ NMR, Methods; Cells Studied By NMR; $^{19}F$ NMR, Applications, Solution State; MRI Applications, Biological; MRI Applications, Clinical; MRI Applications, Clinical Flow Studies; MRI Instrumentation; $^{31}P$ NMR; Solvent Suppression Methods in NMR Spectroscopy.**

## Further reading

Bailey IA, Williams SR, Radda GK and Gadian DG (1981) A $^{31}P$ NMR study of the effects of reflow on the ischaemic rat heart. *Biochemical Journal* **196**: 171–178.

Berliner LJ and Reuben J (eds) (1992) *In Vivo Spectroscopy, Biological Magnetic Resonance,* vol 11. New York: Plenum.

Bottomley PA, Foster TB and Darrow RD (1984) Depth-resolved surface coil spectroscopy (DRESS) for *in vivo* $^1H$, $^{31}P$ and $^{13}C$ NMR. *Journal of Magnetic Resonance* **59**: 338–342.

Duyn JH, Frank JA and Moonen CTW (1995) Incorporation of lactate measurement in multi-spin-echo proton spectroscopic imaging. *Magnetic Resonance in Medicine* **33**: 101–107.

Hoult DI, Busby SJW, Gadian DG, Radda GK, Richards RE and Seeley PJ (1974) Observation of tissue metabolites using $^{31}P$ nuclear magnetic resonance. *Nature (London)* **252**: 285–287.

Ordidge RJ, Connelly A and Lohman JAB (1986) Image-selected *in vivo* spectroscopy (ISIS) – a new technique for spatially selective NMR spectroscopy. *Journal of Magnetic Resonance* **66**: 283–294.

Pettegrew JW (ed) (1990) *NMR: Principles and Applications to Biomedical Research.* New York: Springer-Verlag.

Ross BD, Radda GK, Gadian DG, Rocker G, Esiri M and Falconer-Smith J (1981) Examination of a case of suspected McArdle's syndrome by $^{31}P$ nuclear magnetic resonance. *New England Journal of Medicine* **304**: 1338–1342.

Sillerud LO and Shulman RG (1983) Structure and metabolism of mammalian liver glycogen monitored by $^{13}C$ nuclear magnetic resonance. *Biochemistry* **22**: 1087–1094.

van Zijl PCM, Chesnick AS, Despres D, Moonen CTW, Ruiz-Cabello J and van Gelderen P (1993) *In vivo* proton spectroscopy and spectroscopic imaging of (1-$^{13}C$)-glucose and its metabolic products. *Magnetic Resonance in Medicine* **30**: 544–551.

# Indium NMR, Applications

*See* **Heteronuclear NMR Applications (B, Al, Ga, In, Tl)**

# Induced Circular Dichroism

**Kymberley Vickery** and **Bengt Nordén**, Chalmers University of Technology, Gothenburg, Sweden

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Applications**

Induced circular dichroism (ICD) is the CD observed in an optically inactive (achiral) chromophore due to its interaction with an optically active (chiral) moiety. Early applications exploited ICD just as evidence for such an interaction. For example, the ICD in the absorption band of an achiral drug may be used to prove its binding to a (chiral) protein or nucleic acid. More recent use of ICD includes the structural information that theoretical analysis of the inducing mechanism can provide. ICD has been an important complement to linear dichroism (LD) for the assessment of binding geometries in drug–nucleic acid systems.

The CD induced in an achiral chromophore upon interaction with a chiral substrate can be ascribed to one or several mechanisms, each corresponding to a more or less idealized situation: the chromophore molecule may be perturbed, by mainly steric interactions, to adopt a chiral (equilibrium) conformation or it may gain optical activity through electric and/or magnetic interactions with surrounding chromophores.

## Different types of ICD

Strictly, ICD encompasses all kinds of circular dichroism effects except for those of inherently chiral chromophores. This is so because most systems (molecules) may be more or less arbitrarily divided into one achiral chromophore in which the ICD is observed and a remaining component that includes the chirality. Thus, the CD exhibited by a chiral molecule may be considered ICD if the chromophoric entity itself is achiral. With this broader definition of ICD, molecules such as carbonyl compounds (in which the carbonyl chromophore is achiral) display an intramolecular ICD in the first allowed $n \rightarrow \pi^*$ transition around 300 nm due to perturbation from a surrounding chiral (e.g. steroid) skeleton. Distinction between this type of molecule and one which is an inherently chiral chromophore could be difficult to make as the boundaries of the chromophore are often diffuse. For example, when a diene is in immediate proximity of a carbonyl, a composite chromophore is obtained, which, if skewed, is inherently chiral. Therefore, it is practical to restrict the definition of ICD to cases in which the achiral part is distinctly separable from the chiral part, generally by being a separate molecular species.

ICD could arise in an achiral chromophore mainly through two mechanisms: (1) ICD of an electric dipole-allowed (eda) transition due to coupling with other eda transitions of the surrounding (chirally arranged) chromophore(s) or (2) ICD of a magnetic dipole-allowed (mda) transition due to mixing with eda transitions in the same chromophore under the perturbation of the electronic states of the chiral surrounding. The latter mechanism generally gives more easily detectable ICD as a combined effect of typically strong CD and weak absorption of mda transitions (CD/absorbance sets the limit for the signal-to-noise ratio of CD detection). In addition to carbonyl groups in organic compounds, d–d transitions of transition metals provide examples where strong ICD from interaction with chiral surroundings are frequently observed. For example, the first d–d transition of $[Co(NH_3)_6]^{3+}$ displays ICDs of varying sign and magnitude due to hydrogen-bonded outer-sphere coordination with tartrate, amino acids and other chiral oxoanions. These ICD effects have been exploited to assess binding constants of the outer-sphere complexes formed, the strength of the ICD and of the binding constants suggesting that they are due to relatively well ordered structures. However, attempts towards more detailed structural assignments have been unsuccessful owing to the complicated nature of this mechanism of induced CD. For this reason, the focus here will be on the first (electric dipole) ICD mechanism, which, despite its weaker ICD signals and consequently more demanding experiments, can be more simply related to structural models.

## Interactions between electric dipole-allowed transitions

This mechanism of ICD is related to the strong exciton CD that is observed for a chirally arranged set of identical intense absorbers such as in a chiral homodimer (or polymer) species. A characteristic feature is the strong bisignate CD doublet observed at the frequency of the transition. The dominant CD bands centred around 200 nm in proteins and at

260 nm in nucleic acids are both of exciton-type $\pi \rightarrow \pi^*$ transitions. Exciton CD will not be considered an ICD as the transition is not localized on a single achiral chromophore but is spread over the whole dimer (polymer) system and thus constitutes an inherently chiral chromophore. Instead, the ICD arises from the interaction between two different (nondegenerate) chromophores. It is generally of a much weaker magnitude than the exciton CD, but becomes gradually more intense as the energy between the two transitions decreases. This is seen from the energy difference in the denominator of the expression (Eqn [1]) for the ICD rotational strength $R$(A) for an eda transition, with transition moment $\mu$(A) occurring at frequency $\nu$(A) on chromophore A, due to interaction with another transition $\mu$(B) at frequency $\nu$(B) on chromophore B of the chiral environment. A requirement for non-vanishing $R$(A) is that the transition moments $\mu$(A) and $\mu$(B) are chirally arranged relative to each other, that is, they must not be coplanar (**Figure 1**).

$$\boldsymbol{R}(\mathrm{A}) = \frac{-2\pi\nu_\mathrm{A}\nu_\mathrm{B}}{hc(\nu_\mathrm{B}^2 - \nu_\mathrm{A}^2)} V_\mathrm{BA} O_\mathrm{BA} \qquad [1]$$

where $V_{\mathrm{BA}}$ is the corresponding dipole–dipole interaction energy:

$$V_\mathrm{BA} = \frac{1}{4\pi\varepsilon_0}\frac{1}{|\boldsymbol{R}_\mathrm{BA}|^3}\left[(\boldsymbol{\mu}_\mathrm{A}\cdot\boldsymbol{\mu}_\mathrm{B}) - \frac{3}{|\boldsymbol{R}_\mathrm{BA}|^2} \times (\boldsymbol{\mu}_\mathrm{A}\cdot\boldsymbol{R}_\mathrm{BA})(\boldsymbol{\mu}_\mathrm{B}\cdot\boldsymbol{R}_\mathrm{BA})\right] \qquad [2]$$

where $\mu_\mathrm{A}$ and $\mu_\mathrm{B}$ are the two (unperturbed) electronic transition moments, $\varepsilon_0$ is the permittivity of vacuum, and $R_{\mathrm{BA}}$ is the vector from the centre of A to the centre of B. $O_{\mathrm{BA}}$ in Equation [1] is an 'optical factor' which depends on the geometrical arrangements of the two electronic transition moments relative to each other:

$$O_\mathrm{BA} = \boldsymbol{R}_\mathrm{BA}\cdot(\boldsymbol{\mu}_\mathrm{A}\times\boldsymbol{\mu}_\mathrm{B}) \qquad [3]$$

A decreased difference between $\nu$(A) and $\nu$(B) in the denominator of Equation (1) would make $R$(A) increase; consequently, ICD of bands close to the inducing band would be strong. The ICD further depends on the geometry of the A–B pair, both through $V_{\mathrm{BA}}$ (Eqn [2]) and $O_{\mathrm{BA}}$ (Eqn [3]), which could be described by four coordinates: the length of

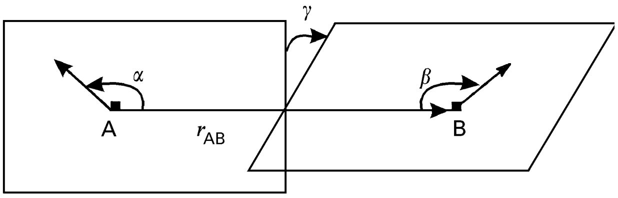



**Figure 1** Geometry and coordinates of the A–B system. Transition moments $\mu_\mathrm{A}$ and $\mu_\mathrm{B}$ of chromophores A and B and the separation vector $\boldsymbol{R}_{\mathrm{BA}}$.

the separation vector $R_{\mathrm{BA}}$ and three angles $\alpha$, $\beta$, and $\gamma$ defining how $\mu$(A) and $\mu$(B) are oriented relative to $R_{\mathrm{BA}}$. As seen from Equations [2] and [3], the interaction factor $V_{\mathrm{BA}}$ and the optical factor $O_{\mathrm{BA}}$ of Equation [1] are both dependent on all four coordinates. The ICD drops off as $R_{\mathrm{AC}}^{-2}$ as a combined effect of the $R_{\mathrm{BA}}$ dependences in Equations [2] and [3]. Furthermore, the ICD is proportional to the dipole strengths $\mu_\mathrm{A}^2$ and $\mu_\mathrm{B}^2$, i.e. to absorption of both A and B chromophores:

$$\boldsymbol{R}(\mathrm{A}) = -\frac{2\pi\nu_\mathrm{A}\nu_\mathrm{B}(\boldsymbol{\mu}_\mathrm{A}\boldsymbol{\mu}_\mathrm{B})^2}{hc(\nu_\mathrm{B}^2 - \nu_\mathrm{A}^2)\boldsymbol{R}^2{}_\mathrm{BA}}(\sin\alpha\sin\beta\cos\gamma + 2\cos\alpha\cos\beta)\sin\alpha\sin\beta\sin\gamma \qquad [4]$$

Generally, the achiral chromophore A interacts with several chromophores ($\mathrm{B}_i$) of the chiral substrate. If $\nu_\mathrm{A} << \nu_\mathrm{B}$, the ICD may be regarded as a sum of pairwise contributions, each corresponding to one term in Equation (1) with a given set of $\mathrm{B}_i$, $V_{\mathrm{B}_i\mathrm{A}}$ and $O_{\mathrm{B}_i\mathrm{A}}$. In many molecular systems, such as DNA and cyclodextrins, the symmetric distribution of B chromophores makes it possible to deduce simple expressions for the net ICD.

The following examples will focus on studies of DNA-induced circular dichroism; however the principles are general, based on the application of Equations [1]–[4].

## Detection and quantitation of DNA–drug complex

Many DNA-binding molecules (usually referred to as ligands or drugs although they may have little or no therapeutic value) are themselves achiral. However, upon binding to DNA, they acquire an ICD that is characteristic of their interaction. When considering CD applications of proteins and DNA, the ligand ICD can be used at a number of levels. The simplest use is to note that it exists, and therefore that the molecule *does* bind to the chiral substrate; however, more information can usually be extracted.

## Empirical use of ICD for binding site analysis

It is often useful to measure a series of spectra where some variable such as temperature, ionic strength, or mixing ratio (drug: DNA ratio) is changed. If the CD intensity changes but the shape of the spectrum remains constant during such an experiment, this indicates that the ligand binding mode is unchanged although the amount of bound ligand may have changed. If, however, there is a change in the shape of the spectrum as an experimental parameter is changed, then a change in the DNA–drug interaction as a function of that parameter is implied. The change is frequently due to occupancy of more than one binding site as the drug load of the DNA increases. Alternatively, it may also be due to changes in DNA conformation or to drug–drug interactions. In other words, if the experiment is conducted with a constant ligand concentration and varying DNA concentration, then unless the ligand binding mode is changing, there should be no change in the observed ligand ICD. A particularly dramatic change, due to exciton CD, is observed if, as the DNA concentration is decreased, the ligands begin to stack or self-associate with each other.

## Extended treatment for structural analysis

The important case of coupled oscillator (exciton) CD of interacting, identical chromophores (i.e. degenerate transitions) has already been mentioned. Exciton CD was early observed for dye molecules upon association to biopolymers, attributed to interactions between adjacently bound dimolecules with chiral orientation relative to each other. The less exploited case of coupling of electric dipole transition of the ligand molecule A due to interaction with one or several *nondegenerate* transitions of the chiral molecular system B has been considered in terms of Equations [1]–[4]. Because of the nondegeneracy, this CD is generally very small and should be measured at low drug:biopolymer ratios to avoid dominance by the much more intense exciton CD from drug–drug interactions, DNA-induced monomer CD is exhibited by, for example, intercalating dyes such as methylene blue, acridine orange and proflavin. The DNA-induced CD of several of these intercalators depends sensitively on both salt conditions and on base sequence. These effects have been interpreted in terms of varying binding geometries in the intercalation pocket or to varying inducing transitions of the nucleobases. For eda drug transitions which are energetically isolated from the inducing transitions of the DNA bases, a simple relation has been derived between the DNA-induced monomer CD (rotational strength $\boldsymbol{R}$) and the 'azimuthal' angle γ, defining the orientation of the intercalator transition in the plane of the intercalation pocket (angle to dyad axis):

$$\boldsymbol{R} = f|\boldsymbol{\mu}|^2(1 - 2\cos^2\gamma) \qquad [5]$$

Where $f$ is a function comprising contributions of the various transitions of the surrounding nucleobases. The DNA-induced CD of an intercalator has also been calculated using the more rigorous 'matrix method'. The orientation dependence obtained in this manner (**Figure 2**) was found in good comparison with the $(1-2\cos^2\gamma)$ dependence predicted by Equation [2]. The CD amplitudes were also in reasonable agreement with measured ICD spectra of various intercalators which can be seen in **Table 1**.

The situation is more complicated for a nonintercalator than for an intercalator in that there will be more degrees of geometrical freedom (**Figure 2**) and, as a result of the external position, interactions with a larger number of DNA bases must to be considered.

## Groove-bound and intercalator ligand–DNA complex geometries

The relative geometries of DNA and the transition moment of an externally bound chromophore are defined by the parameters $r$, $\alpha$, $\beta$, and $\gamma$, in **Figure 2B**. $r$ and $\gamma$ define the *position* of the external chromophore's transition moment relative to the DNA helical axis and $\alpha$ and $\beta$ describe the *orientation* of the transition moment; $\alpha$ is the angle between the $z$-axis (DNA helical axis) and the transition moment and $\beta$ is the angle between the projection of the transition moment onto the $xy$ plane and the vector $r$ the distance from the helix axis to the position of the transition moment.

In **Figure 3**, an example is shown of the ICD calculated for a nonintercalator ($r$ = 7 Å) sitting outside a GC step in the minor groove, as a function of the $\alpha$ and $\beta$ angles. $\alpha = 45°$ and $\beta = 90°$ correspond to a minor groove binding geometry (indicated by an ×), with a transition moment pointing along the groove. This geometry should give a strong positive ICD. However, $\alpha = 90°$ and $\beta = 0°$, which correspond to a transition moment along the duplex dyad axis of an intercalator which has been withdrawn from the centre, should produce a negative ICD. A strong positive ICD appears to be a common feature for minor groove binders with transition moments parallel with the groove. However, depending on the DNA

**Table 1** Calculated and observed ICD rotational strengths for DNA intercalators and externally bound drugs

| *Drug system*[a] | *Energy of transition* ($10^3$ *cm*$^{-1}$) | *Transition moment angles (Fig. 2)/°* $\alpha$ | $\beta$ | $\gamma$ | *Rotational strength (DBM units)* *Expt.* | *Calc./predicted* |
|---|---|---|---|---|---|---|
| **Intercalators** | | | | | | |
| DAPI–[Poly(dC–dG)]$_2$ | 30 | 80 | 0 | 0 | +0.01 | +0.01 |
| Acridine orange–DNA | 20 | 90 | 0 | 90 | − | − |
| Pseudoisocyanine–DNA | 18 | 90 | 0 | 90 | − | − |
| APF–[poly(dC–dG)]$_2$ | 28 | 80 | 0 | 0 | + | + |
| **Minor groove binders** | | | | | | |
| DAPI–[poly(dA–dT)]$_2$ | 30 | 45 | 0 | 90 | +1.8 | +0.5 |
| Netropsin–DNA | 32 | 45 | 0 | 90 | +1.3 | +0.4 |
| Hoechst 33258–DNA | 26 | 45 | 0 | 90 | +2.8 | +0.4 |
| APF–[poly(dA–dT)]$_2$ | 28 | 45 | 0 | 90 | + | + |
| **Major groove binders** | | | | | | |
| Methyl green–DNA | 23 | 90 | 180 | 0 | −0.6 | −0.1 |
| | 15 | 45 | 180 | 270 | +0.8 | +0.2 |

[a] DAPI = 4,6-diamidino-2-phenylindole and APF = 2,5-bis(4-amidinophenyl)furan.

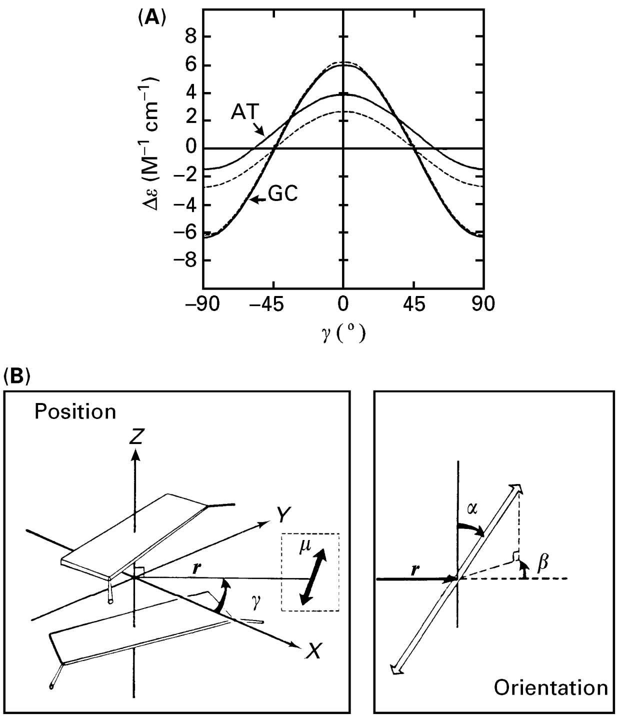



**Figure 2** (A) ICD of a DNA intercalator calculated as a function of the transition moment in the plane of the intercalation pocket and centred on the helix axis. $\gamma$ is the angle between the transition moment and the DNA dyad axis. Reproduced with permission from Lyng, Hard and Nordén (1987) *Biopolymer* **26**: 1327. (B) definition of geometrical parameters of a DNA–ligand system: with $r = 0$, $\alpha = 90°$ and $\beta = 0°$, then $\gamma$ corresponds to the orientation of an intercalator transition moment (see top figure). Typical minor groove binding geometry parameters are $r = 7$ Å, $\alpha = 45°$, $\beta = 0°$ and $\gamma = 0$.

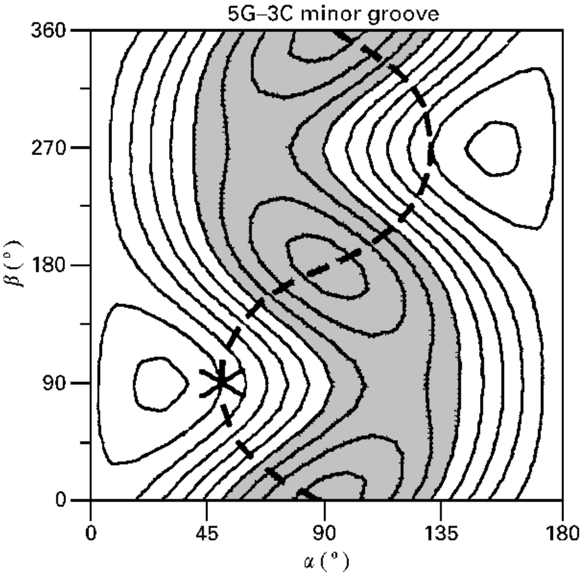



**Figure 3** Example of ICD calculated for a minor groove binder (at 5′G–3′C step) as a function of angles $\alpha$ and $\beta$ (see **Figure 2**). Reproduced with permission of Wiley from Lyng, Hard and Nordén (1991) *Biopolymers* **31**: 1709.

sequence, both the sign and amplitude are predicted to vary for other nonintercalated and semi-intercalated geometries.

As can seen in **Table 1**, DAPI and APF, which are both known to bind preferentially, in the minor groove at AT-rich regions of duplex DNA, can also bind by intercalation in pure GC tracts (this difference is most clearly seen in the flow linear dichroism spectrum, which probes the $\alpha$ angle). Thus, **Figure 4** shows three types of DNA ICD. Spectrum (a) shows the APF [2,5 bis(4-amidinophenyl)furan] ligand bound in the minor groove of [poly(dAdT)$_2$]. Spectrum (f), obtained at a higher drug: DNA binding ratio, also contains a contribution from exciton CD

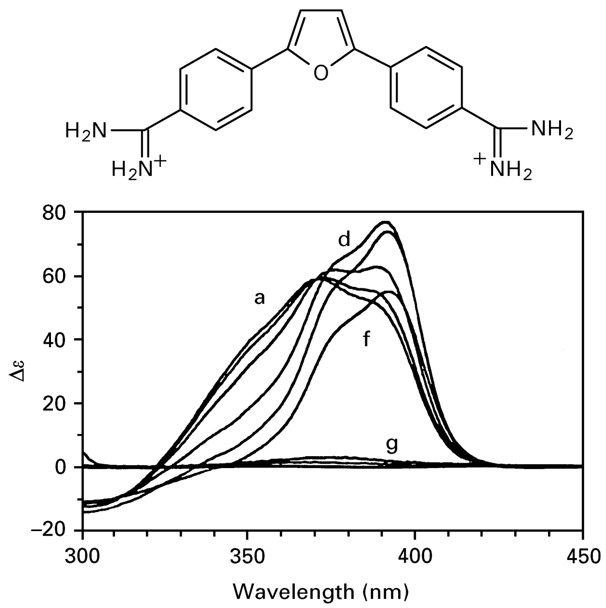



**Figure 4** Measured ICD of APF (see formula) bound to [poly(dAdT)]$_2$ [spectra (a)–(f)] and to [poly(dGdC)]$_2$ [spectrum(g)]. The strongly positive ICD in spectrum (a) (drug: DNA phosphate ratio $r = 0.01$) is consistent with minor groove binding (see **Figure 3**), while increasing binding ratios ($r = 0.02$, 0.04, 0.08, 0.16 and 0.31) give a gradually increasing contribution of CD due to exciton coupling between adjacently bound APF molecules. The weak positive ICD in [poly(dGdC)$_2$] is consistent with intercalation (transition moment along the DNA dyad axis, see **Figure 2**). Reproduced with permission of the American Chemical Society from Jansen, Lincoln and Nordén (1993) *Biochemistry* **32**: 6605.

due to drug–drug interactions between adjacent sites in the minor groove. Spectrum (g), which refers to APF in [poly(dGdC)$_2$], is the weak positive CD corresponding to the intercalated drug with the transition moment preferentially directed along the short (dyad) axis of the intercalation pocket.

## Conformational ICD in DNA systems

When chromophore A could adopt a chiral conformation and B could change its conformation upon binding, some additional sources of ICD induced by chromophore B should be considered. The ICD is generally assumed to be the difference

$$\mathrm{ICD} = \Delta\mathrm{CD} = \mathrm{CD(B+A)} - \mathrm{CD(B)} - \mathrm{CD(A)} = 0 \qquad [6]$$

where A = achiral molecule (drug) and B = chiral molecule (DNA). However, the net ΔCD may also have a contribution due to a change of conformation in B:

$$\Delta\mathrm{CD} = \mathrm{ICD(A)} + \delta\mathrm{CD(conf\ B)} \qquad [7]$$

where $\delta$CD(conf B) = CD change of B due to conformational change of B upon interaction with A.

Since ICD(A) is the true ICD, providing information about the interaction, it is recommended that one corrects, if possible, for any contributions due to conformational change of B[ΔCD − $\delta$CD(conf B)]. Three contributions to the ICD of A may be distinguished:

$$\mathrm{ICD(A)} = \mathrm{ICD(A)_{el}} + \delta\mathrm{CD(conf\ A)} + \delta\mathrm{CD\ (exciton\ A)} \qquad [8]$$

where ICD(A)$_{el}$ = induced CD in A's transitions due to electronic coupling with B's transition's, $\delta$CD(conf A) = CD of A due to adopted chiral conformation, and $\delta$CD(exciton A) = exciton CD of A interaction with neighbour A molecules.

*Special situation (a).* When A is a chiral molecule added in enantiomerically pure form, then ICD(A) = induced CD in A's transitions due to electronic coupling with B's transition + $\delta$CD(chir) of A due to any changed (chiral) conformation of A + $\delta$CD(Abs) of A due to any changed absorption property (hypo-/hyperchromism) of A. The third contribution scales directly as the changed absorption, $\delta$Abs:

$$\delta\mathrm{CD(Abs)} = \frac{\delta\mathrm{Abs} \times \mathrm{CD(A)}}{\mathrm{Abs}}$$

Hence if CD(A) is large, δCD(Abs) will be large, even for small $\delta$Abs.

*Special situation (b).* A is a chiral molecule added as a racemate: then, in addition to the contributions in situation (*a*), there is additionally the effect of any enantiopreferentiality of binding A to B, which adds the following extra term to ICD(A): + $\delta$CD of A due to that the enantiomer bound may have changed its absorption properties (hypo/hyperchromism) or conformation.

*Special situation (c).* A is a chiral molecule subject to inversion. The enantiopreferential binding to B can then affect the equilibrium between the enantio-mers to make the total of A non-racemic (an excess of one enantiomer). Generally, the preferred enantiomer is thermodynamically stabilized. This is the so-called 'Pfeiffer effect'. An example is given by the addition of labile $[Fe(phen)_3]^{2+}$ to DNA where the observed

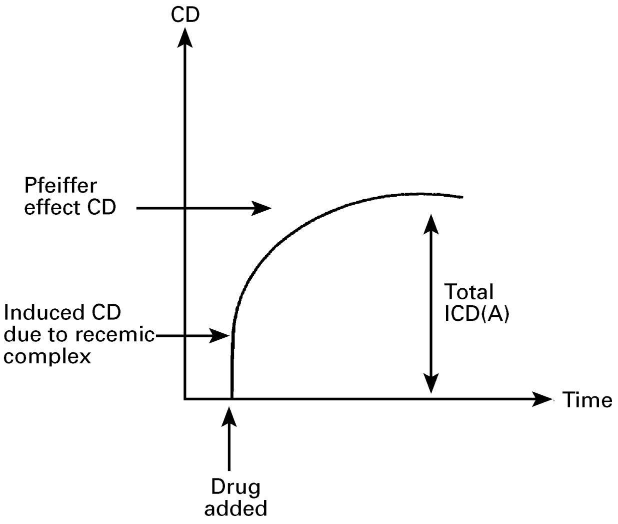



**Figure 5** The 'Pfeiffer' ICD of inversion labile $[Fe(phen)_3]^{2+}$ upon binding to DNA. Following DNA addition, a small ICD is immediately seen due to the preferential association of one enantiomer and electronic perturbation of its CD [*special situations (a) and (b)*]. Therefore, a strong CD is more slowly evolving as a result of a shifted diastereomeric equilibrium giving net excess to one enantiomer.

ICD is comprised of two components. There is an immediate effect according to situation (b) and a slower developing (stronger) Pfeiffer ICD due to an inversion yielding a net excess of the enantiomer in the solution, as shown in **Figure 5**.

## ICD in cyclodextrins

The CD of large, formally rather complicated systems is sometimes as simple as that of much smaller ones, particularly if the system has high symmetry. One such system is a $\beta$-cyclodextrin host–guest complex. Cyclodextrins are $\alpha$-1,4-linked D-glucose oligomers, where $\alpha$-cyclodextrins have six glucose units and $\beta$-systems have seven and so on. Since cyclodextrins have been found to catalyse reactions involving their guests, it is of interest to know the orientation of the guest inside the sugar ring (**Figure 6**). The CD induced into the transitions of known polarization of the guest may show that immediately. Most experiments have been conducted using the $\beta$-cyclodextrin framework, which has a cavity of the correct size to accommodate a single aromatic molecule such as benzene or naphthalene. Another application of cyclodextrin-induced CD is for the assignment of transitions in $C_{2v}$ or $D_{2h}$ symmetric molecules: ICD of opposite signs correspond to transitions with mutually perpendicular polarizations.

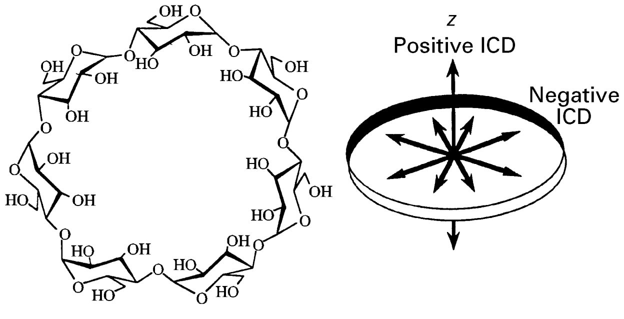



**Figure 6** $\beta$-Cyclodextrin (left) and the signs of the CD observed in the transitions of a guest molecule depending on their polarization directions: axially polarized transitions appear with positive ICD, whereas transitions polarized in the plane of the cyclodextrin ring appear with negative ICD. Reproduced with permission from Roger A and Nordén B (1997) *Circular Dichroism and Linear Dichroism*, Oxford: Oxford University Press.

## List of symbols

$O$ = optical factor; $r$ = distance from helix axis to position of transition moment; $r$ = drug:DNA phosphate ratio; $R$ = rotational strength; $V$ = dipole–dipole interaction energy; $\alpha$ = angle; $\beta$ = angle; $\gamma$ = angle; $\varepsilon_0$ = permittivity of vacuum; $\mu$ = transition moment; $\nu$ = frequency.

*See also:* **Biomacromolecular Applications of Circular Dichroism and ORD; Chiroptical Spectroscopy, Emission Theory; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Chiroptical Spectroscopy, General Theory; Macromolecule-Ligand Interactions Studied By NMR; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Nucleic Acids Studied Using NMR; Pharmaceutical Applications of Atomic Spectroscopy.**

## Further reading

Harada N and Nakanishi K (1983) *Circular Dichroism Spectroscopy: Exciton Coupling in Organic Stereochemistry*. Mill Valley, CA: California University Science Books.

Mason SF (1982) *Molecular Optical Activity and the Chiral Discrimination*. Cambridge: Cambridge University, Press.

Nakanishi K, Berova N and Woody RW (eds) (1994) *Circular Dichroism Principles and Applications*, New York: VCH.

Nordén B, Kubista M and Kurucsev T (1992) *Quarterly Reviews of Biophysics* **25**: 51–170.

Rodger A, and Nordén B (1997) *Circular Dichroism and Linear Dichroism*. Oxford: Oxford University Press.

# Inductively Coupled Plasma Mass Spectrometry, Methods

**Diane Beauchemin**, Queen's University, Kingston, Ontario, Canada

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Inductively coupled plasma mass spectrometry (ICP-MS) combines an ICP with a mass spectrometer. It features a multielemental capability, good precision, a long linear dynamic range, simple spectra, low detection limits and the ability to do rapid isotopic analysis. With such advantages, ICP-MS has found wide application to the analysis of a variety of samples, including water and food, as well as geochemical, environmental and biological samples. However, it has a number of limitations, such as matrix effects, which can be important, and the need to keep the concentration of dissolved solids low to avoid clogging problems. These complicate the analysis by requiring sample pre-treatment and/or more involved calibration strategies. Nonetheless, with the appropriate method development and/or optional accessories, these limitations can be largely circumvented, thereby allowing ICP-MS to be applied to the analysis of virtually any type of sample.

## Description of a basic ICP-MS instrument

**Figure 1** shows a block diagram of an ICP-MS instrument. A 1–2.5 kW Ar ICP is the typical ion source for MS. The ICP is a high-temperature (5000–10 000 K) electrodeless discharge that is sustained in argon flowing within a torch. In general, the ICP uses samples in solution, which presents the advantages of good control over homogeneity and ease of calibration. A sample solution is delivered to a nebulizer using a peristaltic pump to minimize physical interferences (from, for instance, changes in viscosity). The nebulizer converts this solution into an aerosol that is then carried by the nebulizer gas through a spray chamber where large droplets condense and drain out. The resulting fine aerosol is injected into the heart of the plasma, and undergoes several sequential processes as it moves deeper into the plasma: desolvation, vapourization, atomization and ionization.

In an Ar ICP, the degree of ionization of 52 elements is expected to be $\geq 90\%$. Only three elements (He, F and Ne) which possess a first ionization potential greater than that of Ar would not be ionized and could therefore not be determined by ICP-MS with an Ar ICP. Similarly, the highest degree of double ionization is 10% and occurs only for a few elements. The ICP is therefore an efficient elemental ion source for MS since the majority of elements in the periodic table are singly ionized.

A portion of the plasma (which is typically three sampler orifice diameters wide and two sampler orifice diameters deep) is sampled from the central channel of the plasma, and extracted through a differentially pumped interface (through two water-cooled metal cones with orifice diameter ~1 mm, the sampler and the skimmer), which is maintained at approximately 1 torr using a mechanical roughing pump. This portion is then transmitted into the mass spectrometer where ions are separated according to their mass-to-charge ratios ($m/z$) and counted.

Two different categories of ICP-MS instruments are currently commercially available: low-resolution instruments, using either a quadrupole or a time-of-flight (TOF) tube to achieve ion separation, and double-focusing high-resolution instruments that use an electrostatic and a magnetic analyser in series. Both the quadrupole-based and the double-focusing instruments allow a sequential multielement measurement, whereas TOF-ICP-MS allows simultaneous multielement detection.

Selected specifications for these two categories are summarized in **Table 1**. Aside from resolution, one distinctive feature of high-resolution instruments is the negligible background which is observed across the mass range. This results from the curved path that the ion beam must follow, which efficiently prevents photons from reaching the detector, as well as from the operating pressure that is considerably lower than that of low-resolution instruments, and thereby minimizes collisional processes in the ion beam.

## Capabilities of ICP-MS

**Table 2** shows that ICP-MS has numerous advantageous features: large linear dynamic range,

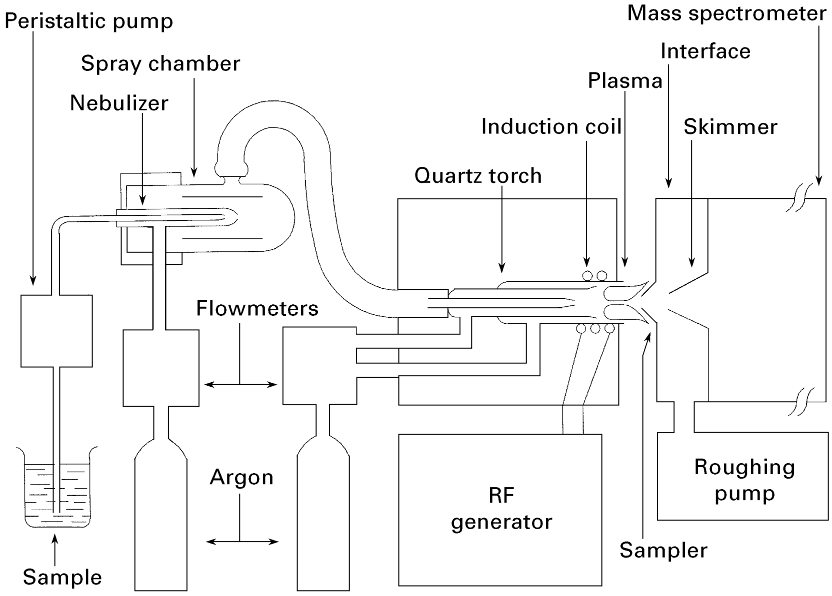



**Figure 1** Schematic representation of a conventional ICP-MS instrument.

**Table 1** Typical specifications of ICP-MS instruments

| | *Type of instrument* | |
|---|---|---|
| *Parameter* | *Low resolution* | *High resolution* |
| Resolution | 1 amu[a] | 300–10 000 amu[b] |
| Mass range | 1–300 amu | 1–500 amu |
| Background level across the mass range | < 2–50 cps | << 0.1 cps |

[a] Peak width at 10% from the peak maximum.
[b] 10% valley between two adjacent peaks.

**Table 2** Attributes of ICP-MS instruments

| | *Type of instrument* | |
|---|---|---|
| *Feature* | *Low resolution* | *High resolution* |
| Orders of linear dynamic range | 5–9 | 9 |
| Short term precision for 10 repeats (%) | 1–3 | 2 |
| Sensitivity (Mcps/ppm) | 1–30 | 20 at $R = 300$<br>1 at $R = 3\,000$ |
| Detection limit[a] (ng/L) | <1–10 | <0.01–1 ($R$ = 300–500) |

[a] Based on three times the standard deviation of the blank.

advantage of working with samples in solution, good precision, low detection limits. The linear dynamic range for many elements is such that major, minor and trace levels may all be determined at once. Furthermore, the multielement capability of ICP-MS allows the determination of, for instance, 70 elements in less than 2 min, using less than 2 mL of solution at an uptake rate of 1 mL $min^{-1}$. Detection limits are typically ≤ 1 ng $L^{-1}$ for most elements with all low-resolution instruments, and even lower with high-resolution instruments. This difference arises from the lower background typically obtained on high-resolution instruments, as well as the higher reliability in peak hopping which results from the flat-top shape of the peaks observed in high resolution. Sensitivity is similar on all instruments in low resolution, but is degraded by an increase in resolution because this is typically achieved by decreasing the slit width of the entrance and exit slits on double-focusing instruments, which in turn leads to a reduction in ion transmission.

In any case, since specific isotopes are detected, isotope ratios can be measured. Also, calibration by the very accurate isotope dilution analysis method can be done (where internal standardization is essentially carried out using an enriched isotope of the analyte). The flat-top peaks, which are typical of high-resolution instruments, should allow the measurement of ratios with better precision than that achievable with quadrupole-based instruments. However, because these mass spectrometers cannot be scanned as fast as quadrupoles, the resulting precision can be degraded compared to that obtained with quadrupole MS. The best precision is expected with TOF instruments since they allow the simultaneous detection of all masses.

All these features explain why even the basic ICP-MS instrument described so far is being applied to analyses in a variety of areas. For instance, **Table 3** shows that ICP-MS is widely used for environmental,

**Table 3** Examples of applications of ICP-MS

| *General category* | *Samples* | *Sample pretreatment* |
|---|---|---|
| Water | Sea water, industrial waste water, pore water, river water, fresh water, rain, ground water, lake water, spring water, snow, tap water, high-purity water | Filtration and any of the following: acidification and dilution; ion exchange separation to remove interferents and/or preconcentrate analytes; preconcentration by evaporation |
| Environmental | Soil, dust, leaves, sediments, sewage, industrial effluents, paint, atmospheric aerosols, suspended particulate matter, tobacco smoke condensate, domestic sludge, glass, spent nuclear fuel, leachates of high-activity waste, rocks, metals, ores, alloys, ceramic materials | Crushing or grinding, and either digestion with high-purity acids or fusion and dissolution, followed by dilution |
| Food | Prawn parts, wheat flour, spinach, cabbage, wine, mussels, beef kidney, milk | Digestion with high-purity acids followed by dilution |
| Biological | Faecal material, urine, human serum, plasma, red blood cells, whole blood, cerebrospinal fluids | Dilution or drying, ashing and/or digestion with high-purity acids followed by dilution or extraction |
| Organic | Crude oil, gasoline | Extraction into aqueous acidic solution |

clinical and geochemical studies, with appropriate sample preparation procedures. These are required not only to put solid samples in solution, but also to avoid or circumvent limitations of ICP-MS by, for instance, removing sources of interference.

## Limitations of ICP-MS

**Table 2** indicates that, whatever the resolution, detection limits vary with the element being determined. There are, indeed, several factors affecting detection limits in ICP-MS. The degree of ionization of the analyte, the natural abundance of the isotope that is used for the determination of a particular element, spectroscopic interferences from the background and/or the sample matrix, and mass discrimination are the most notorious sources of degradation. Obviously, if an element is only 50% ionized in the plasma, then its detection limit will be reduced by a factor of 2 compared to an element which is completely ionized. Similarly, if only one isotope of a multi-isotope element is detected, the resulting detection limit will be less than that of a monoisotopic element.

Furthermore, the sample introduction system is the Achilles' heel of all ICP-based instruments. Indeed, only about 2% of the sample solution actually reaches the plasma with a conventional system (composed of a pneumatic nebulizer and spray chamber, as shown in **Figure 1**), most of it going down the drain. This means, therefore, that only a small fraction of the total possible signal is obtained.

As mentioned earlier, because the mass spectrum is simple, the likelihood of spectroscopic interferences is reduced, particularly with high-resolution instruments. However, when they occur, the number of alternative lines is small (there are indeed a number of monoisotopic elements). One example of an isobaric interference that cannot be resolved even at high resolution is that of $^{40}Ar^+$ on $^{40}Ca^+$, the major isotope of Ca. Other severe interferences arise from Ar-containing polyatomic ions as well as oxides of matrix elements. Several of these can be resolved at higher resolution but with a concurrent sacrifice in sensitivity and detection limits. On low-resolution instruments, a 'cold plasma', where the ICP is operated at lower power and higher nebulizer gas flow rate, can be used to reduce important interferences from the background but, again, a lower sensitivity concurrently results across the mass range.

Mass discrimination is most important for quadrupole-based instruments where the transmission of ions is not uniform over the mass range. In particular, the transmission of light ions is reduced compared to heavier ions. Another parameter which may also affect the detection capability of ICP-MS is the plasma sampling position, which may not be optimal for all the analytes being determined simultaneously. This position is indeed important to maximize the density of analyte ions extracted from the ICP. Even the chemical form (i.e. speciation) of an element in solution may influence where it undergoes the various processes (desolvation, vapourization, etc) in the plasma, ultimately determining where the maximum ion density will be.

The sampling process itself is the source of another limitation: because part of the plasma passes through a small orifice, this orifice can become clogged by

dissolved solids. The maximum concentration of dissolved solids which can be tolerated on a continuous basis is 0.2%. Finally, even if this level is kept reasonably low, effects of concomitant elements (also called matrix effects) can occur, whereby the analyte signal is either suppressed or enhanced by the presence of other elements. Some trends have been observed and partly explained. For example, a heavy matrix element will in general induce a more important suppression than a light matrix element on quadrupole-based instruments because of space–charge effects. However, these matrix effects are still largely unpredictable, which complicates the calibration process for samples that do not possess a relatively simple matrix. Whenever possible, a simple dilution of the sample is recommended to minimize the effects of concomitant elements, since they depend on the absolute amount of concomitant element.

The various calibration strategies available with ICP-MS are summarized in **Table 4**, along with implications of the above-mentioned limitations. (More elaborate chemometric calibration procedures have also been developed by a few groups.) In general, unless the sample matrix is simple, internal standardization (or frequent calibration) will be required for quantitative analysis using external calibration with a series of standard solutions. The choice of the internal standard(s) depends on the instrument, the sample, etc. For the best results, an element with properties similar to that of the analyte (so that it will behave similarly in ICP-MS) is needed, so several internal standards may be required for a multielemental analysis spanning the entire mass range.

**Table 4** Calibration strategies in ICP-MS

| *Method* | *Requirements* | *Quality of results* |
|---|---|---|
| External calibration | One isotope free of spectroscopic interference per analyte<br>Matrix matching and/or correction for drift (using internal standardization or frequent calibration) | Good accuracy and precision |
| Standard additions | One isotope free of spectroscopic interference per analyte<br>May require a correction for drift (using internal standardization) | Good accuracy and precision |
| Isotope dilution | Two isotopes free of spectroscopic interference per analyte<br>Good equilibrium between isotopic spike and sample | High accuracy and precision |

The method of standard additions efficiently compensates for matrix effects but not for drift. If the latter is also present (for instance because of gradual clogging), internal standardization may also be required. The best method to compensate for both matrix effects and drift is isotope dilution analysis. However, it is only applicable to elements with at least two isotopes (stable and/or long-lived radioactive) free of spectroscopic interferences. Because of mass discrimination, this normally standardless technique requires a prior determination of mass bias using standards of known isotopic composition in order to correct the measured isotope ratios for mass discrimination.

## Accessories available to expand the capabilities of ICP-MS

Major modifications can be made to the ICP-MS instrument, such as adding one or two quadrupoles, and operating one quadrupole in an RF-only mode so that it can serve as a cell where gas-phase collisions are performed to break polyatomic ions. However, several more or less expensive accessories that can be readily interfaced to ICP-MS are available to circumvent one or more of the above-mentioned limitations. The most frequently used are summarized in **Table 5**.

The sample introduction system can simply be replaced by a high-efficiency one, such as an ultrasonic nebulizer coupled to a desolvation system. Analyte transport efficiencies of 10–20% are typically achieved using such a system. Other high-performance nebulization systems include direct injection nebulization, which introduces 100% of the sample directly into the plasma (i.e. no spray chamber is needed), thermospray, hydraulic high-pressure nebulization and monodisperse dried microparticulate injection.

The characteristics of the plasma can also be modified by adding another gas to argon. For instance, an addition of $O_2$ to the plasma is beneficial for the analysis of organic solutions as it prevents the deposition of soot on the interface. The addition of gases with a higher thermal conductivity than argon (such as $H_2$) can be used to enhance sensitivity (by improving the ionization efficiency), reduce and/or eliminate spectroscopic and non-spectroscopic interferences, and reduce mass discrimination.

**Table 5** demonstrates that various approaches can be used with a basic ICP-MS instrument to overcome spectroscopic interferences. Many of these approaches consist in separating the problematic element(s) from the analyte during sample preparation

**Table 5** Options widely used to enhance the capabilities of ICP-MS

| *Method* | *Advantages* | *Disadvantages* | *Examples of applications* |
|---|---|---|---|
| Ultrasonic nebulization with desolvation | High sample introduction efficiency | Prone to memory effects | Analysis of river water, lake water, rain and other samples with simple matrix |
| | Reduction of oxides through removal of solvent | Cone blockage and matrix effects from concentrated matrix | |
| Mixed-gas plasmas | Reduced spectroscopic interferences | Higher incidence of torch melting | Analysis of fuel oil, vegetable oil, sea water, urine, slurries |
| | Reduced matrix effects | | |
| Flow injection analysis | Reduced memory effects | System must be modified and optimized for each type of application | Analysis of samples with high level of dissolved solids (sea water, brines, urine, etc), concentrated acids, volume-limited samples |
| | Increased sample throughput | | |
| | Reduction of matrix effects | | |
| | Enhancement of sensitivity | | |
| | On-line pretreatment chemistry | | |
| Chromatographic techniques | Speciation of aqueous and gaseous samples | Different interfaces required for different techniques | Determination of toxic and non-toxic species in biological materials, food, environmental samples |
| | Separation of analyte from matrix | Chromatographic resolution dependent on interface to ICP | |
| Chemical vapour generation | Selective separation of analyte from matrix | Different chemistries and/or conditions required for different analytes | Determination of hydride-forming elements and/or species in a variety of samples |
| | Up to 100% sample introduction efficiency | | |
| Electrothermal vapourization | Small sample volume | Only one or a few elements can be determined at once | Analysis of waters, biological materials, food, volume-limited samples |
| | Direct solid analysis | Matrix modifiers often required to prevent loss of analyte | |
| | *In-situ* pre-treatment of sample | | |
| | Up to 100% sample introduction efficiency | | |
| Slurry nebulization | Eliminates digestion and fusion steps | No available sample blanks | Analysis of rocks, soils, hard-to-dissolve samples (such as coal) |
| | Calibration with aqueous standards | Efficiency of calibration dependent on particle size and sample type | |
| Laser ablation | Direct solid analysis | Limited number of reference materials available for calibration | Analysis of geological samples, ceramic materials, glass |
| | *In-situ* analysis of interstitial fluid | No available sample blanks | |
| | Microscopic profiling of solid samples | | |

(by extraction, chelation, selective precipitation, etc). Although they can be quite effective at reducing both spectroscopic and non-spectroscopic interferences (if the source of the latter is also removed during sample processing), they are nonetheless time-consuming and increase the likelihood of contamination.

Most of these separations can be carried out on-line with ICP-MS using a flow injection manifold. A big advantage of this approach is that all the chemistry is done in a closed system, often on a reduced scale, thereby reducing sources of contamination, as well as sample and reagent consumption. Chemical vapourization approaches can also selectively isolate

the analyte and transform it into a volatile species, in either continuous or flow injection fashion. Alternatively, selected species of the analyte can be determined by coupling gas, liquid or supercritical chromatography, or capillary electrophoresis, to ICP-MS.

As mentioned earlier, solid samples must first be put in solution for analysis by conventional ICP-MS, using digestion, fusion, dry-chlorination, etc, which all introduce potential sources of contamination. Slurry nebulization, which simply consists in grinding the sample and suspending it in a suitable dispersing agent, can significantly reduce these contaminations. Alternatively, solids can be analysed directly using various techniques coupled to ICP-MS: electrothermal vapourization, laser ablation (as well as arc/spark ablation), direct sample insertion, fluidized bed introduction, etc. These techniques are even more interesting for samples which are difficult to dissolve. Some of them (electrothermal vapourization and direct sample insertion) can also be used for the analysis of solutions.

All these approaches enhance the capabilities of ICP-MS but complicate the optimization process. The full potential of some of them (such as ablation techniques, flow injection, etc), which generate short, transient signals, will be better realized if an instrument allowing simultaneous multielement detection (such as TOF) is used. In any case, no single method can be used for the direct analysis of every sample. Some method development is therefore required for each new sample type, whether ICP-MS is used with or without one or more of the available options.

*See also:* **Calibration and Reference Systems (Regulatory Authorities); Chromatography-MS, Methods; Food Science, Applications of Mass Spectrometry; Forensic Science, Applications of Mass Spectrometry; Glow Discharge Mass Spectrometry, Methods; Isotope Ratio Studies Using Mass Spectrometry; Quadrupoles, Use of in Mass Spectrometry; Sector Mass Spectrometers; Time of Flight Mass Spectrometers.**

## Further reading

Brenner IB and Taylor HE (1992) A critical review of inductively coupled plasma mass spectrometry for geoanalysis, geochemistry, and hydrology. Part I. Analytical performance. *CRC Critical Reviews in Analytical Chemistry* **23**: 355–367.

Date AR and Gray AL (eds) (1989) *Applications of Inductively Coupled Plasma Mass Spectrometry*. Glasgow: Blackie.

Gray AL (1988) Inductively coupled plasma source mass spectrometry. In: Adams F, Gijbels R and Van Grieken R (eds) *Inorganic Mass Spectrometry, Chemical Analysis Series*, Vol 95, pp 257–300. London: Wiley.

Hall GEM (1989) Inductively coupled plasma mass spectrometry. In: Thompson M and Walsh JN (eds) *Handbook of Inductively Coupled Plasma Spectrometry*, 2nd edn. Glasgow: Blackie.

Houk RS (1986) Mass spectrometry of inductively coupled plasmas. *Analytical Chemistry* **58**: 97A–105A.

Houk RS (1994) Elemental and isotopic analysis by inductively coupled plasma mass spectrometry. *Accounts of Chemical Research* **27**: 333–339.

Jarvis KE, Gray AL and Houk RS (1992) *Handbook of Inductively Coupled Plasma Mass Spectrometry*. Glasgow: Blackie.

Montaser A (ed) (1998) *Inductively Coupled Plasma Mass Spectrometry*. New York: Wiley.

Montaser A and Golightly DW (eds) (1992) *Inductively Coupled Plasmas in Analytical Atomic Spectrometry*, 2nd edn. New York: VCH.

Olesik JW (1991) Elemental analysis using ICP-OES and ICP/MS. An evaluation and assessment of remaining problems. *Analytical Chemistry* **63**: 12A–21A.

# Industrial Applications of IR and Raman Spectroscopy

**AS Gilbert**, Beckenham, Kent, UK
**RW Lancaster**, Glaxo-Wellcome plc,
Medicines Research Centre, Stevenage, Herts, UK

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
Applications

Extensive use is made of IR and Raman spectroscopy in industry, because these techniques have the capacity to provide information at the chemical molecular level on virtually any material of interest be it solid, liquid or gas. They are well understood, can be easily applied and results can be obtained quickly and cheaply. Instruments utilized range from simple single-channel IR filtometers to sophisticated computerized interferometers and spectrographs supported by flexible data handling software. In addition to process functions they are employed in quality assurance and troubleshooting in the quality control (QC) environment and in analysis (both qualitative and quantitative) in research and development (R&D). Areas of application are pre-eminently in the chemical, oil and gas, plastics and pharmaceutical industries, but there is also considerable use in materials industries, mining, agriculture and food processing.

Applications of near-infrared (NIR) spectroscopy will not be described in this article but can be found elsewhere in the Encyclopedia. The examples of IR and Raman spectroscopy given below represent, of course, only a tiny fraction of what is done in industry and are necessarily a rather arbitrary selection.

## Brief history

With the introduction of commercially available, automatically recording, spectrometers during World War 2, IR spectroscopy began to be utilized in industrial laboratories and by the 1960s dedicated instruments had reached the factory floor. The advent of FT-IR machines designed specifically for the analytical market and a much wider range of accessories then saw a considerable expansion in use from the early 1980s onwards, particularly in QC and development laboratories.

Expense and technical difficulty meant that it was not until the 1970s that anything more than a handful of Raman spectrometers were employed in industry. Take-up, even by industrial research laboratories, continued to be slow until the arrival of relatively cheap computerized spectrographs, FT-Raman and advanced methods of sampling during the 1980s.

## The special conditions of the industrial environment

Employment of instrumentation is much the same in industrial research laboratories as it is in academic institutions, although of course the aims of the; former are necessarily more focused to particular objectives. But elsewhere in industry both time and money, within the constraints of safety, are of prime importance and the usefulness of information provided by any technique must be judged primarily by these criteria. The nearer the production line, the greater is the degree of rigour with which these criteria are applied. Rapid measurement, high degree of automation and reliability, all at minimum cost, are therefore demanded of manufacturers of spectroscopic equipment.

A major cost element of any activity is personnel. This predicates dedicated operation with as much 'black box' control of instrumentation as possible and incorporating robust protocols.

The factors described above tend inevitably to induce conservative attitudes in industrial practitioners so that it is important that any technique and its particular area of application be well understood, widely accepted and validated by longtime usage. The activities of organizations such as ASTM (American Society for Testing Materials) which have overseen the establishment of standard procedures for many common but specific spectroscopic measurements have been highly important in cementing acceptance of such methodology. Scientific forums that bring together both industrial and academic workers are extremely helpful in circulating ideas. The long-standing British-based IRDG (Infrared and Raman Discussion Group) is a good example of a body that organizes regular lecture meetings, often at industrial locations.

## Instrumentation

As vibrational spectroscopy was originally a research technique, many spectrometers used in industry in the past have been derivatives of research-oriented

machines. However, most equipment is now produced for the 'analytical' market so this relationship (even for Raman spectrometers) has been turned around and now many so-called 'research' machines are just the standard 'analytical' instrument with add-on features. Outside R&D most instrumentation is designed for dedication to specific tasks and a great many IR machines sold are single-channel filtometers.

Such instruments are non-dispersive and utilize a narrow bandpass filter to monitor one compound of interest, commonly carbon dioxide or carbon monoxide, in a process environment. The bandwidth of the filter is closely matched to the bandwidth (for carbon dioxide the unresolved envelope of a number of rotational sidebands) of the analyte and a reference signal is obtained from another filter which passes radiation from a transparent area of the spectrum.

Full-spectrum machines are now, in the case of IR, mostly interferometers. The benefits of full computerization, high sensitivity, ease of operation and 'quantitative analysis' software packages have rendered the classical dispersive spectrometers obsolete. Portable operating software has meant that industry standard PCs and workstations can be used as 'front ends' on most if not all commercial instruments now sold.

Until relatively recently, Raman spectroscopy was not generally suited to industrial use with the exception of specific problems in R&D. Although the classical double grating dispersive scanning spectrometers could be computerized, they were slow and often subject to problems such as sample fluorescence. This has been changed by the advent first of simple and rugged spectrographs with multichannel (usually charge-coupled device (CCD)) detectors and second of FT-Raman modules that can be 'bolted on' to standard FT-IR machines. Spectrographs have addressed the speed problem while FT-Raman, which uses near-IR excitation, is not troubled by fluorescence.

## Sampling methods

Many of the sampling accessories now available for vibrational spectroscopy were developed for specific industrial use or with industry in mind. Some of the most important recent developments have been in the production of fibre-optic attachments that can allow data to be obtained *in situ* with remote spectrometers. Optical fibres are available for the Raman region but though suitable materials (e.g. chalcogenides) that are transparent in parts of the mid-IR region can now be manufactured, fibre-optic probes in this region are somewhat limited in usefulness. **Figure 1** demonstrates the quality of a Raman spectrum that can be obtained from inside a container, in this example an amber vial holding acetone. The flexible fibre-optic probe was merely placed on the outside of the glass so that sampling was completely non-invasive.

Raman microscopy imaging, the spectroscopic mapping of a sample placed on a computer-controlled motorized stage, has applications in many areas. A TV attachment gives visual guidance to the

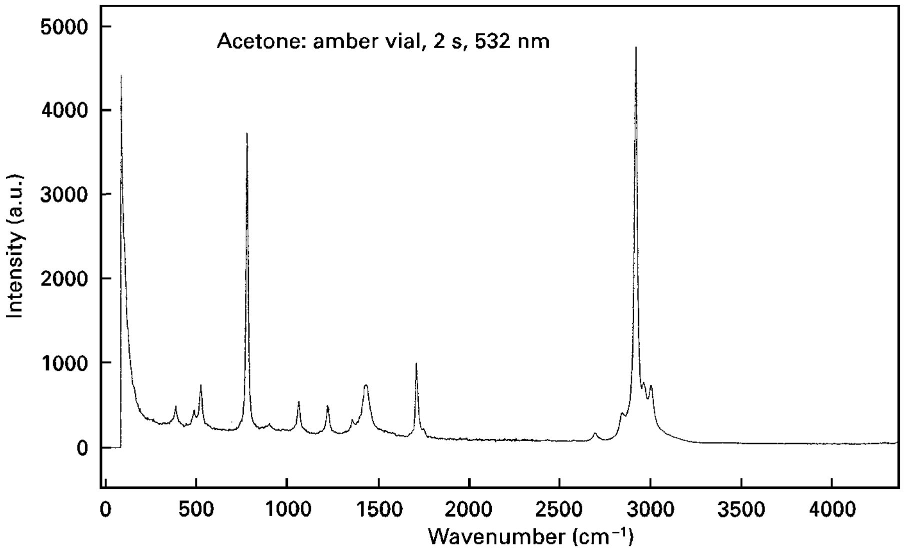



**Figure 1** The Raman spectrum of acetone contained within an amber vial. A HoloProbe (Kaiser Optical Systems, USA) fibre-optic probe transmitting a visible excitation line at 532 nm was utilized. The length of the fibre optic can be 100 m or more while, by changing attachments, the head can collect the Raman signal from distances of between about 0.03 and 40 cm.

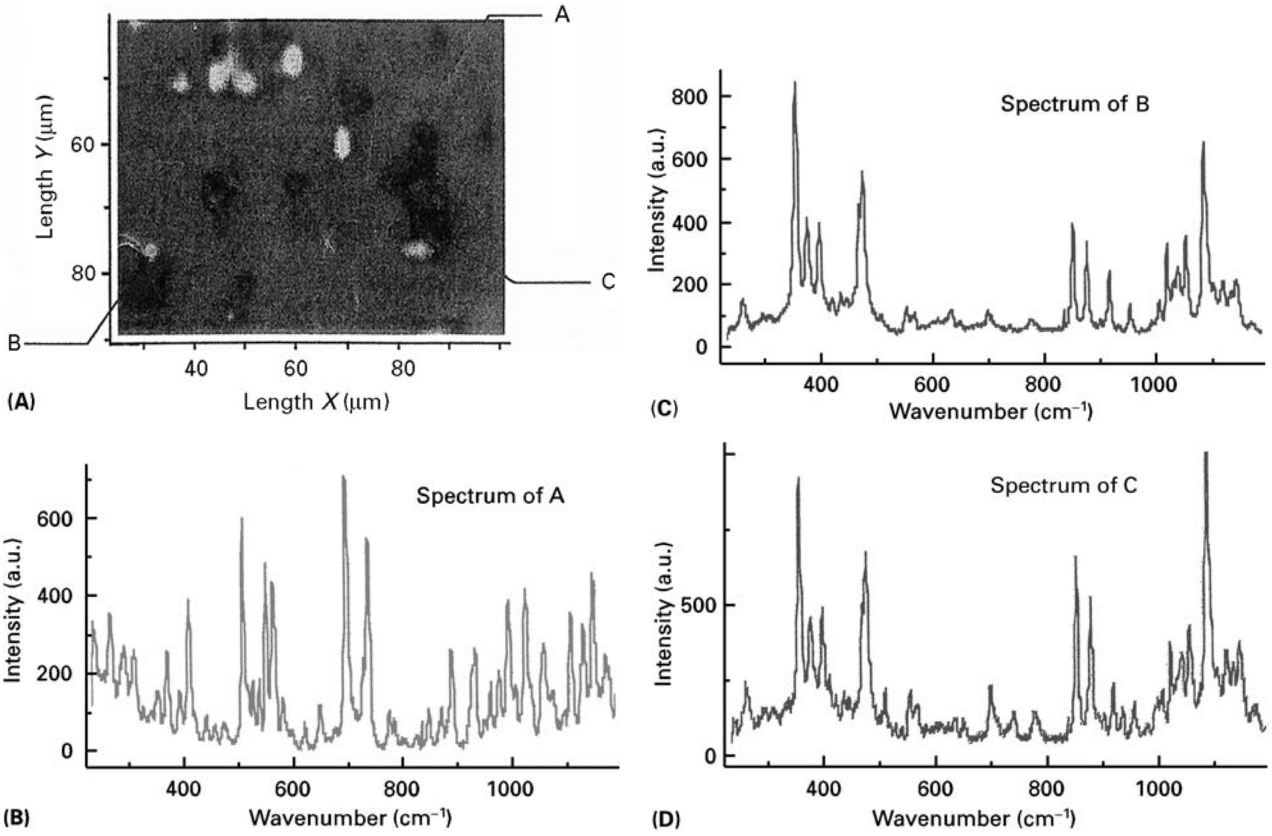



**Figure 2** Raman microscope image and associated spectra of a pharmaceutical powder. (A) Raman image that gives the relative importance of each compound (A, B, C), and spectra of (B) compound A, (C) compound B and (D) compound C. (See Colour Plate 23)

operator for location of an area of interest and real-time spectra are produced. A particular application, in this example from the pharmaceutical industry, is the examination of powders, particles down to 1 μm (or less) being distinguished. This is an order of magnitude better than IR microscopy. **Figure 2A** illustrates a colour-coded Raman map that offers chemical information on the disposition of different components of a mixture, the three distinct spectra shown in **Figures 2B**, **C** and **D** being obtained by ‘specifying’ three particular points in the image. The Raman image analysis software can also be used to assess particle size distribution.

Attenuated total reflection (ATR) probes can be used to follow chemical reactions in situ. An example is the react-IR (Applied Systems) instrument. Although ZnSe is commonly used as the optical element in ATR, it is not robust enough to deal with certain chemical reactions, e.g. bromination, so a diamond tip is utilized instead. **Figure 3** shows the spectrum of a fused-ring β-lactam which yields a carbonyl band at 1785 $cm^{-1}$. This can be seen reacting via an intermediate to give a less strained β-lactam with a carbonyl vibration near 1760 $cm^{-1}$.

The production of potassium bromide discs for IR spectroscopy has never been automated, although at least one scheme has indicated the possibility. When dispersive spectrometers were used, this did not matter so much as spectrum acquisition was of the same order of duration as sample preparation, but FT-IR machines are much more rapid. Thus diffuse reflectance has recently become a very popular technique for monitoring the IR spectra of organic compounds that are soluble in volatile solvents as it only requires the placing of a drop of solution on a small layer of ground alkali metal halide powder.

For dealing with large numbers of compounds from chromatographic or other types of separation, the so-called ‘hyphenated’ methods are used. The most popular are the combinations with gas chromatography (GC-IR), thermogravimetry (TG-IR) and gel-permeation chromatography (GPC-IR). **Figure 4** is the IR chromatogram of part of a GC separation of fractions in a typical gasoline (petrol). One of the aims was to detect aromatic components the levels of which are strictly controlled by various ‘clean fuel’ legislations (in the USA and elsewhere). Only IR absorption from the 900–600 $cm^{-1}$ region has been

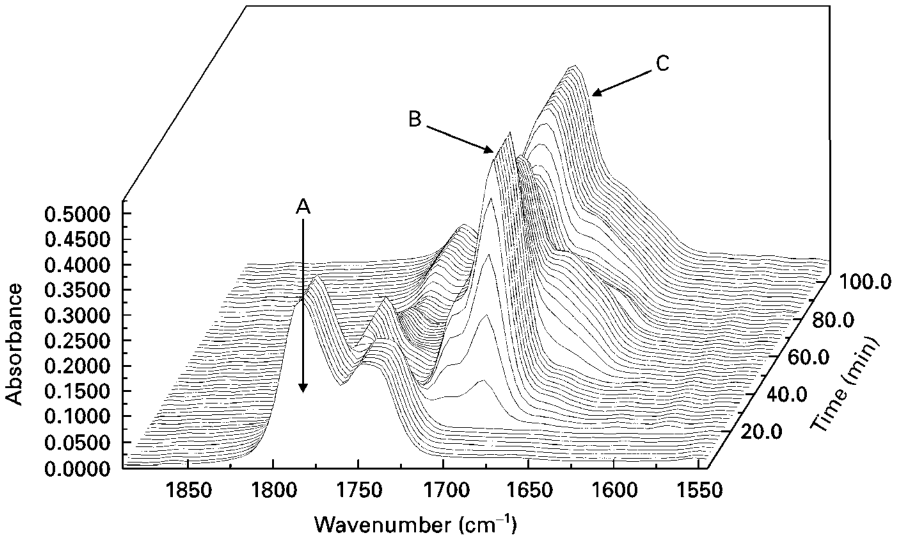



**Figure 3** IR stack plot of the progress of a reaction of a β-lactam reactant A giving intermediate B and product C.

utilized in this example, this being the region of the characteristic C–H out-of-plane bending modes of aromatic compounds. The level of benzene in this sample, which was of the order of 0.02% by mass, could be determined by the GC-IR method to an accuracy of within 1% of the measured value.

## Data handling

Data transfer to computer networks such as LIMS (laboratory information management systems) is now straightforward using the JCAMP world standard data format and modern operational software for vibrational spectrometers is usually set up for this. Sophisticated quantitative analysis packages are generally available for use in analytical laboratories and designed for largely 'black box' operation, which is obviously attractive to industry. Some of the chemometrics methodology (e.g. partial least squares (PLS)) behind the software has only been developed relatively recently, however, and still draws criticism from some quarters, which might cause problems when regulatory authorities are involved.

## Research and development

Many substances of interest to industry are highly complex but vibrational spectra can still provide useful insights into their nature. **Figures 5A**, **B** and **C** show the IR spectra of a series of coal tar fractions which were obtained by carbonization of a wet charge of coal in an industrial coke oven at a mean flue temperature of 1220°C. The fractions are principally composed of complex multi-ring aromatics. While the spectra clearly yield many bands characteristic of aromatic systems, for instance the substitution sensitive C–H out-of-plane bending vibrations below 1000 cm$^{-1}$, other functionalities are obviously present, such as O–H and N–H. Interestingly, spectrum (B) shows a band at about 2220 cm$^{-1}$

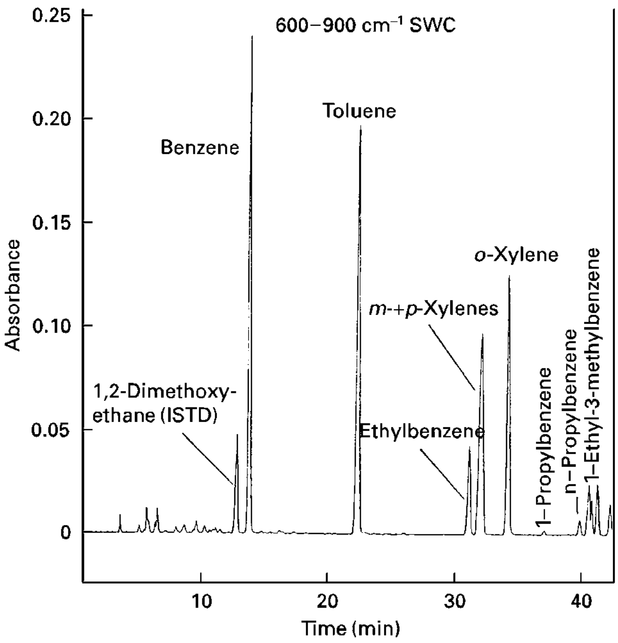



**Figure 4** IR chromatogram of the first 40 min of a GC-IR separation of aromatic components in a sample of gasoline. The peaks represent the total integrated absorbance intensity in the region 900–600 cm$^{-1}$. Reprinted with permission from Diehl JW, Finkbeiner JW and DiSanzo FP (1995) Determination of aromatic hydrocarbons in gasolines by gas chromatography/Fourier transform infrared spectroscopy. *Analytical Chemistry* **67**: 2015–2019. Copyright 1995 American Chemical Society.

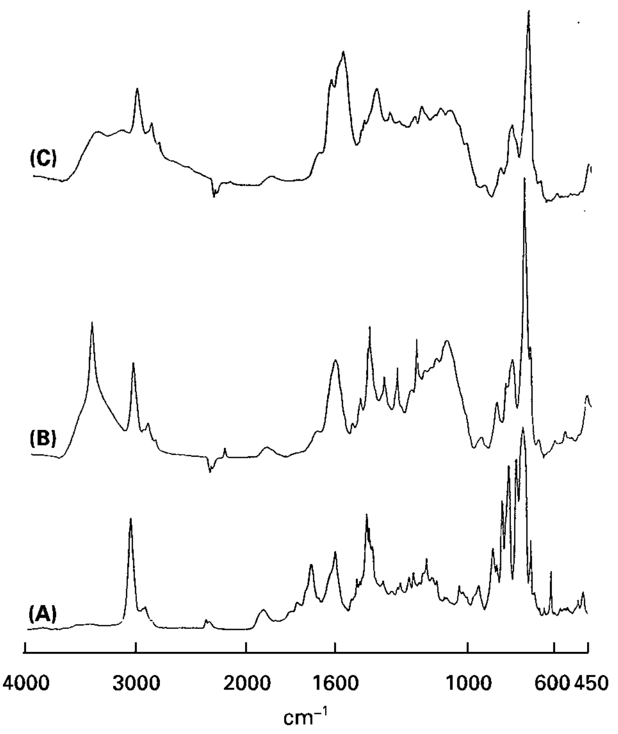
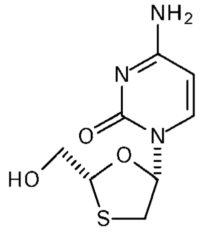



**Figure 5** FT-IR spectra of coal tar fractions examined as either (A) a film or (B, C) KBr discs. Reprinted from Diez MA, Alvarez R, Gonzalez AI, Menendez R, Moinelo SR and Bermejo J (1994) *Fuel* **73**: 139–142, Copyright 1994, with permission from Elsevier Science.

(just below the negative, instrument-uncompensated, $CO_2$ peak at about 2300 $cm^{-1}$) indicating nitrile, possibly from a naphthocarbonitrile.

Vibrational bands are very sensitive to intermolecular effects, thus different crystal modifications of a single compound usually yield discrete spectra. The occurrence of differing crystal forms is known as polymorphism and is of interest as the physical properties are dissimilar. This can have considerable ramifications in the pharmaceutical industry as the dissolution and absorption rates of polymorphic forms of a drug may differ significantly, with serious consequences for therapeutic action and patient tolerance.

**Figures 6A** and **B** display the IR spectra of forms I and II, respectively, of lamivudine, a 1,3-oxathiolane nucleoside which has anti-viral activity:

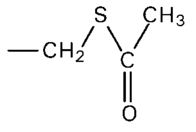



The two forms are not strict polymorphs because I is a hydrate (one $H_2O$ per five lamivudines), but have different crystal structures and spectra. The O–H stretching mode from the hydrate water in I can clearly be seen at 3545 $cm^{-1}$. The strong bands just below 3000 $cm^{-1}$ are from the hydrocarbon mulling agent used to prepare the material for spectroscopic measurement.

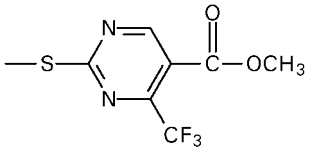
## Combinatorial chemistry

Schemes for combinatorial syntheses of compounds, particularly candidate pharmaceuticals, now attract great interest because of the potential for greatly speeding up the discovery of new drug molecules. They are based on carrying out chemical reactions on solid artifacts, the starting material, intermediates and final product being attached to the solid phase; on completion the product is cleaved away into solution. Typically the solid substrate is a mass of small (of the order of tens or 100 μm or so) porous beads made of polymer or silica resin. They are suspended, by agitation, in reactant solution, which is removed on completion by filtration. The beads are washed with solvent between stages of synthesis.

A considerable challenge is thereby presented to analytical technology if the progress of the reaction is to be monitored without going through the time-consuming business of analyte removal. Mass spectrometry cannot be used and NMR only with difficulty because of sensitivity problems. However, IR spectroscopy is fairly easy to utilize, either by diffuse reflectance or, as in the cases below of two drug intermediates, ATR sampling methods. **Figure 7** is the spectrum of




attached to polystyrene beads, the carbonyl stretching band being clearly seen at 1690 $cm^{-1}$. There is little interference with this highly characteristic band from the polymer. The silica resin (Tentagel) presents a greater problem, although as shown in **Figure 8**, a number of bands from attached



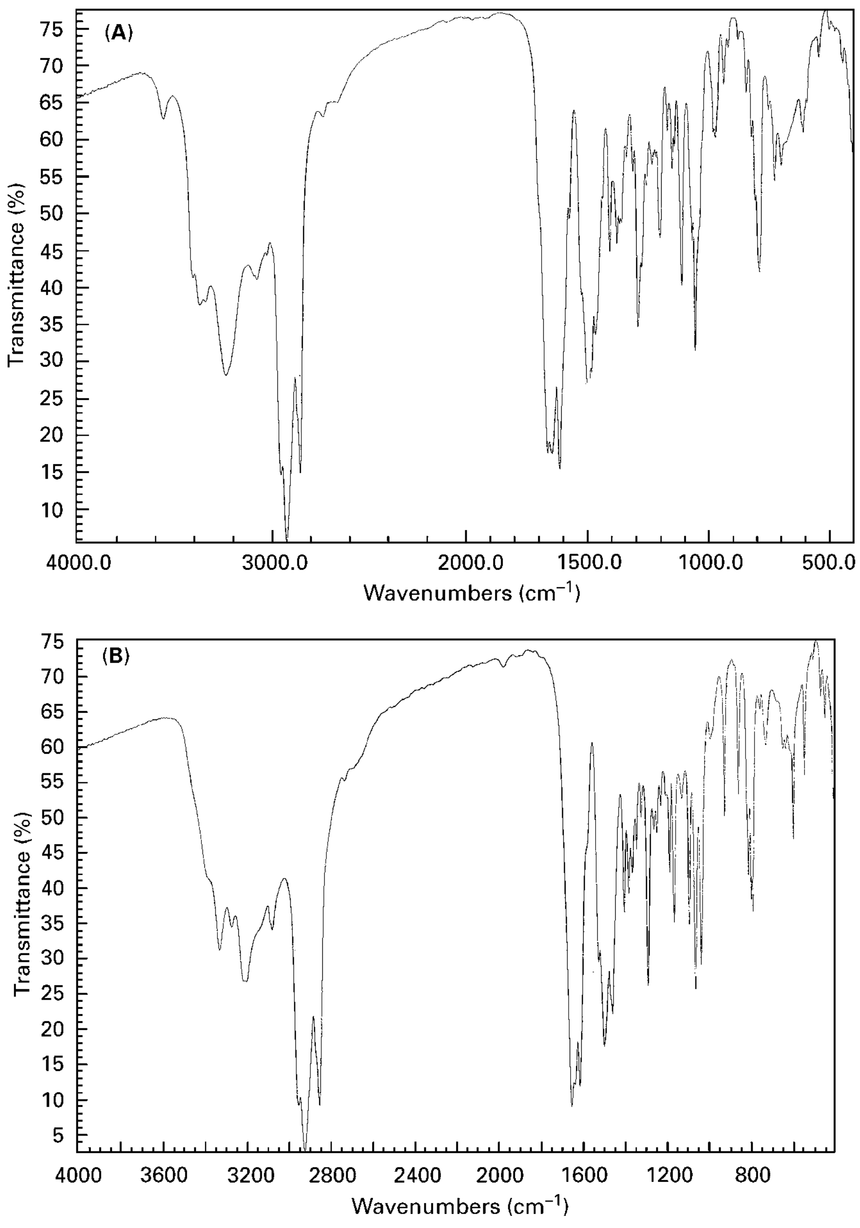



**Figure 6** FT-IR spectra of forms (A) I and (B) II of lamivudine as Nujol mulls. Resolution, 2 cm$^{-1}$. Reproduced with permission from Harris RK, Yeung RR, Lamont RB, Lancaster RW, Lynn SM and Staniforth SE (1997) *Journal of the Chemical Society, Perkin Transactions* **2**: 2653–2659.

can still be observed, e.g. the carbonyl vibration at 1741 cm$^{-1}$. The very strong band near 1100 cm$^{-1}$ is, of course, from Si–O stretching. These spectra were obtained from a small cluster of beads, but it is possible using microscopy techniques to examine a single bead at a time.

## Quality assurance

### Pharmaceuticals

As visible radiation is unaffected by passage through clear, colourless glass (apart from inducing minor fluorescence), Raman spectroscopy can be used to monitor gases contained within the headspace of sealed vials. **Figure 9** is a photograph of equipment that performs automated headspace analysis of vials (1 mm thick borosilicate glass) containing the pharmaceutical product Flolan (sodium prostacyclin) that is freeze-dried and sealed (with rubber stoppers secured by aluminium caps) under nitrogen. The product, although buffered, is highly sensitive to degradation by ingress of carbon dioxide if admission of air occurs, leading to a severe reduction in shelf-life. Any leakage, which might very occasionally occur before final capping, is detectable, however, by

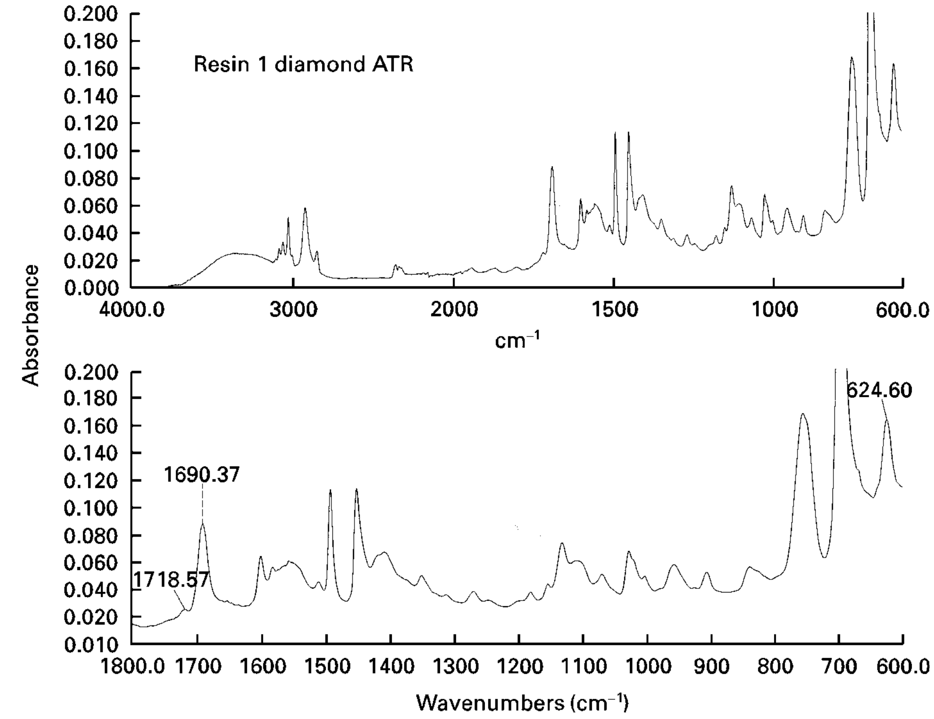



**Figure 7** Part of the FT-IR/ATR spectrum of polystyrene beads with attached compound (see text). Courtesy of Perkin-Elmer Ltd.

assaying for oxygen with the nitrogen as reference, using the Raman bands from the stretching modes of these molecules at about 1550 and 2330 $cm^{-1}$, respectively (**Figure 10**).

The analyser is built around a spectrograph with a CCD multichannel detector and the spectra are generated by laser radiation at 488 nm from an argon ion laser. Oxygen percentages can be measured to

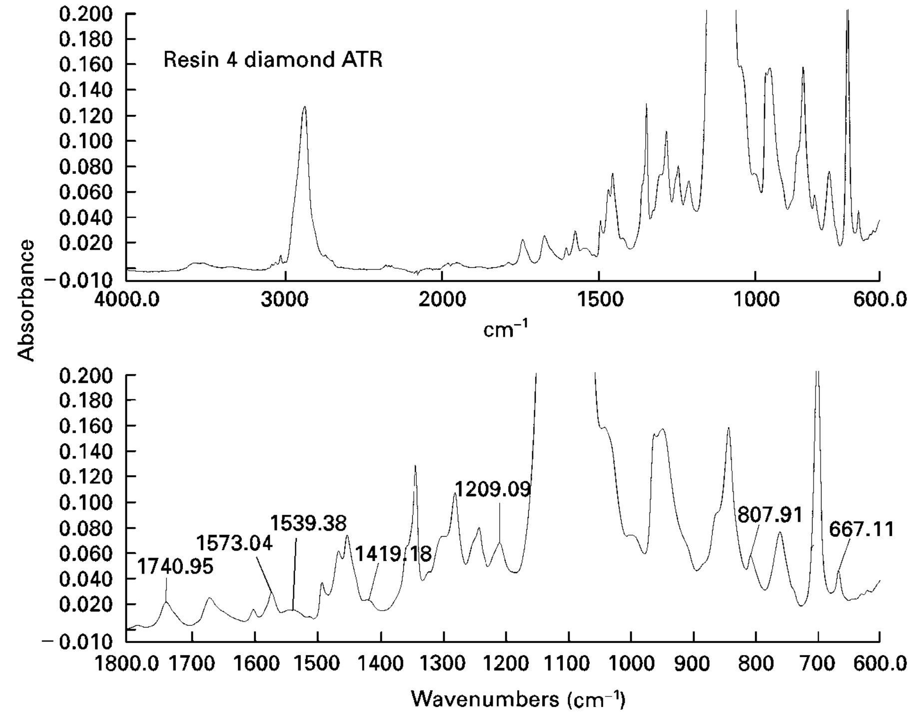



**Figure 8** Part of the FT-IR/ATR spectrum of Tentagel beads with attached compound (see text). Courtesy of Perkin-Elmer Ltd.

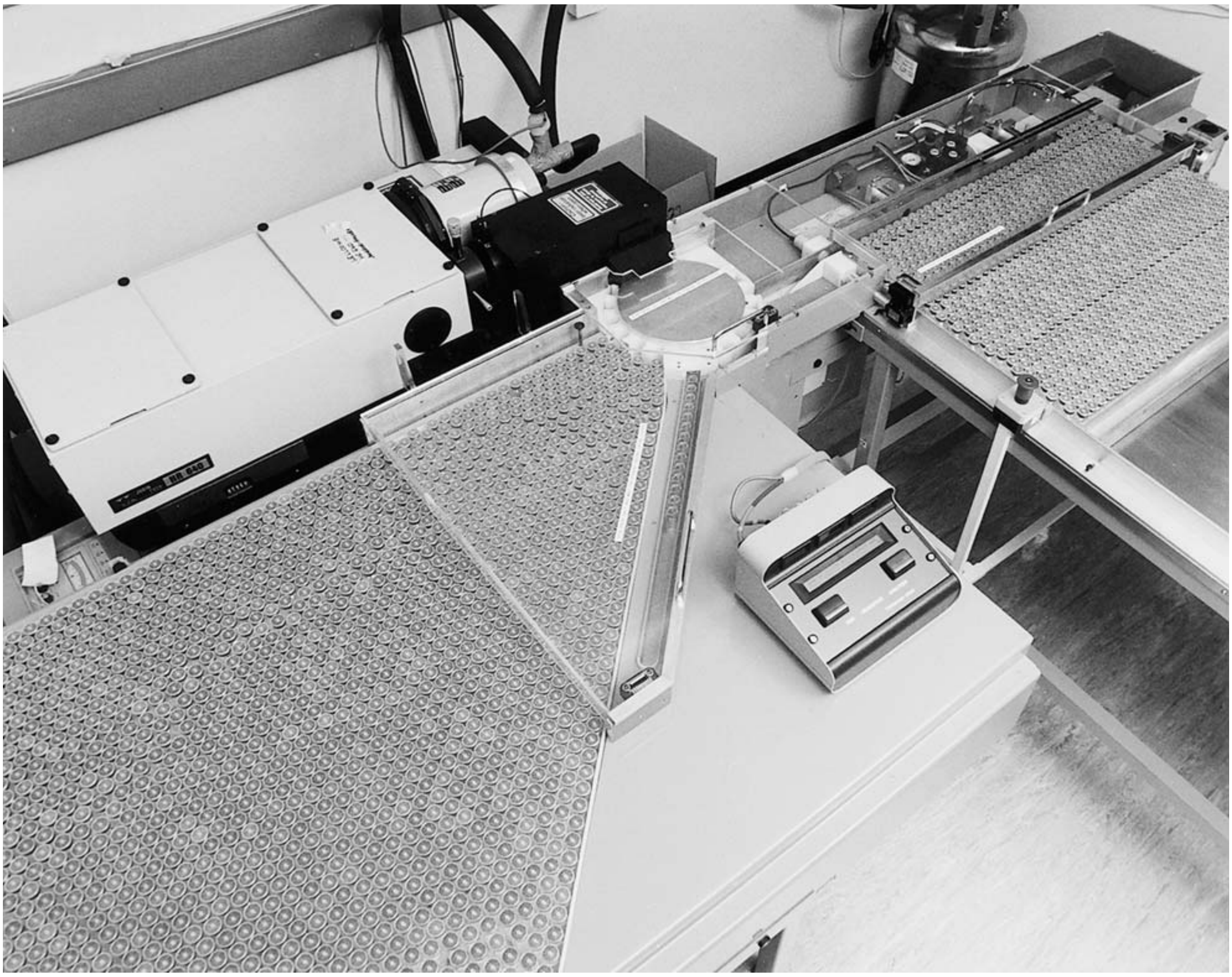


**Figure 9** Automated headspace analyser. A Model HR640 MSL (Instruments SA) Raman spectrograph with an Astromed CCD multichannel detector is utilized. Laser excitation at 488 nm. Courtesy of Hobbs KW, Glaxo-Wellcome plc.

±0.3% $O_2$ (± one standard deviation) for a measurement time of 1 s. Mechanical restrictions mean that sample throughput is slower than this, but still sufficiently fast that total batch inspection can be performed.

## Semiconductors

An important requirement in the semiconductor industry is to be able to measure various undesirable impurities in silicon wafers which can interfere with their electrical properties. Two significant impurities are oxygen which can incorporate interstitially and carbon which can substitute into lattice sites. Although pure silicon itself yields a very weak IR spectrum, fairly thick samples can be examined in transmission and therefore trace levels of the impurities can be detected and routinely quantified. Standard methods for IR spectroscopic determination of the impurities have been published by ASTM.

Oxygen forms Si–O–Si bonds which yield a characteristic band at 515 $cm^{-1}$ from Si–O–Si deformation (a stronger band from Si–O stretching occurs at 1107 $cm^{-1}$). Carbon can be detected by the Si–C stretching band at 607 $cm^{-1}$. **Figure 11** shows some room temperature difference spectra of CZ silicon using FZ silicon as reference (CZ and FZ refer to two different methods of growing silicon crystals, Czochralski process and float zone, respectively). It is interesting to note that the high-carbon CZ sample contains so much oxygen that weak bands from carbon–oxygen complexes are observed. Quantitative determination of carbon and oxygen is possible down to low ppm levels.

Dedicated FT-IR spectrometers are extensively employed to measure the thickness of epitaxial layers

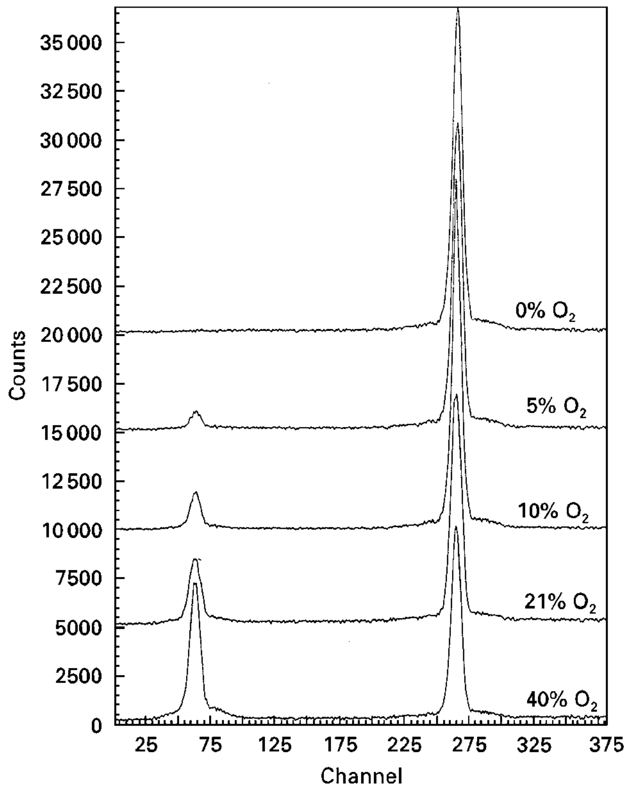



**Figure 10** Raman spectra of various mixtures of nitrogen and oxygen obtained with the spectrograph in **Figure 9**. For the sake of clarity the spectra have been vertically displaced from one another. The $N_2$ peak is on the right. Band intensity is essentially from the pure vibrational mode. The rotational–vibrational side bands form weak overlapped 'wings' on either side of the main peaks. Reproduced with permission of the Society of Photo-Optical Instrumentation Engineers (SPIE) from Gilbert AS, Hobbs KW, Reeves AH and Jobson PP (1994) Automated headspace analysis for quality assurance of pharmaceutical vials by laser Raman spectroscopy. *Proceedings of the SPIE – Society of Photo-Optical Instrumentation Engineers* **2248**: 391–398.

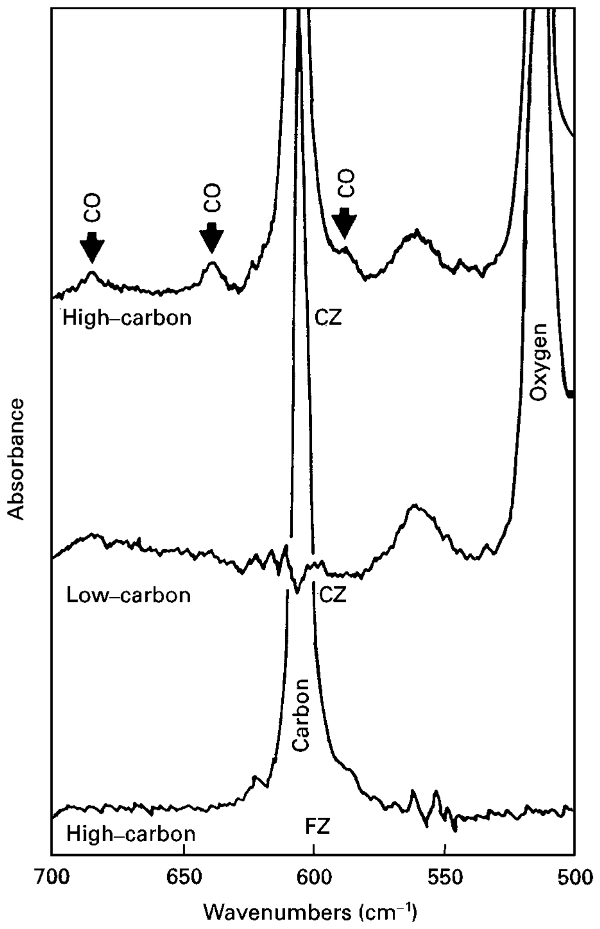



**Figure 11** IR difference spectra, CZ–FZ, using FZ (bottom) as reference. Samples were 2 mm thick. Carbon–oxygen complex bands are indicated by arrows. See text for definition of CZ and FZ. Reproduced by permission of Academic Press from Krishnan K, Stout PJ and Watanabe M (1990) Characterisation of semiconductor silicon using Fourier transform infrared spectrometry. In: Ferraro JR and Krishnan K (eds) *Practical Fourier Transform Infrared Spectroscopy,* pp 286–351. New York: Academic Press.

on silicon wafers. The measurements are made in reflectance mode and what is observed is not the spectrum but the interference pattern from multiple back and forth reflections within the layer. The regular distance between the maxima (or minima) gives the thickness and the fringes can be directly perceived from the interferogram as two 'side-bursts' on either side of the centre.

## Process control

Analysis of edible oils and fats is an important activity in the food industry. Quantification of cis and trans unsaturation levels is one part of this activity. Many edible oils are found in the cis form and are liquids. They are often partially hydrogenated to improve stability and hardness (e.g. for 'healthy' alternatives to butter), but in the process some cis bonds are converted to trans. Unfortunately, trans unsaturated fatty acids have been implicated in heart disease, but they can easily be detected by the presence of the characteristic out-of-plane olefinic C–H bending vibration at 980–965 $cm^{-1}$. Cis bonds do not interfere as they yield the corresponding vibrational mode at 730–650 $cm^{-1}$.

The standard IR spectroscopic procedure (American Oil Chemists Society) involves dissolution in carbon disulfide with chemical modification if trans levels are < 15%. Peak height only is used for calculation and the overall procedure is slow. By taking advantage of modern multivariate chemometric methods, a more rapid and convenient means of analysis has recently been developed utilizing a

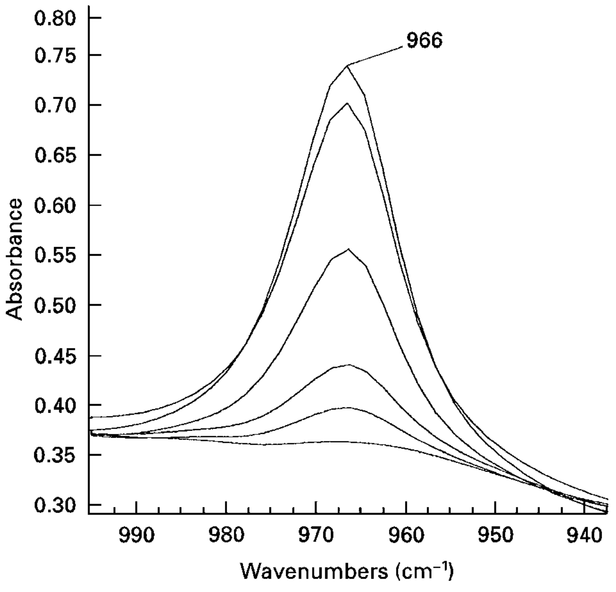



**Figure 12** FT-IR calibration standard spectra of the '*trans*' region. The standards are a '*trans*' free base oil with various known amounts of trielaidin, a triglyceride of carbon chain length 18, containing a *trans* double bond at the 1-position. Reproduced with permission of the AOCS from Sedman J, van de Voort FR, Ismail AA and Maes P (1998) *Journal of the American Oil Chemists Society* **75**: 33–39.

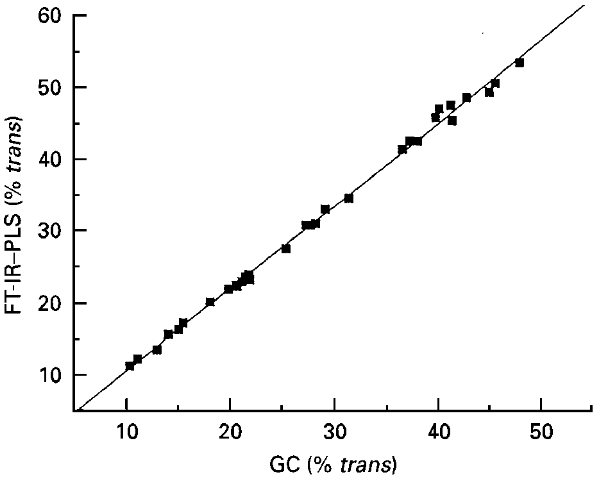



**Figure 13** Comparison of results from FT-IR and GC methods for determination of the '*trans*' contents of edible oils. Reproduced with permission of the AOCS from Sedman J, van de Voort FR, Ismail AA and Maes P (1998) *Journal of the American Oil Chemists Society* **75**: 33–39.

heated flow cell (thermostated to 80±0.2°C) to analyse the neat sample in its liquid state. **Figure 12** is the IR spectrum of the trans region for a series of calibration standards and **Figure 13** is a plot comparing measured values by the new IR method with gas chromatography, demonstrating high correlation between the two techniques. Calculation for the new IR procedure is carried out by the PLS method using multiple spectroscopic points and is thus very 'black box'. The sample measurement time is claimed to be 2 min, sufficiently rapid, therefore, for process monitoring.

A mundane but very common requirement is the on-line measurement of the water content of industrial solvents and feedstocks in process streams. Chemical methods (e.g. Karl Fischer) are not suited to continuous monitoring. Relatively high levels (> 0.1%) can be dealt with using NIR methodology, but for very low levels the much stronger (fundamental) O–H stretching bands in the mid-IR region need to be observed. **Figure 14** shows the IR spectrum of traces of water in 1,2-dichloroethane (ethylene dichloride, EDC). Because of the low concentration there is no significant hydrogen bonding between water molecules so that the 'free' asymmetric and symmetric stretching modes appear at 3685 and 3595 cm$^{-1}$, respectively.

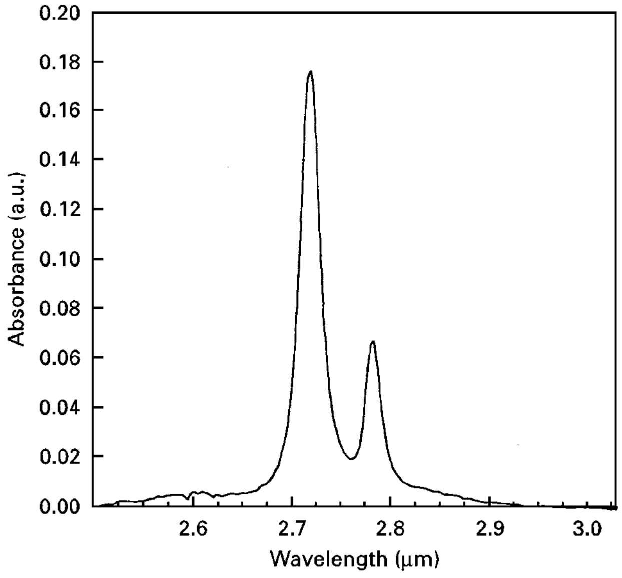



**Figure 14** IR transmission spectrum of traces of water in 1,2-dichloroethane. Note that the scale is in microns, which can be converted to wavenumbers by taking the reciprocal and multiplying by 10 000; e.g. 2.7 μm is equivalent to 3703.7 cm$^{-1}$. Reproduced with permission of the Royal Society of Chemistry from Bruce SH and Dhaliwal HK (1991) On-line moisture analysis by IR, in Davies AMC and Creaser CS (eds) *Analytical Applications of Spectroscopy II*, pp 46–52. Cambridge: Royal Society of Chemistry.

EDC is the feedstock for the manufacture of PVC, being first cracked in a furnace to produce vinyl chloride followed by polymerization. The moisture measurement is usually done in the transmission

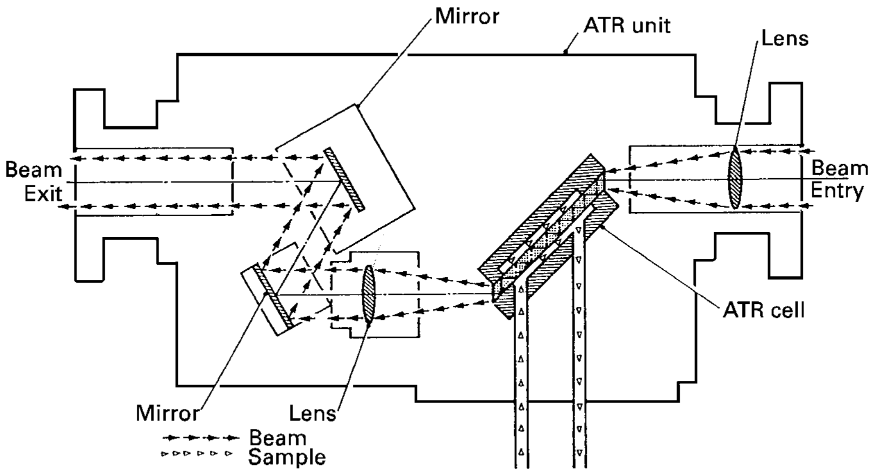



**Figure 15** Diagram of process ATR cell for on-line moisture analysis. Reproduced with permission of the Royal Society of Chemistry from Bruce SH and Dhaliwal HK (1991) On-line moisture analysis by IR, in Davies AMC and Creaser CS (eds) *Analytical Applications of Spectroscopy II*, pp 46–52. Cambridge: Royal Society of Chemistry.

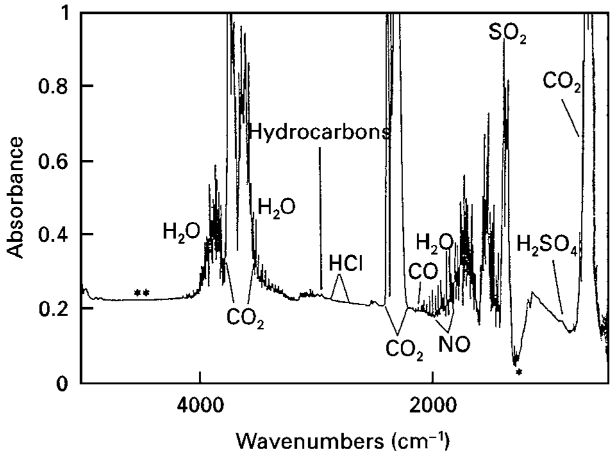



**Figure 16** FT-IR spectrum of a flue gas stream, resolution 4 cm$^{-1}$. Reproduced with permission of the Society of Photo-Optical Instrumentation Engineers from Solomon PR, Morrison PW Jr, Serio MA *et al* (1992) Fourier transform infrared spectroscopy for process monitoring and control. *Proceedings of the SPIE – Society of Photo-Optical Instrumentation Engineers* **1681**: 264–273.

mode, with a pathlength of around 2 mm, and as there are no other bands in the O–H stretching region a simple IR filtometer can be used. In other cases where the pathlength is too restrictive, for instance if the solvent/feedstock is too viscous, an ATR method can be employed instead. **Figure 15** shows an on-line ruggedized (150 psig, 150°C) ATR cell for this purpose (Servomex, UK) which can be fitted with a variety of IR crystals (for chemical compatibility) and adjusted to give the requisite effective pathlength.

The high sensitivity of FT-IR to many small gas molecules can be exploited to monitor flue gases. **Figure 16** shows the make-up of exhaust gas from a coal-fired boiler. To see the minor components it is necessary to expand the spectral scale as in **Figure 17**. Of importance are obviously the levels of the acid gases, NO and (unburned) hydrocarbons, which can all be measured at ppm levels.

Most industrial polymers contain additives which are there as stabilizers (e.g. cross-linkers, antioxidants) or property modifiers (e.g. slip agents). The diagram of a system (Automatik, Germany) for molten polymer process control using FT-IR to monitor additives on-line is given in **Figure 18**. The flow cell can withstand high pressures (300 bar) and temperatures (400°C) and has an adjustable pathlength. The spectrum shown in **Figure 19** is of various amounts of the slip agent oleamide in low-density polyethylene the monitored band being the amide carbonyl stretch near 1715 cm$^{-1}$.

## Conclusion

The wide application of vibrational spectroscopy to industrial problems is long established and its use is growing, as evidenced by steadily increasing sales of instrumentation. In many instances other techniques can do the same job but are either less sensitive or slower or more expensive, or are ruled out through other circumstances. In certain cases vibrational spectroscopy may be the only method practicable, such as in the semiconductor applications described above.

The fact that vibrational spectroscopy is amenable to almost any chemical substance is responsible for its general utility, but this ‘non-specificity’ can be a

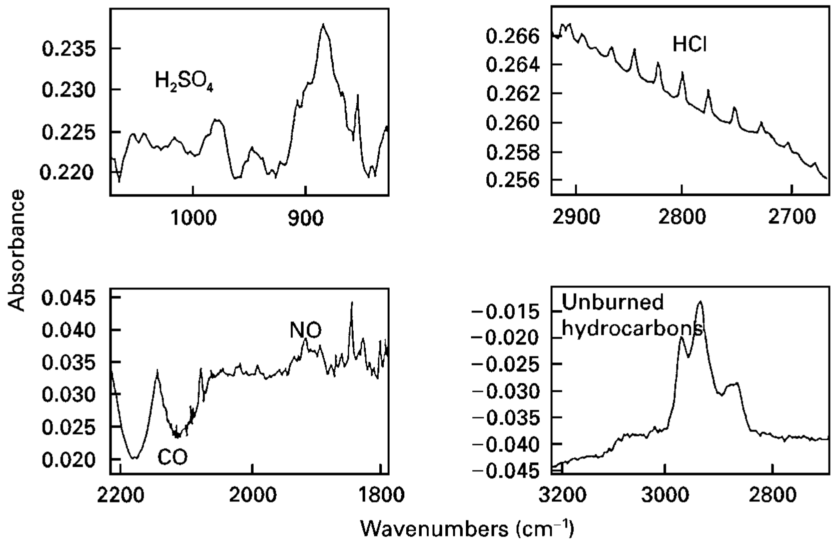



**Figure 17** Expanded regions of spectrum in **Figure 16** to show trace gases. Reproduced with permission of the Society of Photo-Optical Instrumentation Engineers from Solomon PR, Morrison PW Jr, Serio MA *et al* (1992) Fourier transform infrared spectroscopy for process monitoring and control. *Proceedings of the SPIE – Society of Photo-Optical Instrumentation Engineers* **1681**: 264–273.

weakness where it is necessary to 'pick out' one component from a background. There are one or two special techniques in Raman spectroscopy where this can be done (e.g. resonance Raman and surface-enhanced Raman spectroscopy), but they are not of general applicability.

A particular strength compared with other main-line analytical techniques is in the monitoring of gases from the points of view of both the sensitivity of IR spectroscopy and the non-invasive capability of Raman spectroscopy, coupled with the capacity of both techniques for remote viewing.

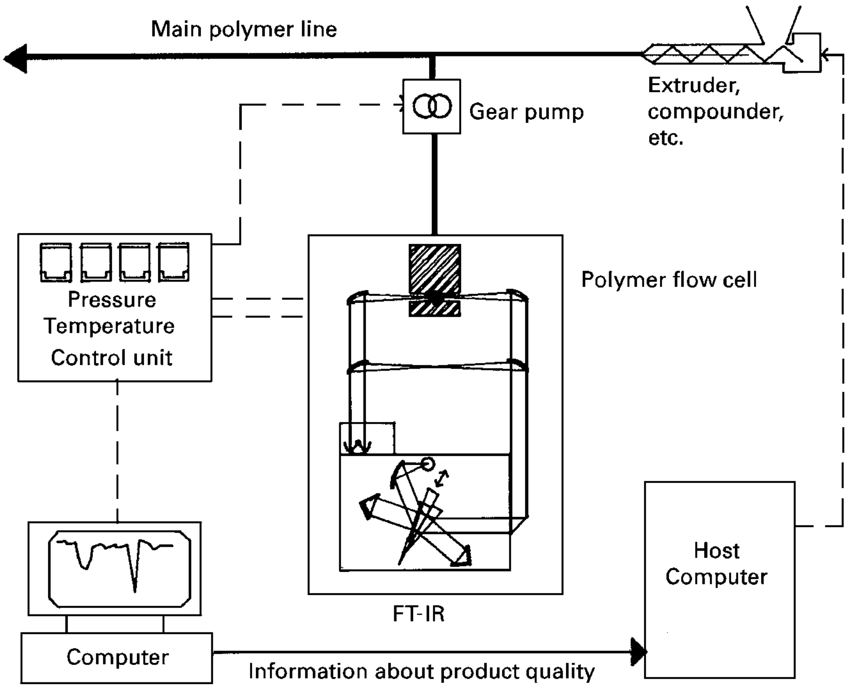



**Figure 18** Diagram of an on-line IR process control system for polymer production. Reproduced with permission of the Society of Photo-Optical Instrumentation Engineers from Stengler RK and Weis G (1992) Infrared process control on molten polymers using a high pressure, high temperature flow cell. *Proceedings of the SPIE – Society of Photo-Optical Instrumentation Engineers* **1681**: 33–38.

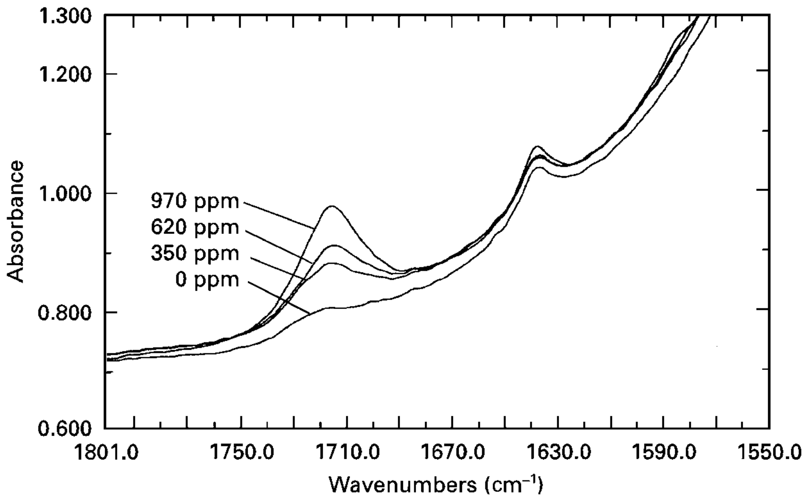



**Figure 19** FT-IR spectra of various concentrations of oleamide in molten low-density polyethylene. Reproduced with permission of the Society of Photo-Optical Instrumentation Engineers from Stengler RK and Weis G (1992) Infrared process control on molten polymers using a high pressure, high temperature flow cell. *Proceedings of the SPIE – Society of Photo-Optical Instrumentation Engineers* **1681**: 33–38.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Biochemical Applications of Raman Spectroscopy; Food Science, Applications of Mass Spectrometry; Food Science, Applications of NMR Spectroscopy; Fourier Transformation and Sampling Theory; FT-Raman Spectroscopy, Applications; IR Spectrometers, IR Spectroscopy Sample Preparation Methods; IR Spectroscopy, Theory; IR Spectral Group Frequencies of Organic Compounds; Nonlinear Optical Properties; Raman Optical Activity, Spectrometers; Raman Spectrometers.**

## Further reading

Bower DI and Maddams WF (1989) The Vibrational Spectroscopy of Polymers. Cambridge: Cambridge University Press.

Chalmers JM and Dent G (1997) Industrial Analysis with Vibrational Spectroscopy. Cambridge: Royal Society of Chemistry.

Durig JR (1985) Chemical, Biological and Industrial Applications of Infrared Spectroscopy. New York: Wiley.

Ferraro JR and Krishnan K (1990) Practical Fourier Transform Infrared Spectroscopy. New York: Academic Press.

White R (1990) Chromatography/Fourier Transform Infrared Spectroscopy and its Applications. New York: Marcel Dekker.

# Inelastic Neutron Scattering, Applications

**Stewart F Parker**, Rutherford Appleton Laboratory, Didcot, UK

Copyright © 1999 Academic Press

HIGH ENERGY SPECTROSCOPY

Applications

Inelastic neutron scattering (INS) spectroscopy was pioneered in the 1950s by the American physicist Bertram Brockhouse. So it is not surprising that most of the early applications were physics based; particularly the determination of phonon dispersion curves of metals which are critical to the understanding of their thermodynamic properties. However, like many other spectroscopic methods, of which nuclear magnetic resonance is probably the best example, the use of the technique has greatly broadened, and it is now used by researchers in virtually any field that involves the condensed state. The wide energy (or, equivalently, timescale) range available to INS means that different kinds of motion can be probed in the same system, although not usually simultaneously. The applications span from biology through to materials science and to engineering, in addition to physics and chemistry. Particularly active areas include: catalysis, polymers, magnetism, hydrogen-in-metals, hydrogen bonding, glasses, geology and quantum fluids.

## Coherent and incoherent scattering

Experimentally, the quantity that is usually measured is the scattering function, $S(Q, \omega)$ [$Q$ is the momentum transfer (Å$^{-1}$) and $\omega$ the energy transfer (meV)]. The scattering can be coherent, $S^{coh}(Q, \omega)$, or incoherent, $S^{inc}(Q, \omega)$. The definition of $S^{coh}(Q, \omega)$ involves the correlation between the positions of different nuclei at different times and thus gives interference effects. This gives information on collective motions in the sample. $S^{inc}(Q, \omega)$ only involves the correlation between the position of the same nucleus, so there are no interference effects, and the motions of a single particle are probed. This is a powerful method for studying the local environment of the scattering atom.

Whether the scattering is coherent or incoherent depends on the relative size of the coherent and incoherent scattering cross sections of the scattering nuclei; see **Table 1** for a selection of elements. In practice, the signal is a mixture of coherent and incoherent scattering, unless the experiment is designed to select one type of scattering. The cross section is both element and isotope dependent. The most important feature is the enormous incoherent cross section of hydrogen; in contrast, deuterium and most other elements have a modest incoherent cross section and a significant coherent cross section. This means that the scattering from a protonic sample will be mostly incoherent, while that from its deuterated isotopomer will be largely coherent. Isotopic substitution therefore provides a very powerful method to change the scattering characteristics of a sample.

**Table 1** The coherent and incoherent scattering cross sections of selected elements

| | | *Scattering cross section (barns = $10^{-28}$ m$^{-2}$)* | |
|---|---|---|---|
| *Element* | *Mass$^a$ (amu)* | *Coherent* | *Incoherent* |
| Hydrogen | 1.0078 | 1.758 | 79.7 |
| Deuterium | 2.0141 | 5.597 | 2.04 |
| Carbon | 12.011 | 5.6 | 0.001 |
| Nitrogen | 14.007 | 11.0 | 0.49 |
| Oxygen | 15.999 | 4.23 | 0.000 |
| Aluminium | 26.982 | 1.492 | 0.0085 |
| Silicon | 28.066 | 2.163 | 0.015 |
| Chlorine | 35.453 | 11.531 | 5.2 |
| Iron | 55.847 | 11.4 | 0.22 |
| Copper | 63.546 | 7.486 | 0.52 |
| Platinum | 195.08 | 11.65 | 0.13 |

[a] The atomic mass given refers to the natural isotopic composition of the element except for hydrogen and deuterium where the mass is that of the isotopes $^1$H and $^2$H respectively.

## The energy range ±10 meV (±80 cm$^{-1}$): quasielastic scattering and tunnelling spectroscopy

### Quasielastic scattering

Quasielastic scattering is concerned with scattered neutrons that have undergone small energy transfers, typically ≤ 2 meV (≤ 16 cm$^{-1}$). This originates from neutrons scattered by atoms that are undergoing diffusion or reorientation on a timescale of $10^{-12}$–$10^{-7}$ s. Since the motions are not quantized, they form a continuous distribution, and the effect is manifested as a broadening of the peak due to the elastically scattered neutrons. This is entirely analogous to the Doppler broadening of an audio signal by reflection from a moving object. The neutron has a magnetic moment so it can be scattered magnetically and the random

fluctuations of unpaired electron spins of stationary atoms can also cause quasielastic broadening.

In general, a molecule in a condensed phase will be simultaneously undergoing translation, rotation and vibration and the scattering law is a convolution of the scattering laws corresponding to the different types of motion. In practice, the vibrational term only affects the intensity as a Debye–Waller factor, $\exp(-Q^2\langle u^2\rangle)$, where $\langle u^2\rangle$ is the mean square displacement in the vibrational modes. For hydrogenous materials the scattering will be incoherent and only the motion of the hydrogen atoms need be considered. In this case the scattering law can be written as the sum of an elastic and an inelastic component:

$$\begin{aligned} S(Q,\omega) &= \text{Elastic} + \text{Inelastic} \\ &= A_0(Q)\delta(\omega) + S^{\text{inc}}(Q,\omega) \end{aligned} \quad [1]$$

Provided that the experimental data can be resolved into elastic and inelastic contributions, the $A_0(Q)$ term can be determined from:

$$A_0(Q) = \frac{\text{Elastic intensity}}{\text{Total intensity}} \quad [2]$$

and is called the elastic incoherent structure factor (EISF). The elastic intensity can be considered to arise from motions that are localized in space and the inelastic arise from motions that are not localized. Thus for purely translational motion, the EISF is zero since there is no elastic peak. For rotational motion the EISF is unity at $Q = 0$ and falls to a minimum at a $Q$ value that is related to the radius of gyration. At higher $Q$ the EISF is oscillatory and the shape gives information on the geometry of the rotational process. This is most characteristic beyond the first minimum and shows the importance of obtaining data at as wide a range of $Q$ as possible.

Rotational motions in many different systems have been studied. In particular, methyl group rotations in polymers have been extensively studied. The rotational motion is represented by instantaneous jumps between three equivalent positions on a circle of radius $R$. When all orientations are equally likely, Equation [1] may be written as:

$$S(Q,\omega) = A_0(Q)\delta(\omega) + \frac{1}{\pi}[1 - A_0(Q)]L(\omega) \quad [3]$$

with the EISF given by:

$$A_0(Q) = \frac{1}{3}\left[1 + 2J_0(QR\sqrt{3})\right] \quad [4]$$

where $J_0(x)$ is the zeroth-order spherical Bessel function. $L(\omega)$ is a Lorentzian line shape:

$$L(\omega) = \frac{2\tau/3}{1 + \omega^2(2\tau/3)^2} = \frac{\Gamma}{\Gamma^2 + \omega^2} \quad [5]$$

whose full width at half maximum is given by $2\Gamma$ or $3/\tau$ ($\tau$ being the average time between jumps). **Figure 1A** shows the decomposition of the experimental data into an elastic and a quasielastic part for selectively deuterated poly(methyl methacrylate) at 25°C. The deuterated part of the molecule gives considerable coherent scattering and after correction for this, the experimental data, **Figure 1B**, shows excellent agreement with the EISF calculated from Equation [4].

Recent work has shown that the data can be better modelled by a distribution of jump rates, since for amorphous polymers the methyl groups are in a variety of environments.

For translational diffusion, the simplest model is to consider that the diffusing particle starts at the origin at time zero. Since this is the dynamics of a single particle, the scattering is incoherent. In this case the scattering law takes the form:

$$S(Q,\omega) = \frac{1}{\pi}\frac{D_t Q^2}{(D_t Q^2)^2 + \omega^2} \quad [6]$$

where $D_t$ is the tracer diffusion coefficient. This is valid at small $Q$ ($\leq 0.5$ Å$^{-1}$) where the diffusion is being studied over long (tens of Å) distances, since in this case the details of the individual jumps ($\leq 2$ Å) are lost by averaging. The function is a Lorentzian with full width at half maximum of $D_t Q^2$, allowing $D_t$ to be extracted from the width of the quasielastic peak at low $Q$.

For coherent scattering, there will only be intensity at low $Q$ if the scattering atoms are in a compressible fluid. The intensity arises because at $t = 0$ there is an atom at the origin and hence excess concentration. This diffuses away in the same manner as for $D_t$, except that it is a concentration distribution that is studied. This gives rise to the chemical or Fick's law diffusion coefficient, $D_{chem}$, which is generally larger than $D_t$.

The diffusion of methane in the zeolite ZSM-5 can be treated using a more sophisticated version of the same idea that assumes that the jumps can be described by a Gaussian distribution of jump lengths. Physically, this corresponds to the molecule making jumps along the channel segments as well as across the channel intersections. Using this model,

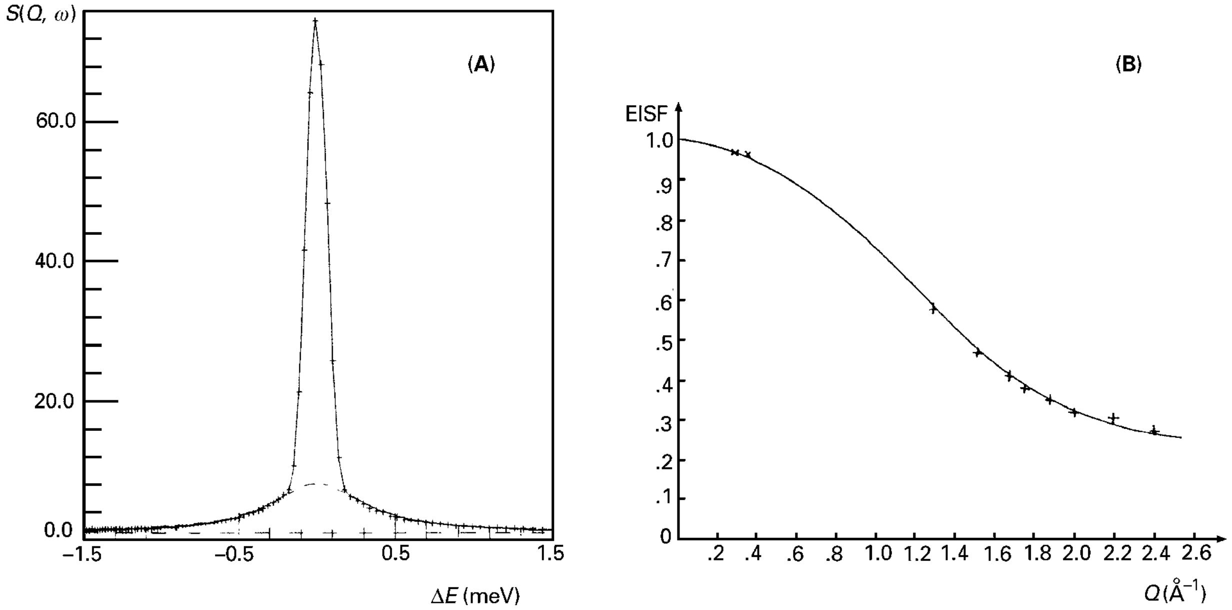



**Figure 1** (A) Decomposition of the quasielastic scattering at $Q = 2.06$ Å$^{-1}$ from selectively deuterated syndiotactic poly(methyl methacrylate) at 25°C into a delta function (sharp component) superimposed on a Lorentzian (broad component), due to the ester methyl rotation. Both components are resolution broadened. (B) EISF for the same sample. The data points have been corrected for coherent scattering and the solid line is a fit using Equation [4]. (Data recorded with IN5 and IN6 at the ILL). Reproduced with permission from Gabrys B, Higgins JS, Ma KT and Roots JE (1984) Rotational motion of the ester methyl group in stereoregular poly(methyl methacrylate): A neutron scattering study. *Macromolecules* **17**: 560–566.

there is a more complex $Q$ dependence, as shown in **Figure 2**. At small values of $Q$ Equation [6] is obtained and at large values of $Q$ the width tends asymptotically towards the mean jump rate. From the initial slope of the data in **Figure 2**,

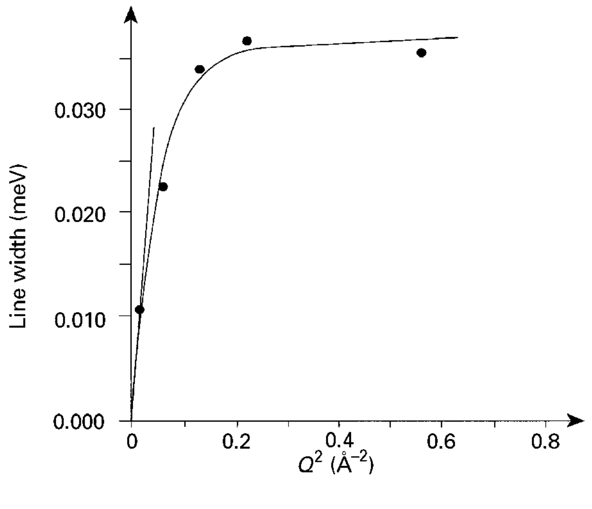



**Figure 2** The broadening of the elastic peak as a function of $Q^2$ for methane at a loading of 1.5 molecules per unit cell in ZSM-5 at 250 K. (Data recorded with IN5 at the ILL). Reproduced with permission from Jobic H, Bée M and Kearley GJ (1989) Translational and rotational dynamics of methane in ZSM-5 zeolite: A quasielastic neutron scattering study. *Zeolites* **9**: 312–317.

$D_t = 5 \times 10^5$ cm$^2$ s$^{-1}$ and the mean time between jumps is $3.6 \times 10^{-11}$ s.

A traditional strength, and still a major application, of INS is the study of hydrogen-in-metal systems. In these systems hydrogen occupies a well-defined lattice site, but is often free to move between sites. This proton mobility has led to numerous applications in battery technology, particularly for $LaNi_5H_x$ and the $AB_2$ Laves phase compounds. The simplest model is to consider that: the jumps are all of the same length ($l$), that jumps in any direction are equally probable, and that the jump time is negligibly small compared to the residence time ($\tau$). With these assumptions Equation [6] becomes:

$$S(Q,\omega) = \frac{1}{\pi}\frac{F(Q)}{[F(Q)]^2 + \omega^2} \qquad [7]$$

with

$$F(Q) = \frac{1}{\tau}\left(1 - \frac{\sin Ql}{Ql}\right) = \frac{6D_t}{l^2}\left(1 - \frac{\sin Ql}{Ql}\right) = \frac{6D_t}{l^2}(1 - \operatorname{sinc} Ql) \qquad [8]$$

since $D_t = l_2/6\tau$. As shown in **Figure 3A** for $ZrTi_2H_{3.6}$, the scattering function is still Lorentzian

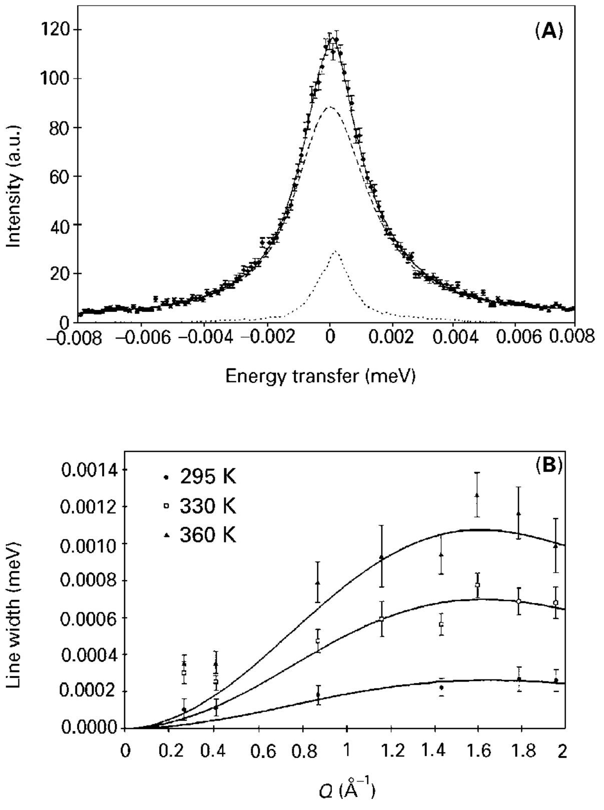



**Figure 3** (A) Scattering at $Q = 0.87$ Å$^{-1}$ and 360K from $ZrTi_2H_{3.6}$. The fit (solid line) to the data (•) is the sum of the quasielastic scattering with a Lorentzian line shape, (dashed line) and elastic scattering from the metal atoms (dotted line). (B) Line width of the Lorentzian component as a function of $Q$ for $T = 295$, 330 and 360 K. The solid lines are fits from the Chudley–Elliott model (Eqn [8]) with $l = 2.8$ Å. (Data recorded with IN10 at the ILL). Reproduced with permission from Fernandez JF, Kemali M, Johnson MR and Ross DK (1997) Quasielastic neutron scattering measurements on the $ZrTi_2H_{3.6}$ C-15 Laves phase compound. *Physica B* **234–236**: 903–905.

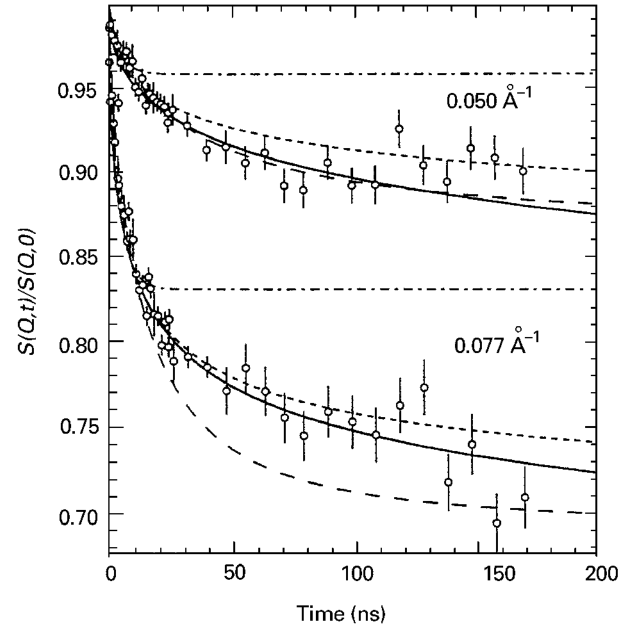



**Figure 4** Plot of the normalized intermediate scattering function [$S(Q,t)/S(Q,0)$] versus time for hydrogenous polyethylene (12% w/w) in deuterated polyethylene at 509 K for $Q = 0.050$ and 0.077 Å$^{-1}$. Also shown are the predictions from the reptation model (solid line) and competing (dashed and dot-dashed lines). (Data recorded with IN15 at the ILL). Reproduced with permission from Schleger P, Farago B, Lartigue C, Kollmar A and Richter D (1998) Clear evidence for reptation in polyethylene from neutron spin echo spectroscopy. *Physical Review Letters* **81**: 124–127.

but the line width has a more complex $Q$ dependence **(Figure 3B)**. This model, known as the Chudley–Elliott model, was originally developed for the case of diffusion in liquids. It has since been extended to take account of sites that are different both crystallographically and energetically.

Neutron spin echo has been largely used to look at the main chain dynamics of polymers in concentrated solution or in the melt. In these systems the entanglements between chains determine key physical properties, particularly the viscosity, which is an area of central concern to the processing of polymers. The most successful theory to explain the properties of such systems is that of reptation. This considers the polymer to be confined to a 'tube' formed by its neighbours and that the only motion is a snake-like creep along the tube, lateral motion being prevented by the walls of the tube. However, there are competing theories and there was no direct experimental evidence that the reptation model is correct. For the case of hydrogenous polyethylene (12% w/w) in deuterated polyethylene, recent improvements in instrumentation have allowed the models to be tested. **Figure 4** compares the experimental data and the predictions of the various models. Clearly only the reptation model (solid line) accurately describes the data over the entire time range, thus providing compelling evidence that reptation is the best model.

## Tunnelling spectroscopy

Tunnelling is the movement of a particle from one site to another at temperatures where, classically, it does not have sufficient energy for the transition. Thus tunnelling is intrinsically a quantum phenomenon and is essentially restricted to motion of hydrogen (and its isotopes). It can be rotational or translational in origin. **Figure 5** illustrates a model for the hindered rotation in a system with three equivalent hydrogens such as $-CH_3$ or $NH_3$. The crucial factors are the barrier height, $V$, and the rotational constant, $B$, of the molecule. As the ratio $V/B$ decreases, the overlap of the ground-state wavefunctions increases. Since the

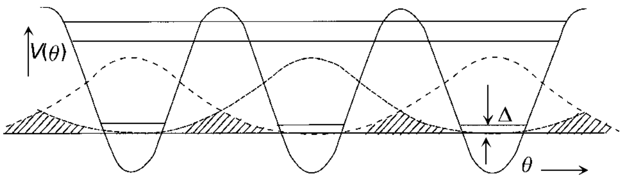



**Figure 5** Model of the tunnel splitting (Δ) of the rotational ground state and of the first excited torsional state. The splitting is due to the overlap of the wavefunctions (hatched region) in neighbouring wells.

energy levels of the protons in each well are the same, they are degenerate and they split to remove the degeneracy. The degree of splitting is the tunnelling splitting. Since the wavefunctions are approximately Gaussian, the splitting depends exponentially on the barrier height, making the tunnel splitting a very sensitive probe of the immediate environment of the molecule. INS is sensitive to tunnel splittings in the range 0.1 μeV ($10^{-4}$ $cm^{-1}$) to 5 meV (40 $cm^{-1}$).

Intercalation of graphite with small molecules is used to modify its physical and chemical (including catalytic) properties. **Figure 6A** shows the rotational tunnelling of ammonia in a caesium–graphite intercalate. $C_{28}Cs(NH_3)_x$. As $x$ varies from 0.0 to 1.90 the tunnelling peaks can be seen to grow and decay as the layers of caesium atoms in the layers become coordinatively saturated by ammonia ligands. The peaks occur in pairs at ± energies because even at the low temperature of the experiment, the first excited state is populated, thus both neutron energy gain and neutron energy loss peaks are observed. The staging behaviour is reproduced in the diffraction pattern of the intercalate which was recorded simultaneously (**Figure 6B**). The ability to obtain structural and dynamic information simultaneously is a unique aspect of neutron studies. Tunnelling is very strongly temperature dependent and is only usually observed at very low temperatures. As the temperature is raised the intensity and transition energy of the lines decrease, while simultaneously a broad quasielastic feature appears.

One of the major discoveries in chemistry in recent years is the characterization of a series of complexes where dihydrogen, $H_2$, acts as a ligand. For the compounds $M(CO)_3(\eta^2\text{-}H_2)(PR_3)$ (M = Mo, R = cyclohexyl; M = W, R = cyclohexyl, isopropyl) the tunnel spectra are shown in **Figure 7**. The observation of the tunnel splitting is proof that the H–H bond is intact. It can be seen that changing the metal has a large (300%) effect on the tunnel splitting (compare **Figure 7A** and **Figures 7B** and **7C**), whereas changing the phosphine has a minor (20%) effect (compare **Figure 7B** and **Figure 7C**), indicating that the electronic component of the M–$H_2$ bonding is more significant than the steric influences of the phosphines. (Two peaks in the gain and loss spectra are observed because of crystallographic disorder.)

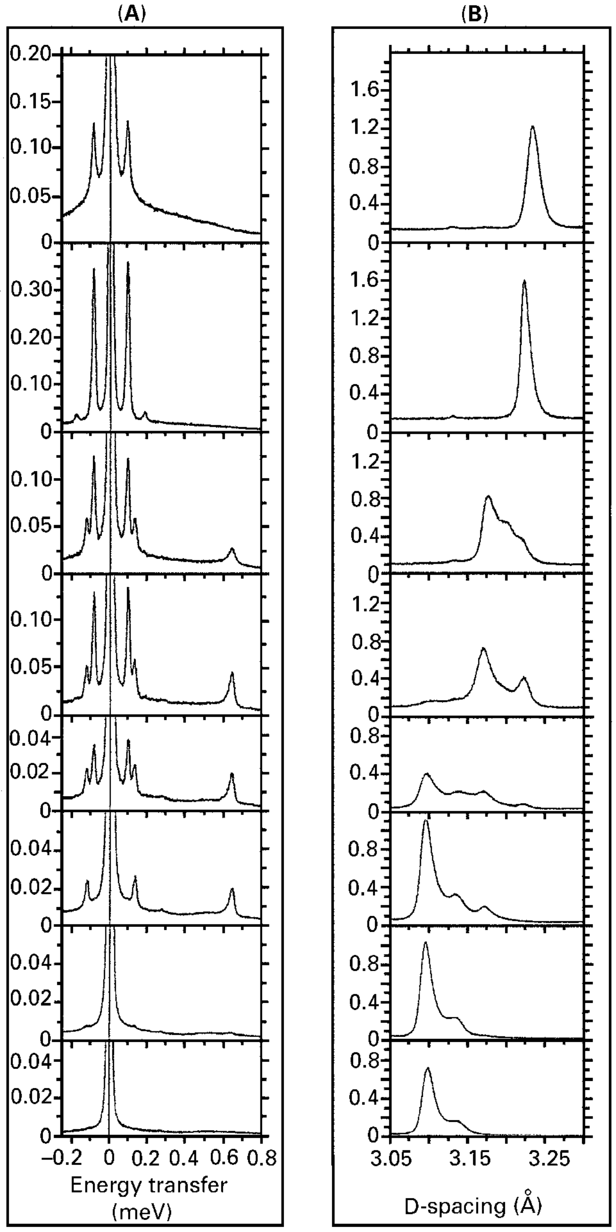



**Figure 6** Neutron studies of $C_{28}Cs(NH_3)_x$ (from top to bottom $x$ = 1.90, 1.54, 1.25, 0.94, 0.74, 0.42, 0.20, 0.0). (A) Rotational tunnelling spectra and (B) neutron diffraction patterns recorded simultaneously at 4.2 K as a function of $x$. (Data recorded with IRIS at ISIS). Reproduced with permission from Carlile CJ, Jamie IM, Lockhart G and White JW (1992) Two-dimensional caesium–ammonia solid solutions in $C_{28}Cs(NH_3)_x$. *Molecular Physics* **76**: 173–200.

Translational tunnelling is conceptually different since the hydrogen may be in a different crystallographic site after the tunnelling event. Hydrogen-in-metal systems may show this type of

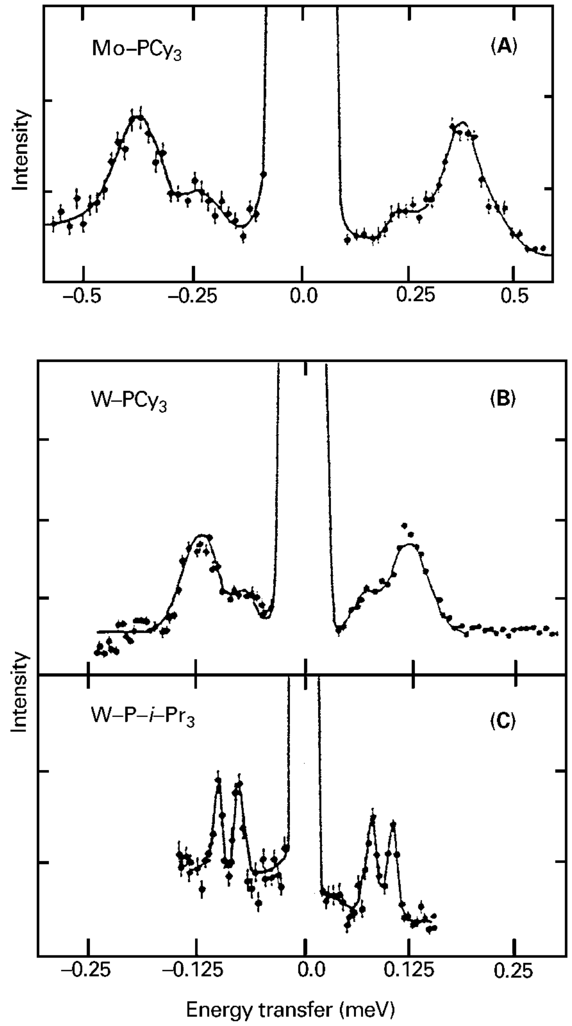



**Figure 7** Rotational tunnel spectra of dihydrogen complexes at 4 K: (A) $Mo(CO)_3(\eta^2\text{-}H_2)[P(C_6D_{11})_3]$, (Mo–$PCy_3$), (B) $W(CO)_3(\eta^2\text{-}H_2)[P(C_6D_{11})_3]$, (W–$PCy_3$) and (C) $W(CO)_3(\eta^2\text{-}H_2)$ $[P(\mathit{iso}\text{-}C_3D_7)_3]$, (W–P-*i*-$Pr_3$). Note that the energy scale in (B) and (C) is different from that in (A). (Data recorded with IN5 at the ILL). Reproduced with permission from Eckert J, Kubas GJ, Hall JH, Hay J and Boyle CM (1990) Molecular hydrogen complexes 6. The barrier to rotation of $\eta^2$-$H_2$ in $M(CO)_3(PR_3)(\eta^2\text{-}H_2)$ (M = W, Mo; R = Cy, *i*-Pr) inelastic neutron scattering, theoretical and molecular mechanics studies. *Journal of the American Chemical Society* **112**: 2324–2332.

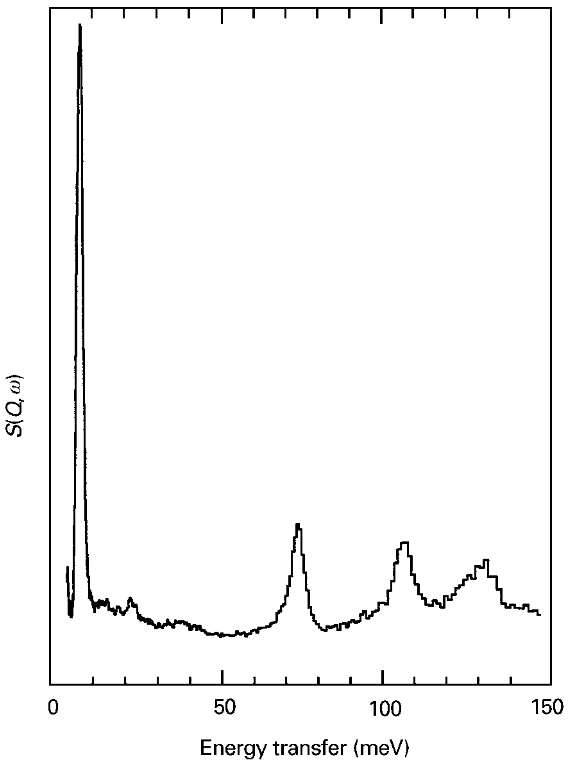



**Figure 8** Hydrogen tunnelling in $\alpha$-Mn. The intensity of the tunnelling line at 6.3 meV (51 $cm^{-1}$) is unusually large when compared to the optic modes at 80–120 meV (645–968 $cm^{-1}$). (Data recorded with TFXA at ISIS).

tunnelling. **Figures 8** shows the tunnelling transition of hydrogen in α-Mn. The tunnelling in this instance is between two sites of the same symmetry separated by 0.68 Å. This is a very unusual system in that the tunnelling energy is 6.3 meV (51 $cm^{-1}$), which is ~15 times higher than commonly found for hydrogen-in-metal systems, and the integrated intensity is very large when compared to the optic modes at 80–120 meV (645 – 968 $cm^{-1}$).

## The energy range 0–1 eV (0–8000 $cm^{-1}$): vibrational spectroscopy and magnetic excitations

### Coherent INS spectroscopy

Coherent INS spectroscopy gives information on collective motions in crystals. These are important in understanding phenomena as diverse as magnetism, phase transitions and thermodynamic properties. Neutron diffraction is coherent elastic scattering from a notionally static lattice; coherent one-phonon neutron scattering can be considered to be ‘diffraction’ from a lattice where the atoms are undergoing sinusoidal motion of frequency given by the phonon frequency. This means that the scattering occurs in specific directions at specific energies. It is to be emphasized that coherent INS is the most general method of measuring the excitations (which can be vibrational or magnetic in origin) across the entire Brillouin zone.

Palladium hydride has been extensively studied because of its ready take-up and large capacity for hydrogen. This has lead to interest in its use as a

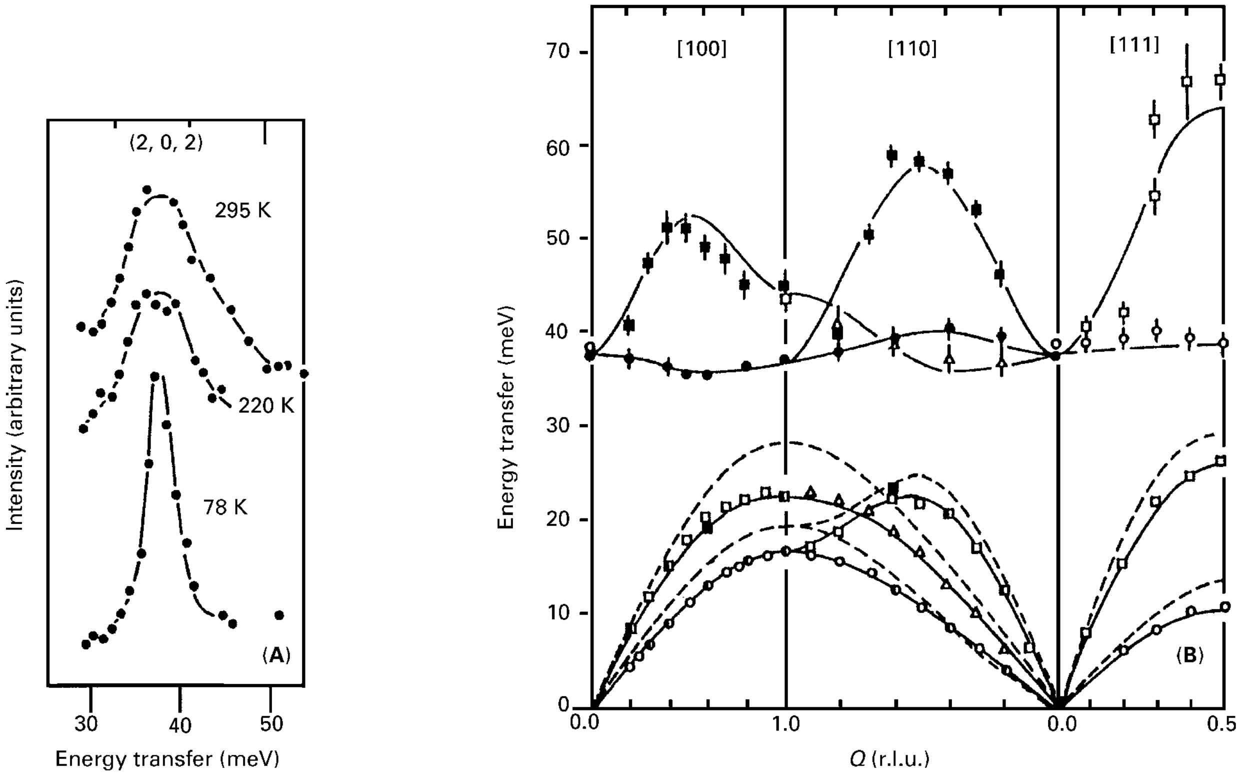



**Figure 9** (A) Example of a phonon group in $PdD_{0.63}$ and (B) the dispersion curves for $PdD_{0.63}$. The dashed lines are the dispersion curves for pure Pd (r.l.u. = reciprocal lattice unit) (data recorded with HB3 at Oak Ridge). Reproduced with permission from: Rowe JM, Rush JJ, Smith HG, Mostoller M and Flotow HE (1974) Lattice dynamics of a single crystal of $PdD_{0.63}$. *Physical Review Letters* **21**: 1297–1300.

hydrogen storage material. **Figure 9A** shows one of the measured transitions for $PdD_{0.63}$ and **Figure 9B** the dispersion curves derived from the data. The lower set of curves are the acoustic modes, where all the atoms in the unit cell move in the same direction (as is the case for sound waves) and largely reflect the dynamics of the metal atoms. The higher energy curves are the optic modes and different types of atom move in different directions. This generates a changing dipole, and these modes would give rise to infrared or Raman active modes. Infrared and Raman spectroscopies only observe modes at approximately zero wave vector and from **Figure 9B** only the optic modes have nonzero frequencies at $Q = 0$. It is one of the great strengths of INS that modes at all wave vectors can be studied.

The comparatively low energy of the optic modes indicates that the Pd–D force constant is small. For $PdD_{0.63}$ the variation of energy transfer with wave vector (dispersion) of the optic modes shows that there are strong interactions between the H atoms. To fit the data fourth-nearest neighbours have to be included. Comparision of the acoustic modes for $PdD_{0.63}$ and Pd (dashed lines) shows that the former are lower in energy so the bonding is weaker. This is consistent with the lattice expansion observed on hydrogen (deuterium) take-up.

The unique properties of water are essential to life but are still not understood in detail. They derive from the hydrogen bonding present in water which is manifested in the complex phase diagram and anomalously high melting point of ice. Coherent INS studies of ice $I_h$ provide a stringent test of models of the dynamics. **Figure 10A** shows one of the measured spectra for a single crystal of $D_2O$ ice $I_h$ and **Figure 10B** shows the derived dispersion curves. It can be seen that while the longitudinal acoustic (LA) and transverse acoustic (TA) modes, which mainly involve oxygen motion are sharp, the transverse optic (TO) mode is broad. The TO mode is largely due to deuteron motion and the width probably reflects the highly disordered arrangement of deuterons (protons) in ice $I_h$.

Virtually all of our understanding of magnetism at the atomic level is based on neutron scattering; neutron diffraction allows the orientation of the spins in the lattice to be determined and inelastic scattering allows their dynamics to be studied.

Magnetoresistance is a phenomenon observed in some materials where the electrical resistivity

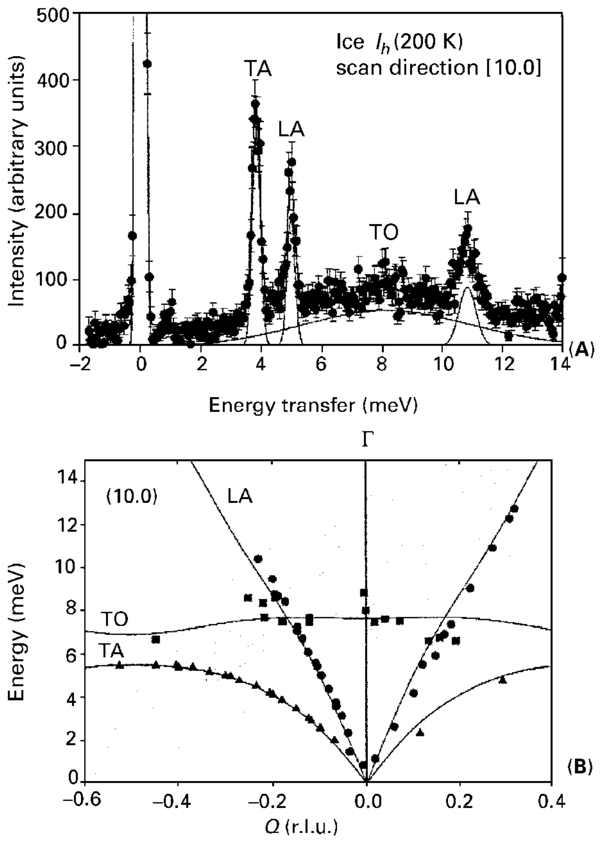



**Figure 10** (A) One of the measured spectra of the phonon dispersion along the (10.0) direction about the (11.0) reciprocal lattice point of $D_2O$ ice $I_h$. (B) The derived dispersion curves; the solid lines are fits to the TA and LA modes. The dashed lines show the $Q, \omega$ trajectories available in this experiment (r.l.u. = reciprocal lattice unit) (data recorded with PRISMA at ISIS).

changes in an applied magnetic field. In recent years there has been a huge revival in the study of manganese perovskites which, depending on the addition of other elements, can change their resistance by several orders of magnitude in fields of a few tesla; this is known as colossal magnetoresistance (CMR). The interest in these materials is not just because of possible technological applications, particularly read-out heads for hard disks that will allow much higher storage densities and faster access, but also because of fundamental questions relating to the transition between metallic and insulating phases.

The CMR manganites are metallic and derived from the insulating ionic compound $LaMnO_3$. Partial substitution of $La^{3+}$ by a divalent metal ion e.g. $Ba^{2+}$, $Sr^{2+}$, $Pb^{2+}$, introduces mobile charge carriers and these are responsible for the unusual behaviour. The materials are ferromagnets in which the magnetic moments associated with the manganese ions are aligned in the same direction below the Curie temperature. In 1930 Bloch showed that the lowest energy excitations in a ferromagnet are spin waves; coherent precessions of the spins around their average orientation. **Figure 11** shows an image of the inelastic magnetic scattering by a single crystal of $La_{0.7}Pb_{0.3}MnO_3$. The spin wave is clearly seen as a ring of scattering around the (1,0,0) reciprocal lattice point. From the spectra the dispersion relation for the spin wave can be obtained by plotting the position of maximum intensity along the main directions of crystal symmetry. Fitting the data shows that the dynamics can be described by only considering nearest neighbours, a surprising result for such a complex material.

## Incoherent INS spectroscopy

Incoherent INS spectroscopy probes the local environment of the scattering atom and is thus complementary to infrared and Raman spectroscopies. However there are important differences as illustrated in **Figure 12A–C** which shows the infrared, Raman and INS spectra respectively, of *N*-phenylmaleimide, a model compound for bismaleimide composites that are used in aerospace applications. It can be seen that all three spectra are very different. There are two main reasons for this. Firstly, whether a mode is infrared and/or Raman active (or is inactive in both) is determined by the symmetry of the molecule. Note that even when a mode is allowed it may still have little or no intensity. In contrast, there are no selection rules for INS spectroscopy and all modes are allowed.

The strongest bands in the infrared and Raman spectra are the out-of-phase carbonyl stretch at 212 meV (1707 $cm^{-1}$) and the in-phase carbonyl stretch at 219 meV (1770 $cm^{-1}$) respectively. In contrast, the carbonyl bands are completely absent in the INS spectrum, but there are several strong bands that do not have any obvious counterparts in the infrared and Raman spectra.

The differences arise because the intensity of the $i$th INS band at a frequency $\omega_i$ is proportional to:

$$S(Q,\omega_i) \propto \sum_d \frac{|QU_d^i|^2}{m_d\,\omega_i}\sigma_d^{\rm inc}\exp(-Q^2U_{\rm total}^2) \qquad [9]$$

The exponential term is the Debye–Waller factor, $U_{\rm total}^2$ is the mean square displacement of the molecule and its magnitude is in part determined by the thermal motion of the molecule. This can be reduced by cooling the sample and so spectra are typically recorded below 30 K.

$U_d^i$ is the amplitude of vibration of the $d$th atom in the $i$th mode, $\sigma_d^{\rm inc}$ and $m_d$ are its incoherent cross section and mass respectively. From **Table 1** it can be seen that if hydrogen is present, the combination

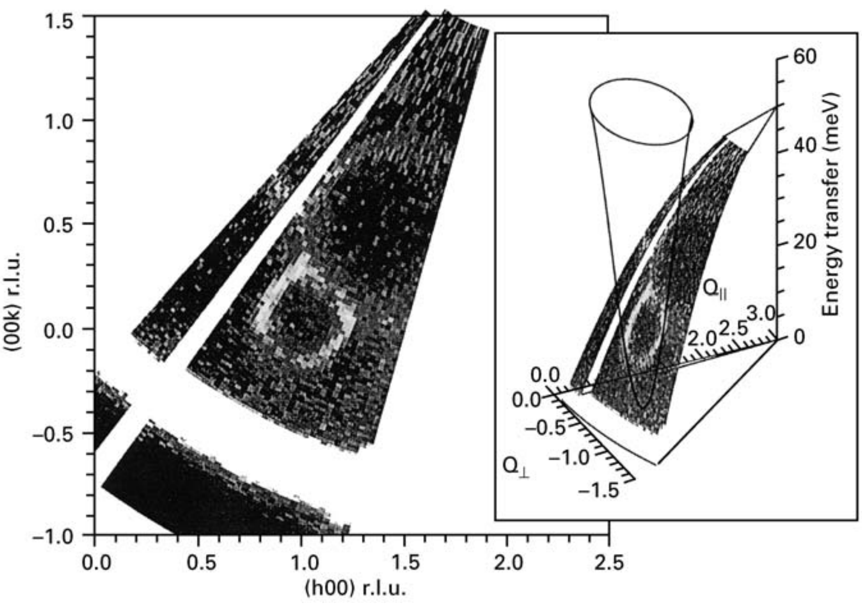



**Figure 11** (Left) Image of the magnetic excitations in $La_{0.7}Pb_{0.3}MnO_3$. (Right) Data plotted in a three-dimensional space defined by two coordinates of momentum in the horizontal plane and the energy vertically. The shaded surface indicates the possible energies and momenta of the scattered neutrons that are accessible in this experiment; maximum intensity occurs where this surface intersects the dispersion surface of the spin waves (r.l.u. = reciprocal lattice unit) (data recorded with HET at ISIS).

of large incoherent cross section and small mass means that the scattering from hydrogen will dominate the spectrum. For hydrogenous materials the scattering from all other atoms can be neglected. This dependence on the cross section is why the INS spectrum is so different from the optical spectroscopies. There, the intensity derives from changes in the electronic properties of the molecule that occur as the vibration is executed (the dipole moment and the polarizability for infrared and Raman spectroscopy respectively).

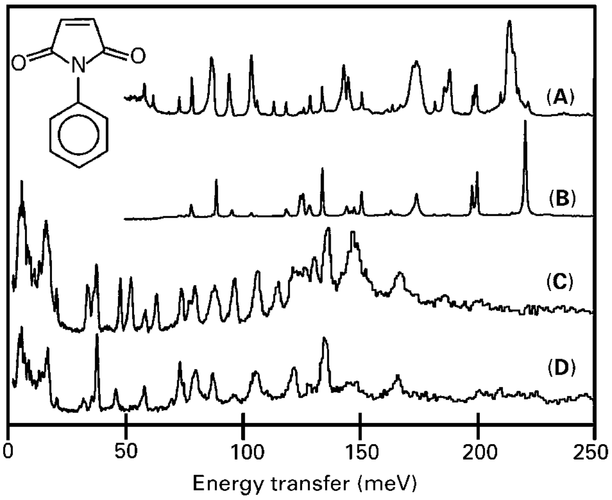



**Figure 12** Vibrational spectra of *N*-phenylmaleimide: (A) infrared, (B) Raman, (C) INS; (D) INS spectrum of *N*-(perdeuterophenyl)maleimide (INS data recorded with TFXA at ISIS).

The cross-section dependence can be exploited to 'eliminate' parts of the molecule from the spectrum. **Figure 12D** is again *N*-phenylmaleimide but with the phenyl ring deuterated. This results in a considerable simplification of the spectrum and all of the modes can be assigned to vibrations of the maleimide ring with the phenyl group treated as a point mass.

Vibrational spectroscopy is a commonly used tool for the study of catalysts. The neutron is a particle with zero charge which means that it is highly penetrating and is not a surface-sensitive probe. This can be circumvented by using high surface area materials and hydrogenous adsorbates that offer contrast to the substrate. The advantages are that common catalyst-supports such as silica, alumina and carbon are weak scatterers, see **Table 1**, and the entire 0–500 meV (0–4000 $cm^{-1}$) 'mid-infrared' range is accessible, and in particular the low-energy range below 125 meV (1000 $cm^{-1}$). This range is very difficult to study with optical methods; the supports tend to be highly absorbing and the important metal–adsorbate vibrations are often very weak.

Formate (HCOO) species adsorbed on a copper-based catalyst are a key intermediate in the industrial synthesis of methanol from $CO_2$ and $H_2$. INS spectra were recorded for a number of different copper surfaces; spectra were taken before and after formic

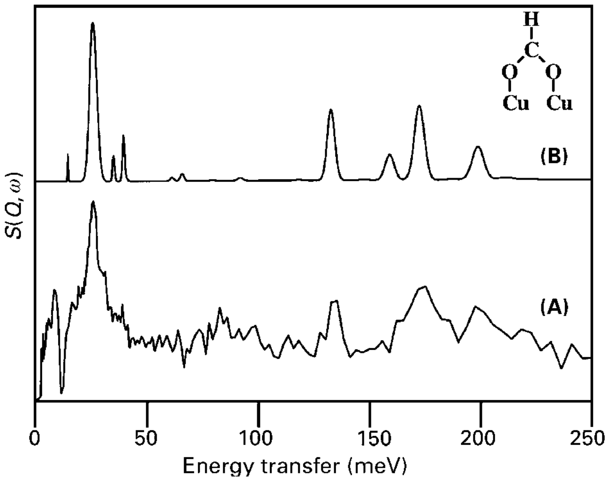



**Figure 13** Experimental (A) and calculated (B) spectra of bridging formate on CuO. The experimental spectrum is the difference spectrum [CuO + formate] – [CuO]. (Data recorded with TFXA at ISIS).

acid adsorption. In all the spectra adsorbed, formate could be clearly distinguished even though it represented less than 1% of the total mass in the beam. The difference spectrum ([CuO + formate] – [CuO]) is shown in **Figure 13A**.

Equation [9] is noteworthy in that the intensity of a mode is determined only by the cross section, the momentum transfer and the amplitude of vibration, hence it is purely dynamic. From a conventional Wilson–Decius–Cross normal coordinate analysis it is possible to calculate both the energies (from the eigenvalues) and intensities (from the eigenvectors) for INS spectra of molecular species. The requirement to fit the INS intensities is a stringent test of any force field. The spectrum calculated for a bidentate configuration of formate bound to two Cu atoms (**Figure 13B**) gives good agreement with the experimental data, supporting this as the mode of binding on this surface. It also demonstrated that the mode at 26.8 meV (216 $cm^{-1}$) is the Cu–O torsion, which was the first time this mode was detected on a surface.

The rapid growth in computer power means that increasingly complex systems can be treated by *ab initio* quantum chemistry techniques. The vibrational spectrum of the object of interest is often calculated, since this offers a check of whether the structure is a true minimum or a saddle point on the potential surface, as well as being a rigorous test of the results. The amplitudes of vibration are calculated as part of this process and it is relatively straightforward to derive the INS spectrum from these via Equation [9]. In contrast, the infrared and Raman intensities are poorly reproduced for anything but the simplest systems. The use of *ab initio* methods is undoubtedly the way that INS data will be routinely analysed in the near future. The power of the method is illustrated in **Figure 14** where it allows an unambiguous differentiation between two possible modes of water binding to the Brönsted acid sites of a zeolite. Clearly, the hydrogen-bonded water model fits the data better than the hydroxonium ion, $H_3O^+$, model.

## The energy range >1 eV (> 8000 $cm^{-1}$): neutron Compton scattering

Neutron Compton scattering (NCS) is the measurement of atomic momentum distributions, $n(p)$, in condensed matter systems by high-energy inelastic neutron scattering. NCS measurements are particularly important in the study of quantum fluids.

For protons, $n(p)$ is related to the Fourier transform of the proton wavefunction and an NCS measurement of $n(p)$ can be used to determine the wavefunction in a way analogous to the determination of a real space structure from a diffraction pattern. If $n(p)$ is known then, in principle, both the proton wavefunction and the exact form of the potential energy well can be reconstructed.

For accurate NCS measurements energy transfers much greater than the maximum vibrational frequency are required for the impulse approximation to be valid. In the impulse approximation, momentum and energy are conserved in the scattering event and from a measurement of the momentum and energy change of the neutron, the momentum of the proton before the collision can be determined. This requires neutrons with energies >1 eV and only spallation sources provide a sufficient flux of high-energy neutrons for the measurement to be feasible. At lower energies the impulse approximation is no longer valid and $n(p)$ is not related in a simple way to the observed intensities.

An important feature is that the neutron energy loss depends on the atomic mass of the scattering atom and hence different masses occur at different times-of-flight. In the harmonic approximation, the peak width depends on the Gaussian average of the vibrational frequencies. The peak intensity depends only on the number of scatterers and their cross section. Thus NCS functions as a 'neutron mass spectrometer', albeit with poor mass resolution. Hydrogen is usually well separated from other elements and allows its quantification, even in the presence of deuterium.

This aspect of NCS has been exploited in the study of the battery material $LaNi_5$. For a well-annealed polycrystalline sample of $\alpha$-$LaNi_5$ after hydrogen

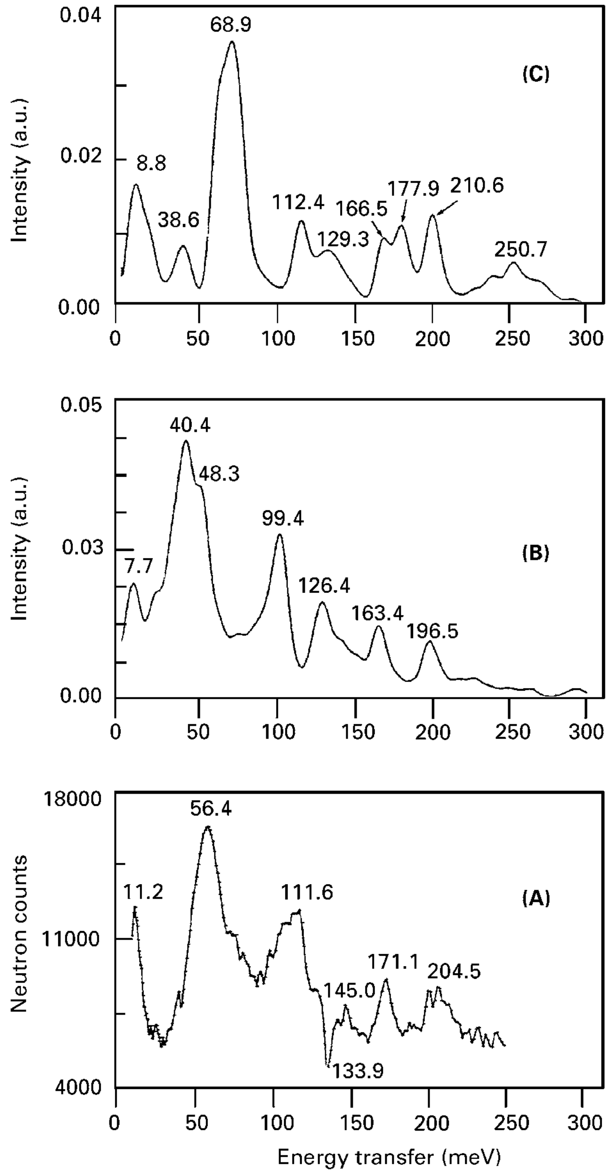



**Figure 14** INS spectrum of water adsorbed at low loading on H-ZSM5. (A) Experimental spectrum. *Ab initio* simulation for: (B) water hydrogen-bonded to a bridging hydroxyl and (C) a hydroxonium ion, $H_3O^+$. (Data recorded with IN1BeF at ILL). Reproduced with permission from Jobic H, Tuel A, Krossner M and Sauer J (1996) Water in interaction with acid sites in H-ZSM-5 zeolite does not form hydroxonium ions. A comparison between neutron scattering results and *ab initio* calculations. *Journal of Physical Chemistry* **100**: 19545–19550.

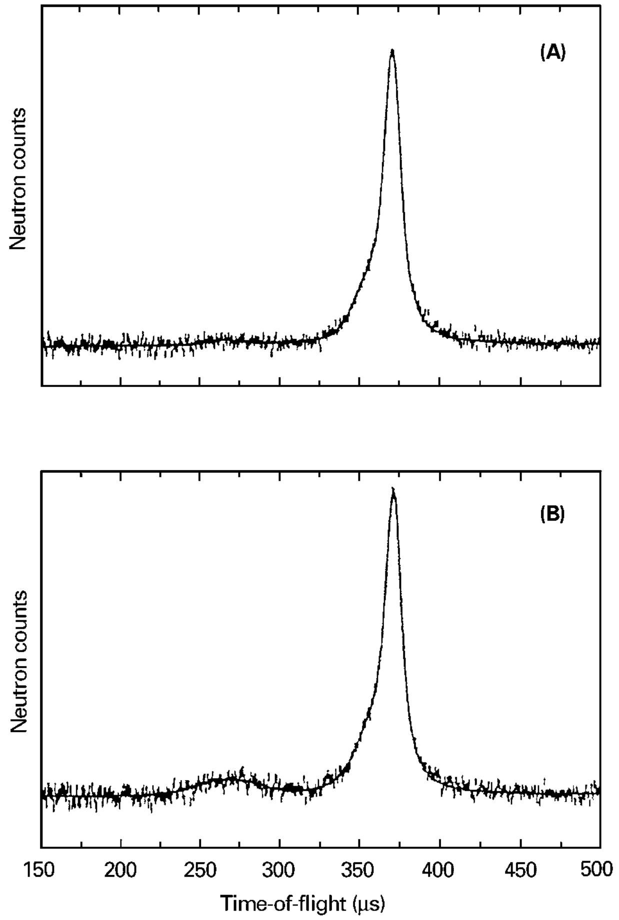



**Figure 15** Time-of-flight spectra for $\alpha$-$LaNi_5$: (A) nominally hydrogen-free (however, note the small peak at 275 μs due to hydrogen), (B) after absorption of hydrogen. (Data recorded with eVS at ISIS). Reproduced with permission from Gray EMA, Kemali M, Mayers J and Norland J (1997) Interstitial site occupation in $\alpha$-$LaNi_5$ studied by deep inelastic neutron scattering. *Journal of Alloys and Compounds* **253–254**: 291–294.

desorption, there should be no residual hydrogen. Figure 15A shows an intense peak at 375 μs due to the metal atoms and a weak feature at 275 μs assigned to hydrogen. This is confirmed by the absorption of hydrogen, resulting in an increase of the scattering at 275 μs (Figure 15B). Analysis shows that the nominally hydrogen-free material has a hydrogen-to-metal ratio of 0.017 ± 0.002, indicating that there is a significant number of trap sites, which are probably detrimental to battery performance.

## List of symbols

$A_0(Q)$ = elastic incoherent structure factor (EISF); $B$ = rotational constant; $D_{chem}$ = chemical diffusion coefficient; $D_t$ = tracer diffusion coefficient; $J_0(x)$ = zeroth-order spherical Bessel function; $L(\omega)$ = Lorentzian line shape; $m_d$ = mass of $d$th atom; $n(p)$ = atomic momentum distribution; $Q$ = momentum transfer; $S^{coh}(Q, \omega)$ = coherent scattering function; $S^{inc}(Q, \omega)$ = incoherent scattering function; $\langle u^2 \rangle$ = mean square displacement in the vibrational modes; $U^2_{total}$ = mean square displacement; $U^i_d$ =

amplitude of vibration of the *d*th atom in the *i*th mode; $V$ = barrier height; $\sigma_d^{inc}$ = incoherent cross section of *d*th atom; $\tau$ = average time between jumps; $\omega$ = energy transfer.

*See also:* **Inelastic Neutron Scattering, Instrumentations; IR Spectral Group Frequencies of Organic Compounds; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; Scattering Theory; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading


Bée M (1988) *Quasielastic Neutron Scattering*. Bristol, UK: Adam Hilger.

Eckert J and Kearley GJ (eds) Spectroscopic applications of inelastic neutron scattering: Theory and practice (1992) *Spectrochimica Acta, Part A* **48**: 269–478.

Furrer A (ed) (1994) *Neutron Scattering from Hydrogen in Materials*. Singapore: World Scientific.

Higgins JS and Benoît HC (1994) *Polymers and Neutron Scattering*. Oxford: Oxford University Press.

Newport RJ, Rainford BD and Cywinski R (eds) (1988) *Neutron Scattering at a Pulsed Source*. Bristol, UK: Adam Hilger.

Prager M and Heidemann A (1997) Rotational tunnelling and neutron spectroscopy. *Chemical Reviews* **97**: 2933–2966.

Ross DK (1997) Neutron scattering studies on metal hydrogen systems In: Wipf H (ed) *Hydrogen in Metals III: Properties and Applications (Topics in Applied Physics. Vol. 73)*, pp. 153–214. Berlin: Springer-Verlag.

Squires GL (1996) *Introduction to the Theory of Thermal Neutron Scattering*. New York: Dover Publications.


# Inelastic Neutron Scattering, Instrumentation


**Stewart F Parker,** Rutherford Appleton Laboratory, Didcot, UK



Copyright © 1999 Academic Press


**HIGH ENERGY SPECTROSCOPY**
**Methods & Instrumentation**

## Introduction

The prime requirement for inelastic neutron scattering (INS) spectroscopy is a source of neutrons. The only methods that produce a sufficient flux of neutrons (> $10^{14}$ neutrons $cm^{-2}$ $s^{-1}$) for spectroscopy are: $^{235}U$ fission in a nuclear reactor and spallation. The latter process involves accelerating a beam of protons to near-light-speed and impacting them onto a heavy-metal target, usually made of either depleted uranium or tantalum. The target atoms absorb the protons, generating highly excited nuclei that decay, in part, by emission of neutrons. In the case of a uranium target, fission also occurs, increasing the neutron yield. Typically, a reactor yields one neutron per fission event that is available for spectroscopy, while spallation yields up to 30 neutrons per proton absorbed. At present, the most powerful reactor source is the Institut Laue Langevin (ILL) (Grenoble, France) and the most powerful spallation source is the ISIS Facility at the Rutherford Appleton Laboratory (Chilton, UK). Usually, fission sources operate continuously and spallation sources are pulsed, although this is not always the case. The reactor at the Franck Laboratory of Nuclear Physics (Dubna, Russia) consists of two pieces of sub-critical uranium, one of which is mounted on a rotating disc; as the two pieces come into proximity, the mass goes critical, producing neutrons, and as it goes sub-critical again, neutron production ceases. The net result is a pulsed reactor. The SINQ facility at the Paul Scherrer Institute (Villigen, Switzerland) is a continuous spallation source.

By whatever process they are produced, the ‘newborn’ neutrons are very energetic, $> 2 \times 10^9$ meV (1 meV = 8.067 $cm^{-1}$ = 0.241 THz) and must be brought to useful energies by multiple inelastic collisions in a moderator; this is usually a hydrogenous material. By suitable design it is possible to establish a quasi-thermal equilibrium between the temperature of the moderator and the energy of the neutrons, which provides a means of tailoring the neutrons’ energy to the requirements of the experiment.

Neutrons are quantum entities, so they exhibit both particle-like and wave-like properties. For INS spectroscopy, the particle-like properties are relevant since the neutron–sample interaction occurs on a femtosecond timescale and the interaction can be considered to be analogous to billiard-ball scattering.

**Table 1** Relationship between neutron properties and typical values

| Neutron energy | | | | Wavelength | Wave vector |
|---|---|---|---|---|---|
| meV | $cm^{-1}$ | THz | K | (Å) | ($Å^{-1}$) |
| $E$ | | | $=k_BT$ | $=\frac{h^2}{2m\lambda^2}$ | $=\frac{\hbar^2k^2}{2m}$ |
| $E$ | $=8.066\,7E$ | $=\frac{E}{4.135\,4}$ | $=0.086\,165T$ | $=\frac{81.787}{\lambda^2}$ | $=2.071\,7k^2$ |
| 0.01 | 0.081 | 0.002 42 | 0.116 | 90.4 | 0.069 5 |
| 0.1 | 0.867 | 0.024 2 | 1.16 | 28.6 | 0.220 |
| 1.0 | 8.067 | 0.241 8 | 11.16 | 9.04 | 0.695 |
| 10.0 | 80.7 | 2.418 | 116.1 | 2.86 | 2.197 |
| 100.0 | 807 | 24.18 | 1 161 | 0.904 | 6.948 |
| 1 000.0 | 8 067 | 241.8 | 11 606 | 0.286 | 21.97 |
| 10 000.0 | 80 667 | 2 418 | 116 056 | 0.090 | 69.48 |

meV = millielectronvolts, $cm^{-1}$ = wavenumbers, THz = terahertz ($10^{12}$ Hz), K = kelvins, Å = angstroms ($10^{-10}$ m), $h$ = Planck's constant, $\hbar = h/2\pi$, $m$ = neutron mass, $k_B$ = Boltzmann's constant.

However, the selection of the incident energy or the analysis of the scattered neutrons' energy often uses Bragg diffraction from a single crystal, making use of the wave-like properties. **Table 1** shows the relationships between the units commonly used.

INS spectroscopy spans an enormous energy range, which is equivalent to a wide range of timescales. **Figure 1** illustrates the ranges available and some of the applications. It is convenient to divide the whole energy range into three sectors, while recognizing that there is significant overlap between them. In the lowest energy range, ±10 meV, the applications are principally quasielastic scattering and tunnelling spectroscopy. The second range, 0–1000 meV, covers the regions of vibrational spectroscopy and magnetic excitations. The highest range above 1000 meV is the province of neutron Compton scattering.

The major difference between INS spectroscopy and the photon-based spectroscopies that cover the same energy range is that the neutron has a significant mass (1.009 amu). Thus an inelastic collision

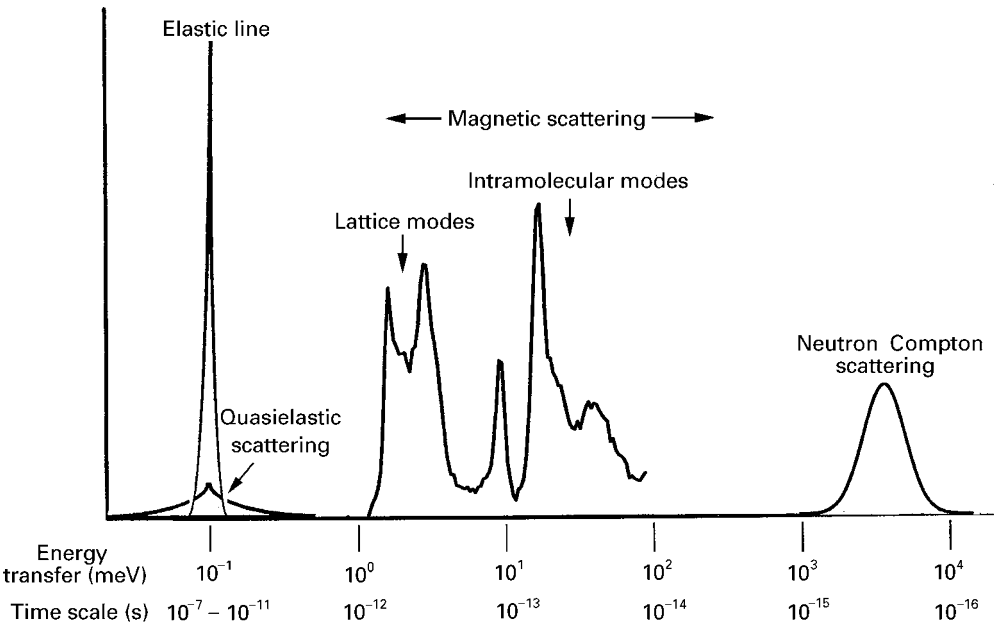



**Figure 1** Illustration of the energy range and the equivalent timescales and length scales available with inelastic neutron scattering. The type of information available is also indicated.

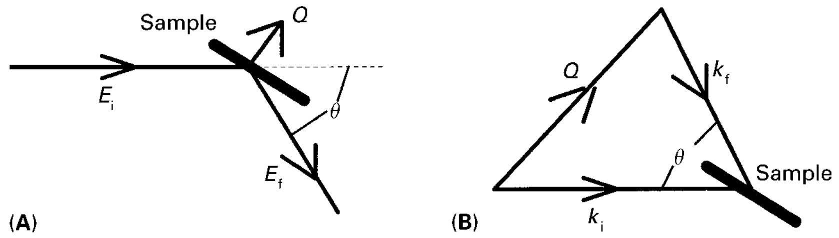



**Figure 2** Neutron scattering in (A) real space and (B) reciprocal space.

results in transfer of both momentum ($Q$, Å$^{-1}$) and energy ($E$, meV). In contrast, optical processes result in essentially zero momentum transfer. The aim of an INS experiment is, usually, to measure scattered intensity at each energy transfer as a function of the momentum transfer. In **Figure 2** the scattering process is shown in both real space and reciprocal space.

The energy transfer, $\hbar\omega$, is given by the difference in the incident (i) and final (f) neutron energies:

$$\hbar\omega = E_\mathrm{i} - E_\mathrm{f} \qquad [1]$$

where $E$ is the kinetic energy of the neutron and

$$E = mv^2/2 \qquad [2]$$

where $m$ is the neutron mass and $v$ its velocity. The momentum transfer, $Q$, is given by:

$$Q = \sqrt{(k_\mathrm{i}^2 + k_\mathrm{f}^2 - 2k_\mathrm{i}k_\mathrm{f}\cos\theta)} \qquad [3]$$

where $\theta$ is the scattering angle and $k$ (Å$^{-1}$) is the wavevector:

$$k = 2\pi/\lambda \qquad [4]$$

associated with the wavelength, $\lambda$ (Å). Thus INS offers the possibility of exploring both momentum and energy transfer and can be considered to be a 'two-dimensional' form of spectroscopy. This is illustrated in **Figure 3** which shows the $Q$, $E$ coverages of some of the instruments at ISIS and the ILL.

## The energy range ±10 meV: quasielastic scattering and tunnelling spectroscopy

Quasielastic scattering is a very low energy inelastic process. The term is usually taken to mean a broadening of the elastic line and is most commonly the result of diffusional (translational or rotational) motion of atoms. (It can also be caused by the randon fluctuations of unpaired electronic spins of

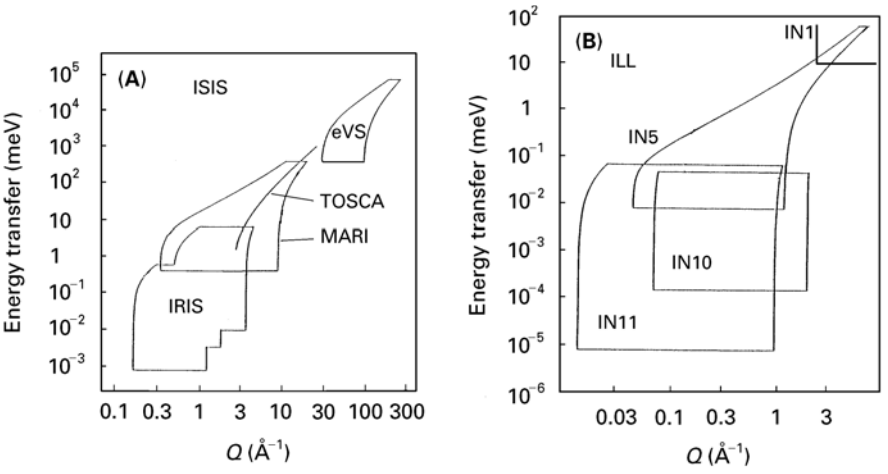



**Figure 3** The $Q$, $E$ coverages of some of the instruments at (A) the ILL and (B) ISIS.

stationary atoms). Translational and rotational diffusion occur simultaneously but on different timescales. The energy resolution of a spectrometer, $\Delta E$, and the timescale $\tau$ of the motion, are related by the Heisenberg uncertainty principle:

$$\Delta E \times \tau \sim h/2\pi \qquad [5]$$

Thus rapid motions require relaxed resolution, while slower motions require high resolution. Note that these are relative terms; typically an energy resolution of better than 1% $\Delta E/E$ and 100 μeV is required. The timescale to be probed can be separated into three regimes, each of which uses a different technique: for $\tau \sim 10^{-11}$ s, $\Delta E$ is 10-100 μeV and direct-geometry time-of-flight is used, for $\tau \sim 10^{-9}$ s, $\Delta E$ is 0.3-20 μeV and a backscattering crystal analyser is used, for $\tau \sim 10^{-7}$ s, $\Delta E$ is 0.005-1 μeV and neutron spin echo is used. Examples of each type of spectrometer will be considered.

## Direct-geometry time-of-flight

For time-of-flight INS spectroscopy it is necessary to determine either the incident energy or the final energy of the neutron. Spectrometers that define the neutron's final energy are indirect-geometry instruments and those that determine the neutron's incident energy are direct-geometry instruments. In a direct-geometry quasielastic spectrometer, a monochromatic (single energy), pulsed beam of neutrons is produced by a rotating mechanical chopper system. This pulse of neutrons is incident upon the sample and, after interacting with the atoms of the sample, the scattered neutrons are timed over a known distance to the detectors which are set at a range of scattering angles. The time-of-flight, $T_1$, from the sample to the detectors is given by $T_1 = L_1/v_1$, where $L_1$ is the sample-to-detector distance. Since the incident neutron velocity, $v_0$, is known it follows from Equations [1] and [2] that the energy transfer, $\hbar\omega$ is:

$$\hbar\omega = \frac{m}{2}(\nu_1^2 - \nu^2) = \frac{mL_1^2}{2}\left(\frac{1}{T_1^2} - \frac{1}{T^2}\right) \qquad [6]$$

where $T_0$ is the time-of-flight of the elastically scattered neutrons. Using Equation [3] and the known angular position of the detectors the momentum transfer $Q$ of the detected neutrons can also be determined.

**Figure 4** illustrates the principle for the multichopper time-of-flight spectrometer IN5 at the ILL. The incoming neutron distribution spreads from ~36 meV (1.5 Å) to ~0.25 meV (18 Å) with a maximum flux at 3.3 meV (5 Å). As chopper 1 opens, a pulse of neutrons is admitted; the only neutrons that pass through the last chopper (number 4) are those with the correct velocity to arrive at chopper 4 as it opens. Thus the energy of the neutrons incident on the sample is determined by the phase angle between the first and the last choppers. Neutrons with velocities that are one-half of the desired value, would also pass through chopper 4

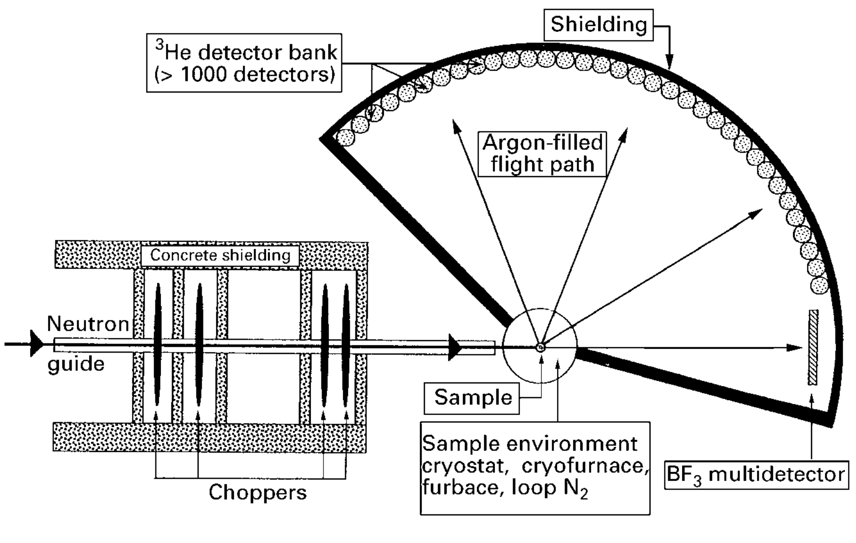



**Figure 4** Layout of the cold-neutron multi-chopper spectrometer IN5 at the ILL. Reproduced with permission of Adenine Press from Ferrand M (1997) Neutron instrumentation in studying dynamics of biomolecules. In: Cusack S, Büttner H, Ferrand M, Langan P and Timmins P (eds). *Biological Macromolecular Dynamics*. Schenectady: Adenine Press.

after a further revolution; these are eliminated by chopper 2. Chopper 3 spins at an integral fraction of the speed of the other choppers and eliminates a certain number of pulses, e.g. 1 in 2, or 3 in 4. This is to prevent frame-overlap, where fast neutrons from a following pulse overtake slow neutrons from the previous pulse. This is a problem since the time-of-flight method depends on knowing when a given neutron was created; from Equation [6] any uncertainty in the time-of-flight translates into a corresponding uncertainty in the energy transfer.

### Backscattering crystal analyser

The backscattering crystal analyser is an example of an indirect-geometry spectrometer and is used when higher resolution is required. Neutrons with a band of energies whose width (in energy) is greater than the maximum energy transfer to be measured, are incident on the sample. From the scattered neutrons, one neutron energy is selected by Bragg diffraction from a single crystal. Bragg's law states:

$$\lambda = 2d \sin\theta \qquad [7]$$

where $\lambda$ is the incident neutron wavelength, $d$ is the interplanar spacing in the analyser crystal and $\theta$ is the angle of incidence. The resolution is determined by the differential:

$$\frac{\Delta E}{E} = 2\frac{\Delta\lambda}{\lambda} = 2\left([\Delta\theta \cot\theta]^2 + \left[\frac{\Delta d}{d}\right]^2\right)^{0.5} \qquad [8]$$

The dominant term in most cases is cot $\theta$ since for most single crystals the uncertainty in the lattice parameter is ~$10^{-4}$ and can be ignored. In the limit as the Bragg angle approaches 90°, cot $\theta$ tends to zero and the energy resolution of the analyser becomes extremely good, being 0.5 μeV for silicon and 10 μeV for graphite. Hence, the term 'backscattering spectrometer'.

The implementation of a backscattering spectrometer is somewhat different at a reactor than at a spallation source. **Figure 5** shows a schematic view of the spectrometer IN10 at the ILL and **Figure 6** shows the equivalent spectrometer IRIS at ISIS.

IN10 sits at the end of a cold-neutron guide. For low-energy (long-wavelength) neutrons, total external reflection occurs with a corresponding increase in transmission as compared to a tube of the same diameter. This provides a means of transporting neutrons long distances, several tens of metres, without the intensity decrease that would result from solid-angle considerations. High-energy neutrons do

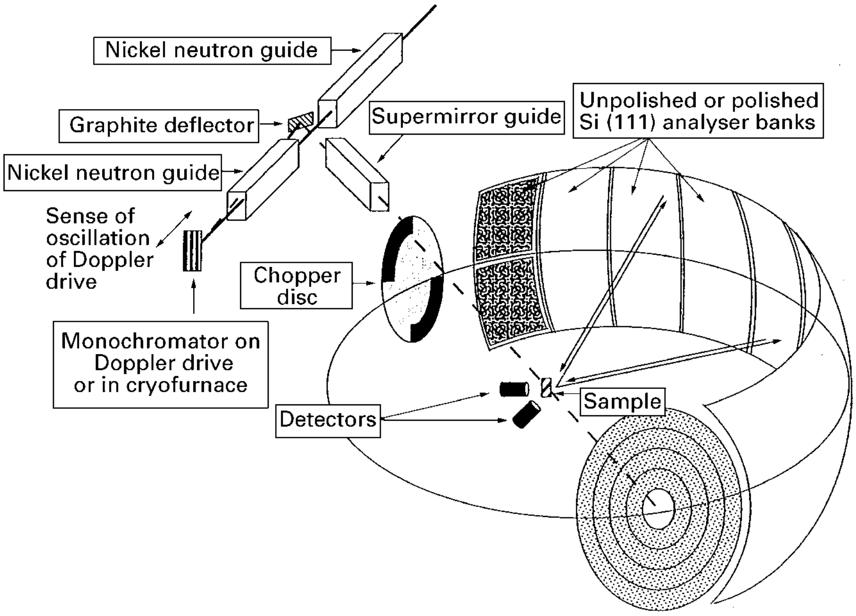



**Figure 5** Schematic view of the cold-neutron spectrometer IN10 at the ILL. The incident energy is changed either by Doppler-shifting the neutron wavelengths (IN10A) or by thermal expansion of the monochromator (IN10B). Reproduced with permission of Adenine Press from Ferrand M (1997) Neutron instrumentation in studying dynamics of biomolecules. In: Cusack S, Büttner H, Ferrand M, Langan P and Timmins P (eds). *Biological Macromolecular Dynamics*. Schenectady: Adenine Press.

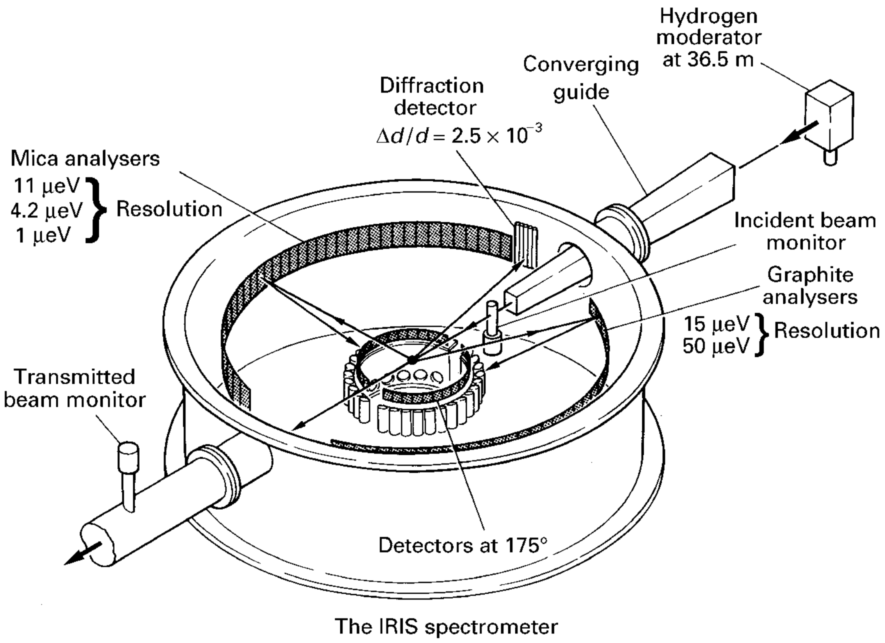



**Figure 6** Schematic of the cold-neutron backscattering spectrometer IRIS at ISIS. Reprinted from *Physica*, 182B, Carlile CJ and Adams MA, 431–440, 1992, with kind permission of Elsevier Science – NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

not undergo total reflection and pass through the walls of the guide. By making the guide gently curved, with a radius of curvature of the order of a kilometre, a guide can also provide a means of selecting the neutrons of interest.

On IN10 the energy scan is performed by varying the energy of the incident neutrons. One may do this scan, which typically covers ±15 μeV for 2 meV incident neutrons, by oscillating the monochromator to produce a Doppler shift in the velocity of the neutrons impinging on the sample. An alternative scanning procedure is to slowly vary the temperature and hence the lattice constant of the monochromator.

It can be seen in **Figure 6** that the analyser system on IRIS is similar to that of IN10 but that the incident white beam is generated differently. From the neutron pulse emitted by the liquid-hydrogen moderator, a broad band of neutron energies is selected by a rotating disc chopper at 6.4 m. The required energy resolution in the incident beam is achieved by allowing this band of neutrons to disperse in time over a long distance as it drifts along a curved neutron guide to the sample situated at 36 m from the moderator. Neutrons whose energies change to precisely match the analyser energy after scattering from the sample are selected by Bragg scattering from the pyrolytic graphite crystal array in near-backscattering geometry. A continuous range of scattering angles from 15° to 165° is covered simultaneously by the analyser-detector array. Using the 002 reflection from the graphite analyser an elastic resolution of 15 μeV is provided over an energy transfer range from +4 meV to - 1 meV.

An advantage of all indirect-geometry instruments at spallation sources is that a relatively broad range of neutron wavelengths is incident on the sample. This means that it is possible to add a diffraction capability simply by the addition of detectors that have a clear view of the sample; see **Figure 6**. This allows information on both structure and dynamics to be collected simultaneously. Inelastic scattering is much weaker than elastic scattering, so over the timescale of the experiment a few diffraction detectors will be sufficient to record a diffraction pattern of good quality.

## Neutron spin echo

The highest energy resolution is obtained by using the neutron spin echo technique. In this type of instrument, the velocities of neutrons incident on, and scattered from, a sample are coded as the number of Larmor precessions that the neutron spins undergo in a well-defined applied magnetic field. The method requires a beam of polarized neutrons with a distribution of velocities. As indicated in **Figure 7**, these neutrons enter the precession-field region with their spins perpendicular to the field. As a neutron

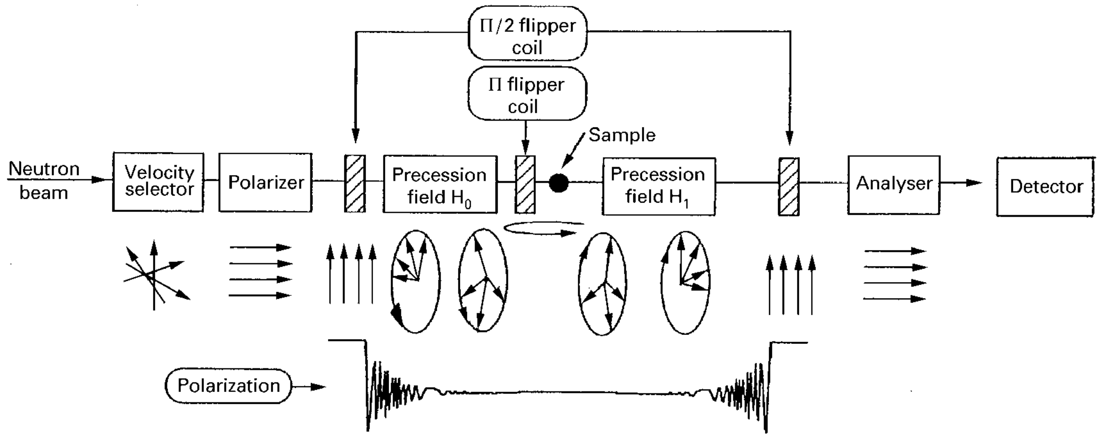



**Figure 7** Schematic diagram of a neutron spin echo spectrometer. The direction of the neutron spin is shown as it traverses the spectrometer. Reproduced with permission of Adenine Press from Ferrand M (1997) Neutron instrumentation in studying dynamics of biomolecules. In: Cusack S, Büttner H, Ferrand M, Langan P and Timmins P (eds) *Biological Macromolecular Dynamics*. Schenectady: Adenine Press.

traverses this field, it precesses and the number of precessions is directly proportional to the time it spends in the magnetic field and thus inversely proportional to its velocity. After scattering by the sample, the spins 'unwind' in a reverse-precession-field region. The method is unique in using the magnetic moment of the neutron to detect the velocity change rather than the wave (as in the backscattering spectrometer) or particle properties (as in the direct-geometry spectrometer).

If the scattering at the sample is elastic and the precession fields before and after the sample are identical in magnitude and spatial extent, all neutrons will be in phase at the point in space where they leave the second precession region. This alignment of spins is the 'echo'. However, if the scattering is not purely elastic, the spins will not be completely in phase, weakening the echo. The distribution of phases is then a measure of the distribution of the atomic velocities in the sample.

The high resolution of this technique stems from the large number of precessions the neutron undergoes, $\sim 10^5$ precessions $m^{-1}$. As a result, the velocity can be measured to an accuracy of better than $10^{-5}$ and hence the energy exchange can be measured to similar accuracy.

Another feature that is unique to neutron spin echo is that the signal, which is the neutron polarization at the echo point, is directly proportional to the intermediate scattering function, $I(Q, t)$, which is the Fourier transform of the scattering function $S(Q, \omega)$. This is in contrast to the other types of spectrometer that measure $S(Q, \omega)$. It is often advantageous to calculate $I(Q, t)$ rather than $S(Q, \omega)$.

## The energy range 0–1000 meV: vibrational spectroscopy and magnetic excitations

This energy range corresponds to the mid- and near-infrared regions of the electromagnetic spectrum. This is the province of vibrational spectroscopy where the forces between atoms in condensed matter can be investigated directly. It is also the region where magnetic excitations are observed. These are electronic transitions between the energy levels of a metal atom (or ion) and usually involve d or f orbitals. They are often optically forbidden and the transitions can be observed by INS because of their interaction with the magnetic moment of the neutron.

Neutrons in this energy range are known as thermal neutrons, because typical energies correspond to room temperature; see **Table 1**. Thermal-neutron INS spectroscopy was the first type to be studied and the instrument invented by Bertram Brockhouse in 1952, the triple-axis spectrometer, is still a mainstay of inelastic instrumentation at reactor sources. In 1994, Brockhouse, together with the pioneer of elastic neutron scattering (diffraction) Clifford Shull, received the Nobel prize for physics for his invention and subsequent work with it.

The triple-axis spectrometer is designed to study coherent excitations, which gives information about the relative motions of the nuclei. This approach mandates the use of single crystals and these need to be very large, of the order of several cubic centimetres. The scattering from powders and liquids is largely incoherent, which gives information about the

motions of the individual nuclei. For these types of materials different spectrometers are used. Spectrometers for coherent and incoherent INS spectroscopy will be considered separately, although this division is one of convenience only.

### Spectrometers for coherent INS spectroscopy

The principle of the triple-axis spectrometer is displayed in **Figure 8**. The prerequisite is a neutron source with a constant flux, which can be a steady reactor or a continuous spallation source. A monochromatic neutron beam is selected from the neutron source by using Bragg reflection from a single crystal, the monochromator. This beam is then incident on the sample. In a selected direction another single crystal, the analyser, is positioned in such a way that only a given wavelength can be reflected; a neutron detector is placed after this analyser. If neutrons are counted in the detector, this means that there has been some process in the sample which has induced the specific changes of trajectory and wavelength of the incident neutrons.

The important point is that a given configuration of the spectrometer corresponds to a single point in $(Q, \omega)$ space. Changing step-by-step either $k_i$ or $k_f$, the crystal is rotated and the scattering angle varied so as to scan the $(Q, \omega)$ space and detect all possible neutron-scattering processes, which can be vibrational (phonons) or magnetic (magnons) in origin. Furthermore, a given $(Q, \omega)$ point may be obtained in different ways (varying, for instance, either $k_i$ or $k_f$). This gives the experimentalist the possibility to adjust conditions to the specific requirements of a given problem.

The great strength of triple-axis spectroscopy is simultaneously its Achilles heel. The instrumentation is enormously flexible and, in principle, any point in $(Q, \omega)$ space is accessible. However, it is a point-by-point method and this means that it is slow and to map the phonon dispersion in one sample can require weeks of measurement time.

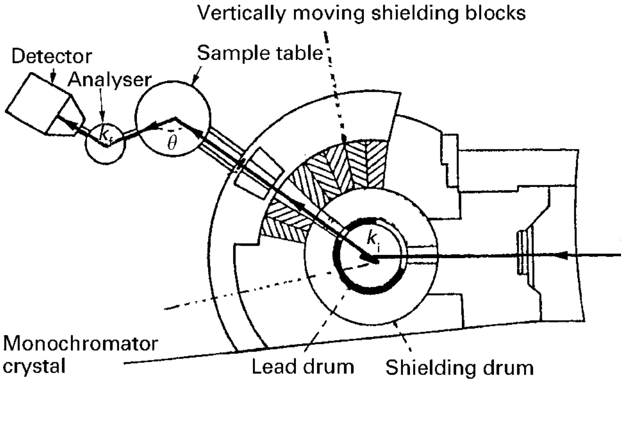



**Figure 8** General layout of the hot-neutron triple-axis spectrometer IN1 at the ILL.

The pulsed-source equivalent of a triple-axis spectrometer is demonstrated by PRISMA at ISIS. The difference is that a white beam of neutrons is incident on the sample, rather than a monochromated beam. As in the triple-axis spectrometer, Bragg reflection from analyser crystals is used to select the neutrons scattered by the sample that will ultimately be detected. However, PRISMA has an array of 16 independent analyser and detector arms instead of just one detector. Each arm of PRISMA measures along a parabolic path in $(Q, \omega)$ space, so generating a two-dimensional scan through $(Q, \omega)$ space for each ISIS pulse, in a single setting of the instrument and sample.

### Spectrometers for incoherent INS spectroscopy

The simplest form of instrument is a filter spectrometer which consists of a monochromator to define the incident neutron energy on the sample and a filter after the sample that only allows neutrons of a given energy to reach the detector. This type of instrument is used at a reactor and is exemplified by IN1BeF at the ILL. The monochromator is the same as for the triple-axis instrument IN1, see **Figure 8**, but the analyser is replaced by a cooled beryllium filter. Beryllium only transmits neutrons with energy of less than 5 meV; higher-energy neutrons are Bragg-scattered out of the beam. The transmission and the sharpness of the cut-off are improved by cooling the filter to below 100 K and it is routinely operated at liquid-nitrogen temperature. Since both the incident and final energies are known, the energy transfer is obtained from Equation [1]. The resolution of the instrument is determined by the bandpass of the filter and the detector response at low energies and by the monochromator at higher energies. Typically the resolution increases from ~3 to ~8% of the energy transfer between 20 and 300 meV. The resolution can be improved by using a graphite filter which has a narrower bandpass of ~1.5 meV but at the cost of a large decrease in detected flux.

At a spallation source, both indirect- and direct-geometry instruments are used. A schematic of the indirect-geometry spectrometer TOSCA at ISIS is shown in **Figure 9**. A small fraction of the neutrons from the incident white beam (which extends from 2 meV to beyond 2500 meV) are inelastically scattered by the sample; those that are backscattered through an angle of 135° impinge on a graphite crystal. Only one wavelength (and its harmonics) are Bragg-scattered by the crystal; the remainder pass through the

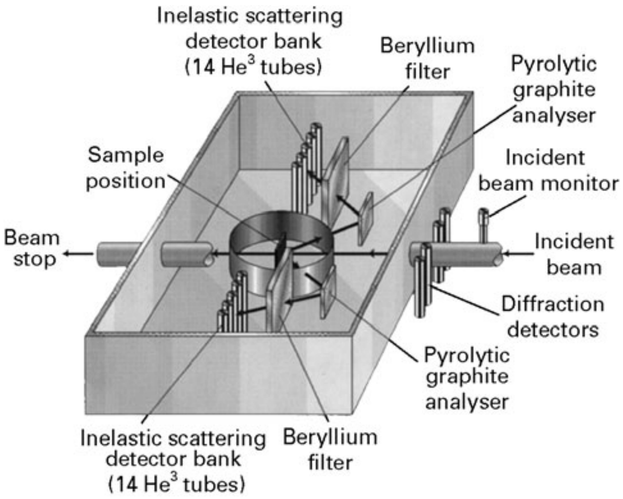



**Figure 9** Principle of operation of the indirect-geometry spectrometer TOSCA at ISIS. Only two of the 10 analyser/detector modules are shown.

graphite crystal to be absorbed by the shielding. The neutrons at multiples of the fundamental wavelength are removed by the beryllium filter, so only neutrons of ~4 meV reach the detectors. The net effect of the combination of the graphite crystal and beryllium filter is to act as a narrow bandpass filter.

The use of a small final energy has two important consequences. Since for most energies $k_i >> k_f$ it follows that both the magnitude and direction of $Q$ are largely determined by $k_i$. Thus $Q$ is almost parallel to the incoming neutron beam and perpendicular to the sample plane across virtually the entire energy transfer range. The significance of this is that in order to observe an INS transition, the vibration must have a component of motion parallel to $Q$. For a randomly oriented (i.e. polycrystalline) sample this condition will be satisfied for all the vibrations and all will be observed. This is not the case for an oriented sample and experiments directly analogous to optical polarization measurements are possible. Further, because $Q$ is independent of $k_f$ there is a unique value of $Q$ for each energy transfer and $Q^2(\text{Å}^{-2}) \approx E(\text{meV})/2$. Thus TOSCA (and IN1BeF since it also uses the same final energy) trace out a parabola in $(Q, \omega)$ space as shown in **Figure 3B**.

The resolution function varies with energy transfer but is typically 2% of the energy transfer between 2 and 500 meV. The advantages of this system are that it is mechanically simple (there are no moving parts) and it offers excellent resolution across a wide energy range with very low background. Further, the reflection geometry means that the spectrometer is relatively insensitive to multiple scattering (where the neutron undergoes more than one inelastic collision in the sample) so quite large samples (several grams) can be used to improve the count rate.

The pre-eminent direct-geometry spectrometer for this energy range is MARI at ISIS. **Figure 10** shows a schematic of the instrument. Two background-suppression choppers are used, the first is a Nimonic chopper which is used to suppress the prompt pulse of very high energy neutrons and gamma rays that are produced when the proton pulse hits the target. The second is a disc chopper made of borated resin with a single hole. This is designed to suppress the flux of delayed neutrons that are a significant fraction of the background when using a depleted-uranium target.

The core of the instrument is the Fermi chopper. This is a metal drum with a series of slots cut through it. It is magnetically suspended in a vacuum perpendicular to the beam and able to rotate at speeds up to 600 Hz. Thus for most of a rotation, the incoming neutrons are blocked, but at one particular time the slots are parallel to the incoming neutrons and they pass through to the sample. The incident neutron energy is selected by phasing the opening time of the slots with respect to the neutron pulse from the target station. Incident energies in the range 10 to 2000 meV can be selected. The detector bank continuously covers the angular range from 3° to 135° and so is able to map large regions of $(Q, \omega)$ space in a single measurement. The energy resolution

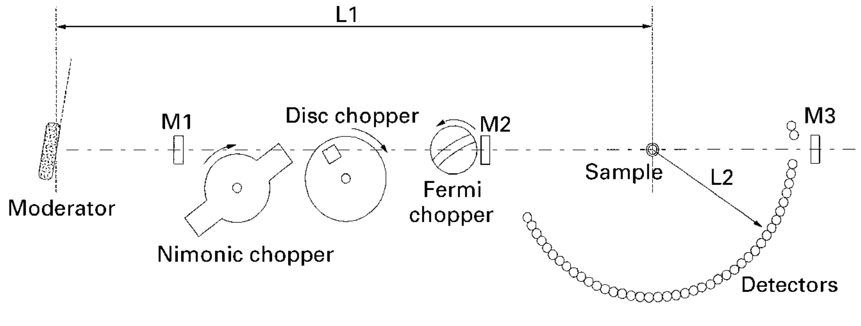



**Figure 10** Schematic diagram of the direct-geometry spectrometer MARI at ISIS. M1, M2 and M3 are incident beam monitors.

is between 1-2% of the incident energy, and with all the detectors at the same secondary flight path, this resolution is constant for all the detector banks. The large range in $Q$ is essential because phonons and magnons have different $Q$-dependencies and this is the best way to distinguish between the two sorts of scattering processes.

Three low-efficiency (because of the high neutron flux) detectors are placed in the main beam. The first is placed before the background choppers to monitor the incident flux for the purposes of normalization. The second and third are placed just after the Fermi chopper and behind the sample respectively. These are used to accurately determine the incident energy of the neutrons.

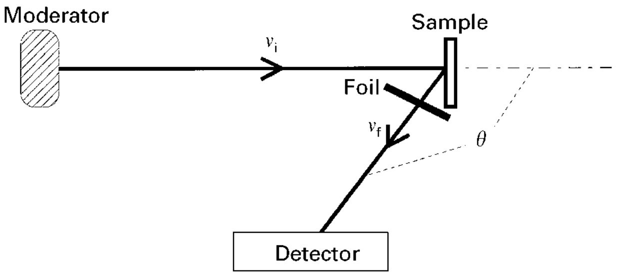



**Figure 11** Schematic diagram of the eVS spectrometer at ISIS. The incident neutrons have velocity $v_i$, that of the scattered neutrons is $v_f$ and the scattering angle is $\theta$. The moderator-to-sample distance is $L_i$ and the sample-to-detector distance is $L_f$. The final neutron energy $E_f$ is defined by the analyser foil, which can be either gold or uranium.

## The energy range >1000 meV: neutron Compton scattering

In this energy range only spallation sources can provide a sufficient flux of high-energy neutrons. The electron volt spectrometer (eVS) at ISIS is a neutron spectrometer, which uses the high intensity of neutrons in the electronvolt energy range to measure atomic-momentum distributions in a variety of condensed matter systems. When the momentum transfer in the scattering process is sufficiently large (at least 10 times that of the mean atomic momentum), the scattering can be interpreted within the impulse approximation. In this approximation, the scattering is essentially single-atom billiard-ball scattering, which is determined entirely by conservation of momentum and kinetic energy of the atom and the neutron. By measuring the energy and momentum change of the neutron, one can determine the momentum component of the target atom along the direction of the scattering vector. This process is neutron Compton scattering.

The instrument is shown schematically in **Figure 11**. eVS is an indirect-geometry inelastic spectrometer. Energy transfers in the 1000-30000 meV region and momentum transfers between 30 and 200 Å$^{-1}$ are achieved using a filter-difference technique. The filter is a thin foil which has a strong nuclear resonance that absorbs neutrons over a narrow band of energy centered at $E_1$. Two time-of-flight spectra are taken, one with the foil between the sample and detectors and the second with the foil removed. The analyser foils are moved in and out of the scattered beam by computer-controlled pneumatic pistons. One complete cycle is performed every 10 minutes, i.e. 5 min with foil in and 5 min with foil out. This is to avoid effects due to long-term changes in the relative efficiency of the detectors, since typical measurement times are 2-12 hr.

The difference between these spectra is due to those neutrons absorbed in the foil and is effectively the spectrum for neutrons scattered with energy $E_1$. If $E_1$ is known, the momentum and energy transfer can be determined from a measurement of the total time-of-flight $t$ of the neutron from moderator to detector. The velocity $v_1$ of the scattered neutrons can be calculated from the kinetic energy. The velocity $v_i$ of the incident neutron is determined from a measurement of the neutron time-of-flight $t$:

$$t = \frac{L_i}{v_i} + \frac{L_f}{v_f} + t_0 \qquad [9]$$

where $L_i$ is the distance from source to sample and $L_f$ that from sample to detector. $t_0$ is an electronic time delay constant. The energy transfer is then given by Equation [6] and the momentum transfer is:

$$Q = m\left(v_i^2 + v_f^2 + 2v_i v_f \cos\theta\right)^{0.5} \qquad [10]$$

where $\theta$ is the scattering angle.

## Future developments

In terms of instrumentation, developments in guide technology will lead to modest gains (factors of two to five) in the flux at the sample. The prevent trend is towards increasing the area covered by the detectors with corresponding gains in sensitivity and data rate.

INS is flux-limited technique. Thus major instrumental advances can only be expected when more-intense neutron sources become available. These will

be spallation sources, since reactor sources, such as the ILL, are already close to the limit of the heat load generated by the fission process in the reactor core that can be handled. In addition, concerns about nuclear-weapon proliferation (enriched uranium is required for the reactor core) mean that these are politically and environmentally contentious. In the near future, it is planned to increase the ISIS proton beam current by 50% with a corresponding flux increase. For the year 2005, the USA is proposing to build a spallation source with several times the ISIS flux. In the longer term, the most powerful source will be the European Spallation Source, which will have a time-averaged flux equal to that of the ILL and a peak intensity 30 times that of ISIS.

## List of symbols

$d$ = analyser crystal interplanar spacing; $E$ = neutron kinetic energy; $h$ = Planck's constant; INS = inelastic neutron scattering; $I$ = intermediate scattering function; $k$ = wave vector; $k_B$ = Boltzmann's constant; $L$ = sample–source/detector distance; $m$ = neutron mass; $Q$ = neutron momentum transfer; $S$ = scattering function; $T$, $t$ = time of flight; $\hbar = h/2\pi$; $\lambda$ = wavelength; $v$ = neutron velocity; $\theta$ = scattering angle; $\tau$ = timescale; $\omega$ = frequency.

*See also:* **Inelastic Neutron Scattering, Applications; Neutron Diffraction, Theory; Structure Refinement (Solid State Diffraction).**

## Further reading

The most up-to-date and useful source of information on neutron scattering centres and their instrumentation is the World Wide Web. The site:
http://www.neutron.anl.gov/neutronf.htm
provides an extensive list of websites.
The ISIS website is at: http://www.isis.rl.ac.uk
and the ILL website is at: http://www.ill.fr/

Bée M (1988) *Quasielastic Neutron Scattering*. Bristol: Adam Hilger.

Eckert J and Kearley GJ eds (1992) Spectroscopic applications of inelastic neutron scattering: theory and practice. *Spectrochimica Acta* **48A**: 269–478.

Newport RJ, Rainford BD and Cywinski R eds (1988) *Neutron Scattering at a Pulsed Source*. Bristol: Adam Hilger.

Nicholson LK (1981) The neutron spin-echo spectrometer: a new high resolution technique in neutron scattering. *Contemporary Physics* **22**: 451–475.

Squires GL (1978) *Introduction to the Theory of Thermal Neutron Scattering*. Mineola: Dover Publications.

Windsor CG (1981) *Pulsed Neutron Scattering*. London: Taylor and Francis.

# Inorganic Chemistry, Applications of Mass Spectrometry

**Lev N Sidorov**, Moscow State University, Russia

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Applications |

As a new research field, high-temperature mass spectrometry was established in the early 1950s. Of special value was the combined use of the classic Knudsen effusion method and mass spectral study of the evaporation products of inorganic compounds (the KCMS technique). The first review appeared as early as 1959 and contained data for approximately 120 systems and compounds which had been studied by that time.

Currently, high temperature mass spectrometry is one of the most powerful methods in high temperature chemistry. A particular feature of this method is a high temperature molecular beam source, namely a Knudsen (or effusion) cell. A Knudsen cell is an isothermal enclosure with a small orifice of precisely known dimensions. Vapour molecules escape through the orifice at a rate assumed to equal the equilibrium rate of collisions with the wall of the cell. Total vapour pressure inside the Knudsen cell has to be less than 10 Pa to ensure molecular flow conditions. The effusion cell with the substance under investigation is placed inside the mass spectrometer near an ionization chamber. By increasing the temperature of the Knudsen cell, the equilibrium pressure of the sample also increases, and usually a measurable signal can be detected from a pressure of $10^{-7}$ Pa inside the cell. Thus, the working range of the sample pressures in high temperature mass spectrometry is $10–10^{-7}$ Pa. After passing collimating slits, the molecular beam goes through the

ionization chamber where electron impact ionization takes place. As usual in mass spectrometry, the electron energy is approximately 40–60 eV and can be changed to measure appearance energies.

Measurements of the ion current intensities provides the partial pressures of individual species inside the Knudsen cell and equilibrium constants of gas-phase and heterogeneous chemical reactions. Thermal ionization of samples inside the Knudsen cell also takes place at high temperatures. These thermal ions are in equilibrium with the neutral species and can be drawn out from the cell by a weak electrostatic field and their partial pressures measured. This provides the equilibrium constants of gas-phase ion molecule reactions. As a result, the following thermodynamic data can be obtained: molecular and ion composition of the vapour; enthalpies and Gibbs free energy of gas-phase and heterogeneous chemical reactions; enthalpies of formation, Gibbs free energies, electron affinities and ionization energies for vapour species (molecules, positive and negative ions); activities of the components in mixtures, partial and integral Gibbs free energies of mixing and Gibbs free energies of formation of solid compounds in the mixture.

## Knudsen cell–mass spectrometer equipment

### Mass analysers

In principle, any type of mass analyser can be coupled with the Knudsen cell system. As a rule, a mass range up to 1000 a.m.u. and resolution near 800 are sufficient for refractory material investigations. A single focusing magnetic sector field is used in most cases for mass analysis of effusion beams. Quadrupole and time-of-flight mass spectrometers are also applied. More expensive double focusing instruments provide high resolution and help to avoid interfering background ion intensities arising from ionization of organic impurities.

### Knudsen cells, heating and temperature measurement systems

Knudsen cells are mostly heated by resistance heaters or by electron bombardment. In the first case, the limiting working temperature of the cell does not exceed 1900 K and can be measured by thermocouples. Platinum–platinum/rhodium is applied up to 1300 K and tungsten–rhenium for higher temperatures. When electron bombardment is used the upper temperature is limited by melting and evaporation of the cell material. Optical pyrometers of various types are the instruments of choice for temperature measurements in the range 2200–3000 K. Various types of Knudsen cell are used depending on the particular aim of the investigation. Stainless steel, quartz, nickel, graphite, tungsten or molybdenum may be applied as the cell materials. Protecting layers on the inner walls of a cell are also used to avoid chemical interactions between the sample and the cell material.

Some of the cells are shown in **Figure 1**. Two-compartment Knudsen cells (**Figure 1D**) are used if the vapour species formed from two different compounds with very different vapour pressures have to be investigated. They can also be used to get unsaturated vapours. The second compartment is replaced by a gas inlet system if gases are involved in the chemical reactions to be studied. ‘Russian doll’ type and a twin cell with a connecting channel (**Figures 1B** and **1C**) allow generation of saturated and unsaturated vapours of the sample.

Multiple Knudsen cells (**Figures 1E** and **1F**) are used for the determination of chemical activities. One cell contains a reference compound, whereas the others contain mixtures having different compositions. A comparison of the vapour pressures in the different cells gives the activities of the components and differences between the sublimation enthalpies of the samples.

Local heating of the sample to higher temperatures can be provided by a laser. Containerless heating and free evaporation take place under such conditions. Local heating is widely used in cluster formation studies.

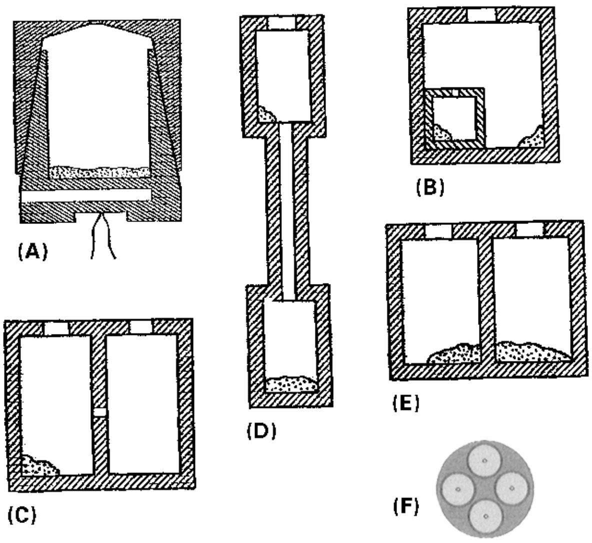



**Figure 1** (A) Knudsen cell; (B) Russian doll; (C) twin cell with connecting channel; (D) two compartment cells; (E) twin cell; (F) four Knudsen cell units (multiple cells).

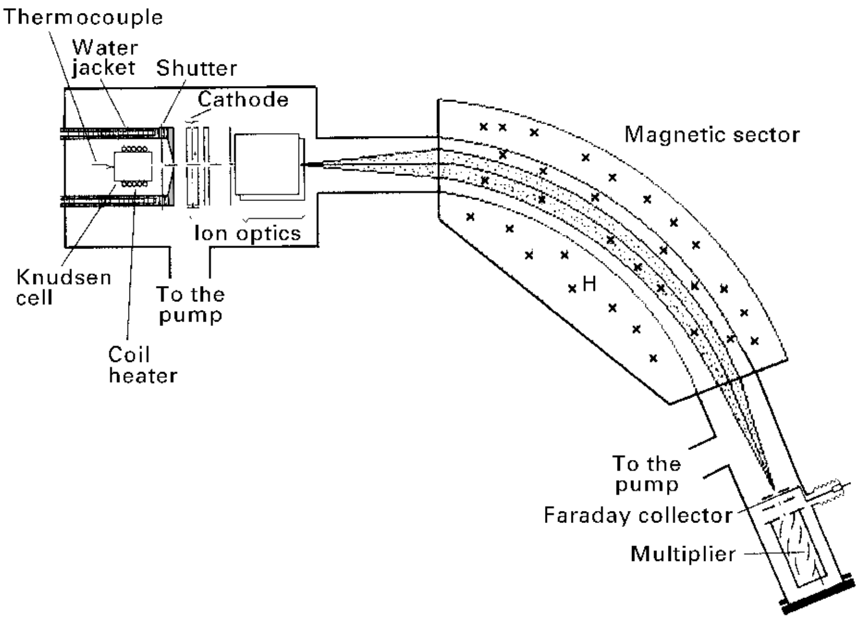



**Figure 2** Knudsen cell-sector magnetic mass spectrometer with coaxial ion and molecular beams.

## Ion sources

Conventional ion sources with electron impact ionization are utilized. Photo-ionization and different types of electron energy selectors are applied to improve the accuracy of ion appearance energy measurements. The electron beam is always perpendicular to the effusion beam. The ion beam can be coaxial with the effusion direction or perpendicular to both the electron and the effusion directions. Both types of KSMC instruments are shown in **Figures 2** and **3**.

Thermal ionization of samples inside the Knudsen cell takes place at high temperatures and the Knudsen cell can itself be a source of thermal ions. The charge and quantity of escaping ions depend on the draw-out potential in front of the effusion hole. A combined ion source can operate in two modes, to register either neutral or charged particles escaping from the cell.

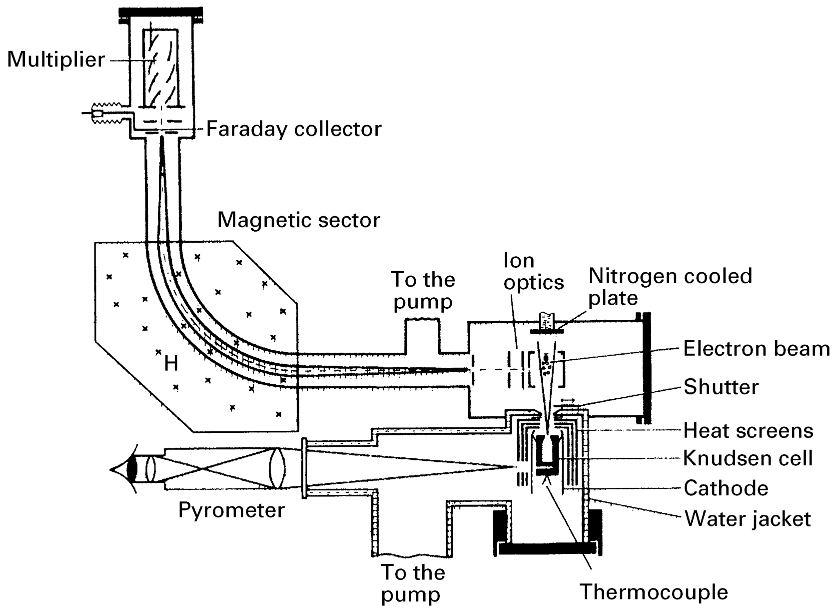



**Figure 3** Knudsen cell-sector magnetic mass spectrometer with perpendicular ion and molecular beams.

At an ion pressure of less than $10^{-8}$ Pa, the Debye shielding radius becomes larger than the linear dimension of the cell and the concentrations of positive and negative ions in the cell are not equal. Changing the electron work function of the inner surface of the cell can increase the concentration of the ions of interest.

Deflection plates are placed behind the exit slit of the ion source. The vector of the electrostatic field is coaxial with the sector magnetic field vector. The plates permit measurement of the ion kinetic energy distribution in the direction coinciding with the vector of the magnetic analyser field strength.

### Background and legitimate signal

A problem which is equally important for both neutral and thermal ion registration is separation of the legitimate signal from the background. In neutral species studies, the legitimate signal is separated from the background by use of a movable shutter impeding the passage of the molecular beam to the ionization cell. In thermal ion studies, the signal is separated and located by means of deflecting plates. Because the electrostatic draw-out field has an axial symmetry and the axis of symmetry passes through the centres of the effusion hole and of the draw-out lens, the ions generated away from the axis (in the space between the cell and the lens) have an additional velocity in the direction normal to the axis of symmetry. This additional velocity component can be readily fixed by using the deflection plates. The use of these plates allows one to discriminate between the ions drawn out from the effusion hole (the legitimate signal) and those appearing on the outer surface of the effusion cell, protective screens, etc. (the background). Moreover, the method makes it possible to monitor the surface and to locate the place of sample evaporation (**Figures 4** and **5**).

Due to the high temperature in the surroundings of the Knudsen cell some additional units are installed. There is a water jacket around the heating assembly and a target cooled by liquid nitrogen to condense the effusion beam escaping from the Knudsen cell.

A valve between the Knudsen cell compartment and the ion source is useful and makes possible a fast change of sample in the cell without changing the parameters of the ion source. The Knudsen cell and heating assembly are fixed on a moving plate and can be shifted relative to the ionization chamber entrance slit. This allows one to find the appropriate positions of the electron and effusion beams which give the highest ionization efficiency.

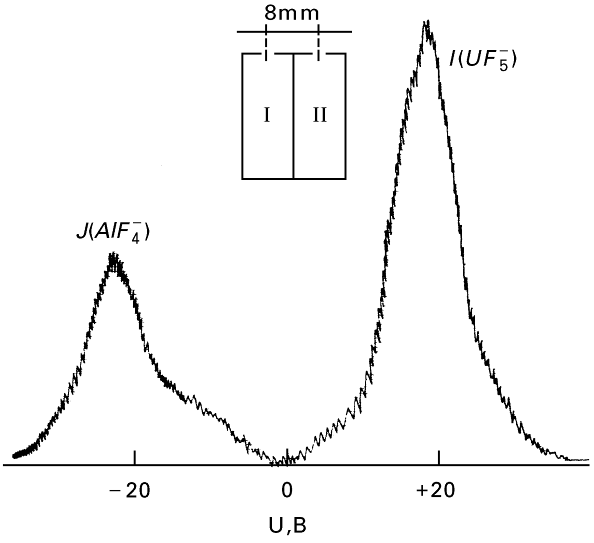



**Figure 4** Ion current intensity as a function of the deflecting plates potential. Twin cell with NaF–$AlF_3$ in one section and KF–$UF_4$ in the other. The distance between effusion holes is 8 mm. Distance resolution is 5.2 V $mm^{-1}$.

## Ionization efficiency curves and cross-sections

Ionization efficiency curves are plotted in the electron energy range 5–50 eV and ion appearance energies are determined in most cases. The threshold law is written as $I = \text{const}(E–E_{\text{th}})^{n-1}$ where $n$ is the number of electrons leaving the collision area. In the case of the formation of singly charged positive ions,

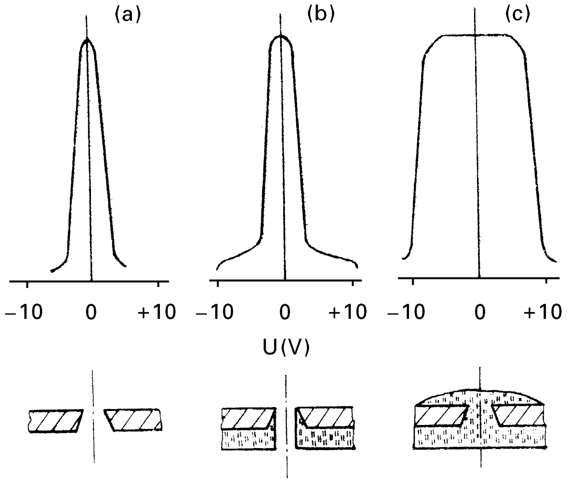



**Figure 5** Ion current intensity as a function of the deflecting plates potential. (A) clean effusion hole; (B) condensation on the inner surface around the effusion hole; (C) condensation on the outer surface around effusion hole (creeping).

the linear extrapolation of the linear part of the ionization efficiency curves is widely used to find the ion appearance energy. The ionization cross-sections of atoms can be measured for different ionization energies between 0 and 220 V in steps of 1 eV and their maximal values are located in the range 40–60 eV. These data are used in Knudsen effusion mass spectrometry; the calculated maximum values are considered to be in agreement within 20% with the experimental data if the ionization cross-sections are compared at the same energy.

Ionization cross-sections of molecules can be obtained from those of atoms by applying the additivity rule postulated by Otvos and Stevenson in 1956. The ionization cross-section calculated in such a way is a total for the electron energy in the range 40–60 eV. Fragmentation always takes place at such electron energies, and the mass spectrum of a molecule gives the ratios of partial cross-sections with respect to the base peak. The total ionization cross-section is the sum of the partial ones. It is generally accepted that the above additivity rule gives a value which differs from the experimental one by not more than a factor of 2. This rough estimate is only correct for the total ionization cross-section. There are two limiting cases in Knudsen cell mass spectrometry: the first is when the molecular ion peak is the base peak and its partial ionization cross-section with respect to the total is more than 80%. In the second case, the molecular ion peak is small and its partial ionization cross-section is less than 10% of the total. Treatment of the experimental data and the experimental approaches are essentially different in these cases. In the first case, measurements of the molecular ion intensities are sufficient for the reliable interpretation of the experimental data. In the second case, one has to know the fragmentation pattern (mass spectrum) of each vapour species and calculate the total ion current intensities for each vapour species. This task is rather complicated because a particular feature of high temperature vapour is the existence of vapour species which cannot be isolated as individual compounds and whose mass spectra cannot be measured in the previous runs. The superposition of the same fragments formed from different species takes place and ion current intensities have to be apportioned between each species.

The total ionization cross-section is a function of the electronic states of a molecule. It does not significantly depend on temperature in the one- to two-hundreds degree temperature range. It is not a case of partial cross-sections which depend on vibrational states. The temperature dependence of the mass spectrum of the individual molecules is well known and has to be taken into account before the experimental data are interpreted.

## Relationships between ion intensities and partial pressures of vapour species

The relation between the measured ion current intensities and partial pressures of the neutral species inside the effusion cell can be written as follows:

$$p_j = \frac{k}{\sigma_{ij} I_{ij} T} \qquad [1]$$

$$p_j = \frac{k_i}{\sigma_j \sum I_{ij} T} = \frac{k}{\sigma_j I_j T} \qquad [2]$$

where $p_j$ is the partial pressure of molecule $j$, $k$ is the sensitivity constant, $\sigma_j$ is the total ionization cross-section of molecule $j$, $\sigma_{ij}$ is the partial ionization cross-section of molecule $j$, $I_{ij}$ is the measured ion current of ion $i$ originating from molecule $j$ ($I$ is calculated allowing for the isotopic composition of the elements and the secondary multiplier gain, $I_j = \sum_i I_{ij}$ is the total ion current from molecule $j$, $T$ is the temperature of the effusion cell. While the thermal ions escaping the effusion cell are recorded the corresponding equation relating the measured ion currents and ion pressures inside the effusion cell is given as:

$$p_i = C\, I_i\, M_i^{1/2}\, T^{1/2} \qquad [3]$$

where $P_i$ is the partial pressure of ion $i$, $C$ is the sensitivity constant and $M_i$ is the mass of ion $i$.

## Mass spectra and gaseous species

The chemical formula of the ions in the mass spectrum is obtained from the mass of the ions and their isotopic distribution. The next step is the assignment of the observed ions to their neutral precursors in the equilibrium vapour and the calculation of the vapour species, partial pressures. Equations [1] to [3] reveal a set of tasks which have to be completed if calculated partial pressure values are to be reliable:

determination of sensitivity, $k$ and $C$ in Equations [1], [2] and [3] (calibration of an instrument);

determination of molecular weights of vapour species, $j$;

apportioning the measured ion currents $I_i = \sum_i I_{ij}$, i.e. determination of values $I_{ij}$;
calculation of total $\sigma_j$ or partial $\sigma_{ij}$ ionization cross-sections.

The last three points relate to neutral species only.

### Molecular weight of neutral species

Molecular ions and ion fragments have to be identified. The main approach is based on efficiency curve measurements and appearance potential estimates. For example, consider bismuth oxide. The ionization energy (IE) of Bi is 7.3 eV, IE of $Bi_2O_3$ is 9.5 and IE of $Bi_4O_6$ is 9.0 eV. The expected IE of molecular ions with atomic composition shifted to bismuth have to be in the range 7.3–9.0. There are 6 ions in the mass spectrum of saturated bismuth oxide vapour of $Bi_4O_5$, $Bi_4O_4$, $Bi_4O_2$, $Bi_3O_4$, $Bi_2O_2$ and $Bi_2O$. The first three ions have IE 9.1, 10.0 and 10.3 eV and are treated as fragments, the next have IE 7.5, 7.9 and 7.5 eV and are considered to be molecular ions.

### Apportion of ion currents

When molecular ions give the base peaks in mass spectra, one can decrease the electron energy and obtain a mass spectrum which mainly consists of molecular ions. The estimated total ionization cross-section can be utilized and the partial pressures of vapour species calculated. When molecular ions give only a minor peak in the mass spectrum, decreasing the electron energy does not give satisfactory results and therefore some thermodynamic approaches have been developed. These are based on the variation of ion current intensities with the isothermal change of vapour composition.

The isothermal change of vapour composition is governed by the thermodynamic laws, and their analytical expressions can be included in the equation that relates measured and apportioned ion currents. A variation of partial pressures at constant temperature can be achieved by changing the composition of the solid (liquid) phase equilibrated with the vapour. When an individual compound is under study, Knudsen cells of **Figures 1B**, **C** and **D** allow one to compare the saturated and unsaturated vapours at the same temperature and apportion the measured ion current.

## Instrument calibration

For thermal ions, the ratio of ion partial pressures is measured without calculation of absolute values. In the case of neutral species, the ratio and absolute values of partial pressures have to be determined in most cases. Two methods of calibration can be used: (1) the sensitivity $k$ is determined by comparison with a reference compound having a known vapour pressure. The ratio of the ionization cross-sections of a sample and a reference compound must be calculated according to the additivity rule and theoretical values of atomic cross-sections; (2) The sensitivity $k_j$ is determined directly by complete evaporation of a weighed amount of a sample, and if complete evaporation of two samples is carried out the ratio of their ionization cross-sections can be determined. In the first case

$$p_j = \frac{k}{\sigma_j} I_j T \qquad [4]$$

$$p_{\mathrm{ref}} = \frac{k}{\sigma_{\mathrm{ref}}} I_{\mathrm{ref}} T \qquad [5]$$

and

$$p_j = p_{\mathrm{ref}} \frac{\sigma_{\mathrm{ref}}}{\sigma_j} \frac{I_j}{I_{\mathrm{ref}}} \qquad [6]$$

where $p_{\mathrm{ref}}$ and $p_j$ are the pressures of the reference and investigated samples respectively.

In the second case, the Herz–Knudsen equation is utilized:

$$q[\mathrm{mol}] = s_{\mathrm{eff}}(2\pi M_j R)^{-1/2} \int_0^t p_j T^{-1/2}\,\mathrm{d}t = B_j \int_0^t p_j T^{-1/2}\mathrm{d}t \qquad [8]$$

where $s_{\mathrm{eff}}$ is the orifice area, $q$ = weight of sample; $M_j$ the mass of vapour specimen and $R$ is the gas constant. Taking into account Equation [2] it can be written as follows:

$$q[\mathrm{mol}] = B_j k_j \int_0^t I_j T^{1/2}\,\mathrm{d}t \qquad [9]$$

Measurements of temperature, ion current intensities as a function of time, weight of sample and the orifice area make it possible to calculate $k_j$ and $P_j$ at any time during evaporation. If there are several species in the vapour, Equation [9] has to be written as

$$q = \sum_j B_j k_j \int_0^t I_j T^{1/2}\,\mathrm{d}t \qquad [10]$$

### Dynamic range

The upper limit of vapour pressure in the cell is determined by the molecular flow through the effusion orifice. When the orifice diameter is 0.3–1.0 mm, the upper limit lies in the pressure range 1–10 Pa. A vapour pressure equal to $10^{-7}$ Pa is accepted as the lower limit for neutral species. In the case of thermal ions, the sensitivity of the instrument is 5–6 orders of magnitude higher, and the data fall in the range $10^{-6}$–$10^{-12}$ Pa. On the other hand, the concentration of ions in the vapour is 6–7 orders of magnitude lower than that of the neutral species, and as a result the ion currents from neutral species and thermal ions are of the same order.

## Second and third law calculations

When partial pressures are measured, the equilibrium constants of the gas phase and heterogeneous reactions can be calculated and their temperature dependence plotted.

The equation $d/dT(\ln K_p) = \Delta H^o/RT^2$ allows the determination of reaction enthalpies from the slope of the curve $\ln K_p$ versus $1/T$ (second law calculation). As $K_p = \prod_i P_j^{\nu_j} = \text{const} \prod_i I_j^{\nu_j} T^{\Delta\nu_j} = \text{const}\ K_p(I)$ one can determine the enthalpy change from the slope of the curve $\ln K_p(I)$ versus $1/T$.

The equation $\Delta H_o^o = -RT \ln K_p + \Delta\Phi^\circ T$ allows one to calculate reaction enthalpies from the absolute value of $K_p$ (third law calculation), where $\Phi_T^o = -(G_T^o - H_o^o)/T$.

## Determination of activities

According to the definition of activity, $a_j$ is the ratio $p_j/p_j^0$, and putting ion current intensities $I_j$ instead of partial pressures one has $a_j = I_j/I_j^0$. Even in the case of a binary mixture, not less than 10 samples of different composition have to be tested to obtain reliable data for the activities of both components in the complete range of composition. Two approaches have been developed to increase the efficiency of this type of measurement.

One of them is the multiple Knudsen cell which allows one to obtain activities for several mixture compositions during a run. Another is the isothermal evaporation method (IEM) which gives activities as a continuous function of the melt composition. The isothermal evaporation process should be carried out rather slowly (1–3% per hour) and can be considered a sequence of equilibrium states under such conditions. Simultaneously, ion current intensities are recorded as a function of time. Experimental data taken in such a way give ion current intensities as a function of time, integrals $\int_0^t I_{ij} T^{1/2} dt$ and derivatives of two types $dI_i(t)/dt$ and $dI_i(t)/dI_u(t)$. These data can be transposed into partial pressure and activity depending on the melt composition. The mathematical part of IEM which makes this transformation possible is based on the Herz–Knudsen and Gibbs–Duhem equations.

There are many binary systems whose vapour contains mixed associates and a large number of vapour species. The virtue of the KSMC method is the possibility of directly determining the partial pressures of vapour species, rather than the total pressure in the system. These advantages of mass spectrometry are widely used in thermodynamic research into two-component systems. If there are mixed species $A_kB_l$ and $A_rB_s$ in the saturated vapour of system A–B, the Gibbs–Duhem equation $\Sigma_j N_j d \ln a_j = 0$ can be written as

$$\frac{d\ln P(A_rB_s)}{d\ln P(A_kB_l)} = \frac{rN_B - sN_A}{kN_B - lN_A}$$

where $N_A$ ($N_B$) is the mole fraction of A(B) in the solution. Partial pressures can be replaced by ion currents and any pair of vapour species can be treated in accordance with the Gibbs–Duhem equation. In the simplest cases ($r = 1$, $s = 0$, $k = 0$, $l = 1$) it will be

$$\frac{d\ln p_A}{d\ln p_B} = -\frac{N_B}{N_A} = \frac{d\ln I_A}{d\ln I_B}$$

and for $r = 1$, $s = -1$, $k = 0$, $l = 1$

$$N_A d\ln\frac{p_A}{p_B} + d\ln p_B = 0 = N_A d\ln\frac{I_A}{I_B} + d\ln I_B$$

Different forms of Gibbs–Duhem equation are used and the choice of independent variables is determined by the vapour species being easily registered. The replacement of low intensities by the ratio of measured ones is a general approach in KCMS. Any type of equilibrium involving molecules $j$ can be used for this purpose. For example, consider the reactions; $A + B \leftrightharpoons AB$

$$K_p = \frac{p_{AB}}{p_A\, p_B}$$

and $p_A$ may be replaced by the ratio $p_{AB}/p_B$.

## Applications

### Vapour species and their thermodynamic properties

During the last four decades all major classes of inorganic compounds have been studied by KSMC. The discovery of new gas phase molecules and ions is considered a main achievement of KCMS. Thermodynamic properties of the newly discovered vapour species were also obtained and most of them can be found in reference books. Metal vapours consist of atoms and small amounts of dimers. The main vapour species of carbon is a trimer and a series of $C_n$ ($n = 1$–8) molecules were found in the saturated vapour. The main components of P, As and Sb vapours are tetramers. In the vapour of S and Se, a series of polymers with $n = 1$–12 were observed with the most abundant species being $S_8$ and $Se_6$. Heteroatomic molecules are always present in the vapour of mixtures of elements. The saturated vapour of inorganic compounds that evaporate in the temperature range 700–3000 K, generally contains species formed by the dissociation of the base compound and a number of associates. Many interesting and unexpected examples can be found in the extensive database on the species in inorganic vapours. The existence of the dimers and trimers of oxides $V_2O_5$, $WO_3$ and $MoO_3$ has been reported. Salts of oxygen-containing acids can be evaporated without decomposition to give small amounts of dimers. Mixtures of halides of alkali metals and halides of polyvalent metals give associates in the vapour and the electrical properties of such associates have been investigated.

### Heterogeneous reactions and gas phase reactions

Gas phase reactions of any type can be successfully studied. The magnitude of the equilibrium constant is important. *Isomolecular* reactions $A + B \leftrightharpoons C + D$ do not require an absolute instrument calibration which makes them the most convenient and frequently observed type of reaction. This also concerns ion–molecule and electron exchange reactions. A bond dissociation energy is calculated from the equilibrium constants of the reaction $MX_n \leftrightharpoons MX_{n-1} + X$. The sublimation enthalpy and dissociation enthalpy are determined from the equilibrium constants of the heterogeneous reactions $A_{solid} \leftrightharpoons A_{gas}$ and $AB_{solid} \leftrightharpoons B_{solid} + A_{gas}$.

### Ion/molecule equilibrium in the saturated vapour

A study of the complex halogenides and salts of oxygen-containing acids by KCMS modified for ion/molecule equilibrium study (KCMS-IME) showed a very high concentration of ions in the saturated vapour (only 5 orders of magnitude less than the vapour pressure). These salts of alkali metals have large anions and as a result the enthalpies of heterolytic dissociation of the complex salts are much lower than those of alkali metal fluorides. For example, for the system KF–33 mol% $UF_4$, $T = 1000$ K, $p(KUF_5) = 0.35$ Pa; $p(K^+) = p(UF_5) = 10^{-6}$ Pa. At such ion pressures the Debye shielding radius $R_d$ (0.2 mm) is significantly smaller than the linear dimension of the Knudsen cell (10 mm). An electronless plasma is formed and the main vapour species are:

Molecules KF, $K_2F_2$, $KUF_5$, $UF_4$
Positive ions $K^+$, $K_2F^+$, $K_2UF_5{}^+$
Negative ions $KF_2^-$, $UF_5^-$, $UF_6^-$, $U_2F_9^-$

### Electron affinity determination

Addition of complex alkali metal salts to a high temperature system leads to negative ion formation in the vapour. KCMS-IME can be applied to electron affinities and the estimation of negative ion enthalpies of formation. The EA of fluorides and oxides of transition metals in their higher oxidation states were determined. Almost all of these compounds have EA in the range 3.5–7.0 eV. Application of this method to fullerene vapours has yielded the electron affinities for the series of higher fullerenes $C_n$ ($n = 70, 72, ..., 106$).

### Activities in mixtures and the Gibbs free energy of formation of solid compounds

The determination of activities of components allows calculation of the partial and integral Gibbs free energy for the mixing of liquid and solid solutions. If solid compounds $A_nB_m$ are formed in the mixture A–B, the Gibbs free energy of formation of these compounds from A and B may be calculated.

### P–T–X diagram study

Measurements of the dependence of partial pressures on temperature along the liquidus (two-phase equilibrium between liquid and solid phases) and solidus (solid–solid) lines and partial pressure dependences on system composition at constant temperature give the total pressure and vapour brutto composition as function of temperature and system composition. After these data are obtained, a T–X diagram can be transposed to a P–T–X diagram. Such diagrams are used in applied science e.g. the growth of crystals and films by the vapour deposition method.

### Homogeneous range, Partial pressures of oxygen

KCMS is often used for the study of the homogeneous range of solid compounds. The partial pressure change can be easily selected. Isothermal evaporation is sometimes used to identify solid phase composition changes. Diffusion in the solid phase and enrichment of a surface layer by a hard volatile component demands special attention. When oxides, glasses or ceramics are studied the oxygen partial pressure is of particular interest. One advance in the last decade is that very low partial pressures of oxygen can be calculated and controlled by measurement of the ratio of two negative ions (KCMS-IME). Oxides of variable valency metals are the source of these two ions and the equilibrium constant of the reactions $MeO_n^- + \frac{1}{2}O_2 \leftrightharpoons MeO_{(n+1)}^-$ have to be previously obtained. The same approach is used for activity determinations and is based on similar reactions $A + B^- \leftrightharpoons AB^-$ and the measurements of negative ion ratios $I(AB^-)$ to $I(B^-)$.

## List of symbols

$A_kB_l$ and $A_rB_s$ are vapour species in the binary system A–B; $B_j = \rho_{eff}(2\pi M_jR)^{-1/2}$; $E_{th}$ = threshold energy; $I_{ij}$ = measured ion current of ion $i$ originating from molecule $j$; $I_i = \sum_j I_{ij}$ = measured current of ion $i$; $I_j = \sum_i I_{ij}$ = total ion current from molecule $j$; $M_i$ = mass of ion $i$; $M_j$ = mass of vapour specimen $j$; $N_A(N_B)$ = mole fraction of A(B); $p_i$ = partial pressure of ion $j$; $P_i$ = partial pressure of ion $i$; $p_j$ = partial pressure of molecule $j$; $p_{ref}$, $p_j$ = pressure of reference and investigated samples; $q$ = weight of sample; $s_{eff}$ = orifice area; $\sigma_j$ = total ionization cross-section of molecule $j$; $\sigma_{ij}$ = partial ionization cross-section of molecule $j$; $\Phi_T^o = -(G_T^o - H_o^o)/T$; $\nu$ = stoichiometric coefficient of chemical reaction.

*See also:* **Cluster Ions Measured Using Mass Spectrometry; Flame and Temperature Measurement Using Vibrational Spectroscopy; Gas Phase Applications of NMR Spectroscopy; Ionization Theory; Ion Molecule Reactions in Mass Spectrometry; Negative Ion Mass Spectrometry, Methods; Photoionization and Photodissociation Methods in Mass Spectrometry; Quadrupoles, Use of in Mass Spectrometry; Time of Flight Mass Spectrometers.**

## Further reading

Boltalina OB, Ponomarev DB and Sidorov LN (1997) Thermochemistry of fullerene anions in the gas phase. *Mass Spectrometry Reviews* **16**(N6): 333–352.

Droward J (1964) In: Mass spectrometric studies of the vaporation of inorganic substances at high temperature. Rutner E, Goldfinger F and Hirtz JP (eds) *Condensation and Evaporation of Solids*, pp 255–310. New York: Gordon and Breach.

Droward J and Goldfinger P (1967) Investigation of inorganic systems at high temperature by mass spectrometry. *Angewandte Chemie* **79**: 589–604.

Droward J (1985) Mass spectrometric studies at high temperature. In: Todd JFJ (ed) *Advances in Mass Spectrometry, Part A*, pp 195–214. New York: Wiley.

Hastie JW (1984) New techniques and opportunities in high temperature mass spectrometry. *Pure and Applied Chemistry* **56**: 1583–1600.

Hilpert K (1990) Chemistry of Inorganic Vapors. *Structure and Bonding* **73**: 97–198.

Inghram MG and Droward J (1960) Mass spectrometry applied to high temperature chemistry. In *Proceedings of an International Symposium on High Temperature Technology*, pp 219–240. New York: McGraw-Hill.

Sidorov LN and Sholtz VB (1972) Mass spectrometric investigation of two component systems of complex vapour composition by the isothermal evaporation method. *International Journal of Mass Spectrometry and Ion Physics* **8**: 437–458.

Sidorov LN, Zhuravleva LV and Sorokin ID (1986) High temperature mass spectrometry and studies of ion–ion, ion–molecule and molecule–molecule equilibrium. *Mass Spectrometry Reviews* **5**: 73–97.

Stolyarova VL and Semenov GA (1994) *Mass Spectrometry Study of the Vaporization of Oxides Systems*. New York: Wiley.

# Inorganic Chemistry, Applications of NMR Spectroscopy

*See* **Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules.**

# Inorganic Compounds and Minerals Studied Using X-Ray Diffraction

**Gilberto Artioli**, Università degli Studi di Milano, Italy

Copyright © 1999 Academic Press

HIGH ENERGY SPECTROSCOPY
**Applications**

## Introduction

X-ray diffraction from the solid state is a fundamental technique for the characterization of synthetic and natural materials. Since the discovery of Bragg-type X-ray diffraction from periodic crystal lattices, the technique has proven an essential tool for (1) the identification of crystalline compounds, (2) the quantitative analysis of polyphasic mixtures, (3) the study of the long-range atomic structure of crystals, including detailed charge density studies, and (4) the physical analysis of crystalline aggregates, including orientation texture, crystallite size distribution, and lattice microstrain effects. Such routine applications have developed during the first half of the 20th century into standard analysis procedures, and during the last decades they became so widely utilized that data are now available for most known inorganic compounds and minerals, to the extent that electronic databases are now accessible for automatic identification and rapid retrieval of crystallographic and structural information of crystalline substances.

The focus of the modern crystallographic research on inorganic solids is the interpretation of the physical and chemical properties of materials in terms of their ideal or defective atomic structure, and the transfer of the acquired crystal chemical knowledge to the engineering of solid-state compounds with novel technological properties. The years since the mid 1980s have also witnessed the development of dedicated second and third generation synchrotron radiation sources in the region of hard X-rays. The availability of brilliant, polarized and collimated synchrotron radiation beams with a wavelength that is tunable over a wide spectral range has made a number of interesting advances possible in materials characterization. State-of-the-art investigations of inorganic compounds and minerals by synchrotron X-ray diffraction include: *in situ* dynamic diffraction of kinetic processes and phase transformations, structural characterization of compounds at ultra-high pressures, and use of resonant scattering effects for element-selective structural analysis.

## X-ray diffraction analysis in the laboratory

X-ray diffraction (XRD) data collection and analysis is routinely performed in any laboratory dealing with inorganic or mineral compounds. Samples are commonly in the form of small single crystals, polycrystalline aggregates (powders), or oriented specimens (fibres, thin films, polished surfaces). The experimental techniques and the data analysis methods to be applied may therefore vary depending on the physical state and the chemical composition of the sample, besides of course the nature of the requested information. Typical laboratory applications in inorganic chemistry, solid-state physics and mineralogy may involve: (1) identification of unknown crystalline phases, (2) quantification of phase abundance in polycrystalline mixtures, (3) crystal structure determination and refinement, (4) physical analysis of the sample in terms of crystallite size, lattice microstrain, or orientation texture of the crystal domains. On one hand the mineralogist usually employs X-ray diffraction to identify and classify mineral specimens, to quantify mineral phases in raw materials for industrial use, or to characterize the temperature–pressure–time history of a rock through the detailed analysis of the crystal structure of the phases present in the mineral assemblage. On the other hand the inorganic chemist frequently characterizes by X-ray diffraction the products of synthesis reactions in an attempt to refine the synthesis processes or produce new materials with novel structural and physical properties.

The experimental apparatus commonly involve a sealed-tube or rotating-anode X-ray source, a 2-circles (for powders) or a 4-circles (for single crystals) automated diffractometer for sample orientation with respect to the incident beam, and one or more detector banks for detection of the diffracted signal. The angular position, the integrated intensity, and the peak profile shape of the diffracted signal are important for the subsequent analysis.

### Phase identification

Powder X-ray diffraction patterns are the fingerprint of all crystalline substances. The technique is a rapid and powerful tool for identification of all crystalline compounds, and it is especially valuable in the presence of polymorphs or quasi-isochemical compounds, where chemical techniques cannot be applied. The data collection time of a medium-resolution powder diffraction pattern for identification purposes is typically of the order of 20–30 min. Identification of the unknowns in a compound composed of a few phases is a straight-forward automatic procedure in which the Bragg diffraction peaks in the observed pattern are matched against the reported powder diffraction patterns of all known crystalline substances in the powder diffraction file. Advanced computer software is available for such applications.

If the investigated material is composed of a large number of phases, then the identification process can be more complicated due to substantial overlap of the Bragg peaks related to the different phases. Chemical information on the compound of interest may in this case help in defining subsets of phases from the database containing particular chemical elements. Subsets search may be more successful than the search performed on the whole database, although it requires prior knowledge of information on the sample.

### Quantitative phase analysis

Quantification of the relative abundance of crystalline phases in a multiphase mixture is an everyday problem in a wide range of applications. Common examples are: evaluation of the yield in inorganic synthesis and catalytic processes, characterization of raw mineral materials for industrial processes, quality check of fired ceramic products, and many more. While in most cases the required accuracy level of the analysis is a few percent at best, in particular cases such as in the quantification of phase contaminants in technologically important materials, or of hazardous and toxic phases in environmentally dispersed aerosols, the required level of accuracy must be substantially lower than 1 wt% relative abundance. Accuracy levels of 2–3 wt% are commonly reached if standard procedures of quantitative phase analysis by diffraction data are properly performed. Generally employed analytical methods include: the internal or external standard method, the matrix flushing method, and the reference intensity ratio method. Very recently, the availability of analysis techniques of powder diffraction data based on full-profile (Rietveld method), originally developed for structural analysis, had a great impact in the area of quantitative analysis: if the crystallographic structure of the phases present in the polycrystalline mixture are known, then the complete simulation of the diffraction pattern provides an easy standardless method for accurate quantification of all phases. The major advantage is that only one powder diffraction pattern is needed, with no need for standard addition or laborious external calibration of the peak intensities. **Figure 1** shows the observed, calculated and difference powder diffraction patterns relative to the quantitative analysis of a polycrystalline specimen composed of three standard phases, performed using the standardless Rietveld technique. Estimated standard deviations of relative proportions of all phases are in the range 0.1–0.2 wt%.

If the diffraction data are of above-average quality, the full-profile method can easily allow quantification of crystalline phases well below the 1 wt% accuracy level. Advanced nonroutine application of the full-profile technique involve modelling of anisotropic Bragg peak broadening for the quantification of defective or disordered crystal phases. The

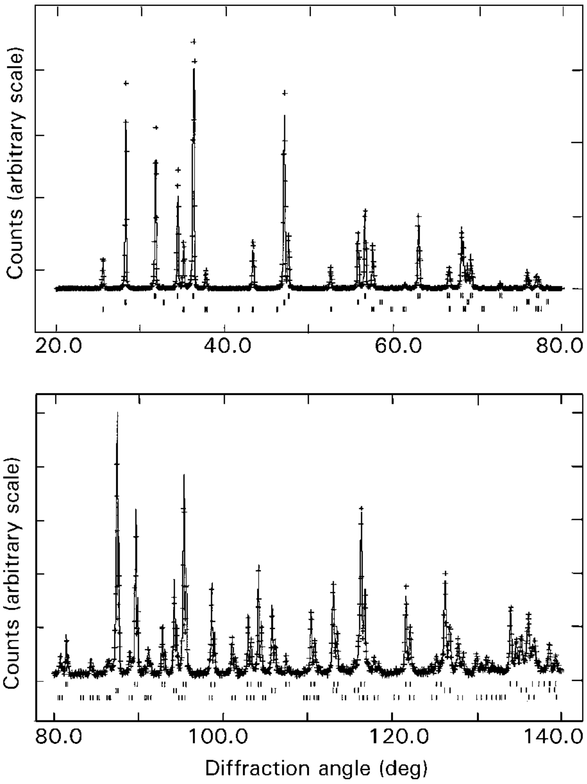



**Figure 1** Quantitative phase analysis of a polycrystalline mixture containing three phases performed by the Rietveld method. Selected portions of the observed (crosses) and calculated (solid line) X-ray powder diffraction patterns are shown. The vertical marks indicate the position of the single Bragg peaks.

combination of the full-profile Rietveld analysis with the addition of a known quantity of an internal standard in the mixture provides a simple and reliable method for estimation of the amorphous phase in samples containing noncrystalline components. The accuracy and reliability levels reached by modern methods of quantitative phase analysis based on powder diffractometry commonly exceed the specifications required by phase quantification in traditional and advanced industrial materials.

### Crystal structure analysis

The analysis of the atomic structure of crystals by diffraction is one of the more developed applications in the whole field of diffraction. Although often used as a simple, rapid technique for molecular structure determination, the full analysis of the crystal structure by X-ray diffraction indeed represents a powerful tool for the detailed interpretation of the relationship between the long-range atomic arrangement and the physicochemical properties in any solid-state material. Recent advances in the preparation, engineering and characterization of high-technology materials such as high-$T_c$ superconducting oxides and fullerenes, could hardly have been achieved without the accurate study of their crystal structure; and the long list of presently known synthetic microporous compounds widely used in catalysis would be much shorter were it not for the detailed structural data obtained on aluminosilicate zeolite minerals.

The standard crystallographic description of the atomic arrangement in a crystal includes the definition of a translation unit cell with its symmetry properties (Bravais lattice type and space group symmetry), the fractional coordinates of each crystallographic position together with the occupancy factor of each chemical species located on it, and its atomic displacement parameters (related to the probability density function of the electron distribution). Standard crystallographic descriptions as such are now available for most known solid-state compounds: the primary reference sources for inorganic crystals and minerals are the Structure Reports (in printed form) and the Inorganic Crystal Structure Database (in CD-ROM form). A number of comprehensive monographs review the essential concepts of chemical bonding in crystals and the classification of basic structural types. The vast majority of structural information was derived from single-crystal diffraction studies performed with standard laboratory apparatus at room temperature and pressure, although in recent years a growing number of inorganic crystal structure data has resulted from diffraction studies performed on polycrystalline samples using the full-profile Rietveld techniques of analysis. Equilibrium studies at nonambient temperature or pressure, and studies of slow kinetic processes are also accessible in modern X-ray laboratories, both using single-crystal and powder specimens.

The standard crystal structure description of a compound is commonly aimed at the geometrical analysis of the long-range atomic arrangement in terms of interatomic bond distances and angles, coordination polyhedra, and topological bond connectivity. This is the minimal information required for the correct interpretation of the macroscopic physicochemical properties of the compound at the atomic level. Common applications include the understanding of the solid-state reactivity of inorganic compounds, and the reconstruction of the crystallization and transformation processes, i.e. the geological record, in minerals.

Sometimes subtle or elusive physical effects in crystals require a deeper understanding of the chemical bonding and the crystallochemical role of the atoms. Examples are the oxygen site vacancies and cation substitutions in high-$T_c$ superconducting materials, the atomic or molecular polarization effects in nonlinear optical materials, the cation partitioning among sublattice sites in microporous framework structures, or the distinction between atomic static (i.e. site disorder) and dynamic (i.e. anisotropic motion) effects in minerals exhibiting nonideal solid solution behaviour. To face such complex problems, nonstandard single crystal structure analysis techniques are available. For example, detailed mapping of the electron density in a crystal can be performed by means of charge density studies, commonly performed by low-temperature (down to a few K) X-ray diffraction in order to minimize atomic thermal vibrations smearing out the density effects due to localized electrons. **Figure 2** shows the detailed mapping of the electron density of the H–O–H plane in the water molecule enclosed in the channel of a porous hydrous sodium aluminosilicate, the natural zeolite natrolite, whose complete structure is shown in **Figure 3**. The charge density distribution is derived from the experimental structure factor amplitudes measured by single-crystal X-ray diffraction. Also shown is the map of the Laplacian of the electron charge density in the same plane. On the basis of a Bader-type analysis the map allows detailed interpretation of the character of the chemical bonds in the crystal structure.

Accurate crystal structure studies can be performed on the same crystal in a wide temperature range and allow extrapolation of the true atomic

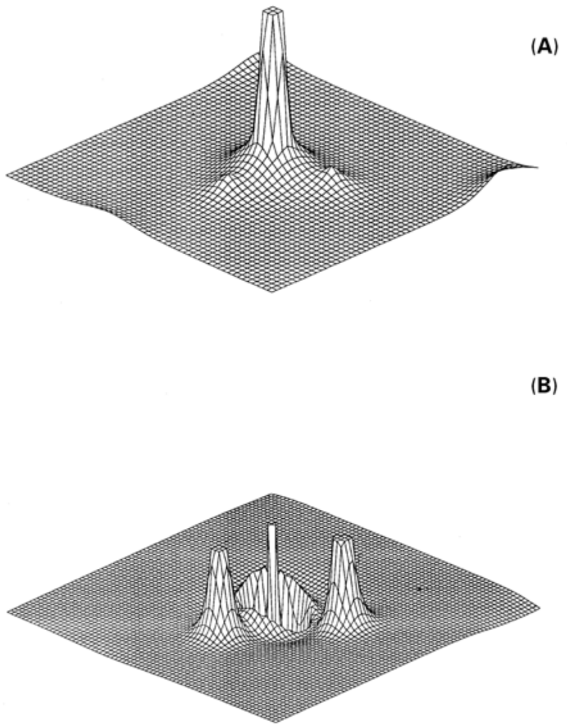



**Figure 2** Experimental distribution of the electron charge density (A) in the H–O–H plane of the water molecule in the zeolitic channel of natrolite obtained from single-crystal X-ray diffraction data. Detailed interpretation of the chemical bond features, including the position and character of the critical points, is possible through the analysis of the maps of the Laplacian of the charge density (B).

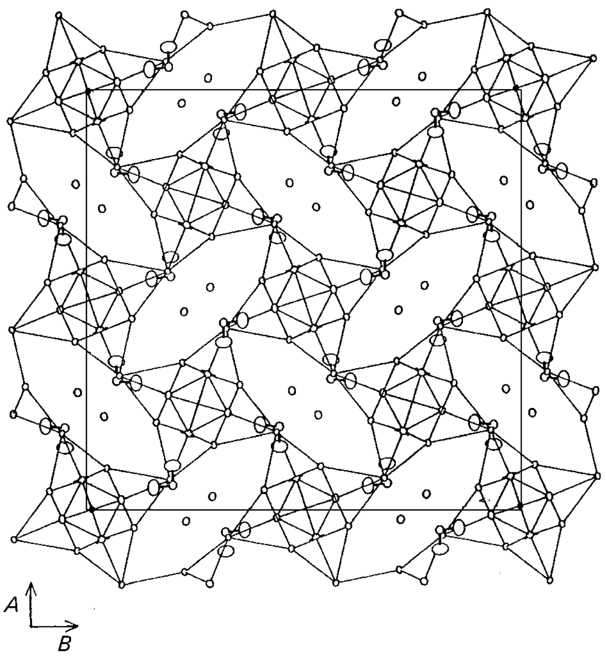



**Figure 3** Projection of the structure of the zeolite natrolite along the orthorhombic *c* axis, showing the zeolitic channels containing the Na atoms (small circles) and the water molecules forming hydrogen bonds to framework oxygen atoms.

displacement parameters down to the 0 K temperature region, evaluation of the zero-point energy of the atom, and control of the presence of static subsite disorder effects. **Figure 4** shows selected experimentally determined values of the atomic displacement parameters for the Mg, Si, Al and O atoms in a synthetic pyrope garnet of composition $Mg_3Al_2Si_3O_{12}$. The deviation of the values at 30 K from the linear fit performed on the atomic displacement values measured above 298 K is due to the contribution of the zero-point energy to the atomic motion. Accurate diffraction data of this kind allow separation of the static and dynamic contributions to the measured atomic displacement, and many also indicate the presence of anharmonic vibrational components, although these effects are more appropriately evaluated using temperature-dependent neutron diffraction data.

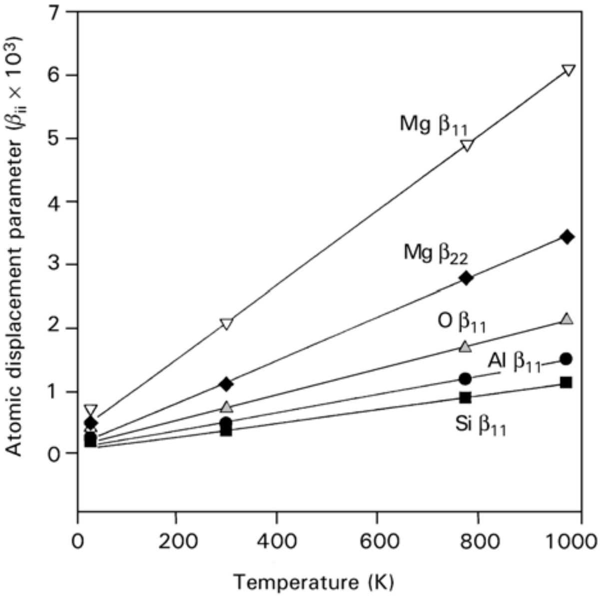



**Figure 4** Temperature dependance of selected atomic displacement parameters for the Mg, Si, Al and O atoms in pyrope garnet. Analysis of the displacement parameters obtained from accurate single-crystal diffraction data allow evaluation of the zero-point energy contribution, separation of static and dynamic effects, and detection of anharmonic vibrational contributions to the atomic motion.

### Physical analysis of crystals and crystal aggregates

The basic information concerning the nature of the phases present in a synthetic or natural material, and

the ideal models of their crystallographic structures are frequently known. The information of interest for the characterization of the material and the understanding of its crystallochemical behaviour then reside in the physical status of the crystal or crystal aggregate. It is known that many material properties of interest in technological applications are directly related to what we might call the deviation from the 'ideal' crystal, i.e. the presence in the material of a variety of structural defects at the atomic level, such as doping elements, chemical modulations and stacking faults. These perturbations of the periodic crystal lattice produce effects of various kinds observable in the diffraction signal, i.e. diffuse scattering, superlattice reflections, commensurate or noncommensurate Bragg peak modulations, peak shifts, peak profile distortions, etc. A substantial part of X-ray crystallography deals with the measurement of the diffraction effects caused by lattice disorder, and their interpretation in terms of structural defects. The whole branch of 'quasicrystallography' is devoted to alloys and solid materials exhibiting diffraction effects incompatible with the standard three-dimensional space group symmetry.

A rapidly developing area of study is the measurement of lattice microstrain in polycrystalline aggregates. It is based on the interpretation of the angular dependence of the broadening of Bragg peak profiles in a powder diffraction pattern. The surface or volume mapping of the distribution of lattice strain gradients in the material yields a variety of information on the presence of residual stresses, developing tensions due to microcracking, local imperfections due to coprecipitated phases, and so on. Furthermore the analysis of the peak profile broadening as a function of crystallographic direction (anisotropy of the broadening effects) allows determination of the mean particle size and shape of the crystallites in the compound.

The analysis of the statistical or preferred orientation of the crystallites in solid polycrystalline materials is commonly referred to as texture analysis. Again, the diffraction technique allows the definition of the relationship between a microscopic property, i.e. the orientation of the crystallites defined as coherent diffraction domains, and the macroscopic physical properties of the crystal aggregate. Texture studies are of course crucial in the characterization of oriented synthetic materials such as cold-rolled metals or oxide thin films, but they are also of great relevance in the study of the formation processes of mineral assemblages. As an example, the texture features of olivine or pyroxene minerals in meteoritic chondrules yield information on the early condensation sequence and cooling history of the meteorite body they belong to. **Figure 5** shows the heavily textured diffraction pattern of chondrule from the Mafra L-4 chondritic meteorite collected using synchrotron radiation and a charge-coupled device (CCD) area detector. The pattern displays at the same time sharp diffraction spots produced by the olivine and pyroxene microcrystals present within the chondrule, and the inhomogeneous intensity distribution on the diffraction rings produced by the polycrystalline matrix. The example indicates how hard X-rays from synchrotron sources can effectively be used to investigate precious materials in a nondestructive way.

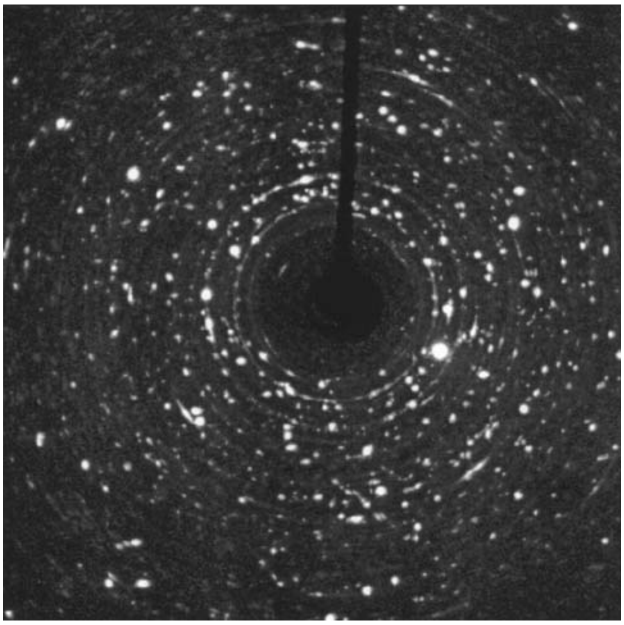


**Figure 5** Textured diffraction pattern from a meteoritic chondrule collected using short wavelength synchrotron radiation ($\lambda$ = 0.5 Å) and a CCD area detector. The image displays sharp diffraction spots from olivine and pyroxene single crystals, and powder diffraction rings from the polycrystalline matrix of the chondrule.

## Synchrotron X-ray diffraction

The availability in recent years of the intense, collimated and energy-tunable X-ray beams emitted by synchrotrons and electron storage rings has had a spectacular impact on most fields related to materials characterization. As a consequence, a number of theoretical and instrumental aspects of spectroscopic techniques are rapidly developing in order to take full advantage of the peculiar properties of synchrotron radiation. All kinds of laboratory X-ray diffraction experiments described above can be performed at synchrotron sources as well, and the

results are of significantly higher quality in terms of angle- or energy-resolution and signal-to-noise ratio. Furthermore, since the incident radiation beam is several orders of magnitude more brilliant than the laboratory X-ray beams, the experiments can usually be performed using a much smaller volume of sample (down to few $\mu m^3$) or, alternatively, by collecting diffraction data with extremely fast acquisition times (down to the ms time range). Finally, several completely new types of experiment, such as those outlined in the following sections, are only possible at synchrotron sources.

### Dynamic diffraction

The study of the evolution of the diffraction signal in a sample during a phase transformation or a chemical reaction involving crystalline phases is defined as dynamic diffraction. The term implies that the sample is studied *in situ* under operating conditions of temperature, pressure, vacuum or controlled atmosphere; and that the diffraction pattern is sampled using acquisition times much shorter than the time the process takes to reach completion. Diffraction experiments under nonambient conditions can be performed on laboratory instrumentation only at thermodynamic equilibrium. Direct determination of the metastable and transient phases present during the reaction path, and measurement of fast kinetics are only possible if dynamic diffraction using synchrotron radiation is employed. Commonly the resolution in reciprocal space (defined as $\Delta d/d$, or the ability to distinguish two closely spaced Bragg diffraction peaks) is not a critical experimental parameter when studying the qualitative aspects of the phase transformations, or for the quantification of phase proportions during the process (for example in kinetics studies). On other hand, resolution is a crucial factor in the investigation of structural phase changes, where high-quality powder spectra collected in time-resolved mode are required.

A typical dynamic diffraction experiment involves (a) a high X-ray flux incident on the sample, (b) a conditioning chamber around the specimen controlling the environmental variables, and (c) a fast data acquisition system (i.e. position sensitive or area detectors). Depending on the require $\Delta d/d$ resolution, at synchrotron sources these experimental conditions are easily met either using angle-dispersive or energy-dispersive geometries, i.e. using monochromatic or polychromatic incident radiation, respectively. The flexible design of a synchrotron experiment also allows simultaneous *in situ* experiments to be carried out, such as XRD and small-angle X-ray scattering, XRD and differential scanning calorimetry, or XRD and X-ray absorption spectroscopy in fast scanning mode. Combined experiments of this kind provide a wealth of information, as they furnish the means of simultaneously investigating the process at different space- or momentum-transfer scales. For example, simultaneous time-resolved X-ray powder diffraction and X-ray absorption spectroscopy performed during a solid-state reaction catalysed by a specific chemical element can provide identification and quantification of all the crystalline phases involved, and at the same time also provide short-range information on the coordination environment of the catalyser element at each stage of the reaction.

Recent applications of the dynamic diffraction experimental techniques were mostly in the areas of temperature-induced dehydration processes in natural zeolites, hydrothermal crystallization of aluminosilicate and aluminophosphate microporous materials, thermal decomposition of layer silicate minerals, high-temperature synthesis of advanced ceramics, and hydrothermal ion-exchange and conversion processes in synthetic molecular sieves. The time-resolved powder diffraction patterns relative to the isothermal nucleation and growth process of zeolite-A by the hydrothermal treatment of activated metakaolinite is shown in **Figure 6**.

Although most of the cited studies only imply the qualitative analysis of the transformation, or the simple quantification of the crystalline phases involved in the process in order to properly extract the kinetic parameters, a number of full structural studies have been performed, which allow the complete determination of the structure response to the external gradient applied to the sample.

The case study of the temperature-induced phase transformation in the Ca-rich zeolite mineral garronite is illustrated here to show the large amount of detailed information to be gained by dynamic diffraction studies. The time evolution of the experimental powder diffraction pattern is shown in **Figure 7**. The image was obtained by synchrotron radiation using a gas flow heater, a polycrystalline sample in a glass capillary, and an image plate detector installed on a translating camera. Constant-temperature integration of the image pixels yield a number of powder diffraction histograms for the structure analysis. **Figure 8** shows the Rietveld analysed powder spectra of phase A at room temperature, and phase B at 170°C. Full structural analysis of all powder data sets produce a detailed picture of the structure response to dehydration, involving the positional shift and the change in coordination of the Ca cation in zeolitic channels, the progressive distortion of the aluminosilicate tetrahedral framework

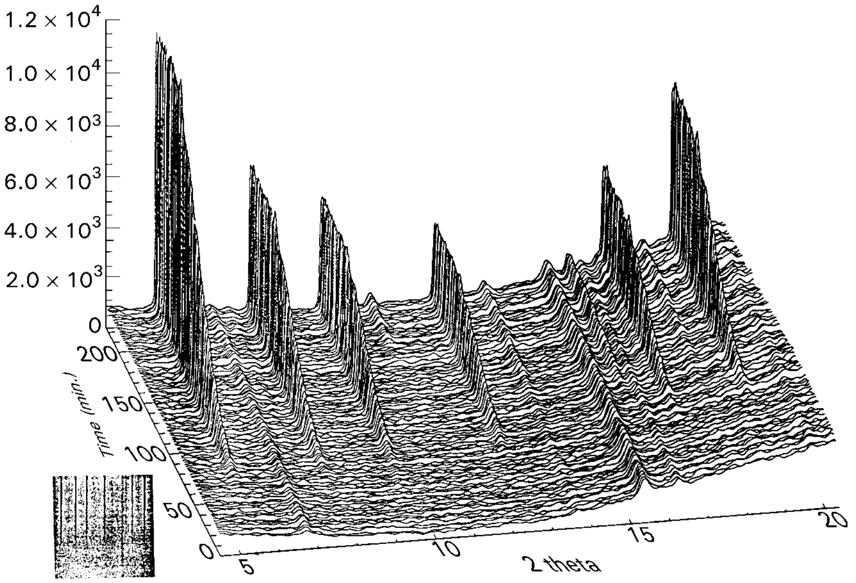



**Figure 6** Time evolution of the powder diffraction pattern of zeolite-A crystallizing from activated metakaolinite under hydrothermal conditions. The original image plate picture is shown as a small figure at the bottom left.

inducing a space-group change from monoclinic to tetragonal, and ultimately the collapse and amorphization of the structure. **Figure 9** describes the progression of the process by the measured unit cell parameters and volume, and **Figure 10** shows the distortion of the zeolite channels caused by the contraction of the tetrahedral framework along the *c* tetragonal direction in phase B, which corresponds to the *b* monoclinic axis of phase A.

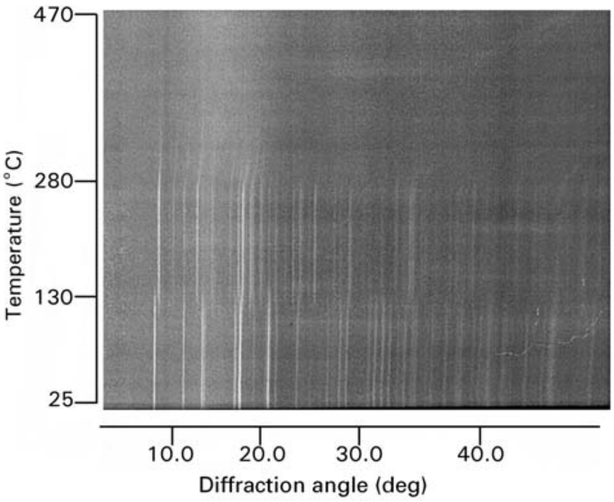



**Figure 7** Evolution of the X-ray powder diffraction pattern of the natural zeolite garronite during the temperature-induced dehydration process. The date were collected using synchrotron radiation, a gas flow heater, a capillary sample, and an image plate detector installed on a translating camera.

### High pressure studies

Small volume samples are required to reach ultrahigh pressures (i.e. above $10^2$ GPa). Therefore, high-brilliance synchrotron radiation is needed to get a useful diffraction signal out of microsamples. This is the reason why the new techniques of high-pressure diffraction were mostly developed with synchrotron radiation sources. Physics, materials science, and especially mineral physics greatly benefited from such developments, because of the possibility of studying high-pressure-stabilized crystal structures using *in situ* diffraction techniques. Since the earth's and planetary interiors are not directly reachable for sampling, we need laboratory techniques to stabilize and characterize the materials satisfying the observed large-scale geophysical (density distribution within the planetary interior, seismic waves velocity) and geochemical (abundance of chemical elements in the solar system, element distribution within the planetary bodies) parameters governing the global planetary processes, such as the thermal flow, the mantle convection regime and the crust plate tectonics.

High-pressure mineral physics therefore is concerned with the characterization of the physical properties and stability fields of planetary materials under extreme pressure and temperature conditions. For example, conditions present at the centre of the earth are thought to be in excess of 300 GPa and 5000 K. Such conditions can at present be obtained by laser-heated diamond anvil cells or multianvil presses. Emphasis of current *in situ* diffraction

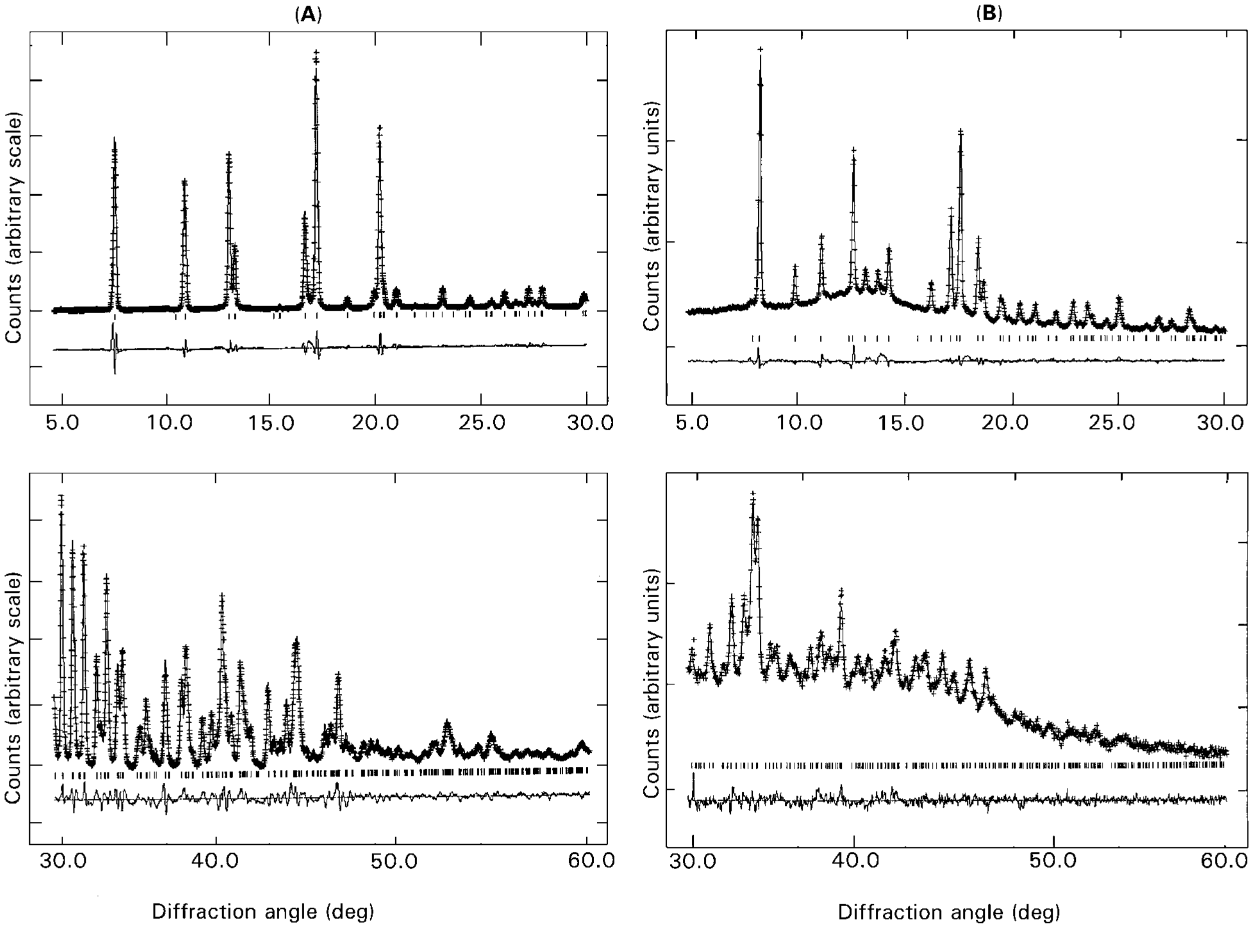



**Figure 8** Selected portions of the Rietveld-analysed powder diffraction pattern of the zeolite garronite at 25 (A) and 170 (B)°C. The patterns were integrated from the image plate data shown in **Figure 2**. A full structure analysis of the zeolite as a function of temperature was performed.

studies using synchrotron radiation under extreme conditions is on the definition of the thermodynamic phase stability fields of the crystalline phases possibly present in the earth's mantle and core, on the experimental measurement of the equations of state in P–V–T space, and in the crystal structure determination of the high-pressure phases.

## Resonant scattering experiments

A noticeable advantage of synchrotron radiation is the possibility of fine tuning the incident energy over a wide range, typically 5–50 keV for bending magnet radiation from high-energy third-generation synchrotrons (up to 100 keV for wiggler radiation from the same machines) and 2–20 keV for bending magnet radiation from low-energy third-generation synchrotrons. The available energy range covers absorption K-edges of chemical elements from P to the rare earth elements, and L-edges from Rb to the actinides. The wavelength tunability is not only of obvious advantage for X-ray absorption spectroscopy, but it also offers diffraction an increased flexibility through proper exploitation of the resonant scattering effect. The excitation of resonant scattering takes place when diffraction is performed using an incident energy close to that of an absorption edge of a chemical element present in the sample. Under these conditions the atomic scattering factor of the element is modified by a factor (called the anomalous dispersion correction) amounting to several electrons. It is obvious that by performing diffraction studies at different wavelengths in the vicinity of an absorption edge, it is possible to change to scattering contrast between the chemical elements present in the sample, and therefore increase the ability of diffraction to (1) determine the location of the particular element in the structure, or (2) quantify the statistical amount of that element in specific crystallographic sites.

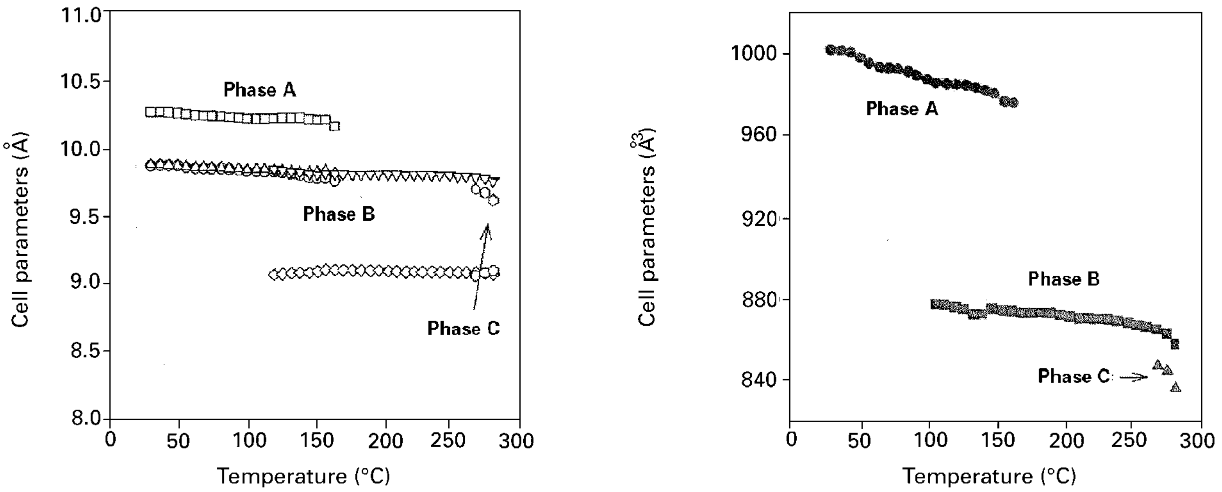



**Figure 9** Structure evolution of the zeolite garronite during the dehydration process. The unit cell parameters and volume resulted from the Rietveld refinements of all powder diffraction data shown in **Figure 2**. Phase A is monoclinic *I*2/*a* (circles = cell *a*; squares = cell *b*; triangles = cell *c*), phase B is tetragonal $P4_12_12$ (triangles = cell *a* = *b*; diamonds = cell *c*) and phase C is also tetragonal $P4_12_12$ (hexagons = cell *a* = *b*; circles = cell *c*). It may be noted that the cell volume contraction during the monoclinic–tetragonal phase transition is related to the shortening of the monoclinic *b* axis alone (corresponding to the *c* axis of the tetragonal cell). The structure analysis indicates that the framework distortion is related to the change in bond coordination of the Ca cation in the zeolitic cages.

Since an energy shift of the absorption edge is observed depending on the oxidation state of the element, the accurate determination of the energy dependence of the anomalous dispersion term yields information on the valence state of the element present in a crystallographic site, much the same way this information is obtained from the chemical shift of the edge in X-ray absorption spectroscopy. However, the information obtained from absorption spectroscopy is averaged over the sample volume, and cannot discriminate between different valence states of the same element simultaneously present on different structural sites, whereas diffraction under resonant conditions intrinsically yields site-resolved information. The resonance effect is also exploited for the resolution of crystal structures, especially organic macromolecules, using the multiwavelength anomalous dispersion technique.

Diffraction exploiting the resonant scattering effect is generally used in inorganic chemistry and mineralogy to determine the cation partitioning of specific elements among different crystallographic sites, or to evaluate the oxidation state of chemical elements present in specific structural sites. A promising

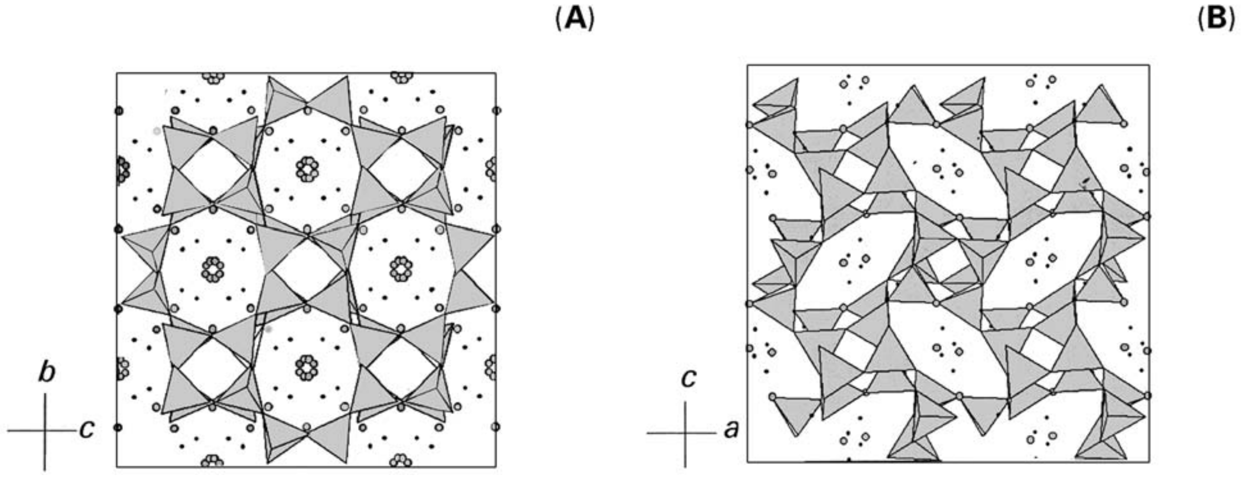



**Figure 10** Projections of the crystal structures of garronite: (A) phase A at 25°C along the [100] monoclinic direction; (B) phase B at 170°C along the [010] tetragonal direction. The distortion of the zeolitic channels caused by the change in coordination of the Ca atoms is evident.

extension of the technique is called diffraction anomalous fine structure spectroscopy, and it involves measurement of the energy dependence of the integrated intensity of particular Bragg diffraction peaks across an absorption edge.

*See also:* **Fibres and Films Studied Using X-Ray Diffraction; Materials Science Applications of X-Ray Diffraction; Neutron Diffraction, Instrumentation; Neutron Diffraction, Theory; Powder X-Ray Diffraction, Applications; Quantitative Analysis; Small Molecule Applications of X-Ray Diffraction; Structure Refinement (Solid State Diffraction).**

## Further reading

Bader RFW (1994) *Atoms in Molecules, a Quantum Theory*. Oxford: Clarendon Press.
Coppens P (1992) *Synchrotron Radiation Crystallography*. London: Academic Press.
Guinier A (1963) *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies*. San Francisco: W.H. Freeman.
Hyde BG and Andersson S (1989) *Inorganic Crystal Structures*. New York: John Wiley.
ICSD, *Inorganic Crystal Structure Database. Karlsruhe*: FIZ Fachinformationszentrum.
Materlik G, Spark CJ and Fischer K (eds) (1994) *Anomalous X-ray Scattering. Theory and Applications*. Amsterdam: North Holland.
*PDF, Powder Diffraction File*. Swarthmore, PA: ICDD International Center for Diffraction Data.
Structure Reports (1940–1990) *Metals and Inorganic Sections*. IUCr International Union of Crystallography Geneva: IUCr; Dordrecht: Kluwer.
Wells AF (1984) *Structural Inorganic Chemistry*. Oxford: Clarendon Press.
Willis BTM and Pryor AW (1975) *Thermal Vibrations in Crystallography*. Cambridge: University Press.
Young RA (ed) (1993) *The Rietveld Method. IUCr Monographs on Crystallography*. Oxford: Oxford University Press.
Zevin LS and Kimmel G (1995) *Quantitative X-ray Diffractometry*. New York: Springer-Verlag.

# Inorganic Condensed Matter, Applications of Luminescence Spectroscopy

**Keith Holliday**, University of Strathclyde, Glasgow, UK

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Applications**

## The Interaction of impurities in condensed matter with light

The optical spectroscopy of impurities in solids and liquids is similar to that of free particles except that interpretation requires an extra interaction to be taken into account, i.e. the influence of the host on the energy levels of the impurity. In other words, the presence of the impurity particle in a particular environment adds an extra term to the Hamiltonian: the crystal field.

There are two components of the crystal field; static, referring to the average position of nearest neighbour particles, and dynamic, referring to their motion. The static crystal field shifts energy levels and lifts degeneracies to cause splittings whilst the dynamic crystal field allows coupling of electronic energy levels to vibrations. In terms of luminescence spectroscopy, the static crystal field is principally responsible for shifting emission wavelengths whilst the dynamic component broadens emission lines into bands of widely varying widths and shapes. Both static and dynamic effects can influence the oscillator strength of the transition and thus the luminescence lifetime. Excited state lifetimes can also be influenced by factors such as nonradiative decay and concentration quenching.

Insulators provide a useful trap for optically active ions due to their transparency in the visible region. Nuclear (vibrational) resonances with electromagnetic radiation take place in the infrared part of the spectrum whilst resonances with bound electrons (from the valence band to the conduction band) are usually in the ultraviolet. Visible radiation therefore

interacts with the impurities alone. The role of the host is simply to influence the optical behaviour of the impurity.

## Spectroscopic characteristics of impurities in condensed matter

### Lanthanide ion doping

The optically active electrons in lanthanide ions are 4f electrons and they are shielded from the influence of the crystal field by the complete 5s and 5p subshells that surround them. Therefore, the static crystal field has only a small effect on the energy levels of lanthanides and is treated as a perturbation to the free ion states. This perturbation is significant, however. Electric dipole transitions between pure 4f states are parity forbidden (though some of the much weaker magnetic dipole transitions are allowed). The addition of the crystal field term causes weak mixing of 4f states, most importantly with 5d states, so that some electric dipole transitions become weakly allowed. Reducing the symmetry of the crystal field tends to cause greater mixing thus making the optical transitions more intense. At the same time, the *J*-degeneracy of the state is lifted so that the luminescence spectrum of a particular transition appears as a number of closely separated lines such as those shown in **Figure 1A**. The dynamic crystal field is also weakly interacting with lanthanide ion impurities so that phonon-assisted transitions are rarely significant features.

Luminescence studies are particularly useful for lanthanide-activated matter as transitions are often weak thus rendering absorption difficult to measure. Each lanthanide ion has a unique energy level structure so that excitation of such a material will result in a luminescence spectrum characteristic of that impurity ion. Only a small proportion of energy levels luminesce, however. Most decay nonradiatively through multiphonon emission to states at slightly lower energies. The excited electron rapidly tumbles between states until it reaches one that is separated by more than about 1000–3000 cm$^{-1}$ (1 eV ≈ 8000 cm$^{-1}$), depending on the system, from the next lowest level. The probability of multiphonon decay is an exponential function of separation so that states with wide gaps to next lowest levels are unlikely to decay nonradiatively. Luminescence is thus emitted from such states. As an example, the energy structure of $Nd^{3+}$ is shown in **Figure 1B**.

### Transition metal ion doping

The active electrons in transition metal ions are not shielded by outer subshells and this leads to a much

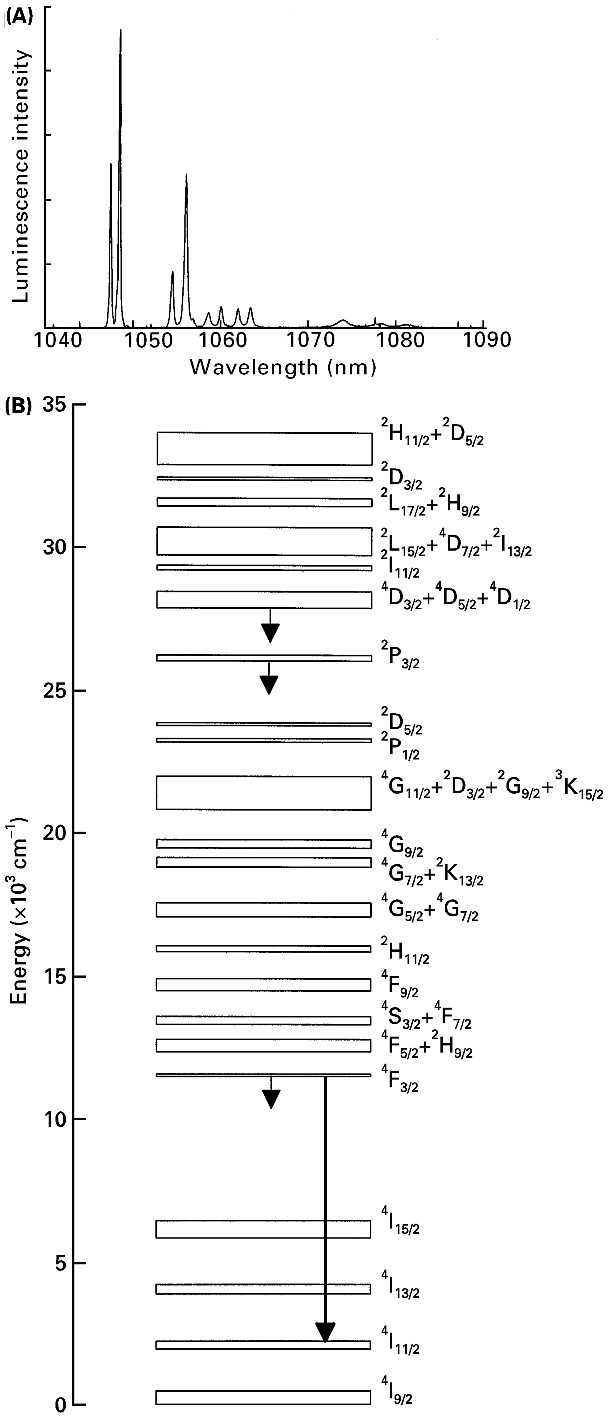



**Figure 1** Luminescence from a typical lanthanide-doped crystal, $Nd^{3+}$-doped $KLiYF_5$. The $^4F_{3/2} \rightarrow {}^4I_{11/2}$ luminescence multiplet, most often used as the gain transition in Nd-based lasers, is shown in (A) and the energy level diagram is shown in (B). The total $Nd^{3+}$ multiplet splittings for this host are indicated by the thickness of the boxes, levels that radiate significantly are denoted by small arrows and the $^4F_{3/2} \rightarrow {}^4I_{11/2}$ transition is indicated by a long arrow. This luminescence multiplet would normally be expected to be split into no more than six components but the 12 observed indicates that $Nd^{3+}$ substitutes into two different sites in this crystal. The similarity of the spectra due to different site substitution is typical for lanthanides and the energy level diagram in (B) varies very little from host to host.

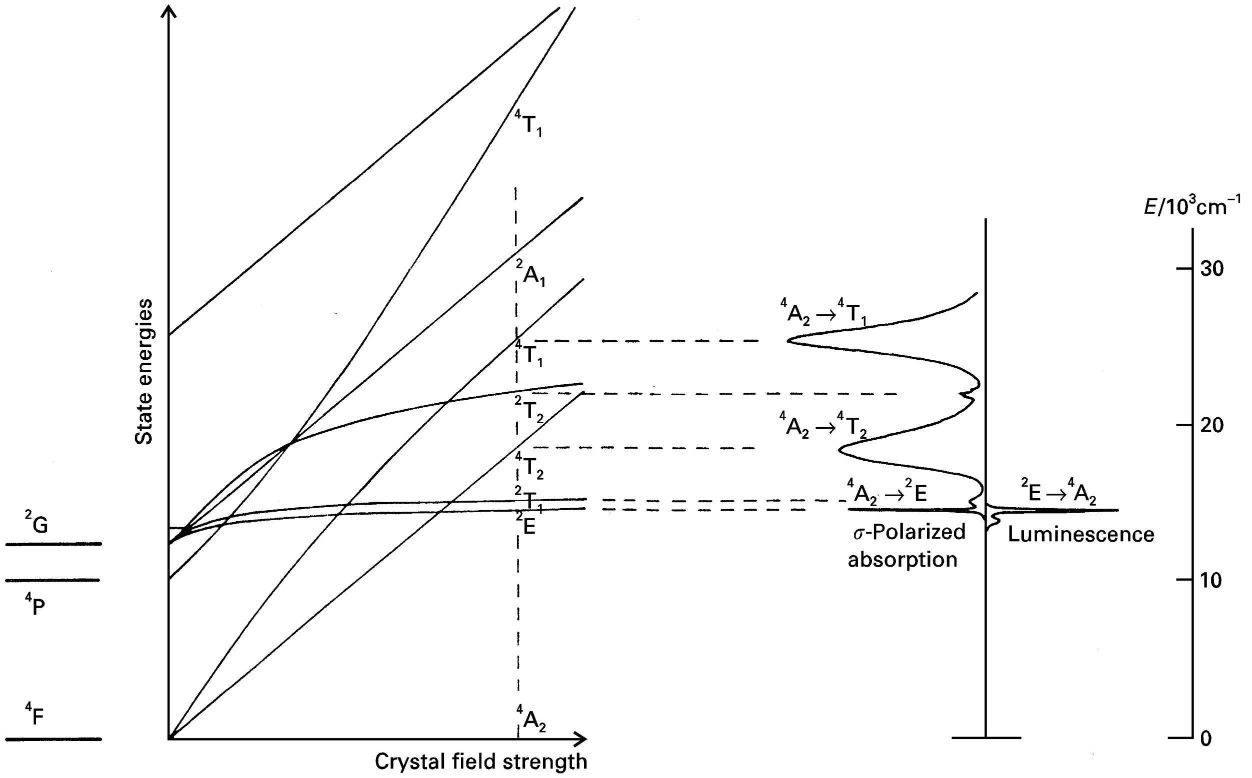



**Figure 2** A Tanabe–Sugano diagram for a 3d³ impurity ion such as $Cr^{3+}$. The octahedral crystal field increases along the *x*-axis and the energy levels split and shift as indicated. For zero crystal field the free ion LS states are indicated on the left. A dashed line indicates the crystal field strength relevant to ruby. The energy levels indicated appear in the absorption spectrum shown on the right. Luminescence is only from the lowest lying excited state, $^2E$. For materials that provide weaker field environments the $^4T_2$ state is lower and much broader bandwidth emission is obtained. Reproduced with permission of Oxford University Press from Henderson B and Imbusch GF (1989) *Optical Spectroscopy of Inorganic Solids.* Oxford: Clarendon Press.

stronger interaction with the crystal field. The Hamiltonian contains the same components as for lanthanides but now the crystal field cannot be treated as a perturbation to the free ion states as this term is similar in strength to the electron–electron interactions and stronger than spin–orbit coupling terms. The resulting wavefunctions are denoted by a nomenclature that derives from group theory.

Plots, known as Tanabe–Sugano diagrams, show how the state energies of transition metal ions vary as a function of crystal field strength. From these plots absorption and luminescence characteristics can be predicted or explained. **Figure 2** shows the Tanabe–Sugano diagram for $Cr^{3+}$. The dashed line indicates the crystal field strength for aluminium oxide ($Al_2O_3$:$Cr^{3+}$ is ruby) and it indicates where the excited state energy levels lie relative to the ground state ($^4A_2$). Absorption bands appear resonant with the energy levels, broad in the case of states that are strongly vibronically coupled and narrow for weak electron–phonon coupling. The emission spectrum consists of a single line with some weak vibronically assisted structure regardless of which absorption transition is excited. This is because, as for lanthanides, rapid nonradiative decay prevents luminescence from states that have other levels at energies close below. In this case, narrow band luminescence is seen (the ruby laser transition) but for materials that provide a weaker crystal field environment for $Cr^{3+}$ ions, the $^4T_2$ state may be at a lower energy than the $^2E$ state making it the emitting level. Much stronger electron–phonon coupling occurs for this transition (as for the reverse $^4A_2 \to {}^4T_2$ (absorption) transition indicated in **Figure 2**) and broadband emission results.

## Colour centres

Colour centres (or F centres) result from the presence of defects in the structure of a solid, often intentionally induced, for instance by high-energy-photon irradiation. Such treatment causes disruption to the lattice and leads to the production of particular defects. For instance, a single electron can become trapped in an anion vacancy. The Hamiltonian for such a system is considerably different from that of an impurity ion but the resulting eigenenergies give rise to transitions in the optical region of the

spectrum. Such a centre is influenced very strongly by the crystal field and so transitions are always strongly coupled to vibrations resulting in very broad absorption and luminescence spectra (~1000–4000 $cm^{-1}$). There is a wide range of different optically active defects and determining the structure of colour centres is difficult using only optical techniques. However, by applying external fields and stresses and making measurements of properties such as polarization and magnetic circular dichroism much can be understood. Of great assistance are magnetic resonance measurements, in particular electron paramagnetic resonance. Such techniques reveal information on the number of electrons present and the symmetry of the colour centre from which structure can be deduced.

## Crystals, glasses and liquids as hosts

It might be expected that the various types of condensed matter host will give rise to quite different optical properties for the impurities that they contain. In fact, the impurity has a much stronger influence than the host. This is especially true for lanthanide ions. These are not greatly influenced by their surroundings so that a change of environment does not make very much difference to optical behaviour. Splittings between *J*-levels will vary between hosts and transition strengths will also vary whilst remaining relatively weak.

Transition metal ions are more strongly influenced by environment, for instance $Cr^{3+}$ ions decay via entirely different transitions depending on crystal field strength. Generally, transition metal ions will have an energetically preferred orientation. The energy of solvation of the $Cr^{3+}$ ion is lowest when octahedrally coordinated, in solids and liquids. In aqueous solution six water molecules are arranged about the chromium ion. In glasses lack of long-range order means that the solvation energy of the $Cr^{3+}$ ion can be minimized without influencing the overall structure of the host so that, again, $Cr^{3+}$ ions become octahedrally coordinated, this time surrounded by the dominant anions (often fluoride or oxide ions). In crystals the situation is somewhat different. The lattice energy of the host is minimized by long-range order so that considerations of the solvation energy of the $Cr^{3+}$ ion are of much less significance. If there is a suitable octahedral site (or somewhat distorted octahedral site) for the $Cr^{3+}$ ion then substitution can occur. If not, then the chromium ion may enter as an alternative valence (e.g. $Cr^{4+}$ in a tetragonal site) or may not enter the lattice at all. In fluids and glasses it is more likely that a transition metal ion can find its energetically preferred orientation due to the impurity ion's greater influence on its own neighbourhood. In crystals, transition metal ions often enter more distorted sites according to availability and this sometimes leads to more unusual optical properties such as large energy level splittings or transition strengths.

The main difference between spectra obtained in different classes of hosts is in the broadening of features. In a perfect crystal every impurity ion would experience the same potential due to neighbour particles so that each ion would behave in a spectroscopically identical way. Spectra would not be broadened, other than by homogeneous and vibrational effects, and would consist of very narrow zero phonon lines and phonon side bands. Perfect crystals do not exist and inhomogeneous broadening due to variation in crystal field contributions to the Hamiltonian is always present (see **Figure 3**). This is not always apparent at room temperature when homogeneous broadening dominates in most systems but below about 100 K, in solids, inhomogeneous broadening has a significant influence. In glasses and liquids inhomogeneous broadening is usually more significant. The local environment of the impurity ion is influenced less strictly by the host and the lack of order leads to a much greater range of environments and thus greater broadening of spectral features.

Lanthanide doped crystals typically give rise to luminescence multiplets with spreads in energy of around 100–1000 $cm^{-1}$. Line widths of individual features at low temperatures tend to be around 100 times narrower than this. In glasses and liquids, however, the effect of variation in environment causes luminescence to be emitted in bands that encompass the whole of the multiplet and individual lines cannot usually be resolved. Such bandwidths are not greatly broadened beyond the multiplet widths of crystals, however.

The zero phonon lines of transition metal ions behave similarly to those of lanthanides but phonon side bands are more evident. Such is the extent of broadening due to electron–phonon coupling that the influence of environment is proportionally much smaller. It is unusual to see disorder-induced broadening of phonon-assisted bands by more than a factor of two.

The greatest broadening effects are actually observed in substitutionally disordered crystals. Substitutional disorder in crystals occurs when two or more constituent ions have interchangeable lattice positions. Whilst the stoichiometry and periodicity of the crystal is maintained, the local environment varies throughout the structure. Optically active impurity ions experience different crystal fields as a

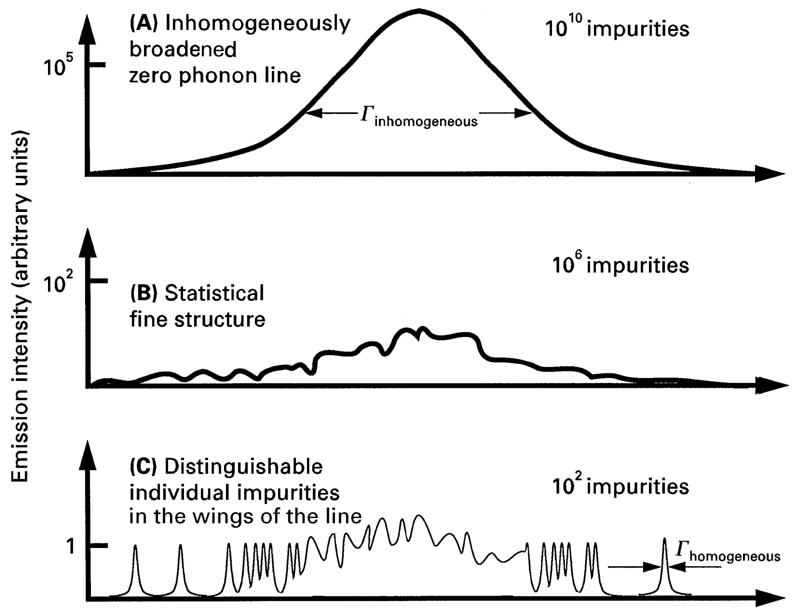



**Figure 3** A typical zero phonon luminescence line from a bulk sample would appear as in (A) with a typical inhomogeneous width ($\Gamma$) of 1–10 cm$^{-1}$. As the number of emitting impurities decreases, the smooth line shape typical of a large distribution is lost (B) and eventually the contributions of individual homogeneously broadened contributions appear in the wings (C). The homogenous width is exaggerated by a factor of about $10^4$ for temperatures of less than about 4 K.

result of this local variation. Inhomogeneous broadening through substitutional disorder is strongest when it is the nearest neighbour ions that are randomly occupied. For instance, the $^7F_0 \rightarrow {^5D_0}$ transition of $Sm^{2+}$ in BaFCl is broadened by a factor of 30 at low temperatures through the addition of bromine to the melt to produce $BaFCl_{0.5}Br_{0.5}$. Such is the magnitude of the effect that inhomogeneous broadening dominates homogenous broadening up to temperatures in excess of 300 K. This discovery prompted a successful search for room temperature persistent spectral hole-burning materials (see below).

## Techniques used to study optically activated condensed matter

Very simple methods can be used to obtain useful information on the optical behaviour of impurity doped condensed matter. The first step in the characterization of such a material is usually to measure its absorption spectrum in the region of about 0.3–1 μm. Such measurements are often sufficient to identify the active impurities and to provide some information on their environment. For crystalline hosts, the polarization of absorption spectra can be analysed to draw conclusions on site symmetries and crystal field parameters (see **Figure 4**).

As well as ground state absorption measurements, it is often also useful to know the absorption characteristics of excited metastable states, particularly for potential laser gain media. Such measurements are considerably more difficult than ground state measurements because it is necessary to maintain a significant excited state population. The experimental arrangement required to obtain such spectra requires two excitation sources, a powerful beam to excite the material and a weaker probe to measure the absorption immediately after the excitation pulse.

There are some materials for which even ground state absorption is too weak to measure. A method that provides equivalent information is zero-order excitation spectroscopy. Emission at all wavelengths is directed to a detector without dispersion. A time delay is required to prevent excitation light from swamping the signal. Ideally, this is achieved using a pulsed laser and gated detector but other systems also work, such as asynchronous choppers in a continuous beam. By varying the wavelength of the exciting beam and recording the strength of emission at all wavelengths a spectrum equivalent to that of the absorption of all luminescent centres is recorded.

The zero-order excitation spectrum reveals the wavelengths at which the material can be excited to obtain luminescence. The luminescence spectrum can be recorded by exciting at one of these wavelengths and dispersing the emission using a monochromator. When emission is weak a chopper can be placed in the excitation beam to produce a modulated intensity that a lock-in amplifier can detect.

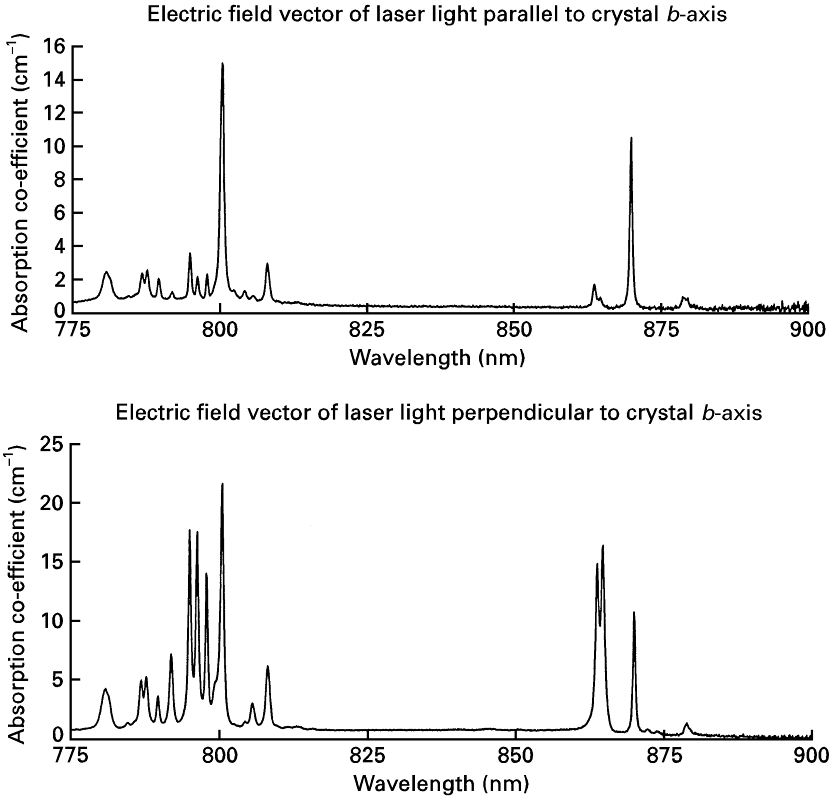



**Figure 4** The absorption spectrum of the $^4I_{9/2} \rightarrow {}^4F_{5/2} + {}^2H_{9/2}$ and $^4I_{9/2} \rightarrow {}^4F_{3/2}$ transitions (see **Figure 1B**) in $Nd^{3+}$-doped $KLiYF_5$ measured in two different polarizations at 77 K. Similar variations in intensities within multiplets are observed in luminescence spectra.

Generally, excited ions step between energetically nearby states nonradiatively until a large gap is encountered from which radiative transitions occur. In a crystal in which all impurity ions occupy equivalent sites similar spectra are obtained regardless of which absorption transition is excited, unless there are two or more widely separated emitting states. In the latter case a different luminescence spectrum is obtained for each emitting state when excitation is resonant with levels between it and the next highest emitting state (see **Figure 1**).

Occasionally, the same impurity ion may substitute into two or more crystallographically different impurity sites within a single host. Each will normally have its own luminescence characteristics so that excitation at particular wavelengths may excite one set of ions and not the other when absorption bands do not overlap. This occurs most often for materials with narrow band absorption spectra, in particular lanthanides, and different luminescence spectra will result depending on the selected absorption feature (see **Figure 5**). Energy transfer between sites may cause both sites to emit under some circumstances.

In hosts in which a distribution of sites exist, shifts in emission bands result from excitation at different energies within an absorption band. The site distribution may be broadened by variation in crystal field strength or in electron–phonon coupling strength and measurements of the shift of the emission band as a function of excitation energy allow the contributions of each distribution to be quantified. **Figure 6** shows the luminescence spectrum of $Cr^{3+}$-doped strontium gallogermanate, a substitutionally disordered crystal that gives rise to luminescence via both the $^2E \rightarrow {}^4A_2$ and $^4T_2 \rightarrow {}^4A_2$ transitions due to a large variation in crystal field strength that causes the emitting state to vary between sites (see **Figure 2**).

Selective excitation techniques may be used to make high-resolution measurements. At low temperatures where inhomogeneous broadening dominates, it is possible to use a highly monochromatic excitation source such as a single frequency dye laser to select a subset of impurities, each of which contributes to the same portion of a spectral line. The resolution of spectral measurements is then, in principle, limited by the homogeneous line width. Below 10 K this is typically of the order of $10^6$ times smaller than the inhomogeneous line width.

In the absence of energy transfer, exciting a subset of impurities will result in luminescence from only this subset. Single mode dye lasers can routinely

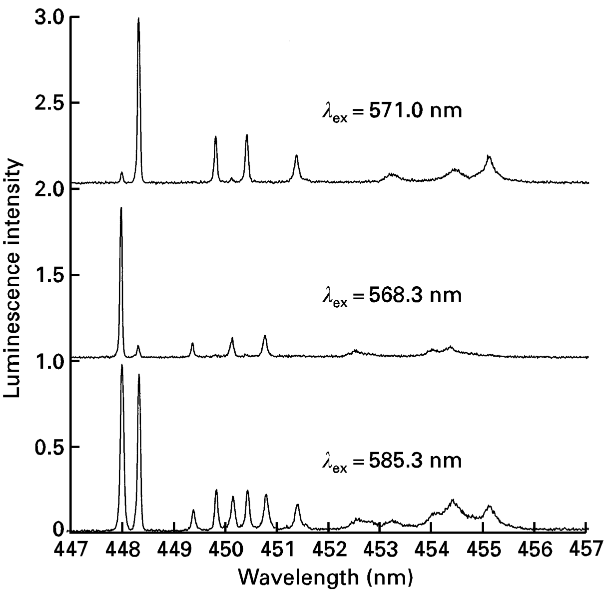



**Figure 5** Selective excitation spectra for the $^2P_{3/2} \rightarrow {}^4I_{13/2}$ transition in $Nd^{3+}$-doped $KLiYF_5$. The top two spectra are produced by exciting the material at wavelengths that result in luminescence from only one of the two $Nd^{3+}$ substitutional sites whereas the bottom spectrum is excited at a wavelength that excites both. This is an example of upconversion as the exciting photon energy is lower than the output photon energy. Site selection in this case occurs during excited state absorption from the $^4F_{3/2}$ state. Reproduced with permission of Elsevier Science from Russell DL, Henderson B, Chai BH, Nicholls JFH and Holliday K (1997) *Optics Communications* **134**: 398–406.

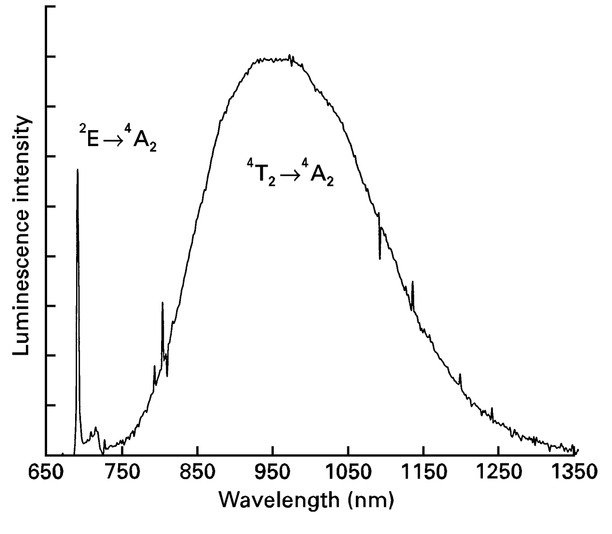



**Figure 6** The luminescence spectrum of $Cr^{3+}$-doped $Sr_3Ga_2Ge_4O_{14}$ which decays via two different transitions due to the large variation in crystal fields experienced by the $Cr^{3+}$ ions (see **Figure 2**). The structure in the $^4T_2 \rightarrow {}^4A_2$ band is instrumental noise.

provide excitation line widths of 1 MHz and solid state tunable lasers can do considerably better. A typical monochromator or scanning Fabry–Perot etalon cannot reach such resolutions and so the detected emission is limited by the resolution of the dispersion apparatus. Nevertheless, fluorescence line narrowing (see **Figure 7E**) can reveal information on, for instance, ground state splittings, that are normally hidden beneath the inhomogeneous broadening of the transition. Fluorescence line narrowing is usually employed on purely electronic transitions but is quite generally applicable to these.

For certain materials the excitation process results in a subsequent decrease in absorption at the same wavelength. This selective bleaching may persist indefinitely and can occur through a variety of mechanisms, all of which rely on the removal of a proportion of absorbing impurities from resonance. The spectral hole line shape may then be recorded by scanning the excitation wavelength and the spectral resolution is determined by the laser line width alone, making possible measurements of homogenous line widths. Persistent spectral hole-burning (see **Figure 7C**) is not a universal phenomenon but does occur for a wide range of materials. Amongst the most common mechanisms for impurities to be removed from resonance are oxidation (e.g. $Sm^{2+} \rightarrow Sm^{3+}$), motion of the impurity ion within the lattice (or alternatively a redistribution of the local environment about the impurity ion) and redistribution of ground state nuclear spins. The persistence of the spectral hole allows perturbation experiments (e.g. Stark effect, Zeeman effect, stress effects) to be undertaken with much greater sensitivity than is possible when observing the effects of very high fields on entire inhomogeneously broadened spectral lines.

The ultimate way to unambiguously examine the behaviour of impurities and defects in host matrices is to detect single chromophores (see **Figure 3**). The principle underpinning this luminescence excitation technique relies, like spectral hole-burning, on the reduction of the homogenous line width of the zero phonon electronic transition with decreasing temperature. Inhomogeneous broadening is insensitive to temperature and so the overlap of spectral lines from individual chromophores decreases with decreasing temperature. If the sample is then diluted until the concentration is sufficiently low, the superposition of absorption contributions is no longer smooth. This is known as statistical fine structure. As the sample is diluted further, single chromophores will have no effective absorptive spectral overlap with other chromophores in the wings of the absorption line and any emission detected will correspond to this single atom, ion, molecule or defect centre. With continuing dilution, single chromophores may be detected closer to the inhomogeneous line centre. To date, this technique has only been

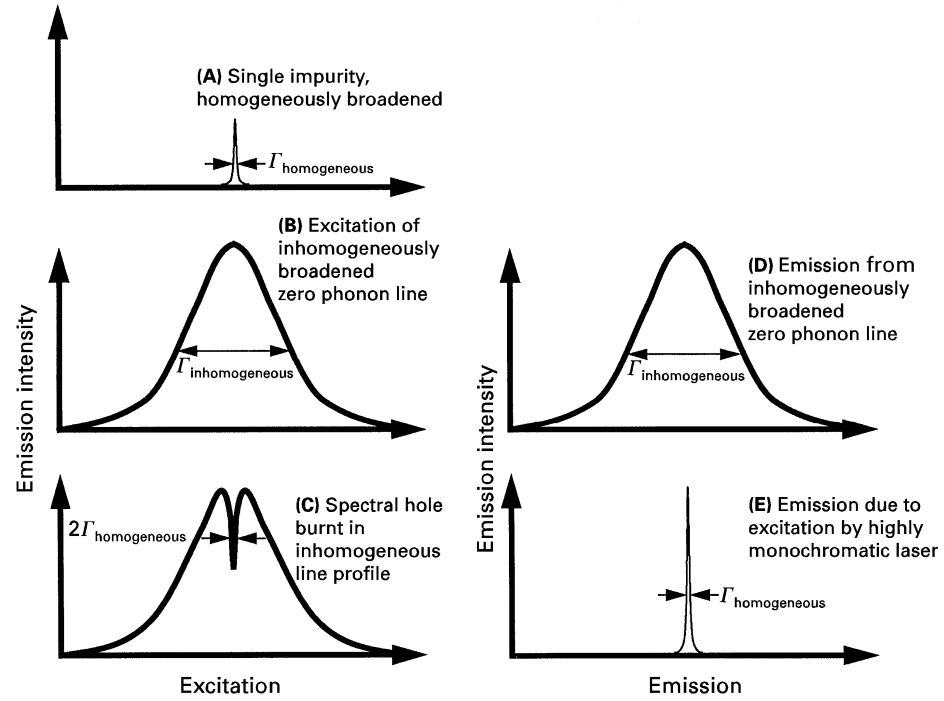



**Figure 7** Schematic illustration of two high-resolution laser techniques. On the left is a technique that relies on excitation (or absorption) spectroscopy, spectral hole-burning, which is therefore instrumentally limited only by laser line width. On the right, fluorescence line narrowing is indicated, a technique that involves recording an emission spectrum which is therefore limited both by laser line width and dispersion resolution.

applied to dye molecules in organic hosts but could in principle be applied to any impurity doped solid that has a zero phonon electronic transition.

## Luminescence phenomena studied in optically activated condensed matter

The interaction between impurity (or defect) and host determines the energy levels of the optically active species and so the most obvious application of luminescence spectroscopy to condensed matter is determination of site structure. The magnitude of the splittings of electronic levels is determined by the strength of the crystal field and the extent to which degeneracies are lifted is determined by the site symmetry of the crystal field. Simple absorption, excitation or luminescence spectra can thus reveal such information through application of group theoretical analytical techniques and this can be enhanced with data from perturbation experiments. Clearly these methods are most useful for materials in which the electron–phonon coupling is weak so that zero phonon lines dominate and hence lanthanide (and some transition metal) doped crystals are most appropriate for such a study. Doping impurities into hosts is not a sensible method to study the pure host as the presence of the impurity can distort the local lattice structure, but more general conclusions on the host (such as the extent of crystallinity) can be reached by studying parameters such as the broadening of transitions.

When electron–phonon coupling is strong, so that zero phonon lines are weak, other techniques can be used to determine information about site structure. Polarized measurements of features in absorption, excitation and luminescence spectra provide direct evidence of the symmetries of the wavefunctions involved in these transitions and this leads to information on site structure. Usually this information is best used in conjunction with other spectroscopic techniques such as electron paramagnetic resonance (detected either directly or optically). The form of a phonon-assisted sideband can itself reveal information on site structure, especially the extent of the local deformation that occurs when the impurity is excited to a higher electronic state. **Figure 8** compares the luminescence spectra of $Cr^{3+}$ impurity ions in $LiCaAlF_6$ and $LiSrAlF_6$, crystals that are isomorphous but with different until cell sizes. The decreased bandwidth and increased structure associated with the $LiCaAlF_6$ spectrum indicates weaker electron–phonon coupling and a smaller excited state deformation. Again, group theoretical analysis can be applied to such results to obtain detailed information. For

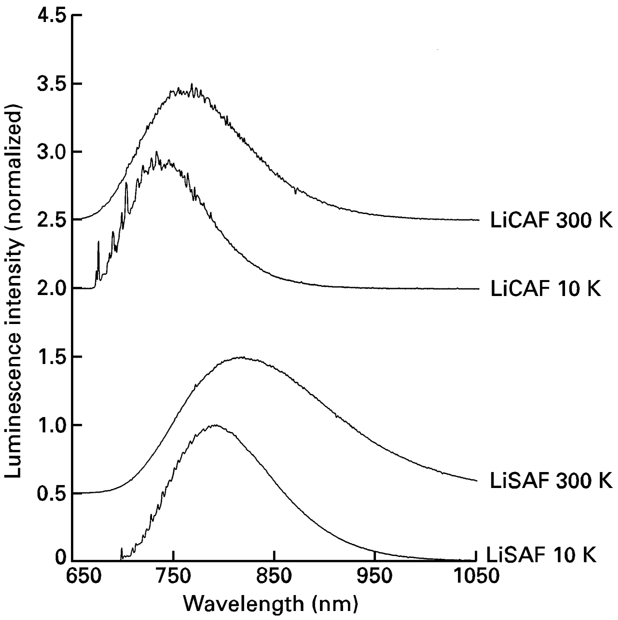



**Figure 8** Luminescence spectra of the $^4T_2 \rightarrow {}^4A_2$ transition at indicated temperatures in $Cr^{3+}$-doped $LiCaAlF_6$ (LiCAF) and $LiSrAlF_6$ (LiSAF), host crystals that are isostructural but with slightly different unit cell sizes.

$LiCaAlF_6$ specific phonon transitions can be identified at low temperatures. Application of polarized spectroscopy to these features and comparison with Raman spectra can provide further information on the details of electron–phonon coupling and site structure.

Upconversion refers to a process in which a material emits photons of higher energy than those with which it is excited (see **Figure 5**). It is a phenomenon most commonly associated with lanthanide-doped materials and can take place through a variety of mechanisms including cross relaxation, energy transfer, excited state absorption and two-photon absorption. Studying the dynamics of such systems using time resolved techniques is an excellent method for exploring these processes within solids. Variation of doping levels and pump intensities help to determine exactly which processes dominate and the way in which they operate.

Fluorescence line narrowing is also a useful technique for studying energy transfer, in particular when measurements are time resolved. When a subsection of the impurities in a material is excited via energy selection it is the same impurities that luminesce, except when energy transfer takes place. If time-resolved spectra are taken the rate at which energy transfer processes progress can be determined as the line broadens. The line narrowing that results from energy selection can also be used to resolve small splittings in either the ground or excited state of the material. $Cr^{3+}$-doped strontium gallogermanate provides a nice example of this. **Figure 9** shows spectra obtained by exciting the $^4A_2 \rightarrow {}^2E$ electronic transition and recording emission from the reverse transition. By using relatively low resolution dispersion (a monochromator) the relatively large excited state splitting (~ 20 $cm^{-1}$) can be measured, but when detecting using a Fabry–Perot etalon the ground state splitting (~ 2 $cm^{-1}$) is revealed.

Spectral hole-burning has been used to investigate a wide variety of phenomena that includes impurity and defect structure, electron–nuclear interactions, spin-spin cross relaxation, tunnelling processes, optical dephasing, spectral diffusion in glasses, electron–phonon coupling and the behaviour of special materials such as biological compounds (involved in photosynthesis for example) and thin films. Much fuller details are given elsewhere in this encyclopedia.

A range of phenomena can be studied in solution, though generally only the simpler techniques discussed above are available due to the dynamics of the liquid phase. Measurements of luminescent lifetime can be interpreted in terms of the coordination number of the optically active solvated ion. Quenching and sensitizing of luminescence by the manipulation of the solution (e.g. addition of other species, change of acidity, etc.) is another common area of investigation. Generally, such effects are chemical in nature. The luminescent ion forms a bond, either loosely with surrounding solvent molecules or more strongly with ligand species. Quenching or sensitization takes place depending on the interaction between the optically active ion and its immediate surroundings. Much of this area of study is related to organic chemistry in that large organometallic ions are formed in solution and it is their behaviour that is being probed in optical experiments. This is discussed further elsewhere in the encyclopedia.

## Technological applications of optically activated condensed matter

To investigate materials such as laser gain media only simple experimental techniques are usually required. To evaluate the potential of a new material in a technological application it is often sufficient to measure absorption and emission spectra along with luminescence lifetimes, though it is desirable to also have an indication of relevant loss mechanisms. These include nonradiative decay, excited state absorption and energy transfer. The result of significant nonradiative decay is indicated by variations in luminescence intensity and lifetime as a function of temperature

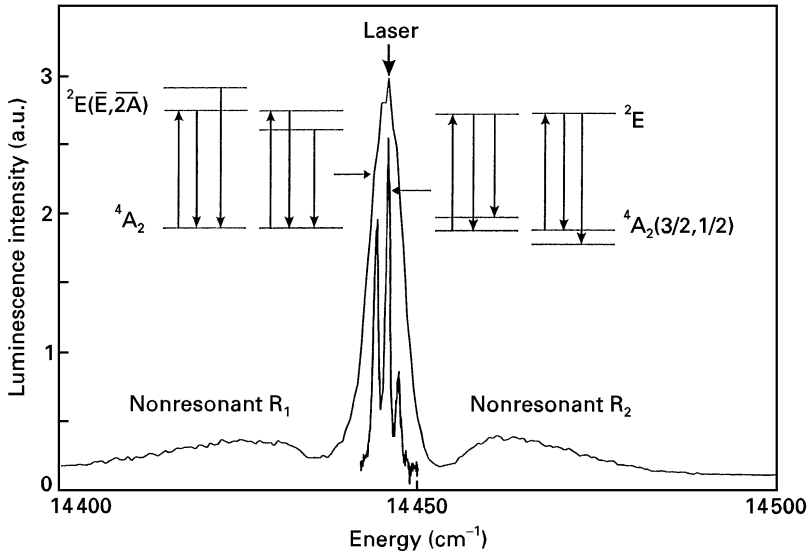



**Figure 9** Fluorescence line narrowed spectra of the $^2E \rightarrow {}^4A_2$ transition in $Cr^{3+}$-doped $Sr_3Ga_2Ge_4O_{14}$ (see **Figure 6**). When dispersing emission using a monochromator (long spectrum) the excited state ($^2E$) splitting of ~20 $cm^{-1}$ is observed. When dispersing emission using a Fabry–Perot etalon (internal spectrum) the ground state ($^4A_2$) splitting of ~2 $cm^{-1}$ is observed. Reproduced with permission of the American Physical Society from Grinberg M, Macfarlane PI, Henderson B and Holliday K (1995) *Physical Review* B **52**: 3917–3929.

whereas energy transfer can be investigated using techniques such as time-resolved fluorescence line narrowing and upconversion. The results of studies of laser gain media and related materials has led to the concept of 'crystal field engineering' where the ideal combination of impurity and crystal field for any particular application can be calculated and then created through careful choice of host. In reality there is always feedback between spectroscopic interpretation and technological implementation.

The best known materials that have been developed over the last four decades are for implementation in lasers, either as gain media or as saturable absorbers. The first laser gain medium was ruby (see **Figure 2**), providing single wavelength output in the visible part of the spectrum. Neodymium based lasers have now largely replaced ruby as the industrial laser of choice. The $^4F_{3/2} \rightarrow {}^4I_{11/2}$ electronic transition (see **Figure 1**) provides laser emission just beyond 1 μm in a number of hosts, most notably in yttrium aluminium garnet ($Y_3Al_5O_{12}$). More recently, tunable laser gain media such as $Ti^{3+}$-doped sapphire ($Al_2O_3$) and $Cr^{3+}$-doped $LiSrAlF_6$ (**Figure 8**) have been developed and these provide variable wavelength outputs across the near-infrared. Both systems are based on vibronically assisted luminescence bands that arise due to strong electron–phonon coupling. Many other laser systems have been developed, including those that use lanthanide ion doped optical fibres as the gain medium and others that are able to sustain oscillation on upconversion transitions.

The much greater inhomogeneous broadening that results from disorder in glasses has also been exploited, for instance in erbium-doped fibre amplifiers. In these devices a small part of an optical fibre communication system is doped with erbium. A diode laser is coupled into the fibre and pumps the erbium ions to an excited metastable state. As the signal passes through this part of the fibre it is amplified through stimulated emission. Using a glass host for this purpose broadens the transition and allows broad bandwidth operation.

There are more speculative applications associated with some of the more advanced techniques. For instance, spectral hole-burning has been suggested as a high density optical storage medium whereby binary digits are indicated by the presence or absence of a hole at any given wavelength thus adding wavelength to spatial position in a compact disc-like memory. Extensions to this idea include the storage of images as spectral holes through the production of wavelength multiplexed holographic gratings and the combination of these images through Stark effect splittings of spectral holes to produce an optical parallel processor. These ideas have been demonstrated in principle but, to date, are only feasible when the sample is maintained at very low temperatures.

Amongst the more important applications of luminescence spectroscopy in the liquid phase is the detection of small amounts of impurities in naturally occurring solutions such as rivers and groundwater. A full knowledge of quenching and sensitization

effects are required to be able to solve these problems. Of particular importance in this area is the detection of pollution and especially radioactive pollution that can be dangerous in very small quantities. Actinides are generally strongly luminescent and so spectroscopic techniques are very effective.

*See also:* **EPR, Methods; Fluorescence Microscopy, Applications; Fluorescent Molecular Probes; Hole Burning Spectroscopy, Methods; Laser Magnetic Resonance; Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Light Sources and Optics; Luminescence Theory; Near-IR Spectrometers; Raman Optical Activity, Applications; Symmetry in Spectroscopy, Effects of; UV-Visible Absorption and Fluorescence Spectrometers; Zeeman and Stark Methods in Spectroscopy, Applications.**

## Further reading

Bartram RH and Henderson B (1999) *Crystal Field Engineering of Solid State Laser Materials*. Cambridge: Cambridge University Press.

Choppin GR and Peterman DR (1998) Applications of lanthanide luminescence spectroscopy to solution studies of coordination chemistry. *Coordination Chemistry Reviews* **174**: 283–299.

Gan F (1995) *Laser Materials*. Philadelphia: World Scientific.

Henderson B and Imbusch GF (1989) *Optical Spectroscopy of Inorganic Solids*. Oxford: Clarendon Press.

Holliday K and Wild UP (1993) Spectral hole-burning. In: Schulman S (ed) *Molecular Luminescence Spectroscopy: Methods and Applications Part III*, Chapter 5, pp 149–228.

Jorgensen CK and Reisfeld R (1982) Uranyl photophysics. *Structure and Bonding* **50**: 121–171.

Kaminskii AA (1990) *Laser Crystals* Berlin: Springer.

Kaplyanskii AA and Macfarlane RM (eds) (1987) *Spectroscopy of Solids Containing Rare Earth Ions*. Amsterdam: North-Holland.

Sugano S, Tanabe Y and Kamimura H (1970) *Multiplets of Transition Metal Ions in Crystals*. New York: Academic Press.

Wybourne BG (1965) *Spectroscopic Properties of Rare Earths*. London: Interscience.

# Interstellar Molecules, Spectroscopy of

**AGGM Tielens**, Kapteyn Astronomical Institute, Groningen, The Netherlands

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| Applications |

## Introduction

The origin and evolution of life has long fascinated mankind. It is directly tied to the abiotic formation and chemical history of the biogenic elements, in particular H, C, N, O, P and S. The evolution of most elements starts with nucleosynthesis inside the fiery cauldrons of stellar interiors. Eventually, these elements are injected into the interstellar medium, either through violent supernova explosions or through more gentle stellar winds which terminate the life of most stars. Thus, stellar formation, evolution and death are intimately connected, as the ashes of one generation of stars become the building blocks of the next generation. Much of the expelled material is in the form of molecules, ranging from the simple diatomics such as molecular hydrogen ($H_2$) and carbon monoxide (CO) to more complex species such as polycyclic aromatic hydrocarbons (PAHs) and fullerenes ($C_{60}$), or in the form of small (≈ 100 nm) dust grains – carbon-based (i.e. soot and silicon carbide) and silicates (aluminium oxide, olivine, enstatite). In their evolution from their birth sites, through the interstellar medium, until their incorporation into new planetary systems, these elements undergo a complex chemical history. Various processes cycle the material from the gas phase to the solid phase and from one chemical compound to

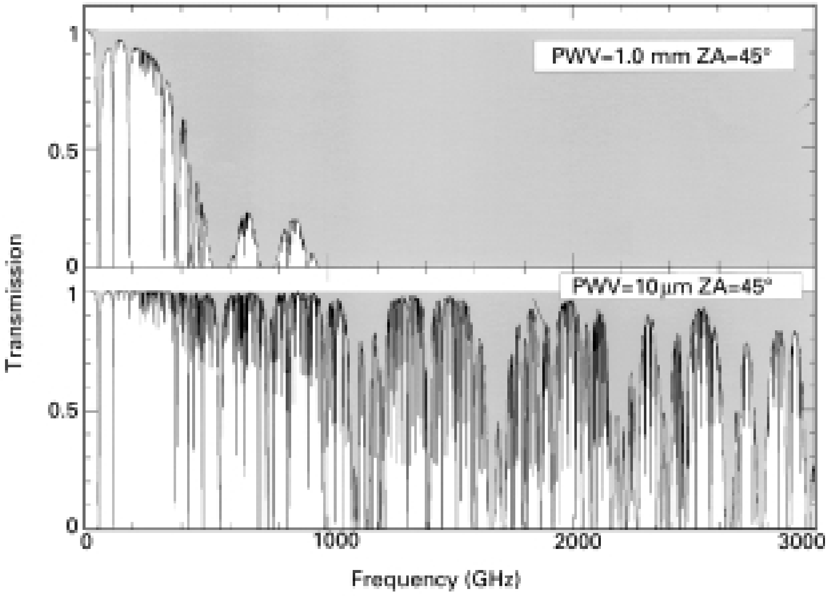



**Figure 1** Atmospheric transmission in the submillimetre and far-infrared for typical observing conditions on a dry mountain top with 1 mm of precipitable water vapour and under a zenith angle of 45 degrees (top panel) and on an airborne platform with 10 μm of precipitable water vapour flying at an altitude of about 40 000 feet also under a zenith angle of 45 degrees (bottom panel). Courtesy of TG Philips.

another. Understanding this history and its relation to the origin of life is one of the key questions of modern astrophysics.

The most important tool for astronomers to understand the origin and evolution of the molecular universe is spectroscopy. This includes studies of the rotational emission spectrum of the cold interstellar gases in the microwave wavelength region. Near bright stars, vibrational modes can be excited through fluorescence. Electronic spectra are predominantly measured through absorption spectroscopy against a bright background source. Each of these wavelength regions has its own techniques and provides unique information on interstellar molecules. Together they point to a chemically diverse and active molecular universe. Each of these techniques and their results will be briefly summarized.

## Rotational spectroscopy

Hundreds of pure rotational lines of a variety of molecules have been detected in the interstellar medium, most of them in emission but some also in absorption. These lines occur throughout the millimetre, centimetre and decametre wavelength region of the spectrum and are detected using radio antennas and heterodyne detection techniques at spectral resolutions of typically 300 000. At submillimetre and far-infrared wavelengths, the atmosphere is not very transparent (**Figure 1**) and telescopes are located on high, dry sites, on airborne platforms or in space. **Figure 2** shows the results of a line survey of a star-forming region in the Orion constellation, illustrating the richness of the spectrum and the diversity of the molecules contributing. Many of these rotational lines are bright enough to map the spatial distribution of the molecular emission. Such maps can be intercompared to study the interrelationship of various molecules. Also, they can be compared to maps of (newborn) stars to determine the interaction of stars with the surrounding (natal) molecular gas.

Similar spectral and spatial studies have been performed for the ejecta of stars in the later stages of their life to probe their molecular composition and evolution. Because the central star heats the gas to warm temperatures (up to 1000 K), rotational lines in the far-infrared can be excited as well. **Figure 3** shows the far-infrared spectrum of such a star, IRC 10216, obtained by the Infrared Space Observatory. For these objects, molecules can also be detected through their rovibrational transitions against the bright far-red, and near- and mid-infrared continuum of the star.

Rotational transitions are very characteristic for molecules and, at the high resolution allowed by heterodyne techniques, precise molecular identifications can be made using laboratory measured rotational constants. Over a hundred different molecules have been detected in interstellar space (**Table 1**) and new ones are found at a rate of a few per year. Some of these species are very simple well-known molecules such as water and ammonia; however, most are

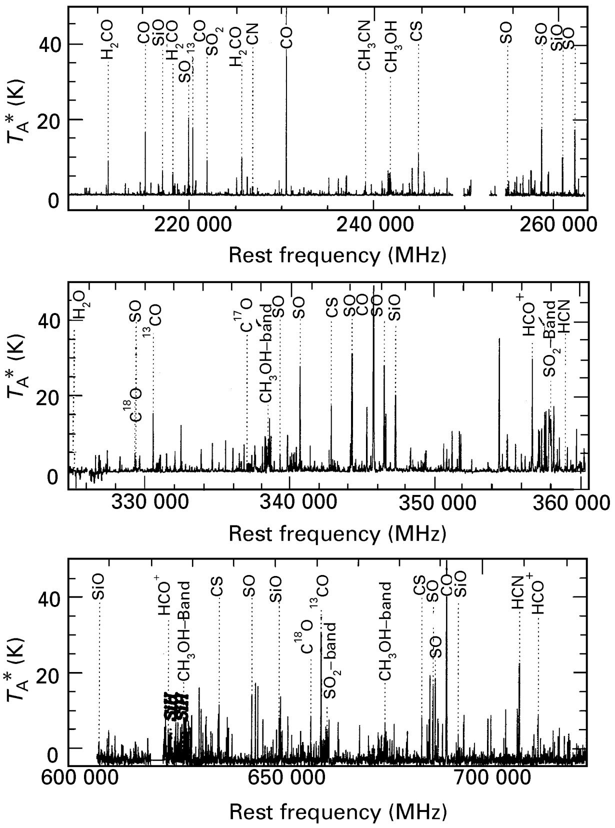



**Figure 2** Ground-based line surveys of a star-forming region in the Orion constellation in three atmospheric windows centred at 240, 330 and 700 GHz. Note the richness of the spectrum. Some of the more prominent lines are labelled by the molecule responsible. Courtesy of TG Philips.

**Table 1** Identified interstellar and circumstellar molecules

| | | | | |
|---|---|---|---|---|
| *Simple hydrides, oxides, sulfides, halogens and related molecules* | | | | |
| $H_2$ (IR,UV) | CO | $NH_3$ | CS | NaCl*[a] |
| HCl | SiO | $SiH_4$* (IR) | SiS | AlCl* |
| $H_2O$ | $SO_2$ | $C_2$ (IR) | $H_2S$ | KCl* |
| $N_2O$ | OCS | $CH_4$ (IR) | PN | AlF* |
| HF | | | | |
| *Nitriles and acetylene derivatives* | | | | |
| $C_3$ (IR,UV) | HCN | $CH_3CN$ | HNC | $C_2H_4$* (IR) |
| $C_5$* (IR) | $HC_3N$ | $CH_3C_3N$ | HNCO | $C_2H_2$ (IR) |
| $C_3O$ | $HC_5N$ | $CH_3C_5N$ ? | HNCS | |
| $C_3S$ | $HC_7N$ | $CH_3C_2H$ | HNCCC | |
| $C_4Si$* | $HC_9N$ | $CH_3C_4H$ | $CH_3NC$ | |
| | $HC_{11}N$ | $CH_3CH_2CN$ | HCCNC | |
| | $HC_2CHO$ | $CH_2CHCN$ | | |
| *Aldehydes, alcohols, ethers, ketones, amides and related molecules* | | | | |
| $H_2CO$ | $CH_3OH$ | HCOOH | $CH_2NH$ | $CH_2CC$ |
| $H_2CS$ | $CH_3CH_2OH$ | $HCOOCH_3$ | $CH_3NH_2$ | $CH_2CCC$ |
| $CH_3CHO$ | $CH_3SH$ | $(CH_3)_2O$ | $NH_2CN$ | |
| $NH_2CHO$ | $(CH_3)_2CO$ | $H_2CCO$ | $CH_3COOH$ | |
| *Cyclic molecules* | | | | |
| $C_3H_2$ | $SiC_2$ | c-$C_3H$ | $CH_2OCH_2$ | |
| *Molecular ions* | | | | |
| $CH^+$(VIS) | $HCO^+$ | $HCNH^+$ | $H_3O^+$ | $HN_2^+$ |
| $HCS^+$ | $HOCO^+$ | $HC_3NH^+$ | $HOC^+$ | $H_3^+$(IR) |
| $CO^+$ | $H_2COH^+$ | $SO^+$ | | |
| *Radicals* | | | | |
| OH | $C_2H$ | CN | $C_2O$ | $C_2S$ |
| CH | $C_3H$ | $C_3N$ | NO | NS |
| $CH_2$ | $C_4H$ | HCCN* | SO | SiC* |
| NH (UV) | $C_5H$ | $CH_2CN$ | HCO | SiN* |
| $NH_2$ | $C_6H$ | $CH_2N$ | MgNC | CP* |
| HNO | $C_7H$ | NaCN | MgCN | |
| $C_6H_2$ | $C_8H$ | $C_5N$* | | |

[a] Species denoted with * have only been detected in the circumstellar envelope of carbon-rich stars, notably IRC 10216. Most molecules have been detected at radio and millimetre wavelengths, unless otherwise indicated (IR, VIS or UV). Isotopomers of many of these species – including D, $^{13}C$, $^{15}N$, and $^{18}O$ – have been detected as well.

rather exotic, unsaturated carbon chain molecules, radicals and ions. To a large extent this reflects observational selection effects. The detection of these species is heavily favoured because of their large dipole moments and their relatively small partition function, which allows detection of even trace amounts. Also, while in terrestrial laboratories many of these species are very transient, they can persist much longer under the extreme vacuum conditions of the interstellar medium when properly shielded from stellar ultraviolet photons by the ubiquitous dust grains. Many of the ions are protonated versions of stable molecules such as water and carbon monoxide. Finally, isotopomers of many of these species have been discovered as well.

The intensity of rotational line emission depends on the density and temperature of the gas as well as the abundance of the molecular species. Because many lines of the rotational ladder of a species can be measured, these spectra provide excellent probes of the physical conditions of the emitting gas. Most of the molecules are present in so-called molecular clouds, which are dense for interstellar standards (1000 molecules $cm^{-3}$), cold (10 K), large scale (30 light-years) structures. About 40% of the interstellar gas is in such molecular clouds. The large column densities of dust associated with these molecular clouds shields the molecules from the dissociating radiation from surrounding bright stars and allows

them to survive. At the same time, these molecules allow the clouds to cool down to low temperatures and be compressed. Gravity can then overcome the thermal and magnetic support of these clouds and new stars can form. Molecular clouds are indeed the exclusive sites for the formation of new stars. The molecular inventory of these clouds is therefore directly fed into newly formed stars and their planetary systems.

The most abundant species in molecular clouds is $H_2$, simply because hydrogen is the most abundant element. Carbon monoxide is second at an abundance relative to hydrogen of $10^{-4}$. Most of the other species in **Table 1** are trace species at abundances of $10^{-8}$ or less. Many processes contribute to the formation and evolution of these molecules. Molecular hydrogen is thought to be formed on the surfaces of dust grains (mainly because gas-phase routes towards this molecule are measured to be extremely slow). Once molecular hydrogen is present, other molecules can be formed through gas-phase ion–molecule reactions. These reactions are fast and, unlike neutral–neutral reactions, often do not have reaction barriers because of the Coulomb interaction. A trace amount ($10^{-7}$) of ionization is provided by penetrating high energy (MeVs) cosmic ray

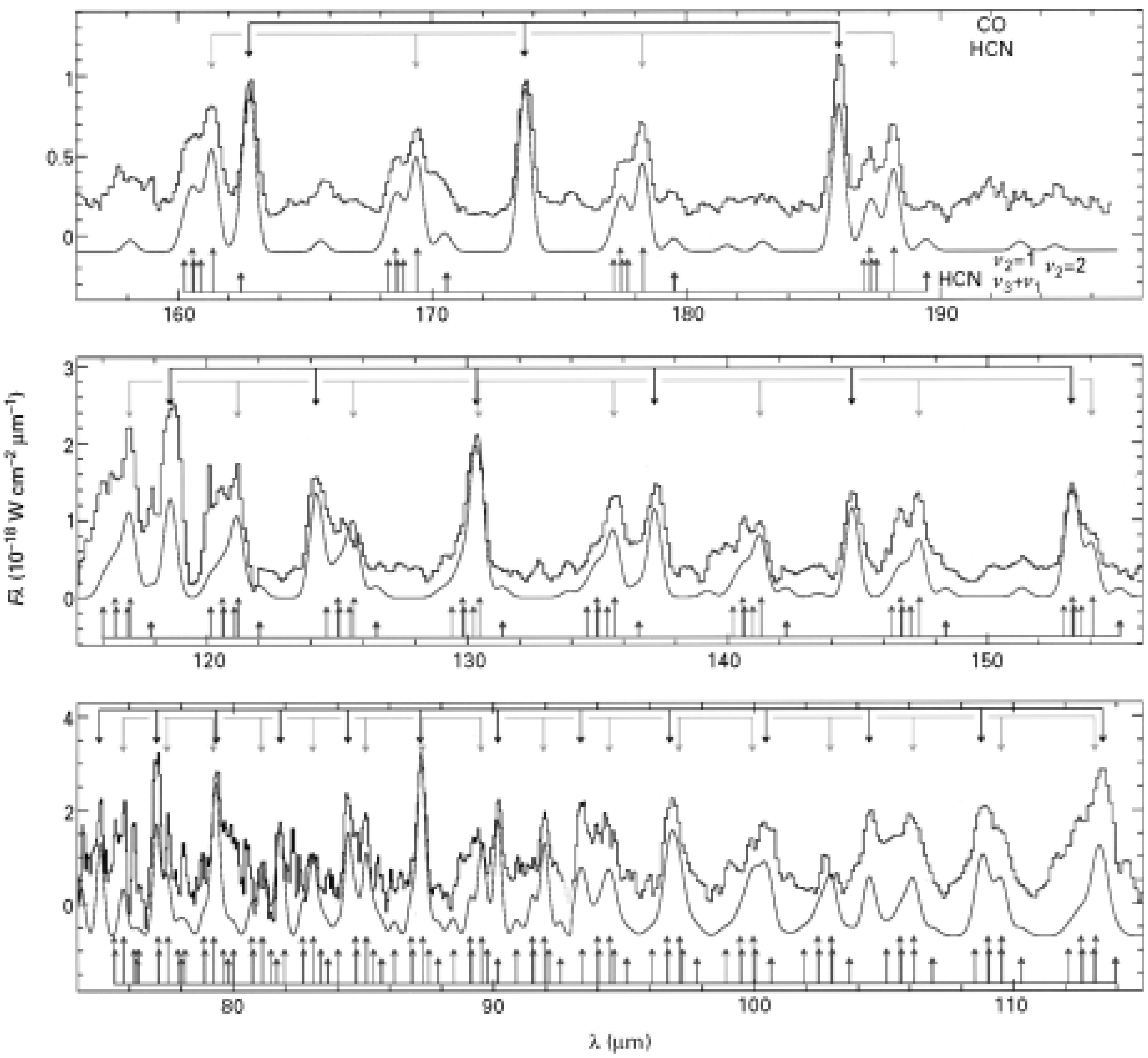



**Figure 3** Far-infrared spectrum of the ejecta from the star IRC 10216 obtained with the Long-Wavelength Spectrometer on board the Infrared Space Observatory at a resolution of about 100. Most of the lines in this spectrum can be identified with pure rotational transitions in CO, HCN and their isotopomers (top arrows). Rotational transitions associated with the lowest vibrational states of HCN are also present (bottom arrows). Reproduced with permission from Cernicharo J, Barlow MJ, Gonzaléz-Alfonso *et al* (1996) *Astronomy and Astrophysics* **315**: L201.

protons. The cosmic rays ionize $H_2$ producing $H_2^+$ which reacts with another $H_2$ to form protonated molecular hydrogen, $H_3^+$. This species can transfer its proton to other species, in particular atomic oxygen, to start the chain that will lead to the formation of OH, $H_2O$ and their protonated forms. Reaction of these species with atomic C and $C^+$ leads then eventually to the formation of CO. This very stable molecule locks up most of the available carbon, but trace amounts of atomic C and $C^+$ present can drive the build up of more complex carbon molecules including the carbon chains. Finally, the low temperatures (10 K) of molecular clouds will enhance isotopic fractionation. In particular, because of the zero-point energy difference between hydrogenated and deuterated molecules (typically 100–500 K), deuterium fractionation is important and the abundance of deuterated molecules can reach levels comparable to their hydrogenated counterparts despite the very low elemental abundance of deuterium ($10^{-5}$ relative to H).

## Vibrational spectroscopy

### Ices in star-forming regions

Gas and dust in molecular clouds are very cold and hence vibrational spectroscopy is limited to absorption spectroscopy, where a newly formed star inside the cloud or a chance superposition of the cloud against a background star provides the continuum against which the spectra can be measured. Such spectra show a variety of broad absorption features due to simple molecular species: $H_2O$, $CH_3OH$, CO, $CO_2$, OCS and $CH_4$. **Figure 4** shows as an example the 2–20 μm spectrum of the luminous protostar W33A. These species are very volatile and are therefore never seen in emission in the infrared. At this resolution, it is easy to separate absorption due to gaseous molecules from that of solid ices (**Figure 5**). **Table 2** summarizes the observed composition of interstellar ice towards this source. $H_2O$ dominates the column density of ice by a large factor. The carbon-bearing species, CO, $CO_2$ and $CH_3OH$, are less abundant by a factor 10 to 20, while $CH_4$ is at the 1% level compared to $H_2O$. This

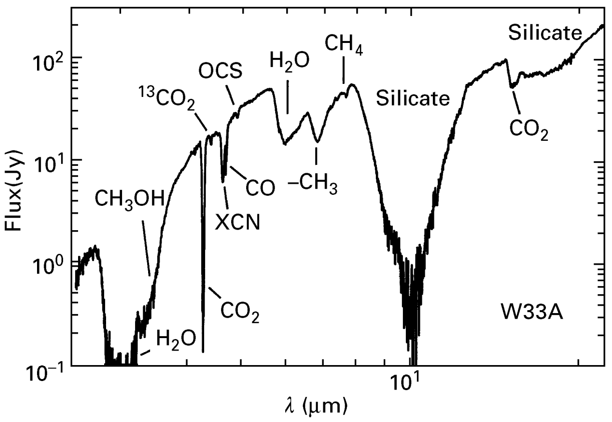



**Figure 4** The 2–20 μm spectrum of the protostar W33A obtained at a spectral resolution of 250 using the Short-Wavelength Spectrometer on board the Infrared Space Observatory. Except for the 10 μm silicate band, all features are due to simple molecules in ice mantles.

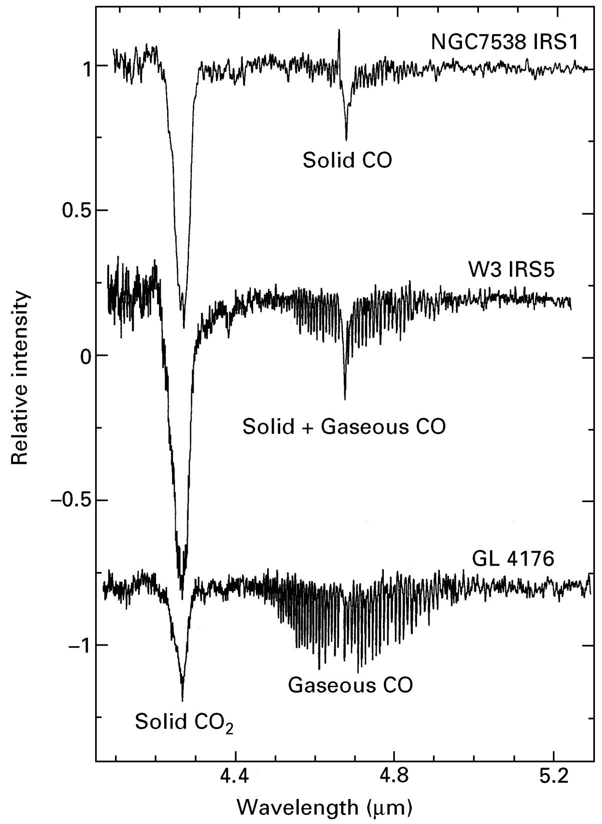



**Figure 5** Absorption spectra of a number of protostars in the stretching region of CO and $CO_2$ obtained by the Short-Wavelength Spectrometer on board the Infrared Space Observatory. The broad band at about 4.25 μm is due to the stretching mode of solid $CO_2$. The narrow features in the 4.5–4.9 μm range are the rovibrational transitions of gaseous CO. The broader feature in the Q-band gap of this linear molecule is the stretching mode of solid CO. The narrow emission feature at about 4.66 μm in the spectrum of NGC 7538 IRS 1 is a hydrogen recombination line. These different sources are characterized by different temperatures of the absorbing gas and ices. The relative amounts of the rather-volatile molecule CO in the gas phase and the ice reflect this rather strikingly. Reproduced with permission of the Royal Society of Chemistry from van Dishoeck EF (1998) *Faraday Discussions* **109**: 31.

**Table 2** Observed abundances in interstellar ices

| *Species* | *W33A* | *NGC 7538-IRS 9* | *Comets* |
|---|---|---|---|
| $H_2O$ | 100 | 100 | 100 |
| $NH_3$ | – | 12 | 1 |
| $CH_4$ | 0.4 | 2 | 0.2–1.2 |
| $CH_3OH$ | 4 | 5 | 0.3–5 |
| CO | 2 | 12 | 6 |
| $CO_2$ | 3 | 15 | 3 |
| OCS | 0.04 | – | 0.1 |
| HCOOH | 3 | 3 | 0.2 |

Observed abundances normalized to that of water ice (=100). W33A and NGC 7538-IRS 9 are two luminous protostars which span the observed range in interstellar ice composition. The abundances for the comets are an average of those observed for comets Hyakutake and Hale–Bopp.

reflect the timescale of this thermal processing; i.e. the lifetime of the newly formed star. The $CO_2$ stretching mode is not sensitive to these effects and neither the laboratory nor the interstellar spectra show much variation upon warm up. However, the stretching mode of the $^{13}C$ isotope of solid $CO_2$ does show these effects and the interstellar spectra are in good agreement with the conclusions derived from the bending mode of $^{12}CO_2$.

Interstellar ices are formed through the accretion of simple gas-phase species on surfaces of small dust grains injected into the interstellar medium by stars. Typical accretion rates are about one species a day.

dominance of $H_2O$ is directly evident from the complete IR spectrum of this source **(Figure 4)**, since the $H_2O$ absorption features are more conspicuous than any of the other ice features and yet the intrinsic strength of absorption features are very similar for most molecules.

The most striking aspect of the IR spectrum of interstellar ices is the apparent simplicity of the composition. Nevertheless, this global ice inventory is somewhat misleading. The peak position and profile of absorption bands are sensitive to the molecular environment of the absorbing species in the ice. Thus, $H_2O$ is mainly located in an $H_2O$-rich ice, as evidenced by the width and peak position of the 3 and 6 μm bands representing OH stretching and bending modes. While some of the solid CO is embedded in the $H_2O$ ice, most of the absorption is due to an apolar ice component dominated by $O_2$, $N_2$ or CO itself. Solid $CO_2$ studies also reveal the presence of multiple ice mantle components along many lines of sight. **Figure 6** shows the bending mode of $CO_2$ in the spectra of a number of protostars embedded in their natal molecular clouds. The spectra show weak structure in the depth of the band at 15.1 and 15.25 μm as well as a shoulder at about 15.4 μm. These spectra are well fitted by laboratory spectra of mixtures of $H_2O$, $CH_3OH$ and $CO_2$ heated to about 100 K. At about that temperature, these mixtures segregate into separate $H_2O$, $CH_3OH$ and $CO_2$ ices and this is reflected in the absorption spectra. Thus, the features at 15.1 and 15.25 μm are due to the bending mode in pure solid $CO_2$. This mode is doubly degenerate and splits when the axial symmetry of the molecule is broken in the $CO_2$ matrix. The observed peak positions are slightly shifted from those of a pure $CO_2$ ice film because, in these warmed-up mixtures, the $CO_2$ ice clusters are still embedded in the $H_2O$ and $CH_3OH$ ices. The strength of the 15.1 and 15.25 μm features in the interstellar spectra

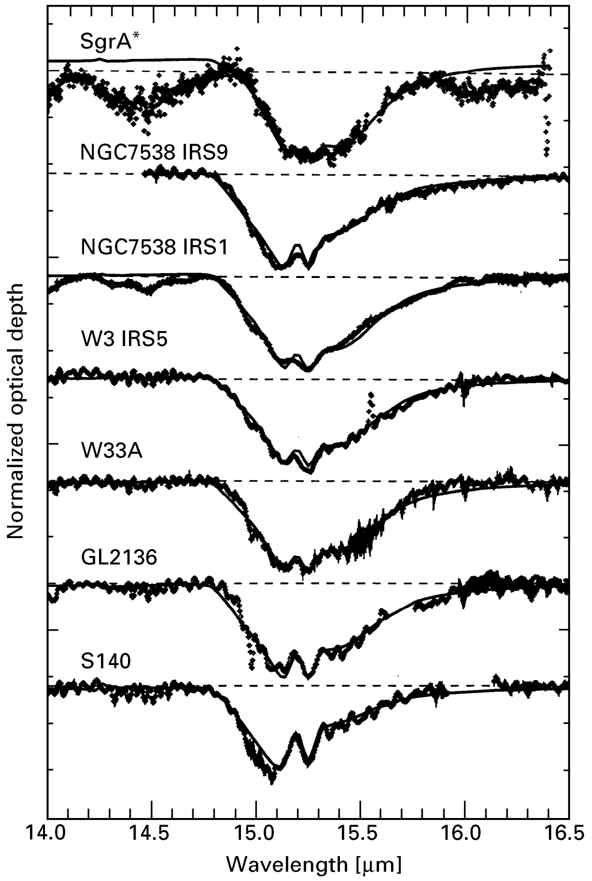



**Figure 6** Absorption spectra of the bending mode of $CO_2$ ice towards a number of protostars embedded in their natal molecular clouds obtained by the Short-Wavelength Spectrometer on board the Infrared Space Observatory at a spectral resolution of about 2000. The observed profiles show weak features at about 15.1 and 15.25 μm as well as a shoulder at about 15.4 μm. The solid lines are laboratory spectra of the $H_2O$:$CH_3OH$:$CO_2$ = 1:1:1 mixture deposited at 10 K and then warmed to temperatures in the range of 114 to 118 K. The sharp structure near 14.97 μm in some sources is due to the rovibrational transitions in the Q branch of gaseous $CO_2$ which at this resolution pile up in one spectral resolution element. The individual P and R branch lines are too weak to be observable at this resolution.

Interstellar grain surface chemistry is dominated by hydrogenation and oxidation reactions. Of particular interest are those reactions involving the dominant gas-phase molecule, CO, leading to $H_2CO$, $CH_3OH$ and $CO_2$. Atomic oxygen, carbon and nitrogen are readily hydrogenated to $H_2O$, $NH_3$ and $CH_4$. In the general interstellar medium, the omnipresent stellar radiation photodesorbs these species keeping the grain surfaces clean. However, the shielding provided by molecular clouds allows thin (10 s of nm) ice mantles to form. During the gravitational collapse associated with the formation of a new star, the accretion rate of gas-phase species on a grain becomes so high that essentially all species can freeze out in these ice mantles. Such ice grains entering the cold outskirts of protoplanetary disks can survive and agglomerate into large bodies, leading to the formation of comets. The great similarity in molecular inventory and abundances of interstellar ices and comets (cf. **Table 2**) supports such a scenario. Besides grain surface chemistry, the composition of interstellar ices is also affected by thermal processing when the ices are too close to a newborn star, much like the outgassing of comets when they approach the inner solar system. The difference in volatility between apolar molecules, such as CO and $CH_4$, and hydrogen-bonding species such as $H_2O$ and $CH_3OH$ can then lead to thermal fractionation.

### Interstellar and circumstellar PAHs

Ground-based, airborne and space-borne observations have shown that the IR spectra of bright sources with associated dust and gas are dominated by relatively broad emission features at 3.3, 6.2, 7.7, 8.6 and 11.3 μm, which always appear together. Because the carriers of these bands remained unidentified for almost a decade, these bands have become collectively known as the unidentified infrared (UIR) bands. These bands are now unequivocally identified with an aromatic carrier and this name is somewhat of a misnomer. Nevertheless, the abbreviation has stuck and is still widely used. As an example, **Figure 7** shows three mid-infrared spectra of carbon-rich outflows from stars in the latest stages of their evolution. It is now known that these emission features are not limited to stellar outflows but are also present in the spectra of disks of newly formed

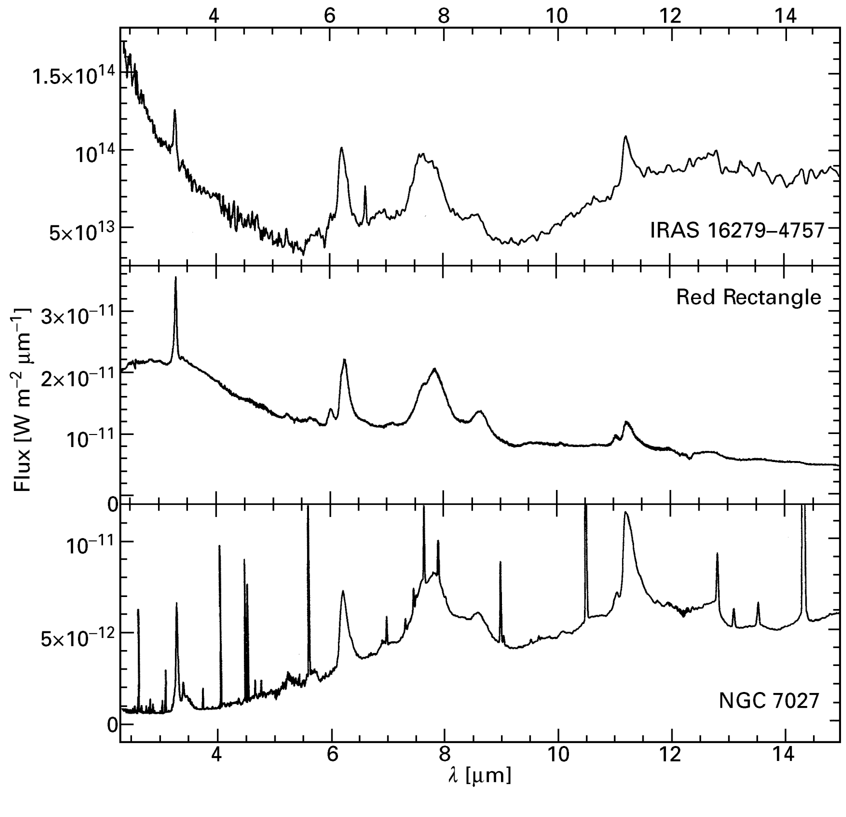



**Figure 7** The 3–15 μm emission spectra of three stars in the latest stages of their evolution illustrating the ubiquitous nature and the richness of the UIR spectrum. These spectra were obtained with the Short-Wavelength Spectrometer on board the Infrared Space Observatory at a spectral resolution ranging from 250 to 2000.

stars, the spectra of the surfaces of interstellar gas clouds, the spectra of reflection nebulae, the spectra of ionized nebulae around massive stars, and the spectra of the general interstellar medium of the Milky Way and other nearby galaxies. The carrier is thus ubiquitous and able to withstand the harsh environment of interstellar space.

These bands are very characteristic for aromatic hydrocarbon species in which carbon atoms are arranged in hexagons decorated by hydrogens at the edges. These features are observed far from illuminating sources where interstellar gases are cold (10–100 K) and exciting collisions infrequent (once a day). The emission is therefore attributed to infrared fluorescence from a collection of vibrationally excited PAH species electronically pumped through the absorption of one stellar UV photon (about 10 eV). During the vibrational cascade (about 1 s), the molecule emits in its C–H and C–C vibrational modes (the UIR bands) and, afterwards, the molecule remains cold (about 10 K) until the next UV photon is absorbed (about a day). Analysis of this process leads then to a determination of the size of the emitting species of about 50 carbon atoms. Thus the carriers of the UIR bands are PAH molecules such as hexabenzocoronene ($C_{48}H_{24}$) and circumcoronene ($C_{54}H_{18}$).

Typical abundances of PAHs are of the order of $10^{-7}$ relative to hydrogen, which makes these species the most abundant polyatomic molecules present in space. Vibrational spectroscopy is eminently suited for detecting the presence of classes of molecules, but it is much more difficult to identify specific molecules within a collection of species. However, while the gross characteristics of these stellar and interstellar spectra are very similar, when examined in detail, they vary from source to source. Analysis of these variations may well provide us with a tool to identify specific molecules within the circumstellar and interstellar PAH family and such efforts, supported by extensive laboratory studies, are now underway.

The 11–15 μm region of the spectrum is rich in details showing bands at 11.0, 11.3, 11.9, 12.8, 13.6, 14.2 and 14.8 μm plus a number of weaker shoulders (**Figure 8**). Emission features in this spectral region are due to a C–H out-of-plane deformation mode of aromatic hydrocarbons and the observed pattern is very characteristic for the edge structure of the emitting aromatic species. Due to coupling between the vibrating H atoms, the exact peak position of these modes depends on the number of directly adjacent H atoms (i.e. bonded to neighbouring C atoms on a ring). The 11.3 μm band is attributed to this mode for isolated Hs in PAHs while the 11.9 and 12.8 μm bands are due to 2s and

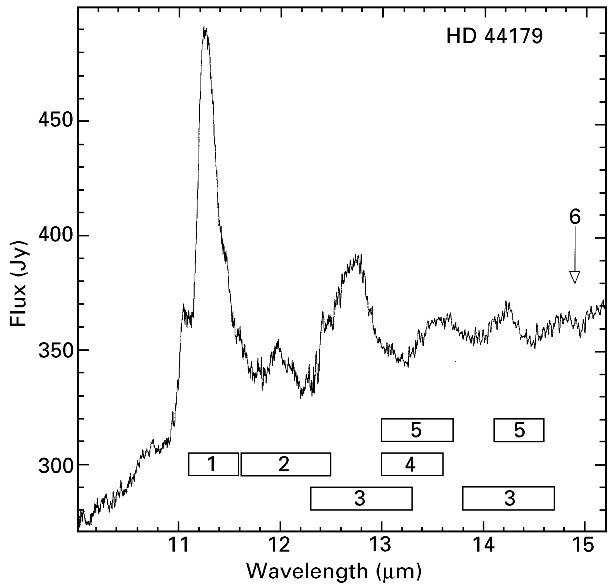



**Figure 8** The 10–15 μm spectrum of the outflow from a star known as the Red Rectangle because of the rectangular appearance of its associated reflection nebula on red plates. Bands in this spectral region are largely due to the C–H out-of-plane bending mode whose exact peak position is sensitive to the number of adjacent H atoms involved. The boxes underneath the spectra give the ranges where PAHs with the indicated number of adjacent H atoms emit. The arrow labelled 6 indicates the position for benzene. These spectra were obtained with the Short-Wavelength Spectrometer on board the Infrared Space Observatory at a spectral resolution ranging from 250 to 2000.

3s. The weak bands at 13.6 and 14.2 μm are likely due to phenyl (5s) groups; i.e. one ring connected by a single bond to the rest of the PAH. In that case, the 14.2 μm band is actually a ring-bending mode.

These PAHS are thought to form in carbon-rich stellar winds through chemical processes akin to those occurring in sooting flames and other terrestrial combustion environments. Acetylene is the precursor molecule from which larger species, and eventually soot, are formed. This process starts with the formation of the first aromatic ring (e.g. benzene). Rapid chemical growth of this ring through the addition of acetylene then forms larger PAHs. Further chemical growth as well as clustering of these PAHs leads to the formation of small (≈100 nm) soot particles. These soot particles and the gas are driven out by the radiation pressure of the luminous star, thereby enriching the interstellar medium.

## Electronic spectroscopy

### Elemental abundances in the galaxy

With a few notable exceptions, most atomic species possess strong absorption lines at far-ultraviolet

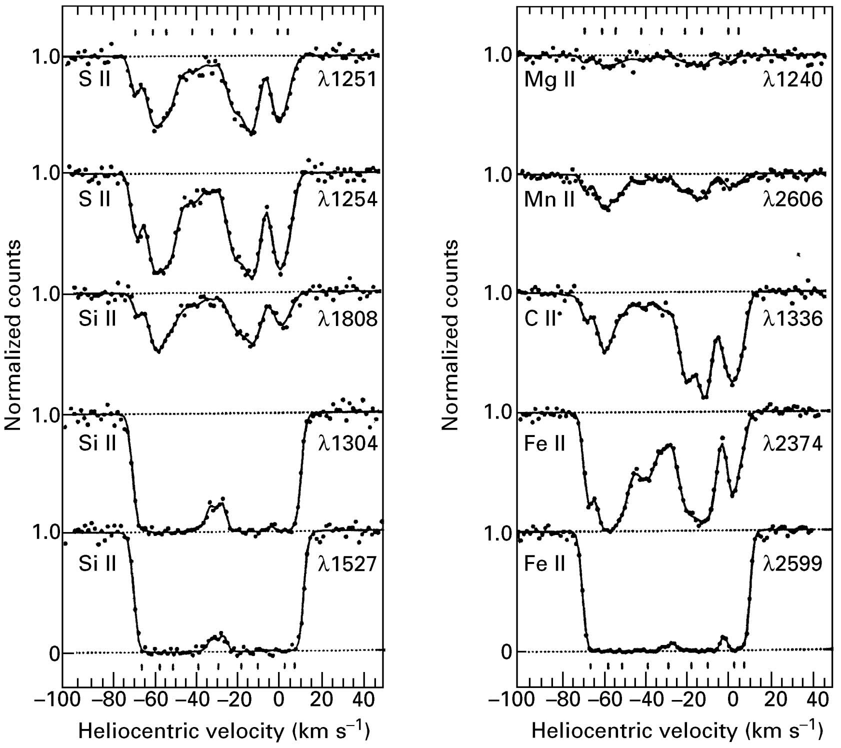



**Figure 9** Normalized intensity versus heliocentric velocity for a number of lines of different interstellar atoms towards the star HD 93521. These were obtained with the Goddard High-Resolution Spectrograph on the Hubble Space Telescope at a resolution of about 100 000. Tick marks at the top and bottom of the panels indicate individual absorption components present along the line of sight to this star.

wavelengths and their detection requires space-borne instrumentation. Since the 1970s, a number of missions have measured electronic absorptions of many elements as well as a few diatomic molecules. These lines are due to gas in the interstellar medium measured in absorption against a bright background star. A sample of such is shown in **Figure 9** obtained by the Goddard High-Resolution Spectrograph on the Hubble Space Telescope. At the high spectral resolution of this instrument, these absorption lines show components due to many individual clouds along the line of sight which have different velocities with respect to the Sun because of galactic rotation. The lines shown cover a range of line strengths, due to differences in intrinsic oscillator strength and column density of the absorbing species involved. These observations allow the determination of the abundance of elements in the interstellar medium with respect to atomic hydrogen. Typically, the measured abundances fall short of the elemental abundances measured in the Sun, in solar system objects such as meteorites, and in stellar photospheres (**Figure 10**). Because the Sun and other stars are formed from interstellar gas and, hence, should have very similar abundances, these large differences are attributed to the depletion of these elements from the gas phase in the form of small solid dust grains. This is supported by the loose correlation of the amount of material depleted from the gas phase with condensation temperature of the likely condensate of these elements (**Figure 10**). These dust grains are formed in the high-density, high-temperature environments of stars and injected into the interstellar medium in a supernova explosion or a stellar wind.

## The diffuse interstellar bands

The diffuse interstellar bands (DIBs) are weak absorption features which typically appear in the visual range of stellar spectra. These bands were discovered over 60 years ago and their origin in interstellar material along the line of sight towards these stars was quickly established, because they do not share the stellar velocity variations in spectroscopic binaries and their strength increases with distance and the column of absorbing interstellar material along the line of sight. **Figure 11** shows a compilation of these absorption features. The richness of the DIB spectrum is apparent: over 200 DIBs have been discovered. Most DIBs have a width in the range 0.5–3 Å,

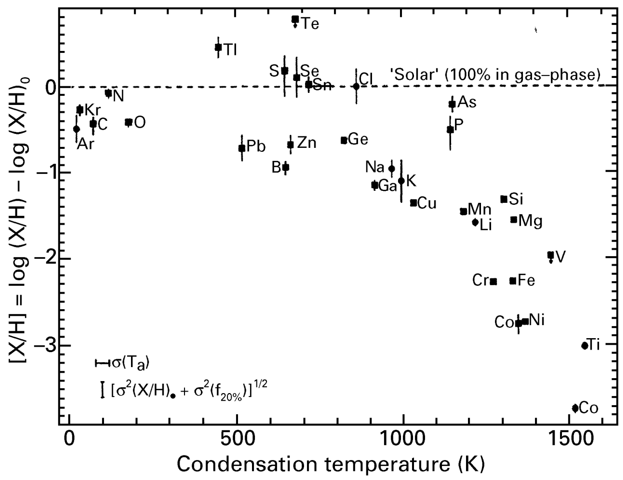



Figure 10 Gas-phase abundances measured towards the star z Oph of various elements relative to atomic hydrogen (X/H) and normalized with respect to solar abundances $(X/H)_0$ plotted versus condensation temperature of these elements. The condensation temperature is the temperature at which 50% of that element is depleted from the gas phase into the solid phase. The error bars on the squares indicate measurement errors only. Representative uncertainties in the solar abundances, oscillator strength and condensation temperature are indicated in the lower left hand corner.

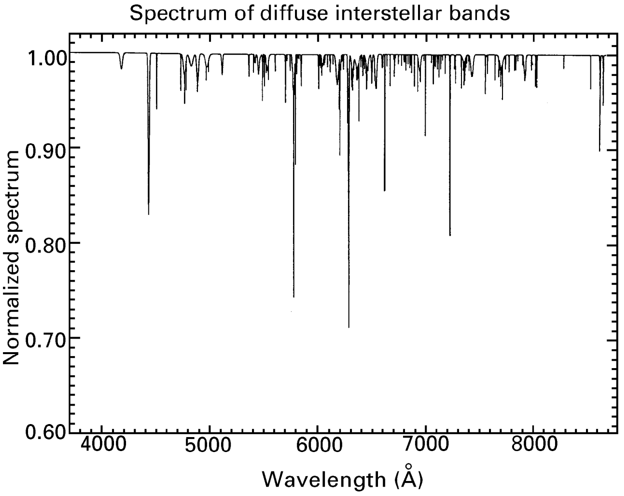



Figure 11 A compilation of the diffuse interstellar bands spectrum. Individual DIBs have been smoothed to a spectral resolution of 2 Å and their strength normalized to a column corresponding to line-of-sight dust reddening equal to unity. Reproduced with permission from Jenniskens P and Désert F-X (1994). *Astronomy and Astrophysics Supplement Series* **106**: 39.

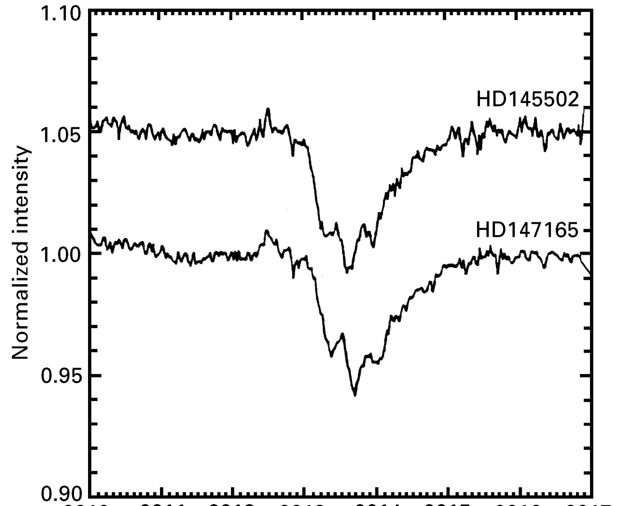



Figure 12 Line profile for the 16 613 Å DIB at a resolution of about 100 000 revealing a three-peaked structure characteristic for rotational contours of electronic transitions in molecular species. Reproduced with permission from Ehrenfreund P and Foing B (1996) *Astronomy and Astrophysics* **307**: L25

but bands as broad as 20 Å exist as well. The strengths of individual DIBs correlate reasonably well with each other but not perfectly, indicating that more than one carrier is involved. Some DIBs show substructure at high spectral resolution (**Figure 12**). In recent years, a few DIBs have been observed in the ultraviolet and far-ultraviolet part of the spectrum. But such studies are severely hampered by atmospheric absorption and by interstellar absorption due to solid dust grains in the interstellar medium, which increases rapidly with increasing frequency.

At present, none of the DIBs has been identified with certainty – not for lack of trying, though. The DIBs are too broad to be due to electronic transitions in atoms. For decades, the DIBs were thought to be associated with small (100 nm) dust grains which are responsible for the reddening and extinction of starlight. In that view, the DIBs were just thought to represent a small substructure in the extinction properties of the grains likely due to electronic transitions in impurities in an otherwise dielectric material such as silicates or oxides. However, the constant peak position and width of the DIBs in all directions argues against that since, theoretically, the profile is expected to be very sensitive to the grain size as well as the detailed composition of the matrix material. Also, these interstellar dust grains are observed to cause a small amount of polarization of starlight but the DIBs are unpolarized.

Instead, it is now generally accepted that the DIBs are due to absorption by interstellar gaseous molecules. Absorption by molecules (CH, CN) in the visible part of the spectrum was detected only shortly after the discovery of the DIBs, but these species are too small and have too low an abundance

to be responsible for the DIBs. This and the perceived difficulty to form an abundant concentration of molecules in the harsh environment of the interstellar medium has long driven the field away from a molecular carrier. However, presently, there is abundant evidence from other wavelength regions for the presence of large and complex molecules such as PAHs in the interstellar medium and this argument has lost much of its persuasion. There is also direct observational support for a molecular origin of the DIBs. One of the key pieces of evidence for this was found in the visible spectrum of the red rectangle. The spectra of the nebulosity associated with this object (known to be bright in PAH emission features in the infrared) show emission features near the wavelengths of prominent DIBs, including those at 5797 and 6613 Å, and, with increasing distance from the central illuminating star, their peak position shifts closer and closer to that of the interstellar DIB absorption features. These visual emission bands also become progressively narrower with distance from the star. This is the behaviour expected for fluorescence of electronic transitions in molecules excited by the radiation of the central star. The gas flowing away from the star cools and, consequently, fluorescence is dominated by lower and lower rotational lines. In the cold interstellar medium, absorption only occurs from the lowest rotational levels. A molecular origin is also supported by the substructure observed at high resolution in some DIBs, which is characteristic for rotational contours of electronic transitions in molecules (cf. **Figure 12**). Analysis of these line profiles shows that, depending on the class of species responsible (i.e. the rotational constants of the electronic states involved), the carrier contains between 10 and 60 C atoms. Finally, two new DIBs have been detected in the far-red (9577 and 9632 Å), which are close to laboratory measured absorption features of the fullerene cation, C , in a neon matrix. Unfortunately, no gas-phase spectra of this cation are presently available to confirm this identification unambiguously.

The carrier of the DIBs has still not been identified unequivocally. Part of this is, likely, that there is not one molecular carrier, or even one class of molecular carriers, responsible for all the DIBs, but rather a variety of species contribute. Among the likely candidates are PAHs, fullerenes and carbon chains (such as $C_7$). Vigorous studies on the visible absorption properties of such species are underway at various laboratories.

*See also:* **Cosmochemical Applications Using Mass Spectrometry; Microwave and Radiowave Spectroscopy, Applications; Rotational Spectroscopy, Theory; Solid State NMR, Rotational Resonance; Vibrational, Rotational and Raman Spectrocopy, Historical Perspective.**

## Further reading

Allamandola LJ, Tielens AGGM and Barker JR (1989) Interstellar polycyclic aromatic hydrocarbons: the infrared emission bands, the excitation/emission mechanism, and the astrophysical implications. *The Astrophysical Journal Supplement Series* **71**: 733–775.

Bakes ELO (1997) *The Astrochemical Evolution of the Interstellar Medium*. Vledder: Twin Press.

Kaler JB (1997) *Cosmic Clouds: Birth, Death, and Recycling in the Galaxy*. New York: Scientific American Library.

Latter WB, Radford SJE, Jewell PR, Mangum JG and Bally J (eds) (1997) *CO: Twenty-Five Years of Millimeter-Wave Spectroscopy*. Dordrecht: Kluwer.

Millar TJ and Williams DA (1993) *Dust and Chemistry in Astronomy*. Bristol: Institute of Physics.

Savage BD and Sembach KR (1996) Interstellar abundances from absorption line observations with the Hubble space telescope. *Annual Reviews of Astronomy and Astrophysics* **34**: 279–329.

Tielens AGGM and Snow TP (eds) (1995) *The Diffuse Interstellar Bands*. Dordrecht: Kluwer.

van Dishoeck EF (ed) (1997) *Molecules in Astrophysics*. Dordrecht: Kluwer.

# Ion Beam Analysis, Methods

*See* **High Energy Ion Beam Analysis.**

# Ion Collision Theory

**Anil K Shukla** and **Jean H Futrell**, University of Delaware, Newark, DE, USA

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Theory |

## Introduction

Collisions between an ion and neutral species result in a number of possible outcomes depending upon the chemical and physical properties of the two reactants, their relative velocities, and the impact parameter of their trajectories. These include elastic and inelastic scattering of the colliding particles, charge transfer (including dissociative charge transfer), atom abstraction, complex formation and dissociation of the colliding ion. Each of these reactions may be characterized in terms of their energy-dependent rate coefficients, cross sections and reaction kinematics. This article outlines a theoretical framework for discussing these processes that emphasizes simple models and classical mechanics. We divide the discussion of collision processes into the two categories of low-energy and high-energy collisions. Experiments under thermal or quasi-thermal conditions – swarms, drift tubes, chemical ionization and ion cyclotron resonance – are strongly influenced by long-range forces and often involve 'capture collisions' in which atom exchange and extensive energy exchange are common characteristics. High-energy collisions are typically impulsive, involve short-range intermolecular forces and are direct, fast process.

## Low-energy collisions

For historical reasons, and because the majority of investigations of ion–molecule reactions have involved quasi-thermal collisions of relatively low-mass ions and neutrals, we begin with the simplest case of a point charge interacting with a polarizable neutral. It is recognized that this Langevin model applies only at low collision energies where long-range forces dominate collision dynamics. Despite these restrictions, this simplistic picture provides a near-quantitative rationalization for the rates of thousands of ion–molecule reactions.

Ion–molecule collisions at low energies are dominated by the attractive long-range polarization force described in the simplest case by

$$V(r) = -\frac{\alpha q^2}{2r^4} \qquad [1]$$

where the ion is tacitly assumed to be a point charge, $q$, and the neutral is a point polarizable atom or molecule having a polarizability $\alpha$ at a distance $r$ from the ion. At relative velocities where repulsive forces can be neglected (that is, they are sensed at much closer distances than the impact parameter for orbiting collisions, as discussed below), Equation [1] is a plausible approximation. Conservation of angular momentum in a collision is imposed by expressing the effective potential of the ion–molecule system as the sum of the central potential energy $V(r)$ and the centrifugal potential energy:

$$V_{\text{eff}}(r) = -\frac{q^2\alpha}{2r^4} + \frac{L^2}{2\mu r^2} \qquad [2]$$

where $L$ is the classical orbital angular momentum of the two particles and $\mu$ is the reduced mass of the system. Here $L$ is given by $\mu vb$ where $b$ is the impact parameter and $v$ is relative velocity. The total relative energy of the system is given by

$$E_{\text{r}} = E_{\text{trans}}(r) + V_{\text{eff}}(r) \qquad [3]$$

where $E_{\text{trans}}(r)$ is the translational energy along the line of centres of the collision.

A plot of $V_{\text{eff}}(r)$ versus $r$ at constant $E_{\text{r}}$ for several values of the impact parameter is shown in **Figure 1** for the collisions of $N_2^+$ with $N_2$. For the special case in which the centrifugal barrier height equals $E_{\text{r}}$ (at $b = b_{\text{c}}$ in **Figure 1**), $E_{\text{trans}}(r) = 0$, and the particles will orbit the scattering centre with a constant separation distance $r_{\text{c}}$. For $b > b_{\text{c}}$ the distance of closest approach $\geq b_{\text{c}}/\sqrt{2}$ and for $b < b_{\text{c}}$ the particles spiral into small radii before they separate. The cross section for orbiting collisions is

$$\sigma(v) = \pi b_{\text{c}}^2(v) \qquad [4]$$

The rate coefficient corresponding to the orbiting or capture cross section is given by

$$k_{\text{c}} = v\sigma_{\text{c}} = v\pi q(2\alpha/E_{\text{r}})^{1/2} = 2\pi q(\alpha/\mu)^{1/2} \qquad [5]$$

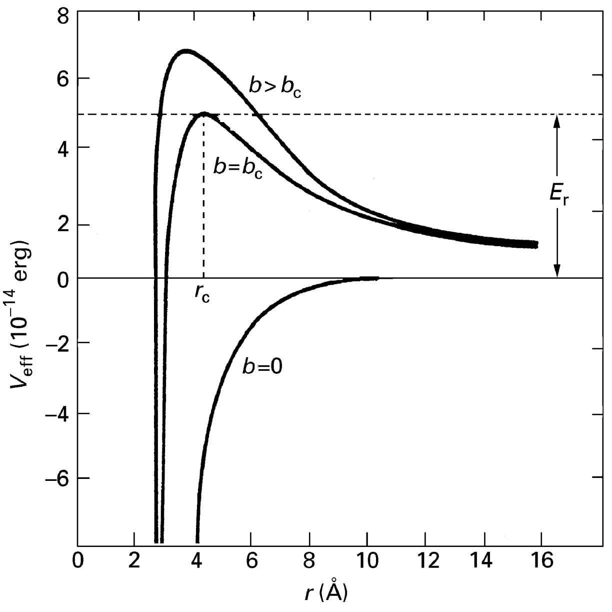



**Figure 1** Plot of effective potential, $V_{eff}$, versus $r$ for $N_2^+$ colliding with $N_2$ at a fixed energy of $E_r$ and several values of the impact parameter. Reproduced with permission of Academic Press from Su T and Bowers MT (1979) In Bowers MT (ed) *Gas Phase Ion Chemistry*, Vol. 1.

which is the commonly cited energy (velocity)-independent Langevin rate constant for ion–molecule reactions. Use of these expressions implicitly assumes there is a reactive cross section $\sigma_{rx}$ for real systems which is smaller than the capture cross section given by Equation [4].

The reaction cross section also sets a conceptual limit of relative velocity beyond which treatments of only long-range critical potentials cannot apply. For a hypothetical example of $\sigma_{rx} = 50$ Å$^2$, polarizability of 10 Å$^3$ and reduced mass of 50 amu, the capture cross section equals the $\sigma_{rx}$ at a centre-of-mass (CM) collision energy of approximately 0.7 eV. We may infer that collision energies of the order of 1 eV roughly define the boundary between low-energy collisions – e.g., reactions in high-pressure ion sources, drift tubes, ion cyclotron resonance and plasma swarm experiments – and higher-energy collisions, considered later. Considering these limitations, it is remarkable that this simplistic theory accurately correlates collision cross section and rate coefficients for hundreds of reactions of ions with nonpolar molecules.

## Ion–dipole and ion–quadrupole interactions

Over time, increasingly sophisticated modifications to the Langevin capture cross-section model have been developed to rationalize cross sections larger than Langevin and the temperature/energy dependences of cross sections. Considering first ion–dipole interactions, the classical expression for the effective potential is

$$V(r,\theta) = \frac{L^2}{2\mu r^2} - \frac{\alpha q^2}{2r^4} - \frac{\mu_D q}{r^2}\cos\theta \qquad [6]$$

where $\mu_D$ is the dipole moment of the molecule and θ is the angle between the dipole and the line of centres of the collision. The first such treatment by Hamill and co-workers assumed that the dipole 'locks in' on the ion, i.e. $\theta = 0$. This upper limit, locked-dipole approximation is obviously unrealistic for reactions at and above room temperature, for which rotation of the polar molecule tends to average out the dipolar potential.

Addressing this limitation, Su and Bowers developed the average dipole orientation (ADO) theory in which the orientation angle of the dipole with respect to the line of centres of collision is averaged as a function of ion–molecule separation distance. This concept leads to a 'locking constant', $c$, which is multiplied by the $\cos\theta$ term; this constant has been parametrized as a function of $\mu_D/\alpha^{1/2}$ for the range of temperatures that are of general interest in low-energy ion–molecule reactions. Further elaboration of this approach by Su and Bowers modified the treatment of angular momentum to formalize angle-averaged dipole orientation (AADO) theory. These workers also considered the interaction with a point charge of the quadrupole moment of molecules with $D_{\infty h}$ symmetry and formulated the average quadrupole orientation (AQO) theory for these types of long-range interaction.

Barker and Ridge developed a statistical model for ion–polar molecule collisions based on similar concepts. They considered the average interaction energy for a statistical ensemble of ion–neutral pairs as a function of their separation. The effective potential for such an ensemble is

$$V(r) = -\frac{\alpha q^2}{2r^4} - \frac{\mu q}{r^2}\mathscr{L}\left(\frac{\mu q}{r^2 k_B T_R}\right) + \frac{Eb^2}{r^2} \qquad [7]$$

where $k_B$ is the Boltzmann constant, $T_R$ is the rotational temperature of the neutral, $E$ is the relative translational energy, $b$ is the impact parameter and $\mathscr{L}(X)$ is the Langevin function defined as $\mathscr{L}(X) = \text{cotanh}(X) - 1/X$. The momentum transfer collision frequencies calculated from this model or from the

family of ADO theories described above are in reasonable agreement with experimental measurements of rate coefficients. They rationalize both the larger cross sections found for ion–molecule reactions involving highly polar neutrals and the increase in reaction rate coefficients for polar molecules at low temperatures.

## Quantum-mechanical approaches

Quantum-mechanical calculations of ion–dipole capture rate constants have been modelled successfully for low-temperature experiments and correlate well with the classical collision theory and capture cross sections just described. Clary has calculated rate constants of several ion–molecule reactions using the adiabatic capture centrifugal sudden approximation (ACCSA). Since the rotational motion of the neutral is strongly hindered by the ion at low velocity, he chose collinear reaction geometry, leading to a basis set that localizes the rotational motion of the neutral about this collinear configuration. The partial wave Hamiltonian for the entrance reaction channel is given by

$$H = -\frac{\hbar^2}{2\mu r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r + Bj^2 + \frac{|J-j|^2}{2\mu r^2} + V(r,\theta) \qquad [8]$$

where $\hbar$ is Planck's constant divided by $2\pi$, $\mu$ is the reduced mass, $B$ is the rotational constant, $j$ is the rotational angular momentum operator, $J$ is the total angular momentum operator, $V$ is the ion–molecule potential energy surface, $\theta$ is the orientation angle and $r$ is the position vector of the particle measured from the centre-of-mass. As an approximation, $|J-j|^2$ is replaced by the diagonal value $[J(J+1)\hbar^2+j^2-2\hbar^2\Omega^2]$, where $\Omega$ is the projection of both rotational and total angular momentum along the $Z$ axis (centre-of-mass vector). Thus for a fixed value of $r$, the Hamiltonian becomes

$$H = Bj^2 + \frac{1}{2\mu r^2}(j^2 - 2\Omega^2\hbar^2) + V(r,\theta) \qquad [9]$$

The reaction cross section is obtained from partial wave expansion and the capture approximation discussed earlier; that is, the reaction takes place if there is enough energy to surmount the centrifugal barrier. It follows that

$$\sigma(j) = \frac{\pi}{k_j^2(2j+1)} \sum_{J=0}^{J_{\max}(j,\Omega)} \sum_{\Omega=-\min(J,j)}^{\min(J,j)} (2J+1)P_j^{J\Omega} \qquad [10]$$

where $P_j^{J\Omega}$ is the reaction probability. When $J_{\max}(j,\Omega)$ is larger than $j$, the sum over $J$ in this equation gives the relationship

$$\sigma(j) = \frac{\pi}{k_j^2(2j+1)}\left[\left(\sum_{\Omega=-j}^{j}(J_{\max}(j,\Omega)+1)^2\right) - (2j+1)(j+1)^{(j/3)}\right] \qquad [11]$$

with the condition that $\sigma(j) = 0$ if the energy is below the centrifugal barrier. Rate constants calculated by this formalism are in excellent agreement with experimental room-temperature constants for such systems as positive and negative ion–molecule reactions of HCN with $H^-$, $D^-$ and $H_3^+$ ions.

Troe has used a statistical adiabatic channel model to calculate thermal ion–molecule capture rates using a pure long-range potential. The adiabatic channel potentials $V(r)$ are calculated from perturbation theory, and analytical expressions for channel threshold energies, activated complex partition functions and capture rate constants are obtained. His results are in good agreement with trajectory calculations at temperatures above 10 K for such systems as $H_3^+ + HCN$. Additionally, Sakimoto has developed a time-independent quantum-mechanical approach for collinear triatomic systems A+BC in the energy range from thermal to several eV. Using hyperspherical coordinates, the time-independent Schrödinger equation is solved numerically using the discrete variable representation algorithm for collision energies of 1–6 eV. Partial cross sections for reactive scattering and ion dissociation for $He + H_2^+$ (plus isotopes) were calculated that generally agree with limited experimental data for this system.

## High-energy ion–neutral collisions: Beam scattering

Low-energy ion–neutral collisions considered thus far often result in random angular scattering with extensive mixing of orbital and angular momentum. Beam experiments at the low-energy limit demonstrate that persistent complexes are formed but provide no further information on reaction mechanisms. In contrast, conservation of angular momentum in high-energy collisions leads to specific angular scattering, which can provide detailed information on reaction mechanisms. Deducing this information is a complex exercise and we begin with a description of elastic scattering. After these scattering characteristics have been described – both classically and quantum mechanically – we shall outline briefly how

reactive scattering is treated in this framework. A collision is better described by CM coordinates, which permit us to describe kinetic energy $T$ as the motion of the centre-of-mass of the system in the laboratory frame and the relative motion of the particles in the CM frame

$$T = \frac{1}{2}MV_{\mathrm{CM}}^2 + \frac{1}{2}\mu v^2 \quad [12]$$

where $M$ is the sum of the masses, $V_{\mathrm{CM}}$ is the velocity of the CM and the reduced mass and relative velocity have been defined previously. Since the motion of the CM is conserved in any collision, we can focus our attention on the relative motion in which the two particles are described as a single particle having mass $\mu$, angular momentum $\mu(\mathbf{r} \times \mathbf{v})$ and kinetic energy $\frac{1}{2}\mu v^2$.

For central potentials considered previously, the angular momentum is constant and the total CM energy at separation $r$ is

$$E = \frac{1}{2}\mu\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + V_{\mathrm{eff}}(r) \quad [13]$$

where the effective potential

$$V_{\mathrm{eff}}(r) = V(r) + \frac{L^2}{2\mu r^2} \quad [14]$$

includes the centrifugal potential $L^2/2\mu r^2$ described previously. The centrifugal potential is included in the equation for radial motion to account for angular motion of the system.

The observable results of a collision are the velocity of the particle after collision (deduced from its kinetic energy) and its direction of motion. For ordinary collisions between spherical particles, total kinetic energy must be conserved, so the initial and final velocities must be equal at distances large enough that $V(r)$ is negligible. **Figure 2** depicts how the deflection angle – the quantity of interest in these elastic and nonreactive collisions – relates to the interaction potential. Here a particle of reduced mass $\mu$ approaches parallel to the $x$ axis with kinetic energy $\frac{1}{2}\mu v^2$ and angular momentum $\mu vb$. When the effective potential energy $V_{\mathrm{eff}}$ equals the initial kinetic energy, the particle reaches its distance of closest approach, $r_c$, the turning point of the trajectory. The particle then recedes, and at a large distance from the scattering centre no longer senses the interaction potential. The change in direction of motion is the deflection angle $\theta$, and is measured from the negative $x$ axis in **Figure 2**. We note parenthetically that deflection through $+\theta$ and $-\theta$ are experimentally indistinguishable and that deflection by $\pm 2\pi N\theta$ (where $N$ is an integer) is indistinguishable from the deflection $\theta$.

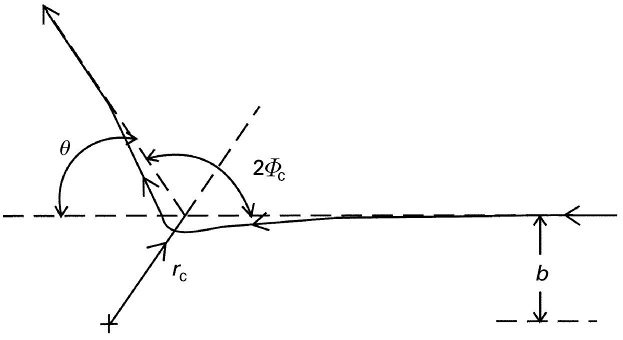



**Figure 2** Deflection of particles interacting through a potential $V(r)$. A single particle of reduced mass $\mu$ and velocity $v$ moves toward the stationary scattering centre with impact parameter $b$ and is scattered by a central potential $V(r)$ through an angle $\theta$. Reproduce with permission of John Wiley from Shirts RB (1986) In Futrell JH (eds) *Gaseous Ion Chemistry and Mass Spectrometry*.

**Figure 2** illustrates that the deflection is symmetric about the turning point; that is, the deflection of the particle in the outgoing trajectory equals its deflection in the approach trajectory for elastic scattering. Consequently, the polar scattering angle $2\phi_c$ and the final scattering angle $\theta$ is $\pi - 2\phi_c$ (see **Figure 2**). This leads to the expression for the deflection angle as a function of initial energy and impact parameter:

$$\phi(b, E) = \pi - 2b\int_{r_c}^{\infty} \frac{\mathrm{d}r}{r^2}\left(1 - \frac{V(r)}{E} - \frac{b^2}{r^2}\right)^{-1/2} \quad [15]$$

where $r_c$ is the outermost solution of $E = L^2/2\mu r_c^2 + V(r_c)$. Note that if $b = 0$ (head-on collision), then $\theta = \pi$ (unless $V = 0$). Further, as $b$ increases from zero, $\theta$ decreases from $\pi$, and $\theta$ decreases to zero as $b$ increases without limit.

A reasonable central force potential for ion–neutral elastic scattering is the Morse potential given by

$$V(r) = D(\mathrm{e}^{-2x} - 2\mathrm{e}^{-x}) \quad [16]$$

where $x = \alpha(r - r_e)$, with $\alpha$ specifying the curvature of the potential at its minimum. This potential has an attractive well of depth $D$ at the equilibrium

separation, $r = r_e$. For $r < r_e$, the repulsive part of the potential rises steeply to a high value, for $r > r_e$ the attractive part of the potential is dominant. Because the Morse potential goes to zero exponentially, it does not correctly describe the potential at large $r$. If necessary, it can be modified by adding a term for long-range polarization forces. It nicely describes the short-range behaviour and the balance between attractive and repulsive forces.

## Differential scattering

To illustrate the relationship between potential and scattering, **Figures 3** to **5** show respectively the effective potential, deflection function and scattering intensity for the Morse potential, (Eqn [16]), parameterized to describe scattering of low-energy protons by argon atoms. The potential for zero impact parameter (hence zero angular momentum, $L = 0$) in **Figure 3** has the parameters $r_e = 1.28$ Å, $D = 4.17$ eV and $\alpha = 1.85$ Å$^3$; the Morse potential with these parameters matches the experimentally determined scattering of protons by argon to better than 1% accuracy over the interval from 0 to 5 Å.

The curves in **Figure 3** are labelled with the angular momentum, $L$, in units of $\hbar$. At high $L$, the centrifugal potential overpowers the attractive well and the scattering is always repulsive. At $L = 75.95\hbar$ there is an inflection point at $V = 1.153$ eV, and for lower angular momenta collisions the effective potential has two regions of repulsion separated by an attractive well. The outermost root is the classical turning point for the potential and a plot of the turning point $r$ as a function of impact parameter (or, equivalently, of $L$) leads to a singularity where $\sqrt{c} = b$ for all collision energies. This corresponds to the point where the potential $V(r) = 0$ [~1 Å for the potential shown in **Figure 3**] and is a measure of the 'size' of the interacting particles, analogous to but different from a hard-sphere radius for the collision.

**Figure 4** shows the deflection function calculated from Equation [15] for the effective potential given in **Figure 3** for a number of collision energies. For zero impact parameter, the deflection angle is $\pi$, just as it would be for hard-sphere collisions, with positive scattering angle corresponding to repulsive collisions. The deflection angle goes through zero and becomes negative as the collision senses the attractive potential at longer distances. As $b$ increases, the deflection angle goes smoothly to zero. For moderate energy collisions, the centrifugal potential dominates scattering at all impact parameters. For very low energies, < 1.153 eV for the potential illustrated, there are three roots of the deflection function. (The reader is reminded that positive and negative deflections are experimentally equivalent.) The 'piling up'

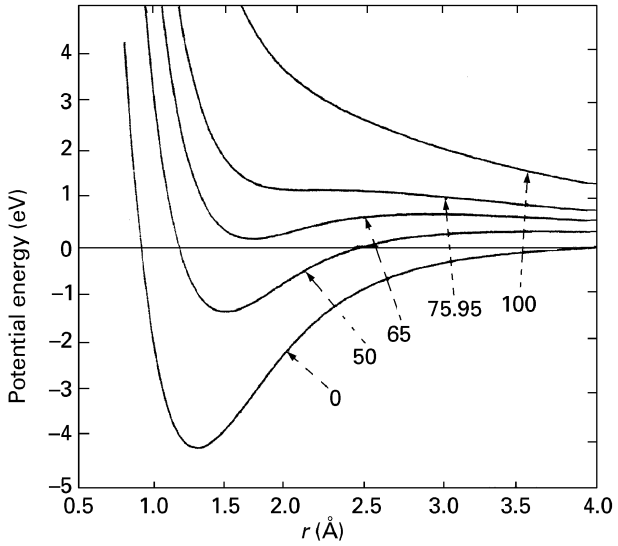



**Figure 3** Plot of effective potentials derived using the Morse potential for several values of angular momentum, $L$, as marked on each curve. See text for the parameters used for the Morse potential. Reproduced with permission of John Wiley from Shirts RB (1986) In Futrell JH (eds) *Gaseous Ion Chemistry and Mass Spectrometry.*

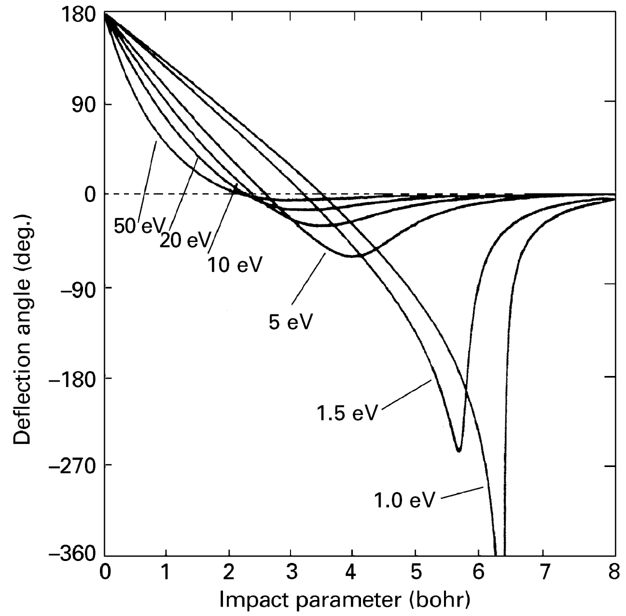



**Figure 4** The relationship between deflection function and impact parameter for interactions governed by the Morse potential of **Figure 3** at the indicated collision energies. 1 bohr = 0.529 Å. Reproduced with permission of John Wiley from Shirts RB (1986) In Futrell JH (eds) *Gaseous Ion Chemistry and Mass Spectrometry.*

of intensity at a particular scattering angle is called the rainbow angle. Measurement of rainbow scattering is an important experimental parameter for deducing the scattering potential. For specific $L$ values below this critical energy, two of the roots coincide at the maximum of the effective potential. For these trajectories, orbiting occurs and the deflection angle becomes infinite. This phenomenon is readily detected in careful beam scattering experiments; it corresponds to the capture cross section in our description of low-energy collisions.

**Figure 5** illustrates the connection of crossed-beam measurements of ion scattering with the interaction potential of the ion–neutral collision, again using the Morse potential. The quantity measured experimentally is the differential scattering cross section for a given CM collision energy. The differential cross section, $I(\theta)$, at a specified energy is given by

$$I(\theta)\,\mathrm{d}\Omega = \frac{I(b)}{I_0} = I(\theta)[2\pi \sin\theta\, \mathrm{d}\theta] = 2\pi b\, \mathrm{d}b \qquad [17]$$

$$I(\theta) = \frac{b}{|\sin\theta\,(\mathrm{d}\theta/\mathrm{d}b)|} \qquad [18]$$

where $\mathrm{d}\Omega$ is the differential of solid angle. The cross section calculated for the Morse potential shown in

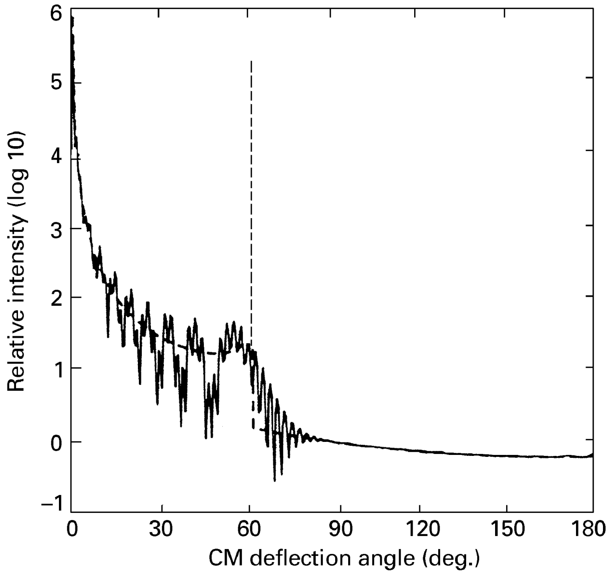



**Figure 5** Plot of scattering intensity (differential cross section) versus centre-of-mass scattering angle for interactions with Morse potential of **Figure 3** using classical treatment. Reproduced with permission of John Wiley from Shirts RB (1986) In Futrell JH (eds) *Gaseous Ion Chemistry and Mass Spectrometry*.

**Figure 5** illustrates for several energies the singularities determined by the two terms in the denominator of Equation [18]. The sin $\theta$ term gives infinity at CM angles of zero and $\pi$. The rainbow angle gives rise to a singularity when $\mathrm{d}\theta/\mathrm{d}b$ goes to zero; this occurs when several impact parameters give the same deflection angle. At higher collision energies this occurs at small angles; with decreasing collision energy the rainbow angle increases smoothly with decreasing energy and, for 1.5 eV is greater than 180°. Numerically it is 254° for this potential, leading to an observed scattering angle of 360° −254° = 106°. Two details distinguish this class of trajectories experiment from high energy collisions – namely, the singularity at 180° and the reversal of the 'bright side' of the rainbow. The backscattering of particles corresponding to the sin $\theta$ singularity at 180° is the experimental demonstration that particles are scattered by a potential sufficiently attractive at that collision energy to hold them together for at least half a revolution.

## Reactive encounters

We have explained in some detail the relationship of measured variables in elastic scattering to the interaction potential between the ion and neutral. This framework is the starting point for describing inelastic and reactive scattering. Inelastic scattering may be introduced as an instantaneous conversion of kinetic energy of motion into internal energy of either or both of the collision partners, while reactive collisions involve both energy exchange and reaction. It is plausible that these physical and chemical conversions occur at or near the distance of closest approach in the respective trajectories. This breaks the symmetry of the elastic scattering formalism and the retreat trajectory of products involves both a different energy inventory in the collision partners and a change in the potential governing the trajectory as they separate from the collision centre. This approach has been reasonably successful in rationalizing the differential scattering cross sections of a variety of proton transfer reactions. The time-dependent quantum-mechanical treatment of collisions described in the next section automatically takes into account energy transfer and reaction – at least in principle.

## Newton diagrams

As noted previously, describing collision processes in the CM frame is more informative than in the laboratory (LAB) frame. Explicitly the motion of the CM is conserved in collisions and is unavailable to the

reactants. The proper framework for displaying collisions is the CM relative velocity diagram, frequently called the Newton diagram. The collision process (for simplicity assumed to occur at right angles) of

$$M_1^+ + M_2 \rightarrow M_3^+ + M_4 \quad [19]$$

with $M_1^+$ and $M_2$, having velocities of $V_1$ and $V_2$, is depicted in **Figure 6**. The collision centre is the origin of the laboratory reference frame. In the CM frame, $M_1^+$ and $M_2$ move collinearly towards each other with velocities $U_1$ and $U_2$ and collide at the CM located on their relative velocity vector at the point defined by conservation of momentum such that $M_1U_1 = M_2U_2$. After collision, the two particles recoil from the CM in opposite direction with velocities $U_1'$and $U_2'$. The velocity of the CM is given by

$$V_{\mathrm{CM}} = \frac{M_1 V_1}{M_1 + M_2} + \frac{M_2 V_2}{M_1 + M_2} \quad [20]$$

and is constant. The relative kinetic energy of the collision partners, $E_{\mathrm{CM}}$, is available for the reaction and is given by

$$E_{\mathrm{CM}} = \frac{1}{2} M_1 U_1^2 \left(1 + \frac{M_1}{M_2}\right) \quad [21]$$

Similarly, the postcollision relative kinetic energy is given by

$$E'_{\mathrm{CM}} = \frac{1}{2} M_1 U_1'^2 \left(1 + \frac{M_1}{M_2}\right) \quad [22]$$

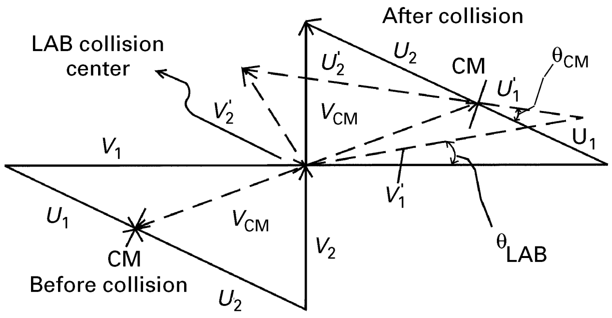



**Figure 6** Newton diagram showing pre- and postcollision velocity vectors in the laboratory and centre-of-mass reference frames for the collision of $M_1^+$ and $M_2$ having velocities of $V_1$ and $V_2$.

The difference $\Delta T = E'_{\mathrm{CM}} - E_{\mathrm{CM}}$, the translational exoergicity of the collision process, is a direct measure of the energy used in the process. Consequently, experimental measurements and transformation of mass, intensity, energy and angular distributions of product ions as Newton diagrams provides quantitative information on the energetics and dynamics of the collision process.

Two limiting cases of ion–molecule reaction dynamics will serve to illustrate the utility of Newton diagrams. If a collision between an ion and a molecule proceeds via an intermediate whose lifetime is larger than several rotational periods (~$10^{-12}$ s), internal energy redistribution is usually complete and the complex loses its memory of the reactant velocity vectors. The products will separate with equal probability on both sides of a plane passing through the CM and normal to the collision axis, giving rise to the forward–backward symmetry. Such orbiting complexes are usually formed at relative kinetic energies that are of the order of or less than the well depth of the interaction potential governing the collision.

As relative kinetic energy increases, the lifetime of the intermediate complex decreases and there is a transition to a direct reaction. The interaction time is too short for the intermediate to rotate and the Newton diagram contour plot becomes asymmetric with respect to the CM. The dependence of reaction probability on impact parameter and the effective potential for the approach and retreat trajectories determine the scattering angle and the form of the Newton diagram.

## Semiclassical theory of collisions, the WKB phase shift

In three dimensions and for central potentials, the Schrödinger equation separates in spherical coordinates analogously to corresponding classical equations of motion. The incoming particle is described by a wavefunction with a local wavenumber $k(r)$. Repulsive potentials decrease the velocity of the wave packet and the resulting scattered wavefunction has fewer nodes, resulting in a negative phase shift. Conversely, attractive potentials result in an increase in velocity; the wavefunction acquires more nodes and this results in a positive phase shift. The overall phase shift for scattered particles is

$$\delta(E, b) = \int_{\text{actual path}} k(r)\,\mathrm{d}r - \int_{\text{unscattered path}} k_0(r)\,\mathrm{d}r \quad [23]$$

Integrating this expression from the distance of closest approach to infinity (this gives half the total difference in phase for the entire trajectory; this is the convention for phase shift) gives

$$\delta(E,b) = \frac{\mu v}{h}\lim_{R\to\infty}\left[\int_{r_c}^{R}\left(1-\frac{V(r)}{E}-\frac{b^2}{r^2}\right)^{1/2} \mathrm{d}r - \int_{b}^{R}\left(1-\frac{b^2}{r^2}\right)^{1/2}\mathrm{d}r\right] \quad [24]$$

For both integrals, the lower limit of integration is the (outermost) zero of the associated integrand. Since both integrals are divergent, $\delta$ depends on cancellation between the two. This difficulty is repaired by subtracting $\int_{r_c}^{R}\mathrm{d}r$ from both terms:

$$\delta(E,b) = \frac{\mu v}{\hbar}\left[\int_{r_c}^{\infty}\left(1-\frac{v(r)}{E}-\frac{b^2}{r^2}\right)^{1/2} - 1\right]\mathrm{d}r - r_c + \frac{\pi b}{2} \quad [25]$$

This may be recast into a form incorporating the definition of phase shift from Equation [23]:

$$\delta(E,b) = \int_{r_c}^{\infty}(k(r)-k)\,\mathrm{d}r - k\left(r_c - \frac{\pi b}{2}\right) \quad [26]$$

where the integral term gives the phase difference resulting from the interaction potential and a correction for the difference in turning points of the unscattered and scattered particle includes the effect of the centrifugal potential.

Finally, it can be shown that the phase shift method leads to expressions for the collision time (different between $V(n) \neq 0$ and $V(n) = 0$) and for scattering angle.

$$\tau(E,b) = 2\hbar\left(\frac{\partial\delta(E,b)}{\partial E}\right)_b \quad [27]$$

$$\theta(E,b) = \frac{2\hbar}{\mu\nu}\left(\frac{\partial\delta(E,b)}{\partial E}\right)_E \quad [28]$$

This makes the precise connection we seek between WKB scattering theory and the classical model developed previously.

## Quantum scattering

Moving from a semiclassical towards a full quantum treatment we note that, at large distances, the wavefunction for a scattered particle is a combination of an incoming plane wave and a scattered wave:

$$\psi(r,\theta) = \mathrm{e}^{\mathrm{i}kz} + \frac{f(\theta)}{r}\mathrm{e}^{\mathrm{i}kr} \quad [29]$$

where the $z$ direction is the direction of travel. The first term describes a plane wave and the second term is an outgoing spherical wave whose amplitude is $f(\theta)$. This simple form of the wavefunction is valid for $r$ large enough that no interference between the incoming and outgoing waves is important. This applies experimentally to a collimated beam of finite width.

Since the wavefunction in Equation [29] obeys the Schrödinger equation, it can be expanded in any complete set of orthogonal functions, for example Legendre polynomials:

$$\psi(r\ \theta) = \sum_{l=0}^{\infty}\frac{\psi_2(r)}{kr}P_l(\cos\theta)(2l+1)i^l \quad [30]$$

where $P_l$ is Legendre polynomial. Multiplying by $P_{l'}(\cos\theta)$, and integrating over θ we obtain for large values of $r$ the asymptotic result

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}r^2}+k^2\right)\psi_l(r) = 0 \quad [31]$$

which has solutions $\psi_1(r) = a_1\sin(kr - l\pi/2 + \delta_l)$. Here the phase shift, $\delta_l$, for the $l$th partial wave contains all the information about the potential, and the term $l\pi/2$ comprises the centrifugal potential. The expression for scattering amplitude is obtained as

$$f(\theta) = \sum_{l=0}^{\infty}\frac{2l+1}{k}\mathrm{e}^{\mathrm{i}\delta_1}\sin\delta_l\,P_l(\cos\theta) \quad [32]$$

The differential cross skection $I(\theta)$ is the probability density for observing a particle at an angle $\theta, |f(\theta)|^2$ and the total cross section is obtained by integration

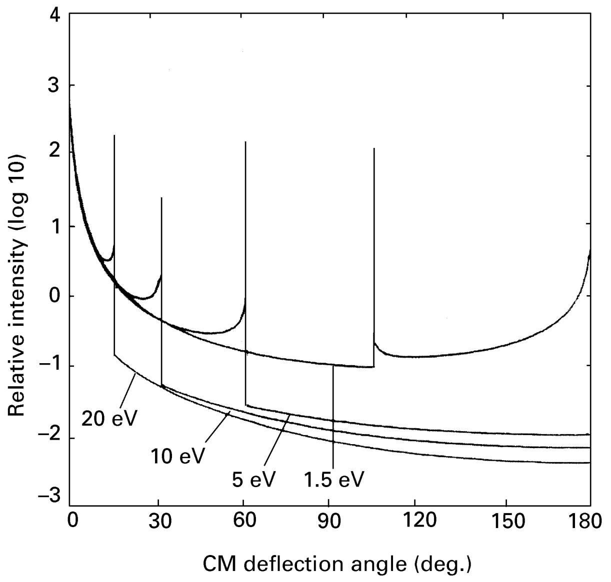



**Figure 7** Plot of scattering intensity versus centre-of-mass scattering angle at 5 eV collision energy using quantum-mechanical treatment. The calculations were performed using 500 phase shifts and the Morse potential of **Figure 3**. Reproduced with permission of John Wiley from Shirts RB (1986) In Futrell JH (ed) *Gaseous Ion Chemistry and Mass Spectrometry.*

of $I(\theta)$ over all angles as given below.

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l \qquad [33]$$

The phase shifts for calculating scattering as a function of angle are obtained by starting at $r = 0$ and increasing $r$ until $V(r)$ becomes negligible.

**Figure 7** illustrates the differential cross section for the Morse potential (Eqn [17]) calculated using Equation [32] and 500 phase shifts that were calculated using Equation [26]. The relative energy is 5 eV. The classical differential cross section from **Figure 5** is also shown. The quantum-mechanical form of $I(\theta)$ removes the classical singularity. The additional peaks occurring at angles smaller than the rainbow angle are interferences from the three different values of the impact parameter that contribute to scattering into these angles. The quantum result tends to oscillate about the classical mechanical result, which cannot recover the interferences of the quantum result.

For noncentral potentials – e.g. molecular scattering – the Schrödinger equation does not separate into partial waves. The solution can still be expanded in partial waves, but then the individual partial-wave radial Schrödinger equations will have additional terms that couple them together. Sophisticated computer programs are now available to solve these complex equations and calculate the detailed dynamics of a collision process.

## List of symbols

$b$ = impact parameter; $B$ = rotational constant; $c$ = locking constant; $D$ = depth of attractive potential well; $E$ = energy; $f()$ = function; $H$ = Hamiltonian; $I(\theta)$ = differential cross section; $j$ = rotational angular momentum operator; $J$ = total angular momentum operator; $k$ = rate coefficient; local wavenumber; $k_B$ = Boltzmann constant; $\mathscr{L}$ = Langevin function; $L$ = classical orbital angular momentum; $P$ = Legendre polynomial reaction probability; $q$ = ionic charge; $r$ = distance; magnitude of position vector; polar coordinate; $r$ = position vector; $r_e$ = equilibrium distance; $T$ = kinetic energy; $T_R$ = rotational temperature; $v$ = relative velocity vector; $v$ = magnitude of relative velocity vector; $V(r)$ = central potential energy; $V$, $V'$, $U$, $U'$ = velocity in centre-of-mass frame; $V_{CM}$ = velocity of centre-of-mass; $\alpha$ = atomic/molecular polarizability; curvature of potential energy curve at minimum; $\delta$ = phase shift on scattering; $\theta$ = angle between dipole and line of centres; final scattering angle; polar coordinate; $\mu$ = reduced mass; $\mu_D$ = dipole moment; $\sigma$ = collision cross section; $\tau$ = collision time; $\varphi_c$ = polar scattering angle; $\psi$ = wavefunction; $\Omega$ = projection of rotational and total angular momentum on $Z$ axis; solid angle.

## Subscripts

c = capture; c = closest approach; CM = centre of mass; e = equilibrium; eff = effective; r = relative; trans = translational

*See also:* **Ionization Theory; Quadrupoles, Use of in Mass Spectrometry.**

## Further reading

Chesnavich WJ, Su T and Bowers MT (1979) Ion-dipole collisions: recent theoretical advances. In Ausloos P (ed) *Kinetics of Ion Molecule Reactions,* pp. 31–53. New York: Plenum.

Child MS (1973) *Molecular Collision Theory.* New York: Academic Press.

Fluendy MAD and Lawley KP (1973) *Chemical Applications of Molecular Beam Scattering.* London: Chapman and Hall.

Franklin JL (ed) (1972) *Ion Molecule Reactions.* New York: Plenum Press.

Futrell JH (ed) (1986) *Gaseous Ion Chemistry and Mass Spectrometry.* New York: Wiley.

Levine RD and Bernstein RB (1987) *Molecular Reaction Dynamics and Chemical Reactivity*. Oxford: Oxford University Press.
McDaniel EW (1964) *Collision Phenomena in Ionized Gases*. New York: Wiley.
McDaniel EW, Cermak V, Dalgarno A, Ferguson EE and Friedman L (1970) *Ion–Molecule Reactions*. New York: Wiley-Interscience.
Shirts RB (1986) Collision theory and reaction dynamics. In Futrell JH (ed) *Gaseous Ion Chemistry and Mass Spectrometry*, pp. 25–57.New York: Wiley.
Steinfeld JI, Francisco JS and Hase WL (1989) *Chemical Kinetics and Dynamics*. Englewood Cliffs, NJ: Prentice-Hall.
Su T and Bowers MT (1979) Classical ion-molecule collision theory. In Bowers MT (ed) *Gas Phase Ion Chemistry*, Vol. 1, pp. 83–118. New York: Academic Press.
Vestal ML, Wahrhaftig AL and Futrell JH (1976) Application of a modified elastic spectator model to proton transfer reactions in polyatomic system. *Journal of Physical Chemistry*, **80**: 2892–2899.

# Ion Dissociation Kinetics, Mass Spectrometry

**Bernard Leyh**, F.N.R.S. and Université de Liège, Belgium

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Theory |

The ionization process taking place in the source of a mass spectrometer or the activation step of a tandem mass spectrometry experiment leads to ions with a range of internal energies extending usually well above the first dissociation threshold, i.e. the molecular ion has enough energy to dissociate. Dissociation is a dynamic process characterized by its associated lifetime. The whole story of kinetics in general and of ion dissociation kinetics in particular is to measure how fast a dissociation proceeds under given conditions (state selection, well-defined energy or temperature, for example) and to determine the reaction mechanisms. Both aspects will be addressed.

## Unimolecular rate constant

Dissociation kinetics of state-selected reactant ions has been investigated only in a limited number of relatively small systems. However, a large amount of data are now available on ionic reactions investigated with internal energy selection.

The dissociation of an energy-selected molecular ion is a unimolecular process, the rate of which, $R(E,t)$ is defined by the following equation, where $N_0$ is the initial number of ions formed:

$$\mathrm{ABC^+}\ (\text{internal energy} = \mathrm{E}) \rightarrow \mathrm{AB^+} + \mathrm{C}$$

$$\begin{aligned} R(E,t) &= \frac{\mathrm{d}N_{\mathrm{AB^+}}}{\mathrm{d}t} = k_{\mathrm{AB^+}}(E)N_{\mathrm{ABC^+}}(E,t) \\ &= k_{\mathrm{AB^+}}(E)N_0 \exp\Big(-k_{\mathrm{AB^+}}(E)t\Big) \end{aligned}$$

This represents the most simple situation where a single dissociation process occurs for ions possessing a single total energy, $E$. $k(E)$ is called the unimolecular rate constant. However, in practice, the parent ion can be formed with a broad distribution of internal energies, $P(E)$, and competitive reactions can take place. If the different competitive channels are characterized by rate constants denoted $k_i(E)$, then the following equation holds for the rate of production of the $j$th fragment ion:

$$R_j(E,t) = \frac{\mathrm{d}N_j}{\mathrm{d}t} = N_0 \int_E \mathrm{d}E\, k_j(E)P(E)\exp\left(-\sum_i k_i(E)t\right)$$

If further dissociations occur, the rates must be calculated by solving the set of differential equations resulting from the appropriate kinetic scheme.

If the dissociating ion is characterized by a welldefined internal energy, i.e. if $P(E)$ reduces to a single value, then each ion of the ionized sample can be considered to be one of the replicas of a microcanonical ensemble. For this reason, $k(E)$ is often referred to as the ‘microcanonical rate constant’. If, however, $P(E)$

is a thermal Boltzmann distribution characterized by a temperature $T$, then the corresponding rate constant, $k(T)$ is called the 'canonical rate constant'.

One of the most widespread approaches to model the experimental microcanonical rate constant is the RRKM-QET statistical theory, which postulates that a rapid internal energy randomization takes place before dissociation and that a transition state can be defined. This leads to the following well-known equation:

$$k(E) = \frac{\sigma N^{\ddagger}(E - E_0)}{h\rho(E)}$$

where $\sigma$ is the degeneracy of the reaction path (e.g. $\sigma = 4$ for the loss of H from $CH_4$), $N^{\ddagger}(E - E_0)$ is the number of states of the transition state up to an energy $E - E_0$ above the critical energy $E_0$ and $\rho(E)$ is the density of states of the parent ion at the internal energy $E$.

If $k(E)$ is averaged over a thermal internal energy distribution at temperature $T$, the equation of transition state theory for the canonical rate constant is recovered. It can be expressed as a function of the activation entropy, $\Delta S^{\circ\ddagger}$, and of the activation enthalpy, $\Delta H^{\circ\ddagger}$.

$$k(T) = \frac{k_B T}{h} \exp\left(\frac{\Delta S^{\circ\ddagger}}{R}\right) \exp\left(-\frac{\Delta H^{\circ\ddagger}}{RT}\right)$$

$k_B$ and $h$ are respectively Boltzmann's and Planck's constants. The value of the activation entropy gives insight into the looseness or the tightness of the transition state. Loose transition states are usually associated with simple bond cleavages, whereas rearrangement processes are expected to involve tighter transition states. Competitive processes, multistep fragmentations, nonadiabatic processes and tunnelling effects affect the variation of the rate constant with internal energy, which then forms a diagnostic tool to investigate such effects. Furthermore, as in neutral physical organic chemistry, kinetic isotope effects (mostly upon deuteration) are one of the most powerful tools to investigate the mechanisms of a chemical reaction.

## Experimental determination of the microcanonical rate constant

Various experimental techniques are now available to measure the unimolecular rate constant over quite an extended range, typically from $10^2$ to $10^{10}$ s$^{-1}$. Broadly speaking, most experiments can be classified in one or the other of the three following groups, depending on the rate constant range accessible: microsecond time window ($10^5 < k < 10^7$ s$^{-1}$), nanosecond time window ($10^7 < k < 10^{10}$ s$^{-1}$) and millisecond time range ($10^2 < k < 10^5$ s$^{-1}$). Trapping techniques are instrumental in this last range. The techniques pertaining to the first two groups are based on the fact that the kinetic energy acquired by a fragment ion produced within an accelerating electric field depends on the position (or, equivalently, on the time) at which dissociation took place. The rate constant can then be extracted from the shape of an ion kinetic energy spectrum or of a time-of-flight distribution. The rate constant range accessible in a given experiment depends on the magnitude of the electric field.

### Microsecond time window

Ions dissociating in the microsecond time window, i.e. with a rate constant in the $10^5$–$10^7$ s$^{-1}$ range, have received considerable attention. This range is very often referred to as the metastable time frame and processes occurring within this time window are called metastable Photoelectron–photoion coincidence (PEPICO) spectroscopy is the technique of choice in this range. The most useful version of this technique, called threshold PEPICO (TPEPICO), initiated by Stockbauer and by Baer, uses a variable-wavelength photon source, nowadays in most cases a synchrotron or a tunable laser, and detects ions in delayed coincidence with zero-kinetic-energy electrons (threshold electrons). Under such conditions, the ion internal energy is equal to the photon energy less the adiabatic ionization energy of the molecule. By combining angular and time-of-flight discriminations of energetic electrons, a resolution of about 10 meV can be obtained. Even better resolutions are reached by using recent zero-kinetic-energy techniques. The data have to be corrected for the analyser apparatus function, for the thermal distribution of the parent ions and for the possible release of translational energy to the dissociation fragments. Extracting rate constant information in the presence of an appreciable translational energy release is somewhat hazardous.

The upper limit in the rate constant range accessible by threshold PEPICO is dictated by the extraction field which cannot exceed a few V cm$^{-1}$ to maintain a reasonable energy resolution.

Another way of selecting ions with a well-defined internal energy involves multiphoton ionization (MPI) where the vibrationless ground state of the molecular ion is first prepred and then photo-

dissociated by absorption of an additional photon. In practice, the upper rate constant limit is about the same as in PEPICO. The major advantage of MPI is it makes rotational state selection possible, so that the influence of angular momentum on ionic fragmentations can be investigated.

### Nanosecond time window

In field ionization, much higher electric fields can be achieved (of the order of $10^5$ V $cm^{-1}$): the ions are then much more strongly accelerated, and dissociation decays in the nanosecond to picosecond range can be monitored. No internal energy selection is performed in such experiments. Nontrivial data handling procedures assuming some reasonable form for the internal energy distribution have been developed to obtain $k(E)$ data.

Photodissociation mass-analysed ion kinetic energy spectrometry (PD-MIKE) has been designed by Kim and co-workers to monitor nanosecond range decays. Laser photodissociation of mass-selected ions is induced in the presence of a strong electric field within the region located between the magnet and the electrostatic analyser of a reverse-geometry sector mass spectrometer (a region which would be otherwise field-free). The internal energy range of the parent ion prior to photodissociation is relatively broad but can be narrowed by collisional cooling inside the ion source.

### Millisecond time window

Slow dissociation rates ($10^2$–$10^4$ $s^{-1}$) have been measured in Dunbar's laboratory by time-resolved photodissociation, which consists of trapping ions in an ICR cell during a variable delay time after a photodissociating photon pulse. The technique called 'time-resolved photoionization mass spectrometry', developed by Lifshitz, consists of trapping photoions in a cylindrical trap at very low pressure to avoid bimolecular collisions, and then ejecting them into a mass filter after a variable delay covering the microsecond to millisecond range. When the dissociation rate constant becomes lower than *ca.* $10^3$ $s^{-1}$, competition with infrared fluorescence takes place and limits the lifetime of the decomposition process. This has to be taken into account to extract the dissociation rate constant from the experimental data.

The rate constant for the $C_6H_5Br^{+\bullet} \rightarrow C_6Hc + Br^{\bullet}$ reaction has been measured by all these experimental techniques, so that a broad range could be sampled (**Figure 1**). Note that very good agreement is obtained between independent measurements in the metastable domain.

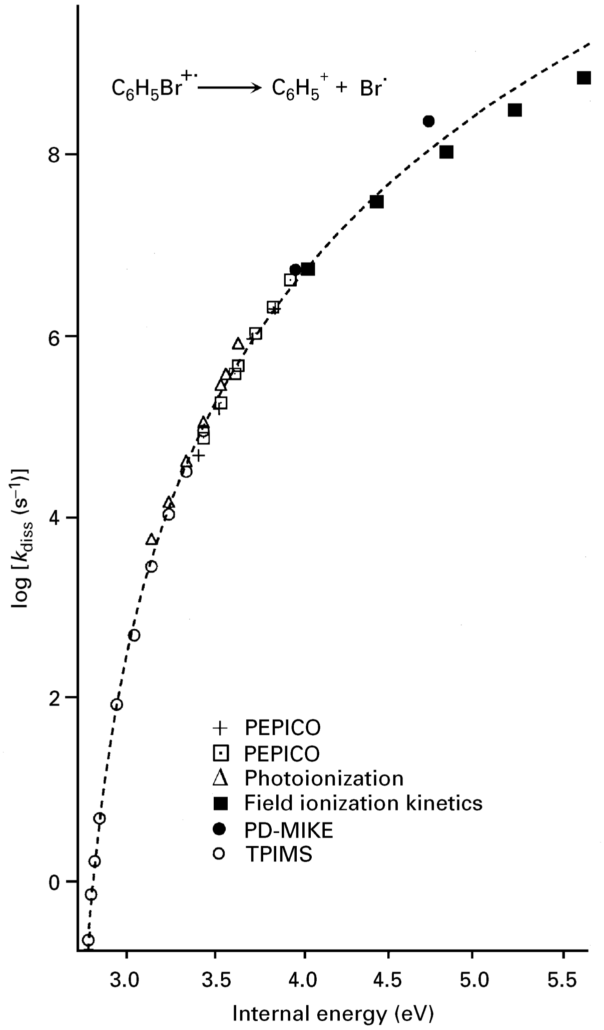



**Figure 1** Microcanonical rate constant for the $C_6H_5Br^{+\bullet} \rightarrow C_6H_5^+ + Br^\bullet$ decomposition obtained using various experimental techniques. (+) Baer *et al* (1976) (+) Baer *et al* (1982) (⊡) Rosenstock *et al* (1980) (Δ) Pratt and Chupka (1981) (■) Brand *et al* (1984) (●) Lim *et al* (1998) (○) Malinovich *et al* (1985).

## Canonical rate constant

Flow tube techniques (flowing afterglow, selected ion flow tube) have been used extensively to study ion–molecule reactions and to measure their rate coefficients. There are far fewer studies on the thermal decomposition of gaseous ions. Recently, cluster ions and noncovalent complexes have, however, received considerable attention.

Dissociation rate constants of ions at atmospheric pressure have been obtained by ion mobility spectrometry. The time frame of such experiments is in the millisecond range. Dissociation rates of proton-bound amine dimer ions, for example, have been measured recently.

In the technique called 'blackbody infrared radiation dissociation' (BIRD), developed in E.R. Williams' laboratory, ions trapped in an ICR cell at very low pressure absorb infrared photons emitted by the cell walls. Dissociation decays in the second to minute time frame can be followed. In both techniques, the analysis of the temperature dependence of the rate constant leads to activation parameters. BIRD has been applied to biologically important non-covalent complexes like proton-bound dimers of amino acids or double-stranded DNA complexes.

## Kinetic isotope effects

Isotope effects, mostly involving hydrogen–deuterium, are one of the most powerful tools for chemical kinetics investigations. The kinetic isotope effect is defined as the ratio of the rate constants corresponding to the unlabelled and labelled species at a given internal energy. A distinction is made between intermolecular and intramolecular kinetic isotope effects. An intramolecular isotope effect refers to the competition, from a unique partially labelled molecular ion, between two channels differing only by the location of the heavier isotope, e.g.:

$$\mathrm{CHDOH^+} \rightarrow \begin{cases} \mathrm{DCO^+ + H_2} \\ \mathrm{HCO^+ + HD} \end{cases}$$

An intermolecular isotope effect refers to two equivalent dissociation pathways from two differently labelled molecules. As an example:

$$\mathrm{C_2H_5Cl^{+\bullet} \rightarrow C_2H_4^{+\bullet} + HCl}$$
$$\mathrm{C_2D_5Cl^{+\bullet} \rightarrow C_2D_4^{+\bullet} + DCl}$$

For intramolecular isotope effects, as one starts from a unique precursor, the two processes correspond to the same internal energy distribution, so that useful information can be inferred even from ionabundances measured without internal energy selection (like metastable dissociations or field ionization kinetics). This is not the case for intermolecular effects, where there is no warranty that the two internal energy distributions are the same. Within the RRKM framework, the intramolecular kinetic isotope effect depends only on the transition state properties (critical energy, rotational constants and vibrational frequencies) and not on the reactant properties.

In primary isotope effects, the isotopically substituted atom is directly involved in bond breaking or bond formation. In secondary isotope effects, the substituted atom is remote from the incriminating bond. Primary isotope effects occur largely because of changes in the critical energy $E_0$ upon labelling. Secondary isotope effects result from changes in the densities of states. Intermolecular secondary isotope effects can be larger than primary effects. The larger the critical energy, the larger the secondary isotope effect. A kinetic isotope effect, where $k_H/k_D$ is larger than unity, is said to be normal. If $k_H/k_D < 1$, the effect is called inverse.

Isotope effects increase when internal energy decreases. Large effects are therefore expected in the metastable time range which corresponds to relatively low internal energies. Isotope effects are instrumental in (i) localizing the particular atoms involved in H-shift reactions, (ii) providing evidence of hydrogen scrambling, (iii) distinguishing between stepwise or concerted processes, (iv) postulating a reasonable structure for the transition state, (v) deciding whether tunnelling is involved or not. Examples will be given in the next sections.

## Reaction mechanisms from kinetic data

We will consider, successively, reactions involving only one potential energy surface and those where atleast two electronic states of the same or of a different multiplicity interact. These latter reactions are called nonadiabatic. Among the adiabatic reactions, we will move from the simple to the more complex, from one-step dissociations to competitive processes and finally to multistep fragmentations. While convenient, this classification is oversimplified. A multistep fragmentation, for example, can involve both adiabatic and nonadiabatic intermediate steps.

### One-step dissociations

Dissociations of substituted benzene ions are probably among the most extensively studied ionic reactions. Most of them were satisfactorily modelled by the RRKM equation. The looseness or tightness of the transition state is usually inferred from the value of the activation entropy (at a somewhat arbitrary temperature, usually 1000 K) resulting from the RRKM fit. A clear-cut positive value for $\Delta S^{\circ\ddagger}_{1000\ \mathrm{K}}$ (usually not exceeding 35 J K$^{-1}$ mol$^{-1}$) indicates a loose transition state, whereas a slightly positive or a negative value (usually not smaller than −30 J K$^{-1}$ mol$^{-1}$) is the signature of a tighter transition state. **Table 1** illustrates this for a few selected reactions,

**Table 1** Entropies of activation for selected one-step dissociations

| *Reaction* | $\Delta S^{\circ\ddagger}_{1000\,K}$ (J K$^{-1}$ mol $^{-1}$) |
|---|---|
| $C_6H_5Br^{+\bullet} \rightarrow C_6H_5^+ + Br^\bullet$ | 33.8[a] |
| $C_6H_5I^{+\bullet} \rightarrow C_6H_5^+ + I^\bullet$ | 31.1[b] |
| $C_6H_5OH^{+\bullet}$(phenol)$\rightarrow$ c-$C_5H_6^{+\bullet}$ + CO | 9.2[c] |
| p-$C_6H_4Cl_2^{+\bullet}$ + $C_6H_4Cl^+$ + $Cl^\bullet$ | 7.5[d] |
| $C_6H_5CN^{+\bullet} \rightarrow C_6H_4^{+\bullet}$+ HCN | 1.7[e] |
| $C_8H_8^{+\bullet}$(styrene) $\rightarrow C_6H_6^{+\bullet} + C_2H_2$ | –26.8[f] |

[a] Malinovich Y, Arakawa R, Haase G and Lifshitz C (1985) Time-dependent mass spectra and breakdown graphs. 6. Slow unimolecular dissociatioon of bromobenzene ions at near threshold energies. *Journal of Physical Chemistry* **89**: 2253–2260.

[b] Malinovich Y, and Lifshitz C (1986) Time-dependent mass spectra and breakdown graphs. 7. Time-resolved photoionization mass spectrometry of iodobenzene. The heat of formation of $C_6H_5^+$. *Journal of Physical Chemistry* **90**: 2200–2203.

[c] Malinovich Y and Llifshitz C (1986) Time-dependent mass spectra and breakdown graphs. 8. Dissociative photoionization of phenol. *Journal of Physical Chemistry* **90**: 4311–4317.

[d] Olesik S, Baer T and Morrow JC (1986) Dissociation rates of energy-selected dichloro- and dibromobenzene ions. *Journal of Physical Chemistry* **90**: 3563–3568.

[e] Rosenstock HM, Stockbauer R and Parr AC (1980) Photoelectron-photoion coincidence study of benzonltrile. *Journal de Chimie Physique* **77**: 745–750.

[f] Dunbar RC (1989) Time-resolved unimolecular dissociation of styrene ion. Rates and activation parameters. *Journal of the American Chemical Society* **111**: 5572–5576.

which are believed to involve a single dissociative step, i.e. to take place on a single-well potential energy surface.

At first sight, it might seem inconsistent to use a thermal quantity to characterize a microcanonical rate constant. However, the canonical rate constant is related, via a thermal average, to the microcanonical rate constant with the same transition state. The advantage of the canonical formulation is twofold. First, the essential information is compacted into two factors, the critical energy and the entropy of activation, rather than being scattered among many individual vibrational frequencies. Second, it makes a comparison possible with thermal data for the neutral counterpart of an ionic reaction. Alternatively, some authors calculate the microcanonical entropy of activation at infinite internal energy, so that the critical energy becomes negligible, according to:

$$\Delta S^{\ddagger}_{\mu}(E \rightarrow +\infty) = R \ln \frac{N^{\ddagger}(E - E_0 \rightarrow +\infty)}{h\nu_{\mathrm{R}}\rho(E \rightarrow +\infty)}$$

$\nu_R$ is the frequency associated with the reaction coordinate. For the $C_6H_6^{\bullet+} \rightarrow C_6H_5^+ + H^\bullet$ reaction, e.g. $\Delta S^{\circ\ddagger}_{1000K}$ = 11.3 J K$^{-1}$ mol$^{-1}$ and $\Delta S^{\ddagger}_{\mu}(E \rightarrow +\infty)$ = 13.2 J K$^{-1}$ mol$^{-1}$. For the $C_2H_4^{+\bullet} \rightarrow C_2H_3^+ + H^\bullet$ reaction, the canonical and microcanonical entropies of activation amount to 32 and 36 J K$^{-1}$ mol$^{-1}$ respectively.

The large activation entropies observed for the dissociations of the bromo- and iodobenzene cations are compatible with a maximum entropy analysis of the product energy distributions, which leads to a nearly complete (about 80%) phase sample sampling. As illustrated in **Figure 1**, an excess energy of about 0.7 eV, called the 'kinetic shift' is necessary to bring the rate constant above $10^5$ s$^{-1}$ so that dissociation can be detected in the metastable window.

The dissociation of the styrene ion involves a transition state very similar to the reactant. The large negative entropy of activation results from the fact that the internal energy content of the transition state is reduced, compared to the reactant ion, by an amount equal to the critical energy. The transition state structure for the CO loss from the phenol cation is also close to the precursor but some vibrational modes loosen up, resulting in a slightly positive activation entropy.

Activation entropies result from a fitting procedure including the critical energy $E_0$, too. In many cases, the limited range of internal energies investigated, as well as the lack of information on the structure and/or the thermochemistry of the fragments, and on a possible reverse activation barrier, preclude any reliable control over $E_0$. One should therefore be aware of energy–entropy trade-offs.

## Competitive reactions

Ions generally have low isomerization barriers so that decomposition channels leading to different product ions with the same molecular formula can compete. Two examples will illustrate this point.

In contrast to the simple C–I bond cleavage observed with the iodobenzene cation, the corresponding reaction with the iodotoluene cation is much more complex, leading to a mixture of three $C_7H_7^+$ isomers with tolyl, benzyl and tropylium structures. The experimental data available have been analysed by Dunbar and co-workers using a two-channel model (**Figure 2**). The lowest energy channel leads to the benzyl and tropylium structures via a rearrangement characterized by a tight transition state ($\Delta S^{\circ\ddagger}_{1000\,K}$ = –29 J K$^{-1}$ mol$^{-1}$), whereas the highest energy channel is a direct bond cleavage ($\Delta S^{\circ\ddagger}_{1000\,K}$ = 32 J K$^{-1}$ mol$^{-1}$). As expected, the branching ratio between the two channels depends dramatically on the internal energy content.

Loss of HF from the 1,1-difluoroethylene cation at low internal energy leads to the fluoroacetylene fragment ion with a rate constant in the metastable range and a large reverse activation barrier of nearly 1 eV. At higher internal energy, collisional activation experiments and TPEPICO data show that a competitive channel becomes open, leading probably to the fluoroethenylidene cation. This is associated with a decrease in the kinetic energy released to the fragments.

## Multistep fragmentations

A dissociation can very seldom be described as a one-step process starting from a unique potential energy minimum and leading to the fragments via a unique transition state. One or more reversible isomerization steps prior to dissociation have been frequently proposed to interpret experimental data in the framework of a classical kinetic scheme coupled with a RRKM evaluation of the appropriate rate constants. *Ab initio* calculations have revealed the importance of unconventional species like bridged structures and of distonic ions as stable ionic structures. Distonic ions are radical ions for which the ionic site and the radical site are formally separated, like in $^{\bullet}CH_2CH_2OH_2^+$. Furthermore, ion–neutral complexes have been suggested in many instances to play an important role in the disso-ciation mechanism. The production of $HCO^+$ from $CH_3OH^{+\bullet}$ at threshold has been shown to proceed via the following multistep sequence:

$$CH_3OH^{\bullet+} \rightarrow \overset{\bullet}{C}H_2\,\overset{+}{O}H_2 \rightarrow [CH_2OH\ldots H]^{\bullet+}$$
$$\rightarrow [CH_2O\ldots H_2]^{\bullet+} \rightarrow CH_2O^{\bullet+} + H_2$$
$$\downarrow$$
$$HCO^+ + H^{\bullet} + H_2$$

where both a distonic ion and two ion–neutral complexes (bracketed species) are involved. The typical ion–neutral separation in such a complex is in the 3–5 Å range. Rearrangements or ion–neutral reactions can take place within this entity. They are favoured at low internal energy, whereas simple cleavages become more probable at higher internal energy. A spectacular ion–molecule reaction within an ion-neutral complex has been discovered by Botter and Longevialle: it results in the transfer of a proton from the D side to the A side of a steroid molecular ion.

The decomposition of the propanol cation is a nice example of a concerted investigation by coincidence methods, including selective deuteration, and *ab*

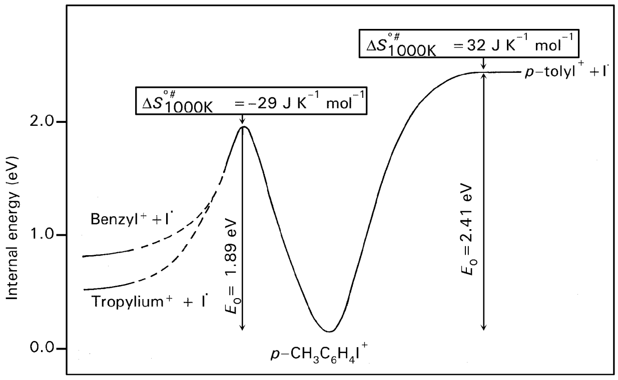



**Figure 2** Reaction mechanism for the loss of iodine from the *p*-iodotoluene cation, after the work of Lin and Dunbar: Lin CY and Dunbar RC (1994) Time-resolved photodissociation rates and kinetic modelling for unimolecular dissociations of iodotoluene Ions. *The Journal of Physical Chemistry* **98**: 1369–1375.

*initio* calculations. **Figure 3**, adapted from work of Baer and Booze, describes the multistep process involved in the competition between the H loss and $H_2O$ loss channels. Intermolecular isotope effects were studied with partially deuterated propanol cations, $CD_3CH_2CH_2OH^+$. A very small isotope effect is observed for the H loss, which, according to the model, is a secondary effect. Water loss, on the other hand, involves a 1,4-hydrogen shift as an intermediate step. Deuteration brings about a primary isotope effect, characterized by a 0.15 eV increase in the critical energy.

The intermediate steps which have just been mentioned correspond to successive minima of the potential energy surface. What about the transition state? It is not unique either. To find its position is a particularly critical problem if no saddle point appears along the dissociation pathway. As suggested by Bowers and co-workers, two transition states exist in this case, respectively at small and large values of the reaction coordinate. The internal transition state results from the competition between the conversion of kinetic into potential energy and the decrease of the vibration frequencies as the system moves along the reaction coordinate. The second transition state, which is called the orbiting transition state, is associated with the centrifugal barrier appearing at larger values of the reaction coordinate. The internal transition state is expected to play a significant role at high internal energy, whereas the external transition state governs the dissociation rate at energies closer to the threshold. The experimental observation of this transition state switching is still sought for single-well potential energy surfaces in ionic systems. Other kinds of transition states can also be invoked. A strong

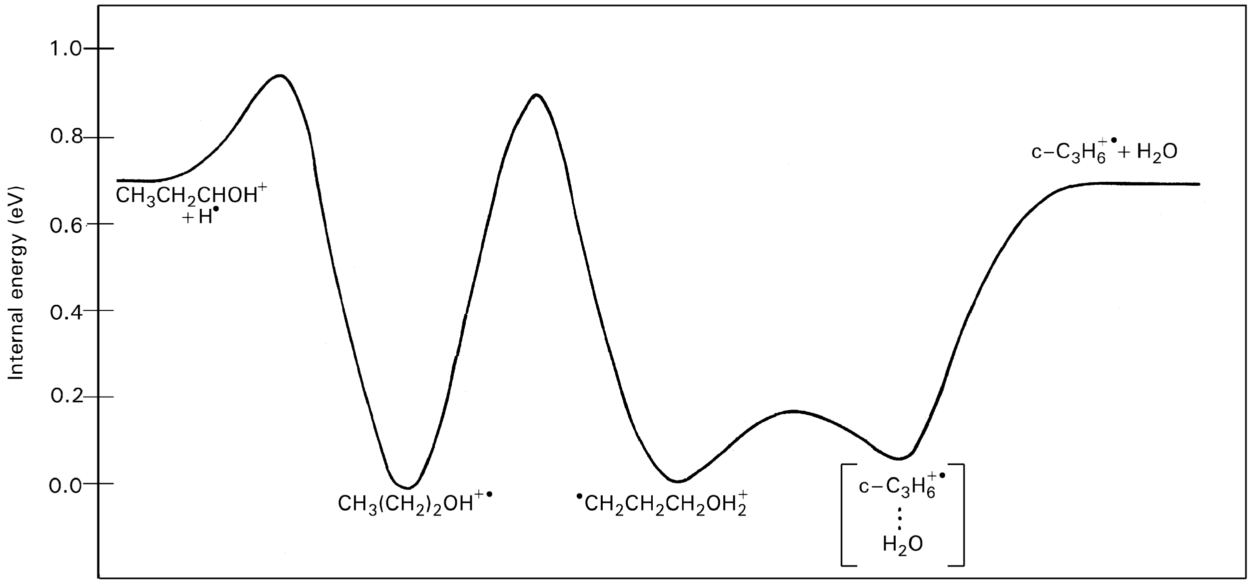



**Figure 3** Multistep reaction pathway for the loss of H and of $H_2O$ from ionized propanol. The figure is adapted from a study by Booze and Baer: Booze JA and Baer T (1992) Dissociation dynamics of energy selected propanol ions from a $\sigma$-type ion structure. *The Journal of Physical Chemistry* **96**: 5715–5719.

curvature of the reaction path brings about a bottleneck in the energetically allowed phase space and thus a decrease of the reaction rate. An avoided crossing between two potential energy surfaces limits also the reactive flux due to a branching between the two electronic states. This latter aspect will be discussed in the next subsection.

### Nonadiabatic reactions

The excited electronic states of the molecular ion which are populated by the ionization process in general decay very rapidly (lifetime of the order of 0.1 ps) to the ground electronic state by internal conversion of electronic energy into vibrational energy. If the electronic ground state does not correlate adiabatically to a given pair of fragments, the production of these fragments requires a transition to a potential energy surface correlating to the appropriate dissociation asymptote, at some stage of the decomposition process. A reaction requiring such an interstate transition is referred to as nonadiabatic. Nonadiabatic behaviour is always expected when a competition involves two close-lying dissociation asymptotes, differing only by the location of the positive charge on one or the other fragments. In general, this will always be the case for charge-transfer reactions. The statistical theories can be adapted to accommodate nonadiabatic transitions occurring along the reaction path. Briefly, the nonadiabatic transition probability, i.e. the probability of switching from one potential energy surface to the other, is introduced as a kind of transmission coefficient in a RRKM-like equation.

The two electronic states involved in a nonadiabatic transition have either the same or a different multiplicity. In the former case, the coupling of the electronic and nuclear motions is involved. If multiplicity changes, the transition is due to the spin–orbit coupling. A recent example deals with the unimolec-ular loss of singlet $H_2$ from triplet $CH_3O^+$ to produce singlet $HCO^+$. Based on the measurement of intramolecular kinetic isotope effects and on *ab initio* calculations, Schwarz and co-workers rejected the usually accepted stepwise mechanism in favour of a mechanism in which loss of $H_2$ and intersystem crossing are concerted.

### Tunnelling

If a reverse activation barrier exists, tunnelling is possible. The probability of tunnelling depends dramatically on the reduced mass. The smaller the ‘reduced mass, the larger the probability. Furthermore, the largest H–D isotope effects are expected to occur near the threshold. This corresponds generally to the metastable time scale, which is the most easily accessible to mass spectrometric experiments. If very large H/D kinetic isotope effects are observed, sometimes larger than 1000, then tunnelling is often proposed as the most probable explanation. To take tunnelling into account in the RRKM equation, a transmission coefficient is introduced. It can be calculated analytically provided an appropriate

model is used for the potential energy barrier. The asymmetric Eckart potential is often chosen.

*Ab initio* calculations performed by different groups on the loss of $H_2$ from the ethane cation agree that the appearance energy for this rearrange-ment lies below the top of the barrier. On the other hand, any RRKM calculation neglecting tunnelling leads to a rate constant at threshold larger than $10^8$ s$^{-1}$, which is inconsistent with the metastability of this reaction. RRKM calculations by Weitzel including tunnelling are in good agreement with $k(E)$ measurements by the TPEPICO technique.

The intermolecular kinetic isotope effect $k_H/k_D$ for the loss of HCl or DCl from $C_2H_5Cl^{+\bullet}$ and $C_2D_5Cl^{+\bullet}$ respectively lies between $10^2$ and $10^5$. It has been well accounted for by RRKM calculations including tunnelling (Booze and co-workers).

As shown by Lifshitz, Schwarz and co-workers, the methane elimination from the acetone cation proceeds via two successive intermediates, a hydrogen-bridged complex and a ketene–methane ion–neutral complex. The barrier separating these two entities lies 14.2 kJ mol$^{-1}$ above the $CH_2CO^{+\bullet} + CH_4$ asymptote. Tunnelling through this barrier is required to account for the appearance of the $CH_2CO^{+\bullet}$ fragment at its thermochemical threshold.

### Isolated state decay

In most ionic unimolecular reactions, internal conversion to the ground electronic state is rapid compared to dissociation. In a few cases, however, evidence has been found of a rapid and specific dissociation process occurring on the potential energy surface of an excited electronic state. When such an 'isolated state decay' occurs, a correlation exists between the branching ratio for the specific channel and the photoelectron spectrum. Ionized difluoroethylenes for example, dissociate statistically below the $\tilde{C}$ state but this latter state favours strongly the F loss channel.

## Future prospects

A fairly important amount of kinetic data is nowadays available, which in the vast majority of cases, have been accounted for by a classical kinetic scheme coupled to a RRKM evaluation of the rate constants involved, possibly including tunelling. The story is, however, far from being complete.

1. The rate constant is often known only in a limited range so that being able to fit the data to a statistical equation like the RRKM-QET one is not a definitive proof of statisticity by itself. The translational energy distribution of the fragment ions is complementary piece of information which is more sensitive to any deviation from a purely ergodic situation. Rather few detailed analyses of experimental distributions are available for the time being.
2. Nonadiabatic reactions are probably much more widespread than usually thought. In particular the role of spin–orbit coupling in unimolecular chemistry deserves careful investigation. In addition to the experimental component, this raises nontrivial *ab initio* computational problems.
3. The interpretation of kinetic experiments is almost always performed on the basis of a classical kinetic scheme. A quantum approach studying the decay of quasibound states, called resonances, could be more appropriate under given circumstances. Large fluctuations of nearly two orders of magnitude have been found, for example, for the decay rate of state-selected formaldehyde.

## List of symbols

$E$ = total energy; $h$ = Planck constant; $k_B$ = Boltzmann constant; $k(E)$ = unimolecular microcanonical rate constant; $k(T)$ = canonical rate constant; $N_0$ = initial number of ions formed; $N^{\ddagger}(E - E_0)$ = number of states of the transition state up to $E - E_0$ above the critical energy $E_0$; $P(E)$ = distribution of internal energies; $R(E,t)$ = rate of dissociation; $T$ = temperature; $\Delta H^{\circ\ddagger}$ = activation enthalpy; $\Delta S^{\circ\ddagger}$ = activation entropy; $\Delta S^{\ddagger}_{\mu}$ = microcanonical entropy of activation; $\nu_R$ = reaction coordinate frequency; $\rho(E)$ = density of states of the parent ion at internal energy $E$; $\sigma$ = degeneracy of reaction path.

*See also:* **Fragmentation in Mass Spectrometry; Metastable Ions; Multiphoton Excitation in Mass Spectrometry; Photoelectron–Photoion Coincidence Methods in Mass Spectrometry (PEPICO); Statistical Theory of Mass Spectra.**

## Further reading

Baer T and Hase WL (1996) *Unimolecular Reaction Dynamics. Theory and Experiments.* Oxford: Oxford University Press.

Baer T, Booze J and Weitzel KM (1991) Photoelectron–photoion coincidence studies of ion dissociation dynamics. In: Ng CY (ed), *Vacuum Ultraviolet Photoionization and Photodissociation of Molecules and Clusters*. Singapore: World Scientific.

Derrick PJ and Donchi KF (1983) Mass spectrometry. In: Bamford CH and Tipper CFH (eds) *Comprehensive Chemical Kinetics*. Vol. 24. *Modern Methods in Kinetics*. Amsterdam: Elsevier.

Forst W (1973) *Theory of Unimolecular Reactions*. New York: Academic Press.
Gilbert RG and Smith SC (1990) *Theory of Unimolecular and Recombination Reactions*. Oxford: Blackwell Scientific Publications.
Lifshitz C (1992) Recent developments in applications of RRKM-QET. *International Journal of Mass Spectrometry and Ion Processes* 118/119: 315–337.
Longevialle P (1992) Ion–neutral complexes in the unimolecular reactivity of organic cations in the gas phase. *Mass Spectrometry Reviews* **11**: 157–192.
Lorquet JC and Leyh B (1984) Potential energy surfaces and theory of unimolecular dissociation. *Ionic Processes in the Gas Phase. NATO ASI SERIES* **C118**: 1–6
Steinfeld JI, Francisco JS and Hase WL (1999) *Chemical Kinetics and Dynamics*, 2nd edn. Upper Saddle River: Prentice Hall.

# Ion Energetics in Mass Spectrometry

**John Holmes**, University of Ottawa, Canada

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| **Theory** |

## Neutral thermochemistry

For neutral organic compounds there are only some 3000 good quality data for molecular heats of formation, $\Delta_f H^0$[M]. These have been carefully selected by Pedley and co-workers over an extended period. The standard heat of formation of an entity, $\Delta_f H^0$, is defined as the heat absorbed or released when 1 mole of the species is formed from its constituent elements at 298 K and at 1 atm pressure. This is represented by an equation, e.g.

$$C(s) + 2H_2(g) \rightarrow CH_4(g), \Delta_f H^0 = -17.78\,\text{kcal mol}^{-1}$$

The terms in brackets represent the phase in which the reactants and products are stable in their standard state at 298 K and 1 atm pressure (1 cal = 4.184 J).

Although so few organic compounds have accurate experimental $\Delta_f H^0$ values, the knowledge that $\Delta_f H^0$ for homologous organic compounds can be represented as an additive property has resulted in the production of additivity coefficients, the terms which are most closely associated with the work of Benson. The coefficients now available allow one to estimate $\Delta_f H^0$ for a very wide range of organic compounds via simple arithmetical calculations. Sufficient data exist for the estimation of many geometric effects such as ring strain energies, geometric isomers, etc.

## Ionization energies and ionic heats of formation, $\Delta_f H^0[M^{\bullet+}]$

The $\Delta_f H^0$ value for a molecular ion $M^{\bullet+}$ is obtained from the measurement of the ionization energy (IE) of neutral M. The reaction

$$M \rightarrow M^{\bullet+} + e^-$$

has

$$\Delta H^0_{\text{reaction}} = \text{IE[M]} = \Delta_f H^0[M^{\bullet+}] - \Delta_f H^0[M] \quad [1]$$

$\Delta_f H^0$[M] is obtained from reference data, from an additivity estimation or, more recently, from high level *ab initio* molecular orbital theory calculations. The IE is derived from an experimental measurement. In Equation [1] the electron is given an enthalpy of zero (i.e. it is assumed to be stationary). Ionization energies can be obtained from photoelectron spectra or from direct measurement of the ionization threshold using monochromated photons or energy-selected electrons. Whatever method is chosen, the most important factor in determining a molecular IE value is the relative geometry of the molecular ion and the neutral molecule, both in their ground states. The terms vertical IE and adiabatic IE are defined in **Figure 1**.

If the geometry difference between the molecule and its ion is particularly large then the adiabatic IE

(that necessary for the determination of $\Delta_f H^0$ [$M^{\bullet +}$] in Eqn [1]) will be difficult to measure. Photoionization is only a vertical process; electron impact is less so because the approach of the ionizing electron and its effects on the molecule are not instantaneous. For moderate geometry differences, the adiabatic IE can be measured using energy-selected electron beams, as a reasonably well-defined asymptote in the ionization efficiency curve, see **Figure 1**.

Molecular adiabatic ionization energies are to be found in the NIST database (see Further reading section) and other reference books.

A number of attempts have been made to produce empirical equations for the estimation of molecular ion enthalpies, notably by Mouvier and co-workers and by Holmes, Fingas and Lossing. The latter is a simple combination of Benson-type additivity terms for the neutral's $\Delta_f H^0$ and a reciprocal term that incorporates the IE; This arises from the observation that for homologous molecules their IE values fall linearly with the reciprocal of the ion size, represented as $1/n$, where $n$ is the number of atoms in the molecule. The equation

$$\Delta_f H^0[M^{\bullet +}] = A - Bn + (C/n)$$

where $A$, $B$ and $C$ are derived from experimental data for a given homologous series works well for various organic compounds that contain C, H, O and halogen. The secondary empirical relationship, that IE for homologous molecules are proportional to $1/n$, has proved useful in many applications where an unmeasured IE needs to be estimated or where an experimental value appears to be anomalous.

The empirical equation of Mouvier is more complex:

$$\log\left[\frac{\mathrm{IE}(R^1XR^2) - \mathrm{IE}_\infty}{\mathrm{IE}_0 - \mathrm{IE}_\infty}\right] = 0.106[\mathrm{I}(R^1) + \mathrm{IR}^2)]$$

where X is a functional group and $R^1$ and $R^2$ are attached alkyl groups, $\mathrm{IE}_\infty$ is a constant for each compound type and $\mathrm{IE}_0$ is that for a reference compound. $\mathrm{I}(R^n)$ are empirical constants for each R group. The IE values derived and the $\Delta_f H^0$[ion] values obtained using the appropriate neutral enthalpies provide constant ionic heats of formation.

Perhaps the most important use of all such schemes is their ability to reveal misfits in reference data or to highlight an ion having unusual stabilizing or destabilizing properties.

The $\Delta_f H^0$ values for even-electron ions (i.e. ionized free radicals) can be obtained from direct measurement provided that a clean source of the appropriate free radical can be devised. The most versatile method has been that pioneered by Lossing in which the flash pyrolysis of nitrite esters was used to generate the radicals, e.g.

$$(CH_3)_2CHCH_2ONO \xrightarrow{\text{Heat}} (CH_3)_2CH^\bullet + CH_2O + NO^\bullet$$

Direct IE measurements by these means remove the uncertainties which arise in appearance energy (AE) measurements (described below).

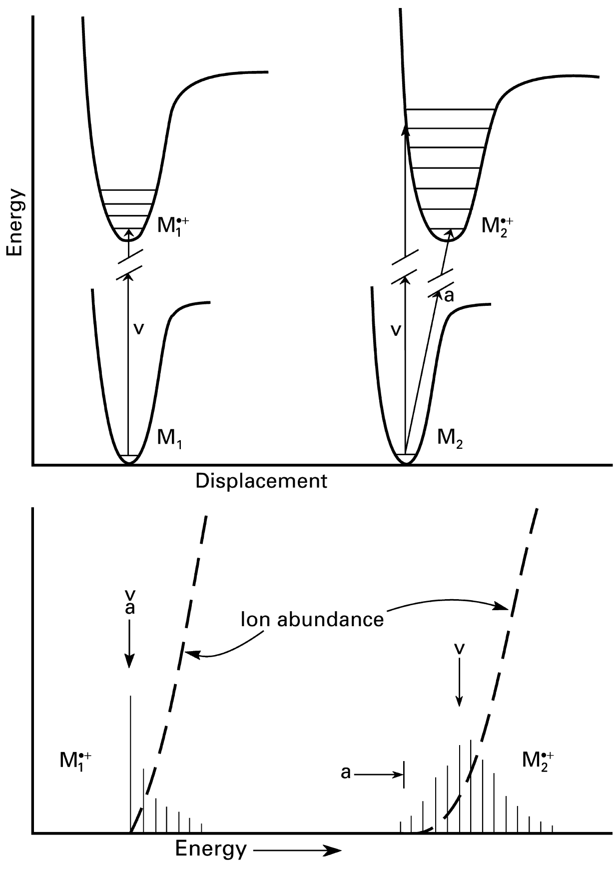



**Figure 1** Vertical (v) and adiabatic (a) ground state to ground state ionization energies. Upper curves are the potential energy diagrams for two molecules, $M_1$ and $M_2$, and their molecular ions. For $M_1$ there is no significant geometry change on ionization and v and a have the same value. For molecule $M_2$, the geometry of the ion is not the same as that of the neutral and so v > a. The lower figure shows the ion onset curves (broken lines) and the vibrational modes accessed in the ion with the lines' length showing their relative importance in the ionization process. Note, the more poorly defined onset for $M_2^{\bullet +}$.

## Appearance energies

The measurement of the $\Delta_f H^0$ for a fragment ion is an important aid to ion structure determination. The energy change associated with the general equation

$$M_1 \rightarrow M_2{}^+ + M_3{}^{\bullet} + e^-$$

is called the appearance energy (AE) (for the fragment ion $\overset{+}{M_2}$). The simple thermochemical relation, Equation [2], corresponding to this reaction:

$$\Delta H_2^0 = AE[M_2^+] = \Delta_f H^0[M_2^+] + \Delta_f H^0[M_3{}^{\bullet}] - \Delta_f H^0[M_1] \quad [2]$$

is insufficient to describe the full result of the AE measurement for several reasons:

(i) The reverse reaction may have an energy barrier, $E_{rev}$, and so the measured AE will exceed the thermochemical minimum for the reaction.

(ii) The measured AE may be too high because the dissociation of $M_1^{\bullet+}$ involves a significant kinetic shift. A kinetic shift arises as follows. The timescale of the experiment for measuring the AE is usually very short, typically $10^{-4}$ to $10^{-6}$ s, and so for the unimolecular dissociation of $M_1^{\bullet+}$ ions to achieve rate constants of $10^4$–$10^6$ s$^{-1}$ they must contain excess energy above the minimum required for the reaction to proceed. For many dissociations, particularly simple bond cleavages in small ions, the kinetic shift effect is negligible. For dissociations of large ions or molecular elimination reactions, the effect can be appreciable, as much as 0.5 eV. In addition, for small molecular ions which lie in deep potential wells, the large density of states at the dissociation limit may give rise to a significant kinetic shift. A good example is provided by the phenyl halides, where on the $\mu$s time-scale the mean kinetic shift for the simple bond cleavage reaction

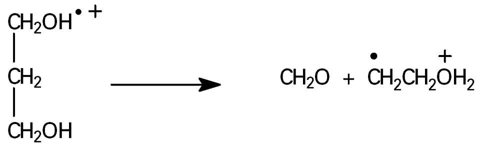
$$C_6H_5Br^{\bullet+} \longrightarrow C_6H_5{}^+ + Br^{\bullet}$$

is 0.7 eV. Note that if the AE values for fragment ions from *metastable ion* precursors are measured, the limiting observational rate constant reduces to ~$10^4$ s$^{-1}$ and the above kinetic shift is halved. Delayed observations, on the ms time-scale, reduce the kinetic shift even further.

(iii) In general, the internal energy of the reactant molecule is taken to be that corresponding to the ambient temperature of the ion source and that of the products is essentially unknown. In photoionization AE measurements, Traeger has supposed that the products are at 0 K, perhaps an unlikely outcome for polyatomic species. Proposing that the products' energy differs little from those of the neutral reactant does not appear to introduce major errors.

(iv) The caveats regarding $\Delta_f H^0[M_1]$ apply also for $\Delta_f H^0[M_3{}^{\bullet}]$, which as a free radical may not have a well-assigned enthalpy of formation. Indeed, radical $\Delta_f H^0$ values can reliably be obtained from AE measurements when the ionic heat of formation $\Delta_f H^0[M_2{}^+]$ is well established and $\Delta_f H^0[M_3{}^{\bullet}]$ is the unknown. The results of this approach equal those derived from activation energies determined from the kinetics of neutral species' reactions.

(v) Hidden rearrangements: in many cases the structure of the ion and the accompanying neutral fragment are not in doubt. However, a significant number of molecular ions undergo unpredicted rearrangements before they fragment. If, therefore, there is any doubt as to the structure of the fragmentation products, then appropriate experiments should be performed. Two examples suffice to illustrate the problem: the $C_2H_6O^{\bullet+}$ ion generated by loss of $CH_2O$ from ionized propane-1,3-diol is neither $CH_3CH_2OH^{\bullet+}$ nor $CH_3OCH_3^{\bullet+}$ (molecular ions of the only two stable $C_2H_6O$ compounds) but has the structure $\dot{C}H_2CH_2\overset{+}{O}H_2$. This ion behaves as an electrostatically bound water molecule – ethene ion; its $\Delta_f H^0$ value is *lower* than that of either of the above two conventional ions and, indeed, is the lowest energy $C_2H_6O^{\bullet+}$ isomer.




The dissociation of ionized methyl acetate appears to be trivial

$$CH_3CO_2CH_3{}^{\bullet+} \longrightarrow CH_3\overset{+}{C}O + [C, H_3, O]^{\bullet}$$

with the formation of the acetyl ion and a methoxy radical as the products. Astonishingly, the ion is as expected but the lowest energy dissociation of this

molecular ion produces the isomeric hydroxymethyl radical, $^{\bullet}CH_2OH$, instead of the predicted $CH_3O^{\bullet}$ species. In this case, profound rearrangements have preceded dissociation.

In summary, AE measurements, although relatively easy to perform, may be difficult to evaluate because of the above uncertainties. However, a reverse energy barrier will often be manifest by the magnitude of the kinetic energy release observed in the metastable peak accompanying the fragmentation. A small kinetic energy release of say < 10 meV likely indicates neither a reverse energy barrier nor a significant kinetic shift.

Except for the existence of a reverse energy barrier, most of the above problems are avoided in photoelectron photoion coincidence experiments, where the dissociation of only internal energy selected species are observed.

## Estimation of ion enthalpies

It was described above how empirical expressions can be used to estimate $\Delta_f H^0$ values for odd-electron (molecular) ions.

No such simple formulae have been derived for even-electron (fragment) ions, but some useful general semiquantitative relationships have been found. The simplest is that for successive substitution of a single group (e.g. $CH_3$, OH, etc.) at a formal charge-bearing site. $\Delta_f H^0$[Ion] decreases linearly with log(ion size), the latter usually being represented by the number of atoms in the ion. This is illustrated in **Figure 2** for successive substitution of H by X, beginning with the methyl cation. Note that the OH substitutions (steep lines) and $CH_3$ substitutions produce linear plots of similar slope.

## Charge density in ions

It is worth emphasizing that thermochemical data give insights as to the distribution of charge in related ions. Thus, for example, the effect of methyl substitution at carbon in $[CH_2OH]^+$ is much greater than for the substitution at oxygen; $\Delta_f H^0$ $[CH_2OH]^+$ = 169 kcal mol$^{-1}$. $\Delta_f H^0[CH_3CHOH]^+$ = 139 kcal mol$^{-1}$ [$\Delta(\Delta_f H^0)$ = –30 kcal mol$^{-1}$] whereas $\Delta_f H^0[CH_2OCH_3]^+$ = 157 kcal mol$^{-1}$ [$\Delta(\Delta_f H^0)$ = –12 kcal mol$^{-1}$]. Thus the hydroxymethyl cation (protonated formaldehyde) is better represented as $^+CH_2OH$ (a $^+$C hydroxy substituted carbocation) than as $CH_2=O^+H$ (an oxonium ion). This is also borne out by the ion's reactivity and its behaviour in ion–molecule complexes.

**Table 1** The effect of OH substitution at various positions in alkenes and their ions (energies in kcal mol$^{-1}$)

| *Alkene* | $\Delta_f H^0$[M] | $\Delta_f H^0[M^{\bullet+}]$ | $\Delta(\Delta_f H^0)$ *Neutral* | $\Delta(\Delta_f H^0)$ *Ion* |
|---|---|---|---|---|
| $CH_2=CHCH_2CH_3$ | 0 | 222 | – | – |
| $HOCH=CHCH_2CH_3$ | –42 | 150 | –42 | –72 |
| $CH_2=C(OH)\ CH_2CH_3$ | –43 | 150 | –43 | –72 |
| $CH_2=CHCH(OH)CH_3$ | –38 | 181 | –38 | –41 |
| $CH_2=CHCH_2CH_2OH$ | –36 | 184 | –36 | –38 |
| $CH_2=C(CH_3)_2$ | –4 | 209 | – | – |
| $HOCH=C(CH_3)_2$ | –45 | 145 | –41 | –64 |
| $CH_2=C(CH_3)CH_2OH$ | –38 | 176 | –34 | –33 |

The effects of, for example, OH substitution at, near and away from a formal charge-bearing site are clearly and significantly different, a good example being provided by the isomeric ions shown in **Table 1**. Note that the heat of formation of the ion, when OH is *not* at the formal charge-bearing site (here the π-system), is governed by the difference in $\Delta_f H^0$ values for the neutral alkene and alkenol. When OH is at the charge site, the stabilization effect is much greater.

## Proton affinities

It is predicted that there should be a linear relationship between the proton affinities of homologous molecules and their ionization energies. A simple thermochemical cycle gives

$$\mathrm{PA[M] = HA[M^+] + IE[H] - IE[M]}$$

If the hydrogen atom affinity (HA) of homologous ions is a constant, then

$$\mathrm{PA[M]} \approx \mathrm{constant - IE[M]}$$

This equation indeed works quite well for homologous alcohols, ethers, carbonyl compounds, amines, etc.

## Electron affinities

The electron affinity (EA) of a molecule is defined as

$$\mathrm{EA[M]} = \Delta_f H^0\mathrm{[M]} - \Delta_f H^0\mathrm{[M^-]}$$

Again the problems of adiabatic and vertical values

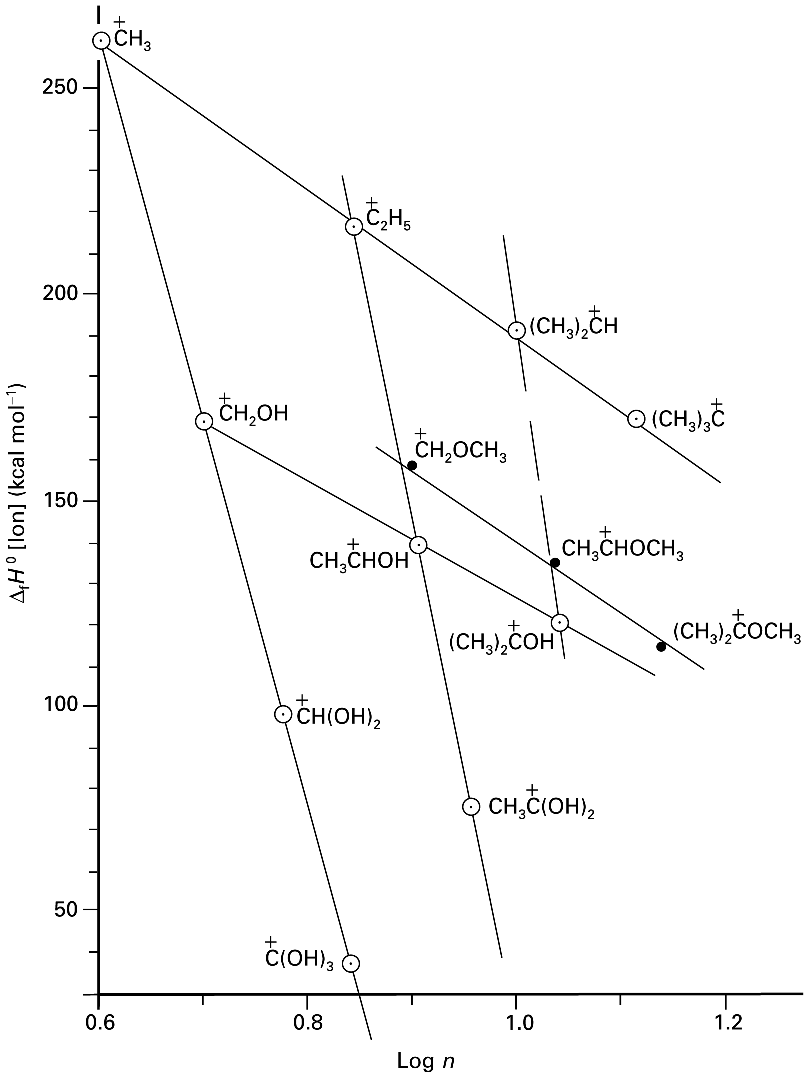



**Figure 2** Plots of $\Delta_f H^0$ [ion]$^+$ versus log (ion size) the latter being represented by *n*, the number of atoms in the ion. The steep curves show the effect of successive OH substitution in alkyl ions and the shallow curves are for $CH_3$ substitution therein.

arise. The best values probably are determined by laser photoelectron spectroscopy and laser photodetachment. Both of these methods directly measure a threshold energy for the removal of an electron from the anion. EA values are listed in the NIST inventory of data. In general, electron-binding energies to molecules and radicals are less precisely known than cationic $\Delta_f H^0$ values. In thermochemical calculations, the term 'gas phase acidity' $\Delta_{acid} H^0[AH]$ is often employed

$$\Delta_{acid} H^0[AH] = BDE(A-H) - EA[A] + IE[H]$$

where BDE(A–H) is the homolytic bond strength (A–H). Alternatively

$$\Delta_{acid} H^0[AH] = \Delta_f H^0[A^-] + \Delta_f H^0[H^+] - \Delta_f H^0[AH]$$

## List of symbols

$E_{rev}$ = energy barrier of reverse reaction; *n* = number of atoms in a molecule; $\Delta_f H[M]$ = molecular heat of formation.

*See also:* **Ion Dissociation Kinetics, Mass Spectrometry; Ion Structures in Mass Spectrometry; Metastable Ions; Photoelectron–Photoion Coincidence Methods in Mass Spectrometry (PEPICO); Proton Affinities; Thermospray Ionization in Mass Spectrometry.**

## Further reading

Bachiri M, Mouvier G, Carlier P and Dubois PE (1980) *Journal de Chimie Physique* **77**: 899- 914.

Bensen SW (1976) *Thermochemical Kinetics*, 2nd edn. New York: Wiley.

Holmes, JL, Fingas M and Lossing FP (1981) *Canadian Journal of Chemistry* **59**: 80- 93.

Hunter PL and Lias SG (1988) Evaluated gas phase basicities and proton affinities of molecules. *Journal of Physical Chemistry Reference Data* **27**: 413- 656.

Lias SG, Bartmess JE, Liebman JF, Holmes JL, Levin RD and Mallard WG (1988) Gas phase ion and neutral thermochemistry. *Journal of Physical Chemistry Reference Data* **17**(suppl 1).

*NIST Chemistry WebBook*: **http://webbook.nist.gov/chemistry/**

Pedley JB, Naylor RD and Kirby SP (1986) *Thermochemical Data of Organic Compounds*. London: Chapman & Hall.

Traegar JC, McLoughlin RG (1981) *Journal of the American Chemical Society* **103**: 3647- 3652.

# Ion Imaging Using Mass Spectrometry

**Albert JR Heck**, University of Utrecht,
The Netherlands, UK

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Applications |

Ion imaging, as described here, is a multiplex detection technique used in gas-phase molecular reaction dynamics to visualize the breaking and formation of chemical bonds. The primary aim of the ion imaging technique is to reveal quantitatively the way in which energy associated with the processes that govern the breaking and formation of chemical bonds is released into the reaction products. The ion imaging technique is based on a combination of two more established techniques: optical spectroscopy (resonance-enhanced multiphoton ionization, REMPI) and two-dimensional time-of-flight mass spectrometry. Ion imaging measures the three-dimensional recoil (velocity, i.e. speed and angular distribution) of the ensuing fragments or reaction products and uni- and bimolecular reactions. Using resonance-enhanced multi-photon ionization, the recoil of individual quantum states (electronic, vibrational and/or rotational) of a selected fragment or reaction product can be measured. More extensive reviews covering various aspects of the ion imaging technique are available.

## Principal aspects

One of the major goals of molecular reaction dynamics, and of ion imaging, is to elucidate how the energy deposited into products of a molecular dissociation or bimolecular reaction is distributed over all available degrees of freedom. Consider a unimolecular dissociation initiated by absorption of a photon. Following dissociation, the nascent fragments will recoil into space, away from the point where the reaction was initiated. The speed of the products will depend on the amount of energy deposited in the translational degrees of freedom, and on the masses of the products. The recoil of the fragments or reaction products does not have to be isotropic. In a laser-induced dissociation, the angular distribution of the photofragments depends on the polarization vector of the photolysis source, on the symmetry of the electronic states involved, and on the lifetime of the activated complex. In the case of a bimolecular reaction, the products will be scattered at various angles with respect to the centre-of-mass velocity vector. Measurement of such recoil will reveal directly whether reaction products are, for example, forward or backward scattered. Additionally, reaction products can be formed with excess internal energy distributed over the electronic, vibrational and rotational degrees of freedom. In ion imaging, the internal energy distributions are typically measured using laser-based techniques, i.e. resonance-enhanced multiphoton ionization, although laser-induced fluorescence imaging has been employed as well.

Simultaneous measurement of all the above quantities, as achieved by ion imaging, is important so that correlated quantities (e.g. the quantum state of one product correlated with that of the other, or the velocities of fragments correlated with the particular quantum state populated) can be determined. A further elegant aspect of the imaging technique is that the recoil of the fragments and/or reaction products is visualized by the ion images appearing on a

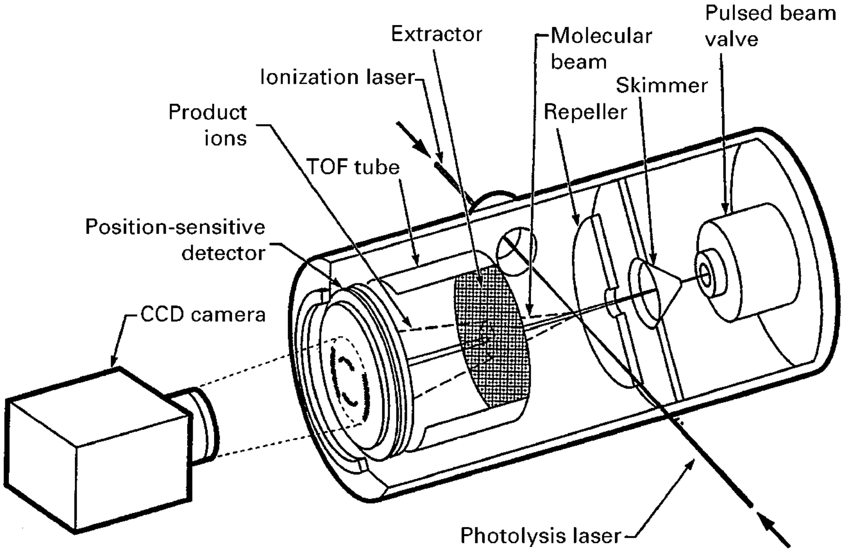



**Figure 1** Schematic diagram of an ion imaging apparatus (TOF = time of flight). Reproduced with permission from Heck AJR and Chandler DW (1995) Imaging techniques for the study of chemical reaction dynamics. *Annual Review of Physical Chemistry* **46**: 335–378. © 1995 Annual Reviews Inc.

phosphor screen. In this way the ion imaging technique provides instructive and direct 'snapshots' of the breaking and making of chemical bonds.

## Experimental aspects

**Figure 1** shows a schematic of an ion imaging apparatus. The basic experimental components of an ion imaging instrument are: (1) molecular beams used to collisionally cool the reactants and to generate reactants with a well-defined velocity; (2) lasers used to photolyse or generate the primary reactants, and to ionize the reaction product quantum-state-selectively; (3) two-dimensional time-of-flight mass spectrometry used to measure the recoil of the reaction products; (4) a position-sensitive detector used to detect two-dimensional time-of-flight profiles. The position-sensitive detector typically consists of a pair of chevron-type microchannel plates (MCP) coupled to a phosphor screen and a charge-coupled-device (CCD) camera. The CCD camera records the images appearing on the phosphor screen.

A typical photofragment imaging experiment starts when a jet-cooled molecule absorbs a laser photon, eventually leading to dissociation. The usual short (~3 ns) and focused laser pulse provides a well-defined starting point in time as well as in space for the experiment. Following the bond cleavage, the nascent fragments start to fly away from that point on three-dimensional spheres. Before the fragments recoil far away, the products of interest are selectively ionized using REMPI. Ionization does not change the recoil of the fragments significantly. The ions are extracted from the reaction region and transferred to the position-sensitive detector through a field-free time-of-flight region. During their flight to the detector, the spheres expand significantly. The flight-times of the fragments, and thus also the physical size of the spheres upon reaching the detector, can be adjusted experimentally by varying the extraction field strengths in the time-of-flight region. On arrival at the detector, the three-dimensional distribution of the ionized fragments is crushed onto the plane of the two-dimensional detector. The position where the ion impinges on the detector is directly fibreoptically coupled to a small spot on the phosphor screen. The light emitted by the phosphor screen is detected by a CCD camera. The measured ion images are therefore two-dimensional projections of the three-dimensional velocity distributions.

## Extraction of information from the images

Several methods can be used to analyse the ion images. The most straightforward approach, analysing the 'raw' images, already provides insight into the dynamical processes involved in the photofragmentation. The two-dimensional projections may reveal different channels involved in the reaction; however, branching ratios cannot be determined quantitatively when contributions of competing channels spatially overlap in the image. Two-dimensional projections show qualitatively the angular distributions of the recoiling fragments. Knowledge of the time of flight of the ionized fragment, the masses of the two

concomitant fragments, the photolysis laser wavelength and the bond energy in the parent molecule provides information about the translational and internal energy of the complementary fragment.

More quantitative information can be obtained from the images when the full three-dimensional speed and angular distributions are reconstructed using mathematical transformations of the crushed two-dimensional images, or alternatively by using forward convolution simulation techniques. If the initial three-dimensional distribution has cylindrical symmetry, a unique transformation - the inverse Abel transform - can be used to reconstruct the initial three-dimensional velocity distribution. As the photolysis laser vector defines automatically an axis of cylindrical symmetry, the inverse Abel transformation can usually be used, as long as the plane of the position-sensitive detector is placed parallel to the laser polarization vector.

The angular distribution of photofragments depends particularly on the symmetry of the electronic transition and the direction of the photolysis laser polarization vector. The angular distribution of photofragments is usually described by the anisotropy parameter $\beta$. In particular, the recoil distribution, which is related to the velocity of the fragment $v$ and the polarization vector of the electromagnetic field $E$, is given by

$$I(\theta) = (4\pi)^{-1}[1 + \beta P_2(\cos\theta)] \qquad [1]$$

where $\theta$ is the angle between $v$ and $E$, $P_2$ is the second-order Legendre polynomial and $\beta$ is the anisotropy parameter. This anisotropy parameter $\beta$ may range from - 1 for a perpendicular transition, i.e. $\mu$ perpendicular to $E$, resulting in a $\sin^2(\theta)$ distribution of the photofragments and 2 for a parallel transition (which results in a $\cos^2(\theta)$ distribution). $\beta = 0$ corresponds to an isotropic distribution of the fragments. Figure 2 shows schematic representations of three-dimensional distributions of photofragments resulting from prompt dissociations following a parallel transition (Figure 2A) and a perpendicular transition (Figure 2B). Following a parallel transition, the fragments will recoil primarily along the laser polarization vector, whereas following a perpendicular transition, the fragments recoil perpendicular to the laser polarization vector.

Figure 3 shows a series of reconstructed ion images of nascent D (deuterium) atoms generated via the photolysis of DI, at three different photolysis wavelengths. These images can be used to illustrate many of the points made above. The upper panels show the inverse Abel-transformed ion images, the

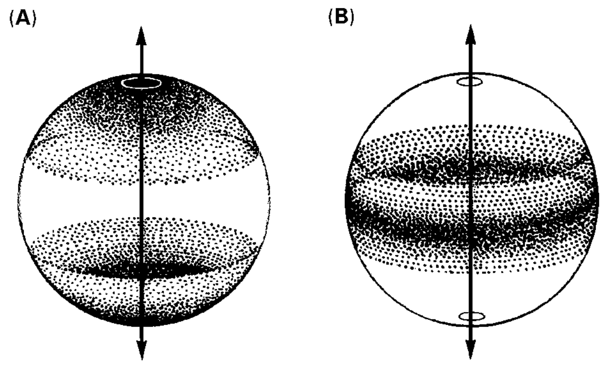



**Figure 2** Schematic diagram of the three-dimensional distributions of photofragments formed following a (A) pure parallel transition in the precursor molecule and (B) a pure perpendicular transition in the precursor molecule. The arrows indicate the direction of the laser polarization vector.

lower panels the corresponding D atom speed distributions. What information is inherent in the ion images of Figure 3? If the nascent D fragments had not acquired any kinetic energy following the photodissociation the ionized $D^+$ ions would have shown up at the centre of the image. From the measurement of the size of the rings and arrival time of the ions at the detector, the speed of the D atoms may be determined. It is immediately clear that the speed of the nascent D atoms is much higher when DI is photolysed using $\lambda = 205$ nm photons instead of $\lambda = 308$ nm photons, since the ion image obtained at $\lambda = 308$ nm is smaller. From these measurements and knowledge of the photolysis energy (or laser wavelength), the bond dissociation energy of the DI molecule may be determined. The photodissociation of HI or DI is possibly one of the simplest, but also most-studied, dissociation processes. At ultraviolet wavelengths two product channels are energetically accessible:

$$\begin{aligned} DI &\rightarrow D + I\,(^2P_{3/2}) \\ DI &\rightarrow D + I^*(^2P_{1/2}) \end{aligned} \qquad [2]$$

where I ($^2P_{3/2}$), or I, is ground-state atomic iodine and I ($^2P_{1/2}$), or I*, is the spin- orbit excited state of iodine, which is 0.94 eV higher in energy than the ground state. Therefore, two distinct D atom translational energies are possible corresponding to the concomitant formation of the two different spin- orbit states of the iodine atom. From analysis of the inverse Abel-transformed ion images we can directly obtain the I/I* branching ratios for the dissociation, and moreover the angular distributions of the D atoms

for the two channels individually. The three images presented in Figure 3 reveal that the branching ratio between the I and I* channel changes quite dramatically with photolysis energy. At low energy, only the I channel is observed and a single ring is seen that corresponds with a pure $\sin^2(\theta)$ angular distribution. As the energy is increased, the I* channel becomes energetically accessible and an additional component, corresponding to a $\cos^2(\theta)$ distribution, is seen in the images. This example of imaging of D atoms from the photolysis of DI reveals the general capabilities of the technique. In a single measurement, many valuable parameters that characterize a molecular dissociation, such as product branching ratios, angular distributions of the nascent fragments and bond dissociation energies, are revealed. At the time of writing, photofragment ion imaging has already been applied to more than 30 molecular systems, in most cases revealing exciting new aspects of the dynamics governing molecular photodissociations.

## New developments

Ion imaging is now, some ten years after its introduction, a well-established tool for the study of molecular reaction dynamics. The most recent years have seen some remarkable progress in the field, in particular in the areas of velocity resolution and photoelectron imaging, complex photodissociations of polyatomic molecules exhibiting many competing reaction pathways, and reaction product imaging of bimolecular reactions.

### Energy resolution, 'velocity mapping'

Parker and Eppink have introduced a modification to the ion optics in ion imaging instruments employing a set of electrostatic lenses. The modified setup does not require grids, which were typically used to cover the large holes (of approximately 20 mm diameter) in the extractor plates, to ensure parallel field lines (see Figure 1). Such grids may cause 'grid

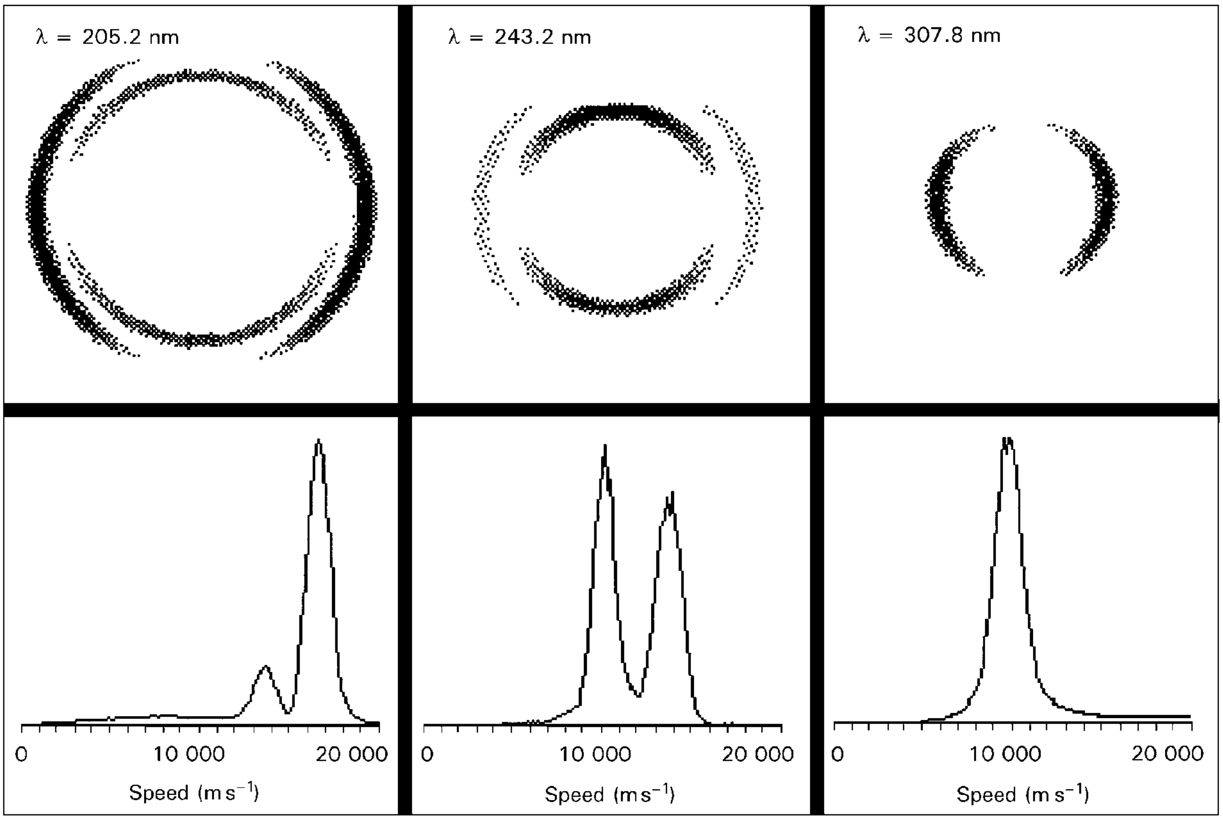



**Figure 3** Inverse Abel-transformed ion images and speed distributions of D atoms formed in the UV photolysis of DI at three different photolysis wavelengths. The faster fragments will have recoiled farther from the centre of the image. For the images obtained at photolysis wavelengths λ = 205 nm and 243 nm, the outer rings correspond with the formation of D atoms concomitantly with ground-state I atoms, whereas the inner rings correspond with the formation of D atoms concomitantly with spin–orbit excited-state I* atoms. At a photolysis wavelength λ = 308 nm, only ground-state I is formed concomitantly with the D atoms. Intensities are shown in a linear grey scale (darker corresponds to higher intensities). Reproduced with permission from Heck AJR and Chandler DW (1995) Imaging techniques for the study of chemical reaction dynamics. *Annual Review of Physical Chemistry* **46**: 335–378. © 1995 Annual Reviews Inc.

distortions' in the measured ion images. Another problem encountered in the original photofragment ion imaging instruments stemmed from the interaction region between the molecular beam and the photolysis laser often being not, as desired, a point source but rather more a line source defined by the physical size of the interaction volume of the molecular beam and the laser beam. Line sources may cause 'blurring' of ion images. To reduce 'blurring' of the images, several computational 'deblurring' algorithms have been introduced. The experimental approach introduced by Parker and Eppink greatly diminishes blurring as it focuses all ions with the same initial velocity to the same spot (i.e. pixel) on the MCP detector and therefore overcomes the need for computational deblurring routines. Focusing conditions are defined by the time-of-flight path length and the ratio of the voltages applied to the repeller and extractor plates. Parker and colleagues demonstrated the improved capabilities of the modified ion imaging apparatus investigating complex competing pathways in photoionization and dissociation of molecular oxygen, excited to the v=2, n = 2 level of the 3dp $^3\Sigma^-_{1g}$ Rydberg state. By absorbing a subsequent photon, these excited intermediate states of molecular oxygen may (a) dissociate generating ground-state and excited-state oxygen atoms (the latter being ionized by a subsequent photon), (b) autoionize forming $O_2^+$ ions in ground or excited states, (c) autoionize and dissociate forming $O^+$ and ground-state and excited-state O atoms, or (d) autoionize and dissociate after absorption of one more photon. In total, more than ten different channels are expected, which normally would hamper any sort of analysis.

An interesting advantage of imaging instruments is that they may be converted, without major modifications, to photoelectron imaging instruments as described by Helm and co-workers, providing an alternative method of angle-resolved photoelectron spectroscopy. Owing to the enhanced resolution and by combining results of the $O^+$ images with the images of the corresponding photoelectrons, Parker and Eppink were able to unravel the relative importance of most of the anticipated channels, all exhibiting different kinetic energy releases and different angular recoil. Figure 4 shows on the left the ion image of the $O^+$ ions and on the right the corresponding image of the photoelectrons. The positions of the rings corresponding to the formation of O* atoms are indicated, as well as the positions of the rings involving intermediate $O_2^+$ ($\nu$) states. Channels separated by less than 100 meV in kinetic energy are still quite well resolved. This improvement, brought about by a relatively simple modification, increases the resolution in the images tremendously and is expected to enhance the capabilities of photofragment ion imaging ever more.

### Ion imaging of complex polyatomic photodissociations

The first ion imaging study, by Chandler and Houston, focused on the molecular dissociation of $CH_3I$. Although, this is a relatively large polyatomic molecule, its dissociation behaviour is comparatively simple as far as possible reaction products are concerned, in that the $H_3C$- I bond cleavage is by far the most dominant process. One of the most notable features of ion imaging is the capability to distinguish different competing reaction pathways even when they produce the same products. The study of complex vacuum ultraviolet (VUV) photolysis of methane benefited greatly from the multiplex detection capabilities of ion imaging described by Zare. Considering only spin-conserving dissociation channels, reactions [3] are energetically accessible following VUV absorption:

$$CH_4(^1A_1) \rightarrow \begin{cases} CH_3\,(^2A_2'') + H \\ CH_2\,(a\,^1A_1) + H_2 \\ CH_2\,(^3B_1) + H + H \\ CH_2\,(a\,^1A_1) + H + H \\ CH\,(^2\Pi) + H_2 + H \end{cases} \qquad [3]$$

Many of these five channels lead to one or more identical reaction products, and the contribution of all these pathways to the VUV photolysis could only be unravelled using the ion imaging technique. $H_2$(v, J) photofragment images revealed that the two channels leading to molecular hydrogen in the photolysis of methane were indeed active. The rovibrational distributions of the $H_2$(v, J) products were determined using REMPI, and by combining REMPI with the imaging technique for each of these two channels individually. Similarly, H atom photofragment images revealed the relative contributions of the pathways leading to atomic hydrogen products.

The group of Reisler has employed ion imaging to study the dynamics evolving on the $S_0$ and $S_1$ excited states of isocyanic acid (HNCO). The unimolecular decomposition of HNCO is equally complex as it involves coupling between different electronic surfaces and competition among different reaction pathways. Following ultraviolet absorption, the HNCO molecule dissociates via the pathways in Equation [4].

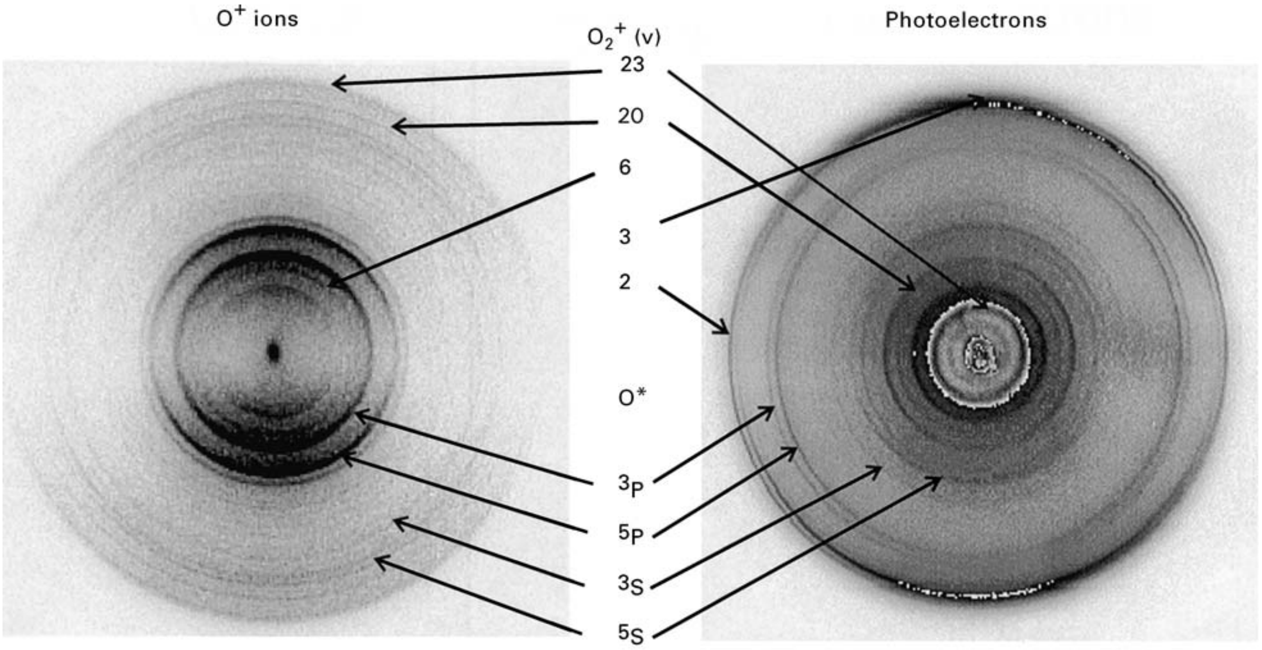



**Figure 4** Photolysis and ionization of $O_2$ through the $v = 2$, $n = 2$ level of the $3dp\,^3\Sigma_{1g}^-$ Rydberg intermediate state. The figure on the left shows a raw $O^+$ ion image, whereas the figure on the right shows the corresponding photoelectron image. The positions of rings that correspond with $O_2^+$ vibrational levels and excited O* atoms are indicated in both images. The laser polarization vector was vertical for both these images. Intensities are shown in a linear grey scale (darker corresponds to higher intensities). Reproduced with permission from Parker DH and Eppink AJJB (1997) Photoelectron and photofragment velocity map imaging of state-selected molecular oxygen dissociation/ionization dynamics. *Journal of Chemical Physics* **107**: 2357–2362.

$$\mathrm{HNCO} \rightarrow \begin{cases} \mathrm{H}(^2\mathrm{S}) & + \mathrm{NCO}\,(\mathrm{X}\,^2\Pi) \\ \mathrm{NH}(\mathrm{X}\,^3\Sigma^-) & + \mathrm{CO}\,(\mathrm{X}\,^1\Sigma^+) \\ \mathrm{NH}(\mathrm{a}\,^1\Delta) & + \mathrm{CO}\,(\mathrm{X}\,^1\Sigma^+) \end{cases} \qquad [4]$$

Although the formation of triplet NH is the least endothermic reaction path, this channel is formally not allowed. However, this spin-forbidden channel has been observed. Initially, the two latter channels contributing to the formation of ground-state CO molecules, concomitant with triplet and singlet NH respectively, were identified using laser-induced fluorescence of the singlet and triplet NH separately. However, such an approach does not allow the identification of the relative branching ratios of these two channels, unless the transition strengths of the probed transitions are known quantitatively. Therefore, Reisler and colleagues studied the formation of ground-state CO by photofragment ion imaging, using state-selective REMPI on CO (v = 0,J). Probing only the CO fragments, the two channels are not resolved spectroscopically, but they can be resolved in velocity space using ion imaging: the two different reaction pathways lead to different velocities of the CO molecules. Figure 5 shows an ion image obtained for CO with zero quanta of vibration and 10 quanta

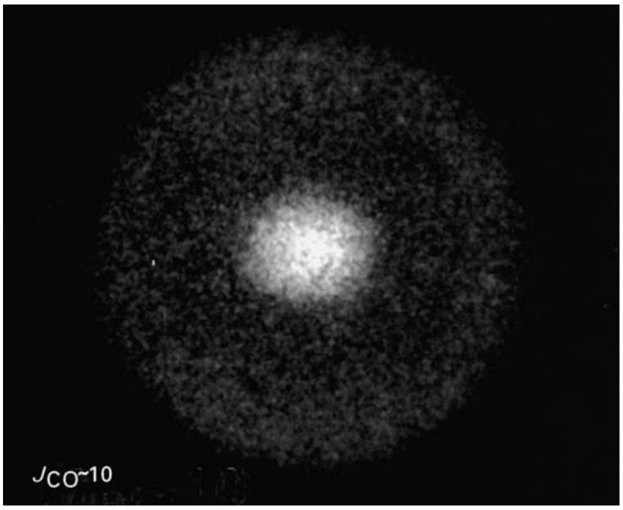



**Figure 5** Ion image of CO (v = 0, $J = 10$) formed in the photolysis of HNCO at $\lambda = 230.1$ nm. The outer part of the image originates from CO molecules formed concomitantly with triplet NH, whereas most of the CO molecules in the inner, more dense area of the image originate from CO molecules formed concomitantly with singlet NH. Intensities are shown in an inverse linear grey scale (darker corresponds to lower intensities). Reprinted with permission of Elsevier Science-NL from Droz-Georget T, Zyrianov M, Reislev H and Chandler DW (1997) Correlated distributions in the photodissociation of HNCO to NH($a^1\Delta$ + CO($X_1\Sigma^+$) near the barrier on $S_1$. *Chemical Physics Letters* **276**: 316–324.

of rotation. CO (v = 0, J = 10) fragments are formed concomitantly with triplet as well as singlet NH. Singlet NH is approximately 12 500 $cm^{-1}$ higher in energy than triplet NH. The singlet pathway leads to the smaller, somewhat more intense image, as less energy is available following this decomposition pathway, whereas the triplet channel leads to the bigger image, i.e. higher average speed of the CO molecules. As in the methane study, inverse Abel transformation of the images allowed the determination of rotational distributions of the CO fragments formed via these separate channels.

## Reaction product ion imaging

Photofragment ion imaging is now an established method for the measurement of photofragment velocity (speed and angle) distributions from unimolecular dissociation processes. Recently, the domain of the ion imaging technique has been expanded to investigate bimolecular reactions as well, with the aim of determining differential cross sections in a single measurement. Houston and colleagues have studied the inelastic scattering reaction of NO molecules off argon atoms. In this study, a molecular beam of NO was crossed with a molecular beam of argon. The scattering of NO molecules was investigated for NO products in selected rotational and spin- orbit levels. The scattering of NO was visualized directly using the reaction product imaging technique, through selective ionization of the selected NO quantum state, and appeared to vary quite dramatically depending on the J level of the NO product.

Chandler and colleagues studied the bimolecular reaction $H + D_2 \rightarrow D + HD$. In their study, an atomic hydrogen beam (generated through the photolysis of HI) of well-defined velocity was crossed with a molecular beam of $D_2$. The velocity distribution of the deuterium atoms formed in the reaction was studied at two different centre-of-mass collision energies of 0.54 and 1.29 eV. Figure 6 shows the D

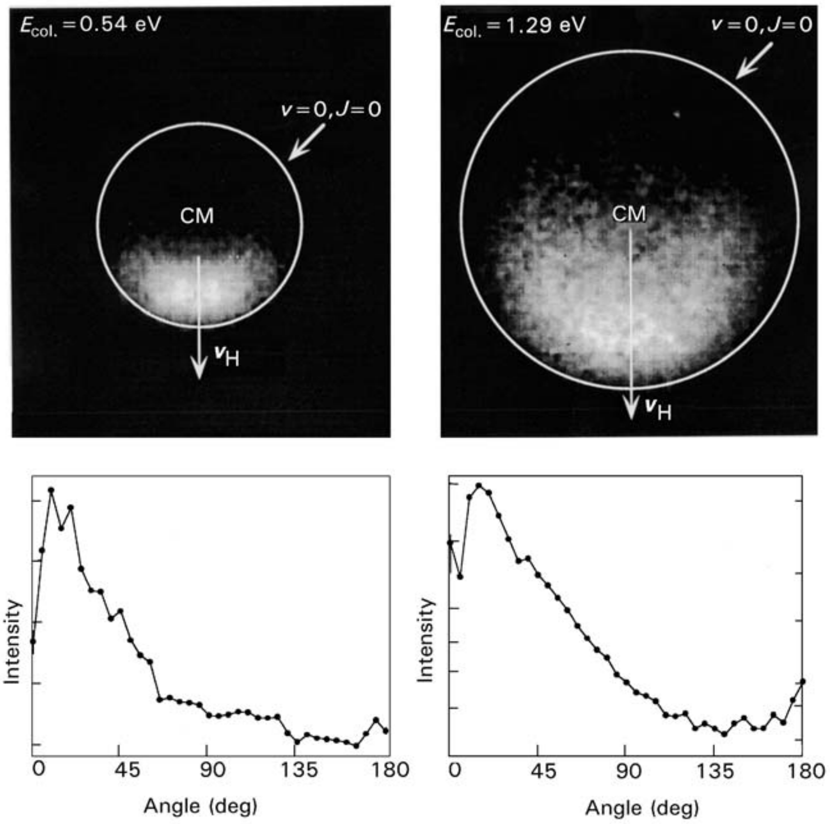



**Figure 6** Raw images of D atoms formed in the reaction of $H + D_2$. Intensities are shown in an inverse linear grey scale (darker corresponds to lower intensities). The left and right columns show results obtained for this reaction at centre-of-mass collision energies of 0.54 eV and 1.29 eV, respectively. In the images, the centre-of-mass (CM) of the reaction and the direction of the H atoms are indicated. The circles indicate the calculated positions of the fastest D atoms corresponding to formation of HD in the *v* = 0, *J* = 0 state. At the bottom of the figure the differential cross-sections, obtained from the inverse Abel-transformed images are displayed. Reproduced with permission from Heck AJR and Chandler DW (1995) Imaging techniques for the study of chemical reactions dynamics. *Annual Review of Physical Chemistry* **46**: 335–378. © 1995 Annual Reviews Inc.

atom reaction product images at these two collision energies. The circles represent the maximum speed the D atoms could have, corresponding to the concomitant formation of HD in its rotational and vibrational ground states. It is directly obvious from these images that the nascent D atoms are predominately forward scattered with respect to the direction of the H atom beam velocity. The differential cross sections obtained from the inverse Abel-transformed images for the $H+D_2 \rightarrow D+HD$ reaction at these two different collision energies agreed fairly well with theoretically predicted differential cross sections. Reaction product imaging offers several advantages over more traditional methods used to study bimolecular reactive scattering in that the differential cross sections of the products, possibly in a single selected quantum state, are obtained in a single measurement.

## List of symbols

$I(\theta)$ = photo-fragment angular distribution; J = rotational quantum number; $P_n$ = nth order Legendre polynomial; $\beta$ = anisotropy parameter; $\mu$ = mass; $\nu$ = vibrational quantum number; $\theta$ = angle between $\nu$ and E; $\lambda$ = photon wavelength.

*See also:* **Fragmentation in Mass Spectrometry; Gas Phase Applications of NMR Spectroscopy; Ion Dissociation Kinetics, Mass Spectrometry; Ion Energetics in Mass Spectrometry; Photoionization and Photodissociation Methods in Mass Spectrometry; Time of Flight Mass Spectrometers.**

## Further reading

Bontuyan LS, Suits AG, Houston PL and Whitaker BJ (1993) State-resolved differential cross-sections for crossed-beam Ar-NO inelastic scattering by direct imaging. Journal of Physical Chemistry 97: 6342-6350.

Bordas C, Paulig FH, Helm H, Huestis DL (1996) Photoelectron imaging spectrometry: Principle and inversion method. Review of Scientific Instruments 67: 2257-2268.

Chandler DW and Houston PL (1987) Two-dimensional imaging of state-selected photodissociation products detected by multiphoton ionization. Journal of Chemical Physics 87: 1445-1447.

Droz-Georget T, Zyrianov M, Reisler H and Chandler DW (1997) Correlated distributions in the photodissociation of HNCO to $NH(a^1\Delta) + CO(X_1\Sigma^+)$ near the barrier on $S_1$. Chemical Physics Letters 276: 316-324.

Eppink ATJB and Parker DH (1997) Velocity map imaging of ions and electrons using electrostatic lenses: Application in photoelectron and photofragment ion imaging of molecular oxygen. Review of Scientific Instruments 68: 3477-3484.

Heck AJR (1997) Throwing light on molecules falling apart: photofragment ion imaging. European Mass Spectrometry 3: 171-183.

Heck AJR and Chandler DW (1995) Imaging techniques for the study of chemical reaction dynamics. Annual Reviews of Physical Chemistry 46: 335-372.

Heck AJR, Zare RN and Chandler DW (1996) Photofragment imaging of methane. Journal of Chemical Physics 104: 4019-4030.

Houston PL (1995) Snapshots of chemistry: product imaging of molecular reactions. Accounts of Chemical Research 28: 453-460.

Houston PL (1996) New laser-based and imaging methods for studying the dynamics of molecular collisions. Journal of Physical Chemistry 100: 12757-12770.

Kitsopoulos TN, Buntine MA, Baldwin DP, Zare RN and Chandler DW (1993) Reaction product imaging: The $H + D_2$ reaction. Science 260: 1605-1610.

Parker DH and Eppink ATJB (1997) Photoelectron and photofragment velocity map imaging of state-selected molecular oxygen dissociation/ionization dynamics. Journal of Chemical Physics 107: 2357-2362.

Whitaker BJ (1993) Photo-ion imaging and future directions in reactive scattering. In Compton RG and Hancock G (eds) Research in Chemical Kinetics, Vol 1, pp 307-346. Oxford: Elsevier.

# Ion Molecule Reactions in Mass Spectrometry

**Diethard K Böhme**, York University, Toronto, Ontario, Canada

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Mass spectrometry has played a dominant role in measurements of ion–molecule reactions. These measurements have had broad impact, in both fundamental and applied areas of science. For example, they have provided extensive thermodynamic information about ions and molecules including standard enthalpies of ion formation and the electron affinities, ionization energies and proton affinities of neutral molecules. They have found important applications in analytical and biomedical mass spectrometry and in plasma chemistry. In physical-organic chemistry they have provided fundamental insight into the intrinsic efficiencies of ion–molecule reactions with analogues in solution (e.g. proton transfer (acid–base), nucleophilic substitution ($S_N2$) and metal–ion coordination reactions) and into the influence of solvation on these efficiences. Also, they have provided a deeper understanding of the chemistry of the earth's ionosphere and the ionospheres of other planets, of chemical synthesis in interstellar and circumstellar clouds, of flame chemistry, of radiolysis and of ion-induced polymerization.

## Ion–molecule reaction measurements

Ion–molecule reactions have been studied using a large variety of mass-spectrometric techniques for over 3 decades over a large range in pressure and temperature.

Low pressure (from ~$10^{-5}$ to $10^{-7}$ torr) measurements have been performed with Fourier transform (FT) and ion-cyclotron resonance (ICR) mass spectrometers and with ion traps. Moderate pressure (from 0.1 to 5 torr) techniques used to investigate ion–molecule reactions include high-pressure mass spectrometry, the flowing afterglow (FA) technique, the selected-ion flow tube (SIFT) technique and the variable-temperature selected-ion flow-drift tube (VT-SIFDT) technique. The latter technique operates over a temperature range from 85 to 550 K and at average kinetic energies up to at least 1 eV. Measurements at atmospheric pressure have been restricted largely to ion mobility spectrometry and flame-ion mass spectrometry. The latter provides data in the high-temperature range from 1800 to 2400 K. Measurements at temperatures below 160 K have been achieved in a cooled Penning trap, a drift tube and a flow reactor that utilizes cooling by adiabatic expansion. A free jet-flow reactor has been employed to study ion–molecule reaction rate coefficients at temperatures as low as 0.3 to 30 K.

## Rates of ion–molecule reactions

Reactions between ions and molecules are among the fastest chemical reactions known. This is because the electrostatic interaction between the charge on the ion and the polarizable (or polar) molecule draws these reactants together at long range. Furthermore, the resulting interaction energy at short range is often sufficiently large to overcome intrinsic energy barriers to chemical change. As a consequence, exothermic ion–molecule reactions often proceed very rapidly at room temperature *at the capture rate*, without an activation energy. This is in sharp contrast to molecule–molecule reactions that normally involve an activation energy. Generalized profiles for the dependence of reactant energy on the reaction coordinate for ion–molecule and molecule–molecule reactions are shown in **Figure 1**. Ion–molecule reactions are often best described by a double-minimum, rather than a single-minimum, potential energy profile.

The reaction rate coefficient, $k_r$, for an exothermic ion–molecule reaction proceeding without an activation energy is equal to the capture or collision rate coefficient, $k_c$, which is typically ~$1 \times 10^{-9}$ cm$^3$ molecule$^{-1}$ s$^{-1}$. This is some one to two orders of magnitude larger than that for an exothermic molecule–molecule reaction proceeding without an activation energy. Several theories are available that allow calculation of the capture or collision rate coefficient, $k_c$, for an ion–molecule collision. According to the most popular theory, the 'average dipole orientation' (ADO) theory, the magnitude of $k_c$ is determined by the physical properties of the reactants: the polarizability, $\alpha$, and permanent dipole, $\mu_D$, of the reacting molecule, the charge on the ion, $q$, and the reduced mass, $\mu$, of the colliding reactants. The collision-rate coefficient $k_c$ can be calculated using ADO theory from the following equation:

$$k_c = \frac{2\pi q}{\mu^{1/2}}\left[\alpha^{1/2} + C\mu_D \frac{2^{1/2}}{(\pi k_B T)^{1/2}}\right]$$

where $C$ is a correction factor that is a function of $\alpha$ and $\mu_D$, and $k_B$ is Boltzmann's constant. The ratio of the reaction rate coefficient to the collision rate coefficient, $k_r/k_c$, is a measure of reaction probability per collision or reaction efficiency. Although many exothermic ion–molecule reactions proceed with unit reaction probability at room temperature, others exhibit reaction probabilities much less than unity.

## Temperature/kinetic energy dependence

Ion–molecule reactions that are rapid and efficient at room temperature frequently remain rapid as the temperature of the reactants is varied and follow the temperature dependence of the collision rate. Slow reactions at room temperature almost always increase in rate with decreasing or increasing temperature or kinetic energy (see **Figure 2**). This can be understood in terms of the double-minimum potential shown in **Figure 1** and the following reaction mechanism:

$$A^+ + B \rightleftharpoons (AB)^{+*} \rightarrow C^+ + D$$

As the reactants approach each other they are attracted by electrostatic ion–induced dipole and ion–dipole forces and form the 'entrance channel' complex $(AB)^{+*}$. Further advance along the minimum potential energy path ultimately leads to an increase of potential energy due to bond redisposition. Then the complex exits along the 'product channel' and finally dissociates into products. The two potential energy minima (or 'wells') are separated by a barrier that controls the overall reaction efficiency, even though the barrier often lies below the initial energy of the reactants. A rate of dissociation back to reactants that decreases more rapidly with decreasing temperature than the rate of the dissociation of $(AB)^{+*}$ into products leads to an increase in overall rate with decreasing temperature. An increase in rate at higher temperatures often results from the emergence of new vibrational, electronic or chemical channels.

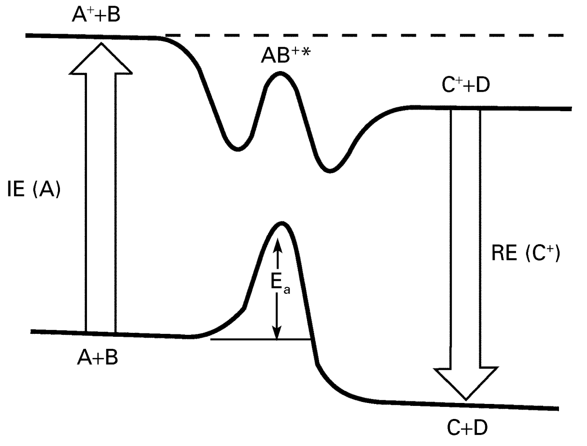



**Figure 1** Generalized profiles for the dependence of reactant energy on the reaction coordinate for ion–molecule and molecule–molecule reactions. IE and RE represent ionization energy and recombination energy, respectively. Some ion–molecule reactions are better represented by a single- rather than a double-minimum potential.

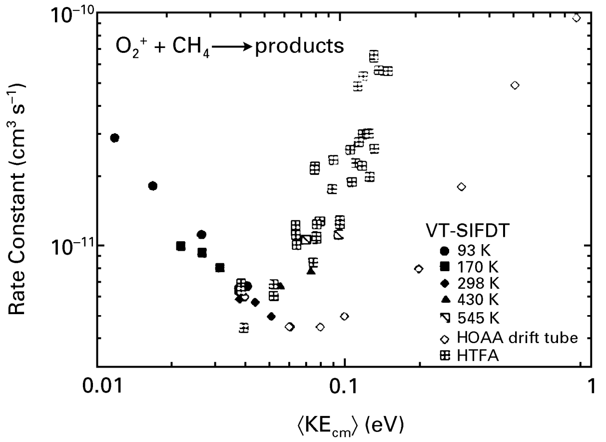



**Figure 2** Rate coefficients for the reaction of $O_2^+$ with methane measured as a function of average kinetic energy at several temperatures in a variable-temperature selected-ion flow drift tube (VT-SIFDT). Also shown are data taken with a high-temperature flowing afterglow (HFTA) and a drift tube. Reproduced with permission from Viggiano AA and Morris MA (1996) *Journal of Physical Chemistry* **100**: 19227–19240.

Endothermic ion–molecule reactions and ion–molecule reactions with an activation energy have been examined over a wide range of the kinetic energy of the reactant ion (from ~0.1 to 600 eV in the laboratory frame). They exhibit a reaction threshold with increasing ion kinetic energy, rapidly increase in reaction rate and then slow down again at higher kinetic energies (see **Figure 3**). An analysis of the kinetic energy dependence of the reaction cross-section can provide thermodynamic information about the reacting ions and neutral molecules.

Rotational and translational energy have been found to be equally effective in driving endothermic reactions. For exothermic reactions large rotational effects appear mainly when one or both of the

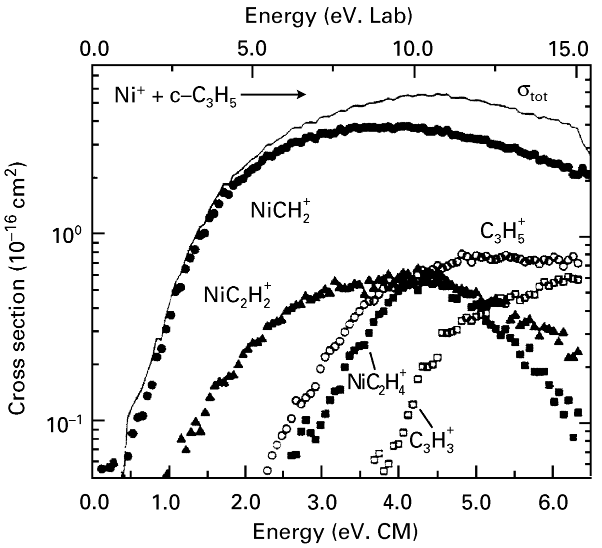



**Figure 3** Observed variations of product cross-sections with translational energy in the centre-of-mass frame of reference (lower scale) and the laboratory frame (upper scale) for reactions of cyclopropane with $Ni^+$. Reproduced with permission from Fisher ER and Armentrout PB (1990) *Journal of Physical Chemistry* **94**: 1674–1683.

reactants have a large rotational constant. Vibrational effects are more varied: in some reactions vibrational excitation in the anticipated reaction coordinate can strongly affect reaction probability.

## Types of ion–molecule reaction

Reactions between ions and molecules can lead to a variety of consequences ranging from the simple transfer of an electron between an ion and a molecule to the nominal transfer of a large molecular fragment accompanied by extensive chemical bond redisposition in both the molecule and the ion. **Table 1** summarizes various types of ion–molecule reaction that have been investigated involving singly charged ions, but recently many also have been observed with multiply charged ions.

**Table 1** Types of ion–molecule reaction

| *Reaction type* | |
|---|---|
| Electron (charge) transfer | $A^{\pm} + B \rightarrow B^{\pm} + A$ |
| Proton transfer | $AH^+ + B \rightarrow BH^+ + A$ |
| | $A^- + BH \rightarrow AH + B^-$ |
| H-atom transfer | $A^{\pm} + BH \rightarrow AH^{\pm} + B$ |
| $R^+$ transfer | $AR^+ + B \rightarrow BR^+ + A$ |
| Nucleophilic displacement | $A^- + RX \rightarrow AR + X^-$ |
| Bond redisposition | $AB^{\pm} + C \rightarrow AC^{\pm} + B$ |
| Radiative association | $A^{\pm} + B \rightarrow AB^+ + h\nu$ |
| Collisional association | $A^{\pm} + B + M \rightarrow AB^{\pm} + M$ |
| Switching | $A^{\pm}.B + C \rightarrow A^{\pm}.C + B$ |
| Associative detachment | $A^- + B \rightarrow AB + e$ |

### Electron-transfer reactions

Electron transfer between ions and molecules at thermal energies is subject to energy resonance and Franck–Condon effects. When small ions are involved, electron transfer is generally inefficient unless there exists a near-resonance between an energy level of the product ion and an available recombination energy of the reactant ion. In addition there must be large Franck–Condon factors connecting the upper and lower states involved in the transition. Energy resonance is necessary, but not always sufficient. For electron-transfer reactions involving triatomic or polyatomic species, the density of states is much larger so that an energy match and Franck–Condon requirement can almost always be obtained. Vibrational-energy dependence is often observed in electron-transfer reactions presumably due to energy resonance and Franck–Condon effects. Low energy electron-transfer collisions often produce products with large amounts of internal energy that can lead to the dissociation of the product ion. The dissociation products provide a signature of the reactant molecule so that electron transfer is useful in the analysis of gases by chemical ionization mass spectrometry. Electron transfer with multiply charged ions leads to charge separation according to the following reaction:

$$A^{n+} + B \rightarrow A^{(n-1)+} + B^+$$

Such reactions involve a ‘reverse activation energy’ resulting from the Coulombic repulsion between the product ions and as a consequence exhibit a delayed onset.

### Proton-transfer reactions

Proton transfer, either from a protonated molecule to a neutral molecule or from a neutral molecule to a negative ion, often proceeds at or near the collision rate when exothermic (see **Figure 4**). When the exothermicity is large, internal energy appearing in the protonated product may be sufficient to cause dissociation with the loss of one or more neutral fragments. Thus, as was the case with dissociative electron-transfer reactions, the product ions again provide a signature of the reactant molecule so that these reactions also are useful in the analysis of gases by chemical ionization mass spectrometry. Virtually any molecule can be chemically ionized with an appropriate proton donor and often this is possible

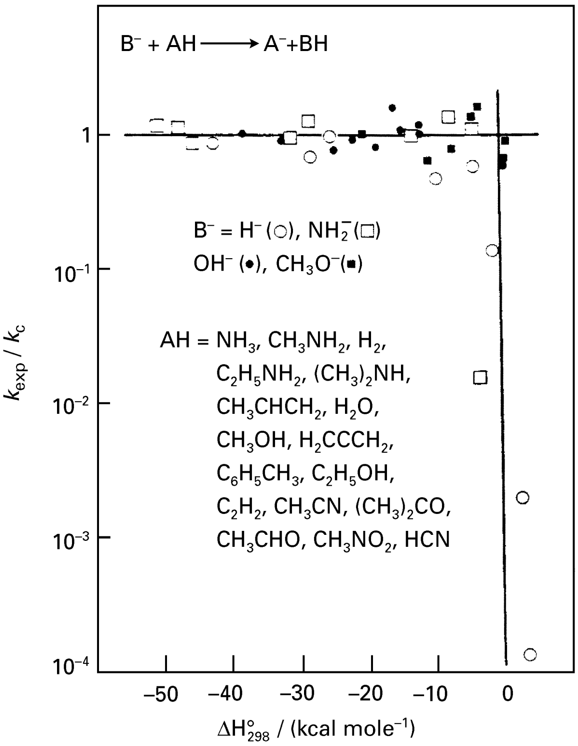



**Figure 4** A correlation between the efficiency of proton transfer, $k_{exp}/k_c$, and its overall change in enthalpy, $\Delta H^o$, at room temperature. The measurements were taken using the flowing-afterglow technique. Reproduced with permission from Bohme DK (1981) *Transactions of the Royal Society of Canada* **19**: 265–284.

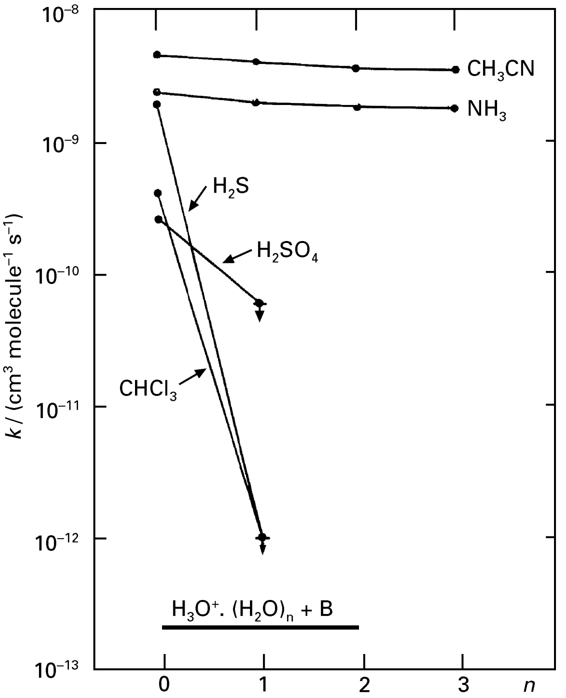



**Figure 5** Observed variations with the number of water molecules in the rate coefficient for proton-transfer reactions between hydrated hydronium ions and various molecules B at room temperature. The measurements were taken using the flowing-afterglow technique. Adapted with permission from Bohme DK, Mackay GI and Tanner SD (1979) *Journal of the American Chemical Society* **101**: 3724–3730.

with high efficiency and selectivity. Solvation can influence the rate of proton transfer as is illustrated in **Figure 5**, but proton transfer generally remains fast until solvation renders the reaction endothermic. There is a growing interest in the ion chemistry of multiply protonated biological ions such as multiply protonated peptides and proteins. Proton transfer from multiply protonated cations may occur according to the following reaction:

$$AH_n^{n+} + B \rightarrow AH_{n-1}^{(n-1)+} + BH^+$$

Again, a 'reverse activation energy' is involved in such reactions as a consequence of the Coulombic repulsion between the product ions.

## H-atom and $R^+$ transfer and bond redisposition reactions

Successive H-atom transfer reactions can lead to the hydrogenation of unsaturated positive or negative ions. They find importance, for example, in the conversion of atomic ions to hydrogenated polyatomic ions in interstellar clouds rich in molecular hydrogen. There appears to be no systematic or easy way to predict the efficiency of such ion–molecule reactions other than to say that hydrogen-atom transfer is rarely observed when electron-transfer or proton-transfer channels are exothermic. Similar comments apply to the transfer of larger $R^+$ species and more general bond redisposition reactions.

## Nucleophilic displacement reactions

Nucleophilic displacement reactions, reactions that formally also involve the transfer of $R^+$, have been explored in great depth in the gas phase with the majority of available mass spectrometric techniques. In part this has been because of the importance of $S_N2$ reactions in solution chemistry and their extreme sensitivity to solvent, both in solution and in the gas phase. Unsolvated $S_N2$ reactions are characterized by double minima in the reaction energy profile as shown in **Figure 1** and the following reaction mechanism:

$$A^- + RX \rightleftharpoons A^- \ldots RX \rightleftharpoons AR \ldots X^- \rightarrow AR + X^-$$

The central barrier in the double-minimum potential energy profile, a Walden-inversion transition state, may lie above or below the initial energy of the reactants. The rate coefficients for $S_N2$ reactions show a slight negative temperature dependence.

Gas-phase measurements also have provided information on the dependence of reaction efficiency on the stepwise molecular solvation of the nucleophile $A^-$ in reactions of the following type, where X can be a halogen atom, R=$CH_3$ and S a solvent molecule:

$$A^-S_n + RX \rightarrow AR + X^-S_m + S_{n-m}$$

Such reactions are also characterized by double minima in the reaction-energy profiles. Because of the lower stabilization of the charge-delocalized transition state relative to the charge-localized reactant and product ions, these profiles converge to a single-maximum profile typical for $S_N2$ reactions in solution as the number of solvent molecules on the reactant and product ions is increased. Gas-phase measurements of the influence of stepwise solvation on reaction efficiency have shown that solvation by even one or two molecules can reduce the specific rate for nucleophilic displacement precipitously (see **Figure 6**). Retention of all the solvent molecules in the product ion appears to be the exception rather than the rule. Limited measurements indicate that increasing solvation increases the magnitude of the negative temperature dependence.

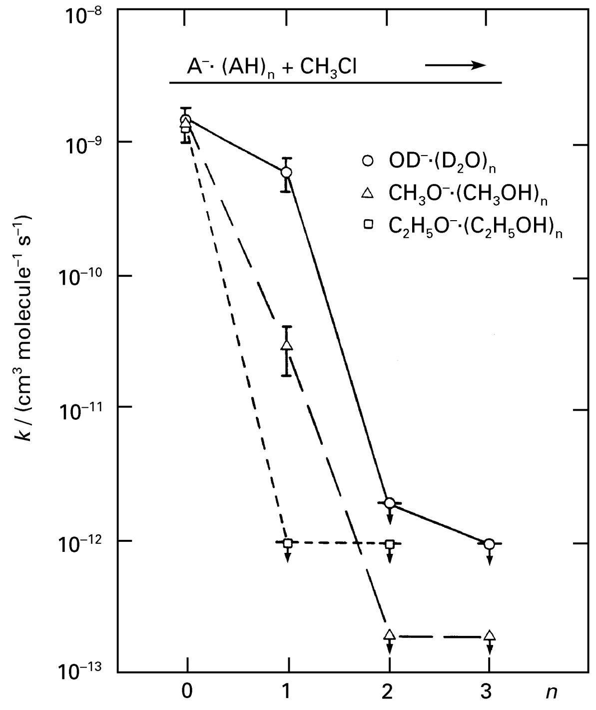



**Figure 6** Observed variations with the number of solvent molecules in the rate coefficient for nucleophilic-displacement reactions between various solvated nucleophiles and methyl chloride at room temperature. The measurements were taken using the flowing-afterglow technique. Reproduced with permission from Bohme DK and Raksit AB (1984) *Journal of the American Chemical Society* **106**: 3447–3452.

## Association reactions

Association reactions lead to the growth of an ion though the formation of new bonds ranging from weak electrostatic to strong covalent bonds. For example, association reactions may lead to the solvation of an ion by weak electrostatic or hydrogen bonding, to ion ligation involving bonds of intermediate strength, or even to strong chemical bond formation. Two steps are required for association as shown below: the formation of an intermediate excited adduct ion $(AB)^{\pm *}$ followed by its de-excitation.

$$A^{\pm} + B \rightleftharpoons (AB)^{\pm *}$$
$$(AB)^{\pm *} \rightarrow AB^{\pm} + h\nu$$
$$(AB)^{\pm *} + M \rightarrow AB^{\pm} + M$$

Radiative association occurs when the intermediate becomes stabilized by the emission of a photon ($h\nu$), usually in the infrared spectral region. Collisional association results when the intermediate is stabilized by collision with a third molecule M. Although radiative association is a bimolecular and therefore second-order process, termolecular collisional association is third-order at low pressures and becomes second-order at high pressures when the frequency of collisions with M exceeds the frequency of the unimolecular dissociation of the intermediate ion $(AB)^{\pm *}$. The rate coefficient for collisional association exhibits a *negative* temperature dependence. Also, FT-ICR measurements have shown that the rate coefficient for radiative association has a strong negative temperature dependence.

Radiative association finds importance in high-vacuum ion-trapping instruments and the low-pressure environments of interstellar space. In contrast, collisional association predominates in high-pressure, low-temperature environments. When they occur *sequentially*, collisional association reactions can play a vital role in gas-phase ion solvation (as occurs naturally, for example, in the earth's lower ionosphere), in ion-cluster formation generally, and in the coordination or multiple derivatization of the type important in organometallic chemistry or in ion-

induced polymerization, respectively. The following are examples of ion solvation, ion coordination and ion-induced polymerization, respectively:

$$H_3O^+(H_2O)_n + H_2O + M \rightarrow H_3O^+(H_2O)_{n+1} + M$$
$$Fe^+(C_2H_2)_n + C_2H_2 + M \rightarrow Fe^+(C_2H_2)_{n+1} + M$$
$$C_{60}^{2+}(\text{allene})_n + \text{allene} + M \rightarrow C_{60}^{2+}(\text{allene})_{n+1} + M$$

In the latter reaction involving a multiply-charged reactant ion, intramolecular charge separation is expected to occur and, in this case, may even act to direct the polymerization away from the surface of $C_{60}$.

### Switching reactions

Adduct ions formed by radiative or collisional association may be transformed by bimolecular switching reactions in which one solvent or ligand molecule is replaced by another. Such reactions may occur at any extent of solvation or ligation and are usually efficient at room temperature when they are exothermic.

### Associative detachment reactions

Associative detachment reactions are unique to negative ions. The energy profiles for such reactions involve one or more curve-crossings between one or more of the attractive potential curves of the reactants and the potential curve for the product AB. AB is formed by the rapid autodetachment of an electron from $(AB)^{\pm *}$ at short interaction distances. Exothermic associative detachment for atomic reagents is inevitably fast. Inhibiting factors can arise for molecular systems.

## Sources of information

For up-to-date accounts in the rapidly evolving field of ion chemistry, the reader is directed to volumes 1 to 3 and future volumes in the series *Advances in Gas Phase Ion Chemistry* edited by NG Adams and LM Babcock. For earlier reports on various aspects of ion–molecule reactions see the NATO ASI Series edited by P Ausloos, by P Ausloos and SG Lias and by MA Almoster Ferreira. Another source of earlier information is the *Gas Phase Ion Chemistry* series edited by MT Bowers.

There have been a number of tabulations of rate coefficients for gas-phase ion–molecule reactions that include leading references to individual ion–molecule reactions. For the most part they are restricted to limited areas of interest. They include tabulations of rate constants measured with flow reactors (DL Albritton), rate constants of interest in modelling the chemistry of planetary atmospheres, cometary comae and interstellar clouds (VG Anicich), rate constants for reactions involving organic ions containing only C and H (LW Sieck) and those containing more than C and H (LW Sieck and SG Lias). An early, more general, tabulation has been published by Talroze. The most comprehensive tabulation of rate constants for both bimolecular and termolecular reactions determined up to 1986 has been published by Ikezoe and co-workers.

*See also:* **Chemical Ionization in Mass Spectrometry; Cosmochemical Applications using Mass Spectrometry; Interstellar Molecules, Spectroscopy of; Ion Energetics in Mass Spectrometry; Ion Trap Mass Spectrometers; Mass Spectrometry, Historical Perspective; SIFT Applications in Mass Spectrometry.**

## Further reading

Adams NG and Babcock LM (eds) (1992) *Advances in Gas Phase Ion Chemistry*, Vol 1. Greenwich, CT: JAI Press.

Adams NG and Babcock LM (eds) (1996)*Advances in Gas Phase Ion Chemistry*, Vol 2. Greenwich, CT: JAI Press.

Adams NG and Babcock LM (eds) (1998) *Advances in Gas Phase Ion Chemistry*, Vol 3. Greenwich, CT: JAI Press.

Albritton DL (1978) *Ion–Molecule Reaction-Rate Coefficients Measured in Flow Reactors Through 1977.* Atomic Data Nuclear Data Tables 22: 1–101. Academic Press.

Almoster Ferreira MA (ed) (1984) *Ionic Processes in the Gas Phase,* NATO ASI Series C, Vol 118. Dordrecht: Reidel.

Ausloos P (ed) (1975) *Interactions between Ions and Molecules,* NATO ASI Series B, Vol 6. New York: Plenum.

Ausloos P (ed) (1979) *Kinetics of Ion-Molecule Reactions,* NATO ASI Series B, Vol 40. New York: Plenum.

Anicich VC (1993) A survey of bimolecular ion-molecule reactions for use in modeling the chemistry of planetary atmospheres, cometary comae and interstellar clouds. *Astrophysical Journal. Supplement Series* **84**: 215–315.

Anicich VC (1993) Evaluated bimolecular ion-molecule gas phase kinetics of positive ions for use in modeling the chemistry of planetary atmospheres, cometary comae and interstellar clouds. *Journal of Physical Chemistry Reference Data* **22**: 1469–1569.

Ausloos P and Lias SG (eds) (1987) *Structure/Reactivity and Thermochemistry of Ions,* NATO ASI Series C, Vol 193. Dordrecht: Reidel.

Bowers MT (ed) (1979) *Gas Phase Ion Chemistry*. New York: Academic.

Sieck LW and Lias SG (1976) Rate coefficients for ion-molecule reactions: ions containing C and H. *J. Phys. Chem. Ref. Data* **5**, 1123–1146.

Sieck LW (1979) *Rate Coefficients for Ion-Molecule Reactions: Ions Other Than Those Containing Only C and H.* National Standards Reference Data System – National Bureau of Standards 64, pp 1–20.

Ikezoe Y, Matsuoka S, Takebe M and Viggiano A (1987) *Gas Phase Ion–Molecule Reaction Rate Constants Through 1986*. Tokyo: Maruzen (Distributors).

Virin LI, Dzhagatspanyan RB, Karachevtsev GV and Talroze VL (1979) *Rate Coefficients for Ion-Molecule Reactions* (in Russian). Moscow: Nauka.

# Ion Structures in Mass Spectrometry

**Peter C Burgers**, Hercules European Research Center B.V., Barneveld, The Netherlands
**Johan K Terlouw**, McMaster University, Hamilton, Ontario, Canada

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
| --- |
| Theory |

## Motivation and methods

During the early days in the development and application of mass spectrometry to the analysis of organic molecules, electron ionization (EI) was the only ionization method available. Today, many other ionization techniques are available for the analysis of large (bio)molecules, but EI remains a popular method for the ionization of organic molecules that are not too involatile; it is, for example, the method of choice for identifying components in complex mixtures using the powerful technique of gas chromatography–mass spectrometry (GC-MS) in conjunction with fast searches in large EI spectral databases. Although EI produces both positively and negatively charged ions, usually only the singly charged positive ions, which make up conventional EI mass spectra, are used for interpretive purposes. It is not surprising, then, that most of the work published on ion structures concerns the properties of singly charged positive ions.

Almost all organic molecules contain an even number of electrons and so removal of an electron by EI leads to a radical cation, $M^{\bullet+}$, the molecular ion. The conventional mass spectrum reflects the competitive and consecutive fragmentations with rate constants $k > 10^6\ s^{-1}$ of the molecular ion induced by EI. The interpreter's task then is to deduce the structure of the original neutral molecule from the occurrence or absence of ionic fragmentations. Obviously, a good understanding of which ion structures can be formed and how they dissociate is critical to the interpreter's success.

All experimental methods for ion structure elucidation give information only about the ion's constitution or connectivity, i.e. which atoms are joined together and the probable (formal) locations of charge (and radical) sites. Details of the ion's configuration, such as bond lengths and angles, are not yet accessible to experiment for polyatomic ions. In the mass spectrometer, ions are gaseous isolated species: they are free from solvent molecules and this makes them especially amenable to molecular orbital (MO) calculations. Indeed, the growing synergism between experiment and theory in gas-phase ion chemistry has greatly facilitated our understanding in this area. State-of-the-art *ab initio* MO calculations can be as accurate as experiment and as often as not disagreements that emerge are solved by reinterpretation of the experimental data.

The most important methods for ion structure determination are:

- Ion dissociation characteristics obtained by tandem mass spectrometry: both unimolecular, as observed in the metastable ion (MI) mass spectrum and collision-induced, as observed in the collision-induced dissociation (CID) mass spectrum.
- Reactivity, as studied by ion–molecule reactions.
- Ion thermochemistry, that is heat of formation ($\Delta H_f$) data obtained by ionization energy (IE) and appearance energy (AE) measurements, often in conjunction with theoretical calculations.

These may be complemented by isotope labelling experiments, a time-honoured method of unravelling fragmentation mechanisms. With these methods,

ions of different internal energies and ion lifetimes are sampled. MI characteristics, as obtained with a sector mass spectrometer, reflect ions that decompose in a field-free region with a narrow range of internal energies just above the threshold for decomposition, ~10 μs after their formation. CID mass spectra, also measured with a sector mass spectrometer, and ion–molecule reactions as studied in an ion cyclotron resonance (ICR) cell involve nondecomposing ions with internal energies up to the lowest threshold for decomposition; these ions have lifetimes >10 μs and >1 ms, respectively. Finally, $\Delta H_f$ data and MO calculations characterize ions in their electronic and vibrational ground states having, in principle, infinite lifetimes. These different techniques sample ions of different internal energies and lifetimes and, depending on the heights of possible isomerization barriers, they may also sample ions of different structure.

## Isomerization of solitary ions

Structure elucidation of an organic ion may be rendered difficult by isomerization reactions; this is especially true for the low-energy (metastable) ions. This is illustrated in **Figure 1** for the isomeric ions $ABC^+$ and $ACB^+$ whose dissociations of lowest energy requirement from the unrearranged species lead to $AB^+$ and $AC^+$, respectively.

In **Figure 1A** the isomerization barrier is higher than either dissociation threshold and the isomeric ions will yield different ion products in their mass spectra, so that differentiation is possible on the basis of MI spectra alone. In the case of **Figure 1B** the isomerization barrier is lower than either dissociation threshold and now, prior to dissociation, ions $ACB^+$ and $ABC^+$ interconvert. As a consequence, prior to dissociation, the metastable ions will consist of an equilibrium mixture $ACB^+ \rightleftharpoons ABC^+$ that will choose the exit channel of lowest energy, namely to $AB^+$. In this case the isomers will yield identical MI spectra. Ions with energies below the transition state for isomerization will have retained their original structure, but these do not dissociate. These ions can be induced to decompose by collisional activation with an inert gas, such as helium, to dissociation products of higher enthalpy, resulting in the CID mass spectrum. These dissociations take place rapidly (<0.1 μs) and so interconversion $ACB^+ \rightleftharpoons ABC^+$ will be slow by comparison and the isomers can be differentiated on the basis of their respective CID spectra. The situation pertinent to ion $ACB^+$ in **Figure 1B** is frequently observed: because they dissociate slowly, metastable ions can undergo more or less extensive rearrangement reactions to form the most 'economical', but usually non-structure-characteristic, set of products. In CID ('sledge-hammer' approach), the energized species fragment rapidly and so more of the structure-characteristic direct bond cleavages are seen. Finally, the isomerization barrier can be intermediate between the dissociation levels of $AB^+ + C^•$ and $AC^+ + B^•$. In this case $ACB^+$ isomerizes slowly to $ABC^+$ which then fragments rapidly. Now the metastable ions will yield the same products but with different internal energies and kinetics. In particular, the kinetic energy released ($T$ value) upon dissociation will be larger for $ACB^+$, resulting in broadening of the metastable peak.

Usually CID experiments serve to increase the internal energy of the ions, which then fragment. Under certain conditions, however, such collisions may lead not only to dissociation but also to association. Thus collisions of $C_{60}^{•+}$ with He at high translational energies (~$3 \times 10^4$ m s$^{-1}$) leads to the capture complex $C_{60}He^{•+}$. This complex survives a neutralization–reionization (NR) experiment, showing that the ionic and neutral complexes have an endohedral structure He@$C_{60}$. Complexes

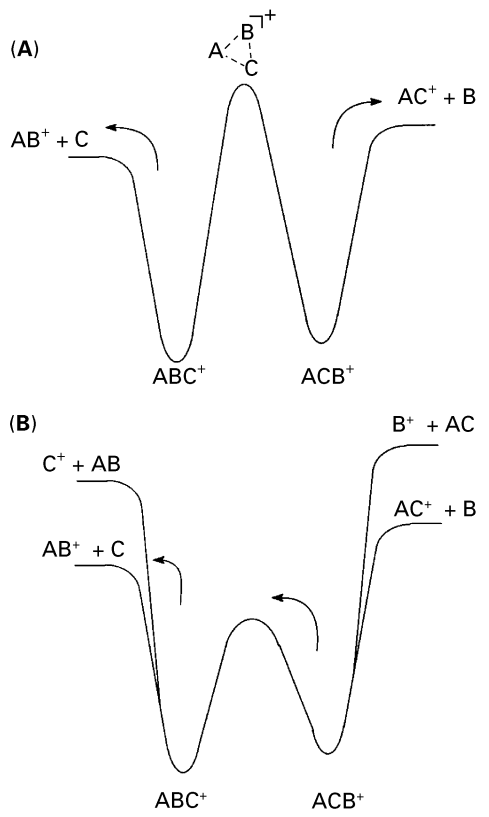



**Figure 1** Schematic potential energy diagrams showing the relationship between the heights of isomerization barriers and observed products.

containing two noble-gas atoms have also been prepared by collision experiments.

### Ion structures

The molecular radical cation, an odd-electron ion, can dissociate in two basic ways. It can undergo a direct homolytic cleavage to produce an even-electron ion and a radical, or it can produce another odd-electron ion by multiple cleavage (as in a cyclic compound) or via rearrangement (as in the McLafferty rearrangement).

## Even electron ions

### Ionized radicals

Even-electron ions are formed by loss of a radical from the molecular ions or by addition of a reagent ion to a molecule in an ion–molecule reaction. Most, but not all, even-electron ions formed by dissociative ionization can be viewed as ionized radicals. Such species may be stabilized by resonance or aromaticity. Thus $C_3H_5^+$ ions are abundant ions in many mass spectra, even in cases where the $C_3H_5$ substructure is not present in the original molecule, and this can be attributed to stabilization by resonance: $CH_2{=}CH{-}CH_2^+ \Longleftrightarrow {}^+CH_2{-}CH{=}CH_2$. Similarly, the $C_7H_7^+$ ion, $m/z$ 91, is a ubiquitous peak in the mass spectra of aromatic compounds. Elegant $^{13}C$ labelling experiments performed in the 1950s have shown that this species in the spectrum of toluene and substituted aromatics is not the expected benzyl ion, $C_6H_5CH_2^+$ (even though allylically stabilized) but rather the thermodynamically more stable (by ~40 kJ mol$^{-1}$) cyclic tropylium ion. The reverse situation holds for radicals; here the benzyl structure is the more stable one. The tropylium ion obeys Hückel's $4n$+2 rule ($n$ = 1) on aromaticity, whereas the tropyl radical does not, and indeed it was Hückel who predicted 70 years ago that the cation should enjoy considerable stability. A similar, but more dramatic, situation occurs for the $C_3H_3$ structures. Here, for the neutrals, the cyclopropenyl radical is ~96 kJ mol$^{-1}$ less stable than the propargyl radical, $^{\bullet}CH_2{-}C{\equiv}CH$ whereas, for the ions, the cyclopropenium ion ($n = 0$) is the more stable isomer (by 105 kJ mol$^{-1}$). For the cations, the large destabilization due to ring strain is more than offset by the energy gain due to aromaticity. These examples serve to illustrate an important phenomenon in mass spectrometry: upon ionization a stability reversal of isomeric structures may take effect. This becomes particularly important for odd-electron ions (see below).

The $C_2H_3O^+$ system is one of the earliest examples where experiment and theory have gone hand in hand to produce a detailed description of the potential energy surface. Four stable $C_2H_3O^+$ ions have been generated and identified by MI, CID and isotope labelling experiments and all are calculated by theory to be stable species. These are (with relative calculated $\Delta H_f$ values in kJ mol$^{-1}$ in parentheses):

(0) (181) (216) (244)

In particular the isomer $CH_3{-}O^+{=}C$ presented an experimental challenge, but after theory predicted it should be stable, it was subsequently made. Not surprisingly, the most stable ion, the acetylium ion $CH_3C{=}O^+$, is the base peak in the mass spectra of methyl ketones. As shown by CID experiments, the ion is also abundantly formed from radical cations that do not contain the acetyl substructure, for example from ionized vinyl ethers, $CH_2{=}C(H)OR^{\bullet+}$, which do not produce the intrinsically unstable $\alpha$-formylcarbenium ion $^+CH_2C(H){=}O$ via a simple bond-breaking reaction but rather $CH_3C{=}O^+$. Such 'pseudo-simple cleavage reactions' – which proceed via a 'hidden' hydrogen migration in the molecular ion – often occur, and if they are not recognized for what they are, various erroneous conclusions may be drawn from the experiments. This is exemplified by the behaviour of the dimethyl sulfoxide (DMSO) radical cation, $CH_3{-}S({=}O)CH_3^{\bullet+}$, which is a case par excellence of the situation sketched in **Figure 1B**, where $ACB^+$ represents the DMSO radical cation. Loss of $^{\bullet}CH_3$ had always, quite reasonably, been viewed as a direct bond-cleavage reaction to produce $CH_3S{=}O^+$, and indeed CID experiments on the product ions formed in the ion source (i.e. those ions formed at relatively high internal energies) showed that this is indeed the case. From AE measurements a $\Delta H_f$ = 736 kJ mol$^{-1}$ was derived for the product ions and, logically then, this value was proposed to correspond to the structure $CH_3S{=}O^+$. However, it followed from CID experiments on the metastably generated product ions (i.e. those ions formed at lower internal energies, corresponding more closely to the AE measurements) that these ions did not have the $CH_3S{=}O^+$ connectivity, but rather the carbenium structure $^+CH_2{-}S{-}OH$, formed via a low-lying ('hidden') 1,3-H shift in the molecular ions, yielding the intermediate $CH_2{=}S(OH)CH_3^{\bullet+}$ ions, $ABC^+$ in **Figure 1B,** and the measured $\Delta H_f$ was now assigned

to the structure $^{+}CH_2$–S–OH. Interestingly, these CID experiments were inspired by the *ab initio* MO calculations which correctly predicted that $CH_2{=}S^{+}OH$ should have a lower $\Delta H_f$ than $CH_3S{=}O^{+}$ (by 84 kJ $mol^{-1}$) in sharp contrast to the carbon analogues $CH_2{=}C^{+}OH$ and $CH_3$–C=$O^{+}$ (see above). In contrast to ketene, $CH_2{=}C{=}O$, where protonation takes place at the carbon to produce the stable structure $CH_3C{=}O^{+}$, the protonation of sulfine $CH_2{=}S{=}O$, takes place at the oxygen to produce $^{+}CH_2$–S–OH.

### Proton-bound molecule pairs

Even-electron proton-bound molecule pairs ($M_1 \cdots H^{+} \cdots M_2$) can conveniently be made by ion–molecule reactions and their properties (for example, bond dissociation energies) can be obtained by ICR-based experiments. Such experiments on symmetric ($M \cdots H^{+} \cdots M$) proton-bound dimers and asymmetric ($M_1 \cdots H^{+} \cdots M_2$) proton-bound molecular pairs indicate a remarkable constancy in hydrogen bond energies. It appears that symmetric proton-bound dimers have hydrogen bond energies of 130 ± 10 kJ $mol^{-1}$ and so $\Delta H_f$ of a proton-bound dimer is given by $\Delta H_f = \Delta H_f[MH^{+}] + \Delta H_f[M] - 130$; for asymmetric proton-bound molecule pairs, the hydrogen bond energy depends on the difference in proton affinities (PA) of the bases involved, and for the case $PA[M_1] > PA[M_2]$ the following empirical equation (Eqn [1]) applies for their $\Delta H_f$ :

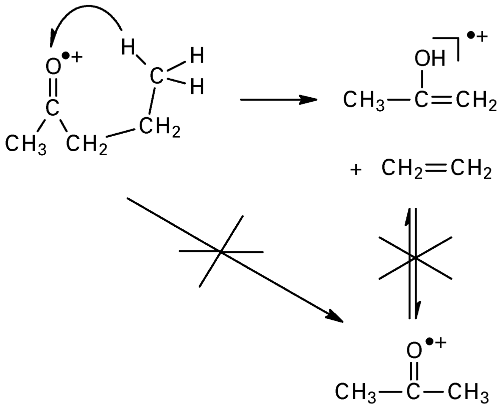
$$\Delta H_f = 0.54\{\Delta H_f[M_2H^{+}] + \Delta H_f[M_1]\} + 0.46\{\Delta H_f[M_1H^{+}] + \Delta H_f[M_2]\} - 130 \quad [1]$$

Thus, the constancy in hydrogen bond energy allows the assessment of $\Delta H_f$ of the proton-bound molecule pair, provided that the PA values of $M_1$ and $M_2$ are known. Alternatively, the relative PA values of $M_1$ and $M_2$ may be assessed from MI experiments of the proton-bound molecule pair using a sector-type instrument: the intensity ratio of the observed $M_1H^{+}$ and $M_2H^{+}$ in the MI spectrum is a measure of the relative PA values of the molecules $M_1$ and $M_2$.

## Odd-electron Ions

### Enol and related radical cations

The McLafferty rearrangement, discovered in 1952, is the only named reaction in mass spectrometry. For 2-pentanone, the reaction is represented by Equation [2].

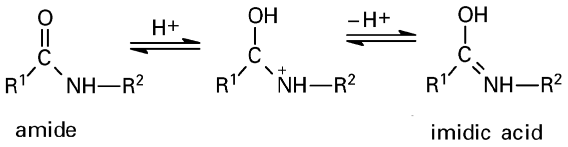



[2]

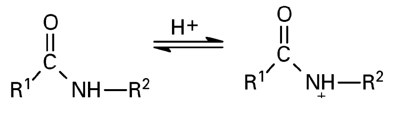
At first it was thought that the driving force of this reaction was the expulsion of a stable neutral species (such as $CH_2{=}CH_2$) rather than formation of a stable neutral (keto) ionic product, but from subsequent AE measurements it was concluded – and these findings have been amply corroborated – that the product ions had an enol and not the tautomeric keto structure and that, in contrast to the corresponding neutral molecules, enol radical cations are more stable thermodynamically than the keto tautomers. Thus, as with the $C_7H_7$ and $C_3H_3$ structures discussed above, enol structures undergo a stability reversal upon ionization; this is illustrated in **Figure 2** for the acetone radical cation and its enol structures.

Because of their enhanced stability, enol radical cations are produced from many compounds (*m*/*z* 58 from methyl ketones; *m*/*z* 59 from amides; *m*/*z* 60 from acids) and they are thus highly structure-characteristic. They provide a rich source for the generation of neutral enols via NR (neutralization–reionization) mass spectrometry. It is found from such experiments that gaseous neutral enols do not isomerize to the more stable keto counterparts, in agreement with results from pyrolysis mass spectrometry. The class of enol radical cations can be regarded as the harbinger of all ions of unexpected stability (see below). Imidic acids RC(OH)=NR (iminols) are the ‘enol’ equivalents of amides, RC(=O)NHR, and the prototype imidic acid radical cation, $HN{=}C(H)OH^{\bullet+}$, ionized formimidic acid, has been generated by an appropriate dissociative ionization and identified by CID type experiments. Neutral imidic acids have been proposed as intermediates in the acid-catalysed proton exchange of

amides, peptides and proteins, a process of fundamental importance in biochemistry (Eqn [3] as opposed to the classical mechanism, Eqn [4]). Simple imidic acids have not yet been seen in solution.




[3]




[4]

NR experiments on ionized formimidic acid showed that a significant fraction of the ions survive neutralization, but, because of interference from neutral products arising from dissociation of the neutralized species, it was not possible, from straightforward NR experiments alone, to assign a unique structure to the neutralized species. Such problems plague other systems too, but neutral identification may still be feasible by selectively analysing the ions recovered after the NR process. Such a NR-CID type of multiple collision experiment is demanding in terms of sensitivity, but nevertheless revealing. Such experiments showed that the formimidic acid structure was maintained throughout the NR experiment, and so neutral formimidic acid is a stable species on the microsecond timescale, which makes proton exchange via the imidic acid mechanism (Eqn [3]) in principle possible. NR-CID experiments are also a good check of the (isomeric) purity of the primary ion beam.

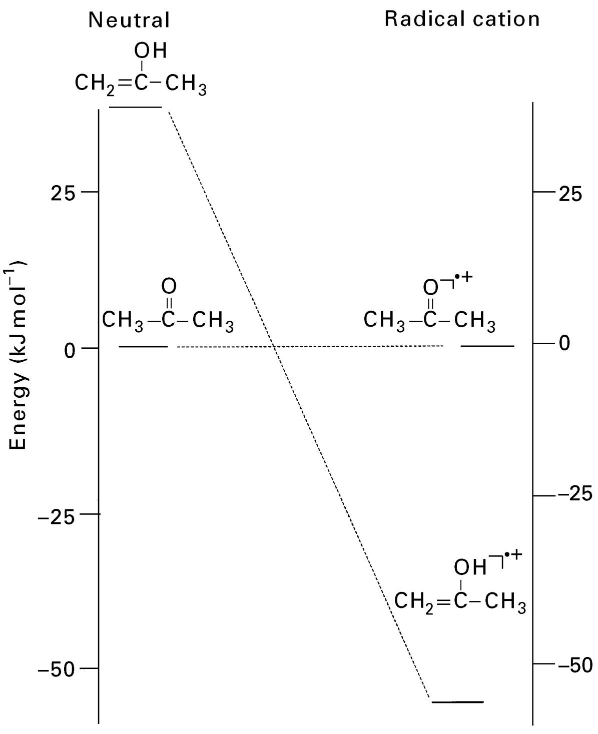



**Figure 2** Heats of formation of neutral and ionized acetone and its enol showing stability reversal upon ionization. Arbitrary energy scales. $\Delta H_f\,[CH_3C(=O)CH_3] = -218$ kJ mol$^{-1}$, $\Delta H_f\,[CH_3(=O)CH_3]^{\bullet+} = 720$ kJ mol$^{-1}$.

Another isomer of ionized formamide is ionized aminohydroxycarbene, $H_2NC^{\bullet+}OH$. This ion was also generated in the gas phase by an appropriate dissociative ionization and distinguished from its isomers by CID experiments on multiply labelled ions. This followed earlier results indicating that carbene radical cations are remarkably stable species.

## Carbene radical cations

In the early days of mass spectrometry little attention was paid to carbene-type radical cations, $R_1CR_2^{\bullet+}$, probably because it was thought that such species, like their neutral counterparts, would be thermodynamically too unstable to be formed. For example, neutral hydroxycarbene, $HC^{\bullet\bullet}OH$ lies 218 kJ mol$^{-1}$ above its isomer formaldehyde, $CH_2{=}O$. However, from experiments and theoretical calculations from the early 1980s, it appeared that the radical cations of simple carbenes could be as stable thermodynamically as their better-known isomers: for example, $HC^{\bullet+}OH$ is only 13 kJ mol$^{-1}$ less stable than $CH_2{=}O^{\bullet+}$, compared to 218 kJ mol$^{-1}$ for the neutral system. This behaviour parallels that of keto–enol structures (see above). The generation of $HC^{\bullet+}OH$ is simple: it is formed by a 1,1-$H_2$ loss from ionized methanol, but other ionized carbenes are not so easily made. Several oxacarbene radical cations, $R_1C^{\bullet+}OR_2$($R$ = H or alkyl) have been generated and they all possess a surprising property: direct cleavage of $R_2$ leads to stable acylium ions (see above), but without exception this reaction has a large barrier (~100 kJ mol$^{-1}$) for the reverse reaction. This is illustrated in **Figure 3** for the prototype oxacarbene $HC^{\bullet+}OH$.

Many ionized carbenes have been generated by appropriate dissociative ionizations and these ions provide a convenient source for the generation of neutral carbenes by NRMS, for example the prototype dioxacarbene, $HOC^{\bullet\bullet}OH$.

## Keto–enol tautomerization

In the rarefied gas phase, keto–enol tautomerization of neutral species can only occur via intramolecular

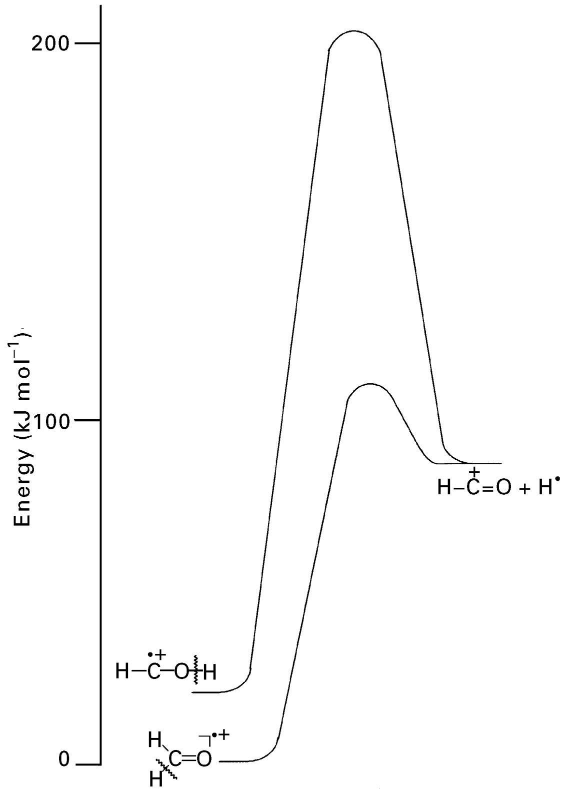
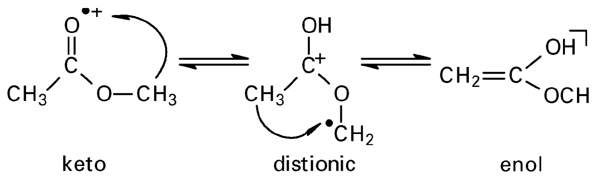



**Figure 3** Energy level diagram for H• loss from the isomers $CH_2$=O•+ and HC•+–OH showing a larage reverse energy for the direct O–H cleavage.

hydrogen transfers. For simple systems, such as methyl acetate, $CH_3C(=O)OCH_3$, tautomerization is only possible via an energy-demanding 1,3-H shift ($CH_3C(=O)OCH_3 \longrightarrow CH_2=C(OH)OCH_3$) and thus, in the gas phase, enol structures, despite their unfavourable heats of formation, remain as they are (see above). The situation is much different for the corresponding radical cations. Here keto–enol tautomerization is mediated by a surprisingly stable third structure, for which no neutral counterpart exists. Instead of a one-step energy-consuming and thus slow 1,3-H shift, the tautomerization takes place via two less energy-demanding fast 1,4-H shifts. This is why deuterium-labelled methyl acetate molecular ions show H/D randomization reactions ('scrambling'). The stable intermediate structure is called a 'distonic ion'.




[5]

## Distonic ions

In *Tetrahedron Report* number 280 (1990) the question was addressed: 'One electron more, one electron less. What does it change?' The authors critically evaluated the perception that 'comparing the reactivity of molecules whose only difference was the electron number was like comparing apples and oranges'. However, if one thing has become clear over the past decade, it is that for gas phase species 'one electron less' can change the chemistry of the species beyond recognition. One of the great discoveries in the 1980s with respect to ion structures was the finding that certain ion structures that did not correspond to those of stable, neutral molecules could nevertheless be very stable, both thermodynamically and kinetically. For example, even a simple radical cation such as $CH_3OH^{\bullet+}$ has a stable isomer, $CH_2OH_2^{\bullet+}$, the methyleneoxonium radical cation. It has been said, not without exaggeration, that $CH_2OH_2^{\bullet+}$ represents a triumph for theory. This is because, theory (*ab initio* MO calculations) predicted not only that $CH_2OH_2^{\bullet+}$ is more stable than $CH_3OH^{\bullet+}$ but also that a large barrier separates these ions; this is shown in **Figure 4A**.

Soon after this prediction was made, the ion was indeed generated and identified. These new species all have in common that, in contrast to conventional isomers, the charge and radical site are formally at separate atoms and as such they are referred to as 'distonic' ions from the Greek 'diestos' and Latin 'distans' meaning 'separate'. If the radical and charge sites are on adjacent atoms such as in $CH_2OH_2^{\bullet+}$, the distonic ion can be viewed as an ionized ylid, and as such they are also referred to as 'ylidion'. Distonic

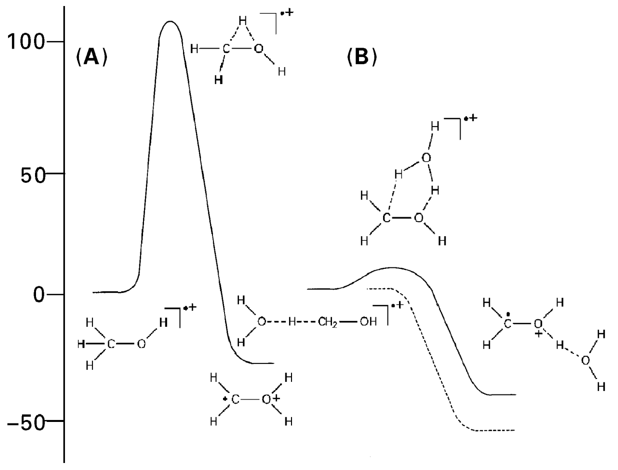
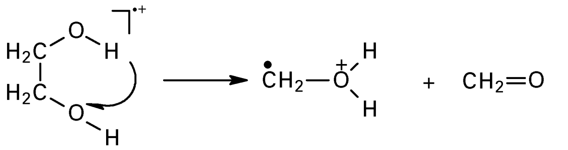



**Figure 4** (A) Energy diagram for the (unassisted 1,2-H shift) isomerization $CH_3OH^{\bullet+}$ to $CH_2OH_2^{\bullet+}$. (B) Energy diagram for the rearrangement $CH_3OH^{\bullet+}$ to $CH_2OH_2^{\bullet+}$, catalysed by water. Broken lines indicate catalysis by formaldehyde. Energies in kJ mol$^{-1}$.

ions can also be viewed as charged or protonated radicals, but, because they form a well-defined group of radical cations having characteristic properties, the term distonic ion is a useful one. Distonic ions can be made by judiciously chosen dissociative ionizations and they can be differentiated from their conventional isomers by MI and CID experiments and by ion–molecule reactions. In **Figure 5** are shown the CID spectra of ionized methanol, $CH_3OH^{\bullet+}$ and of the $CH_2OH_2^{\bullet+}$ product ion generated by the loss of $CH_2{=}O$ from ionized 1,2-ethanediol (Eqn [6]):




[6]

The major dissociation products in the spectrum of **Figure 5B** are $CH_2^{\bullet+}$ (*m/z* 14) and $H_2O^{\bullet+}$ (*m/z* 18), and this is compatible only with the $CH_2OH_2^{\bullet+}$ structure. This spectrum also reveals that the doubly charged ion $CH_2OH_2^{++}$, *m/z* 16, which can be viewed as doubly protonated formaldehyde, also is a remarkably stable species, despite the large coulombic repulsion of the charges. Such doubly charged ions have been invoked in the field of super-acid chemistry pioneered by Olah, and in solution these species behave as ‘super electrophiles’. Thus $^{+}CH_2OH_2^{+}$ is proposed to be a key intermediate in the formation of methane from formaldehyde under ‘super-acid’ conditions.

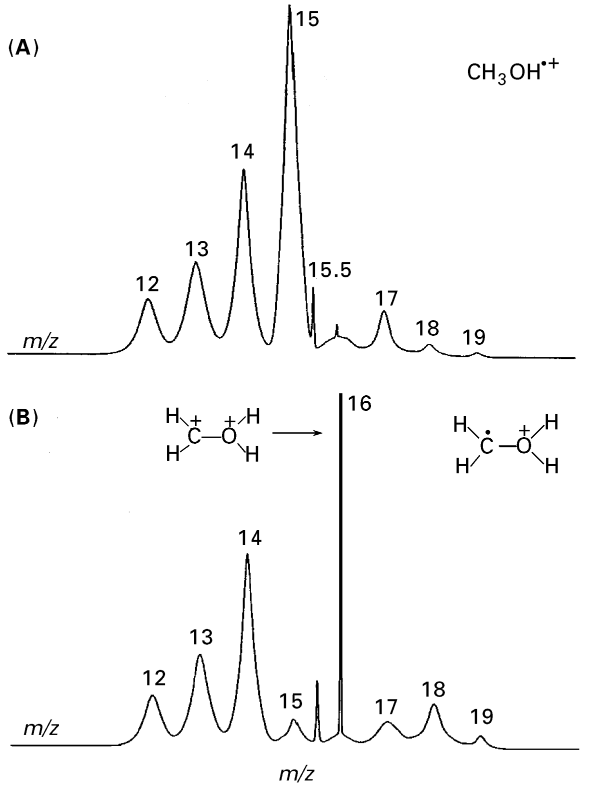
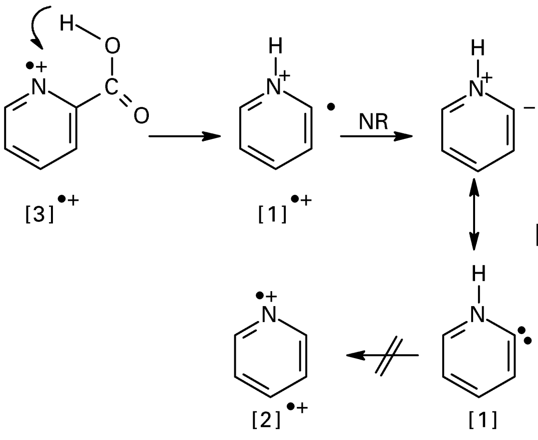



**Figure 5** Collision induced dissociation mass spectra of (A) $CH_3OH^{\bullet+}$ and (B) $CH_2OH_2^{\bullet+}$. Note the intense signal for the doubly charged ion in (B).

In the mass spectrometer, many initial molecular ions rearrange to distonic ions by way of a hydrogen shift; these intermediate distonic ions may rearrange further and their reactions involve both the radical and charge sites: often a charged moiety can easily migrate to the radical site, thereby generating an isomeric distonic ion that in turn can shift its ionized part to the new radical site. Such a sequence can rationalize many otherwise unintelligible reactions. Distonic ions can serve as precursors for the genesis of novel neutral species via the techniques of NRMS. This is illustrated by the generation of the Hammick intermediate [1], a previously unobserved species that has been held responsible for the accelerated decarboxylation of 2-picolinic acid [3] (Eqn [7]).




[7]

In the mass spectrometer, the molecular ions of 2-picolinic acid eliminate $CO_2$ and CID experiments indicate that the resulting *m/z* 79 product ions are the ylidion isomer of pyridine $[1]^{\bullet+}$ and not ionized pyridine $[2]^{\bullet+}$ itself. (The possibility of differing descriptive formalisms, i.e. ylidion or ionized carbene, exists for $[1]^{\bullet+}$). NR experiments showed that a significant fraction of the ylidions survive neutralization experiments, but because of interference from neutral products arising from dissociation of the neutralized species, it was not possible to assign a structure to the neutralized species. The same problem plagues the formidic acid experiment (see above). However, the NR-CID mass spectrum clearly showed that the ylid structure was main-

tained throughout the NR experiments and so the Hammick intermediate is a stable species on the microsecond timescale.

The first homologue of the methyleneoxonium ion is the distonic ion $^{\bullet}CH_2CH_2OH_2^+$, an isomer of ionized ethanol. *Ab initio* calculations show that the distonic ion is indeed a low-energy species lying ~40 kJ mol$^{-1}$ beneath ionized ethanol, but also that the C–O bond readily stretches to form the ion–dipole (ion–neutral) complex $CH_2{=}CH_2^{\bullet+}/H_2O$. This species can best be viewed as a positively charged ethene rod to which the water molecule is attached by its dipole, but which can be moved more or less freely along and about the positive rod.

### Ion–neutral complexes

Traditionally, unimolecular rearrangements of gas-phase (radical) cations have been rationalized in terms of transformations within the covalently bound ion. Increasingly, ion–neutral complexes with a relatively strong electrostatic bond between the incipient fragments are being invoked as intermediates to account for reactions that would otherwise be considered implausible for the intact ion. In an ion–neutral complex the charged and neutral components are sufficiently separated that they show reactivities similar to those expected for the isolated species. Two distinct but related phenomena may be interpreted by means of ion–neutral interactions: the ionic component may react individually by isomerization to a more stable structure (Eqn [8a]) or the ionic and neutral species may react with each other, for example by proton transfer (Eqn [8b]).

$$I_1N^+ \rightarrow I_1^+ \cdots N \rightarrow I_2^+ \cdots N \rightarrow I_2^+ + N \qquad [8a]$$

$$\rightarrow [I_1{-}H] \cdots HN^+ \rightarrow [I_1{-}H] + HN^+ \qquad [8b]$$

Although the earliest suggestion of a complex-mediated dissociation dates back to 1956, it was not until the early 1980s that it became evident that their involvement in the unimolecular dissociations of odd and even electron ions is a widespread phenomenon.

The formation of $CH_3CH^+CH_3$ from the onium ion $CH_2{=}O^+CH_2CH_2CH_3$ is illustrative of isomerization (via a 1,2-hydride shift) of the cation within the complex followed by dissociation, that is: $CH_2{=}O\cdots{}^+CH_2CH_2CH_3 \rightarrow CH_2{=}O\cdots{}^+CH(CH_3)_2$.

The rearrangement of carbocations, such as the archetypal cation isomerization $^+CH_2CH_2CH_3 \rightarrow {}^+CH(CH_3)_2$, is a central theme of organic chemistry; the ion in an ion–neutral complex occupies a position between that of the 'bare' ion and that corresponding to the ion in solution. Thus the influence of a single solvent molecule on the rearrangement may be studied.

A classical case of proton transfer between the formal components of a complex concerns the dissociation of steroidal diamines, for example 3,20-diaminopregnane. Here, a long-distance intramolecular and interfunctional proton transfer was proposed to occur in an ion–neutral complex, where upon separation the incipient fragments rotate more or less freely, thus allowing proton transfer between functional groups that originally were distant in the intact molecular ion (Eqn [9]).

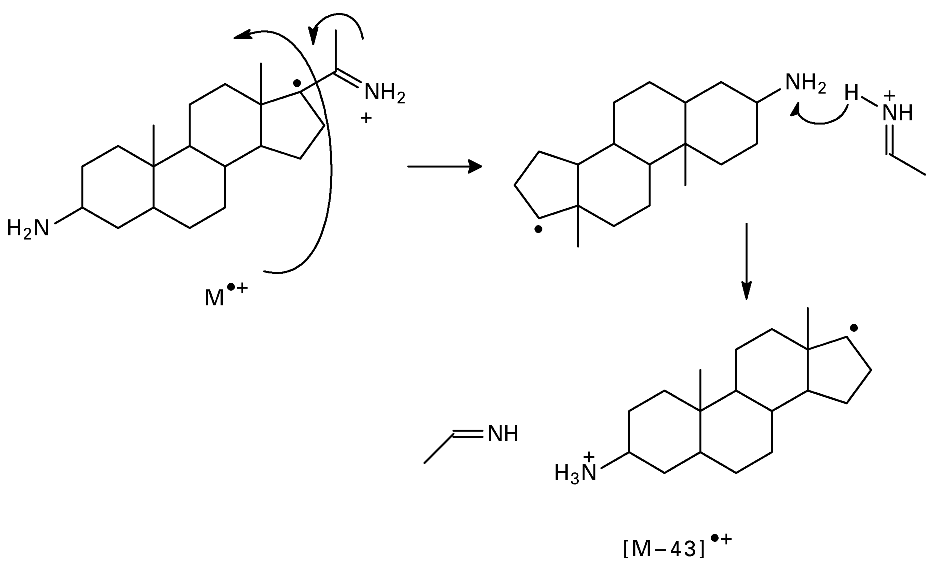
The maximum stabilization energy (SE) of an ion–dipole complex is achieved when the dipole moment vector points away from the charge; SE (kJ mol$^{-1}$) = 288 $\mu/r^2$ where $\mu$ is the dipole moment (in Debye) and $r$ is the distance (in Å) between the point charge and the dipole. For typical values of $\mu$ (1.5 D) and $r$ (2 Å) one obtains SEs of ~100 kJ mol$^{-1}$ with concomitantly smaller values for lower $\mu$. Thus, a strong ion–dipole bond is energetically on a par with a weak covalent bond.

Criteria have been proposed for the intermediacy of ion–neutral complexes, but probably the only good criteria follow from the proposal that such complexes dissociate exceedingly close to threshold. These are (1) very small kinetic energy releases and (2) unprecedented isotope effects. Thus, extreme isotopic fractionation due to differences in vibrational frequencies in the products has been identified in ion–dipole dissociations.

One of the puzzling observations seen in MI spectra is that metastable ions invariably dissociate to the products of lowest enthalpy, despite the obvious mechanistic complexities involved from a conventional point of view; the particular virtue of intermediate ion–neutral complexes is that they facilitate the formation of such low-energy products.

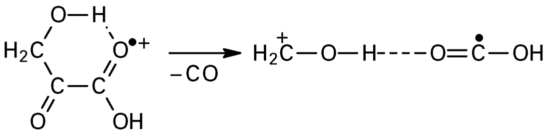
### Hydrogen-bridged radical cations

Hydrogen-bridged radical cations are the odd-electron counterparts of proton-bound molecule pairs, $M_1\cdots H^+M_2$. Formally they can be represented as proton-bound molecule–radical pairs ($M\cdots H^+\cdots R^{\bullet}$), but they are better viewed as H-bridged ion–neutral complexes of the type $MH^+\cdots R^{\bullet}$ or $M\cdots HR^{\bullet+}$, with the bridging H closer to the partner of higher PA, where most of the stabilization energy is provided by ion–dipole attractions. The H-bridge furnishes additional stabilization by about 20 kJ mol$^{-1}$ but its main function is to direct the course of isomerization by allowing a facile proton transfer. Hydrogen-bridged radical cations are harder to make than




[9]

proton-bound molecule pairs via ion-molecule reactions. They are usually generated by elimination of a stable molecule from a molecular ion, such as loss of CO from ionized $\beta$-hydroxypyruvic acid (Eqn [10]).




[10]

There remains an inherent problem with regard to the identification of a H-bridged radical cation as a stable product ion: the ion may have isomers with closely similar reactivity and dissociation characteristics. This is a well-recognized problem in general, but it becomes particularly acute for H-bridged radical cations because they often dissociate very much like distonic isomers. A case in point is the structure of the ion generated by the loss of $C_2H_4$ from ionized 1,4-butanediol, $HOCH_2CH_2CH_2CH_2OH^{\bullet+}$. Its dissociation characteristics rule out a simple extrusion reaction yielding $HOCH_2CH_2OH^{\bullet+}$ (ionized 1,2-ethanediol), but they are compatible with the formation of species comprising vinyl alcohol and water: i.e. with both the H-bridged ion $H_2O{\cdots}HOC(H){=}CH_2^{\bullet+}$ and the distonic ions $H_2O^+CH_2C^{\bullet}HOH$ and $^{\bullet}CH_2CH(OH)OH_2^+$. Analysis of *ab initio* computational results on the isomers, stabilities and conversion barriers in conjunction with energetic information from experiment (AE measurements) resolved the ambiguity and led to the identification of the H-bridged ion.

In some cases, H-bridged radical cations have been generated by ion–molecule reactions. Electron impact ionization of a mixture of $H_2O$ and $CO_2$ leads to the H-bridged species $HOH{\cdots}O{=}C{=}O^{\bullet+}$. Although the ion could be differentiated from its covalently bounded isomer ionized carbonic acid, $(HO)_2C{=}O^{\bullet+}$, through its different CID spectrum, the NR mass spectrum provided definitive evidence of its structure. As expected, and in contrast to $(HO)_2C{=}O^{\bullet+}$, the H-bridged species does not survive neutralization and dissociates completely to $CO_2$ and $H_2O$.

Hydrogen-bridged radical cations, like ion–neutral complexes, consist of two separate entities and as such they may be considered the smallest type of clusters. Dimer radical cations $M_2^{\bullet+}$ constitute an interesting class of complexes within the mainstream of cluster chemistry. They can be made by ligand exchange reactions at low pressures in an FT-ICR cell. Noble gas dimer radical cations such as $Xe_2^{\bullet+}$ can exchange, in a three-body collision, a Xe atom for a molecule M to form a $XeM^{\bullet+}$ radical cation. The latter can exchange, in a similar process, the second Xe atom for another molecule to form the dimer radical cation $M_2^{\bullet+}$. In this way, the water dimer radical cation $(H_2O)_2^{\bullet+}$ has been generated. With this technique it is possible to measure thermochemical properties of such complexes. Thus, bracketing of the electron transfer processes $(H_2O)_2^{\bullet+} + M \longrightarrow M^{\bullet+} + H_2O + H_2O$ leads to an energy difference of 1030–1038 kJ mol$^{-1}$ between $(H_2O)_2^{\bullet+}$ and $H_2O + H_2O$. *Ab initio* calculations show that the H-bridged radical cation $H_2O{-}H^+{\cdots}OH^{\bullet}$ is 38 kJ mol$^{-1}$ more stable than the 'true' ion-dipole complex $H_2O^{\bullet+}{\cdots}OH_2$ and that it lies 1021 kJ mol$^{-1}$ below $H_2O + H_2O$, indicating that the observed species is in fact H-bridged.

Hydrogen-bridged radical cations, in particular O···H···O bonded ions and their less stable C···H···O

bonded counterparts, are important intermediates in the dissociation chemistry of many oxygen-containing radical cations. They can interconvert via a so-called proton-transport catalysis. It was mentioned above that distonic radical cations are often more stable than their conventional counterparts and that they may be separated from those isomers by large barriers; see **Figure 4A** for the 1,2-H shift separating the conventional ion $CH_3OH^{\bullet+}$ and its distonic isomer $CH_2OH_2^{\bullet+}$. However, experiments and *ab initio* calculations show that interaction with an appropriately basic molecule, such as water or formaldehyde, greatly lowers the barrier. Thus the isomerization $CH_3OH^{\bullet+} \rightarrow CH_2OH_2^{\bullet+}$, which does not occur for the bare ions, is greatly accelerated by the addition of water; see **Figure 4B**. Here the water molecule attracts one of the methanol C-protons and then via a 5-membered ring donates one of its original protons to the O atom, thus producing $CH_2OH_2^{\bullet+}$. The barrier, which for the unassisted reaction is 108 kJ mol$^{-1}$ is reduced to a mere 10 kJ mol$^{-1}$ for the water-catalysed rearrangement. Proton transport catalysis has been shown to occur in many ionic rearrangement/dissociation processes. For example, the rearrangement $CH_2OH_2^{\bullet+} \rightarrow CH_3OH^{\bullet+}$, but now catalysed by formaldehyde, plays a key role in the low-energy rearrangement of ionized 1,2-ethanediol. These are cases par excellence of effects induced by a single solvent molecule on the intrinsic chemistry of the ion.

### Charge transfer complexes

Although charge (or electron) transfer complexes are not stable entities as such, *ab initio* calculations indicate that they do play a crucial role in the rearrangement of a variety of molecular ions. The situation often arises that a molecular ion $M\text{–}DH^{\bullet+}$ (D = dipole) rearranges to an ion–dipole complex of the type $M^{\bullet+} \rightarrow D\text{–}H$, where the arrow represents the dipole vector, and that a hydrogen needs to be transferred from the dipole molecule D–H to $M^{\bullet+}$ in order to obtain the observed products $MH^+ + D^{\bullet}$. This, however, is not possible because in order to transfer its hydrogen to $M^{\bullet+}$, D–H must rotate to such an extent that the resulting ion–dipole repulsion would lead to dissociation rather than to hydrogen transfer. In such cases charge (electron) transfer through orbital interaction ($M^{\bullet+} \rightarrow D\text{–}H \rightleftharpoons M \leftarrow D\text{–}H^{\bullet+}$) provides an alternative because now, and in contrast to the situation before charge transfer, the charged $D\text{–}H^{\bullet+}$ moiety is more or less free to rotate to produce $M\cdots H\text{–}D^{\bullet+}$ after which proton transfer takes place to produce $MH^+ + D^{\bullet}$. Thus, the charge transfer process kills two birds with one stone: it allows rotation of D–H and subsequent proton transfer rather than hydrogen atom transfer.

### Dications

Historically, dications were regarded as a curiosity and were observed only incidentally in mass spectrometric studies. Small polyatomic dications are remarkable species indeed; usually they are thermodynamically unstable with respect to dissociation into two monocations, but significantly kinetic stability may result if large barriers impede dissociation. This is illustrated by the methyleneoxonium dication $CH_2OH_2^{++}$ discussed above. This species is less stable than its dissociation products $CH_2OH^+ + H^+$ (by 105 kJ mol$^{-1}$), but a barrier of 250 kJ mol$^{-1}$ makes the ion experimentally accessible, see **Figure 5**. In many cases, deprotonation of a dication is best viewed as a two-stage process. Initially, the departing unit is a hydrogen atom rather than a proton, and later an electron transfer takes place to form the products: $AH^{++} \rightarrow A^{++} - H \rightarrow A^{+}\cdots H^{+} \rightarrow A^{+} + H^{+}$. The first step, a homolytic cleavage, is responsible for the large barrier. Hence, a dication is truly a tiger in a cage.

## List of symbols

$k$ = rate constants.

*See also:* **Fragmentation in Mass Spectrometry; Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy; Ion Collision Theory; Ion Energetics in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Metastable Ions; Negative Ion Mass Spectrometry, Methods; Neutralization–Reionization in Mass Spectrometry; Proton Affinities; Sector Mass Spectrometers; Stereochemistry Studied Using Mass Spectrometry.**

## Further reading


Bouchoux G (1988) Keto-enol tautomers and distonic ions: the chemistry of $[C_nH_{2n}O]$ radical cations. *Mass Spectrometry Reviews* **7**: 1–39, 203–255.

Burgers PC and Terlouw JK (1989) Structures and reactions of gas-phase organic ions. In Rose ME (ed) *Specialist Periodical Reports Mass Spectrometry* **10**, Chapter 2, pp 35–74. Cambridge, UK: The Royal Society of Chemistry.

Freiser B (1996) Gas-phase metal ion chemistry. *Journal of Mass Spectrometry* **31**: 703–715.

Hammerum S (1988) Distonic radical cations in gaseous and condensed phase. *Mass Spectrometry Reviews* **7**: 123–202.

Heinrich N and Schwarz H (1989) Ion/molecule complexes as central intermediates in unimolecular

decompositions of metastable radical cations of some keto/enol tautomers: theory and experiment in concert. In Maier JP (ed) *Ion and Cluster Ion Spectroscopy and Structure*, pp 329-370. Amsterdam: Elsevier.

Holmes JL (1985) Assigning structures to ions in the gas-phase. *Organic Mass Spectrometry* **20**: 169-183.

Longevialle P (1992) Ion-neutral complexes in the unimolecular reactivity of organic cations in the gas phase. *Mass Spectrometry Reviews* **11**: 157-192.

McLafferty FW and Turecek F (1993) *Interpretation of Mass Spectra*. Mill Valley, CA: University Science Books.

Morton TH (1992) The reorientation criterion and positive ion-neutral complexes. *Organic Mass Spectrometry* **27**: 353-368.

Radom L (1991) Chemistry by computer: a theoretical approach to gas-phase ion chemistry. *Organic Mass Spectrometry* **26**: 359-373.

Splitter JS (1994) Mass-spectral intermediate-ion structures. In Splitter JS and Turecek F (ed) *Applications of Mass Spectrometry to Organic Stereochemistry*, Chapter 3, pp 39-81. Weinheim: Verlag Chemie.

# Ion Trap Mass Spectrometers

**Raymond E March**, Trent University, Peterborough, Ontario, Canada

Copyright © 1999 Academic Press

MASS SPECTROMETRY
**Methods & Instrumentation**

## Introduction

An ion trap mass spectrometer functions both as a mass spectrometer of considerable mass range and variable mass resolution and as an ion store in which gaseous ions can be confined. Unlike other mass spectrometers that operate at pressures $< 10^{-6}$ Torr, the ion trap operates at 1 mTorr of helium. As a storage device, the ion trap acts as an 'electric-field test-tube' for the confinement of gaseous ions either positively charged or negatively charged. The confining capacity of the ion trap arises from a trapping potential well formed when appropriate potentials are applied to the ion trap electrodes. The ion trap functions as a mass spectrometer when the trapping field is changed, so that the trajectories of simultaneously trapped ions of consecutive specific mass/charge ratio become sequentially unstable, and ions leave the trapping field in order of mass/charge ratio. Upon ejection from the ion trap, ions strike a detector and provide an output signal.

With the advent of new methods by which gaseous ions can be formed from polar molecules and injected into an ion trap, a wider range of ion trap applications is possible. The coupling of liquid chromatography (LC) with electrospray (ES) ionization and with mass spectrometry (MS) in the early 1980s led to the development of new ion trap instruments for the analysis of nonvolatile, polar and thermally labile compounds. In 1995, new ion trap instruments (Finnigan's LCQ and GCQ, and Bruker-Franzen's ESQUIRE) were introduced, which employ external ion sources with injection of ions into the ion trap. The major focus for the application of these instruments, using LC-ES/MS, is the examination of high-molecular-mass biopolymers such as proteins, peptides and oligodeoxyribonucleotides.

## The quadrupole ion trap mass spectrometer

The quadrupole ion trap consists of three electrodes which are shown in open array in **Figure 1**. Two of the electrodes are virtually identical and, while having hyperboloidal geometry, resemble small inverted saucers; these saucers are end-cap electrodes and are distinguishable by the number of holes in the centre of each electrode. One end-cap electrode has a single aperture through which electrons and/or ions can be gated, while the other has several apertures arranged centrally and through which ions pass to a detector. The third electrode, also of hyperboloidal geometry but of two sheets rather than one, is the ring electrode; it resembles a napkin holder and is of similar size, since the radius, $r_0$, in the central plane is

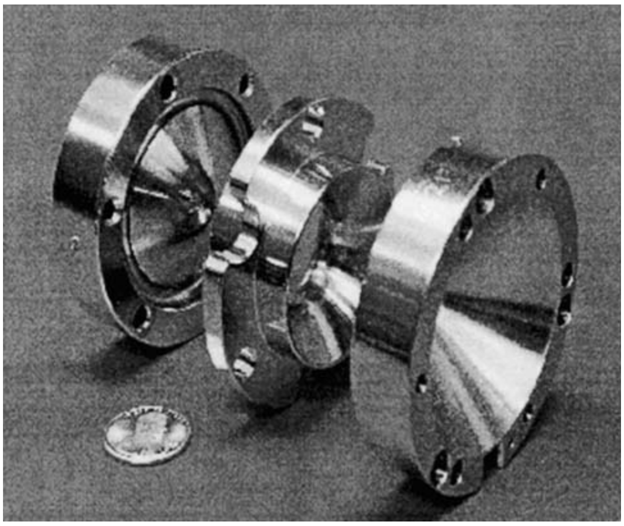


**Figure 1** The three electrodes of the quadrupole ion trap shown in open array.

~1 cm. The ring electrode is positioned symmetrically between two end-cap electrodes as shown in **Figure 2**; **Figure 2A** is a photograph of an ion trap cut in half along the axis of cylindrical symmetry, while **Figure 2B** is a cross-section of an ideal ion trap showing the asymptotes and the dimensions $r_0$ and $z_0$, where $2z_0$ is the separation of the end-cap electrodes along the axis of the ion trap.

The electrodes in **Figure 2** are truncated for practical purposes but, in theory, they extend to infinity and meet the asymptotes shown. The electrode geometries are defined so that, when an RF potential is applied to the ring electrode with the end-cap electrodes grounded, a near-ideal quadrupole field is produced, which creates a parabolic potential well for ion confinement, **Figure 3**. As shown in **Figure 3**, the potential well in the axial direction is of depth $\bar{D}_z$, while that in the radial direction $\bar{D}_r$; since $\bar{D}_z \approx 2\bar{D}_r$, the potential well resembles more a flower vase than a bowl.

Mass-selective ejection of ions from the potential well is accomplished by linearly ramping the amplitude of the RF potential; each ion species is ejected at a specific RF amplitude and, since the initial amplitude and ramping rate are known, the mass/charge ratio can be determined for each ion species. This method for measuring mass/charge ratios of confined ions was developed by Stafford and is known as the 'mass-selective axial instability mode'; this method made possible the commercialization of the ion trap in the early 1980s. A prerequisite is that ions be focused initially to the ion trap centre by momentum-dissipating collisions with helium atoms.

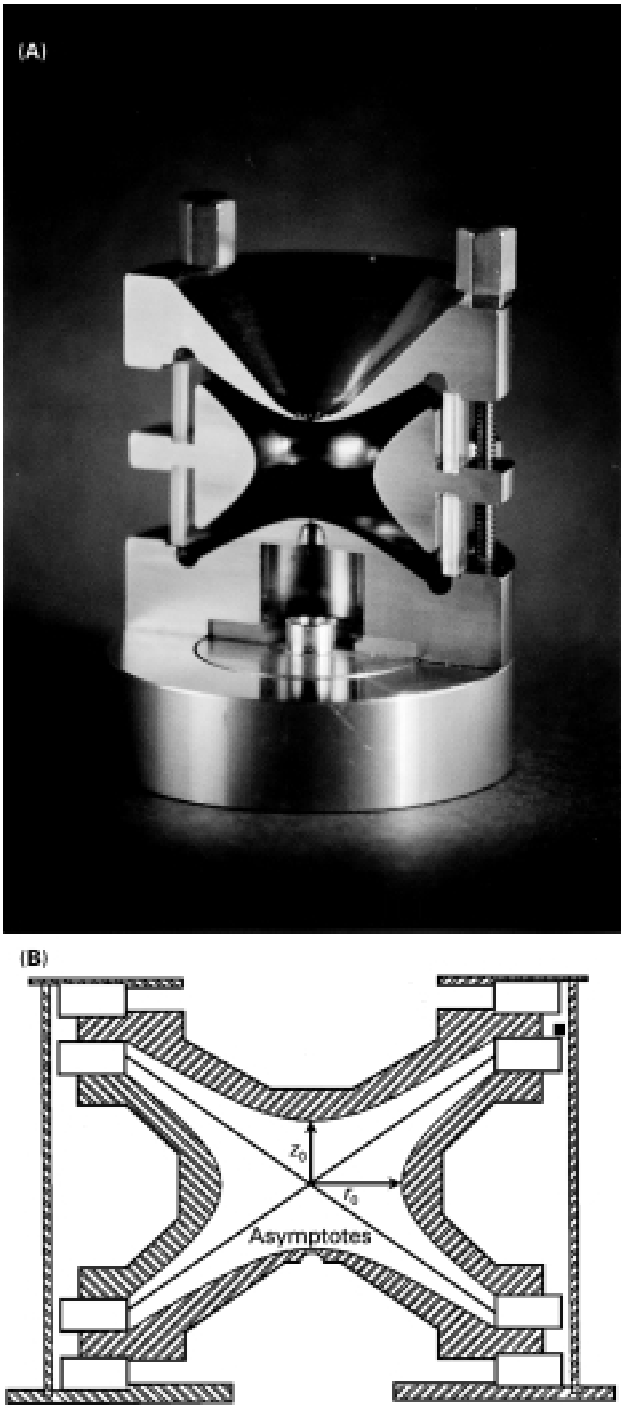



**Figure 2** Quadrupole ion trap. (A) Photograph of an ion trap cut in half along the axis of cylindrical symmetry. (B) Schematic diagram of the three-dimensional ideal ion trap showing the asymptotes and the dimensions $r_0$ and $z_0$.

## History and literature

The history of the quadrupole ion trap originates in the pioneering work of Paul and Steinwedel in the mid-1950s; their work was recognized by the award

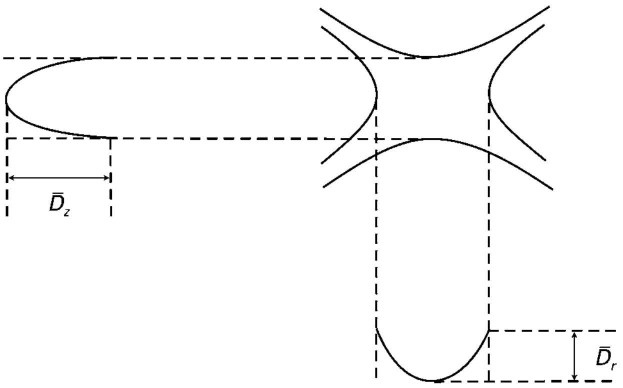



**Figure 3** Representation of the parabolic trapping potential wells of depths $\bar{D}_z$ and $\bar{D}_r$.

of the 1989 Nobel Prize in Physics to Wolfgang Paul. Yet the basis of the theory of operation of quadrupole devices was laid down over 140 years ago by Mathieu, from his investigation of vibrating stretched skins.

Those papers that are landmarks in this field are listed in the bibliography together with the five texts in which are described a wide variety of applications of the ion trap. The history of this device has been expressed as several 'ages', beginning with mass-selective detection; this age, which covered the 1950s and began with disclosure of quadrupolar devices by Paul and Steinwedel, included the storage of microparticles by Wuerker, Shelton and Langmuir and the first use of the ion trap as a mass spectrometer by Fischer. The second age is described as mass-selective storage and covers the period 1962-82. Detailed accounts of the development of the quadrupole-type devices during the early part of this period have been published by Dawson and Whetten and by Dawson; this latter publication has been reprinted in the series of 'American Vacuum Society Classics' by the American Institute of Physics under ISBN 1563964554. The third age (1983-88) is described as mass-selective ejection; this age ushered in the commercialization of the ion trap as a mass-selective detector for gas chromatograph and was a period of intense activity. In 1989, the present age of recent advances and developments commenced with the recognition of the work of Wolfgang Paul and Hans G. Dehmelt by the award, in part, of the Nobel Prize in Physics. In the same year, an account of the ion trap together with a full treatment of ion trap theory appeared in the text Quadrupole Storage Mass Spectrometry by March, Hughes and Todd; the historical account in this text by Todd was expanded into a full-scale review.

Other reviews have been contributed by Cooks and co-workers and a special collection reporting upon recent developments has also appeared. Ion trap mass spectrometry was reviewed for the 12th International Mass Spectrometry Conference in 1991. In 1995, three volumes entitled Practical Aspects of Ion Trap Mass Spectrometry were published in the CRC Series, Modern Mass Spectrometry. Volume 1 of this series, subtitled 'Fundamentals of Ion Trap Mass Spectrometry', covers the history of the ion trap, nonlinear ion traps, ion activation, ion-molecule reactions and ion trajectory simulations; the reader is referred to chapter 1 for a discussion of the Ages of the Quadrupole Ion Trap and to chapter 2 for an exposition of the mathematical basis of ion trap operation. Volume 2, subtitled 'Ion Trap Instrumentation', deals with enhancement of ion trap performance, confinement of externally generated ions, ion structure differentiation, ion photodissociation, lasers and the ion trap, and ion traps in the study of Physics. Volume 3, subtitled 'Chemical, Environmental and Biomedical Applications', includes a revisitation of fundamentals and expositions on gas chromatography-ion trap tandem mass spectrometry (GC-MS/MS) and liquid chromatography-ion trap tandem mass spectrometry (LC-MS/MS). Recently, an introduction to the quadrupole ion trap written in a tutorial form has been published. An excellent source of information is the publication entitled Proceedings of the Annual Conference of the American Society for Mass Spectrometry and Allied Topics; almost 1000 extended abstracts on ion traps can be found in the Proceedings from the past 15 years.

## The trapping potential well

The trapping potential well created within the electrode assembly of an ion trap is of parabolic cross section; ion species are confined in layers in the well rather like an exotic drink of several liqueurs arranged carefully in horizontal layers according to their density, as shown in **Figure 4A**. However, the ions of lowest mass/charge ratio reside at the ion trap centre (bottom of the well) surrounded, like the centre of an onion, by layers of ions of increasing mass/charge ratio. Upon tilting the bowl to the right, equivalent to ramping the RF trapping potential, the layer of least density, corresponding to ions of highest mass/charge ratio, will be poured from the well. To withdraw the layer of greatest density, that is, ions of lowest mass/charge ratio, a 'straw' is introduced and the bottom layer is sucked out as shown in **Figure 4B**; the straw represents axial modulation (see below) and the liqueur glass in **Figure 4B** corresponds to the detector.

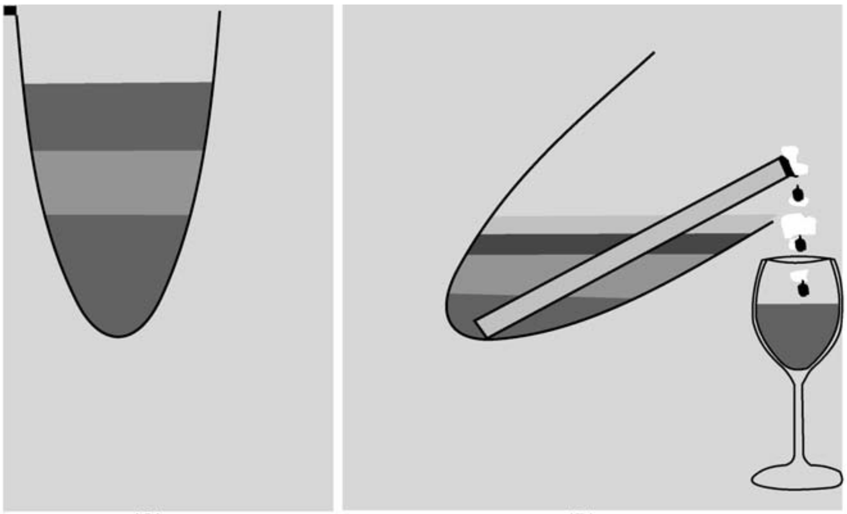



**Figure 4** (A) Schematic presentation of a trapping parabolic potential well where the three liquids differing in density represent ions differing in mass/charge ratio. (B) The tilting of the well corresponds to ramping of the RF potential while the straw, with which ions are withdrawn in order of increasing mass/charge ratio, represents axial modulation.

For an ideal quadrupole field, the following identity is given,

$$r_0^2 = 2z_0^2 \qquad [1]$$

so that once the magnitude of $r_0$ is given the sizes of all three electrodes and the electrode spacings are fixed; in the majority of commercial ion traps, $r_0$ lies in the range 0.7- 1.0 cm.

## The theory of ion trap operation

The motions of ions in quadrupole devices differ markedly from those in magnetic and electrostatic sectors. The quadrupole ion trap is described as a dynamic instrument since ion trajectories are influenced by time-dependent forces.

### An ion in a quadrupole field

An ion in a quadrupole field experiences strong focusing in that the restoring force, which drives the ion back towards the centre of the device, increases as the ion deviates from the centre. The resulting ion trajectories resemble Lissajous figures (figures-of-eight) and the motion of ions is described mathematically by the solutions to the second-order linear differential equation described originally by Mathieu from his investigation of the mathematics of vibrating stretched skins. He described solutions in terms of regions of stability and instability; these solutions and the criteria for stability and instability also describe the trajectories of ions confined in quadrupole devices and define the limits to trajectory stability.

### The Mathieu equation

The canonical form of the Mathieu equation is

$$\frac{d^2u}{d\xi^2} + (a_u - 2q_u \cos 2\xi)u = 0 \qquad [2]$$

where u represents the coordinate axes r and z, $\xi$ is a dimensionless parameter equal to $\Omega t/2$, $\Omega$ (for the ion trap) is the radial frequency of the RF potential applied to the ring and $a_u$ and $q_u$ are dimensionless trapping parameters. For the quadrupole ion trap, the trapping parameters are expressed as

$$a_r = \frac{4eU}{mr_0^2\Omega^2}; \quad q_r = \frac{-2eV}{mr_0^2\Omega^2} \qquad [3]$$

$$a_z = \frac{-8eU}{mr_0^2\Omega^2}; \quad q_z = \frac{4eV}{mr_0^2\Omega^2} \qquad [4]$$

where U is a DC potential and V is the amplitude of the RF potential of the form $V\cos \Omega t$. It is seen that $a_z = -2a_r$ and $q_z = -2q_r$.

Since $U = 0$, $a_r$ and $a_z$ are equal to zero and the common mode of ion trap operation corresponds to

operation on the $q_z$ axis of the stability diagram. The expression for $q_z$ contains the mass/charge ratio for a given ion, the size of the ion trap, $r_0$, the amplitude V of the RF potential and the radial frequency *Ω*, that is, all of the parameters that are needed to understand the operations of the ion trap.

### The 'stretched' ion trap

The ion trap electrodes are truncated in order to obtain a practical instrument, but this truncation introduces higher-order multipole components to the potential. To compensate for these multipole components, the electrodes of commercial devices prior to 1995 were assembled with a 'stretched' separation of the end-cap electrodes; the value of $z_0$ was increased by 10.6%. An account of the 'stretching' of the ion trap is given by Syka in chapter 4 of volume 1 in the CRC books, while in chapter 3 is presented an account by Franzen, Gabling, Schubert and Wang of nonlinear ion traps. The immediate consequences of stretching are that the asymptotes to the end-cap electrodes no longer coincide with those for the ring electrode, $r_0^2 \neq 2z_0^2$ and the values of the trapping parameters are changed. The trapping parameters are now expressed as

$$a_r = \frac{8eU}{m(r_0^2 + 2z_0^2)\Omega^2}; \quad q_r = \frac{-4eV}{m(r_0^2 + 2z_0^2)\Omega^2} \qquad [5]$$

and

$$a_z = \frac{-16eU}{m(r_0^2 + 2z_0^2)\Omega^2}; \quad q_z = \frac{8eV}{m(r_0^2 + 2z_0^2)\Omega^2} \qquad [6]$$

For the ion trap in the LCQ and GCQ instruments, $r_0 = 0.707$ cm and $z_0 = 0.785$ cm, such that the geometry has been stretched by ~57%.

## Regions of ion trajectory stability

Quadrupole ion trap operation is concerned with the criteria that govern the stability of an ion trajectory in both r and z directions within the field, that is, the experimental conditions that determine whether or not an ion is stored.

The solutions to Mathieu's equation are of two types, (i) periodic but unstable, and (ii) periodic and stable. Solutions of type (i) form the boundaries of unstable regions on the stability diagram and correspond to those values of a trapping parameter, $\beta_z$ that are integers, that is, 0, 1, 2, 3, ... ; $\beta_z$ is a complex function of $a_z$ and $q_z$ that is approximated as

$$\beta_z = \sqrt{a_z + \frac{q_z^2}{2}} \qquad [7]$$

for $q_z < 0.4$. The boundaries represent, in practical terms, the point at which an ion trajectory becomes unbounded. Solutions of type (ii) determine ion motion in an ion trap. The stability regions corresponding to stable solutions in the z-direction are shaded and labelled z-stable in **Figure 5A**, while those corresponding to stable solutions in the r-direction are shaded and labelled r-stable in **Figure 5B**; the latter are doubled in magnitude along the ordinate and inverted.

Ions are confined in the ion trap provided their trajectories are stable in the r and z directions simultaneously; such trajectory stability is obtained in the region closest to the origin, that is, region A in **Figure 6**, which is plotted in $a_u, q_u$ space, that is, where $a_u$ is plotted against $q_u$. Regions A and B are referred to as stability regions; region A is shown in detail in **Figure 7**. The coordinates of the stability region in **Figure 7** are the parameters $a_z$ and $q_z$. In **Figure 7**, the $\beta_z = 1$ stability boundary intersects with the $q_z$ axis at $q_z = 0.908$; this intersection is the working point of the ion of lowest mass/charge ratio that can be stored.

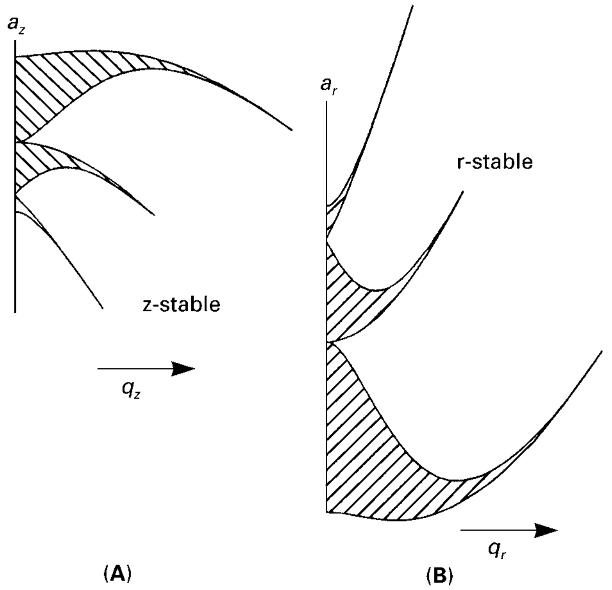



**Figure 5** Several Mathieu stability regions for the three-dimensional quadrupole field. (A) Diagrams for the *z* direction of ($a_z$,$q_z$) space. (B) Diagrams for the *r* direction of ($a_z$,$q_z$) space.

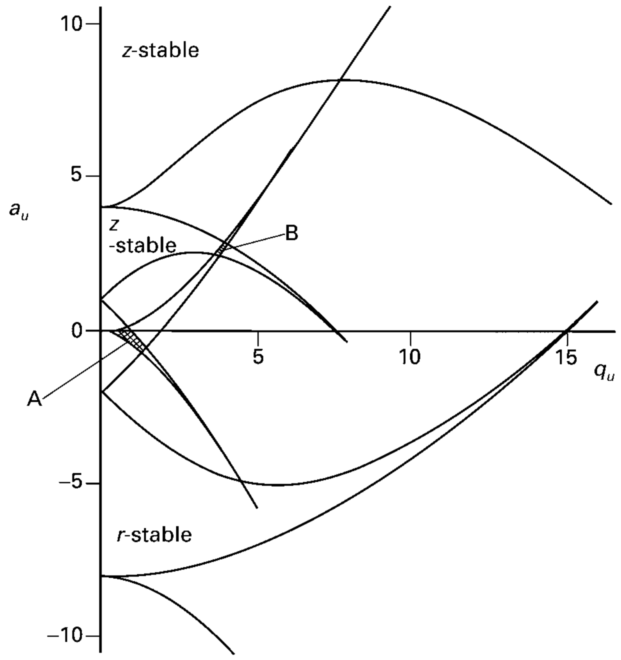



**Figure 6** The Mathieu stability diagram in $(a_z,q_z)$ space for the quadrupole ion trap in both the $r$ and $z$ directions. Regions of simultaneous overlap are labelled A and B.

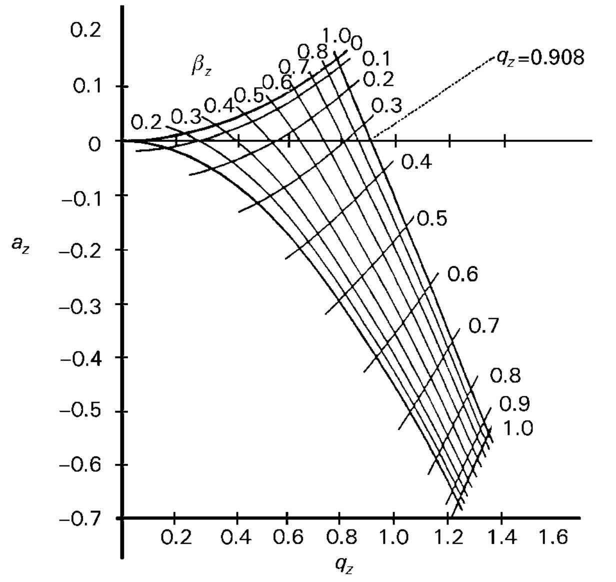



**Figure 7** Stability diagram in $(a_z,q_z)$ space for the region of simultaneous stability in both $r$ and $z$ directions near the origin for the three-dimensional quadrupole ion trap; the iso-$\beta_r$ and iso-$\beta_z$ lines are shown. The $q_z$ axis intersects the $\beta_z$ = 1 boundary at $q_z$ = 0.908, which corresponds to $q_{max}$ in the mass-selective instability mode.

## Secular frequencies

A three-dimensional representation of an ion trajectory, shown in **Figure 8**, has the general appearance of a Lissajous curve composed of two fundamental frequency components, $\omega_{r,0}$ and $\omega_{z,0}$ of the secular motion. Higher-order (n) frequencies exist and the family of frequencies is described by $\omega_{r,n}$ and $\omega_{z,n}$ as given by

$$\omega_{u,n} = (n + \beta_u/2)\Omega; \quad 0 \leq n < \infty \qquad [8]$$

and

$$\omega_{u,n} = -(n + \beta_u/2)\Omega; \quad -\infty < n < 0 \qquad [9]$$

The simulated ion trajectory shown in **Figure 8** resembles a roller-coaster ride and depicts the motion of an ion on the potential shown in **Figure 9**. The oscillatory motion of the ion results from the undulations of the potential surface, which can be envisaged as rotating. The simulation of the ion trajectory was carried out using ITSIM simulation program, while the potential surface was generated by calculating the potential for increments of 1 mm in both radial and axial directions.

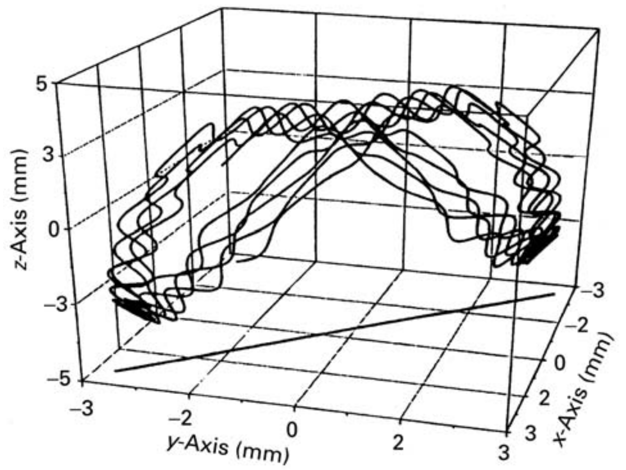



**Figure 8** Trajectory of a trapped ion of *m*/*z* 105. The initial position was selected randomly from a population with an initial gaussian distribution (FWHM of 1 mm); $q_z$ = 0.3; zero initial velocity. The projection onto the *xy* plane illustrates planar motion in three-dimensional space. The trajectory develops a shape that resembles a flattened boomerang. Reproduced from Nappi *et al* (1997) *International Journal of Mass Spectrometry and Ion Processes* **161**: 77–85.

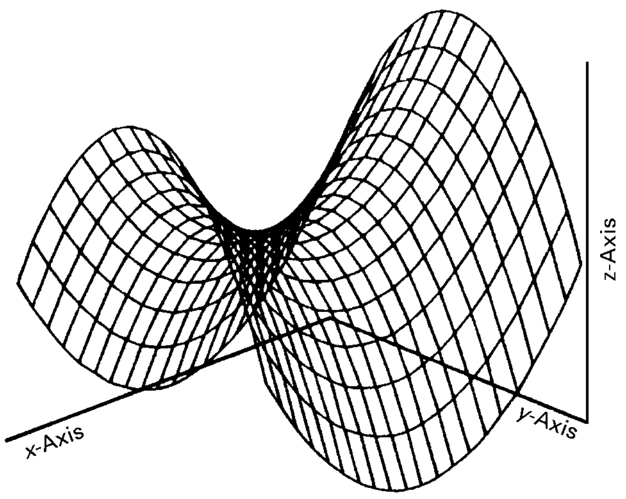



**Figure 9** Pure quadrupole field, or potential surface, for a quadrupole ion trap. Note the four poles of the surface and the similarity of the field shape to the trajectory in **Figure 8**.

## Resonant excitation

The motion of confined ions can be excited upon resonant irradiation at $\omega_z$; resonant excitation is a powerful technique in ion trap mass spectrometry since predetermined waveforms composed of specified frequencies or frequency ranges can be used. Irradiation is effected by applying a supplementary potential of some hundreds of millivolts across the end-cap electrodes. Prior to resonant excitation, ions are focused collisionally to the ion trap centre by collisions with helium atoms. This process is described as 'ion cooling' in that ion kinetic energies are reduced to ~0.1 eV, corresponding to ~800 K. Resonant excitation, or 'tickling' of cooled ions causes ions to move away from the ion trap centre so that they experience the trapping field and are accelerated to kinetic energies of tens of electronvolts.

Resonant excitation is used to increase ion kinetic energy for the following purposes. (1) To eject unwanted ions during ionization and to isolate a narrow range of mass/charge ratios. (2) To promote endothermic ion/molecule reactions. (3) To increase ion internal energy through momentum-exchange collisions with helium atoms; in the limit, ions dissociate. (4) To move ions towards an end-cap electrode where an image current can be detected for their non-destructive measurement and re-measurement. (5) To eject ions either for ion isolation or for mass-selective ejection while the applied frequency is swept. (6) To eject ions while the amplitude V of the main RF potential is ramped up; this mode, known as axial modulation, uses a fixed frequency to eject ions just before their trajectories become unstable. In axial modulation, the resonant frequency is slightly less than half the main drive frequency $\Omega$. Resonant excitation at lower frequencies has been used with great success to extend the normal mass range of the ion trap.

## Operation of the ion trap as a mass spectrometer

In the ion trap, gaseous molecules are bombarded with 50-80 eV electrons emitted from a heated filament and gated into the trap, as shown in **Figure 10A**. Under automatic gain control (AGC); the number of ions formed during 200 µs is used to scale the ionization time and produce the required number of ions. During ionization, an RF voltage $V_0$ is applied to the ring electrode so as to confine ions in a given range of mass/charge ratio. Nascent ions are subjected to about 20 collisions per millisecond with helium at a pressure of $10^{-3}$ torr; those ions that are not ejected become focused near the trap centre. In **Figure 10B**, an RF amplitude is ramped over the period 30-85 ms, during which mass-selective ion ejection and mass analysis occur.

Each ion species confined within the ion trap is associated with a $q_z$ value that lies on the $q_z$ axis on the stability diagram; ions of high mass/charge ratio have $q_z$ values near the origin, while ions of lower mass/charge ratio have $q_z$ values that extend towards the $\beta_z = 1$ stability boundary, as shown diagrammatically using stick-people of various sizes in **Figure 11A**. At the intersection of the $\beta_z = 1$ stability boundary and the $q_z$ axis, where $q_z = 0.908$ (see **Figure 7**), the trajectories of trapped ions become unstable axially and ions leave the ion trap. Once the

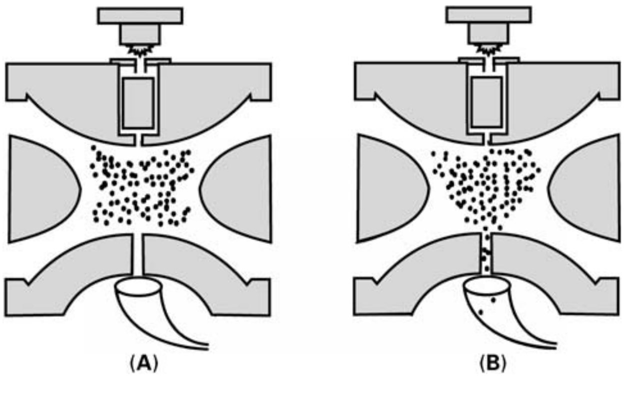



**Figure 10** An overview of MS-in-time. (A) Step 1: a trapping RF amplitude is applied for 0–30 ms during which ions are formed from sample molecules and stored. (B) Step 2: an RF amplitude is ramped over the period 30–85 ms during which mass-selective ion ejection and mass analysis occurs. Reproduced from *An Overview of MS-in-Time.*

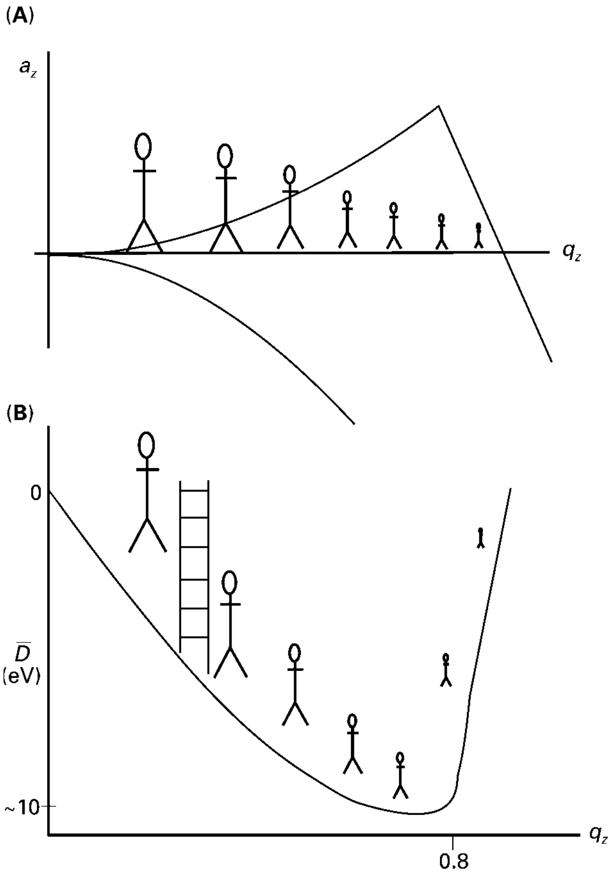



**Figure 11** (A) Schematic representation of working points (that is, coordinates in $a_z, q_z$ space) in the stability diagram for several species of ions stored concurrently. The arrangement of working points with respect to mass/charge ratio is depicted by figures that differ in size. (B) Ions are shown residing near the bottom of their respective axial potential wells of depth $\bar{D}_z$; the ladder represents the opportunity for resonant ejection of an ion species.

ion cloud has been focused collisionally to the ion trap centre, the amplitude of the RF potential is ramped; this operation, which is described as an analytical ramp or analytical scan, increases the $q_z$ values of all ion species and, when $q_z = 0.908$ for each ion species, the ions are ejected axially through the end-cap electrodes. This mass-selective axial instability method of ion ejection has been supplanted by the use of axial modulation. Originally, axial modulation described resonant ejection of ions at a frequency of 485 kHz and at a $q_z$ value slightly less than 0.908. When the RF amplitude is ramped, ions come into resonance at 485 kHz as their $q_z$ values approach 0.908 and are ejected axially in order of increasing mass/charge ratio. Since ions are focused near the ion trap centre in the fashion of an onion, resonance ejection has the effect of removing the ions of low mass/charge ratio residing in the innermost layer of the onion from the influence of space-charge perturbations induced by other ion species, so that ions are ejected free of space-charge and with enhanced mass resolution. Resonant ejection, which can be carried out over a wide frequency range, is depicted pictorially in **Figure 11**B, which shows ions residing near the bottom of their respective axial potential wells of depth $\bar{D}_z$; the ladder represents resonant ejection of an ion species. For axial modulation, the ladder is positioned at a $q_z$ value slightly less than 0.908. Resonantly ejected ions pass through holes in both end-cap electrodes so that ~50% impinge upon an electron multiplier located behind one of the end-cap electrodes; ion signals are created that produce a mass spectrum in order of increasing mass/charge ratio.

A mass spectrum is obtained by running the scan function (see below) a number of times specified by the number of microscans; the signals from each microscan are averaged and yield a mass spectrum.

## Scan function

The above operation can be expressed succinctly in a scan function, which shows the temporal variation of the potentials applied to the electrodes such that a scan function is a visual representation of the sequence of program segments in the software that controls ion trap operation. The scan function for the above mass spectrometric operation is shown in **Figure 12**.

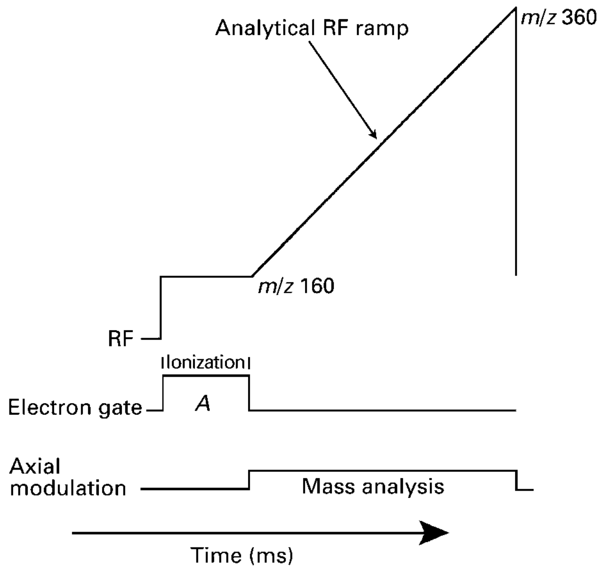



**Figure 12** Scan function for obtaining an EI mass spectrum. The scan function shows the ionization period, *A*, followed immediately by the analytical ramp with concurrent axial modulation. Note that the pre-scan for the automatic gain control algorithm is not shown.

## Collision-induced dissociation

Collision-induced dissociation (CID) of an isolated ion species in a quadrupole ion trap is a powerful technique for both the determination of ion structures and the analytical identification of compounds with high specificity. Under the influence of the tickle voltage, ions are moved from the centre to a region of higher potential, whereupon they are accelerated and their kinetic energies are increased. Subsequent collisions with helium atoms lead to enhancement of ion internal energy. In analytical applications, where the objective is to dissociate all isolated ions and to maximize the trapping of fragment ions produced, ion kinetic energy uptake must be balanced with incremental accumulation of internal energy so that ejection of isolated ions and fragment ions is avoided.

## Tandem mass spectrometry

Tandem (Latin: "at length") mass spectrometry, MS/MS, is the practice of performing one mass-selective operation after another, much as the riders are seated on a tandem bicycle. The first mass-selective operation isolates an ion species designated as the parent ion, while the second determines the mass/charge ratios of the fragment, or product, ions formed by CID of the parent ions. MS/MS with a quadrupole ion trap, where successive mass-selective operations are carried out in time, offers a number of advantages. First, since the ion trap operates in a pulsed mode, mass-selected ions can be accumulated over time. Second, since CID is wrought by many collisions of mass-selected ions with helium atoms wherein the energy transferred per collision is small, dissociation channels of lowest activation energy are accessed almost exclusively. Third, all isolated ions can be dissociated and fragment ions arising from some 90% of them can be confined. Fourth, a sequence of several mass-selective operations can be performed as in MSn.

When gas chromatography is interfaced with an ion trap tandem mass spectrometer, individual compounds can be detected at the hundreds of femtograms level. The high specificity or informing power obtainable with GC-MS/MS is achieved by observation of specific fragment ion signals from an isolated molecular ion species $M^{\cdot+}$ formed from M that elutes within a specified retention-time window.

## Chemical ionization and ion molecule reactions

In the quadrupole ion trap, several types of reactions can and do occur simultaneously and spontaneously once electron ionization of a compound has occurred. Ion-molecule reactions involving charge transfer, proton transfer and clustering occur sequentially, resulting in the formation of stable even-electron ions. Proton transfer chemical ionization (CI) involves the transfer to a neutral species of a proton (or other even-electron charged particles) from an ion that has been formed in an ion-molecule reaction. Common CI reagents are $CH_5^+$ and, $C_2H_5^+$, which are formed rapidly and can be isolated prior to reaction. CI reagent ions can be created externally and injected into the ion trap, isolated mass-selectively and allowed to react with sample molecules.

## Conclusions

Ion trap mass spectrometry is a versatile technique of high sensitivity and high specificity. The relatively low cost of commercial instrumentation has permitted a substantial growth in the practice of mass spectrometry. The theory of ion trap operation differs from those of other mass spectrometers and presents an exciting challenge to the mass spectrometry community.

## List of symbols

$a_u$ = dimensionless trapping parameter (u = r or z); $\bar{D}_r$ = depth of potential well in radial direction; $\bar{D}_z$ = depth of potential well in axial direction; m = ion mass [amu]; n = order of frequency component; $q_u$ = dimensionless trapping parameter (u = r or z); $r_0$ = radius of ring electrode; u = collective coordinate axes r and z; U = DC potential; V = 0-to-peak amplitude of the RF potential; $z_0$ = half the separation of the end-cap electrodes, along the axis of cylindrical symmetry; $\beta_z$ = trapping parameter = $(a_z + q_z^2/2)^{1/2}$; $\xi$ = a dimensionless parameter = $\Omega t/2$; $\omega$ = fundamental frequency component; $\Omega$ = radial frequency of the applied RF potential.

*See also:* **Chemical Ionization in Mass Spectrometry; Chromatography-MS, Methods; Proton Affinities.**

## Further reading

Cooks RG and Kaiser RE Jr (1990) Quadrupole ion trap mass spectrometry. Accounts of Chemical Research **23**: 213- 224.

Dawson PH (1976) Quadrupole Mass Spectrometry and Its Applications. Amsterdam: Elsevier.

Dawson PH and Whetten NR (1969) Radiofrequency quadrupole mass spectroscopy. Electronics and Electron Physics **27**: 58- 158.

Fischer E (1959) Three-dimensional stabilization of charge carriers in a quadrupole field. Zeitschrift für Physik **156**: 1- 26.

Glish GL and McLuckey SA (eds) (1991) Quadrupole Ion Traps (International Journal of Mass Spectrometry and Ion Processes **106**).

March RE (1992) Ion trap mass spectrometry. International Journal of Mass Spectrometry and Ion Processes **118/119**: 71.

March RE (1997) An introduction to quadrupole ion trap mass spectrometry. Journal of Mass Spectrometry, **32**: 351.

March RE, Hughes RJ and Todd JFJ (1989) Quadrupole Storage Mass Spectrometry. New York: Wiley Interscience.

March RE and Todd JFJ (eds) (1995) Practical Aspects of Ion Trap Mass Spectrometry, Modern Mass Spectrometry Series; Vol. 1, Fundamentals, ISBN 0-8493-4452-2; Vol. 2, Instrumentation, ISBN 0-8493-8253-X; Vol. 3, Chemical, Biomedical, and Environmental Applications, ISBN 0-8493-8251-3. Boca Raton, FL: CRC Press.

Mathieu E (1868) Journal de Mattematiques Pures et Applies (J. Liouville) **13**: 137. (See also McLachan NW (1947) Theory and Applications of Mathieu Functions. Oxford: Clarendon Press and Campbell R (1955) Théorie Générale de l'Equation de Mathieu. Paris: Masson.

Nappi M, Weil C, Cleven CD, Horn LA, Wollnik H and Cooks RG (1997) Visual representations of simulated three-dimensional ion trajectories in an ion trap mass spectrometer. International Journal of Mass Spectrometry and Ion Processes **161**: 77- 85.

Nourse BD and Cooks RG (1990) Aspects of recent developments in ion trap mass spectrometry. Analytica Chimica Acta **228**: 1- 11.

Paul W (1990) Electromagnetic traps for charged and neutral particles (Nobel Lecture). Angewandte Chemie **29**: 739- 748.

Paul W and Steinwedel H (1960) Apparatus for separating charged particles of different specific charges. German Patent 944 900, 1056; US Patent 2 939 952.

Reiser HP, Kaiser RE Jr and Cooks RG (1992) A versatile method of simulation of the operating of ion trap mass spectrometers. International Journal of Mass Spectrometry and Ion Processes **121**: 49- 63.

Stafford GC Jr, Kelley PE, Syka JEP, Reynolds WE and Todd JFJ (1984) Recent improvements in and analytical applications of advanced ion trap technology. International Journal of Mass Spectrometry and Ion Processes **60**: 85- 98.

Todd JFC (1991) Mass Spectrometry Reviews **10**: 3.

Wuerker RF, Shelton H and Langmuir RV (1959) Journal of Applied Physics **30**: 324.

# Ionization Methods in MS

*See* **Atmospheric Pressure Ionization in Mass Spectrometry**; **Chemical Ionization in Mass Spectrometry**; **Fast Atom Bombardment Ionization in Mass Spectrometry**; **Ionization Theory**; **Plasma Desorption Ionization in Mass Spectrometry**; **Thermospray Ionization in Mass Spectrometry.**

# Ionization Theory

**C Lifshitz**, The Hebrew University of Jerusalem, Israel
**TD Märk**, Leopold Franzens Universität, Innsbruck, Austria

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
| --- |
| Theory |

## Introduction: Electron ionization, photoionization

At the heart of mass spectrometry lies the formation of the ions to be analysed. A thorough understanding of the production of the ions is extremely important, since the method used to ionize a sample markedly affects its mass spectrum (fragmentation pattern). Besides electron impact ionization (EI) and photoionization (PI), a number of different methods are used in mass spectrometry including chemical ionization (CI), field ionization, fast atom bombardment, surface ionization, electrospray ionization, laser ablation and other plasma methods. EI is by far the most common method as it is simple to use, easy to set up and very effective. For instance the maximum ionization cross section (see definition below) for the hydrogen atom is about $10^{-20}$ m$^2$ (at 100 eV) for electrons as compared to about $5 \times 10^{-22}$ m$^2$ (at 14 eV) for photons. On the other hand, PI has yielded basic and accurate data required for the understanding of the energetics and dynamics of ionization and fragmentation.

Because of its overwhelming importance in mass spectrometry, only EI (and, to a lesser extent, PI) will be treated here. We will consider the ionization mechanism, the various types of ions produced and the kinetics and dynamics of the ionization process and subsequent fragmentation reactions.

Electron ionization (EI) is the process by which an atom or a molecule M is ionized by electron impact to form a positive ion (Eqn [1]).

$$\mathrm{M} + \mathrm{e}^- \rightarrow \mathrm{M}^{\bullet +} + 2\mathrm{e}^- \qquad [1]$$

Photoionization (PI) is the process by which M is ionized through photon impact.

$$\mathrm{M} + h\nu \rightarrow \mathrm{M}^{\bullet +} + \mathrm{e}^- \qquad [2]$$

The *ionization energy* (IE) is the minimum electron energy or photon energy required to produce ionization from M and is related to the binding energy of the most loosely bound valence electron in the molecule. When the electron energy or photon energy is varied continuously there is a threshold energy below which the ion current is zero and above which it rises with increasing energy, according to a particular threshold law (see below). This threshold is an experimental measure of the IE. The IE often (but not always) can be determined spectroscopically as the limit of a series of Rydberg states in which the principal quantum number $n$ of the electron reaches increasingly higher values. The IE can be determined mass spectrometrically even in those cases for which a Rydberg series has not been observed spectroscopically.

The majority of neutral, chemically stable, molecules possess a closed shell. As a result, ionization leads to a radical cation, i.e. to an odd-electron ion ($\mathrm{OE}^{\bullet +}$). In spectroscopic terms, a neutral singlet state leads to an ionic doublet state.

## Ionization mechanism

EI or PI involves the collision of an electron or photon of sufficient energy with a neutral (or ionized) target particle and the subsequent production of an ion and the respective ejected electron(s). The term ‘electron impact ionization’ is somewhat misleading, because an electron is small compared to the size of a molecule and thus would have difficulty ‘hitting’ any part of an atomic target. It is better to think of the electron as passing close to or even through the atomic target, while in quantum-mechanical terms the wave of the electron interacts with and distorts the electric field of the atomic system. Electrons accelerated through a potential of several tens of volts have a de Broglie wavelength of ~0.1 nm. In this case the wavelength and the molecular dimensions are similar and mutual quantum effects (distortions) occur. The wavelength of a photon required for PI is about 100 nm which is much larger than the usual molecular dimension and almost no molecular distortion occurs during ionization. Because of this difference in interaction in the case of EI, the ionizing transition is not strictly vertical (see below) and more states (also taking into account spin conservation from the target to the product ion) can be reached relative to PI.

In addition, electron impact can also give rise to the birth of negative ions (anions). In contrast to the formation of positive ions (cations), direct attachment of the incident electron to a polyatomic target to give a stable anion is rather improbable. The reason is that the energy of the attaching electron and the binding energy (electron affinity) must be taken up (accommodated) in the emerging product (anion). Usually, the excess energy leads either to fragmentation of the anion or to shake-off of the electron (autodetachment); radiative stabilization of this excited anion takes place only with very low probability. One way to produce stable anions by direct electron attachment is to have sufficient gas pressure in the ion source for three-body reactions to occur that remove the excess energy (three-body stabilization).

### Ionization channels

Ionization of atoms results only in the production of singly and multiply charged atomic ions. As the energy deposited by the ionizing agent is increased, the variety and abundance of the ions produced from a specific molecular target will increase, because the ionization process may proceed via different reaction channels, each of which gives rise to characteristic ionized and neutral products. For the simple case of a diatomic molecule AB the corresponding reaction channels are as shown in Equations [3]–[13] where $e_s$ is a 'scattered' electron and $e_e$ is an 'ejected' electron.

$$\mathrm{AB} + \mathrm{e} \rightarrow$$

$$\mathrm{AB}^{\bullet+} + \mathrm{e_s} + \mathrm{e_e} \quad \text{(single ionization)} \quad [3]$$

$$\mathrm{AB}^{2+} + \mathrm{e_s} + 2\mathrm{e_e} \quad \text{(double ionization)} \quad [4]$$

$$\mathrm{AB}^{z+} + \mathrm{e_s} + z\,\mathrm{e_e} \quad \text{(multiple ionization)} \quad [5]$$

$$\mathrm{AB}^{\mathrm{K}+} + \mathrm{e_s} + \mathrm{e_e} \quad \text{(K-shell (inner) ionization)} \quad [6]$$

$$\mathrm{AB}^{**} + \mathrm{e_s} \rightarrow \mathrm{AB}^{\bullet+} + \mathrm{e_s} + \mathrm{e_e} \quad \text{(autoionization)} \quad [7]$$

$$\mathrm{AB}^{+*} + \mathrm{e_s} + \mathrm{e_e} \rightarrow \begin{cases} \mathrm{A}^{\bullet+} + \mathrm{B}^{\bullet} + \mathrm{e_s} + \mathrm{e_e} & \\ \text{(metastable fragmentation)} & [8] \\ \mathrm{AB}^{2+} + \mathrm{e_s} + 2\mathrm{e_e} & \\ \text{(autoionization)} & [9] \\ \mathrm{AB}^{\bullet+} + \mathrm{e_s} + \mathrm{e_e} + h\nu & \\ \text{(radiative ionization)} & [10] \end{cases}$$

$$\mathrm{A}^{+} + \mathrm{B} + \mathrm{e_s} + \mathrm{e_e} \quad \text{(prompt dissociative ionization)} \quad [11]$$

$$\mathrm{AB}^{-*} \rightarrow \mathrm{A}^{\bullet-} + \mathrm{B}^{\bullet} \quad \text{(dissociative attachment)} \quad [12]$$

$$\mathrm{AB}^{**} + \mathrm{e_s} \rightarrow \mathrm{A}^{+} + \mathrm{B}^{-} + \mathrm{e_s} \quad \text{(ion pair formation)} \quad [13]$$

Similar reaction channels are possible for PI but there are no scattered electrons and dissociative attachment is only possible for EI. Different and more complex reactions may occur when more complex targets are involved, i.e. polyatomic molecules or clusters (even including multiple electron collisions and subsequent reactions within the molecular target). Clusters and fullerenes have also been demonstrated to undergo a process of delayed ionization which is akin to thermionic emission in bulk material.

Most of the ionization reactions shown above (e.g. Eqns [3]–[6], [11]) can be classified as being due to a direct ionization mechanism in which the ejected and the scattered electrons leave the ion within $10^{-16}$ s of each other. Conversely, there exist alternative ionization channels (competing with direct ionization) in which either the electrons are ejected one after the other (autoionization) or the molecular ion decays at a later stage (see **Figure 1** and below). The autoionization event (e.g. Eqns [7] and [9]) can be described as a two-step reaction: first, a neutral molecule (or atom) is raised to a superexcited state, which can exist for some finite time. Then, radiationless transition into the underlying ionization continuum occurs. Superexcited states exist in principle for every target system except for the hydrogen atom. For molecules, the upper autoionization rate (and hence the ionization cross section) is limited by the characteristic energy-storage mode frequency. In addition, if predissociation (into two neutrals) is faster than autoionization, the latter will not occur at an appreciable rate.

Autoionization by photons is a resonance process. Autoionization complicates the ionization cross section function (see definition below) at low as well as at high electron energies. PI has been extremely valuable in the quantitative study of autoionization processes both in atomic and in molecular systems. Classical examples involve the photoionization mass spectra of noble gas atoms such as xenon (**Figure 2**) and molecular hydrogen, $H_2$ (**Figure 3**). Xenon has filled 5s and 5p shells. Ionization is initially observed at $h\nu = 12.130$ eV with the ejection of a 5p electron and the formation of the ground $^2P_{3/1}$ state of $Xe^+$. In the interval between $^2P_{3/1}$ and its spin–orbit partner $^2P_{1/2}$ (12.12–13.436 eV), autoionization structure is prominent (see **Figure 2**) owing to Rydberg states converging to $^2P_{1/2}$. One notes sharp $p \rightarrow ns$

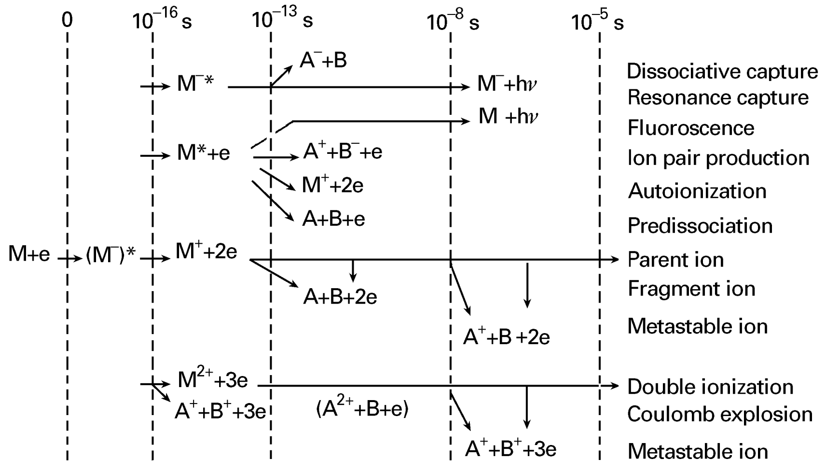



**Figure 1** Schematic time evolution of EI of a molecule.

transitions and broad ones corresponding to p → *n*d transitions. Autoionization takes place by virtue of a 'spin–flip' mechanism in which *J* changes from $\frac{1}{2}$ to $\frac{3}{2}$ and Rydberg electron is ejected. Autoionization in the case of molecular hydrogen (**Figure 3**) involves conversion of vibrational energy into additional electronic excitation of the Rydberg electron. Predissociation competes with autoionization of $H_2$.

## Franck–Condon principle

Inelastic collisions between electrons/photons and molecules involve transitions between defined (electronic, vibrational and rotational) molecular states. The energy deposited into the excitation of molecular vibration and rotation is usually rather small compared with the energy of the electronic transition, at least if one assumes that only one vibrational quantum is excited in a single collision. For example, the greatest separation between a ground and a first excited vibrational state of any molecular ion is that of $H_2^+$ (0.27 eV), which is small compared with the ionization energy of 15.426 eV. The changes in the vibrational levels from $v''$ to $v'$ that result in ionization can be described in terms of the Franck–Condon principle.

Qualitatively, the Franck–Condon principle may be summarized as follows. During an electronic transition no (or only negligible) changes occur in the nuclear separation and the velocity of relative nuclear motion. Owing to the large ratio of nuclear to electronic mass and the short interaction time (~$10^{-17}$ s, as compared with ~$10^{-13}$ s for a vibration), the point on the upper potential-energy curve (corresponding to the configuration after the transition) lies directly above the starting point on the initial potential-energy curve (vertical transition). This leads to a number of possible electronic transitions, which depend on the relative shapes of the potential-energy curves available in a specific system. Several

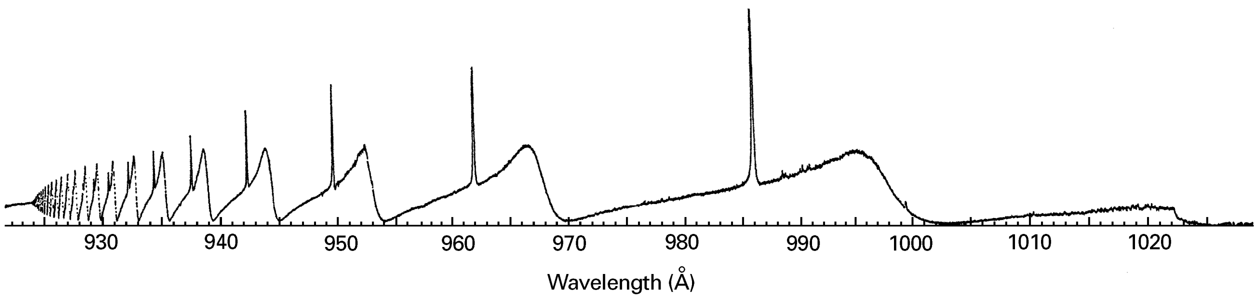



**Figure 2** Photoionization mass spectrum of xenon between the $^2P_{3/2}$ (ground) and $^2P_{1/2}$ (excited) ionic states, displaying sharp 's' like and broad 'd'-like resonances. Reproduced with permission of Academic Press from Berkowitz J (1979) *Photoabsorption, Photoionization, and Photoelectron Spectroscopy*, Chapter VI.

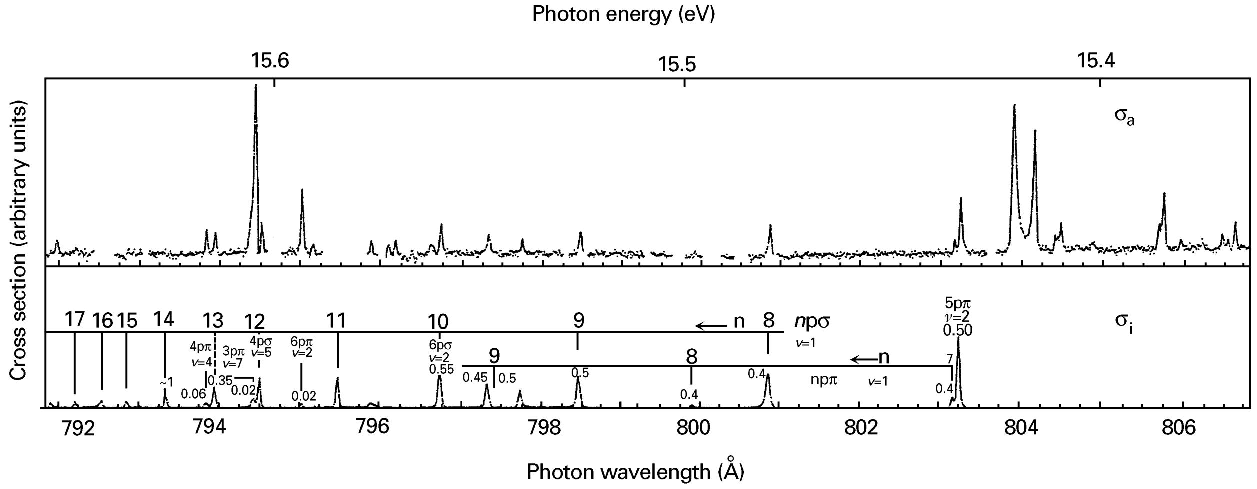



**Figure 3** Relative photoionization and photoabsorption cross sections $\sigma_i$ and $\sigma_a$ respectively, for parahydrogen at 78 K. The numbers above peaks in the photoionization data give the value of $\eta_i$ the quantum yield of ionization. Reproduced with permission of Academic Press from Berkowitz J (1979) *Photoabsorption, Photoionization and Photoelectron Spectroscopy*, Chapter VI.

cases can be distinguished. (1) The final accessible level lies within the region of discrete vibrational states of the upper potential-energy curve. The probability that the vibrational quantum number will change depends on the relative positions of the potential-energy curves. (2) The final accessible level not only lies within the region of discrete vibrational states, but includes some part of the continuum. Hence, some of the transitions will lead to dissociation. (3) The final accessible level lies within the continuum of a repulsive state and all transitions lead to dissociation.

The Franck–Condon principle can be used to treat quantitatively the fragmentation of diatomic or pseudodiatomic molecules at low energies. The cross section for ionization from the vibrational level $v''$ of the neutral molecule to a vibrational level $v'$ in the ionized molecule is given by the Franck–Condon factor (FCF) which is equal to the square of the overlap integral between the respective vibrational wavefunctions, i.e. FCF = $|\langle\psi_{v''}|\psi_{v'}\rangle|^2$. In polyatomic molecules, the two-dimensional potential-energy curves have to be replaced by multidimensional potential-energy hypersurfaces. Although EI and PI proceed without nuclear displacement (Franck–Condon principle), the resulting (excited) polyatomic ion will usually subsequently undergo further internal transitions (e.g. radiationless transfer of energy, that is transitions from one surface to another surface), possibly leading to subsequent unimolecular decomposition (see below). Even with an accurate knowledge of the potential energy surfaces, a detailed description of the evolution of the ionized system is virtually impossible; it is necessary to use statistical theory methods.

## Koopman's theorem; photoelectron spectroscopy

The photoionization process (Eqn [2]) may be studied by using a fixed photon wavelength $h\nu$ and recording the kinetic energy distribution of the emitted photoelectrons. This is called photoelectron spectroscopy (PES). The photoelectron carries away the difference between the photon energy and the binding energy of the ejected electron (Eqn [14]).

$$\varepsilon = h\nu - \mathrm{IE(M)} \qquad [14]$$

where $\varepsilon$ is the kinetic energy of the electron and IE(M) is the ionization energy of the molecule, equal to the binding energy of the photoionized electron. Each ionization energy is approximately equal to the negative eigenvalue of the molecular orbital (MO) from which the electron is ejected. This approximation is known as Koopman's theorem, which is based on the self-consistent field (SCF) model.

The photoelectron spectrum of $N_2$ (**Figure 4**), obtained with $h\nu = 21.21$ eV, demonstrates what happens upon ionization from the three outermost MOs of the molecule, whose electronic structure is $(1\sigma_g)^2(1\sigma_u)^2(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2$. Ionization of an electron from a nonbonding orbital leads to an ionic state whose equilibrium internuclear distance is not shifted appreciably with respect to the neutral. The vertical transition, according to the Franck–Condon principle, leads to very little, if any vibrational excitation. The outermost $\sigma_g$ orbital is very weakly bonding. As a result, the strongest transition observed is from $v'' = 0$ of the neutral to $v' = 0$ of the ion.

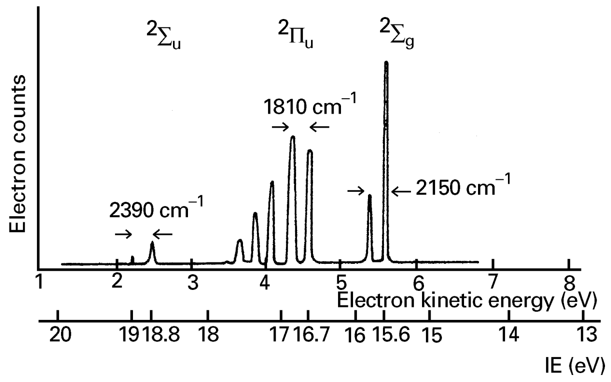



**Figure 4** Photoelectron spectrum of $N_2$ obtained with 21.21 eV incident radiation; electron counts versus electron kinetic energy (upper) and versus IE of $N_2$ (lower). Reproduced with permission of Oxford University Press from Atkins PW (1986) *Physical Chemistry*, 3rd edn, p 483.

Furthermore, the vibrational spacing in the $^2\Sigma_g$ ionic state that is produced is 2150 cm$^{-1}$, only slightly less than in the neutral (2345 cm$^{-1}$). The second band envelope reached, which is due to the $^2\Pi_u$ state, has a much richer vibrational structure, because the $\pi_u$ electron is ionized from a strongly bonding orbital. This causes the position of the minimum in the potential energy curve to be shifted to a larger internuclear distance in the ion with respect to the neutral, to excitation of a number of vibrational levels in the ion within the vertical Franck–Condon region, and to a rather small vibrational spacing of 1810 cm$^{-1}$ in the ion. Finally, the third band, which is due to formation of the $^2\Sigma_u$ state of the ion, demonstrates the ionization of a weakly antibonding electron, with a strong propensity for excitation of the 0–0 transition and with a slightly larger vibrational spacing (2390 cm$^{-1}$) in the ion as compared to the neutral.

## Types of ions produced

### Parent ions

Parent ions are positively charged ions produced by reaction [3] through removal of one electron from the neutral precursor. The production of these parent ions relative to that of other ions originating from the same neutral precursor (molecule) depends on the electron or photon energy and on the properties of the neutral molecule. At and just above the IE, only singly charged (parent) ions are produced, but at higher electron energies, other ions (see below) are also observed.

In general, for small molecules the parent ion is the dominant ion at all impact energies, although there are exceptions, such as $CCl_4$ and $CF_4$ which have no stable parent ions. Conversely, for large molecules the relative parent-ion abundance usually decreases with increasing molecular mass and increasing incident energy. Again there are exceptions to this rule, the most notable one concerning the fullerenes, where the singly charged $C_{60}^{\bullet+}$ parent ion is the most abundant ion in the mass spectrum (see **Figure 5**).

The parent-ion abundance depends also on the temperature of the molecular gas target. At constant ion-source gas pressure a decrease in all ion intensities is noted with increasing temperature, while an increase in vibrational energies of the molecular ion (due to the increase in the vibrational energies of the

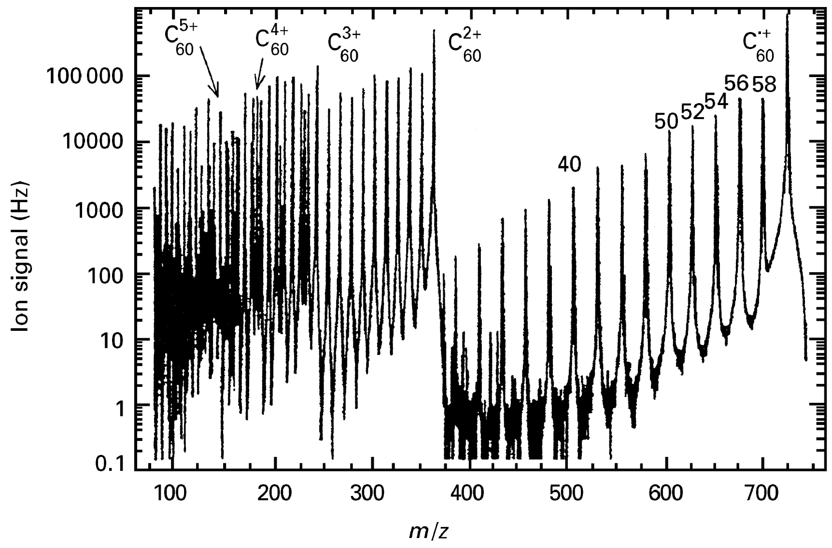



**Figure 5** Mass spectrum of $C_{60}$ ionized with 200 eV electrons.

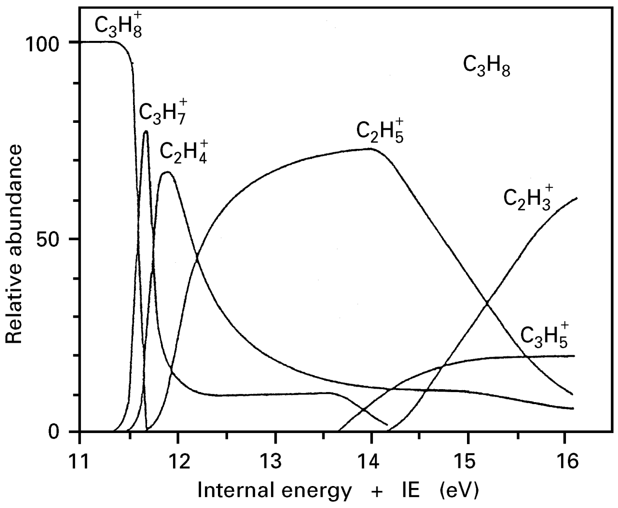



**Figure 6** Breakdown graph of propane. The relative abundances of the ions indicated are plotted as functions of the sum of the IE of propane and the internal energy of the parent propane ion.

neutral precursor) may lead to an appreciable decrease in the relative parent-ion intensity. For polyatomic molecules, this latter effect can be explained in terms of statistical theories.

## Fragment ions

If the incident energy is increased above the IE of the molecule, fragment ions appear owing to reaction paths discussed above in relation to the Franck–Condon principle. In the case of polyatomic molecules, a wide variety of fragmentation channels is available; that is, the molecular ion can immediately decay into a fragment ion with an even or odd number of electrons (primary fragment ion). These primary fragment ions may also be produced in (unstable) excited states and immediately decay into further fragments. The number of fragment ions and their relative cross-section functions are characteristic of the corresponding parent molecule. The relative importance of the various ions as a function of internal energy can be demonstrated with the help of breakdown curves (see **Figure 6** showing the breakdown graph of propane).

In the case of diatomic parent molecules, the fragmentation can be treated quantitatively in terms of the Franck–Condon principle. For the dissociation of small polyatomic molecules, spectroscopic and quantum-theoretical ideas (correlation rules) can be used to determine the dissociation path and predict the resulting electronic states. With the advent of *ab initio* CI (configuration interaction) calculations, detailed information on the potential energy surfaces of the different states and on their interactions is becoming available. However, to deal with large polyatomic molecules, statistical theories have to be used, i.e. the so-called quasiequilibrium (QET) theory or an almost identical approach, the so-called RRKM theory (named after Rice, Ramsperger, Kassel and Marcus).

In general, the relative abundance of any fragment ion is related to its rate of formation and its rate of dissociation by unimolecular decomposition. Hence, the measured fragmentation pattern of a molecule is a record in time of the 'quasiequilibrium' balance of these rates. In other words, because the fragmentation pattern is a slice of the three-dimensional plot of ion current as a function of incident energy and mass-to-charge ratio, the respective partial ionization cross-section functions will depend on the time after formation of the primary ion. Trapped-ion mass spectrometry has been used to investigate this phenomenon for EI and PI. Typical results for short and long delay times are shown for $C_6H_6^{\bullet+}$, $C_6H_5^+$ and $C_6H_4^{\bullet+}$ in **Figure 7**. The increased fragmentation of $C_6H_6^{\bullet+}$ at long delay times (due to the presence of metastable ions, see below) is obvious.

In this context it is interesting to point out that the reason for running mass spectra at relatively high electron energies (50–100 eV) is that fragmentation patterns do not vary very much with electron energy in this energy range and that the partial ionization cross-section functions (and thus the detection efficiency) have their respective maxima in this energy range (see below). On the other hand, much energy can be transferred to the molecular ion in this energy range, resulting in extensive fragmentation (see below), sometimes making it difficult to identify the parent ion. Because of this, EI is considered a 'non-soft' ionization technique compared with PI or CI. However, if similar energies are used for PI and EI, very similar mass spectra are often observed.

Finally, it should be pointed out that dissociative ionization yields fragment ions with small to moderate kinetic energies. To describe the dissociative ionization process completely, these kinematic properties must be known.

## Metastable ions

Ions produced with lifetimes longer than those of excited ions that decompose in the ion source (prior to about $10^{-6}$ s) are called metastable ions. The existence of metastable ions can be explained by different mechanisms depending on the size and property of the precursor ion. Normally, dissociation of an excited ion occurs during the first vibration (prompt dissociation), or predissociation that involves a transition from one potential energy hypersurface to another, which also occurs rapidly;

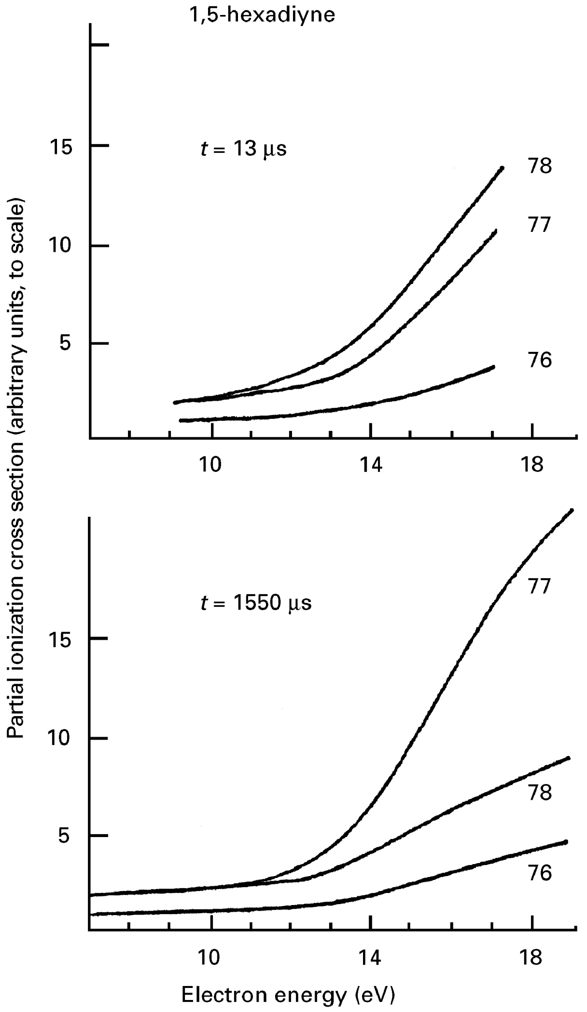



**Figure 7** Partial ionization cross-section functions for ions of 1,5-hexadiyne (a linear isomer of benzene) at two different detection times, *t*.

therefore, the product ions are produced in the ion source (see fragment ions). However, if crossings of potential hypersurfaces are spin-forbidden, some ions will nevertheless undergo electronic predissociation but will do so at a reduced rate. Thus, electronic (forbidden) predissociation via repulsive states is one origin of metastable ions. A second possible mechanism is dissociation by tunnelling. In this case the initial excitation is to bound levels of an excited state of the molecular ion that are above the dissociation limit of this state, but below the top of either an intrinsic or centrifugal barrier. The existence of large metastable ions can be rationalized in the framework of QET and is termed vibrational predissociation. After ionization, radiationless transitions yield highly vibrationally excited ground-state ions with a distribution of internal energies and hence a distribution of decay rate constants

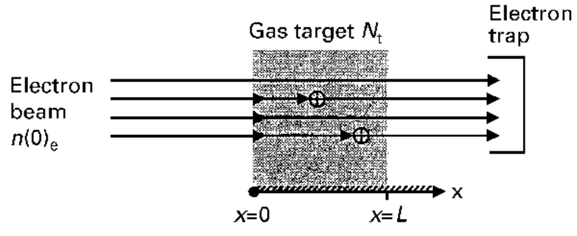



**Figure 8** Schematic view of an electron impact ionization experiment.

including those that lead to unimolecular decomposition in the metastable time range.

If the metastable decomposition occurs in flight before the ion reaches the analyser of a mass spectrometer, a typical metastable peak may be observed in the mass spectrum depending on the type of mass analyser used. Moreover, using appropriate methods, it is also possible to deduce the kinetic energy distribution of the nascent fragment ions at these late times after the initial excitation.

### Multiply charged ions

Multiply charged atomic and molecular ions were observed and identified as early as 1912. Subsequent observation of numerous doubly charged ions followed. Triply charged molecular ions have been detected in low abundances in the mass spectra of some species (e.g. $CO_2$, $CS_2$, $C_2N_2$ and aromatic compounds). Moreover, EI of free van der Waals clusters can lead to multiply charged ions with up to five and more elementary charges. Conversely, with certain molecules (e.g. $H_2O$ and $CH_4$), it is not possible to detect any stable doubly charged molecular parent ions by EI.

The partial ionization cross sections for the production of multiply charged atoms and molecular ions rarely exceed 1–5% of that of the dominant singly charged ion. Certain atoms and compounds, however, have been found to possess an increased ability to sustain two or even more positive charges, an especially intriguing example being $C_{60}$ (see **Figure 5**).

## Ionization cross sections

### Definitions

Consider, as shown in **Figure 8**, a parallel, homogeneous and monoenergetic beam of electrons crossing a semi-infinite medium containing $N_t$ target particles per cubic centimetre at rest. If $n(0)_e$ represents the initial electron intensity (number of electrons per area

and per second), the intensity of the electron beam at depth $x$ is given by the exponential absorption law (Eqn [15]):

$$n(x)_{\mathrm{e}} = n(0)_{\mathrm{e}} \exp(-N_{\mathrm{t}}\sigma x) \qquad [15]$$

If $N_t\sigma x << 1$ (single-collision condition), the number of ions generated per second along the collision interaction path $x = L$ (over which the ions are collected and analysed) is

$$n(L)_{\mathrm{i}} = n(0)_{\mathrm{e}} N_{\mathrm{t}}\sigma_{\mathrm{c}} L \qquad [16]$$

where $\sigma_c$ is the counting ionization cross section. The total positive-ion current $i_t$ produced in this interaction volume is given by

$$i_{\mathrm{t}} = n(0)_{\mathrm{e}} e N_{\mathrm{t}}\sigma_{\mathrm{t}} L \qquad [17]$$

where $\sigma_t$ is the total ionization cross section. If the ions produced are analysed with respect to their mass $m$ and charge $ze$, the respective individual ion currents measured with a Faraday cup are given by

$$i_{\mathrm{ms}} = n(0)_{\mathrm{e}} ze N_{\mathrm{t}}\sigma_{zi} L \qquad [18]$$

where $i_{ms}$ is the mass selected ion current and $\sigma_{zi}$ is the partial ionization cross section for the production of a specific ion $i$ with charge $ze$.

Total and counting ionization cross sections of a specific target system are the weighted and the simple sum of the various single and multiple partial cross sections, respectively:

$$\sigma_{\mathrm{t}} = \Sigma\sigma_{zi} z \qquad \text{and} \qquad \sigma_c = \Sigma\sigma_{zi} \qquad [19]$$

Sometimes the macroscopic cross section $s = \sigma N_t$, which represents the total effective cross-sectional area for ionization of all target molecules in 1 $cm^3$ of the target medium, is used.

Partial and total ionization cross sections are of great importance for practical applications, but they give only a limited insight into the ionization process itself. Conversely, a great deal of information about the kinematic aspects can be obtained from differential electron ionization cross sections, including the single, double and triple differential cross sections. The single differential cross section $d\sigma(E,W)/dW$, where $W$ is the energy of the electrons after the ionizing collision, measured at the collision energy $E$ represents the electron energy distribution (integrated over all angles) of the two outgoing electrons. The energy distribution is in principle symmetrical with respect to $\frac{1}{2}(W_1 + W_2) = \frac{1}{2}(E - \mathrm{IE})$, since for each fast electron a corresponding slow electron of complementary energy is ejected. If the incident energy $E$ is large enough, the low-energy part of the distribution is mainly due to 'ejected' electrons and the high-energy part is due to 'scattered' electrons. Although in principle the two outgoing electrons are indistinguishable, this interpretation is supported by the different angular behaviour of these two groups of electrons. Usually, the single differential cross section has a minimum at $W = \frac{1}{2}(W_1 + W_2)$ and maxima close to $W = 0$ and $W = (E - \mathrm{IE})$. At very small incident energies, the cross section becomes independent of $W$ (at least for atoms).

The double differential cross section $d^2\sigma(E,W,\delta)/dW\,d\Omega$ is higher by one order with respect to the number of kinematic parameters, i.e. the energy $W$ and the polar angle $\delta$ of the outgoing electrons have to be measured. In practice, one of the three experimental parameters $E$, $W$ or $\delta$ is varied while the others are kept constant. This allows the determination of either angular distributions ($E$ and $W$ constant) of ejected electrons or energy loss spectra. In general, the higher the energy of the scattered electron, the more their intensity is peaked forward, whereas the angular distribution of very low-energy electrons is nearly isotropic. The triple differential cross section $d^3\sigma(E,W_1,\delta_1,\delta_2,\varphi_2)\,dW\,d\Omega_1\,d\Omega_2$ contains all available kinematic information. It has to be measured in coincidence ('e,2e experiments') and full graphical representation is impossible.

## Ionization threshold law

Following the postulates of Wigner that the ionization probability for EI to a first approximation depends critically on the likelihood of the separation step in the exit channel (that is, separation of the collision complex $AB^{-*}$ into a stable ion and two free electrons depends on the probability of disposing of the excess energy), Wannier, using a very simple phase space argument, showed that this probability depends on the number of degrees of freedom $n$ for sharing the excess energy between the electrons, thus giving a threshold law

$$\sigma = k(E - \mathrm{IE})^n \qquad [20]$$

For single ionization and thus emission of two electrons, $n = 1$ and the coefficients $k$ is determined by the electronic transition probability, which is

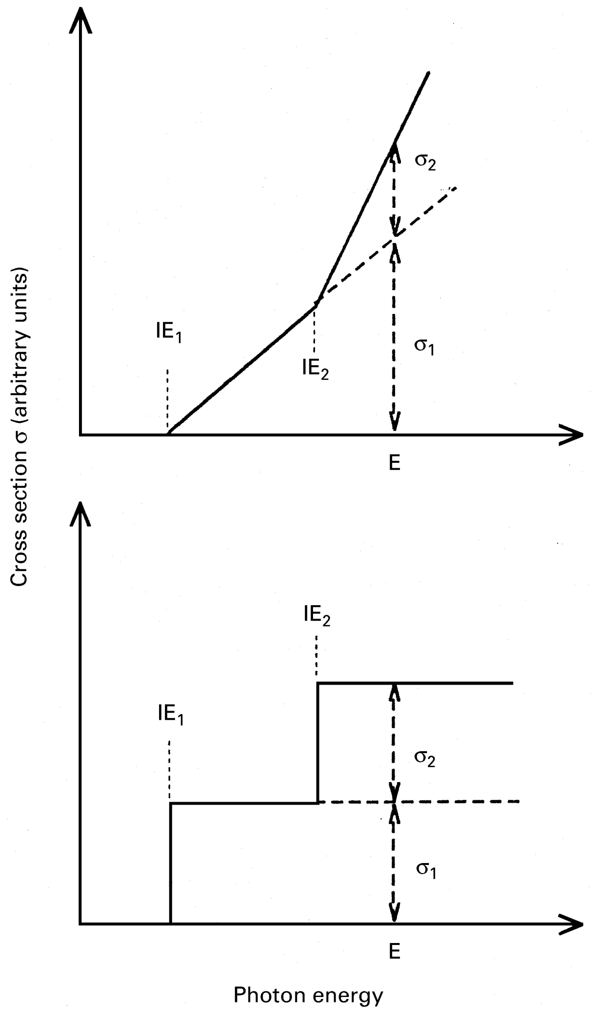



**Figure 9** Idealized threshold laws for electron impact ionization and photoionization.

proportional to the overlap of the vibronic wavefunctions for the neutral ground state and the state of the ion produced. This argument can be extended to $z$-fold ionization, giving $n = z$. More detailed calculations predict that these power laws should all be increased by 0.127. Using the same arguments for PI, it is then possible to conclude that

$$\frac{d\sigma_{\text{electron}}}{dE} = \sigma_{\text{photon}} \qquad [21]$$

Therefore, for single ionization a linear increase of the cross section with electron energy is expected, whereas for direct PI a step-function cross-section law is predicted. This simple theory gives no prediction of the energy range of validity and in practice the threshold laws are slightly different for several reasons, including influence of competitive channels, tailing due to rotational levels, and autoionization

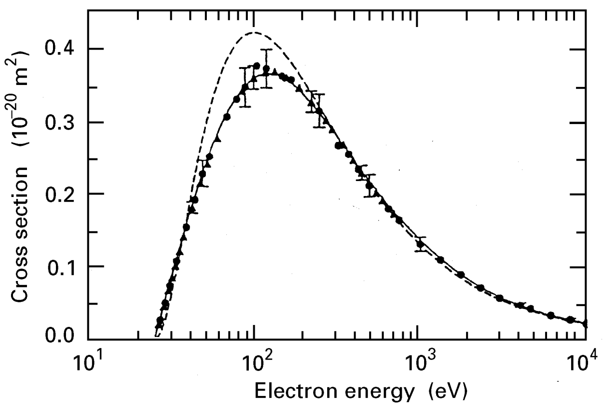



**Figure 10** Experimental (solid symbols) and theoretical (dashed line, distorted-wave Born approximation; full line, Deutsch–Maerk formula and BEB (Binary–Encounter–Bethe) approximation) electron ionization cross section for helium.

which will obscure the true threshold behaviour. Nevertheless, it is customary to assume in accordance with many experimental examples, that for each state of the ion there exists a separate ionization probability described by Equation [20] and that to a first approximation these probabilities are additive (see **Figure 9**). Finally, it is interesting to note that the shape of the threshold law is especially important for the determination of the ionization and appearance energies, because owing to the low ion signal close to the onset it is necessary to invoke an extrapolation method and knowledge of the threshold behaviour facilitates this procedure.

## Partial and total ionization cross sections

As discussed above, there exist different threshold laws for EI and PI. This difference pertains also for higher energies. Whereas in the case of EI the

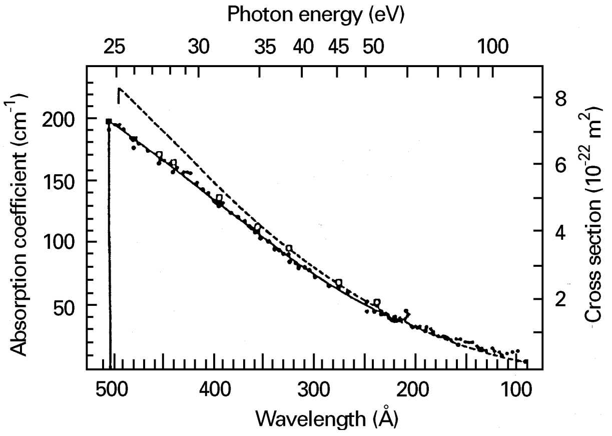



**Figure 11** Experimental and theoretical (dashed line) photoionization cross sections of helium.

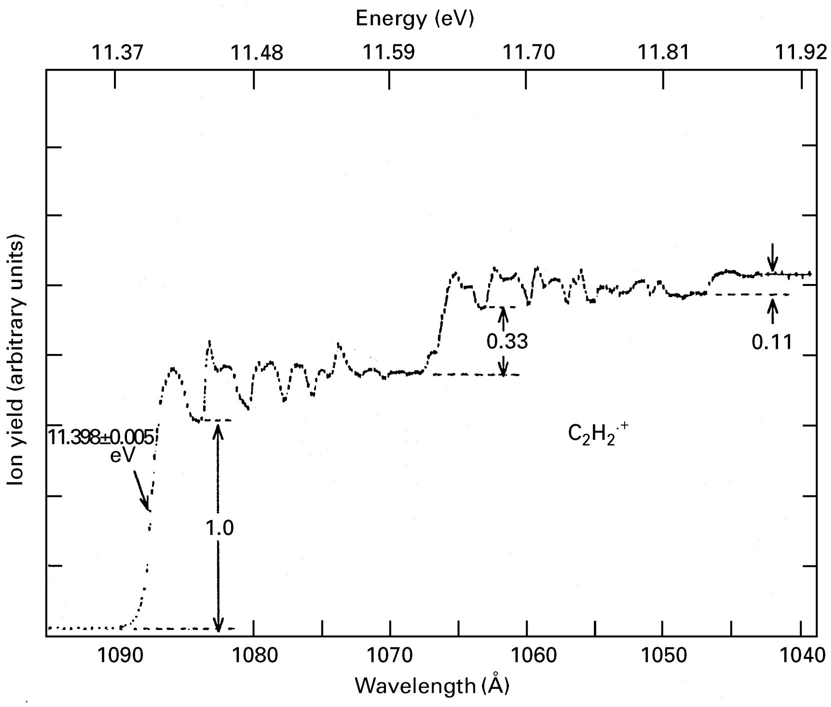



**Figure 12** Mass-selected photoion yield curve of $C_2H_2^{\cdot+}$ in the vicinity of the ionization threshold. Three steps are observed owing to direct ionization from the vibrational ground state of the neutral to $v' = 0$, 1 and 2 of the C–C stretch of the ion. Autoionization structure is superimposed upon the steplike vibrational structure. Reproduced with permission of Academic Press from Berkowitz J (1979) *Photoabsorption, Photoionization, and Photoelectron Spectroscopy*, Chapter VI.

ionization cross-section function (**Figure 10**) rises more or less linearly over some tens of eV, reaches a maximum at an energy of about 10 eV and then decreases monotonically over some thousands of eV (a logarithmic decrease is predicted in the limit of the Born–Bethe approximation), in the case of PI the abrupt rise at threshold is followed by a monotonic, more or less exponential decline to higher energies (**Figure 11**).

In the limit of zero excess photon energy above threshold, the details of the absorption probability characteristic of the individual targets do not enter and the cross-section function is only dependent on the long-range interaction of the departing particles (electron and positive ion). The Coulomb force then leads to the abrupt onset to a finite crosssection value with a post-threshold behaviour independent of excess photon energy. But even for atomic targets this step function description is inappropriate (see as an example helium in **Figure 11**, which has been studied both experimentally and theoretically). For the heavier noble gases, autoionization fine structures due to spin–orbit interaction immediately set in beyond the threshold and thus prevent examination of the isolated continuum (see above and **Figure 2**). For more complicated targets, such as molecular species, additional effects thwart the observation and analysis of the unperturbed continuum. Even the abrupt rise at threshold is not always apparent and therefore step-function behaviour, while a convenient idealization, has little correspondence with reality. A nice example where the step-function behaviour does occur is in the case of the vibrational staircase behaviour of acetylene PI (**Figure 12**).

Inelastic collisions of the electrons with atomic targets are usually divided into two regimes, fast collisions and slow collisions, depending on the

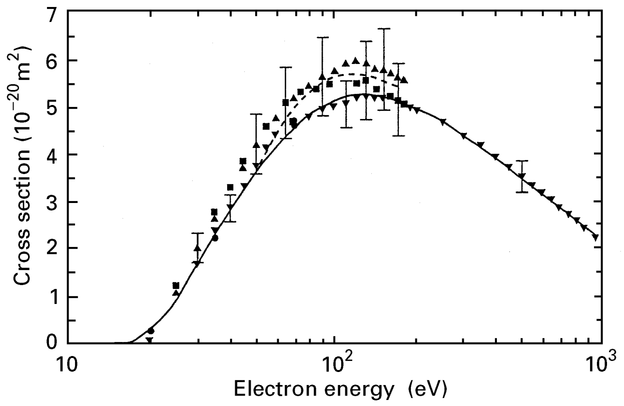



**Figure 13** Experimental (solid symbols) and theoretical (dashed line, Deutsch-Maerk formula; full line, Binary–encounter–Bethe approximation) total electron ionization cross section for $CF_4$.

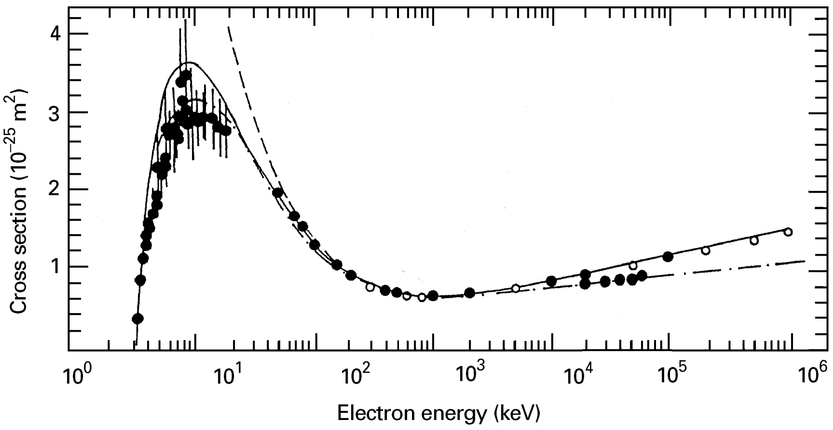



**Figure 14** Experimental (solid circles and full line) and theoretical K-shell electron ionization cross section for argon: dashed line, Born–Bethe approximation; dash-dotted line, Deutsch–Maerk formula; open circles, Born–Bethe including relativistic corrections.

relation between the impacting electron velocity to the mean orbital velocity of the electron(s) in the (sub)shell responsible for the inelastic process under consideration. For fast collisions, the influence of the incoming electron upon the target can be viewed as a sudden and small external perturbation. This concept leads to a cross-section function derived in the Born–Bethe approximation, where the ionization cross section is given in good accordance with experimental determinations asymptotically by the simple expression

$$\sigma = (4\pi a_0{}^2/u)\ M_\mathrm{i}{}^2 \ln(4cu) \qquad [22]$$

with $a_0$ the Bohr radius, $u = E/R$ the reduced electron energy, $R$ the ionization energy of H, $M_i^2 = \int(\mathrm{d}f/\mathrm{d}E(1/u)\,\mathrm{d}E$ with $\mathrm{d}f/\mathrm{d}E$ the differential oscillator strength and $c$ a collision parameter. It is possible to extract $M_i^2$ from the slope of a plot of $\sigma u$ versus $\ln(u)$ (Fano-plot).

Conversely, for slow collisions the combined system of incoming electron and target molecule has to be considered, leading in the exit channel to a full three-body problem. Quantum-mechanical (approximate) calculations are difficult and have been carried out only for a few selected examples. Therefore, other methods have been developed with the goal of obtaining reasonably accurate cross sections using either classical or semiclassical theories and by devising semiempirical formulae. Some of these concepts are based on the Born–Bethe formula [22] and on the observation that the ejection of an atomic electron with quantum numbers $(n,l)$ is approximately proportional to the mean-square radius of the electron shell $(n,l)$. This leads also to proposed correlations of the ionization cross section with polarizability, diamagnetic susceptibility, number of electrons, and the additivity of atomic cross sections to give molecular cross sections. Moreover, there exist simple empirical relations such as between the energy position $u_{\mathrm{max}}$ of the maximum cross section and the ionization energy $\sigma_{\mathrm{max}} u_{\mathrm{max}} \approx (R/\mathrm{IE})^2$, which is the high-energy limit of the classical Thomson formula.

A particularly useful, simple and versatile formula is the semi-classical Deutsch–Maerk (DM) approach giving the ionization cross section for single ionization $\sigma_i$ in terms of $\xi_{n,l}$ (the number of equivalent electrons in the $(n,l)$ subshell), $r_{n,l}$ (the electron radius of the $(n,l)$ shell), and some weighting factors $g_{n,l}$. It has been demonstrated recently that this formula may be applied not only to calculate ionization cross sections for ground-state atoms (**Figure 10**), but also for the calculation of cross sections for excited atomic species, for molecules (**Figure 13**), for radicals, for inner-shell ionization (**Figure 14**), for ionization of ions and even clusters, thereby spanning a range of cross-section values from about $10^{-34}$ m$^2$ for 14-fold ionization of Si all the way up to about $5\times10^{-19}$ m$^2$ for $C_{60}$ and in the case of inner-shell ionization energies from threshold up to $10^9$ eV.

## List of symbols

$a_0$ = Bohr radius; $\mathrm{d}f/\mathrm{d}E$ = differential oscillator strength; $E$ = electron collision energy; $i$ = ionization current (t, total; ms, mass selected); IE = ionization energy; $N_\mathrm{t}$ = number of target particles cm$^{-3}$; $n(0)_\mathrm{e}$ = initial electron intensity [cm$^{-2}$ s$^{-1}$]; $u$ (= $E/R$) = reduced electron energy [$R$ = H ionization energy]; $W$ = electron energy after ionization collision; $\sigma$ = ionization cross section (t, total; c, counting; $zi$ for ion $i$ of charge $ze$); $\delta$ = polar angle of outgoing electron.

*See also:* **Chemical Ionization in Mass Spectrometry; Fragmentation in Mass Spectrometry; Metastable Ions; Photoelectron Spectroscopy; Photoionization and Photodissociation Methods in Mass Spectrometry.**

## Further reading

Berkowitz J (1979) *Photoabsorption, Photoionization and Photoelectron Spectroscopy*. New York: Academic Press.

Beckey HD (1977) *Principles of Field Ionization and Field Desorption Mass Spectrometry*. Oxford: Pergamon Press.

Christophorou LG (1984) *Electron–Molecule Interactions and Their Applications*. Orlando: Academic Press.

Christophorou LG, Illenberger E and Schmidt WF (1994) *Linking the Gaseous and Condensed Phases of Matter: The Behaviour of Slow Electrons*. New York: Plenum.

Duncan ABF (1971) *Rydberg Series in Atoms and Molecules*. New York: Academic Press.

Field FH and Franklin JL (1970) *Electron Impact Phenomena and the Properties of Gaseous Ions*. New York: Academic Press.

Futrell JH (1986) *Gaseous Ion Chemistry and Mass Spectrometry*. New York: Wiley-Interscience.

Hasted JB (1972) *Physics of Atomic Collisions*. London: Butterworths.

Illenberger E and Momigny J (1992) *Gaseous Molecular Ions*. Darmstadt: Steinkopff.

Levsen K (1978) *Fundamental Aspects of Organic Mass Spectrometry*. Weinheim: Verlag Chemie.

Maerk TD and Dunn GH (1985) *Electron Impact Ionization*. Wien: Springer

Marr GV (1967) *Photoionization Processes in Gases*. New York: Academic Press.

Massey HSW, Burhop EHS and Gilbody HB (1969) *Electronic and Ionic Impact Phenomena*. Oxford: Clarendon Press.

Rose ME and Johnstone RAW (1982) *Mass Spectrometry for Chemists and Biochemists*. London: Cambridge University Press.

# IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds

**Claudio Pettinari** and **Carlo Santini**, Università di Camerino, Italy

Copyright © 1999 Academic Press

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
|---|
| **Applications** |

Infrared and Raman spectroscopy techniques are being employed to an ever-increasing extent for the recognition and the quantitative analysis of structural units in unknown compounds. The infrared region of the electromagnetic spectrum extends from the red end of the visible spectrum to the microwave region. It includes radiation at wavelengths between 14 000 and 20 cm$^{-1}$, but the spectral range used most is the mid-infrared one, which covers frequencies from 4000 to 200 cm$^{-1}$. Infrared spectrometry involves examination of the twisting, bending, rotating and vibrational motions of atoms in a molecule. Upon interaction with infrared radiation, portions of the incident radiation are absorbed at specific wavelengths. The multiplicity of vibrations occurring simultaneously produces a highly complex absorption spectrum that is uniquely characteristic of the functional groups that make up the molecule and of the overall configuration of the molecule as well.

When monochromatic radiation is scattered by molecules, a small fraction of the scattered radiation is observed to have a different frequency from that of the incident radiation; this is known as the *Raman effect*. Since its discovery in 1928, the Raman effect has been an important method for the elucidation of molecular structure, for locating various functional groups or chemical bonds in molecules and for the quantitative analysis of complex mixtures. Although vibrational Raman spectra are related to infrared absorption spectra, a Raman spectrum arises in a different manner and thus often provides complementary information. Vibrations that are active in Raman scattering may be inactive in infrared, and vice versa. A unique feature of Raman scattering is that each line has a characteristic polarization, and polarization data provide additional information about molecular structure.

Raman spectra can be used to study materials in aqueous solutions, a medium that transmits infrared radiation very poorly. Another advantage is the ability to examine the entire vibrational spectrum with one instrument, unlike infrared spectroscopy in

which the far-infrared is usually scanned separately from the mid-infrared. For quantitative determinations, the Raman scattering power is directly proportional to concentration, whereas, in infrared spectroscopy, it is the logarithm of the ratio of incident to transmitted power, the absorbance, that is proportional to concentration. Furthermore, the application of infrared spectroscopy to the identification of inorganic and coordination compounds has been somewhat unsuccessful because many simple inorganic compounds such as the borides, nitrides and oxides do not exhibit absorption in the region between 4000 and 600 $cm^{-1}$, which for many years was the range of the infrared region covered by most commercially available spectrometers.

In addition, the Raman technique has proved to be particularly valuable in the study of single crystals where the infrared technique has greater limitations on sample size and geometry. Polarization data obtained from Raman spectra allow unambiguous classification of fundamentals and lattice modes into the various symmetry classes. Although Raman spectroscopy will never challenge X-ray diffraction as a tool for quantitative structural analysis, it is the preferred technique when qualitative information is sufficient because it is faster and less expensive.

## Experimental aspects

The spectrum obtained for a given sample of inorganic, coordination or organometallic compounds depends upon the physical state of the sample. Gaseous samples usually exhibit a rotational fine structure which is damped in solution because collisions of molecules in the condensed phase occur before a rotation is completed. In addition to the difference in resolved fine structure, the number of absorption bands and the frequencies of the vibrations vary in the different states. Often there are more bands in the liquid state than in the gaseous state of a substance, and frequently new bands below 300 $cm^{-1}$ caused by *lattice vibrations*, i.e. translational and torsional motions of the molecules in the lattice (*phonon modes*) appear in the solid state spectrum. The stronger intermolecular forces that exist in the solid and liquid states compared with those in the gaseous state are often the cause of slight shifts in the frequencies.

An additional complication arises if the unit cell of the crystal contains more than one chemically equivalent molecule. When this is the case, the vibrations in the individual molecules can couple with each other. This intermolecular coupling can give rise to frequency shifts and band splitting.

Sample handling also presents a number of problems in the infrared region. For example, there is no rugged window material for cuvettes that is transparent and also inert over this region. The alkali halides are widely used, particularly sodium chloride, which is transparent at wavelengths as long as 625 $cm^{-1}$. For frequencies less than 600 $cm^{-1}$, polyethylene cells are frequently used.

Samples that are liquid at room temperature are usually scanned in their neat form or in solution. Unfortunately, not all substances are soluble to a reasonable concentration in a solvent not absorbing in the regions of interest. In some cases use is made of solvents susceptible to hydrogen-bonding effects which can alter the vibrational frequency characteristic of the compound examined; the stronger the hydrogen bonding, the more the fundamental frequency is lowered.

Powders or solids can be examined as a thin paste or mull (mineral oil, nujol, hexachlorobutadiene, perfluorokerosene) or through the pellet technique which allows quantitative analyses to be performed since accurate measurements can be made of the ratio of weight of sample to internal standard in each disk. Modern technique using diamond ATR is removing the need for mulls and disks.

The use of laser excitation allows Raman spectroscopy to be performed on specimens in almost any state: liquid, solution, transparent solid, translucent solid, powder, pellet or gas. Water is a weak scatterer and therefore an excellent solvent for Raman work. This has important consequences in studies of biochemical interest and in the pharamaceutical industry. Water is a bad solvent in FT Raman because the Raman light is self absorbed.

## Group frequencies and band assignments

The infrared spectrum of a compound is essentially the superposition of absorption bands of specific functional groups: even small interactions with the surrounding atoms of the molecule impose the stamp of individuality on the spectrum of each compound. For qualitative analysis, one of the best features of an infrared spectrum is that the absorption or the *lack of absorption* in specific frequency regions can be correlated with specific stretching and bending motions and, in some cases, with the relationship of these groups to the rest of the molecule.

In the near-infrared region (NIR), which meets the visible region at about 12 500 $cm^{-1}$ and extends to about 4000 $cm^{-1}$, generally there are many absorption bands that result from harmonic overtones

of fundamental and combination bands often associated with hydrogen atoms.

Many useful correlations have been found in the mid-infrared region. This region is divided into the 'group frequency' region, 4000–1300 $cm^{-1}$, and the fingerprint region, 1300–650 $cm^{-1}$. In the group frequency region the principal absorption bands are assigned to vibration units consisting of only two atoms of a molecule: in the interval from 4000 to 2500 $cm^{-1}$, the absorption is characteristic of hydrogen stretching vibrations with elements of mass 19 or less. The intermediate frequency range, 2500–1540 $cm^{-1}$ (*unsaturated* region) contains triple bond frequencies which appear from 2500 to 2000 $cm^{-1}$ and double bond frequencies from 2000 to 1540 $cm^{-1}$. In the region between 1300 and 650 $cm^{-1}$ there are single bond stretching frequencies and bending vibrations (skeletal frequencies) of polyatomic systems that involve motions of bonds linking a substituent group to the remainder of the molecule.

The region 667–10 $cm^{-1}$ contains the bending vibrations of carbon, nitrogen, oxygen and fluorine with atoms heavier than mass 19, and additional bending motions in cyclic or unsaturated systems. The low-frequency molecular vibrations in the far-infrared are particularly sensitive to changes in the overall structure of the molecule; thus the far-infrared bands often differ in a predictable manner for different isomeric forms of the same basic compound.

The far-infrared frequencies of organometallic compounds are often sensitive to the metal ion or atom, and this can be used advantageously in the study of coordination bonds. Moreover, this region is particularly well suited to the study of organometallic or inorganic compounds whose atoms are heavy and whose bonds are inclined to be weak.

In Raman spectroscopy those vibrations that originate in relatively nonpolar bonds with symmetrical charge distributions and that are symmetrical in nature produce the greatest polarizability changes and are the most intense. Vibrations from –C=C–, –C≡C–, –C≡N, –C=S, –C–S–, –S–S–, –N=N– and –S–H bonds are readily observed. Absorption bands due to the skeletal vibrations of finite chains and rings of saturated and unsaturated hydrocarbons are generally sharper in Raman than in infrared spectra. Aromatic compounds exhibit particularly strong spectra with a strong ring deformation mode at $1600 \pm 30$ $cm^{-1}$. Skeletal motions are very characteristic and highly useful for cyclic and aromatic rings, steroids and long chains of methylenes.

In solid samples, sharpening and intensifying of certain bands appear to be a function of crystallinity. For all molecules that have a centre of symmetry, a band allowed in the infrared is forbidden in the Raman, and vice versa. In molecules with symmetry elements other than a centre of symmetry, certain bands may be active in the Raman, infrared, both or neither. For a complex molecule that has no symmetry, all of the normal vibrational modes are allowed in both the infrared and Raman spectra. One other strong difference is the tendency for peaks in a Raman spectrum to have a greater range of intensities, from very weak to very strong. These factors, plus the contrasting relative intensities for a given group, are the main basis for making more confident assignment of chemical structure through the combination of infrared and Raman data. For example, in infrared spectra the 3300 $cm^{-1}$ region is badly obscured by intense OH absorption, whereas the same region in Raman is very helpful in the assignment of NH and CH stretchings because the OH absorption is weak.

## Applications

### Selection rules

Molecular symmetry is very important in determining the infrared activity and degeneracy of a molecular vibration. When a molecule is present in a crystal, the symmetry of the surroundings of the molecule in the unit cell, the so called *site symmetry*, determines the selection rules. Often, bands forbidden in the gaseous state or in solution appear in the solid, and degenerate vibrations in the gaseous state are split in the solid. To show this effect, let us consider the infrared spectra of carbonate derivatives. The infrared (**Figure 1A**) and Raman spectra of $CaCO_3$ in calcite, where the carbonate ion is in a site of $D_3$ symmetry, contain the following bands (in $cm^{-1}$): $\nu_1$, 1087 (R); $\nu_2$, 879 (IR); $\nu_3$, 11 432 (IR, R); $\nu_4$, 710 (IR, R). The infrared spectra of $CaCO_3$ (**Figure 1B**) in aragonite, where the site symmetry for the carbonate is $C_8$ is different: $\nu_1$ is infrared active and both $\nu_3$ and $\nu_4$ split into two bands.

In **Figure 2** some of the possible structures for a compound with an empirical formula $NSF_3$ are shown whereas the calculated number and symmetry of bands for the illustrated structures are listed in **Table 1**. In the infrared spectrum of this compound, six intense bands are found, in agreement with a $C_{3v}$ structure but certainly not in conflict with other possible structures. It was found that the four bands $\nu_1$, $\nu_2$, $\nu_3$ and $\nu_5$, have *P*, *Q* and *R* branches indicating that the molecule investigated is a symmetric top molecule and supports a $C_{3v}$ structure. The spectral data are contained in **Table 2.**

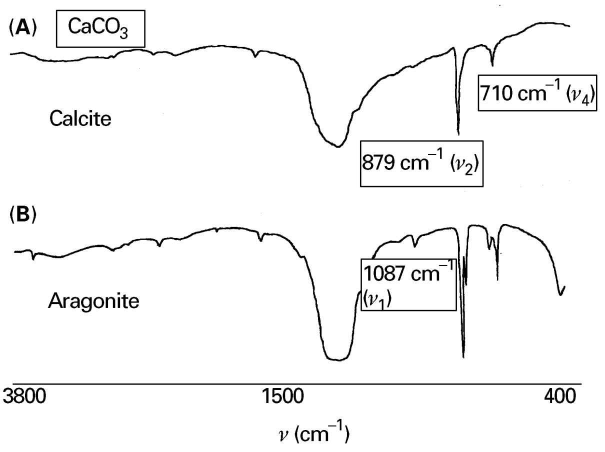



**Figure 1** Infrared spectra of $CaCO_3$ in calcite and in aragonite.

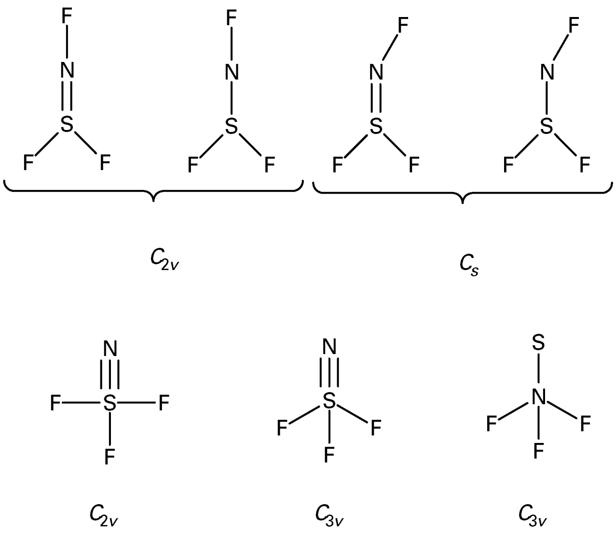



**Figure 2** Possible structures for $NSF_3$.

**Table 1** Calculated number of bands for some of the possible structures for $NSF_3$

| $C_{3v}$ | $C_{2v}$ | $C_s$ |
|---|---|---|
| $3A_1$ (IR active ) | $4A_1$ (IR active) | $7A'$ (IR active) |
| | $3B_1$ (IR active) | $2A''$ (IR active) |
| $3E$ (IR active ) | $2B_2$ (IR active) | |

**Table 2** Fundamental vibration frequencies (cm$^{-1}$) for $NSF_3$

| | | | |
|---|---|---|---|
| $\nu_1$ | 1515 | $\nu_4$ | 811 |
| $\nu_2$ | 775 | $\nu_4$ | 429 |
| $\nu_3$ | 521 | $\nu_6$ | 342 |

These data, combined with an NMR study, have been employed to support the $C_{3v}$ structure $F_3S$–N. The $F_3NS$ structure has not been eliminated by these data; however, the other observed bands support $F_3SN$: the S–N stretch is the only fundamental vibration that could be expected at a frequency of 1515 cm$^{-1}$. The NF stretching vibrations in $NF_3$ occur at 1031 cm$^{-1}$ and would not be expected to be as high as 1515 cm$^{-1}$ in $F_3N$–S. The $F_3SN$ structure has been demonstrated on the basis of microwave and mass spectroscopic studies.

### Change in the spectra of donor molecules upon coordination

In the infrared spectrum of *N*,*N*-dimethylacetamide in $CCl_4$, a strong absorption band at 1662 cm$^{-1}$ has been found due to a highly coupled carbonyl absorption. The low frequency with respect to that found for acetone (1715 cm$^{-1}$) is attributed to a resonance interaction with the lone pair on the nitrogen (**Figure 3**).

A decrease in the frequency of this band has generally been found upon complexation of the molecule with Lewis acids. This has been attributed to coordination of the oxygen to the acid. Oxygen coordination produces a decrease in the carbonyl force constant by draining the $\pi$ electron density out of the carbonyl group, which causes the observed decrease in the carbonyl frequency. The absence of any absorption in the carbonyl region on the high-frequency side of the uncomplexed carbonyl band further supports coordination of this molecule through the oxygen. On the other hand, in $Pd(NH_2CONH_2)_2Cl_2$, and $Pt(NH_2CONH_2)_2Cl_2$, coordination of the ligand $NH_2CONH_2$ through the nitrogen occurs, which results in a decrease of the C–N vibration frequency and in a higher frequency carbonyl absorption.

Analogously, a decrease in the S–O stretching frequency, indicative of oxygen coordination, is observed when dimethyl sulfoxide or tetramethylene sulfoxide is complexed to many metal ions, iodine and phenol, whereas the S–O stretching frequency increases in the palladium complex of dimethyl sulfoxide, compared to free sulfoxide, which suggests sulfur coordination.

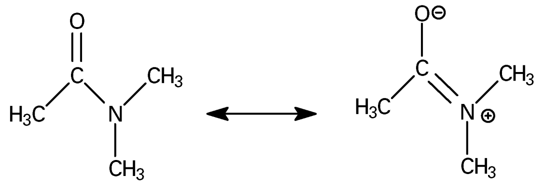


**Figure 3** Resonance interaction between the carbonyl and the lone pair on the nitrogen in *N,N′*-dimethylacetamide.

The infrared spectra of ethylenediaminetetraacetic acid metal complexes have been used to distinguish between tetradentate, pentadentate or hexadentate coordination of the ligand on the basis of the absorption bands in the carbonyl region corresponding to free and complexed carbonyl groups.

The change in C≡N stretching frequency of nitriles and metal cyanides resulting from their interaction with Lewis acids has also attracted considerable interest. When such molecules are coordinated to Lewis acids not generally involved in $\pi$ back-bonding, the C≡N stretching frequency increases. A combined molecular orbital and normal coordinate analysis of acetonitrile and some of its adducts indicates that a slight increase is to be expected from this effect, but that the principal contribution to the observed shift arises from an increase in the C≡N force constant. This increase is mainly due to an increase in the C≡N $\sigma$ bond strength from nitrogen rehybridization. In those systems where there is extensive $\pi$ back-bonding from the acid into the $\pi^*$ orbitals of the nitrile group, the decreased $\pi$ bond energy accounts for the decreased frequency.

### Change in symmetry upon coordination

Some of the most useful applications of infrared spectroscopy in the area of coordination and organometallic chemistry originate from the change in the symmetry of a ligand upon coordination. For example, when small molecules (e.g. $N_2$, $O_2$ and $H_2$) are linked to transition metal ions a symmetry change occurs which has a strong influence on the infrared spectra. The stretching vibration of free $N_2$ is infrared inactive but Raman active (2331 cm$^{-1}$). If the azide ion is added to $Ru(NH_3)_5H_2O^{3+}$, $Ru(NH_3)_5N_2^{2+}$ is obtained, which exhibits in infrared a sharp absorption due to $\nu$(N–N) at 2130 cm$^{-1}$.

Infrared spectroscopy is a very effective tool for determining the nature of the bonding of the sulfate group in its derivatives. The sulfate ion ($T_d$ symmetry) has two infrared bands in the 1200–600 cm$^{-1}$ region, one assigned to $\nu_3$ at 1104 cm$^{-1}$ and one to $\nu_4$ at 613 cm$^{-1}$. In the complex $[Co(NH_3)_5OSO_3]Br$ the coordinated sulfate group has lower symmetry, $C_{3v}$, with six new bands at 970 ($\nu_1$), 438 ($\nu_2$), 1032 to 1044, and 1117 to 1143 (from $\nu_3$), and 645 and 604 cm$^{-1}$ (from $\nu_4$). In a bridged sulfate group, the symmetry is lowered to $C_{2v}$ and even more bands appear: the $\nu_3$ band is split into three peaks and the $\nu_4$ band into three other peaks.

Five-coordinate addition compounds such as $(CH_3)_3SnCl\cdot(CH_3)_2SO$ had a structure in which the three methyl groups were in the equatorial positions: in fact, the symmetric Sn–C stretch present in $(CH_3)_3SnCl$ at 545 cm$^{-1}$ disappeared in the addition compound where a single Sn–C vibration due to the asymmetric Sn–C stretch was detected at 551 cm$^{-1}$. This is in accordance with an isomer with three methyl groups in the equatorial position: in fact, a small dipole moment change is associated with the symmetric stretch in compounds having this structure. All the other possible structures possess at least two different Sn–C vibrations.

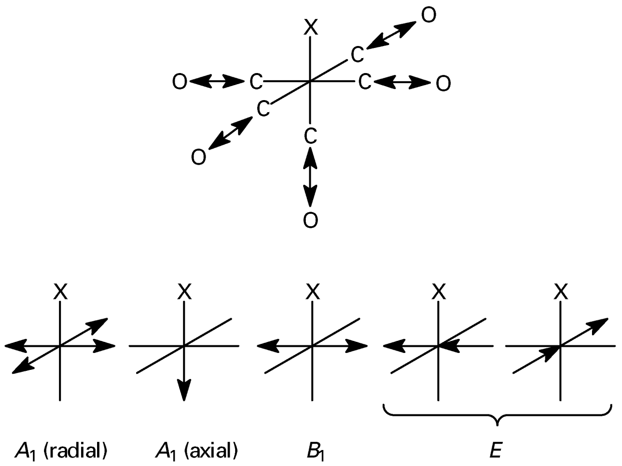



**Figure 4** Carbonyl stretching vibrations of $M(^{12}CO)_5X$ derivatives.

Also, the carbonyl stretching vibrations of $M(CO)_5$ X (and their $^{13}$C-substituted) derivatives can be employed to study change in symmetry upon coordination. The operations or the $C_{4v}$ point group for all-$^{12}$CO compounds such as those in **Figure 4** produce $2A_1$ (one radial and one axial), $B_1$ (radial) and $E$ (radial) vibrations, three of which are infrared active and four Raman active.

If the molecule has an axial $^{13}$CO, the symmetry is still $C_{4v}$ and the $B_1$ and $E$ modes (which have no contribution from the axial group) will be unaffected by the mass change. The two $A_1$ modes having the same symmetry will mix, so isotopic substitution could affect both; but it will have a major effect on $A_1$ (axial). If a $^{13}$CO is located in the radial position, the symmetry is lowered to $C_s$, leading to the possible vibrations shown in **Figure 5**.

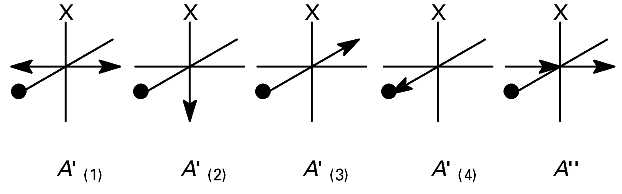



**Figure 5** Possible vibrations for a $M(^{12}CO)_4(^{13}CO)X$ derivative, containing the $^{13}$CO in the radial position.

Only the A″ will be unaffected by the mass change, and it resembles the *E* mode in the all-$^{12}CO$ molecule. All of the other four modes can mix and could be shifted by distributing the mass effect over all four modes. This would shift the vibrations to lower frequencies than those in the all-$^{12}CO$ molecule. Several normal coordinate analyses on such systems, employing a large number of isotopes, have been made which indicates that as the formal positive charge on the central atom increases in forming $M(CO)_5Br$ from $M(CO)_6$, the $\sigma$-bonding increases and the extent of π back-bonding decreases.

### Metal-isotope spectroscopy

The metal–ligand vibrations appear in the low-frequency region (600–100 cm$^{-1}$) and provide direct information about the structure of the coordination sphere and the nature of the metal–ligand bond. It is difficult to make unequivocal assignments of metal–ligand vibrations since the interpretation of the low-frequency spectrum is complicated by the appearance of ligand vibrations, as well as lattice vibrations in the case of solid-state spectra. However, different methods can be employed: for example, the metal–ligand vibrations are absent in the spectrum of the free donor (it should be taken into account that in some cases new ligand vibrations, activated by complex formation, can appear).

Metal–ligand vibrations are also metal sensitive and are shifted by changing the metal or its oxidation state: this method is applicable only to isostructural metal complexes and it does not provide definitive assignments since some ligand vibrations (for example the chelate ring deformations) are also metal sensitive.

If the ligand is isotopically substituted, the metal–ligand vibrations present an isotope shift: the $\nu$(Ni–N) in $[Ni(NH_3)_6]Cl_2$ at 334 cm$^{-1}$ is shifted to 318 cm$^{-1}$ upon deuteration of the ammonia groups. It is possible, with this method, to assign the metal–ligand vibrations of chelate compounds such as oxamido ($^{14}N/^{15}N$) and acetylacetonato ($^{16}O/^{17}O$) complexes. However isotopic substitution of the α-atom (an atom directly bonded to the metal) causes shifts not only of metal–ligand vibrations but also of ligand vibrations involving the motion of the α-atom. Finally, the frequency of a metal–ligand vibration may also be predicted if the metal–ligand stretching and other force constants are known as a priori.

The ‘metal-isotope technique’ is the best one for obtaining reliable metal–ligand assignments. Isotope pairs such as (H/D) and ($^{16}O/^{17}O$) had been used in the past routinely by many spectroscopists. Isotopic pairs of heavy metals such as ($^{58}Ni/^{62}Ni$) and ($^{104}Pd/^{110}Pd$) were not employed until 1969 when the first report on the assignments of the Ni–P vibrations of *trans*-$Ni(PEt_3)_2X_2$ (X = Cl and Br) was made. The delay in their use was probably due to the high cost of the pure metal isotopes.

The magnitudes of metal isotope shifts are generally of the order of 2–10 cm$^{-1}$ for stretching modes and 0–2 cm$^{-1}$ for bending modes, whereas the experimental error in measuring the frequency could be as small as ±0.2 cm$^{-1}$.

In highly symmetrical molecules ($T_d$, $O_h$, etc.), the central atom does not move during the totally symmetric vibration and no metal-isotope shifts are expected. When the central atom is coordinated by several donor atoms, multiple isotope labelling is necessary to distinguish different coordinate bond-stretching vibrations. For example, complete assignments of bis(β-diketonato)nickel(II) derivatives require $^{16}O/^{18}O$ as well as $^{58}Ni/^{62}Ni$ isotope shift data. The metal-isotope technique is often indispensable not only in assigning the metal–ligand vibrations but also in refining metal–ligand stretching force constants in normal coordinate analysis. The presence of vibrational coupling between metal–ligand and other vibrations can also be detected by combining metal-isotope data with normal coordinate calculations, since both experimental and theoretical isotope shift values will be smaller when such couplings occur.

The metal-isotope technique is very important in biological molecules such as haem proteins: structural and bonding information about the active site (iron porphyrin) can be obtained through definitive assignments of coordinate bond-stretching vibrations around the iron atom. Using resonance Raman techniques, it is possible to observe iron porphyrin and iron–axial ligand vibrations without interference from peptide chain vibrations. Thus, these vibrations can be assigned by comparing a resonance Raman spectrum of a natural haem protein with that of a $^{54}Fe$-reconstituted haem protein.

## Inorganic compounds

This and the following paragraphs are devoted to a general resumé of the more common and important spectroscopic and structural information on inorganic compounds: the assignment of the fundamental vibrations are based on point group symmetry.

Diatomic molecules have only one vibration along the chemical bond: in homopolar AA($D_{\infty h}$) molecules, this vibration is Raman active but not infrared active, whereas the stretching vibration is both Raman and infrared active in heteropolar AB($D_{\infty v}$) molecules.

**Table 3** Observed frequencies of diatomic molecules (cm$^{-1}$)

| *Molecule* | $\nu$ | *Molecule* | $\nu$ |
|---|---|---|---|
| $N_2$ | 2360 | HF | 4138 |
| CO | 2138 | HCl | 2991 |
| NO | 1880 | HBr | 2650 |
| $O_2^+$ | 1865 | HI | 2309 |
| $O_2$ | 1580 | AlH | 1593[a] |
| $O_2^-$ | 1097 | $[OH]^-$ | 3637[b] |
| $O_2^{2-}$ | 766 | $[CN]^-$ | 2080[c] |

[a] Matrix.
[b] Solid state.
[c] Solution.

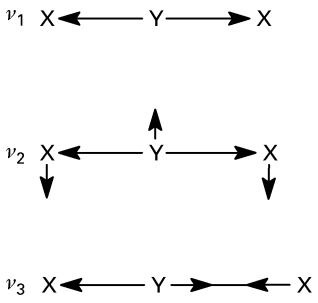
**Table 3** lists some relevant observed frequencies for diatomic molecules, ions and radicals: the effect of the charge on the frequency in the series of the dioxygen species is worthy of note and also the vibrational frequencies characteristic of hydrogen halides which generally polymerize in the condensed phases. The HX stretching bands are shifted markedly to lower frequencies upon polymerization.

IR and Raman spectra can be properly used to select between bent and linear structures in triatomic inorganic molecules (**Table 4**): the linear $X_3(D_{\infty h})$ and XAX ($D_{\infty h}$) type molecules have three normal modes (**Figure 6**), $\nu_1$, Raman active but not infrared active, and $\nu_2$ and $\nu_3$, infrared active but not Raman active. In the case of bent $X_3$ ($C_{2v}$) and XAX ($C_{2v}$) type molecules all three vibrations are infrared and Raman active.

Pyramidal $XY_3$ ($C_{3v}$) molecules have four normal modes of vibration (**Figure 7**) (**Table 5**), all infrared and Raman active, which have been used, for example, to confirm the presence of the hydronium ($H_3O^+$) ion in hydrated acids. Substitution of one of the Y atoms of a pyramidal $XY_3$ molecule by a Z atom lowers the symmetry from $C_{3v}$ to $C_s$. Then the degenerate vibrations split into two bands [$\nu_d(XY) \rightarrow \nu_s(XY) + \nu_a(XY)$; $\delta_d(YXY) \rightarrow \delta_s(YXY) + \delta_a(YXZ)$] ($\delta$ = bending) and all six vibrations become infrared and Raman active (**Table 5**).

**Table 4** Vibrational frequencies of triatomic molecules (cm$^{-1}$)

| *Compound* | *Structure* | $\nu_1$ | $\nu_2$ | $\nu_3$ |
|---|---|---|---|---|
| $BeCl_2$ | Linear | (680) | 345 | 1555 |
| $MgCl_2$ | Linear | 327 | 93 | 601 |
| $SnCl_2$ | Bent | 354 | (120) | 334 |
| $[CuCl_2]^-$ | Linear | 300 | 109 | 405 |
| $BaCl_2$ | Bent | 225 | – | 260 |
| $O_3$ | Bent | 1135 | 716 | 1089 |
| $S_{3(g)}$ | Bent | 585 | 490/310 | 651 |
| $ClOCl_{(s)}$ | Bent | 631 | 296 | 671 |
| SCS | Linear | 658 | 397 | 1533 |
| $O^{13}CO$ | Linear | – | (649) | 2284 |
| NNO | Linear | 2224 | 589 | 1285 |



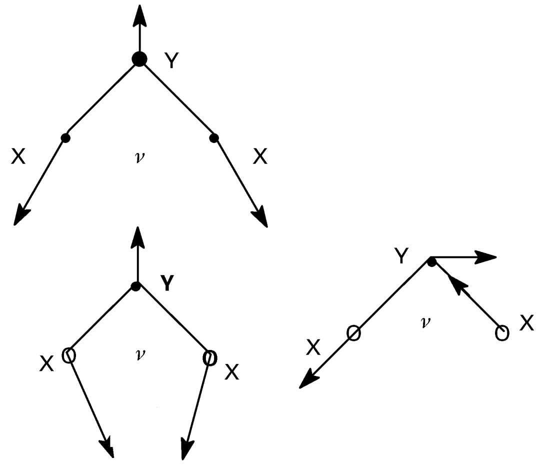




Figure 6 Normal modes of vibration for triatomic inorganic molecules.

The four normal modes of vibration of planar $XY_3$ ($D_{3h}$) molecules are shown in **Figure 8**, whereas significant vibrational frequencies of planar $XY_3$ molecules are listed in **Table 5**. Pyramidal and planar

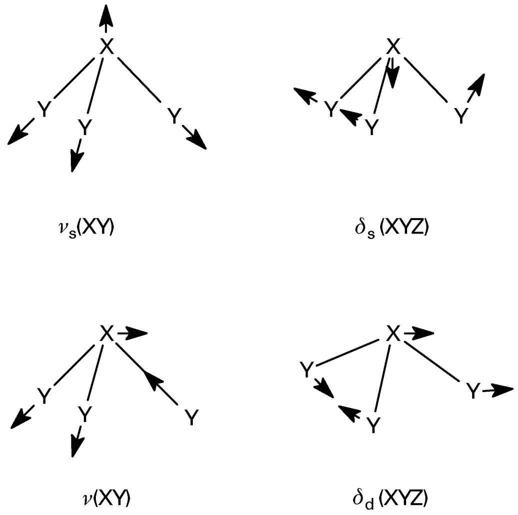



**Figure 7** Normal modes of vibration for a pyramidal $XY_3$ molecule.

**Table 5** Vibrational frequencies of tetratomic molecules (cm$^{-1}$)

| *Compound* | *Structure* | *State* | $\nu_1$ | $\nu_2$ | $\nu_3$ | $\nu_4$ |
|---|---|---|---|---|---|---|
| $NH_3$ | Pyramidal | Solid | 3223 | 1060 | 3378 | 1646 |
| $[OH_3]^+$ | Pyramidal | Solution | 3560 | 1095 | 3510 | 1600 |
| $PH_3$ | Pyramidal | Gas | 2327 | 990 | 2421 | 1121 |
| $NF_3$ | Planar | Gas | 1035 | 649 | 910 | 500 |
| $[SnCl_3]^-$ | Planar | Solution | 297 | 128 | 256 | 103 |
| $BiCl_3$ | Planar | Gas | 342 | 123 | 322 | 107 |
| $BiBr_3$ | Planar | Gas | 220 | 77 | 214 | 63 |
| $BiI_3$ | Planar | Solid | 146 | 90 | 115 | 71 |

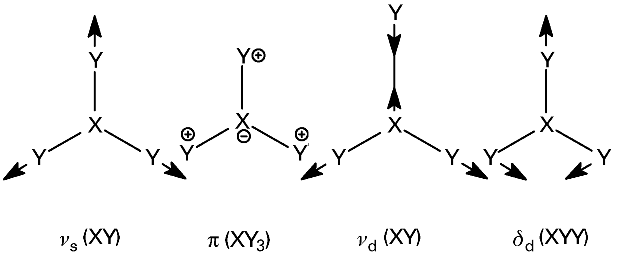



**Figure 8** Normal modes of vibration for a planar $XY_3$ molecule.

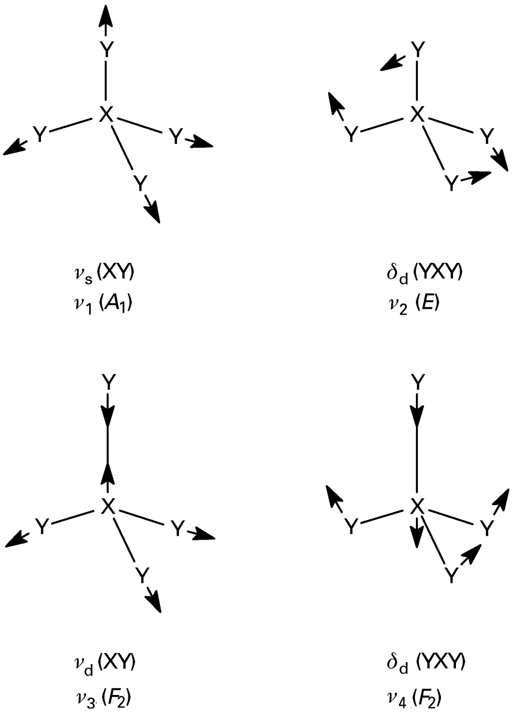



**Figure 9** Normal modes of vibration for a tetrahedral $XY_4$ molecule.

structures can be distinguished easily on the basis of the difference in selection rules.

Tetrahedral $XY_4$ molecules have four normal modes of vibration (**Figure 9**) which are all Raman active, whereas only $\nu_3$ and $\nu_4$ are infrared active. In $XH_4$-type molecules (**Table 6**) the following trend is observed: $\nu_3 > \nu_1$ and $\nu_2 > \nu_4$. In $MX_4$ derivatives the MX stretching frequency generally decreases on going from F to I. The average values of $\nu$(MBr)/$\nu$(MCl) and $\nu$(MI)/$\nu$(MCl) calculated from selected compounds are 0.76 and 0.62, respectively, for $\nu_3$, and 0.61 and 0.42, respectively, for $\nu_1$. These values are very useful in the assignment of the MX stretching bands in halogeno complexes.

The effect of changing the oxidation state on the MX stretching frequency has been extensively studied with $[FeX_4]^-$ and $[FeX_4]^{2-}$ (X = Cl and Br)

**Table 6** Vibrational frequencies of tetrahedral compounds (cm$^{-1}$)

| *Compound* | $\nu_1$ | $\nu_2$ | $\nu_3$ | $\nu_4$ | $\nu_5$ | $\nu_6$ |
|---|---|---|---|---|---|---|
| $CF_4$ | 908 | 434 | 1283 | 631 | | |
| $SiCl_4$ | 423 | 145 | 616 | 220 | | |
| $[MnCl_4]^{2-}$ | 256 | – | 278, 301 | 120 | | |
| $[BrO_4]^-$ | 801 | 331 | 878 | 410 | | |
| $FSiCl_3$ | 465 | 948 | 239 | 640 | 282 | 167 |
| $FCCl_3$ | 538 | 1080 | 351 | 848 | 243 | 395 |
| $ONF_3$ | 743 | 1691 | 528 | 883 | 558 | 400 |
| $[IF_4]^{+a}$ | 728 | 614 | 345 | 263 | – | – |

[a] Distorted, $\nu_7$: 388 ; $\nu_8$: 719; $\nu_9$: 319.

species: it has been found that the MX stretching frequency increases as the oxidation state of the metal increases. In tetrahedral $MO_4^-$, $MS_4^-$ and $MSe_4^-$ compounds, the $\nu_1/\nu_3$ ratio increases as the negative charge of the anion increases; for anions having the same negative charge and the central atom belonging to the same group of the periodic table, the $\nu_1/\nu_3$ ratio increases with increasing mass of the central atom, and finally the $\nu_1/\nu_3$ ratio increases with increasing negative charge of the anion in isoelectronic ions, with the mass of the central atom approximately constant.

**Figure 10** shows the seven normal modes of vibration of square-planar $XY_4$ molecules. Vibra-tions $\nu_3$, $\nu_6$ and $\nu_7$ are infrared active, whereas $\nu_1$, $\nu_2$ and $\nu_4$ are Raman active. Some vibrational frequencies for selected square-planar $XY_4$ molecules are reported in **Table 7**.

$XY_5$ molecules may be trigonal bipyramidal ($D_{3h}$) or tetragonal pyramidal ($C_{4v}$). Trigonal bipyramidal have eight normal modes of vibration, six of which ($A_1'$, $E'$ and $E''$) are Raman active and five ($A_2''$ and $E'$) are infrared active. Normal coordinate studies have demonstrated that in neutral trigonal bipyramidal $MX_5$ compounds, the equatorial bonds are stronger than axial ones.

In tetragonal $XY_5$ molecules, the axial stretching frequency ($\nu_1$) is generally higher than equatorial

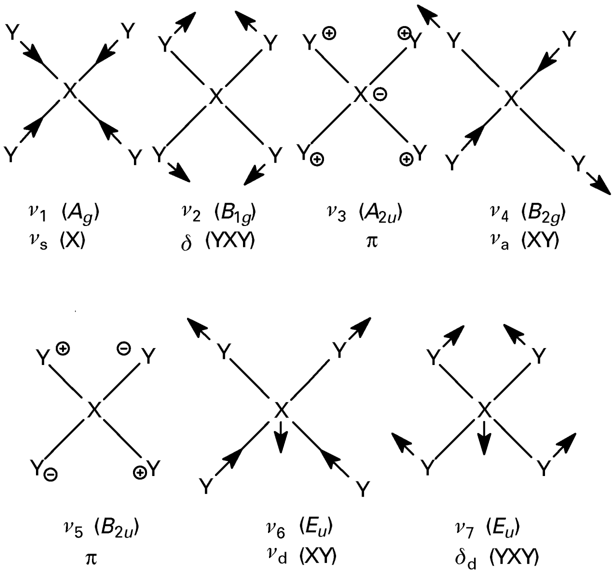



**Figure 10** Normal modes of vibration for a square planar $XY_4$ molecule.

stretching frequencies ($\nu_2$, $\nu_4$ and $\nu_7$). This is the opposite of what is found for trigonal bipyramidal $XY_5$ molecules.

**Figure 11** shows the six normal modes of vibration of an octahedral $XY_6$ molecule. Vibrations $\nu_1$, $\nu_2$ and $\nu_5$ are Raman active, whereas only $\nu_3$ and $\nu_4$ are infrared active; $\nu_6$ is inactive in both, and its frequency is estimated from an analysis of combination

**Table 7** Vibrational frequencies of square planar $XY_4$ molecules (cm$^{-1}$)

| *Compound* | $\nu_1$ | $\nu_2$ | $\nu_3$ | $\nu_4$ | $\nu_6$ | $\nu_7$ |
|---|---|---|---|---|---|---|
| $XeF_4$ | 554 | 218 | 291 | 524 | 586 | (161) |
| $[AuCl_4]^-$ | 347 | 171 | – | 324 | 350 | 179 |
| $[PtCl_4]^-$ | 330 | 171 | 147 | 312 | 313 | 165 |
| $ClF_4$ | 505 | 288 | 425 | 417 | 680–500 | – |

and overtone bands. Vibrational frequencies of several octahedral $XY_6$ molecules are reported in **Table 8**. The order of the stretching frequencies can be $\nu_1 > \nu_3 >> \nu_2$ or $\nu_1 < \nu_3 >> \nu_2$, depending on the compound.

The order of the bending frequencies is $\nu_4 > \nu_5 > \nu_6$ in most cases. In the same group of the periodic table, the stretching frequencies decrease as the mass of the central atom increases.

Several $XY_6$-type ions show splitting of degenerate vibrations due to lowering of symmetry in the crystalline state, an effect which has been attributed to a static Jahn–Teller effect.

Characteristic infrared frequencies and their assignment for some other important anions are summarized in **Table 9**.

In some cases where there is high point group symmetry, vibrations normally infrared inactive often appear as a weak band and doubly and triply degenerate vibrations can split into two and three compo-

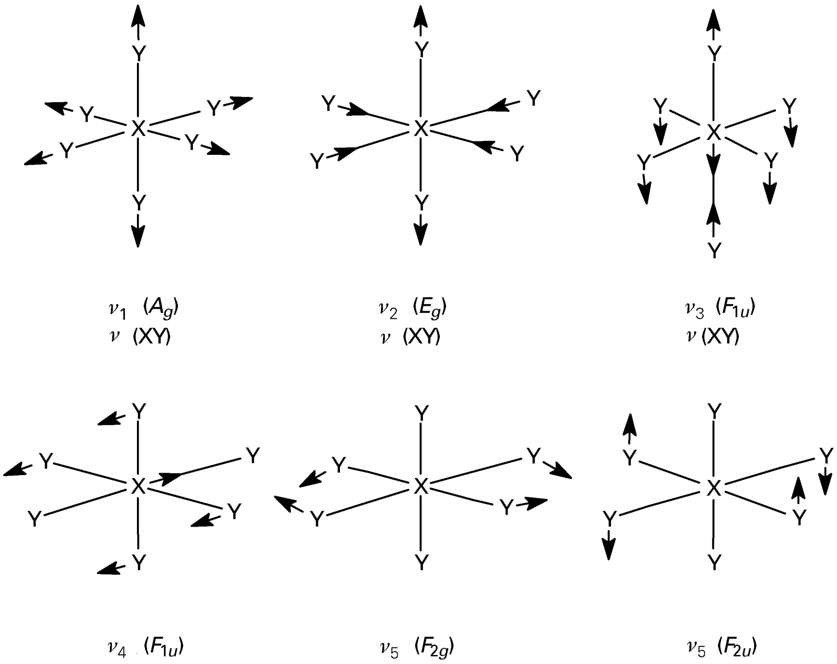



**Figure 11** Normal modes of vibration for an octahedral $XY_6$ molecule.

**Table 8** Vibrational frequencies of octahedral $XY_6$ molecules ($cm^{-1}$)

| *Compound* | $\nu_1$ | $\nu_2$ | $\nu_3$ | $\nu_4$ | $\nu_5$ | $\nu_6$ |
|---|---|---|---|---|---|---|
| $SF_6$ | 775 | 643 | 939 | 614 | 534 | (347) |
| $[BiCl_6]^{3-}$ | 259 | 215 | 172 | 130 | 115 | – |
| $[SnI_6]^{2-}$ | 311 | 229 | 303 | 166 | 158 | – |
| $[TiCl_6]^{2-}$ | 320 | 271 | 316 | 183 | 173 | – |
| $[BrF_6]^+$ | 658 | 660 | – | – | 405 | – |

**Table 9** Fundamental vibrations of selected inorganic ions ($cm^{-1}$)

| *Formula* | *Ion* | *Point group* | *Vibrations* |
|---|---|---|---|
| $AlH_4^-$ | Tetrahydroaluminate | $T_d$ | 1790 ($\nu_1$), 799 ($\nu_2$), 1740 ($\nu_3$), 764 ($\nu_4$) |
| $CN^-$ | Cyanide | | 2080–2239 ($\nu_{C\equiv N}$) |
| $SCN^-$ | Thiocyanate | $C_{\infty v}$ | 1293 ($\nu_1$), 470 ($\nu_2$), 2066 ($\nu_3$) |
| $CO_3^{2-}$ | Carbonate | $D_{3h}$ | 1087 ($\nu_1$), 874 ($\nu_2$), 1432 ($\nu_3$), 706 ($\nu_4$) |
| $PO_4^{3-}$ | Orthophosphate | $T_d$ | 935 ($\nu_1$), 420 ($\nu_2$), 1080 ($\nu_3$), 550 ($\nu_4$) |
| $MoO_4^{2-}$ | Molybdate(III) | $T_d$ | 940 ($\nu_1$), 220 ($\nu_2$), 895 ($\nu_3$), 365 ($\nu_4$) |
| $SnCl_6^{2-}$ | Hexachlorostannate | $O_d$ | 311 ($\nu_1$), 229 ($\nu_2$), 158 ($\nu_3$) |
| $IO_3^-$ | Iodate | $C_{4v}$ | 779 ($\nu_1$), 390 ($\nu_2$), 826 ($\nu_3$), 330 ($\nu_4$) |
| $ClO_4^-$ | Perchlorate | $T_d$ | 935 ($\nu_1$), 460 ($\nu_2$), 1050–1170 ($\nu_3$), 630 ($\nu_4$) |
| $PtCl_4^{2-}$ | Tetrachloroplatinate | $D_{4h}$ | 335 ($\nu_1$), 164 ($\nu_2$), 304 ($\nu_4$) |
| $NO_2^-$ | Nitrite | $C_{\infty h}$ | 1320–1365 ($\nu_1$), 807–818 ($\nu_2$), 1221–1251 ($\nu_3$) |
| $NO_3^-$ | Nitrate | $D_{3h}$ | 1018–1050 ($\nu_1$), 807–850 ($\nu_2$), 1310–1405 ($\nu_3$), 697–716 ($\nu_4$) |
| $MnO_4^-$ | Permanganate | $T_d$ | 840 ($\nu_1$), 340–350 ($\nu_2$), 900 ($\nu_3$), 387 ($\nu_4$) |

nents, respectively. These effects derive either from lowering of the point group symmetry or from factor group splitting as a result of different crystalline environments.

# Coordination compounds

## Ammine complexes

The $NH_3$ stretching frequencies in hexammine complexes (**Figure 12**) are generally lower than those in the free $NH_3$ molecule due to the fact that upon coordination the N–H bond is weakened: the stronger the M–N bond, the weaker is the N–H bond and the lower are the $NH_3$ stretching frequencies. For this reason, the $NH_3$ stretching frequencies may be used as a rough measure of the M–N bond strength. The effect of the counterion should be considered: the $NH_3$ stretching frequencies in chloride derivatives are much lower than those in the perchlorate ones and the weakening of the N–H bond is due to the formation of the N–H···Cl-type hydrogen bond in the former. The main effect of coordination and hydrogen bonding is the shift to higher frequencies of the $NH_3$ deformation and rocking modes. The $NH_3$ rocking mode is most sensitive whereas the degenerate deformation is least sensitive to these effects.

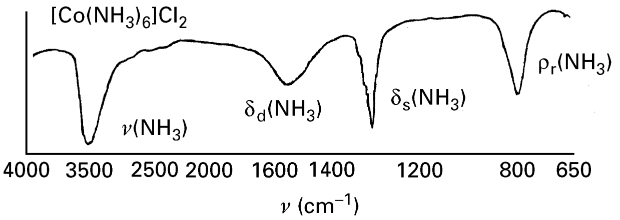



**Figure 12** Infrared spectrum of the hexammine cobalt complex $[Co(NH_3)_6]Cl_2$.

## Nitro, nitrito and nitrato complexes

The $NO_2^-$ ion is able to coordinate metals in a variety of ways (**Figure 13**), and vibrational spectroscopy is very useful in distinguishing the possible structures.

If the $NO_2$ group is bonded to a metal through its N atom (nitro form) it exhibits $\nu_a(NO_2)$ and $\nu_s(NO_2)$ in the 1470–1370 and 1340–1320 $cm^{-1}$ regions, respectively. The free $NO_2^-$ ion exhibits the same modes at 1250 and 1335 $cm^{-1}$, respectively; thus $\nu_a(NO_2)$ shifts markedly to a higher frequency, whereas $\nu_s(NO_2)$ changes very little upon coordination.

On the other hand, if the $NO_2$ group is bonded to a metal through one of its O atoms (nitrito form) it shows two $\nu(NO_2)$ well separated at 1485–1400 $cm^{-1}$, $\nu(N{=}O)$, and 1110–1050 $cm^{-1}$, $\nu(NO)$, respectively. The nitrito complexes lack the wagging modes near 620 $cm^{-1}$ which appear in all nitro derivatives and exhibit $\nu(MO)$ (M = Cr(III), Rh(III) and Ir(III)) in the 360–340 $cm^{-1}$ region.

If the $NO_2$ group is $O_2$-chelating, both antisymmetric and symmetric $NO_2$ stretching frequencies are lower and the ONO bending frequency is higher than that for unidentate N-bonded nitro complexes. In this case, $\nu_a(NO_2)$ depends on the degree of asymmetry of the coordinated nitro group: it is lowest when two N–O bonds are equivalent and becomes higher as the degree of asymmetry increases.

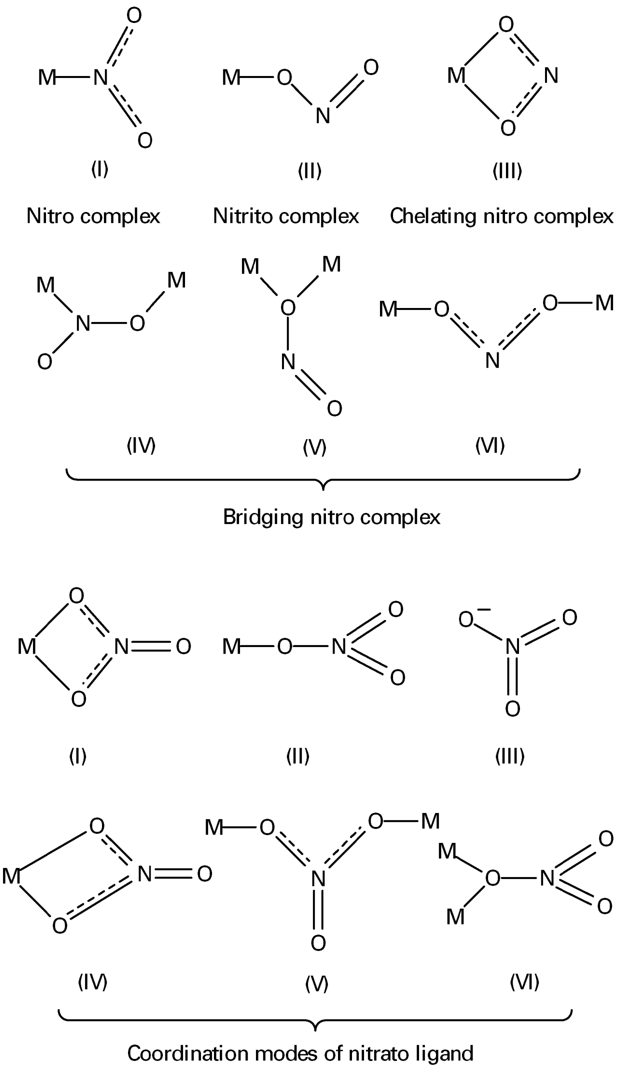



**Figure 13** Coordination modes of the $NO_2^-$ and $NO_3^-$ ions.

The nitro group is also known to form a bridge between two metal atoms (**Figure 13**) with $NO_2$ stretching bands appearing at 1492 and 1180 $cm^{-1}$. Upon isotopic substitution of the bridging oxygen, the latter is shifted by –10 $cm^{-1}$ while the former is almost unchanged. For this reason, these bands have been assigned to the $\nu$(N=O) (outside the bridge) and $\nu$(N–O) (bridge), respectively.

The $NO_3^-$ ion is able to coordinate in unidentate, symmetric and asymmetric chelating bidentate, and bridging bidentate fashion (**Figure 13**). It is very difficult to distinguish the possible structures by vibrational spectroscopy since the symmetry of the nitrate ion differs very little among them ($C_{2v}$ or $C_s$). The unidentate $NO_3$ group exhibits three NO stretching bands, as expected due to its $C_{2v}$ symmetry. For example, in $[Ni(en)_2(NO_3)_2]$ containing unidentate nitrato groups the three bands at 1420, 1305 and 1008 $cm^{-1}$ have been identified as $\nu_a(NO_2)$, $\nu_s(NO_2)$ and $\nu$(NO), respectively; whereas $[Ni(en)_2NO_3]ClO_4$ which contains a chelating bidentate nitrate, exhibits the three bands at 1476 $cm^{-1}$, 1290 $cm^{-1}$ and 1025 $cm^{-1}$, due to $\nu$(N=O), $\nu_a(NO_2)$ and $\nu_s(NO_2)$, respectively.

The separation of the two highest frequency bands is larger for bidentate than for unidentate coordination, a rule which is not applicable if the complexes are markedly different. For structural diagnosis, it is possible to use the combination band, $\nu_1+\nu_4$, of free $NO_3^-$ ion which appears in the 1800–1700 $cm^{-1}$ region. Upon coordination the $\nu_4$ ($E'$, in-plane bending) near 700 $cm^{-1}$ splits into two bands, and the magnitude of this splitting is expected to be larger for bidentate than for unidentate ligands. In some cases, this produces the separation of two ($\nu_1+\nu_4$) bands in the 1800–1700 $cm^{-1}$ region, the $NO_3^-$ ion being bidentate if the separation is 70–20 $cm^{-1}$ and unidentate if it is 25–5 $cm^{-1}$.

### Lattice water

It is possible to classify water in inorganic salts and coordination compounds such as lattice or coordinated water, but no definite borderline exists between the two forms. Vibrational spectra are often useful for providing information which allows the two forms to be distinguished.

Lattice water exhibits strong absorptions at 3550–3200 $cm^{-1}$ (antisymmetric and symmetric OH stretchings) and at 1630–1600 $cm^{-1}$ (HOH bending). In the region 600–200 $cm^{-1}$, lattice water exhibits 'libration modes' that are due to rotational oscillations of the water molecule.

### Aquo and hydroxo complexes

The structure of aquo complexes is often assigned on the basis of their vibrational spectra: $TiCl_3 \cdot 6H_2O$ has been formulated as *trans*-$[Ti(H_2O)_4Cl_2]Cl \cdot 2H_2O$ since it exhibits one TiO stretching (500 $cm^{-1}$, $E_u$) and one TiCl stretching (336 $cm^{-1}$, $A_{2u}$) mode.

In the Raman spectra of aqueous solutions of nitrate and sulfate zinc(II) and mercury(II) salts, the polarized Raman bands in the 400–360 $cm^{-1}$ region have been assigned to the metal–O stretching modes of hexacoordinated aquo complex ions.

The hydroxo group can be distinguished from the aquo group since the former lacks the HOH bending mode near 1600 $cm^{-1}$, and hydroxo complexes exhibit MOH bending modes below 1200 $cm^{-1}$ (for example, 1150 $cm^{-1}$ for the $[Sn(OH)_6]^{2-}$ ion). The OH group is able to bridge two metal atoms: in $[(bipy)Cu(OH)_2Cu(bipy)]SO_4 \cdot 5H_2O$ the bridging OH bending mode is at 955 $cm^{-1}$ and is shifted to 710 $cm^{-1}$ upon deuteration.

## Perchlorato complexes

Infrared and Raman spectroscopy techniques have been used extensively to determine the mode of coordination of the $ClO_4^-$ ligand. For example: the ionic $K[ClO_4]$ shows only two characteristic bands at 1170–1050 and 935 $cm^{-1}$, whereas $Cu(ClO_4)_2 \cdot 2H_2O$, where the perchlorate is coordinated in unidentate fashion, exhibits three strong absorptions at 1158, 1030 and 920 $cm^{-1}$. In the bidentate perchlorate complexes, the first absorption is split into three bands (1270–1245, 1130 and 948–920 $cm^{-1}$) whereas the second band appears at 1030 $cm^{-1}$.

## $\beta$-Diketones

$\beta$-Diketones form metal chelates of type I, II, III and IV (**Figure 14**). The $\nu$(MO) of these chelates, assigned by using the metal-isotope technique, provides direct information about the M–O bond strength. Vibrational spectroscopy can be employed to differentiate *cis* and *trans* octahedral complexes $M(acac)_2X_2$. For example, $Ti(acac)_2F_2$ is '*cis*' with two $\nu$(TiF) at 633 and 618 $cm^{-1}$ whereas *cis*- and *trans*-$Re(acac)_2Cl_2$ can be isolated. In the infrared spectrum the *trans* isomer exhibits $\nu$(ReO) and $\nu$(ReCl) at 464 and 309 $cm^{-1}$, respectively, while each of these bands splits into two in the *cis* isomer: 472 and 460 $cm^{-1}$ for $\nu$(ReO) and 346 and 333 $cm^{-1}$ for $\nu$(ReCl).

In some derivatives, the keto form of acac forms a chelate ring (II). This kind of coordination has been observed in $M(acacH)Cl_2$ (M=Co and Zn) which exhibits a strong $\nu$(C=O) band near 1700 $cm^{-1}$.

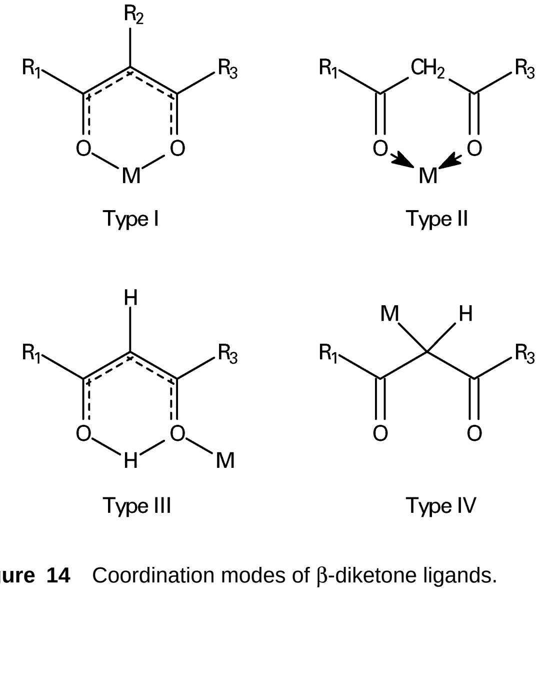



**Figure 14** Coordination modes of β-diketone ligands.

In $Mn(acacH)_2Br_2$, the acacH coordinates in the unidentate enolic form through only one O atom (III). The CO and CC stretching bands of the enol ring were found at 1627 and 1564 $cm^{-1}$, respectively.

In $[Pt(acac)_2Cl_2]^{2-}$, the metal is bonded to the γ-carbon atom of the acetylacetonato ion (IV). The infrared spectra show two $\nu$(C=O) at 1652 and 1626 $cm^{-1}$, two $\nu$(C–C) at 1350 and 1193 $cm^{-1}$, and one $\nu$(Pt–C) at 567 $cm^{-1}$.

## Carbonyl complexes

Carbonyl complexes generally exhibit strong sharp $\nu$(CO) bands between 2100 and 1800 $cm^{-1}$. Studies of $\nu$(CO) provide valuable information about the structure and bonding of carbonyl complexes because $\nu$(CO) is generally free from coupling with other modes and is not obscured by the presence of other vibrations. The $\nu$(CO) is generally at lower frequencies in the complexes than in the free CO (2155 $cm^{-1}$). However, the opposite trend is observed when CO is bonded to metal halides with the metals in a relatively higher oxidation state. Frequencies for bridging CO groups are much lower (1900–1800 $cm^{-1}$) than those of the terminal CO group (2100–2000 $cm^{-1}$), and extremely low $\nu$(CO) (~1300 $cm^{-1}$) are observed when the bridging CO group forms an adduct through its O atom.

Vibrational studies on $Fe(CO)_5$ and $M(CO)_6$ (M = Cr, Mo and W) derivatives, including their $^{13}C$ and $^{18}O$ species, have indicated that the M–C bond strength increases in the order Mo < Cr < W, an order also supported by a Raman study.

It has been found that in the $M(CO)_4$ complexes ($T_d$), $\nu$(CO) decreases and $\nu$(MC) increases on going from $Ni(CO)_4$ to $[Co(CO)_4]^-$ to $[Fe(CO)_4]^{2-}$. This result indicates that the M→CO π back-donation increases in the order Ni(0) < Co(-I) < Fe(-II). Highly reduced species such as $[W(CO)_4]^{4-}$ exhibit very low $\nu$(CO) values (1530–1460 $cm^{-1}$).

Vibrational spectroscopy has been a subject useful for elucidating structures of polynuclear carbonyls. For $CO_2(CO)_8$, which has a structure containing six terminal and two bridging CO groups, five terminal and two bridging $\nu$(CO) are expected to be infrared active: the former are in fact observed at 2075, 2064, 2047, 2035 and 2028 $cm^{-1}$, whereas the latter are at 1867 and 1859 $cm^{-1}$.

In $MX(CO)_X$ derivatives (X = Cl, Br or I), $\nu$(CO) is highest for the chloro compound and lowest for the iodo compound (the bromo compound being between the two) in accordance with their electronegativities.

### Dioxygen complexes

Dioxygen ($O_2$) adducts of metal complexes have been extensively investigated with vibrational spectroscopy because of their importance as models for oxygen carriers in biological systems and as catalytic intermediates in oxidation reactions of organic compounds.

The bond order of the O–O linkage decreases as the number of electrons in the antibonding 2pπ* orbital increases: the bond order in $[O_2^+]AsF_6$, $O_2$, $K[O_2^-]$ and $Na_2[O_2^{2-}]$ is 2.5, 2.0, 1.5 and 1.0, respectively. This decrease causes an increase in the O–O distance and a decrease in the $\nu(O_2)$, which are at 1858, 1555, 1108 and 760 $cm^{-1}$, respectively.

An interesting example of dioxygen adducts of transition-metal complexes is the *trans*-planar $[IrCl(CO)(PPh_3)_2]$ (Vaska's salt), which binds dioxygen reversibly to form a trigonal bipyramidal complex. The O–O distance of this compound (1.30 Å) is close to that of the $O_2^-$ ion (1.28 Å): its $\nu(O_2)$ (857 $cm^{-1}$) is typical of peroxo adducts.

### Halogeno complexes

Halogens (X) are the most common ligands in coordination chemistry. Terminal MX stretching bands appear in the regions of 750–500 $cm^{-1}$ for MF, 400–200 $cm^{-1}$ for MCl, 300–200 $cm^{-1}$ for MBr, and 200–100 $cm^{-1}$ for MI. The $\nu$(MX) generally is higher with a higher oxidation state of the metal. In some cases, such as the $[M(dias)_2Cl_2]^{n+}$, $\nu$(MCl) changes rather drastically on going from Ni(III) ($d^7$) to Ni(IV) ($d^6$) (240 and 421 $cm^{-1}$, respectively) while very little change is observed between Fe(III) ($d^5$) and Fe(IV) ($d^4$) (384 and 390 $cm^{-1}$, respectively). This was attributed to the presence of one electron in the antibonding $e_g^*$ orbital in the Ni(III) complex.

The SnX stretching force constants of halogenotin compounds are approximately proportional to the oxidation number of the metal divided by the coordination number of the complex. It is interesting to note that the $\nu$(SnCl) of free $SnCl_3^-$ ion [289 ($A_1$) and 252 ($E$) $cm^{-1}$] are shifted to higher frequencies upon coordination to a metal: the $\nu$(SnCl) of $[Rh_2Cl_2(SnCl_3)_4]^{2-}$ are at 339 and 323 $cm^{-1}$.

The number of $\nu$(MX) vibrations observed is often very useful for determining the stereochemistry of the complex: in the vibrational and Raman spectra of planar $M(NH_3)_2X_2$ [M = Pt(II) and Pd(II)] the *trans* isomer ($D_{2h}$) exhibits one $\nu$(MX) ($B_{3u}$), whereas the *cis* isomer ($C_{2v}$) exhibits two $\nu$(MX) ($A_1$ and $B_2$) bands in the infrared.

Some derivatives, such as $Ni(PPh_2R)_2Br_2$ (R = alkyl) exist in two isomeric forms, tetrahedral and trans-planar. Distinction between these two can be made easily since the numbers and frequencies of infrared-active $\nu$(NiBr) and $\nu$(NiP) are different for each isomer: in fact, $\nu$(NiBr) and $\nu$(NiP) are at ~330 and 260 $cm^{-1}$, respectively, for the planar form, and at ~270–230 and 200–160 $cm^{-1}$, respectively, for the tetrahedral form.

IR data allow us to differentiate between *fac* and *mer* $[MX_3(L)_3]$ isomers: for example *fac*-$[RhCl_3(Py)_3]$ gives two bands at 341 and 325 $cm^{-1}$ whereas *mer*-$[RhCl_3(Py)_3]$ shows three bands at 355, 322 and 295 $cm^{-1}$.

Bridging MX stretching frequencies [$\nu_b$(MX)] are generally lower than terminal MX stretching frequencies [$\nu_t$(MX)].

### Complexes containing metal–metal bonds

Complexes containing metal–metal bonds show $\nu$(MM) in the low-frequency region (250–100 $cm^{-1}$) because the M–M bonds are relatively weak and the masses of metals are relatively large. In the infrared, $\nu$(MM) is forbidden if the complex is perfectly centrosymmetric with respect to the M–M. If the derivative is not centrosymmetric, weak $\nu$(MM) appear.

In heteronuclear complexes the $\nu$(MM′) vibration is more strong than $\nu$(MM) because of the presence of a dipole moment along the M–M′ bond.

In Raman spectroscopy both $\nu$(MM) and $\nu$(MM) appear strong since large changes in polarizabilities are expected as a result of stretching long, covalent M–M bonds. Several $\nu$(MM) of polynuclear carbonyls have been reported, and the MM stretching force constants obtained from normal coordinate analysis have been used to discuss the nature of the M–M bond. For example, $(CO)_5Mn–W(CO_5)$ exhibits the $\nu$(M–M) in the Raman spectra at 153 $cm^{-1}$.

Compounds containing unusually short M–M bonds exhibit unusually high $\nu$(MM). The Mo–Mo distance in $Mo_2(OAc)_4$ is 2.09 Å, and the $\nu$(MM) appear at 406 $cm^{-1}$ because the Mo–Mo bond consists of one $\sigma$-bond, two π-bonds and a δ-bond (bond order 4).

## Organometallic compounds

### Metal alkanes

The methyl group bonded to a metal ($M–CH_3$) exhibits the six normal vibrations expected for tetrahedral $ZXY_3$ molecules. In addition, CMC bending and $CH_3$ torsional modes are expected for $M(CH_3)_n$ ($n \geq 2$) compounds. In $M(CH_3)_4$ derivatives (M = Si or Sn) the $CH_3$ rocking, MC stretching and CMC

bending frequencies are very sensitive to the change in metals, and the number of these skeletal infrared or Raman active modes provides direct information about the structure of the MC skeleton.

Some metal alkyls are polymerized in the solid state: $Li(CH_3)$ forms a tetramer containing $Li–CH_3–Li$ bridges in the solid state and its $CH_3$ frequencies are lower than those of nonbridging compounds. The infrared spectra of $Li[Al(CH_3)_4]$ and $Li_2[Zn(CH_3)_4]$ have been interpreted on the basis of linear polymeric chains in which the Al (or Zn) atom and the Li atom are bonded alternately through two $CH_3$ groups.

### Metal alkenes and alkynes

A vinyl group $\sigma$-bonded to a metal ($M–CH{=}CH_2$) exhibits, in addition to the MC stretching and CMC bending modes, strong C=C stretching vibrations in the Raman spectra. Their intensity and position in infrared spectra depend on the metal: for example $Hg(CH{=}CH_2)_2$ exhibits $\nu$(C=C) at 1603, and $\nu$(MC) at 541 and 513 $cm^{-1}$.

Also, the C≡C stretching frequencies of acetylenic compounds are strong in the Raman, but vary from strong to weak in the infrared, depending on the metal involved. In contrast to $\sigma$-bonded complexes, the C=C and C≡C stretching bands of π-bonded complexes show marked shifts to lower frequencies relative to those of free ligands.

Alkenes form π-complexes with transition metals. Zeise's salt, $K[Pt(C_2H_4)Cl_3]\cdot H_2O$ has been extensively investigated, but assignments have been often controversial. Two types of bonding are involved in the $Pt–C_2H_4$ bond: a $\sigma$-bond formed by the overlap of the filled 2pπ bonding orbital of the alkene with the vacant $dsp^2$ bonding orbital of the metal and a π-bond formed by the overlap of the 2pπ* antibonding orbital of the alkene with a filled dp hybrid orbital of the metal. If the $\sigma$-bond is predominant, bonding scheme I (**Figure 15**), which predicts one Pt(II)-alkene stretching mode, should be preferred.

Some workers assigned the absorption at 407 $cm^{-1}$ band of Zeise's salt to this mode. On the other hand, if bonding scheme II is operating, a five coordinate Pt atom is formed and the two bands at 491 and 403 $cm^{-1}$ can be assigned to the symmetric and antisymmetric PtC stretching modes, respectively. The real bonding is somewhere between I and II, and the latter may become more significant as the oxidation state of the metal decreases. In Zeise's salt X-ray analysis and MO calculations, together with infrared studies suggests bonding scheme I.

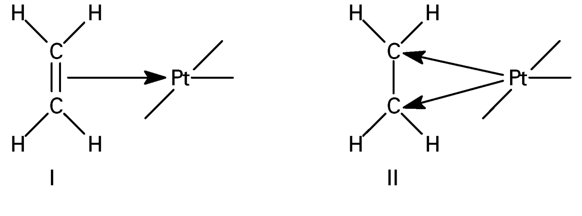



**Figure 15** Possible bonding schemes for the $Pt–C_2H_4$ bond in Zeise's salt.

Free HC≡CH exhibits a CC stretching band at 1974 $cm^{-1}$. When coordinated as in $(HC{\equiv}CH)Co_2(CO)_6$, the CC stretching band was observed at 1402 $cm^{-1}$, which is ~570 $cm^{-1}$ lower than the value for free acetylene. The spectrum of the coordinated acetylene in this complex is similar to that of free acetylene in its first excited state, at which the molecule takes on a *trans*-bent structure. Free $HC{\equiv}C(C_6H_5)$ exhibits the C≡C stretching band at 2111 $cm^{-1}$. In the case of $\sigma$-bonded complexes, this band shifts slightly to a lower frequency (2036–2017 $cm^{-1}$), while in $M[–C{\equiv}C(C_6H_5)]_2$ [M = Cu(I) and Ag(I)], it shifts to 1926 $cm^{-1}$ due to the formation of both $\sigma$-and π-type bonding.

*See also:* **Biochemical Applications of Raman Spectroscopy; Far-IR Spectroscopy, Applications; IR Spectroscopy, Theory; IR Spectrometers; IR Spectroscopy Sample Preparation Methods; Raman Spectrometers; Rayleigh Scattering and Raman Spectroscopy, Theory.**

## Further reading

Degen IA and Newman GA (1993) Raman spectra of inorganic ions. *Spectrochimica Acta Part A (Molecular Spectroscopy)* **49**: 859–887.

Ferraro JR and Basile LJ (eds) (1978) *Fourier Transform Infrared Spectroscopy*. New York: Academic Press.

Greenwood NN (ed) (1968) *Spectroscopy Properties of Inorganic and Organometallic Compounds*. London: The Chemical Society.

Herzberg G (1945) *Infrared and Raman Spectra of Polyatomic Molecules*. Princeton, New York: Van Nostrand.

Maslowsky E Jr (1977) *Vibrational Spectra of Organometallic Compounds*. New York: Wiley.

Nakamoto K (1969) Characterization of organometallic compounds by infrared spectroscopy. In: Tsutsui (ed) *Characterization of Organometallic Compounds*, Part I. New York: Wiley Interscience.

Nakamoto K (ed) (1986) *Infrared and Raman Spectra of Inorganic and Coordination Compounds*, IVth edn. Toronto: Wiley.

Nyquist RA and Kagel RO (1971) *Infrared Spectra of Inorganic Compounds*. San Diego: Academic Press.

Nyquist RA, Kagel RO, Putzig CL and Lugers MA (eds) (1996) *The Handbook of Infrared Spectra of Inorganic Compounds and Organic Salts*. San Diego: Academic Press.

# IR and Raman Spectroscopy Studies of Works of Art

*See* **Art Works Studied Using IR and Raman Spectroscopy.**

# IR Microspectroscopy

*See* **Raman and IR Microspectroscopy.**

# IR Spectral Group Frequencies of Organic Compounds

**AS Gilbert**, Beckenham, Kent, UK

Copyright © 1999 Academic Press

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
|---|
| Applications |

The vibrational spectra of organic compounds are typically complex, possessing many individual bands of variable intensity and shape. It is observed that certain bands are associated with specific chemical groups, that is, certain vibrational motions are largely concentrated within a portion of the molecule. Such bands are identifiable primarily by their occurrence within a narrow range of frequencies and their presence or absence from a spectrum allows empirical judgments to be made concerning the chemical structure. No chemical group, however, is completely isolated in vibrational or electronic terms, and therefore the exact frequency and intensity is determined by and can provide information about the adjacent chemical environment. Examination of many thousands of organic compounds has enabled numerous correlations to be established between elements of IR spectra and chemical structure and extensive tables of group frequency data are available in many publications. Although it is rarely possible to assign most of the bands in an IR spectrum, sufficient information can often be gleaned that, when allied with data from Raman spectra and other techniques, allows complete structural determination to a high degree of confidence. The concept of group frequencies underpins much of vibrational spectroscopy, though it is not necessary for quantitative analysis.

## Initial considerations and definitions

### Frequency and wavenumber

IR absorption (or Raman shift) is caused by a quantum transition between two energy levels, and so band positions ought to be defined in terms of energy units. However, atoms can be treated as oscillating under the influence of spring-like bonds obeying Hooke's law, so the terms vibration and vibrational frequency are habitually used. The vibrational frequency is considered to be equal to the frequency of absorbed radiation (or frequency difference in the

case of Raman) and is proportional to the energy difference, but it is customary to define its value not in hertz but in a frequency-dependent unit, the wavenumber, $cm^{-1}$, which is the reciprocal of the spectral wavelength as measured in centimetres.

### Differences between IR and Raman spectroscopies

For a variety of reasons, IR spectroscopy is more fruitful than Raman for applying group frequencies to structural analysis. Some of these reasons are technical, for instance access to IR equipment is much easier and many compounds are not amenable to Raman spectroscopy with visible excitation (as used in the majority of instruments) owing to fluorescence or decomposition under laser illumination.

Chemical moieties that yield good group IR frequencies tend to be polar in nature and such entities tend not to give prominent bands in Raman spectra. Conversely, some elements of structure such as C=C bonds, particularly if situated in largely symmetrical and nonpolar surroundings, absorb very weakly in the IR, but in contrast are relatively strong in the Raman.

This complementary feature of Raman spectroscopy arises because it is change in *polarizability* during vibration that is important for determining intensity, rather than dipole change as in the IR. Confidence in assigning group frequencies is therefore often considerably strengthened by utilizing Raman spectroscopy if possible.

Overall molecular symmetry is rarely of importance in studying organic compounds as the vast majority of compounds do not possess any. However, local group symmetry usually has a bearing on the relative intensities of certain bands.

### Physical state

In what follows, the listed frequencies are mostly for samples measured in liquid or solid phases. These can often be found to vary by a few wavenumbers owing to environmental conditions so that it is usually pointless to quote exact values. Care should be exercised with gas-phase frequencies as they can differ quite substantially, owing to the removal of intermolecular association effects. Hydrogen-bonded groups such as hydroxyl are particularly susceptible.

## The basis of vibrational frequency

For simple harmonic oscillation of a diatomic molecule the value of the vibrational frequency ν of the fundamental mode, in $cm^{-1}$, is given by the relationship

$$\nu = 130.3\left(f\left[\frac{1}{m_1}+\frac{1}{m_2}\right]\right)^{1/2} \qquad [1]$$

$$= 130.3(f/\mu)^{1/2} \qquad [2]$$

where $f$ is the force constant (stiffness) of the bond in newtons per metre, $m_1$ and $m_2$ are the individual atomic masses (in atomic mass units) and μ is the 'reduced' atomic mass, which is given by

$$\mu = \frac{m_1\, m_2}{m_1 + m_2} \qquad [3]$$

A lone X–H bond (where X represents C, N, O etc.) in a polyatomic compound vibrates almost as if it were a diatomic molecule. As μ in this case is always close to 1 and X–H bonds have broadly similar values of $f$, then the X–H moiety will yield a band in the same general region. For C–H, $f$ is about 490 N $m^{-1}$, which yields a frequency of 3000 $cm^{-1}$; actual values are usually 3100–2800 $cm^{-1}$. Frequencies for O–H and N–H are usually slightly higher owing to greater values of $f$.

The reduced mass of X–X′ is rather larger than that of X–H, for instance for C–O μ is about 6.86 and since $f$ here is roughly similar to that of C–H (both being single bonds) ν can be expected to be about √6.86 times less than 3000 $cm^{-1}$, i.e. about 1150 $cm^{-1}$. Double bonds are approximately twice as strong (and stiff) as single bonds, so C=O should 'vibrate' approximately √2 times more rapidly than C–O in the region around 1600 $cm^{-1}$. In fact, the 'average' carbonyl frequency is around 1700 $cm^{-1}$ and for a variety of reasons, discussed below, can be found at least 100 $cm^{-1}$ or so on either side of this value.

Likewise, triple bonds have values of $f$ roughly three times greater than single bonds, thus C≡C and C≡N groups would be expected around 2000 $cm^{-1}$; they are usually seen between 2300 and 2100 $cm^{-1}$ or so.

The simplest case of a deformation or bending mode (given the symbol δ as opposed to ν which refers to stretching modes) involves a three-atom group with the atoms undergoing a scissors-like motion. Geometric considerations mean that the expression for μ is more complicated than, but not too different from, that for the diatomic oscillator,

but the force constants are rather lower. Thus δ($XH_2$) vibrations are broadly positioned in the region around 1500 $cm^{-1}$ while δ(XYH) occurs from 1500 to 1000 $cm^{-1}$ and beyond. Deformations of XYZ (all C, N, O or heavier atoms) are below 1000 $cm^{-1}$.

The frequencies of stretching vibrations can usually be roughly predicted from knowledge of the bond order and the atomic masses. However, the values of the force constants involved in deformations are not obvious, so that their frequencies often appear to be rather arbitrary. The total number of vibrations for a nonlinear molecule is given by the relation, 3*N*-6, where *N* is the number of atoms.

These simple considerations explain the general appearance of the vibrational spectra of organic chemicals. Essentially, most bands are found in two regions. First, from about 3600 to 2800 $cm^{-1}$, wherein lie features resulting from X–H stretching modes. Second, from about 1800 to below 400 $cm^{-1}$, a region densely occupied by bands from most other vibrations. The upper part, 1800 to about 1500 $cm^{-1}$, is often referred to as the double bond region, the rest as the fingerprint region. The region between 2800 and 1800 $cm^{-1}$ is generally quite barren unless triple (or fused double) carbon–carbon or carbon–nitrogen bonds or one or two less common groups are present. Strongly hydrogen-bonded O–H and N–H can also produce features here. **Figure 1** provides an overview in chart form.

The IR spectra of inorganic compounds have a rather different aspect, consisting mostly of broad features at low frequencies. Although O–H and N–H are sometimes present, the spectrum down to about 1300 $cm^{-1}$ is usually empty except for weak overtones and combination bands.

Finally, it will be obvious from the equations above that ν(X–D), δ($XD_2$) and δ(XYD), where D is deuterium, will be considerably different in frequency from the corresponding vibrations with hydrogen.

## Vibrational interactions

No bond vibrates independently of any other in a polyatomic molecule, and where adjacent oscillators vibrate with similar or identical frequency, strong interaction takes place. This can be illustrated by considering the simple case of a symmetric linear triatomic molecule or group Y–X–Y, joined together by two equivalent bonds. When analysed in terms of

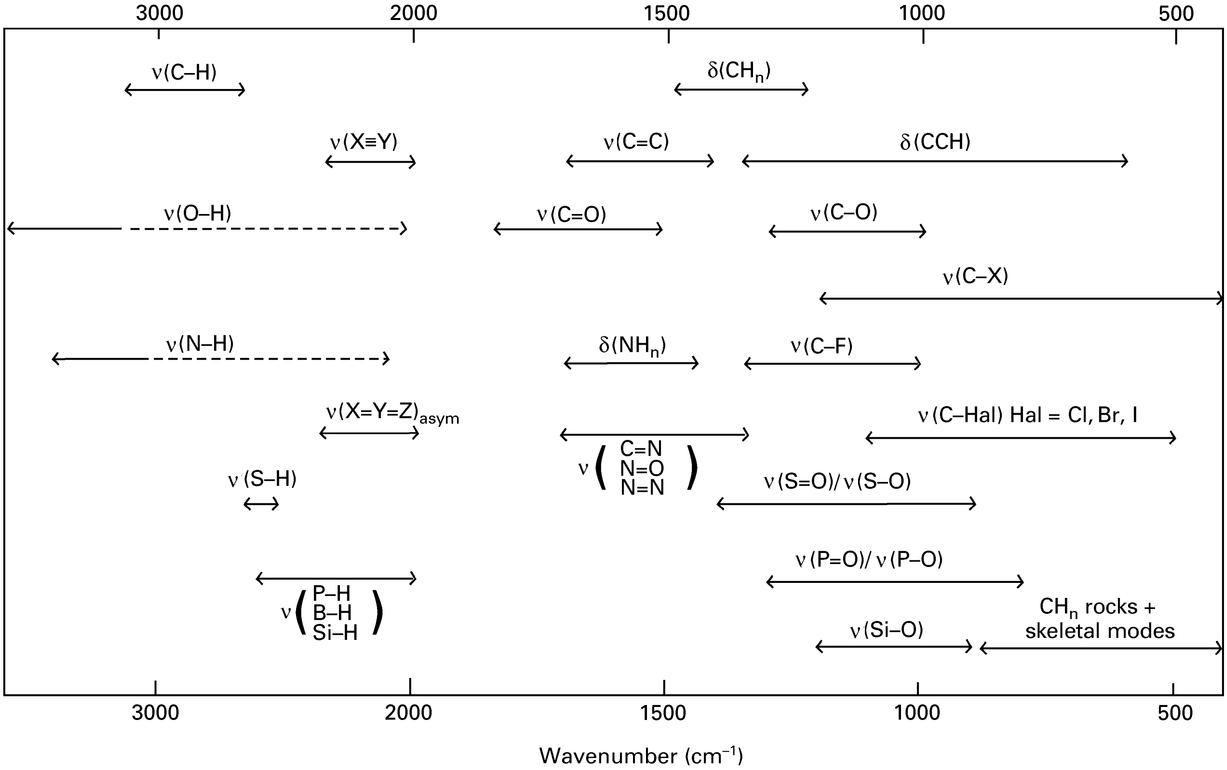



**Figure 1** Correlation chart for vibrational spectra of organic compounds. The symbols ν and δ and X, Y and Z are defined in the text.

**Table 1** Observed and calculated frequencies ($cm^{-1}$) for triatomic molecules Y=X=Y: carbon disulfide ($CS_2$) and allene ($C_3H_4$)

| *observed* | ν*(asym) observed* | ν*(sym) observed* | ν (asym)[a] / ν (sym) | ν (asym)[b] / ν (sym) | ν *(X=Y)*[c] | ν *(X=Y) observe* |
|---|---|---|---|---|---|---|
| $CS_2$ | 1522 | 650 | 2.34 | 2.52 | 1158, 1254 | 1250–1020 |
| $C_3H_4$ | 1980 | 1071 | 1.85 | 1.83 | 1596,1576 | 1680–1630 |

[a] Observed ratio.
[b] Theoretical ratio.
[c] Calculated from $\nu$(asym) and $\nu$(sym) using theoretical ratios.

classical mechanics the two individual oscillators couple mechanically together to yield two distinct modes of vibration with different frequencies. Representations of these modes are given in **Figure 2**.

The frequencies of each mode can be derived in a simplified fashion as follows. For the in-phase (or symmetric) mode, the central mass remains fixed in space so that each end atom is effectively joined to an infinite mass, thus:

$$\nu = 130.3\left(f\frac{1}{m_{\mathrm{Y}}}\right)^{1/2} \qquad [4]$$

For the out-of-phase (antisymmetric or asymmetric) mode, the central mass is shared so that each end atom effectively vibrates independently with a central atom of half mass, thus:

$$\nu = 130.3\left(f\left[\frac{1}{m_{\mathrm{Y}}}+\frac{2}{m_{\mathrm{X}}}\right]\right)^{1/2} \qquad [5]$$

The frequency of the isolated diatomic fragment Y–X is given by,

$$\nu = 130.3\left(f\left[\frac{1}{m_{\mathrm{Y}}}+\frac{1}{m_{\mathrm{X}}}\right]\right)^{1/2} \qquad [6]$$

By substituting some real values into the equations, it may be seen that the individual oscillators couple to produce two new frequencies, one considerably above and one considerably below the isolated oscillator frequency. **Table 1** gives the relevant bands of carbon disulfide (S=C=S) and allene ($H_2C{=}C{=}CH_2$) and the theoretical ratios of their frequency values as calculated from Equations [4], [5] and [6]. For the case of allene, Y is treated as $CH_2$ with a mass of 14.

The strong coupling shown above takes place because the individual vibrators are identical. The less similar they are in terms of frequency, the lower will be the degree of coupling. This is revealed by consideration of the frequencies of simple triatomic linear cyanide molecules shown in **Table 2**. For HCN and the halogenides, the 'natural' (isolated) C–H or C–halogen frequencies are very different from ν(C≡N) and there is little interaction, as demonstrated by the roughly constant values of the latter. But this is not so for DCN and TCN (T = tritium), where by extrapolation from ν(C–H) the C–D and C–T oscillators would have natural frequencies of about 2430 and 2050 $cm^{-1}$ respectively.

## The reality of group frequencies

It should be evident from the above that where two or more oscillators are strongly coupled, they will be highly sensitive to replacement or substitution of any individual units. The associated frequencies will

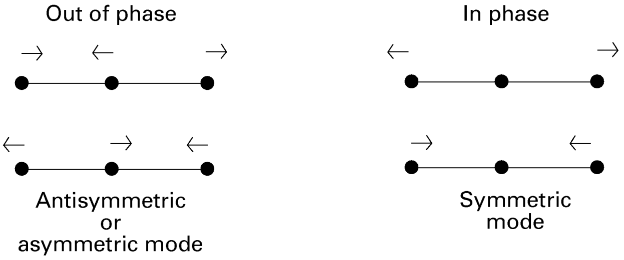



**Figure 2** Representation of the two stretching modes of vibration of a linear triatomic molecule.

**Table 2** Stretching frequencies of linear triatomic nitriles

| *X* | ν (Ξ–X) ($\chi\mu^{-1}$) | ν (X≡N) ($\chi\mu^{-1}$) |
|---|---|---|
| H | 3312 | 2089 |
| D | 2629 | 1906 |
| T | 2460 | 1724 |
| F | 1077 | 2290 |
| Cl | 729 | 2201 |
| Br | 580 | 2187 |
| I | 470 | 2158 |

(The D and T values are bracketed together in both columns.)

[a] Strongly coupled and therefore mixed modes.

therefore not be reliable indicators of particular items of chemical structure.

The observation that many simple groups are associated with bands that occur regularly in fairly narrow regions of the spectrum implies that they are largely isolated in vibrational terms. This means that the frequencies are insensitive to gross changes in molecular structure elsewhere. This is understandable for a carbonyl group or a C=C bond positioned in a singly bonded chain of carbon atoms, for instance. While individual vibrators within a group may couple strongly with each other, as long as they only couple weakly with external moieties, they retain their status as group frequencies.

Such characteristic and reliable frequencies provide a powerful means for identification of molecular subunits. By definition, they are largely insensitive to mechanical (vibrational) interactions and electronic effects from different adjacent atoms or groups. However, these perturbations can produce consistent shifts that are not so great as to destroy the characteristic nature of the group frequency, so that they provide information on the nature of the immediate chemical environment.

### Mixed-character vibrational modes

The modes of the linear triatomic molecule, described in the previous section, are purely bond stretching in character. In nonlinear molecules, such as $H_2O$, bond stretching leads to small changes in bond angles through electron redistribution, and the converse occurs for deformation modes. In these simple cases the degree of mixing of bend and stretch is not significant, so that the modes are easily distinguishable from each other.

However, where atoms are joined together in a small ring, neither bond stretching nor bond angle bending can take place without each significantly affecting the other so that many of the modes are highly mixed in character. Such vibrations of the molecular framework are often termed ‘skeletal’ modes. Some idea of their complexity can be gained from **Figure 3** which shows how the individual atoms oscillate for the various modes of the benzene framework.

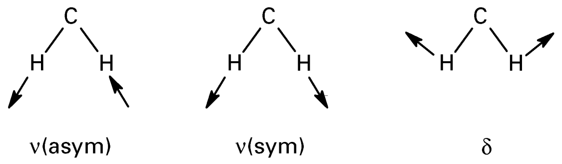
There are few good identifiable group frequencies below about 1000 cm$^{-1}$, spectra here are usually dominated by very mixed carbon skeletal and C(H)$_n$ rocking modes.

## Some examples of group frequencies

A few of the more common submolecular units are now discussed below. Space permits only a cursory treatment and the frequency values given are necessarily a compromise between figures quoted in various publications, which tend to differ slightly.

### Hydrocarbons

**The methylene group** This is shown schematically in [1] together with a representation of its associated




[1]

local modes of vibration. The two C–H bonds are coupled to give two discrete stretching modes of vibration. Simple analysis shows that typically both overall dipole fluctuation and frequency are greater for ν(asym), so that in IR spectra the higher component of such a band doublet has the greater intensity.

The exact position of each band and its relative intensity is dependent on the nature of the atoms to which the methylene group is connected. In simple hydrocarbons the bands are situated in very narrow ranges, 2925 ± 10, 2855 ± 10 and 1465 ± 15 cm$^{-1}$, respectively. Shifts occur from the presence of other groups and Fermi resonance can often bring some confusion to the region.

When methylene is part of a cyclopropane ring, distortion of bond angles leads to increasing sp$^2$ character in the C–H bonds and thereby increase in force constant. This leads to a rise in both ν(asym) and ν(sym), to over 3000 cm$^{-1}$.

Methylene also engages in external vibrations where the group, as an effectively rigid unit, oscillates with respect to other atoms to which it is connected. One such mode, called a ‘rock’, is shown below, and where more than four adjacent $CH_2$ groups form an extended hydrocarbon chain, a characteristic band appears in the IR spectrum near to 725 cm$^{-1}$. The band is essentially due to a mode where all the $CH_2$ groups are rocking in phase, see [2].

**The methyl group** There are three stretching modes, two normally identical (degenerate) asymmetric modes, and one symmetric, with a similar relation of frequency and intensity as for methylene but displaced to higher wavenumber. Like methylene, again, they are situated in narrow ranges, 2960 ± 10 and 2870 ± 10 cm$^{-1}$, respectively, in simple hydrocarbons.

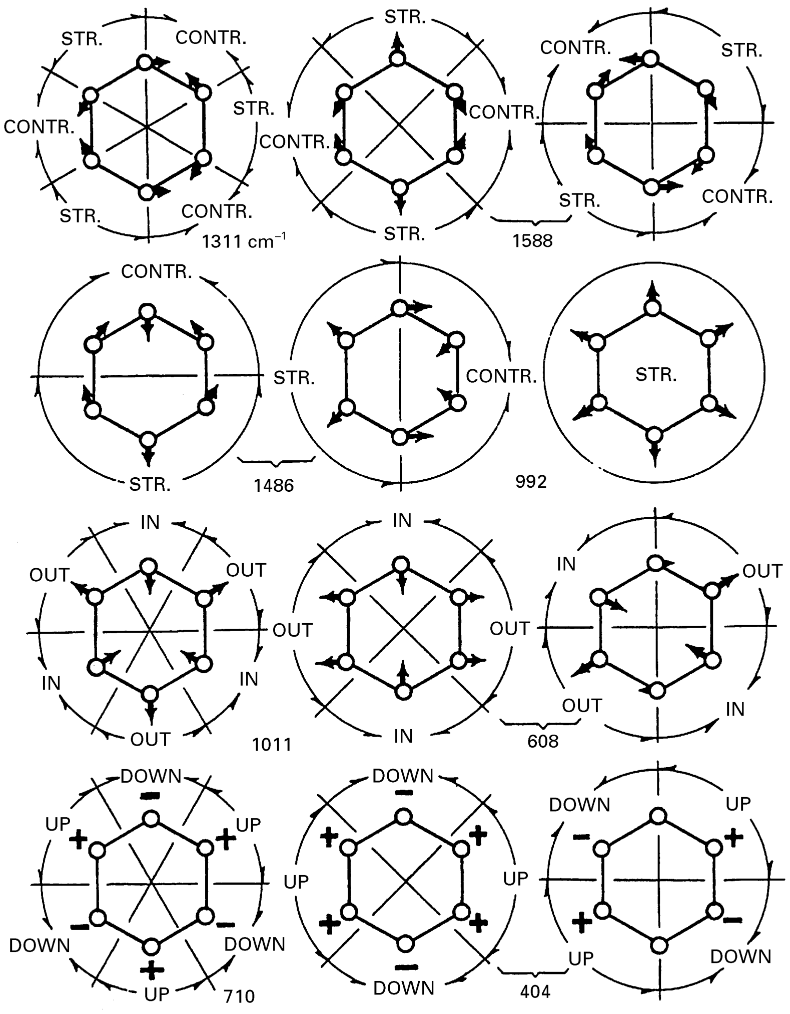
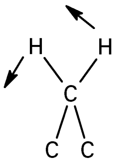



**Figure 3** Representation of the skeletal modes of the benzene molecule. Reproduced with permission from Colthup NB, Daly LH and Wiberley SE (1975) *Introduction to Infrared and Raman Spectroscopy.* New York, San Francisco and London: Academic Press.




[2]

As with the stretching modes, there are two normally or nearly degenerate asymmetric deformation modes that usually overlap with δ($CH_2$) when both groups are present, and a symmetric bending mode in the region 1380–1370 cm$^{-1}$ for simple hydrocarbons. Substituents other than carbon can shift the latter frequency quite markedly (**Table 3**), while carbonyl instead raises the intensity.

**Table 3** Symmetrical δ($XH_3$) frequencies for various X–$CH_3$

| *X* | *δ ($CH_3$)*(cm$^{-1}$) |
|---|---|
| F | ~1475 |
| O | 1455–1435 |
| N | 1430–1370 |
| $CF_3$ | ~1410 |
| $CCl_3$ | ~1385 |
| Cl | ~1355 |
| B | 1320–1290 |
| Br | ~1305 |
| S | 1325–1300 |
| P | 1310–1290 |
| Se | ~1280 |
| Si | 1280–1250 |

**Table 4** Group frequencies of triatomics

| Group[a] | *ν (asym)* (cm$^{-1}$) | *ν (asym)* (cm$^{-1}$) | *δ* (cm$^{-1}$) | *External vibrations* (cm$^{-1}$) |
|---|---|---|---|---|
| $NH_2$ | 3500–3350 | 3400–3300 | 1650–1590 | <900 |
| aph $NO_2$ | 1550–1530 | 1380–1350 | } 630–610 | <700 |
| aro $NO_2$ | 1550–1470 | 1350–1310 | | |
| $CO_2^-$ | 1610–1550 | 1420–1380 | ? | |
| aph $SO_2$ | 1330–1300 | 1150–1110 | } 610–545 | <525 |
| aro $SO_2$ | 1370–1350 | 1170–1150 | | |
| R–$SO_2(NH_2)$ | 1380–1310 | 1180–1140 | ? | |

[a] aph = aliphatic; aro = aromatic.

## Triatomic groups

The presence of these groups is signalled by a number of characteristic bands, as shown in **Table 4**. The respective bands are well separated and stably positioned in their respective regions, and therefore provide easily identifiable group frequencies. Typically, in IR spectra, the highest frequency band, from ν(asym), is the strongest or is roughly equal in intensity to ν(sym), while the deformation mode or modes are the weakest. In Raman spectra, however, both δ and ν(asym) are weak while ν(sym) is very prominent.

## The carbonyl group

The carbonyl group provides one of the best indicators for the IR spectrum yielding a strong band typically in the region between 1800 and 1650 cm$^{-1}$, though occasionally up to 100 cm$^{-1}$ more on either side. It is also an exemplar for the various electronic factors that influence group frequencies.

The C=O bond is polarized owing to the tendency of the oxygen to attract electrons from carbon (inductive effect) and this gives the group some single-bond character, ($^+$C–O$^-$) when the substituents are, for instance, methyls. If one of these is replaced by an electron-withdrawing atom such as chlorine, then the effect is largely cancelled out so that bond order is more nearly 2. The bond is in consequence stiffer and its vibrational frequency higher (see [3]).

Nitrogen, like chlorine, is also more electronegative than carbon but it affects the carbonyl frequency in the opposite sense because of the additional and

ν(C=O): ~1720 cm$^{-1}$ ~1790 cm$^{-1}$

[3]

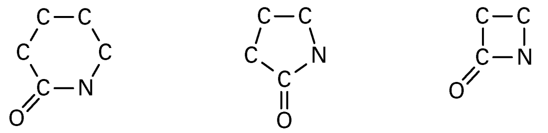
dominant effect of mesomerism. Here the C=O loses some of its double-bond character, and is thereby weakened, owing to redistribution of electrons into the C–N bond, ($^-$O–C=N$^+$).

In the case of esters, the mesomeric effect is weak so that the inductive effect of electronegative oxygen dominates to raise the frequency, and for vinyl and aryl substituents on oxygen it is higher still.

Conjugation reduces frequency because of reduction of double-bond character. Single conjugation usually generates a lowering of about 20–30 cm$^{-1}$, double conjugation a little more; for example, diaryl ketones yield ν(C=O) near 1660 cm$^{-1}$.

Noncyclic amides yield ν(C=O) between 1690 and 1620 cm$^{-1}$ and, if primary or secondary, have also one, slightly weaker, band from N–H bending at 1650–1620 cm$^{-1}$ or 1550–1510 cm$^{-1}$, respectively.

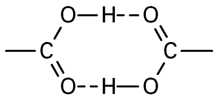
Where carbonyl is part of a strained ring, its frequency is raised, probably as a consequence of several different effects, not just that of electron redistribution. For instance, simple lactams display in the regions illustrated in [4].



1670–1640 cm$^{-1}$ 1750–1700 cm$^{-1}$ 1760–1730 cm$^{-1}$

[4]

**Table 5** lists the basic frequencies for the carbonyl group in some of its more commonly encountered chemical forms. It can be seen that the carboxylic acid frequency observed depends on the technique. This is because they are almost always (in liquid and solid states) dimerized (see [5]) so that the IR band observed is from the coupled asymmetric mode while that in the Raman is from the symmetric one.



[5]

Where two carbonyls are coupled together as in anhydrides or imides the symmetric stretching mode is, unexpectedly, at higher frequency and the relative intensities of the IR bands are strongly dependent on the relative orientation of the groups. For noncyclic anhydrides the higher mode is the most intense, but for cyclic anhydrides and imides it is the weaker.

**Table 5** The carbonyl bond as a component of various chemical groups

| *Group*[a] | ν (C=O) (cm$^{-1}$) |
|---|---|
| R–(C=O)–OR | 1750–1730 |
| Ar–(C=O)–OR | 1730–1715 |
| R–(C=O)–O–C=C | 1775–1750 |
| R–(C=O)–O–(C=O)–R | 1825–1815, 1755–1745 |
| 5-membered ring cyclic anhydride | 1870–1845, 1800–1775 |
| 5-membered ring cyclic imide | 1790–1735, 1745–1680 |
| Lactone (6 mem ring) | 1750–1735 |
| Lactone (5 mem ring) | 1795–1765 |
| Lactone (4 mem ring) | > 1800 |
| R–(C=O)–$NH_2$ | 1670–1650 |
| R–(C=O)–NH– | 1690–1620 |
| R–(C=O)–N | 1670–1630 |
| R–(C=O)–R | 1725–1705 |
| Ar–(C=O)–Ar | 1670–1660 |
| R–(C=O)–H | 1740–1730 |
| Ar–(C=O)–H | 1715–1695 |
| R–(C=O)–OH | 1725–1700 (IR) |
| R–(C=O)–OH | 1670–1640 (Raman) |
| Ar–(C=O)–OH | 1715–1680 (IR) |
| Ar–(C=O)–OH | 1690–1620 (Raman) |
| HN–(C=O)–OR | 1740–1680 |
| N–(C=O)–N | 1670–1610 |

[a] R = alkyl; Ar = aromatic.

## Alcohols, ethers and esters

The C–O stretching bands of these units, while coupled with and potentially obscured by other vibrations, can be identified by their relatively high intensity in IR spectra. For ethers, ν(C–O) occurs over the range 1140–1080 cm$^{-1}$; aliphatic alcohols can be anywhere between 1100 and 1000 cm$^{-1}$; while aromatic examples are higher in frequency, but usually weaker in intensity, at 1260–1180 cm$^{-1}$. For esters, ν(C–O) is between 1300 and 1150 cm$^{-1}$, the exact position being dependent on the nature of the substituents both on oxygen and on the carbon. A second band is also observed around about 1100 cm$^{-1}$ or so and both can probably be regarded as having C–O stretching character coupled with adjacent C–C stretching vibrations. With regard to the higher frequency, simple acetates are generally 1250–1230 cm$^{-1}$, while longer-chain homologues are 1200–1190 cm$^{-1}$; conjugation gives greater values, so that unsaturated compounds such as phthalates and benzoates yield a strong band between 1300 and 1260 cm$^{-1}$.

## Multiple bonds

**Table 6** gives information for some multiple bond groups. Bands from the CC and CN groups, where isolated, show rather variable intensity in IR spectra and, particularly those from C=C, are often difficult to identify. By contrast, these moieties feature strongly in the Raman. The asymmetric stretching mode band of a fused double-bond group, however, is always strong in the IR and occurs in the triple-bond region; for example, $CO_2$ which absorbs at just above 2300 cm$^{-1}$.

**Table 6** Group frequencies of some unconjugted multiple bonds

| *Mode* | *Frequency* (cm$^{-1}$) |
|---|---|
| ν(C=C) | 2140–2100 |
| ν(C≡N) | 2260–2240 |
| ν(C=C) | 1680–1630[a] |
| ν(C=N) | 1690–1630 |
| ν(S=O) | 1060–1040[b] |
| ν(R–N≡C) | 2145–2135[c] |
| ν(R–N=C=N–R) asym | 2150–2100 |
| ν(–N=N=N) asym | 2170–2080 |
| ν(N=N) | 1580–1550 |
| ν(P=O) | 1300–1140 |

[a] Olefin. [b] Sulfone. [c] Isonitrile.

Frequencies are rarely lowered by more than 50 cm$^{-1}$ on conjugation, but an exception is the azo group, for which it can be reduced by almost 200 cm$^{-1}$.

Usually P=O is present as part of a larger group such as a phosphorous acid or a phosphate ester and this is often the case as well for S=O as in sulfonic acids. The double bond can be delocalized somewhat into adjacent P–O or S–O single bonds (particularly on ionization in salts) and coupling of their vibrations with neighbouring C–O or C–C bonds leads to several bands between 1300 and 800 cm$^{-1}$. These often provide good diagnostic information.

Hydrogens attached directly to doubly or triply bonded carbons give stretching bands between 3100 and 3000 cm$^{-1}$, owing to the $sp^2$ nature of the C–H bond, and are therefore a good indicator of unsaturation. In olefins, out-of-plane δ(C=C–H) modes are substitution dependent, *trans* substitution yields a strong IR band at 980–965 cm$^{-1}$, while for *cis* an IR band is seen at 730–650 cm$^{-1}$.

## The benzene ring

Although the benzene ring is a relatively large entity, it nevertheless yields a number of very useful group frequencies. As in olefins, C–H stretching absorptions appear from 3100 to 3000 cm$^{-1}$. Between 1600 and 1450 cm$^{-1}$ a number of bands appear; often as many as four can be distinguished. They are produced by what are essentially stretching modes of

conjugated C=C bonds with some contributions from in-plane C–C–H bending motions.

The most prominent IR bands are substitution sensitive and occur below 1000 $cm^{-1}$. **Table 7** lists the relevant frequencies. Bands above 730 $cm^{-1}$ are from out-of-plane C–C–H bending (C–H wag) modes, while those below are from ring bending.

Raman spectra of benzene rings show a strong very characteristic band in the 1000 $cm^{-1}$ region. For benzene itself this comes from a ring breathing mode, but in substituted rings it is derived from other modes.

Heterocyclic aromatic compounds show some similarity to benzene and its homologues.

### Halogens

Halogenated organics do not give good group frequencies. Bands from carbon–halogen stretching are spread across quite wide ranges; ν(C–F) is anywhere from 1350 to 1000 $cm^{-1}$, while for C–Cl, C–Br and C–I stretching bands are 1100–500 $cm^{-1}$. Where carbon is aromatic, the frequencies tend to be higher than for aliphatic in each case. With the exception of fluorine (in the IR), the bands are of medium intensity so that it is difficult to differentiate them from the multitude of others in these regions.

### Hydroxyl and N–H stretching bands

Both O–H and N–H moieties generally form intra- and intermolecular hydrogen bonds and their bands therefore occur over fairly wide range of wavenumbers below the position of the unbonded frequencies. Usually this is no more than 200–300 $cm^{-1}$, but in extreme cases such as carboxylic acids or ionized N–H or where tautomerism occurs, the shift can be rather larger.

In this sense they are not good group frequencies, however, hydrogen bonding broadens the width and raises the intensity of the bands in the IR so they can normally be recognized though not definitively assigned to one or other. **Figure 4** shows part of theIR spectrum of *meta*-chlorobenzoic acid where ν(O–H) is represented by the diffuse underlying absorption between about 3500 and 2500 $cm^{-1}$.

**Table 7** Substitution sensitive bands of benzene below 1000 $cm^{-1}$

| *Substitution pattern* | *Band (s)* ($cm^{-1}$) |
|---|---|
| *Mono* | 770–730 and 710 – 690 |
| *Ortho* | 765–735 |
| *Meta* | 790–770 and 705–685 |
| *Para* | 860–800 |

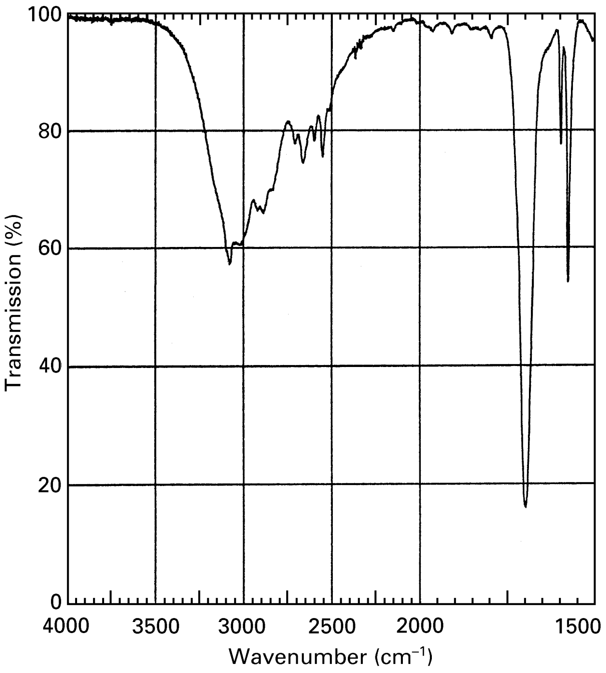



**Figure 4** Part of the IR spectrum of *meta*-chlorobenzoic acid. The bandlets between 2700 and 2500 $cm^{-1}$ are a typical part of such broad bands and are derived from interactions (Fermi-resonance) with combinations and overtones from lower-frequency vibrations.

## Group frequencies in spectral interpretation and structure determination

Some IR spectra are presented below as examples. Band assignments are given in the figure captions.

**Figure 5** shows the spectrum of isobutyl methyl ketone. The position of the carbonyl, while suiting a ketone, is also consistent with conjugated ester and carboxylic acid, but other characteristic bands from these two groups are not present.

**Figure 6** is from benzyl alcohol (Ph–$CH_2$–OH). The hydroxyl bands are dominant despite only one O–H to seven C–H vibrators; this is due mainly to the effects of hydrogen bonding. The aromatic bands are all consistent with or indicative of single substitution, including four from combinations of lower-frequency modes (not labelled) between 2000 and 1700 $cm^{-1}$. These latter are usually too weak to be of use in any other than simple molecules.

**Figure 7** is from cinnamonitrile (Ph–CH=CH–C≡N). The position of the nitrile band at 2220 $cm^{-1}$ is indicative of conjugation.

**Figure 8** shows *para*-nitroaniline. The polar substituents dominate, but some aromatic bands can still be seen.

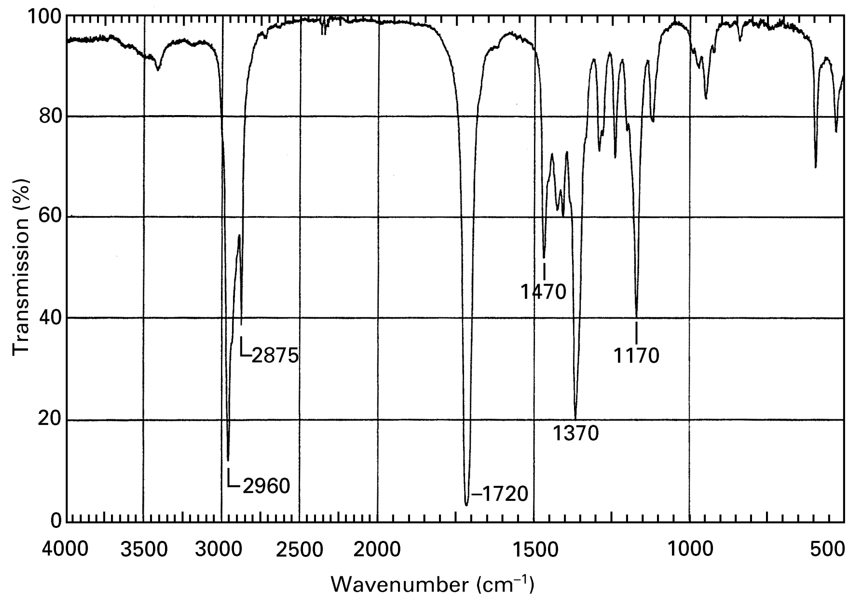



**Figure 5** Band assignments for the IR spectrum of isobutyl methyl ketone: 2960, 2875, ν($CH_3$) asym, sym; 1720, ν(C=O)$_g$; 1470, δ($CH_3$) asym, δ($CH_2$); 1370, δ($CH_3$) sym; 1170, ν(CCC) asym of $-CH_2-C(=O)-CH_3$ unit.

Finally, **Figure 9** shows acetanilide (Ph–NH–C(=O)–$CH_3$). The 3300 cm$^{-1}$ band could suit ν(O–H) but there is no obvious ν(C–O) band. The strength of the 1600 cm$^{-1}$ band is typical of anilino N–H bending rather than a purely benzene ring mode.

In these cases, structure has been assumed and band assignment is therefore definitive. Where a compound is unknown, assignments must initially be tentative and an iterative procedure has to be followed until all attributions are consistent with one another and with any other relevant information. With luck, confirmation may be possible by matching to a spectrum in a database, bearing in mind minor but not significant spectral differences due to different measurement conditions.

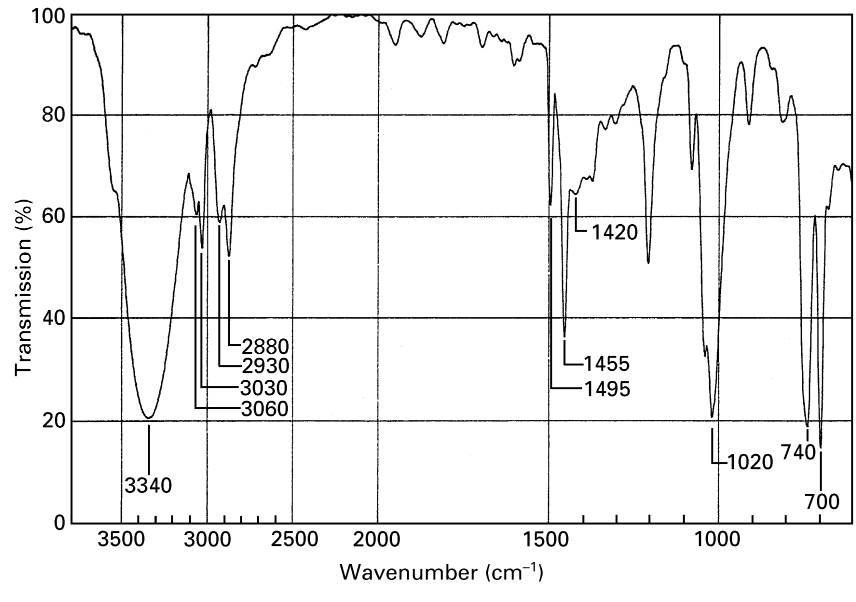



**Figure 6** Band assignments for the IR spectrum of benzyl alcohol (Ph–$CH_2$–OH): 3340, ν(O–H); 3060, 3030, ν(C–H) benzene ring; 2930, 2880, ν($CH_2$) asym, sym; 1495, ν(C=C) benzene ring; 1455, δ($CH_2$); 1420, δ(COH); 1020, ν(C–O); 740, 700, δ(CCH) monosubstituted benzene ring.

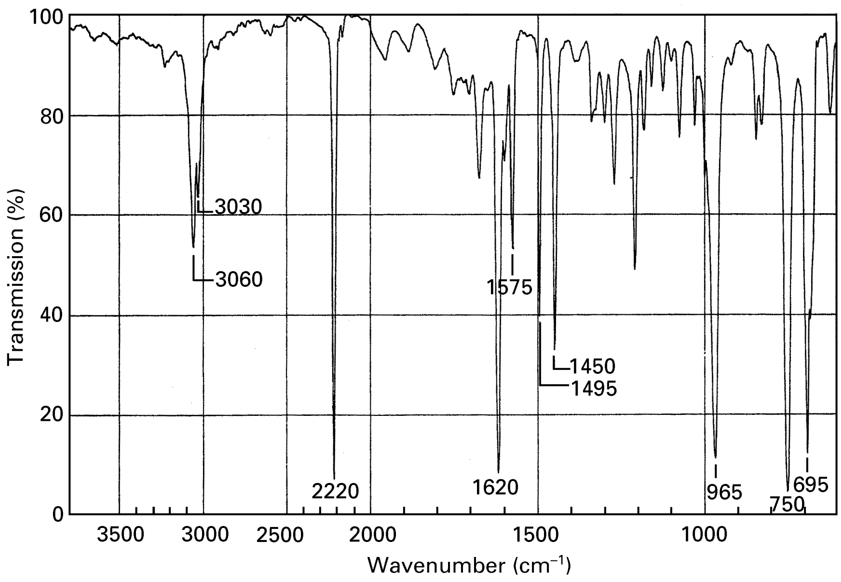
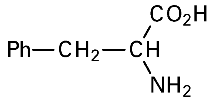



**Figure 7** Band assignments for the IR spectrum of cinnamonitrile(Ph–CH=CH–C≡N): 3060, 3030, ν(C–H) benzene ring; 2220, ν(C≡N); 1620, ν(C=C) exocyclic; 1575, 1495, 1450, ν(C=C) benzene ring; 965, ν(C=C) *trans* substituted; 750, 695, δ(CCH) monosubstituted benzene ring.

There are number of commercial databases with facilities for computer searching. The Sadtler collection of organic compounds contains nearly 100 000 IR and several thousand Raman spectra.

## Some other uses of group frequencies

Group frequencies are also useful in examining secondary effects and structure.




[6]

A simple example of physicochemical influence is provided by phenylalanine, whose IR spectrum is shown in **Figure 10**. This compound is a β-aminocarboxylic acid [**6**], but while the general appearances of the 3300–2500 cm$^{-1}$ region could be suggestive of the broad diffuse absorption yielded by –$CO_2H$ (**Figure 4**), there is no obvious carbonyl frequency in the right place nor are there any distinct ν($NH_2$) bands in the 3500–3350 cm$^{-1}$ region. Instead, carboxylate is indicated by bands at 1565 and 1410 cm$^{-1}$. The broad higher-frequency absorptions around 3000 cm$^{-1}$ are thus consistent with, and evidently from, ionized N(H) groups, in this case $NH_3^+$. The compound therefore exists as an internal salt, or zwitterion.

A more complex case is that of secondary structure in large polypeptides and proteins. These molecules can assume a wide variety of conformations that often include large sections of regular structure where the backbone is arranged either in the form of the alpha-helix or beta-sheet. Vibrational spectra of proteins typically show that the amide carbonyl group is represented by several broad and overlapping bands that are spread over a range of frequencies, 1690–1610 cm$^{-1}$, rather larger than can be explained by the presence of different amino acids. Vibrational coupling/electronic effects from the different side chains would only be expected to shift ν(C=O) by a few wavenumbers.

This multiplicity of bands arises instead because of coupling along the polypeptide backbone, dipole–dipole interactions between carbonyl groups and hydrogen bonding between C=O and N–H groups, all of which are modified by differences in the secondary structure. Thus, amidic ν(C=O) from alpha-helix is found about 1660–1645 cm$^{-1}$ and from beta-sheet between 1640 and 1625 cm$^{-1}$. Disordered structures produce bands above 1670 and below 1630 cm$^{-1}$.

## Comparison of IR and Raman spectroscopy with other techniques

Vibrational spectroscopy provides particular insights into molecular structure but rarely a completely clear view. X-ray crystallography can provide absolute determination of structure but is expensive and

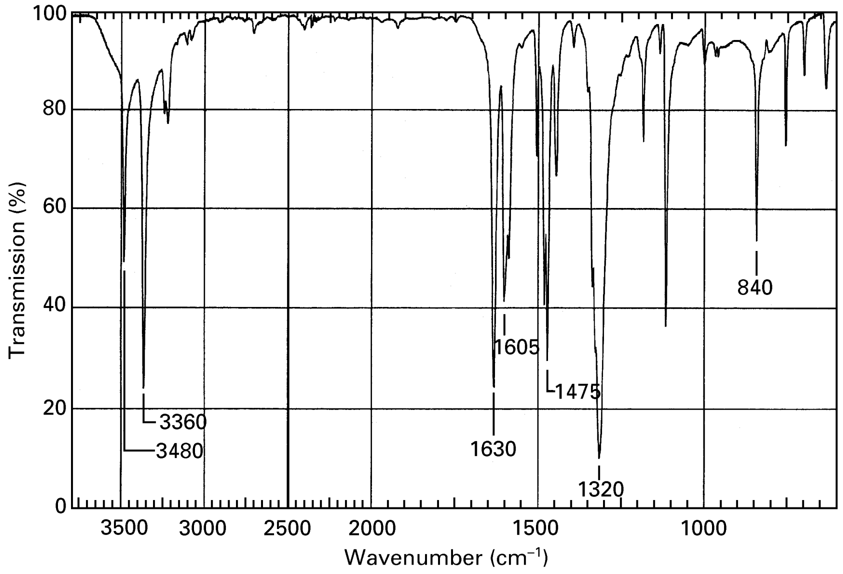



**Figure 8** Band assignments for the IR spectrum of *p*-nitroaniline: 3480, 3360, ν($NH_2$) asym, sym; 1630, δ($NH_2$); 1605, ν(C=C) benzene ring; 1475, 1320, ν($NO_2$) asym, sym; 840, δ(CCH) *para*-disubstituted benzene.

limited to small compounds that will crystallize. Mass spectrometry (MS) can give the empirical formula, but identification of groups is dependent on the somewhat arbitrary way the molecule fragments. Nuclear magnetic resonance spectroscopy (NMR), the most powerful and popular structural technique nowadays, while very good at determining many of the finer points of structure, is limited in practical terms by the types of nuclei that will give useful spectra.

Fortunately, these spectroscopic techniques complement each other and it is therefore sensible to utilize them all for complete structural determination of unknowns. A particular strength of IR spectroscopy is that it provides easily the best method for initially surveying a completely unknown material. Physical state or solvation properties are not a problem (as they can be for MS or NMR in obtaining a spectrum) and it is immediately apparent whether

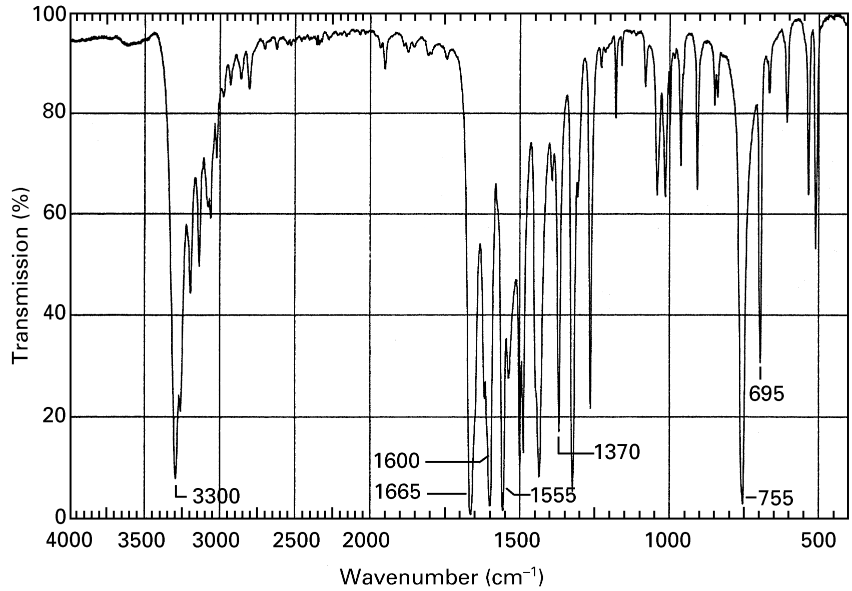



**Figure 9** Band assignments for the IR spectrum of acetanilide (Ph–NH–C(=O)–$CH_3$): 3300, ν(N–H); 1665, 1555, ν(C=O), δ(CNH) secondary amide; 1600, δ(CNH); 1370, δ($CH_3$) sym; 755, 695, δ(CCH) monosubstituted benzene ring.

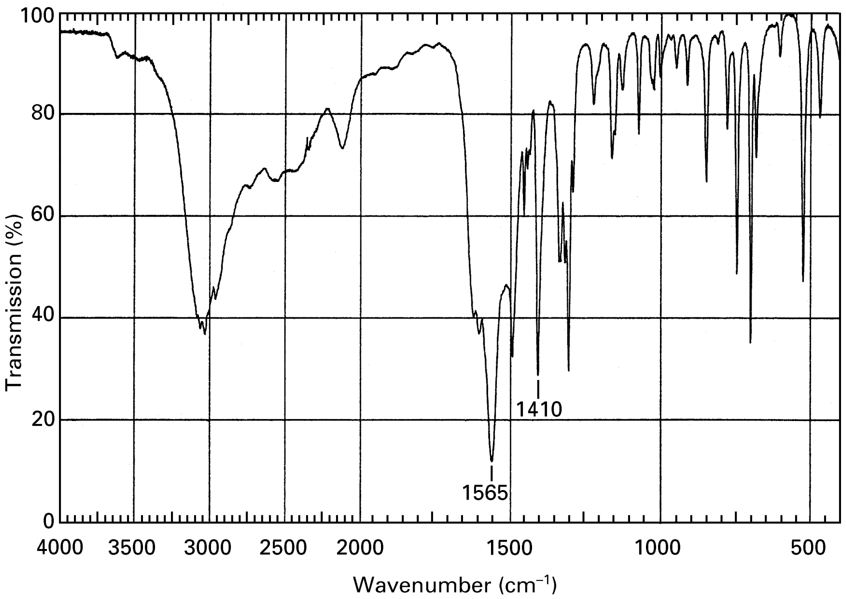



**Figure 10** IR spectrum of phenylalanine. Band assignments are given in the text.

the sample is organic or inorganic (or a mixture) and what the major groups are likely to be.

## List of symbols

$f$ = bond force constant (N m$^{-2}$); $\delta$ = bending mode frequency (cm$^{-1}$); $\mu$ = reduced mass; $\nu$ = stretching mode frequency (cm$^{-1}$).

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Chromatography-IR Methods and Instrumentation; Far IR Spectroscopy, Applications; Forensic Science, Applications of IR Spectroscopy; High Resolution IR Spectroscopy (Gas Phase), Applications; High Resolution IR Spectroscopy (Gas Phase), Instrumentation; Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy; Industrial Applications of IR and Raman Spectroscopy; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; IR Spectrometers; IR Spectroscopy Sample Preparation Methods; Medical Science Applications of IR; Near IR Spectrometers; Polymer Applications of IR and Raman Spectroscopy; Surface Studies By IR Spectroscopy.**

## Further reading

Bellamy LJ (1975) *The Infrared Spectra of Complex Molecules*. London: Chapman and Hall.

Bellamy LJ (1975) *Advances in Infrared Group Frequencies*. London: Chapman and Hall.

Colthup NB, Daly LH and Wiberley SE (1975, 1990) *Introduction to Infrared and Raman Spectroscopy*, 2nd and 3rd editions. New York, San Francisco and London: Academic Press.

Dollish FR, Fatelely WG and Bentley FF (1974) *Characteristic Raman Frequencies of Organic Compounds*. New York: Wiley.

Lin-Vien D, Colthup NB, Fateley WG and Graselli JG (1991) *The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules*. London: Academic Press.

Roeges NPG (1994) *A Guide to the Complete Interpretation of Infrared Spectra of Organic Structures*. New York: Wiley.

Silverstein RM and Webster FX (1998) *Spectrometric Identification of Organic Compounds*. New York: Wiley.

Socrates G (1994) *Infrared Characteristic Group Frequencies*. New York: Wiley.

# IR Spectrometers

**RA Spragg**, Perkin-Elmer Analytical Instruments, Beaconsfield, Bucks, UK

Copyright © 1999 Academic Press



## Introduction

From the 1980s onwards Fourier-transform infrared (FT-IR) spectrometers almost entirely replaced dispersive instruments. The most significant feature of FT spectrometers is that radiation from all wavelengths is measured simultaneously. This is much more efficient than the operation of the dispersive spectrometer, in which different wavelengths are measured successively. This difference, known as the Fellgett advantage, leads to much faster or more sensitive measurement. For example, a dispersive spectrometer typically takes at least 3 min to scan from 4000 to 400 $cm^{-1}$, while most FT instruments can acquire a spectrum in 1 s or less. The first commercial mid-IR Fourier transform spectrometer was introduced in 1969, but initially the cost of the necessary computing power confined the use of such instruments to research applications. More recently, FT-IR spectrometers have become less expensive in real terms than the dispersive instruments they have replaced. As this article will deal with instruments that are capable of recording complete spectra, it will not consider laser-based measurements.

## The IR region

The IR region covers wavelengths from about 800 nm to 1 nm (**Figure 1**). However, spectroscopists generally divide this into three, near IR (NIR) from the visible to 2.5 μm, mid IR from 2.5 to 25 μm, and far IR (FIR) beyond 25 μm. The main significance of this division is that most fundamental molecular vibrations occur in the mid IR, making this region the richest in chemical information. The near-IR region contains overtones and combinations of the fundamental vibrations, especially those involving hydrogen atoms. Some electronic absorptions are also found in this region. The far-IR region is mostly of interest for vibrations involving heavy atoms, lattice modes of solids and some rotational absorptions of small molecules.

This article will concentrate on the mid-IR region, which covers the widest range of applications.

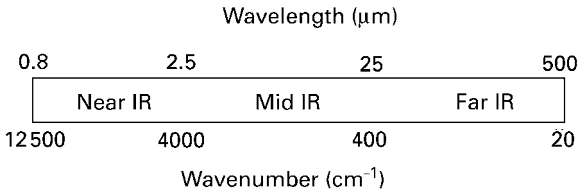



**Figure 1** The divisions of the IR region.

However, it is worth noting that interest in the near-IR region has increased steadily in recent years.

## Dispersive IR spectrometers

The principle of operation of a dispersive spectrometer is very simple. A beam from a broad-band source of radiation is dispersed by a prism or grating. The dispersed radiation falls on a slit through which a narrow range of wavelengths passes to reach a detector. Rotating the prism or grating brings successive wavelengths onto the slit, so scanning the spectrum. For the mid IR, the source is a ceramic rod or wire heated to around 1300 K. In practice, instruments use one or more gratings to cover the mid-IR region, with filters to separate different orders of diffraction. The usual detector is a thermocouple. Instruments operate in a double-beam mode (**Figure 2**) switching rapidly between two equivalent optical paths, one of which contains the sample. The transmission of the sample is measured from the ratio between the detector signals from the two beams directly (ratio recording) or by moving an attenuator into the reference beam until the detector signal matches that through the sample.

## Fourier-transform spectrometers

Most commercial Fourier-transform (FT) spectrometers are based on the Michelson interferometer (**Figure 3**). A collimated beam of radiation falls on a beam splitter that divides it into two beams with approximately equal intensities. These two beams strike mirrors that return them to the beam splitter where they recombine, creating interference, and

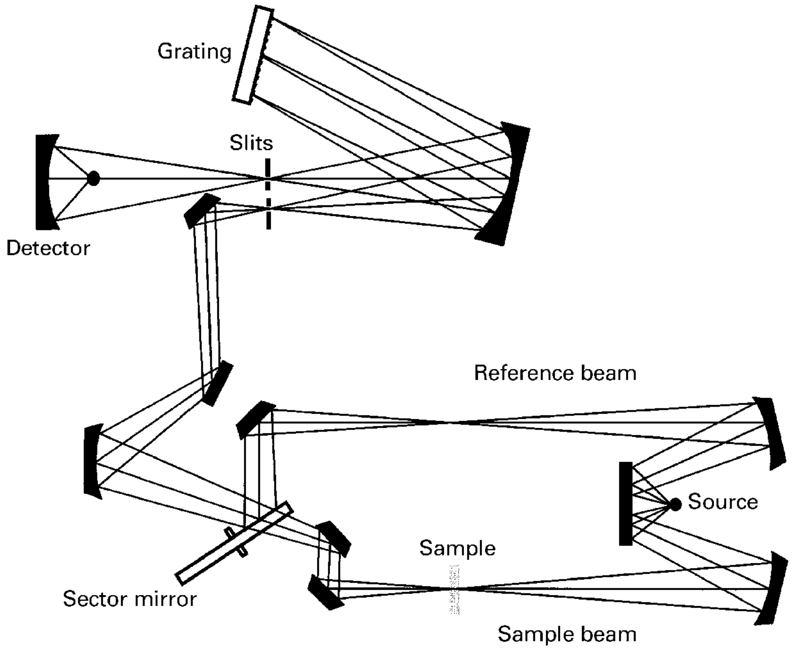



**Figure 2** The optical arrangement of a double-beam dispersive IR spectrometer.

then proceed to a detector. Moving one of the mirrors creates a varying optical path difference (OPD) between the two beams. As the path difference changes, the two beams move in and out of phase with each other, so that there is alternating constructive and destructive interference.

Monochromatic radiation produces a sinusoidal signal, the frequency of which is directly proportional to the wavenumber. The signal from a broadband source consists of a combination of overlapping sinusoidal signals, called an interferogram. A typical interferogram has a strong signal, called the centre burst, at the point where the paths of the two beams are equal, and decays as the path difference increases (**Figure 4**). The interferogram can be converted into the required spectrum of intensities at different wavenumbers by a Fourier transformation (**Figure 5**).

Unlike dispersive spectrometers, FT instruments operate in a single-beam mode. The sample is normally placed between the interferometer and the detector, in contrast to the arrangement in dispersive spectrometers where the sample is placed between the source and the monochromator. Two spectra are recorded, first of the instrument alone and then with the sample in place. The ratio of these, calculated automatically, is the transmission spectrum of the sample (**Figures 6** and **7**).

## Range and resolution in FT-IR spectrometers

The detector in a FT spectrometer sees a range of frequencies that is limited by the output of the source

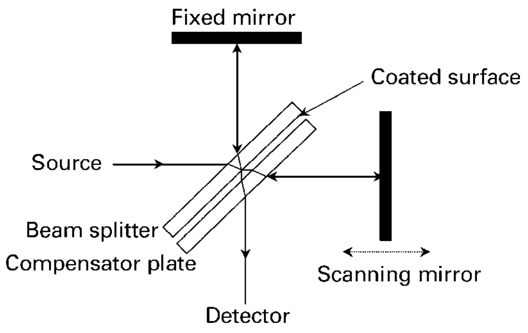



**Figure 3** The basic optics of a Michelson interferometer.

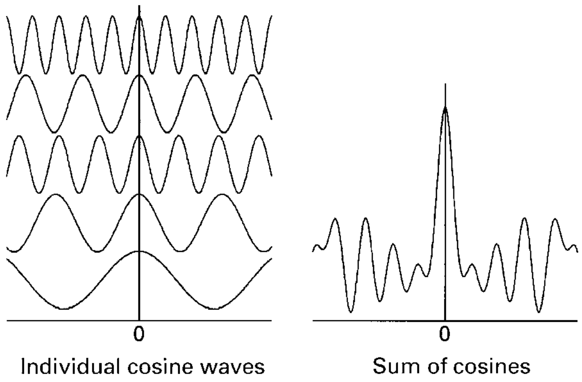



**Figure 4** The interference of overlapping cosine waves.

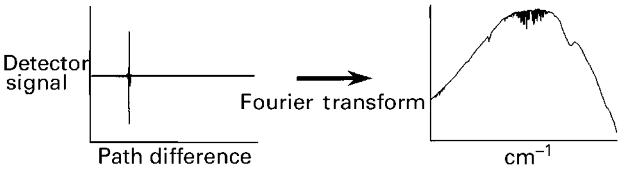



**Figure 5** The interferogram from a broad-band source and the corresponding spectrum.

and the transmission of the optics. The interferogram is digitized prior to the Fourier transformation. The range of frequencies that can be identified in the Fourier analysis depends on the spacing of the data points in the digitized signal. There must be at least two data points for each cycle of the highest frequency in the analysis. This is known as the Nyquist criterion. If there are fewer than two data points per cycle, the signal is indistinguishable from that of another, lower frequency, a phenomenon known as aliasing. Optical or electrical filtering is therefore needed to remove any frequencies that are above the Nyquist limit (**Figure 8**).

The resolution in the spectrum depends on the spacing of the frequencies provided by the Fourier transform. This is determined by the length of the interferogram. The frequencies used in a Fourier transform are those such that an exact number of cycles fits the length of the interferogram from zero to the maximum OPD, $\Delta$ (in cm). Two adjacent frequencies in the analysis, corresponding to wavenumbers $\nu_1$ and $\nu_2$ ($cm^{-1}$), have $n$ and ($n$+1) cycles in this length. The wavenumbers are given by

$$\nu_1 = n/\Delta \qquad \text{and} \qquad \nu_2 = (n+1)/\Delta \qquad [1]$$

From this it follows that the difference between the wavenumbers is

$$(\nu_2 - \nu_1) = 1/\Delta \qquad [2]$$

Thus, the spectrum generated by the Fourier transform contains wavenumbers separated by the inverse of the maximum path difference in centimetres. Instruments for routine analytical work typically have highest resolution of 1 or 0.5 $cm^{-1}$, requiring a maximum path difference of 1 or 2 cm. This resolution is constant across the spectrum. Research instruments have been built with resolutions as high as 0.0009 $cm^{-1}$, which requires a path difference in excess of 10 m.

In a practical instrument, the beam passing through the interferometer cannot be perfectly parallel. There are rays at different angles and these experience different changes in path length for a given displacement of the moving mirror. This

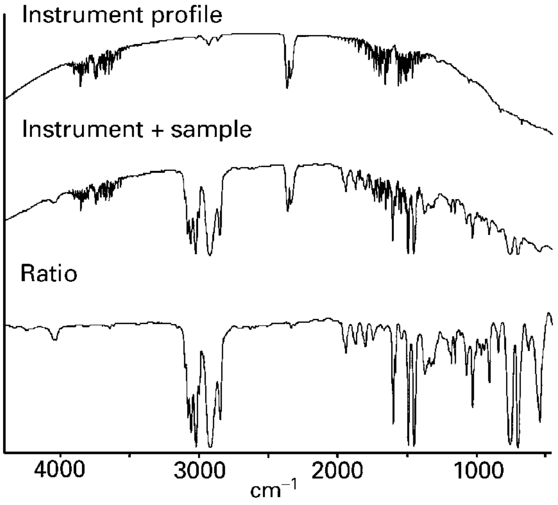



**Figure 7** Single-beam spectra and the resulting sample spectrum.

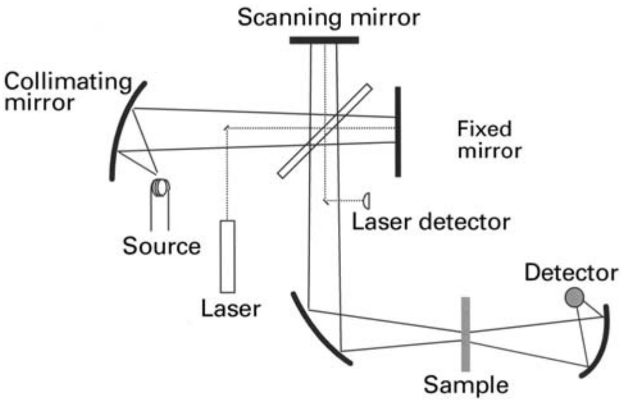



**Figure 6** The optical arrangement of a FT-IR spectrometer.

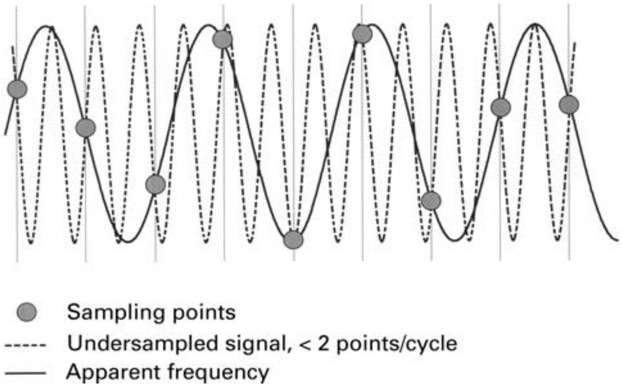



**Figure 8** A signal sampled at less than two points per cycle and a lower frequency with identical values at the sampling points.

causes line broadening by an amount that is proportional to the wavenumber. The range of angles admitted to the interferometer has to be restricted so that the required resolution is achieved. Incorporating an aperture, called a Jacquinot stop, at an appropriate point in the optical system controls this. In some ways this aperture is analogous to the slit in a dispersive spectrometer. For a given resolution, the optical throughput of a FT spectrometer, limited by a circular aperture, can be higher than that of a dispersive spectrometer with a rectangular slit. This is referred to as the Jacquinot advantage and contributes to the higher sensitivity of FT instruments. The size of this aperture depends on the resolution and the wavenumber range. Lowering the resolution by a factor of 2 allows the energy reaching the detector to be doubled. However, in most spectrometers it is not possible to increase the energy above that permitted for 4 $cm^{-1}$ resolution because of other limitations such as the range of the analogue-to-digital converter.

## Measuring the interferogram

To establish the wavelength scale of the spectrometer, the interferogram has to be measured at precisely known intervals of the path difference. Generally this is achieved by using a helium- neon laser at 0.6328 μm (15 800 $cm^{-1}$) to provide a reference signal. The laser radiation passes through the interferometer to a separate detector. This produces a sinusoidal signal with a period equal to the wavelength of the laser (Figure 9). Typically the interferogram signal is measured at each point where the laser signal passes through zero. With this spacing of the data points in the interferogram, all wavenumbers below 15 800 $cm^{-1}$ can be distinguished. The laser wavelength is known very accurately and is very stable. As a result, the wavelength scale of FT spectrometers is much more accurate than that of dispersive instruments. This superiority is known as the Connes advantage after the originator of the laser referencing technique. The absolute accuracy of the scale depends on the spectrometer resolution, while the scan-to-scan precision is of the order of 0.01 $cm^{-1}$.

The time taken for an interferometer to scan from zero to the maximum path difference is typically of the order of 1 s. For most routine measurements, the interferogram signal is averaged over several scans to reduce random noise that is largely generated by the detector. By averaging $N$ scans the noise level in the final spectrum can be reduced by a factor of $N^{1/2}$ relative to that from a single scan. For the signal averaging to work properly, the interferogram must be digitized with sufficient precision to record the noise

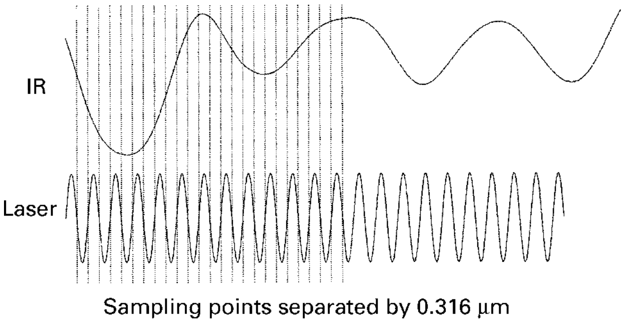



**Figure 9** A short length of an interferogram with the sampling points corresponding to zero-crossings of the laser signal.

in a single scan. The analogue-to-digital converter must be able to handle the very large centre burst and the much smaller noise component. The precision needed for this is at least 16 bits.

In an ideal interferometer, the complete spectral information could be obtained by scanning the path difference from zero to the value needed to achieve the required resolution. However, data have to be measured for a limited distance on both sides of the nominal point of zero path difference in order to carry out phase correction. For signal averaging, the data points corresponding to the same path difference for successive scans have to be combined. The laser signal measures changes in the path difference but not its absolute value. In early FT instruments this was measured by including a broad-band white light source, the interferogram from which is a very sharp spike at zero path difference. The optics were arranged so that zero path difference for the white light was offset from that for the IR radiation. The white-light signal was used to trigger data collection (Figure 10). The interferometer scanned in one direction and collected a complete interferogram on

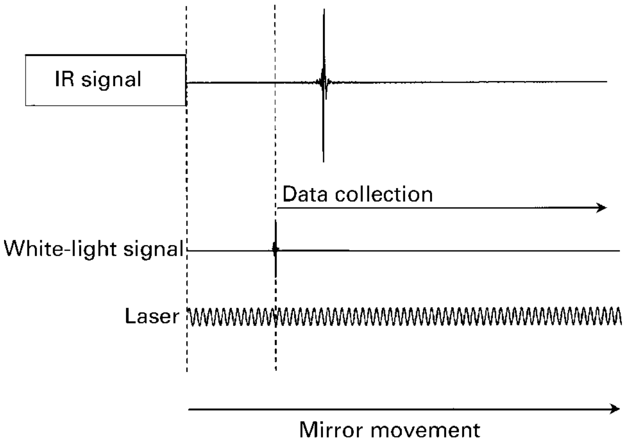



**Figure 10** The use of a white-light interferogram to trigger data collection.

one side of zero path difference only. After each scan, the interferometer would move back past the start position to begin the next scan. Such a scheme is described as unidirectional single-sided scanning (Figure 11). It was necessary to trigger each scan separately because the path difference could not be tracked through the change of direction at the end of the scan. Most systems now use two detectors with a phase shift between them to monitor the laser signal, allowing forward and reverse motion to be distinguished. This quadrature detection arrangement can track the path difference continuously and so control the co-addition of successive scans without individual triggering. It also allows the interferometer to collect data from scans in both directions. This bidirectional scanning achieves more efficient data collection as there is no fly-back period between scans (Figure 11).

## Processing the interferogram signal

The signal processing that converts the interferogram into a spectrum usually involves phase correction, zero-filling and apodization as well as the Fourier transform. An ideal interferogram would be perfectly symmetrical about zero path difference. However, in practice there is some asymmetry because of phase differences between the signals from different wavenumbers. A phase correction routine is applied before the Fourier transform. Phase correction can be avoided by scanning the full length of the interferogram on both sides of zero path difference. This allows a magnitude spectrum to be calculated without phase correction. The Fourier transform itself uses the fast Fourier transform (FFT) algorithm devised by Cooley and Tukey. This algorithm requires that the number of points to be transformed is equal to a power of 2. However, the number of data points collected by scanning to a path difference of 1 cm, for example, does not correspond to a power of 2. The interferogram is therefore extended by additional data points with the value of zero to give the required number of points. This process is called zero-filling. It is a common practice to apply further zero-filling to double or quadruple the number of points used. This improves the appearance of the final spectrum by increasing the density of the data points, although it does not improve the resolution. In some systems, interpolation in the spectrum is used to give results that are equivalent to the additional zero-filling.

Apodization is an operation that is necessary because of the finite length of the interferogram. Monochromatic radiation produces a sinusoidal interferogram, but the length of this is limited by the maximum path difference in the interferometer scan. The Fourier transform of such a signal is a line of finite width with oscillations on either side. This line has the form of a sinc function ((sin $x/x$)) (Figure 12). The oscillations (or side-lobes) that

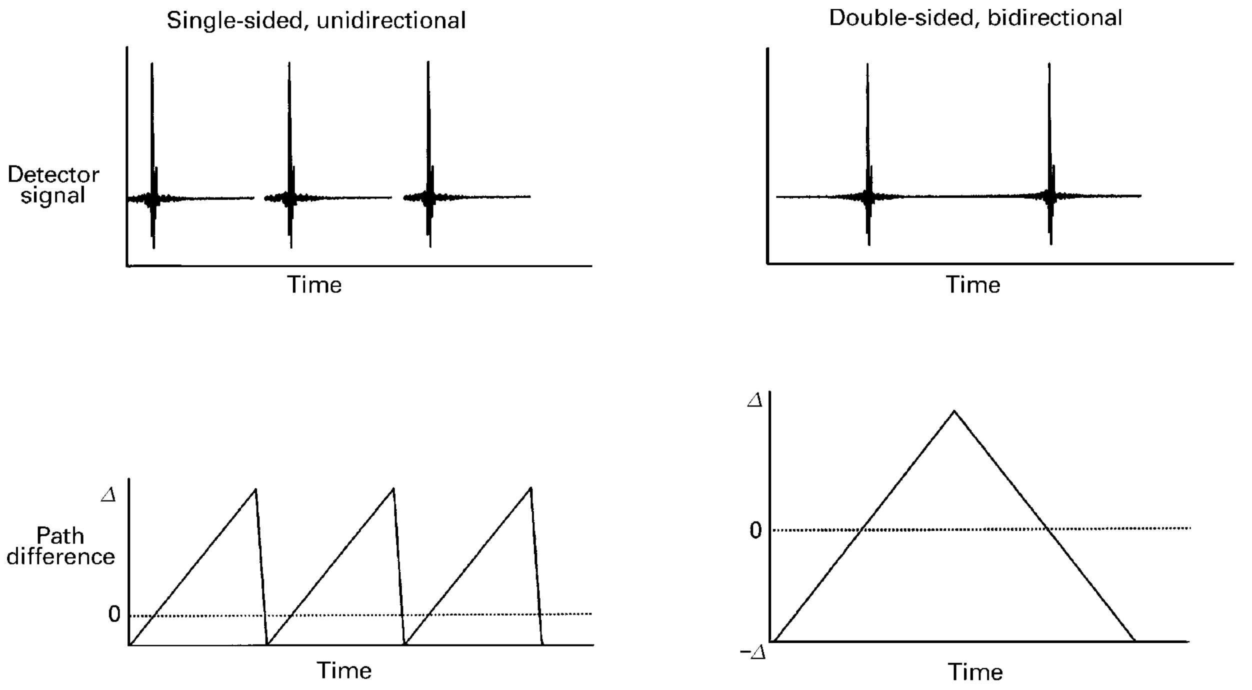



**Figure 11** Alternative scanning and data collection schemes.

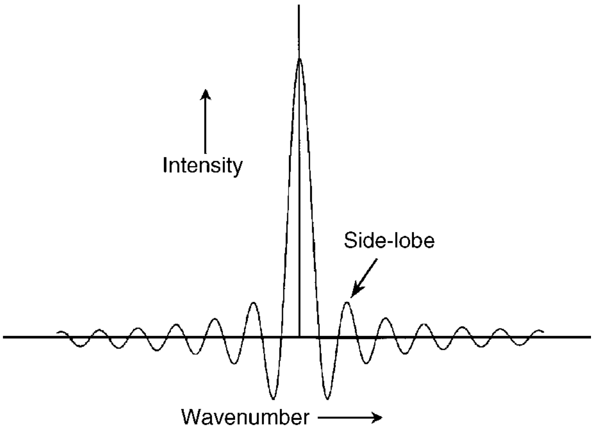



**Figure 12** The sinc function generated by Fourier transformation of a cosine wave of finite length.

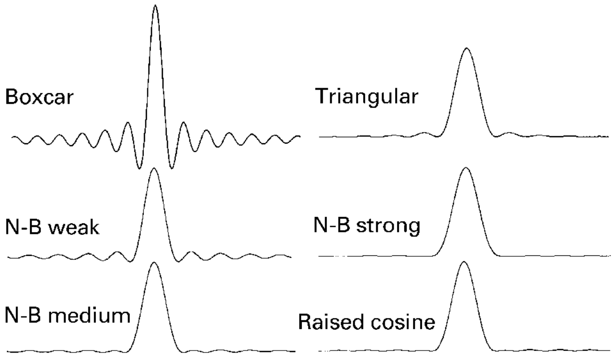



**Figure 13** Some common apodization functions and the resulting line shapes.

accompany each line are obviously a problem because they could obscure smaller features in the spectrum. They can be reduced at the expense of some broadening of the lines by multiplying the interferogram with a function that decreases with increasing path difference. This process is called apodization. Various functions are used for this, representing different compromises between line broadening and the reduction of the side-lobes. Historically, a simple triangular function was often used, but other functions can reduce the side-lobes further with less broadening. By examining a very large number of alternative functions, Norton and Beer established empirical limits for the reduction in side-lobe intensity that could be achieved for a given increase in line width. Three functions identified by Norton and Beer are in common use, weak, medium and strong (Figure 13), corresponding to increasing degrees of side-lobe reduction and line broadening. The use of no apodization in order to achieve the narrowest possible lines is called boxcar truncation.

## Scanning mechanisms

To create interference, the beams returned from the two mirrors in the interferometer must overlap exactly as they meet at the beam splitter. Any change in the alignment of the mirrors reduces the efficiency of the interference by an amount that increases with the square of the wavenumber. Small mechanical changes, for example associated with thermal expansion, can cause the alignment to change slowly. Changes that occur between measurement of the background and sample spectra result in a curved baseline (Figure 14).

The scanning mechanisms of interferometers have to be designed to minimize any change in alignment during the scan that would introduce noise or distortion into the spectrum. In systems where a plane mirror moves linearly, the problem is to avoid any tilting of the moving mirror. In some spectrometers the alignment is monitored and corrected continuously in a process called dynamic alignment. Multiple detectors measure the laser reference signal from different points in the beam. Any mirror tilt introduces a phase shift between the signals from the detectors because the path differences are not equal. This is automatically corrected in real time by using piezoelectric transducers to tilt the non-scanning mirror. The problem can also be addressed by using a cube-corner reflector instead of a plane mirror. Tilting the cube corner does not alter the direction of the reflected radiation but lateral movement would affect the alignment. There are also scanning arrangements that completely avoid alignment errors

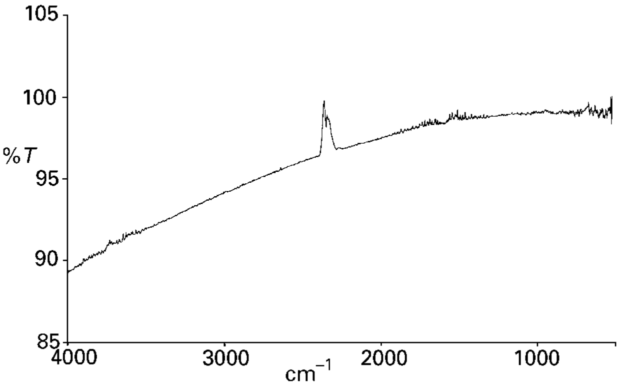



**Figure 14** A parabolic baseline caused by misalignment of the interferometer.

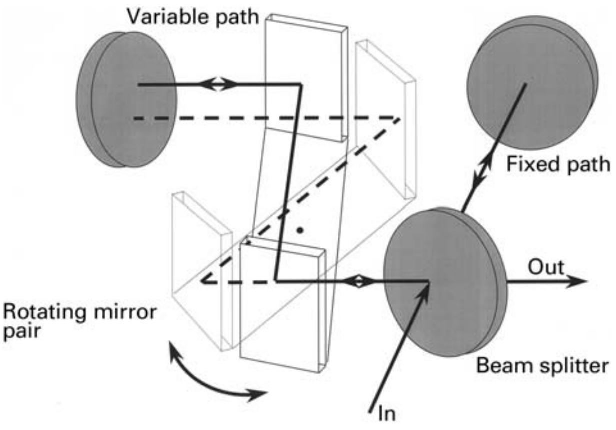



**Figure 15** A scanning mechanism with a rotating mirror pair.

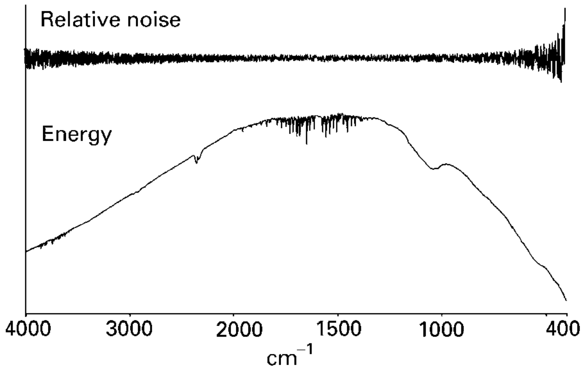



**Figure 16** The instrument profile for a mid-IR spectrometer and the resulting variation of the noise level.

during scanning. One of the simplest generates the path difference by rotating a pair of parallel mirrors in one arm of the interferometer. The mirror pair does not alter the direction of radiation passing through it, whatever its orientation, so that it cannot affect the alignment as it rotates (Figure 15).

Some commercial instruments use a scanning mechanism that does not involve moving mirrors. Changing the refractive index instead of the physical length can alter the optical path. By moving a wedge of a material such as potassium bromide (KBr) into one arm of an interferometer and displacing air, the optical path is increased. Because the mechanical tolerances are less severe than for simple moving-mirror systems, these instruments are mostly used in non-laboratory applications. One disadvantage is that the refractive index varies with wavelength so that the wavenumber scale is not inherently linear as it is in conventional instruments.

The range covered by a spectrometer depends primarily on the properties of the source and the beam splitter, and is sometimes limited by the detector. The sources used in mid-IR spectrometers typically operate at temperatures around 1300 K. Their output has a fairly similar profile to a black body and so provides significant energy across the entire mid-IR region. An ideal beam splitter would provide uniform reflectivity of 50% and no absorption across the region of interest. In practice, a single beam splitter made from KBr with a multilayer coating can be used from the near IR to about 400 cm$^{-1}$, where the KBr absorbs strongly. Most interferometer designs incorporate a compensator plate so that both beams pass through the same thickness of KBr. A FT spectrometer can cover the entire mid-IR region in a single measurement. As can be seen from Figure 16 the single-beam energy falls off towards the ends of the range. In consequence, the noise level in the final spectrum increases at each end.

## Detectors

For routine measurements the most common detector material is deuterated triglycine sulfate (DTGS). This is a thermal detector, the signal being generated in response to the change in temperature caused by absorption of the IR radiation. It is generally operated at a temperature close to ambient. The response time of these detectors is less than 1 ms, allowing them to follow changes occurring at rates up to several kilohertz. In spectrometers with DTGS detectors, the OPD is typically scanned at about 0.3 cm s$^{-1}$ because faster scanning would reduce the response of the detector. The time taken to scan the 0.25 cm needed for a spectrum at 4 cm$^{-1}$ resolution is therefore approximately 1 s. These detectors have excellent linearity over a wide range of signal levels. Lithium tantalate is an alternative room-temperature detector material with somewhat lower performance but also lower cost.

When higher speed or sensitivity is needed, liquid nitrogen-cooled MCT (mercury cadmium telluride) detectors are used. These are photon detectors in which electrons are excited directly by the absorption of radiation. Cooling to liquid-nitrogen temperature is needed to avoid excitation of electrons by thermal motion. These detectors do not cover the complete IR range because there is a minimum energy needed to excite an electron, corresponding to a long-wavelength limit. Changing the composition can vary this limit, but extending the range reduces the sensitivity. The highest sensitivity is obtained with narrow-band detectors, which have a limit of about 750 cm$^{-1}$, while wide-band detectors

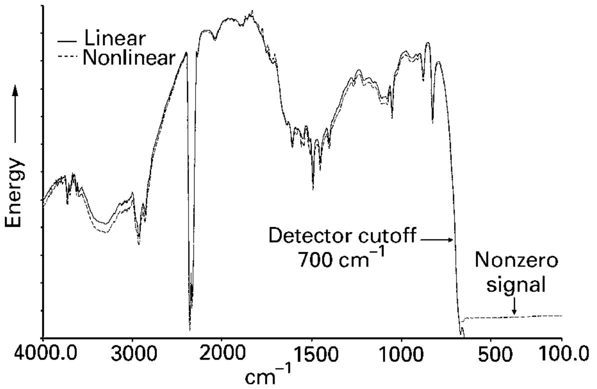



**Figure 17** The instrument profile with a cooled MCT detector showing linear and nonlinear behaviour.

**Table 1**

| *Region* | *Range* ($cm^{-1}$) | *Source* | *Beam splitter* | *Detector* |
|---|---|---|---|---|
| Mid IR | 7 000–400 | Ceramic, wire,Sic | KBr, ZnSe | DTGS, LiTa, MCT, InSb |
| Near IR | 15 000–4 000 | Tungsten–halogen | $CaF_2$, quartz | DTGS, PbS, InGaAs, InSb |
| Far IR | 400–20 | Ceramic, wire, mercury arc | Mylar, wire grid, CsI (>200 $cm^{-1}$) | DTGS, Si |

reach 450 $cm^{-1}$, close to the limit for KBr optics. Scan rates are not limited by the detector response, and indeed the detector performance improves as the scan rate is increased. This allows data to be acquired at rates up to 100 scans per second in commercial instruments. The response of MCT detectors is inherently nonlinear (see Figure 17). This restricts their use to low signal levels, typically less than 10% of the energy available. However this is not a problem in applications where the signal level is restricted, such as microscopy. Because of the nonlinearity the centre burst signal is compressed relative to the rest of the interferogram. After Fourier transformation, the primary consequence is an offset in transmission, so that totally absorbing bands do not appear at zero. The presence of nonlinearity can be detected by examining the single-beam signal in the region beyond the long-wavelength limit of the detector. Any nonlinearity causes this signal to be nonzero. Various methods, both in electronics and in software, have been employed to minimize the effects of nonlinearity. None of these is entirely successful. For this reason, room-temperature detectors are preferred for quantitative measurements.

The combinations of components found in use for the three regions are summarized in Table 1.

## Step-scan instruments

In most spectrometers the path difference is scanned continuously, but so-called step-scan instruments have become important for some research applications (see Figure 18). In these instruments the path difference is held constant at one point while the interferogram signal is measured before being moved to the next point. The IR signal has to be modulated to match the frequency response of the detector. This is achieved by applying a small-amplitude oscillation to the fixed mirror. In practice, a constant path difference can be maintained either by moving the two mirrors synchronously or by holding both stationary. Step-scanning instruments can address a much wider range of dynamic measurements than continuously scanning systems because the signal modulation frequencies are not restricted by the velocity of the moving mirror.

The principal applications have been to very rapid kinetics, photoacoustic spectroscopy and spectroscopic imaging with array detectors. The kinetic applications involve processes that are repeated for each step of the path difference. The signal can be measured at time intervals limited only by the detector response and electronics. At each path difference measurements are made at the same times after initiation of the process. The data are then reshuffled to create a series of interferograms in which each point was measured at the same time into the process. This approach has allowed FT-IR measurements to reach the nanosecond timescale. In photoacoustic spectroscopy there are advantages because the modulation frequency is constant across the spectrum, making it feasible to measure the spectra of subsurface layers. In the imaging application, the sample is focused onto an array detector. The signal from each pixel in

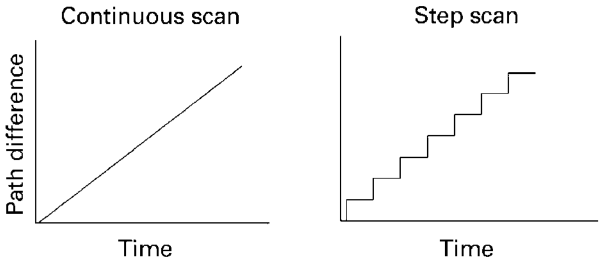



**Figure 18** The difference between continuous and step scanning.

the array can be read and averaged at each path difference. This is not feasible with continuous scanning spectrometers because the array cannot be read out sufficiently fast. With a step-scan spectrometer a 64 × 64-pixel image containing 4096 spectra can be generated in a few minutes.

## Commercial instruments

Commercial spectrometers can be grouped into those intended for routine measurements and research instruments. In general, the routine instruments are designed to simplify operation while research instruments aim to offer versatility. In virtually all cases the instrument is controlled by a personal computer that handles an increasing amount of the digital signal processing in the system.

A typical routine mid-IR spectrometer has KBr optics, best resolution of around 1 $cm^{-1}$, and a room-temperature DTGS detector. Noise levels below 0.1%*T* peak-to-peak can be achieved in a few seconds. To reduce the problem of fluctuating levels of atmospheric water vapour and carbon dioxide, much of the optical path is sealed and desiccated. The sample compartment will accommodate a variety of sampling accessories such as those for ATR (attenuated total reflection) and diffuse reflection. In most cases the output from the interferometer can be directed through an external port, principally for coupling to an IR microscope. Other external accessories provide coupling to a gas chromatograph or thermogravimetric analyser. These instruments often have versions for near-IR operation in which the source, beam splitter and detector are different.

Research spectrometers are generally larger, permitting higher optical throughput and hence better signal-to-noise performance, and higher resolution. There may be multiple external ports for the interferometer output so that several experiments can be installed simultaneously. The major components - source, beam splitter and detector - are generally exchangeable, so that a wide spectral range can be covered. A common arrangement is to have alternative sources and detectors permanently installed and selectable by switching mirrors. Different ranges can then be selected simply by changing the beam splitter. For far-IR operation, the beam splitters are generally Mylar films, which have a narrow range, or metal grids on a polymer substrate. Cooled silicon bolometers are the detectors of choice for high resolution and sensitivity. Some instruments have a vacuum version that removes the very strong water absorption in this region, allowing measurement out to 5 $cm^{-1}$. Although there have been relatively few reported applications, some spectrometers are also capable of measuring throughout the visible region.

The range of scan speeds in research instruments is greater than in routine instruments. The higher speeds provide better performance with cooled detectors and for kinetics, and the lower speeds are used with photoacoustic detectors. Digital signal processing is increasingly being used to facilitate experiments involving external modulation in which two channels of information have to be acquired simultaneously. Examples include polarization modulation in reflection- absorption of species on metal surfaces and in measuring the dynamic dichroism associated with polymer stretching. Research instruments generally have both continuous-scanning and step-scan modes as discussed earlier.

Since the introduction of the first commercial FT instrument in the late 1960s, there have been great improvements in speed and sensitivity. Routine measurements using the common sampling techniques now take only seconds, so that further improvements in performance will bring little benefit. Future advances are more likely to be seen in the development of smaller and more robust spectrometers. There are areas where improved performance is desirable, such as the sensitivity of chromatographic interfaces and the spatial resolution of IR microscopes. However, there has been little recent progress and there is little prospect of significant advances in these areas. The advent of step-scan spectrometers has led to great advances, but these have been in rather specialized applications, apart from imaging. Thus FT-IR instrumentation can be classified as mature. There is little prospect of any alternative methods of measuring IR spectra replacing FT instruments for most purposes.

## List of symbols

$n$ = number of cycles in $\Delta$; $N$ = number of scans that are averaged; $\%T$ = percent transmission; $\Delta$ = maximum optical path difference; $\nu$ = frequency.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Chromatography-IR Applications; Chromatography-IR, Methods and Instrumentation; Far IR Spectroscopy, Applications; Fourier Transformation and Sampling Theory; High Resolution IR Spectroscopy (Gas Phase), Applications; High Resolution IR Spectroscopy (Gas Phase), Instrumentation; IR Spectral Group Frequencies of Organic Compounds; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; IR**

**Spectroscopy, Theory; Near-IR Spectrometers; Surface Studies by IR Spectroscopy.**

## Further reading

*Analytical Chemistry* (1995) 67: 381A- 385A, (1998) 70: 273A- 276A. (Reviews of currently available commercial laboratory spectrometers appear every few years in this journal. These are the most recent.)

Chalmers JM and Dent G (1997) *Industrial Analysis with Vibrational Spectroscopy*, RSC Analytical Spectroscopy Monographs. Cambridge: Royal Society of Chemistry. (More practically oriented than Griffiths and de Haseth, this describes the optical layouts of both dispersive and FT commercial spectrometers.)

Griffiths PR and de Haseth JA (2000) *Fourier Transform Infrared Spectroscopy*. 2nd edn. New York: Wiley. (This is the standard book on the subject with detailed discussion of the principles of operation and design of FT-IR spectrometers.)

Mackenzie MW (ed) (1989) *Advances in Applied Fourier Transform Spectroscopy*. Chichester: Wiley.

Persky MJ (1995) *Review of Scientific Instruments* 66: 4763- 4797. (A review of spaceborne Fourier transform spectrometers for remote sensing. This is an excellent review of non-laboratory instrumentation with a very clear summary of the difficulties. Numerous designs are described.)

# IR Spectroscopic Studies of Hydrogen Bonding

*See* **Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy.**

# IR Spectroscopy in Forensic Science

*See* **Forensic Science, Applications of IR Spectroscopy.**

# IR Spectroscopy in Medicine

*See* **Medical Science Applications of IR.**

# IR Spectroscopy of Surfaces

*See* **Surface Studies By IR Spectroscopy.**

# IR Spectroscopy Sample Preparation Methods

**RA Spragg**, Perkin-Elmer Analytical Instruments, Beaconsfield, Bucks, UK

Copyright © 1999 Academic Press



## Introduction

One of the strengths of IR spectroscopy as an analytical technique is the ability to obtain spectra from a very wide range of materials. However, in many cases some form of sample preparation is required in order to obtain a good quality spectrum. For direct transmission spectra, samples have to be no more than a few tens of micrometres thick. This means that reflection measurements are often needed if materials are to be measured in their original form. Fortunately, the choice of methods for obtaining spectra by reflection has expanded significantly in recent years. When neither transmission nor reflection spectra are satisfactory, spectra can generally be obtained by photoacoustic spectroscopy, although this is a relatively expensive option.

Modern spectrometers achieve very low noise levels and have a range of data processing options that can improve the appearance of spectra. However, these are not sufficient to generate a good spectrum from a sample that has not been prepared satisfactorily. In the context of this article, a good spectrum is considered as one that is suitable for comparison with reference data. This requires that the relative intensities, shapes and positions of both strong and weak bands should be observable. With modern spectrometers the intensities of bands at 1% transmission can be measured adequately for qualitative purposes, although for quantitative analysis it is wise to avoid bands at below 10% transmission. Although noise levels can be well below 0.1%, it should be remembered that the measurement of very weak bands is more likely to be hindered by interferences such as matrix impurities or instrument artefacts than by high-frequency noise.

## Transmission measurements

### Liquids

Liquids are measured as thin films between IR-transmitting windows. The choice of window material depends on solubility, chemical compatibility and the range needed. The properties of common windows are listed in **Table 1**. For organic materials KBr is the most common choice. Zinc selenide has advantages for aqueous solutions because of the wide range, but the high refractive index leads to interference fringes that can be a problem for quantitative work, and calcium fluoride and barium fluoride therefore continue to be preferred. KBr and other alkali metal halide windows deteriorate with continued exposure to moist atmospheres and are easily scratched, but can be repolished fairly easily. Gloves should be worn to handle them to avoid fogging. The windows used for aqueous samples are more robust and are not generally repolished.

There are three main ways to measure liquids, in cells with spacers that may be fixed or demountable, or as a film directly between windows. In so-called sealed cells the spacer is a metal amalgam which provides a fixed pathlength. These are generally preferred for quantitative work. Demountable cells use a spacer made from metal foil or PTFE. The windows can be separated for cleaning or to change the pathlength. However, the pathlength generally changes slightly when the cell is reassembled, creating problems for quantitative work. When assembling demountable cells it is important to avoid dust particles on the spacer since these affect the spacing and can cause leaks. The pathlengths used for mid-IR measurements range from about 15 μm for aqueous samples to 500 μm or more for measuring components at low concentrations in weakly absorbing solvents. The pathlength of an empty cell is easily determined by measuring interference fringes. The thickness is equal to $(2 \times \text{fringe spacing})^{-1}$.

Care must be taken in filling short-pathlength cells because it is possible to generate enough internal pressure to damage the windows. For this reason it is better practice to pull the liquid into the cell by attaching a reservoir to one port and a syringe to the other, than to push liquid into the cell from a syringe. The cell should be tilted for filling to avoid trapped bubbles that cause stray light. A demountable cell is shown in **Figure 1**.

Neat liquids can be measured by placing one drop on a window and putting a second window on top of this. The pathlength can be varied by applying pressure. This is very convenient for qualitative work, but cannot be used for very volatile materials.

**Table 1** Properties of IR-transmitting materials

| *Material* | *Range* (cm$^{-1}$)$^a$ | *Refractive index* | *Properties* |
|---|---|---|---|
| NaCl | >600 | 1.52 | Soluble in water and some alcohols, easily polished |
| KBr | >400 | 1.54 | As NaCl, hygroscopic |
| CsI | >200 | 1.74 | Soluble in water and some alcohols, very hygroscopic |
| $CaF_2$ | >1100 | 1.40 | Insoluble, attacked by $NH_4^+$ salts, robust |
| $BaF_2$ | >750 | 1.45 | As $CaF_2$, but sensitive to shock |
| ZnSe | >500 | 2.43 | Insoluble but attacked by acids and strong bases |
| ZnS | >750 | 2.25 | Insoluble, attacked by acids and strong oxidizing agents, good thermal shock resistance |
| Diamond | >2200, <1900 | 2.38 | Highest chemical and mechanical resistance |
| KRS–5 | >250 | 2.38 | Slightly soluble, attacked by bases, very soft, toxic |
| Sapphire | >1600 | 1.62 | Very hard, slightly soluble in acids and alkalis |
| Ge | >600 | 4.01 | Hard and inert, brittle |
| Si | >650, <450 | 3.42 | Very hard and inert |
| AMTIR | >725 | 2.5 | Insoluble, attacked by bases, brittle |
| Polyethylene | <625 | 1.5 | Insoluble, swells in some organic solvents |
| Fused silica | >2500 | 1.46 | Insoluble |
| AgBr | >300 | 2.30 | Insoluble, attacked by acids and $NH_4^+$ salts, soft |

$^a$ This is for windows. The longer pathlength in ATR elements generally reduces the range, e.g. to 650 cm$^{-1}$ for ZnSe.

For non-volatile liquids porous polymer films are a convenient alternative to conventional windows. These '3M cards' (**Figure 2**) are made of polyethylene or PTFE. The polyethylene films absorb in rather narrow regions and, except for the region near 3000 cm$^{-1}$, their absorptions can be removed by subtraction. The PTFE cards are useful for obtaining spectra in the CH stretching region around 3000 cm$^{-1}$.

## Solids

The two most common forms of sample preparation for solids both involve grinding the material to a fine powder and dispersing it in a matrix. Scattering and reflection effects are reduced when the particle size is less than the wavelength of incident radiation, which in this case is a few micrometres. Surrounding the particles by a matrix with a similar refractive index further reduces the effects. Most samples can be ground either mechanically with a small vibrating mill or by hand with a pestle and mortar, the choice being a matter of personal preference. For manual grinding the amount of sample should be no more than around 20 mg. The sample should be ground until it is adhering to the sides of the mortar and has a glossy appearance. This process may take several minutes.

The ground material can be dispersed in a liquid to form a mull. The most commonly used liquid is

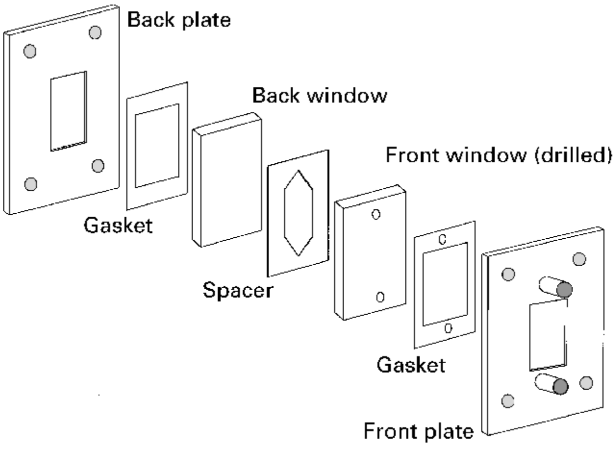



**Figure 1** Construction of a demountable cell for liquids.

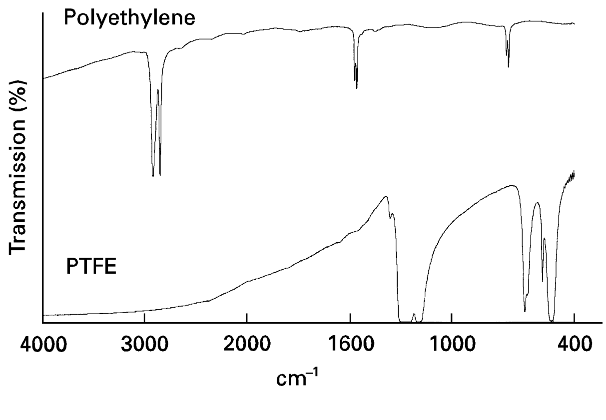



**Figure 2** Spectra of 3M cards.

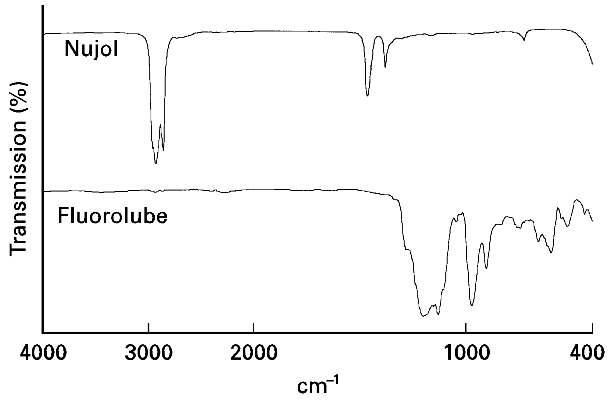



**Figure 3** Spectra of mulling agents.

mineral oil (Nujol), while a fluorinated hydrocarbon (Fluorolube) can be used to observe the 3000 cm$^{-1}$ region (**Figure 3**). No more than one or two drops are needed to produce a paste that can be spread between windows. The consistency of this paste is a compromise between the ease of spreading and the interference from the absorptions of the mulling agent.

Potassium bromide (KBr) is probably the most widely used matrix material, even though preparing a good KBr disc is time consuming and requires practice. Between 1 and 3 mg of ground material is mixed thoroughly with about 350 mg of ground KBr. The mixture is transferred to a die that has a barrel diameter of 13 mm. This is placed in a suitable press, evacuated and then pressed at around 12 000 psi for 1–2 min. Recrystallization of the KBr results in a clear glassy disc about 1 mm thick. Discs can be made with manual presses and without evacuation, but the best results are obtained with the more elaborate procedure.

Provided that the sample, the KBr and the disc are weighed, this technique can be used for quantitative work. It has the advantage that there are no matrix absorptions above 400 cm$^{-1}$, although adsorbed water in the KBr has to be avoided. Potential limitations are that the grinding and pressing can modify the crystalline form of the sample and that ion exchange can occur. To avoid ion exchange it is common to use potassium chloride for amine hydrochlorides. Caesium iodide is used for the region to 200 cm$^{-1}$ while for far-IR measurements polyethylene is a suitable matrix. Grinding is the main issue in disc preparation (**Figure 4**). Inadequate grinding results in sloping baselines caused by light scattering and in asymmetric bands. The asymmetry, known as the Christiansen effect (**Figure 5**) results from refractive index differences.

Very few solid samples arrive in a suitable form for direct transmission measurement, the principal

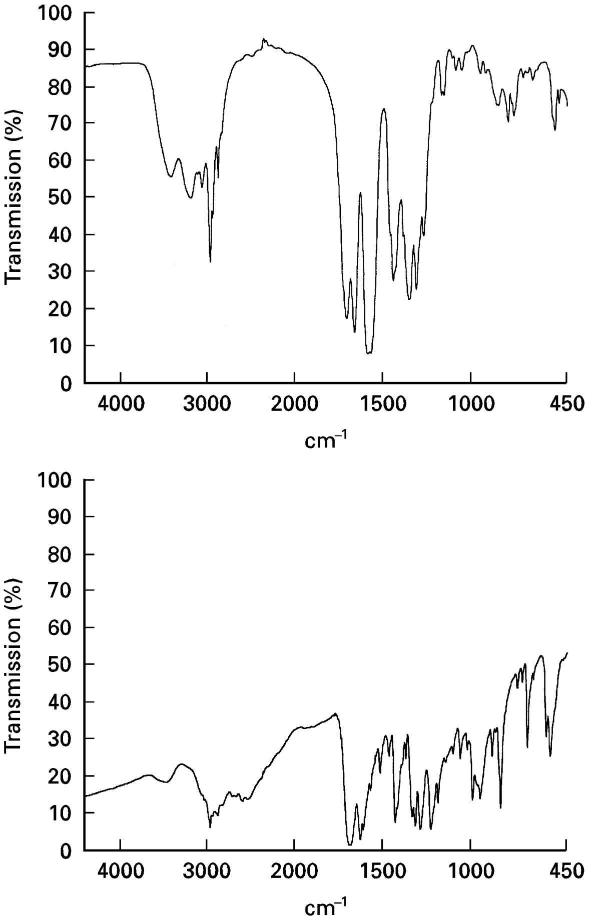



**Figure 4** Spectra from poorly (bottom) and well (top) ground discs.

exception being polymer films. Unfortunately, interference fringes often distort the spectra. There are several techniques for eliminating interference fringes by reducing the surface reflections. The simplest are to abrade the surface or to sandwich the film between windows. Another possibility is to smear one surface with mineral oil, provided that this does not obscure

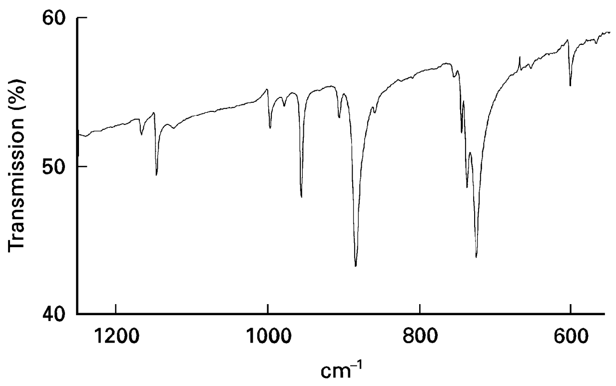



**Figure 5** Christiansen effect.

important regions of the spectrum. Polymer samples that are too thick for direct measurements can often be pressed to a suitable thickness. Heated platens for making films of controlled thickness are available from accessory suppliers.

Simple compression cells with diamond windows have appeared recently. Although they generate lower pressures than traditional diamond anvil cells, they can provide excellent spectra from a range of samples that are too thick for direct measurements.

Soluble materials can be prepared as cast films by evaporating them from solution on a suitable substrate such as a KBr window. An infrared lamp is a convenient means of warming the film to remove residual solvent. This approach is especially useful for those polymers that can be cast on to a glass block from which they can be removed for measurement without a support. A wide range of organic compounds can be prepared on the porous polymer substrates known as 3M cards. The best spectra are obtained when the compound forms an amorphous deposit. The spectra of crystalline samples may be distorted by scattering and reflection effects.

### Gases

Measurements on gases require much longer pathlengths than those for condensed phases; 10 cm is usual for compounds at relatively high concentrations. For trace concentrations multiple reflection cells provide pathlengths of several metres in compact designs. Detection limits are well below 1 ppm for many compounds.

## Internal reflection methods

In recent years, internal reflection methods have developed into perhaps the most versatile approach to obtaining IR spectra. Total internal reflection can occur at the interface between media with different refractive indices (**Figure 6**). If the angle of incidence exceeds a critical value, radiation cannot pass into the medium with the lower refractive index; instead, it is totally reflected. Although there is total reflection, an electric field, sometimes referred to as an evanescent wave (**Figure 7**), extends a short distance into the second medium. This distance is of the same order of magnitude as the wavelength. In an internal reflection measurement the radiation enters a high refractive index material and emerges after undergoing internal reflection. The high refractive index element is usually called the crystal, even though the material may be a glass. An absorbing sample placed on the surface of the crystal interacts with the electric field to generate a spectrum. This is called an attenuated total reflection or ATR spectrum. The band positions and shapes are slightly different from those in transmission spectra, but not enough to affect qualitative identification.

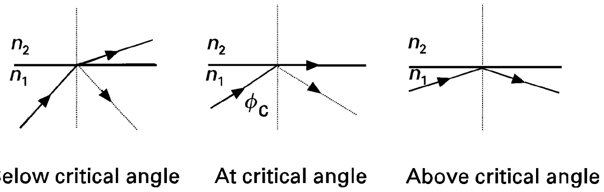



**Figure 6** Total internal reflection at interface where $n_1 > n_2$.

The penetration depth, $d_p$, of the light into the lower refractive index material is defined as the depth at which the field amplitude has fallen to 1/e of its value at the surface. For two materials with refractive indices $n_1$ and $n_2$ it is given by

$$d_p = \frac{\lambda_1}{2\pi(\sin^2\alpha - n_{21}^2)^{1/2}}$$

where $\lambda_1$ is the wavelength in the higher refractive index material and is related to the incident wavelength $\lambda$ by the expression $\lambda_1 = (\lambda/n_1)$, $\alpha$ is the angle of incidence and $n_{21} = n_2/n_1$.

The intensity of an ATR spectrum depends on the distance that the electric field extends into the sample and on the number of reflections. The most commonly used materials are zinc selenide and diamond, both with a refractive index of 2.3, and germanium, with a refractive index of 4.0. The higher refractive index reduces the distance that the electric field penetrates into the sample, giving a weaker spectrum. The depth of penetration is proportional to the wavelength, so that when ATR spectra are compared with transmission the intensities at longer

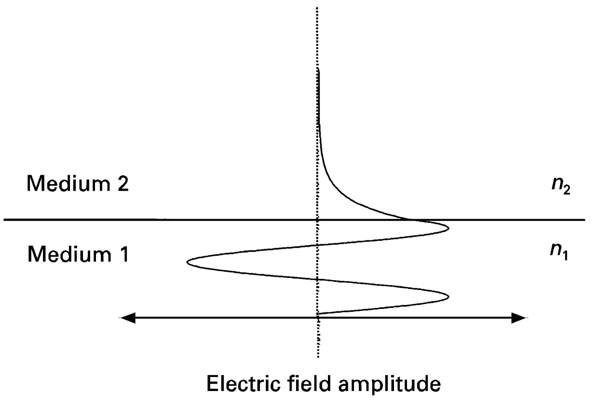



**Figure 7** Evanescent wave at an interface where $n_1 > n_2$.

wavelengths appear enhanced. For an ATR crystal with a refractive index of 2.2 and a 45° angle of incidence, the depth of penetration in a typical organic sample is about 4μm at 1000 $cm^{-1}$.

Because ATR measurements are obtained from a thin surface layer, the spectra may not be representative of the bulk material. For example, spectra of polymer laminates such as packaging materials generally show only the surface layer. However, the spectra of emulsions and suspensions may be unrepresentative of the bulk material because of segregation at the surface.

There are two basic designs of ATR accessory in common use. These have replaced traditional designs in which a thin sample was clamped against the vertical face of the crystal. In horizontal ATR (HATR) units the crystal is a parallel-sided plate, typically about 5 cm by 1 cm, with the upper surface exposed (**Figure 8**). Radiation traverses the length of the crystal, undergoing internal reflection at the upper and lower surfaces several times. The number of reflections at each surface is usually between five and ten, depending on the length and thickness of the crystal and the angle of incidence. Liquids are simply poured on to the crystal, which is recessed within a metal plate to retain the sample. Pastes and other semisolid samples are readily measured by spreading them on the crystal. Horizontal ATR units are often used for quantitative work in preference to transmission cells because they are easier to clean and maintain.

When measuring solids by ATR, there is the problem of ensuring good optical contact between the sample and the crystal. The accessories have devices that clamp the sample to the crystal surface and apply pressure. This works well with elastomers and other deformable materials, and also with fine powders, but many solids give very weak spectra because the contact is confined to small areas. The effects of poor contact are greatest at shorter wavelengths where the depth of penetration is lowest.

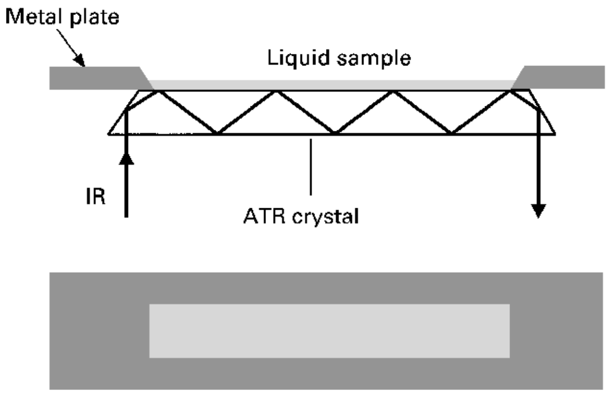



**Figure 8** Schematic of an HATR top plate for liquids.

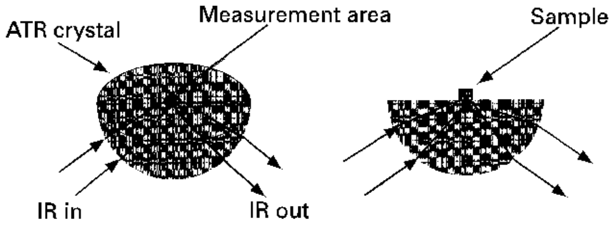



**Figure 9** Schematic of a single-reflection ATR accessory.

The problem of contact has been overcome to a great extent by the introduction of accessories with very small crystals, typically about 2 mm across. The crystal is often diamond, which offers durability and chemical inertness. Germanium and silicon are lower cost alternatives. The accessories generally provide only a single reflection (**Figure 9**), but this is generally sufficient, given the very low noise levels of modern instruments. Because of the small area, much higher pressure can be generated with limited force. A much smaller area of contact is required than in conventional HATR units. As a result, spectra can be obtained from a wide variety of solid samples, even from minerals using systems designed for high pressures. Some accessories use a thin diamond element in combination with zinc selenide. This approach can provide multiple reflections, giving spectra equivalent to those from HATR accessories from very small sample volumes. Because of their versatility, accessories of this type are sometimes described as universal sampling devices. However, it must always be remembered that ATR only measures the surface of the sample.

The ATR technique is also employed with the unit inserted into bulk liquids, for example for process monitoring. The principle is exactly the same, the difference being that the form of the crystal is a prism so that radiation enters and leaves in the same direction. Radiation is coupled into the prism with light guides or optical fibres. There are also ATR probes in which the fibre itself forms the ATR element. The fibres are either chalcogenide glasses or mixed silver halides. One design of contact probe has a U-shaped section of exposed fibre that is pressed against the sample. This approach has been used for *in vivo* measurements of skin.

## External reflection methods

Reflection measurements can be conveniently divided into surface or specular reflection and diffuse

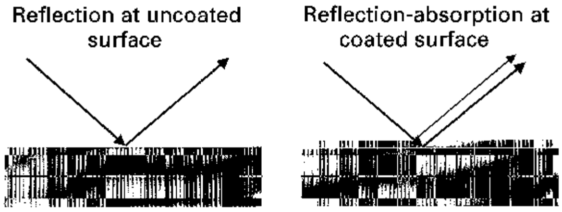



**Figure 10** Reflection and reflection–absorption at surfaces.

reflection. The reflection spectrum from the surface of a continuous solid, such as a large crystal or a block of polymer, does not resemble an absorption spectrum, being determined by the refractive index according to the Fresnel equations. However, it can be converted into an absorbance spectrum by a mathematical process called the Kramers–Kronig transform which is provided in the software from instrument suppliers. This is a common approach for obtaining spectra of single crystals and from polished surfaces of minerals. It is also very effective for carbon-filled polymers. Very simple accessories using two flat mirrors can be used to obtain reflection spectra from flat surfaces. The same accessories can be used to measure the spectra of coatings on metal surfaces such as protective or decorative polymer films on food and beverage cans. In this case what is measured is a transmission spectrum of the coating, the radiation having passed twice through the coating and having been reflected at the metal surface. This is sometimes called a transflectance measurement. Examples of the above are shown in **Figures 10** and **11**.

Diffuse reflection is seen when radiation penetrates below the surface before being scattered or reflected back from within the sample. Spectra of this kind have the character of transmission spectra since there is absorption as the radiation travels inside the sample. Samples for diffuse reflection are generally powders (**Figure 12**). Accessories for this kind of measurement have to collect the light that emerges from the sample over a wide range of angles. Integrating spheres are little used because suitable large detectors are not available for the mid-IR region. Instead, commercial accessories focus the radiation on to a small region, typically 1–2 mm across, and collect over a wide but limited range of angles (**Figure 13**). The spectra obtained with these accessories represent a mixture of reflection from the surface and from within the sample. They can be used to measure pure surface reflection from samples that have no internal reflection or scattering.

In spectra obtained directly from powdered organic compounds, the stronger bands often appear flattened and are distorted by surface reflection. The diffuse reflection spectrum can be thought of as a transmission spectrum in which there is a range of different pathlengths. This results in enhancement of

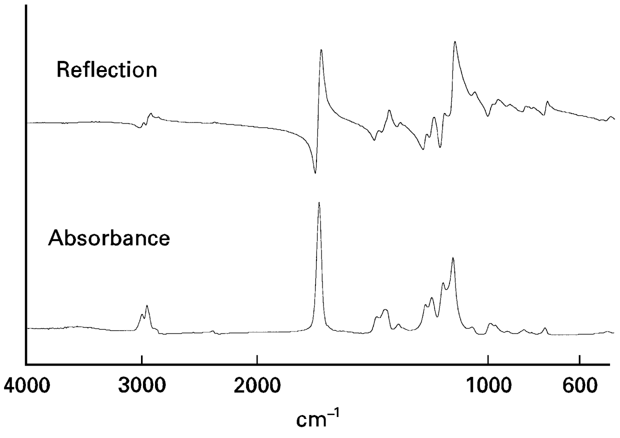



**Figure 11** Top, reflection spectrum of poly(methyl methacryiate); bottom, absorbance calculated by Kramers–Kronig transformation.

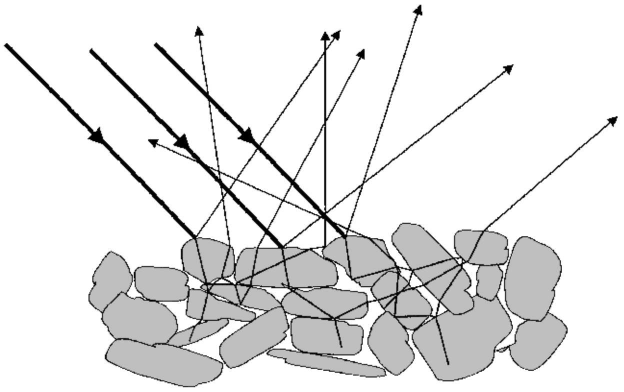


**Figure 12** Representation of diffuse reflection by a powder.

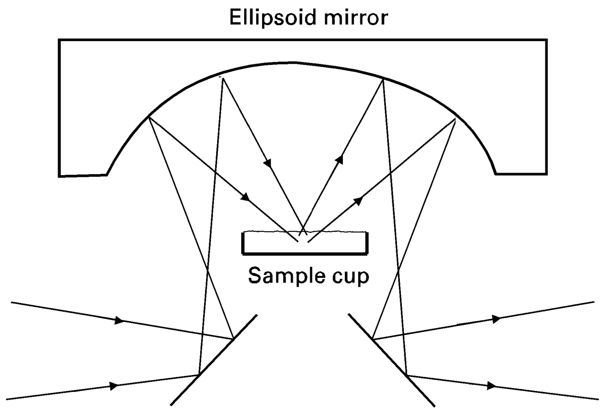



**Figure 13** Optics of diffuse reflection accessory.

the relative intensities of the weaker bands. In general, the aim is to obtain a result that resembles a true transmission spectrum. The standard technique is to grind the sample to a fine powder, ideally to a particle size below 10 μm, and to mix this with a non-absorbing matrix such as finely ground potassium bromide at a concentration of about 1%. The small particle size and dilution together ensure that absorption is weak for all pathlengths. This means that relative band intensities in transmission are similar for all contributing pathlengths and so are correct in the final spectrum. The presence of the non-absorbing matrix reduces reflection at the surfaces of the particles by better refractive index matching, and so minimizes the surface reflection component. Sample preparation is quicker than for a potassium bromide disc. It is also gentler, making it less likely to induce changes in the crystalline form. However, it has proved difficult to achieve sufficiently consistent sample preparation to allow good quantitative measurements. Representative spectra of finely ground aspirin are shown in **Figure 14**.

Diffuse reflection spectra are often transformed to Kubelka–Munk units as defined below, just as spectra measured in transmission are converted to absorbance:

$$f(R_\infty) \propto (1 - R_\infty)^2 / R_\infty$$

where $R_\infty$ is the ratio of the sample reflectance to that of a reference such as potassium bromide. The reason is that under certain conditions the band intensities in these units are proportional to concentration. However, the necessary conditions are not met in typical mid-IR measurements, so there is little reason to prefer this presentation to the alternative log(1/*R*), which is a direct equivalent of absorbance (**Figure 15**). It is worth noting that near-IR diffuse

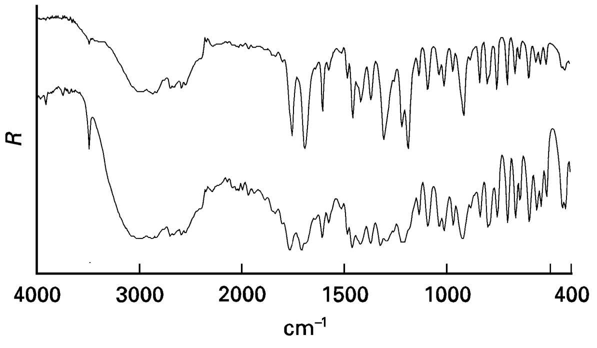



**Figure 14** Diffuse reflection spectra of finely ground aspirin. Top, diluted in KBr; bottom, neat.

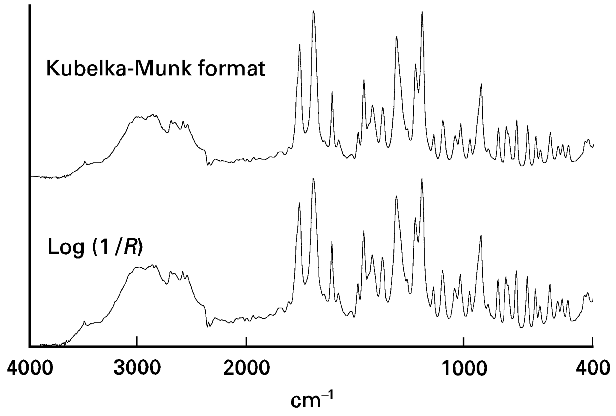



**Figure 15** Diffuse reflection spectra of aspirin in Kubelka–Munk and log(1/*R*) units.

reflection spectra are used extensively for quantitative work, almost invariably using log(1/*R*) rather than Kubelka–Munk units.

Another way of presenting samples for diffuse reflection is as a thin film of fine particles. This also achieves the objective of ensuring weak absorptions. Samples can be taken from the surface of large objects by abrading them with silicon-carbide-coated paper. The spectrum of the material removed can be measured directly on the abrasive substrate. This method has proved ideal for painted surfaces and moulded plastic objects. It can be extended to very hard materials by using abrasive devices such as metal rods tipped with diamond powder. The main limitation to this technique is that abrading soft and rubbery materials generally does not produce particles that are small enough to give good spectra.

## Photoacoustic spectroscopy

Photoacoustic measurements are unique in that they depend directly on the energy absorbed by the sample, rather than on what is transmitted or reflected (**Figure 16**). The sample is placed in a sealed chamber. When radiation is absorbed it is converted into heat which can be detected as a change of pressure in the gas above the sample. This heat transfer process is sufficiently fast to follow the modulation of the radiation produced in an FT-IR spectrometer. The change in pressure is detected with a microphone. A spectrum is obtained by ratioing the detector signal to that generated with a totally absorbing material such as carbon black. As with ATR, the spectrum comes from a surface layer, but in this case the effective penetration can be more than 100 μm. This can result in total absorption for strong bands, so that relative intensities are distorted. The effective

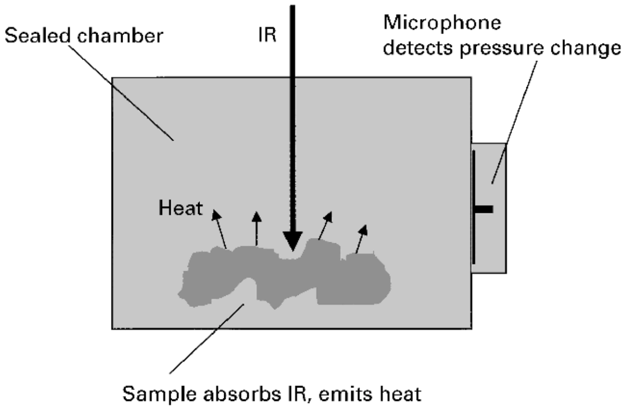



**Figure 16** The principle of a photoacoustic measurement.

depth of penetration depends on the modulation frequency which varies with wavelength in rapid-scanning FT-IR spectrometers. In step-scan spectrometers the modulation frequency is constant for all wavelengths, so simplifying interpretation of the spectra. Some information about the depth distribution of different components can be obtained by varying the modulation frequency. However, depth profiling, measuring a concentration quantitatively as a function of depth, has not been achieved.

In practice, the main value of photoacoustic measurements has been in obtaining spectra from strongly absorbing samples such as carbon-filled polymers and for materials that cannot be ground to a fine powder or be prepared with a flat surface. Sample preparation is minimal, requiring only that the specimen is small enough to fit in the chamber. Commercial accessories can accommodate samples several millimeters across.

## Microsampling

In recent years, the approach to measuring small samples has been transformed by the widespread use of IR microscopes. The advantages over traditional beam condensers lie in much greater convenience and versatility, as well as in the ability to measure much smaller samples. The region for measurement can be selected while a relatively large area of the sample is viewed. Spectra can be obtained from isolated samples that are only a few micrometres across. However, in extended samples diffraction limits spatial resolution to about 10 μm. Because of this, it is generally preferable to remove small inclusions from a matrix rather than to try to measure them *in situ*. One consequence of having very small samples is that it is often easy to compress them to a suitable thickness for transmission measurement. Compression cells are therefore more useful than they are with large samples. Sample preparation is greatly helped by having standard microscopy tools and a microtome available. Reflection measurements are more generally useful than they are with larger samples. Because of the small sample area and large collection angle, it is easier to find an area that gives a pure surface reflection spectrum. This can then be converted into an absorbance spectrum by the Kramers–Kronig transform.

## Miscellaneous techniques

Emission spectra can be obtained without any sample preparation, but they are of very limited use. If the sample is opaque, the information provided is identical with that in the reflection spectrum since the sum of transmission, reflection and emission at any wavelength adds up to one. The reflection spectra are usually measured more easily than emission. Emission is very useful to obtain spectra from thin coatings on uneven metal surfaces when diffuse reflection measurements would require a small sample.

Pyrolysis can provide qualitative information from intractable samples without requiring expensive accessories. The simplest approach is to heat the sample in the bottom of a long test-tube so that decomposition products condense at the top of the tube and can be measured by the usual methods. Commercial devices also allow gaseous products to be collected. There are collections of the spectra of decomposition products from polymers. The most sophisticated version of this method is to couple a gas cell in the spectrometer to a thermogravimetric analyser and measure the products evolved at different temperatures.

Separation techniques are needed in order to identify individual components in mixtures. The most important of these is gas chromatography. To maximize sensitivity, the sample has to be confined to a small area. The usual approach is to pass the gas flow from the GC column through a light pipe ~1 mm in diameter and 10 cm long. This achieves detection limits of ~10 ng for strongly absorbing compounds. For higher sensitivity a more elaborate deposition method is available. The effluent from the column is directed on to a moving cold window where the peaks in the chromatogram can be condensed as a series of spots about 100 μm across. The detection limits are below 100 ng.

Coupling to liquid chromatography is much less successful because of the problem of eliminating the solvent and modifiers. In the most widely used system the eluate is deposited in a circular track on a rotating heated window which is subsequently

transferred to the spectrometer. This is used especially in combination with size-exclusion chromatography.

## List of symbols

$d_p$ = penetration depth; $n$ = refractive index; $R_\infty$ = ratio of sample reflectance to that of a reference; $\alpha$ = angle of incidence; $\lambda$ = wavelength.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Chromatography-IR, Applications; Chromatography-IR, Methods and Instrumentation; High Resolution IR Spectroscopy (Gas Phase), Applications; High Resolution IR Spectroscopy (Gas Phase), Instrumentation; Photoacoustic Spectroscopy, Applications; Photoacoustic Spectroscopy, Theory; Surface Studies by IR Spectroscopy.**

## Further reading

Chalmers JM and Dent G (1998) *Industrial Analysis with Vibrational Spectroscopy*. Cambridge: Royal Society of Chemistry.

Griffiths PR and de Haseth JA (1986) *Fourier Transform Infrared Spectroscopy*. New York: Wiley.

Harrick NJ (1977) *Internal Reflection Spectroscopy*. Ossining, NY: Harrick Scientific.

Messerschmidt RG and Harthcock MA (1988) *Infrared Microspectroscopy*. New York: Marcel Dekker.

Nishikida K, Nishio E and Hannah RW (1995) *Selected Applications of Modern FT-IR Techniques*, Tokyo: Kodanasha.

# IR Spectroscopy, Theory

**Derek Steele**, Royal Holloway College, Egham, UK

Copyright © 1999 Academic Press

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
|---|
| Theory |

## Factors which control absorption

Absorption in the infrared region arises predominantly from excitation of molecular vibrations. For $N$ atoms there are $3N$–6 vibrational degrees of freedom ($3N$–5 for a linear molecule), which is equivalent to stating that all vibrational distortions can be described as the sum of $3N$–6 '*fundamental*' vibrational modes. Where any symmetry exists in a molecule, the vibrations will reflect that symmetry in that the vibrational wavefunctions will be symmetric or antisymmetric with respect to any simple symmetry element. This statement has to be modified slightly for situations where degeneracy of rotations occurs. For a detailed discussion of groups and their relevance in spectroscopy the reader is referred to the Further reading section. In the event that a rigid point group is inappropriate, as for free rotors (e.g. $CH_3BF_2$), then the invariance of energy to certain distortions requires that permutation group theory be used. What we are seeking to explore with group theory is the invariance of energy and the consequences to certain transformations of the molecular configuration. In free rotors, by definition, the energy is unchanged by rotation about a bond. This then makes rigid group theory inappropriate.

For excitation of any single mode $i$, the energy, $E_i$, can be expressed as

$$\frac{E_i}{hc} = \omega_i(v+\tfrac{1}{2}) - \omega_i\, x_i\, (v+\tfrac{1}{2})^2 + \omega_i\, y_i\, (v+\tfrac{1}{2})^3 + \cdots \quad [1]$$

where $v$ is the vibrational quantum number and $\omega$, the wavenumber, is $\nu/c$. Restricting the discussion initially to a harmonic oscillator with no molecular rotation, we can then examine the conditions required for vibrations to be excited through interaction with electromagnetic radiation. For absorption the following conditions must be satisfied:

(i) the photon energy, $h\nu_i$, must be equal to the change in vibrational energy. We should qualify this by giving due allowance for the Doppler effect. In a low-density gas the spread of molecular velocities relative to the probing beam leads to a natural line

width, $\Delta\nu$, at a temperature $T$ of

$$\Delta\nu = \frac{\nu_i}{c}\left(\frac{2\,k_B\,T\ln 2}{M}\right)^{\frac{1}{2}}$$

where $M$ is the relative molecular mass. Such widths are very small and relevant only when probing fine structure. Thus, for methane at room temperature and for a band at $\omega = \nu_i/c = 1500\ \mathrm{cm}^{-1}$ then $\Delta\omega = 0.003\ \mathrm{cm}^{-1}$. By using molecular beams moving at right-angles to the probing beam resolutions substantially less than this are attainable. Other techniques exist for exploring intrinsic line widths which are less than the Doppler width. These include two, or more, photon excitation and saturation spectroscopy.

(ii) Interaction of electromagnetic radiation and matter can only occur through electric and magnetic fields. For vibrational spectroscopy the electric field interaction is by far the greater. In classical terms there must be a change in dipole moment for the vibration to be excited. Thus for stretching of the bond of a homonuclear diatomic no absorption at the vibrational frequency occurs. In general, the absorption is proportional to the square of the dipole change. In terms of group theory only those vibrations which form representations of the same classes as the electric dipole vector (or simple translations) can be excited by EMR. Thus, for $CF_3CF_3$ ($D_{3h}$ symmetry), only vibrations of the $A_2''$ and $E'$ species can absorb electromagnetic energy of the appropriate frequency since a dipole vector forms such representations. We must note that it is unnecessary for an electric dipole to exist in the equilibrium configuration.

According to quantum mechanics, the probability of an induced transition per unit time per unit radiation density between states $|i\rangle$ and $|f\rangle$ is given by Fermi's 'Golden Rule':

$$B_{if} = B_{fi} = 2\pi^2/h^2\varepsilon_0|\langle f|\hat{e}_r\cdot\mu|i\rangle|^2\ \delta(\nu-\nu_f+\nu_i) \qquad [2]$$

where $\varepsilon_0$ is the permittivity of free space, $\hat{e}_r$ is a unit vector in the direction of the electric field vector of the radiation and $\delta(\ )$ is the Kronecker delta, which takes the value zero unless its argument is zero, in which case it has a value of unity. In the above equation the Kronecker delta requires that the radiation frequency is exactly equal to the difference in energies of the initial and final states divided by Planck's constant.

For an ensemble of randomly oriented molecules, the power absorption is given by

$$\mathrm{d}I_\nu = (N_iB_{if} - N_fB_{fi})\rho_\nu\ h\nu_{fi} \qquad [3]$$

where $N_i$ is the number of molecules in state $|i\rangle$ and $\rho_\nu$ is the radiation density. Defining the intensity over the $k$th band as

$$\Gamma_k = \frac{1}{Cl}\oint \ln\frac{I_0^\nu}{I^\nu}\frac{\mathrm{d}\nu}{\nu} \qquad [4]$$

and assuming harmonic oscillator wavefunctions, then Equations [2], [3] and [4] yield

$$\Gamma_k = \frac{N\pi}{3\nu_k}\left(\frac{\partial\mu}{\partial Q_k}\right)^2$$

where $C$ is the concentration of the absorbing species, $l$ is the pathlength, $I^\nu$ and $I_0^\nu$ are the intensities of radiation at the frequency $\nu$ with and without the sample absorption and $\partial\mu/\partial Q_k$ is the dipole change with respect to the vibrational motion in normal, or fundamental, mode $k$. $\nu_k$ is the fundamental transition frequency. The definition of band intensity in Equation [4] differs from the usual integral (the integrated absorbance, $A$) in the presence of $\nu$ in the denominator, but is much more satisfactory in relation to theory and to isotopic intensity sum rules. Since the band widths are generally much smaller than the frequency band centre, the integrated absorbance, $A$, is related to $\Gamma$ by $A \approx \Gamma\nu$.

(iii) The dipole moment can be written as a Taylor series expansion in the vibrational motion $Q$:

$$\mu = \mu_0 + \frac{\partial\mu}{\partial Q}Q + \frac{1}{2}\frac{\partial^2\mu}{\partial^2 Q}Q^2 + \text{higher terms}$$

For a harmonic oscillator, quantum theory requires that the transition integral $\langle f|\,Q\,|i\rangle$ disappears unless the vibrational quantum numbers $v_i$ and $v_f$ are related by $v_i - v_f = 0$ or $\pm1$. This requires that overtones are forbidden if only the linear term in $Q$ exists. It is known, however, that the vibrational spectrum contains weak features due to overtones and combination bands in which two quantum numbers change simultaneously. These occur because the dipole expansion in general has finite terms quadratic in the deformation and because true wavefunctions are nonharmonic. Occasionally combination bands appear with considerable strength

due to mixing with a nearby fundamental transition of the same symmetry (Fermi resonance). Interaction leads to mutual repulsion of the mixed energy states. In the event that the combination transition has zero or much less intensity than the fundamental, then the unperturbed fundamental transition can be estimated as the intensity weighted mean of the two bands. A classical example of Fermi resonance is seen in the infrared spectrum of $CCl_4$. The very strong doublet at 768/797 $cm^{-1}$ is due to the triply degenerate CCl stretch interacting with the combination state $\nu_1 + \nu_4$ (305 + 458 $cm^{-1}$).

There is considerable interest in the factors that determine the dipole derivative. The stretching of a highly polar bond, such as C–F, might well be expected to lead to a large dipole change. In practice, the dipole change is much greater than would be expected for the charges on the atoms in their equilibrium positions. Charge flows as the nuclei move, in this case towards the situation where the fluorine atom carries a complete extra electron. In certain cases where a *trans* lone pair of electrons exist, such as for the hydrogen atom in the aldehyde unit, extension of the bond results in a flow of electrons from the lone pair into an antibonding orbital of the C–H bond, thereby greatly enhancing the dipole change. This is the same phenomenon that is responsible for bond weakening in such cases. In bending motions, nonbonding interactions, as well as rehybridization of the electrons at the nonterminal atom, cause substantial deviations from the dipole change expected from the equilibrium charge distribution. This explains the remarkably high intensity associated with movement of a hydrogen atom from the plane of benzene or ethene derivatives. As the electrons rehybridize around the carbon (moving from $sp^2$ towards $sp^3$), so electrons build up on the opposite side of the skeletal plane to the hydrogen, thereby enhancing its positive charge.

Many studies have been made of the absorption intensities associated with bands arising from group vibrations. It is recognized that the transition moments are far more sensitive to neighbouring group effects than are the frequencies themselves, and intensities are not particularly useful in evaluating the concentration of functional groups. However, some qualification of this statement is required in view of the widespread application of principal factor (or principal component) analysis (PFA) in industrial spectroscopy. This has been mainly used with near-infrared spectra, but this is probably more to do with ease of generating suitable spectra than to fundamental limitations. The application to the mid-infrared spectra of coals is an excellent example. The relevance to group intensities is that PFA is a statistical technique of fitting data to a progressively increasing number of unknown components until the fit reaches an acceptable level. As such it requires the data to consist of a finite number of linearly additive components. The application to very complex mixtures such as coals, wheat and petroleum, both refined and unrefined, strongly suggests additivity is much better than used to be believed. This has been rationalized by demonstrating that group spectra are characteristic provided that the neighbouring groups are the same.

## Band shapes in liquids

In low-density vapours an absorption band consists of lines arising from a large number of transitions arising from rovibrational transitions, perhaps from excited vibrational states as well as the ground state. In high-density fluid states the interaction between different molecules will affect the spectrum by causing perturbations of the rovibrational energy states and by reducing the lifetimes of the excited states through energy transfer processes. The individual transitions initially broaden and then merge into a smooth band contour. It is sometimes stated that the rotational transitions are suppressed, but this cannot be so. The equipartition theorem states that energy must be equally distributed between each degree of freedom, and this must extend to the condensed state. It has been demonstrated that the contours still contain the rotational contribution. Thus for a diatomic molecule of moment of inertia $I$, the integrated second moment of the intensity ($\Gamma$ as defined in Eqn [1]) about the vibrational band centre, divided by the integrated intensity itself, is, apart from a vibrational relaxation contribution, $k_B T / I_X$, where $I_X$ is the moment of inertia. This is the same for vapour and for liquid. It is simply the relaxation of the excited states that has changed. There is a vibrational contribution to the width (see below). It is possible in certain situations to separate these contributions using Raman spectroscopic methods. If we designate the transition dipole moment at a time $t$ of a given transition as $\mu(t)$ then its development as a function of the time lapse $t = t'' - t'$ is related to Fourier transform about the band centre of the intensity as

$$C(t'' - t') = \langle \mu(t'') \cdot \mu(t') \rangle = \langle \mu(0) \cdot \mu(t'' - t') \rangle$$
$$= \int_{-\infty}^{+\infty} \Gamma(\nu_0 + \Delta\nu) \exp(i2\pi\, \Delta\nu t)\, \mathrm{d}\nu$$

The brackets ⟨ ⟩ designate an ensemble average and, since the transition dipole is a vector, this transform

yields the ensemble time decay of the transition dipole moment from an arbitrary time of origin. The time scale of the decay will be of the order of the inverse of the band width. Thus for a band of half-band width 10 cm$^{-1}$ the half-life of the decay will be ~1/(3 × 10$^{10}$ × 10) ≈3 × 10$^{-12}$ s. If we discount vibrational relaxation and consider the transition moment as being fixed to the molecule, then the correlation function $\langle\mu(0)\cdot\mu(t'-t'')\rangle$ will describe the rotation of the molecule as defined by the transition dipole axis. In the condensed state after an initial period in which the motion is governed by the rotational energy, then decay becomes exponential due to collisional processes. In the vapour the dipole correlation shows a minimum, perhaps becoming negative, in a period after which most molecules are facing in the opposite direction to that in which they started (see **Figure 1**). As in magnetic resonance, the decay time will consist of isotropic and anisotropic contributions. In a fluid the transition dipole will also decay by transfer of energy between molecules and by multi-body collision deactivation of the vibration. These are responsible for the isotropic component. Owing to the very short time-scale of these processes, separation of the relaxation mechanisms is generally impossible. However, some progress has been made with laser optical nutation experiments and with Raman studies.

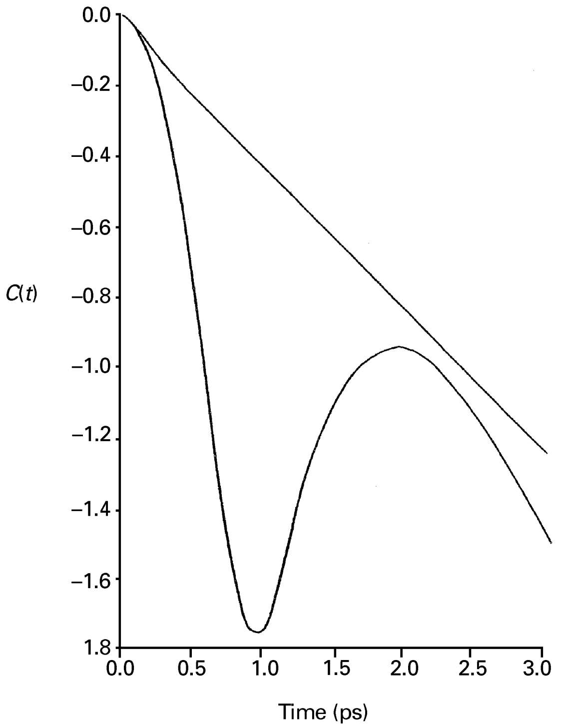



**Figure 1** The dipole autocorrelation function of the $\nu_{11}$ band of benzene. The top line is the dipole correlation function for benzene dissolved in cyclohexane and the bottom line is that for benzene vapour. Hill IR and Steele D, unpublished work.

## Calculation of vibrational spectra

Although energetic aspects of vibrational spectroscopy, such as zero point energy and crystal lattice energies, are strictly quantum phenomena, much can be achieved by classical-type calculations based on inter-atomic potential functions. Assuming that the potential energy for deformation about an equilibrium configuration is strictly quadratic in deformations then the vibrational frequencies are given as the eigenvalues of the matrix product *GF* where *F* is the matrix of quadratic force constants and *G* (so-called inverse kinetic energy matrix) is such that

$$G_{ij} = \sum_k \frac{\partial R_i}{\partial \xi_k}\frac{\partial R_j}{\partial \xi_k}\frac{1}{m_k} \quad \xi = x,\, y,\, z$$

where *R* represents an internal coordinate of the molecule, usually chosen from bond stretches, angle bends and bond torsions. The sum is over all atoms. The methodology of carrying out these classical calculations are well documented. One of the principle advantages of such calculations is that the eigen-vectors of *GF* yield the nature of the molecular vibration.

Up to this point we have assumed harmonic oscillations. In reality we cannot ignore the impact of cubic and higher terms in Equation [1]. For a change in quantum number of $\Delta v$, then from Equation [1] we have for a transition $|0\rangle \rightarrow |v\rangle$

$$\frac{\Delta E}{hc} = \omega_i \Delta\nu - \omega_i x_i (2\nu + 1)\Delta\nu + \text{higher terms}$$

A fundamental transition ($v$ = 0→1) then has a wavenumber of $\omega_i - 2\omega_i x_i$. The anharmonicity constant $x$ is much less than 1. For the C–H vibrations it is relatively large because of the considerable amplitude of vibration, and it has a value of 0.021. This means that for a transition seen at 3000 cm$^{-1}$ the fundamental frequency is at 3126 cm$^{-1}$. The first overtone would be expected at $2\omega_i - 6\omega_i x_i$ = 5874 cm$^{-1}$. In fact the C–H stretching overtone and combination frequencies deviate from these expectations. The cause of this lies in the large amplitudes of motion which result in a breakdown of local symmetry. A methylene group in an environment which conserves the equivalence of the two bonds will have vibrations which are symmetric and antisymmetric with respect to the bonds, and in the limit of small amplitude vibrations point group symmetry applies. As the amplitudes of all the vibrations increase, so the symmetry breaks down and the bonds tend towards

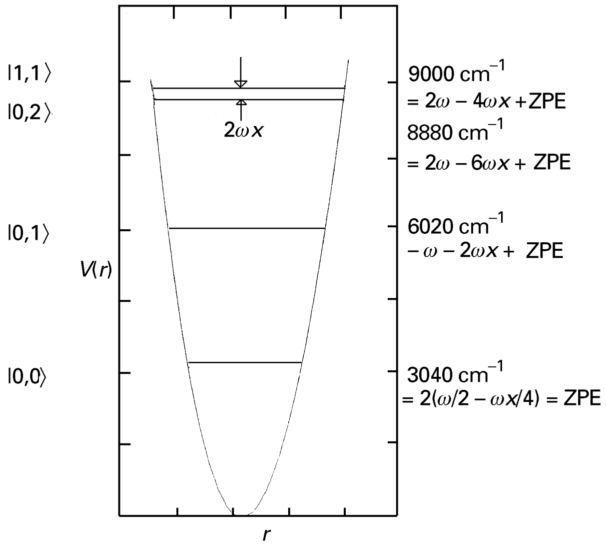



**Figure 2** Relative energies in the local mode model of ground, fundamental, first harmonic and combination states for a system with two degenerate vibrators. A fundamental harmonic wavenumber of 3100 $cm^{-1}$ and a vibrational anharmonicity of $\omega x = 60$ $cm^{-1}$ have been assumed. ZPE represents the zero point energy.

independent oscillators. In this event the theoretical model needed becomes that of the local mode. Application of Equation [1] shows that the separation of states in which one of two equivalent bonds is doubly excited and that in which both bonds are singly excited is $2x_i\omega_i$. The higher transition will be at twice the observed fundamental frequency (see **Figure 2**).

## Quantum theoretical calculations of spectra

Applied quantum theory has reached the stage where infrared and Raman spectra can be predicted to a very useful degree of accuracy for molecules of fairly substantial size. The ability to programme the calculations so that force constants and dipole gradients are calculated analytically from the wavefunctions is one of the features which has greatly enhanced these computations. The calculations as routinely performed are of the harmonic wavenumbers and the band intensities. At the Hartree–Fock level the wavenumbers are too high when compared with experiment owing to the harmonic approximation, but also due to lack of electron correlation and the restriction of electron pairing. The latter leads to dissociation to states in which electron pairing is still present, and these are generally of much higher energy. Very good results are obtained by simply multiplying the wavenumbers (frequencies) by a scaling factor. Improved results are generated by scaling the force constants, keeping one factor for C–H stretching, another for the C–H bending constants, etc. Improved results, with scaling constants nearer to unity, are obtained with perturbation calculations of the electron correlation correction (Möller–Plesset theory). Density functional theory (DFT) now permits very good frequencies to be obtained in times which are close to the Hartree–Fock times. Both Raman and infrared intensities are much more difficult to compute. This is not surprising when it is seen that dipole moments are extremely sensitive to the basis sets. Although there are cases where the results are as yet disappointing, DFT appears to be yielding much improved intensities and dipole moments over earlier Hartree–Fock and Möller–Plesset calculations. Although these calculations have led to much improved force fields and spectra, it must be said that some difficulties still exist.

## Vibration rotation spectra

When a vibrational quantum state changes, other quantum states of lower energy will also generally change. Thus, for a heteronuclear diatomic molecule the $|0\rangle$ to $|1|$ transition is accompanied by a change in rotational quantum number of ±1 leading to the well known P and R branches. For polyatomic molecules the selection rules required so as to allow the nondisappearance of the transition integral are more complex, and for details the reader is referred to a specialist text. In the situation where individual rovibrational transitions cannot be resolved in an asymmetric top, then the overall contour can be related to the direction of the transition moment with respect to the inertial axes and to the moments themselves. Thus transitions in which the dipole change occurs along the direction of greatest moment of inertia are characterized by a strong Q branch (type C bands), corresponding to no change in the rotational quantum number change (see **Figure 3**).

The contours have been calculated as a function of the moments of inertia and presented in easily used figures. With laser and high-resolution Fourier instruments, an increasing amount of detail of rovibrational information is becoming available. Here we shall refer to a few general points.

The parameters of fitting include the moments of inertia. Even with small molecules and isotopically substituted molecules it is very difficult to generate accurate molecular geometries. The inertial constants refer to the ground and vibrationally excited states, but even in the ground state at the absolute zero of temperature each quantum state has a half quantum of energy (zero point energy, see **Figure 2**). The average of the root mean square bond length

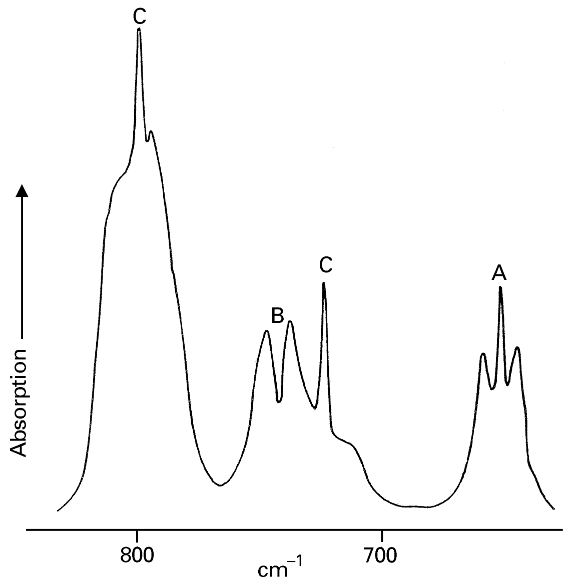



**Figure 3** Part of the infrared absorption spectrum of 2,6-difluoropyridine in the vapour phase showing typical A, B and C contours (transition moments along axes of least, intermediate and greatest moments of inertia, respectively.) Reproduced with permission from Bailey RT and Steele D (1967) *Spectrochimica Acta, Part A* **23**: 2997.

during a vibration differs significantly from the equilibrium bond length. *Ab initio* calculations, such as referred to in the previous section, may be used to estimate the necessary corrections.

One feature unique to vibration rotation interaction is due to Coriolis interaction. There is a component of force on a moving particle due to rotation about a perpendicular axis equal to

$$\omega \otimes (\omega \otimes m_i r_i)$$

where ⊗ represents the vector product operator, $\omega$ is the molecular angular velocity and $m_i$ and $r_i$ are the mass and coordinate of the *i*th particle. Such coupling leads to interaction between two states, the product of whose symmetry representations belongs to the same species as a molecular rotation. For degenerate vibrations, this can be a first-order effect and has a dramatic impact on the overall vibrational rotational contour. The same phenomenon, of course, controls the contrary direction of rotation of air masses in the northern and southern hemispheres.

## List of symbols

$A$ = integrated absorbance; $B_{\text{if}}$ = probability of an induced transition per unit time per unit radiation density; $c$ = velocity of light; $C$ = concentration; $E_i$ = energy of $i$th mode; $F$ = quadratic force constant matrix; $G$ = inverse kinetic energy matrix; $h$ = Planck's constant; $k_B$ = Boltzmann's constant; $l$ = pathlength; $M$ = molecular mass; $Q$ = normal mode; $R$ = internal coordinate; $v$ = vibrational quantum number; $\Gamma$ = integrated absorption intensity; $\mu$ = dipole moment; $\nu$ = frequency; $\rho$ = radiation density; $\omega$ = wavenumber.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Far IR Spectroscopy, Applications; Fast Atom Bombardment Ionization in Mass Spectrometry; High Resolution IR Spectroscopy (Gas Phase), Applications; High Resolution IR Spectroscopy (Gas Phase), Instrumentation; Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy; IR Spectral Group Frequencies of Organic Compounds; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; IR Spectrometers; IR Spectroscopy Sample Preparation Methods; Near-IR Spectrometers; Rotational Spectroscopy, Theory; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading

Allen HC and Cross PC (1963) *Molecular Vib-Rotors*. New York: Wiley.

Badger RM and Zumvalt LR (1938) *Journal of Chemical Physics* **6**: 711.

Berne BJ and Pecora R (1976) *Dynamic Light Scattering*. New York: Wiley.

Brewer RG and Shoemaker RL (1971) *Physical Review Letters* **27**: 631.

Bright Wilson E, Decius JC and Cross PC (1955) *Molecular Vibrations*. New York: McGraw-Hill.

Cagnac B (1982) *Philosophical Transactions of the Royal of the Society of London, Series A* **307**: 633.

Cotton, FA (1990) *Chemical Applications of Group Theory*, 3rd edn. New York: Wiley.

Crawford B Jr (1952) *Journal of Chemical Physics* **20**: 977.

Fredericks PM, Lee JB, Osborn PR and Swinkels DAJ (1985) *Applied Spectroscopy* **39**: 303, 311.

Gordon RG (1965) *Journal of Chemical Physics* **42**: 3658.

Gordon RG (1965) *Journal of Chemical Physics* **43**: 1307.

Jennings DE (1978) *Applied Physics Letters* **33**: 493.

Longuet Higgins HC (1963) *Molecular Physics* **6**: 445.

Martens H and Naes T (1989) *Multivariate Calibration*. New York: Wiley.

Mills I and Robiette AG (1985) *Molecular Physics* **56**: 743.

Parker ME, Steele D and Smith MJC (1997) *Journal of Physical Chemistry, Section A* **101**: 9618.

Steele D (1971) *Theory of Vibrational Spectroscopy*. Philadelphia: WB Saunders.

Wexler AS (1967) *Applied Spectroscopy* **1**: 29.

# Iridium NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# Iron NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# Isotope Ratio Studies Using Mass Spectrometry

**Michael E Wieser**, University of Calgary, Alberta, Canada
**Willi A Brand**, Max-Planck-Institute for Biogeochemistry, Jena, Germany

Copyright © 1999 Academic Press.

| MASS SPECTROMETRY |
|---|
| Applications |

## Introduction

Isotopes of an element contain the same number of protons but different numbers of neutrons. Whereas the former means that isotopically different compounds undergo the same reactions, the latter means that they differ in mass (e.g. ${}^{12}_{6}C$ and ${}^{13}_{6}C$ are 12 and 13 atomic mass units, respectively). As a result of their different masses, isotopes of an element participate in chemical, biological and physical processes at different rates. Hence, the isotope composition of the element in a given compound depends on the history and origins of the sample. Isotope abundance data can provide information concerning the source of a material or processes responsible for its synthesis and conversion. Variations in the isotope abundance ratios of many elements of biogeochemical importance are subtle, but significant. To resolve these differences, isotope ratio mass spectrometers must accurately measure variations of 20 to 50 parts per million. In the case of carbon, the average ${}^{13}C/{}^{12}C$ isotope abundance ratio of 0.011 200 ranges over ±0.000 450 with biogeochemically important variations of ±0.000 000 5. Such requirements have resulted in a branch of mass spectrometry that has developed its own specialized instrumentation and analytical methods.

Isotope ratio studies are employed in a variety of multidisciplinary research projects encompassing chemistry, physics, biology, medicine, geology, archaeology and environmental technology. This article will describe mechanisms responsible for isotope abundance variations, the essential components of the isotope ratio mass spectrometer (with particular reference to electron impact ionization) and isotope abundance variations of H, C, N, O and S. Finally, the development of recent isotope ratio monitoring techniques to extract isotope information from transient signals is presented.

## Natural isotope ratio variations

To understand how mass differences give rise to isotope abundance variations, it is useful to consider chemical bonds as harmonic oscillators with fundamental vibrational frequencies inversely proportional to the square root of the molecular masses. The zero point energy of a molecule is defined as the finite energy (above $h\nu$ where $h$ is Planck's constant and $\nu$ is the frequency of the oscillation) possessed by that molecule at 0 K. Isotope substitution changes the mass of the molecule and results in a shift in the zero point energy. A molecule containing a

lighter isotope of an element will have a higher vibrational frequency and hence higher zero point energy compared to the molecule with the heavier isotope. The bonds in the molecule with the higher zero point energy are more easily broken. For all elements except hydrogen, the vibrational terms have the largest isotope effect. The magnitudes of these effects are proportional to the relative mass difference between the isotopic species.

Processes that result in isotope ratio variations can be divided into three general categories: (1) isotope exchange reactions such that isotopes of an element are redistributed among different molecules or different phases of the same molecule at thermodynamic equilibrium; (2) unidirectional processes where the reaction rates of chemical or physical reactions of isotopic species differ; and (3) radiogenic decay.

### Equilibrium isotope effects

Equilibrium isotope exchanges involve the redistribution of an element's isotopes among different chemical compounds or different phases of a molecule. While chemical equilibrium is necessary in order to attain isotopic equilibrium, isotope exchange will occur without any net change in the distribution of the chemical species in the system. Generally, the heavier isotope is favoured in the molecule that has the lowest zero point energy or the more condensed phase. The equilibrium constant of the isotope exchange reaction is related to the partition function ratios of the products and reactants. Partition functions control the distribution of internal energies within molecules. It is possible to calculate reduced partition function ratios and hence equilibrium constants for simple gaseous molecules (and to some extent condensed phases) based on spectroscopic and thermodynamic data. The dependence of the equilibrium constant on temperature is such that at higher temperatures the extent of isotope fractionation decreases.

The relationship between the isotope abundance ratios of the different species in the reaction is expressed as the fractionation factor '$\alpha$'. For example, consider the oxygen isotope exchange of $H_2O$ between condensed and vapour phases:

$$\mathrm{H_2{}^{18}O_{vapour} + H_2{}^{16}O_{liquid} \leftrightarrow H_2{}^{16}O_{vapour} + H_2{}^{18}O_{liquid}}$$

The corresponding isotope fractionation factor $\alpha$ is 1.0092 at 25°C, i.e. the liquid phase is enriched in $^{18}O$ over the gaseous phase by 0.92 percent (Eqn [1]):

$$\alpha_{\mathrm{liquid-vapour}} = \frac{(^{18}\mathrm{O}/^{16}\mathrm{O})_{\mathrm{liquid}}}{(^{18}\mathrm{O}/^{16}\mathrm{O})_{\mathrm{vapour}}} = 1.0092 \qquad [1]$$

### Kinetic isotope effects

Kinetic isotope effects are observed in unidirectional processes (such as diffusion, evaporation, and bacterial conversions) where the reaction rate depends on the masses of the reacting molecules. Given the same amount of kinetic energy, molecules of lighter mass $m_1$ will react at a rate $\surd(m_2/m_1)$ faster than the heavier molecules of mass $m_2$.

Biochemical reaction pathways are often dependent on the mass of an element at a particular position in a molecule. The result mostly is to favour the lighter isotopic species in the formation of the product. Often, exact information about the intermediate activated complex between reactants and products is not available. This makes calculation of rate constants and hence isotope fractionations from first principles difficult.

### Radiogenic isotopes

The transmutation of elements due to radioactive decay is another important cause of isotope abundance variations. Over time, the numbers of radioactive parent isotopes will decrease as they decay into the daughter products. The abundance determination of radiogenic daughter isotopes is applied extensively for age determinations of geological and organic materials. The ages of minerals and rocks ($t$) can be calculated from the measurement of the number of daughter nuclides formed in a mineral that decayed under closed conditions ($D^*$) and the parent atoms ($N$) where the half-life of the parent (probability of a decay, $\lambda$) is known (Eqn [2]).

$$D^* = N[\exp(\lambda t) - 1] \qquad [2]$$

Successful age determinations are possible providing (1) the rock or mineral has remained closed and changes in the numbers of parent and daughter atoms are due to nuclear decay only; (2) the decay constant is known accurately; (3) the number of daughter nuclides originally present (at time $t = 0$) is known; and (4) the numbers of daughter and parent atoms can be measured accurately.

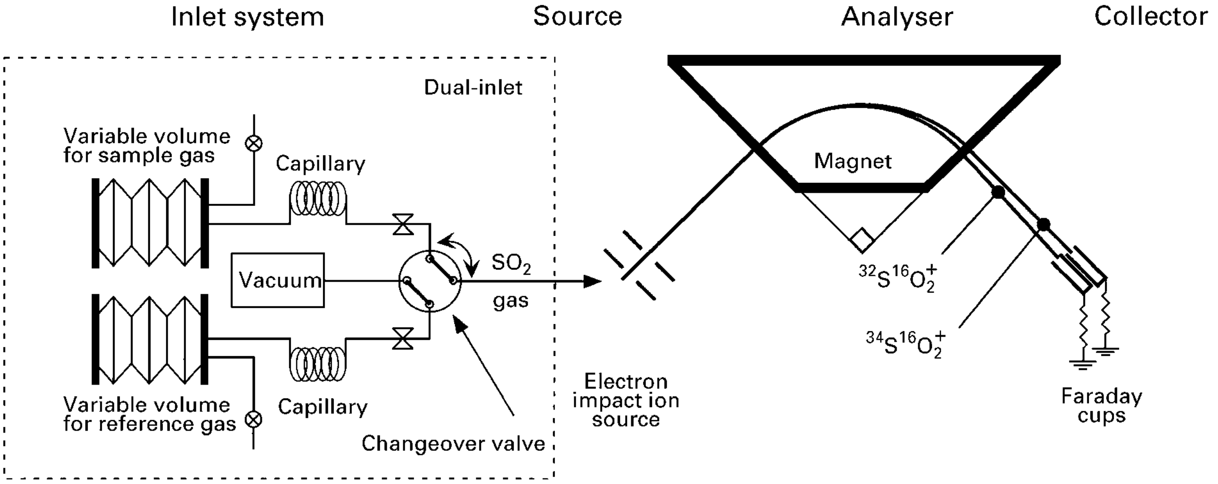



**Figure 1** Typical mass spectrometer and inlet system for isotopic comparison of an unknown sample with a reference gas of precisely known isotopic composition. Clean gases are needed for such systems.

## Isotope ratio mass spectrometer

The majority of isotope abundance determinations are performed using magnetic sector mass spectrometers based on a design from Alfred Nier dating back to the 1940s. The isotope ratio mass spectrometer consists of an ion source, magnetic sector and Faraday cup ion detector(s) (**Figure 1**). The source is responsible for the production of ions from either solid or gaseous material. The magnetic sector separates ions according to their mass to charge ratio and brings them into focus at the collector. The detector records and converts the charge of the ions into (computer readable) electrical signals. Neutral species are removed from the interior of the mass spectrometer (maintained at less than $10^{-8}$ mbar or $10^{-6}$ Pa) to minimize the number of collisions between ions and residual molecules in the ionization and analyser regions. These interactions would otherwise cause the ions to lose variable amounts of kinetic energy, resulting in a defocused ion beam.

Isotope ratio mass spectrometers tend to be of low resolution (~ 100, 10% valley) to resolve the isotopomers in the simple gases introduced (i.e. in the case of $CO_2$ gas; $^{12}C^{16}O_2^+$, $^{13}C^{16}O_2^+$ and $^{12}C^{16}O^{18}O^+$ have masses 44, 45 and 46). The mass spectrometer provides high precision isotope abundance data, typically of the order of 0.001%. This is achieved by optimizing the number of ions collected at the detector to improve the counting statistics.

### Magnetic sector

The task of the ion optics of the mass spectrometer (of which the magnetic sector is the critical component) is to bring all of the ions produced at the ion source into focus at the collector. Ions formed in the source are collimated and accelerated through an electric potential difference (V) from 2000 to 10000 V, depending on instrument type. The accelerated ions enter the magnetic field and are separated along radial paths dependent on their charge to mass ratio and kinetic energy. The radius (r) of the path traced by an ion while in the magnetic field is related to its mass (M), charge (q), the potential of the source (V), and the magnetic field flux density (B) according to Equation [3]:

$$r = \sqrt{\frac{2MV}{qB^2}} \qquad [3]$$

It is possible to change either the accelerating potential of the source or the magnetic field strength to bring an ion of a given mass into focus at the collector. A unique feature of the magnetic sector isotope ratio mass spectrometer is the trapezoidal peak shape produced by the selection of ion source (entrance) and collector (exit) slit widths (**Figure 2**). As the ion beam is scanned across the exit slit of the mass spectrometer, the ion current increases until the entire width of the ion image is contained inside the collector. The resulting flat-top peak ensures that the measured ion current does not change if there are small fluctuations in the accelerating potential (V) and/or magnetic flux density (B).

### Faraday cup detectors

The ion currents realized with most applications are of the order of $10^{-9}$ A for the major (most abundant) isotopic species and can be measured by a Faraday

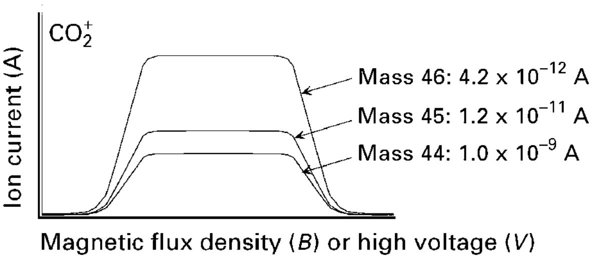



**Figure 2** Flat-topped peaks ensure that small fluctuations in the magnet current or high voltage regulation do not hamper precision. For clarity, the peaks are recorded with different sensitivity factors (1:100 : 300 for 44, 45 and 46).

cup type detectors (**Figure 3**). The Faraday cup is connected to ground via a high ohmic resistor ($10^8$ to $10^{12}$ Ω). The incident ions produce a current (I) resulting in a voltage drop (V) across the resistor (R) proportional to the ion current incident at the collector by Ohm's law (V = IR). Either voltage to frequency converters or analog to digital converters digitize the voltage for computer-assisted processing of the data.

The accurate Faraday cup detector must neutralize the incoming singly charged particle with exactly one electron. However, due to the energy of the ion (2 to 10 keV), the surface of the detector is sputtered and secondary electrons and ions are ejected. These must be suppressed to prevent a false measurement of the ion current. Therefore, Faraday cup detectors are mostly thin, deep boxes with graphite surfaces and an electric field (~100 V) at their entrance to force secondary electrons back into the cup. The noise of the detector system can be no greater than $1.5 \times 10^{-16}$ A or 1000 ions per second in order to measure the ratio of $^1H^2H$ to $^1H_2$ with a precision of 0.01% (where a sample of ocean water contains about 156 ppm $^2H$ and the $^1H_2^+$ signal is $5 \times 10^{-9}$ A).

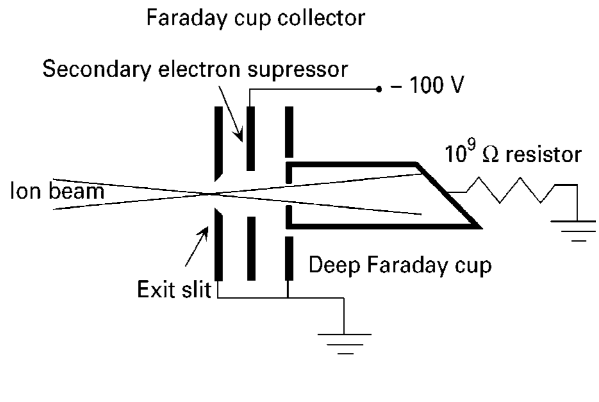



**Figure 3** Schematic layout of a Faraday cup assembly. The cups are deep and have secondary electron suppression in order to precisely record one charge per incoming singly charged ion. Different resistor values are used depending on the ion current in question.

As suggested by the name isotope ratio mass spectrometry, more than one ion current is measured in order to determine isotope abundance ratios. Most modern isotope ratio mass spectrometers have from 3 to 8 collectors carefully positioned along the focal plane of the instrument to measure several isotopic ion currents simultaneously, typically a few mass units apart. This configuration is known as simultaneous collection and has several advantages over designs employing a single detector. Since multiple detectors are measuring the different isotopic ion currents at the same time, any changes in the total ion current affect the signal in all detectors. Fluctuations in the production of ions by the ion source effectively cancel and the stability of the ion current is not a limiting factor.

### Ion source

Two ion production methods are employed for isotope abundance ratios. Thermal ionization is best suited for the analysis of nanogram quantities of metals. Electron impact is suitable for the measurement of gaseous samples and is the most widely used ionization process.

**Thermal ionization** Thermal ionization sources produce ions by heating sample-coated metallic filaments. Samples in solution are first deposited as microlitre drops on the filaments and dried at low temperatures forming thin salt or oxide layers (**Figure 4**). The prepared filament is transferred into the mass spectrometer source, the source is evacuated, and the filament is then heated to temperatures ranging from 800 to 2000°C. The probability of forming an ion in thermal equilibrium with the surface of the hot metal filament depends on the temperature (T), work function of the filament (W), and the ionization potential associated with the production of a given ion species (IP). This is described by the Langmuir-Saha equation as the relative number of ionized particles ($N^+$) to neutral species ($N^0$) evaporated from the filament (Eqn [4])

$$\frac{N^+}{N^0} \propto \exp\left[\frac{(W - \mathrm{IP})}{kT}\right] \qquad [4]$$

Typically, a filament material is chosen such that the work function of the metal is higher than the ionization potential of the sample. The filament is just heated to temperatures that allow efficient ionization without melting the metal.

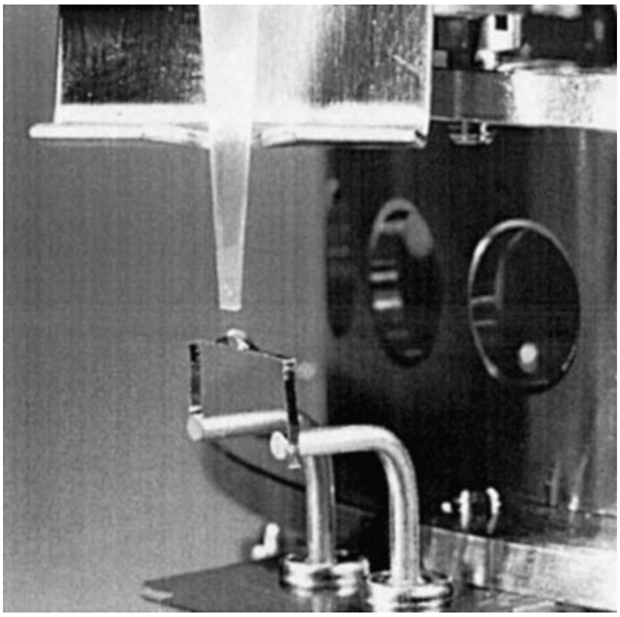


**Figure 4** Deposition of sample on a Re filament for thermal ionization analysis.

Using similar techniques, it is also possible to produce and analyse negative ions. This method has the advantage of allowing a number of elements with high ionization potentials to be more easily ionized as negative atomic or oxide species (i.e. B, Se and Re). Again, the number of ions relative to the number of neutral species produced can be described by the Langmuir- Saha equation. In this instance, the relative number of negative ions ($N^-$) produced depends on the electron affinity of the sample (EA) as well as the filament's temperature (T) and work function (W) (Eqn [5]:

$$\frac{N^-}{N^0} \propto \exp\left[\frac{(\mathrm{EA} - W)}{kT}\right] \qquad [5]$$

Mass-dependent fractionation processes that occur during the ionization process limit the ultimate accuracy and precision of isotope ratio measurements. As the sample is heated and ions are formed, the lighter isotopic species will evolve from the filament at a faster rate. The remaining (unionized) sample becomes relatively more enriched in the heavier isotopes and no longer has a representative isotope composition. Corrections to this fractionation are based on empirical calculations according to exponential or Rayleigh distillation models.

Principal applications of thermal ionization mass spectrometry (TIMS) are in the earth sciences to measure radiogenic isotopes for age determinations (i.e. U-Pb, Pb-Pb, Rb-Sr, Re-Os, Nd-Sm), absolute isotope abundance measurements, tracing of Pb and other metals in the environment, and isotope dilution mass spectrometry. The high efficiency of the ion source for many elements (i.e. 1 ion collected for every 200 rubidium atoms on the filament) allows very small samples ($10^{-9}$ to $10^{-12}$ g) to be analysed.

**Electron impact** Electron impact sources ionize gases by collisions with transverse electron beams of up to 1 mA current and 100 eV energy (**Figure 5**). The mostly singly charged ions produced are extracted from the ionization region by an applied electric field that further serves to collimate and accelerate the ions. The resulting ion current depends on the amount of sample gas inside the ion source. The influx of the sample gas into the ion source is of the order of $10^{-12}$ mol $s^{-1}$ under viscous flow conditions. The electron impact ion source will result in 1 ion being collected for every 1000 to 10 000 molecules introduced, depending upon the sensitivity of the ion source and the focusing properties of the mass spectrometer. Collisions in the ion source between ions and neutral species can lead to a nonlinear dependence of the ion current ratios with the source pressure. To prevent this, ions must be extracted immediately after formation. For isotope abundance measurements, simple gases (e.g. $CO_2$, $N_2$, $H_2$, $SO_2$, CO) are prepared from raw materials being careful to preserve the isotope composition of the sample.

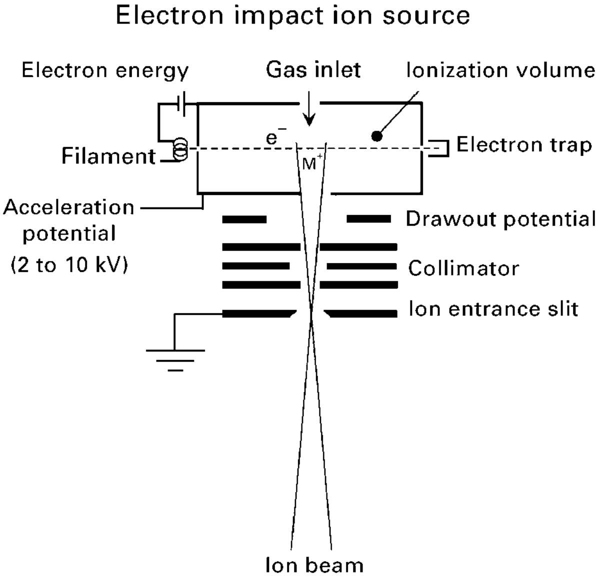



**Figure 5** Schematic diagram of electron impact ion source for the ionization of gaseous samples. The whole ion source is closed, leaving only small holes for the electrons to enter and the ions to exit. This ensures the highest sensitivity by having the gas molecules pass through the ionization region several times before being pumped away.

*Inlet systems: the changeover concept* Mass spectrometers do not measure isotope abundance ratios accurately due to inherent biases caused by the ion optical design, ion current detectors and electronics, and method of sample preparation. In practice, to correct for machine bias, the measured ion current ratios of the sample are compared to those from a reference gas (of known isotopic composition) analysed under identical operating conditions. This can be achieved using a dual inlet system with two stainless steel variable volumes (1 to 100 mL) for reference and sample gases (Figure 1). The volume of these reservoirs is adjusted to change the pressure of the gas (10 to 1000 mbar) and obtain matched major ion current intensities of both reference and sample gases. Each measurement of the sample ion currents is immediately followed by an analysis of the standard's ion currents.

Stainless steel capillaries (1 m length and 0.1 to 0.2 mm i.d.) transport the gas from the variable volumes to the ion source. The changeover valve alternately connects the reference and sample bellows to either the ion source or vacuum and the isotope compositions of the two gases are measured in turn. This method ensures a constant flow of gas through the capillaries at all times. To maintain viscous flow conditions, a pressure of ~20 mbar in the inlet system is required. This constitutes a lower limit of gas suitable for analysis with the dual inlet system. The smallest practical volume of an inlet system is ~250 μL. This represents 200 nmol of gas at 20 mbar. In some cases, there is not enough sample material available to produce sufficient quantities of gas and alternative analytical methods must be employed, e.g. isotope ratio monitoring.

## Delta (δ) values

In order to realize inter- and intra-laboratory consistency, measured isotope abundance ratios of the sample and reference gases are compared using a δ (delta) scale. In the case of carbon isotope ratio measurements, δ values are calculated as shown in Equation [6]. The δ values for other elements are calculated based on the isotope abundance ratios listed in Table 1.

$$\delta^{13}\,\mathrm{C} = \left[\frac{(^{13}\mathrm{C}/^{12}\mathrm{C})_{\mathrm{sample}}}{(^{13}\mathrm{C}/^{12}\mathrm{C})_{\mathrm{reference}}} - 1\right] \times 1000 \qquad [6]$$

This quantity is the deviation of the isotope abundance ratio of the sample from that of a standard in parts per thousand or per mill (‰). Positive δ values indicate that the sample is relatively enriched in the heavier isotope of the element with respect to the reference. Negative δ values represent samples that are relatively depleted in the heavier isotope compared to the reference material. The National Institute of Standards and Technology (NIST) in Gaithersburg, Maryland, USA and the International Atomic Energy Agency (IAEA) in Vienna, Austria both distribute and continuously calibrate isotope reference materials for a number of elements. The reference materials chosen are representative chemical compounds and reflect nominal terrestrial isotope abundances.

**Table 1** The isotopes that are most commonly compared for the determination of δ values for the elements H, C, N, O and S

| *Element* | *Isotope abundance ratio* | *δ value* |
|---|---|---|
| Hydrogen | D/H or $^2H/^1H$ | δΔ or $\delta^2H$ |
| Carbon | $^{13}C/^{12}C$ | $\delta^{13}C$ |
| Nitrogen | $^{15}N/^{14}N$ | $\delta^{15}N$ |
| Oxygen | $^{18}O/^{16}O$ | $\delta^{18}O$ |
| Sulfur | $^{34}S/^{32}S$ | $\delta^{34}S$ |

## Analysis of H, C, N, O and S isotope ratio variations

### Hydrogen and oxygen

Hydrogen exhibits the largest natural isotope abundance variations due to the relatively large mass difference between its two stable isotopes; $^1_1H$ and $^2_1H$ (also written commonly as H and D, respectively). H is approximately 99.985 % and D 0.015 % naturally abundant. The three isotopes of oxygen $^{16}_8O$ $^{17}_8O$ and $^{18}_8O$ are ~99.763%, ~0.0375% and ~0.1995 % abundant, respectively. Examples of isotope abundance variations measured for H and O are summarized in Figures 6 and 7.

Hydrogen and oxygen combine to form water and the isotopic species of $H_2O$ vary in mass from 18 to 22 (i.e. $H_2{}^{16}O$, $HD^{16}O$, $D_2{}^{18}O$). The decreasing vapour pressure with increasing mass of the different isotopomers causes evaporated water vapour in equilibrium with the condensed phase to be relatively enriched in the lighter isotopes of both oxygen and hydrogen. Conversely, the condensation of water from the vapour phase results in a precipitate enriched in the heavier isotopes of O and H (higher $\delta^{18}O$ and δD values). The condensation of water vapour in equilibrium with its liquid phase - a situation that occurs in clouds - can be described by

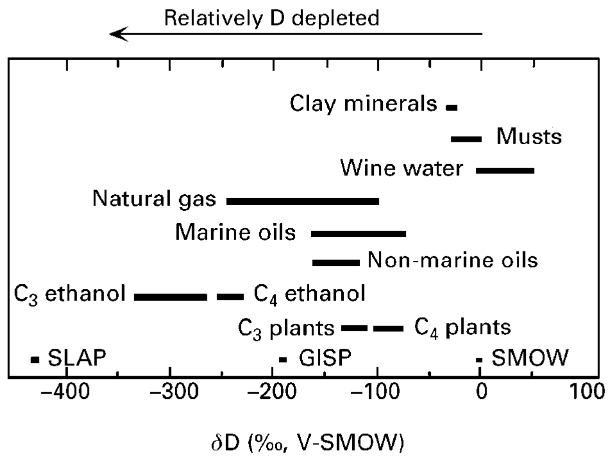



**Figure 6** Natural variations in hydrogen isotope abundances. V-SMOW, SLAP, and GISP are international reference water samples available from the International Atomic Energy Agency (IAEA) in Vienna, Austria.
V-SMOW = Vienna Standard Mean Ocean Water
SLAP = Standard Light Antarctic Precipitation
GISP = Greenland Ice Standard Precipitation.

the Rayleigh fractionation law (Eqn [7]).

$$R/R_0 = f^{\alpha-1} \qquad [7]$$

Here, $R_0$ is the initial $^{18}O/^{16}O$ ratio of the gas phase, R is the actual ratio as a function of the remaining molar fraction in the gas phase (f). The term $\alpha_{\text{liquid-vapour}}$ is the fractionation factor between the liquid and vapour phases at a given temperature. With the increasing amount of $^{18}O$-enriched liquid formed and removed from the vapour, the gas phase and liquid phase formed from it becomes isotopically lighter. The $\delta^{18}O$ values for liquid and vapour as a result of this process are plotted in Figure 8. As the precipitation event continues, the immediate removal of the condensed phase from the cloud without any isotope exchange or reevaporation continuously depletes the vapour in D and $^{18}O$. As a result, precipitation formed at higher latitudes becomes progressively more depleted in the heavier isotopes of hydrogen and oxygen.

Because the extent of isotope fractionation between the condensed and vapour phases is temperature dependent, variations in the δD and $\delta^{18}O$ values recorded at a particular site reflect seasonal and possibly long-term fluctuations. Ices cores from Greenland, Antarctica, and elsewhere are excellent archives of past climate change, the information recorded mainly in the $^{18}O/^{16}O$ ratio of the ice. $\delta^{18}O$ measurements indicate that very rapid changes from cold to warm and vice versa within a decade have occurred over the past 160 000 years. Such data have challenged our understanding of the underlying processes and caught the attention of the media because of the social and political implications changing climatic events could have today.

Hydrogen and oxygen are also found in many minerals and are fractionated during geochemical processes. The hydrogen and oxygen isotope compositions of the mineral are controlled by the isotope

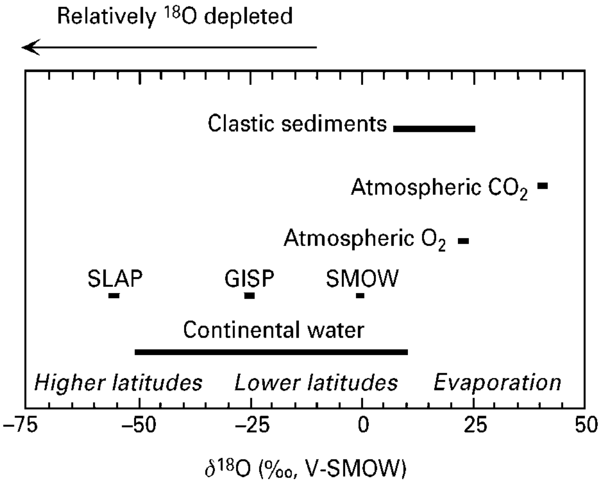



**Figure 7** Natural variations in oxygen isotope abundances. The $\delta^{18}O$ values in ancient precipitation (Greenland Ice, Antarctic Ice, etc.) and in foramineferal carbonate from sediments provide an excellent record of the temperatures of the past.

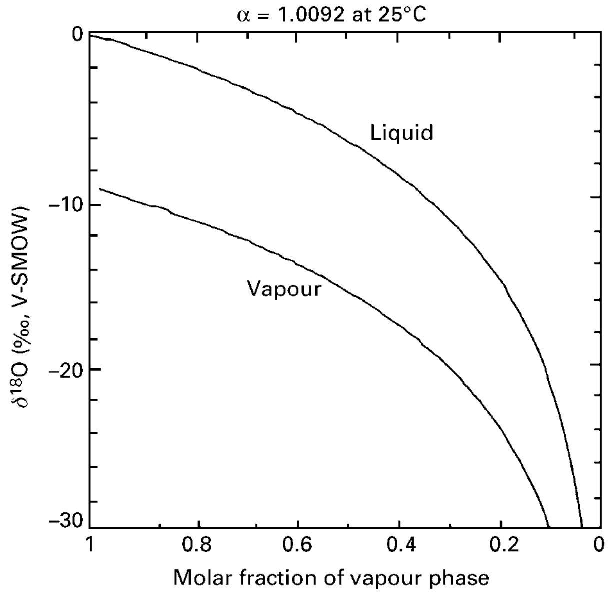



**Figure 8** Rayleigh fractionation during water condensation. Condensation starts with the liquid enriched in $\delta^{18}O$ by 9.2‰ over the gas phase. As the condensation process proceeds, the remaining vapour becomes lighter and, consequently, also the precipitate formed from it.

composition of the fluid and the fractionation factor appropriate for the temperature during formation. For example, the $\delta^{18}O$ value of oxygen in $CaCO_3$ differs from that of the water as the carbonate precipitates under equilibrium conditions. The extent of fractionation is temperature dependent and at 25°C the mineral is enriched in $^{18}O$ by about 28‰ compared to the water. The $\delta^{18}O$ value of carbonates and fluids are used to assess water-rock interactions. Whereas the $\delta D$ value of hot spring waters may be similar to that of local precipitation, the $\delta^{18}O$ value may be enriched in $^{18}O$ indicating oxygen isotope exchange with nearby rocks.

Standards The standard for both hydrogen and oxygen is V-SMOW (Vienna Standard Mean Ocean Water) distributed by the IAEA in Vienna, Austria. A secondary standard often employed is SLAP (Standard Light Antarctic Precipitation) which has a $\delta D$ values of -428‰ and $\delta^{18}O$ values of -55.5‰ with respect to V-SMOW.

Sample preparation Hydrogen isotopes in water are measured from $H_2$ gas formed by reducing water with a suitable reducing agent like U, Zn, Cr or C. It is also possible to exchange hydrogen isotopes between $H_2$ gas and $H_2O$ in the presence of Pt catalysts.

Ratios of HD to $H_2$ (mass 3 to 2) of the sample hydrogen gas are measured and reported as $\delta D$ values. Ion-molecule reactions in the mass spectrometer source will result in the production of $H_3^+$, which is isobaric with $HD^+$ at mass 3. The number of $H_3^+$ ions produced is proportional to the $H_2$ pressure and to the number of $H_2^+$ ions in the source, necessitating an empirical correction factor to the measured ratios.

Oxygen isotope compositions of water samples are determined from the measurement of $^{12}C^{16}O_2^+$ and $^{12}C^{16}O^{18}O^+$ at masses 44 and 46, respectively. $CO_2$ gas is equilibrated with 5 mL of water at 25° C for 8 h and isotope exchange between $CO_2$ gas and water makes the oxygen isotope composition of the $CO_2$ gas reflect that of the water.

### Carbon

Carbon has two stable isotopes ($^{12}_{6}C$ and $^{13}_{6}C$) of ~98.89% and ~1.11% abundance, respectively. Carbon compounds are exchanged between the oceans, atmosphere, biosphere and lithosphere. Significant isotope ratio variations exist within these groups due to both kinetic and equilibrium isotope effects. Representative carbon isotope abundance variations are shown in Figure 9.

During photosynthesis, the metabolized products become depleted in $^{13}C$ compared to the $CO_2$. The initial assimilation and intracellular diffusion of

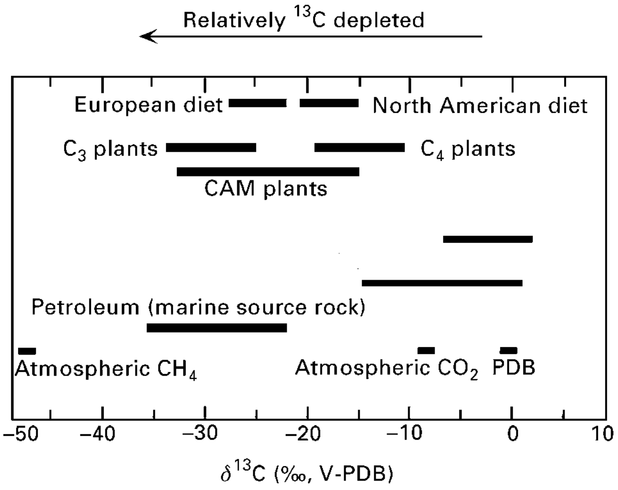



**Figure 9** Natural variations in carbon isotope abundances. North American diet to a large extent is derived from maize, a $C_4$ plant. In countries like Japan, where fish is a major part of the average diet, $\delta^{13}C$ values are in between the European and North American values.

$CO_2$ in the plant produces a 4‰ depletion in its $\delta^{13}C$ value. Isotope fractionation during the decarboxylation of the absorbed $CO_2$ is dependent on the plant species and causes a much larger shift in the $\delta^{13}C$ value. Plants that form a 3-carbon phosphoglyceric acid as the first product of carbon fixation according to the Calvin cycle (commonly referred to as $C_3$ plants) have $\delta^{13}C$ values averaging -28‰ V-PDB. Examples of $C_3$ plants include cereal grains, peanuts, rice, tobacco, beans, sugar beets and all evergreen and deciduous trees. Plants that photosynthesize dicarboxylic acid by the Hatch-Slack mechanism (referred to as $C_4$ plants) are more enriched in $^{13}C$ and have $\delta^{13}C$ values closer to − 13% V-PDB. $C_4$ plants include corn, millet, sorghum, sugar cane and many grasses such as crab grass and bermuda grass. Crassulacean acid metabolism (CAM) plants have intermediate $\delta^{13}C$ values. This arises because they metabolize $CO_2$ via $C_3$ pathways in light and $C_4$ in darkness. The large isotope fractionations that arise in plants are generally retained as the organic material is incorporated into sediments and eventually as hydrocarbon deposits.

The carbon isotope composition of food ingested by animals is generally reflected in the metabolized products. For example, it is possible to differentiate between animals that have primarily $C_3$ versus $C_4$ diets based on the measurement of the $\delta^{13}C$ value of hair, finger nails, bone collagen, etc. This information is useful in food web studies to determine primary nutrient sources and to discover who is the prey and who the hunter. Nitrogen isotopes are

often combined with $\delta^{13}C$ data for a comprehensive trophic level analysis.

Oceans play a significant role in the exchange of $CO_2$ between the atmosphere and biosphere. Carbon isotope exchange between $CO_2$ in the air and $HCO_3^-$ in the oceans results in $^{13}C$ enrichment of the latter. $CO_2$ in air is well mixed globally and currently has a $\delta^{13}C$ value of -8‰ V-PDB compared to marine limestones at ~0‰ V-PDB. Freshwater carbonates have much more variable and lower $\delta^{13}C$ values due to the oxidation of $^{13}C$ depleted organic matter and subsequent formation of $HCO_3^-$.

The $CO_2$ content of the atmosphere has increased dramatically due to the combustion of fossil fuels and oxidation of soil and organic matter. Fossil fuels are generally depleted in $^{13}C$ because they were derived from organic material. Consequently, the global $\delta^{13}C$ value of atmospheric $CO_2$ has seen a change from 6.5‰ in preindustrial times to about 8.0‰ today.

Standards The standard for carbon isotope abundance measurements is based on a Cretaceous belemnite sample from the Peedee formation in South Carolina, USA. The original material is no longer available. It has been replaced by the convention that NBS 19, a carbonate material, has a value of +1.95‰ versus PDB. This new scale is termed V-PDB (Vienna-PDB). The IAEA distributes a number of secondary standards including graphite (USGS24) with a $\delta^{13}C$ value of -15.99‰ V-PDB, oil (NBS-22) at -29.74‰ V-PDB, and calcium carbonate (NBS-18) with a value of -5.01‰ V-PDB.

The oxygen isotope composition of carbonates is also commonly referenced to the $^{18}O/^{16}O$-isotope ratio of V-PDB. It is not used as a reference for $\delta^{18}O$ analyses of igneous or metamorphic rocks. The difference between V-SMOW and V-PDB $\delta^{18}O$ values is approximately 30‰ ($\delta^{18}O_{V\text{-}SMOW} = 1.03091 \times \delta^{18}O_{V\text{-}PDB} + 30.91$).

Sample preparation $CO_2$ gas is produced from carbonate minerals by reaction with 100% $H_3PO_4$ or by thermal decomposition. The carbon and oxygen isotope compositions of the carbonate can be determined simultaneously from the $CO_2$ produced. $\delta^{13}C$ values from organic samples are analysed by combusting the material at high temperature with an oxygen source (commonly CuO) in sealed quartz tubes. Ion currents at mass 44, 45 and 46 produced from $CO_2$ gas are measured in order to determine a $\delta^{13}C$ value. The isobaric interference at mass 45 from $^{17}O$ (i.e. both $^{13}C^{16}O^{16}O^+$ and $^{12}C^{17}O^{16}O^+$ will be collected by the same Faraday detector) requires an empirical correction to enable the calculation of the $^{13}C/^{12}C$ ratio from the measured mass 45 ($^{17}O$ correction).

## Nitrogen

Nitrogen is present primarily as $N_2$ gas in the atmosphere and dissolved $N_2$ in oceans. In terrestrial systems, nitrogen is found in minor amounts bound to H, C and O in both reduced and oxidized states. Nitrogen has two stable isotopes $^{14}_{7}N$ and $^{15}_{7}N$ of ~99.64% and ~0.36% abundance, respectively. Isotope abundance variations measured for nitrogen are summarized in Figure 10.

Biochemical reactions mediated by bacteria are responsible for many nitrogen isotope fractionations. The fixation of $N_2$ gas to compounds such as $NH_3$ at the nodules of plant roots requires substantial energy to break the $N_2$ bonds and results in negligible isotope fractionation. However, the conversion of ammonia to nitrates (nitrification) initially to $NO_2^-$ by Nitrosomonas ssp. and then to $NO_2^-$ by Nitrobacter yields nitrates that can have $\delta^{15}N$ values 20‰ more negative than the ammonia. The bacterial conversion of nitrates to $N_2$ under anaerobic conditions (denitrification) by Pseudomonas denitrificans or Thiobacillus denitrificans enriches the $^{15}N$ content of the residual nitrate increasing the $\delta^{15}N$ value by as much as 30‰ as the reaction progresses. Generally, the organic matter of the biosphere is enriched in $^{15}N$ compared to the atmosphere and $\delta^{15}N$ values for marine environments are greater than those measured in continental regions.

Increasing populations and agricultural activities have a significant influence on the amount of nitrate in groundwater. Within a given region, nitrate concentrations and $\delta^{15}N$ values can be used to trace the source of anthropogenic nitrogen compounds into a system and monitor its occurrence over time.

The relationship between the nitrogen isotope composition of an animal and its position in the food

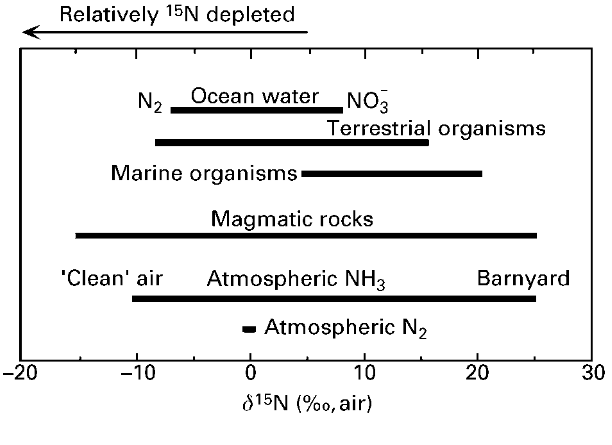



**Figure 10** Natural variations in nitrogen isotope abundances. $\delta^{15}N$ values are often used to establish the position of a particular species within the food chain of an ecosystem (trophic level).

chain is such that one observes an increase in the $\delta^{15}N$ value of approximately 3‰ per trophic level. Compared to the food source, excreted nitrogen is relatively depleted in $^{15}N$ due to fractionation during urea formation. The analysis of contemporary and archaeological specimens of feathers, bone, hair or skin can provide information about the diets of people and animals and changes that may have occurred in response to past climatic or historical events.

Standards Nitrogen in the atmosphere is well mixed and $\delta^{15}N$ values are reported relative to the isotope composition of $N_2$ in air.

Sample preparation Nitrogen isotope abundances are determined by measuring ions at masses 28 ($^{14}N^{14}N^+$) and 29 ($^{14}N^{15}N^+$). Nitrogen compounds are converted to $N_2$ gas by high temperature combustion in an elemental analyser. $N_2$ gas introduced into the ion source must be free from CO or $CO_2$ that produce an isobaric interference at masses 28 and 29.

### Sulfur

Sulfur has four stable isotopes; $^{32}_{16}S$, $^{33}_{16}S$, $^{34}_{16}S$ and $^{36}_{16}S$ which are ~95.02%, ~0.75%, ~4.21% and 0.02% abundant, respectively. Sulfur is ubiquitous in the global environment and major sulfur reservoirs include sulfate in the oceans, evaporites, sulfide ore deposits, and organic sulfur compounds. Sulfur isotope abundance variations are shown in Figure 11.

Large kinetic isotope effects occur during the reduction of sulfate by anaerobic bacteria such as Desulfovibrio ssp. resulting in the product sulfide having $\delta^{34}S$ values ranging from +3 to -46‰ with respect to the remaining sulfate. Another potential mechanism for sulfur isotope fractionation includes exchange reactions among sulfur compounds with S in different valence states. For example, sulfur isotope exchange between sulfate and sulfide (and resulting change in oxidation state from +6 to −2) is predicted to yield fractionations as large as 75‰ between the $H_2S$ and $SO_4^{2-}$ at 25°C. However, sulfur isotope exchange between sulfates and sulfides proceeds very slowly and probably occurs at high temperature where the equilibrium constant is close to unity. Sulfur oxidation, mineral precipitation, and high temperature processes result in less extensive isotope fractionations.

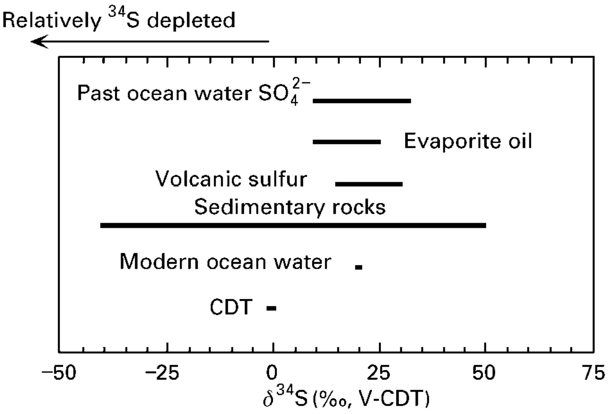



**Figure 11** Natural variations in sulfur isotope abundances. Large sulfur isotope fractionations are possible during the reduction of sulfates to sulfides by anaerobic bacteria that produce $^{32}S$ enriched $H_2S$ relative to the original sulfate.

Sulfur isotope compositions are employed to assess the effect of industrial activities in pristine environments, particularly in regions where the pollutant sulfur has a distinct $\delta^{34}S$ value compared to that of the environmental receptors. For example, in the province of Alberta, Canada, emissions from sour gas processing plants have $\delta^{35}S$ values 20‰ higher compared to sulfur isotope compositions of unaffected soils. The distinct isotope composition of the anthropogenic sulfur allows it to be followed as it enters the environment via biogeochemical reactions.

Standards The international sulfur isotope scale is based on a troilite sample from the Canyon Diablo meteorite, termed CDT. The International Atomic Energy Agency (IAEA) in Vienna has promoted the use of a more accurate sulfur isotope scale, V-CDT, to replace CDT because of isotopic inhomogenities of ±0.4‰ in the original sample. A new $Ag_2S$ standard, IAEA S-1, has a defined $\delta^{34}S$ value of -0.30‰ V-CDT.

Sample preparation Sulfur isotope abundances are measured commonly from isotopic ion currents of $SO_2^+$ at masses 64, 65 and 66. $SO_2$ gas is generated by reacting sulfur or sulfide minerals with $V_2O_5$ and $SiO_2$ in a ratio of 1 : 10 : 10 in sealed quartz tubes heated to 900°C. Isobaric interference from oxygen isotopes necessitates the preparation of $SO_2$ from standards and samples with identical oxygen sources. $SO_2$ is easily adsorbed onto glass and metal surfaces and care is required in its handling to avoid fractionation and memory effects among different samples. An alternative (and lesser used) approach is to form the more stable $SF_6$ gas and measure $SF_5^+$ ions.

## Isotope ratio monitoring

A recent analytical technique for high precision, high sensitivity isotope ratio measurements is isotope ratio

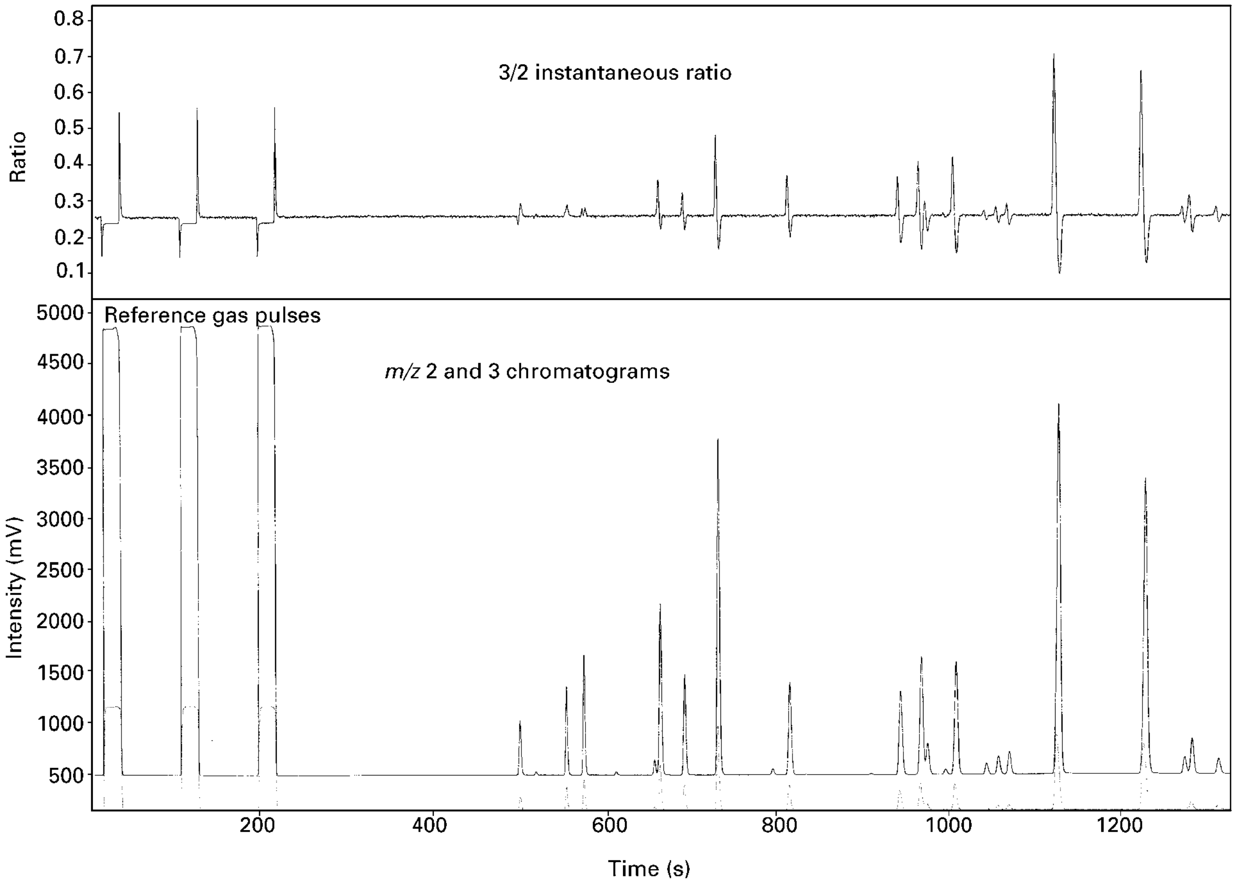



**Figure 12** *irm*-GCMS chromatogram from a crude oil extract. The isotope ratio of deuterium versus hydrogen from a post column high temperature (1400°C) pyrolysis reaction is studied. The lower trace shows the mass 2 intensity whereas the upper trace exhibits the instantaneous intensity ratio of 3/2 as a function of time. The deuterated compounds elute earlier than the nondeuterated ones with a time shift between 1 and 5 s. This gives rise to the marked isotope swing in the upper 'ratio' trace. The first 3 peaks are reference gas pulses injected behind the GC, thus they do not exhibit this behaviour. Here, differences in the amplifier time constants give rise to the sharp edges displayed on the ratio trace.

monitoring (irm). The irm instrument is a further development of the classical isotope ratio mass spectrometer including improvements in source design, vacuum pump technology, high speed detection electronics and powerful personal computers for data processing. This method for isotope analysis does not employ a dual inlet (Figure 1) and hence there is no requirement to generate large amounts of sample gas to maintain viscous flow conditions in the capillaries. Instead, isotope abundance data are extracted from time-varying ('transient') signals as opposed to measurements of prepared gases stored in reservoirs that produce near constant signal intensities. The entire sample entrained in a helium carrier gas enters the ion source as a transient and is ionized for isotope abundance measurements. A typical irm chromatogram for D/H analysis from GC eluates is shown in Figure 12. All ion currents of interest are integrated as the He carrier sweeps the sample through the source. A pulse of reference gas is admitted to the source and its isotope ratio determined in a similar manner. From the two measurements the $\delta$ value of the respective sample peak is calculated. Advantages of the irm technique include the smaller amount of sample that must be prepared for analysis (1 nmol compared to 200 nmol). In addition, reduced sample handling and preparation is possible because raw materials are converted to simple gases for analysis by automated sample preparation devices (such as elemental analysers). It should be noted that there are a number of terms in current use in the literature to describe this method, including continuous flow isotope ratio mass spectrometry or CF-IRMS.

## Configuration of *irm* inlet systems

The irm system consists of three components; the sample preparation device, isotope ratio mass spectrometer and the gas and electronic interface between the two. Depending on the nature of the

sample and type of analysis required, there are two approaches to assembling an irm inlet system (Figure 13). For isotope abundance ratio determinations from bulk samples, the entire sample is either oxidized or pyrolysed, producing a mixture of reaction products ($CO_2$, $N_2$, $SO_2$ and $H_2O$ for the oxidation, CO and $H_2$ for the pyrolysis) that must be separated prior to introduction into the mass spectrometer. In contrast, compound specific analysis is the separation of a particular component from some complex mixture followed by conversion to a gas suitable for isotope analysis (by oxidation or pyrolysis) and subsequent introduction into the mass spectrometer ion source. Here, reaction products are isolated by chemical means on-line when possible.

The amount of sample and carrier gas that enters the ion source is limited to a maximum of about

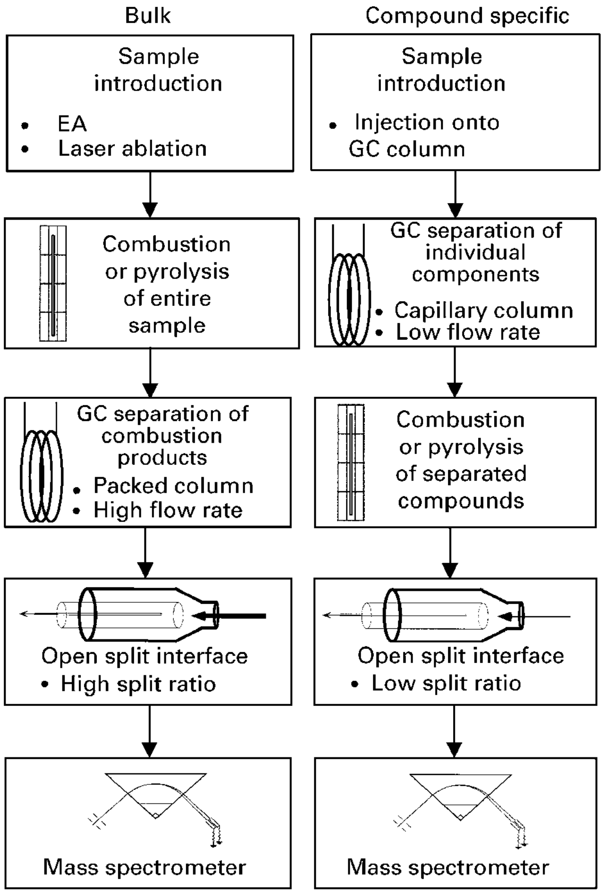



**Figure 13** Comparison of isotope ratio monitoring (*irm*) sample preparation and introduction for bulk versus compound specific sample analysis. In bulk sample analysis, the GC follows the combustion step whereas in the compound specific *irm* system $CO_2$ and $N_2$ from a single peak both enter the mass spectrometer at the same time. In both cases the amount of gas flowing to the ion source is comparable. The difference in sensitivity comes from the difference in carrier gas flow rate (about 100 ml $min^{-1}$ versus ~1.5 ml $min^{-1}$).

0.5 mL $min^{-1}$. Therefore, only a portion of the gas from the sample preparation device is admitted, requiring an open split. A high gas flow from the preparation system will require a higher split ratio and reduce the relative amount of sample available for ionization. To maximize sensitivity the split ratio should be as low as practically possible thereby maximizing the amount of sample ionized.

Bulk composition isotope ratio monitoring The determination of isotope compositions from bulk samples is one of the most widespread applications of irm techniques primarily because of its ease of use, possibility for automated analysis, and the variety of sample types that can be analysed. Raw samples are most often converted to simple gases by an elemental analyser (EA) where submilligram quantities of sample are packed into Sn or Ag capsules ($2 \times 5$ mm) and loaded into the autosampler of the EA. Carrier gas flow rates are typically 80 to 120 mL $min^{-1}$, necessitating a high split ratio at the mass spectrometer. Individual samples are dropped into an oxygen-enriched He stream sweeping a reactor at 1050°C where they are combusted instantaneously. A second copper-filled furnace at 600°C in-line reduces nitrogen oxides produced in the combustion process to $N_2$ gas. It also scavenges excess oxygen not used in the combustion. Water is removed from the combustion products and the $CO_2$, $N_2$ and $SO_2$ (if any) are separated by a packed column (Figure 14).

Compound specific isotope ratio monitoring The isotope composition of individual components constituting a complex mixture can be determined by

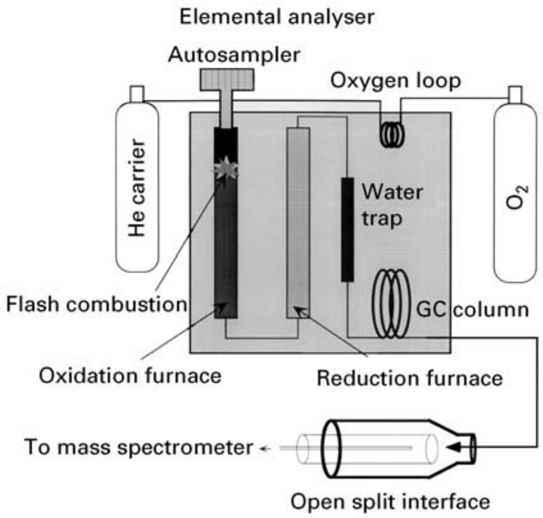



**Figure 14** Bulk sample isotope ratio monitoring interface. The means for reference gas injection and for effluent dilution are not shown.

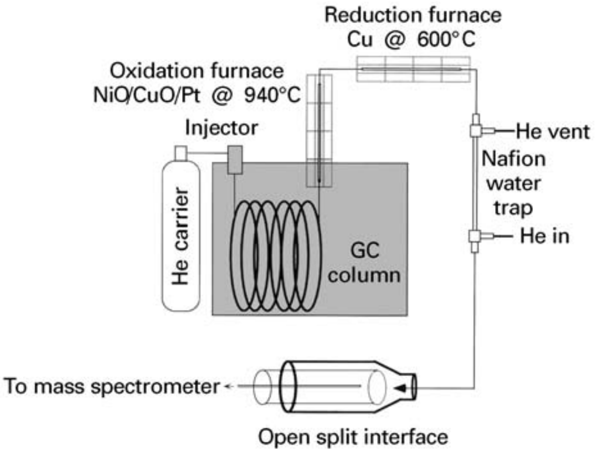



**Figure 15** Compound specific isotope ratio monitoring interface. For clarity, the devices for backflushing the solvent, for effluent diversion from the ion source, and for reference gas injection are not shown.

separating out the individual fractions followed by conversion into simple gases suitable for analysis. The basic arrangement is a gas (or liquid) chromatograph followed by an oxidation (or pyrolysis) furnace and open split to the mass spectrometer (Figure 15). A capillary chromatographic column first separates the mixtures.

For oxidation and carbon or nitrogen isotopic analysis the components are then swept sequentially through a furnace converting them to $N_2$, $CO_2$ and $H_2O$. Water is removed and $CO_2$ is removed cryogenically if $\delta^{15}N$ measurements are to be made. For $\delta^{13}C$ analysis alone, nitrogen, if present, also enters the ion source. It does not interfere with the measurement.

For pyrolysis at high temperatures ($\delta D$ or $\delta^{18}O$ analysis) the gaseous reaction products $H_2$ and CO are allowed to enter the mass spectrometer simultaneously. For D/H it is possible to remove CO cryogenically using, for instance, a piece of capillary molecular sieve column immersed in liquid nitrogen.

The low carrier gas flow (<3 mL min$^{-1}$) allows for a low split ratio or high sensitivity. This technique termed irm-GCMS is sometimes referred to as gas chromatography-combustion mass spectrometry (GCCMS).

### Characteristics of the ***irm*** mass spectrometer

The irm mass spectrometer must be able to handle the resulting high pressures in the ion source region ($10^{-5}$ mbar) while processing data from transient signals with precisions comparable to, or better than, dual-inlet performance. Differential pumping of the source and analyser regions enables the ion source of the irm instrument to operate under a high continuous He load.

To measure nanomoles of gas, the mass spectrometer ion source must maximize the number of ions detected per number of gas molecules introduced. Typical irm instruments are capable of producing 1 ion for every 1500 molecules admitted to the source. In practice, this is achieved by using a ‘closed’ source design and allowing a large angle of the ion beam to enter the analyser region of the mass spectrometer. The ion optical design of the instrument ensures that the vertical height of the ion beam is identical at the entrance and exit slits so that loss of ions due to collisions with the walls of the flight tube is minimized.

The transient signals encountered in irm mass spectrometry vary over several orders of magnitude during an analysis. In order to measure accurate isotope abundance ratios, the measured ion current ratios must not depend on the ion current intensities. This characteristic is defined as source linearity. Note that with dual-inlet systems, major ion current intensities of the sample and reference gases were matched by varying the pressure of the gas in the ion source to counter errors introduced by nonlinear source behaviour. Two major causes of nonlinear source behaviour are the buildup of space charge due to the large numbers of He atoms in the source and the reaction of neutrals with protonating agents leading to isobaric interference.

The Faraday cups receive a constant stream of time-varying ion currents that must be processed without loss of data. This capability was only possible recently with the development of high speed, low cost personal computers. Ion currents are converted to voltages by high gain, low noise operational amplifiers. The voltages are then digitized into signals that can be processed by a computer. Typically, the voltage is integrated over 0.25 s time slices and converted to a frequency pulse such that the number of pulses within the interval is directly proportional to the measured ion current. The time constants of the ion current amplifiers (i.e. how quickly the device responds to a changing signal) are optimized and matched for all detectors.

### Examples of ***irm*** applications

Doping in athletics  Testosterone is a potent performance-enhancing drug that is used in both human and equestrian sports. The current method to detect testosterone abuse is based on measurement of the testosterone to epi-testosterone ratio using GC or LC/MS techniques. However, this method is not reliable because some ‘clean’ individuals may naturally register a ‘positive’ result and the measurement must

be made no more than 5 h after ingestion of the drug. Successful detection of drug abuse by irm is possible because commercially available testosterone has a lower $\delta^{13}C$ value than that metabolized by the body (approximately 5‰ lower). The isotope ratio determination of a subject's testosterone and metabolites by irm-GCMS can detect this shift in $\delta^{13}C$ values and detect testosterone abuse 24 h after the initial dose.

Food adulteration Quality assurance of food products is an important activity in the food industry where the substitution of lower cost alternatives constitutes fraud as well as a public health risk. Stable isotope measurements have been successfully applied to many instances where 'all-natural' fruit juices, honey and maple syrup have been artificially sweetened. This is possible when the added sugar is from a plant species which follows a different photosynthetic pathway from that of the bulk material ($C_3$ versus $C_4$ plants), leading to a $\delta^{13}C$ composition of the food product intermediate to the 12‰ difference between the two plant types.

The determination of the $\delta^{13}C$ value of individual fatty acids of adulterated and unadulterated maize oils by irm-GCMS allows as little as 5% of an adulterant $C_3$ oil to be detected. More recently, the measurement of $\delta^{18}O$ value of olive oils that were pyrolysed over nickelized carbon demonstrated a geographical dependence between $^{18}O$ and latitude. Lower $\delta^{18}O$ values were measured for olive oils coming from Italy as compared to Greece, allowing the origin of the product to be verified.

Detection of Heliobacter pylorii The Heliobacter pylorii bacterium is found in the mucous membrane of the stomach and is the major cause of gastritis, gastric ulcers and cancer of the stomach and duodenum. It produces urease by cleaving urea to form ammonia and $CO_2$. The detection of this dangerous bacterium is possible by administering a small amount (100 mg) of $^{13}C$-labelled urea to the patient and measuring the $\delta^{13}C$ value of breath $CO_2$. If after 30 min an increase in the $\delta^{13}C$ of the breath-$CO_2$ is detected, the bacterium is present and actively forming $^{13}C$-enriched $CO_2$ from the labelled urea. This noninvasive procedure is simple to administer and is widely applied as a $^{13}C$ urea breath test by a number of medical laboratories.

Trace gas analysis in air Greenhouse gases such as methane, $CO_2$ and $N_2O$ are causing the temperature of the planet to increase, or they are interfering with ozone chemistry in the stratosphere. Isotope analysis of these gases may help to understand their cycling in the atmosphere or between global pools. For $CH_4$ and $N_2O$, dual inlet techniques require in excess of 70 L of air to achieve the precision (±0.2‰ for $\delta^{13}C$, $\delta^{18}O$ and $\delta^{15}N$) required to discriminate among different sources. Such large samples are difficult to process and are of low spatial and temporal resolution. To obtain meaningful, high resolution data from modern and archived air samples, irm techniques have been developed which are capable of analysing the isotope composition of $CH_4$ (1.7 ppm) and $N_2O$ (0.3 ppm) in 25 mL and 100 mL air samples, respectively.

## List of symbols

B = magnetic field flux density; h = Planck's constant; I = current; m = mass of molecule; N = number of parent nuclides; q = charge; R = path radius; t = time; T = temperature; W = work function; $\alpha$ = isotope fractionation factor; $\nu$ = frequency of oscillation; $\lambda$ = probability of decay; ‰ = per mille; $\delta^{13}C$, $\delta^{18}O$, $\delta D$, $\delta^{15}N$ and $\delta^{34}S$ = Deviations of isotope ratios from a reference.

*See also:* **Biochemical Applications of Mass Spectrometry; Cosmochemical Applications Using Mass Spectrometry; Food Science, Applications of Mass Spectrometry; Food Science, Applications of NMR Spectroscopy; Forensic Science, Applications of IR Spectroscopy; Forensic Science, Applications of Mass Spectrometry; Inductively Coupled Plasma Mass Spectrometry, Methods; Interstellar Molecules, Spectroscopy of; IR Spectroscopy, Theory; Isotopic Labelling in Mass Spectrometry; Mass Spectrometry, Historical Perspective; Medical Applications of Mass Spectrometry; Scattering Theory; Sector Mass Spectrometers.**

## Further reading


Bowen R (1988) Isotopes in the Earth Sciences. Essex, England: Elsevier.

Bowen R (1991) Isotopes and Climate. London: Elsevier.

Brand WA (1996) High precision isotope ratio monitoring techniques in mass spectrometry. Journal of Mass Spectrometry 31: 225–235.

Brand WA (1998) Isotope ratio mass spectrometry: precision from transient signals. In: Advances in Mass Spectrometry. Karajalainen EJ, Hesso AE, Jalonen JE and Karjalainen UP (eds), Vol 14, pp 655–680. Amsterdam: Elsevier.

Dickin AP (1995) Radiogenic Isotope Geology. Cambridge: Cambridge University Press.

Faure G (1986) Principles of Isotope Geology, 2nd edn. New York: Wiley.

Fritz P and Fontes J Ch (eds) (1980) *Handbook of Environmental Isotope Geochemistry*. Amsterdam: Elsevier.
Heumann KG (1988) Isotope dilution mass spectrometry. In: Adams F, Gijbels R and van Grieken R (eds) *Inorganic Mass Spectrometry (Chemical Analysis Series, vol. 95)*, pp 301-376. New York: Wiley.
Hoefs J (1997) *Stable Isotope Geochemistry (Minerals and Rocks; 9)*, 4th edn. Berlin: Springer-Verlag.
Krouse HR and Grinencko VA (eds) (1991) *Stable Isotopes: Natural and Anthropogenic Sulphur in the Environment (Scope/Scientific Committee on Problems of the Environment, No 43)*. Wiley.
Platzner IT (ed) (1997) *Modern Isotope Ratio Mass Spectrometry*, Vol 145 in Chemical Analysis. Chichester: Wiley.

# Isotopic Labelling in Biochemistry, NMR Studies

*See* **Labelling Studies in Biochemistry Using NMR.**

# Isotopic Labelling in Mass Spectrometry

**Thomas Hellman Morton**, University of California, Riverside, CA, USA

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| **Applications** |

As is well known, an atomic nucleus is characterized by its positive charge $Z$ (an integer that designates the number of fundamental electric charges) and its mass. Nuclear masses are so close to being multiples of the proton rest mass that they are also characterized by integers, $A$. The integers $A$ and $Z$ characterize a nuclide. Most elements as found in nature contain mixtures of nuclides with different values of $A$, known as isotopes.

Pure isotopes have masses that differ measurably from the integers $A$ (based on a scale on which $^{12}$C has a mass defined as 12.00000 atomic mass units or amu). None of the stable nuclides (i.e. those with nuclear half-lives $>10^9$ years) deviates from integer mass by more than 0.1 amu. The nominal mass of a molecule is the sum of the $A$ values of its constituent atoms.

## Labelling nuclides

In 1914 F.W. Aston demonstrated mass spectrometric separation of the two most abundant isotopes of neon. Since that time, careful measurement has established that 20 chemical elements are monoisotopic (i.e. exhibit only one naturally occurring stable nuclide to any significant extent). From the remaining members of the periodic table, the relative abundances for hundreds of isotopic pairs (same $Z$, different $A$) have been determined by mass spectrometry. The 163 stable nuclides for elements lighter than xenon give rise to slightly more than 300 isotopic pairs.

Labelling embraces those experimental designs in which two nuclides have been mixed under conditions in which they might be expected to behave identically (or nearly so), followed by an analysis that distinguishes them. Usually the mass spectrometer acts as the analyser, but it can play the role of the mixer instead. A measurement of the $^{44}$Ti half-life (a piece of data pivotal for investigating astronomical cataclysms) illustrates one such application in nuclear physics. Bare nuclei with the same $A/Z$ ratio will coincide in a mass separator. At gigavolt kinetic energies, where ions undergo nuclear reactions with target atoms and emerge without any bound electrons, ions from different elements having $A = 2Z$ will not be separated but can be isolated from those for which $A \neq 2Z$. This technique has been used to implant mixtures of $^{22}$Na and $^{44}$Ti simultaneously into solid matrices, where the radioactive sodium provides a standard for measuring the decay rate of the radioactive titanium. In this example, one

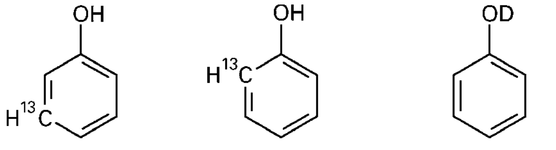
nuclide has been used to label another that has very different chemistry, so as to provide a well-calibrated mixture of the two.

More conventionally, mass spectrometry is used to separate isotopes of a given element or isotopically tagged compounds. Four stable isotopic pairs have found application in the vast majority of molecular labelling experiments examined by mass spectrometry - $^{1}H/^{2}H$, $^{12}C/^{13}C$, $^{14}N/^{15}N$, and $^{16}O/^{18}O$. In each of these cases, the heavier isotope is the rarer, with natural abundances of 0.15% ($^{2}H$), 1.1% ($^{13}C$), 0.37% ($^{15}N$), and 0.20% ($^{18}O$). The survey below will confine itself to molecular labelling experiments that utilize hydrogen, carbon, nitrogen and oxygen isotopes.

## Isomers, isotopomers and isobars

The term isomer has two different meanings, only one of which will pertain here. The meaning that is not relevant is employed by physicists, to whom *nuclear* isomers designate states of atomic nuclei with the same $A$ and $Z$ (generally having different net spins), which have different masses. For example, two nuclear isomers of cadmium, $^{113}Cd$ and $^{113m}Cd$, differ in mass by 0.0003 amu. Nuclear isomers will be omitted from further mention. Isomerism will refer below only to molecules that possess the same connectivities among their atoms but which are not superimposable.

The term isotopomer designates a variant of a molecule with a different net isotopic content. Except for the isotopic difference, the two molecules must otherwise have identical structures. Three isotopomers of phenol provide an example. The ($^{13}C$)phenol drawn as structure [1] represents an isotopomer of phenol containing only $^{12}C$. But [1] is not an isotopomer of structure [2], which also contains one $^{13}C$. Instead [1] and [2], properly speaking, are *isotopic isomers*, differing only by the placement of the labelled atom.




[1] [2] [3]

Ions with different elemental compositions may display the same nominal mass-to-charge ratio ($m/z$), such as the $m/z$ 28 positive ions $N_2^{\bullet+}$, $CO^{\bullet+}$, $C_2H_4^{\bullet+}$, $HCNH^{+}$, $Fe^{2+}$, and $Ce^{5+}$. These are sometimes called isobars. Among physicists, *isobars* denote two nuclides with the same value of $A$ but different values of $Z$. Chemists use the term less restrictively. The monodeuterated phenol [3] is an isotopomer of [1] and [2]. If all of its carbons are $^{12}C$, then the mass of [3] will have the same integer value as the masses of [1] and [2], and all three molecular ions will have $m/z$ 95. [3] is an *isobaric isotopomer* of [1] and [2]. The deuterated ion does not have the same exact mass as the $^{13}C$-containing ions, and they can be distinguished by high-resolution mass spectrometry (HRMS).

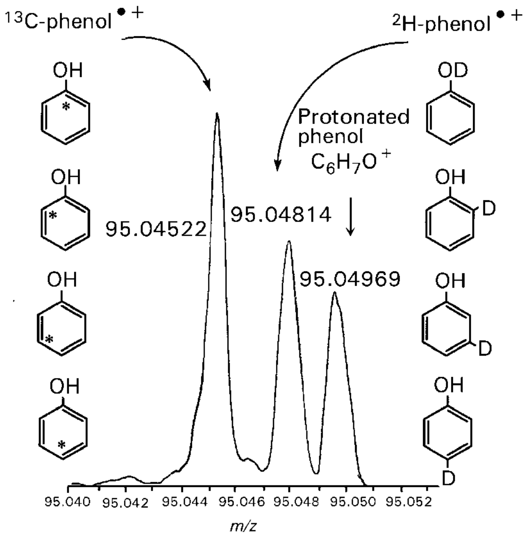



**Figure 1** HRMS of $m/z$ 95 ions from 70 eV electron ionization of $C_6H_5OCH_2CD_2OH$, illustrating the resolution of isobaric ions on a double focusing reverse Nier–Johnson (B-E) instrument with resolving power of 70 000. Each of the peaks that contains a rare isotope is a radical ion (odd number of electrons), which is a mixture of isomeric structures, all the possibilities for which are drawn. Asterisks designate $^{13}C$-containing positions of the ring. The unlabelled ion is an even-electron species. Reproduced a from Nguyen V, Bennett JS and Morton TH (1997) *Journal of the American Chemical Society* **119**: 8342 with permission of the American Chemical Society with modifications.

Mass spectrometry cannot separate the molecular ions of isomers, although two isotopic isomers might well give rise to different fragmentation patterns (see below). Mass spectrometry can separate isotopomers, even if they happen to have the same nominal mass. **Figure 1** illustrates a published example, an experimental measurement of the abundances of $m/z$ 95 fragment ions from a larger molecule. Phenol ($C_6H_5OH$) tagged with a single $^{13}C$ atom corresponds to any one of four isomers (including [1] and [2]) with identical masses, or a mixture thereof. Monodeuterated phenol has the same nominal mass as ($^{13}C$)phenol. The positive-ion mass spectrum reproduced in **Figure 1** displays the resolution of the two isotopomers (which differ by 0.003 amu) as well as protonated phenol ($C_6H_7O^+$), which has a different elemental composition.

At the time of writing, the techniques that give the highest resolution (such as Ion Cyclotron Resonance (ICR) MS) do not enjoy a reputation for accurate quantitation of isobars. Double focusing sector mass spectrometers (such as was used for the mass spectrum in **Figure 1**) appear to be the current method of choice for experiments that require high-resolution measurements of relative ion abundances, even though they have lower resolving power. It is important to bear in mind the distinction between high-resolution mass spectrometry (as exemplified by **Figure 1**) and exact mass determination (for which ICR gives excellent results based on accurate frequency measurements).

## Variations of natural abundances

Precise measurement of relative natural abundances finds application in chemistry. The food industry routinely assays $^{12}C/^{13}C$ ratios of flavourings in order to ascertain their origin. For instance, that isotopic ratio is slightly greater for synthetic vanillin than for the identical molecule from the vanilla bean. As the commercial value of natural vanillin is approximately one million times greater than that of the synthetic, isotope ratio mass spectrometry (IRMS) plays a major role in quality control.

Isotopes do not fractionate identically among the products of most chemical reactions. If a reaction proceeds in 100% yield, the net balance of isotopes in the products ought to be identical to that of the reactants, though the distribution may differ between two products. When a reaction does not run to completion, the natural abundance in the products may vary appreciably from that of the reactants. Measurement of the distribution of natural-abundance isotopes represents an important tool in studying reaction pathways, but the partitioning does not always depend solely upon isotopic mass. In 1978 B. Kraeutler and N.J. Turro used mass spectrometry to determine the fractionation of $^{13}C$ in carbon monoxide from decomposition of the radical pair produced by photolysis of dibenzyl ketone (Eqn [1]).

$$(C_6H_5CH_2)_2CO \xrightarrow{h\nu} [C_6H_5CH_2\cdot \quad C_6H_5CH_2CO\cdot] \rightarrow (C_6H_5CH_2)_2 + CO \qquad [1]$$

The nuclear magnetism of the heavier carbon isotope perturbed the branching ratio under the reaction conditions, favouring return of the intermediate caged radical pair to the starting material. Hence, recovered CO was measurably depleted in $^{13}C$. In part, the elegance of this experiment derives from the fact that it required no synthesis or separation of isotopically enriched starting material. In 1980 R.G. Lawler pointed out that conclusive demonstration of nuclear spin isotope effects demands bracketing of the magnetic isotope between two nonmagnetic ones (e.g. depletion of $^{13}C$ relative to $^{14}C$ as well as depletion relative to $^{12}C$); however, few experiments have been reported that fulfil this rigorous requirement.

## Isotopic dilution and tracer techniques

Isotopic dilution and tracer techniques represent different methodologies, but they overlap to a degree that obviates any clear-cut categorization. Both are techniques for quantitative analysis. In an isotopic dilution experiment one adds a known amount of an isotopomer (sometimes called the ‘spike’) of a known compound (usually containing a rare isotope) and measures the abundance of ions derived from it relative to those from the analyte (which is usually the more common isotopomer). If, for example, one wished to quantitate phenol ($C_6H_5OH$) in a mixture, one might add a known quantity of pentadeuterated phenol ($C_6D_5OH$) as a spike and measure the m/z 94:m/z 99 intensity ratio by mass spectrometry.

This conceptually simple approach, called a mass isotopomer abundance ratio (MIAR) measurement, requires a number of refinements. First of all, other components of the mixture might give rise to ions of the same mass - for instance, all alkyl phenyl ethers larger than anisole give very prominent m/z 94 ions. Therefore, MIAR experiments ordinarily require chromatographic separation prior to mass spectrometry (such as GC-MS), an approach sometimes referred to as isotope ratio monitoring (IRM) chromatography MS. Since different isotopic variants may have slightly different retention times, one must monitor intensity ratios over the entire elution profile of the component under analysis. Second, isotope effects may alter the relative abundances of fragment ions; hence, calibration with authentic samples is necessary. Finally, MIAR measurements give best results when the peaks to be compared have comparable intensities. When one ion has abundance more than an order of magnitude greater than the other, interferences and nonlinearities begin to introduce uncertainty and the measured MIAR may depend upon the amount of sample introduced into the mass spectrometer. C.K. Fagerquist and J.-M. Schwarz have surveyed the origins of systematic

deviations and list ion-molecule reactions within the mass spectrometer and detector nonlinearities as possible sources of error.

The most reliable abundance measurements involving stable isotopes make use of carbon, nitrogen or oxygen labelling, with chromatographic separation followed by quantitative conversion to light gases prior to mass spectrometric measurement. GC-combustion-isotopic dilution MS (GC-C-IDMS) depends upon the high precision with which isotopic abundances for $CO_2$ and $N_2$ can be measured by means of isotope ratio mass spectrometry. Oxygen has three stable isotopes, and $^{13}C$ natural abundance interferes if one wishes to analyse $^{17}O$ content in $CO_2$. For such purposes, $CO_2$ can be converted to molecular oxygen with 5% molecular fluorine in helium. The basic equation for IDMS (regardless of whether the mass analysis is performed on the intact molecule or on its degradation products) is

$$N^{\text{sample}} = N^{\text{spike}} \frac{(R^{\text{spike}} - R^{\text{blend}})\Sigma R_i^{\text{sample}}}{(R^{\text{blend}} - R^{\text{sample}})\Sigma R_i^{\text{spike}}} \qquad [2]$$

where $N^{spike}$ stands for the (known) number of moles of spike; $N^{sample}$ stands for the number of moles of analyte (to be quantitated); $R^{spike}$ stands for the observed peak intensity ratio for the pure spike substance; $R^{sample}$ stands for the observed peak intensity ratio for the pure analyte; $R^{blend}$ stands for the observed peak intensity ratio for the mixture of spike and sample; and the $R_i$ stand for the observed ratios for all isotopomers.

The following worked example illustrates how to apply Equation [2]. Consider an unknown number of moles of analyte ($N^{sample}$) for which the observed $^{13}C$:$^{12}C$ peak intensity ratio is $R^{sample} = 1/9$, to which one adds a known number of moles of spike ($N^{spike}$), for which the observed $^{13}C$:$^{12}C$ peak intensity ratio is $R^{spike} = 2/1$. Suppose the observed $^{13}C$:$^{12}C$ peak intensity ratio for the mixture is $R^{blend} = 23/37$. The values to be summed in the numerator are $\sum R_i^{sample} = 1/9 + 1$, the sum of the $^{13}C$:$^{12}C$ peak intensity ratio and the $^{13}C$:$^{13}C$ peak intensity ratio for the pure analyte. The values to be summed in the denominator are $\sum R_i^{spike} = 2/1 + 1$, the sum of the $^{13}C$:$^{12}C$ peak intensity ratio and the $^{13}C$:$^{13}C$ peak intensity ratio of the pure spike. Inserting these values into Equation [2] gives

$$N^{\text{sample}} = N^{\text{spike}} \left[ \frac{(2/1 - 23/37)(10/9)}{(23/37 - 1/9)(3)} \right] = N^{\text{spike}}[1] = N^{\text{spike}} \qquad [3]$$

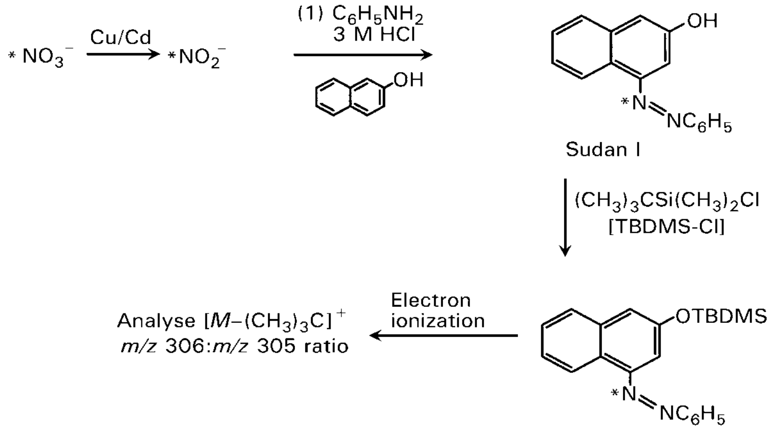
In other words, the example corresponds to the outcome for equal amounts of sample and spike.

The above instance might obtain when the sample has nine carbon atoms (containing natural-abundance $^{13}C$) and the spike has one partially labelled atom (though the spike might contain a mixture of isotopic isomers). Quantitation becomes more reliable if naturally occurring isotopomers do not overlap the peaks from the spike, so that $R^{sample} \to 0$. Multiple labelling can adequately address that issue. Also, precision increases as the isotopic purity of the spike is increased, so that $R^{spike} \to \infty$. Strictly speaking, it is not the values of $R^{sample}$ and $R^{spike}$ that introduce error, but rather the uncertainty in measuring those ratios when they do not have their limiting values of zero and infinity.

In tracer studies, the analyte is not present at the time the spike is added. Diagnosis of human infection by Helicobacter pylori (the bacteria implicated in digestive tract ulcers) exemplifies such a technique. This microorganism survives at low pH in the stomach by generating locally high concentrations of ammonia to buffer its immediate environment. When a patient swallows a sample of $^{13}C$-labelled urea, a colony of these bacteria (if present) will hydrolyse the urea to $NH_3$ and $CO_2$, giving a characteristic rate of production of $^{13}C$-carbon dioxide in the patient's breath.

The distinction between tracer and isotope dilution methodologies is often blurred. For instance, IDMS sometimes requires a series of chemical transformations. A recently published protocol for quantitating nitrate in environmental samples dilutes the sample with $^{15}NO_3^-$ followed by a series of steps that reduce it to nitrite and then convert it to an azo dye, Sudan I, as **Scheme 1** summarizes. Derivatization of label-containing Sudan I to its t-butyldimethylsilyl (TBDMS) ether then renders the labelled molecule sufficiently volatile to be introduced into a mass spectrometer. The ratio of M−57 fragment ion intensities quantitates the $^{15}N/^{14}N$ ratio, once correction has been made for $^{29}Si$ (natural abundance 5% that of $^{28}Si$) and $^{13}C$ natural abundance of the 18 carbons in the ion.

A physiological labelling technique that applies tracer and isotopic dilution methodologies concurrently involves the use of doubly labelled water (containing $^{18}O$ in addition to a heavy isotope of hydrogen) for metabolic studies. After an animal is dosed with doubly labelled water, part of the labelled oxygen turns up in exhaled $CO_2$. From the amount of $^{18}O$ tracer in the $CO_2$ and the proportions of $^{18}O$ and isotopic hydrogen in excreted water (the isotopic dilution), one can infer the rate of energy expenditure of a living organism.




**Scheme 1**

Quantitative analysis of tracer experiments can be more complicated than for isotopic dilution. Consider a hypothetical outcome for pure $H_2{}^{18}O$ mixed thoroughly with a 10-fold excess of pure $H_2{}^{16}O$, where all of the oxygen atoms (and only those oxygen atoms) turn up in recovered $CO_2$. Statistically, the carbon dioxide should exhibit a distribution of zero, 1, or 2 labelled oxygens in the proportions m/z 44:m/z 46:m/z 48 = 100:20:1. It is clear that Equation [2] is not the right expression for analysing that result. Moreover, experiment has shown that observed ratios often deviate from statistical proportions because of fractionation factors (i.e. isotope effects that preferentially partition the heavier isotope into one chemical form versus another). The mathematical analysis for doubly labelled water studies, which includes the effects of fractionation factors, is well documented elsewhere.

## Accelerator mass spectrometry

The sensitivity of tracer experiments increases as the naturally occurring isotopic background decreases. The natural abundance of $^{14}C$ in the biosphere (standardized to the year 1950) is approximately 1.2 parts per trillion (a level designated as Modern), a factor of $10^{-10}$ lower than the natural abundance of $^{13}C$. Moreover, when an organism dies, it ceases to equilibrate with atmospheric $CO_2$ and its $^{14}C$ decays away. Many commercially available compounds are petroleum derived and routinely contain $^{14}C$ levels < 1% Modern. Therefore, mass spectrometric determination of $^{14}C$ has much greater sensitivity for tracer experiments than $^{13}C$, with a usable detection limit of 0.001 fmole for $^{14}C$ versus 0.25 fmole for $^{13}C$.

Accelerator mass spectrometry (AMS) is designed to analyse unstable isotopes, and a greatly simplified schematic is drawn in **Figure 2**. Historically, AMS measurements of isotopic ratios have been employed to measure half-lives of radionuclides by monitoring the disappearance of an unstable nuclide relative to a stable one (e.g. $^{44}Ti$ relative to $^{46}Ti$) over a period of months. Contemporary applications address more chemical problems. The general technique involves two stages of ionization alternating with two stages of m/z selection. In the first ionization, atomic negative ions are produced from a solid sample, and the first mass spectrometer selects all ions with a given nominal mass. The negative ion is then accelerated to high velocity and transmitted through a foil stripper (a thin sheet of solid material), a passage that removes the outer electrons from the projectile ions. This second stage of ionization converts atomic ions with a single negative charge into positive ions with multiple charges, which are analysed by a second mass spectrometer.

When applied to $^{14}C$, AMS makes use of a sample that has been converted to graphite. Since all chemical elements are present in the sample in concentrations greater than one part in $10^{12}$, a major obstacle is to avoid isobaric ions. Bombarding the graphite with a fast ion beam produces $C^-$ among the stable atomic ions. Nitrogen atoms do not form stable negative ions, so the only atomic negative ion produced having m/z 14 is $^{14}C^-$. Along with this are produced a number of isobaric m/z 14 molecular ions, such as $^{13}CH^-$ and $^{12}CH_2{}^-$. The m/z 14 negative ions are mass-selected by a low-energy mass spectrometer. The ion beam is then accelerated to megavolt kinetic energies and passed through a thin carbon foil, which strips off all the valence electrons. The molecular ions fragment as a consequence, while the atomic carbon ions simply acquire up to six positive charges. The mass filtering of AMS relies upon selection of m/z 3.5 positive ions ($^{14}C^{4+}$) formed from previously selected m/z 14 negative ions. One might

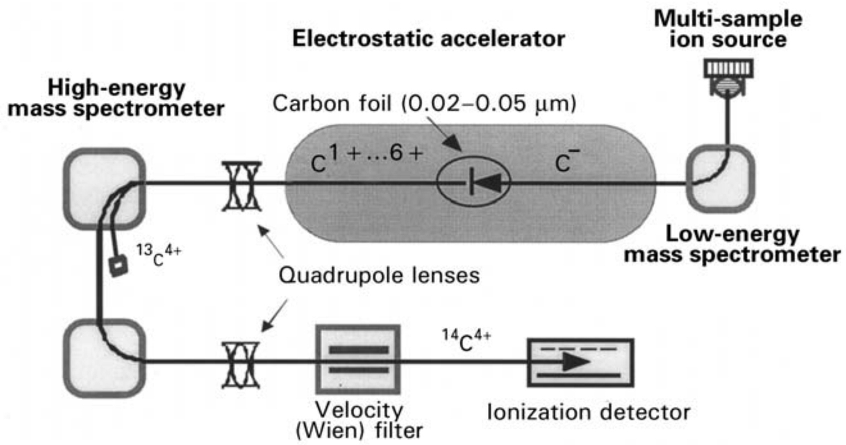



**Figure 2** Schematic of the accelerator mass spectrometer at the Centre for AMS at Lawrence Livermore National Laboratory. Negative ions from a multi-sample ion source are separated by a low-energy mass spectrometer, accelerated to 6 MeV in an electrostatic accelerator, converted to positive ions by passage through a foil stripper, and accelerated again. Quadruply charged carbon atomic ions (30 MeV kinetic energy) are focused by quadrupole lenses and resolved by high-energy mass spectrometers, followed by velocity selection and identification of charge state in an ionization detector.

think that two stages of mass analysis would suffice to separate the $^{14}C$ atomic ions from all other isobars, but there is a substantial interference from $^{7}Li_2^-$ in the source that is converted to $^{7}Li^{2+}$ by passage through the foil stripper. Therefore, the final stage of identification of $^{14}C$ requires that the ion beam be velocity selected and then passed through an ionization chamber, where collisions with neutral gas molecules produce secondary ions with a rate that distinguishes between a doubly charged projectile ($^{7}Li^{2+}$) and a quadruply charged projectile ($^{14}C^{4+}$). The high kinetic energy of the ions in AMS permits sensitive single-ion counting, and samples containing $< 10^6$ atoms of $^{14}C$ in $< 1$ mg total carbon can routinely be analysed.

In practice AMS is used to measure $^{14}C/^{13}C$ isotopic ratios. For most applications the variation of $^{13}C$ abundance in the sample does not significantly affect the accuracy of an AMS measurement. AMS has been widely used for samples whose $^{14}C$ content is below Modern (as in carbon dating) and is beginning to find major applications for samples where $^{14}C$ content greatly exceeds Modern (as in biomedical research). The latter is exemplified by a recent study of covalent modification of proteins. When the enzyme aldolase is treated with acetoacetic ester, the protein irreversibly attaches a small but reproducible amount of the ester. Such a low level of tagging is sometimes referred to as background labelling, since (even with carrier-free ($^{14}C$)acetoacetic ester) there is not enough incorporated radioactivity to give a decay rate very much higher than the background for detecting nuclear disintegrations. The isotopic label stands out well above natural $^{14}C$ levels, but decay counting is an intrinsically inefficient method of detection. No matter what type of counter is used, only fourteen $^{14}C$ nuclei will decay per hour for every $10^9$ that are present in the sample. In the case of tagged aldolase, the amount of incorporated radioactivity was too low for decay counting to be used for fragments of the tagged enzyme, and AMS presented the only method for analysing $^{14}C$ content after cleavage of the protein.

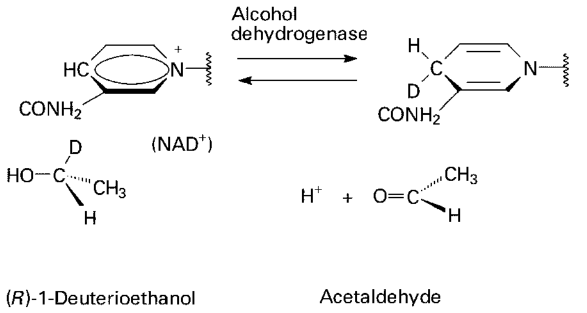
AMS takes advantage of the inherent efficiency of mass spectrometry and can readily measure the amount of label in protein fragments purified by analytical gel electrophoresis. Site-specific chemical cleavage of tagged aldolase gave four electrophoresis bands, each containing $\leq 0.05$ mg of polypeptide. Individual gel slices (containing 5-10 mg of polyacrylamide each) were converted to $CO_2$ and then reduced to pure graphite. AMS gave analyses that showed $^{14}C/^{13}C$ ratios in the range from 5 to 50 times Modern reproducibly, depending on the identity and quantity of labelled polypeptide in the gel slice, which (when corrected for the great excess of stable carbon isotopes contributed by the polyacrylamide) corresponded to 0.03-3 fmol $^{14}C$ per polypeptide fraction. The extent of labelling was found to be nonuniform: it did not correlate with the molecular mass of the fragments but was proportional to the number of tyrosine residues. No correlation was observed for any other amino acid. Therefore, what had hitherto been dismissed as background labelling turns out to be an amino acid-specific covalent modification.

## Tracer studies for characterizing reaction pathways

Mass spectrometry is typically used to look at products from reactions where a reactant has intentionally been enriched in a rare stable isotope. This is illustrated by a 'crossover experiment' reported by P.D. Bartlett and T.G. Traylor in 1963, who reported the labelling of molecular oxygen liberated by chain termination of free t-butyl hydroperoxy radicals. When a radical doubly labelled with $^{18}O$ reacted with an unlabelled radical to yield di-t-butyl peroxide (Eqn [4]), the predominant gaseous product gave m/z 34 ($^{16}O^{18}O$). That outcome indicates that each radical contributes one oxygen atom and implicates a tetroxide intermediate, which could not be directly observed.

$$(CH_3)_3C^{18}O^{18}O\cdot + (CH_3)_3C^{16}O^{16}O\cdot \rightarrow [(CH_3)_3C^{18}O^{18}O^{16}O^{16}OC(CH_3)_3] \rightarrow {}^{16}O^{18}O + (CH_3)_3COOC(CH_3)_3 \quad [4]$$

An experiment that fails to show crossover can be equally meaningful. The 1953 report by F.A. Loewus, F.H. Westheimer and B. Vennesland on the enzyme alcohol dehydrogenase represents one of the earliest and most profound applications of isotopic labelling. These experiments demonstrated the facial selectivity of the enzyme-catalysed reaction, which is depicted schematically in Equation [5]. They performed mass spectrometric analyses of mixtures of HD and $H_2$ produced by reduction of water from combusted samples of ethanol and acetaldehyde, and they established that the (R) enantiomer of deuterated ethanol transfers only deuterium to the coenzyme $NAD^+$ (a partial structure of which is drawn in Eqn [5]). The reduced coenzyme transfers only deuterium back to acetaldehyde. Under conditions where the reaction goes back and forth many times, no deuterated acetaldehyde ($CH_3CD{=}O$) is detected. If, on the other hand, $CH_3CD{=}O$ reacts with the unlabelled, reduced coenzyme (known as NADH), the reduction gives (S)-1-deuterioethanol. No label transfers between deuterated acetaldehyde and the coenzyme. The remarkable conclusion (confirmed for many enzymes subsequently) is that biological catalysts distinguish the two faces of planar carbon atoms that bear three nonidentical substituents (e.g. the top and bottom faces of the planar ring of $NAD^+$ and of the carbonyl carbon of acetaldehyde).



[5]

## MS/MS – tandem versus ion storage

In the gas phase, a neutral molecule can be ionized by removal or addition of an electron via an electron beam (electron ionization or EI), by reaction with another gaseous ion (chemical ionization or CI), or by interacting with ionizing radiation. Most methods of depositing charge are so energetic that they lead to fragmentation. By close examination of fragment ions, one can sometimes draw conclusions regarding the placement of isotopic label and the mechanism of ionic decomposition. Quite often the source mass spectrum is confusing, since isotopic label frequently randomizes under typical ionization conditions. Sequential mass spectrometry separates a portion of the ion beam (usually a single nominal m/z) and reanalyses it after it has undergone subsequent ion decompositions. Two general types of MS/MS apparatus enjoy wide usage: those that have mass spectrometers in spatial sequence (tandem instruments) and those that examine ions using a single mass spectrometer, in which the two stages of mass selection occur in temporal sequence (ion storage instruments).

Ions that fragment between two stages of mass selection are said to undergo metastable ion decompositions. These often show patterns different from a source mass spectrum. The spectrum from which **Figure 1** is taken provides an example. Electron ionization of $C_6H_5OCH_2CD_2OH$ shows a $C_6H_7O^{\cdot+}$:$C_6H_5DO^{\cdot+}$:$C_6H_6DO^+$ ratio of about 0.75:1:0.6. Metastable ion decomposition of the corresponding molecular ion $C_6H_5OCH_2CD_2OH^{\cdot+}$ shows a ratio $< 0.05{:}1{:}2$. The high-resolution spectrum in **Figure 1** required a two-sector instrument. In metastable studies, the same two sectors were used in tandem. Since isobaric m/z 95 ions cannot be resolved by a single sector, the $C_6H_7O^{\cdot+}$: $C_6H_5DO^{\cdot+}$ abundance ratio from metastable ion decompositions had to be inferred from a comparison of several additional isotopic variants, $C_6H_5OCH_2CD_2OD$, $C_6H_5OCD_2CH_2OH$, $C_6H_5OCD_2CD_2OD$ and $C_6D_5OCH_2CH_2OH$.

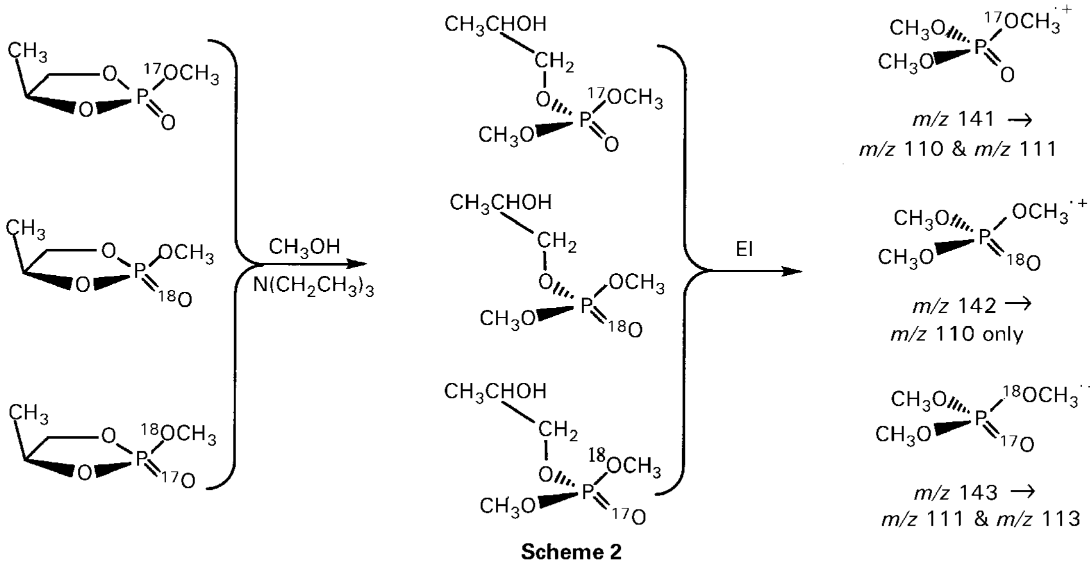
## Determination of isotopic stereochemistry

In favourable cases, configuration of an isotopically substituted stereogenic centre can be determined by means of mass spectrometric fragmentation. The labelled analyte must contain a ring or else must pass through a cyclic transition state in the course of decomposition. Acyclic secondary alcohols with a single deuterium adjacent to the oxygen-bearing carbon (general formula $R^1CHOHCHDR^2$) have two diastereomeric forms as a consequence of isotopic substitution. While the mass spectra of the alcohols present major difficulties of interpretation, several derivatives undergo mass spectrometric eliminations by which relative stereochemistry can be assigned. The ionized phenyl ethers, for instance, expel neutral alkene. The $C_6H_5OD^{\bullet+}$: $C_6H_5OH^{\bullet+}$ ratios differ from one stereoisomer to the other to such an extent that ion storage techniques can be used to assay the proportions of isotopic diastereomers. While chromatography cannot separate isotopic isomers, GC-MS/MS for quantitating mixtures of isotopic diastereomers has been reduced to practice.

In 1979 Knowles, McLafferty and their co-workers made dramatic use of metastable ion mass spectrometry in solving the stereochemistry of phosphate monoesters $ROP^{16}O^{17}O^{18}O^{2-}$ that are chiral by virtue of substitution with oxygen isotopes. Reaction with optically active 1,2-propanediol affords the mixture of three isotopomeric cyclic triesters shown to the left in **Scheme 2**. One $ROP^{16}O^{17}O^{18}O^{2-}$ enantiomer gives the mixture depicted; the other $ROP^{16}O^{17}O^{18}O^{2-}$ enantiomer gives products in which $^{17}O$ and $^{18}O$ are transposed. Ring opening with methanol gives a mixture of three isotopomeric triesters (middle of **Scheme 2**) that exhibits mass spectrometric fragment ions corresponding to trimethyl phosphate. The isotopomeric trimethyl phosphate radical ions display prominent metastable ion decompositions from loss of formaldehyde. For one $ROP^{16}O^{17}O^{18}O^{2-}$ enantiomer the m/z 142 fragment loses only $CH_2{}^{16}O$, as **Scheme 2** portrays, while m/z 141 loses either $CH_2{}^{16}O$ or $CH_2{}^{17}O$ and m/z 143 loses either $CH_2{}^{16}O$ or $CH_2{}^{18}O$. These metastable ion decompositions are summarized beneath their respective trimethylphosphate ions to the right in **Scheme 2**. The other $ROP^{16}O^{17}O^{18}O^{2-}$ enantiomer leads to ions (not shown in **Scheme 2**) for which m/z 141 loses only $CH_2{}^{16}O$; m/z 142 loses $CH_2{}^{16}O$ or $CH_2{}^{18}O$; and m/z 143 loses $CH_2{}^{16}O$ or $CH_2{}^{17}O$.

## Elucidation of fragmentation pathways and isotope effects

Isotopic labelling provides an excellent method for probing how ions decompose in the mass spectrometer. Frequently the pathway followed by an ion resembles a cognate reaction from solution chemistry, but sometimes the outcome differs quite unexpectedly. Many elimination reactions - both in solution and in the gas phase - operate via loss of groups attached to adjacent atoms. But EI on $(CH_3)_2CDCH_2CH_2Br$ yields an M- DBr daughter ion that is much more abundant than the M- HBr fragment. The structure of the ion (inferred from subsequent surface neutralization and analysis of the $C_5H_{10}$ neutral) is $(CH_3)_2C{=}CHCH_3$. Clearly some




Scheme 2

transposition of hydrogen has taken place. Investigating the rearrangements of gaseous ions represents a major area of research in mass spectrometry.

Deuterium substitution often affects the competition between two otherwise identical fragmentation pathways. EI on 1,4-dimethoxybutane shows no molecular ion and an intense peak from loss of methyl radical (a fragmentation not seen in the shorter or longer homologous $\alpha,\omega$-dimethoxy-n-alkanes). When ionizing energy is lowered, the contribution from methyl loss increases until it represents > 70% of the total ionization. Deuterating one methyl ($CD_3OCH_2CH_2CH_2CH_2OCH_3$) exerts a comparatively small isotope effect: loss of $CD_3$ is 0.7 as intense as loss of $CH_3$. Deuterating the nearest methylene as well ($CD_3OCD_2CH_2CH_2CH_2OCH_3$) lowers the loss of $CD_3$ to 0.5 the intensity of loss of $CH_3$. In other words, replacing two H atoms with D on the other side of the oxygen has as great an effect as replacing all three H atoms with D on the methyl itself. This has been taken to imply a hidden hydrogen rearrangement, in which a proton transfers to the remote oxygen simultaneously with a C- O bond cleavage.

If two reactions exhibit different isotope effects in a product-determining step, they cannot be proceeding via a common intermediate. For example, elimination of propene produces the most intense fragment ion in the EI and CI mass spectra of n-propyl phenyl ether, with extensive scrambling of all of the hydrogens in the side-chain. Could it be that the n-propyl ether ions isomerize to isopropyl ether ions prior to dissociating? If so, photoionization of $CD_3CH_2CH_2OPh$ and $CH_3(CD_3)CHOPh$ should manifest the same isotope effects on propene expulsion. But the isotope effect (measured as the $C_6H_5OH^{\cdot+}/C_6H_5OD^{\cdot+}$ intensity ratio) is about 1.2 for n-propyl, significantly less than the value (1.4) for isopropyl. Therefore, the two isomeric ions dissociate via distinct pathways.

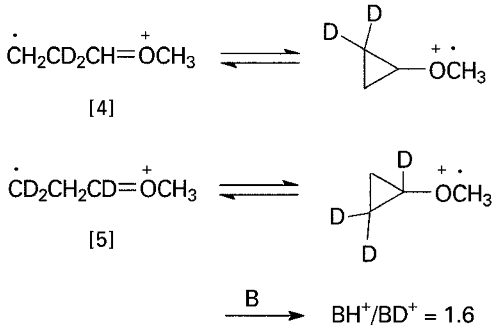
## Discrimination of isobars and isomers

Isotopic labelling often shifts the mass of one isobar relative to another for the purposes of identification. In ionized mixtures of methyl fluoride with a trace of ammonia, the weakly hydrogen-bonded cluster ion $CH_3F\text{-}HNH_3^+$ (m/z 52) has the same mass as a much more stable ammonia cluster ion, $(NH_3)_3H^+$. With $^{15}NH_3$ the mass-to-charge ratios become m/z 53 and m/z 55, respectively. Moreover, the isotopic substitution identifies the product of the m/z 52 $\rightarrow$ m/z 53 fragmentation as $FCH_2^+$.

Ion structures can be deduced from isotopic labelling. One of the first distonic ions ever demonstrated, portrayed in Equation [6],

$$\overset{D^+}{CH_3OCH_2\dot{C}H_2} \xrightarrow{B} BD^+ \qquad [6]$$

was characterized by its reaction with gaseous base to transfer only $D^+$. If the ion had had the ionized methyl ethyl ether structure, it would have transferred $H^+$ at least as often as $D^+$. A similar type of experiment demonstrated that the distonic ion in Equation [7] converts to (or interconverts with) ionized methoxycyclopropane, since the two isotopomers [4] and [5] transfer $H^+$ and $D^+$ to gaseous bases in the same proportions.




## Nuclear magnetic resonance and mass spectrometry

Neutral products from decomposition of gaseous ions can be isolated and subsequently characterized by gas chromatography or magnetic resonance spectroscopy. Fluorine NMR has been especially useful in this regard, since the $^{19}F$ chemical shift scale is so wide that isotopic shifts of NMR peak positions permit assessment of the extents and positions of deuteration in isomers/isotopomer mixtures.

**Figure 3** reproduces $^{19}F$ NMR spectra of fluoroethylenes recovered from EI of two different isomers of dideuterated β-fluorophenetole. The isomers give slightly different mass spectrometric fragmentation patterns, but the decomposition mechanism could not be fully understood without the analysis of the expelled neutrals. As the NMR shows, the distribution of deuterated fluoroethylenes is not the same for the two isomers. One isomer yields a product not formed by the other, and vice versa. $C_6H_5OCD_2CH_2F$ produces CHD=CHF, but $C_6H_5OCD_2CH_2F$ does not. $C_6H_5OCH_2CD_2F$ produces CHD=CDF, but $C_6H_5OCD_2CH_2F$ does not. Hence hydrogens cannot be randomizing within ionized β-fluorophenetole prior to elimination.

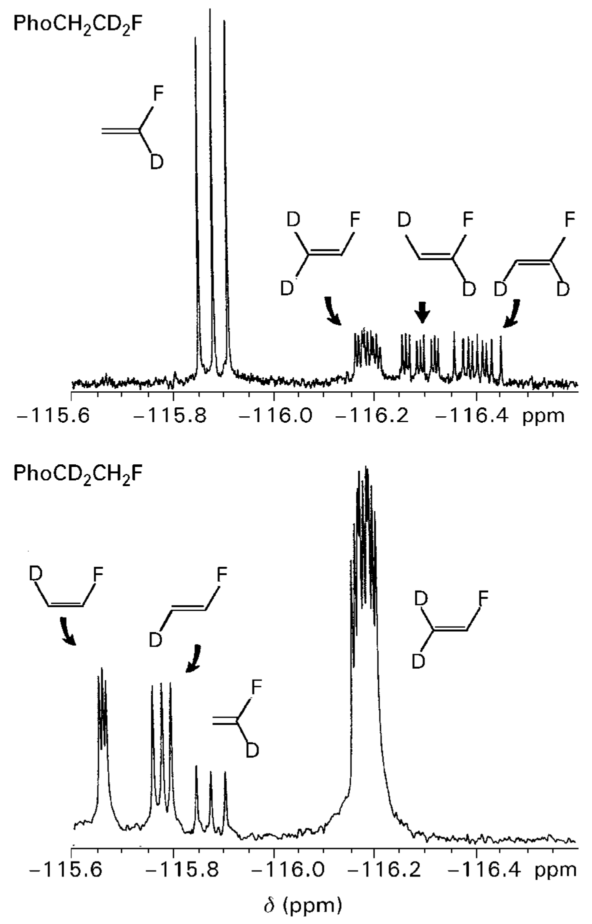
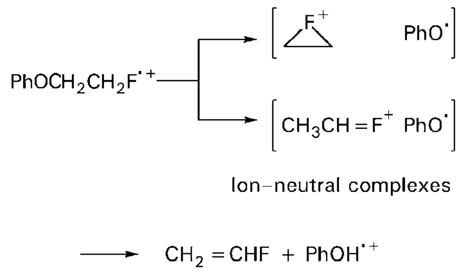



**Figure 3** $^{19}F$ NMR spectra (proton noise-decoupled) of deuterated fluoroethylenes expelled from the indicated dideuterated β-fluorophenetoles after 70 eV electron ionization. Observed peak splittings come from deuterium–fluorine spin–spin coupling, by which isomers/isotopomers are identified. Reproduced with permission of the American Chemical Society from Nguyen V, Cheng X and Morton TH (1992) *Journal of the American Chemical Society* **114:** 7127.

It turns out that ionized β-fluorophenetole dissociates via two competing pathways, illustrated in Equation [8], which are indistinguishable but for the isotopic label. Both involve cleavage of the $sp^3$-carbon oxygen bond to form ion- neutral complexes, which live of the order of a nanosecond before a proton transfers from the fluorine-containing cation to the phenoxy radical. The upper pathway drawn in Equation [8] has the fluorine bridging between two carbons, yielding a symmetrical cation structure. The lower pathway involves a hydrogen shift to form the most stable of the possible $C_2H_5F^+$ structures. The two competing mechanisms operate in about a 1:2 ratio.




[8]

Chemists have long dreamed of combining mass spectrometry with NMR, so as to probe the structures of gaseous ions directly. While that vision remains to be fulfilled, NMR analysis of neutral fragments expelled by gaseous ions represents a step towards that goal.

## List of symbols

$N^{sample}$ = number of moles of sample; $N^{spike}$ = number of moles of spike; $R^{sample}$ = peak intensity ratio for pure analyte; $R^{spike}$ = peak intensity ratio for pure spike substance; $R^{blend}$ = peak intensity ratio for mixture of sample and spike.

*See also:* **Biochemical Applications of Mass Spectrometry; Chromatography-MS, Methods; Cosmochemical Applications Using Mass Spectrometry; Food Science, Applications of Mass Spectrometry; Fragmentation in Mass Spectrometry; Isotope Ratio Studies Using Mass Spectrometry; Labelling Studies in Biochemistry Using NMR; Medical Science Applications of IR; MS–MS and MS$^n$; Metastable Ions; Sector Mass Spectrometers; Stereochemistry Studied Using Mass Spectrometry.**

## Further reading

Bennett JS, Bell DW, Buchholz BA, Kwok ESC, Vogel JS and Morton TH (1996) Accelerator mass spectrometry for assaying irreversible covalent modification of aldolase by acetoacetic ester. International Journal of Mass Spectrometry 179/180: 185- 193.

Brenna JT (1997) High-precision continuous flow isotope ratio mass spectrometry. Mass Spectrometry Reviews 16: 227- 258.

DeBièvre P (1993) Isotope dilution mass spectrometry as a primary method of analysis. Analytical Proceedings 30: 328- 333.

Holmes JL (1975) The elucidation of mass spectral fragmentation mechanisms by isotopic labelling. In: Buncel E and Lee CC (eds) Isotopes in Organic Chemistry, Vol 1. Amsterdam: Elsevier, pp 61- 133.

McAdoo DJ and Morton TH (1993) Gas phase analogues of cage effects. Accounts of Chemical Research 26: 295–302.

Morton TH and Beauchamp JL (1975) Chainlength effects upon the interaction of remote functional groups. Journal of the American Chemical Society 97: 2355–2362.

Shaler TA, Borchardt D and Morton TH (1999) Competing 1,3- and 1,2-hydrogen shifts in gaseous fluoropropyl ions. Journal of the American Chemical Society 121: 7907–7913.

Speakman JR (1997) Doubly Labelled Water. London: Chapman and Hall.

Taphanel MH, Morizur JP, Leblanc D, Borchardt D and Morton TH (1997) Stereochemical analysis of monodeuterated isomers by GC/MS/MS. Analytical Chemistry 69: 4191–4196.

# L

# Labelling Studies in Biochemistry Using NMR

**Timothy R Fennell** and **Susan CJ Sumner**,
Chemical Industry Institute of Toxicology,
Research Triangle Park, NC, USA

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

The elucidation of the pathways involved in the metabolism of exogenous and endogenous materials has been advanced considerably by the use of labelled chemicals. The transformation of these labelled intermediates can be followed by a variety of techniques. Much of the early work in the elucidation of pathways of intermediary metabolism involved the use of radioactive isotopes, and $^{14}$C in particular, as tracers. More recently, the focus of research has been on the control of metabolism, and flux through pathways. Nuclear magnetic resonance (NMR) spectroscopy has been used in the study of metabolic processes to follow the conversion of compounds with natural abundance isotopes such as $^{1}$H, $^{19}$F, or $^{31}$P or isotopically labelled compounds containing $^{2}$H, $^{15}$N, or $^{13}$C. This discussion focuses on the application of isotope labels to track the fate of chemicals in biological systems. The majority of the NMR applications involving labels use $^{13}$C-enriched chemicals and in general are designed to describe the pathways of metabolism or to evaluate flux through pathways. Although less frequently used, $^{2}$H and $^{15}$N can also provide useful information. An isotope effect on metabolism resulting from the increase in mass with $^{2}$H compared with $^{1}$H may be a significant limitation. $^{15}$N labels are limited to the study of nitrogen-containing compounds and have generally been limited by the low sensitivity for detection of $^{15}$N. Two types of applications can be identified: those in which the metabolism of endogenous materials is assessed by addition of labelled materials, with the fate of the label indicating the fate of the pool of both labelled and unlabelled material and those in which the metabolism of labelled exogenous materials is assessed.

## Detection of $^{13}$C-labelled metabolites

$^{1}$H NMR spectroscopy, which is widely used for structure elucidation, is very sensitive as a result of both the high natural abundance and NMR sensitivity of the proton. As a means of tracking the fate of chemicals in biological systems, $^{1}$H NMR is widely used and works well with chemicals with functional groups whose NMR signals can be distinguished readily from endogenous materials (see also other sections). However, $^{1}$H NMR is, in general, not very selective, and frequently lacks the ability to discriminate among compounds of interest in complex biological mixtures. In addition, water in biological systems causes dynamic range issues that must be circumvented by suppression techniques during the acquisition of $^{1}$H NMR spectra. Carbon-13 occurs with a low natural abundance (1.1%), and has a spin quantum number of 1/2; $^{13}$C has an NMR sensitivity 1.6% that of $^{1}$H, but is highly selective. Signals in $^{13}$C spectra are dispersed over a wide frequency range, depending on the substituents of the carbon atom from which the signal arises. The wide dispersion of $^{13}$C NMR signals is one of the major reasons for using $^{13}$C-enriched chemicals in the study of metabolism. Using $^{13}$C-enriched materials permits detection over the low natural abundance of unlabelled materials. Compared with natural abundance of 1.1%, enrichment of a carbon atom to 99%

would cause an increase in the signal-to-noise ratio of 90-fold. The use of NMR for the detection of labelled materials has the advantage over other techniques that the enrichment at a particular carbon atom in a molecule can readily be determined without degradation of the molecule.

Proton-decoupled $^{13}C$ NMR signals for natural abundance materials occur as singlets since the probability of two adjacent $^{13}C$ nuclei is 0.0123%. An isolated enriched carbon atom in an otherwise natural abundance material gives rise to a single resonance. Adjacent enriched carbon atoms give multiplet patterns as a result of carbon–carbon coupling. Two adjacent $^{13}C$ atoms each give rise to a doublet as shown in **Figure 1**. The NMR spectrum of three adjacent $^{13}C$ atoms ($C_a$–$C_b$–$C_c$) will give rise to a doublet for each of the external carbons ($C_a$ and $C_c$) and a doublet of doublets for the centre carbon atom ($C_b$), excluding the effects of long-range $^{13}C$–$^{13}C$ coupling. This type of splitting pattern permits the distinction of exogenous and endogenous compounds. In mixtures, a singly labelled compound may not readily be distinguished from endogenous signals, whereas the multiplet patterns associated with multiply labelled materials readily stand out.

The $^{1}H$-decoupled 1D $^{13}C$ NMR spectrum is the general starting place for the detection of labelled metabolic intermediates or products. In the study of intermediary metabolism, assignments can frequently be made by comparison of the chemical shifts of signals with those from known standards. With multiply labelled metabolites, more detailed structural information can be obtained by the acquisition of additional spectral information. The measurements of carbon–carbon coupling constants from the 1D spectrum may be used to determine the connection between carbon resonances. In mixtures, however, coupling patterns for carbon atoms from different compounds can be similar. Carbon–carbon connectivity experiments (e.g. incredible natural abundance double quantum transfer experiment; INADEQUATE) provide an unambiguous determination of carbon–carbon connectivity. The two-dimensional spectra generated from these experiments contain a dimension for chemical shift and a dimension for double quantum frequency. The presence of two contours with the same double quantum frequency (i.e. aligned horizontally along the $F_1$ axis) but with different chemical shifts ($F_2$ axis) indicates coupling between the carbons giving rise to these signals (**Figure 2**). These experiments can be conducted to determine direct and long-range carbon–carbon coupling.

A number of techniques permit the detection of the number of protons attached to specific carbon atoms. The attached proton test (APT) experiment and related pulse sequences can be used to distinguish between CH, $CH_2$, and $CH_3$ carbon signals. However, they are not always able to distinguish between CH and $CH_3$ groups. The 2D carbon proton experiment (e.g. heteronuclear 2D *J*-resolved or HET2DJ) permits the identification of the number of hydrogen atoms attached to each carbon atom. The 2D spectrum generated from these pulse sequences contains a dimension for carbon chemical shift and a dimension indicating the multiplets produced by the attached protons ($2nI+1$ rule). Carbons with three, two, or one attached protons will have four, three, or two contour peaks, respectively, aligned at the chemical shift of the carbon signal in the 2D spectrum (**Figure 3**).

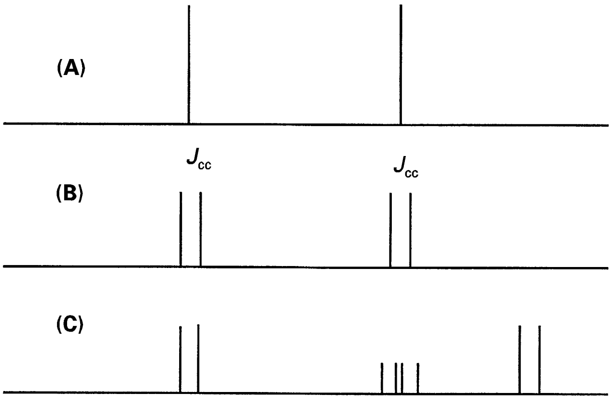



**Figure 1** Diagrammatic representation of $^{13}C$ NMR spectra illustrating carbon–carbon coupling. (A) The spectrum of two connected natural abundance carbon atoms. (B) The spectrum of two directly connected $^{13}C$ atoms. The separation between each doublet is equal to the carbon–carbon coupling constant ($J_{CC}$). (C) The spectrum of three directly connected $^{13}C$ atoms.

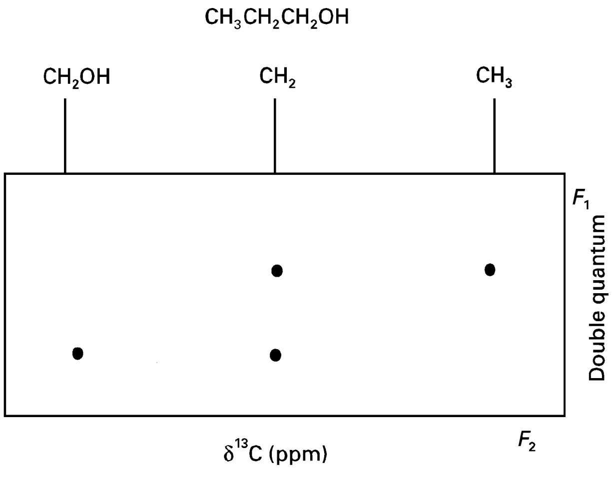



**Figure 2** Diagrammatic representation of an INADEQUATE spectrum illustrating carbon–carbon coupling.

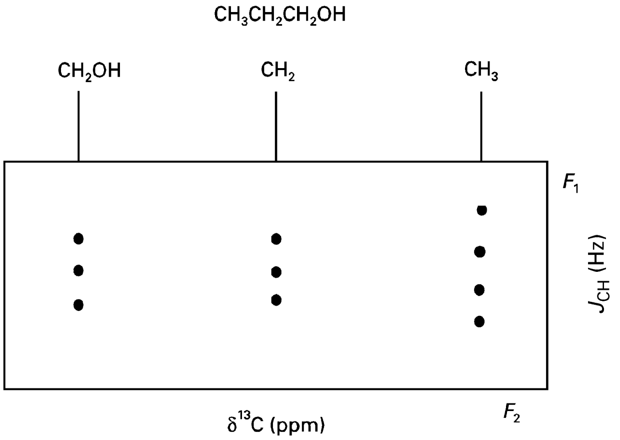



**Figure 3** Diagrammatic representation of a HET2DJ spectrum illustrating carbon–proton coupling.

Two-dimensional $^{13}C$–$^{1}H$ heteronuclear shift correlation allows the detection of the proton signals associated with the labelled carbon atoms. Indirect detection methods are heteronuclear techniques with detection of the nucleus with higher gyromagnetic ratio. Detection of carbon signals through attached protons provides a considerable increase in sensitivity over that obtained in the 1D $^{1}H$-decoupled $^{13}C$ spectrum. For biological samples, the increase in sensitivity can be outweighed by other factors, as recently described by Gemmecker and Kessler. In a complex mixture, for example, a large number of increments (second dimension of a 2D experiment) is required to obtain sufficient resolution of carbon cross peaks to allow the identification of the carbon chemical shift. Also, water suppression techniques must be employed with biological samples.

The chemical shifts of carbon atoms in various environments can be estimated through the use of substituent increments that have been tabulated over many years. Computational programs are also available, some of which use the same logic in the estimation of shift. In addition, database programs can be used to search existing spectral data and to calculate chemical shifts by interpolating from compounds in the database.

In the investigation of intermediary metabolism, the position of labelling in a newly synthesized molecule provides important information about the pathway of its formation. The potential labelling possibilities and the impact that these will have on the NMR spectra need to be considered. For a molecule with $N$ carbon atoms, the number of carbon isotopomers (isomers with differing distributions of $^{13}C$ and $^{12}C$) is $2^N$. In biochemical studies, the potential number of isotopomers possible may not be realized. However, even a small number of possibilities can add considerably to the complexity of the signals detected. For many studies, critical information is extracted from the analysis of isotopomer distribution by NMR and requires the interpretation of complex signals for individual metabolites. Frequently computer simulation of the expected spectra is conducted to enable interpretation.

## Detection of other labels

Deuterium-labelled materials are readily available and can be applied to studying metabolic processes. However, $^{2}H$ has a spin quantum number of 1 and is relatively insensitive compared with $^{1}H$ ($9.7 \times 10^{-3}$). It has a low natural abundance of 0.016%.

Of the two isotopes of nitrogen, $^{15}N$ with a spin quantum number of 1/2 is the most widely used. It is very insensitive ($1 \times 10^{-3}$), and has a low natural abundance of 0.37%. While some studies have used the direct detection of $^{15}N$-labelled compounds, the indirect detection of $^{15}N$ through $^{1}H$ by techniques such as INEPT (insensitive nuclei enhanced by polarization transfer) and HMQC (heteronuclear multiple quantum coherence), which considerably enhance sensitivity, have increased the applicability of $^{15}N$ in biological systems.

## Applications of labels

### Exogenous metabolites

$^{13}C$-labelled tracers have been used in the elucidation of the pathways of metabolism of a number of exogenous chemicals, with labels incorporated in one or several carbon atoms of the materials of interest. The metabolism of [$^{13}C$]formaldehyde, a chemical with a single labelled carbon, provides an uncomplicated illustration of the use of a labelled material in the study of metabolism. The metabolism of formaldehyde to formate, methanol, and several additional metabolites could be determined from the 1D $^{1}H$-decoupled $^{13}C$ spectrum of *Escherichia coli* incubated with [$^{13}C$]formaldehyde (**Figure 4**). From $^{1}H$-coupled spectra, two of these metabolites with signals at 64 and 68 ppm appeared to contained labelled carbons present as $CH_2$ groups. $^{13}C$–$^{1}H$ correlation spectra indicated that these additional metabolites contained protons in the range 3.4–3.7 ppm, with nonequivalent protons on each labelled carbon atom. The metabolism of formaldehyde to produce 1,2-propanediol, 1,3-propanediol, and glycerol was established by culturing *E. coli* in medium containing 14.3% uniformly labelled [$^{13}C$]glucose, administering 90% labelled [$^{13}C$]formaldehyde, and analysing

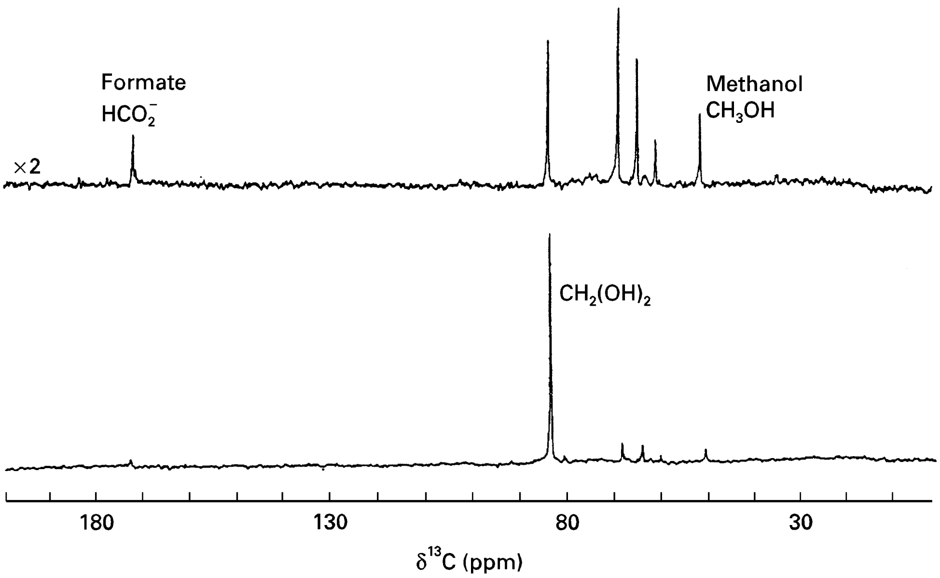



**Figure 4** 1D $^1H$-decoupled $^{13}C$ spectrum of *E. coli* incubated with [$^{13}C$]formaldehyde. Reprinted with permission from Hunter BK, Nicholls KM and Sanders JKM (1984) Formaldehyde metabolism by *Escherichia coli. In vivo* carbon, deuterium, and two-dimensional NMR observations of multiple detoxifying pathways. *Biochemistry* **23**: 508–514. Copyright 1984 American Chemical Society.

the products formed by $^{13}C$ NMR spectroscopy. The $^{13}C$–$^{13}C$ couplings produced permitted the detection of the signals from the glucose-derived carbon atoms adjacent to the formaldehyde-derived carbons. An additional metabolite was determined to be *S*-[$^{13}C$]hydroxymethylglutathione, formed by addition of [$^{13}C$]formaldehyde to glutathione.

$^{13}C$-labelled materials have been used in the elucidation of pathways of metabolism of many chemicals. Information can be obtained from chemicals containing more than one carbon atom but only a single labelled carbon. Following administration of [3-$^{13}C$]-1,2-dibromo-3-chloropropane (81 mg $kg^{-1}$) to rats, 15 biliary metabolites and 12 urinary metabolites were observed. Metabolites were assigned based on chemical shift, proton multiplicity, $^{13}C$–$^1H$ correlation spectra, and comparison with synthetic standards. The metabolic profile proposed for this material is shown in **Figure 5**.

The incorporation of multiple labels into a compound with multiple carbon atoms provides an opportunity for using additional methods for characterizing metabolites. Following administration of [1,2,3-$^{13}C$]acrylonitrile to rats and mice, a number of metabolites in urine were detected and assigned structures based on chemical shift, carbon–carbon multiplicity, carbon–carbon connectivity, and proton multiplicity. Examples of the spectra obtained in urine from rats administered [1,2,3-$^{13}C$]acrylonitrile are shown in **Figure 6**. The multiplets assigned to the labelled carbons of the metabolites result from carbon–carbon coupling (**Figure 6C**). The HET2DJ spectrum shown in **Figure 7** illustrates the determination of the number of protons attached to each labelled carbon atom. The HET2DJ spectra obtained from multiply-labelled compounds provide an additional means of distinguishing signals derived from labelled compounds from natural abundance carbon signals in that the carbon–carbon coupling displaces the multiplets from the centre of the $F_1$ axis. For example, each peak of the doublet labelled 3a (**Figure 7**) has three contours aligned along the *x*-axis, indicating that the signals are from a labelled methylene carbon. The nature of the labelled carbon atoms in metabolites derived from acrylonitrile could be assigned. The identity of the unlabelled substituents of the metabolites could be inferred from the substituent effects on the chemical shifts of the labelled carbons. Most of the metabolites detected were derived from conjugation of acrylonitrile or its epoxide metabolite, cyanoethylene oxide, with glutathione and further metabolism of the glutathione-derived portion of the metabolites. The metabolism of acrylonitrile concluded from this study is shown in **Figure 8**. Quantitation of metabolites was conducted with dioxane added as an internal standard by measuring peak areas under conditions that ensured adequate relaxation of the carbon signals of interest. This combination of techniques has been used to study the metabolism of a number of compounds.

The metabolism of mixtures has received little attention, in part because of the complexity of the problem. NMR spectroscopy with labelled materials provides an ideal tool for studying the metabolism of

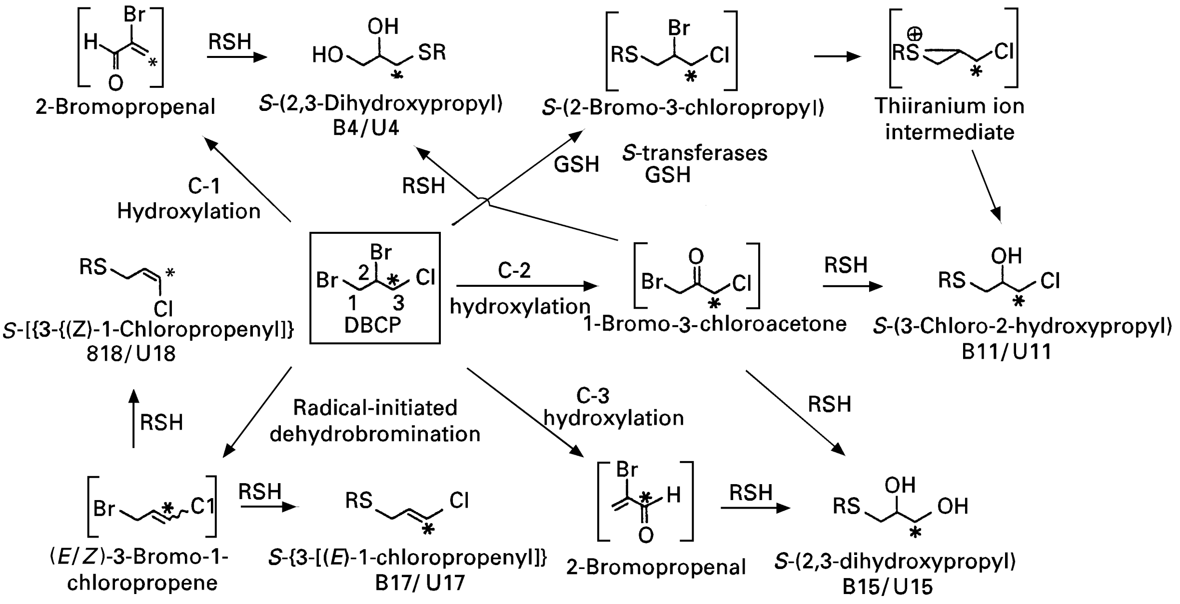



**Figure 5** Metabolism of [3-$^{13}$C]-1,2-dibromo-3-chloropropane in rats. Reprinted with permission from Dohn DR, Graziano MJ and Casida JE (1988) Metabolites of [3-$^{13}$C]-1,2-dibromo-3-chloropropane in male rats studied by $^{13}$C and $^{1}$H–$^{13}$C correlated two-dimensional NMR spectroscopy. *Biochemical Pharmacology* **37**: 3485–3495. Copyright 1988, Elsevier Science.

several components of a mixture. On coadministration of [1,2,3-$^{13}$C]acrylonitrile and [1,2,3-$^{13}$C]acrylamide to rats and mice, the metabolites of both compounds could be resolved readily in the urine.

### Endogenous metabolites

NMR spectroscopy has been widely used in the investigation of metabolism. Therefore, only a very general review of this area can be provided. For additional information, the reader is directed to the references contained in the Further reading section. Most of the recently reported studies involve investigation of the regulation of metabolism. Generally, one or more labelled starting materials are administered in a biological system: subcellular fractions, cells, isolated perfused organs, or whole animals. NMR spectroscopy has been applied to the investigation of metabolism in a wide range of experimental systems from plants to animals and people. Analysis can be conducted at the level of cell extracts, whole cells, body fluids, whole tissues, and *in vivo*. Mammalian tissues that are widely studied include brain, liver, muscle, and heart. The use of multiply labelled compounds aids in distinguishing high levels of naturally occurring unlabelled materials (which give singlet signals) from low levels of added labelled material (which yield multiplet signals) and can provide a means for distinguishing added from endogenous material in determining the size of pools by reverse isotope dilution.

Many pathways of metabolism have been studied using labelled chemicals, including the conversion of glucose to glycogen, glycolysis, the Krebs cycle, fatty acid metabolism, amino acid metabolism, and the pentose phosphate pathway. Among the considerations in conducting studies involving the use of labels in metabolism is the introduction of the label into the biological system, the concentration of label required, distinction of pathways, and isotopomer distribution. Some of these issues can be illustrated with consideration of specific examples.

Oxaloacetate, malate, fumarate, and $\alpha$-ketoglutarate are intermediates in the Krebs cycle that are present in very low concentration and are generally not directly detectable by NMR spectroscopy. Measurement of surrogates that accumulate to measurable levels such as glutamate, alanine, aspartate, and lactate is used for the determination of isotope distribution. Oxaloacetate, malate, and fumarate exist virtually in equilibrium. Glutamate and $\alpha$-ketoglutarate are in fast equilibrium via transamination. Determination of the labelling pattern in glutamate is thought to be reflective of that of $\alpha$-ketoglutarate and, in turn, of oxaloacetate. Similarly, aspartate and oxaloacetate are in equilibrium by transamination, and the labelling pattern of aspartate is thought to reflect that of oxaloacetate.

Succinate and fumarate are symmetrical molecules, and label that passes through these intermediates can become scrambled. Introduction of label from α-ketoglutarate into succinate (**Figure 9**)

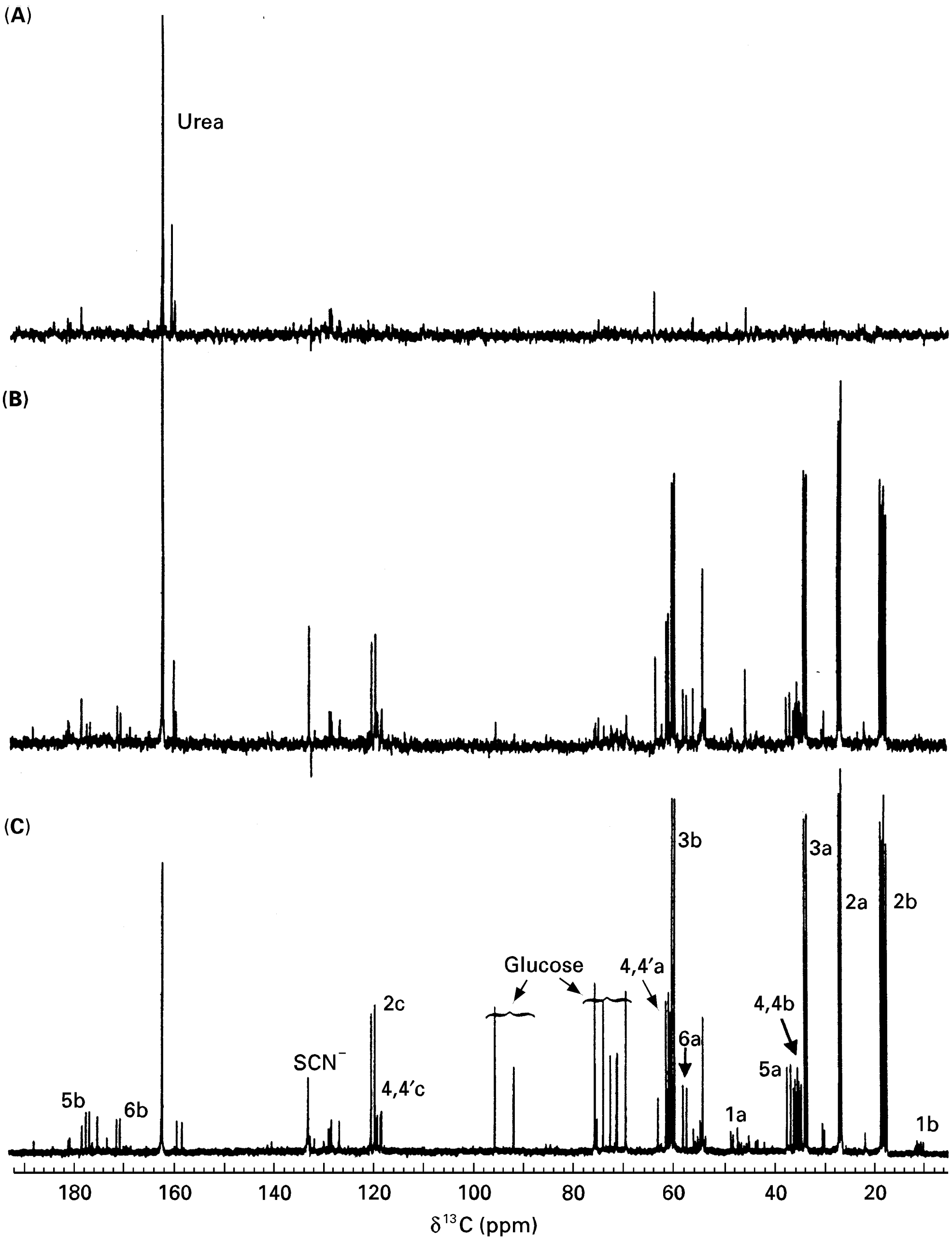



**Figure 6** The $^{1}$H-decoupled $^{13}$C NMR spectrum of (A) control rat urine and urine collected for 24 h following administration of (B) 10 or (C) 30 mg kg$^{-1}$ [1,2,3-$^{13}$C]acrylonitrile. Signals are labelled according to metabolite number (see Figure 8, and the letter of carbon-derived form acrylonitrile ($_{a}CH_2{=}_{b}CH_2{-}_{c}CN$). Reprinted with permission from Fennell TR, Kedderis GL and Sumner SC (1991) Urinary metabolites of [1,2,3-$^{13}$C]acrylonitrile in rats and mice detected by $^{13}$C nuclear magnetic resonance spectroscopy. *Chemical Research in Toxicology* **4**: 678–687. Copyright 1991, American Chemical Society.

would result in the presence of label at more than one site in a molecule. Such mixing can be indicative that the label has passed through a symmetrical intermediate.

Quantitative information can be obtained from isotopomer analysis in one of several ways: modelling with differential equations, modelling with infinite convergent series, and modelling with input–output equations (see the review by Künnecke). Simplified models that focus on a limited number of isotopomers can be used to estimate the contribution of competing pathways.

The incorporation of [1-$^{13}$C]glucose into glycogen can be followed by NMR spectroscopy. Glucose and glycogen give prominent signals in the natural abundance $^{13}$C NMR spectrum of liver. The C-1

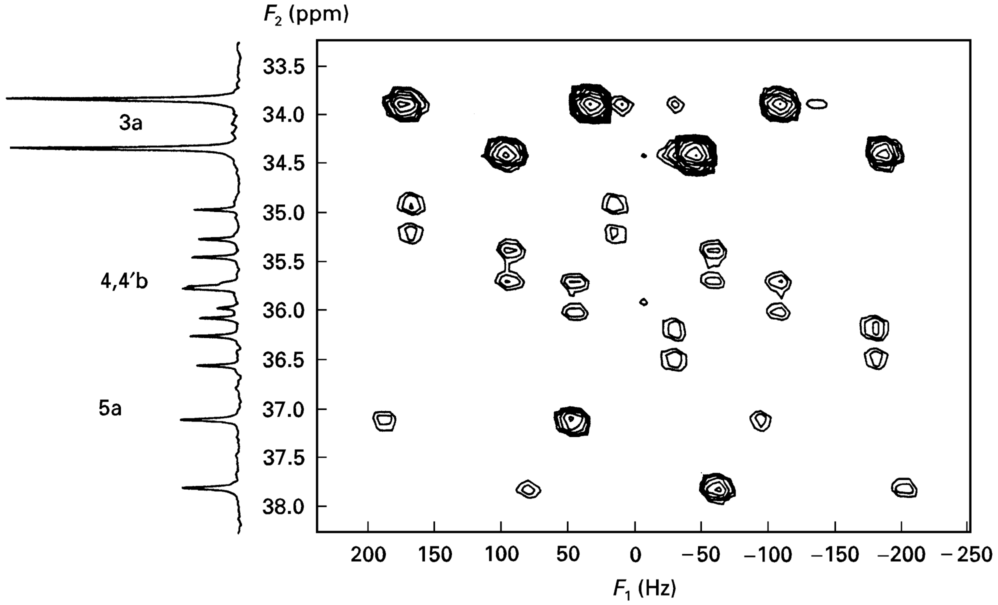



**Figure 7** HET2DJ spectrum of urine from a rat administered [1,2,3-$^{13}$C]acrylonitrile. Reprinted with permission from Fennell TR, Kedderis GL and Sumner SC (1991) Urinary metabolites of [1,2,3-$^{13}$C]acrylonitrile in rats and mice detected by $^{13}$C nuclear magnetic resonance spectroscopy, *Chemical Research In Toxicology* **4**: 678–687. Copyright 1991, American Chemical Society.

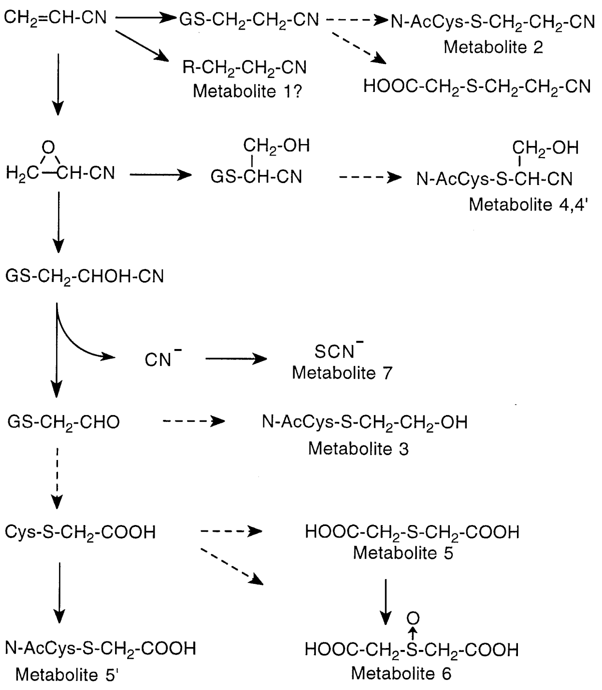



**Figure 8** Proposed metabolism scheme for acrylonitrile in the rat and mouse. GS– represents a glutathionyl residue and Cys–S– and NAcCys–S– a cysteinyl and an *N*-acetylcysteinyl residue, respectively. The broken arrows represent processes that involve several transformations. Reprinted with permission from Fennell TR, Kedderis GL and Sumner SC (1991). Urinary metabolites of [1,2,3-$^{13}$C]acrylonitrile in rats and mice detected by $^{13}$C nuclear magnetic resonance spectroscopy. *Chemical Research in Toxicology* **4**: 678–687. Copyright 1991, American Chemical Society.

carbon signal of glycogen at 100.4 ppm can be readily distinguished from the C-1 signals of $\alpha$- and $\beta$-glucose at 92.6 and 96.7 ppm. Although glycogen is a high molecular mass ($10^9$) molecule with extensive branching, the resonances of glycogen can be readily detected. The internal motion of the glucopyranose

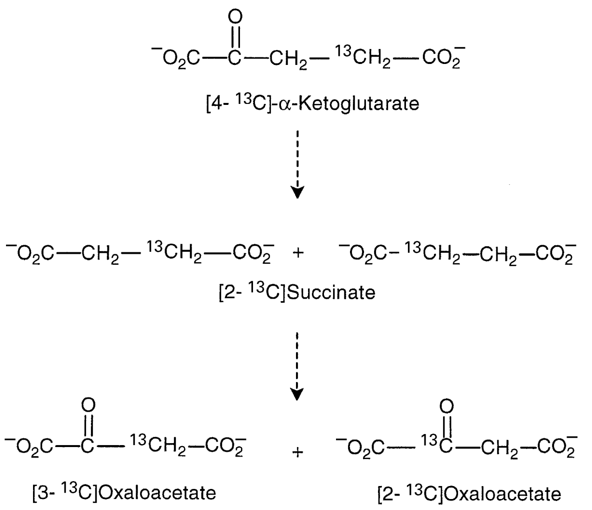



**Figure 9** An illustration of the scrambling of label by passage through the symmetrical intermediate succinate. Incorporation of $^{13}$C in the 4-position of $\alpha$-ketoglutarate results in the formation of [2-$^{13}$C]succinate, and in turn [2-$^{13}$C]- and [3-$^{13}$C]oxaloacetate. Since no molecules labelled at both the 2- and 3-position would be produced, no coupling would be observed between the resonances.

residues of glycogen is thought to be virtually unrestricted resulting in the observation of narrow line widths for glycogen signals. The incorporation of glucose into glycogen can take place by several routes, and not just by the direct conversion of glucose → glucose 6-phosphate → glucose 1-phosphate → UDP glucose → glycogen. Additional pathways involving the conversion of glucose → $C_3$ units → glucose 6-phosphate → glycogen had been extensively studied by NMR spectroscopy. Injection of [1,2-$^{13}$C]glucose into fed rats resulted in the formation of liver glycogen labelled at both the C-1 and C-2 carbons, with spin–spin coupling (**Figure 10**).

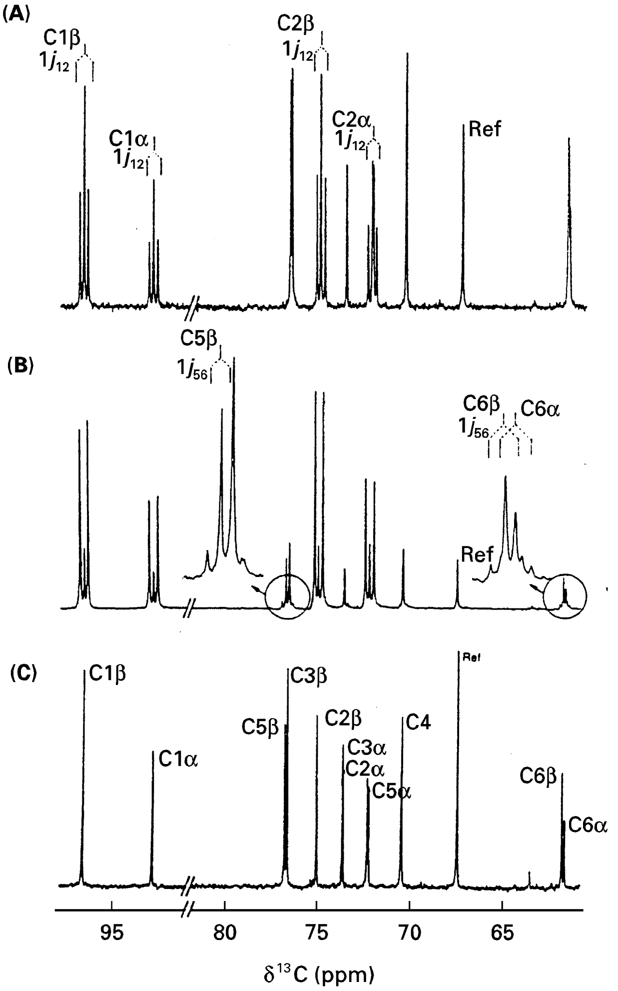



**Figure 10** High-resolution $^{13}$C NMR spectra (100.16 MHz) of enzymatically hydrolysed liver glycogen. Extracted and hydrolysed liver glycogen of (A) a fed and (B) a fasted rat after injection of [1,2-$^{13}C_2$]glucose. Reprinted with permission from Künnecke B and Seelig J (1991). Glycogen metabolism as detected by *in vivo* and *in vitro* $^{13}$C-NMR spectroscopy using [1,2-$^{13}C_2$]glucose as substrate. *Biochimica et Biophysica Acta* **1095**: 103–113. Copyright 1991, Elsevier Science.

In fasted rats, the additional presence of label at C-5 and C-6 indicates the synthesis of glycogen from newly formed glucose. The absence of 1-$^{13}$C- and 1,3-$^{13}$C-isotopomers ruled out the involvement of the pentose phosphate pathway in the synthesis of glucose. In the indirect pathway for glycogen synthesis, glucose is converted in several steps to dihydroxyacetone phosphate, which is rapidly isomerized to glyceraldehyde 3-phosphate. Glyceraldehyde 3-phosphate, in turn, can be used for gluconeogenesis via fructose 6-phosphate or via pyruvate and oxaloacetate. In either eventuality, the $^{13}$C labels can appear simultaneously at the 5- and 6-positions of glucose in glycogen, giving rise to [1,2-$^{13}$C]-, [5,6-$^{13}$C]-, and [1,2,5,6-$^{13}$C]glucose.

## Areas of development

Although few studies have used indirect detection methods for the analysis of labels in biological systems, their use will become more widespread as the technology develops. Examples of the application of these techniques can be found in the literature. Metabolites of [1-$^{13}$C]glucose have been detected in maize root tips using heteronuclear multiple quantum coherence (HMQC) NMR. Labelling of the amide group of glutamine could be detected by $^1$H–$^{15}$N HMQC in root tips labelled with $^{15}NH_4^+$. Following infusion of $^{15}NH_4^+$, $^1$H–$^{15}$N HMQC NMR has been applied to the detection of [5-$^{15}$N]glutamine protons in rat brain.

*See also:* **Biofluids Studied By NMR; $^{13}$C NMR, Methods; $^{13}$C NMR, Parameter Survey; Carbohydrates Studied By NMR; Cells Studied By NMR; Drug Metabolism Studied Using NMR Spectroscopy; $^{19}$F NMR Applications, Solution State; *In Vivo* NMR, Applications, Other Nuclei; Nucleic Acids Studied Using NMR; Perfused Organs Studied Using NMR Spectroscopy; $^{31}$P NMR; Proteins Studied Using NMR Spectroscopy; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals; Tritium NMR, Applications; Two-Dimensional NMR, Methods.**

## Further reading

Beckmann N (1995) *Carbon-13 NMR Spectroscopy of Biological Systems.* San Diego: Academic Press.

Brainard JR, Downey RS, Bier DM and London RE (1989) Use of multiple $^{13}$C-labeling strategies and $^{13}$C NMR to detect low levels of exogenous metabolites in the presence of large endogenous pools: measurement of glucose turnover in a human subject. *Analytical Biochemistry* **176**: 307–312.

Cohen JS, Lyon RC and Daly PF (1989) Monitoring intracellular metabolism by nuclear magnetic resonance. In: Oppenheimer NJ and James TL (eds) *Nuclear Magnetic Resonance. Part B. Structure and Mechanism*, pp. 435–452. San Diego: Academic Press.

Cohen SM (1989) Enzyme regulation of metabolic flux. In: Oppenheimer NJ and James TL (eds) *Nuclear Magnetic Resonance. Part B. Structure and Mechanism*, pp. 417–434. San Diego: Academic Press.

Gemmecker G and Kessler H (1995) Methodology and applications of heteronuclear and multidimensional $^{13}C$ NMR to the elucidation of molecular structure and dynamics in the liquid state. In: Beckmann N (ed.) *Carbon-13 NMR Spectroscopy of Biological Systems*, pp. 7–64. San Diego: Academic Press.

Jans AW and Kinne RK (1991) $^{13}C$ NMR spectroscopy as a tool to investigate renal metabolism. *Kidney International* **39**: 430–437.

Jeffrey FM, Rajagopal A, Malloy CR and Sherry AD (1991) $^{13}C$-NMR: a simple yet comprehensive method for analysis of intermediary metabolism. *Trends in Biochemical Science* **16**: 5–10.

Kanamori K, Ross BD and Tropp J (1995) Selective, *in vivo* observation of [5-$^{15}N$]glutamine amide protons in rat brain by $^{1}H$-$^{15}N$ heteronuclear multiple-quantum-coherence transfer NMR. *Journal of Magnetic Resonance B* **107**: 107–115.

Künnecke B (1995) Application of $^{13}C$ NMR spectroscopy to metabolic studies on animals. In: Beckmann N (ed.) *Carbon-13 NMR Spectroscopy of Biological Systems*, pp. 159–257. San Diego: Academic Press.

London RE (1988) $^{13}C$ Labeling in studies on metabolic regulation. *Progress in Nuclear Magnetic Resonance Spectroscopy* **20**: 337–383.

Lundberg P, Harmsen E, Ho C and Vogel HJ (1990) Nuclear magnetic resonance studies of cellular metabolism. *Analytical Biochemistry* **191**: 193–222.

Nicholson JK and Wilson ID (1987) High resolution nuclear magnetic resonance spectroscopy of biological samples as an aid to drug development. *Progress in Drug Research* **31**: 427–479.

# Laboratory Information Management Systems (LIMS)

**David R McLaughlin**, Eastman Kodak Company, Rochester, NY, USA
**Antony J Williams**, Advanced Chemistry Development Inc., Toronto, Ontario, Canada

Copyright © 1999 Academic Press

**FUNDAMENTALS OF SPECTROSCOPY**
**Methods & Instrumentation**

The purpose of an analytical laboratory is to provide measurement support, generate information, and solve problems. This often involves processing many samples and requests and produces large numbers of test results and reports. In an attempt to increase efficiency, many analytical laboratories have computerized their logbooks, focusing on tracking jobs, samples, tests and results. With this information accessible via a computer, it is then possible to provide many additional functions such as methods management, test results, archiving, calculation, comparison to specifications, control charting and verification of instrument performance. The information is also readily available and valuable for reference, problem solving and modelling. Inclusion of spectroscopic data in such systems has been hampered by a lack of standards and cooperation between instrument vendors and software developers. Some companies have developed successful systems by building customized interfaces and data manipulation software. Advances in Web-based user interfaces suggest that spectroscopy laboratories may soon be able to take full advantage of integrated spectroscopy software, chemical structure handling software and information systems.

## The product of an analytical laboratory

The main product of an analytical laboratory is information. Some of this information is in the form of new analytical techniques or methods that extend a laboratory's ability to analyse new materials. For most laboratories, however, the largest amount of information is in the form of measurements and reports. These measurements are key to understanding and controlling manufacturing processes, solving

manufacturing and development problems and developing new high-performance materials. Preparing samples, making measurements, analysing data and generating reports is the common process analytical laboratories use to generate information. Properly managing this process and tracking the flow of work through the laboratory is critical to the efficient operation of a laboratory.

When computers became available, analytical laboratories began to use them to help automate many of their processes, including the information-handling and information-generating processes. Analytical chemists and information specialists working to improve their own operations were the first to develop laboratory information management systems, or LIMS. They typically supported only one technique or small unit of a laboratory and did not communicate with other information systems.

The first commercial LIMS arrived on the market in about 1980. Today, there are dozens of LIMS vendors providing packages with many capabilities. There is even a World Wide Web site devoted to providing information on LIMS resources, analyses of the latest technology, and the products and services available from vendors and consultants. In 1994 the American Society for Testing and Materials published a standard guide on LIMS. The guide is intended for anyone interested in LIMS, including users, developers, implementers and laboratory managers. It defines standard terminology, a concept model, primary functions, an implementation guide and a checklist. These are useful for discussing and understanding LIMS as well as developing specifications and cost justification. The guide is a good reflection of what it was thought a traditional LIMS should be and of what most vendors have marketed.

## Types of laboratories

The information-handling requirements of analytical laboratories vary from the research to the manufacturing environments. A standard testing laboratory that supports manufacturing quality control and regulatory compliance uses analysis procedures that are well defined and strictly followed. These laboratories analyse relatively large numbers of samples and produce numerical or tabular results that are often plotted against time. Spectral analysis is usually used in this environment to determine concentration or verify identification relative to known reference spectra.

In contrast, a research analytical laboratory supporting the discovery of new materials uses analysis procedures that are defined in the mind of the analyst. While the analyst may use analytical techniques that are well defined, the process applied to a particular sample is not predetermined, but is developed in response to the questions that need to be answered. These laboratories analyse fewer samples and regularly produce results that combine graphics and images with textual reports. In this environment, spectroscopy is also used to elucidate chemical structure and determine three-dimensional spatial relationships.

Other types of analytical laboratory environments vary between these two idealized cases. Process control environments model a fully automated standard testing laboratory. High-throughput screening is similar, but often includes special processing for unusual results or potentially interesting materials. Like research, problem solving does not follow a fixed analytical process, but often results must be expressed in manufacturing terms. Development environments reflects the need to move research testing for understanding to standard testing for manufacturing. Problems arise when a LIMS vendor tries to sell one generalized solution, customized to fit all environments.

## Selecting a LIMS

Many factors drive analytical laboratories to implement LIMS. These include the need to demonstrate value, comply with detailed regulatory requirements, manage large amounts of samples and tests from automated analysis and screening systems, automate laboratory processes, improve accuracy and deal with increased demands for efficiency and documentation. Implementing a LIMS can be expensive and have significant impact on laboratory workflow and productivity. For successful implementation of a LIMS, it is important to know which factors are most important and where productivity will be affected, both positively and negatively. It is also extremely important that the laboratory workflow is supported by the LIMS in a simple and appropriate manner. It may be appropriate to utilize a re-engineering process to simplify a laboratory's workflow before implementing a LIMS.

Even with useful guides and books and the availability of highly functional LIMS and capable consultants, successful installation of a LIMS is a difficult and frustrating task for some laboratories. Vendors have generally focused on developing generic systems that provide a great deal of functionality but have grown to be quite complex. These products have worked quite well in some laboratory environments and have completely failed in others. They work best in laboratories where the operation is well

understood and standardized and where the processes do not have to be changed to match the LIMS. This is generally not the case in research and development environments, where microscopy and molecular spectroscopy laboratories usually reside. The need to manage images, various types of spectra and chemical structures adds to the problem. Since LIMS vendors generally do not consider these types of data, significant customization and software development is required. This task is easier with LIMS that utilize recent improvements in computer technology.

Managers may still have a hard time finding a system that will work for their operations and meet the expectations of users. Often, the laboratory will have to change its workflow to match that of a LIMS package or deal with additional expense when significant customization is required. The business case for the buy-and-customize route versus the build route can be unclear. While the general trend has been towards purchasing commercial systems, recent surveys examining potential year-2000 problems indicate that most LIMS are still internally developed and maintained.

## Functions of a LIMS

Whether they support manufacturing, research, or problem solving, all analytical laboratories receive and prepare samples, make measurements, analyse data, and report results and information. A simplified illustration of the analytical support process is shown in **Figure 1**. Improving the overall efficiency of this process is the primary function of a LIMS. It can be used to support only a few aspects of this process or it can support all aspects. The most common and perhaps most important use of a LIMS is to maintain proper records of samples and the tests requested on them. This electronic logbook or sample management type function assigns sample numbers or bar codes and records administrative information, often including an indication of the work requested. Most LIMS also allow numerical and tabular results to be entered for tests performed on the samples. Once data have been collected and properly recorded, the LIMS may be used for data manipulation and automated report generation. Significant productivity enhancements are usually obtained from this last step. The following sections describe how a LIMS is used to support the steps in **Figure 1**.

### Definition of problem or job

For the laboratory supporting research, problem solving or method development, the definition of a problem or job is the first step of the analysis process. Job requests could include determining what is happening to a material as it degrades, how the properties of a material might be improved, or whether a fabrication process is producing devices with the desired composition and geometry. When a new chemical is to be commercialized, the request may be to develop tests and fitness for use criteria needed for production monitoring. The job usually defines the context for investigative analytical work performed on many samples. This information is usually stored in laboratory notebooks with only a job title stored in LIMS. When the work is complete and a report is written, this information may be made available in an electronic library. There is an opportunity for the field of knowledge management to improve productivity in this area.

In a standard testing laboratory, the problems or jobs are predefined by the time materials are analysed. Given a type of material and its point-in-process, the suite of required tests, testing methods, specification limits and resulting actions and reports are clearly defined. This information is often stored in LIMS as master or static data and used to automate the workflow for these materials.

### Sample log-in

A primary function of the LIMS is to track the samples of materials to be analysed by a laboratory. This can be as simple as tracking when a sample enters and leaves a laboratory or as detailed as tracking every action that is performed on the sample while it is in the laboratory. The process normally begins with the client or analyst recording general descriptive information about each new sample or batch of samples and the work request. This is ideally entered into or created by information systems used by the clients of the laboratory and interfaced to LIMS, but is normally entered manually on an input screen. Client systems include electronic laboratory notebooks used by researchers or manufacturing scheduling and control systems. Typical information includes the name of the requester of the work, an account number to be charged for the work, the date the request is made, the client's identifier for the sample, the date the sample is received, the job or project to which this sample belongs, the priority, and the hazard information. Test requests vary from a general description of the problem to be solved to a suite of specific point-in-process tests using very specific methods.

### Sample receipt and review

Samples may be logged into a LIMS before they arrive in the laboratory. This is quite common in

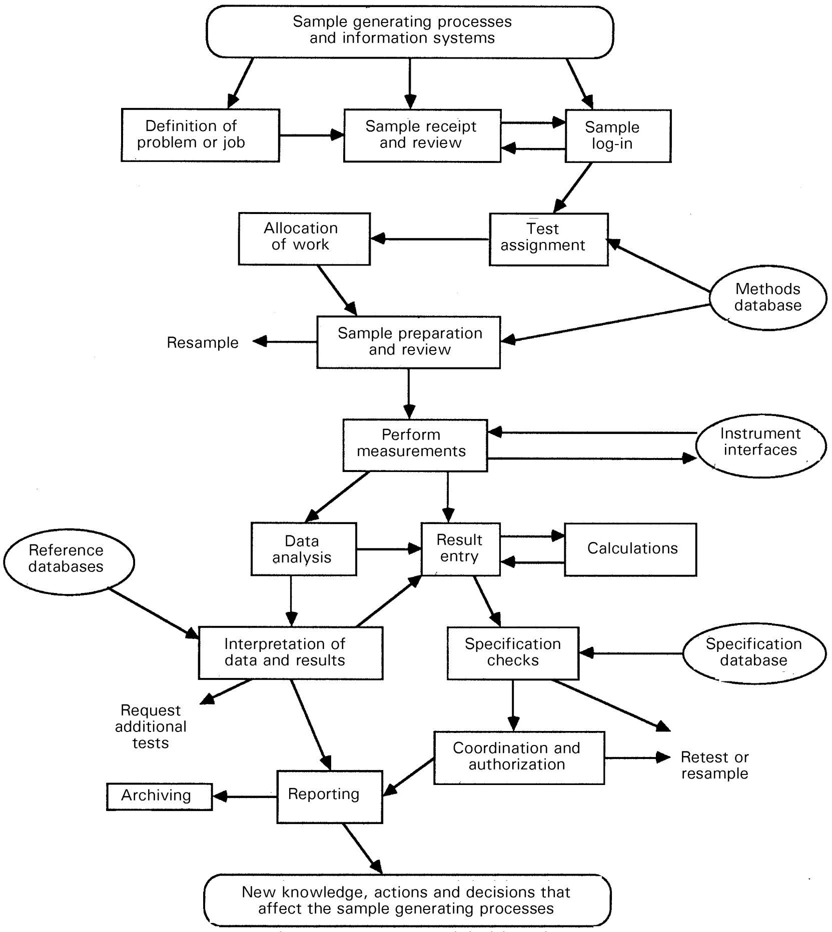



**Figure 1** Generic analytical laboratory workflow supported by LIMS.

standard testing laboratories with a LIMS that is integrated with their client's manufacturing resource-planning system. Good-quality operating procedures dictate that when a sample arrives it should be reviewed to ensure it is suitable for analysis, that the requested work is appropriate and that it has not visibly degraded. The LIMS sample receipt function allows the time a sample arrives and any anomalies to be recorded. If anomalies are observed in the sample, then a decision to proceed, resample or request other tests is made, often after consultation with the client.

### Test assignment

The client may request a specific collection of tests when the sample is logged in. In a standard testing environment, the suite of tests is often assigned automatically by the LIMS. If a general test is requested or the request is for some problem to be solved or investigated, then the appropriate specific tests and course of investigation are selected by the analyst.

### Allocation of work

Work may be allocated among specific instruments, analysts or support groups. This is done to match the test requirements to appropriately skilled analysts and equipment and ensure the best utilization of resources. The differing priority levels of samples received in a laboratory can also be accommodated in this step. When work is allocated, it is recorded in LIMS so that it can generate work lists for the analysts and instruments. LIMS can often make default assignments automatically. In other cases, analysts simply check out the next pending sample from the queue in LIMS. In a research or problem-solving area, this may not be recorded in a LIMS until the results are entered.

### Method management

Methods are the procedures and protocols by which materials are analysed. LIMS designed for standard testing laboratories often provide some support for

managing methods. This can allow complete automation of the analytical process, as exemplified by on-line analysis.

### Sample preparation

Most samples and analysis techniques require some preparation such as dissolving, diluting or mounting before measurements can be made. This step and the subsequent measurement step are a typical focus for laboratory automation using robotics. Such systems follow detailed procedures that may vary for every sample. Associating the correct procedure with each sample and supplying that information to the robotic systems is logically a job for LIMS. Most automated sample preparation and measurement systems were initially isolated, performing very focused tasks and requiring manual data transfer at the beginning and end of a run. This is changing as automation and LIMS continue to mature.

### Perform measurements

Measurements are performed using instruments. In an automated laboratory, the LIMS can download data acquisition instructions to the instruments and receive data and analysis results for entry into its database. Most LIMS provide some capabilities for interfacing with instruments. This has been particularly successful in chromatography, where good standards have been developed and followed by both LIMS and instrument vendors.

### Result entry

After the tests have been performed, the results are often entered into the LIMS. This is a manual process, unless a specific instrument interface has been built or purchased. Many instrument vendors now provide such interfaces, but this has been limited in spectroscopy. Results entered into LIMS are generally derived from processing raw analysis data on the instruments. There can be simple yes or no phrase results to tables of amounts of various materials. Usually these tables have to be predefined by the method that is used to perform the analysis. Interfacing efforts in spectroscopy have been limited partly by the lack of widely accepted standard spectral formats and partly by the lower demand for general spectroscopy applications in standard testing areas.

### Calculations and data analysis

Most LIMS systems provide simple capabilities for performing calculation on the results that are entered. These calculations may simply be transformations into standard units or the generation of averages. More sophisticated analysis is generally performed using instrument vendor data systems or specialized data analysis software. Most LIMS provide customization functions to facilitate entry of these results.

### Specification checking

Many standard testing laboratories run analyses to determine whether a material is within manufacturing specification limits. The manufacturing process and analysis method usually define these limits. Most LIMS provide a mechanism for these specifications to be stored or retrieved and compared to the results. The specifications are often stored in a database controlled by the clients of the analytical laboratory. Materials that are within specifications can be automatically released. Those that are not can be flagged or operators can be automatically notified regarding the test results.

### Coordination and authorization

In some environments, it is important to validate results before they are released. This usually involves a person other than the analyst checking the results and verifying the procedures. The LIMS can enforce such rules to help ensure that the business processes are followed. In environments where several tests are performed on one material, this role may be to coordinate reporting to the client and to identify and resolve inconsistencies. This activity may result in requests for materials to be retested or resampled.

### Interpretation of data and results

The result of research processes is new knowledge. This is generally in the form of written documents with graphics and images that may involve many samples. A LIMS that supports analytical research needs to support the storage of these kinds of results as well as numerical and tabular results.

### Report generation

The output of an analytical laboratory is information that is reported to the client and often affects further actions and decisions. A major function of a LIMS is to improve the efficiency of report generation. In order to facilitate this process, report templates can be established that are automatically triggered when a step in the workflow is completed. Analysis results, billing, turnaround time, instrument calibration and control charts, justification and inventory are just a few of the common reports that can be triggered. Because of the variety of clients that analytical laboratories interact with, report

generation is an area where significant customization is invested to make a LIMS effective.

### Archiving

The results and reports from an analytical laboratory can be of great economic value. Reference information, historical understanding of processes, solutions to problems and assigned spectra from structure elucidation efforts represent a few of the valuable resources that need to be maintained, sometimes indefinitely. This information is usually best stored in a knowledge management system specifically designed for this purpose. Laboratories also produce information needed to support regulatory compliance. Organizations may be required to save this information for many years. Archiving of this data is generally supported by LIMS.

### Summary

The analytical support process exists to aid and enhance the sample generation process. If it cannot be measured, then it is hard to improve it. Making the information produced by this process available electronically can significantly enhance its value to an organization. Having measurements available and logically organized is required for any material or process modelling. Quick and easy access avoids duplication of work and speeds research and problem solving. Computerization of the process requires consideration of interfaces to other information-storing or information-generating systems. These include methods management, manufacturing resource planning and control, electronic laboratory notebooks, instrument interfaces and reference databases.

The power of LIMS lies in its ability to automate mundane tasks and to work with other systems to enhance the access of authorized users to information and the proper maintenance of that information for future reference.

## LIMS for spectroscopy

Ever since computers became of a practical size to fit in a spectroscopy laboratory, they have had an ever more intimate relationship with the spectrometers themselves. Spectrometer vendors have applied them to instrument control, data collection, library searching and information management. The information management functions of these systems could be considered a LIMS with a focus limited to a particular type of testing. With the exception of chromatography, these systems are limited in scope and do not interface well with laboratory-wide LIMS. Vendors of spectral analysis packages such as *Grams* (Galactic) and *Spectacle* (LabControl) similarly provide some limited LIMS functions or the ability to link with spreadsheets where such functions may be built.

To aid in sharing spectra obtained using equipment from different vendors, a number of efforts have been made to establish standard exchange formats. In 1987, the Joint Committee on Atomic and Molecular Physical Data (JCAMP) published the JCAMP-DX format as a standard for exchange of infrared spectra. This general format was subsequently extended to include mass and NMR spectra. Although JCAMP-DX files are created with small variations from vendor to vendor, it is supported as an export format by most infrared and mass spectroscopy instrument vendors.

The Analytical Instrument Association (AIA) created a netCDF-based Analytical Data Interchange (ANDI) format for chromatography that received widespread acceptance. After this success, the AIA adopted a standard for mass spectroscopy in 1993 and began definitions for infrared and NMR. This mass spectroscopy standard, although supported by a few vendors, has not received wide usage and the infrared and NMR definitions have not been implemented.

There is currently no generally used spectral format in NMR. However, the International Union of Pure and Applied Chemistry (IUPAC) has a working party that expects to produce a JCAMP-based multidimensional NMR data standard by March 2000.

Despite the existence of these standard formats, spectrometer vendors have been slow to fully support them as export options. This has been a hurdle for spectral analysis and database software vendors, who have been required to deliver format converters. Spectrometer manufacturers often consider these vendors as competition. Until spectrometer users demand better integration by the vendors, progress will be slow and the lack of widely accepted formats for spectral data will remain a barrier to the general integration of spectroscopy into LIMS.

### Some corporate solutions

In the early 1970s, Wolfgang Bremser and others at BASF began to develop software to aid in spectral interpretation and chemical structure elucidation. Their efforts were focused on improving speed, cost and reliability by using databases of reference spectra to provide reasonable suggestions and dialogue for the interpreting spectroscopist. This work resulted in the very useful SpecInfo system, which is still being enhanced at BASF. This system and associated databases were the basis for formation of the

company Chemical Concepts in 1990, which now sells a variety of advanced interpretation tools.

Eastman Kodak Company began a similar effort in the late 1970s. The initial goal was to improve the efficiency and quality of structure elucidation using NMR, IR, MS and UV-visible data with the ultimate aim of completely automating the analysis of routine samples. Such a system would automatically determine whether the structure a chemist proposed for a sample was consistent with the spectral data obtained. It was soon recognized that high-quality databases relevant to Kodak materials were needed and that interpretation software operated in isolation would be of limited utility. The proposed structure and resulting spectral data needed to be available from a LIMS. Integrated sample management software that incorporated chemical structures and spectral data was developed to solve this problem and to help minimize the task of building an in-house database. The complete system, called QUANTUM (**Figure 2**), combines the spectral and structural analysis software with a sample management system (SoftLog) and a spectral database (SDM).

SoftLog was probably the first LIMS designed specifically for spectroscopists. It incorporated several important features not typically available in LIMS. These included full and substructure searches, spectral display and search, easy incorporation of results into the reference database and a logical interface to fully automated analysis software. In addition, the user was automatically informed of previous sample or material analyses, of inconsistent results, of reference data and of related compounds such as impurities, models, by-products or precursors. The notification of previous analyses of the same material included data generated by spectroscopy laboratories dispersed across the entire corporation. This was particularly useful, because it made analysis information initially obtained in a research environment available to analytical laboratories supporting development.

## The Web is the way

Since the mid-1990s the world has experienced explosive growth of the World Wide Web or Internet. This growth was primarily caused by the introduction of graphical web browsers that provide a simple and intuitive point-and-click environment for access to information. The Web is based on widely accepted and robust standards that are open and simple enough that everyone can participate. It supports most kinds of information (text, graphics, audio, video, binary, etc.) and provides a simple mechanism for adding specialized support. Inexpensive or free graphical browsers are available on virtually all computer platforms commonly used in

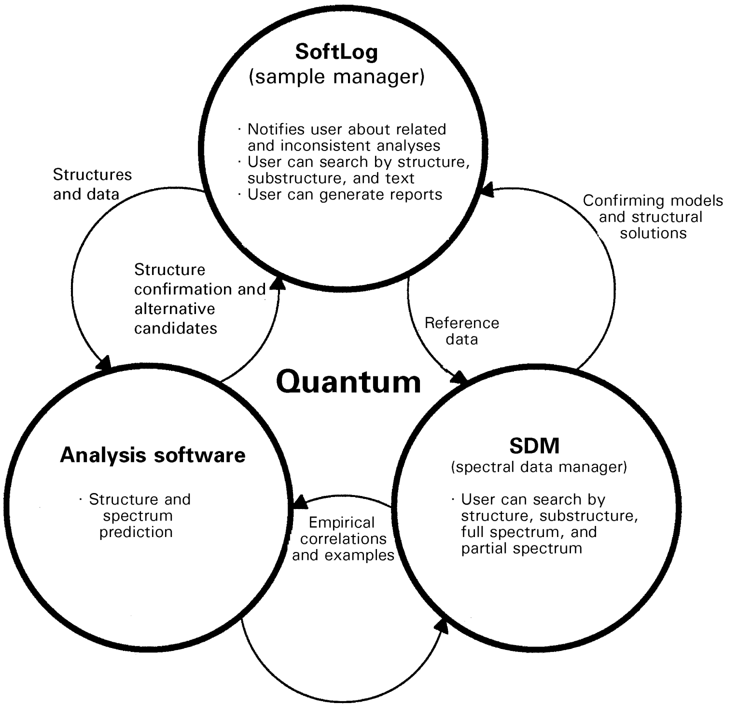



Figure 2 Components of an integrated spectroscopy information management environment.

analytical laboratories. This is valuable in spectroscopy or other laboratories that typically utilize multiple computer platforms to control spectrometers and other instruments. With this paradigm shift in user interface design has come a new level of comfort for chemists, researchers, and analysts in managing information electronically.

A few companies have reported on their efforts to automate their LIMS and spectroscopy environments. Scientists in the analytical research laboratory at Parke-Davis Pharmaceuticals provide NMR and mass spectral analysis support for synthetic chemistry by combining robotic automation with internally developed interfaces to their Beckman Lab Manager LIMS. The LIMS provides preparation conditions to the sample preparation robot and a work list for batch operation of the spectrometer. These workers have used Web technology to provide information on the status of sample analyses and other functions of the analytical laboratory. Web technology has also been used to develop sample tracking and data analysis in support of genome mapping. SmithKline Beecham Pharmaceuticals, Eli Lilly, NASA Lewis Research Center and some oil companies have reported on integration of LIMS with spectroscopy.

The analytical laboratories at Kodak have converted a major portion of its LIMS to a Web-based system. This work began with the development of their Web-based Information Management System, WIMS, in the molecular spectroscopy areas. This integrated Web-based system provides access to sample information including reports, structures and spectra. Sample information is integrated with a Web interface to QUANTUM for structure management needs. Spectra from various instruments are copied as they are processed to a server. Structures and spectra are retrieved and displayed as GIF images by simple C programs and are available for download to the desktop for further analysis by local applications. Reports of analysis can be attached as text, as HTML or as standard word-processor or spreadsheet documents that can be viewed using helper applications. Numerical and tabular results are stored in a relational database linked to the rest of the system with a Web interface developed with one of the several good tools available.

Using the many distinct advantages of Web-based technology as well as its fast growth curve, commercial LIMS vendors are presently developing Web interfaces for their LIMS. The one product with a focus on spectroscopy laboratories is the $S^3$LIMS product from Advanced Chemistry Development. This LIMS, inspired by WIMS, is a fully Web-based integration of sample, structure and spectral management utilizing Java-based structure drawing and spectral display applets. The Java implementation and compatibility with Netscape and Microsoft browsers ensures a platform independence for the system and marries the system with the proton NMR, carbon NMR and other structure-based prediction engines also available from Advanced Chemistry Development. Chemical Concepts, now owned by John Wiley & Sons, has also developed a fully web-based 'version' of their interpretation tools.

## Conclusions

When selecting or developing a LIMS, a laboratory should be clear why the system is needed and what workflow and business functions it must support. Keeping the installation simple in scope, design and, most importantly, user interface will help ensure success. As the laboratory's business changes over time, so must its LIMS. Continued investment in information management must be expected.

LIMS is still an evolving concept. Most commercial LIMS have focused on standard testing, perhaps because the processes are more clearly defined or standardized and perhaps because the market is larger. Recent developments in knowledge management and electronic notebooks systems may help develop better systems for research environments. Web technology allows information system interfaces to be developed more simply and more quickly. There is increasing interest from pharmaceutical companies in using spectroscopy with high-throughput screening. With these technology developments and increased market demand, LIMS with better integration and simpler interfaces for spectroscopy may become ubiquitous.

*See also:* **Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals.**

## Further reading

ASTM (1994) *E1578-93e1 Standard Guide for Laboratory Information Management Systems (LIMS).* West Conshohocken, PA: American Society for Testing and Materials.

Hinton MD (1995) *Laboratory Information Management Systems: Development and Implementation for a Quality Assurance Laboratory.* New York: Marcel Dekker.

LIMSource (1998) *LIMSource: the best site for LIMS and lab data management system info.* http//:www.limsource.com.

Mahaffey RR (1990) *LIMS: Applied Information Technology for the Laboratory*. New York: Van Nostrand Reinhold.
McDowall RD (ed) (1987) *Laboratory Information Management Systems, Concepts, Integration, Implementation*. Wilmslow, UK: Sigma Press.
Nakagawa AS (1994) *LIMS: Implementation and Management*. Cambridge, UK: The Royal Society of Chemistry.
Stafford J (ed) (1995) *Advanced LIMS Technology: Case Studies and Business Opportunities*. Andover, UK: Chapman & Hall.

# Lanthanum NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# Laser Applications in Electronic Spectroscopy

**Wolfgang Demtröder**, University of Kaiserslautern, Germany

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
| --- |
| Applications |

## Introduction

The applications of lasers to atomic and molecular spectroscopy has revolutionized this field. Although methods in classical spectroscopy of electronic transitions using large grating spectrographs or interferometers have brought a wealth of information on molecular structure and dynamics, the new techniques of laser spectroscopy have by far surpassed the capabilities of former methods with regard to sensitivity and spectral or time resolution. This article reviews these new spectroscopic techniques, their applications and some of the results obtained so far; it will be restricted to the spectroscopy of free atoms and molecules and does not include solid-state spectroscopy.

Various techniques of sensitive absorption spectroscopy, including nonlinear techniques, which allow a spectral resolution below the Doppler width are described first. These techniques are termed sub-Doppler-spectroscopy and include linear spectroscopy in collimated molecular beams, nonlinear saturation and polarization spectroscopy, and Doppler-free two-photon spectroscopy. Emission spectroscopy, which covers laser-induced fluorescence as well as stimulated emission methods, is described next. The assignment of complex molecular spectra is greatly facilitated by optical-double resonance methods, which are treated together with optical pumping. The development of ultrashort laser pulses in the picosecond to femtosecond time domain has opened a wide area for investigations of molecular dynamics, intramolecular and intermolecular energy transfer, fast collision-induced relaxations, and detailed real-time studies of chemical reactions by observation of the forming and breaking of chemical bonds. This fascinating and rapidly expanding field of femtosecond chemistry is covered briefly. The final section discusses some aspects of coherent spectroscopy with pulsed and CW lasers that rely on the coherence properties of the laser light and on the coherent preparation of atomic or molecular states. Techniques such as quantum beat spectroscopy, correlation spectroscopy and STIRAP are described and their relevance for the electronic spectroscopy of atoms and molecules is outlined.

## Sensitive absorption spectroscopy

In classical absorption spectroscopy, the intensity $I_T$ of a light beam transmitted over a path length $L$ through a sample of absorbing molecules with absorption coefficient $\alpha$ ($cm^{-1}$) is compared with the

incident intensity $I_0$. According to Beer's law of linear absorption,

$$I_r = I_0 \exp(-\alpha L) \approx I_0(1 - \alpha L) \qquad \text{for } \alpha L \ll 1 \quad [1]$$

the absorption can be written as

$$\alpha L \approx \frac{I_0 - I_T}{T_0} \quad [2]$$

The smallest still detectable absorption is limited by intensity fluctuations of the incident intensity $I_0$ and by detector noise. The absorption coefficient $\alpha = N_i\sigma_{ik}$ is the product of number density $N_i$ of molecules in the absorbing level $|i\rangle$ and absorption cross section $\sigma_{ik}(\nu)$ for the transition $|k\rangle \leftarrow |i\rangle$ The minimum detectable number $N_i$ of absorbing molecules

$$N_i = \frac{\Delta I_{min}}{I_0\, L\, \sigma} \quad [3]$$

is then given by the minimum difference $\Delta I_{min} = I_0 - I_T$ that still can be safely measured. For a given value of $\Delta I_{min}$, the detection sensitivity can be increased by making the product $I_0 L\sigma$ as large as possible. The various sensitive techniques of laser absorption spectroscopy attempt to minimize $\Delta I_{min}$ as well as to maximize $I_0 L\sigma$.

### High-frequency modulation spectroscopy

For high-resolution laser absorption spectroscopy, a narrow-band tuneable CW laser can be used. Possible candidates are dye lasers, Ti:sapphire or other vibronic solid-state lasers, and the variety of tuneable semiconductor diode lasers, which cover a spectral range from the blue to the mid-infrared. Of particular importance are the optical parametric oscillators that have been brought to reliable operation in the pulsed as well as in the CW mode with single-mode performance. They can now span the spectral region from 500 to 5000 nm.

When the output of such a tuneable monochromatic laser is sent through an electrooptic modulator, the phase and thus the frequency of the transmitted laser wave is modulated and the frequency spectrum consists of the carrier and two side-band frequencies, where the side-bands have opposite phases (**Figure 1**). At a modulation frequency of $\Omega = 1$ GHz, the separation of the sidebands from the carrier is about equal to the Doppler width of the absorption lines. If this modulated laser beam is detected after transmission through the absorption cell, the

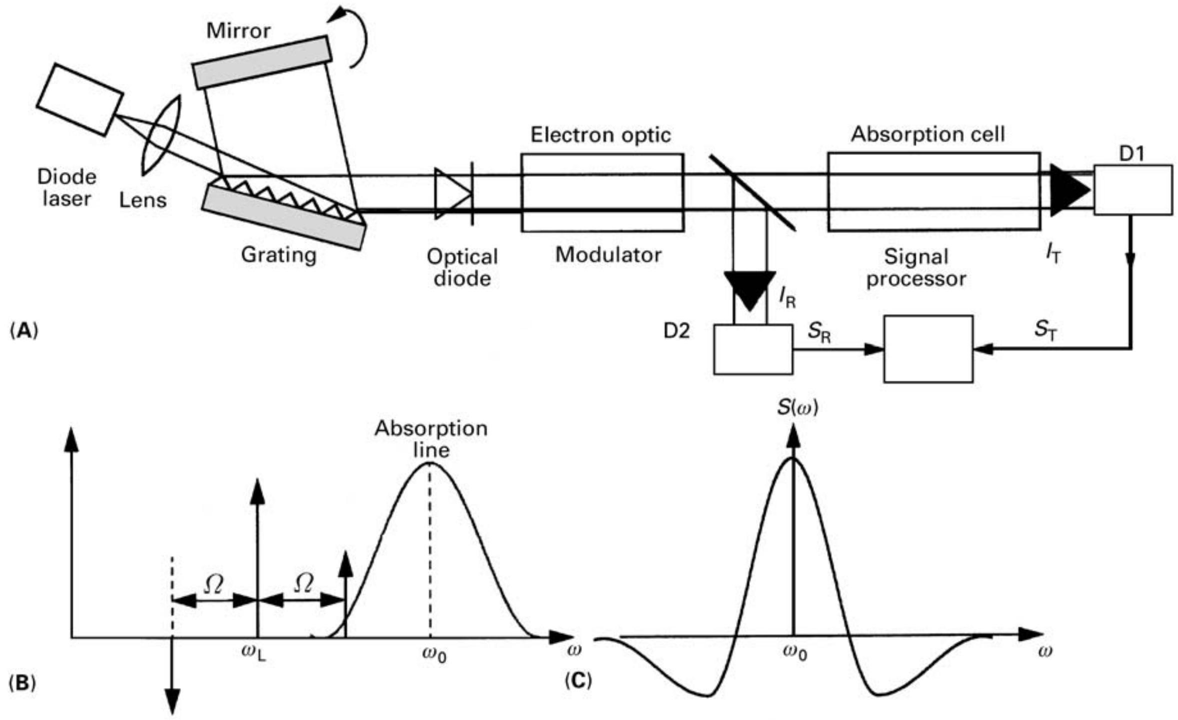



**Figure 1** Principle of modulation spectroscopy. (A) apparatus; (B) sidebands in of phase modulation; (C) line profile of the signal $S(\omega)$.

output of a lock-in, tuned to the frequency $\Omega$, will not show any signal as long as neither the carrier nor the sidebands overlap with an absorption line, since the beat signals between carrier and the two sidebands exactly cancel. Note also that any fluctuations of the laser intensity are automatically subtracted. If, however, one of the three frequencies coincides with a molecular absorption line, the exact balance is perturbed and a detector signal is generated. When the laser frequency $\omega$ is tuned over an absorption line, the measured signal profile $S(\omega)$ is similar to the second derivative of the absorption profile $\alpha(\omega)$.

Since lock-in detectors now available cannot measure frequencies above 100 MHz, the output of the detector is heterodyned with a stable local oscillator of frequency $\Omega_0$ and the beat signal with frequency $(\Omega - \Omega_0)$ is sent to the lock-in.

The sensitivity of this modulation spectroscopy is two to three orders of magnitude higher than in conventional spectroscopy and reaches a minimum detectable absorption of $\alpha L < 10^{-7}$. Using multiple-path absorption cells (Herriot cells) and balanced detectors for detection of reference and transmitted beams, absorption coefficients as low as $\alpha \leq 10^{-10}$ cm$^{-1}$ can still be measured.

### Excitation spectroscopy

Instead of measuring the transmitted laser intensity, the photons absorbed by the sample molecules can be directly detected because they excite the absorbing molecules into a state $|k\rangle$ that emits fluorescence (**Figure 2**). The rate of detected fluorescence photons is

$$\dot{n}_{\text{Fl}} = \dot{n}_{\text{abs}} \cdot \eta_{\text{m}} \cdot \varepsilon \cdot \eta_{\text{k}} \qquad [4]$$

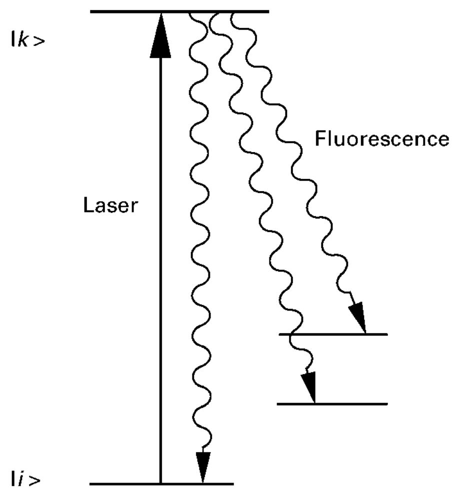



**Figure 2** Level scheme of excitation spectroscopy.

where $\dot{n}_{\text{abs}}$ is the rate of absorbed laser photons, $\eta_{\text{m}} \leq 1$ is the quantum efficiency of the excited atom or molecule, $\varepsilon$ is the collection probability of the emitted fluorescence and $\eta_{\text{k}}$ is the quantum efficiency of the detector. With 1W incident laser power at $\lambda = 500$ nm, a collection efficiency of $\varepsilon = 0.1$, quantum efficiencies $\eta_{\text{m}}=1$ and $\eta_{\text{k}} = 0.2$ and an absorption probability of $\alpha L = 10^{-15}$, we obtain a photon counting rate of 60 photons s$^{-1}$, which gives a signal-to-noise ratio of 6 at a time constant of 1 s and a dark current of 10 photons s$^{-1}$.

For selected atomic transitions, the upper level may decay only into the initial lower level (true two-level system). In this case the same atom can be excited many times during its flight time T through the laser beam. If its upper-state lifetime is $\tau$, the number of fluorescence photons emitted per second may be as high as $T/\tau$ (photon burst), if the laser is intense enough to saturate the transition.

### Ionization spectroscopy

The absorption of a photon in the transition $|i\rangle \rightarrow |k\rangle$ can be also detected by ionizing the molecule in the excited state $|k\rangle$ by a second photon (**Figure 3**). If the intensity of the second laser is sufficiently high, the ionizing transition may be saturated, which implies that every excited molecule is ionized. Since the ions are extracted from the ionization volume by an electric field that focuses them onto the cathode of an ion multiplier, the detection efficiency can reach 100%.

This makes the technique very sensitive and in favourable cases single absorbed photons can be monitored. Since the first transition $|i\rangle \rightarrow |k\rangle$ can be readily saturated, this also implies that single molecules can be detected.

In case of pulsed lasers, the laser power is sufficiently high to allow saturation of the two transitions even without focusing or with only weak focusing. If the ionization energy is high, two or more photons might be required for the ionizing step (REMPI, resonant multiphoton ionization). With CW lasers, tight focusing is required to reach the necessary high intensity.

The further advantage of ionization spectroscopy is the mass-selective detection of the ions by means of a mass spectrometer. This allows the detection of specific isotopes in the presence of other species. With pulsed laser ionization, a time-of-flight mass spectrometer offers the best choice; for CW lasers, a quadrupole mass spectrometer is suitable.

### Absorption spectroscopy in molecular beams

The spectral resolution for absorption spectroscopy in the gas phase is generally restricted by the Doppler

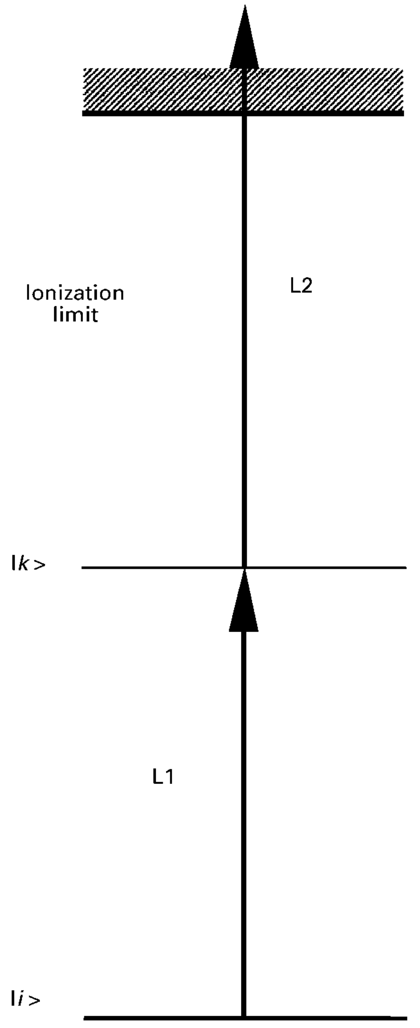



**Figure 3** Level scheme of ionization spectroscopy.

width of the absorption lines. If the laser beam crosses perpendicularly a molecular beam that is formed by molecules passing from a reservoir through a narrow nozzle A into vacuum and is collimated by a slit of width b at a distance d from the nozzle (**Figure 4**), the velocity components of the beam molecules parallel to the direction of the laser beam are reduced, owing to the geometric collimation. This reduces the Doppler width by a factor $\varepsilon = b/d$, which equals the collimation ratio $b/d \approx 0.01$ of the molecular beam.

Besides the increase in spectral resolution, supersonic molecular beams have the additional advantage that during the adiabatic expansion the molecules transfer a large fraction of their internal energy $E_{int}$ into directed flow energy. This means that the molecules are cooled and their population distribution $N(E_{int})$ is compressed into the lowest rotational-vibrational levels. This reduces the number of absorbing levels and therefore simplifies the absorption spectra considerably and facilitates the assignment of complex spectra of larger molecules.

## Nonlinear sub-Doppler spectroscopy

The Doppler limit in spectral resolution can be also overcome by nonlinear spectroscopy, which is based on saturation of absorbing transitions resulting in a partial or complete depletion of the population in the absorbing level.

The scheme of saturation spectroscopy is depicted in **Figure 5**. The output beam of a single-mode tuneable laser is split into a strong pump beam and a weak probe beam, which pass in opposite ± z directions through the sample cell with length L. Only those molecules with velocity components $v_z$ can absorb the pump photons on the transition $|i\rangle \to |k\rangle$ with centre frequency $\omega_0$, which are Doppler shifted into resonance with the laser frequency $\omega$. This gives the relation

$$|\omega_0 - \omega \pm kv_z| \leq \gamma$$

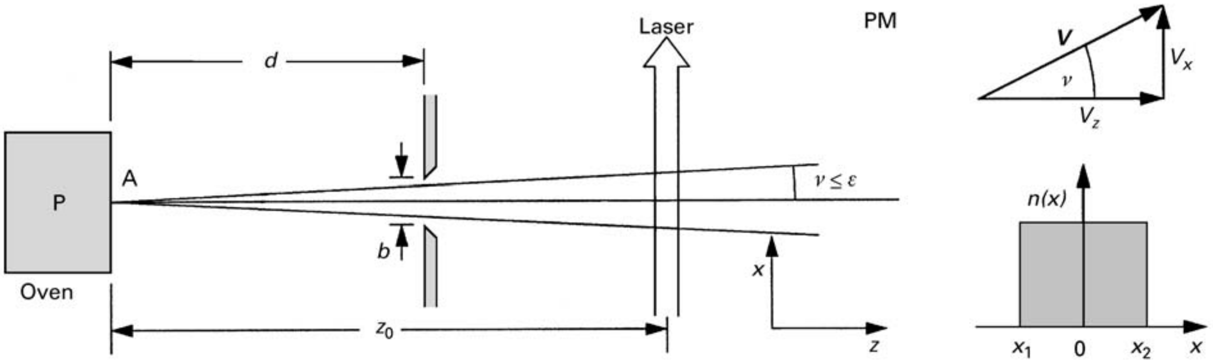



**Figure 4** Sub-Doppler spectroscopy in a collimated molecular beam.

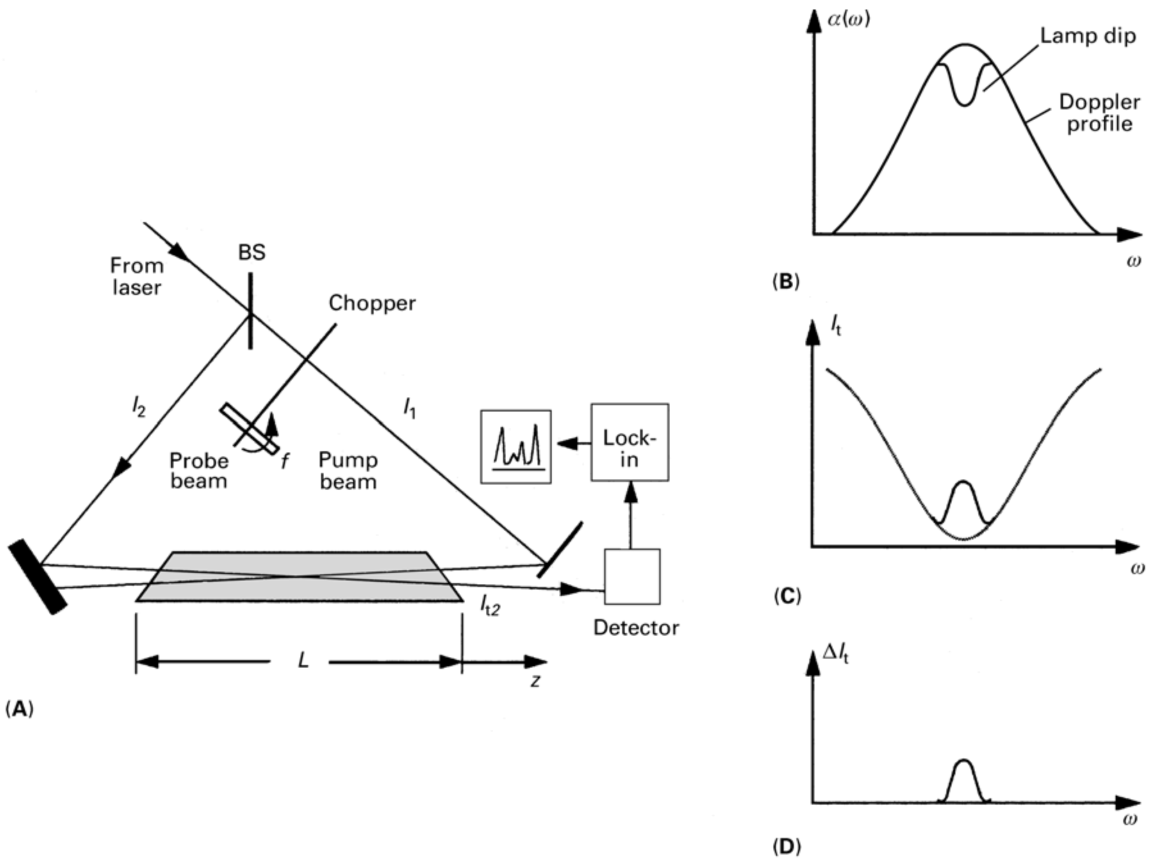



**Figure 5** Saturation spectroscopy: (A) experimental scheme; (B) Lamb dip at the centre of the Doppler-broadened absorption profile; (C) Lamb peak; (D) difference of the transmitted probe intensity with and without pump.

where $\gamma$ is the homogeneous line width of the transition. Since the wave vector $k$ of the probe is opposite to that of the pump beam, the Doppler shifts for pump or probe absorption are opposite. A molecule can only interact with both beams if its Doppler shift is smaller than the homogeneous width $\gamma$, i.e. $|k \cdot v_z| < \gamma/2$. In this case the probe absorption is decreased due to the saturation by the pump, which depletes the number $N_i$ of absorbing molecules. When the laser frequency $\omega$ is tuned over the Doppler broadened absorption profile, the absorption $\Delta I = \alpha L$ of the probe intensity shows a dip at the line centre (Lamb dip) with a spectral width $\gamma$ that may be two orders of magnitude smaller than the Doppler width (Figure 5B). This results in a corresponding peak in the transmitted probe intensity $I_T \approx I_0(1 - \alpha L)$ (Figure 5C).

The disadvantage of this simple arrangement is its low sensitivity, since a tiny change of the transmitted intensity has to be detected. If the pump intensity is chopped and the probe transmission is recorded with a lock-in detector, tuned to the chopping frequency, the difference of the transmitted intensities with and without saturation is monitored. This eliminates the Doppler-broadened background and yields Doppler-free profiles (i.e. the Lamb peaks) with higher sensitivity (Figure 5D).

Still higher sensitivities can be reached with polarization spectroscopy, in which the sample is placed between two crossed polarizers (Figure 6). The circularly polarized pump beam with intensity $I_2$ induces transitions with $\Delta M = 1$ (M = quantum number of the projection of the total angular momentum $J$ on the laser beam axis), causing a partial orientation of the otherwise randomly orientated molecules, owing to non-uniform saturation of the different M-sublevels. This makes the otherwise isotropic gaseous medium birefringent and causes a rotation of the polarization vector of a linearly polarized probe wave with intensity $I_1 \ll I_2$, passing in the opposite direction through the sample. However, this happens only if the probe wave interacts with the same molecules as the pump wave, i.e. if these molecules have velocity components $|v_z| \leq \gamma/2k$, just as in the case of saturation spectroscopy.

Since without saturation no signal is detected through the crossed polarizers, the method suppresses the background and therefore reduces

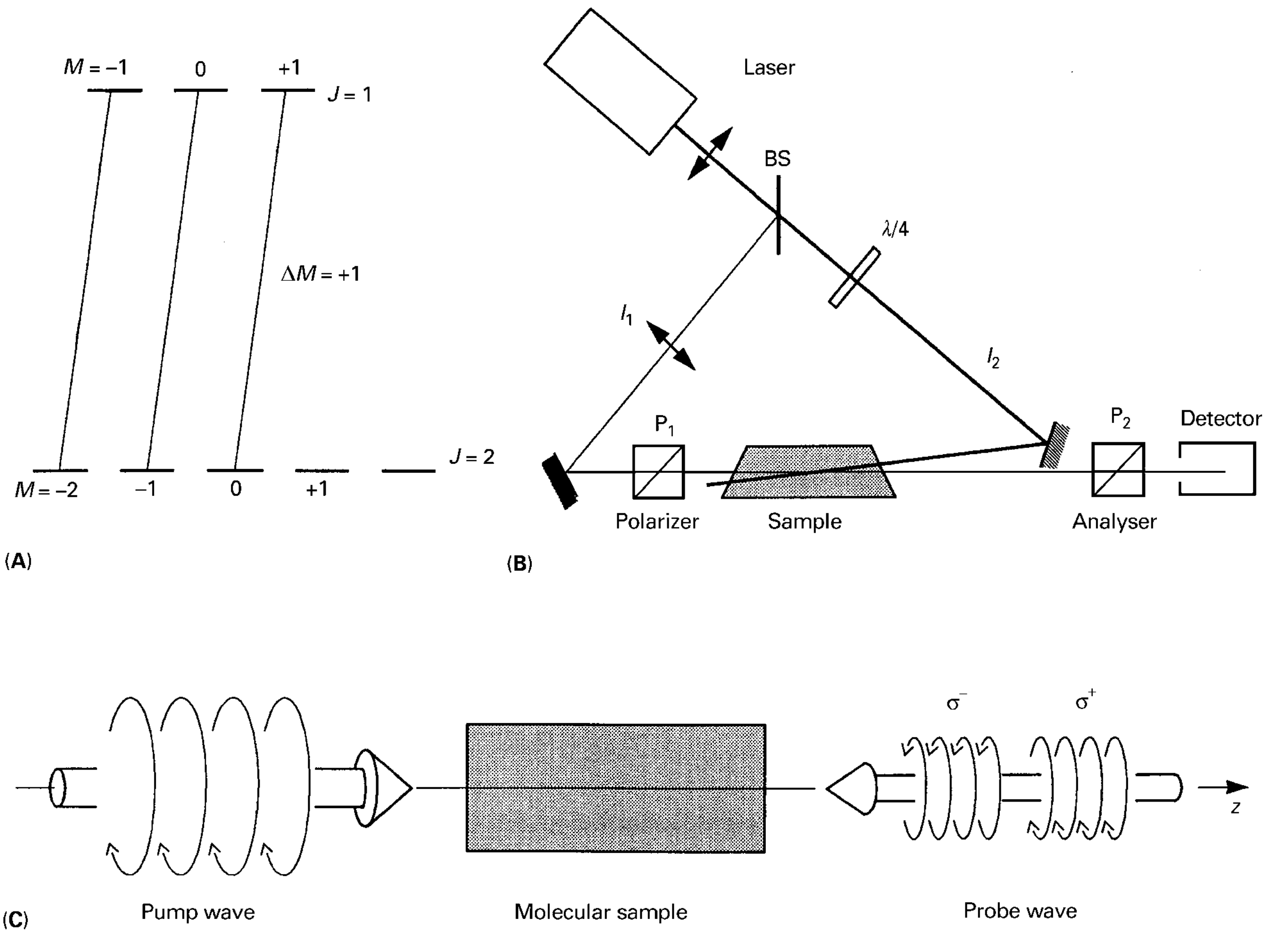



**Figure 6** Polarization spectroscopy.

much of the noise, resulting in a higher signal-to-noise ratio.

## Doppler-free two-photon spectroscopy

When two photons interact simultaneously with a molecule, a two-photon transition $|i\rangle \xrightarrow{h\nu_1 + h\nu_2} |k\rangle$ can occur if the two levels $|i\rangle$ and $|k\rangle$ have the same parity and differ in their angular momenta J by $\Delta J = 0, \pm 2$. This allows the excitation of high-lying electronic states in atoms or molecules, such as Rydberg states. Since the transition probability for two-photon transitions is much smaller than for one-photon transitions, generally pulsed lasers with high peak powers are used. However, two-photon transitions have been observed even with CW lasers. If the two photons come from two opposite beams of the same tuneable lasers, the two-photon transition in a molecule moving with velocity component $v_z$ requires the energy conservation

$$\begin{aligned} E_K - E_i &= (\hbar\omega_1 - kv_z) + (\hbar\omega_2 + kv_z) \\ &= 2\hbar\omega \quad \text{for} \quad \omega_1 = \omega_2 = \omega \end{aligned} \qquad [5]$$

independent of the molecular velocity $v_z$! This means that all molecules in level $|i\rangle$, independent of their velocities, contribute to the two-photon signal at a laser frequency $\omega = (E_k - E_i)/2\hbar$. This is different from saturation or polarization spectroscopy where only a small subgroup of molecules with $|v_z| \leq \gamma/2$ can absorb both laser beams. Despite the smaller transition probability, it therefore yields comparable signal magnitudes as for the other nonlinear techniques. This is illustrated by Figure 7, which shows the Doppler-free two-photon transition $5S_{1/2} \leftarrow 3S_{1/2}$ of sodium-atoms, measured by Cagnac and co-workers.

# Emission spectroscopy

Classical emission spectroscopy is based on excitation of atoms or molecules into higher electronic states by electron impact (in gas discharges), photon absorption or thermal excitation at high temperatures (in star atmospheres). Excitation by narrow-band lasers may result in the selective

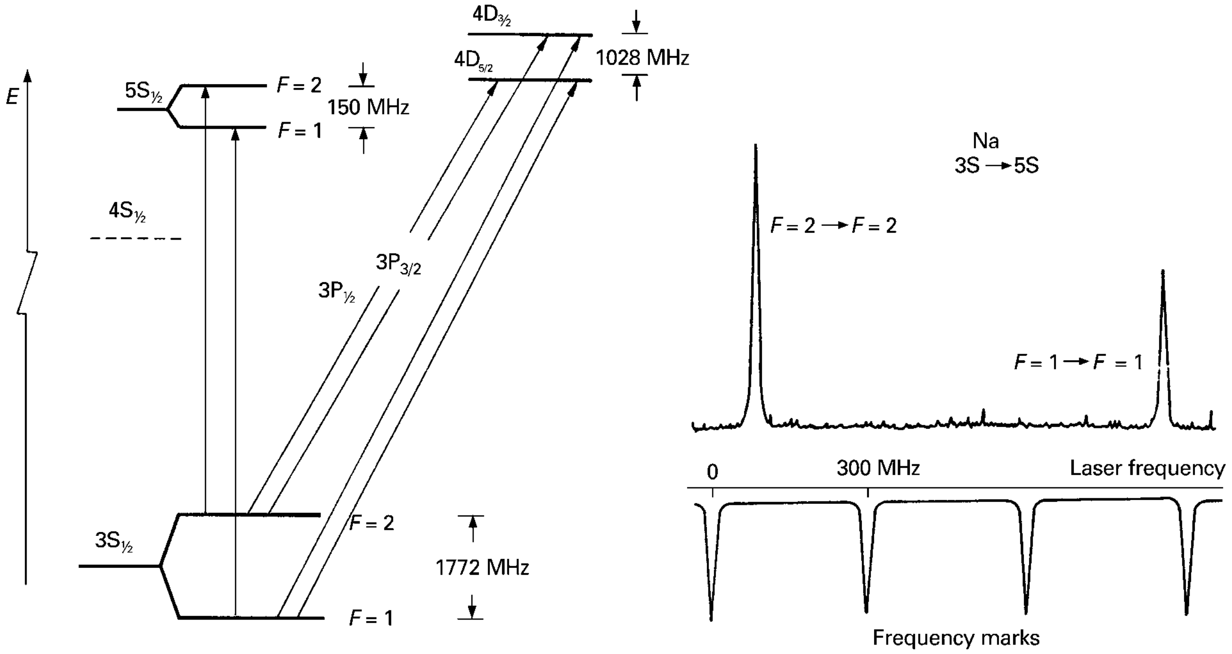



**Figure 7** Doppler-free two-photon spectroscopy of the $5S_{1/2} \leftarrow 3S_{1/2}$ transition of the sodium atom. Reproduced from Cagnac and co-workers.

population of wanted levels, which emit their excitation energy as fluorescence photons. Such a laser-induced fluorescence (LIF) spectrum is much simpler than the emission spectrum of a gas discharge, where the superposition of fluorescence from many emitting levels is observed.

Measurements of spectrally resolved LIF allows the determination of the final levels into which the fluorescence is emitted (Figure 8A). This gives access, for example, to high vibrational-rotational levels in the electronic ground state that are not thermally populated and are therefore not accessible to absorption spectroscopy. Examples are levels closely below the dissociation energy, where the vibrating atoms reach very large internuclear distances at their outer turning point.

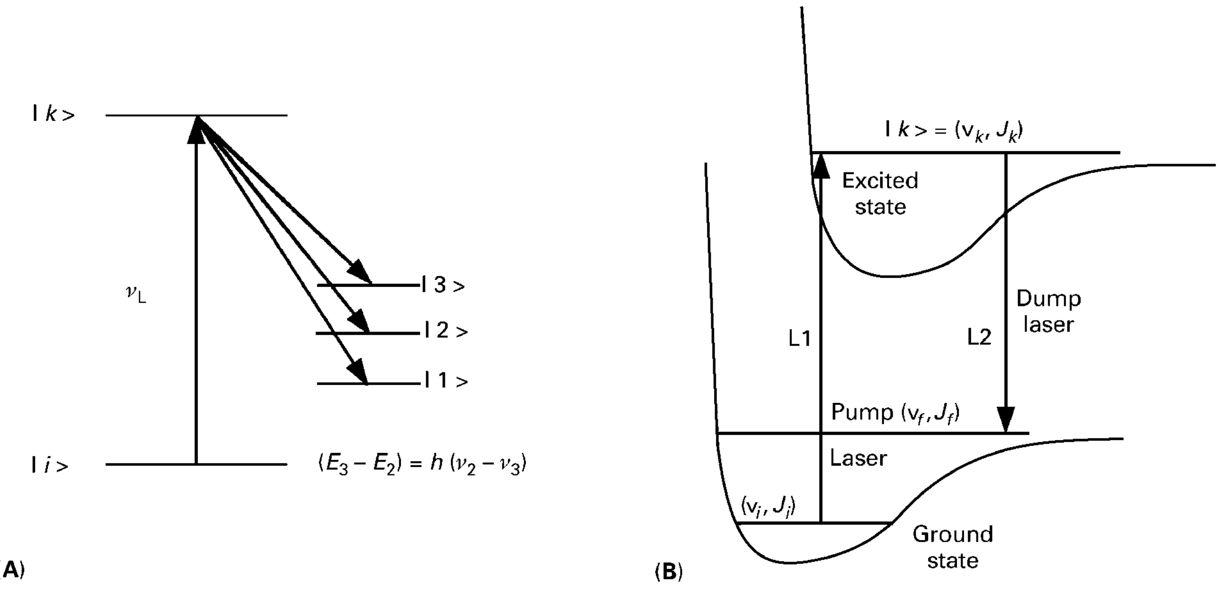



**Figure 8** (A) Laser-induced fluorescence and (B) stimulated emission pumping.

The accuracy of wavelength measurements is limited by that of the spectrograph or interferometer used for dispersing the LIF spectrum. Higher spectral resolution can be achieved by stimulated emission pumping, where a second tuneable laser L2 is used to induce stimulated emission from the excited level $|f\rangle$ (Figure 8B). The transition can be detected by the corresponding decrease of the population N(k), which may be monitored by LIF or by photoionization of level $|k\rangle$.

## Double-resonance spectroscopy

Double-resonance spectroscopy is based on the simultaneous resonant interaction of two electromagnetic waves, which induce two transitions sharing a common level (Figure 9). In the V-type double resonance (Figure 9A) this common level of both transitions is the lower level $|i\rangle$ of the two transitions. The laser L1, called the pump laser, depletes the population of level $|i\rangle$ and increases that of $|k\rangle$. If L1 is chopped, the populations $N_i$ and $N_k$ are modulated with opposite phases. The probe laser L2 is tuned through the spectrum of allowed transitions. Each time its wavelength $\lambda_2$ coincides with a transition $|i\rangle \rightarrow |m\rangle$ or $|k\rangle \rightarrow |n\rangle$, its absorption will be modulated. This can be monitored either on the transmitted intensity $I_T$(L2), or with LIF or RTPI. Since the starting level, marked by the pump, is known, these optical double resonance signals can be more readily assigned than the much more abundant lines in a conventional absorption spectrum. The OODR technique therefore is a very convenient method for facilitating the assignment of complex and perturbed spectra.

The Λ-type OODR scheme (Figure 9B) can be regarded as stimulated resonance Raman scattering. It can be performed in cells or in molecular beams and yields Doppler-free signals if the two lasers involved are operated as single-mode lasers. The resonance signal appears when L2 is tuned to resonance $\nu_1 - \nu_2 = (E_i - E_f)/h$. The Λ-type scheme can be used, for instance, to investigate with sub-Doppler resolution the structure and dynamics of levels closely below the dissociation limit. One example is the mixing of gerade and ungerade singlet and triplet states at large internuclear distances, where the energy differences between the two electronic states becomes comparable to the hyperfine splittings of the atomic dissociation products. In these cases the indirect coupling of the nuclear spins by the electron spins can mix both states. This process may become important in cold atomic clouds kept in magnetooptical traps where collisions between atoms can lead to spin flips, which correspond in the molecular picture to transitions between triplet and singlet states.

The double-resonance scheme of Figure 9C corresponds to stepwise excitation, which may be regarded as the special resonance case of two-photon absorption. The first laser populates the excited level $|k\rangle$ and the second laser induces further excitation into high energy levels, such as Rydberg states. They can be monitored by field ionization or photoionization with high efficiency. These Rydberg levels may also be autoionizing levels above the ionization limit, which can be directly detected through the ions. Since all transitions induced by L2 start from the known level $|k\rangle$, their assignment is greatly facilitated. This enables detailed investigation of structure and dynamics of Rydberg levels and their autoionization probabilities.

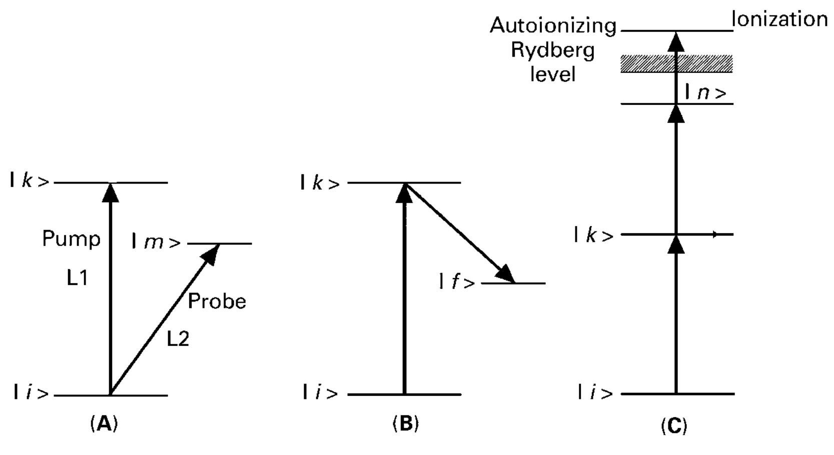



**Figure 9** Double-resonance schemes: (A) V-type; (B) Λ-type; (C) stepwise excitation.

## Time-resolved electronic spectroscopy

The development of ultrashort laser pulses down to pulse durations of $5 \times 10^{-15}$ s has opened access to studies of extremely fast transient phenomena. Examples are the relaxation of electrons in semiconductors after their excitation by a short light pulse. The electrons excited with a definite energy $\hbar\omega$ into the conduction band thermalize within $10^{-13}$ s by electron-phonon collisions (Figure 10). With a much longer decay constant, they recombine with holes in the valence band before they can be excited again by the next pulse. Such time-resolved studies give important information on the limiting processes for the maximum speed of computers.

In molecular physics, elementary processes such as the vibrations of molecules, the formation or breaking of chemical bonds or electronic isomerization processes can now be visualized by femtosecond spectroscopy. This will be illustrated by two examples. The first is concerned with the wavepacket dynamics of vibrating diatomic molecules. The schematic principle is shown in Figure 11. The molecule is excited by a short pulse with duration T from its vibrational ground state into vibrational levels of an electronically excited state. If the spectral width $\Delta\nu \approx 1/T$ is large compared to the vibrational spacings, several vibrational levels are simultaneously and coherently excited. This coherent superposition of vibrational wavefunctions produces a vibrational wavepacket in the upper state that moves with the average velocity of the vibrating nuclei towards the outer turning point, where it is reflected. When a second laser pulse with a variable time delay $\Delta t$ brings the excited molecule into a higher electronic state, the transition probability depends on the location of the wavepacket, because the overlap with the vibrational wavefunctions of this higher state depends on the internuclear distance. For the example of Figure 11, transitions from the inner turning point produce the molecular ion $M^+$ plus the ejected electron $e^-$, which carries away the excess energy. From the outer turning point, however, the repulsive ion potential can be reached and the molecular ion $M^+$ dissociates into $A^+ + B$ or $A + B^+$. When the time delay between pump and probe is continuously increased, the ratio $M^+/A^+$, or respectively $M^+/B^+$, will show a periodic modulation, where the period reflects the periodic movement of the wavepacket. The important point is that this pump-and-probe technique allows a time resolution on the femtosecond scale, although the detectors may be much slower. With such controlled time delay of femtosecond pulses, the resultant reaction products may be selected, and in such favourable cases channel-selective photochemistry can be realized.

A second example illustrates the dissociation of a molecule following optical excitation into an electronic state with a repulsive potential (Figure 12). The dissociating molecule is excited by a second laser pulse into a fluorescing bound state. The wavelength of this second transition depends on the internuclear separation R. Choosing the wavelength $\lambda(R_1)$, the correct time delay between pump and probe for exciting the fluorescence reflects the time that the dissociating molecule needs to reach $R_1$. This method is not restricted to diatomic molecules but has been demonstrated also for polyatomic molecules.

Measurements of lifetimes of excited states of atoms and molecules give direct information on the

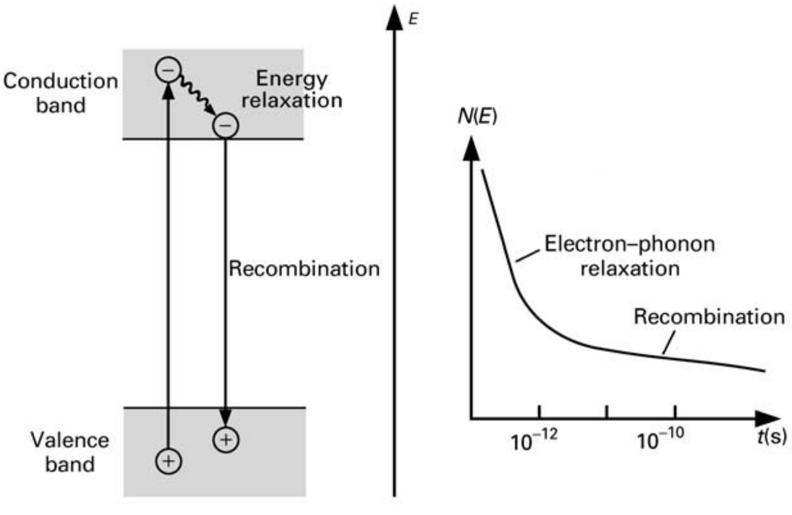



**Figure 10** Thermalization and recombination of electrons after excitation into the energy *E* within the conduction band in a semiconductor with a femtosecond laser pulse.

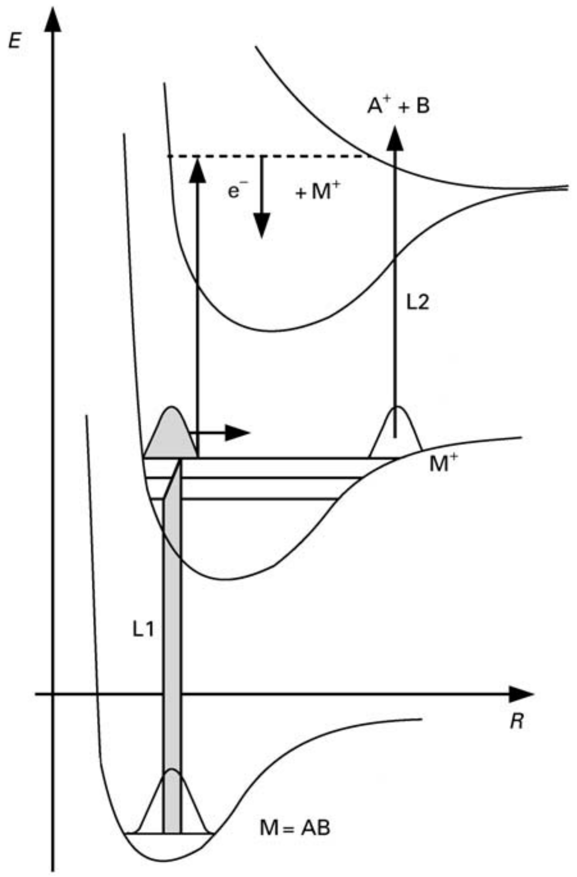



Figure 11 Wavepacket dynamics of a vibrating diatomic molecule.

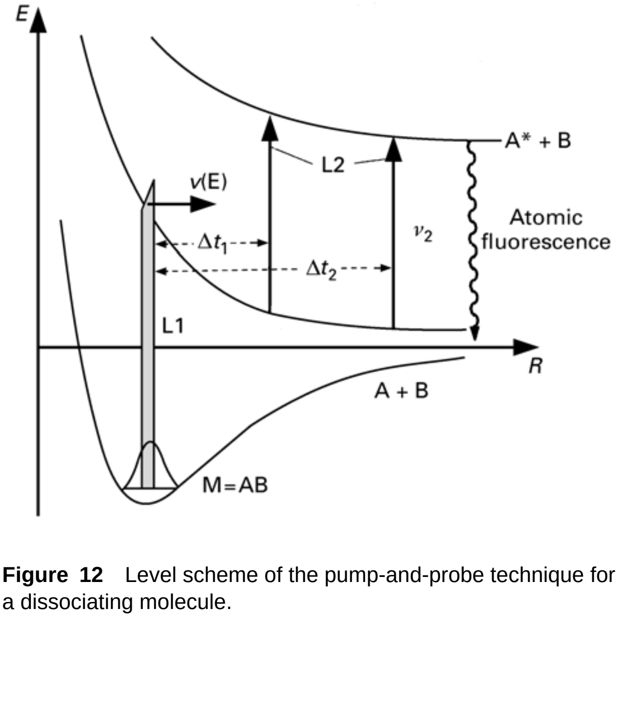



Figure 12 Level scheme of the pump-and-probe technique for a dissociating molecule.

absolute transition probabilities and are therefore important for many fields in atomic physics, chemistry and astrophysics. Accordingly, several experimental methods have been developed that cover different time domains. In the nanosecond to microsecond range, the delayed coincidence photon counting technique has proved to be very useful. The molecules are excited by a short laser pulse, which also triggers a linear voltage ramp. The first fluorescence photon emitted by the excited molecules produces an output voltage in the photomultiplier that stops the voltage ramp. The output voltage of this time-to-pulse height converter is proportional to the time delay between the exciting laser pulse and the emission of the fluorescence photon. Measuring many excitation fluorescence cycles yields a voltage pulse height distribution that is stored in a multichannel analyser (or a computer) and reflects the exponential decay of the excited molecules.

Since the probability per cycle of detecting a fluorescence photon has to be smaller than unity, the cycle frequency should be large and is only limited by the requirement that the time between successive excitation pulses should be larger than the lifetime of the excited molecules. Therefore, mode-locked synchronously pumped cavity-dumped CW dye lasers are ideal excitation sources. For the excitation of higher electronic states, the visible output pulses can be converted into ultraviolet pulses by optical frequency doubling in optically nonlinear crystals.

In the picosecond and femtosecond range, the pump-and-pulse technique is more suitable where the time delay between pump and probe can be accurately realized by an optical delay line (Figure 13) where one of the laser pulses travels a variable path length before it is again superimposed on the pump pulse direction.

Time-resolved investigations have been applied to measurements of fast relaxation processes, such as the primary steps in the visual process or in photosynthesis, or to collision-induced relaxation of excited states in liquids or gases under high pressures.

## List of symbols

$E_{int}$ = internal energy; $E_i$, $E_k$ = energies of levels i, k; $\hbar$ = Planck constant/$2\pi$; I = light intensity; J = total angular momentum quantum number; $\boldsymbol{k}$ = wave vector, k = $|\boldsymbol{k}|$; L = path length; M = quantum number of the projection of the total angular momentum on the laser beam axis; $\dot{n}_{abs}$ = rate of absorption of laser photons; $N_i$ = number density of molecules in level $|i\rangle$; R = internuclear separation; $S(\omega)$ = signal profile; T = flight time, pulse duration; $v_{i,k}$ = vibrational quantum number of level i, k; $v_z$ = z direction

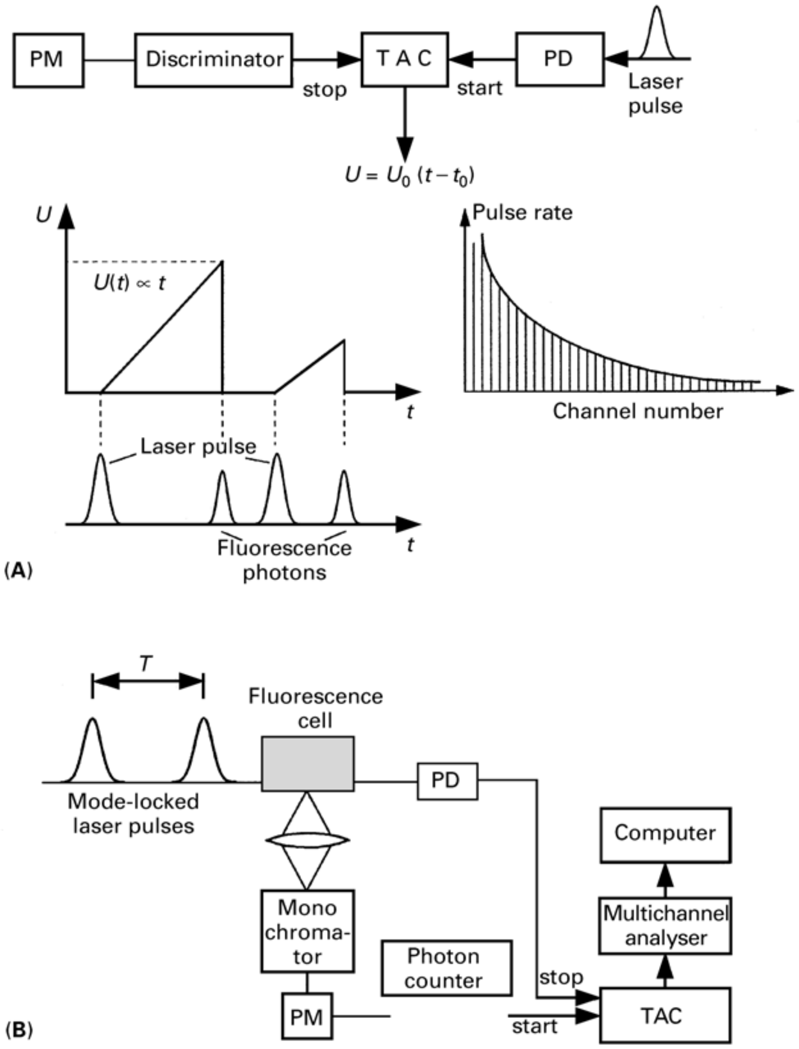



**Figure 13** Delayed coincidence photon counting: (A) principal scheme; (B) experimental setup for lifetime measurements.

velocity component; $\alpha$ = absorption coefficient; $\alpha(\omega)$ = absorption profile; $\Delta t$ = laser pulse time delay; $\gamma$ = homogeneous line width; $\varepsilon$ = collection probability of emitted fluorescence; $\eta_k$ = quantum efficiency of detector; $\eta_m$ = quantum efficiency of excited species; $\lambda$ = wavelength; $\nu$ = transition frequency; $\sigma_{ik}(\nu)$ = cross section for transition $|k\rangle \leftarrow |i\rangle$; $\tau$ = upper-state lifetime; $\omega$ = laser (pump) frequency; $\Omega$ = modulation frequency; $\Omega_0$ = local oscillator frequency.

*See also:* **Laser Spectroscopy Theory; Multiphoton Spectroscopy, Applications; Nonlinear Optical Properties; X-Ray Emission Spectroscopy, Applications; X-Ray Emission Spectroscopy, Methods.**

## Further reading

Demtröder W (1998) Laser Spectroscopy, 2nd edn. Berlin: Springer-Verlag.

Hurst GS and Payne MG (1988) Principles and Applications of Resonance Ionization Spectroscopy. Bristol: Hilger.

Levenson MD (1986) Introduction to Nonlinear Laser Spectroscopy, 2nd edn. New York: Academic Press.

Scoles G (ed) (1991, 1992) Atomic and Molecular Beam Methods, Vols 1 + 2. Oxford: Oxford University Press.

Ye Jun, Ma-Long-Sheng and Hall JL (1988) Ultrasensitive detections in atomic and molecular physics. Journal of the Optical Society of America B15: 6.

Zewail AH (1994) Femtochemistry, Vols I, II. Singapore: World Scientific.

# Laser Induced Optoacoustic Spectroscopy

**Thomas Gensch**, Katholieke Universiteit of Leuven, Heverlee, Belgium
**Cristiano Viappiani**, Università di Parma and Instituto Nazionale per la Fisica della Materia, Italy
**Silvia E Braslavsky**, Max-Planck-Institut für Strahlenchemie, Mülheim an der Ruhr, Germany

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
| --- |
| **Applications** |

## Introduction

Time-resolved laser-induced optoacoustic spectroscopy (LIOAS) is a sensitive photobaric method for nondestructive probing of materials that monitors the pressure changes induced in a liquid sample after photoexcitation with a pulsed laser. LIOAS delivers information generally not available from optical studies and therefore complements other methods. After excitation, commonly with nanosecond pulsed lasers, the time evolution of the pressure pulse is monitored by fast piezoelectric transducers frequently located in a plane parallel to the direction of the laser beam (right-angle arrangement). Front-face geometry has also been used, especially with strongly absorbing samples. This article concentrates mainly on the former arrangement.

So far, time-resolved information can be used only in the case of pseudo-first-order reactions, either parallel or sequential, including thermal equilibria, since the LIOAS data analysis has been developed for schemes complying with a time-dependent pressure evolution represented by a sum of single-exponential decays.

Measurement of the heat evolved is used for determination of energy levels of intermediates (including the energy associated with chromophore–solvent interactions), provided that the quantum yields for the reactions are known. Conversely, if the energy level of the intermediates is known from other spectroscopic techniques, the determination of quantum yields is possible, for example for isomerization, charge transfer, fluorescence, intersystem crossing and energy transfer processes.

The structural volume changes are essentially density changes and reflect movements in the chromophore such as photoinduced changes in charge distribution and in the strength of the interactions with solvent molecules. Time-resolved LIOAS data also afford the decay rate constants of the transients, and therefore provide a check for the assignment of the observed events to a particular intermediate. A further identification of the transients involved is offered by the action spectrum of the pressure evolved.

LIOAS is an interesting tool for investigating photoinduced 'optically silent' intermediates, i.e. displaying no absorption or emission. A phenomenological approach to the equations is employed here and a description is given of the equipment and data handling as well as a few examples of applications of LIOAS to small molecules. Applications to biological complex systems can be found in the work by Schulenberg and Braslavsky listed in the Further reading section.

## Phenomenological approach

Pressure pulses in the illuminated sample arise from the volume changes produced by radiationless relaxation ($\Delta V_{\mathrm{th}}$) and structural solute–solvent rearrangements ($\Delta V_{\mathrm{r}}$). Relaxation originates either from nonradiative decay of excited states or release of heat (enthalpy change) in photoinitiated reactions, including the heat involved in the rearrangement of the solvent. The structural volume changes reflect movements of the photoexcited molecules and/or the surrounding solvent in response to such events as dipole moment change, charge transfer, and photoisomerization. The overall pressure change is given by Equation [1].

$$\Delta p = -\frac{1}{\kappa_T}\left(\frac{\Delta V}{V_0}\right)_T = -\frac{1}{\kappa_T}\left(\frac{\Delta V_{\mathrm{th}} + \Delta V_{\mathrm{r}}}{V_0}\right)_T \qquad [1]$$

where $p$ is the pressure change, $V_0$ is the volume of the illuminated sample, $\kappa_T$ is the isothermal compressibility (see below). The pressure change induces an electrical signal $H$ from the transducer, proportional to the pressure change through an

instrumental factor $k$ (Eqn [2]):

$$H = k \cdot \Delta p \qquad [2]$$

The voltage is given by

$$H = \frac{1}{\kappa_T}\frac{k}{V_0}(\Delta V_{\rm th} + \Delta V_{\rm r}) = k'(\Delta V_{\rm th} + \Delta V_{\rm r}) \qquad [3]$$

Since the signal is proportional to the number of moles of absorbed photons, $n$,

$$n = \frac{N_{\rm ph}}{N_{\rm A}}(1 - 10^{-A}) \qquad [4]$$

where $A$ is the absorbance of the solution at the excitation wavelength $\lambda$, $N_{\rm A}$ is the Avogadro number, and $N_{\rm ph}$ is the number of incident photons, the total signal is given by

$$H^{\rm S} = k'n\left(\alpha\frac{\beta}{c_{\rm p}\rho}E_\lambda + \Delta V_{\rm e}\right) \qquad [5]$$

where $E_\lambda = N_{\rm A}$ $(h.c/\lambda)$ is the energy of an Einstein of wavelength $\lambda$, $\alpha = q/E_\lambda$ is the fraction of absorbed energy $E_\lambda$ released as heat $q$, $\Delta V_{\rm e} = \Delta V_{\rm r}/n$ is the structural volume change per absorbed Einstein, $c_{\rm p}$ is the specific heat capacity, $\rho$ is the density, and $\beta$ is the cubic thermal expansion coefficient of the solution [for dilute solutions (< 1 mM) these properties are those of the solvent]. To obtain $\alpha$ and $\Delta V_{\rm e}$, the instrumental constant, $k'$, must be determined by comparing the signal of the sample with that of a calorimetric reference, i.e. a system that releases all the absorbed energy as heat in a time shorter than the instrumental response. The calorimetric reference and the sample solutions should be measured under identical experimental conditions such as solvent, temperature, excitation wavelength and geometry.

Since the isothermal compressibility $\kappa_T$ is related to the adiabatic compressibility, $\kappa_S$ by

$$\kappa_T = \kappa_S + T\frac{\beta^2}{c_{\rm p}\rho} = \frac{1}{\rho\nu_{\rm a}^2} + T\frac{\beta^2}{c_{\rm p}\rho} \qquad [6]$$

(with $v_{\rm a}$ the sound velocity in the medium) care should be taken to use the proper values of $\kappa_T$ in Equation [3], especially when determining the thermoelastic parameters using a calorimetric reference in a different solvent. In neat water $\kappa_T = \kappa_S$ and the situation is simpler.

The signal for the calorimetric reference is given ($\alpha$=1) by

$$H^{\rm R} = k'n\frac{\beta}{c_{\rm p}\rho}E_\lambda \qquad [7]$$

When permanent products or transient species living much longer (~5 times) than the pressure integration time are the only products of the photoreaction, the amplitudes of the LIOAS signals serve to determine the heat and structural volume changes through application of Equation [8] (obtained from the ratio of Eqns [7] and [5]) relating the fluence-normalized LIOAS signal amplitudes for sample and reference solutions, $H_{\rm n}^{\rm S}$ and $H_{\rm n}^{\rm R}$, respectively:

$$\frac{H_{\rm n}^{\rm S}}{H_{\rm n}^{\rm R}} = \alpha + \frac{\Delta V_{\rm e}}{E_\lambda}\frac{c_{\rm p}\rho}{\beta} \qquad [8]$$

For transient species with lifetimes within the pressure integration time, kinetic information is obtained by considering that the signal is a time convolution between the transfer function $H^{\rm R}(t)$ of the instrument, obtained with the calorimetric reference, and the time derivative $H(t)$, of the time-dependent total volume change:

$$H^{\rm S}(t) = \int_0^t H^{\rm R}(t)\,H(t-t')\,{\rm d}t' \qquad [9]$$

Numerical reconvolution has been used to retrieve the kinetic, enthalpic and volumetric parameters. Direct deconvolution methods have been developed as well. The reconvolution methods assume a sum of single-exponential decay functions for the time evolution of the pressure:

$$H(t) = \sum_i \frac{\varphi_i}{\tau_i}{\rm e}^{-t/\tau_i} \qquad [10]$$

where $\tau_i$ is the lifetime of transient $i$ and $\varphi_i$ is its fractional contribution to the measured volume change. This assumption bears no implications regarding the mechanism involved, since several mechanisms lead to a function such as Equation [10].

A quadratic approximation to the convolution integral has proved to be the best method for optimizing the results. A Levenberg–Marquardt algorithm performing a least-squares minimization is used.

The number of transients is successively increased until no further (10–20%) improvements is observed in the value of $\chi^2$ and in the residuals trend. A single-exponential decay did not correctly describe the signal of benzophenone in acetonitrile, whereas the sum of two single-exponential decays yielded a perfect fit (**Figure 1**). Upon addition of a third exponential no significant improvement in $\chi^2$ was observed (not shown).

Up to three lifetimes have been resolved in pulsed photoacoustic spectroscopy experiments, provided that they were separated by at least a factor of 4. In the case of closer values, a strong correlation between amplitudes and lifetimes appears, preventing a reliable recovery of the parameters. Lifetimes below a few nanoseconds are integrated in the prompt response, for which it is only possible to obtain the fractional amplitude.

Lifetimes between a few nanoseconds and some microseconds can readily be deconvoluted. By applying simulations based on a theoretical treatment of the signal, restoration limits have been derived for amplitudes and lifetimes in the case of a sum of two single-exponential decays. The influences of variable geometry, energy distribution and different deconvolution methods, as well as that of the signal-to-noise ratio on the restoration of the amplitudes and lifetimes, have been analysed. Since a general feature of the microsecond transients is a strong correlation between their lifetime and amplitude, their evaluation is more critical. Simulations show that the front-face geometry in principle covers a broader frequency range than does the perpendicular arrangement at the same signal-to-noise level. However, the polymeric detector used in this scheme has a smaller sensitivity compared to the piezoceramic detectors, which defeats the purpose of a higher signal-to-noise ratio. Experimental validation of these restoration limits has not yet been done. The high absorbance needed and other technical problems constitute a problem with the front-face geometry.

The fractional amplitudes $\varphi_i$ in Equation [10] contain the heat release $q_i$ and the structural volume change $\Delta V_{r,i}$ of the ith step. Both contributions can be considered additive provided the time evolution of both is the same (Eqn [11]). This assumption has been validated in several systems.

$$\varphi_i = \alpha_i + \frac{\Delta V_{e,i}}{E_\lambda}\frac{c_p\rho}{\beta}, \quad \alpha_i = \frac{q_i}{E_\lambda}, \quad \Delta V_{e,i} = \Phi_i \Delta V_{r,i} \qquad [11]$$

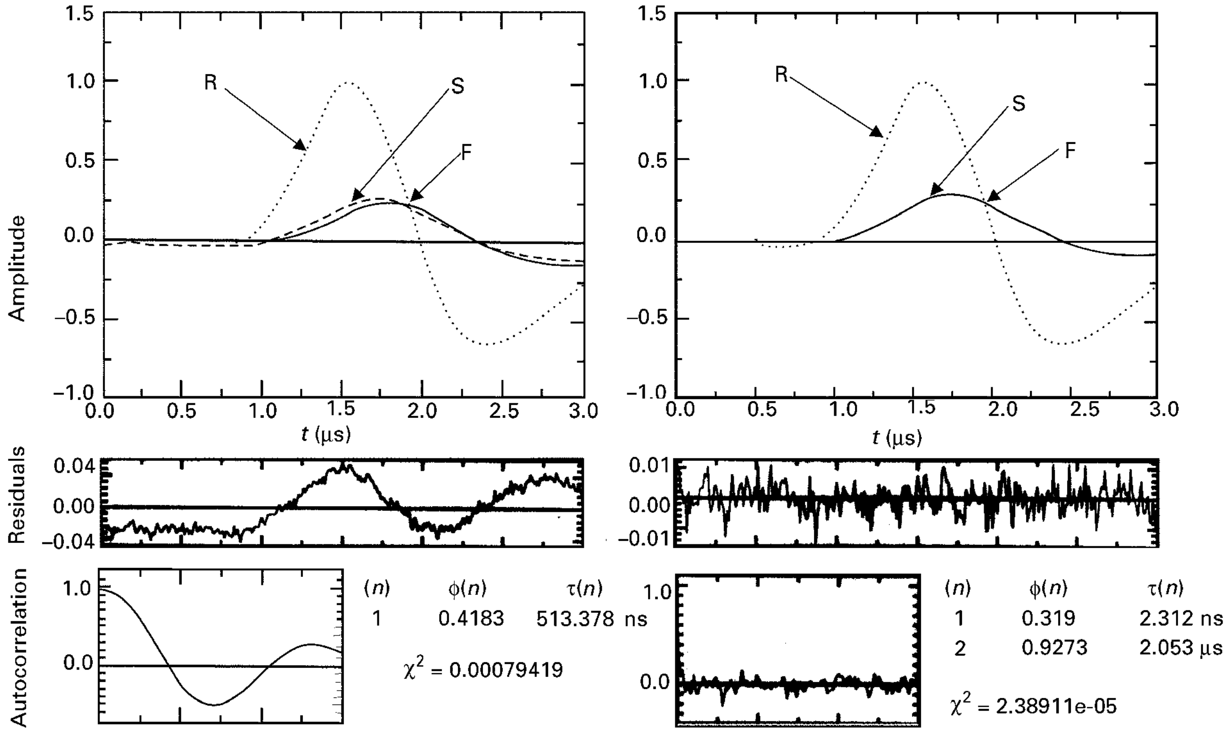



**Figure 1** LIOAS signals of benzophenone (S, -----) and 2-hydroxybenzophenone (R, ·····) in $N_2$-bubbled acetonitrile. Fitted curves (F,——) obtained by reconvolution of the reference signal R with single-exponential (left) and two-exponential (right) decay model functions *H*(*t*). Residuals and autocorrelation function are given for inspection of the goodness of the fits. Index *n* corresponds to subindex *i* in text.

where $\Phi_i$ is the quantum yield of step i. $q_i$ is related to the enthalpy of the species produced, including the reorganization energy of the medium and to the value of $\Phi_i$. $\Delta V_{r,i}$ represents the molar structural volume change of the ith step, which includes intrinsic changes in the chromophore as well as solute–solvent interactions. Equation [8] is thus a particular case of Equation [11] for i = 1.

A mechanism needs to be determined for the proper interpretation of $\varphi_i$. In the case of a sequential reaction scheme such as depicted in equation [12],

$$A \xrightarrow{h\nu} A^* \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} D \qquad [12]$$

the amplitudes $\varphi_i^{\mathrm{app}}$ derived from the reconvolution analysis are related to the parameters of the elementary steps (i.e., the weight of each step $\varphi_i$ and the lifetimes $\tau_i$) by Equations [13]–[15]:

$$\varphi_3 = \frac{(\tau_3 - \tau_1)(\tau_3 - \tau_2)}{\tau_3^2}\varphi_3^{\mathrm{app}} \qquad [13]$$

$$\varphi_2 = \frac{\tau_2 - \tau_1}{\tau_2}\varphi_2^{\mathrm{app}} + \frac{\tau_2}{\tau_3 - \tau_2}\varphi_3 \qquad [14]$$

$$\varphi_1 = \varphi_1^{\mathrm{app}} + \frac{\tau_1}{\tau_2 - \tau_1}\varphi_2 + \frac{\tau_1^2}{(\tau_2 - \tau_1)(\tau_1 - \tau_3)}\varphi_3 \qquad [15]$$

It is obvious that for a system with $\tau_3 >> \tau_2 >> \tau_1$, the equations simplify and $\varphi_i = \varphi_i^{\mathrm{app}}$.

The equations for the reaction scheme [16], including a thermal equilibrium,

$$A \xrightarrow{h\nu} A^* \xrightarrow{k_1} B \underset{k_{32}}{\overset{k_{23}}{\rightleftarrows}} C \xrightarrow{k_{34}} D,\ k_1 < 0.1\ \mathrm{ns} \qquad [16]$$

are given by Equations [17]–[19]:

$$\varphi_3 = \frac{1}{k_{23}}\left[\left(k_{23} - \frac{1}{\tau_2}\right)\varphi_2^{\mathrm{app}} + \left(k_{23} - \frac{1}{\tau_3}\right)\varphi_3^{\mathrm{app}}\right] \qquad [17]$$

$$\varphi_2 = \frac{1}{k_{23}}\left(\frac{1}{\tau_2}\varphi_2^{\mathrm{app}} + \frac{1}{\tau_3}\varphi_3^{\mathrm{app}}\right) \qquad [18]$$

$$\varphi_1 \approx \varphi_1^{\mathrm{app}} \qquad [19]$$

## Separation of heat release and structural volume changes

Separation of the heat dissipation and structural terms in Equation [11] is achieved by measuring the LIOAS signals as a function of the solvent ratio $c_p\rho/\beta$. In aqueous solutions the ratio $c_p\rho/\beta$ depends strongly on temperature (mainly owing to the changes in $\beta$). Thus, the separation of the contributions is obtained by performing measurements at various temperatures in a relatively small range (e.g. 8–20°C). A prerequisite is that the temperature does not significantly affect the kinetics of the various processes. The ratio $(c_p\rho/\beta)_T$ for neat water is obtained from tabulated values. In particular, $\beta_{\mathrm{water}} = 0$ to 3.9°C ($T_{\beta=0}$), is positive above and negative below this temperature. At $T_{\beta=0}$, heat release gives rise to no signal. This peculiar feature allows the straightforward assessment of the existence of structural volume changes. For aqueous solutions containing salts or other additives at millimolar concentrations or higher, $(c_p\rho/\beta)_T$ and $T_{\beta=0}$ must be determined by comparison of the signal obtained for a calorimetric reference in the solvent of interest and in water.

The precision of the recovered slope of plots with Equation [11], $\Delta V_{e,i}/E_\lambda$, is generally good, whereas the intercept $\alpha_i$ is affected by larger errors. In case the main goal is the determination of the structural volume change, a two-temperature method may be applied. Essentially, the LIOAS wave is measured for the sample at $T_{\beta=0}$ and at a slightly higher temperature $T_{\beta\neq 0}$ for which $\beta \neq 0$. The parameters $\Delta V_{e,i}/E_\lambda$ and $\alpha_i$ are then calculated with Equations [20] and [21]:

$$\Delta V_{c,i} = \varphi_i(T_{\beta=0})\left(\frac{\beta}{c_p\rho}\right)E_\lambda \qquad [20]$$

$$\alpha_i = \varphi_i(T_{\beta\neq 0}) - \varphi_i(T_{\beta=0}) \qquad [21]$$

These equations hold only for values of $T_{\beta=0}$ and $T_{\beta\neq 0}$ that are near to each other, such that the rise time of the LIOAS signal and the compressibility of the solution are similar. The signal of the reference at $T_{\beta\neq 0}$ is used together with that of the sample at $T_{\beta=0}$ and at $T_{\beta\neq 0}$ to obtain the values of $\varphi_i(T_{\beta=0})$ and $\varphi_i(T_{\beta\neq 0})$. A time shift of the sample waveform with respect to the reference waveform at $T_{\beta\neq 0}$ is necessary to compensate for the change in the speed of sound, which leads to a slightly different arrival time at both temperatures. Thermal instability of the sample holder and trigger arrival fluctuations make the shifting an important option in the fitting routines.

The temperature dependence of the $c_p\rho/\beta$ ratio in organic solvents is poor, rendering the above procedures inapplicable in these solvents. An alternative approach has been used that relies on the determination of the LIOAS signals in a series of homologous solvents such as linear alkanes, cycloalkanes, or, alternatively, in solvent mixtures. This procedure is valid provided that no significant perturbations are introduced in the photophysics and photochemistry of the solute and that the solvation enthalpy and structural volume change do not vary with change of solvent. These are strong assumptions that need verification in each case.

## Calorimetric calculations

The heat released in each step, $q_i$, is interpreted in terms of enthalpy changes by using the energy balance

$$E_\lambda = \Phi_f E_f + \sum_i q_i + E_{stored} \qquad [22]$$

where the input energy, $E_\lambda$, is equal to the sum of the energy released as emission (e.g. fluorescence; $\Phi_f$ and $E_f$ are the quantum yield and energy constant of the emitting state) and the heat dissipated (changes in the states' energy content plus energy associated with solute–solvent interactions) in each of the steps. $E_{stored}$ represents the total energy trapped in species living much longer than the pressure-integration time of the experiment. Analysis of the calorimetric information requires knowledge of the reaction scheme and the parameters associated with the process such as quantum yields and energy levels derived, for example, from optical measurements.

## Instrumentation

**Figure 2** shows a typical LIOAS setup. A calibrated beam splitter guides a fraction of each laser pulse to an energy meter for normalization purposes. The sample holder contains a normal quartz cuvette inside a thermostated metal block.

The unfocused beam is shaped by passing it through a slit (200 µm–2 mm) placed in front of the cuvette. With fast transducers, the beam size determines the rise time of signal and, therefore, the time resolution. The piezoelectric transducer is pressed against the cuvette (**Figure 3**).

The higher sensitivity of the slit-shaping geometry allows the use of the lower photon densities. The transducer is either a crystal such as lead zirconate titanate (PZT) or a film such as β-poly(vinylidene difluoride) (PVF). PZT is more sensitive, whereas the wide band PVF films provide a better temporal resolution. PZTs are generally resonant, damped transducers, the highest usable peak frequencies being about 100 MHz, with typical values around 1 MHz. Very high impedance is common to both types of transducers. After impedance matching, the electrical signal is fed into a high gain amplifier (typically $\times$100 and of bandwidth compatible with the transducer in use). The amplified signal is stored by a transient recorder.

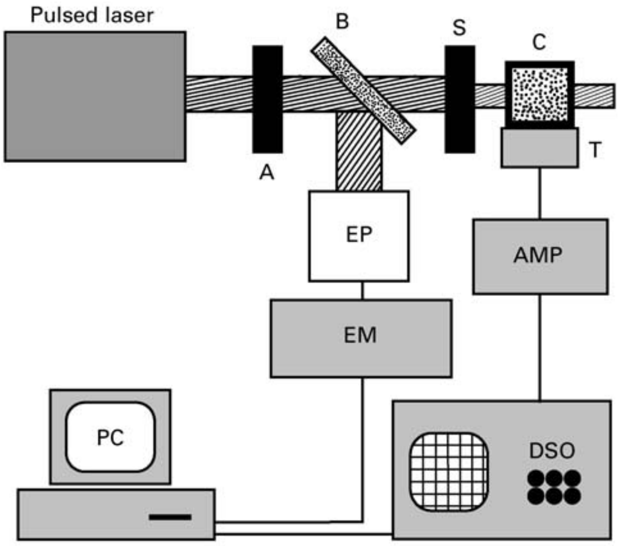



**Figure 2** Block diagram of a typical LIOAS setup. A, variable attenuator; B, beam splitter; S, slit; C, quartz cuvette; T, transducer; AMP, amplifier; EP, energy probe; EM, energy meter; DSO, digital sampling oscilloscope; PC, personal computer.

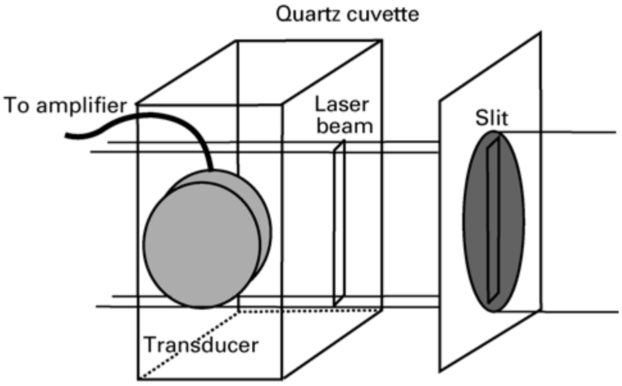



**Figure 3** Close-up of right-angle detection geometry with slit-shaping of the laser beam.

## Applications to selected systems

### The triplet state of a porphyrin ($TPPS_4$)

The high triplet yield of porphyrins is important in a number of photobiological and photochemical

problems, for example in the photodestruction of the photosynthetic apparatus and in photodynamic therapy (PDT). 5,10,15,20-Tetrakis(4-sulfonatophenyl)porphyrin ($TPPS_4$) is a water-soluble porphyrin derivative used as standard compound in PDT research. The LIOAS signal upon excitation of $TPPS_4$ in 10 mM Tris buffer (pH 7.8) at 30°C in the presence of $O_2$ (**Figure 4**) was fitted with a sum of three exponential decays (Eqn [10]), corresponding to the porphyrin triplet formation within a few nanoseconds, its decay (lifetime of about 2 μs) and the decay of singlet molecular oxygen ($O_2(^1\Delta_g)$), lifetime 3–3.5 μs) formed by energy transfer from triplet $TPPS_4$ to ground-state $O_2$ (see **Figure 6**). Measurements at various temperatures (5–30°C) and in $D_2O$ showed that the triplet formation and decay exhibit large structural volume changes, whereas $O_2(^1\Delta_g)$ does not.

Owing to the closeness of the values of the triplet and the $O_2(^1\Delta_g)$ lifetimes, it was not possible to fit the signal with a three-exponential function in the whole temperature range. A sum of two single-exponential functions was used instead. The slow component (fixed at 2.5 μs) represented the triplet plus the $O_2(^1\Delta_g)$ decay. The plot of the parameters $\varphi_i$ (Eqn [12]) vs $c_p\rho/\beta$ (**Figure 5**) shows that triplet formation ($\varphi_1$) is accompanied by a contraction $\Delta V_{r,1} = -10 \pm 3$ mL mol$^{-1}$, whereas triplet decay ($\varphi_2$) is concomitant with an expansion of similar magnitude, $\Delta V_{r,2} = 10 \pm 3$ mL mol$^{-1}$. This structural volume change indicates that triplet $TPPS_4$ has a stronger hydrogen-binding ability than the ground state, inducing a contraction of the surrounding hydration shell. Further LIOAS experiments with $TPPS_4$ and other mesosubstituted porphyrins enabled attribution of the structural volume changes to the photoinduced changes in the interactions with water (via hydrogen bonds) of the N atoms in the pyrrol units of the porphyrin ring.

The molar energy absorbed by the system, $E_\lambda$, is dissipated throughout the four possible processes undergone by the excited molecules, i.e. fluorescence, prompt heat by vibrational relaxation and internal conversion ($S_1 \rightarrow S_0$), intersystem crossing ($S_1 \rightarrow T_1$; $T_1 \rightarrow S_0$), loss during energy transfer to molecular oxygen, and stored heat (Eqn [23]). Phosphorescence from $TPPS_4$ triplet and from $O_2(^1\Delta_g)$ are known to

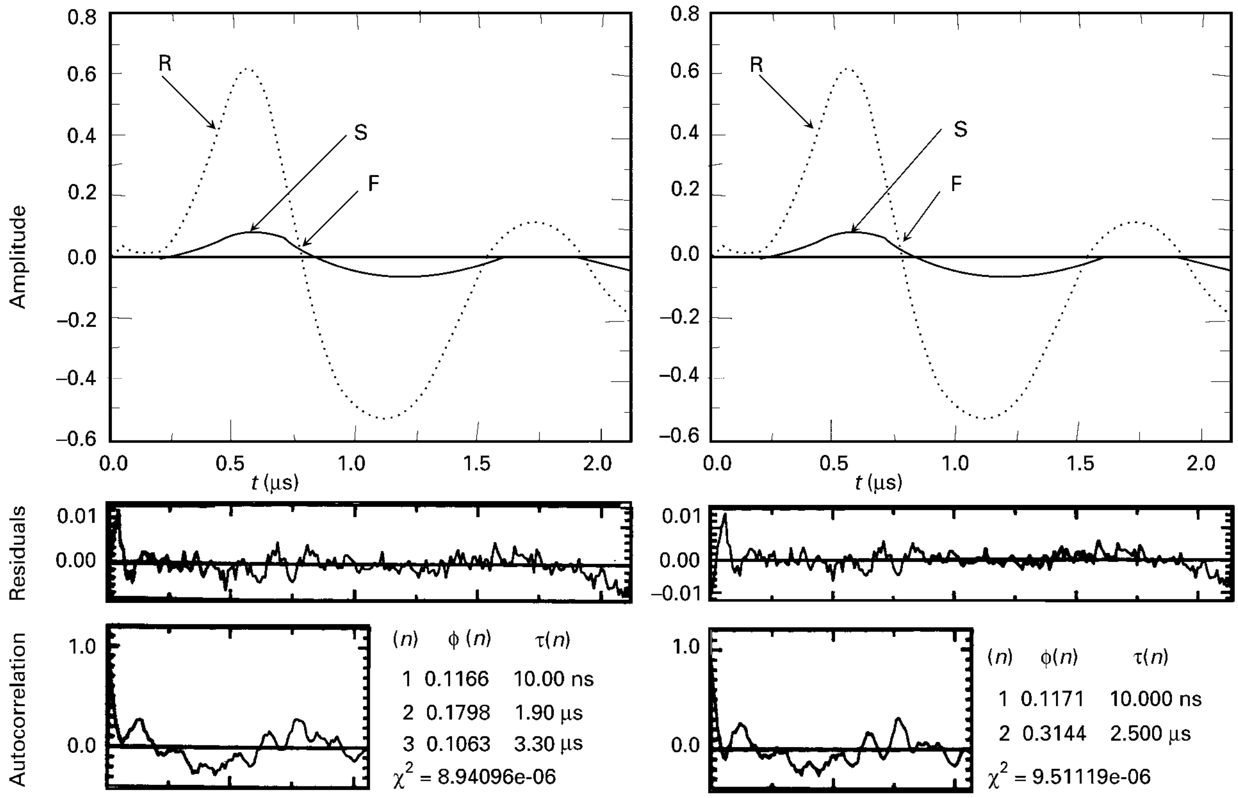



**Figure 4** Fitted curves (F, ——) obtained from deconvolution of the LIOAS signal of $TPPS_4$ (S, -----) in air-satured 10 mM Tris buffer (pH 7.8) at 303 K with two-exponential (right) and three-exponential (left) decay model functions. The reference compound was bromocresol purple (R, ·····). Residuals and autocorrelation function are given for inspection of the goodness of the fits. Index *n* corresponds subindex *i* in text.

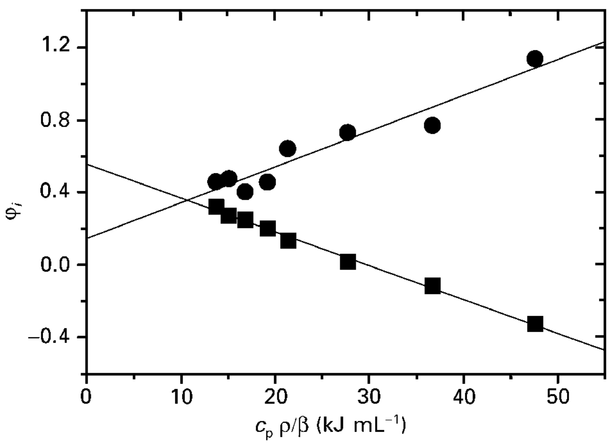



**Figure 5** Plot of $\varphi_1(\tau_1 < 10\,\text{ns})$ and $\varphi_2$ ($\tau_2$ fixed at 2.5 μs) as a function of $c_p\rho/\beta$ (equation [11]) for $TPPS_4$ in air-satured 10 mM Tris buffer (pH 7.8) with bromocresol purple as a calorimetric reference. Reproduced with permission from Gensch T and Braslavsky SE (1997) *Journal of Physical Chemistry B* **101**: 101–108. Copyright 1997, American Chemical Society.

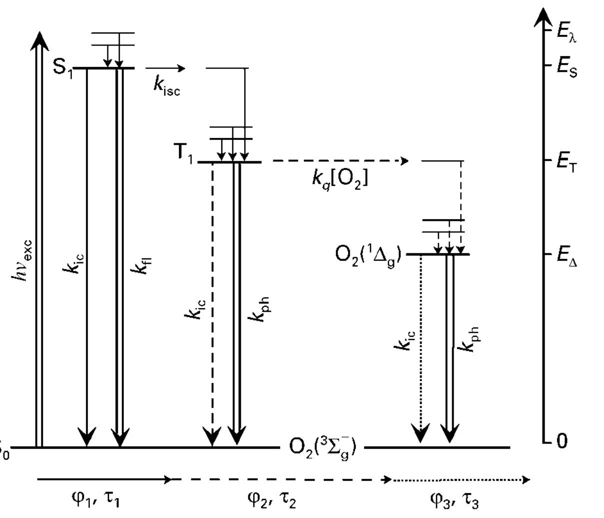



**Figure 6** Term diagram for the photophysical processes occurring after excitation of $TPPS_4$ in aerated aqueous solutions. Double arrows denote light absorption and radiative decay process not detectable with LIOAS. Single full, dashed and broken arrows denote radiationless processes contributing to $\varphi_1$, $\varphi_2$, and $\varphi_3$. Subscript fl denotes fluorescence, ic internal conversion, isc intersystem crossing, ph phosphorescence, $K_q\,[O_2]$ is the quenching constant of triplet $T_1$ by oxygen.

have yields below $10^{-3}$ and were neglected.

$$E_\lambda = \Phi_f \cdot E_f + q_1 + q_2 + q_3 \qquad [23]$$

$$\begin{aligned} q_1 &= E_\lambda - E_S + \Phi_f(E_S - E_f) \\ &\quad + (1 - \Phi_f - \Phi_T)E_S + \Phi_T(E_S - E_T) \\ q_2 &= (\Phi_T - \Phi_\Delta)E_T + \Phi_\Delta(E_T - E_\Delta) \\ q_3 &= \Phi_\Delta E_\Delta \end{aligned}$$

The heat released during the formation of the triplet state as calculated from the intercept of the plot (**Figure 5**, Eqn [12]), $\alpha_1 = 0.58 \pm 0.05$, was used to determine the triplet quantum yield $\Phi_T = 0.61$, by using the literature values $E_f = 174$ kJ mol$^{-1}$, $E_S = E_{0-0} = 186$ kJ mol$^{-1}$ $\Phi_f = 0.06$, and $E_T = 139$ kJ mol$^{-1}$ (at 77 K) and a simple energy balance equation (see **Figure 6** for the meaning of the symbols). Alternatively, the average $\Phi_T = 0.67$ (from literature) can be used to calculate $E_T = 126$ kJ mol$^{-1}$, which is the first report of a triplet energy value for this porphyrin at room temperature.

## Porphyrin to quinone electron transfer

The interest in understanding the molecular basis of the biologically fundamental electron transfer process from the chlorophyll derivative pheophytin to a quinone in photosynthesis has led to the development of a large number of models, such as synthetic systems with porphyrin derivatives as donor and quinones as acceptor. Feitelson and Mauzerall demonstrated the power of LIOAS for the determination of rate constants, quantum yields and structural volume changes in solution of such systems. The triplet yields and electron transfer efficiencies were determined in various solvents. Electrostriction after formation of the ion pair was suggested as the source of structural volume change leading to a contraction in the charge-separated state. Electrostriction describes the contraction of the solvent around a charged molecule. The magnitude of the effect is inversely proportional to the radius of the molecule and to the dielectric constant of the solution. A problem in this connection is that the theoretically derived values using a solvent continuum model are far to small compared to the measured effects. This problem needs further attention.

The structural volume changes for triplet formation of metal-free uroporphyrin (−0.45 cm$^3$ mol$^{-1}$) and of its zinc complex (−1.1 cm$^3$ mol$^{-1}$), as well as the structural volume changes due to electrostriction after electron transfer from Zn-uroporphyrin to a quinone (−4.4 cm$^3$ mol$^{-1}$), were estimated in aqueous solutions. The entropy change due to ion formation was then estimated as −202 kJ mol$^{-1}$K for the latter system.

## Ru(II) bipyridyl complexes

Ruthenium(II) bipyridyl (bpy) complexes constitute interesting compounds for LIOAS studies since their photophysical parameters are known and their transient lifetimes are within the integration time of the LIOAS experiment. In all these complexes, the

metal-to-ligand charge transfer (MLCT) state in aqueous solution is formed with unity quantum yield and relaxes within 100–600 ns (depending on the complex) to the ground state.

In particular, the structural volume changes for the formation of the MLCT state of $Ru(bpy)(CN)_4^{2-}$, $Ru(bpy)(CN)_3(CNCH_3)^-$, $Ru(bpy)_2(CN)_2$, and $Ru(bpy)_3^{2+}$ of 15, 10, 5, and –3.5 cm$^3$ mol$^{-1}$ correlate with the number of CN groups. Thus, the structural volume changes are interpreted as expansion of the strong second-sphere donor–acceptor interaction of the CN groups with water after weakening of the π backbone in the Lewis-base cyano substituents concomitant with the ruthenium photooxidation. The complex $Ru(bpy)_3^{2+}$ (with no CN groups) shows a small contraction of –3.5 cm$^3$ mol$^{-1}$ due to intrinsic Ru–bpy bond shortening. The lifetimes derived by deconvolution were similar to those obtained by emission.

For the cyano-Ru(II) complexes in neat water, the heats derived from the intercepts of plots with Equation [11] afforded values of the energy content of the MLCT states similar to those derived from emission data. Both for intra- and intermolecular electron transfer reactions of cyano-Ru(II) complexes in aqueous solutions in the presence of salts, the structural volume change correlates with the entropy change of the reaction. An enthalpy–entropy compensation effect occurs upon perturbation of the water structure by the addition of the salts.

$Ru(bpy)(CN)_4$ inside the water pools of sodium bis(2-ethylhexyl)sulfo-succinate (AOT) reverse micelles with a large water pool inside the micelle showed a structural volume change similar to that in neat water, whereas in micelles with a small water pool the structural volume changes were smaller.

These findings support the model of a change in the CN–water interaction as the origin of the expansion upon MLCT formation. In a large water pool $Ru(bpy)(CN)_4$ interacts with free water (i.e. the $Ru(bpy)(CN)_4$ environment is similar to that in neat water), whereas in a small water pool the interactions involve 'bound water' (water molecules interacting with the charged detergent molecules), and the expansion is impaired and reduced.

### Proton transfer and the reaction volume for water formation

The structural volume changes photoinduced in solution during the proton transfer between o-nitrobenzaldehyde and hydroxyls in water have been recently studied by means of LIOAS. Upon photoexcitation, o-nitrobenzaldehyde is converted to o-nitrosobenzoic acid and protons are detached with a quantum yield of about 0.4. At pH around and below neutrality, solvation of the charges is accompanied by a fast, subresolution contraction of the medium of approximately –5.2 ± 0.4 cm$^3$ mol$^{-1}$.

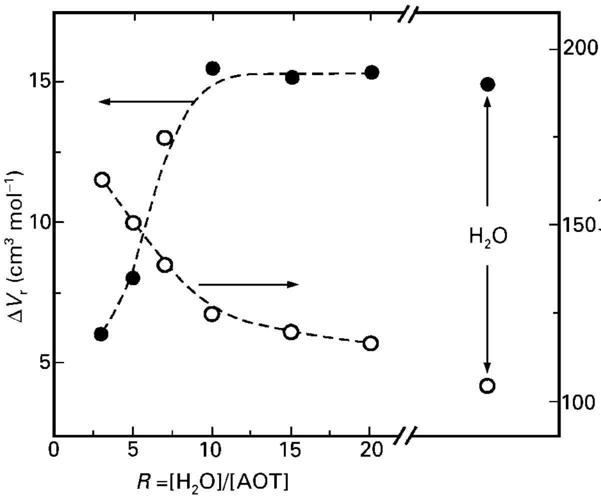



**Figure 7** Structural volume change ($\Delta V_r$) associated with the formation of the MLCT state of $Ru(bpy)(CN)_4^{2-}$ and its decay (lifetime $\tau_2 = \tau_{MLCT}$) in 0.1 M AOT–water–*n*-alkane reverse micelles at various values of *R* = [water]/[AOT]. Reproduced with permission from Borsarelli CD and Braslavsky SE (1997). Nature of the water structure inside the pools of reverse micelles sensed by laser-induced optoacoustic spectroscopy. *Journal of Physical Chemistry B* **101**: 6036–6042 Copyright 1997 American Chemical Society.

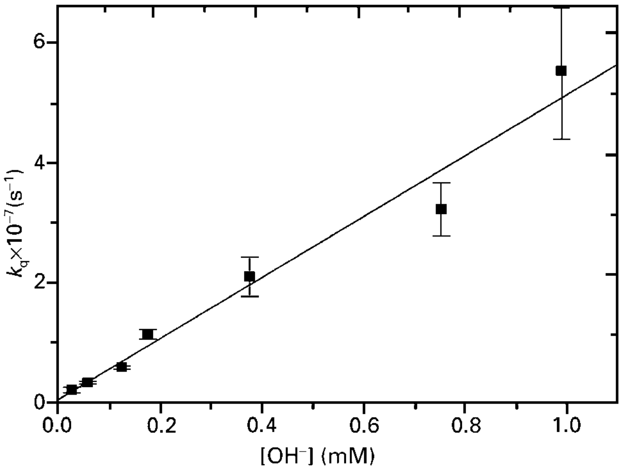



**Figure 8** Rate constant for the reaction between photodetached protons and hydroxyls in water at various $OH^-$ concentrations. Reproduced with permission of the Editors of *Chemical Physics Letters* from Bonetti G, Vecli A and Viappini C (1997) Reaction volume of water formation detected by time-resolved photoacoustics – photoinduced proton transfer between *o*-nitrobenzaldehyde and hydroxyls in water. *Chemical Physics Letters* **209**: 268–273.

This contraction is at least qualitatively accountable for on the basis of the electrostrictive effects described by the Drude–Nernst equation, considering the negative charge localized on the carboxylic oxygen atom of o-nitrosobenzoic acid and using the molar volume of the proton in water. At pH > 9, the reaction with hydroxyls becomes relevant and the formation of water molecules produces an expansion of $24.5 \pm 0.4\ cm^3\ mol^{-1}$. The plot of the quenching rate constant ($k_q$, **Figure 8**), derived from LIOAS data vs the concentration of hydroxyls is linear with a slope of $(4.9 \pm 0.4) \times 10^{10}\ M^{-1}\ s^{-1}$, indicating that the process obeys pseudo-first-order kinetics.

## List of symbols

A = absorbance; $c_p$ = specific heat capacity; $E_\lambda$ = energy of an Einstein of wavelength λ; H = transducer response; k = instrumental factor; n = number of moles of absorbed photons; $N_A$ = Avogadro number; $N_{ph}$ = number of incident photons; q = energy released as heat; T = temperature; $v_a$ = sound velocity; V = volume; $V_0$ = volume of illuminated sample; $\alpha_i = q_i/E_\lambda$; β = cubic expansion coefficient; Δp = pressure change; $\Delta V_e$ = structural volume change per absorbed Einstein; $\Delta V_r$ = volume change due to solute–solvent rearrangements; $\Delta V_{th}$ = volume change due to radiationless relaxation; $\kappa_S$ = adiabatic compressibility; $\kappa_T$ = isothermal compressibility; λ = excitation wavelength; ρ = density; $\tau_i$ = lifetime of transient i; $\Phi_i$ = quantum yield of step i; $\varphi_i$ = fractional contribution to volume change of transient i.

*See also:* **Laser Applications in Electronic Spectroscopy; Laser Microprobe Mass Spectrometers; Laser Spectroscopy Theory; Photoacoustic Spectroscopy, Applications; Photoacoustic Spectroscopy, Theory.**

## Further reading

Bonetti G, Vecli A and Viappiani C (1997) Reaction volume of water formation detected by time-resolved photoacoustic – photoinduced proton transfer between o-nitrobenzaldehyde and hydroxyls in water. Chemical Physics Letters **269**: 268–273.

Borsarelli CD and Braslavsky SE (1997) Nature of the water structure inside the pools of reverse micelles sensed by laser-induced optoacoustic spectroscopy. Journal of Physical Chemistry B **101**: 6036–6042.

Borsarelli CD and Braslavsky SE (1999) Enthalpy, volume, and entropy changes associated with the electron transfer reaction between the $^3$MLCT state of $Ru(bpy)_3^{2+}$ and methyl viologen in aqueous solutions. Journal of Physical Chemistry A **103**: 1719–1727.

Braslavsky SE and Heibel GE (1992) Time-resolved photothermal and photoacoustic methods applied to photoinduced processes in solution. Chemical Reviews **92**: 1381–1410.

Churio MS, Angermund KP and Braslavsky SE (1994) Combination of laser-induced optoacoustic spectroscopy (LIOAS) and semiempirical calculations for the determinations of molecular volume changes: the photoisomerization of carbocyanines. Journal of Physical Chemistry **98**: 1776–1782.

Crippa PR, Vecli A and Viappiani C (1994) Time resolved photoacoustic spectroscopy: new developments of an old idea. Journal of Photochemistry and Photobiology B: Biology **24**: 3–15.

Feitelson J and Mauzerall D (1996) Photoacoustic evaluation of volume and entropy changes in energy and electron transfer. Triplet state porphyrin with oxygen and naphthoquinone-2-sulfonate. Journal of Physical Chemistry **100**: 7698–7703.

Gensch T and Braslavsky SE (1997) Volume changes related to triplet formation of water soluble porphyrins. A laser-induced optoacoustic spectroscopy (LIOAS) study. Journal of Physical Chemistry B **101**: 101–108.

Gensch T, Strassburger JM, Gärtner W and Braslavsky SE (1998) Volume and enthalpy changes upon photoexcitation of bovine rhodopsin derived from optoacoustic studies by using an equilibrium between bathorhodopsin and blue-shifted intermediate. Israel Journal of Chemistry **38**: 231–236.

Mauzerall D, Feitelson J and Prince R (1995) Wide-band time-resolved photoacoustic study of electron-transfer reactions: difference between measured enthalpies and redox free energies. Journal of Physical Chemistry **99**: 1090–1093.

Puchenkov OV (1995) Photoacoustic diagnostics of fast photochemical and photobiological processes. Biophysical Chemistry **56**: 241–261.

Schmidt R and Schütz M. (1996) Determination of reaction volumes and reaction enthalpies by photoacoustic calorimetry. Chemical Physics Letters **263**: 795–802.

Schmidt R and Schütz M (1997) Methodical studies on the time resolution of photoacoustic calorimetry. Journal of Photochemistry and Photobiology A: Chemistry **103**: 39–44.

Schulenberg P and Braslavsky SE (1997) Time-resolved photothermal studies with biological supramolecular systems. In: Mandelis A and Hess P (eds) Progress in Photothermal and Photoacoustic Science and Technology, pp 58–81. Bellingham WA: SPIE Press.

Small JR, Libertini LJ and Small EW (1992) Analysis of photoacoustic waveforms using the non-linear least square method. Biophysical Chemistry **42**: 24–48.

Wegewijs B, Verhoeven JW and Braslavsky SE (1996) Volume changes associated with intramolecular exciplex formation in a semiflexible donor-bridge-acceptor compound. Journal of Physical Chemistry **100**: 8890–8894.

# Laser Magnetic Resonance

**AI Chichinin**, Institute of Chemical Kinetics and Combustion, Novosibirsk, Russia

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
**Methods & Instrumentation**

Laser magnetic resonance (LMR) is a sensitive technique for studying rotational, vibrational-rotational, or electronic Zeeman spectra of paramagnetic atoms and molecules, using fixed-frequency infrared lasers. High resolution and sensitivity are important characteristics of LMR. In a pioneering experiment in 1968, Evenson and his co-workers succeeded in demonstrating the feasibility of this technique when they detected a rotational transition of $O_2$ using a HCN laser. Several years later LMR was extended to the mid-infrared by using CO and $CO_2$ lasers with similar success. Since then, throughout 15 years, much of the increase in understanding of the structure and properties of short-lived paramagnetic molecules has come from the application of LMR. The method has allowed the detection of more than 100 free radicals, including elusive species such as $HO_2$, $CH_2$, FO and Cl($^2P_{1/2}$) which cannot easily be detected by other means. The high sensitivity and versatility of LMR has resulted in numerous applications of the method to chemical kinetics.

## Theory of LMR

### Atoms

As a simple illustration let us consider the LMR spectra of Cl atoms. The energy level diagram at the top of **Figure 1** shows the Zeeman effect in Cl atoms.

In a magnetic field, the $|J, I, F\rangle$ levels of the atom split into components labelled by $M_J$ and $M_F$ quantum numbers (here $F = J + I = L + S + I$). The magnetic field is swept, and at suitable values of the field the frequencies of the transitions $|J, M_J, M_F\rangle \rightarrow \langle J', M'_J, M'_F|$ come into coincidence with the frequency of a fixed-frequency laser. These resonances result in a decrease in detected laser power. In practice, a small modulation (up to 50 G) is added to the magnetic field and the change in transmitted laser power is recorded as a first derivative signal by processing the detector output in a phase-sensitive amplifier.

LMR spectra of Cl atoms are shown at the bottom of **Figure 1**. Note that $^2P_{3/2}$–$^2P_{1/2}$ magnetic dipole transitions with $\Delta M_F = \pm 1$ and $\Delta M_F = 0$ occur at $E \perp B$ and $E \parallel B$ polarizations, respectively; here $E$ is the electric field of the laser radiation and $B$ is the magnetic field of the electromagnet.

The Hamiltonian may be written as

$$H = A_{\text{so}}\boldsymbol{LS} + H_{\text{hfs}}(a, b, \boldsymbol{J}, \boldsymbol{I}) + \mu_0(g_J\boldsymbol{JB} + g_I\boldsymbol{IB})$$

where $H_{\text{hfs}}$ is the hyperfine interaction operator, $A_{\text{so}}$ is a spin-orbit interaction constant, $a$ and $b$ are hyperfine constants, $\mu_0 = e\hbar/2mc$ is the Bohr magneton, and $g_J$ and $g_I$ are electron and nuclear $g$-factors. All these atomic constants can be obtained from the analysis of LMR spectra; the precise values for $A_{\text{so}}$ have been determined in this way for the first time.

### Diatomic radicals

A more common application is the measurement of LMR spectra of linear radicals with both spin and orbital angular momentum (SnH, NiH, GeH, CH, SD, SeH, … etc). The effective Hamiltonian is a summation of spin-orbit ($H_{\text{so}}$), vibrational ($H_{\text{v}}$), rotational ($H_{\text{rot}}$), spin-rotational ($H_{\text{sr}}$), spin-spin ($H_{\text{ss}}$), lambda-doubling ($H_{\text{ld}}$), hyperfine ($H_{\text{hfs}}$) and Zeeman ($H_{\text{z}}$) terms:

$$H = H_{\text{so}} + H_{\text{v}} + H_{\text{rot}} + H_{\text{sr}} + H_{\text{ss}} + H_{\text{ld}} + H_{\text{hfs}} + H_{\text{z}} \qquad [1]$$

The choice of significant terms and representation of the terms are determined by the particular type of radical. Some usual representations are presented below

$$\begin{aligned}
H_{\text{so}} &= A_{\text{so}}\,\boldsymbol{LS}\\
H_{\text{v}} &= \hbar\omega_{\text{e}}(v + 1/2) - \omega_{\text{e}}x_{\text{e}}\hbar(v + 1/2)^2\\
H_{\text{rot}} &= B_v\boldsymbol{N}^2 - D_v\boldsymbol{N}^4\\
H_{\text{sr}} &= \gamma\boldsymbol{NS}\\
H_{\text{ss}} &= \frac{2}{3}\lambda(3S_z^2 - \boldsymbol{S}^2)\\
H_{\text{z}} &= \mu_0(g_{\text{s}}\boldsymbol{S} + g_L\boldsymbol{L} + g_{\text{r}}\boldsymbol{N})\boldsymbol{B}
\end{aligned}$$

Here $v$ denotes the vibrational quantum number and $N = J - L - S$ is the nuclear rotational angular

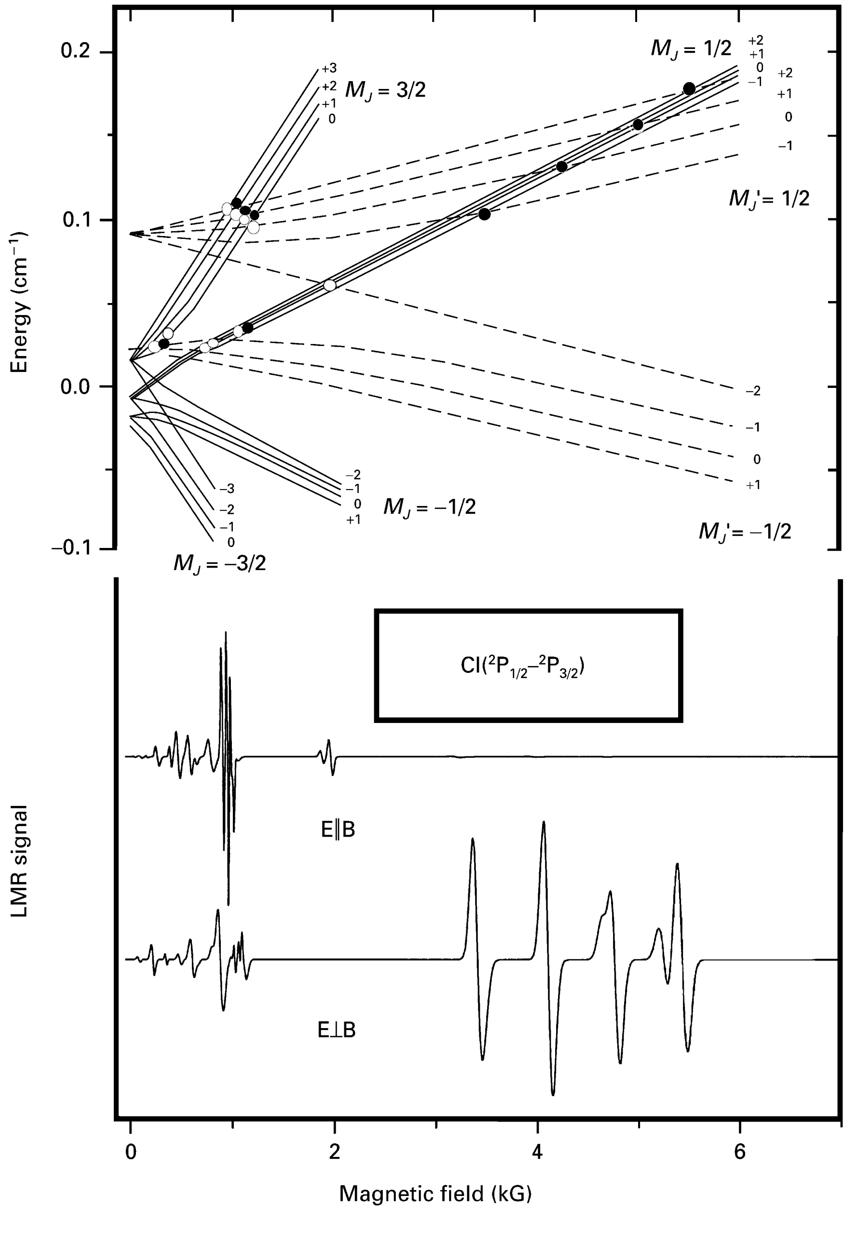



**Figure 1** Top: Zeeman effect in $^{35}$Cl atoms. All sublevels are labelled by $M_F$ values. The lower $^2P_{3/2}$ state is represented by solid lines. Upper $^2P_{1/2}$ state (dashed lines) is downshifted by an amount of the laser photon energy. Intersections of solid and dashed lines with $\Delta M_F = \pm 1$ (open circles) or with $\Delta M_F = 0$ (solid circles) represent possible LMR transitions. Bottom: LMR spectra of Cl atoms (both $^{35}$Cl and $^{37}$Cl).

momentum. Centrifugal corrections to some parameters are often used as well as a tensor expression for $H_z$.

### Polyatomic radicals

The most often encountered type of radical observed by LMR is the three-atomic asymmetric-top radical, such as $NH_2$, $PH_2$, HCO, $HO_2$, etc., Hund's case (b). In this case, the Hamiltonian [1] is used. Without spin-orbit and lambda-doubling terms, $H_{\rm rot}$, $H_{\rm ss}$ and $H_{\rm sr}$ are usually expressed as

$$H_{\rm rot} = A_v N_a^2 + B_v N_b^2 + C_v N_c^2 + H_{\rm cd}(\boldsymbol{N}, \Delta_{\rm N}, \Delta_{\rm NK}, \Delta_{\rm K}, \delta_{\rm N}, \delta_{\rm K})$$

$$H_{\rm ss} = \frac{1}{3} D_v (2S_a^2 - S_b^2 - S_c^2) + E_v (S_b^2 - S_c^2)$$

$$H_{\rm sr} = \varepsilon_{aa} N_a S_a + \varepsilon_{bb} N_b S_b + \varepsilon_{cc} N_c S_c$$

where $H_{cd}$ is the Watson's centrifugal distortion operator; the molecule-fixed components of *N*, *S*, and *ε* are labelled by *a*, *b*, and *c*.

Finally note that the representation of the Hamiltonian is very dependent on the particular radical. Some terms are often omitted, and some new ones are added. For example, Jahn–Teller distortion should be taken into account for symmetric top radicals, such as $CH_3O$ and $SiH_3$.

The LMR detection of polyatomic radicals with heavy (not hydrogen) atoms is hindered for several reasons: (1) the population of the lower state is low, it decreases with rotational partition function; (2) Zeeman splitting is small, it decreases both with moments of inertia and rotational quantum numbers of the radical; hence LMR spectra are observable only at low magnetic fields because of a weak coupling of *S* with *N*. Moreover, the spectra are often unresolved because of a large number of components.

An important exception was found for the first time by Uehara and Hakuta who have observed high-field spectra in the $\nu_1$ band of $ClO_2$ in spite of its small spin-orbit interaction. Later, similar high-field far-infrared LMR spectra of $FO_2$, ClSO, FSO and $NF_2$ were detected. These spectra were assigned to transitions induced by avoided crossings between Zeeman levels having the same value of $M_J$ but differing by one in *N*. This type of transition offers a new means to study radicals which might otherwise be inaccessible to the LMR technique.

## LMR spectrometer

### Basic features

The essential features of an LMR spectrometer are illustrated schematically in **Figure 2**, which shows three of the most commonly encountered experimental configurations.

A conventional intracavity LMR spectrometer is shown in **Figure 2A**. Placing a free-radical absorption cell inside the laser cavity results in a gain in sensitivity due to multipassing of the laser radiation. A further sensibility gain may also be obtained when the modulation frequency is close to the frequency of the laser intensity relaxation oscillations. Another benefit of the intracavity configuration is that it favours the observation of saturation Lamb dips. This is especially important in the mid-infrared region, where Doppler widths are greater and the high resolution of saturation spectroscopy is needed more. The laser power is coupled out by the zeroth order of the grating and is detected by a photoresistor. Coupling through the hole in one of the mirrors or by a variable coupler inserted into the laser cavity are also used.

An intracavity LMR spectrometer based on an optically pumped laser is shown in **Figure 2B**. The only difference is the change of the CO (or $CO_2$) laser in **Figure 2A** by an optically pumped laser. The pump radiation is multiply reflected between two metallic-coated flats parallel to the far-infrared laser axis. Longitudinal pumping in which the pump radiation is introduced along the laser axis through a hole in one of the far-infrared laser mirrors is also employed. The cavity is divided by a beam splitter (e.g. thin polypropylene).

**Figure 2C** gives the block diagram for the Faraday LMR setup, in which there is an extracavity arrangement and detection of paramagnetic species is via their polarization effects. The laser light propagates along the *z*-axis which coincides with the magnetic field direction. The polarization, determined by the polarizer $P_x$ points along the *x*-axis and the polarizer $P_y$ points along the *y*-axis. The multipass absorption cell is placed between these crossed polarizers. The absorbing radicals produce a change in the polarization; this effect is used for sensitive detection of the radicals.

In essence, the Faraday arrangement is extracavity LMR with two crossed polarizers. These polarizers reduce the laser power by a large factor and the LMR signal by the square of that factor. Hence the signal-to-noise ratio for the Faraday arrangement is significantly better (more than two orders of magnitude) than that for conventional extracavity LMR.

Note that time resolution of the intracavity configurations is determined by the medium of the laser: it is usually ~2–4 μs; the time resolution of the Faraday arrangement is considerably better.

The Faraday configuration requires the magnetic field of the electromagnet (solenoid) to be parallel to the Poynting vector, $B \parallel S_p$; An alternative (Voigt) configuration is proposed in which $B \perp S_p$ and a polarization angle relative to *B* is 45°. Although the Voigt configuration is expected to have similar sensitivity to that of Faraday LMR, it is rarely encountered.

Usually, the configurations in **Figure 2A** and **2C** are used in mid-infrared spectroscopy. The configuration in **Figure 2B** is used in far-infrared spectroscopy only. In intracavity configurations, the polarization of the laser with respect to the magnetic field is determined by the intracavity beam splitter in **Figure 2B** or by windows in **Figure 2A** set at Brewster's angle. The splitter (or the windows) are rotatable about the laser axis so the polarization can be rotated. When the electric vector of the laser lies perpendicular to the magnetic field, $E \perp B$, then

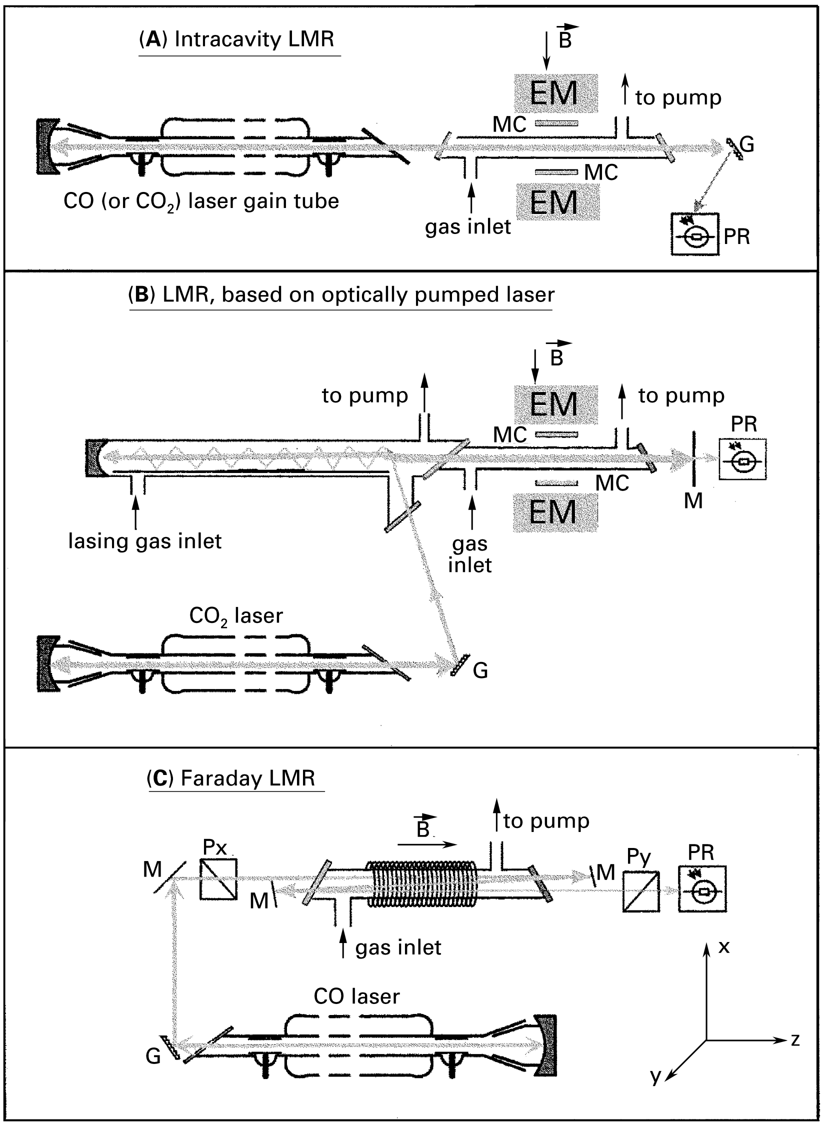



**Figure 2** Schematic representation of a LMR spectrometer: (**A**) mid-infrared intracavity arrangement, (**B**) far-infrared intracavity arrangement with optically pumped laser, (**C**) mid-infrared Faraday extracavity arrangement. EM = electromagnet pole, MC = modulation coil, G = diffraction grating, PR = photoresistor, M = mirror, $P_x$ and $P_y$ = polarizers.

$\Delta M_J = \pm 1$ ($\sigma$) electric dipole transitions are induced; the $\Delta M_J = 0$ ($\pi$) transitions appear with parallel polarization, $E \parallel B$. Note that the latter case is impossible in Faraday LMR.

## Infrared lasers

The first far-infrared LMR spectra were recorded by using HCN, $H_2O$ and $D_2O$ lasers. Since then, optically pumped lasers are usually used with a line-tunable continuous-wave $CO_2$ laser as a light source. By use of these lasers, over 1000 laser frequencies are available in the far-infrared region (30–1200 μm); some 500 have been used in LMR.

In medium-infrared, most of the radical spectra are observed with CO (1200–2000 cm$^{-1}$) and $CO_2$ (875–1110 cm$^{-1}$) lasers. In the former case, the spectral region is now strongly extended by using a CO-overtone laser which provides 200 ($\Delta v = 2$) transitions in the region of 2500–3500 cm$^{-1}$. In the latter case, a great increase in capacity of LMR (several hundreds of lines) was attained by using $^{13}C^{16}O_2$, $^{13}C^{18}O_2$, $^{13}C^{16}O^{18}O$ and $C^{18}O_2$ isotope modifications; an $N_2O$ laser has also been used.

A colour centre laser (2–4 μm), and spin-flip Raman laser (which consisted of a InSb crystal pumped by a CO laser) have also been employed.

### Infrared detectors

The first far-infrared LMR spectra were recorded by using Golay cells. These have the advantage of room temperature operation and are easy to use, but their response is slow and hence only low modulation frequencies (approximately 10–100 Hz) can be used. A further gain in sensitivity was realized by use of a helium-cooled Ge bolometer: the much faster response of the bolometer allows much higher modulation frequencies (~1 kHz), and this, together with the lower noise equivalent power, means that the sensitivity is limited mainly by inherent noise in the laser source. By far the most commonly encountered far-infrared detectors are helium-cooled photoresistors (Ge:B, Ge:In, Ge:Ga, Ge:As, nIn:Sb, etc.); modulation frequencies of up to hundreds of kHz can be used.

In mid-infrared, the most commonly used photoresistors are Ge: Au (77 K), Ge:Hg (53 K), Ge: (Zn, Sb) (64 K) and Hg:Cd:Te (77 K) for $CO_2$ laser output and Ge:Ga (77 K), In:Sb (77 K), Hg:Cd:Te (77 K) for CO laser output.

In general the main noise source is the laser rather than the photoresistor. However, in the mid-infrared region attenuation of laser light is often required since detectors are easily saturated by laser light. For example, for InSb (77 K) detector saturation starts near 1 mW, while typical CO laser output power reaches 1 W. Note that in Faraday and Voigt LMR arrangements, the saturation is absent.

### Modulation

The modulation frequency $f$ is determined as a compromise between two factors. First, at low frequencies the noise of the LMR spectrometer increases due to vibrations of the spectrometer. Second, the modulation amplitude decreases inversely with the square root of $f$ at a fixed power of the generator for modulation coils. Finally, for extracavity LMR the best frequency is 15–20 kHz. For intracavity LMR spectrometers, a gain of sensibility can be obtained when the modulation frequency is close to the frequency of the laser intensity relaxation oscillations (50–150 kHz).

A simple method to improve the sensitivity of LMR is to increase the integration time constant of the lock-in amplifier. However, the long-term instability of the LMR spectrometer usually makes useless time constants larger than 3 s. This problem can be solved by using double modulation. The magnetic field is modulated at high (~100 kHz) and low (~2 Hz) frequencies. The high-frequency signal is demodulated by a phase-sensitive lock-in amplifier. The resulting low-frequency output of the lock-in is again demodulated by another phase-sensitive lock-in. As a result, the effective time constant of the system is ~200 s and the LMR sensitivity increases considerably. Note that the best peak-to-peak amplitude of the low-frequency modulation is equal to the spectral line width. The drawback of this double modulation system is the slow response.

Intracavity systems are very sensitive to acoustic cross talk between the modulation unit and the laser resonator, thus increasing the noise just at the modulation frequency $f$. Hence another method of improving the sensitivity of intracavity LMR is to use a lock-in detector at $2f$ (the second harmonic), where the noise is considerably smaller. In some cases, this change can increase the signal-to-noise ratio of LMR.

### Preparation of radicals

The great majority of radicals studied so far by LMR have been generated by atom–molecule reactions in the gas phase, using discharge-flow techniques or by pumping the products of a microwave discharge rapidly into the sample region of the spectrometer. In addition, a time-resolved arrangement allows generation of radicals either directly by UV photolysis (or multiphoton dissociation), or by the reactions of appropriate molecules with the species prepared photolytically.

### Detection of molecular ions

Molecular ions ($DCl^+$, $DBr^+$, etc.) are generated by discharge-flow techniques. Two difficulties arise in experiments with intracavity LMR detection of these ions. First is the rapid deflection of the ions to the reactor walls in the magnetic field of the electromagnet; the second is the very high noise of a DC discharge situated inside the laser cavity. These problems have been solved by using the Faraday extracavity LMR arrangement, which does not suffer to the same extent from modulation pickup via discharge plasma as the intracavity arrangement. In the Faraday arrangement the discharge is stable, since the magnetic field of the solenoid and the electric field of the discharge are collinear.

An additional advantage of this arrangement is the possibility of tracing hot-band transitions not only for open-shell ions but also for free radicals.

### Sensitivity of LMR, comparison with EPR

The technique of LMR is very similar to other magnetic resonance methods such as EPR and NMR. While NMR uses radiofrequency radiation to produce transitions between nuclear spin levels, and EPR uses microwave radiation to produce transitions between electron spin levels, LMR uses

radiation of a laser to produce transitions between rotational (far-infrared) or vibrational–rotational (mid-infrared) levels in paramagnetic molecules. Note that the sensitivity $\gamma_{\min}$ is nearly equal for EPR and LMR: $\gamma_{\min} \simeq 10^{-10}$–$10^{-9}$ cm$^{-1}$. The concentration sensitivity is given by

$$N_{\min} = \frac{\gamma_{\min}}{\sigma d}$$

where $d$ is the difference of relative populations of the upper and lower levels and $\sigma$ is the cross section of the transition. The superiority of LMR over EPR in sensitivity is determined by the $d$ factor (2–3 orders of magnitude) and by the cross section $\sigma$ which is proportional to the transition frequency. The factor in favour of EPR is the high Q-factor of an EPR resonator. When EPR is compared with mid-infrared LMR, an additional factor in favour of EPR is the dipole transition matrix element, which is an order of magnitude higher for rotational transitions in radicals than that for vibrational–rotational transitions. In general, LMR is much more sensitive than EPR.

The sensitivities of far-infrared intracavity LMR and mid-infrared Faraday LMR are as high as that of laser-induced fluorescence (LIF); the sensitivity of mid-infrared intracavity LMR is usually one order of magnitude less. These estimates, however, are very dependent on the particular radical. The sensitivity achieved in practice also depends on the refinement of the apparatus. The typical values for up-to-date spectrometers are listed in **Table 1**.

For applications in chemical kinetics, combined EPR/LMR spectrometers for both far- and mid-infrared regions have been developed. Three advantages are gained from this arrangement: (i) the versatility of the combined spectrometer, (ii) the possibility of the absolute calibration of the LMR spectrometer by comparison of the LMR and EPR signals of the same radicals, (iii) usage of commercial EPR hardware for LMR detection. In the mid-infrared, EPR and LMR make use of a common detection zone. In the far-infrared, the detection zones of EPR and LMR are placed side by side; they cannot coincide, because of the large wavelength of the far-infrared laser.

## Applications of LMR

### Spectroscopy studies

LMR has allowed the spectroscopic study of many free radicals that could not be detected by more conventional infrared techniques.

**Table 1** Typical sensitivities of LMR and EPR

| *Radical* | *Laser* | *Wavelength* (μm) | *Sensitivity* (*cm*$^{-3}$) |
|---|---|---|---|
| Cl | $CO_2$ | 11.3 | 2(9)[a] |
| O | $CH_3OD$ | 145.7 | 1(10) |
| Hg | CO | 5.67 | 5(9) |
| NO | CO | 5.33 | 1(7) |
| | $CH_3OH$ | 1224 | 1(11) |
| OH | $H_2O$ | 118.6 | 1(6) |
| | $D_2O$ | 84.3 | 2(7) |
| | EPR | 30 000 | 1(10) |
| $O_2$ | $CH_3OH$ | 699.5 | 5(10) |
| | EPR | 30 000 | 3(13) |
| ClO | $CD_3I$ | 556.9 | 2(8) |
| | $H_2O$ | 118.6 | 4(8) |
| $HO_2$ | $CO_2$ | 10 | 1(10) |
| | EPR | 30 000 | 1(13) |
| $NO_2$ | $H_2O$ | 118.6 | 1(11) |
| $CH_2$ | $^{13}CH_3OH$ | 157.9 | 3(8) |
| $NF_2$ | $CO_2$ | 10 | 1(10) |
| | EPR | 30 000 | 1(14) |

[a] $a$(b) = $a$ x $10^b$.

In **Table 2** a list of the species detected by LMR to 1998 is presented. However, the pace of developments has slowed, since many of the better-known free radicals accessible for LMR have now been observed. As for lesser-known free radicals, LMR is a difficult method for obtaining initial knowledge (at least for polyatomics) because the Zeeman effect must be analysed as well as the zero-field spectrum. In the mid-infrared, the tunable semiconductor diode lasers are another factor in the development of infrared spectroscopy of high sensitivity and resolution, although LMR provides magnetic parameters inaccessible to the diode laser spectroscopy. In **Figure 3** the rise and the fall of LMR are illustrated.

A typical spectroscopic study of a radical includes: recording of numerous LMR spectra, assignment of the observed resonances to quantum numbers, and determination of the parameters of the Hamiltonian (for both ground and vibrationally excited states in the case of vibrational–rotational spectra). Although the accuracy of these parameters (up to MHz) is somewhat lower than that obtained by microwave spectroscopy, it is not restricted to rotational transitions but includes vibrational, fine structure and hyperfine effects.

The main requirements for a successful LMR study are: (i) the atom or molecule must be paramagnetic ($S > 0$ or $L > 0$); (ii) the closeness of laser and transition frequencies (<1 cm$^{-1}$); (iii) usually, no more than two heavy (not hydrogen) atoms [exceptions are non-hydride triatomics which are either linear (NCO, $N_3$, NCN), or which have

**Table 2** Free radicals detected by LMR[a]

| | |
|---|---|
| $H_2O$, $D_2O$, HCN, InSb[b] lasers | OH, OD, CH, $O_2$, NO, PH, $HO_2$, $NH_2$, $NO_2$, HCO, $PH_2$ |
| Optically pumped lasers | C, O, Si, CH, CD, CF, $O_2$, $O^{17}O$, $O^{18}O$, OH, $^{17}OH$, OD, NH, $^{15}NH$, ND, ClO, SiH, PH, PD, PO, SH, SD, NS, $S_2$, FeH, CoH, GeH, SeH, SeD, NiH, NiD, $CH_2$, $CD_2$, CCH, $CH_2F$, $CH_2Br$, $CD_2Br$, $NH_2$, NCO, NHD, $HO_2$, $DO_2$, $FO_2$, $PH_2$, $PO_2$, ClSO, FSO, $FeD_2$, $CH_3$, $NH_2O$, $CH_3O$, $CH_2OH$, N*, Mg*, NF*, $O_2$*, CO*, CH*, AsH*, $CH_2$*, $OH^+$, $OD^+$, $HCl^+$, $HBr^+$ |
| $CO_2$ laser | Cl, FO, BrO, SH, SD, SO, SiC, NiH, AsO, SeO, NSe, $CH_2$, $^{13}CH_2$, $NH_2$, $ND_2$, $PH_2$, $HO_2$, $DO_2$, HCO, HSO, FCO, $FO_2$, $ClO_2$, $NO_2$, $NF_2$, $SiH_3$, $SiF_3$, $CH_3O$, Kr*, Xe*, SO* |
| CO laser | NO, ClO, CD, CF, MgO, SiH, SiD, PH, PD, SD, FeH, CoH, CrH, NiH, NiD, GeH, GeD, AsH, SeH, SeD, SnH, SbH, TeD, $C_2H$, $C_2D$, NCN, NCO, $NH_2$, $NO_2$, $^{15}NO_2$, $N_3$, HCO, DCO, $HO_2$, $DO_2$, $FO_2$, $FeH_2$, Hg*, CO*, NCO*, $DCl^+$, $DBr^+$, $SD^+$ |

[a] Electronically excited species are marked by *.
[b] Spin-flip Raman laser: InSb crystal pumped by CO laser.

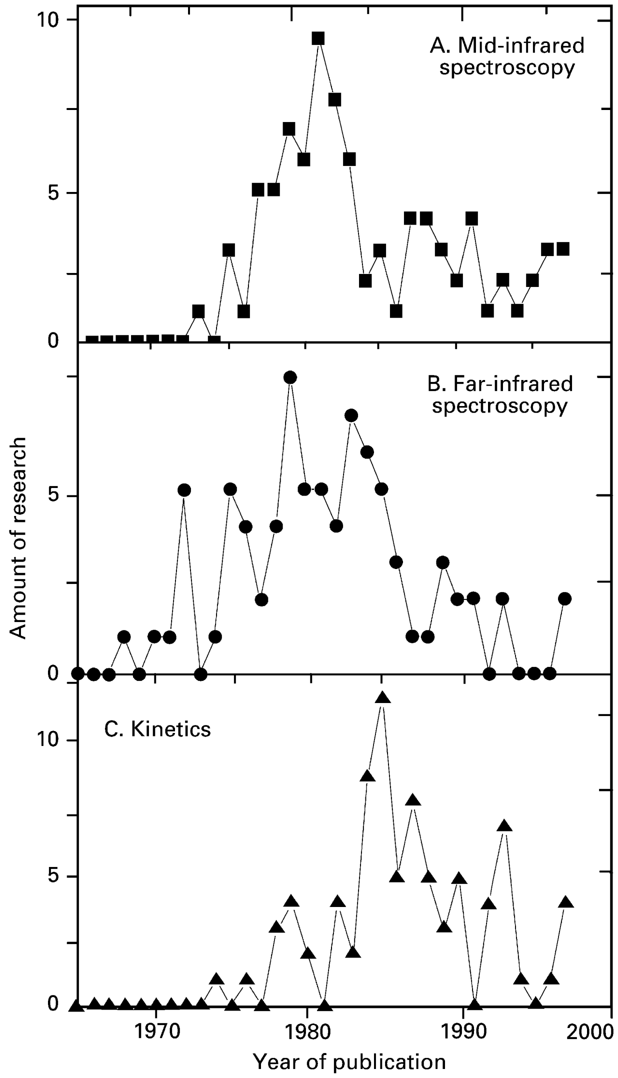



**Figure 3** Amount of research (number of publications) against year of publication. C. Kinetics = chemical kinetics.

high-field LMR spectra due to level anticrossing mechanisms ($FO_2$, ClSO, FSO, $ClO_2$)]; (iv) a moderate transition dipole moment; (v) numerous resonances at numerous laser lines ought to be detected for successful extraction of the Hamiltonian parameters. If the spectra are complex and energy level predictions are not available, the data from other spectroscopic techniques are especially valid.

## LMR in chemical kinetics

The combination of high sensitivity and versatility makes LMR attractive in chemical kinetics studies. Three experimental arrangements of LMR are usually employed.

(i) The first are discharge-flow techniques in which the products of a microwave discharge and consequent chemical reactions are pumped rapidly through the detection zone of a LMR spectrometer. The kinetics are obtained by varying the distance between the radical injector and the detection zone.
(ii) The second is time-resolved LMR in which radicals are produced by either UV photolysis of multiphoton dissociation of a suitable precursor; the LMR signal kinetics is monitored.

These experimental arrangements for kinetics studies are not specific to LMR; they are widely used with other spectroscopic methods, such as EPR, LIF, mass spectrometry, etc.

(iii) The third arrangement is specific to LMR; it is used only to study vibrational relaxation of radicals; an intracavity LMR spectrometer based on a $CO_2$ laser is usually employed. Radicals are prepared in a discharge-flow system; the kinetics of the LMR signal saturation by a $CO_2$ laser radiation field after a fast magnetic field jump are monitored. The magnetic field jump provides fast adjustment to the absorption spectral line of the radical. The saturation kinetics are exponential, the exponent decay time being

$$1/\tau = 2\sigma J_0 + \sum_i k_i[\mathrm{M}_i]$$

where $\sigma$ is the vibrational–rotational transition cross section, $J_0$ is the photon flux density, $[M_i]$ is the concentration of the $i$th gas-relaxator and $k_i$ are the vibrational relaxation rate constants. This relation is used to obtain $k_i$ by measuring $\tau$ at different $[M_i]$.

This method is applicable only to radicals that show sufficient saturation (>10%) of a vibrational transition by the radiation field. The prerequisites for successful detection of the saturation are: radiation intensity in the cavity of the $CO_2$ laser of >100 W $cm^{-2}$, rather a small rotational partition function (≤$10^3$, i.e. nonlinear radicals with low moments of inertia, or linear radicals), and moderate transition dipole moment (~0.1 D).

## List of symbols

$A_{so}$ = spin-orbit interaction constant; $A_v$ $B_v$, $C_v$ = rotational constants of radical; $a$, $b$ = hyperfine constants; $B$ = external magnetic field; $d$ = difference of relative populations; $D_v$ = centrifugal constant for diatomic molecule; $D_v$, $E_v$ = parameters of Hamiltonian $H_{ss}$; $E$ = electric field of laser radiation; $f$ = modulation frequency; $g_J$, $g_I$ = electron and nuclear $g$-factors; $g_s$, $g_L$, $g_r$ = spin, orbital and rotational $g$-factors; $\hbar$ = Planck's constant; $H_{cd}$ = Watson's centrifugal distortion operator; $H_{hfs}$ = hyperfine interaction operator; $H_{ld}$ = lambda-doubling term in Hamiltonian; $H_{rot}$ = rotational term in Hamiltonian; $H_{so}$ = spin-orbit term in Hamiltonian; $H_{sr}$ = spin-rotational term in Hamiltonian; $H_{ss}$ = spin-spin term in Hamiltonian; $H_v$ = vibrational term in Hamiltonian; $H_z$ = Zeeman term in Hamiltonian; $J$, $I$, $F$, $L$, $S$, $N$, $M_F$, $M_J$ = momenta and their projections; $J_0$ = photon flux density; $k_i$ = vibrational relaxation rate constant; $[M_i]$ = concentration of the $i$th gas-relaxator; $N_a$, $N_b$, $N_c$ = components of $N$; $N_{min}$ = concentrational sensitivity; $^2P_{3/2}$, $^2P_{1/2}$ = Cl atomic states; $S_a$, $S_b$, $S_c$, $S_z$ = projections of $S$; $S_p$ = Poynting vector; $v$ = vibrational quantum number; $w_e$, $x_e$ = parameters of vibrational Hamiltonian; $\gamma$ = parameter of spin-rotational term $H_{sr}$; $\gamma_{min}$ = sensitivity in $cm^{-1}$ units; $\Delta_N$, $\Delta_{NK}$, $\Delta_K$, $\delta_N$, $\delta_K$ = centrifugal constants of asymmetric rotor; $\lambda$ = parameter of spin-spin term $H_{ss}$; $\varepsilon_{aa}$, $\varepsilon_{bb}$, $\varepsilon_{cc}$ = components of spin-rotational tensor; $\mu_0 = e\hbar/2mc$ = Bohr magneton; $\sigma$ = cross section of the transition; $\tau$ = decay time for the saturation kinetics.

*See also:* **Atomic Absorption, Theory; Electromagnetic Radiation; EPR, Methods; EPR Spectroscopy, Applications in Chemistry; EPR Spectroscopy, Theory; Far-IR Spectroscopy, Applications; High Resolution IR Spectroscopy (Gas Phase) Instrumentation; IR Spectroscopy, Theory; Laser Spectroscopy Theory; Near-IR Spectrometers; Rotational Spectroscopy, Theory; Spectroscopy of Ions; Zeeman and Stark Methods in Spectroscopy, Applications; Zeeman and Stark Methods in Spectroscopy, Instrumentation.**

## Further reading

Davies PB (1981) Laser magnetic resonance spectroscopy. *Journal of Physical Chemistry* **85**: 2599–2607.

Evenson KM (1981) Far-infrared laser magnetic resonance. *Faraday Discussions of the Chemical Society* **71**: 7–14.

Evenson KM, Saykally RJ, Jennings DA, Curl RF and Brown JM (1980) In: Moore CB (ed) *Chemical and Biochemical Applications of Lasers*, Vol 5, p 95. New York: Academic Press.

Hills GW (1984) *Magnetic Resonance Review* **9**: 15–64.

Russell DK (1983) *Laser Magnetic Resonance Spectroscopy Electron Spin Resonance: A Specialist Periodical Report the Royal Society of Chemistry*, Vol 8, pp 1–30. London: Royal Society of Chemistry, Burlington House.

Russell DK (1991) *Specialist Periodical Report of the Royal Society of Chemistry* **12**B: 64–98.

# Laser Microprobe Mass Spectrometers

**Luc Van Vaeck** and **Freddy Adams**, University of Antwerpen (UIA), Wilrijk, Belgium

Copyright © 1999 Academic Press

MASS SPECTROMETRY
Methods & Instrumentation

## Introduction

The defining attribute of laser microprobe mass spectrometry (LMMS) is the use of a focused laser to irradiate a 5–10 μm spot of a solid sample at a power density above $10^6$ W $cm^{-2}$. The photon–solid interaction yields ions which are mass analysed by time-of-flight (TOF) or Fourier transform (FT) MS. The technique is sometimes referred to as laser probe microanalysis (LPA or LPMA), laser ionization mass analysis (LIMA) and laser microprobe mass analysis (LAMMA).

The use of lasers to ionize solids in mass spectrometry dates back to the early 1960s, when the first high-power pulsed lasers became available. These highly directional and intense monochromatic beams can be easily introduced in the confined space of ion sources without disturbing the electrical fields. Moreover, upon the photon interaction with nonconducting materials, no charging of the sample occurs. First, lasers became an interesting alternative for spark-source MS to quantify elements in dielectrical specimens. Later, the ultrafast heating of the solid by laser irradiation was exploited for the desorption and ionization (DI) of labile organic compounds without thermal decomposition. The growing interest in microanalysis triggered the application of a focused laser in the 1970s. The initial aim was again elemental analysis in nonconducting samples but the main strength of LMMS was found to be the molecule-specific information on the local constituents. Organic and inorganic compounds are characterized by a combination of fragments and adduct ions. The former arise from the structure-specific breakdown of the analyte. The latter simply consist of the intact analyte attached to one or more of these stable fragments.

The understanding of material properties on a local level implies knowledge of the chemistry and, hence, of the molecules present. Therefore the molecular specificity of LMMS is a major advantage in comparison with most micro- and surface analysis techniques, which characterize relative elemental abundances, bonds present or functional groups. Quantitation in LMMS is difficult because adequate reference materials are not always available. However, qualitative identification of local constituents can often be achieved by deductive reasoning only. As a result, LMMS became appreciated as a versatile tool in diverse problem-solving applications in science and technology. The main sample requirement is its stability in the vacuum.

## Instrumentation

The simple construction and operation, excellent transmission, and the registration of full mass spectra made a TOF analyser the obvious choice for the early LMMS instruments. Later it was found that the time definition of the ion production from the laser pulse was less suited to TOF MS than initially assumed. Also, the complex mass spectra required a significantly higher mass resolution and mass accuracy than that available in TOF LMMS. Hence, FT LMMS was developed. Although the laser microbeam ionization was retained, the species detected and their relative intensities differ significantly from TOF LMMS. The reason is that each MS may capture a different fraction of the initial ion population, depending on its acceptance with respect to the initial kinetic energy ($E_{kin}$), their time and place of formation in the ion source. Therefore, a clear distinction between TOF and FT LMMS is mandatory.

### TOF LMMS

**Figure 1** depicts a commercial instrument. The sample is mounted inside the vacuum chamber of the MS. The DI is commonly performed by 266 nm UV pulses from a frequency-quadrupled Nd:YAG (neodymium–yttrium–aluminium–garnet) laser. This instrument allows repositioning of the sample and the optics for analysis in transmission or in reflection. In the former case, the laser hits the lower surface of the sample while the upper surface faces the MS. This suits thin films or particles of about 1 μm on a polymer film. Reflection means that the beam impinges on the sample side facing the MS. The surface of bulk samples can thus be characterized. Micropositioners allow one to move the spot of interest on

the sample into the waist of the ionizing beam, which is visualized by a collinear He-Ne laser.

The UV-irradiated spot is typically 1-3 μm in TOF LMMS so that power densities between $10^6$ and $10^{11}$ W $cm^{-2}$ can be attained with mJ pulses of 10-15 ns. Refractive objectives allow spot sizes down to the diffraction limit of 0.5 μm. Reflective optics (as in **Figure 1**) allow larger working distances and are free from chromatic aberrations so that use of other wavelengths is facilitated. Ionization with a tunable dye laser allows resonant one-step DI or postionization of the laser ablated neutrals. Specificity and sensitivity can thus be improved substantially but wavelength selection becomes cumbersome for unknown compounds.

The principles of a TOF mass analyser are covered in another article. Commercial TOF LMMS permits a mass resolution of about 500, a much lower value than theory predicts, and this figure strongly depends on the analyte. The reason is quite fundamental. The application of TOF implies that ions with different m/z and therefore different velocities must arrive at the same time at the entrance to the drift tube. The pulsed laser ionization does not meet this requirement as ion formation may continue long after the laser pulse, especially for organic compounds but also for inorganic analytes. As a result, mass resolution and mass accuracy may become problematic. The asset of an unlimited m/z range is cancelled by the low mass resolution. A major advantage is the inherent panoramic registration.

## FT LMMS

**Figure 2** illustrates an instrument with an external ion source and shows the different microprobe related devices such as the optical interface, the micropositioners, the sample observation and the exchange system. This set-up features laser irradiation of the sample in reflection with a spot of 5 μm. Electrostatic fields transport ions through a differentially pumped transfer line from the source at $10^{-6}$ torr to the FT MS cell at $10^{-10}$ torr inside a 4.7 tesla magnetic field (B). Inherent advantages of FT MS are the routinely obtainable high mass resolution of over 100 000 at m/z 1000, and up to a few million below m/z 100, and the mass accuracy of better than 1 ppm up m/z 1000. Internal calibration of the m/z scale by adding a reference compound is not necessary as is the case in magnetic high-resolution mass spectrometers. A remarkable feature of FT MS is that better mass resolution means better sensitivity. This contrasts with other types of MS where resolution is increased by selecting a central fraction of the ion beam, thereby sacrificing sensitivity. The direct link of sensitivity and mass resolution in FT MS results from the influence of space charge effects, field imperfections and pressure in the cell. All these cause a faster decay of the coherence in the ion packets after excitation, which decreases mass resolution, but they also limit the quantitative trapping and/or the radius increase of the ion orbit during excitation, which diminish detection sensitivity. FT MS is also limited in the

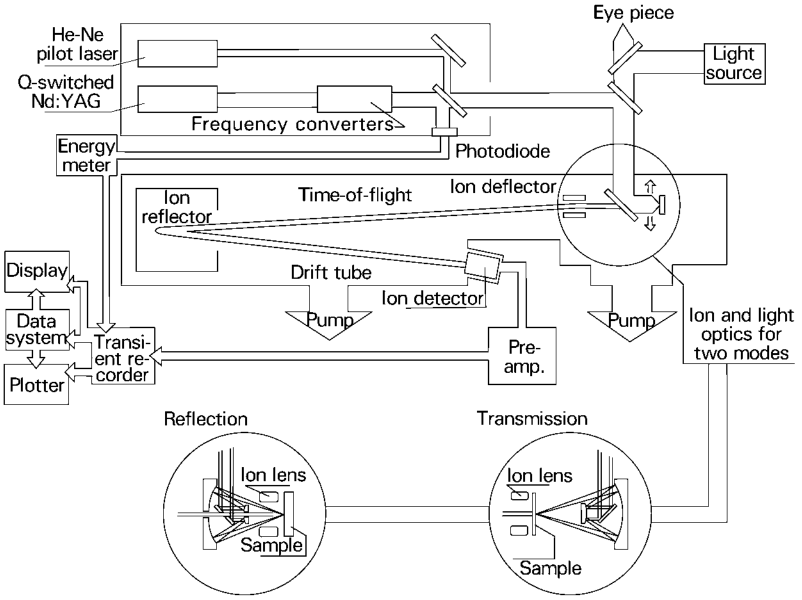



**Figure 1** Schematic diagram of the LIMA 2A TOF LMMS. Reprinted from the LIMA technical documentation with permission of Kratos Analytical.

detection of low and high m/z by its detection bandwidth and B, respectively. A range between m/z 15 and 10 000 is common.

The use of FT MS in microanalysis implies that the technique is operated close to its detection limits. Therefore, the vacuum in the cell becomes a prime reason for using an external ion source be used for LMMS as opposed to a single- or dual-cell instrument, where the sample is placed in the FT MS cell. Additionally, the $E_{kin}$ of laser ions can exceed the energy range imposed by the optimal trapping potentials of typically less than 2 V. Biasing the sample holder or placing additional electrodes around the sample inside the cell disturbs the trapping field. Cooling of the ions by ion-molecule interactions is not always efficient. The use of micropositioners, laser and observation optics inside the narrow bore of the strong magnetic field, is not obvious. However, injection of external ions requires temporary deactivation of the potential gradient in the hole of the first trapping plate. Upon reflection against the second trapping plate, ions come back and may escape unless the entrance is blocked. The use of an electrostatic transfer line causes a time dispersion of ions of different m/z as in TOF MS, i.e. the so-called TOF effect. To trap a high m/z ion together with a low m/z ion, the high m/z ion, which arrives later, must enter the cell before the low m/z ion (which arrived earlier and has been reflected by the second trapping plate) leaves the cell. In practice, the m/z ratio of high-to-low m/z ions that can be trapped simultaneously in the cell is about a factor of 3-4, depending on the time of ion formation and life time of the specific ions. Adapting the time between the laser pulse and the closure of the injection lenses, allows one to shift the m/z window along the m/z range in successive experiments.

## Analytical characteristics

**Table 1** surveys the common techniques for micro- and surface analysis of solids. Both static secondary ion mass spectrometry (SSIMS) and LMMS provide molecular information on local organic and inorganic compounds. However, the primary interaction of keV ions with the sample in SSIMS as opposed to eV-range photons in LMMS makes the direct structural linkage of the detected signals to the sample composition less obvious in SSIMS. Hence, comparison with reference spectra is usually required in SSIMS, while LMMS allows a deductive interpretation. In fact, both methods are complementary with respect to their typical applications. SSIMS can detect high m/z ions from polymers, while LMMS yields more detailed structural information on analytes of up to a few kDa. SSIMS generates primarily ions from the upper monolayer, making surface contamination a problem in real-life applications. In contrast, ions in LMMS originate from the upper 10-50 nm surface layer, although the crater depth goes up to 0.1-1 μm. SSIMS allows imaging, while LMMS performs spot analysis.

The reproducibility of LMMS strongly depends on the positioning of the sample surface in the waist of the UV beam. Hence, mapping by motorizing the

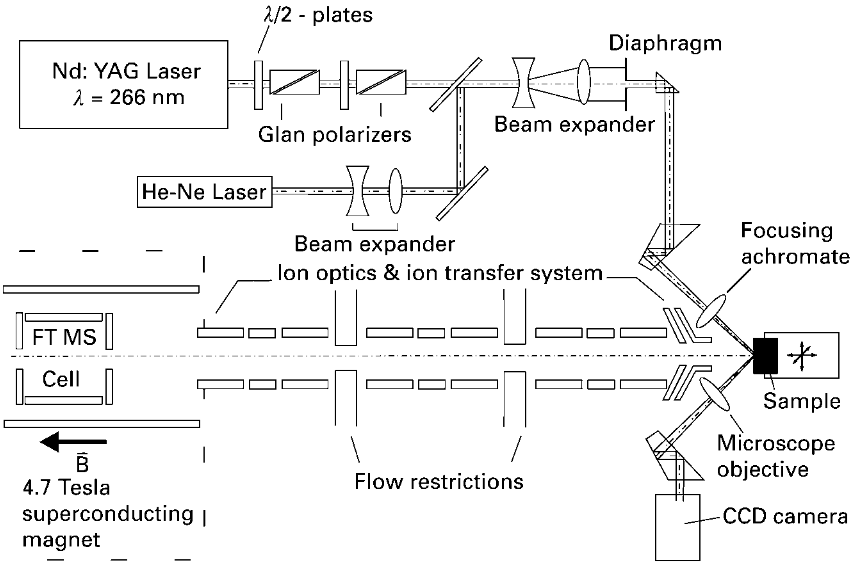



**Figure 2** Schematic diagram of the FT LMMS with external ion source. Adapted from Struyf H, Van Roy W, Van Vaeck L, Van Grieken R, Gijbels R and Caravatti P (1993) *Analytica Chimica Acta*, **283**: 139–151 with permission of Elsevier Science.

**Table 1** Overview of some important microanalytical techniques

| | | | | | | *Dynamic SIMS* | | *Static SIMS* | *LMMS* | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *EPXMA* | *Auger* | *ESCA* | *Raman* | *FT-IR* | *Stigmatic* | *Scanning* | *TOF* | *TOF* | *FT* |
| Probe input beam | 20 keV electrons | Electrons ≤ 3 keV | Photons X-ray, UV | Photons visible | Photons IR | ≤ 20 keV ions | 40–60 keV ions | 20 keV | UV photons (few eV) | |
| Detected beam | X-rays | Electrons | Electrons | Photons | Photons | +/– ions | | +/– ions | +/– ions | |
| Parameter | WDS, EDS | Energy | Energy | Wavelength | Wavelength | *m/z* | | *m/z* | *m/z* | |
| Resolution of detector | WDS 20 eV EDS 150 eV | 1–15 eV | 0.3–1 eV | 0.7 $cm^{-1}$ (sp) 8 $cm^{-1}$ (im) | 0.7 $cm^{-1}$ (IR) 2 $cm^{-1}$ ( µIR) | 400–$10^4$ | 250–500 | $10^3$–$10^4$ | ≤850 | $10^4$–$10^6$ |
| Typically analysed area | ±1 µm | 0.2 µm | ≥150 µm ≥5 µm (µESCA) | 1 µm | 5–10 µm | 2–250 µm | ≥ 20 nm | 0.1 µm | 1–3 µm | 5 µm |
| Depth of information | ≤1 µm | 1–3 nm | 1–10 nm | 10 µm 5–50 nm (PAS) | 10 µm | 0.5 nm | 1 nm | Monolayer | (0.1–1 µm) | n.a. |
| Image resolution | SEM/X ±1 µm STEM/X< µm | 50–100 nm | No Yes (µESCA) | µm | n.a. 1 mm (µIR) | 0.5 mm | 20 nm | 0.5 µm | 1 µm | >1 µm |
| Detection limit | WDS 100 ppm EDS 1000 ppm | ≥ 1 % | ≥ 1 % | Major | ppm | ≤ ppm | 10–100 ppm | Monolayer | $10^{-3}$–$10^{-15}$ | $10^{-11}$–$10^{-12}$g |
| Detection range | WDS $Z \geq 4$ EDS $Z \geq 11$[a] | All but H, He | All but H, He | n.a. | n.a. | H–U | | H-unlimited | H-unlimited | 15–15 000 |
| Direct isotope information | No | No | No | No | No | Yes | | Yes | Yes | |
| Compound speciation | No | No | Yes | Yes | Yes | No | | Yes | Yes | |
| Organic characterization | No | No | No | Yes | Yes | Yes | | Yes | Yes | |
| Destructive | (No) | No | No | No | No | Yes | | Yes | Yes | |
| with sputtering | n.a. | Yes | Yes | n.a. | n.a. | n.a. | | n.a. | n.a. | |
| In-depth profiling | No | No | No | No | No | Yes | | No | Yes | |
| With sputtering | n.a. | Yes | Yes | n.a. | n.a. | n.a. | | n.a. | n.a. | |
| Quantification | Yes | Yes | Yes | (Yes) | Yes | Difficult | | Difficult | Very difficult | |
| Easy analysis of insulators | No | No | Yes | Yes | Yes | No | | No | Yes | |
| Sample in vaccum required | Yes | Yes | Yes | No | No | Yes | | Yes | Yes | |

[a] EDS ranges to lower elements with windowless detector. PAS=photoacoustic single detection; WDS=wavelength dispersive spectrometry; EDS=energy dispersive spectrometry; SEM=scanning electron microscopy; STEM=scanning transmission microscopy; sp=spectrum; im=imaging mode; µIR=microscope FT-IR; µESCA=microESCA; n.a.=not applicable.

sample positioners is only feasible for extremely flat samples. As to the kind of materials analysed, LMMS requires stability in a vacuum of $10^{-6}$ torr but SSIMS needs a pressure of under $10^{-8}$ torr, which prevents the analysis of 'volatile' organic compounds. Charge build-up in dielectrical samples occurs in SSIMS, not in LMMS.

Quantification in MS requires adequate reference samples. For LMMS this means that not only the chemical composition but also the UV absorption, reflective and refractive properties of each microvolume must be comparable to ensure that the energy deposition and ion yield are similar. Hence, preparation of suitable reference materials often becomes the bottleneck. Biological sections or ambient aerosols are chemically and optically heterogeneous, so that quantification is not feasible. On the other hand, element diffusion in the wafers from semiconductor applications can be quantitatively studied.

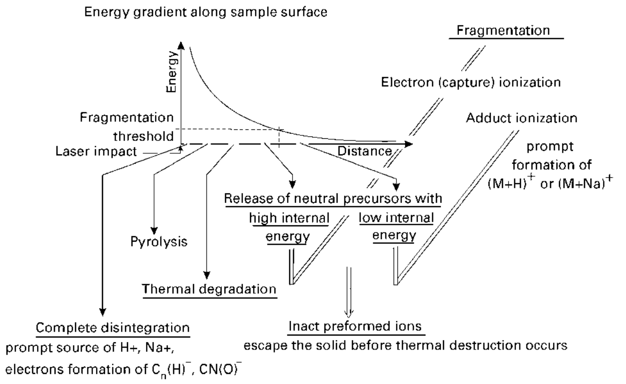



**Figure 3** Relationship between the energy regimes at the surface and the ion formation process according to the tentative model for DI in LMMS. Reprinted from Van Vaeck L, Struyf H, Van Roy W and Adams F (1994) *Mass Spectrometry Reviews*, **13:** 189–208 with permission of John Wiley and Sons.

## Diagnostic use of the mass spectra

### Model of ion formation

Deduction of the analyte composition from the signals requires that one must be able to extrapolate the mass spectra from a limited database and ideally predict the ions from a given analyte. Therefore, practical concepts about ion formation are mandatory. **Figures 3–5** survey a tentative model for DI in LMMS, applicable to both organic and inorganic compounds. Basically, the ionization event is described in terms of the three prime parameters in MS, namely local energy, pressure and time.

First of all, laser impact is assumed to create different local energy regimes in physically distinct regions within and around the irradiated spot. **Figure 3** depicts schematically the occurrence of atomization and destructive pyrolysis in hot spots, direct ejection of fragments and molecular aggregates from surrounding regions. Ultrafast thermal processes allow preservation of the molecular structure of even thermolabile analytes. The initial species will undergo fragmentation or subsequent ion-molecule interactions. Thermionic emission of e.g. alkali ions, yields the necessary flux of ions to form adducts with the initially released neutrals and ion-pairs. This occurs primarily after the laser pulse in the selvedge, i.e. the gas phase just above the sample.

Secondly, the local pressure must be considered. As shown in **Figure 4**, the released neutrals and ions give rise to a gradually expanding microcloud. Its initial density depends on the 'volatility' of the analyte, i.e. the number of species released at a given power density. The importance of ion-molecule interactions depends on the local pressure in this cloud, which determines the collision probability. In low-density regions, ions follow unimolecular behaviour and fragmentation only depends on the internal energy.

Finally, the timescale of the ionization is an essential feature of the approach since it must match the time window of detectable ions by a given mass analyser. The time evolution of the DI process is assumed to include two distinct phases. The schematic picture in **Figure 5** shows the that the ion-generation process comprises two consecutive contributions. An initial intense ion current is produced during the first 10-25 ns, compatible with detection in TOF LMMS. However, selvedge ionization may already start in that initial period but will reach its maximum later and finally continues for microseconds after the end

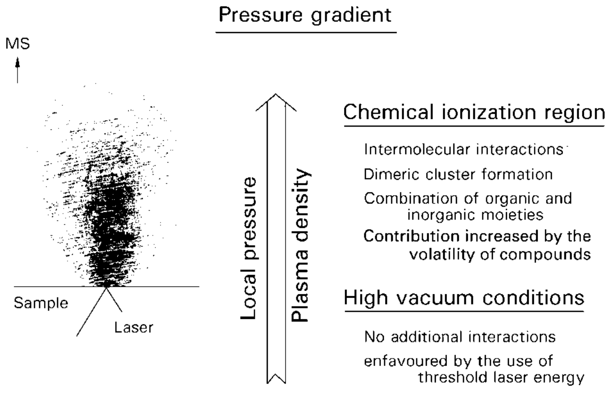



**Figure 4** Effect of the local pressure on the type of ions formed according to the tentative model for DI in LMMS. Reprinted from Van Vaeck L, Struyf H, Van Roy W and Adams F (1994) *Mass Spectrometry Reviews* **13**: 189–208 with permission of John Wiley and Sons.

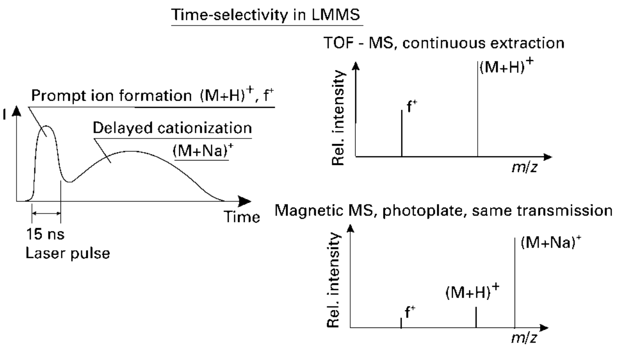



**Figure 5** Influence of the time domain of ion formation and of the analyser on the mass spectra recorded according to the tentative model for DI in LMMS. Reprinted from Van Vaeck L, Struyf H, Van Roy W and Adams F (1994) *Mass Spectrometry, Reviews*, **13**: 189–208 with permission of John Wiley and Sons.

of the laser pulse. This process may be small, but integrated over time, this contribution often prevails over the prompt DI. This rather slow DI component is not compatible with detection in TOF LMMS but is inherently included in the signal when magnetic mass spectrometers or FT MS ion traps are used.

This tentative DI model does not give an accurate description of all the physical processes involved, but up to now, it allows a consistent explanation of the formation of the ions detected. Also, it unifies the processes of organic and inorganic ion formation and allows rationalization of the differences between TOF and FT LMMS spectra as to the species detected and their relative intensities. Moreover, it allows linkage of LMMS to the older literature on the soft ionization of organic compounds by defocused lasers and magnetic analysers, where fragmentation was virtually absent.

## Inorganic speciation

**Figure 6** illustrates how FT LMMS signals can be used for analyte identification. The low m/z fragments and their adducts to the original molecule readily allow tentative identification of an unknown compound by deductive reasoning alone. This represents a major advantage in e.g. industrial materials-science applications, where many not well characterized reagents are possible causes of anomalous behaviour. Recording reference spectra of all possible candidates for the given anomaly could be time consuming if one cannot select deductively the most likely one.

The most demanding speciation task is the distinction between analytes with the same elements but in different ratios, e.g. sodium sulfate, sulfite and thiosulfate. The positive fragments $Na^+$, $Na_2O$, $Na_2O{\cdot}H^+$ and $Na_2O{\cdot}Na^+$ as well as the signals from $Na_2SO_3{\cdot}Na^+$ and $Na_2SO_4{\cdot}Na^+$ are observed by FT LMMS for all three analogues. However, only thiosulfate produces additional peaks for cationized $Na_2S$, which characterize the presence of the

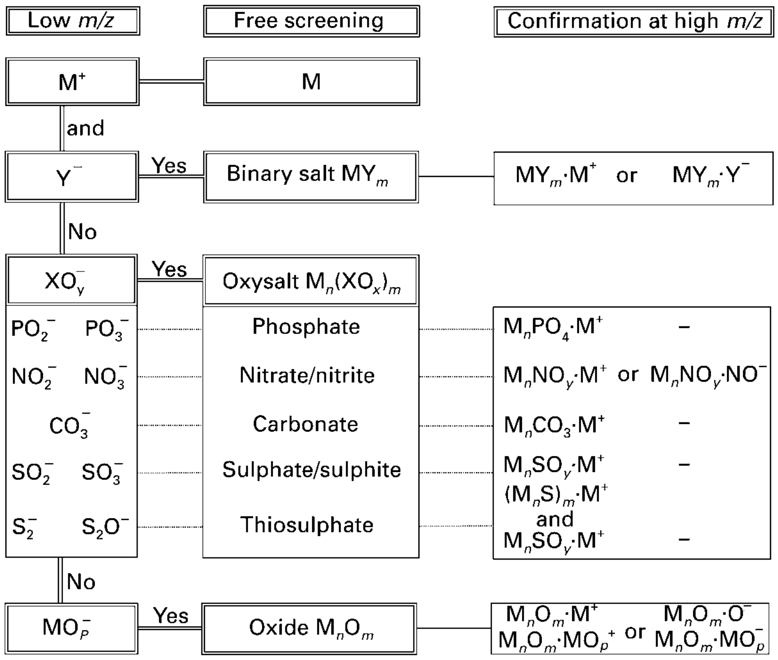



**Figure 6** Scheme for the deductive identification of the analyte molecular composition from mass spectra taken by FT LMMS. Reprinted from Struyf H, Van Vaeck L, Poels K and Van Grieken R (1998) *Journal of the American Society for Mass Spectrometry* **9**: 482–497 with permission of Elsevier Science.

sodium–sulfur bond in this analogue. Adduct ions of $Na_2S_2O_3$ are also seen at low intensity. Distinction can be based on the ratio of the ions $Na_2SO_3 \cdot Na^+$ to $Na_2SO_4 \cdot Na^+$, which is $0.82 \pm 0.13$ for sulfite and $0.30 \pm 0.06$ for sulfate. The production of cationized sulfate from sulfite and vice versa is rationalized by the relatively high-energy regime in the hot spots, where the stability of the formed ion prevails. Nevertheless, the softer regime of the periphery is assumed to produce the cationized analyte and thereby restores the logical connection with the original composition. The major signals in the negative mode are found at the nominal m/z values of 64 and 80 and are common for all three analogues. However, these signals are uniquely due to $SO_2^-$ and $SO_3^-$ for sulfite and sulfate, while thiosulfate produces additional contributions from $S_2^-$ and $S_2O^-$ also at m/z 64 and 80 respectively. The FT LMMS high mass resolution is needed to exploit such distinctive features.

### Identification of organic compounds

Deductive identification is even more important in organic microanalysis because of the numerous structures for each molecular weight. **Figure 7** shows the positive and negative TOF LMMS data of a microscopic residue, obtained from the eluate of a single peak from analytical HPLC. The nanogram amount of available material was not consumed significantly after microprobe analysis although it was insufficient for conventional MS. The latter required the combination of fractions from numerous elutions to yield the mass spectrum of miconazole, the parent drug. In contrast, TOF LMMS detected the protonated molecules by the cluster of isotope peaks around m/z 503. Subsequent loss of HCOOH yields the fragments around m/z 457, while the accompanying lower m/z signals are readily rationalized by the known ring formation and rearrangement mechanisms in organic MS. Complementary information comes with the negative ions. Specifically the fragment at m/z 153 serves to specify further the presence of the imidazole methacrylate function in the molecule. This example highlights two features of the LMMS technique, namely the minute material consumption and the deductive identification of labile compounds without thermal destruction as a result of the ultrafast heating rate of the solid.

### Comparison of mass spectra recorded by TOF LMMS and FT LMMS

Several characteristic differences are observed between the mass spectra, recorded by FT LMMS and TOF LMMS from the same analyte under the same experimental conditions with respect to laser wavelength, pulse duration and power density. While FT LMMS offers superior mass resolution and mass accuracy in comparison to TOF LMMS, the TOF effect in FT LMMS with an external source (see earlier sections on TOF LMMS and FT LMMS) means that only partial mass spectra in given m/z windows can be recorded (TOF LMMS always yields mass spectra that cover the entire m/z range). As well as these directly instrument-related data between FT LMMS and TOF LMMS, the mass spectral patterns still reflect additional distinctive features. Specifically, more intense signals from the adduct ions in comparison to the summed intensities of the fragments are seen in FT LMMS as compared to TOF LMMS. This is related to the increased sampling of the selvedge contribution to the initial ion population, created by the laser impact. Also the existence of the two contributions, i.e. direct ion emission from the solid and the selvedge recombination in the gas phase, could be observed by selectively tuning the ion source in FT LMMS so that only one of the contributions had initial energies within the 1 eV trapping energy of the cell. Note that this increased adduct ion detection in FT LMMS applies to both organic and inorganic compounds. For instance, cation and anion adducts of molecules and even dimers are generally detected in FT LMMS from virtually all salts such as $Na_2SO_4$, $Na_3PO_4$, oxides and binary salts, as opposed to TOF LMMS. Additionally, FT LMMS shows increased signal intensities specifically for fragment ions associated with the fragmentation of adduct ions with low internal energy. Such ions are observed to undergo the loss of small neutral molecules, while ions with high internal energy are subject to more drastic cleavages and rearrangement, which may reduce structural specificity. Finally, TOF LMMS is often handicapped in the analysis of high molecular mass and polar substances. These compounds are hard to desorb from the solid without the aid of matrix-assisted techniques, that require dissolution of the sample in a UV-absorbing matrix and, therefore, are incompatible with the direct local analysis of as-received solid samples. Especially for such strongly adsorbed compounds, the slow release of the analyte over a long time interval (up to several microseconds) occurs (continuing postlaser desorption). The adducts formed upon recombination with for instance $Na^+$ (also thermionically emitted for long periods after the laser pulse) no longer fulfil the TOF LMMS requirement of ion formation within 25 ns. As a result, these ions give rise to broad unresolved peaks or even to a continuous background. However, in FT LMMS, ions formed within a period of 50 μs to several hundreds of μs (depending on the m/z) after the laser pulse are still trapped in the FT LMMS cell

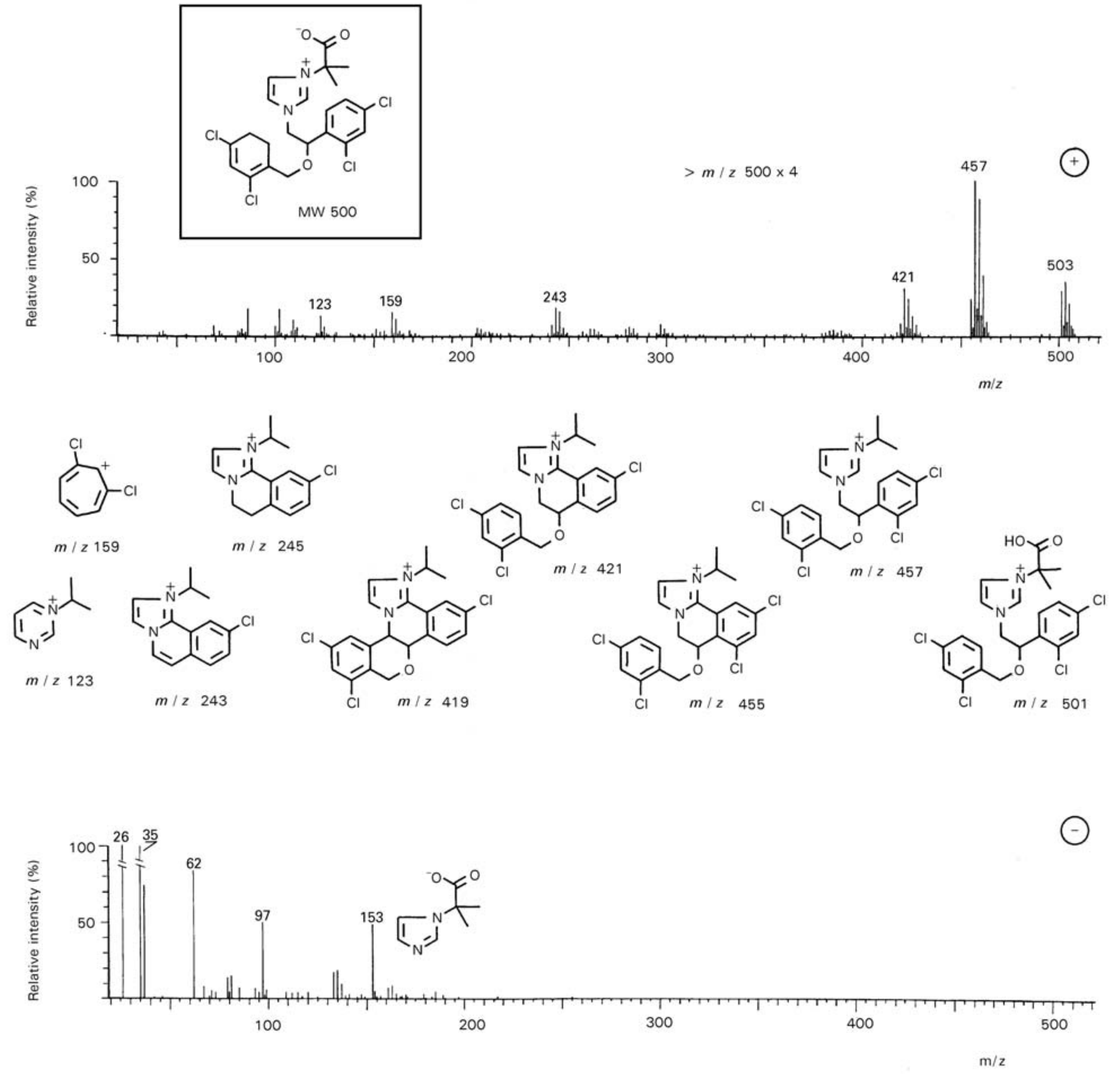



**Figure 7** Positive and negative mass spectrum recorded by TOF LMMS from the residue of the eluate of a single peak in analytical HPLC. Reprinted from Van Vaeck L and Lauwers W (1989), *Advances in Mass Spectrometry*, **11a**: 348–349 with permission of Heyden and Son.

together with the ions formed during the laser pulse. As a result, the rang e of compounds to which FT LMMS can be applied, significantly extends towards higher molecular mass and polarity as compared to TOF LMMS.

## Selected applications

Review of the literature reveals that LMMS is applied to a wide variety of problems in the field of bioscience, environmental chemistry and materials research. Complete coverage is not feasible in this contribution, which only aims to demonstrate the strengths of the method. Specifically, LMMS excels at yielding, within a relatively short period, qualitative information on the organic or inorganic local surface components from the most diverse samples, often with negligible sample preparation.

The transmission type TOF LMMS instruments especially suits biomedical section samples as commonly used for optical and electron microscopy.

The lateral resolution of 0.5 μm allows experiments at the subcellular level. Initial work has concentrated on the localization of elements in specific sites of the tissue. Typical examples involve the assessment of aluminium levels in the bones of chronic haemodialysis patients, the study of lead accumulation in kidneys and calcifications in intraperitoneal soft tissue as a result of chronic lead intoxication, and the detection of several heavy metals in the amalgam tattoos of the oral mucose membrane and human gingiva in direct contact with dental alloys. The relatively low lateral resolution, the lack of automated mapping and the almost impossible quantization strongly hinder the use of LMMS, especially in view of analytical electron microscopy (AEM) with X-ray analysis and the emerging possibilities of nuclear microscopy.

As a result, experiments have become gradually directed towards applications where the direct extraction of molecular information from the biological sector could be exploited. Interpretation of the mass spectra in TOF LMMS with low mass resolution may become problematic when the local microvolume constitutes of a complex multicomponent system. As a result, applications were especially successful when LMMS was used to characterize the micrometre-size microliths and foreign bodies, often found in histological sections. In this way, the spheroliths in the Bowman's membrane of patients suffering from primary atypical bandkeratopathy and the intrarenal microliths, formed after the administration of high doses of cyclosporin were proven to consist of hydroxyapatite. Implants often lead to long-term problems by dispersion of wear particles in tissues, leaching of specific compounds and sometimes subsequent chemical transformation. The pathogenesis of the aseptic loosening of joint prostheses was related to the presence of zirconium oxide in the granular foreign bodies of surrounding tissue. This compound was added to the bone cement to enhance the contrast in later X-ray radiographs.

Examples of characterization of organic molecules include the detection of deposits from an anti-leprosy drug in the spleen of treated mice. The study of biomedical sections is not the exclusive field of TOF LMMS, although the lower lateral resolution of FT LMMS limits the number of applications. However, the much better specificity from the superior mass resolution and mass accuracy pays off with increasing complexity of the local composition. FT LMMS is particularly appreciated because of the separation of isobaric ions in the tissue, such as $CaO^+$ and $Fe^+$ ions, the facile assignment of a signal as an 'organic' ion or an 'inorganic' cluster on the basis of accurate $m/z$ values, and the increased contribution of ions from the selvedge, which decreases the influence of the local laser power density on the mass spectra. Practical applications widen the range, initiated by TOF LMMS. For instance, the foreign bodies in the inflammed tissue around implants were identified as wear particles from a titanium knee-implant. Other experiments involved the verification of the local molecular composition deduced from relative element abundances in AEM. The apoptotic cell death in the tissue around vein grafts was associated with the presence of hydroxyapatite because of the intense Ca and P signals in AEM. FT LMMS clearly showed the erroneous nature of this conclusion since both elements did not belong to the same molecule. This clearly illustrates the importance of molecular information in microanalysis.

**Figure 8** illustrates the application of FT LMMS for the *in situ* identification of pigments in a lichen, *Haemmatomma ventosum*. The interest in such natural products relates to the understanding of their role in the interception of sunlight and their possible interaction in energy transfer. The optical micrograph shows the dark-red dish-like fruiting bodies or apothecia on the light coloured thallus. The combination of positive and negative ion mass spectra taken from the apothecia gives quite a detailed picture of the molecule. Specifically, the molecular mass is available from the potassium adduct and the numerous fragments serve to deduce structural features. The material is analysed as it is found in nature, without any prior sample preparation.

As to materials research, diverse problem-solving examples of LMMS are described. Typical examples involve the identification of local heterogeneities in poorly dispersed rubbers as one of the ingredients of the formulation, the tracing of the origin of occasional organic and inorganic contaminants at the surface of microelectronic devices and the study of segregation and formation of specific compounds in the joints of welded or heat-treated oxide-dispersion-strengthened alloys. Particularly interesting is the study of the dispersion of the magnetic elements inside the polyethylene terephthalate matrix at the surface of faulty floppy discs. Here the capability to detect both inorganic ions and organic structural fragments in the same spectrum is essential for trouble shooting.

A final example illustrates how LMMS can contribute to the fundamental understanding of processing of materials. Specific ally, aluminium strips and plates are produced by hot rolling the primary alloy ingots under high pressure and temperature. The lubricating emulsions contain numerous additives, whose composition is empirically optimized. Sometimes, given constituents adhere or interact with the

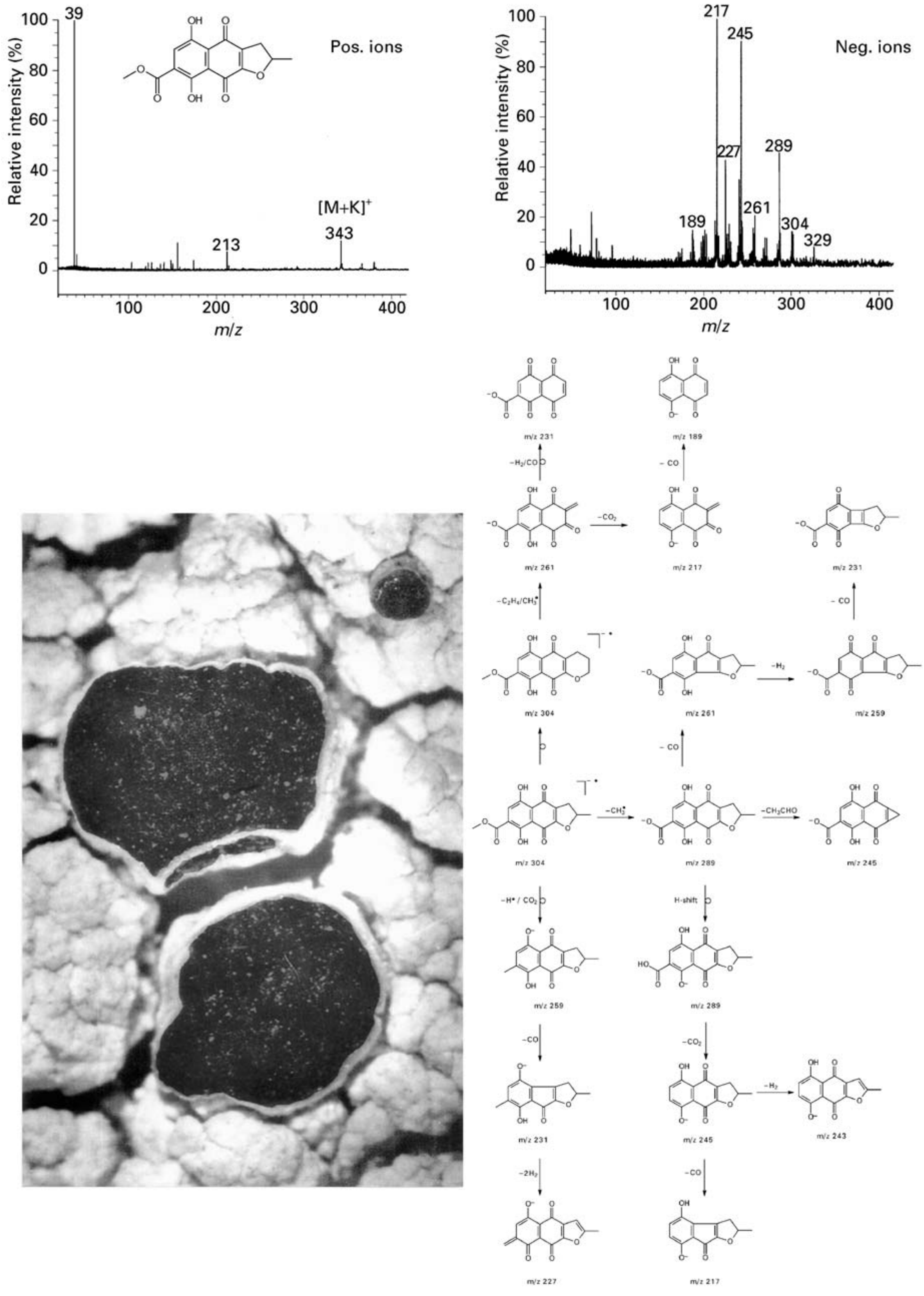



**Figure 8** *In situ* FT LMMS analysis of a pigment in the apothecia of a microlichen. Reprinted from Van Roy W, Matthey A and Van Vaeck L (1996) *Rapid Communications in Mass Spectrometry* **10**: 562–572 with permission of John Wiley and Sons.

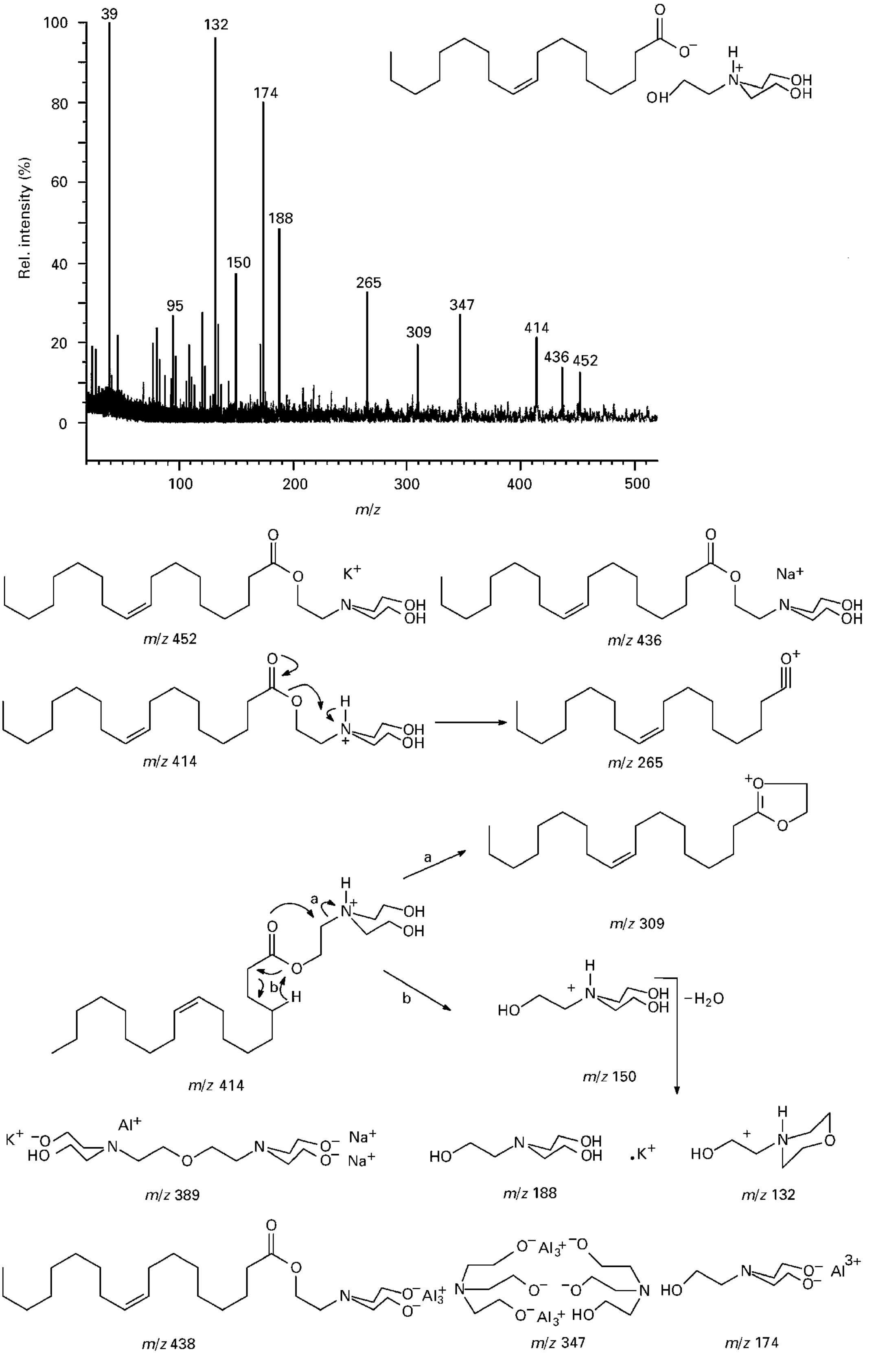



**Figure 9** FT LMMS analysis of an additive coating on aluminium. Reprinted from Poels K, Van Vaeck L, Van Espen P, Terryn H and Adams F (1996) *Rapid Communications in Mass Spectrometry* **10**: 1351–1360 with permission of John Wiley and Sons.

surface and cause optical defects upon subsequent coating or surface treatment. FT LMMS was used to study the possible additive–metal interactions. The data in **Figure 9** for triethanolamine (TEA) oleate on aluminium show the normally expected simple adduct and fragments, as well as particularly interesting signals due to ions, which include multiply charged anions with $Al^{3+}$. Since only singly charged aluminium is formed in laser microbeam DI, these $Al^{3+}$-containing ions must exist as such in the solid and undergo direct ejection. This implies that TEA oleate is likely to bind effectively to the metal, while this additive is often considered as a chemically nonaggressive surfactant in the lubricating emulsions.

To conclude, it should be mentioned that one of the major breakthroughs in the field of organic MS emerged from the research on the initial TOF LMMS instruments. Indeed, the now booming field of matrix assisted laser desorption ionization (MALDI) for the characterization of high molecular mass compounds up to 230 kDa found its roots in the laboratory of Prof. Hillenkamp, which was one of the driving forces behind the development of the initial TOF LMMS instruments. It shows the ingenuity of chemists in exploiting the powerful combination of laser ionization and MS. As to local analysis, the second generation of LMMS, FT LMMS seems to represent a major step in the search for a versatile microprobe, enabling us to characterize the molecular composition of organic and inorganic compounds at the surface of almost any type of solid, electrically conducting or not, with minimal sample preparation.

## List of symbols

$E_{kin}$ = initial kinetic energy of ions.

*See also:* **FT-Raman Spectroscopy, Applications; Time of Flight Mass Spectrometers.**

## Further reading

Eeckhaoudt S, Van Vaeck L, Gijbels R and Van Grieken R (1994) Laser microprobe mass spectrometry in biology and biomedicine. *Scanning Electron Microscopy supplement* **8**: 335–358.

Poels K, Van Vaeck L, Van Espen P, Terryn H and Adams F (1996) Feasibility of Fourier transform laser microprobe mass spectrometry for the analysis of lubricating emulsions on rolled aluminium. *Rapid Communications in Mass Spectrometry* **10**: 1351–1360.

Struyf H, Van Roy W, Van Vaeck L, Van Grieken R, Gijbels R and Caravatti P (1993) A new laser microprobe Fourier transform mass spectrometer with external ion source for organic and inorganic microanalysis. *Analytica Chimica Acta* **283**: 139–151.

Struyf H, Van Vaeck L, Poels K and Van Grieken R (1998) Fourier transform laser microprobe mass spectromery for the molecular identification of inorganic compounds. *Journal of the American Society for Mass Spectrometry* **9**: 482–497.

Van Roy W, Matthey A and Van Vaeck L (1996) *In situ* analysis of lichen pigments by Fourier transform laser microprobe mass spectrometry with external ion source. *Rapid Communications in Mass Spectrometry* **10**: 562–572.

Van Vaeck L, Gijbels R and Lauwers W (1989) Laser microprobe mass spectrometry: an alternative for structural characterisation of polar and thermolabile organic compounds. In: Longiévalle P (ed) *Advances in Mass Spectrometry*, Vol. 11A, pp. 348–349. London: Heyden.

Van Vaeck L, Struyf H, Van Roy W and Adams F (1994) Organic and inorganic analysis with laser microprobe mass spectrometry. Part 1: Instrumentation and methodology. *Mass Spectrometry Reviews* **13**: 189–209.

Van Vaeck L, Struyf H, Van Roy W and Adams F (1994) Organic and inorganic analysis with laser microprobe mass spectrometry. Part 2: Applications. *Mass Spectrometry Reviews* **13**: 209–232.

Verbueken A, Bruynseels F, Van Grieken R and Adams F (1988) Laser Microprobe Mass Spectrometry. In: Adams F, Gijbels R and Van Grieken R (eds) *Inorganic Mass Spectrometry*, pp. 173–256. New York: Wiley.

Vertes A, Gijbels R and Adams F (eds) (1993) *Laser ionisation mass analysis*. Chemical Analysis Series, Vol. 124. New York: Wiley.

# Laser Spectroscopy Theory

**David L Andrews**, University of East Anglia, Norwich, UK

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
| --- |
| **Theory** |

The theory underlying electronic spectroscopy with lasers is essentially the theory of visible or ultraviolet photon interactions. Any spectroscopic technique based on laser instrumentation might in principle be studied with some other kind of light source, but usually at the expense of data quality in terms of signal-to-noise or resolution. The distinctive features that arise with the deployment of laser light in electronic spectroscopy are principally those that relate to or exploit the qualities of the electric field produced by the laser beam. The customarily high level of monochromaticity affords the means to obtain high-resolution data, and high field strengths offer scope for the study of multiphoton processes. Another widely vaunted attribute of the laser, the coherence of its output, is only indirectly relevant to spectroscopic applications. In the so-called 'coherence spectroscopies' such as self-induced transparency and photon echo, laser coherence is significant only in enabling short and highly intense pulses to be created.

## Electric field and intensity

Laser electronic spectroscopy is primarily based on coupling between the electron clouds of individual ions, atoms, chromophores or molecules of the sample with the electric field of the impinging laser radiation. The coupling is usually of dipolar character, and its detailed involvement in photonic interactions is to be discussed below. First, it is expedient to characterize the electric field, and to relate it to the laser intensity – as the latter is more directly amenable to experimental measurement. The electric field produced by a coherent, parallel, radially symmetric and plane polarized laser beam propagating in the $z$-direction, with a wavelength $\lambda$ and a frequency $\nu$, oscillates sinusoidally in time and space and has an amplitude $A$ which depends on the radial displacement $r$ from the beam centre:

$$E(z,r,t) = A(r)\cos(kz - \omega t + \phi) \qquad [1]$$

Here $k = 2\pi/\lambda$, the 'circular frequency' $\omega = 2\pi\nu$ and $\phi$ is the phase. The vector character of the electric field is in general determined by the state of polarization, as discussed later: in the case of plane polarization it is directed along a unit vector $\mathbf{i}$ perpendicular to the $z$-direction of propagation. In cases where the line width of the radiation is significant the amplitude additionally has a frequency dependence and, particularly in the case of pulsed radiation, it also varies with the time $t$.

The beam irradiance $I$ (the power per element of cross sectional area), to which the electric and magnetic fields of the radiation contribute equally, is in general given by

$$I(r,\nu,t) = \tfrac{1}{2}c\varepsilon_0 A^2(r,\nu,t) \qquad [2]$$

The beam thus has radial, frequency and temporal intensity profiles. The radial profile is determined by the mode structure, reflecting the pattern of standing waves sustained within the laser cavity. In the simplest or ideal uniphase case, radiation tracks back and forth in an exactly axial fashion between the end-mirrors. With no hard edge, the beam is then described by an essentially Gaussian distribution, with a beam diameter $2w$ defined as the transverse distance between points at which the irradiance drops to $1/e^2$ (13.5%) of its central value, $I(0)$:

$$I(r) = I(0)\exp\left[-2(r/w)^2\right] \qquad [3]$$

Accordingly, the electric field amplitude falls away as $\exp[-(r/w)^2]$. For a beam of instantaneous power $W$, the beam centre irradiance $I(0)$ is given by

$$I(0) = 2W/\pi w^2 \qquad [4]$$

Fundamental considerations show that any laser beam can only be focused down to a limiting width whose diameter is of the same order as the wavelength – this is the diffraction limit. The focused beam diameter $2w_0$ is generally determined by the relation

$$w_0 = 2M^2\lambda/\pi\theta \qquad [5]$$

where $\lambda$ is the wavelength and $\theta$ the angle of convergence of the focused beam. The quantity $M^2$ is a measure of beam quality and has a diffraction-limited value of unity for a fundamental Gaussian-mode beam as represented by Equation [3].

The frequency profile of the laser beam is usually characterized by its full-width at half-maximum (FWHM), $\Delta\nu$. The single parameter that most effectively characterizes the degree of monochromaticity is then the quality factor $Q$, defined as the ratio of the laser emission frequency to its line width. It is also expressible as the coherence length $l_c$ (the distance over which photons remain effectively in phase) divided by the wavelength:

$$Q = \nu/\Delta\nu = l_c/\lambda \qquad [6]$$

In absorption-based laser spectroscopy this parameter represents an upper limit on the achievable resolution, which in practice may be reduced by other features of the instrumentation.

The issue of temporal profile is one that primarily relates to pulsed lasers, whose high powers result from their delivery of energy over a short period of time rather than continuously. For such devices the detailed temporal profile of the intensity $I(t)$ is determined by the means of pulsed operation. Individual pulses from giant pulse lasers seldom have the smoothly symmetric shape needed to follow any simple analytic form – and although the more closely symmetric pulses from a mode-locked train are often characterized as having Gaussian shape, that more than anything reflects their characterization by averaging over a large number of pulses. The sech$^2$ intensity profile widely regarded as ideal is, however, achieved in pulses from many of the newer generation of ultrafast lasers.

Even continuous-wave laser light generally exhibits temporal fluctuations as a result of imperfect coherence, and the probability of finding a single photon in the volume of space $V$ occupied by the species of spectroscopic interest generally approximates to a Poisson distribution. With a mean number of photons $M$ given by

$$M = \frac{IV\lambda}{hc^2} \qquad [7]$$

the probability $P_N$ of finding $N$ photons is

$$P_N = \frac{M^N e^{-M}}{N!} \qquad [8]$$

## Interaction of light and matter

In formulating the general theory it is helpful first to establish the dipolar response of each ion, atom, chromophore or molecule $\xi$ to the oscillating electric field of the radiation. This is associated with an interaction operator:

$$V(\xi) = -\mu(\xi) \cdot \boldsymbol{E}(\xi) \qquad [9]$$

where $\mu(\xi)$ is the appropriate electric dipole operator and $\boldsymbol{E}(\xi)$ the electric field vector at the position in space occupied by the species $\xi$. Electric dipole coupling accounts for the vast majority of the observations in electronic spectroscopy, though other kinds of multipolar interaction (such as electric quadrupole or magnetic dipole) can permit otherwise 'forbidden' transitions to occur, through a coupling which is typically weaker by a factor of $10^2$ or $10^3$.

Even with the high laser intensities employed for the study of multiphoton processes, the electric fields they produce are seldom comparable to intramolecular coulombic fields. As such, the interaction operator [9] can be treated as a perturbation on the molecular states. The result of these perturbations is to modify the form of the Schrödinger equation so that its usual eigenfunctions no longer represent stationary states, and hence transitions occur. The quantum amplitude for a transition between a given initial state $i$ and a final state $f$ is given by the following result from time-dependent perturbation theory, expressed in Dirac notation:

$$M_{fi}(t,t_0) = \langle f|U_I(t,t_0)|i\rangle = \left(\frac{1}{i\hbar}\right)^n \langle f| \int_{t_0}^{t} dt_1 \times \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n \tilde{V}(t_1)\tilde{V}(t_2)\cdots\tilde{V}(t_n)|i\rangle \qquad [10]$$

where

$$\tilde{V}(t_i) = \exp\left(\frac{iH_0(t_i - t_0)}{\hbar}\right) V \exp\left(\frac{-iH_0(t_i - t_0)}{\hbar}\right) \qquad [11]$$

and $H_0$ is the normal, unperturbed, Schrödinger operator for the molecule. Except for measurements with the most extreme time resolution, evaluation of the quantum amplitude given in Equation [10] leads

to an essentially time-independent result and yields a corresponding transition rate $\Gamma$ given by Fermi's 'Golden Rule':

$$\Gamma = \frac{2\pi}{\hbar}\left|M_{fi}\right|^2 \rho_f \qquad [12]$$

Here $\rho_f$ is the number of energy levels per unit energy interval, the so-called 'density of states'. Though often overlooked, the size of this latter factor contributes significantly to the ultrafast rates of transitions observed to occur in many large molecules, as a result of their tightly packed quasi-continua of vibronic levels. With increasing $n$, successive orders of the perturbation series [10] represent a progressively diminishing coupling, with a scale corresponding to the ratio of the radiative electric field in relation to intramolecular coulombic fields. Together, Equations [1], [2] and [9]–[12] correspondingly show that each additional photon interaction typically reduces the rate by a factor of the order $(I/J)$, where $J$ is a level of irradiance which would produce ionization or dissociation – its value is generally in the region $10^{18\pm4}$ W m$^{-2}$. This figure is substantially undershot in most laser excitation experiments, and as a result the higher orders of photon interaction demand increasingly sensitive detection.

## Single-photon absorption

For a process of single-photon absorption, the quantum amplitude is simply given by

$$M_{fi} = -\mu^{fi} \cdot \boldsymbol{E} \qquad [13]$$

using the convenient shorthand for the vector transition dipole moment, $\mu^{fi} = \langle f \,|\, \mu \,|\, i \rangle$. The rate $\Gamma$ as given by Equation [12] depends quadratically on the electric field, and hence through Equations [1] and [2] exhibits a linear dependence on the laser intensity.

It follows that, since absorption produces a diminution of the laser intensity on passage through the absorbing sample, the rate of intensity loss is proportional to its instantaneous value, $-\mathrm{d}I(t)/\mathrm{d}t \propto I(t)$. In the case of pulsed irradiation, if the amount of absorption per pulse is small and decay processes from the excited state occur on a timescale that is slow compared with the pulse duration, the net absorption per pulse is proportional to the pulse energy. More generally, as the distance $z$ the radiation travels through the sample is proportional to time $t$ through the speed of light, we also have $-\mathrm{d}I(z)/\mathrm{d}z \propto I(z)$. Thus the amount of absorption depends on the path length through Beer's Law:

$$I(z) = I(0)\exp(-\alpha C z) \qquad [14]$$

where $\alpha$ is the absorption coefficient and $C$ the concentration of (ground state) absorbing species. The former is calculable from secondary effects pursuant on the absorption, or more directly from the relative attenuation of the excitation beam. Weak absorption leads to an intensity loss which has an almost directly linear dependence on distance, as follows from Taylor series expansion of the right-hand side of Equation [14].

Laser excitation here offers few distinctive features at the molecular level, except that at high levels of intensity the increased flux can lead to saturation. Then, a significantly high proportion of the sample molecules undergoes transition to an excited state and, as $C$ varies through depletion of the ground state population, departures from Beer's law arise. This is a phenomenon that is exploited for analytical purposes with pulsed radiation, in concentration-modulated absorption spectroscopy.

The selection rules associated with single-photon absorption stem from its dependence on the transition dipole moment, which must be non-zero for absorption to occur. To satisfy this fundamental requirement, the direct product symmetry species of the initial and final wavefunctions must contain the symmetry species of the dipole operator – the latter transforming like the translation vectors $x$, $y$ and $z$. Probably the most familiar aspect is the Laporte selection rule for centrosymmetric molecules, which allows transitions only between states of opposite parity, gerade ($g$) $\longleftrightarrow$ ungerade ($u$).

For single-photon absorption in isotropic media, there is essentially no dependence on beam polarization. Such dependence as does exist arises only for optically active (chiral) compounds, and is associated with quantum interference between electric dipole and magnetic dipole (or electric quadrupole) interactions. These weak effects produce the characteristic polarization dependence that is manifest in the phenomena of circular dichroism and optical rotation.

## Multiphoton absorption

For processes involving the absorption of $n$ photons from a single laser beam, energy conservation dictates

$$nh\nu = E_f - E_i \qquad [15]$$

Multiphoton absorption of this kind is essentially a concerted rather than a multistep process, as there is no physically identifiable intermediate stage between the absorption of successive photons. It is, therefore, necessary for the absorbing molecule to be intercepted by the necessary number of photons almost simultaneously, a condition which can seldom be satisfied other than by the use of pulsed laser radiation.

Determination of the rate of $n$-photon absorption from the Fermi rule [12] invokes the quantum amplitude for the process as given by Equations [9]–[11]. In the case of two-photon absorption, for example, we obtain a quadratic coupling with the impinging electric field:

$$M_{fi} = -S^{\mathrm{fi}}_{\alpha\beta}(\omega) : E_\alpha E_\beta \qquad [16]$$

a tensor counterpart to the scalar product of the single-photon result shown in Equation [13], in which $\alpha$ and $\beta$ here represent the Cartesian indices $x$, $y$ and $z$. Equation [16] is written in this concise form (with obvious extension to higher orders) through adoption of the Einstein convention for summation over repeated tensor indices, i.e. both $\alpha$ and $\beta$ in the above expression are implicitly summed over $x$, $y$ and $z$. The molecular response exhibited in Equation [16] is cast in terms of a two-photon tensor whose explicit structure is as follows

$$S^{\mathrm{fi}}_{\alpha\beta}(\omega) = \sum_r \left[ \frac{\langle f|\mu_\alpha|r\rangle\langle r|\mu_\beta|i\rangle}{E_r - E_i - \hbar\omega - \mathrm{i}\Gamma_r} \right] \qquad [17]$$

in which there is summation over a set of contributions from each molecular state $r$, of energy $E_r$ and FWHM Lorentzian line width $2\Gamma_r$. Again the extension to higher orders follows, each additional photon interaction introducing one further transition dipole in the numerator and also an additional energy denominator term. For all but the simplest cases the exact form of these is best determined through quantum electrodynamical methods using Feynman time-ordered diagrams.

In general the $n$-photon quantum amplitude exhibits an $n$th power dependence on the electric field and accordingly the rate given by Equation [12] depends on the $n$th power of the laser irradiance $I$. The lack of coherence which is manifest as intensity fluctuations in the beam generally means that the time-average value of $I^n$ differs from $\bar{I}^n$, where $\bar{I}$ is the mean irradiance. Consequently the result is usually cast instead in terms of $g^{(n)}\bar{I}^n$, where $g^{(n)}$ is the degree of $n$th order coherence – having the value of unity for a perfect Poissonian source.

In the case of pulsed radiation the extent of absorption depends on the integral $\int I^n(t)\mathrm{d}t$ and so involves not only the energy of the pulse but also the temporal profile of the latter. Equally, the radial profile of the beam has a bearing on the extent of absorption. In particular, the nonlinear dependence on intensity means that if the laser beam is focused within a medium exhibiting multiphoton absorption, most of the absorption occurs at the focus. For example a Gaussian beam focused from a beam waist $w$ to the diffraction-limited value $w_0$ as given by Equation [5] generates across its focus a net absorption which is $(w/w_0)^{2n}$ times larger than across the unfocused beam. This principle is exploited in imaging techniques based on the detection of fluorescence resulting from multiphoton absorption.

Through the nonlinear intensity dependence, Beer's law condition totally fails to be satisfied in processes involving the absorption of two or more photons, and results cannot be given in terms of conventional absorption coefficients. Again, for two-photon absorption the counterpart to Equation [14] takes the form

$$I(z) = I(0)/(1 + \beta C z) \qquad [18]$$

where $\beta$ determines the strength of the two-photon transition. However, the weakness of the absorption generally means that the leading terms of a Taylor series for the right-hand side of Equation [18] provide a very good approximation for the beam attenuation $\Delta I = I(0) - I(z)$, which then displays the same essentially linear dependence on distance as weak single-photon absorption.

It is evident from the form of Equation [17] that the two-photon tensor and hence the extent of two-photon absorption, is significantly enhanced if the molecule possesses an electronic excited state of an energy close to $E_i + \hbar\omega$, for then, in the sum over $r$, the term that state contributes has an absolute minimum value for the denominator. This situation, typical of multiphoton processes, is known as resonance enhancement, and its physical basis can be understood in terms of the time–energy uncertainty principle:

$$\Delta E \Delta t \geq \tfrac{1}{4}\hbar \qquad [19]$$

In frequency regions well removed from resonance, absorption of the first laser photon leads the molecule into a state with a large $\Delta E$, i.e. one that is energy non-conserving. Accordingly, the molecule can exist in such a state only if further

absorption restores energy conservation within a correspondingly short time. However, if the first photon is near resonance with a real molecular state, the intermediate state has a small $\Delta E$ and can persist for much longer. This greatly increases the probability that the next photon will arrive within the necessary window. In the limit where exact resonance occurs and $\Delta E$ is zero there is no longer any temporal constraint save for the decay lifetime of the resonant state, as that state can be physically populated by single-photon absorption. Then the two-photon process can be completed in a second stage, without the need to fulfil the exacting conditions for both photons to arrive at the same instant. Under resonance conditions the power law for the intensity dependence is also in general modified to reflect the greater significance of whichever stage is the rate-determining step.

Each photon interaction in a multiphoton process governed by dipole coupling is subject to the Laporte selection rules. However, the product selection rules depart substantially from the normal results; for example, if an even number of photons is involved, as in Raman scattering, two-photon absorption and four-wave mixing, then the selection rules g $\longleftrightarrow$ g and u $\longleftrightarrow$ u apply. In general, for a given transition to occur it is necessary that the molecular response tensor possesses at least some non-zero elements. The criterion for this rule to be satisfied is that the product of the irreducible representations of the molecular initial and final states must be spanned by one or more components of the tensor. The representation of the (electric dipole) molecular response tensor $T^{(n)}$, for a process of $n$-photon absorption from a single beam, has the following general decomposition into irreducible parts or weights $n$, $(n-2)\ldots$;

$$\mathcal{D}(T^{(n)}) = \mathcal{D}^{(n\varphi)} \oplus \mathcal{D}^{((n-2)\varphi)} \oplus \cdots \qquad [20]$$

where for even $n$ the parity $\varphi = +1$ and for odd $\varphi = -1$, and the series terminates with $D^{(0\varphi)}$ or $D^{(1\varphi)}$ respectively. For molecules of reasonably high symmetry, the operation of the multiphoton selection rule leads to far more states being accessible through multiphoton processes than through conventional single-photon absorption.

## Polarization effects

Laser beam polarization plays an obvious role in the orientational effects displayed in photoabsorption by anisotropic media such as crystals, liquid crystals and surfaces. However, strong pumping with a plane polarized laser beam can induce anisotropy in an originally isotropic sample. This results from the fact that the probability of single-photon depends on $\cos^2\theta$, where $\theta$ is the angle between the relevant molecular transition moments and the photon electric field vector, as in Equation [13]. High intensity pulsed laser radiation can thus create a preferentially oriented population of excited molecules, a process termed photoselection. In many cases rotational relaxation nonetheless destroys this anisotropy within a matter of picoseconds after the inducing laser pulse.

The electric field of laser light need not oscillate in a single plane, and often optics are employed to produce other polarizations with a degree of circularity. Circular polarizations are important in forms of laser spectroscopy which exploit angular momentum selection rules, because the photons carry unit quanta of angular momentum. With chiral substances, a small degree of sensitivity to the handedness of the radiation is also manifest in the circular differential response. For two-photon and higher-order processes, however, even the spectra of reasonably symmetrical molecules display a marked dependence on polarization.

For generality the polarization state of the laser beam can be expressed in a form that can accommodate arbitrary plane, circular or elliptical states. Instead of the $i$ which determines the direction of the electric field in Equation [1], we then have a unit polarization vector given by

$$\boldsymbol{e} = \mathrm{e}^{\mathrm{i}\phi}\cos\,\sigma\hat{\boldsymbol{X}} + \mathrm{e}^{\mathrm{i}\zeta}\sin\,\sigma\hat{\boldsymbol{Y}} \qquad [21]$$

where

$$(\sigma,\phi,\zeta) \mapsto \begin{cases} (\pi/4, 0, \pi/2) & \text{R-circular polarization} \\ (\pi/4, 0, -\pi/2) & \text{L-circular polarization} \\ (\pi/2, 0, 0) & \text{p-polarization} \\ (0,0,0) & \text{s-polarization} \end{cases} \qquad [22]$$

where the latter mutually orthogonal plane polarizations are given the labels that would usually be employed in connection with irradiation of a surface with normal vector $Z$.

## List of symbols

$A$ = amplitude of a plane polarized laser beam; $c$ = speed of light; $C$ = ground state concentration;

$E$ = electric field; $E_r$ = energy of level $r$; $g^{(n)}$ = degree of $n$th order coherence; $I$ = beam irradiance; $J$ = irradiance level producing ionization or dissociation; $l_c$ = coherence length; $M$ = mean number of photons; $M_{fi}$ = quantum probability amplitude; $Q$ = quality factor; $r$ = radial displacement from the beam centre; $t$ = time; $T^{(n)}$ = molecular response tensor; $V(\xi)$ = interaction operator for species $\xi$; $V$ = volume; $w$ = beam radius; $w_o$ = focused beam radius; $W$ = instantaneous beam power; $\alpha$ = absorption coefficient; $\varepsilon_o$ = vacuum permittivity; $\Gamma$ = transition rate; $\theta$ = angle of convergence of the focused beam; $\mu^{fi}$ = transition dipole moment; $\lambda$ = wavelength; $\mu(\xi)$ = electric dipole operator of species $\xi$; $E(\xi)$ = electric field vector at the position occupied by species $\xi$; $\rho_f$ = number of energy levels per unit energy interval; $\nu$ = frequency; $\omega$ = circular frequency; $\hbar$ = Planck's constant/$2\pi$.

*See also:* **Chemical Reactions Studied By Electronic Spectroscopy; Environmental Applications of Electronic Spectroscopy; Laser Applications in Electronic Spectroscopy; Multiphoton Excitation in Mass Spectrometry; Multiphoton Spectroscopy, Applications; Nonlinear Optical Properties.**

## Further reading

Andrews DL and Demidov AA (1995) *An Introduction to Laser Spectroscopy*. New York: Plenum Press.

Craig DP and Thirunamachandran T (1984) *Molecular Quantum Electrodynamics: An Introduction to Radiation–Molecule Interactions*. London: Academic Press.

Demtröder W (1996) *Laser Spectroscopy: Basic Concepts and Instrumentation*, 2nd edn. Berlin: Springer-Verlag.

Kettle SFAK (1985) *Symmetry and Structure*. New York: Wiley.

Siegman AE (1986) *Lasers*. Oxford: Oxford University Press.

# Lead NMR, Applications

*See* **Heteronuclear NMR Applications (Ge, Sn, Pb).**

# Light Sources and Optics

**R Magnusson**, The University of Texas at Arlington, TX, USA

Copyright © 1999 Academic Press

**FUNDAMENTALS OF SPECTROSCOPY**
**Methods & Instrumentation**

## Optics

### Light

Light emanates from sources with a wide variety of physical embodiments. Examples include the Sun, a sodium arc lamp, an incandescent light bulb, and a helium- neon laser. The physical processes producing light emission appear to be distinctly different for each type of source. On a fundamental level, however, light is generated as a result of acceleration of electric charges. Thus, a synchrotron generates light by accelerating free electrons in a circular storage ring. An incandescent light bulb radiates as a result of electric-current-induced electronic charge vibrations and subsequent emission of radiation. A laser operates under atomic transitions involving bound charges acting as miniature, light-radiating electric dipole antennas.

### Electromagnetics

Light propagates as an electromagnetic wave. Maxwell's equations govern the propagation of light and other electromagnetic waves in free space, in material media, at interfaces, and through apertures. Visible light has wavelengths, $\lambda$, expressed in nm or µm ($10^3$ nm = 1 µm = $10^{-6}$ m), or frequency, $\nu$, in Hz

or $s^{-1}$, detectable by the human eye typically falling within the $380 < \lambda < 780$ nm range. The electromagnetic spectrum includes the ultraviolet (UV) region often defined to be $10 < \lambda < 380$ nm and the infrared (IR) region 780 nm $< \lambda <$ 1 mm, also of great importance in optics and spectroscopy.

Maxwell's equations yield a wave equation that predicts the propagation of electromagnetic radiation. The most elementary solution is the plane wave expressible as a function of space (x, y, z) and time (t) as

$$e(z,t) = \hat{x}e_o \cos[\omega t - kz + \phi] \qquad [1]$$

This wave, propagating in the +z direction, is monochromatic with angular frequency $\omega = 2\pi\nu$, wavenumber $k = 2\pi/\lambda$, and with $\lambda\nu = c$ where c is the speed of light in the medium in which the wave propagates. An arbitrary phase factor $\phi$ is included. This wave, expressed in terms of the electric field vector $e(z,t)$, has amplitude $e_o$ and is linearly polarized along the x direction, with $\hat{x}$ denoting a unit vector. An electromagnetic plane wave is a transverse wave without any field component along the direction of propagation. With $e(z,t)$ representing the electric field (V $m^{-1}$), Maxwell's equations require that an orthogonal magnetic field (A $m^{-1}$) vector $h(z,t)$ also be associated with the wave as shown in **Figure 1**. Since $e_o$ = constant, the plane wave is transversely infinite and thus carries infinite energy; it is further infinite in time and exists everywhere in space. In spite of these idealizations, the plane-wave model is exceedingly successful in predicting a large range of optical phenomena such as reflection/refraction at metallic and dielectric interfaces, diffraction by spectroscopic gratings, wave propagation in nonlinear and anisotropic media, etc. A practical, transversely finite, laser beam may be 1 mm in diameter which is $2000\lambda$ for $\lambda = 500$ nm; it is well approximated as a plane wave since it is wide on the scale of $\lambda$.

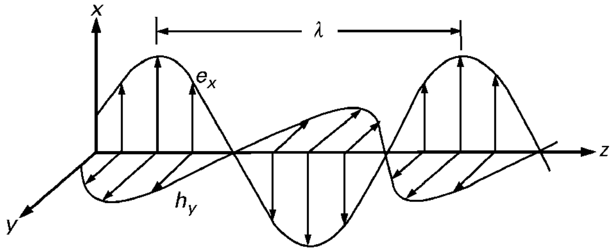



**Figure 1** A snapshot of an electromagnetic plane wave with wavelength $\lambda$ illustrating its spatial characteristics. The electric-field component $e_x$ and the magnetic-field component $h_y$ oscillate in phase but orthogonal to each other and to the direction of propagation. This wave is said to be linearly polarized along the *x*-direction.

## Polarization

The vectorial nature of the electromagnetic fields representing light implies polarization, a fundamental property of light. The particular orientation of the *e*-vector of a plane wave incident on an interface has a profound effect on the detailed interaction that takes place and influences, for example, the amount of light that is reflected or transmitted. If the electric field vector is confined to oscillate in a plane (e.g. the x- z plane in **Figure 1**), the wave is said to be linearly polarized. If the tip of the electric field vector traverses an elliptical path around the direction of propagation, the state of polarization is elliptical. Circular polarization and linear polarization are special cases of the general state of elliptical polarization.

The influence of the state of polarization is illustrated in **Figure 2** where a polarized plane wave in air (refractive index, $n_1 = 1$) is incident on a piece of glass ($n_2 = 1.5$). The index of refraction is defined as $n = c_0/c$, where $c_0$ is the speed of light in vacuum. For s-polarized light, with the e-field oscillating orthogonal to the plane of incidence as shown, the reflectance ($R_s$, defined as the ratio of optical power density (W $cm^{-2}$) reflected to that of the input wave) increases monotonically as the angle of incidence, $\theta_i$, is increased. For a p-polarized incident wave, the reflectance, $R_p$, differs markedly, even exhibiting a reflection zero when $\theta_i = \theta_B$, which defines the Brewster angle given by $\theta_B = \tan^{-1}[n_2/n_1]$. The angle of incidence $\theta_i$ is related to the angle of refraction $\theta_t$ by Snell's law $n_1 \sin\theta_i = n_2 \sin\theta_t$ which holds for both polarization states. Additionally, the angle of reflection is $\theta_r = \theta_i$. If the wave traverses a rare-to-dense interface ($n_1 < n_2$) then $\theta_t < \theta_i$. Conversely, if $n_1 > n_2$, $\theta_t > \theta_i$. In particular, for $\theta_i$ such that $\theta_t = 90°$, total internal reflection occurs and $R_s = R_p = 1$ for all $\theta_i > \theta_c$, which is the critical angle given by $\theta_c = \sin^{-1}[n_2/n_1]$. Light propagation inside optical fibres in telecommunications systems occurs under conditions of total internal reflection.

## Diffraction gratings

Diffraction gratings are widely applied on account of their dispersive properties. Thus a beam of polychromatic light incident on a grating spatially separates according to its spectral content. A diffraction grating with period $\Lambda$ larger than the wavelength generally exhibits multiple diffracted waves excited by a single incident plane wave as illustrated in

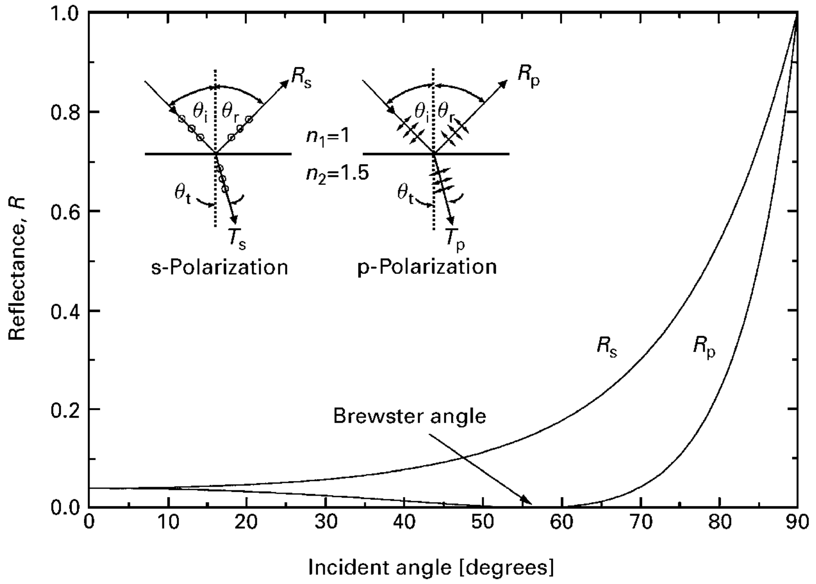



**Figure 2** Light reflectance as a function of the angle of incidence of a planar air–glass interface. The influence of the state of polarization of the incident plane wave is illustrated.

**Figure 3**. Diffraction gratings operate in reflection or transmission. In spectroscopic devices, such as monochromators, reflection gratings play key roles.

With reference to **Figure 3**, the grating equation describes the spatial division of the incident wave of wavelength $\lambda$ into distinct directions $\theta_\iota$ for each order $\iota$ as

$$\Lambda(\sin\theta_i + \sin\theta_\iota) = \iota\lambda \qquad [2]$$

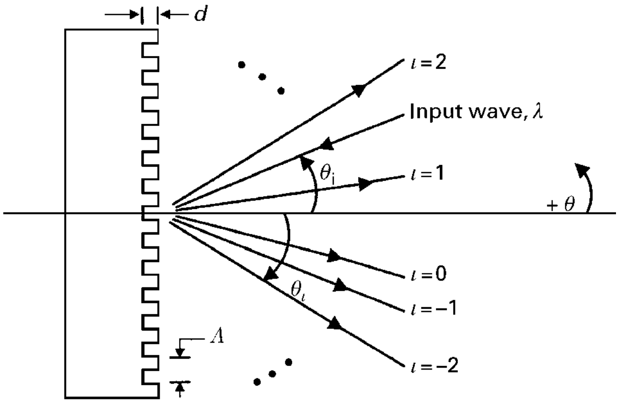



**Figure 3** Illustration of the dispersion properties of a surface-relief reflection grating. The monochromatic incident wave is diffracted into several waves propagating in directions specified by the grating equation. A change in the wavelength, $\lambda$, of the incident wave will alter the directions $\theta_\iota$ of the diffracted orders. Spectroscopic instruments often employ the first order ($\iota = 1$) to characterize the spectral content of the input light.

where $\iota = 0, \pm1, \pm2, \ldots$ labels the order of diffraction. Essentially the directions $\theta_\iota$ correspond to conditions of constructive interference of waves emanating from the grating facets. For a polychromatic incident wave, each component wavelength propagates into directions specified by Equation [2]. Taking $\theta_i$ = constant and differentiating, this angular dispersion is

$$\frac{d\theta}{d\lambda} = \frac{\iota}{\Lambda\cos\theta_\iota} \qquad [3]$$

Thus, the angular spread per unit spectral interval is largest for higher-order ($\iota$) diffracted waves, small grating periods ($\Lambda$), and large diffraction angles ($\theta_\iota$).

The diffraction efficiency, $DE_\iota$, specifies the optical power carried by the diffracted waves normalized with the power of the incident wave. For the surface-relief reflection grating of **Figure 3** with a square-wave profile, the diffraction efficiency is given by

$$DE_\iota = (2/\iota\pi)^2\sin^2(2g) \quad \iota = 1, 3, 5, \ldots \qquad [4]$$

where $g = \pi n_1 d/\lambda \cos\theta_i$. This grating obtains a maximum $DE_1$ of 40.5%. A lossless grating with a sawtooth (blazed) profile has a theoretical maximum $DE_1$ of 100%.

## Monochromators

A typical monochromator is schematically illustrated in **Figure 4**. The input light contains a range of

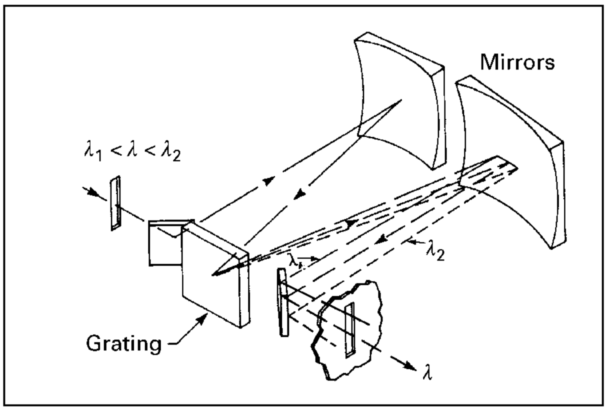



**Figure 4** The basic elements of a monochromator. The input spectrum contains wavelengths in the range $\lambda_1 < \lambda < \lambda_2$. The grating spreads the light within the diffraction orders (say $\iota = 1$) in which the instrument is operated. No other diffracted orders are shown in the drawing for clarity. The width of the slit defines the spectral content of the output light. Reproduced with permission from Oriel Corporation (1989) *Oriel 1989 Catalog*, Volume II.

wavelengths that are spatially separated with the diffraction grating as shown. The exit slit controls the spectral content of the output light by blocking the undesired wavelengths surrounding the central one. The spectrograph is a similar instrument but without a slit. In this case, the output spectrum is picked up by a detector array such that a prescribed spectral band can be associated with a given detector element.

## Thermal sources

Thermal sources supply light by heating the source material to a sufficiently high temperature. If an electric current is passed through a filament made of a metal with high melting point such as tungsten, the filament glows. Thermal radiation exhibits a characteristic, spectrally broad, continuous emission spectrum. The laws of blackbody radiation are useful for describing the emission spectra of thermal, or incandescent, sources. A blackbody consists of an ideal material that completely absorbs all incident radiation independent of wavelength. Conversely, the maximum radiation emitted by a hot solid material at a given temperature is that of a blackbody. The radiation properties of a blackbody are determined completely by the absolute temperature.

In practice, the characteristics of a blackbody can be obtained approximately by a cavity with a small hole in its side. Radiation entering the cavity through the hole is completely absorbed. As the cavity is heated, blackbody radiation emerges from the hole.

Planck's radiation law gives the spectral distribution of blackbody radiation. It may be expressed as

$$M_\lambda(T) = C_1\lambda^{-5}[\exp(C_2/\lambda T) - 1]^{-1} \qquad [5]$$

where $M_\lambda$ is the spectral radiant emittance given in W $cm^{-2}$ per unit wavelength. Thus $M_\lambda$ is the radiated power density per wavelength, $\lambda$. Further, T is the absolute temperature (K), $C_1 = 3.7418 \times 10^{-12}$ W $cm^2$ and $C_2 = 1.4388$ cm K. These coefficients are expressed in terms of fundamental physical constants as $C_1 = 2\pi hc_o^2$ and $C_2 = hc_o/k_B$ where h is Planck's constant and $k_B$ is Boltzmann's constant. **Figures 5** and **6** illustrate Planck's law on logarithmic and linear scales demonstrating the shift of the emission peak to shorter wavelengths with increasing temperature as well as increase of the radiated power density.

Taking $dM_\lambda/d\lambda = 0$ yields the wavelength, $\lambda_{peak}$, at which the radiation peak occurs as $\lambda_{peak}T = 0.28978$ cm K, which is Wien's displacement law traced by the dashed line in **Figure 5**. Finally, calculating the integral

$$M = \int_0^\infty M_\lambda \, d\lambda = \sigma T^4 \qquad [6]$$

gives the total power per unit area, M, emitted by the blackbody, where the Stefan-Boltzmann constant $\sigma = 5.6703 \times 10^{-12}$ W $cm^{-2}$ $K^{-4}$.

As the temperature of the blackbody increases, its colour changes as evident in **Figures 5** and **6**. The colour of a blackbody is thus uniquely dependent on its temperature. Real sources such as tungsten lamps that have spectral distributions approximating those of blackbodies can be assigned a colour temperature at which the spectrum most nearly overlays that of the blackbody. Outside the Earth's atmosphere, the Sun can, for example, be approximated as a blackbody at 6200 K, which is, then, the Sun's colour temperature.

**Figure 7A** shows the spectral distribution of an incandescent lamp fitted with a tungsten filament inside a quartz bulb enclosing a rare gas and a halogen. The spectrum is smooth with a colour temperature exceeding 3000 K containing a significant amount of near-infrared light.

## Electric-discharge sources

Electric discharge denotes the excitation of atomic states in a gaseous medium on passing an electric current through the medium. An ordinary household

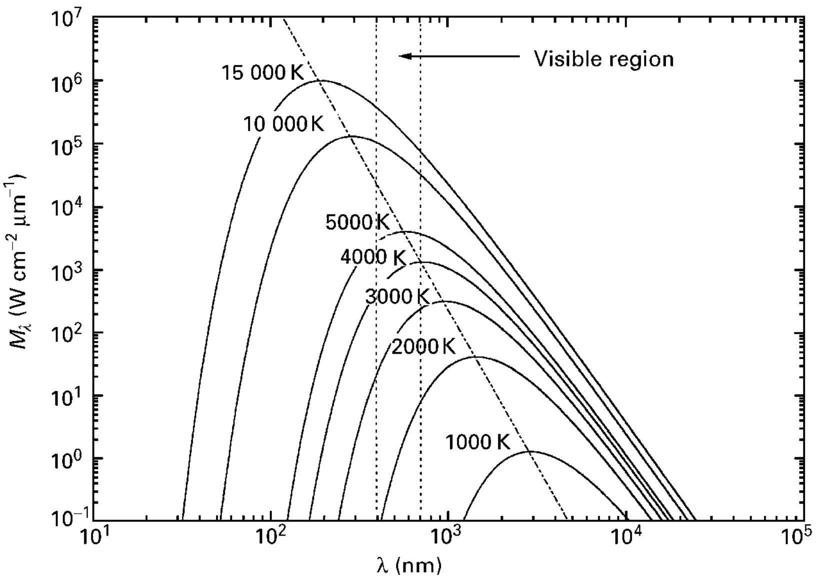



**Figure 5** Spectral characteristics of blackbody radiation. The spectral radiant emittance is plotted as a function of the wavelength for several values of the absolute temperature. The slanted, dashed line indicates Wien's displacement law.

fluorescent lamp is an example of an electric-discharge source. The lamp tube, coated on the inside with fluorescent phosphors, contains a mixture of a rare gas and mercury vapour. On application of sufficiently high voltage across the lamp's terminals, a low-pressure mercury arc forms. Thus, excited Hg atoms emit UV radiation (a particularly strong line exists at 253.7 nm) which is absorbed by the phosphor and re-emitted as visible light. The phosphor radiation spectrum is concentrated in the visible spectral region; it is smooth with several superimposed discrete lines originating in the Hg atoms.

By increasing the lamp's internal operating pressures and temperatures, high-intensity light can be produced. High-intensity discharge lamps are classified as mercury-vapour, metal-halide, or high-pressure-sodium types. In these sources, the discharge is contained in arc-tubes placed inside the lamp envelope filled with inert gas. Additionally, short-arc lamps, such as xenon and mercury-xenon lamps, are capable of power output in the multi-kilowatt region. Typical spectral characteristics of such sources are shown in **Figures 7B** and **C**.

**Table 1** gives wavelength ranges and power levels of selected light sources.

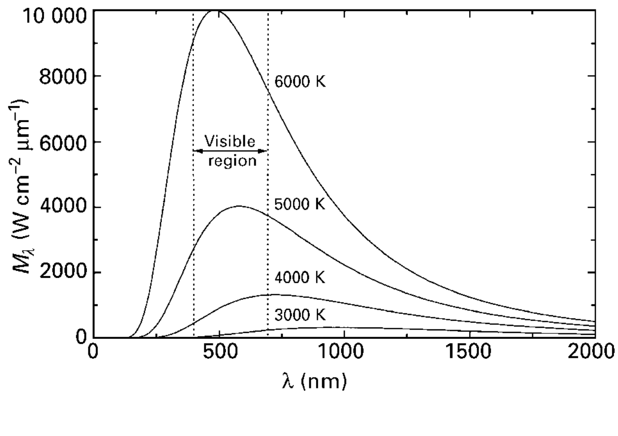



**Figure 6** Similar to **Figure 5** on a linear scale.

## Lasers

### Key characteristics

In contrast to light emitted by traditional thermal or luminescent sources, the laser beam is directional with all of its power confined within a narrow angular range. As the beam expands slowly on propagation, its divergence, $\Delta\theta$, is small, often of the order of a milliradian (mrad). A direct result of this confinement of laser power within a small angular region is the high value of laser brightness, or radiance, measured in watts (W) per unit area (cm$^2$) per solid angle (steradian, sr). A typical argon laser beam with 1 mm diameter and 1 mrad divergence carrying a power of 1 W has brightness of ~$10^8$ W cm$^{-2}$ sr$^{-1}$ whereas a 100 W light bulb radiating uniformly into a solid angle of $4\pi$ obtains only ~1 W cm$^{-2}$ sr$^{-1}$. In addition, laser light is highly monochromatic, meaning that the spectral wavelength spread, $\Delta\lambda$, is small. An associated property of monochromatic sources is temporal coherence; lasers are highly coherent

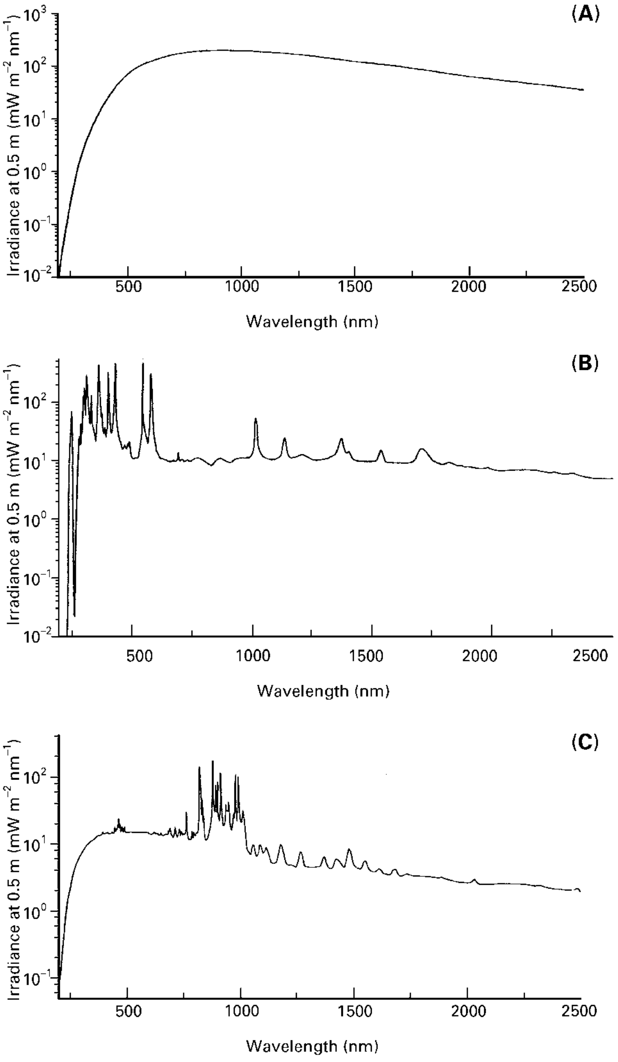



**Figure 7** (A) Spectrum of a 1000 W tungsten–halogen lamp. Reproduced with permission from Oriel Corporation (1994) *Oriel 1994 Catalog*, Volume II. (B) Spectrum of a 200 W mercury arc lamp. Reproduced with permission from Oriel Corporation (1994) *Oriel 1994 Catalog*, Volume II. (C) Spectrum of a 150 W xenon arc lamp. Reproduced with permission from Oriel Corporation (1994) *Oriel 1994 Catalog*, Volume II.

sources. Further, lasers operate in continuous-wave (CW) or pulsed modes, and they may be polarized or unpolarized. While many lasers operate with fixed output wavelengths, some laser types are tunable over relatively large spectral regions. Finally, laser beams can be focused to tiny spots (~1 μm diameter) with extreme power densities obtainable even for CW lasers.

## Applications

These unique characteristics of lasers make them exceedingly useful in numerous applications. Thus,

**Table 1** Representative power levels and wavelength ranges for common light sources

| *Source material* | *Wavelength range* (nm) | *Rated power* (W) |
|---|---|---|
| Mercury | 200–2500 | 50–1500 |
| Mercury–xenon | 200–2500 | 30–5000 |
| Xenon | 200–2500 | 15–30 000 |
| Deuterium | 180–400 | 30 |
| Sodium | 400–800 | 250–400 |
| Tungsten–halogen | 240–2 700 | 50–1 000 |

lasers enable applications in diverse fields such as telecommunications, holography, spectroscopy, integrated optics, lithography, and medicine. Specific functions include alignment, distance measurements, holographic data storage, signal processing, image processing, velocity measurement, industrial process control, inspection, surgery, welding, fusion, and propagation of fibre-optic communication signals.

## Atomic laser processes

The term 'laser' stands for 'light amplification by stimulated emission of radiation'. This stimulated emission occurs on interaction between a photon propagating in the laser's active medium and an atom in an excited state. The incident photon essentially stimulates the emission of an identical photon which joins the incident photon coherently at the same frequency and along the same direction. By viewing the photon as an electromagnetic wave packet, stimulated emission is a coherent wave process wherein the electromagnetic field of the incident photon drives the electrons of the excited atoms into coherent oscillation. The atom then undergoes energy level transition by radiating energy in the form of an optical wavefield that is spatially and temporally coherent with the incident, stimulating wavefield.

The fundamental laser process, thus, relies on quantum-mechanical energy levels in atomic systems. Media sustaining laser action possess discrete energy levels with discrete spectral components produced on transitions between particular levels. If the transition is from an upper level $E_2$ to a lower level $E_1$, the emitted photon will have frequency given by

$$\nu_{21} = (E_2 - E_1)/h \qquad [7]$$

where h is Planck's constant. The transition energy is $\Delta E = h\nu = hc/\lambda$.

Suitable laser energy levels have been discovered in thousands of atomic systems occurring in a variety of media. For example, discrete electron energy levels of isolated Ne atoms, originating in proton-electron Coulomb attraction, enable the ubiquitous HeNe laser. Vibrations of molecules with associated harmonic-oscillator type quantum levels are basic to the $CO_2$ laser. Organic dye molecules exhibit electronic, rotational, and vibrational energy levels on account of their size and complex structure. This provides a continuum of available transitions and yields broadband, tunable dye lasers. Semiconductor lasers operate on transitions based on energy-level bands that form on merging of the discrete levels associated with individual atoms in the solid.

In thermal equilibrium, an atomic system with $N_1$ atoms in level $E_1$ and $N_2$ atoms in $E_2 > E_1$ obeys the Boltzmann principle

$$N_2/N_1 = \exp[-(E_2 - E_1)/k_B T] = \exp[-h\nu/k_B T] \quad [8]$$

Therefore, in thermal equilibrium, $N_2 < N_1$ and, if $\nu$ is in the optical region, $N_2 << N_1$. Such a system is absorptive and not amplifying. To obtain the amplification effect associated with stimulated emission, an external pump mechanism is required to induce population inversion such that $N_2 > N_1$. Population inversion in the active laser medium is, thus, a non-equilibrium condition that enables lasing. The pumping approach employed depends on the laser medium. Optical flashlamps are used to invert some solid-state laser media (Nd:YAG, ruby), electric-current discharge is effective for many gas lasers (HeNe, argon), direct current injection pumps semiconductor lasers (GaAs, InP), and optical pumping with an auxiliary laser (e.g. argon laser) is used to operate dye lasers.

**Figure 8** schematically illustrates a four-level laser system such as that for a Nd:YAG laser. Population inversion and lasing is established between levels $E_2$ and $E_1$. The optical pump populates level $E_3$ that may be a broad range of closely spaced levels in practice. Decay from $E_3$ down to the metastable upper laser level $E_2$ may occur via radiative or nonradiative relaxation processes. Random radiative emission, or spontaneous emission, occurs without a stimulating electric field, in contrast to stimulated emission. Nonradiative relaxation implies energy transfer via lattice vibration also called phonon modes.

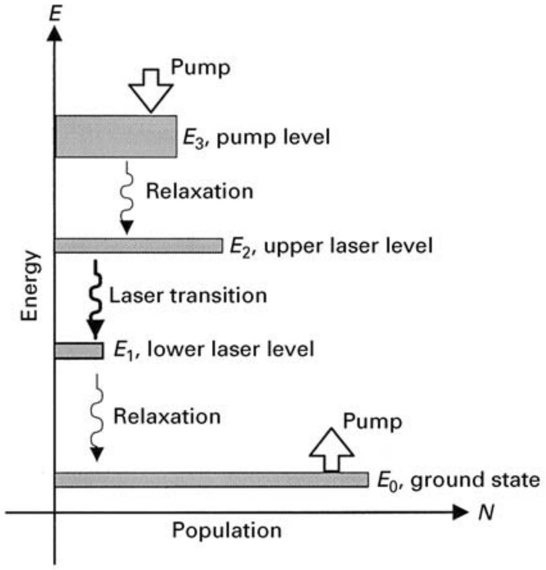



**Figure 8** A four-level atomic energy system.

## Spectral content

Spontaneous transitions in laser media possess a finite spectral bandwidth $\Delta\nu$. The concomitant frequency distribution is described by a line shape function, $g(\nu)$, taken to be normalized as

$$\int_{-\infty}^{\infty} g(\nu)\mathrm{d}\nu = 1 \quad [9]$$

The function $g(\nu)$ determines the spectral characteristics of the laser gain coefficient in inverted media. Light emission and absorption are both spectrally governed by $g(\nu)$.

The functional shape of $g(\nu)$ is defined by the particular line-broadening mechanisms of the atomic systems that constitute the active medium. If the atoms in the collection are indistinguishable, i.e. each atom acts the same, the line broadening is said to homogeneous. Lifetime broadening and collision broadening are examples of homogeneous broadening. Systems with distinguishable atoms suffer inhomogeneous broadening such as Doppler broadening and lattice-strain broadening.

Lifetime broadening is due to the finite lifetime of the excited atomic state. It can be modelled as an exponentially decaying oscillator of the form (see Eqn [1])

$$e(t) = e_0 \exp(-t/\tau)\cos(\omega_0 t), \quad t \geq 0 \quad [10]$$

emitting a wave of frequency $\omega_0$ during the lifetime $\tau$. The frequency content of the decaying oscillator is found by applying a Fourier transform

$$F(\omega) = \int_{-\infty}^{\infty} f(t)\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t \qquad [11]$$

This leads to a normalized, Lorentzian, line shape function in the vicinity of $\omega \cong \omega_0 = 2\pi\nu_0$ expressed as

$$g(\nu) = \frac{\Delta\nu}{2\pi\left[(\nu-\nu_0)^2 - (\Delta\nu/2)^2\right]} \qquad [12]$$

where $\Delta\nu = 1/\pi\tau$ is the line width (full width at half maximum, FWHM) of the lifetime-broadened system. This result applies to homogeneous broadening, where each atom in the collection is characterized by a bandspread $\Delta\nu$. Short pulses ($\tau$ small) yield broad spectra. Atomic collisions that interrupt the decay cause additional broadening by reducing the effective average value of $\tau$.

Doppler broadening is a commonly cited example of inhomogeneous broadening. The atoms in a gas have a statistical distribution of velocities; the atoms are thus distinguishable by their differing velocities that correspond to different atomic resonance, or transition, frequencies. The resulting Gaussian line shape function is given by

$$g(\nu) = \frac{2\sqrt{\ln 2}}{\sqrt{\pi}\Delta\nu_\mathrm{D}}\exp\{-[4\ln 2][(\nu-\nu_o)/\Delta\nu_\mathrm{D}]^2\} \qquad [13]$$

where the FWHM line width is $\Delta\nu_D = 2\nu_0[2k_BT \ln 2/mc^2]^{1/2}$ where m is the atomic mass. Thus, as expected, $\Delta\nu_D$ increases with increasing T and decreasing mass. For example, the HeNe laser is Doppler broadened with $\Delta\nu_D$ ~1.5 GHz whereas the Nd:YAG laser is homogeneously broadened with $\Delta\nu$ ~120 GHz.

The line shape function $g(\nu)$ spectrally governs the rate of spontaneous and stimulated emission. The stimulated transition rate per atom is given by $W_i = \lambda^2 I_\nu g(\nu)/(8\pi n^2 h\nu\tau)$, where n is the refractive index and $I_\nu$ is the intensity (W cm$^{-2}$) of the monochromatic optical wave (at frequency $\nu$) causing stimulated emission. The relation to laser gain is established by considering propagation along the z-direction in the laser medium, which can be described by

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}z} = (N_2 - N_1)\left[\left(\lambda^2 g(\nu)\right)/(8\pi n^2\tau)\right]I_\nu \equiv \gamma(\nu)I_\nu \qquad [14]$$

which defines the exponential gain coefficient. The solution is

$$I_\nu(z) = I_\nu(0)\exp[\gamma(\nu)z] \qquad [15]$$

implying exponential growth of the laser radiation intensity for $\gamma(\nu) > 0$ or $N_2 > N_1$ (inverted medium) and a traditional decay if the system is absorbing, or noninverted, in thermal equilibrium.

## Basic laser components

**Figure 9** illustrates the building blocks of typical lasers. The active medium (gas, solid, or liquid) is placed between two mirrors constituting an optical resonator. The back mirror is highly reflective with reflectance, R, approaching 100%. The front mirror (output coupler) is also highly reflective but with a finite value of transmittance, T, typically a few percent, to allow the useful output laser beam to emerge from the resonator. The resonator-cavity mirrors may be flat or curved. A laser with two flat mirrors is difficult to align as small misalignment causes the oscillating light wave, making multiple round trips, to walk off and leave the cavity. It is thus a marginally stable configuration. Improved design is obtained with curved mirrors confining the laser light energy efficiently within the resonator. In particular, two mirrors with long radii of curvature constitute a practical, stable laser cavity. The laser medium is pumped with an external source, also shown in **Figure 9**, to maintain the inverted atomic population needed for laser power amplification to occur. Pumping methods include optical pumping of solid state lasers with incoherent (e.g. xenon) flash lamps, electric-discharge pumping of gas lasers, and direct-current-injection pumping of semiconductor lasers and light-emitting diodes. Collisions can also play a role

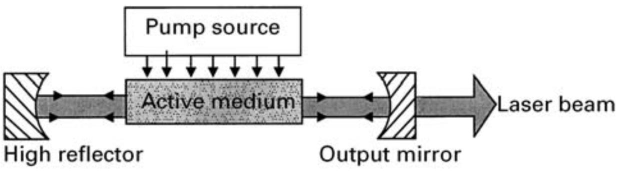



**Figure 9** Chief components of a laser.

in achieving the population inversion. For example, in the HeNe laser, the electrical discharge (a few kV at a few tens of mA current) drives electrons that collide with He atoms, raising them to excited levels (2s level desired). The He atoms then interact with Ne atoms, raising them to $3s_2$ levels by collision energy transfer. The resulting metastable state with a few μs lifetime radiates the familiar red photons with 633 nm wavelength.

### Laser modes

Laser oscillation can build up from noise or from a deliberately injected signal. Before lasing, spontaneous emission generates a flood of photons propagating in random directions. Those taking a path along the resonator axis are reflected back into the active medium (i.e. feedback) and initiate the stimulated emission and amplification process with power provided by the external pump. If the round-trip gain exceeds the round-trip loss, optical energy builds up in the resonator. Since the pump provides finite energy, this oscillating power buildup saturates with steady-state oscillation established when the round-trip gain equals the round-trip loss. Under these conditions, a steady beam may be emitted from the laser, since the useful output light power is taken as part of the loss.

Thus, laser oscillation occurs inside a resonator confining the laser light. The resonator defines the frequency distribution and the spatial distribution of the output light. A resonator with planar mirrors separated by distance L supports longitudinal modes with discrete frequencies

$$\nu_q = q(c/2L), \quad q = 1, 2, 3 \ldots \qquad [16]$$

Adjacent modes are thus separated by c/2L. A laser will support those longitudinal modes of frequency $\nu_q$ that fall under the gain curve of the laser medium.

**Figure 10A** shows an example atomic line and several longitudinal cavity modes with six modes having sufficient gain to oscillate, the rest being below threshold and not surviving. The spectral properties of the individual longitudinal modes can be understood by treating the resonator as a Fabry-Perot etalon. Their spectral width, $\delta\nu$, is given in terms of the resonator's free spectral range, FSR = c/2L, and its finesse, F, as $\delta\nu$ = FSR/F. If each of the cavity mirrors has a reflectance R, the finesse is $F = \pi R^{1/2}/(1-R)$.

**Figure 11** depicts some of the main components in a large-frame laser such as a commercial argon-ion laser. The resonator supports multiple longitudinal modes as suggested by Equation [16]. The active medium, in the case of argon, supports multiple well-separated atomic lines any one of which, as shown in **Figure 10A**, can accommodate several longitudinal cavity modes. To select one of the atomic lines, a prism is inserted in the cavity as shown. By angularly tuning the prism, a principal line, or wavelength, is selected such that it emerges out of the prism at normal incidence on the high reflector. This line can then oscillate in the cavity with other lines cut off. To narrow the laser's spectrum further, an intracavity etalon is applied. The etalon is a planar Fabry-Perot cavity of thickness $d_e$ with its FSR = $c/2d_e$ thus supporting multiple longitudinal modes separated by a much larger frequency spread than the laser cavity modes since L >> $d_e$. The etalon is tuned by rotation and by varying its temperature. When an etalon mode coincides with one of the laser modes in **Figure 10A**, that mode oscillates by itself with the adjacent modes suppressed. The laser then oscillates in a single mode with a corresponding decrease in

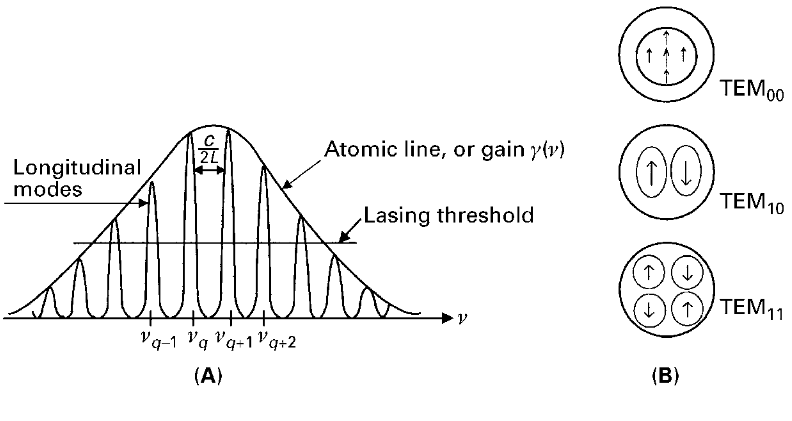



**Figure 10** Spectral and spatial modes of a laser. (A) Atomic line enclosing several longitudinal resonator modes. (B) Transverse spatial modes of a laser beam. The fundamental $TEM_{00}$ mode is the Gaussian laser beam profile. The arrows indicate the direction of the beam's electric field at a given instant of time.

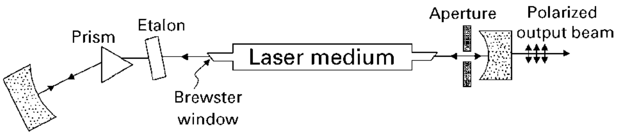



**Figure 11** Principal components used to optimize the performance of a laser. The prism selects a particular atomic line and the wavelength of operation. The etalon selects a longitudinal mode and increases the laser's coherence. Reducing the aperture diameter favours the Gaussian spatial mode. The Brewster windows polarize the beam linearly.

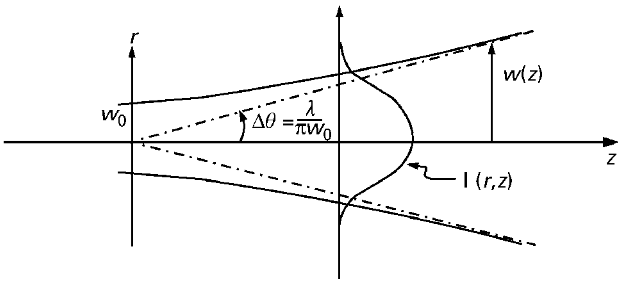



**Figure 12** Parameters of a Gaussian laser beam.

the output spectral line width and associated increase in its level of coherence. The laser in **Figure 11** is equipped with Brewster windows (see **Figure 2**) that force the laser to favour a single polarization state as indicated.

In addition to the longitudinal modes discussed, the laser resonator admits multiple transverse modes that appear as spatial intensity variations across the output laser beam as shown in **Figure 10B**. Such transverse electromagnetic (TEM) modes are labelled with subscripts that count the number of zeros along the (x,y) transverse directions. The most useful mode has no zeros and thus the $TEM_{00}$ mode is preferred. This mode is the ubiquitous Gaussian laser beam obtained when the laser aperture in **Figure 11** suppresses higher-order transverse modes. With reference to **Figure 12**, the intensity of a Gaussian laser beam carrying a total power P is expressible as

$$I(r,z) = \frac{2P}{\pi w^2(z)} \exp\left(-\frac{2r^2}{w^2(z)}\right) \qquad [17]$$

with $w(z) = w_0[1 + (z/z_0)^2]^{1/2}$ where $z_0 = \pi w_0^2/\lambda$, called the Rayleigh range, $w_0$ being the beamwaist radius occurring at z = 0, and $r^2 = x^2 + y^2$. If a thin lens of focal length f, with the lens placed at the waist, focuses this Gaussian beam, the radius of the focal point is

$$w_f = w_0/[1 + (z_0/f)^2]^{1/2} \qquad [18]$$

Often $z_0 >> f$ which gives $w_f \sim f\lambda/\pi w_0$. As an example, with f = 10 mm, $\lambda = 0.5$ μm, and $w_0 = 0.5$ mm, $w_f \sim 3$ μm. If P = 1 W, this yields a power density at focus of ~3 MW cm$^{-2}$ or ~3 kW cm$^{-2}$ if P = 1 mW (e.g. a small HeNe laser). In view of these high power densities, laser safety is an important concern. Eye protective goggles matched to the laser's wavelengths should be worn when working with lasers.

## Pulsed lasers

Generation of intense short pulses of light is of interest in many fields, including nonlinear optics and spectroscopy. Pulse lengths in the femtosecond region with pulse power levels of terawatts have been achieved. Two common methods for short-pulse generation, Q-switching and mode locking, are summarized in this section.

In Q-switching, the laser is prevented from oscillating with the pump running such that a large inverted population builds up in the cavity. Oscillation occurs on removal of the blockage thereby switching the cavity Q back to its normally high value. A short, powerful optical pulse develops as the available gain rapidly depletes, typically on nanosecond timescales. Q-switching is accomplished by placing electro-optic or acousto-optic modulators inside the cavity or, in a passive manner, with a saturable absorber. The passive approach is particularly simple; a saturable absorber (organic dye solution) becomes transparent when sufficiently high gain has developed thus permitting a high-intensity pulse to escape. As an example, a Q-switched Nd:YAG laser may deliver ~1 J energy per pulse with pulse length of $\tau_p$ ~5 ns, power per pulse ~200 MW, and repetition rate of 10 Hz.

An ideal, inhomogeneously broadened laser oscillates on multiple longitudinal modes separated by c/2L in frequency. These free-running modes oscillate independently with random, uncorrelated phases. If the phases ($\phi$ in Eqn [1]) can be locked together, a periodic laser-pulse train is obtained as output light. To accomplish this, an electro-optic or acousto-optic switching element can be placed in the cavity. It is modulated with a period equal to the resonator's roundtrip transit time T = 2L/c. Optimum conditions for light buildup correspond to minimum loss, that is an open switch. This switching action tends to lock the phases of the longitudinal modes such that their interference creates a pulse that passes through the open switch without loss. Wide atomic line widths, $\Delta\nu$, enclosing a large number of

**Table 2** Representative figures characterizing common lasers

| *Laser type* | *Wavelengths* (nm) | *Power* | *Beam dia* (mm) | *Divergence* (mrad) | *Output* |
|---|---|---|---|---|---|
| Helium–neon | 633,1152 | 0.1–50 mW | 0.5–3.0 | 0.5–6.0 | CW |
| Argon–ion | 351–529 several | 5 mW–25 W | 0.5–2.0 | 0.5–1.0 | CW |
| Helium–cadmium | 325, 442 | 1–50 mW | 0.3–1.2 | 0.5–3.0 | CW |
| Dye (Ar pump) | 400–1000 | to 5 W | 0.5–1.0 | 1–2 | CW, tunable |
| Nd:YAG | 1064 | 0.1–500 W (ave) | 1–10 | 1–10 | Pulsed |
| Nd:YAG | 1064 | 0.5–1000 W | 1–10 | 1–10 | CW |
| $CO_2$ | 10 600 | 1–100 W | 3–7 | 1–5 | CW |
| GaAlAs | 750–900 | 1–100 mW | Divergent output | Large, ~30° | CW |
| InGaAsP | 1100–1600 | 1–100 mW | Divergent output | Large, ~30° | CW |
| Ti:sapphire (Ar pump) | 700–1100 | 0.1–1 W | 0.5–1 | 1–2 | CW, tunable |

longitudinal modes produce the shortest mode-locked pulses since $\tau_p \sim \Delta\nu^{-1}$. For example, Ti:sapphire lasers, with their broad line widths, provide short pulses, with $\tau_p \sim 100$ fs, average power ~100 mW, and output pulse rate of ~100 MHz being representative numerical data.

**Table 2** provides numerical parameters for some common lasers.

## List of symbols

c = speed of light in medium; $c_o$ = speed of light in vacuum; $C_l$, $C_2$ = coefficients in Planck's law; d = grating thickness; $d_e$ = etalon thickness; $DE_\iota$ = diffraction efficiency of order $\iota$; ***e*** = electric field vector; $e_o$ = e-field amplitude; $E_l$, $E_2$ = energy levels; f = focal length; FWHM = full width at half maximum; FSR = free spectral range; $F(\omega)$ = Fourier transform of f(t); g = normalized grating modulation strength; $g(\nu)$ = atomic line shape function; ***h*** = magnetic field vector; h = Planck's constant; I(r,z) = intensity of Gaussian laser beam; $I_\nu$ = intensity of light of frequency $\nu$; k = wavenumber; $k_B$ = Boltzmann's constant; L = laser resonator length; m = atomic mass; M = emittance; $M_\lambda$ = spectral radiant emittance; n = index of refraction; $N_l$, $N_2$ = atomic populations; P = power; q = integer mode index; $R_{s,\,p}$ = reflectance of s- or p-polarized light; $r^2 = x^2 + y^2$; t = time; T = absolute temperature; $T_{s,\,p}$ = transmittance of s- or p-polarized light; $w_0$ = radius of a Gaussian beam at the waist; $w_f$ = focal point radius; w(z) = beam radius at arbitrary z; x, y, z = space coordinates; $z_o$ = Rayleigh range; $\gamma(\nu)$ = gain coefficient; $\delta\nu$ = FWHM of longitudinal cavity modes; $\Delta\theta$ = laser beam divergence; $\Delta\lambda$ = spectral line width, wavelength; $\Delta\nu$ = spectral line width, frequency; $\Delta\nu_D$ = Doppler line width; $\tau$ = lifetime; $\theta_i$ = angle of incidence; $\theta_r$ = angle of reflection; $\theta_B$ = Brewster angle; $\theta_c$ = critical angle; $\theta_\iota$ = angle of diffraction; $\theta_t$ = angle of refraction; $\lambda$ = wavelength of light; $\Lambda$ = grating period; $\nu$ = frequency of light; $\nu_{12}$ = transition frequency; $\nu_q$ = frequency of longitudinal cavity mode; $\sigma$ = Stefan-Boltzmann constant; $\iota$ = diffraction-order integer; $\tau_p$ = pulse length; $\omega$ = angular frequency of light; $\phi$ = phase.

*See also:* **Electromagnetic Radiation; Ellipsometry; Fibre Optic Probes in Optical Spectroscopy, Clinical Applications; Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Optical Frequency Conversion; Raman Spectrometers.**

## Further reading

Bass M ed (1995) Handbook of Optics: Fundamentals, Techniques & Design, 2nd edn, Vol I, Part 4: Optical sources. New York: McGraw-Hill.

Born M and Wolf E (1975) Principles of Optics, 5th edn. Oxford: Pergamon Press.

Hecht E (1987) Optics, 2nd edn. Reading, MA: Addison-Wesley.

Hecht J (1992) The Laser Guidebook, 2nd edn. New York: McGraw-Hill.

Proudfoot CN (ed) (1997) Handbook of Photographic Science and Engineering, 2nd edn, Section 1: Radiation sources. Springfield, VA: Society for Imaging Science and Technology.

Saleh BEA and Teich MC (1991) Fundamentals of Photonics. New York: Wiley.

Siegman AE (1986) Lasers. Mill Valley, CA: University Science Books.

Silfvast WT (1996) Laser Fundamentals. New York: Cambridge University Press.

Waynant RW and Ediger MN (eds) (1994) Electro-Optics Handbook. New York: McGraw-Hill.

Yariv A (1997) Optical Electronics in Modern Communications, 5th edn. New York: Oxford University Press.

# Linear Dichroism, Applications

**Erik W Thulstrup**, Roskilde University, Denmark
**Jacek Waluk**, Polish Academy of Sciences, Warsaw, Poland
**Jens Spanget-Larsen**, Roskilde University, Denmark

Copyright © 1999 Academic Press



Light absorption is not only characterized by energy and intensity, it also has directional properties. These are often overlooked, even when they might provide essential information. Such information includes assignments of electronic and vibrational transitions and information on molecular structure, conformation and association. The directional properties of individual molecules cannot be observed in isotropic samples but require samples that are non-isotropic or aligned.

Although a natural (nonpolarized) light beam is nonisotropic in nature, since its electric and magnetic vectors are located exclusively in the plane perpendicular to the direction of the beam, the directional properties of aligned samples may best be studied using linearly polarized light. The term 'linear dichroism' refers to the change in absorption that may often be observed for nonisotropic samples when the direction of polarization of the light is changed by 90°.

In reality, most light in nature is partially polarized – scattering produces polarized light and most light around us is scattered sunlight. Historians believe that the Vikings used the polarization of light from the blue sky as a navigational aid when they travelled to North America about 1000 years ago. Similarly, except for common solutions, most naturally occurring samples are nonisotropic; this is especially true for samples with a biological origin. The first linear dichroism experiment was actually carried out on a dye-stained cell membrane in the 1880s.

## The transition moment

Many observable properties in connection with electronic transitions between state $|0\rangle$ and $|f\rangle$ may be calculated from the two wavefunctions by means of an integral of the type

$$\langle 0|\hat{O}|f\rangle = O_{0f} \quad [1]$$

where $\hat{O}$ is an operator describing the desired property. In many cases $\hat{O}$ is not a scalar but a vector or tensor. The operator that describes absorption probability, the electric dipole operator $\hat{M}$, is a vector. Thus the corresponding observable property, the transition moment $M_{0f}$, will also be a vector.

$M_{0f}$ may vanish for reasons of symmetry; then the transition between $|0\rangle$ and $|f\rangle$ is said to be symmetry forbidden. For example, if the molecule possesses a plane of symmetry, the product of $\langle 0|$, $\hat{M}$ and $|f\rangle$ will be either symmetric or antisymmetric with respect to the plane; if it is antisymmetric, the integral vanishes. The electric dipole operator $\hat{M}$ corresponds to the sum of charges on particles in the molecule times their coordinates and is therefore antisymmetric with respect to any symmetry plane, thus $\langle 0|$ and $|f\rangle$ must have different symmetry properties with respect to the plane in order for the integral to be nonzero.

In molecules with symmetry elements, the direction of $M_{0f}$ is determined by the symmetry of the two states $|0\rangle$ and $|f\rangle$. In the important cases of $C_{2v}$, $D_{2h}$ and $D_2$ symmetric molecules, only three different excited-state symmetries can be reached by an electric dipole transition from a totally symmetric ground state. For example, in the case of $D_{2h}$, each of the three 'allowed' excited state types are antisymmetric with respect to one, and only one, of the three symmetry planes of the molecule. The direction of $M_{0f}$ for the three different allowed transitions is in each case perpendicular to the antisymmetry plane. For transitions to all other excited states, $M_{0f}$ becomes zero; these transitions are said to be symmetry forbidden.

## Production of linear dichroic samples

Although many samples occurring in nature are partially aligned from the start, most linear dichroism experiments are performed on samples that have been aligned in the laboratory. Single crystals may be grown of many molecules: if the crystal structure is known and is suitable, information may be obtained on transition moments for electronic and vibrational transitions. In practice, this is often experimentally difficult – very thin crystals are

required for reasonably strong transitions and the interpretation of the spectra may be very complicated, primarily owing to strong intermolecular interactions in the crystal. Alignment in electric fields may be used, especially for large molecules (see the Kerr effect). Magnetic fields may be used to align molecules with permanent magnetic moments, but the effect is weak (see the Cotton–Mouton effect). Further details may be found in the article on linear dichroism instrumentation.

In the following we shall consider samples aligned either by photoselection or by dissolving the molecules in 'anisotropic solvents'. Ordinary photoselection is not based on a physical alignment of a set of molecules. Instead a partially aligned subset of an isotropic molecular assembly is selected by absorption of natural or polarized light. Since the relative transition probability for molecules with different spatial orientations is known, the orientation distribution of molecules in the photoselected sample will often be known. In order to preserve this alignment, the molecules must be kept in a high viscosity solvent, usually at low temperature. Anisotropic solvents may be single crystals of a suitable host molecule, membranes, a fast-flowing liquid, aligned liquid crystals or stretched polymers. With the possible exception of the first of these solvents, the orientation distribution of the sample molecules is not known, but, as we shall see, a complete interpretation of absorption spectra is possible even when only a few, simple properties of this distribution are known.

## Description of molecular alignment: One-photon processes

The transition probability for linearly polarized light with its electric vector $\boldsymbol{\varepsilon}_U$ along sample (laboratory) axis $U$ at the wavelength of the transition to state $|f\rangle$ in a single molecule is proportional to

$$(\boldsymbol{\varepsilon}_U \cdot \boldsymbol{M}_{0f})^2 = (\boldsymbol{M}_{0f})^2 \ \cos^2(U, \boldsymbol{M}_{0f}) \qquad [2]$$

where $(U, M_{0f})$ is the angle between $U$ and $M_{0f}$. The sample axis $U$ is usually chosen according to the properties of the individual sample; most important samples are uniaxial (see below) and in these $U$ is the unique axis. It is clear that the absorption depends on molecular alignment; if the molecule is aligned so that $M_{0f}$ is perpendicular to $U$, the probability becomes zero, while the maximum absorption probability occurs when $M_{0f}$ is parallel to $U$. In a typical partially aligned sample the angle $(U, M_{0f})$ may take on all possible values between 0° and 90° for individual molecules in the sample. This means that an average value over all sample molecules must be used and we obtain

$$\langle(\boldsymbol{\varepsilon}_U \cdot \boldsymbol{M}_{0f})^2\rangle = (\boldsymbol{M}_{0f})^2 \langle\cos^2(U, \boldsymbol{M}_{0f})\rangle \qquad [3]$$

where the brackets refer to average values for all molecules in the sample.

In the important case of a molecular symmetry ($C_{2v}$, $D_{2h}$ and $D_2$) that only allows transition moments along three different perpendicular $x$, $y$ and $z$ directions within the molecule, we label the corresponding average cosine square values $K_x$, $K_y$ and $K_z$:

$$\langle\cos^2(U, \boldsymbol{M}_{0f})\rangle = K_x, K_y \text{ or } K_z \qquad [4]$$

for $x$-, $y$- or $z$-polarized transitions, respectively. These three $K$ values are not independent; their sum is equal to 1, which often provides a useful test. Note, however, that $K$ values are useful for any direction in a molecule, for example a low-symmetry molecule; the transition to state $|f\rangle$ will, for a given sample alignment, be characterized by the $K$ value $K_f$:

$$K_f = \langle\cos^2(U, \boldsymbol{M}_{0f})\rangle \qquad [5]$$

The molecular direction that corresponds to the largest possible $K_f$ – in other words the direction that, on the average, is best aligned in the molecule with respect to $U$ – is called the orientation axis. Since it is determined by averaging, the position of the orientation axis is not always obvious; as a result, the axis is sometimes called the effective orientation axis. In the symmetrical molecules mentioned above it will be one of the three directions determined by symmetry. It is often labelled $z$. The least well-aligned axis, with the lowest $K$ value, is often called $x$ and the intermediate axis is called $y$. The three axes are those that diagonalize a '$K$-tensor' with elements $\langle\cos(u,U)\cos(v,U)\rangle$, where $u$ and $v$ are directions in the molecule.

The orientational dependence of the two linear dichroism spectra can easily be expressed in terms of the $K$ values. If the electric vector is along the $U$ axis, we obtain for the resulting absorbance $E_U(\nu)$:

$$E_U(\nu) = \sum K_f A_f(\nu) \qquad [6]$$

where the sum is over all relevant transitions $f$ and $A_f(\nu)$ is equal to the contribution to the absorbance from transition $f$ in a sample of perfectly aligned

molecules (this is also equal to three times the contribution from transition $f$ to the absorbance of an isotropic sample). Similarly, if the electric vector is along the $V$ laboratory axis, perpendicular to $U$, we obtain for the resulting absorbance $E_V(\nu)$

$$E_V(\nu) = \sum K'_f A_f(\nu) \qquad [7]$$

where $K_f' = \langle \cos^2 (V, M_{0f}) \rangle$. The linear dichroism is equal to

$$\mathrm{LD}(\nu) = E_U(\nu) - E_V(\nu) \qquad [8]$$

### Uniaxial samples

The equations above are valid for any sample. Each transition moment direction is characterized by two quantities, $K$ and $K'$. If the sample is uniaxial, i.e. all directions perpendicular to a given sample axis are equivalent, the equations simplify further. We chose $U$ as the sample axis. Since $V$ is now equivalent to the third laboratory axis, $W$, it can be shown that for all transitions $f$, $K_f' = (1 - K_f)/2$. We obtain

$$\begin{aligned} E_U(\nu) &= \sum K_f A_f(\nu) \\ E_V(\nu) &= \tfrac{1}{2}\sum (1 - K_f) A_f(\nu) \\ \mathrm{LD}(\nu) &= \tfrac{1}{2}\sum (3K_f - 1) A_f(\nu) \end{aligned} \qquad [9]$$

When the two curves $E_U(\nu)$ and $E_V(\nu)$ differ very little, i.e. when the LD($\nu$) is small, a direct measurement of LD($\nu$) by means of polarization modulation provides a more accurate result. The second (linearly independent) curve is then either the isotropic spectrum

$$\begin{aligned} E_{\mathrm{iso}}(\nu) &= [E_U(\nu) + 2\,E_V(\nu)]/3, \\ & E_V(\nu), \text{or} \\ & [E_U(\nu) + E_V(\nu)]/2 \end{aligned}$$

Many important samples have been shown to be or are assumed to be uniaxial, for example solutes in stretched polyethylene or in nematic liquid crystals. It is now simple to determine $K_f$. If transition $f$ dominates the absorption at $\nu_f$ (if it does not overlap with other transitions), the dichroic ratio at $\nu_f$, $E_U(\nu_f)/E_V(\nu_f)$, can be written

$$\frac{E_U(\nu_f)}{E_V(\nu_f)} = \frac{2K_f}{1 - K_f} \qquad [10]$$

Therefore, $K_f$ can be directly determined:

$$K_f = \frac{E_Z(\nu_f)}{2E_V(\nu_f) + E_U(\nu_f)} \qquad [11]$$

The situation for nonoverlapping transitions is common in IR spectroscopy, but less common in UV spectroscopy. If transition $f$ is overlapped by other transitions, it is still often possible to determine $K_f$. This is done by a trial-and-error method (the TEM method) in which a series of linear combinations of the observed spectra are formed, using values of $K$ near the expected value for $K_f$:

$$\tfrac{1}{2}(1 - K)E_U(\nu) + KE_V(\nu) \qquad [12]$$

The spectral feature (peak, shoulder) due to transition $f$ disappears from the linear combination when $K = K_f$. This may provide a very accurate determination of $K_f$, even in cases of strong overlap. **Figure 1** shows the observed spectra, the dichroic ratio ($E_U(\nu)/E_V(\nu)$), and linear combinations of the former that allow a determination of the two different $K$ values present in the spectrum from the disappearance of the three major peaks in the spectrum.

**Figure 2** shows a plot of $K_z$ values against $K_y$ values for a large number of symmetrical molecules aligned in stretched polyethylene. Obviously, shape is very important for the alignment. The triangle shows the theoretical limits for the $K$ values, assuming that $K_z \geq K_y \geq K_x$. The lower left side corresponds to a rod-like alignment, $K_x = K_y$, and the lower right side to a disc-like alignment, $K_z = K_y$.

It is remarkable that the simple use of squares of directional cosines (the $K$ values; sometimes equivalent quantities are used, the Saupe $S$ matrices and the order parameters) provides a description of all directional properties of the sample, which can be observed in absorption spectroscopy, without any assumptions. Several alternative descriptions (the Fraser–Beer model, the Tanizaki model, the Popov model) involve assumptions that are rarely fulfilled. At the same time, some of these models tend to be mathematically complex. A simple extension of the

$K$ values to averages of fourth powers of directional cosines ($L$ values) makes it possible to describe two-photon processes (luminescence, two-photon absorption, Raman) in aligned samples.

## Linear dichroism spectra: Interpretation

### Photoselection

We shall in the following only consider the photoselection of a subset of molecules from an isotropic, 'stationary' sample. By stationary is meant that the sample molecules do not rotate on the timescale of the experiment. Descriptions of photoselection from an already aligned sample or, from a sample with molecular rotational movement, or of photoselection of each molecule with two photons, are more complicated but are available in the literature. Photoselection is usually performed with a collimated light beam with a well-defined wavelength; the light may be either linearly polarized or natural/circularly polarized, the two latter will produce similar results.

When an aligned subset of sample molecules is selected, another aligned subset, that consisting of the remaining molecules, is also created. When the sample depletion is negligible, the former set of molecules is small in number, but have a high degree of alignment, while the remaining molecules have negligible alignment. When the sample depletion is substantial, both the selected set and remaining set of molecules will be aligned. We shall in the following assume that only one transition is responsible for the photoselection of a small fraction of the sample. Other cases are also of interest and are described in the literature.

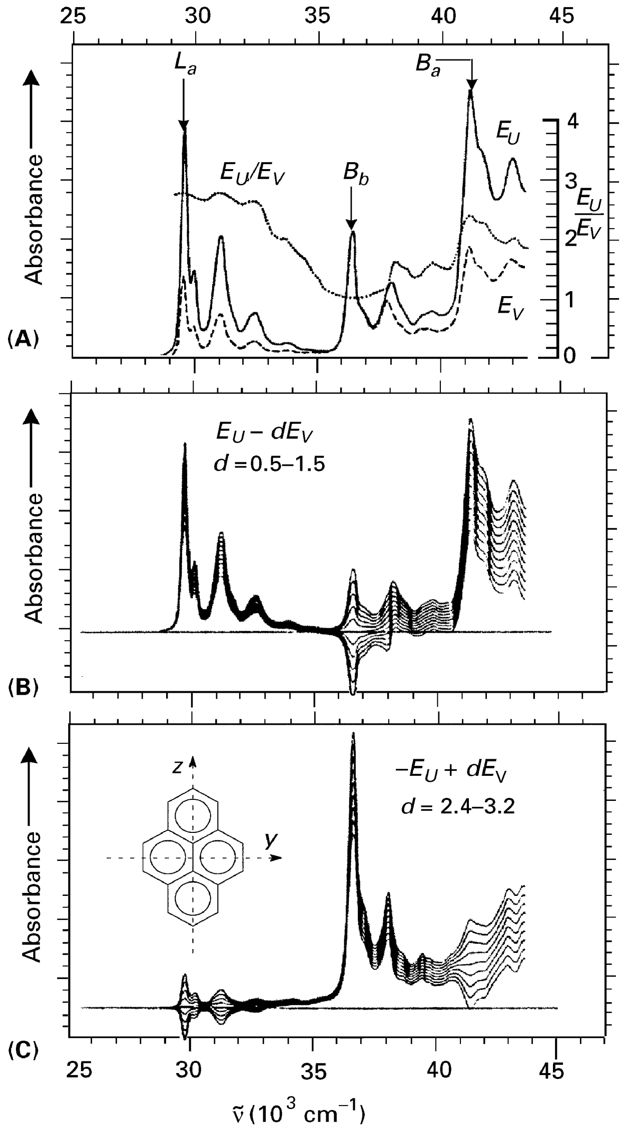



**Figure 1** UV spectra of pyrene in stretched polyethylene at 77K. Absorbance in arbitrary units. (A) $E_U$ and $E_V$ (baseline-corrected) and the dichroic ratio $E_U/E_V$. (B) and (C) Linear combinations of $E_U$ and $E_V$, in search of spectral curves corresponding to $A_z$ and $A_y$, respectively (the TEM trial-and-error process). Adapted by permission from Langkilde FW, Gisin M, Thulstrup EW and Michl J (1983) *Journal of Physics and Chemistry* **87**: 2901–2911.

### Photoselection: negligible depletion; linearly polarized light

The photoselected sample becomes uniaxial with respect to the electric vector of the light, which we label $U$. On the molecular level, the transition moment direction for the transition responsible for the photoselection becomes the orientation axis, which we label molecular axis $z$. All directions perpendicular to $z$ will have the same $K$ value. We obtain a rod-like alignment of the photoselected molecules: $K_z = 3/5$; $K_y = K_x = 1/5$.

### Photoselection: negligible depletion; natural or circularly polarized light

The photoselected sample becomes uniaxial with respect to the direction of light propagation, which we label $U$. On the molecular level, the transition moment of the absorbing transitions now defines the axis that is least aligned with $U$, which we label $x$. We obtain a disc-like alignment of the photoselected molecules: $K_x = 1/5$; $K_y = K_z = 2/5$.

An interesting aspect of these orientation distributions obtained by photoselection is that not only the $K$ values needed for interpretation of absorption spectra but the complete orientation distribution function is known, including the $L$ values.

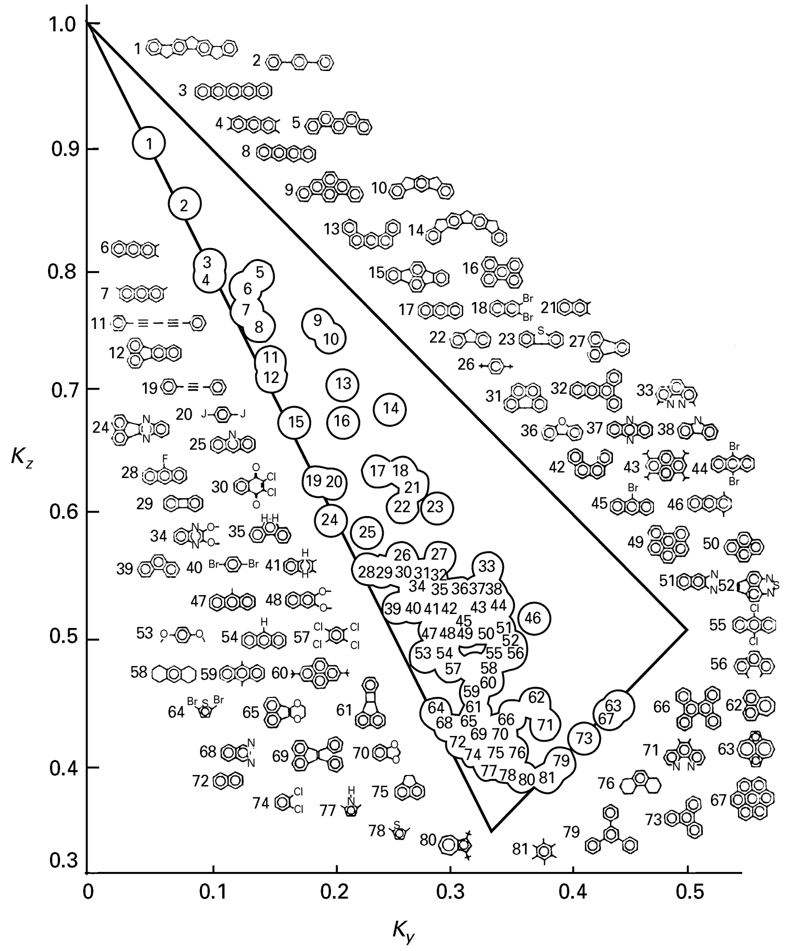



**Figure 2** Values of ($K_z$, $K_y$) for aromatic molecules in stretched polyethylene at 293 K. The molecular *z* axes are horizontal in the structures, the *y*-axes are vertical. Adapted by permission from Thulstrup EW and Michl J (1982) *Journal of the American Chemical Society* **104**: 5594–5604.

## Anisotropic (uniaxial) solvents

In the case of molecules aligned in stretched polymer sheets or nematic liquid crystals, the orientation distribution is unknown. Nevertheless, *K* values for many or all relevant transitions may be easily determined using the TEM method. For molecules with only two different *K* values, for example with a $C_3$ or higher axis, and in spectral regions where only two different *K* values exist, linear combinations of the two observed spectra may be formed that correspond to sums of contributions from transitions with the same *K* value.

One important example is that of π–π* transitions in a molecule of $C_{2v}$ or $D_{2h}$ symmetry. The transition moments for these are restricted to two different, perpendicular directions in the molecule; we shall label these directions *y* and *z*. The sum of contributions from *z*-polarized transitions is

$$A_z(\nu) = [(1 - K_y)E_U(\nu) - 2K_yE_V(\nu)]/(K_z - K_y)$$

while the sum of contributions from *y*-polarized transitions is

$$A_y(\nu) = [2K_zE_V(\nu) - (1 - K_z)E_U(\nu)]/(K_z - K_y)$$

**Figure 3** illustrates the resulting curves for the π–π* transitions in pyrene. The separation of transitions into two groups may help identify excited-state symmetries, separate close-lying, differently polarized transitions and reveal hidden transitions. It thereby simplifies the assignment of both electronic and vibrational transitions considerably.

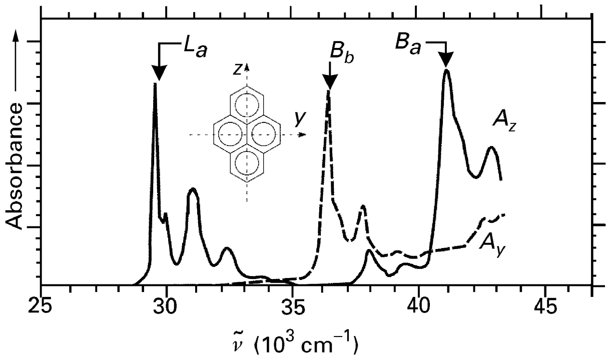



**Figure 3** Curves corresponding to purely *z*-polarized ($A_z$) and purely *y*-polarized absorbance ($A_y$) obtained by the TEM method as illustrated in **Figure 1**. Adapted by permission from Michl J and Thulstrup EW (1995) *Spectroscopy with Polarized Light: Solute Alignment by Photoselection, in Liquid, Polymers, and Membranes.* New York: Wiley-VCH.

If the molecule is of lower symmetry, very useful, but less specific, results can still be obtained. Let us assume that the molecule has a symmetry plane as its only symmetry element and that the out-of-plane direction, $x$, corresponds to the lowest possible $K$ value, $K_x$. Transitions must be polarized either along $x$ or in the $yz$ plane, perpendicular to $x$. We may therefore observe the following $K$ values: $K = K_x$ or $K_y \leq K \leq K_z$.

For a transition $f$ in the ($y$, $z$) plane, the angle between the molecular long axis, $z$, and the transition moment $M_{0f}$ for transition $f$, ($z$, $M_{0f}$), can now be determined from $K_y$, $K_z$ and $K_f$:

$$\tan^2(z, \boldsymbol{M}_{0f}) = \frac{K_z - K_f}{K_f - K_y} \qquad [13]$$

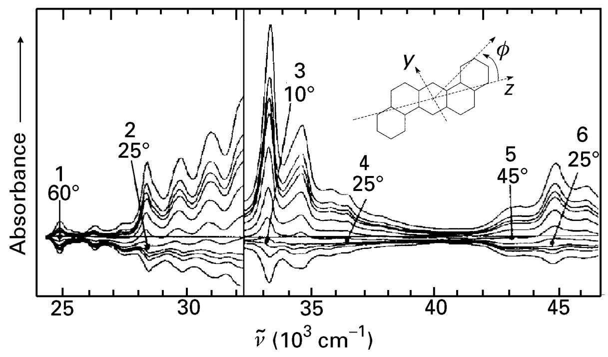



**Figure 4** UV spectra of dibenz[*a*,*h*] anthracene. An example of a planar molecule for which π–π* transition moments may be along any direction in the molecular plane. The curves are linear combinations of $E_U$ and $E_V$ and correspond to transitions that form angles of 0°, 15°, 30°, 45°, 60°, 75°, and 90° with the molecular *z* axis. Adapted by permission from Pedersen PB, Thulstrup EW and Michl J (1981) *Chemical Physics* **60**: 187–198.

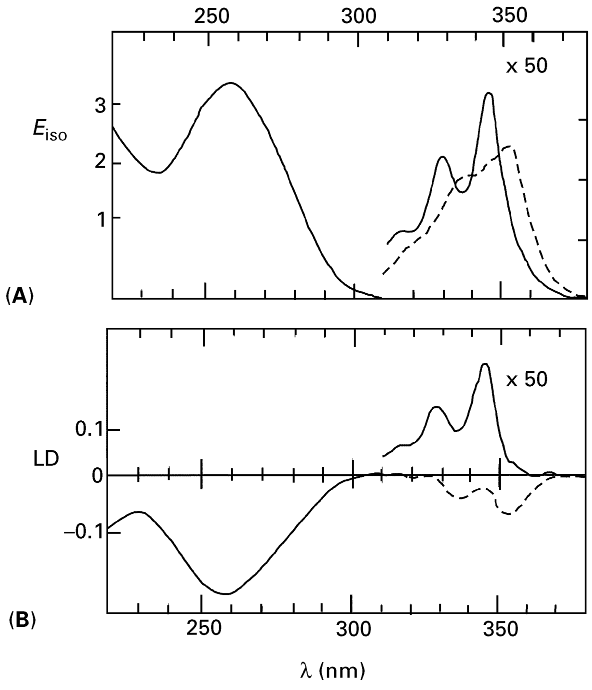



**Figure 5** UV absorption ($E_{iso}$) (A) and LD spectra (B) of two derivatives of benz[a]pyrene, covalently bound to DNA aligned in a flow. Both derivatives are dihydroxyepoxytetrahydrobenzpyrenes, but their two hydroxy groups are positioned on different sides of the molecular plane (*syn*, broken line; anti, full line). The absorption below 300 nm is mainly due to the DNA bases, the planes of which are predominantly perpendicular to the DNA helix axis. This explains the negative linear dichroism below 300 nm. The absorption above 300 nm is mainly due to the two derivatives; it is seen that *anti* has a positive LD, corresponding to a small angle between its molecular plane and the helix axis, while *syn* has a negative LD, corresponding to a large angle between its molecular plane and the helix axis. The biological difference between the two molecules is considerable; *syn* is a weak carcinogen, while *anti* is a very strong one. Adapted by permission from Tjerneld F (1982) Thesis, Chalmers University of Technology, Gothenburg, Sweden.

An example of the determination of such angles in a molecule with only the molecular plane as symmetry element is shown in **Figure 4.**

Macromolecules, especially DNA, that are difficult to dissolve in stretched polymer sheets or nematic liquid crystals have often successfully been aligned in a flow. Flow LD is measured in a solution between a stationary and a fast-rotating cylinder using a Couette or Maxwell cell. It has proved very useful for the study of the interaction between macromolecules and smaller molecules, for example the attachment of carcinogenics to DNA. The long axis of the DNA molecules, the helix axis, becomes the orientation axis. So far such studies have primarily been performed in the UV-visible region. The two

linearly independent spectra obtained are

$$\mathrm{LD}(\nu) = E_U(\nu) - E_V(\nu) \quad \text{and}$$
$$E_{\mathrm{iso}}(\nu) = [E_U(\nu) + 2E_V(\nu)]/3 \qquad [14]$$

The latter curve may be obtained when the flow is stopped. **Figure 5** shows $E_{\mathrm{iso}}(\nu)$ and LD(ν) for two metabolites of benzo[*a*]pyrene, of which one is a strong carcinogen, the other weak. The measurement was done in a Couette cell containing a mixture of a small amount of a metabolite with DNA in a weak solution of sodium cacodylate in water. The strong absorption of the DNA bases dominates the spectrum below 300 nm (with a negative LD because the DNA base absorption is polarized in the molecular planes of the bases, which are almost perpendicular to the helix axis).

The weak LD absorption of the metabolites correspond to an in-plane polarized transition; it can be studied above 300 nm. The stronger carcinogen of the metabolites has a positive LD, the other a negative. This means that the former is positioned with its molecular plane relatively parallel to the helix axis, while the molecular plane of the latter is more perpendicular to the helix axis.

## List of symbols

$A_f(\nu)$ = contribution to the absorbance of transition $f$ in a sample of perfectly aligned molecules; $E_U(\nu)$ = absorbance (at frequency ν) with electric vector in the specified direction ($U$, $V$, etc.); $|f\rangle$ = final state; $K_f$ = average cosine square value = $\langle \cos^2(U, M_{0f})$; $\hat{M}$ = electric dipole operator; $M_{0f}$ = electric dipole transition moment; $\hat{O}$ = generic operator; $u$ = a direction in the molecule; $U$ = laboratory axis, unique axis; $v$ = a direction in the molecule (= $x$, $y$, $z$); $V$ = laboratory axis perpendicular to $U$; $W$ = laboratory axis perpendicular to $U$ and $V$; $|0\rangle$ = initial state; ε = electric vector; ν = light frequency.

*See also:* **Laser Spectroscopy Theory; Liquid Crystals and Liquid Crystal Solutions Studied By NMR; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Nucleic Acids Studied Using NMR; Optical Spectroscopy, Linear Polarization Theory; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading

Kliger DS, Lewis JW, Randall CE (1990) *Polarized Light in Optics and Spectroscopy*. London: Academic Press.

Michl J and Thulstrup EW (1995) *Spectroscopy with Polarized Light: Solute Alignment by Photoselection, in Liquid Crystals, Polymers, and Membranes*. New York: Wiley-VCH.

Rodger A and Nordén B (1997) *Circular Dichroism and Linear Dichroism*, Vol. 106, Oxford Chemistry Masters. Oxford: Oxford University Press.

Samori' B and Thulstrup EW (eds) (1988) *Polarized Spectroscopy of Ordered Systems*, Vol. 242 of NATO ASI Series, Series C: Mathematical and Physical Sciences. Dordrecht: Kluwer Academic.

Thulstrup EW and Michl J (1989). *Elementary Polarization Spectroscopy*. New York: VCH.

# Linear Dichroism, Instrumentation

**Erik W Thulstrup** and **Jens Spanget-Larsen**,
Roskilde University, Denmark
**Jacek Waluk**, Polish Academy of Sciences, Warsaw,
Poland

Copyright © 1999 Academic Press.

ELECTRONIC SPECTROSCOPY
**Methods & Instrumentation**

All linear dichroism (LD) measurements are performed on anisotropic samples. Many samples occurring in nature are partially oriented from the start. Nevertheless, most LD experiments are performed on samples that have been aligned in the laboratory. Therefore, prior to learning the techniques of measurement, one should become familiar with different procedures used to align molecular ensembles.

## Methods of alignment

The majority of LD studies are done on partially aligned samples, which exhibit uniaxial orientation. Such samples have a unique direction (optical axis, unique sample axis) with respect to which the solute molecules are oriented. All directions perpendicular to this axis are equivalent. The most widespread method for obtaining partial alignment is the use of anisotropic solvents: stretched polymers or liquid crystals.

**Stretched polymers** This is the simplest and the least expensive method. Solutes may be introduced into the polymers by (i) swelling the polymer with a solution containing the substrate; (ii) diffusion of the vapour of the substrate into the polymer; (iii) casting a film of the polymer from a solution containing the substrate; or (iv) dissolving the substrate in a molten polymer. The most commonly used polymers are low-density polyethylene, best for nonpolar solutes, and poly(vinyl alcohol), used for polar and nonpolar solutes. Both are transparent in the UV-visible regions, and in some regions of the infrared. Mechanical stretching produces uniaxial orientation, its degree increasing with the stretching ratio. The mechanism of orientation is still under debate, but it is evident that the molecular shape is an important factor: the more elongated molecules align on the average better than those that have more spherical shape.

**Liquid crystals** These are experimentally more demanding than polymers. Usually, thermotropic liquid crystals are used, either nematic or compensated nematic mixtures, obtained from cholesteric components. Many liquid crystals absorb in the UV, but some are transparent down to about 200 nm. The alignment of the liquid crystal sample can be controlled by external electric or magnetic fields, or by cell surface pretreatment such as rubbing or coating. These procedures make it possible to switch the direction of the unique axis from perpendicular to the surface (homeotropic alignment) to parallel (homogeneous alignment). The degree of orientation of the solute depends on its nature and concentration, as well as on the nature and temperature of the sample and the imperfections of the cell surface.

Lyotropic liquid crystals, Langmuir-Blodgett films and even monolayers adsorbed on surfaces can also be used as uniaxial orienting media.

**Single crystals** In principle, this classical technique makes it possible to obtain nearly perfect orientation and absolute polarization assignments. However, both the measurement and the interpretation are often difficult. Very thin crystals are required for electronic absorption studies. Intermolecular interactions (Davydov splitting) often complicate the spectra. The solution is to find a single crystal of a suitable host molecule, transparent in the region of interest. These two conditions are often very difficult to meet simultaneously.

**Electric and magnetic fields** Linear dichroism produced as a result of applying an electric field is known as electric dichroism (Kerr effect). The effect is usually quite small for normal-sized molecules, although polymers such as DNA can be oriented very well. The orientation imposed by a magnetic field is even smaller, and therefore its practical use is limited to large molecular ensembles (Cotton–Mouton effect).

**Flow field** Macromolecules, such as DNA, can become orientated by the hydrodynamic shear associated with liquid flow. This is realized experimentally in a Maxwell cell (sometimes referred to as Couette cell). The cell consists of a fixed cylinder and a rotating cylinder: a flow gradient is produced in

the liquid solution contained in the annular gap between the two cylinders.

**Photoselection and photoorientation** This elegant method allows one to obtain a set of partially oriented molecules in a completely isotropic medium. It makes use of the fact that the photoexcited ensemble of molecules is always anisotropic, and thus exhibits LD (transient dichroism). Moreover, if the excitation is followed by chemical transformation, and the environment is rigid enough to prevent molecular rotation, a permanent alignment of both reactant and product is obtained. Such photooriented samples can be studied by conventional LD techniques. Particularly attractive media for use in photoorientation are low-temperature rare gas matrices, which are inert and transparent in both the UV-visible and IR regions.

## Techniques of measurement

For partially oriented uniaxial samples, two linearly independent spectra can be obtained from an experiment, irrespective of the mode of measurement and the polarization state of the absorbed light. The obvious procedure is to record two absorption spectra, one with the electric vector of light parallel to $U$, the unique axis of the sample ($E_U$), the other with the electric vector perpendicular to the unique axis ($E_V$). The spectra must be baseline-corrected by subtracting the curves recorded on pure solvents. This yields the largest difference between the tow linearly independent spectra. Other alternatives are also possible. A combination of $E_U$ or $E_V$ with the spectrum of an isotropic sample, $E_{\text{iso}}$ ($=[E_U + 2E_V]/3$) is easily obtained in liquid crystal. In cases where the sample geometry precludes the measurement of both $E_U$ and the isotropic spectrum (membranes), the two spectra are $E_V$ and $E_\omega$, the latter recorded with the incidence angle $\pi/2 - \omega$, and with the plane of polarization containing the $U$ axis (the 'tilted plate method'). Another possibility is to combine either $E_U$ or $E_V$ with the unpolarized spectrum measured with a depolarizer in front of the sample and the light beam perpendicular to the $U$ axis. Finally, one can use procedures that do not require the use of a polarizer at all, such as recording first the $E_V$ spectrum for an aligned liquid solution (with light propagating along the $U$ axis), and then $E_{\text{iso}}$ for the same solution with temperature raised above the isotropic melting point.

All these measurements can be performed on a commercial double–beam spectrophotometers supplemented with one or two linear polarizers (**Figure 1**). Calcite Glan prisms are used for measurements down to about 220 nm. Sheet polarizers can tolerate larger apertures and are much cheaper. On the other hand, their transmission is lower and the wavelength range usually limited to the visible range, although the UV-transmitting polarizing sheets have become available.

Various potential sources of errors in LD measurements are: (i) scattering and depolarization by the sample, particularly significant in the UV region; (ii) sample inhomogeneity and irreproducibility, which often makes it difficult to reliably record the baseline; (iii) sample birefringence; and (iv) polarization bias of the instrument. For weakly dichroic samples, the separate recording of two spectra is not practical, and the LD has to be measured directly. This is achieved by modulation methods. The state of polarization of the light beam is modulated in time, using a combination of a linear polarizer and, most often, a photoelastic modulator. Since the transmission of the dichroic sample varies for different polarizations of the light, the detector produces an a.c. signal, superimposed on the d.c. signal, owing to the average intensity of the transmitted beam. The analysis of the two signals allows an evaluation of the LD and of the average optical density of the sample. For measurements of this type, one can adapt a spectropolarimeter normally used for circular dichroism measurements (**Figure 2**).

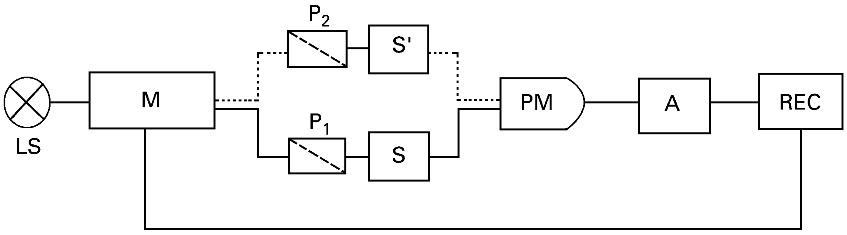



**Figure 1** Schematic diagram of a double-beam spectrophotometer for polarized spectroscopy. LS, light source; M, monochromator; $P_1$, $P_2$, polarizers; S, sample; S′, reference sample (optional); PM, photomultiplier; A, amplifier; REC, recorder. Adapted by permission from Michl J and Thulstrup EW (1995) *Spectroscopy with Polarized Light: Solute Alignment by Photoselection, in Liquid, Polymers, and Membranes*. New York: Wiley-VCH.

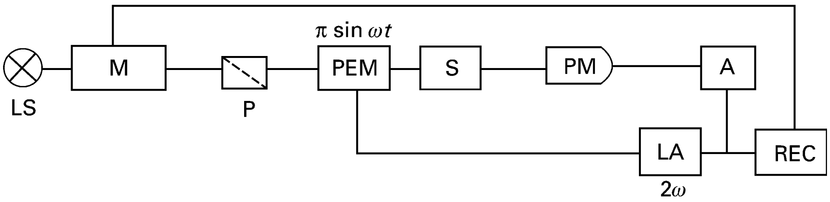



**Figure 2** Schematic diagram of a spectropolarimeter adapted for linear dichroism measurements. LS, light source; M, monochromator; P, polarizer; PEM, photoelastic modulator producing time-dependent retardation $\pi \sin \omega t$; S, sample; PM, photomultiplier; A, d.c. amplifier; LA, lock-in amplifier operating at frequency $2\omega$; REC, recorder. Adapted by permission from Michl J and Thulstrup EW (1995) *Spectroscopy with Polarized Light: Solute Alignment by Photoselection, in Liquid, Polymers, and Membranes.* New York: Wiley-VCH.

Although most LD studies are performed by recording the intensity of transmitted light, other methods are also used. One of them detects the LD via fluorescence, the other by photoacoustic spectroscopy. Both methods are very sensitive and avoid some of the errors inherent in the transition mode, such as light scattering.

Most of the alignment techniques and methods of measurements used for electronic linear dichroism can be adapted for studies in the IR region. The merits of IR-LD make this technique a very attractive choice, complementing and in many respects surpassing the informational content provided by electronic linear dichroism.

## List of symbols

$E_{iso}$ = absorbance in an isotropic sample; $E_U$ = absorbance with the electric field vector parallel to the unique axis, $U$; $E_V$ = absorbance with the electric field vector perpendicular to the unique axis (parallel to the perpendicular axis, $V$); $U$ = the unique optical axis; $\omega$ = angle complementary to the incidence angle in the 'tilted plate method'.

*See also:* **Linear Dichroism, Theory; Linear Dichroism, Applications; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Nucleic Acids Studied Using NMR; Vibrational CD, Applications; Vibrational CD Spectrometers; Vibrational CD, Theory.**

## Further reading

Kliger DS, Lewis JW and Randall CE (1990) *Polarized Light in Optics and Spectroscopy*. San Diego: Academic Press.

Michl J and Thulstrup EW (1995) *Spectroscopy with Polarized Light. Sample Alignment by Photoselection, in Liquid Crystal, Polymers, and Membranes*. New York: Wiley-VCH.

Rodger A and Nordén B (1997) *Circular Dichroism and Linear Dichroism*, Vol. 106. Oxford Chemistry Masters. Oxford: Oxford University Press.

Samori B and Thulstrup EW (1988) *Polarized Spectroscopy of Ordered Systems*, Vol. 242 of NATO ASI Series, Series C: Mathematical and Physical Sciences. Dordrecht: Kluwer Academic.

Thulstrup EW and Michl J (1989) *Elementary Polarization Spectroscopy*. New York: VCH.

# Linear Dichroism Theory

*See* **Optical Spectroscopy, Linear Polarization Theory.**

# Liquid Crystals and Liquid Crystal Solutions Studied By NMR

**Lucia Calucci** and **Carlo Alberto Veracini**,
University of Pisa, Italy

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

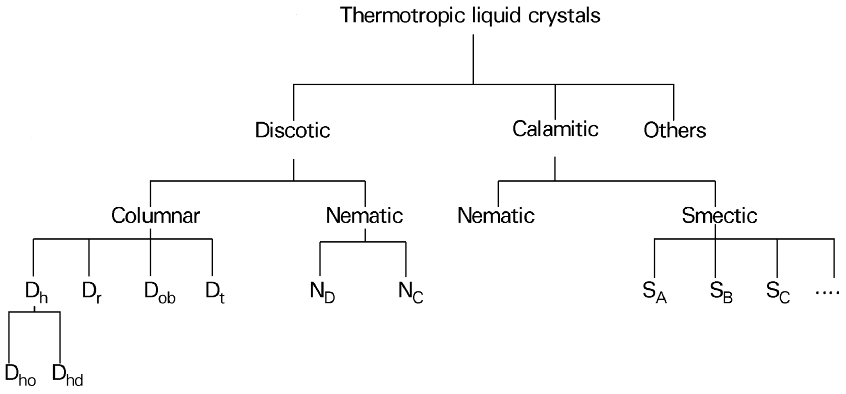
Since the first publications on this subject in 1963, NMR in liquid crystalline systems has been a wide and active field of research in many branches of organic and physical chemistry. In fact, NMR spectroscopy has revealed a powerful means of probing molecular structure, anisotropic magnetic parameters and dynamic behaviour of solute molecules dissolved in liquid crystals. Moreover, this technique has been successfully employed to investigate properties of mesophases themselves, such as their orientational ordering, translational and rotational diffusion and their effects on nuclear relaxation, and molecular organization in different liquid crystalline phases.

NMR of different nuclei has been performed: $^{1}H$, $^{2}H$ and $^{13}C$ NMR have been used extensively, but a number of studies are reported on $^{14}N$, $^{19}F$, $^{23}Na$ and $^{31}P$ also. Recently NMR of active isotopes of noble gases dissolved in liquid crystals has also been employed.

## Liquid crystals

We define a liquid crystal as a state of aggregation that is intermediate between the crystalline solid and the amorphous liquid. In this state the molecules show some degree of orientational order with respect to a preferred direction, *n*, called the director; long-range positional order does not exist or is limited to one or two dimensions. In several liquid crystal phases, termed uniaxial phases, only one director is required to describe the molecular order. In contrast, for biaxial phases the orientational order needs two directors for its description. The principal molecular axes are uniformly aligned only on average. The degree to which the molecules are ordered about the director is described by means of a number of parameters, called order parameters.

Substances that show a liquid crystalline phase, or mesophase, are called mesogens. Several thousands of compounds, both with low molecular mass and polymeric, are now known to form mesophases. They are mainly highly geometrically anisotropic in shape, rodlike or disclike (hence the terms calamitic and discotic liquid crystals), or they are anisotropic in solubility properties, like amphiphilic molecules and, depending on their detailed molecular structure, they can exhibit one or more mesophases between the crystalline solid and the isotropic liquid. Transitions to these intermediate states may be induced by purely thermal processes (thermotropic liquid crystals) or by the action of solvents (lyotropic liquid crystals). Each of these two categories can be further divided according to the structure of the mesophases and/or molecules; **Scheme 1** shows the classification of thermotropic mesophases.




**Scheme 1** Classification of thermotropic liquid crystals.

### Macroscopic alignment

When subjected to an external magnetic field $B_0$ of sufficient magnitude (> 0.5 T) the molecules in a mesophase exhibit a macroscopic orientation so that the director *n* is uniformly aligned in the whole sample. The magnetic field exerts a torque on the molecules, which reorientate cooperatively; the induced reduction in free energy is $\Delta G = \Delta\chi\ B_0^2$ (3 $\cos^2\alpha$–1)/6, where $\Delta\chi = \chi_{\|} - \chi_{\perp}$ is the difference between the component of the diamagnetic bulk susceptibility along ($\chi_{\|}$) and perpendicular to ($\chi_{\perp}$) the director and $\alpha$ is the angle between the director and the magnetic field. The direction of minimal *G*, i.e. the preferred orientation of *n*, depends on the sign of $\Delta\chi$. Thus, for calamitic mesogens with aromatic cores, which have $\Delta\chi > 0$, *n* aligns parallel to $B_0$, while for those formed uniquely by aliphatic units $\Delta\chi < 0$ and *n* aligns perpendicularly to $B_0$. Most discotic liquid crystals have $\Delta\chi < 0$, and thus an orientation of *n* perpendicular to $B_0$ is observed.

The response time for the alignment depends on the relative values of the magnetic torque experienced by the molecules and of the opposing viscous torque. For most nematic liquid crystals it takes only a few seconds; for some lyotropic phases the alignment is relatively slow and takes minutes or hours. Smectic and columnar discotic liquid crystals are sometimes very viscous and a uniform alignment can only be obtained by reaching these mesophases on slow cooling from the nematic or isotropic phases. For cholesteric samples, the presence of magnetic fields can untwist the helical structure, giving rise to a nematic phase.

Molecules dissolved in liquid crystal solvents become partially oriented, the degree of orientation being dependent on the nature and concentration of the solute.

## Observable interactions

The NMR spectra of molecules in a liquid crystalline environment, either mesogenic molecules or solutes, are dominated by second-rank tensorial properties (**T**) such as the shielding tensor ($\sigma_i$), the dipolar and indirect spin–spin interactions ($\mathbf{D}_{ij}$ and $\mathbf{J}_{ij}$) and, for nuclei with $I$>1/2, the quadrupolar interaction ($\mathbf{q}_i$). In most liquid crystals, molecular reorientations and internal motions reduce these interactions to a well-defined average, without the loss of all information. Moreover, molecular translational diffusion averages out to zero the intermolecular interactions, which thus do not complicate the partially averaged spectra. In the usual high-field approximation, in NMR experiments we measure $T_{zz}$, the component of the averaged interaction tensors resolved along the magnetic field direction (*z* in the laboratory frame).

The observed averaged quantities are related to the interactions in their principal axis systems (PAS), fixed on the molecules or on rigid molecular fragments, by several transformations of reference frames. Since the molecules are not fixed in a lattice, another frame, the liquid crystal (director) frame (LC), has to be defined. The transformation between the molecular and the LC frame is time dependent because of molecular reorientational motions and the use of the order parameters is needed in this step. Thus $T_{zz}$ relates the NMR measured quantity to the molecular ordering in the mesophases by a number of order parameters depending on the symmetry of the molecule and phase. Measurements of partially averaged second-rank tensors give access only to second-rank order parameters, which are equal to 0 for completely disordered systems, and range from –0.5 to 1 for partially ordered systems, according to $\langle 3l_\alpha l_\beta - 1\rangle/2$, where $l_\alpha$ and $l_\beta$ are the direction cosines between the PAS and the LC frames. The last step, from the LC frame to the laboratory frame, requires integration over all possible orientations in the case of a powder sample, or the use of a distribution around some preferred orientation in the case of an aligned sample.

### Chemical shift anisotropy

NMR spectroscopy of oriented molecules provides a simple method for measuring chemical shift anisotropies, $\Delta\sigma_i$. The change in chemical shift with respect to a reference signal as a function of the molecular degree of order, which can be varied by changing phase and temperature, determines the elements of the shift tensor. In practice, however, the direct measurement of $\Delta\sigma_i$ poses problems related to the reference signal, which depends on changes in local and bulk effects produced by variable sample temperature, phase transitions and degree of alignment with temperature. The variable angle spinning (VAS) method can be used to avoid some of these problems.

### Dipolar and indirect couplings

The direct dipolar coupling, $D_{ijzz}$, and $J_{ij}^{\text{aniso}}$, the anisotropic component of the indirect spin–spin interaction in the direction of the magnetic field, appear in the Hamiltonian in the same form. The splittings that arise in NMR spectra are the sum of these two quantities, called the experimental anisotropic coupling ($T_{ij} = 2D_{ijzz} + J_{ij}^{\text{aniso}}$). It is therefore important to know experimentally, or to evaluate theoretically, the values of the pseudo-dipolar coupling $J_{ij}^{\text{aniso}}$ in order to accurately determine $D_{ijzz}$, which is related

to the order parameter of the direction joining the two nuclei and to their distance, i.e. to molecular structure. On the other hand, the magnitude of $J_{ij}^{\mathrm{aniso}}$, and in particular the components of $J_{ij}$ expressed in a molecular axis system, are related to the molecular electronic structure. The measurement of a finite value of $J_{ij}^{\mathrm{aniso}}$ is not an unambiguous process, as it does not have a unique effect on the spectrum. Attempts to obtain values of $J_{ij}^{\mathrm{aniso}}$ for different pairs of nuclei have been made, and it has been concluded that some indirect couplings such as $^1$H–$^1$H, $^1$H–$^{13}$C, $^{15}$N–$^1$H have small anisotropic contributions, while other couplings such as $^{13}$C–$^{13}$C, $^{19}$F–$^{19}$F, $^1$H–$^{19}$F and $^1$H–X (where X is Hg, Cd, Si and Sn) may have a significant contribution from the anisotropic term.

### Quadrupolar interaction

Nuclei with spin $I > 1/2$ possess electric quadrupole moments which, in the presence of an electric field gradient, give an additional contribution to the spin Hamiltonian, termed the quadrupolar interaction. In anisotropic media this interaction produces line broadening and gives rise to additional lines in the spectrum. $^2$H has a relatively small quadrupolar moment and is therefore particularly useful for studying anisotropic media. In an X–D bond the field gradients arise almost entirely from the charge on X, and it can be usually assumed that the X–D bond lies parallel to the principal axis for $\mathbf{q}_i$, the quadrupolar interaction tensor. Furthermore, for C–D bonds the quadrupolar coupling constants are found to have typical values depending on hybridization (see **Table 1**).

## NMR of liquid crystalline solutions

### Structural studies of rigid molecules

$^1$H NMR spectroscopy in oriented media has been widely employed for determining the structure of rigid molecules. Proton spectra in nematic solvents are dominated by homonuclear dipolar couplings, $D_{ijzz}$, whose magnitudes are large (up to 3–5 kHz) compared to the line width and the chemical shift differences and which are observable between all active nuclei in a molecule. The spectra can be very complex (see **Figure 1**) even for molecules with a small number of nuclei and various strategies have been devised for their analysis.

For simple spin systems, analytical solutions to the Hamiltonian can be found and transition frequencies and intensities can be written in closed forms. Spectra of complex systems may be simplified by reducing the number of nuclei contributing to them. This may be achieved by partial deuteration of molecules followed either by spin decoupling or spin-echo refocusing. Alternatively, simplified spectra may be obtained by VAS or multiple-quantum experiments. Complex spectra have been analysed by iterative and automatic methods using computer programs; both require good starting parameters, which can be obtained from simplified spectra, and an experienced operator.

Recently, methods for automatic analysis have been improved allowing the treatment of very complex systems; with these procedures it is possible to analyse spectra that depend on up to 27 spectral parameters.

The dipolar couplings obtained from spectral analysis are then used to determine molecular geometries, namely bond lengths and angles. Since these couplings depend on both geometrical factors and order parameters, they never define an absolute length but only relative distances. In contrast, angles can be measured absolutely.

Only for very simple molecular systems with favourable symmetry are there enough dipolar couplings for the determination of both the structural and the order parameters. In more complex cases, the solution must be found iteratively by computer

**Table 1** Quadrupolar coupling constants for deuterium

| *C–D bond hybridization* | *Quadrupolar coupling constant (kHz)* |
|---|---|
| sp | 200 ± 5 |
| sp$^2$ | 185 ± 5 |
| sp$^3$ | 170 ± 5 |

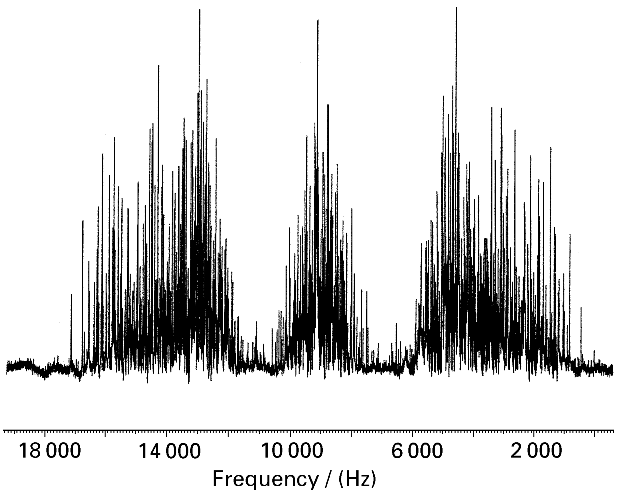



**Figure 1** 300 MHz proton spectrum of a sample of biphenyl dissolved in the nematic solvent ZLI1115.

calculations, and appropriate programs, such as SHAPE, have been written for this purpose. To get a unique and precise geometrical solution in the case of complex systems, it is highly recommended to record several spectra of the same molecule in different solvents. In addition, information on the carbon positions may be derived from the dipolar coupling constants $D$ ($^{13}$C–H) observed as the $^{13}$C satellites in the proton spectra, even though this method has two considerable disadvantages: the spectra are quite complex and the measured $^{13}$C–H couplings are usually less precise than H–H couplings.

The presence of molecular vibrations makes internuclear distances and angles, and therefore the degree of ordering, time dependent. Thus the NMR structural data must be corrected for vibration. The harmonic correction can be computed relatively easily and a computer program, VIBR, is available; the structure corrected for these effects, called the $r_\alpha$ structure, has the advantage of being unaffected by shrinkage effects, and can be compared with the data obtained from different spectroscopies.

The determination of molecular structure by $^1$H NMR spectroscopy in oriented media is apparently extremely precise: dipolar couplings of the order of $10^3$ Hz can be measured with an error of $10^{-1}$ Hz, thus giving a precision of $10^{-5}$ in distance ratios. In practice, however, limitations to structure precision apply because of solvent effects, vibrational corrections and anisotropy of the indirect coupling constants, and the precision that can be easily reached is of the order of $10^{-2}$ Å in bond lengths and $10^{-1}$ degrees in angles. As far as solvent effects on the structure are concerned, correlations of molecular shape with angular orientation have been observed. For example, they are responsible for the observation of dipolar and quadrupolar splittings in NMR spectra of molecules that in their equilibrium conformation possess tetrahedral and higher symmetry. The use of ‘magic’ mixtures has been suggested in order to obtain minimal distortions. Theoretical evaluations of the deformations have been made and used to correct the NMR couplings in structural studies.

$^2$H NMR has been used to determine quadrupolar coupling constants and asymmetry parameters, once the molecular structure and orientational order parameters are known.

NMR in nematic solutions has also been applied in the study of weak molecular complexes. The interactions of systems such as pyridine, pyrimidine and quinoline with iodine have been investigated. $^1$H and $^{13}$C NMR spectra have been recorded and analysed in order to detect whether the complex formation measurably affects the molecular structure, i.e. the proton and carbon positions. Since these effects are small, vibrational as well as deformation corrections have been performed.

### Structural studies of flexible molecules

The presence of large, low-frequency intramolecular motions (such as internal rotation and ring puckering) considerably complicates the interpretation of proton NMR spectra of flexible molecules in nematic solvents. In fact, in this case the effects of two simultaneous motional averagings on the magnetic interactions are present: the internal molecular motions and the overall motions of molecules in the mean potential due to the anisotropic environment. Moreover, owing to possible correlation between these motions, it is not possible to reach the same degree of precision on the orientational and geometrical parameters as can be obtained for rigid molecules.

Information about structure and orientational ordering can be obtained for individual rigid molecular fragments. Problems arise about the relationship between the data relating to different fragments and the interpretation of inter-fragment couplings. In early studies, several approximations were made, case by case, and their degree of soundness was tested by comparison of the results with those from other techniques. Later, systematic approaches were introduced for the analysis of dipolar data from flexible molecules; among which the additive potential (AP) method and the maximum entropy (ME) method have been found to be particularly useful.

### Chiral recognition

A simple and convenient method for enantiomeric analysis by means of NMR has been proposed that consists of using a chiral lyotropic liquid crystal obtained by mixing poly(γ-benzyl-L-glutamate) (PBLG) and various organic solvents. (*R*) and (*S*) enantiomers interact with the chiral centres of PBLG and thus orient differently in the liquid crystal solvent, which implies that all the order-dependent NMR interactions are different for the two enantiomers. $^2$H NMR has been applied successfully to a large number of partially deuteriated chiral molecules bearing various functional groups: enantiomeric excess (ee) up to 98% has been measured by signal integration. In many cases, chemical shift anisotropy of $^{13}$C in natural abundance may be used to discriminate enantiomers and to measure ee through conventional integration with an accuracy of about 5%. In this case, isotopic enrichment is not needed and more sites in the same molecule may be observed in the same spectrum. On the other hand, chemical shift differences between enantiomers can be very small (a few hertz) and, in particular for hydrocarbons, the method may not work. At

present the chiral recognition mechanism in PBLG is not fully understood so that it is not possible to assign absolute configurations.

### Probe studies of liquid crystals

A well-developed approach to the study of structure and dynamics of liquid crystals is based on the use of probes, although the probe can disturb the mesophase. Generally, highly symmetric and rigid molecules are dissolved in mesophases, but, for special purposes, different probes can be chosen; for instance, for lyotropic systems the solvent itself can be used as a probe of the behaviour of aggregates.

The order and dynamic behaviour of probes, revealed by suitable NMR techniques, have been used to study phase transitions and pretransitional phenomena. They are also useful for revealing segregation effects: when highly ordered smectic or columnar discotic phases organize themselves, they may exclude added solutes or confine them in aliphatic or interlayer regions.

The anisotropic potential acting on a probe is a solute–solvent property and therefore the study of probes gives valuable insight into the mechanism governing solute–solvent interactions. It has been found that very often the probe and the solvent can interact more or less specifically and in some cases an exchange between two or more sites has been inferred from large effects on the apparent geometry of the dissolved molecule.

Noble gases have been used as probes and have proved to be very useful for obtaining information about physical properties of liquid crystals. In particular, the shielding of $^{129}$Xe is highly responsive to the structure and temperature of the environment and $^{129}$Xe NMR has been employed to monitor phase transitions of thermotropic and lyotropic liquid crystals and the formation of induced smectic phases. Quadrupolar noble gas nuclei are powerful probes for detecting electric field gradients (EFG); $^{131}$Xe, $^{83}$Kr and $^{21}$Ne quadrupolar splittings and relaxation behaviour in various calamitic liquid crystals have given hints that there are two contributions to the EFG at the site of the nucleus: one is attributed to the distortion of the electron cloud of the atom from spherical symmetry and the other is the external EFG produced by the electric moments of the neighbouring molecules.

## NMR of liquid crystals

### Study of orientational order

Long-range orientational order is a fundamental character of liquid crystals. NMR is perhaps the most powerful technique for studying this property at the molecular level; its results are extremely useful for testing molecular theories that predict the variation of orientational order as a function of temperature, especially in uniaxial mesophases. The order parameters of different molecular segments in a liquid crystal can be obtained from $^{2}$H quadrupolar couplings, $^{2}$H–$^{2}$H, $^{2}$H–$^{1}$H and $^{13}$C–$^{1}$H dipolar couplings, $^{13}$C chemical shifts and, for fluorinated mesogens, from $^{13}$C–$^{19}$F dipolar couplings and $^{19}$F chemical shifts.

The study of deuterium interactions is now a general method in liquid crystal research and, although it requires use of specially deuteriated compounds, it has been widely applied to both thermotropic and lyotropic mesogens. $^{2}$H NMR spectra of aligned samples are well resolved and relatively simple (see **Figure 2**). An analysis of these spectra allows the magnitude of quadrupolar and, in some cases, dipolar splittings to be obtained, which are related to local order parameters for the $^{13}$C–$^{2}$H, $^{2}$H–$^{2}$H and $^{2}$H–$^{1}$H directions.

$^{13}$C–$^{1}$H dipolar couplings of a large number of calamitic liquid crystals have been determined by SLF/VAS, a 2D $^{13}$C NMR technique which combines separated local field spectroscopy (SLF) and variable angle spinning (VAS). In this method, mesogens with $^{13}$C in natural abundance are used; however, special software and hardware as well as prolonged spectrometer time are required. It has been shown that the order parameters of the phenyl rings and the aliphatic C–H bonds in liquid crystals, determined by the SLF/VAS technique, are linearly related to their $^{13}$C chemical shifts. The same linear correlation is applicable to different compounds in a homologous

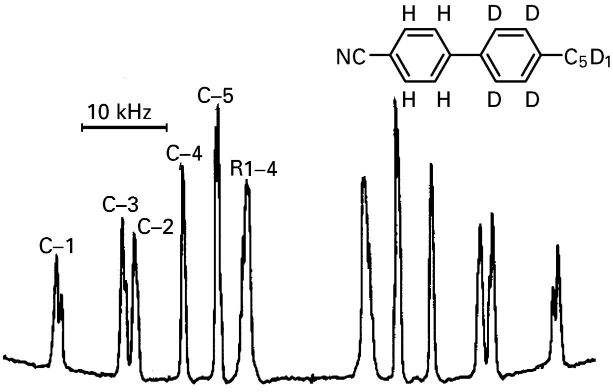



**Figure 2** 30.7 MHz $^{2}$H NMR spectrum of 5CB-$d_{15}$. The peaks are labelled according to their assignment to sites in the molecule. Reproduced with kind permission from Kluwer Academic Publishers from Emsley JW (1985) In Emsley JW (ed) *NMR of Liquid Crystals*, p 397. Dordrecht: Reidel.

series: this provides a convenient means for the determination of the order parameters of liquid crystals from their $^{13}C$ chemical shift data, without the need to know the chemical shift tensors.

The local order parameters obtained by the various methods can be used to describe the orientational ordering of molecules, or fragments of molecules, through reference system transformations that require assumptions about geometry (i.e. bond length and angles) and conformational distributions. Usually, the order parameters for the aromatic core of mesogens can be determined quite accurately and can be assumed as a measure of the molecular ordering in the mesophases: values up to 0.6–0.7 are found for nematic phases, while values up to 0.9–1.0 are typical of highly ordered smectic and columnar discotic phases.

In fortunate cases the measured interactions also allow the calculation of structural parameters of molecular fragments, the structure of mesogenic molecules being too complex for a complete determination.

As far as side chains are concerned, the local order parameters relative to the various $CH_2$ and $CH_3$ groups give a measure of chain disorder, being dependent on chain flexibility and conformation–orientation correlations. Different theories of orientational ordering in flexible chains, such as the AP and the chord models, have been formulated and critically tested by comparison with experimental data (see **Figure 3**).

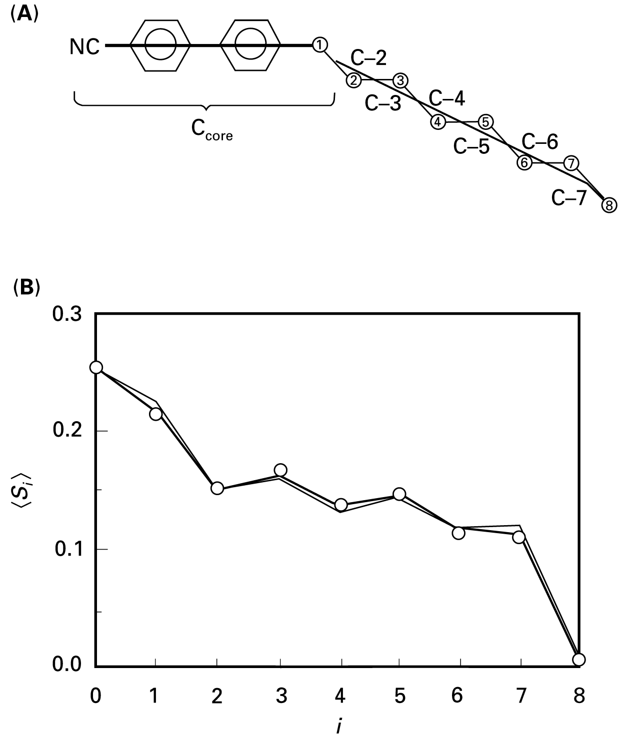



**Figure 3** (A) Molecular geometry and interacting modules of the 8CB molecule in the chord model; (B) orientational order parameter profile of 8CB at $T_c$–$T$ = 4 K. The open circles represent experimental values; the bold line gives the calculated profile using the chord model, while the fine line gives the calculated profile using the AP model. The order parameter of the cyanobiphenyl para-axis ($i$ = 0) is reduced by 50% for plotting convenience. Reproduced with permission of Springer-Verlag from Dong RY (1994) *NMR of Liquid Crystals*, p 107. New York: Springer-Verlag.

### Phase structure determination

NMR spectroscopy can provide valuable information on the structure of mesophases. The most commonly used method for this is the analysis of line shapes in spectra of various nuclei, and, in particular, of $^2H$. For a large number of mesophases, $^2H$ NMR spectra are motionally averaged and the spectral line shape reflects the phase symmetry. Line shape analysis of $^2H$ NMR spectra has been employed successfully to investigate the molecular organization and tilt angle of smectic C phases, as well as the structure of columnar discotic phases and it has been shown to be more sensitive than optical techniques to some aspect of biaxial ordering. Moreover, this technique has been shown to be extremely helpful in the discrimination of lyotropic phase symmetry. As examples, $^2H$ NMR spectra of some lyotropic phases are reported in **Figure 4**.

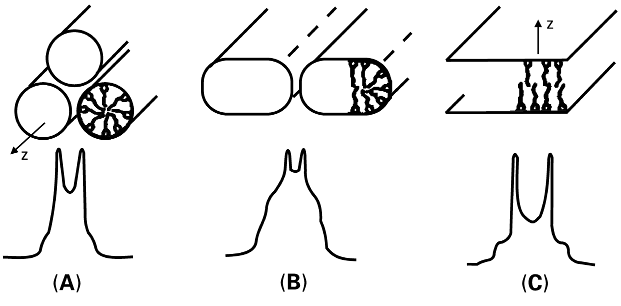



**Figure 4** Schematic view of the cylindrical (A), ribbon (B) and lamellar (C) lyotropic aggregates along with their corresponding $^2H$ NMR spectral patterns recorded for a potassium palmitate-$d_3$ in a mixture with 7 wt% potassium laurate and 30 wt% water. Reproduced with kind permission from Kluwer Academic Publishers from Doane JW (1985) In: Emsley JW (ed) *NMR of Liquid Crystals*, p 414. Dordrecht: Reidel.

Other nuclei can be useful in determining structure of lyotropic liquid crystals. In fact, $^{14}N$ and $^{31}P$ NMR have been applied successfully to investigation of the structure of phospholipid aggregates in water.

## Molecular dynamics

The complex dynamics of liquid crystals is characterized by a superposition of local and collective motions, comprising internal isomerization, overall rotational diffusion (rotation of the molecule about the long axis and reorientation of this axis) and translational diffusion, and collective order fluctuations. Different NMR techniques are designed to follow these motions and to differentiate the various motional modes on the basis of their timescale.

Overall molecular reorientations and internal motions take place in the $10^{-11}$–$10^{-6}$ s time window and information about them can be accessed using relaxation rate measurements. By far the best approach is to use $^2$H NMR experiments where deuterium Zeeman and quadrupolar spin–lattice relaxation times are measured on selectively deuteriated mesogens or on deuteriated probes dissolved in the mesophase. Frequency, orientation and temperature dependence of the spectral densities obtained in these relaxation studies give indications about the various motional modes and are extremely useful for testing orientational diffusion and chain dynamics models. Different deuterium relaxation experiments can be employed to extend the observable dynamic range: the quadrupolar echo sequence gives spin–spin relaxation times sensitive to motions in the range $10^{-8}$–$10^{-4}$ s, while extremely slow motions ($10^{-4}$–10 s) are accessible by the Carr–Purcell–Meiboom–Gill pulse train.

Internal and overall motions that take place on a dynamic scale from $10^{-7}$ to $10^{-4}$ s can affect the line shape of $^2$H NMR powder spectra. This is often the case for phenyl ring rotations or chain isomerizations in polymeric liquid crystals and for rotational diffusion in discotic mesophases. Line shape analysis allows different types of motions to be discriminated and the relative kinetic parameters to be obtained. Moreover multidimensional exchange NMR can be applied to characterize molecular details of dynamical processes when the correlation times range between $10^{-5}$ and $10^2$ s.

Fluctuations in the orientation of the director are better investigated by field cycling experiments that allow frequency-dependent $T_1$ measurements to be performed over a range from 100 Hz to 10 MHz. $^1$H and $^2$H spin relaxation studies have been carried out for numerous nematic and smectic liquid crystals to investigate the frequency dependence of director fluctuations, not accessible by experiments run on standard NMR spectrometers. It has been found that this motion is clearly observable by the relaxation rates only at low frequencies, i.e. below the MHz regime.

The translational diffusion motion is usually monitored on macroscopically aligned samples, by imposing a concentration or field gradient. The range of accessible diffusion rates is determined by the spectral or temporal resolution of the measuring technique. Moreover, translational diffusion constants may be determined indirectly by measuring spin–lattice relaxation times ($T_1$, $T_{1D}$, $T_{1\rho}$) and analysing them in terms of suitable theories.

$^1$H and $^2$H NMR studies of dynamic processes such as isomerization, ring inversion, hindered rotation and intramolecular rearrangement have also been performed on solute molecules in nematic solvents. In the case of proton NMR the dominant anisotropic splittings in the spectra come from the dipole–dipole interaction, and dynamic processes in the range of 10–$10^3$ s$^{-1}$ may be studied by line shape analysis. With $^2$H NMR the dominant quadrupolar interaction allows motions in the range of $10^3$–$10^9$ s$^{-1}$ to be investigated. Besides the considerably larger dynamic intervals that may be achieved, NMR in liquid crystalline solvents is superior to that in isotropic solvents in revealing the details of reaction mechanisms. In this respect 2D exchange deuteron NMR techniques have also been found helpful.

## List of symbols

$B_0$ = applied magnetic field; **D** = dipolar interaction tensor; ee = enantiomeric excess; $G$ = free energy; $I$ = nuclear spin; **J** = spin-spin interaction tensor; $l$ = direction cosine; $n$ = the director, preferred direction of orientation; **q** = quadrupolar interaction tensor; **T** = generic interaction tensor; $T_1$, $T_{1D}$, $T_{1\rho}$ = spin-lattice relaxation times; $T_{zz}$ = component of averaged interaction tensor resolved along the magnetic field ($z$) direction; $\sigma$ = shielding constant; $\chi$ = diamagnetic susceptibility.

*See also:* **$^{13}$C NMR, Methods; Chemical Exchange Effects in NMR; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Diffusion Studied Using NMR Spectroscopy; Enantiomeric Purity Studied Using NMR; $^{19}$F NMR, Applications, Solution State; High Resolution Solid State NMR, $^{13}$C; Nitrogen NMR; NMR Data Processing; NMR in Anisotropic Systems, Theory; NMR Pulse Sequences; NMR Relaxation Rates; $^{31}$P NMR; Rigid Solids Studied Using MRI; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Xenon NMR Spectroscopy.**

## Further reading

Chandrasekhar S (1992) *Liquid Crystals*, 2nd edn. Cambridge: Cambridge University Press.
Courtieu J, Bayle JP and Fung BM (1994) Variable angle sample spinning NMR in liquid crystals. *Progress in NMR Spectroscopy* **26**: 141-169
Diehl P and Khetrapal CL (1969) In: Diehl P, Fluck E and Kosfeld R (eds) NMR - *Basic Principles and Progress*, Vol. 1. Berlin: Springer-Verlag.
Doane JW (1979) NMR of liquid crystals. In: Owens FJ, Poole CP and Farach HA (eds) *Magnetic Resonance of Phase Transitions*. New York: Academic Press.
Dong RJ (1994) *Nuclear Magnetic Resonance of Liquid Crystals*. New York: Springer-Verlag.
Emsley JW (ed) (1985) *Nuclear Magnetic Resonance of Liquid Crystals*. Dordrecht: Reidel.
Emsley JW and Lindon JC (1975) *NMR Spectroscopy in Liquid Crystal Solvents*. Oxford: Pergamon Press.
Jokisaari J (1994) NMR of noble gases dissolved in isotropic and anisotropic liquids *Progress in NMR Spectroscopy* **26**: 1-26.
Luckhurst GR and Gray GW (eds) (1979) *The Molecular Physics of Liquid Crystals*. London: Academic Press.
Luckhurst GR and Veracini CA eds (1994) *The Molecular Dynamics of Liquid Crystals*. Dordrecht: Kluwer Academic.

# Lithium NMR Spectroscopy

*See* **NMR Spectroscopy of Alkali Metal Nuclei in Solution.**

# Luminescence Spectroscopy of Inorganic Condensed Matter

*See* **Inorganic Condensed Matter, Applications of Luminescence Spectroscopy.**

# Luminescence Theory

**Mohammad A Omary** and **Howard H Patterson**,
University of Maine, Orono, ME, USA

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| **Theory** |

## Introduction

Luminescence refers to the emission of light from an excited electronic state of a molecular species. In the case of photoluminescence, a molecule absorbs light of wavelength $\lambda_1$ decays to a lower energy excited electronic state, and then emits light of wavelength $\lambda_2$, as it radiatively decays to its ground electronic state. Generally, the wavelength of emission, $\lambda_2$, is longer than the excitation wavelength but in resonance emission $\lambda_1 = \lambda_2$. Luminescence bands can be either fluorescence or phosphorescence, depending on the average lifetime of the excited state which is much longer for phosphorescence than fluorescence. The relative broadness of the emission band is related to the relative difference in equilibrium

distance in the excited emitting state versus the ground electronic state.

Photoluminescence of a molecular species is different from emission of an atomic species. In the case of atomic emission, both the excitation and the emission occur at the same resonance wavelength. In contrast, excitation of a molecular species usually results in an emission that has a longer wavelength than the excitation wavelength.

If a chemical reaction results in the production of a molecular species in an excited electronic state that emits light, this phenomenon is termed chemiluminescence. Chemiluminescence usually occurs in the gas or liquid phase. In contrast, photoluminescence can occur in the gas, liquid or solid phases.

The next two sections provide a discussion of the basic principles of luminescence spectroscopy, which include the electronic transitions and the important parameters determined from luminescence measurements. Then follow two sections that describe the general characteristics of luminescence measurements and one which provides two case studies for organic and inorganic luminophores. The remainder of the article covers more specific topics and phenomena in luminescence spectroscopy, namely quenching, energy transfer, exciplexes and chemiluminescence. The examples in this article were selected to cover multidisciplinary areas of science.

## Electronic transitions and relaxation processes in molecular photoluminescence

### The Jablonski energy level diagram

The radiative and nonradiative transitions that lead to the observation of molecular photoluminescence are typically illustrated by an energy level diagram called the Jablonski diagram. **Figure 1** shows a Jablonski diagram that explains the mechanism of light emission is most organic and inorganic luminophores. The spin multiplicity of a given electronic state can be either a singlet (paired electrons) or a triplet (unpaired electrons). The ground electronic state is normally a singlet state and is designated as $S_0$ in **Figure 1**. Excited electronic states are either singlet ($S_1$, $S_2$) or triplet ($T_1$) states.

When the molecule absorbs light an electron is promoted within $10^{-14}$- $10^{-15}$ s from the ground electronic state to an excited state that should possess the same spin multiplicity as the ground state. This excludes a triplet excited state as the final state of electronic absorption because the selection rules for

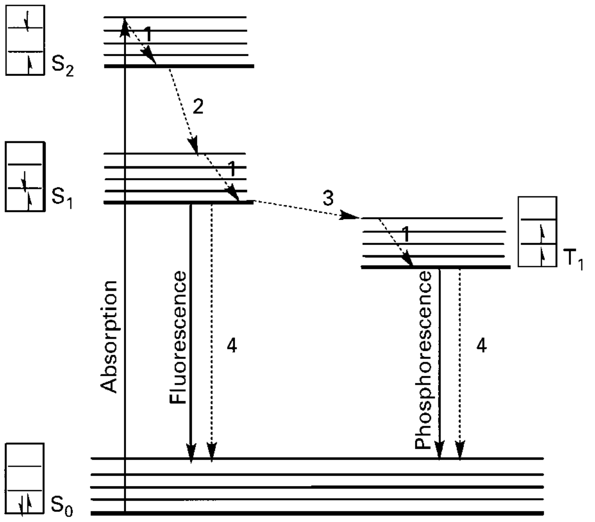



**Figure 1** The Jablonski diagram. Radiative and non-radiative processes are depicted as solid and dashed lines, respectively.

electronic transitions dictate that the spin state should be maintained upon excitation. A plethora of non-radiative and radiative processes usually occur following the absorption of light en route to the observation of molecular luminescence. The following is a description of the different types of non-radiative and radiative processes.

### Non-radiative relaxation processes

**Vibrational relaxation** Excitation usually occurs to a higher vibrational level of the target excited state (see below). Excited molecules normally relax rapidly to the lowest vibrational level of the excited electronic state. This non-radiative process is called 'vibrational relaxation'. Vibrational relaxation processes occur within $10^{-14}$- $10^{-12}$ s, a time much shorter than typical luminescence lifetimes. Therefore, such processes occur prior to luminescence.

**Internal conversion** If the molecule is excited to a higher-energy excited singlet state than $S_1$ (such as $S_2$ in **Figure 1**), a rapid non-radiative relaxation usually occurs to the lowest-energy singlet excited state ($S_1$). Relaxation processes between electronic states of like spin multiplicity such as $S_1$ and $S_2$ are called 'internal conversion'. These processes normally occur on a time scale of $10^{-12}$ s.

**Intersystem crossing** Non-radiative relaxation processes between different excited states are not limited to states with the same spin multiplicity. A process in which relaxation proceeds between

excited states of different spin multiplicity is called 'intersystem crossing'. The relaxation from $S_1$ to $T_1$ in **Figure 1** is an example of an intersystem crossing.

Intersystem crossing is generally a less probable process than internal conversion because the spin multiplicity is not conserved. Because of the lower probability for intersystem crossing processes they occur more slowly (~$10^{-8}$ s) than internal conversions. Intersystem crossing processes become more important in molecules containing heavy atoms such as iodine and bromine in organic luminophores and metal ions in inorganic luminophores (transition-metal complexes). Significant interaction between the spin angular momentum and the orbital angular momentum (spin- orbit coupling) becomes more important in the presence of heavy atoms and, consequently, a change in spin becomes more favourable. In solution, the presence of paramagnetic species such as molecular oxygen increases the probability of intersystem crossing. In transition metal complexes, intersystem crossing processes become increasingly important as one proceeds from complexes of the 3d-block to those of the 4d- and 5d-blocks.

**Non-radiative de-excitation** The excitation energy stored in the molecules following absorption must be dissipated owing to the law of conservation of energy. The aforementioned non-radiative processes occur very rapidly and only release very small amounts of energy. The rest of the stored energy will be dissipated either radiatively, by emission of photons (luminescence), or non-radiatively, by the release of thermal energy. The non-radiative decay of excitation energy which leads to the decay of the excited molecules to the ground electronic state is called 'non-radiative de-excitation'. These processes result in the release of infinitesimal amounts of heat that cannot normally be measured experimentally. The experimental evidence for non-radiative de-excitation processes is the quenching of luminescence. A major route for non-radiative de-excitation processes is energy transfer to the solvent or non-luminescent solutes in the solution. In the solid state, crystal vibrations (phonons) provide the mechanism for non-radiative de-excitation.

### Radiative processes: fluorescence and phosphorescence

The spin selection rule for electronic transitions (both absorption and emission) states that 'spin-allowed' transitions are those in which the spin multiplicity is the same for the initial and final electronic states. Therefore, spin-allowed transitions are more likely to take place than spin-forbidden transitions. 'Fluorescence' refers to the emission of light associated with a radiative transition from an excited electronic state that has the same spin multiplicity as the ground electronic state. Fluorescence is depicted by the radiative transition $S_1 \rightarrow S_0$ in **Figure 1**. Since fluorescence transitions are spin-allowed, they occur very rapidly and the average lifetimes of the excited states responsible for fluorescence are typically < $10^{-6}$ s.

Electronic transitions between states of different spin multiplicity are 'spin-forbidden' which means that they are less probable than spin-allowed transitions. However, spin-forbidden transitions become more probable when spin- orbit coupling increases. The factors that increase the probability of phosphorescence are the same factors discussed above that increase the probability of intersystem crossing. Therefore, if the triplet excited state is populated by intersystem crossing then luminescence might occur from the triplet state to the ground state. 'Phosphorescence' refers to the emission of light associated with a radiative transition from an excited electronic state that has a different spin multiplicity from that of the ground electronic state. Phosphorescence is depicted by the radiative transition $T_1 \rightarrow S_0$ in **Figure 1**. Since phosphorescence transitions are spin-forbidden, they occur slowly and the average lifetime of the excited states responsible for phosphorescence typically range from $10^{-6}$ s to several seconds. Therefore, some textbooks refer to phosphorescence as 'delayed fluorescence'. Photoluminescence refers to both fluorescence and phosphorescence.

## Excited-state distortions and the Franck–Condon principle

Luminescence spectroscopy can be used to gain information about the geometry of a molecule in an excited electronic state. Such information provides an understanding of the difference in the bonding properties of an excited state relative to the ground state. These properties can be better understood when the electronic transitions are discussed in the context of the potential surfaces of the ground and excited states.

Electronic absorption of light occurs within $10^{-15}$ s. Since this time is extremely short the nuclei are assumed to be 'frozen' during the time scale of absorption. Therefore, the transitions between various electronic levels are depicted as 'vertical' transitions in energy level diagrams. This assumption of negligible nuclear displacement during electronic transitions is known as the 'Franck- Condon

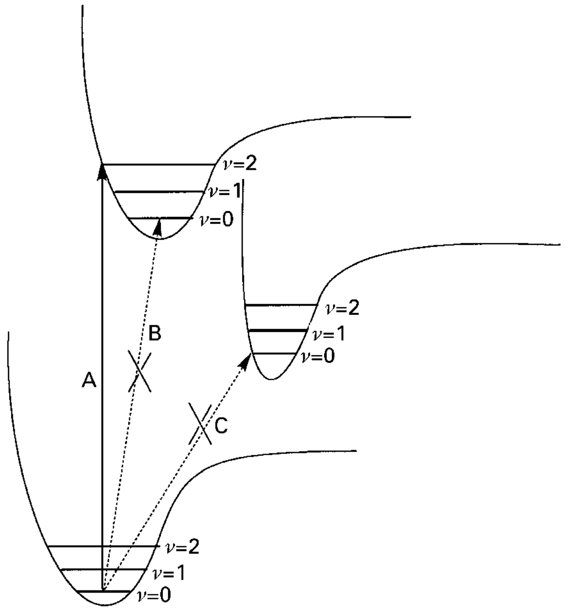



**Figure 2** The Franck–Condon principle. Only vertical electronic transitions are allowed.

principle'. **Figure 2** illustrates the Franck-Condon principle. Electronic excited states usually have different geometries than the ground state and, consequently, different equilibrium distances. Since electronic transitions are vertical, only transition A in **Figure 2** occurs. Transition C involves an excited state that is largely displaced from the ground state and thus no vertical transition is possible to this state. Transition B, on the other hand, terminates in the lowest vibrational level of the excited state. **Figure 2** shows that this transition cannot occur vertically, either. The three transitions depicted in **Figure 2** explain why in the Jablonski diagram (**Figure 1**): (i) the absorption was depicted to a higher vibrational level of the $S_2$ excited state than the v = 0 level, and (ii) no direct excitation to the triplet excited state ($T_1$) was depicted.

Usually, the nuclei of excited molecules are displaced from their ground state positions. The displacement is caused by differences in the bonding properties between the molecular orbitals that represent the ground and excited states, respectively. The extent of the nuclear displacement varies from one case to another. **Figure 3** illustrates two cases (A) and (B) of excited state distortion. The transitions depicted in **Figure 3** are called 'vibronic transitions'. A vibronic transition refers to a transition that involves a change in both electronic and vibrational states.

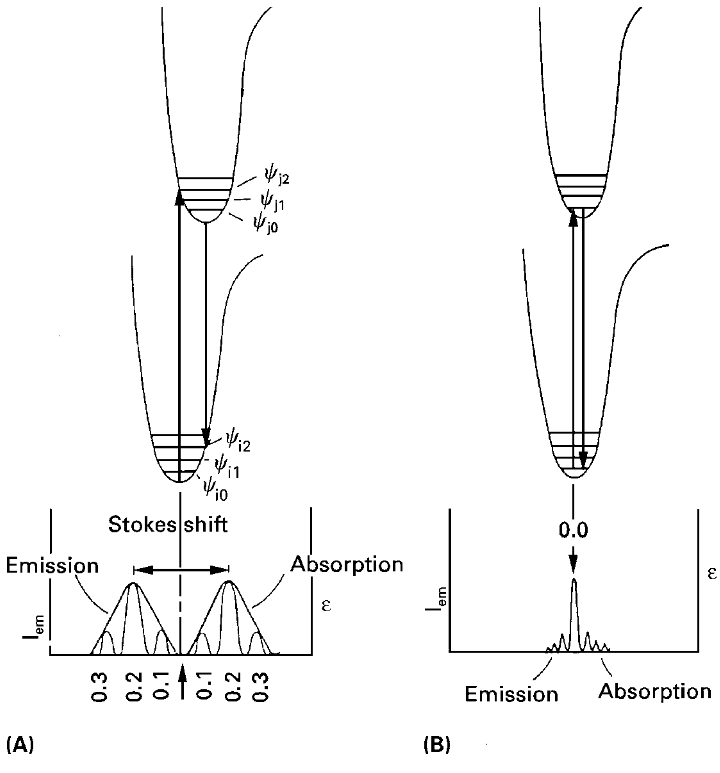



**Figure 3** Potential energy diagrams for two cases of excited state distortions: (A) non-zero distortion and (B) zero distortion. The corresponding absorption and emission spectra are shown below. Reproduced with permission from Adamson AW and Fleischauer PD (1975) *Concepts of Inorganic Photochemistry*, p 4. New York: Wiley-Interscience.

The case shown in **Figure 3B** represents an extreme case in which the excited state distortion is virtually zero. As a result, the electronic transition for which $\nu = 0$ in both the ground and excited states (called the 0-0 vibronic transition) has the greatest probability and, thus, the strongest intensity in both absorption and emission spectra. When excited state distortion is significant, as in **Figure 3A**, the 0-0 transition becomes much less probable and so its intensity becomes much weaker than other 0-$\nu$ vibronic transitions with higher $\nu$ values.

The probability (P) of the occurrence of a vibronic transition is proportional to the square of the Franck-Condon factor, which is defined as the overlap integral of the vibrational wavefunctions, $\Psi(\nu)$. Therefore,

$$P \propto |\langle \Psi(\nu_1)|\Psi(\nu_2)\rangle|^2 \qquad [1]$$

In many instances the absorption and emission maxima correspond to the same ($\nu_1$-$\nu_2$) vibrational pair. For example, **Figure 3A** shows that the 0-2 vibronic transition has the strongest intensity in both the absorption and the emission spectra. This is called the 'mirror image' rule and is followed by luminophores whose excited state distortion is zero or small. However, the mirror image rule may not apply for cases where large excited state distortion exists. Examples of each case are provided in the section dealing with organic luminophores.

Quantitative information about excited state distortion can be obtained from the luminescence spectra. For cases where vibronic structure is observed, the excited state distortion can be probed qualitatively simply by identifying the most intense vibronic transition: the larger the value of the vibrational level of the excited state, the larger the excited state distortion (as illustrated in **Figure 3**). Quantitatively, the displacement ($\Delta q$) of the excited state equilibrium distance from the corresponding ground state distance can be calculated. The procedure involves the calculation of the Franck-Condon factor as a function of $\Delta q$ for each vibronic transition. This results in a calculated emission spectrum, which can be compared with the experimental emission spectrum as a function of $\Delta q$ until a good fit is obtained. The actual $\Delta q$ would be the value that gives the best fit.

The emission (and absorption) spectra in many practical cases do not show resolved vibronic peaks. This is especially the case in solutions of the luminophores at ambient temperatures. The interaction of excited molecules of the luminophore with solvent molecules (especially polar solvents) is responsible for this broadening. There are methods to quantify $\Delta q$ in these cases with the aid of some computer programs but the description of these methods is beyond the scope of this article. In the section on inorganic exciplexes, nevertheless, we provide an example in which $\Delta q$ is evaluated theoretically for a system that exhibits structureless emission bands.

Another quantitative measure of excited state distortion is the Stokes shift, defined as the energy difference between the emission and absorption peak maxima for the same electronic transition. The lower part of **Figure 3A** illustrates how the Stokes shift is evaluated in typical cases. The Stokes shift of the rare case shown in **Figure 3B** is zero because the absorption and emission spectra have the same peak positions. The band width can also be used to quantify excited state distortions. The band width is normally quantified in terms of the full-width-at-half-maximum (fwhm). The value of fwhm is calculated as the energy difference between the band positions that have intensities equal to one half the peak maximum, assuming a gaussian shape is obtained for the emission band. In the absence of excited state distortions the emission bands appear as sharp peaks with very small fwhm values (**Figure 3B**), whereas in more common cases the electronic bands are much broader because of excited state distortions (**Figure 3A**). In those cases where the emission spectra are devoid of vibronic structure, the fwhm values are calculated for the whole broad band. The Stokes shift and fwhm are parameters for excited state distortion because their values are larger for 0-$\nu$ vibronic transitions with higher $\nu$ values. The Stokes shift and fwhm are especially useful in cases where the emission spectra are structureless, because of the difficulty of carrying out the $\Delta q$ calculation in such cases.

## Emission and excitation spectra

In luminescence spectroscopy the emission of the luminophore is monitored. There are two different types of luminescence spectra that can be recorded with modern spectrofluorometers, emission and excitation spectra. Note that although the name implies that these instruments measure only fluorescence, spectrofluorometers can also measure phosphorescence, especially when special accessories are added. Spectrofluorometers contain both an excitation monochromator and an emission monochromator. In emission spectra, the excitation wavelength is fixed and the emission monochromator is scanned. The excitation is usually fixed at a wavelength at which the sample has significant absorbance. In excitation spectra, on the other hand, the emission wavelength is fixed and the excitation monochromator is scanned. The emission is normally fixed at a

wavelength that corresponds to the emission peak of the sample.

Theoretically, the excitation spectra should mimic the absorption spectra of the luminophores. However, the lamp output is wavelength-dependent. For example, a common light source in modern spectrofluorometers is a xenon lamp. The output of this lamp is a continuum spectra in the range ~200- 1200 nm. The radiation curve approximates blackbody radiation with a maximum near 550 nm and a sharp decline at short wavelengths. A correction is, therefore, needed for the lamp background in order for the excitation spectra to correlate with the absorption spectra. Corrected excitation spectra are usually obtained by using a quantum counter, a strong luminophore capable of absorbing virtually all incident light over a wide range of wavelength. Rhodamine B is a commonly used quantum counter. Concentrated solutions of rhodamine B absorb virtually all the incident light in the 220- 600 nm range. The excitation spectrum of rhodamine B monitoring its emission maximum (633 nm) is determined only by the lamp output in the 220- 600 nm range (**Figure 4A**).

Corrected excitation spectra of luminophores are obtained by dividing the uncorrected spectra by the excitation spectrum of the quantum counter. A common improvement in the corrected excitation spectra versus the uncorrected spectra is to eliminate the sharp ‘spikes’ of the xenon lamp between 450 and 500 nm (**Figure 4A**). This makes the excitation spectra of many luminophores very similar to their absorption spectra. However, more dramatic differences may exist, especially for luminophores which absorb in the UV region. An example of such a luminophore is the silver complex $[Ag(CN)_2^-]$. **Figure 4B** shows the corrected and uncorrected excitation spectra of $[Ag(CN)_2^-]$ doped in NaCl crystals. Note that the corrected excitation spectra are significantly different from the uncorrected spectra. This difference is attributed to the low output of the xenon lamp at wavelengths shorter than 280 nm (**Figure 4A**), at which $[Ag(CN)_2^-]$ species absorb strongly.

Absorption and excitation spectra are complementary. The necessity to obtain a correction for the excitation spectra represents a disadvantage. There are advantages, however, for excitation spectra over absorption spectra. The much higher sensitivity of luminescence techniques compared to absorption techniques is an obvious advantage for excitation spectra. The greater sensitivity of luminescence techniques stems from the fact that the luminescence intensity can be enhanced by increasing the intensity of the excitation source, which is not the case in absorption. The greater sensitivity of luminescence techniques is, however, accompanied by less precision

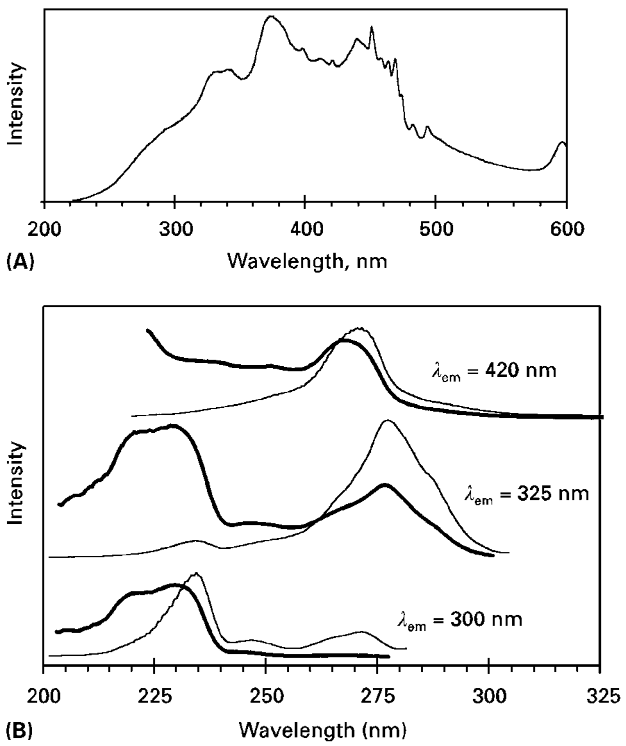



**Figure 4** Correction of the excitation spectra by the quantum counter method: (A) excitation spectrum of rhodamine B; (B) excitation spectra of $[Ag(CN)_2^-]$/NaCl doped crystals monitoring the emission at different wavelengths. Corrected and uncorrected spectra are shown as thick and thin lines, respectively.

compared to absorption techniques. The relative ease of acquiring excitation spectra for some materials such as solids represents another advantage over absorption measurements. Finally, excitation spectra can provide valuable information about excited state processes such as energy transfer (see below) that cannot be obtained by absorption spectra.

## Luminescence lifetimes

### Theory

Consider the simplest case in which a molecule A absorbs light and is in an excited electronic state denoted by *A. If the molecule has a single pathway for decay, say fluorescence, we can write the following equation:

$$^*\mathrm{A} \rightarrow \mathrm{A} + h\nu \qquad [2]$$

This is a first-order process whose rate can be expressed mathematically as:

$$-\mathrm{d}[^*\mathrm{A}]/\mathrm{d}t = k_\mathrm{f}[^*\mathrm{A}] \qquad [3]$$

where $k_f$ is the rate constant of the fluorescence decay (Eqn [2]). The reciprocal of $k_f$ is called the fluorescence lifetime $\tau_f$ ($\tau_f = 1/k_f$). Integration of Equation [3] gives:

$$[^*A] = [^*A]_o \exp(-t/\tau_f) \qquad [4]$$

Hence, a plot of ln [*A] versus time (t) should give a straight line with a slope of $1/\tau_f$. The value of [*A] is determined from the fluorescence intensity. Experimentally, lifetime measurements are obtained using a pulsed laser source. Pulsing leads to the population of the excited state of A, followed by emission of light by $^*A$ with a time profile according to Equation [4]. **Figure 5** shows a schematic description of a luminescence decay curve (A) and the plot used for the determination of the excited state lifetime (B).

Next, consider the case where $^*A$ can exhibit fluorescence (Eqn [2]) and also non-radiative decay:

$$^*A \rightarrow A + \text{vibrational energy} \qquad [5]$$

with a rate constant of $k_{nr}$. The decay of $^*A$ can now be described as:

$$-d[^*A]/dt = k_f[^*A] + k_{nr}[^*A] = (k_f + k_{nr})[^*A] \qquad [6]$$

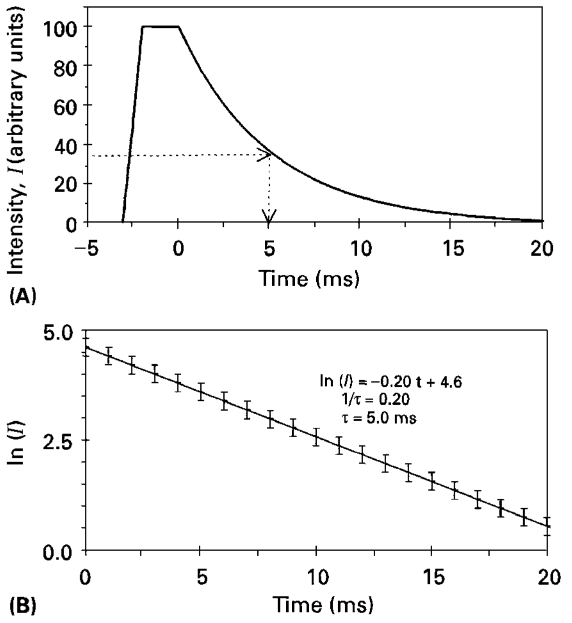



**Figure 5** Experimental determination of excited state lifetimes: (A) a plot of the intensity (*I*) versus time after the laser pulse. The lifetime corresponds to the time at which the intensity decays to 1/e of its maximum value; (B) a plot of ln (*I*) versus time after laser pulse. The lifetime here can be calculated directly from the slope of the linear equation.

In this case the lifetime $\tau$ becomes:

$$\tau = 1/(k_f + k_{nr}) \qquad [7]$$

or the reciprocal of the sum of $k_f$ and $k_{nr}$.

Generally, the non-radiative rate constant decreases at lower temperatures. As the temperature (T) approaches absolute zero, $\tau$ approaches $1/k_f$. A plot of $1/\tau$ versus T extrapolated to T = 0 K allows one to determine $k_f$ as the value of the intercept (see the example in the case study on inorganic luminophores). Therefore, the value of the fluorescence lifetime ($\tau_f$) is determined as $1/k_f$. Note that the treatment described in this section applies for both fluorescence and phosphorescence lifetimes and not just for fluorescence.

Now consider a final lifetime case in which two states are thermally populated at a temperature T. If excited state 1 has a lifetime $\tau_1$ and excited state 2 has a lifetime $\tau_2$, the observed lifetime $\tau_{obs}$ will be a weighted average of the two lifetimes:

$$\tau_{obs} = (n_1/N)\tau_1 + (n_2/N)\tau_2 \qquad [8]$$

with $N = n_1 + n_2$. Assuming that excited states 1 and 2 are non-degenerate and using the Boltzmann distribution function gives:

$$\tau_{obs} = \{[\tau_1 + \tau_2 \exp(-\Delta E/kT)]/[1 + \exp(-\Delta E/kT)]\} \qquad [9]$$

where $\Delta E$ is the energy difference between states 1 and 2 and k is the Boltzmann constant. Fitting $\tau_{obs}$ versus T to Equation [9] allows the determination of the value of $\Delta E$ as well as $\tau_1$ and $\tau_2$.

### Fluorescence versus phosphorescence lifetimes

A major advantage of luminescence lifetime measurements is their use for spectral assignment. Specifically, the assignment of the luminescence bands as fluorescence or phosphorescence is primarily determined via luminescence lifetime measurements. As a rule of thumb, lifetimes on the order of microseconds and longer (milliseconds, seconds) are normally indicative of phosphorescence, while fluorescence lifetimes are normally on the sub-microsecond level (nanoseconds, picoseconds etc). It should be noted, however, that fluorescence and phosphorescence lifetime values vary from one case to another, depending on the system under study as well as

other factors such as the extent of excited state distortion. This causes some subjectivity in the assignment, especially when the measured lifetimes are at borderline levels between the aforementioned levels of fluorescence and phosphorescence lifetimes. The clearest cases are those in which fluorescence and phosphorescence bands are both present in one system. In these cases, the phosphorescence bands exhibit lifetimes that are orders of magnitude longer than the lifetimes of the corresponding fluorescence bands. An example is given for such a case in the section on inorganic luminophores.

### Luminescence quantum yields

The quantum yield is a luminescence property that is related to lifetimes. The luminescence quantum yield ($\Phi$) is the ratio of the number of photons emitted to the number of photons absorbed. Therefore, the maximum value of $\Phi$ is 1. In practice, the quantum yield is less than unity for virtually all luminescent materials. The reason is the large number of non-radiative processes that lead to a decrease in the number of emitted photons. The quantum yield can be defined in terms of the rate constants of the radiative ($k_{rad}$) and non-radiative ($k_{nr}$) processes according to Equation [10]:

$$\Phi = \frac{\sum k_{\mathrm{rad}}}{\sum k_{\mathrm{rad}} + \sum k_{\mathrm{nr}}} \qquad [10]$$

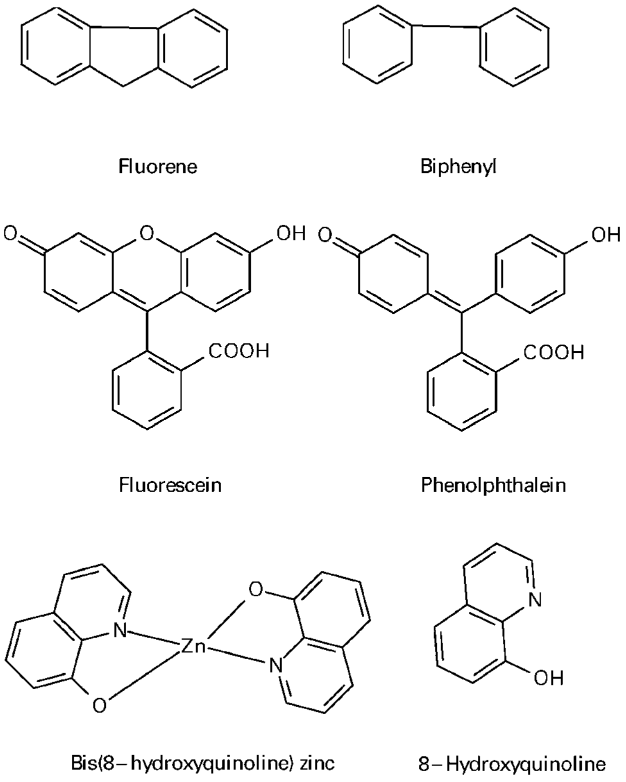
The $k_{rad}$ term includes the rate of fluorescence and phosphorescence while the $k_{nr}$ term includes the rate constants of all the non-radiative processes described previously. Remembering that $\tau_f = 1/k_f$ and also using Equation [7], one can express the quantum yield of the fluorescence ($\Phi_f$) in terms of the luminescence lifetimes as:

$$\Phi_{\mathrm{f}} = \frac{k_{\mathrm{f}}}{k_{\mathrm{f}} + \sum k_{\mathrm{nr}}} = \frac{\tau}{\tau_{\mathrm{f}}} \qquad [11]$$

A variety of factors are believed to influence the luminescence quantum yields of luminophores.

**The transition type** Because luminescence is a technique of electronic spectroscopy, the same selection rules apply for luminescence as those that apply for absorption. For organic compounds, the most common luminescence bands are due to $\pi^*$-$\pi$ and $\pi^*$-n transitions. The $\sigma^*$-$\sigma$ transitions, although strongly allowed, are not normally seen because of their high energies in most organic compounds. The spin selection rules also apply. Consequently, fluorescence in most luminescent organic compounds is much stronger than phosphorescence (see the example in the section on organic luminophores). For inorganic compounds, the luminescence transitions may involve the energy levels of the ligands (intra-ligand transitions), the metal ions (d-d transitions), or both the metal and the ligand (charge transfer transitions). Intra-ligand transitions are similar to the transitions discussed above for organic compounds. The d-d transitions are usually forbidden transitions, so the corresponding absorption and emission bands are generally weak. Charge transfer transitions are strongly allowed and can occur either from the ligand orbitals to the metal orbitals (ligand-to-metal charge transfer, LMCT) or vice versa (metal-to-ligand charge transfer, MLCT).

**Structural rigidity** Molecules that have rigid structures normally exhibit strong luminescence. Three examples are shown in **Scheme 1**. In each pair the fluorophore on the left has the more rigid structure, resulting in greater luminescence quantum yield. Note in the third example that the stronger luminescence intensity is due to the complexation with a metal ion (i.e. the complex is more rigid than the free ligand). The higher luminescence quantum yields for rigid luminophores are probably due to the




**Scheme 1**

inhibition of the internal conversion rates and vibrational motion in these compounds.

**Substitution** The nature of the substituents on, say, an aromatic ring may increase or decrease the quantum yield. For example, halogen and ketone substituents on anthracene generally decrease the fluorescence quantum yield. In contrast, the quantum yield of diphenyl anthracene is nearly unity (compared with $\Phi_f = 0.4$ for unsubstituted anthracene). Phosphorescence becomes an important factor if the substituents contain heavy atoms (for example halogens). In this case the reduction in fluorescence is accompanied by an increase in phosphorescence. This effect (called the heavy atom effect) is specially important in inorganic luminophores because of the involvement of the metal ions. The strong spin-orbit coupling in metal ions leads to the relaxation of the spin selection rules. Therefore, most inorganic luminophores exhibit strong phosphorescence (see below).

**Other factors** Many other factors affect luminescence quantum yields. Among these factors are temperature, solvent, phase and pH. A reduction in temperature suppresses non-radiative processes and thus increases $\Phi$. This is why it is common to run luminescence experiments at cryogenic temperatures. The solvent is involved in many non-radiative processes. An increase in the viscosity of the solvent generally decreases the rate of non-radiative deexcitation. Also, the stretching frequency of the bonds of the solvent molecules is an important factor. For example, one of the common procedures to increase the quantum yield is to run the luminescence measurements in deuterated solvents (e.g. $D_2O$ instead of $H_2O$). The quenching caused by the solvent may be removed by running the luminescence measurements in the solid state instead of solutions (examples are given later). Finally, a change in pH may strongly alter the $\Phi$ value because the structure of many luminophores may be different in acidic and basic media.

## Case studies for photoluminescence

### Organic luminophores

The luminescence properties of anthracene and its derivatives provide an excellent illustration of the relation between the excited state distortion and the profile of the luminescence spectra. **Figure 6** shows the luminescence and absorption spectra of anthracene. Note that the extinction coefficient for the singlet-singlet transition is 8 orders of magnitude higher than the value for the singlet-triplet transition. This observation suggests that the absorption and emission bands in the region with $\lambda < 500$ nm are due to a spin-allowed transition (singlet $\leftrightarrow$ singlet), hence the emission in this region is due to fluorescence. On the other hand, the bands in the region with $\lambda >$ 500 nm are due to a spin-forbidden transition (singlet $\leftrightarrow$ triplet), hence the emission in this region is due to phosphorescence.

**Figure 6** shows that the mirror image rule applies very well to the absorption and fluorescence bands of the $S_0 \leftrightarrow S_1$ transition. Note that among the vibronic bands, the 0-0 and 0-1 transitions have the strongest intensities. These observations suggest that the excited state distortion is very small in anthracene. **Figure 6** also shows that there is correlation even between the absorption and phosphorescence characteristic of the $S_0 \leftrightarrow T_1$ transition. However, the mirror image rule does not apply as strongly as it does for the $S_0 \leftrightarrow S_1$ transition. For example, the intensity is greater for the 0-1 transition than for the 0-2 transition in the phosphorescence band, but the opposite trend is seen in the corresponding absorption band. Moreover, while the absorption and emission peaks are nearly superimposed for the 0-0 vibronic peak of the $S_0 \leftrightarrow S_1$ transition, there is a greater separation between the corresponding peaks characteristic of the $S_0 \leftrightarrow T_1$ transition. These observations are consistent with the excited state distortion for the triplet excited state ($T_1$) being greater than the distortion of the singlet excited state ($S_1$).

Substitution of hydrogen atoms of anthracene may lead to a geometry change in the excited state. The extent of this change can be probed by luminescence spectroscopy. **Figures 7** and **8** provide an illustration. In **Figure 7**, the absorption and fluorescence spectra are shown for 9-anthramide. Similar geometries of the ground state and the fluorescent excited state of 9-anthramide are illustrated by: (i) the applicability of the mirror image rule, (ii) the peaks characteristic of the 0-0 and 0-1 transitions both being strong, (iii) the absorption and emission peaks being superimposed for the 0-0 transition. The situation is not the same when the substituent on anthracene is changed from an amide group to an ester group. **Figure 8** shows the absorption and emission spectra of cyclohexyl-9-anthroate. Note that the emission spectrum of cyclohexyl-9-anthroate is structureless and the mirror image rule is lost. Also note the large values of the Stokes shift (~6000 $cm^{-1}$) and fwhm (~5000 $cm^{-1}$). These observations suggest a largely displaced excited state for cyclohexyl-9-anthroate from the ground state geometry of the molecule. In conclusion, **Figures 7** and **8** show that the luminescent excited state is much more distorted for the ester derivative of anthracene (cyclohexyl-9-

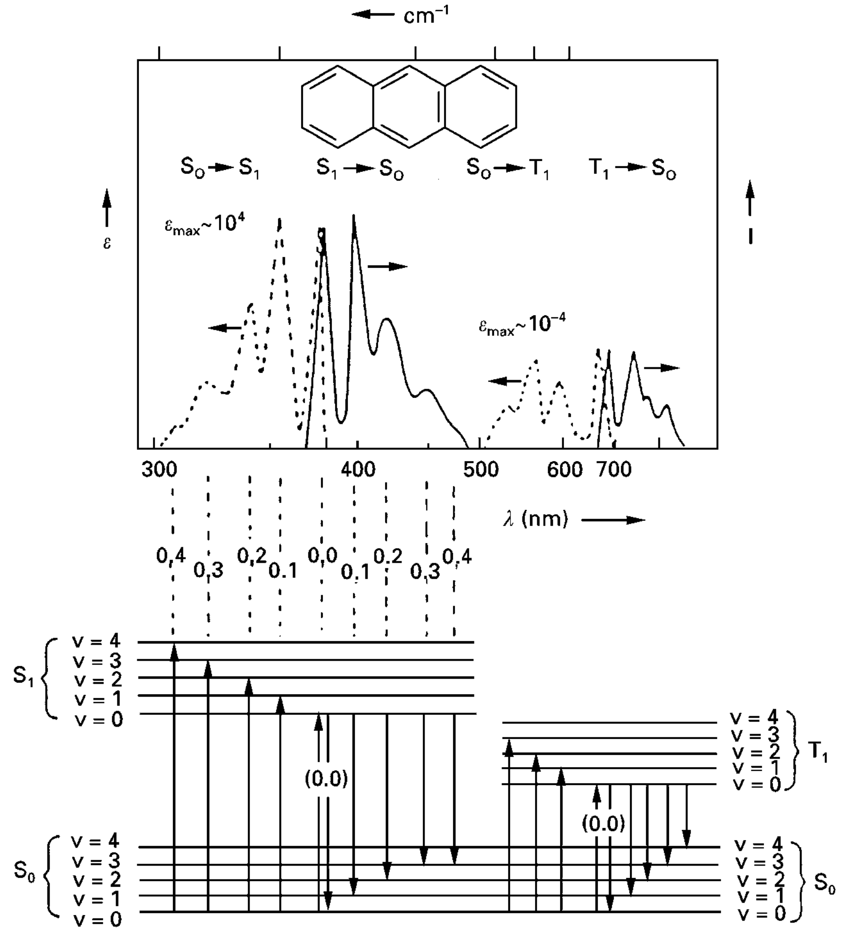



**Figure 6** Emission (solid line) and absorption (dashed line) spectra of anthracene in solution. The assignments of the vibronic transitions are shown in the bottom portion of the figure. Reproduced with permission from Turro NJ (1978) *Modern Molecular Photochemistry*, p 94. Menlo Park: Benjamin/Cummings.

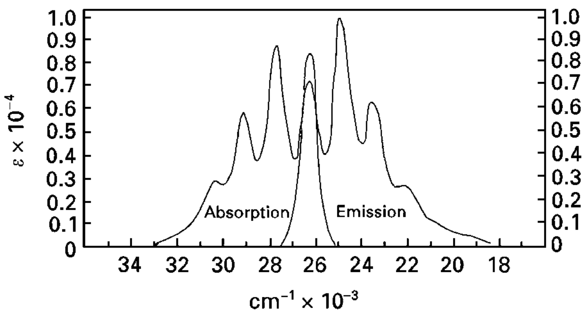



**Figure 7** Emission (right) and absorption (left) spectra of 9-anthramide in tetrahydrofuran. Reproduced with permission from Shon RS-L, Cowan DO and Schmiegel WW (1975) Photodimerization of 9-anthroate esters and 9-anthramide. *The Journal of Physical Chemistry* **79**: 2087–2092.

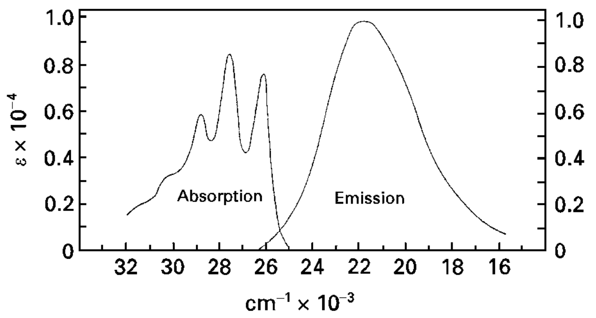



**Figure 8** Emission (right) and absorption (left) spectra of cyclohexyl-9-anthroate in benzene. Reproduced with permission from Shon RS-L, Cowan DO and Schmiegel WW (1975) Photodimerization of 9-anthroate esters and 9-anthramide. *The Journal of Physical Chemistry* **79**: 2087–2092.

anthroate) than for the amide derivative of the same molecule (9-anthramide). One can therefore predict that the ester has a greater tendency to undergo photochemical reactions than the amide.

### Inorganic luminophores

The compound $BaPd(CN)_4.4H_2O$ provides an excellent example for the differentiation between fluorescence and phosphorescence based on lifetimes and

other luminescence properties. Inorganic compounds that exhibit chain structures show interesting luminescence properties in the solid state. The X-ray structure of $BaPd(CN)_4.4H_2O$ consists of $[Pd(CN)_4]^{2-}$ square-planar ions stacked in one-dimensional chains with an intra-chain Pd- Pd distance of 3.37 Å at room temperature. **Figure 9** shows the luminescence spectrum of $BaPd(CN)_4.4H_2O$ at 7 K using a nitrogen pulsed laser for excitation. Two emission bands are observed at ~ $19 \times 10^3$ and $26 \times 10^3$ $cm^{-1}$ (~512 and 382 nm, respectively). These bands are designated the lower energy (LE) and the higher energy (HE) bands, respectively. In **Figure 10** the lifetime of the LE band is plotted versus temperature. A linear fit is obtained. From the resulting equation, the lifetime at 0 K is estimated to be ~ 2.5 ms and decreases by approximately a factor of 2 for each 30 K temperature rise. The shorter lifetimes at higher temperatures are due to the larger values of $k_{nr}$ at higher temperatures (Eqn [7]). In contrast, the HE band has a lifetime shorter than the instrumental resolution of 10 ns. Thus, the lifetime data suggest that the HE band is fluorescence while the LE band is phosphorescence.

Polarized light can be used to study the luminescence bands for $BaPd(CN)_4.4H_2O$. The HE band has polarized emission with the phase of polarization of the fluorescence perpendicular to the $Pd(CN)_4^{2-}$ square planar plane (z-direction). In contrast, the LE band has polarized emission with the phase of the polarized light the same as the $Pd(CN)_4^{2-}$ molecular plane (xy-direction). The metal ions in one-dimensional chained compounds are oriented along one particular direction in crystals of layered compounds (along the xy-direction in $BaPd(CN)_4.4H_2O$). The absorption band of $BaPd(CN)_4.4H_2O$ solid is polarized along the xy-direction. The fact that the HE emission band has the same polarization as the absorption band provides further evidence that the HE band is fluorescence. This is because fluorescence takes place immediately after absorption ($\tau < 10$ ns) and from the same singlet excited electronic state as the one which absorption populates. In contrast, the LE band occurs from a triplet excited state long after absorption ($\tau = 2.5$ ms). Because the absorption and the LE emission bands have different excited states, these bands have different polarization. In accordance with the polarization of the HE band along the same direction as the Pd chains, the emission maximum of the HE band undergoes a progressive red shift (longer wavelength, lower energy) as the temperature is decreased. The red shift occurs because cooling leads to a thermal contraction of the intra-chain Pd- Pd distance, which results in a smaller HOMO- LUMO energy gap (HOMO = highest occupied molecular orbital, LUMO = lowest unoccupied molecular orbital). In contrast, the position of the LE band is independent of temperature because this band is not polarized along the Pd chains.

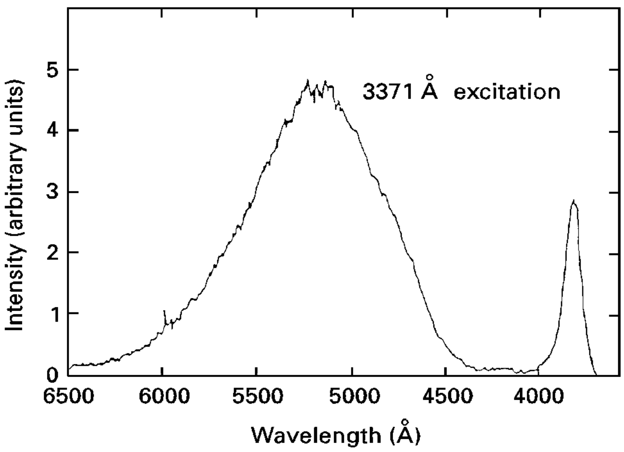



**Figure 9** Luminescence spectrum of $BaPd(CN)_4.4H_2O$ at 7 K. A nitrogen laser was used for excitation. Reproduced with permission from Ellenson WD, Viswanath AK and Patterson HH (1981) Laser-excited luminescence study of the chain compound $BaPd(CN)_4.4H_2O$. *Inorganic Chemistry* **20**: 780–783.

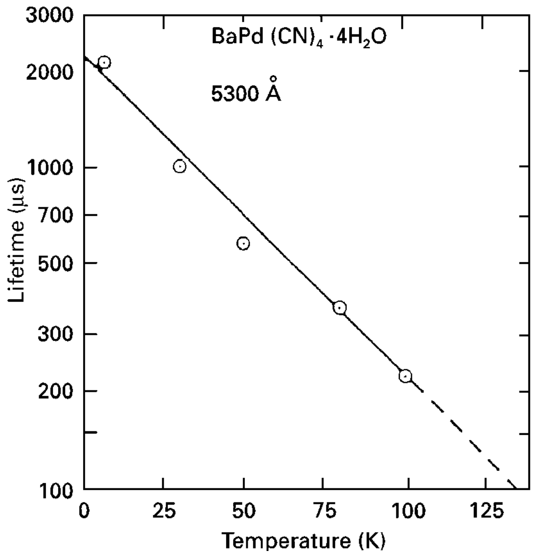



**Figure 10** Lifetime of the $19 \times 10^3$ $cm^{-1}$ luminescence band of $BaPd(CN)_4.4H_2O$ versus temperature. The dashed line is extrapolation to higher temperatures. Reproduced with permission from Ellenson WD, Viswanath AK and Patterson HH (1981) Laser-excited luminescence study of the chain compound $BaPd(CN)_4.4H_2O$. *Inorganic Chemistry* **20**: 780–783.

The preceding results can be compared with the selection rules for all the possible electronic transitions obtained from group theory, to give electronic assignments (term symbols) for each of the luminescence bands. The resulting electronic assignments are $^1A_{2u}$ and $^3A_{2u}$ for the HE and LE bands, respectively. A discussion of the concepts of group theory that lead

to this assignment is beyond the scope of this article. However, it is enough to note that the term symbols of the HE and LE bands of $BaPd(CN)_4.4H_2O$ differ only in the spin multiplicity (superscripts indicate singlet and triplet, respectively). With this in mind, the difference in the excited state distortion between the HE and LE bands can be used to provide further evidence for the band assignment. The lowest energy absorption band of $BaPd(CN)_4.4H_2O$ solid has an energy of ~$31 \times 10^3$ $cm^{-1}$. This gives a Stokes shift of ~5000 and 11 500 $cm^{-1}$ for the HE and LE bands, respectively. **Figure 9** shows that the LE band is much broader than the HE band. The fwhm values of the HE and LE bands are 900 and 3300 $cm^{-1}$, respectively. The higher values of the Stokes shift and fwhm for the LE band are consistent with the assignment that the LE band is phosphorescence while the HE band is fluorescence of the same electronic transition. This is the case because the excited state distortion of phosphorescence is greater than that of the fluorescence of the same electronic transition (with a different spin state). The Jablonski diagram shown in **Figure 1**) illustrates the higher Stokes shift for phosphorescence compared with fluorescence.

In summary, the HE and LE luminescence bands of $BaPd(CN)_4.4H_2O$ are assigned as fluorescence and phosphorescence, respectively. The basis of this assignment is the difference between the two bands in the lifetime values, polarization character, temperature dependence, and excited state distortion (Stokes shift and fwhm values).

Finally, an interesting aspect of the luminescence spectrum of $BaPd(CN)_4.4H_2O$ is the strong phosphorescence intensity (compared with the relative phosphorescence/fluorescence intensity of anthracene). This is a direct consequence of the strong spin-orbit coupling of palladium. This provides an illustration of the heavy atom effect which is generally more important in inorganic luminophores relative to their organic counterparts.

## Quenching of emission

### Theory

In luminescence spectroscopy, quenching refers to any process that leads to a reduction in the luminescence intensity of the luminophore. Static and dynamic quenching are the most common types of luminescence quenching and will be described in the following sections. A third type of quenching also exists, namely 'inner-filter quenching'. An inner-filter effect occurs when the total absorbance of the solution is high (greater than 0.1 au). This leads to a reduction in the intensity of the excitation radiation over the path length. Quenching of this type is not generally categorized among the major quenching process because it is a trivial type of quenching that is not really involved in the radiative and non-radiative transitions in luminescence spectroscopy.

Both static and dynamic quenching require contact of the luminophore with a quencher molecule. This requirement is the basis of the many applications of luminescence quenching. Because there are so many molecules that can act as luminescence quenchers, an appropriate quencher can be selected for any luminophore under study in order to investigate specific properties of the luminophore. An example of biochemical applications of luminescence quenching is given later in this section.

### Static quenching

Static quenching occurs upon the complexation of the luminophore (L) with a quencher (Q):

$$\mathrm{L} + \mathrm{Q} \rightleftharpoons \mathrm{L} - \mathrm{Q} \quad [12]$$

The total concentration of the luminophore, $[L]_0$, is given by:

$$[\mathrm{L}]_0 = [\mathrm{L}] + [\mathrm{L} - \mathrm{Q}] \quad [13]$$

The dependence of the luminescence intensity on the quencher concentration can be derived by considering the association constant for the formation of the L- Q complex, $K_s$:

$$K_\mathrm{s} = \frac{[\mathrm{L} - \mathrm{Q}]}{[\mathrm{L}][\mathrm{Q}]} = \frac{[\mathrm{L}]_0 - [\mathrm{L}]}{[\mathrm{L}][\mathrm{Q}]} \quad [14]$$

Rearrangement gives:

$$\frac{[\mathrm{L}]_0}{[\mathrm{L}]} = 1 + K_\mathrm{s}[\mathrm{Q}] \quad [15]$$

Static quenching occurs because the resulting ground-state complex is not luminescent and, therefore, the concentration of free L decreases upon its complexation. The ratio $[L]_0/[L]$ is a measure of the decrease of the luminescence intensity due to static quenching. In the absence of the quencher (Q), the luminescence intensity, $I_0$, is highest because $[L]_0 = [L]$. In the presence of Q, the luminescence intensity, I, decreases due to static quenching. Therefore, the ratio $[L]_0/[L]$ is the same as $I_0/I$. Substituting in Equation [15] gives the

Stern-Volmer equation:

$$I_0/I = 1 + K_s[\mathrm{Q}] \qquad [16]$$

A Stern-Volmer plot of $I_0/I$ against [Q] should give a straight line with a slope equal to $K_s$, which is also known as the Stern-Volmer constant for static quenching.

### Dynamic quenching (collisional quenching)

Dynamic or collisional quenching occurs when the lifetime of the fluorophore is reduced, thus reducing the luminescence quantum yield. The mechanism of this type of quenching involves a collision of the excited luminophore molecule (*L) with a quencher molecule (Q). As a result, *L returns to its ground state without emitting photons and the excitation energy is transferred to Q. The transfer of the excitation energy to Q leads to a reduction of the excited state lifetime of *L and, therefore, a reduction of the luminescence intensity. It can be shown that the reduction of the luminescence intensity due to dynamic quenching is the same as the corresponding reduction of the excited state lifetime:

$$I_0/I = \tau_0/\tau \qquad [17]$$

where $\tau_0$ and $\tau$ refer to the lifetimes in the absence and presence of the quencher, respectively. The Stern-Volmer equation that governs dynamic quenching can be derived on the basis of the kinetic model shown in **Figure 11**.

The radiative decay rate ($k_{rad}$) is the reciprocal of the excited state lifetime in the absence of the quencher ($1/\tau_0$). The rates of depopulation of *L in the absence and presence of Q are given in Equations [18] and [19], respectively:

$$\mathrm{d}[^*\mathrm{L}]/\mathrm{d}t = w - k_{\mathrm{rad}}[^*\mathrm{L}]_0 = 0 \qquad [18]$$

$$\mathrm{d}[^*\mathrm{L}]/\mathrm{d}t = w - k_{\mathrm{rad}}[^*\mathrm{L}] - k_{\mathrm{Q}}[^*\mathrm{L}][\mathrm{Q}] = 0 \qquad [19]$$

Substituting ($1/\tau_0$) for $K_{rad}$, after rearrangement, gives Equation [20], which is the Stern-Volmer equation for dynamic quenching:

$$I_0/I = 1 + k_{\mathrm{Q}}\,\tau_0[\mathrm{Q}] \qquad [20]$$

The term ($k_Q\,\tau_0$) is the Stern-Volmer quenching

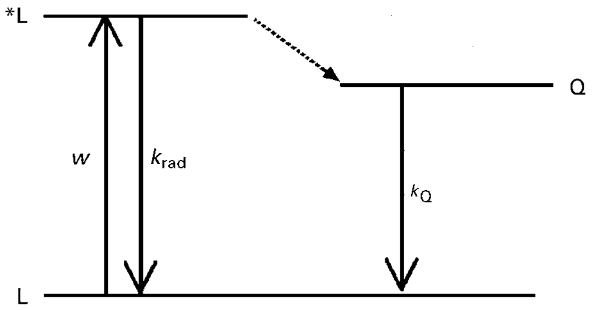



**Figure 11** Kinetic model for dynamic quenching. In the notation used *w*. $k_{rad}$ and $k_Q$ refer to the absorption rate, radiative rate and quenching rate, respectively.

constant for dynamic quenching, $K_D$. Therefore, the Stern-Volmer equation for dynamic quenching can be re-written as:

$$I_0/I = 1 + K_{\mathrm{D}}[\mathrm{Q}] \qquad [21]$$

Note that Equations [16] and [21] are identical in form, and from both equations Stern-Volmer plots can be made in order to obtain the constants $K_S$, and $K_D$, respectively. This can be carried out simply by measurements of the luminescence intensity in the presence and absence of the quencher. However, such a study does not distinguish between the two mechanisms. The easiest and most definitive way to determine the quenching mechanism is to carry out a study of the lifetimes in the presence and absence of the quencher. The excited state lifetime decreases in the presence of quencher if dynamic quenching is the mechanism involved. Note that according to Equation [17] a Stern-Volmer plot can be obtained by plotting ($\tau_0/\tau$) on the y-axis instead of $I_0/I$:

$$\tau_0/\tau = 1 + K_{\mathrm{D}}[\mathrm{Q}] \qquad [22]$$

In contrast, the excited state lifetime is invariant in the presence of the quencher if static quenching is the mechanism involved. Therefore, plotting ($\tau_0/\tau$) versus [Q] will simply yield a horizontal line parallel to the x-axis (y = 1).

### Example

Tryptophan residues in a protein luminesce strongly whether they are on the surface of the protein or in its interior. When a quenching experiment is carried out with ionic quenchers, such as the iodide ion, the resulting quenching involves only surface-localized tryptophan residues. On the other hand, non-ionic quenchers, such as acrylamide, are able to penetrate

into the interior of the protein and provide quenching constants that are really average constants and give information about both surface and interior tryptophan residues.

One way of separating the fluorescence quenching parameters associated with external and internal fluorophores is to use a double-quenching method in which two quenchers are applied simultaneously. The first quencher is used to selectively quench the fluorescence emission of exposed fluorophores. On the other hand, the second quencher is used to quench the fluorescence emission of both surface and interior fluorophores non-selectively.

**Figure 12** shows Stern-Volmer plots of two samples of high density lipoproteins, $HDL_1$ and $HDL_2$ (referring to different densities of high density lipoproteins). These samples were obtained from rats raised on diets containing different amounts of the element manganese. A linear plot was obtained only for the manganese-deficient sample of $HDL_2$. From the equation of the straight line for this sample, the Stern-Volmer constant was determined as 0.88 $M^{-1}$. The plots were non-linear for the other samples, indicating the presence of more than one class of fluorophores, which are unequally accessible to the charged $I^-$ quencher. In **Figure 13**, the fluorescence quenching of $HDL_1$ by iodide in the presence of various amounts of the non-ionc quencher acrylamide are shown using a modified Stern-Volmer method. The plots are all linear and provide fluorescence quenching parameters characteristic of interior and surface-exposed residues. **Table 1** shows a listing of the quenching constants for the different samples.

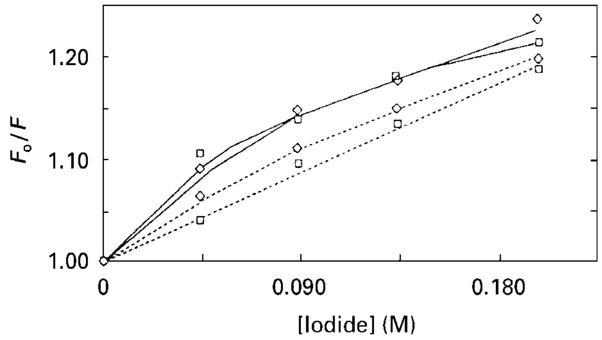



**Figure 12** A Stern–Volmer plot for the fluorescence quenching of $HDL_1$ (solid lines) and $HDL_2$ (dashed lines). The different legends refer to samples obtained from rats fed with manganese-adequate (◊) and manganese-deficient (□) diet. The fluorescence was monitored at 338 nm ($\lambda_{exc}$ = 295 nm) characteristic of tryptophan residues. (Reproduced with permission from Taylor PN, Patterson HH and Klimis-Tavatzis DJ (1997) A fluorescence double-quenching study of native lipoproteins in an animal model of manganese deficiency. *Biological Trace Element Research* **60**: 69–80.

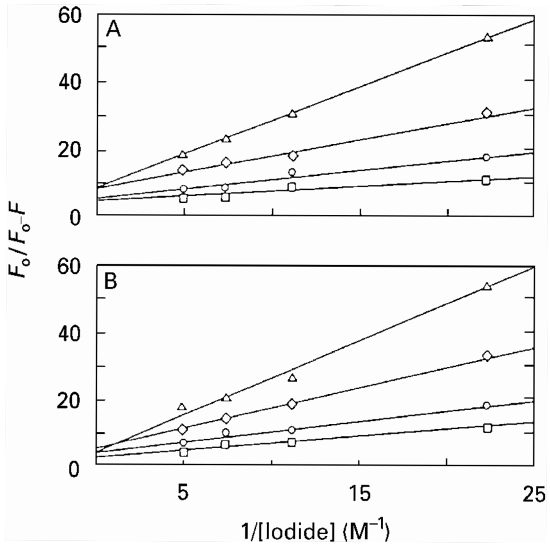



**Figure 13** A modified Stern–Volmer plot for the fluorescence quenching of $HDL_1$ by iodide in the presence of varying amounts of acrylamide. The data are shown for $HDL_1$ samples obtained from rats fed with manganese-deficient (top) and manganese-adequate (bottom) diet. The fluorescence was monitored at 338 nm ($l_{exc}$ = 295 nm) characteristic of tryptophan residues. Reproduced with permission from Taylor PN, Patterson HH and Klimis-Tavatzis DJ (1997) A fluorescence double-quenching study of native lipoproteins in an animal model of manganese deficiency. *Biological Trace Element Research* **60**: 69–80.

**Table 1** Quenching constants for tryptophan residues in different high-density lipoprotein samples

| *Quenching constant* / $M^{-1}$ | *Lipoprotein* | | | |
|---|---|---|---|---|
| | $HDL_1$ | | $HDL_2$ | |
| | MnA | MnD | MnA | MnD |
| $K_I$ | 9.14±0.44 | 17.18 ± 7.72 | 4.63 ± 0.00 | 0.88±0.00 |
| $K_{A1}$ | 5.21±0.88 | 6.15 ± 1.50 | 1.61±0.50 | – |
| $K_{A2}$ | 4.27±0.11 | 5.96 ± 0.05 | 1.51±0.00 | – |

MnA, manganese-adequate sample; MnD, manganese-deficient sample; $K_I$, iodide quenching constant; $K_{A1}$ and $K_{A2}$, acrylamide quenching constant for exposed and partly-exposed fluorophores, respectively.

A statistical analysis of the parameters in **Table 1** indicates that the acrylamide quenching constant for exposed fluorophores is significantly different in manganese-adequate $HDL_1$ compared with manganese-deficient $HDL_1$. In manganese-adequate $HDL_2$, there were two populations of fluorophores accessible to acrylamide, whereas in manganese-deficient $HDL_2$, all fluorophores were accessible to both quenchers. It was concluded based on these results that the local environments of the external fluorophores have different structures and charge distribution in manganese-adequate $HDL_1$ versus

manganese-deficient $HDL_1$. For $HDL_2$ samples, it was concluded that one-third of the fluorophores were accessible to iodide and all external and internal fluorophores were accessible to acrylamide in manganese-adequate samples, whereas in manganese-deficient samples, all fluorophores were accessible to both quenchers.

## Energy transfer

### Theory

Energy transfer refers to a process in which an excited atom or molecule (donor) transfers its excitation energy to an acceptor atom or molecule during the lifetime of the donor excited state. Figure 14 shows that as a result of energy transfer, the donor returns to its ground state while the acceptor is promoted to its excited state. If the acceptor is a luminescent species, it can emit by virtue of energy transfer, i.e. the acceptor luminesces as a result of the excitation of the donor. Such a luminescence is called 'sensitized luminescence', and some textbooks use the terms 'sensitizer' and 'activator' instead of 'donor' and 'acceptor'. Many applications in chemistry, physics, materials science and biochemistry are based on sensitized luminescence. A later section provides an example of the application of energy transfer processes in biological systems. Energy transfer processes occur via radiative or non-radiative mechanisms.

Radiative mechanisms This mechanism involves the absorption of light by a donor atom or molecule (D) followed by emission of a photon by the donor and the absorption of the emitted photon by another molecule called the acceptor (A). The process can be represented as follows:

$$D + h\nu \rightarrow D^* \quad [23]$$

$$D^* \rightarrow D + h\nu' \quad [24]$$

$$A + h\nu' \rightarrow A^* \quad [25]$$

The radiative mechanism is important if the acceptor A absorbs at the wavelength at which the donor emits. The efficiency of the process is determined simply by the quantum yield of the donor luminescence and the absorbance of the acceptor at the donor emission wavelength. No significant interaction between A and D is required in this mechanism and, therefore, radiative energy transfer can occur over extremely large separations of D and A. The radiative mechanism is not important if A and D are similar molecules because of the usually small overlap of the emission and absorption spectra in this case.

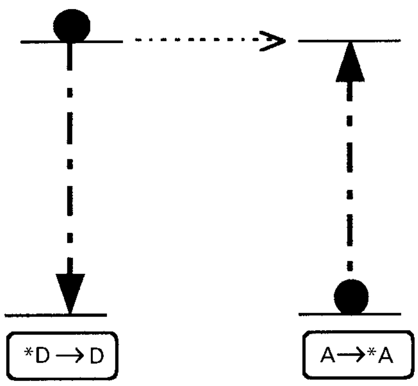



**Figure 14** Energy transfer from a *donor* (D) to an *acceptor* (A).

Non-radiative mechanisms The radiative mechanism is a trivial case of energy transfer because it can be characterized by measuring the donor absorption and the acceptor emission separately from each other. In fact, most textbooks refer to the radiative pathway as a 'trivial mechanism' for energy transfer and do not really consider it as an energy transfer mechanism. Note in Figure 14 that the donor does not emit light. Instead, the excitation energy is transferred non-radiatively to the acceptor. The non-radiative pathway shown in Figure 14 is the mechanism that is typically used in most textbooks to represent energy transfer processes. Non-radiative energy transfer processes occur via two major mechanisms, the Förster resonance mechanism and the Dexter exchange mechanism.

The Förster resonance mechanism involves an electrostatic interaction between D and A. Such an interaction can occur over a long range, i.e. it does not require a very short contact between the donor and the acceptor. Energy transfer via the resonance mechanism may proceed over donor- acceptor distances as long as 50- 100 Å. In order for energy transfer to be efficient via the resonance mechanism, the energies of the donor and acceptor transitions $D^* \rightarrow D$ and $A \rightarrow A^*$ must be nearly identical (hence the name 'resonance mechanism'). Nevertheless, the presence of phonons (vibrational quanta) may provide assistance for energy transfer when small differences exist between the donor and acceptor excited states. Such processes are called phonon-assisted energy transfer processes (Figure 15). The energies of the phonons should be high enough to surmount the difference

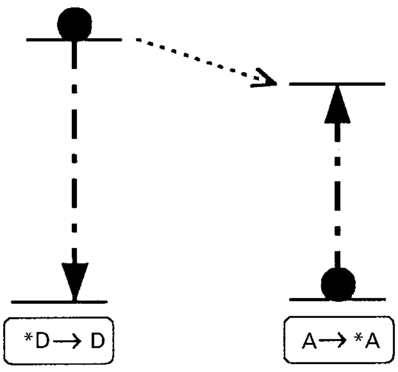



**Figure 15** Phonon-assisted energy transfer. The energy difference between the excited states of D and A is provided by phonons.

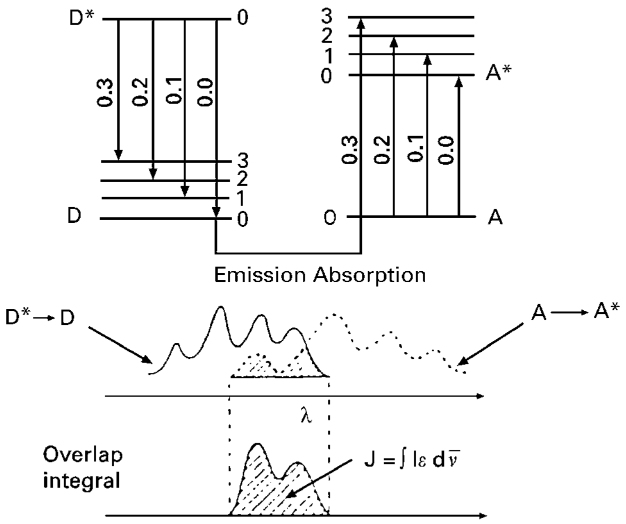



**Figure 16** Experimental determination of the donor–acceptor spectral overlap from emission and absorption spectra. Reproduced with permission from Turro NJ (1978) *Modern Molecular Photochemistry*, p 299. Menlo Park: Benjamin/Cummings.

between the donor and acceptor excited states. However, too high phonon energies increase the likelihood of non-radiative de-excitation (see above), which would take place before energy transfer can take place.

The Dexter Exchange Mechanism involves a direct contact between the donor and the acceptor atoms or molecules. The exchange interaction between the donor D and acceptor A involves a transition state with a D-A distance that is close to the sum of the gas-kinetic collision radii of D and A, respectively. Therefore, information about the energy transfer mechanism can be gained by structural studies of the system under consideration. If the D-A distance is short (normally < 5 Å) then the exchange mechanism is the more likely, but if the D-A distance is long (>> 5 Å) then the resonance mechanism would be more likely.

The energy transfer efficiency is strongly dependent on the spectral overlap between the donor emission and the acceptor absorption. This is valid for both exchange and resonance mechanisms, as shown in the following expression for the energy transfer rate:

$$k_{ET} = f(r_{D-A}) \int I_D \varepsilon_A \, d\bar{\nu} \qquad [26]$$

where the integral represents the spectral overlap between the donor emission $I_D$ and the acceptor absorption $\varepsilon_A$. The factor $f(r_{D-A})$ is a function of the intermolecular distance between the donor and acceptor centres ($r_{D-A}$) and is governed by the relevant energy transfer mechanism (Förster or Dexter mechanism). The donor-acceptor spectral overlap can be determined experimentally by spectroscopic measurements, as illustrated in Figure 16.

### Example

Trivalent lanthanide ions, Ln(III), can be substituted for certain metal ions in proteins to gain structural information about the sites of these metals. The ability of Ln(III) ions to substitute for metal ions such as Ca(II), Zn(II), and Mn(II) in proteins stems from the similarity between Ln(III) and these ions in ionic radii, coordination numbers (6-9) and preference to oxygen donor ligands. The higher charge density of Ln(III) makes these ions substitute with higher affinity than the metal ions in many metal-binding proteins. The quantum yield of luminescence is low for free Ln(III) but increases upon binding in close proximity to an aromatic amino acid within the protein: phenylalanine (Phe), tyrosine (Tyr) or tryptophan (Trp). When the protein is irradiated with wavelengths that correspond to the absorption maxima of these amino acids (~250-300 nm), the luminescence bands of the bound Ln(III) are enhanced by energy transfer. The strong sensitized luminescence of the Ln(III) makes these ions act as 'probes' of detailed and accurate structural information about the metal-binding sites in the proteins studied, as illustrated in the following example.

The protein α-lactalbumin is involved in the regulation of lactose synthesis. The binding of bovine α-lactalbumin (BLA) to Ca(II) is very strong (log K = 8-9). Eu(III) can bind apo-BLA (apo: metal-free) in the Ca(II) binding site. The appearance of sensitized luminescence for Eu(III) is illustrated in Figure 17. Note that owing to energy transfer, the

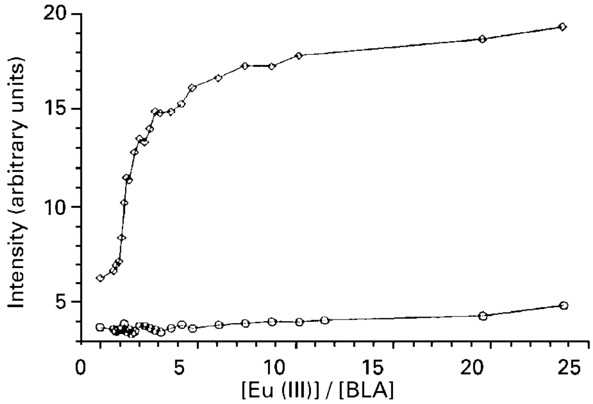



**Figure 17** Luminescence titration of apo-α-lactalbumin (BLA) with europium chloride in $D_2O$. The luminescence intensity is monitored for the $^5D_0 \rightarrow {}^7F_2$ emission line of Eu(III) with $\lambda_{exc}$ = 395 nm. The upper curve shows the data for BLA-bound Eu(III) and the lower curve shows the data for $EuCl_3$ alone. Reproduced with permission from Bünzil JCG and Choppin GR (1989) *Lanthanide Probes in Life, Chemical and Earth Sciences*, p 279. Amsterdam: Elsevier.

luminescence intensity of BLA-bound Eu(III) (upper curve) is much stronger than the intensity of free Eu(III) (lower curve). Figure 17 shows that the increase in the intensity of the $^5D_0 \rightarrow {}^7F_2$ transition is most drastic when the ratio (R) of [Eu(III)/[BLA] is between 1 and 2, and the intensity continues to increase when $R > 2$. This is an indication of the binding of at least two Eu(III) ions to BLA. The Eu(III) luminescence is recorded in Figure 17 for the $^5D_0 \rightarrow {}^7F_2$ hypersensitive transition. A hypersensitive transition is a transition whose intensity is extremely sensitive to changes in the environment of the Ln(III) ion. Hypersensitivity is exhibited because of the $\Delta J = 2$ quadrupolar nature of the transition. The importance of this transition in the BLA-bound Eu(III) is illustrated in Figure 18. The excitation spectra of the Eu(III)-bound BLA are shown in Figure 18 monitoring the $^5D_0 \rightarrow {}^7F_0$ emission. The excitation spectra are resolved into several components, especially for R values > 1. The appearance of bands I and II indicate the presence of two different binding sites for Eu(III) in BLA. Band III is due to non-bonded (solvated) Eu(III) ions. Therefore, it is concluded that Eu(III) ions displace Ca(II) ions from BLA and bind into two sites.

Besides the determination of the number of metal binding sites, luminescence studies of Ln(III) ions in biological systems allow the determination of other structural properties of proteins such as the number of bonded water molecules, the sum of ligand formal charges and the site symmetry of the metal ion sites. For example, lifetime measurements in $H_2O/D_2O$ mixtures allow the determination of the number of water molecules (n) that are coordinated to the metal ions in the different binding sites:

$$n = 1.05\,[1/\tau(H_2O) - 1/\tau(D_2O)] \qquad [27]$$

In the preceding example, the numbers of coordinated water molecules in sites I and II were determined as 2 and 4, respectively.

## Excimers and exciplexes

### Theory

Excimers and exciplexes are excited state complexes. An excited state dimer is called an excimer. Excimer formation can be represented by the following

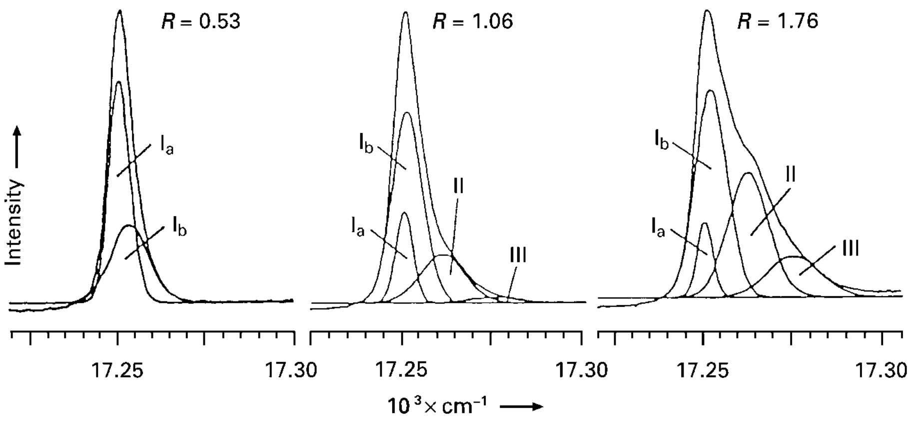



**Figure 18** Curve-resolved Eu(III) excitation spectra of apo-BLA in $D_2O$. *R* represents the [Eu(III)]/[BLA] ratio. Reproduced with permission from Bünzil JCG and Choppin GR (1989) *Lanthanide Probes in Life, Chemical and Earth Sciences*, p 280. Amsterdam: Elsevier.

equation:

$$A + {}^{*}A \rightarrow {}^{*}[A - A] \qquad [28]$$

where the asterisks refer to species in their excited states. According to Equation [28], an excimer is formed when an excited molecule interacts with another identical ground-state molecule. As a result of this interaction, an actual bond forms between the two monomer atoms in the excited state. Excited state interactions are not limited to the formation of homonuclear diatomic complexes. If interaction occurs between an excited molecule of one type and a ground state molecule of a different type, the resulting excited state complex is called an exciplex:

$$A + {}^{*}B \rightarrow {}^{*}[A - B] \qquad [29]$$

An exciplex can also form if more than two atoms are involved in the excited-state bond, whether these atoms are identical or different. Therefore, we can write a general equation to represent exciplex formation:

$$nA + m\,{}^{*}B \rightarrow {}^{*}[A_n - B_m] \qquad [30]$$

From Equation [30], it is clear that an excimer is a special kind of exciplex where A = B and n = m = 1, so it is more appropriate to use the term exciplex when referring to excited state complexes in general, including both homoatomic and heteroatomic species.

It is important to recognize that exciplex formation is a physical phenomenon and not a chemical one. That is, bonding between the molecules only lasts as long as the excited state lifetime of the molecule, and the exciplex bond dissociates upon radiative or non-radiative de-excitation. Nevertheless, there are cases in which exciplex formation leads to photochemical reactions. For example, exciplexes are believed to be the reactive intermediates in the photochemical pathway of the Diels- Alder cycloaddition reactions of unsaturated organic compounds. Since photoluminescence is a photophysical process, the focus here will be on the photophysical aspect of exciplex formation instead of the photochemical reactions that result from exciplex formation.

### Example of an organic exciplex

The characteristics of exciplex emission can be understood from potential energy diagrams of exciplex-forming species. Figure 19 illustrates the spectroscopic features of exciplexes in relation to the potential surfaces of the ground and the excited electronic states. The spectra and potential surfaces are shown for pyrene, the classical example for excimer emission in organic compounds. According to Figure 19, the ground state is repulsive (no potential well) and the excited state is strongly bonding (deep potential well) along the pyrene- pyrene internuclear distance. This explains why the excimer bond forms only in the excited state.

Organic exciplexes are best identified by studying the variation of their emission spectra with concentration. This is illustrated in Figure 19 for pyrene (py). At low pyrene concentrations ($\leq 10^{-5}$ M), the major luminescence band is due to the monomer. The characteristics of the monomer emission include: (i) vibronic structure is present, and (ii) the emission profile is independent of concentration at concentrations $\leq 10^{-5}$ M. As the pyrene concentration is increased above $10^{-5}$ M, the monomer emission is quenched and a new lower-energy emission band appears due to the formation of a *[py- py] excimer. The intensity of the excimer band increases with a concentration increase. The characteristics of the

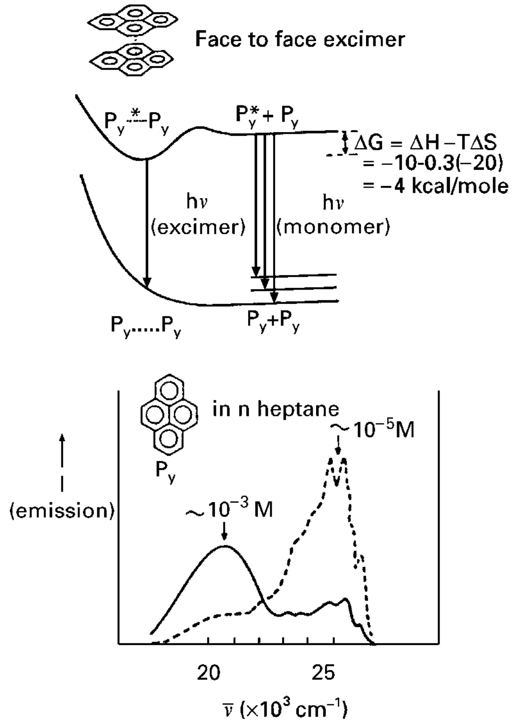



**Figure 19** Excimer emission of pyrene. The upper curves show the potential surfaces of the ground and the luminescent excited electronic state (Reproduced with permission from Turro NJ (1978) *Modern Molecular Photochemistry*, p 141. Menlo Park: Benjamin/Cummings.

excimer emission band of pyrene in Figure 19 are: (i) it is a structureless band with no vibronic structure, (ii) it has a lower energy than the monomer emission band and (iii) its intensity increases relative to the monomer emission band as the concentration is increased above a critical value ($10^{-5}$ M for pyrene). These characteristics of the pyrene excimer bands are valid for the emission bands of organic exciplexes in general.

### Examples of inorganic exciplexes

The formation of inorganic exciplexes has attracted attention only recently. Exciplexes formed in coordination compounds could be either ligand-centred exciplexes, or metal-centred exciplexes. Ligand-centred exciplexes are normally formed from coordinatively saturated complexes (where the metal ion has a high coordination number with no 'vacant' sites) and another species. In this situation, excitation of a ligand in the complex may result in the formation of an exciplex bond between the ligand and another molecule that exists in the solution (such as a solvent molecule). Metal-centred exciplexes, on the other hand, are normally formed from coordinatively unsaturated complexes and another species. In coordinatively unsaturated complexes, the potential coordination site(s) present in the metal ion can be filled by neutral electron-donor species (Lewis bases), anions or another metal ion. The latter type gives rise to the formation of metal-metal bonded exciplexes. This class of inorganic exciplexes is rather interesting because thus far all reported examples are luminescent.

The formation of the silver-silver bonded exciplexes $*[Ag(CN)_2]_n$ ($n \geq 2$) will serve as an illustration. The formation of these exciplexes in the solid state has been reported by the authors of this article. The $[Ag(CN)_2^-]$ complex ion is a good candidate for the formation of metal-metal bonded exciplexes because its low coordination number (2) implies that several potential coordination sites are available.

Single crystals have been grown from a saturated solution of KCl that contains small amounts of $K[Ag(CN)_2]$. The crystals harvested from this solution are called $[Ag(CN)_2^-]$/KCl doped crystals, that is, $[Ag(CN)_2^-]$ guest ions are incorporated into (or doped in) the KCl host lattice. In these doped crystals, the $Ag^+$ and $CN^-$ ions replace the $K^+$ and $Cl^-$ ions, respectively, in some sites in the KCl lattice. Infrared measurements of single crystals of $[Ag(CN)_2^-]$/KCl have shown that multiple peaks exist in the $\nu_{C-N}$ region, indicating the presence of several local environments for the $[Ag(CN)_2]^-$ ions within the KCl lattice. This result has been explained in terms of the presence of monomers, dimers, trimers, etc of $[Ag(CN)_2^-]$ in the doped crystal.

The luminescence spectra of $[Ag(CN)_2^-]$/KCl doped crystals at 77 K are shown in Figure 20. It is interesting to note that at least four emission bands appear in the luminescence spectra of a single crystal of $[Ag(CN)_2^-]$/KCl. The most prominent bands are labelled as A, B, C and D in Figure 20. This observation is consistent with the conclusion based on the infrared spectra that the $[Ag(CN)_2^-]$ ions exist as monomers, dimers, trimers, etc in the KCl lattice. Therefore, $[Ag(CN)_2^-]_n$ oligomer ions in the KCl lattice with different values of n are responsible for the different luminescence bands. The strong dependence of the emission spectra on the excitation wavelength is unusual because emission spectra of luminescent materials are usually independent of the excitation wavelength. By controlling the excitation wavelength, a specific $[Ag(CN)_2^-]_n$ oligomer can be excited independently from the other oligomers so that the luminescence occurs primarily from this

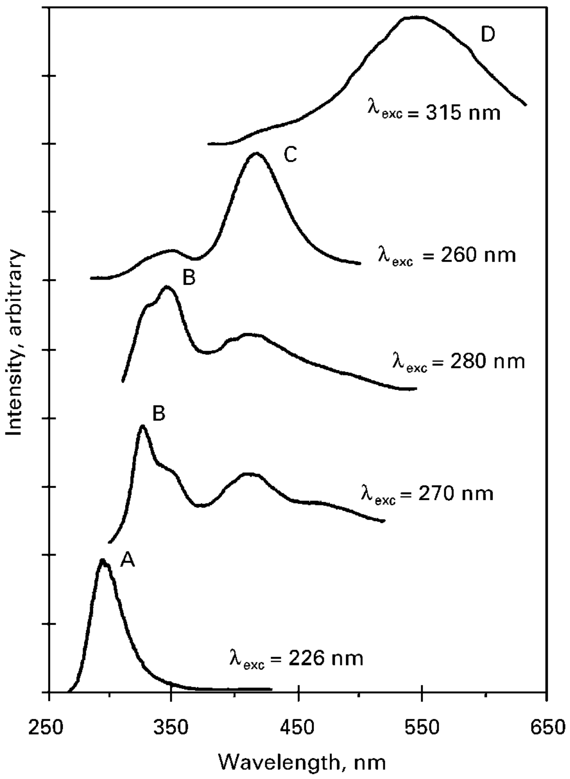



**Figure 20** Photoluminescence spectra of $[Ag(CN)_2^-]$/KCl] doped crystals at 77 K. The letter assignment of the luminescence bands follows the notation used in **Table 2**. Reproduced with permission from Omary MA and Patterson HH (1998) Luminescent homoatomic exciplexes in dicyanoargentate (I) ions doped in alkali halide crystals: 1. Exciplex tuning by site-selective excitation. *Journal of the American Chemical Society* **120**: 7696–7705.

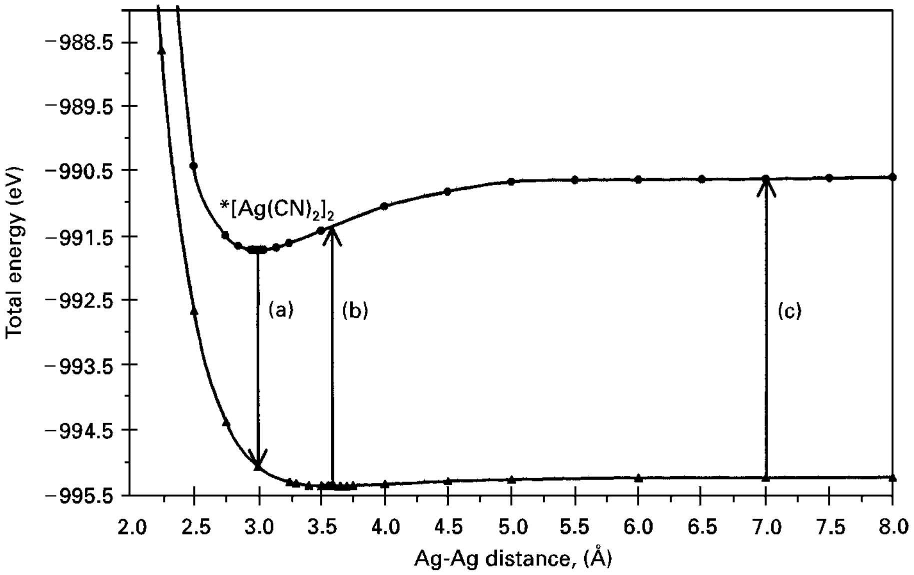



**Figure 21** Potential energy diagram of the ground and the first excited electronic states of $[Ag(CN)_2^-]_2$ (eclipsed configuration) as plotted from extended Hückel calculations. The excimer $[Ag(CN)_2^-]_2$ corresponds to the potential minimum of the excited state. The optical transitions shown are (a) excimer emission, (b) solid state excitation and (c) dilute solution absorption. (Reproduced with permission from Omary MA and Patterson HH (1998) Luminescent homoatomic exciplexes in dicyanoargentate (I) ions doped in alkali halide crystals: 1. Exciplex tuning by site-selective excitation. *Journal of the American Chemical Society* **120**: 7696–7705.

excited oligomer. That is, different $[Ag(CN)_2^-]_n$ oligo-mers in the KCl lattice act as independent luminophores, each of which has a characteristic excitation wavelength. The emission is tuned to any of the bands A- D by changing the excitation wavelength. This is an interesting optical phenomenon that has been called exciplex tuning.

The luminescence bands of $[Ag(CN)_2^-]$/KCl are assigned, as shown in Table 2, to $*[Ag(CN)_2]_n$ exciplexes that differ in the value of n and in their geometry. The exciplex assignment of these bands is based on both experimental and theoretical considerations. The following is an overview of the evidence used for this assignment:

**Table 2** Assignment of the luminescence bands of $[Ag(CN)_2^-]$/KCl doped crystals

| *Band* | $\lambda_{max}^{em}$ *nm* | $\lambda_{max}^{exc}$ *nm* | *fwhm/* ($10^3$ $cm^{-1}$) | *Assignment* |
|---|---|---|---|---|
| A | 285–300 | 225–250 | 3.31 | $*[Ag(CN)_2^-]_2$ (excimers) |
| B | 310–360 | 270–390 | 3.70 | Angular $*[Ag(CN)_2^-]_3$ (trimer exciplexes) |
| C | 390–430 | 250–270 | 3.05 | Linear $*[Ag(CN)_2^-]_3$ (trimer exciplexes) |
| D | 490–530 | 300–360 | 4.01 | $*[Ag(CN)_2^-]_n$ ($n \geq 5$, delocalized exciplexes)[a] |

[a] It is assumed that the stabilization due to Ag–Ag interactions converges in $[Ag(CN)_2^-]_n$ oligomers with $n \geq 5$.

(1) The energies of the luminescence bands of $[Ag(CN)_2^-]$/KCl are extremely low relative to the absorption bands of dilute solutions of $[Ag(CN)_2^-]$ (which represent the $[Ag(CN)_2^-]$ monomer).

The absorption spectra of dilute solutions of $[Ag(CN)_2^-]$ show peak maxima with energies >50 000 $cm^{-1}$. The excitation and emission maxima are red-shifted by as much as 23 000 and 30 000 $cm^{-1}$, respectively, from the monomer transition. These large red shifts must be due to the oligomerization of $[Ag(CN)_2^-]$ units, as suggested by electronic structure calculations (see below).

The fact that the Ag(I) ion has a $4d^{10}$ closed shell electronic configuration forbids the formation of strong Ag- Ag bonds in the ground state. Excited states such as $4d^9$ $5s^1$, however, do not have a closed shell configuration. Hence, Ag- Ag bonding in the excited state is possible, which leads to the formation of $*[Ag(CN)_2^-]_n$ exciplexes. Figure 21 depicts the potential surfaces of the ground state and the first excited state of a $[Ag(CN)_2^-]_2$ dimer, as plotted from electronic structure calculations. The formation of a $*[Ag(CN)_2^-]_2$ excimer is illustrated by the deep potential well in the excited state at a shorter Ag- Ag equilibrium distance than the corresponding ground state distance. The electronic transitions depicted in Figure 21 explain the low emission and excitation energies (transitions (a) and (b)) relative to the monomer absorption (transition (c)).

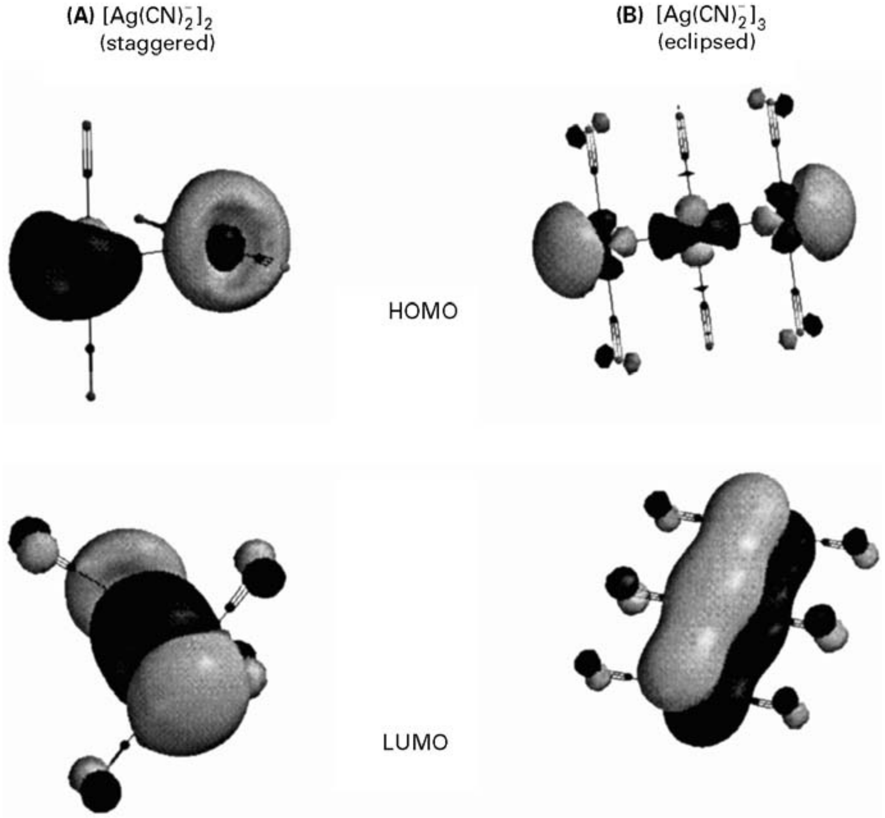



**Figure 22** Surfaces of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) for $[Ag(CN)_2^-]_2$ (staggered isomer) and $[Ag(CN)_2^-]_3$ (eclipsed isomer), as plotted from *ab initio* calculations. Note the Ag–Ag antibonding character for the HOMO and the Ag–Ag bonding character for the LUMO in both cases.

(2) The excited states of $[Ag(CN)_2^-]_n$ oligomers are largely distorted from their ground states.

The Stokes shifts for the luminescence bands of $[Ag(CN)_2^-]$/KCl range from 7000 to 15 000 $cm^{-1}$. The fwhm values are in the 3000–4000 $cm^{-1}$ range (Table 2), which are large values. As explained earlier, large excited state distortions are due to large differences in the bonding properties between the molecular orbitals that represent the ground and excited states, respectively. The ground state of an exciplex is antibonding while its first excited state is bonding. Figure 22 illustrates this fact for the dimer $[Ag(CN)_2^-]_2$ and the trimer $[Ag(CN)_2^-]_3$. In both cases, the HOMO has a Ag–Ag antibonding character while the LUMO has a Ag–Ag bonding character. Similar results were obtained for other $[Ag(CN)_2^-]_n$ oligomers. Photoexcitations from antibonding HO-MOs to bonding LUMOs lead to the formation of $*[Ag(CN)_2]_n$ exciplexes.

(3) The luminescence bands are generally lacking in structure.

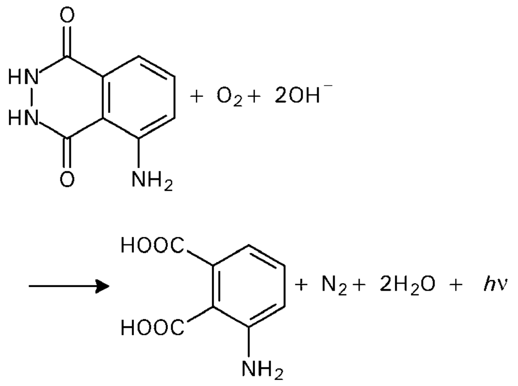
Note that all the luminescence bands shown in Figure 20 have no detailed structure. This is a similar feature to exciplex emission in organic compounds, as illustrated in Figure 19 for the excimer band of pyrene. The apparent structure for band B is due to the existence of two different geometrical isomers of the $[Ag(CN)_2^-]_3$ trimer, not to vibronic structure. The absence of structured emission was obtained for $[Ag(CN)_2^-]$/KCl even though the luminescence measurements were carried out in the solid state for a doped crystal (1.1 mol% Ag) at cryogenic temperature (77 K). Doping and low temperatures normally reduce the chances of band broadening for luminophores that do not exhibit exciplex emission. Hence, the structureless emission of $[Ag(CN)_2^-]$/KCl must be due to exciplex formation.

## Chemiluminescence

When exothermic chemical reactions occur, the product species are usually in their ground electronic states. However, some chemical reactions produce product species in electronically excited states which luminesce. This phenomenon is called chemiluminescence. Incidentally, chemiluminescence occurs in a

number of biological systems (for example, the firefly), where it is called bioluminescence.

A common example of chemiluminescence is the reaction of luminol in basic solution with an oxidizing agent such as oxygen:




[31]

The 3-aminophthalate product ion is in an excited electronic state and shows luminescence in the visible region. Reaction [31] is catalysed by the presence of certain metal ions such as Cr(III) and the intensity of the chemiluminescence has been found to be proportional to the concentration of the specific metal ion, usually in the concentration range of parts per billion (ppb). Thus, the luminol reaction can be used to determine the concentration of selected metal ion species at very low concentrations.

A second common example of chemiluminescence in the gas phase is the reaction of nitric oxide with ozone:

$$NO + O_3 \rightarrow O_2 + NO_2^{*}(\rightarrow NO_2 + h\nu) \qquad [32]$$

Here, the product species $NO_2$ is produced in an excited electronic state and emits light in the visible-near IR region. It has been found that the intensity of the chemiluminescence is proportional to the concentration of NO in the ppm–ppb range. Thus, the reaction shown in Equation [32] can be used as the basis for the development of a chemical sensor for NO. The detection of NO is important because nitric oxide is a chief environmental pollutant, and also because NO plays an important role in human biology.

*See also:* **Biochemical Applications of Fluorescence Spectroscopy; Laser Spectroscopy Theory**.

## Further reading

Adamson AW and Fleischauer PD (1975) Concepts of Inorganic Photochemistry. New York: Wiley-Interscience.

Bünzil JCG and Choppin GR (1989) Lanthanide Probes in Life, Chemical and Earth Sciences. Amsterdam: Elsevier.

Cotton FA (1990) Chemical Applications of Group Theory, 2nd edn. New York: Wiley-Interscience.

Cowan DO and Drisko RL (1976) Elements of Organic Photochemistry. New York: Plenum.

Gilbert A and Baggot J (1991) Essentials of Molecular Photochemistry. Boca Raton: CRC Press.

Lakowicz JR (1983) Principles of Fluorescence Spectroscopy. New York: Plenum.

Omary MA and Patterson HH (1998) Luminescent homoatomic exciplexes in dicyanoargentate (I) ions doped in alkali halide-crystals: 1. Exciplex tuning by site-selective excitation. Journal of the American Chemical Society 120: 7696–7705.

Skoog DA, Holler FJ and Nieman TA (1998) Principles of Instrumental Analysis, 5th edn. Philadelphia: Harcourt Brace.

Turro NJ (1978) Modern Molecular Photochemistry. Menlo Park: Benjamin/Cummings.

Wayne RP (1988) Principles and Applications of Photochemistry. Oxford: Oxford University Press.

# Editor-in-Chief

John C. Lindon, *Biological Chemistry,*
*Division of Biomedical Sciences, Imperial College of Science,*
*Technology and Medicine, Sir Alexander Fleming Building,*
*South Kensington, London SW7 2AZ UK*

# Editors

George E. Tranter, *Glaxo Wellcome Medicines Research,*
*Physical Sciences Research Unit, Gunnells Wood Road,*
*Stevenage, Hertfordshire SG1 2NY, UK*

John L. Holmes, *University of Ottawa,*
*Department of Chemistry, PO Box 450,*
*Stn 4, Ottawa, Canada KIN 6N5*

# M

# Macromolecule–Ligand Interactions Studied By NMR

**J Feeney**, National Institute for Medical Research, London, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| **Applications** |

## Introduction

NMR spectroscopy has proved to be a useful technique for studying interactions between proteins and other molecules in solution. Such interactions are important in biological molecular recognition processes and they have particular significance for studies of drug–receptor complexes where the results can assist in rational drug design. This article indicates how the appropriate NMR data can be extracted and analysed to provide information concerning interactions, conformations and dynamic processes within such protein–ligand complexes. For complexes of moderate size (up to 40 kDa), nuclear Overhauser effect spectroscopy (NOESY) measurements can often be used to determine the full three-dimensional structure of the complex, thus providing detailed structural information about the binding site and the conformation of the bound ligand. For larger complexes (typically up to 65 kDa), ligand-induced changes in protein chemical shifts, dynamic properties, amide NH exchange behaviour and protection from signal broadening by paramagnetic agents can all be used effectively to map out the ligand-binding sites on the protein by reporting on the nuclei influenced by ligand binding. In addition, NMR can sometimes be used to detect bound water molecules within the binding site and to monitor changes in water occupancy accompanying ligand binding.

NMR offers some advantages over X-ray crystallography in that it examines the complexes in solution, does not require crystals and provides a convenient method for defining specific interactions, monitoring changes in dynamic processes associated with these interactions, detecting multiple conformations and identifying ionization states of interacting groups within the protein–ligand complexes. However, unlike X-ray crystallography, NMR can provide full structural determinations only for moderately sized proteins (up to 40 kDa at the present time).

## Equilibrium binding studies

The starting point for studies of protein–ligand interactions often involves determining the equilibrium binding constants for ligands binding reversibly to the protein. These measurements are sometimes made for a series of complexes where either the ligand or the protein is systematically modified in order to measure changes in the binding resulting from the introduction or removal of particular interactions in the complexes. Such investigations need to be accompanied by structural studies on the complexes to see whether the predicted effects have taken place and whether any major conformational perturbations have occurred in the rest of the system. These structural studies need large quantities of purified protein. For a typical sample size of 0.5 mL, the concentrations required vary from 10 μM for one dimensional spectra to 2 mM or greater for some multidimensional experiments. Large quantities of $^{13}C/^{15}N$-labelled proteins are usually prepared by cloning the appropriate gene into an overexpressing bacterial cell line and growing the cells using [$^{13}C$]glucose or [$^{15}N$]ammonium salts as the sole sources of carbon and nitrogen respectively.

## Assignment of protein and ligand signals in the complex

### Fast and slow exchange conditions

Before any detailed structural and dynamic information can be obtained from the NMR spectra of the complexes, the signals need to be assigned to specific nuclei in the ligand or protein. An important first step is to ascertain whether the bound and free species coexist under conditions of fast or slow exchange on the NMR timescale. For a nucleus with chemical shift frequencies $\omega_B$ and $\omega_F$ in the bound and free species respectively, separate signals are seen for the bound and free species for the case where the lifetime of the complex is long compared with $(\omega_B - \omega_F)^{-1}$: this is designated as the slow exchange condition. If the lifetime of the complex is short compared with $(\omega_B - \omega_F)^{-1}$, then conditions for fast exchange prevail and one observes a single averaged signal weighted according to the populations and chemical shifts in the bound and free forms. When the lifetime of the complex is of the same order as $(\omega_B - \omega_F)^{-1}$ then intermediate exchange conditions prevail, giving rise to spectra with broad, complex signals that are more difficult to analyse.

It is necessary to find out whether one is dealing with fast or slow exchange before further work can be attempted. The data can then be analysed to give the chemical shifts of the signals from the bound ligand/or protein. The line widths of the signals can sometimes provide information about the dissociation rate constants of the complex.

### Assignment of protein signals

In making the assignments of the protein resonances, it is important to ensure that the protein is fully saturated with the bound ligand. Using multidimensional NMR methods in combination with $^{2}$H-, $^{13}$C- and $^{15}$N-labelled proteins, it is now possible to obtain almost complete signal assignments for backbone and Cβ protons in proteins of molecular masses up to about 65 kDa. These resonances, once assigned, can be used to monitor ionization state changes, to characterize conformational mixtures and to provide conformational information from NOE measurements for the various complexes.

### Assignment of ligand signals

Assigned signals for nuclei in the ligand are particularly important because these nuclei are obviously well placed to provide direct information about the binding site in the complex. It is easy to assign signals from bound ligands in fast exchange with free ligand if the assignments of the free ligand are known simply by following the progressive shift of the ligand signals during the ligand titration. It is more difficult to assign signals of nuclei in very tightly binding ligands ($K_a > 10^8$ M$^{-1}$) that are in very slow exchange with those in the free ligand. The usual method of assigning signals from tightly bound ligands is to examine complexes formed with isotopically labelled analogues ($^{2}$H, $^{3}$H, $^{13}$C and $^{15}$N). Deuterated ligands can sometimes assist in making $^{1}$H assignments by producing differences between $^{1}$H spectra of complexes formed with deuterated and nondeuterated ligands, since signals from deuterated sites will disappear from the spectra. Complexes formed with $^{13}$C- or $^{15}$N-labelled ligands can also be examined directly by using $^{13}$C or $^{15}$N NMR: only the signals from nuclei at the enriched positions are detected, which simplifies their assignment. Protons directly attached to $^{13}$C or $^{15}$N can be detected using an appropriate editing or filtering pulse sequence. Heteronuclear multiple-quantum (or single-quantum) coherence (HMQC or HSQC) experiments allow the attached protons to be detected selectively and the X nuclei to be detected indirectly. A powerful extension of this approach is the 3D-NOESY-HSQC experiment, which allows selective detection of the NOEs from the ligand protons (attached to $^{15}$N or $^{13}$C nuclei) to neighbouring protons on the protein. The observed $^{1}$H–$^{1}$H NOESY cross peaks are dispersed over the X-chemical shift frequency range. This considerably simplifies the NOESY spectrum at any particular X-frequency and is particularly useful for studying large complexes where there is extensive signal overlap in the normal NOESY spectra.

Complexes formed using less tightly bound ligands ($K_a < 10^6$ M$^{-1}$) can sometimes have spectra showing separate signals for bound and free species in slow exchange that are exchanging sufficiently rapidly to allow their signals to be connected using transfer of magnetization methods. Since the assignments for the free ligand are usually known, these methods give the assignments for the connected signals from the bound ligand.

Other nuclei can sometimes be used effectively for studying protein–ligand interactions. For example, the tritium ($^{3}$H) spectrum of a complex formed with a selectively tritiated ligand shows signals from the ligand only and the chemical shifts of these signals can be directly related to the corresponding protons in the nontritiated ligand. $^{19}$F NMR measurements on complexes formed with fluorine-containing ligands or proteins can also provide useful information. Assignments of $^{19}$F signals from the ligand are often straightforward, since usually only one or two sites are labelled. The simple spectra are ideal for

monitoring multiple conformations and dynamic processes in the complexes. Making $^{19}F$ signal assignments for fluorine-containing proteins is more difficult, but they can be assigned by comparing $^{19}F$ spectra from different proteins where each fluorine-containing amino acid residue has been systematically replaced by a different amino acid using site-directed mutagenesis. Complexes formed with ligands containing phosphorus can be examined directly by $^{31}P$ NMR to provide detailed information about phosphate group ionization states and conformations in the bound state.

## Determination of conformations of protein–ligand complexes

NMR is now able to provide full three-dimensional structures for protein–ligand complexes in solution. The general method involves first making the $^{1}H$ resonance assignments, then estimating the interproton distances from NOE measurements and dihedral angles from vicinal coupling constants and related data, and finally calculating families of structures that are compatible with both these distance and angle constraints and the covalent structure using some optimal fitting method usually, distance geometry-based and/or molecular dynamics simulated annealing-based calculations. Ideally, the structures of the unbound species as well as that of the complex should be determined.

Several workers have reviewed this area, particularly from the perspective of its value in drug design, and there have been many reported studies of ligand–receptor complexes where NMR has provided relevant structural information (see **Table 1**). This present overview will consider only a few examples chosen to illustrate particular aspects of protein–ligand interactions.

**Table 1** Some examples of protein complexes studied by NMR

| |
|---|
| β-Lactamase with substrates |
| β-Lactoglobulin with β-ionone |
| Bcl-x(L) (survival protein) with Bak (cell death protein) |
| Calmodulin with peptides |
| Cyclosporin A with cyclophilin |
| Cytochrome P450 with substrate analogues |
| Dihydrofolate reductase with coenzyme and substrate analogues |
| Elastase with peptides |
| ETS domain of FLI-1 with DNA |
| FK506 binding protein with ascomycin |
| FKBP with immunosuppressants |
| GAT1 domain with DNA |
| Glutathione *S*-transferase with cofactor and substrate analogues |
| Homeodomain proteins with DNA |
| HPr phosphocarrier protein with phosphotransferase domain |
| Integration host factor (*E. coli*) with DNA |
| Lac repressor headpiece with DNA |
| Mu-Ner protein with DNA |
| P53 domain with DNA |
| Pepsin with inhibitors |
| Phospholipase with substrate analogues |
| Pleckstrin homology domain with phosphatidylinositol 4,5-bisphosphate |
| Protease with serpin |
| Protein G (streptococcal) domain with antibody fragment |
| PTB domain of insulin receptor substrate-1 (IRS-1) with phosphorylated peptide from IL-4 receptor |
| Rotamase enzyme FKBP with rapamycin |
| S100B with actin capping protein Cap 2 |
| SHC $SH_2$ domain with tyrosine phosphorylated peptide |
| SRY with DNA |
| Staphylococcal nuclease with substrate analogues |
| Stromelysin domain with N-TIMP-2 inhibitor |
| Stromelysin with nonpeptide inhibitors |
| Thioredoxin with NFκβ peptide |
| Topoisomerase-I domain with DNA |
| Trp repressor with DNA |
| Trypsin with proteinase inhibitors |
| Urbs 1 with DNA |

Many ligands that are flexible in solution adopt a single conformation when bound to a receptor protein. It is important to know the conformation of the bound ligand since this could provide the basis for designing a more rigid and effective inhibitor. Clearly, such information can be obtained directly once the full three-dimensional structure of the complex has been determined. However, in some cases the bound conformation of the ligand can be determined without determining the full structure of the complex if sufficient intramolecular distance and torsion angle constraints can be measured. Several methods based on measurements of intramolecular NOEs in the bound ligand have been proposed. One of these uses the transferred NOE (TrNOE) technique to provide conformational information about the bound ligand. In this method, cross-relaxation (NOE effect) between two protons in the bound ligand is transferred to the free molecule by chemical exchange between bound and free species. Under conditions of fast exchange, the negative NOEs from the bound state can thus be detected in the averaged signals for free and bound ligand. Transferred NOE effects can be detected in 2D-NOESY spectra and this approach has been used, for example, to obtain a set of intramolecular distance

constraints between pairs of ligand protons in the tetrapeptide acetyl-Pro-Ala-Pro-Tyr-$NH_2$ bound to porcine pancreatic elastase and to determine the conformation of the bound peptide.

Other methods of determining the conformation of a bound ligand and details of its binding site involve using isotopically labelled proteins or ligands to simplify the NMR spectra. These approaches are particularly useful for studying tightly binding ligands where transferred NOE methods cannot provide any information. In such cases, it is necessary to measure directly the intramolecular NOEs within the bound ligand. The main problem is one of detecting the relevant NOEs in the presence of a large number of overlapping NOE cross-peaks from protons in the protein. There are several elegant techniques for measuring intra- and intermolecular NOEs in protein ligand complexes by isotopically labelling only one of the partners in the complex. One very direct strategy is to measure intramolecular $^1$H–$^1$H NOEs in unlabelled ligands bound to perdeuterated proteins. Because only the ligand $^1$H signals are detected, the 2D-COSY (correlation spectroscopy) and NOESY spectra are relatively simple. This approach has been used to examine cyclosporin A in its complex with perdeuterated cyclophilin. Another approach is to examine complexes of unlabelled protein with $^{13}$C/$^{15}$N-labelled ligands using NMR isotope-editing procedures that selectively detect only those NOEs involving ligand protons directly attached to $^{13}$C or $^{15}$N. In a $^{15}$N-edited 2D-NOESY experiment on a pepsin/inhibitor (1:1) complex formed with $^{15}$N-labelled inhibitors, NOE cross-peaks between the amide protons attached to $^{15}$N in the ligand and their neighbouring protons in the protein could be detected. Isotope-editing methods have also been used to study $^{13}$C- and $^{15}$N-labelled cyclosporin A bound to cyclophilin. It is also possible to use NMR filter experiments to measure ligand–protein NOEs selectively for complexes containing nonlabelled ligand with $^{13}$C-labelled proteins; this is a useful approach because it is usually easier to obtain labelled proteins than labelled ligands.

## Specificity of interactions

Information about the groups on the protein and ligand that are involved in specific interactions can be obtained by determining the full three-dimensional structure of the complex in solution. More detailed information about specific interactions can often be deduced by monitoring the ionization states of groups on the ligand and protein and noting any changes accompanying formation of the complex. Further information about specific interactions comes from detecting characteristic low-field shifts for NH protons involved in hydrogen bonds.

### Determination of ionization states

NMR is particularly effective for studying electrostatic interactions involving charged residues on the protein or ligand. A change in the charge state of an ionizable group is usually accompanied by characteristic changes in the electronic shielding of nuclei close to the ionizable group. Thus, NMR can monitor the ionization states of specific groups, measure their p*K* values and detect any changes that accompany protein–ligand complex formation.

The p*K* values of histidines in proteins are typically in the range 5.5 to 8.5 and they can easily be studied by carrying out pH titrations of the $^1$H chemical shifts of the imidazole $\varepsilon_1$ protons over a suitable pH range and by fitting the data to the Henderson–Hasselbach equation. Ligand-induced changes in the p*K* behaviour of His residues have been used to monitor interactions in protein complexes formed with novel inhibitors. Protonation states of carboxylate groups in aspartic and glutamic acid residues in proteins have also been studied using $^{13}$C NMR on suitably labelled proteins.

When the ionization state is a protonated species, it is sometimes possible to directly observe the proton involved in the protonation using NMR. If the protonation is at a nitrogen atom, then observation of the selectively labelled $^{15}$NH group provides an unambiguous method of assigning the bonded proton. Such $^{15}$NH proton signals have a doublet splitting (~90 Hz) characteristic of one bond $^{15}$N–$^1$H spin coupling and they can be detected either directly in 1D experiments or by using 2D-HMQC (or HSQC) based experiments. In a $^1$H NMR study examining $^{15}$N-enriched trimethoprim in its complex with dihydrofolate reductase (DHFR), a 90 Hz doublet at 14.79 ppm in the spectrum could be assigned to the N-1 proton of bound trimethoprim (see structure in **Figure 1**). The $^{15}$N chemical shift of the N-1 nitrogen is also characteristic of the protonated species (80 ppm different from the nonprotonated species). Earlier studies using [2-$^{13}$C]trimethoprim had already shown that the N-1 position is protonated in the bound state and that the p*K* value for this protonation is displaced by at least 2 units as a result of formation of the complex in which the protonated N-1 group interacts with the $\gamma$-carboxylate group of the conserved Asp-26 residue.

Ionization states of phosphate groups can be monitored using $^{31}$P NMR and this approach has been used in studies of a coenzyme (nicotinamide–adenine

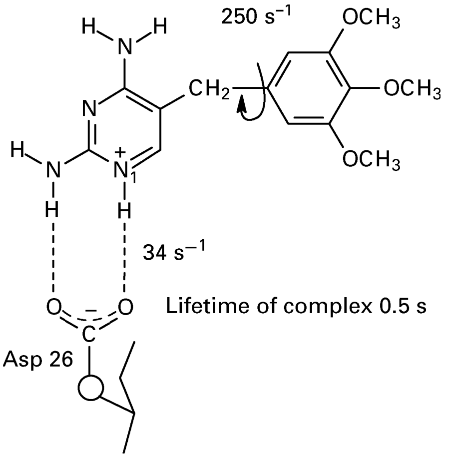



**Figure 1** Dynamic processes in the complex of trimethoprim with *Lactobacillus casei* dihydrofolate reductase measured at 298 K. Reproduced with permission from Searle MS *et al.* (1988) *Proceedings of the National Academy of Sciences of the USA* **85**: 3787–3791.

dinucleotide, NADPH or $NADP^+$) binding to dihydrofolate reductase. In each case the monophosphate group binds in the dianionic form with its p*K* value perturbed by at least 3 units compared to that of the free ligand.

## Hydrogen-bonding interactions involving arginine residues

NMR has proved to be a very effective method for studying hydrogen bonding and electrostatic interactions involving side-chains of arginine residues in protein–ligand complexes. These studies are based on detection of $^1H$ and $^{15}N$ NMR signals from NH groups in $^{15}N$-labelled proteins using gradient-enhanced two-dimensional $^1H/^{15}N$ HSQC NMR experiments where signals for the guanidino $NH^\varepsilon$ and $NH^\eta$ nuclei in arginine residues involved in protein ligand interactions can be detected. Such methods have been used on complexes of $SH_2$ domains formed with phosphopeptides to detect interactions between arginine $NH^\eta$ hydrogens and phosphorylated tyrosines in the protein. Similar interactions have been studied in complexes of *Lactobacillus casei* dihydrofolate reductase formed with antifolate drugs such as methotrexate where four separate $NH^\eta$ signals were observed for the Arg-57 residue, indicating hindered rotation in its guanidino group. Two of the $NH^\eta$ signals had very low-field chemical shifts characteristic of NH hydrogen-bonded protons. From a consideration of the $^1H$ and $^{15}N$ chemical shifts it was possible to deduce that the central pair of $NH^\eta$ protons in the guanidino group of Arg-57 interact with the α-carboxylate group of the glutamic acid moiety of methotrexate in an end-on symmetrical fashion (see **Figure 2**). The rates of rotation about the $N^\varepsilon$—$C^\zeta$ and $C^\zeta$—$N^\eta$ bonds were determined in the binary and ternary complexes of *L. casei* DHFR with methotrexate and NADPH, and their relative values compared with those in free arginine indicate correlated rotation about the $N^\varepsilon$—$C^\zeta$ bond of the Arg-57 guanidino group and the $C'$—$C^\alpha$ bond of the glutamate α-carboxylate group of methotrexate (**Figure 2**).

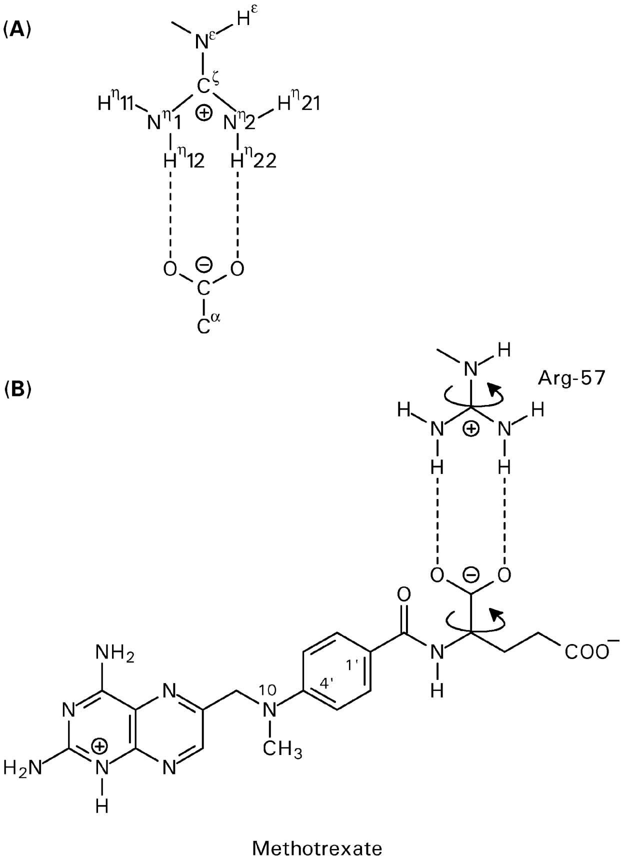



**Figure 2** (A) Symmetrical end-on interaction of a carboxylate group with the guanidino group of an arginine residue. (B) Structure of methotrexate showing interactions of its α-carboxylate group of the glutamic acid moiety interacting in a symmetrical end-on manner with the guanidino group of Arg-57 of *Lactobacillus casei* dihydrofolate reductase and indicating the correlated rotation about the $N^\varepsilon C^\zeta$ bond of the Arg-57 guanidino group and the $C'C^\alpha$ bond of the glutamate α-carboxylate group of methotrexate, which allows the guanidino group to rotate without breaking its hydrogen bonds to the ligand. Reproduced with permission from Nieto PM, Birdsall B, Morgan WD, Frenkiel TA, Gargaro AR and Feeney J (1997) *FEBS Letters* **405:** 16–20. With kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

## Mapping binding sites by ligand-induced chemical shifts

A simple method of mapping the interaction sites in a protein–ligand complex involves measuring the ligand-induced chemical shifts accompanying complex formation using the $^{1}H$ and $^{15}N$ chemical shifts of backbone amide NH groups measured in $^{1}H/^{15}N$ HSQC spectra. This method indicates those residues that undergo a change in environment or conformation on complex formation and it works well, even for the case where the full assignments are available only for the uncomplexed protein. In such cases, lower limits for the shift changes can be estimated and these have proved to be adequate for mapping the binding sites. This method can be used for large protein–protein or protein–DNA complexes (up to 65 kDa).

Using these mapping procedures, a very elegant strategy for designing *de novo* ligands with high-affinity binding for selected target proteins has been developed (the so-called SAR (structure–activity relationships) by NMR approach). Large numbers of ligands were screened for their potential binding to target proteins by measuring $^{1}H/^{15}N$ HSQC spectra of the target $^{15}N$-labelled protein in the presence of batches of ligands. These spectra could be collected relatively quickly and it was possible to screen up to 1000 compounds per day. This method identifies any ligands perturbing the $^{1}H/^{15}N$ chemical shifts (the binding, if any, usually results in conditions of fast exchange). Once a useful binding ligand has been identified, the protein is saturated with this ligand and the screening is continued to find another ligand that binds noncompetitively with the first one. When a suitable second candidate is found, detailed NMR structural work on the ternary complex is undertaken and, based on the structural information obtained, a strategy is developed for chemically linking the ligands to produce a high-affinity binding ligand. This approach has been used successfully to construct inhibitors with high binding affinity for metalloproteinases such as stromelysin.

## Detection of multiple conformations

NMR spectroscopy has proved to be very useful for detecting the presence of different coexisting conformational states in protein–ligand complexes in solution. In some cases the different conformations are in slow exchange such that separate NMR spectra are observed for the different conformations. It is important to characterize the different conformations since each conformation offers a potentially new starting point for the design of improved inhibitors. Recognizing the presence of such conformational mixtures is also important when one is considering structure–activity relationships. NMR is the only method that can provide detailed quantitative information about such conformational equilibria in solution.

Several examples of multiple conformations have been uncovered in NMR studies of complexes of *L. casei* dihydrofolate reductase (DHFR). In many cases the different conformations correspond to a flexible ligand occupying essentially the same binding site but in different conformational states. For example, three conformational states have been detected in the NMR spectra of complexes of the substrate folate with DHFR. Two of the forms have the same pteridine ring orientation as bound methotrexate and their enolic forms can thus bind in a very similar way to the pteridine ring in methotrexate. The other form has its folate pteridine ring turned over by 180°.

Multiple conformations have been detected in several other complexes of *L. casei* DHFR (for example, with $NADP^{+}$ and trimethoprim, and with substituted pyrimethamines) and also in complexes with *S. faecium* DHFR and *E. coli* DHFR: it seems likely that many other protein–ligand complexes will exist as mixtures of conformations. Of course, such conformations are more difficult to detect directly if they are in fast exchange.

## Dynamic processes in protein–ligand complexes

NMR measurements can be used to characterize many of the dynamic processes occurring within a complex: this dynamic information complements the static structural information and provides a more complete description of the complex. Studies using NMR relaxation, line-shape analysis and transfer of magnetization have provided a wide range of dynamic information relating to protein–ligand complexes. The NMR-accessible motions range from fast ($>10^{9}$ $s^{-1}$) small-amplitude oscillations of fragments of the complex to slow motions ($1–10^{3}$ $s^{-1}$) involved in the rates of dissociation of the complexes, rates of breaking and reforming of protein–ligand interactions and rates of flipping of aromatic rings in the bound ligands; several illustrative examples, mainly from studies of dihydrofolate reductase complexes are considered below.

### Rapid motions in protein–ligand complexes

Rapid segmental molecular motions ($>10^{9}$ $s^{-1}$) can be determined by measuring $^{13}C$ relaxation times and useful information about the binding can be

obtained from the changes induced in the motions by the formation of the complex. Protein backbone dynamics are also frequently probed by making $^{15}N$ $T_1$, $T_2$, and $\{^1H\}^{15}N$ heteronuclear NOE measurements on $^{15}N$-labelled proteins and analysing the data using the 'model-free' approach suggested by Lipari and Szabo.

### Dissociation rate constants from transfer of saturation studies

If protons are present in two magnetically distinct environments, for example one corresponding to the ligand free in solution and the other to the ligand bound to the protein, then under conditions of slow exchange separate signals are seen for the protons in the two forms. When the resonance of the bound proton is selectively irradiated (saturated), its saturation will be transferred to the signal of the free proton via the exchange process and the intensity of the free proton signal will decrease. The rate of decrease of the magnetization in the free state as a function of the irradiation time of the bound proton can be analysed to provide the dissociation rate constant. This method has been used to measure the dissociation rate constant for the complexes of $NADP^+$·DHFR (20 $s^{-1}$ at 284 K) and trimethoprim·DHFR (6 $s^{-1}$ at 298 K). 2D-NOESY/EXCHANGE type experiments can also be used for such measurements.

### Rates of ring flipping

Slow and fast rates of aromatic ring flipping have been characterized in ligands bound to proteins. Such studies are facilitated by using $^{13}C$-labelled ligands. For example, $^{13}C$ line-shape analysis on the signals from the enriched carbons in [*m-methoxy*-$^{13}C$]trimethoprim and brodimoprim bound to DHFR has been used to measure the rates of flipping of the benzyl ring in the bound ligand. In all cases these rates are greater than the dissociation rates of the complexes and the flipping takes place many times during the lifetime of the intact complex. Thus the measured rate of flipping is indirectly monitoring transient fluctuations in the conformation of the enzyme structure that are required to allow the flipping to proceed.

### Hydrogen exchange rates with solvent

Extensive NMR measurements of exchange rates between solvent and labile protons on protein or ligand have been reported. These are usually based on line shape analysis or transfer of magnetization methods.

Such measurements have been made for the N-1 proton of bound trimethoprim in complexes of $^{15}N$-labelled trimethoprim with DHFR. The line shape of the N-1 proton signal varies with temperature owing to changes in the exchange rate of this proton with the $H_2O$ solvent. This line-width data can be analysed to estimate the exchange rate. This exchange can be considered as a two step process: in the first step the structure opens to allow access of the solvent, and in the second step the exchange process takes place. In this case, the N-1 proton forms and breaks a hydrogen bond with the carboxylate group of the conserved Asp-26 and the measured exchange rate (34 $s^{-1}$ at 298 K) is thus the rate of breaking and reforming this hydrogen bonding interaction. This provides a further example of a very important interaction in the complex breaking and reforming at a rate much faster than the dissociation rate. Thus, individual protein interactions involving both the pyrimidine ring and the benzyl ring are involved in transient fluctuations during the lifetime of the complex (see **Figure 1**). If these structural fluctuations take place in close succession, they could form part of a sequence of events leading to complete dissociation of the complex.

## Future perspectives

It is clear that advances in NMR methodology, particularly in multidimensional NMR experiments used in conjunction with isotopically labelled molecules, will provide even more detailed information about protein–ligand complexes in solution. Improved methods of structure determination will eventually allow the detection of smaller differences in structure between different complexes. The recently developed approaches for obtaining structural information from dipolar coupling contributions in the spectra resulting from orienting the molecules in solution (either by using high magnetic fields or by using liquid crystal solvents) could have an important impact on structural studies of large protein–ligand complexes. It seems likely that there will be increased input into structure–activity relationship (SAR) studies by use of the 'SAR by NMR' method for designing tightly binding ligands as inhibitors of important target proteins, particularly in industrial pharmaceutical laboratories where suitable libraries of compounds are readily available for screening. Future work should lead to an improved understanding of the implications of the dynamic processes taking place within ligand–protein

complexes. Solid-state NMR studies on ligand complexes of membrane-bound proteins will be undertaken more frequently as the methodology and instrumentation become more widely available: although these studies require demanding isotopic labelling of the ligands, they can provide excellent information about distances and bond orientations that can be used to answer specific questions about the structures of protein–ligand complexes within lipid bilayers. The difficulty of obtaining such information by any other method provides a strong driving force for improving the solid-state NMR approach.

## List of symbols

$T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $\omega_B$ ($\omega_F$) = chemical shift frequency on the bound (free) species.

*See also:* **Drug Metabolism Studied Using NMR Spectroscopy; $^{19}$F NMR Applications, Solution State; Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy; Nitrogen NMR; Nuclear Overhauser Effect; $^{31}$P NMR; Proteins Studied Using NMR Spectroscopy.**

## Further reading

Craik DJ (ed) (1996) *NMR in Drug Design*. Boca Raton, FL: CRC Press.

Emsley JW, Feeney J and Sutcliffe LH (eds) *Progress in NMR Spectroscopy*, Vols 18–33. Oxford: Elsevier. (See articles by C. Arrowsmith (32); M. Billeter (27); G.M. Clore (23); J.T. Gerig (26); A.M. Gronenborn (23); F. Ni (26); G. Otting (31); P. Rosch (18); B.J. Stockman (33); G. Wagner (22); G. Wider (32).)

Feeney J (1990) NMR studies of interactions of ligands with dihydrofolate reductase. *Biochemical Pharmacology* **40**: 141–152.

Feeney J and Birdsall B (1993) NMR studies of protein–ligand interactions. *NMR of Macromolecules* **7**: 183–215.

Fesik SW (1993) NMR structure-based drug design. *Journal of Biomolecular NMR* **3**: 261–269.

Fesik SW, Gampe RT Jr, Holzman TF, *et al* (1990) Isotope-edited NMR of cyclosporin A bound to cyclophilin: evidence for a *trans* 9,10 amide bond. *Science* **250**: 1406–1409.

Handschumacher RE and Armitage IM (eds) (1990) NMR methods for elucidating macromolecule–ligand interactions: an approach to drug design. *Biochemical Pharmacology* **40**: 1–174.

James TL and Oppenheimer NJ (eds) (1989, 1992) Nuclear magnetic resonance. In *Methods in Enzymology*, Vols 176, 177 (1989), Vol 239 (1992). London: Academic Press.

Jardetzky O and Roberts GCK (1981) *NMR in Molecular Biology*. London: Academic Press.

Markley JL (1975) Observation of histidine residues in proteins by means of nuclear magnetic resonance spectroscopy. *Accounts of Chemical Research* **8**: 70–80.

Roberts GCK (ed), (1993) *NMR of Macromolecules: A Practical Approach*. New York: Oxford University Press.

Shuker SB, Hajduk PJ, Meadows RP and Fesik SW (1996) Discovering high-affinity ligands for proteins – SAR by NMR. *Science* **274**: 1531–1534.

Watts A, Ulrich AS and Middleton DA (1995) Membrane protein structure: the contribution and potential of novel solid state NMR approaches. *Molecular Membrane Biology* **12**: 233–246.

Wüthrich K (1976) *NMR in Biological Research: Peptides and Proteins*. Amsterdam: North-Holland.

# Macromolecules Studies By Solid State NMR

*See* **High Resolution Solid State NMR, $^{13}$C.**

# Magnetic Circular Dichroism, Theory

**Laura A Andersson**, Vassar College, Poughkeepsie, NY, USA

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
| --- |
| Theory |

## Introduction and overview

Magnetic circular dichroism (MCD) spectroscopy is a type of electronic spectroscopy, also called the Faraday effect or the Zeeman effect, that can be a particularly useful and effective method for structural analysis. For example, MCD can be used to assign the transitions in the electronic absorption spectrum (UV-visible), with respect to details such as the molecular orbital origins of the transitions. Often, such transitions are not clearly observed in the UV-visible spectra, because they are spin-forbidden and weak, but upon application of the magnetic field, $H_0$, they can be detected. MCD spectroscopy can also be used to determine not only the spin state for a metal such iron, but also the coordination number at the metal.

There is an extensive body of detailed MCD structural data provided for a variety of different biological, organic, and inorganic systems. However, MCD has been surprisingly neglected, given its broad utility, ease of handling, and low sample-concentration requirements relative to many other spectroscopic methods. MCD spectroscopy has only recently begun to be utilized to its full potential.

Biological systems that have been studied by MCD include: (a) a haem (iron porphyrin-containing) proteins and enzymes such as oxygen transport proteins (haemoglobin and myoglobin), electron-transfer proteins (cytochromes), the diverse and ubiquitous P450 enzymes, and peroxide-metabolizing enzymes such as peroxidases and catalases; (b) other biological chromophores such vitamins $B_{12}$ and chlorophylls; (c) tryptophan-containing proteins (this amino acid has a unique and distinguishing MCD signal); (d) non-haem iron proteins; (e) copper- and cobalt-containing proteins (natural or metal-substituted); and (f) a variety of other systems, too diverse to list here.

One particular advantage of MCD spectroscopy is the limited sample requirements, particularly relative to other experimental methods, even in these days of cloning and massive expression of samples. For example, as much as 500 µL of a 1–2 mM solution of haem protein must be used for NMR structural analysis. In contrast, to study the same sample by conventional (electromagnet) MCD, only 2.6 mL of an ~10–25 µM sample is required. Second, the ability to determine key structural information such as spin and coordination states at, or near, biological temperatures is also significant. Whereas the electronic absorption spectrum of a ferric haem protein can generally be used to distinguish high-spin from low-spin systems, more specific information concerning the coordination number was once routinely determined by EPR (also called ESR) spectroscopy. This method not only requires at least 250 µL of an ~250 µM sample, it also requires either liquid nitrogen or liquid helium cooling of the sample to gain the EPR *g* values and make the assignments. In contrast, MCD spectroscopy of a sample under biologically relevant conditions can provide highly detailed and specific data with respect to both the spin and the coordination states of the system (e.g. high-spin pentacoordinate haem vs high-spin hexacoordinate haem). This has recently been illustrated in the case of the haem catalases, which are among the most rapid of all enzymes, converting $H_2O_2$ to $O_2$ and $H_2O$ with a turnover rate of ~100 000 per second per active centre (most catalases have four active centres). A novel set of X-ray crystallographic data for bacterial catalase were published, which were not only in conflict with previous X-ray data for a mammalian catalase, but also appeared inconsistent with the rapidity of normal enzymatic activity. Specifically, the ferric haem of the catalase was suggested to have a water molecule as its sixth ligand that was furthermore stabilized by participation in a hydrogen-bonding network. MCD spectral analysis of the identical bacterial catalase, as well as a mammalian, and a fungal catalase, clearly and unequivocally demonstrated that, under approximately biological conditions, all of the native catalases were always high-spin and pentacoordinate, with NO water ligated at the haem regardless of pH in the range 4–10. Again, this empty coordination site for the haem is of critical significance for the enzymatic reaction of the catalases, where the first step of the reaction requires $H_2O_2$ ligation at the haem.

In part, the increasing employment of MCD spectroscopy in structural analysis derives from a

widening array of modifications that extend the diversity and accuracy of the method. A simple listing of such variations includes:

(1) method of field generation: electromagnet vs superconducting magnet;
(2) spectral region studied: near-infrared (near-IR; NIR) vs UV-visible (ultraviolet and visible);
(3) VT (variable temperature) and VTVH (variable temperature, variable field; [$H_0$] MCD, using a superconducting magnet, with temperature variations to as low as ~1.5 K, and magnetic field variations from ~1 to 50 T (1T = 10 000 G);
(4) 'fast' MCD, which includes nanosecond and picosecond experiments, using an 'ellipsometric approach', also called TRMCD (time-resolved MCD);
(5) VMCD, vibrational MCD, particularly Raman;
(6) the newest modification, XMCD, X-ray detected magnetic circular dichroism.

Fundamentally, MCD spectroscopy can be defined as the differential absorption of left and right circularly polarized light, induced by an external magnetic field ($H_0$) that is parallel (or anti-parallel) to the direction of light propagation. This property is known as the 'Faraday effect', after Michael Faraday who observed (ca. 1845) that *any* substance, when placed in a magnetic field, will rotate the plane of polarized light. Indeed, it was the Faraday effect that was used to establish the electromagnetic nature of light. A general schematic of an MCD instrument is presented in **Figure 1**.

In the case of *degenerate* electronic transitions, for which the components are not resolved in the absorption spectrum, one has access to only limited structural information. However, in the presence of a magnetic field, these degeneracies are lifted (Zeeman effect) and now can be explored in more detail. Using ordinary (conventional) electronic absorption spectroscopy, no detectable spectral difference is observed for such a sample in the presence or absence of the applied external magnetic field. This is because the spectral line width (for most samples) is greater than the splitting of the energy levels. However, using circularly polarized light it is possible to measure and record the differences between these magnetically degenerate states (see *A* and *C* terms below). For a sample that has no nondegenerate energy levels, it is still possible to obtain an MCD spectrum if the nondegenerate energy levels undergo a magnetically induced mixing; this is the origin of MCD *B* terms. A more detailed analysis presented below.

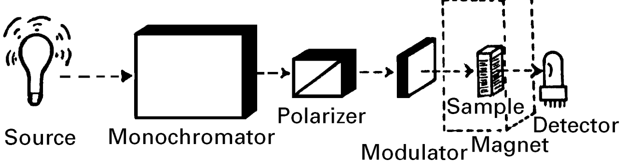



**Figure 1** Optical components of a typical MCD/CD instrument. The modulator, now most commonly a piezoelectrically driven photoelastic device, converts linearly polarized light to a.c. modulated circularly polarized light.

## MCD vs CD spectroscopy

Three aspects of MCD spectroscopy are clearly distinct from those of 'natural' circular dichroism (CD) spectroscopy:

(a) CD requires an optically active, chiral, molecule (essentially one of low molecular symmetry at the chiral centre lacking even a simple mirror plane), whereas MCD has no structural requirements, but rather is a property of *all* matter;
(b) chirality and optical activity (CD) are derived from the presence of *both* electric *and* magnetic dipole transition moments in the sample under study, which furthermore must be parallel (or anti-parallel) to one another, whereas for magnetic optical activity (MCD) only an electric dipole transition moment is required, with the external magnetic field supplying the magnetic component (see **Figure 2**);
(c) CD spectra are sensitive to molecular structure and perturbations of the chiral centre(s) by the physical environment, which is most clearly seen as asymmetry in the chromophore and/or its environment. MCD spectra are representative of the electronic structural properties of a given molecule, such as field-induced perturbations in energy levels. The latter, however, does *not* imply the absence of environmental sensitivity, but rather that molecular perturbations must directly affect the electronic properties. For example, this may include not only a concentration depend-

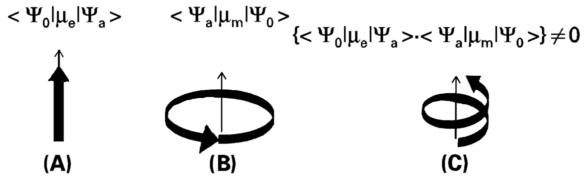



**Figure 2** Cartoon illustrating the photon-induced transitions in a molecule. (A) Electronic absorption from ground to excited state is expressed as shown, where $\mu_e$ is the electric dipole moment operator; (B) magnetic absorption and the mathematical expression, where $\mu_m$ is the magnetic dipole moment operator; and (C) interaction of electronic and magnetic absorption, yielding optical activity.

ence, but also a sensitivity to structural variations, and to precise ligand geometry surrounding the chromophore (often a metal). MCD reflects electronic structural features such as spin and orbital degeneracies – information about spatial and coordination structure.

Note that in the simplest quantum mechanical expression, for a CD spectrum to be observed there must be both electric and magnetic dipole transition moments, for which the cosine between the two transition dipole moments must be non-zero (**Figure 2**). Essentially, this means that the transition dipole moments must have a parallel (or anti-parallel) relationship to one another. Without all three components (the electric dipole transition moment, the magnetic dipole transition moment, and their parallel relationship) there can be no optical activity.

Extensive theoretical discussions of CD spectroscopy focus on the specific origin of CD activity, such as the 'one electron model'. MCD differs specifically here from CD, in that MCD spectroscopy provides the external magnetic field, $H_0$, whereas chiral systems have their own magnetic transition dipole as a consequence of their very low symmetry at the chiral centre.

## MCD experimental details

Fundamentally, then, both magnetic circular dichroism and circular dichroism are phenomena dependent upon the Beer–Lambert law (Eqn [1]) that is to say, upon the *concentration* of the sample, and upon the inherent 'responsiveness' of the sample under study to light, called the extinction coefficient:

$$A_\lambda = \varepsilon_\lambda cb \qquad [1]$$

where $A$ = absorbance (unitless), $\lambda$ = the specific wavelength, $\varepsilon$ = the molar extinction coefficient ($M^{-1}$ $cm^{-1}$); $c$ = molar concentration (mol $L^{-1}$; M), and $b$ = cuvette pathlength (cm). More specifically, this is written as shown in Equations [2] and [3] for circular dichroism and magnetic circular dichroism, respectively, where $\varepsilon_1$ and $\varepsilon_r$ are the specific extinction coefficients for left- and right-circularly polarized light (LCPL and RCPL, respectively):

$$A_l^0 - A_r^0 = \Delta A_{CD}^0 = \Delta A_\lambda^0 = (\varepsilon_l^0 - \varepsilon_r^0)c\,b \qquad [2a]$$

or

$$\Delta\varepsilon^0 = \varepsilon_l^0 - \varepsilon_r^0 \qquad [2b]$$

$$\Delta A_\lambda = A_l - A_r \qquad [3a]$$

$$= \Delta A_{\lambda(CD)}^0 + H\Delta A_{\lambda(MCD)} \qquad [3b]$$

and $\Delta\varepsilon_m = (\Delta\varepsilon - \Delta\varepsilon^0)/H_0 = (\Delta A_\lambda - \Delta A_\lambda^0)/cbH_0$ where $A^0$; $\Delta\varepsilon^0$ etc., with a superscript zero represent those values in the absence of a magnetic field. Thus, the actual MCD experiment requires collection of the MCD spectra for both sample and standard (buffer), and collection of the CD sample for both sample and standard.

The natural CD signal (sample minus buffer) is subtracted from the signal for the sample MCD minus buffer MCD, to yield the 'raw' MCD data. These data are then corrected for field strength (in tesla) and for the molar concentration of the sample under study. Note that the MCD intensity is actually dependent on the strength of the magnetic field, $H_0$, which is a key factor in the type of MCD experiment described as [3] above. This final correction means that it is actually the MCD 'extinction coefficient', $\Delta\varepsilon_M$, that is being reported, and thus one can directly compare the MCD data between different samples in a meaningful manner.

## MCD *A*, *B* and *C* terms; MCD data analysis

In the case of 'natural' CD spectra, each CD spectral band is generally Gaussian in shape, and is associated with a single optically active transition. In contrast, a given electronic spectrum for a sample can result in several MCD spectral features, given the several different mechanisms by which the spectra feature may arise.

Under experimental conditions of temperature such that Zeeman energies are $\ll \kappa T$, and where the Zeeman splitting is small compared with absorption line width, there can be three separate contributions to net ellipticity, imaginatively called $A$, $B$ and $C$ terms. The magnetically degenerate ground or excited states are split by the application of the external field, $H_0$.

The MCD spectral magnitude is directly related to the difference between the LCP (left circularly polarized) and RCP (right-circularly polarized) light, where the (+) and (–) terms below refer to RCP and LCP, respectively:

$$\text{MCD} \propto (\Delta A_\pm \equiv A_- - A_+) \qquad [4]$$

Mathematically, MCD is defined as the difference between two transition moments. It is thus differential,

and the observed MCD spectrum can be either positive or negative for a given absorption:

$$\Delta A_{\pm} = \gamma \beta H_0 c\, b[A_1(-\delta f/\delta E) + (B_0 + C_0/\kappa T) f(E) \quad [5]$$

where $\Delta A_{\pm}$ = change in absorption, $\gamma$ = spectroscopic constant (defined as $N\pi^2\alpha^2\log e)/250hcn$), $N$ is Avogadro's number, $\alpha$ is the proportionality constant between the electric field of the light and the electric field at the absorbing centre, $h$ is Planck's constant, $c$ is the speed of light, $n$ is the refractive index, $\beta$ is the Bohr magneton, $H_0$ is the magnetic field strength, $\kappa$ is the Boltzmann constant, and $T$ is the absolute temperature. The term $f(E)$ is a general Gaussian line shape function, and thus Equation [5] has contributions from both the line shape and derivative of the line shape. $A_1$, $B_0$ and $C_0$ are the MCD parameters ($A$, $B$ and $C$ terms) defining the amount of absorptive ($B_0$ and $C_0$) or derivative ($A_1$) signals in the MCD spectrum.

The $A_1$ term can *only* be non-zero if *either* the ground or the excited state is degenerate and is Zeeman split by longitudinal magnetic field; the derivative line shape function for $A_1$ in Equation [5] results in its characteristic shape. An $A$ term is most easily described for degenerate excited state (**Figure 3**). Because there is a frequency shift between the two transitions, there cannot be cancellation of equal and opposite ellipticity, so the 'derivative shaped' $A$ term is observed. (Note the important distinction that while this feature is shaped like a derivative, it is *not* one in actuality, being a difference spectrum.) Two diagnostic features of an $A$ term are (1) that it is centred at $\nu_0$, which means that the 'zero-point' where the MCD band goes from negative to positive or vice versa, corresponds exactly (in practice, very closely) with the electronic absorption spectral maximum of the sample, and (2) the intensity of the MCD $A$ terms is invariant with temperature, and thus is independent of temperature.

The $C_0$ term is analogous to $A_1$, in that it is only non-zero if there is a degeneracy; however for $C_0$ this must be a ground-state degeneracy. The $C_0$ term is absorptive in nature and also has an associated temperature dependence, which means that the extent of its contribution to the MCD line shape varies with absolute temperature. Assuming thermal equilibrium for the electronic states, the populations are dominated by Boltzmann statistics. Thus as the temperature is lowered, the lower energy level of the degenerate ground state has an increasing population, and the $C$ term increases in intensity. Indeed, the inverse correlation between MCD intensity and absolute temperature is a distinguishing feature of the $C$ term. This is illustrated in **Figure 4**. Again, in comparison with the $A$ term, the $C$ term is not derivative shaped, nor does it have a zero-point that corresponds with the absorption spectral maximum of the correlated electronic transition. For both the $C$ term and the $B$ term (discussed below), the positions of the absorption maximum correlates with the MCD band position, as either a maximum (positive MCD band) or minimum (negative MCD band).

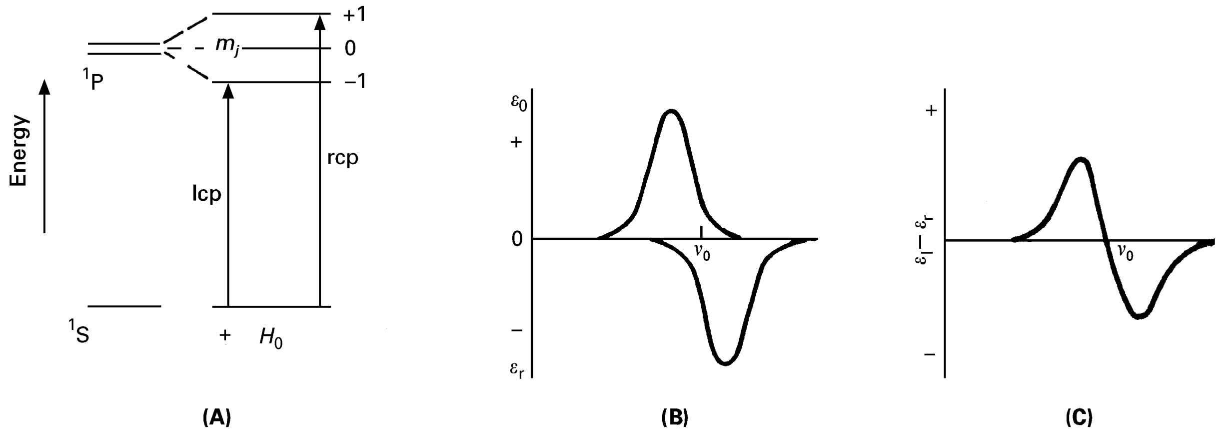



**Figure 3** Origin and typical spectrum of an MCD $A$ term. (A) Electronic transition from ground to degenerate excited state; on the right, in the presence of the magnetic field, $H_0$, the excited state degeneracy is split. This results in separate transitions, corresponding to the differential absorption of left or right circularly polarized light. (B) The separate left ($\varepsilon_l$) and right ($\varepsilon_r$) circularly polarized spectra; $\nu_0$ is the position of the absorption maximum. By convention, the absorption of left circularly polarized light is positive and absorption of right circularly polarized light is negative. (C) The $A$ term MCD spectrum, $\Delta\varepsilon = \varepsilon_l - \varepsilon_r$; note that the position of $\nu_0$ is at the 'crossover' (zero-point) of the derivative-shaped $A$ term. This change of sign for the MCD signal at the position of the absorption maximum is diagnostic for an $A$ term.

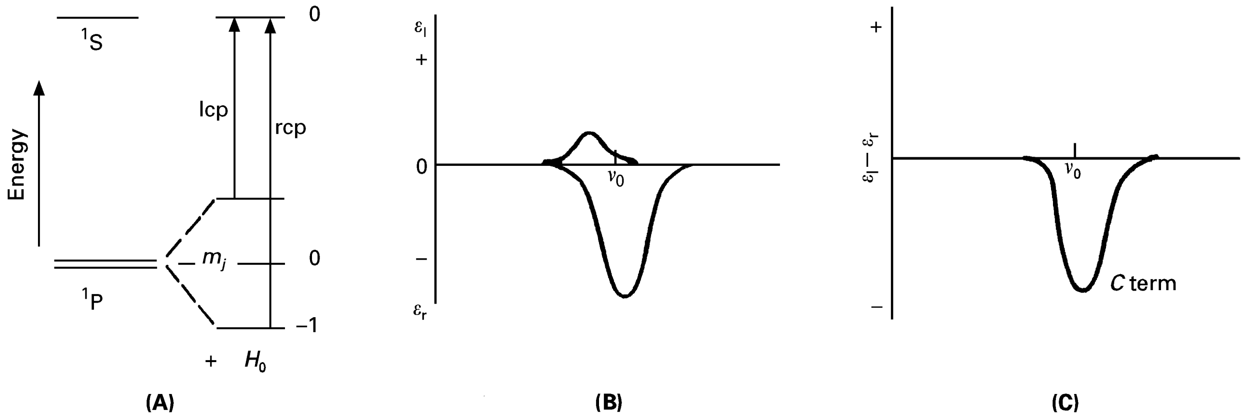



**Figure 4** Origin and typical spectrum of an MCD '*C*' term. (A) Electronic transition from degenerate ground to excited state; on the right, the magnetic field, $H_0$, splits the excited state degeneracy. This results in separate transitions, corresponding to the differential absorption of left or right circularly polarized light. (B) The separate left ($\varepsilon_l$) and right ($\varepsilon_r$) circularly polarized spectra; $\nu_0$ is the position of the absorption maximum. (C) The *C* term MCD spectrum, $\Delta\varepsilon = \varepsilon_l - \varepsilon_r$; note that the $\nu_0$ absorption maximum is not in any particular relationship to the resulting *C* term spectrum. The *C* term intensity is inversely dependent upon temperature, due to Boltzmann population of energy levels; as the temperature decreases, more of the molecules populate the lower of the two degenerate orbitals, and thus the *C* term intensity increases as the temperature is lowered, until the saturation point.

Finally, there is the *B* term of the MCD spectra, which is a complex species deriving from the mixing of electronic states. The quantum-mechanical origin of the $B_0$ term is the magnetically induced mixing of nondegenerate states. The wide spacing of energy levels generally results in a small $B_0$ term. The $B_0$ term is, like the *C* term, Gaussian in shape. However, the *B* term is the temperature-independent portion of the normal absorptive line shape of Equation [5]. To some extent, the MCD spectrum can be considered to be an extension of electronic absorption spectroscopy, revealing a great deal more specific and precise information. This simplification, of course, neglects the information that MCD can provide with respect to magnetic parameters of a transition. As shown in **Figures 3** and **4**, the MCD spectrum not only has band position and intensity for each transition, but also *sign* hence there is more information 'content' provided.

One additional MCD term should be defined: $D_0$, the dipole strength, such that

$$A/E = \gamma D_0 f(E) \qquad [6]$$

Mathematically, the magnetic *g* values for the state involved in the electronic transitions are directly related to the ratios $A_1/D_0$ and $C_0/D_0$. Plotting $\Delta\varepsilon$ vs $\beta H/2\kappa T$ (this graph is called an MCD magnetization curve) permits the determination of the magnetic properties (*g* values) of the ground and excited states, $C_0$ and $A_1$. Although the *g* values determined in this manner are not as precise as those obtained by EPR spectroscopy, they have proved useful. Indeed, the use of group theoretical techniques to calculate MCD parameters is a major factor in the use of the MCD for assignment of energy states. A second key factor, discussed above, is the common ability to use sample concentrations that are an order of magnitude weaker than needed for other structural methods.

## Experimental variations on MCD spectroscopy

The 'basic' MCD instrument is similar to that used for electronic absorption spectroscopy, with the exception that the light must be circularly polarized (see **Figure 1**). A photoelastic modulator (PEM) inputs linearly polarized light, and converts it into circularly polarized light, with RCP and LCP being generated alternatively. The detector observes a signal fluctuation at the ~ 50 Hz frequency and this a.c. signal is input into a sensitive amplifier, capable of detecting very small intensity differences. The sample itself is placed between the two poles of the magnet, and the magnet itself is oriented so that the direction of the field, $H_0$, is parallel (or anti-parallel) to the directions of propagation of the light.

### Variations in method of field generation, electromagnet vs superconducting magnet

Some systems use a simple electromagnet, capable of producing magnetic fields up to 1.5 T (15 000 G). This method is simple and rapid, in that it only

requires cooling water and temperature regulation for the sample. In contrast, other MCD systems use superconducting magnets, which have more demanding experimental constraints (such as liquid helium cooling and a longer 'setup' time prior to onset of experimental data collection), but are capable of generating magnetic fields as large as 50 T.

### Variations in the spectral region studied

Near-infrared (near-IR, NIR) vs UV-visible (ultraviolet and visible): both electromagnets and superconducting magnets can be used successfully non only in the more common ~800–200 nm spectral regions (UV-visible), but also in the near-IR region. There has been considerable success in the assignment of the axial ligands to biological haem systems through study of MCD data for the near-IR region.

### VT and VTVH MCD

These are variable-temperature MCD and/or variable temperature, variable-field MCD, respectively using a superconducting magnet, with temperature variations to as low as ~1.5 K and magnetic field variations from ~1 to 50 T.

VTVH MCD is used to study ground-state electronic structure such as ground-state splittings, especially for non-Kramers ions which do not always have EPR spectra. This approach is a generally useful probe of the ground-state electronic properties of paramagnetic metal ions. MCD is not a 'bulk property technique' like magnetic susceptibility, so is not prone to errors due to paramagnetic impurities. MCD spectroscopy, because it is a type of electronic absorption spectroscopy, can zero in on specific metal ions. VTVH can determine 'single-ion zero field splitting (ZFS)' and the exchange coupling constant, *J*. In comparison, EPR cannot be used in systems that are non-Kramers, even spin systems with large ZFS. Once the ground-state electronic properties are known, then one can determine changes in *J* or in ZFS. This approach, with examination of samples at very low temperatures, down to 1.5 K, is still an absorptive method, so the sample must be optically transparent. This is generally accomplished by dissolution of the sample in, e.g., 50:50 glycerol–buffer (aqueous media).

Low-temperature MCD spectroscopy is an extremely useful tool to distinguish between electronic transitions arising from a diamagnetic (S = 0) ground state and transitions arising from a paramagnetic (S > 0) ground state. This is because the S = 0 level is nondegenerate, and thus cannot provide a temperature-dependent *C* term in the presence of the magnetic field.

### 'Fast' MCD, also called TRMCD (time-resolved MCD)

For reviews of the methodology and applications of nano- and picosecond MCD experiments, see the Further Reading section. This approach has required extensive modification of the equipment used for sample analysis, in particular using elliptically polarized light. This work has led to exciting results, permitting the examination of transient molecular species. Applications of nanosecond MCD have focused primarily on ligand complexes of haem proteins and their photo-produced dissociation intermediates, particularly given the intense absorption maxima of haem systems (typical $\varepsilon$ 100 000 $M^{-1}$ $cm^{-1}$), and their strong MCD signals even at room temperature.

To date, the experimental focus has been on systems with unpaired spins (metal complexes), rotational symmetries (aromatic molecules such as the amino acid trytophan, porphyrins), and metalloporphyrins (haem proteins). An exciting application came from TRMCD of the photodissociated CO adducts of, e.g. haem proteins such as mammalian cytochrome *c* oxidase: the diamagnetic, low-spin, hexacoordinate Fe(II), of the ferrous–CO haem becomes a paramagnetic high-spin, pentacoordinate Fe(II), with a concomitant appearance of a new *C* term. In the case of picosecond TRMCD, picosecond lasers are used. One such application demonstrated that upon photodissociation of the CO ligand bound to the haem protein myoglobin, the change from a hexacoordinate to a pentacoordinate haem occurred, 'within the 20 ps rise time of the instrument'.

### VMCD, vibrational MCD, particularly Raman

Magnetic Raman optical activity determines transitions of electrons among energy levels created by an applied external magnetic field; problems arise here owing to limitations in the field strength. MVCD (magnetic vibrational CD) splits degenerate levels of vibrational transitions and aids in the analysis of bonding.

### X-ray detected magnetic circular dichroism (XMCD)

This technique has only recently evolved into an important method for magnetometry. This technique has unique strengths in that it can be used to determine quantitatively spin and orbital magnetic moments for specific elements, and can also be used to determine their anisotropies through analysis of the experimental spectra. For example, XMCD has been applied to the study of thin films of transient metal multilayers, such as Cu or Fe.

The XMCD method is one where the properties of 3*d* electrons are probed by exciting 2*p* core electrons to unfilled 3d states. The p → d transition dominates the L-edge X-ray absorption spectrum. L-edge X-ray spectroscopy of iron has proved to be useful because the transitions from the 2*p* ground state to 3*d* excited states are strong and dipole allowed, and the small natural line widths also indicate potentially strong MCD spectra. The intense L-edge XMCD spectra of the iron–sulfur protein rubredoxin and of the 2Fe–2S centre of *Clostridium pasteurianum* have also been studied. Both the XMCD sign, and its field dependence, can be used to characterize the type of coupling between magnetic metal ions and the strength of such coupling.

## Conclusion

MCD spectra can profitably separate contributions from multiple metal centres to a protein electronic spectrum, be used to evaluate metallo-biological systems without complications from the protein 'milieu', determine zero-field splitting, assign electronic transitions, provide information about a chromophore's electronic structure, evaluate theoretical models, obtain magnetic properties (*g* values, spin states, magnetic coupling) and be used for structural comparison of 'model' and biological systems.

Modern MCD spectroscopy can only prove to be increasingly useful. Whereas the standard (electromagnetic) instruments available in the 1970s and 1980s could require up to ~45 min per single scan of the data (not counting the buffer, CD, and CD of buffer scans), modern multi-scanning capability permits a significant improvement in signal-to-noise ratio. This has a concomitant advantage in permitting careful and detailed studies to be performed.

Perhaps the greatest utility of MCD spectroscopy is in concert with other methods. No one spectroscopic or structural analysis method can have 'all the answers'. Only a consistent overall structural picture, provided by analysis of data from several methods, with awareness of the shortcomings of each, can lead us closer to the desired 'truth' with respect to the systems under study.

## List of symbols

$A_1$, $B_0$, $C_0$ = MCD parameters ($A$, $B$ and $C$ terms) defining the amount of absorptive ($B_0$ and $C_0$) or derivative ($A_1$) signals in the MCD spectrum; $A$ = absorbance (unitless); $b$ = cuvette pathlength (cm); $c$ = molar concentration (mol $L^{-1}$; M); $c$ = speed of light; $f(E)$ = general Gaussian line shape function; $g$ = EPR $g$ values; $h$ = Planck's constant; $H_0$ = external magnetic field; $n$ = refractive index; $N$ = Avogadro's number; $T$ = absolute temperature; $\alpha$ = proportionality constant between the electric field of the light and the electric field at the absorbing centre; $\beta$ = Bohr magneton; $\gamma$ = spectroscopic constant; $\Delta A_{\pm}$, $A_-$, $A_+$ are the change in absorbance, the negative absorbance, and the positive absorbance, respectively; $\Delta\varepsilon_M$ = MCD 'extinction coefficient'; $\varepsilon$ = molar extinction coefficient ($M^{-1}$ $cm^{-1}$); $\varepsilon_l$ and $\varepsilon_r$ are the specific extinction coefficients for left- and right-circularly polarized light (LCPL and RCPL, respectively); $\kappa$ = Boltzmann constant; $\lambda$ = specific wavelength; $\nu_0$ = the position, in nm, of the electronic absorption maximum for a given transition.

*See also:* **Near-IR Spectrometers; Vibrational CD, Applications; Vibrational CD Spectrometers; Vibrational CD, Theory.**

## Further reading

Andersson LA, Johnson AK, Simms MD and Willingham TR (1995) Comparative analysis of catalases: spectral evidence against haem-bound water for the solution enzymes. *FEBS Letters* **370**: 97–100.

Ball DW (1990) An introduction to magnetic circular dichroism spectroscopy: general theory and applications. *Spectroscopy* **6**: 18–24.

Cheesman MR, Greenwood C and Thomson AJ (1991) Magnetic circular dichroism of hemoproteins. *Advances in Inorganic Chemistry* **36**: 201–255.

Dawson JH and Dooley DM (1989) Magnetic circular dichroism spectroscopy of iron porphyrins and heme proteins. In: Lever ABP and Gray, HB (eds) *Iron Porphyrins, Part III*, pp. 1–133. New York: V. V. H. Publishers.

Goldbeck RA, Kim-Shapiro DB and Kliger DS (1997) Fast natural and magnetic circular dichroism spectroscopy. *Annual Review of Physical Chemistry* **48**: 453–479.

Goldbeck RA and Kliger DS (1992) Natural and magnetic circular dichroism: spectroscopy on the nanosecond time scale. *Spectroscopy* **7**: 17–29.

Holmquist B (1978) The magnetic optical activity of hemoproteins. In: Dolphin D. (ed.) *The Porphyrins*, Vol. III, Chapter 5. New York: Academic Press.

Peng G, van Elp J, Janh H, Que L Jr, Armstrong WH and Cramer SP (1995) L-edge X-ray absorption and X-ray magnetic circular dichroism of oxygen-bridged dinuclear iron complexes. *Journal of the American Chemical Society* 117: 2515–2519.

Solomon EI, Machonkin TE and Sundaram UM (1997) Spectroscopy of multi-copper oxidases. In: Messerschmidt A (ed.) *Multi-Copper Oxidases*, pp. 103–127. Singapore: World Scientific.

Solomon EI, Pavel EG, Loeb KE and Campochiaro C (1995) Magnetic circular dichroism spectroscopy as a probe of the geometric and electronic structures of non-heme ferrous enzymes. *Coordination Chemistry Reviews* **144**: 369–460.

Stohr J and Nakajima R (1997) X-ray magnetic circular dichroism spectroscopy of transition metal multilayers. *Journal de Physique (Paris) IV* **7**: C2–C47.

Sutherland JC (1995) *Methods in Enzymology* **246**: 110–131.

# Magnetic Field Gradients in High-Resolution NMR

**Ralph E Hurd**, G.E. Medical Systems, Fremont, CA, USA

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Theory |

## Introduction

In the 1990s pulsed field gradients became a more common element in multiple-pulse high-resolution NMR methods. Gradients have been incorporated into these sequences to improve water suppression, to spoil radiation damping, to remove undesired signals, and to collect faster or higher-resolution multidimensional spectra. Although the potential for pulsed magnetic field gradients has been known since the early years of NMR, only recently has the performance of gradient systems been sufficient to take full advantage of this tool. There are essentially four ways in which gradients are used: coherence pathway selection, spatial encoding, diffusion weighting and spoiling are all used in modern high-resolution systems. These methods have common and differentiating elements. Coherence selection and diffusion weighting take advantage of the reversible behaviour of the pulse gradient effect. Spoiling is a subset of coherence pathway selection that requires no encoding gradient and hence no read or rephase gradient. Spatial encoding can be used to image and correct $B_0$ inhomogeneity, and can be used to restrict the detected sample volume. The basic elements of $B_0$ field gradients, as used in high-resolution NMR are described.

## Basic properties of gradients

On a typical high-resolution NMR system, a $B_0$ gradient probe can transiently generate a linear change in the otherwise homogeneous $B_0$ field of ±1 mT or moreover the approximately 2 cm $z$ sample length. Many gradient systems can also independently generate linear transverse ($x$ and $y$) gradient fields of similar magnitudes. Linearity, switching speed and gradient recovery times are important gradient performance criteria. The switching time or recovery time was the most significant limitation of early gradient systems. In these early designs, the gradient field was not constrained inside the gradient cylinder, as shown in **Figure 1A**, and the process of generating a transient gradient interval induced undesirable currents in nearby conductors, especially the components of the magnet itself. These induced eddy currents in turn generate magnetic fields that perturb the NMR spectrum. These stray fields cause significant spectral distortions and last for hundreds of milliseconds. It was therefore impossible to maintain reasonable timing in multiple-pulse NMR experiments using this type of gradient. The invention of the actively shielded gradient coil in the late 1980s removed this limitation by constraining the gradient field inside the gradient cylinder, as illustrated in **Figure 1B**. This innovation and the development of dedicated high-resolution NMR gradient probes have made this technology readily available to NMR spectroscopists.

## The gradient pulse effect

A gradient in the $B_0$ field across a sample will cause the spins in that sample to precess at spatially dependent rates. More specifically, a gradient pulse will add a reversible, spatially dependent, and coherence order-dependent, phase to the magnetization:

$$\varphi_r = \gamma prGt \qquad [1]$$

where $\gamma$ is the magnetogyric ratio, $r$ is the distance from gradient isocentre, $G$ is the gradient amplitude, $t$ is the gradient pulse duration and $p$ is coherence

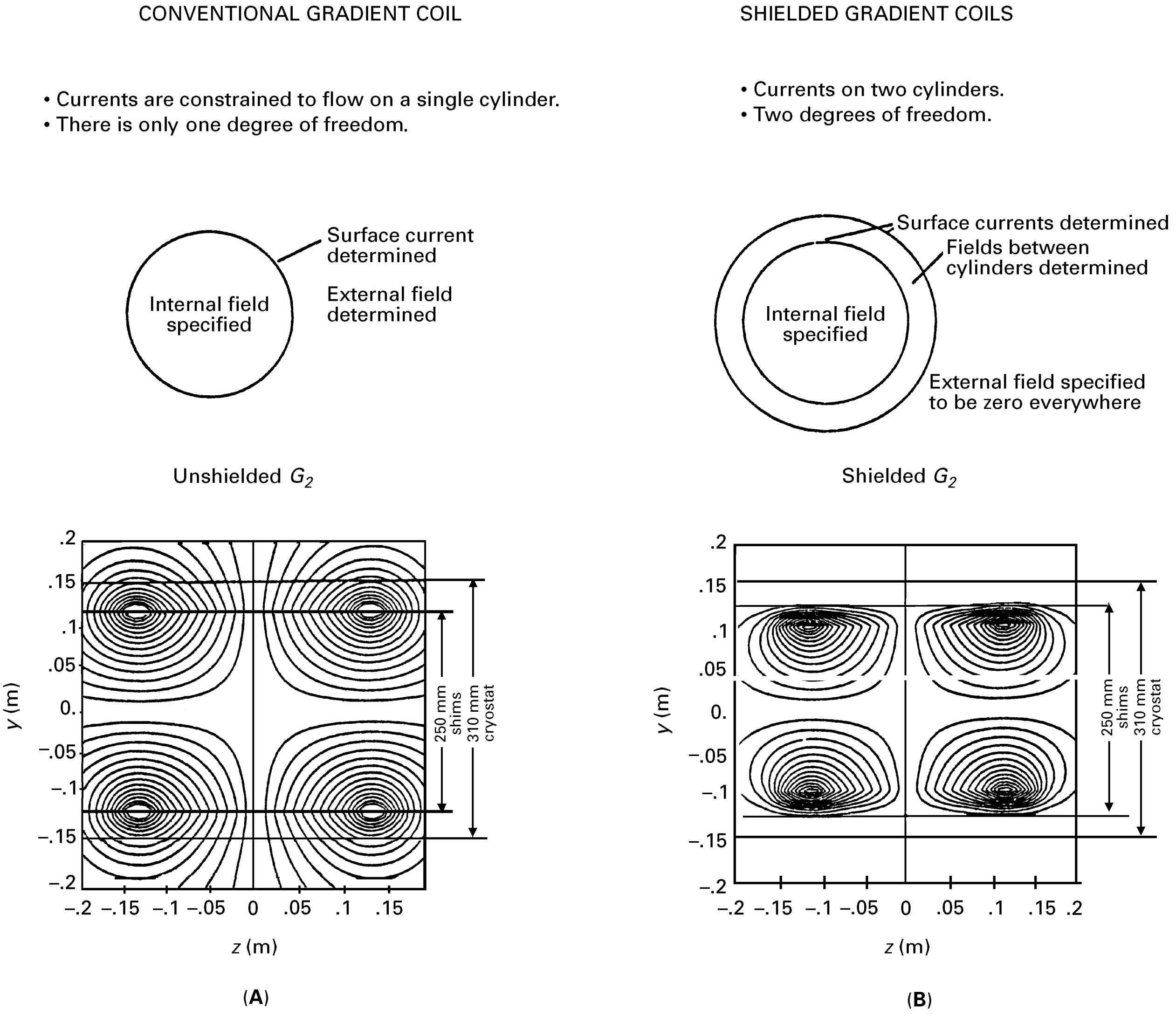



**Figure 1** (A) Current diagram for conventional unshielded *z* gradient coil. (B) Current diagram for actively shielded *z* gradient coil.

order. If not resolved in space, or rephased, the impact of a spatially dependent phase across the detected volume is self-cancellation of signal. In the absence of $B_1$, radiofrequency magnetic field inhomogeneity and susceptibility shifts, in a detected region $\pm\frac{1}{2}\, r_{\max}$, the impact of a gradient pulse on pure *x* magnetization ($p = 1$) will be:

$$M_x(t) = 2\gamma G r_{\max} t \sin\left(\tfrac{1}{2}\, \gamma G r_{\max} t\right) \qquad [2]$$

Under these idealized conditions, perfect cancellation occurs at multiples of $\phi(r_{\max}) = 2\pi$ but practical dephasing requires many cycles of $2\pi$, where residual signal can be approximated as, $2/(\gamma G t r_{\max})$. Thus, a typical 1 ms 0.25 T m$^{-1}$ gradient pulse over a 1.5 cm $B_1$ sample volume would reduce the observable proton signal by a factor of about 1000. For pathway selection, it is the difference in gradient integrals that determine the level of suppression. Of course, practical matters such as gradient linearity (especially the fall-off at the ends of the sample volume), and $B_1$ homogeneity, will determine the actual suppression.

## Coherence, coherence order and pathway selection

Coherence is a generalization of the idea of transverse magnetization. Coherence order is the quantum number difference associated with the *z* component of the rotation generated by the RF excitation, and can only be changed by another RF pulse. Thus, coherence order is conserved in the time periods separating RF pulses, during which the

application of a gradient pulse will encode magnetization according to the coherence order of that interval. The route of the observed magnetization is referred to as the coherence transfer pathway. All pathways start at $p = 0$ (thermal equilibrium) and must end with single-quantum coherence to be detectable. Transverse magnetization is a specific type of coherence characterized by the single-quantum coherence levels $p = +1$ and –1. Both components are detected to distinguish positive and negative frequencies in a quadrature receiver. By convention, $p = -1$ represents the quadrature detected signal, $s^+(t) = s_x(t) + is_y(t)$.

Coherence transfer pathway diagrams are a good way to visualize the need for pathway selection in multiple-pulse NMR experiments. These pathways remind us that each RF pulse transfers magnetization to multiple coherence levels, only one or two of which must be retained to end up with the desired artefact-free spectrum. The traditional way to select a given pathway is to apply phase cycling. With phase cycling, the pulse sequence is repeated, using changes in the phase of the RF pulses, along with addition or subtraction of the corresponding complex signals to retain the desired pathway and cancel all the others. As a difference method, phase cycling can become a problem when the desired pathway is much smaller than the unwanted ones, as is the case in many multiple-quantum experiments. As a nondifference method, pulsed field gradient selection of the pathway is an advantage in these cases.

Coherence transfer pathways are also a convenient way to visualize the action of gradient pulses in a NMR sequence, since the spatial encoding of each interval in the pathway is directly proportional to the product of gradient integral, $Gt$, and coherence order, $p$. Any pathway in which the sum

$$\sum_i p_i(Gt)_i = 0 \qquad [3]$$

will be passed. Pathways where this is not true will retain a spatially dependent phase and will self-cancel.

The pathway for homonuclear correlation spectroscopy (COSY) is shown in **Figure 2** and provides a simple example. The first pulse creates coherence with orders +1 and –1 and leaves some z magnetization as coherence order 0. Thus, there are three pathways by which the coherence can reach the receiver after the second RF pulse, namely [0 → 0 → –1], [0 → + 1 → –1] and [0 → –1 → –1]. If the RF carrier is placed on one side of the F2 spectrum, all of the peaks in the 2D spectrum corresponding to the [0 → –1 → –1] coherence pathway will lie on one side of $F_1 = 0$ and the peaks from [0 → +1 → –1] will lie on the other side of $F_1 = 0$. The [0 → 0 → –1] peaks will occur only at $F_1 = 0$. A single gradient pulse placed between the two RF pulses will spoil coherence that passes through both $p = +1$ and $p = -1$, and will select the [0 → 0 → –1] pathway. The addition of a read gradient interval after the second RF pulse will allow one of the other two pathways to be selected. If the read gradient is equal in sign and integral to the first (encode) gradient, then the pathway that goes through the [0 → +1 → –1] transfer will be selected, while the coherence that remains at –1 during evolution and acquisition [0 → –1 → –1] will be selected by a gradient of equal integral but opposite sign.

## Multiple-quantum coherence transfer selection

A common usage of pulsed field gradients is multiple-quantum coherence transfer selection, which takes advantage of the nondifference filtering of large unwanted signals from the small desired ones. The simplest homonuclear example is the three pulse sequence shown in **Figure 3**. Homonuclear scalar coupled spins will give rise to both double and zero quantum coherences in the mixing time ($t_m$)

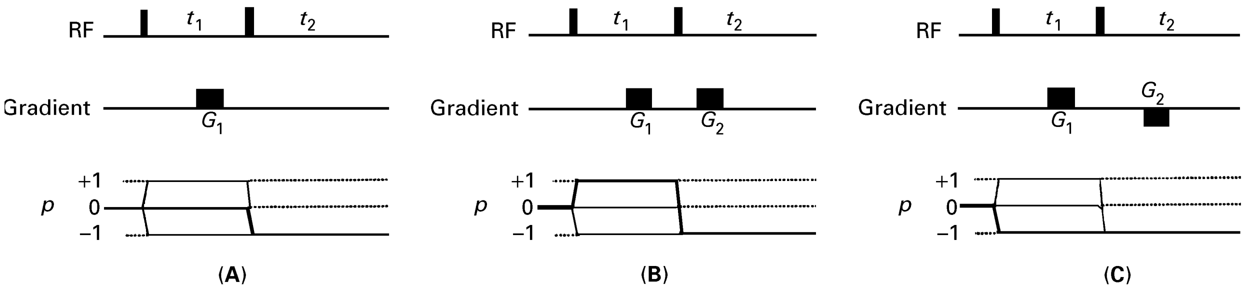



**Figure 2** Coherence-transfer pathway diagrams for COSY, illustrating gradient selection of (A) the $F_1 = 0$ artefacts only, [0 → 0 → –1]; (B) N-type signals, [0 → +1 → –1]; and (C) P-type signals, [0 → –1 → –1].

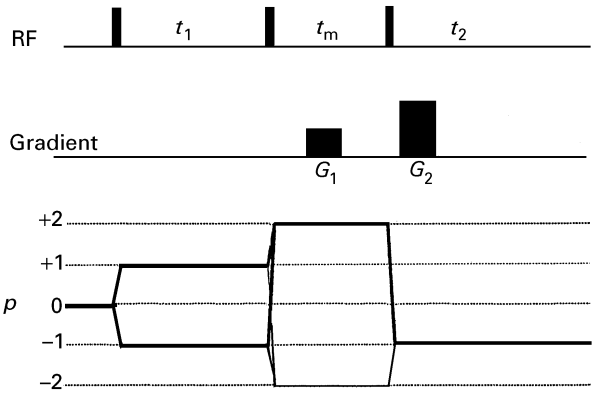



**Figure 3** Coherence transfer pathway diagram for homonuclear double-quantum selection with gradients.

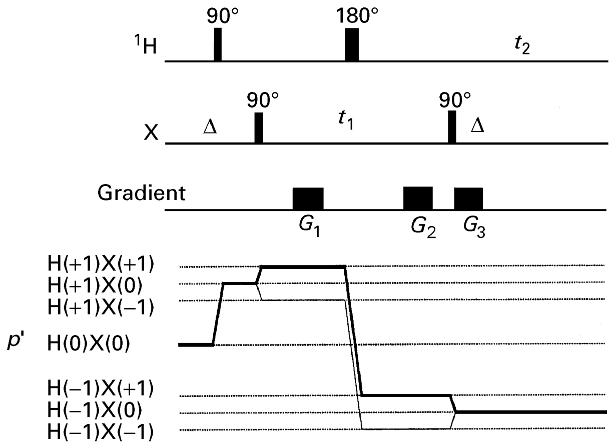



**Figure 4** Coherence transfer pathway diagram for gradient-enhanced HMQC sequence. The pathway illustrated by the solid line selects the pathway through heteronuclear double-quantum [H(+1):X(+1)] and heteronuclear zero quantum [H(–1):X(+1)] levels. For X = $^{13}$C, this pathway can be selected by any of the gradient ratios 4:0:5, 0:4:–3 or 4:4:2.

interval. Uncoupled spins, such as solvent water, will not and thus the use of a 1:2 ratio of gradients ($G_1$:$G_2$) will select only coupled spin coherence that goes through the [0 → +/–1 → +2 → –1] pathways. A spoiler gradient ($G_1$ only) during the mixing time will select for pathways that go though $p = 0$ in the mixing time, selecting the zero-quantum pathway plus any residual $z$ magnetization. A single pathway [0 → +1 → +2 → –1] is selected using a 1:1:3 gradient sequence. This is achieved by adding a gradient of integral $I$ during $t_1$ evolution time and increasing the read gradient at the start of the $t_2$ interval by the same area. These methods are all very good at suppressing the uncoupled water signal and reducing $t_1$-noise. However, both double-quantum and zero-quantum sequences may also pass water or other large solvent signals via the dipolar field effect unless the gradients are oriented at the magic angle 54.7° where triple axis gradients are used and ($G_x = G_y = G_z$). Heteronuclear multiple-quantum selections, often for protons attached to a lower magnetogyric ratio nucleus, are also very common applications of gradient selection. In this case it is often convenient to generate a combined coherence transfer pathway diagram for coupling partners and to use normalized heteronuclear coherence order $p'$, scaled to the proton magnetogyric ratio. The resulting normalized coherence levels are then directly related to the sensitivity to pulsed field gradient integrals. The coherence pathway diagram for the gradient-enhanced heteronuclear multiple quantum correlation (HMQC) experiment is illustrated in **Figure 4**. For X = $^{13}$C, the initial heteronuclear double-quantum level $p'$ [H(+1): $^{13}$C(+1)] = 1.25, and the initial heteronuclear zero-quantum level $p'$ [H(–1): $^{13}$C(+1)] = –0.75. In this example, gradient ratios of 4:0:5 or 0:4:–3 or 4:4:2 would all select for the same pathway through the double → zero-quantum transfer, and spoil the zero → double-quantum transfer pathway, as well as coherence pathways for proton spins not coupled to a $^{13}$C nuclei.

As in the homonuclear case, this method provides excellent water suppression. The suppression of the $t_1$-noise artefacts is so good with these methods that data can be collected under conditions that are not possible with traditional phase cycled methods. This advantage has been exploited especially in long-range proton–carbon correlation studies of polymer branching, as illustrated in **Figure 5**, and for proton–proton correlation at the water chemical shift frequency.

## Spin echoes and gradient pulses

Spin-echo selection with gradient pulses was the first and is probably now the most common use of gradients in magnetic resonance. This element is common to MR imaging, localized spectroscopy, diffusion measurements, water suppression and artefact reduction in multiple-pulse NMR. On high-resolution spectrometers, where all of the $B_1$ sample volume is normally detected, RF refocusing pulses produce a considerable fraction of non-$\pi$ rotation. The placement of equal gradient pulses on either side of the $\pi$ pulse, as illustrated in **Figure 6A**, filters out any coherence that does not refocus ($p \rightarrow -p$ transition). This is also an especially effective method for improving the performance of frequency-selective $\pi$ pulses such as are used in the gradient-enhanced version of spin echo water suppression (SEWS). Gradients of equal integral, but opposite sign, placed on

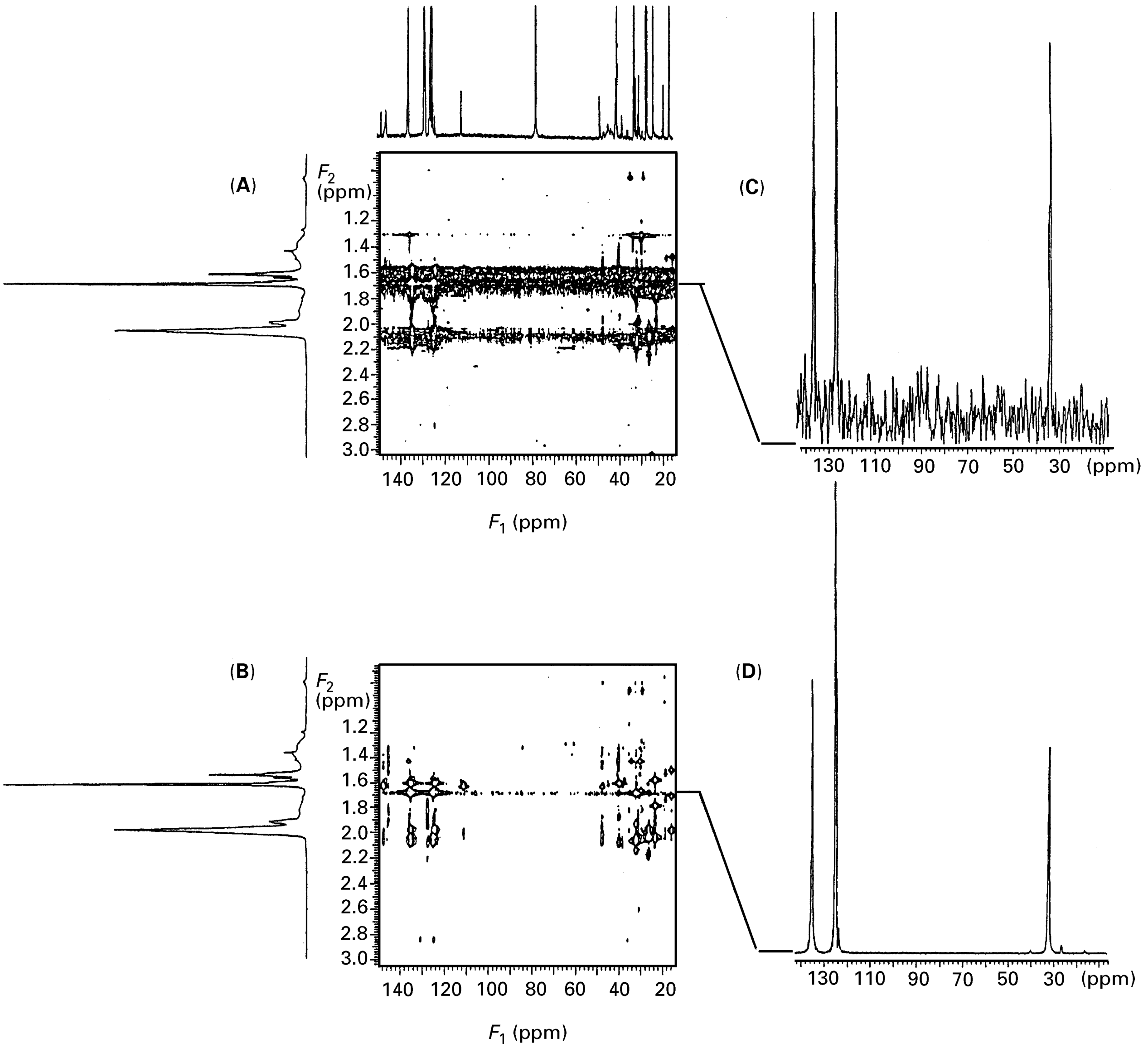



**Figure 5** Gradient (B and D) versus phase-cycled (A and C) HMBC spectra of the polymer PI-b-PS. The comparative traces at $F_2$ = 1.7 ppm show the far superior signal-to-$t_1$-noise achieved by the gradient method (D) relative to the traditional phase-cycled approach (C). Reproduced with permission of The Society of Chemical Industry and John Wiley & Sons, from Rinaldi P, Ray DG, Litman V and Keifer P (1995) The utility of pulse-field gradient–HMBC indirect detection NMR experiments for polymer structure determination. *Polymer International* **36**: 177–185.

either side of a chemical shift selective refocusing pulse, such as the 1–1 binomial example shown in **Figure 6B**, are a powerful way to capture a selective, refocused ($p \rightarrow -p$ transition) bandwidth. This approach can be used to dramatically avoid residual out-of-band signal (e.g. water) relative to the phase-cycled method. Frequency-selective suppression using spin echoes and gradients has also proved very successful in methods such as WATERGATE (water suppression by gradient tailored excitation) and MEGA as illustrated in **Figures 6C and 6D**. In addition to $p \rightarrow -p$ transfer, $\pi$ pulses invert $z$ magnetization, $I_z \rightarrow -I_z$. In this case imperfect $\pi$ pulses will generate transverse magnetization. To select for the $I_z \rightarrow -I_z$ transition and spoil both transverse magnetization and the $p \rightarrow -p$ refocused magnetization, nonequal gradients can be applied before and after a $\pi$ inversion pulse (**Figure 6E**). Each gradient pulse will spoil any transverse magnetization during those intervals, and the nonequal integrals of the gradients will prevent the refocusing of the $p \rightarrow -p$ transition. The selection of $I_z \rightarrow -I_z$ transitions are also useful in multinuclear experiments, in which case the gradient dephasing of $S$ coherences must be avoided. This can

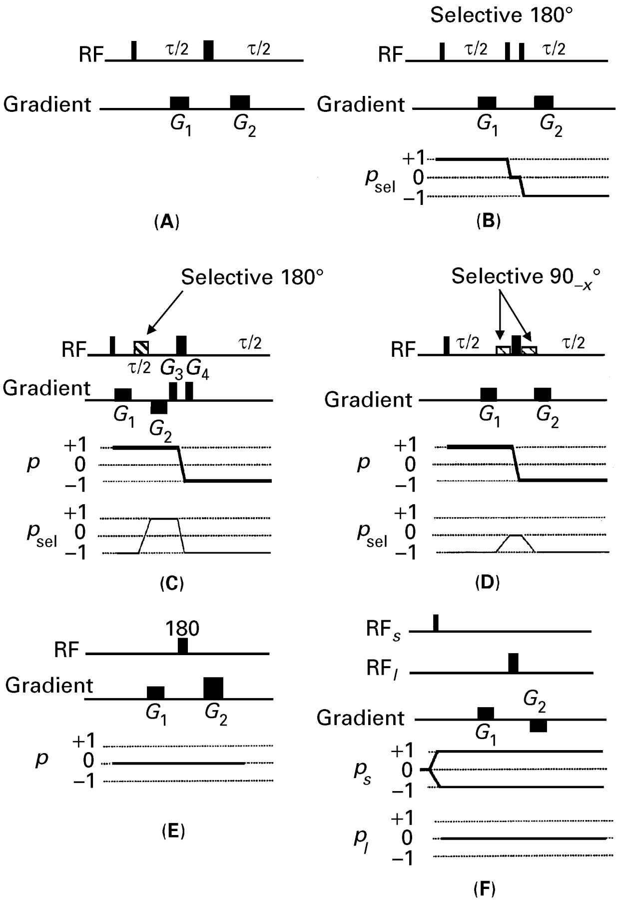



**Figure 6** The use of gradients with RF $\pi$ pulses. (A) Standard spin echo selection, $p \rightarrow -p$ transitions are selected. Any imperfection in the RF refocusing is cancelled. (B) Frequency-selective spin echo selection. Only the spins in the refocused bandwidth are selected. (C) Pathway for MEGA. Spins that are refocused by the selective RF refocusing pulse are dephased by the $G_1$:$G_2$ gradient pair. Outside the frequency-selective bandwidth, $G_2$ reverses the effect of $G_1$. (D) Pathway for WATERGATE. A net zero RF rotation leaves signals in the frequency-selective bandwidth dephased by the $G_1$:$G_2$ pair, while spins outside the selective bandwidth are rephased as in (A). (E) Selection of $I_z$ to $-I_z$ ($p = 0 \rightarrow 0$) transitions uses nonequal gradients prior to the $\pi$ pulse to eliminate any existing transverse magnetization, and after the $\pi$ pulse to eliminate any transverse magnetization generated by RF pulse inhomogeneity. (F) selection of $I_z$ to $-I_z$ ($p = 0 \rightarrow 0$) in a heteronuclear sequence while preserving any nonzero $S$ coherence levels.

be accomplished by using gradients of equal integrals but opposite sign (**Figure 6F**). The second gradient will reverse any accumulated phase for the $S$ spin caused by the first gradient, but will still spoil all $I$ spin coherences except of the $I_z \rightarrow -I_z$ transitions.

## Spoiling

A gradient spoiler pulse can be applied in intervals where the desired signal has coherence order $p = 0$. These applications include gradient-enhanced $z$ and $zz$ filters, stimulated echo selection, multiple-quantum suppression during NOESY (nuclear Overhauser effect spectroscopy) mixing times and the homonuclear zero-quantum methods as previously described. Two examples of this gradient element are illustrated in **Figure 7**.

The gradient-enhanced $z$ filter is a pulse field gradient version of the multiple-acquisition nongradient method. In the original method, magnetization is stored as $I_z$, and multiple delay times are collected to allow non-$I_z$ magnetization to evolve and self-cancel. The gradient method accomplishes this in a single step. As shown in **Figure 7A**, the two-pulse RF filter acts as a $\pi$ pulse for the desired magnetization, which means the spoiler during the $I_z$ interval can be combined with a spin-echo gradient pulse pair outside the $I_z$ interval.

The gradient-enhanced $zz$ filter selects for heteronuclear longitudinal spin order described by the density operator $I_zS_z$, and can be easily integrated into the heteronuclear single quantum correlation (HSQC) type sequences, or as a preparation period for HMQC methods. The gradient version of the $zz$ filter as shown in **Figure 7B** also passes $I_z$ magnetization and is not as efficient at rejecting unwanted pathways as coherence selection.

Spoiler gradients can also be used following frequency-selective excitation to eliminate a narrow band of chemical shift. This approach is often referred to as a chemical shift selective (CHESS) pulse. Optimum performance requires the tip angle of the selective excitation pulse be adjusted for water $T_1$ relaxation that occurs during the excitation–dephase intervals. Multiple excitation–dephase intervals can be concatenated to achieve a moderate level of $B_1$ and $T_1$ insensitivity. Alternatively, $T_1$- recovery time and water excitation flip angle can be adjusted to exploit differences in the solute and the water $T_1$ values and to allow significant recovery of the solute spins during the time it takes water to reach a null. Like many gradient methods, the $T_1$-delayed CHESS pulse inherently eliminates the radiation damping effect and makes it possible to take advantage of the true water $T_1$.

## Diffusion-weighted water suppression

In any experiment where gradients are used to label spins with a spatially dependent phase, that are subsequently rephased with a second gradient pulse, there will be a loss of signal due to any movement of the spins during the time interval between labelling and rephasing. For a spin-echo ($p$ +1 $\rightarrow$ −1) transition, this loss of signal is related to translational

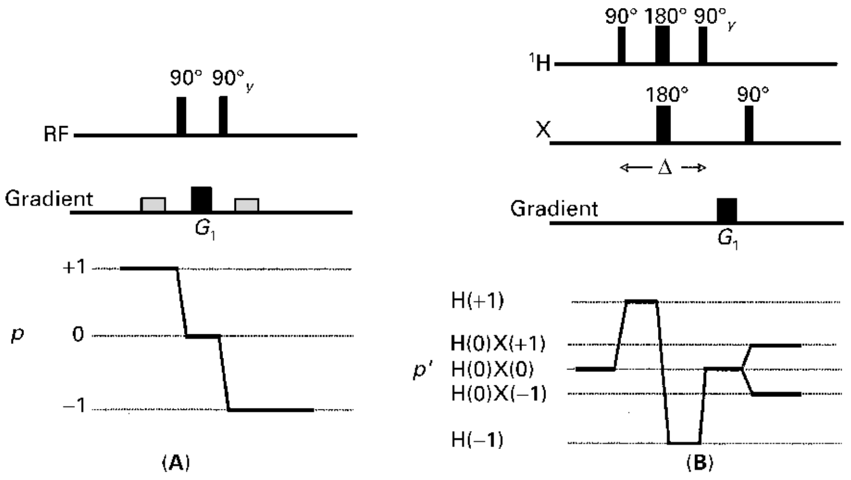



**Figure 7** (A) Gradient-enhanced *z* filter and (B) *zz* filter. As in the inversion examples shown in **Figure 6**, these gradients dephase all but $p = 0$ coherence order.

diffusion by the Stejskal–Tanner equation:

$$\ln(S/S_0) = -\gamma^2 g^2 \tau^2 (\Delta - \left(\tfrac{1}{3}\tau\right) D \qquad [4]$$

where $\gamma$ is the magnetogyric ratio, $g$ is the strength, and $\tau$ the duration of the gradient pulse pair, $\Delta$ is the time between gradient pulses, and $D$ is the diffusion coefficient. Normally, diffusion weighting is minimized by using modest gradient integrals ($g$ and $\tau$) and by keeping the separation ($\Delta$) between the encode and rephase portion of a gradient pair small. However, by increasing both gradients integrals and separation ($\Delta$), it is possible to take advantage of the significant differences in the translational diffusion of solvent water and large solute molecules such as proteins. This is the basis of the DRYCLEAN, diffusion reduced water signals in spectroscopy of molecules moving slower than water. With a modest 20-fold difference in diffusion constant, $D$, a gradient pair could be selected to preserve over 70% of the solute signal, while suppressing water by 1000-fold. It is important to note that, like multiple-quantum coherence pathway selection, this method is also independent of the width and shape of the water signal. The same basic gradient-selected spin-echo methodology is also used to study exchange processes in biomolecules.

## Phase-sensitive methods

Modern multidimensional spectra are almost always recorded in pure absorption mode. The primary reasons are phase sensitivity, improved resolution, and a $\sqrt{2}$ factor increase in SNR compared with magnitude mode. Pure absorption phase is obtained from the amplitude-modulated signal in $t_1$, separating the frequencies of the two mirror image pathways, $p = +1$ and $p = -1$ in an evolution time analogue to quadrature detection. In methods without gradients, or in methods that use gradients only for spoiling, spin echo and/or $I_z$ inversion selection, this is accomplished using a two-step phase cycle for each $t_1$ increment. Both steps contain $p = +1$ and $p = -1$ coherence, and the combination provides frequency discrimination at full signal intensity. Pure absorption line shape with gradient selection during evolution is also a two-step process, but each acquisition contains only $p = +1$ or $p = -1$, leading to a $\sqrt{2}$ factor loss in SNR relative to the phase-cycled selection of quadrature in $F_1$. The trade-off is that signal-to-$t_1$ noise is often better for the gradient methods, as illustrated in **Figure 5**, and in the instances where pure absorption line shape is not required the gradient selection methods are significantly faster. Unlike the single-step selection with gradients, phase-cycled methods require multiple steps to separate $p = +1$ or $-1$. The advantage of a reduction in required phase cycle steps is most evident in three- and four-dimensional NMR studies, where proper sampling of the evolution time alone generates more signal averaging than necessary. It is possible to collect separate $p = +1$ and $p = -1$ pathways in a single acquisition per $t_1$ time by using the switched acquisition time (SWAT) gradient method. In this method the two coherence pathways are alternately and individually acquired on alternate sampling points in the digitizer. Although a doubling in $F_1$ bandwidth still results in the $\sqrt{2}$ factor loss in SNR, this approach offers the ability to collect pure absorption multidimensional data in a minimum total acquisition time.

The method, however, is very demanding on gradient switching time.

## Spatial selection

In the typical high-resolution NMR experiment, the entire $B_1$ volume contributes to the final result. The volume of spins in the transition band where the relatively linear $B_1$ field falls from maximum to zero can be significant. This inhomogeneity and line shape distortions from bulk susceptibility effects also found at the ends of the sample are among the reasons why gradient selection and phase cycling methods are so heavily used for artefact reduction in multiple-pulse NMR. A gradient-based method that can be used to reduce these end effects (transition band suppression or TBS), uses slice select–spoil intervals during the pulse sequence preparation period to avoid this difficult region of the sample. Combination of TBS with $T_1$-delayed CHESS pulses and gradient selection of double-quantum coherence makes it possible to study proton–proton correlations at the water chemical shift in both $F_1$ and $F_2$. The pulse sequence and an example are shown in **Figure 8**.

Another application of spatial localization in high-resolution NMR applies to the specialized field of high-performance liquid chromatography–NMR (HPLC-NMR). With NMR as the detector for a liquid chromatography system, it can be valuable to spatially resolve the NMR sample volume. This can be done by phase encoding, which allows the data to be retrospectively processed to eliminate end effects or to separate partially overlapping HPLC fractions.

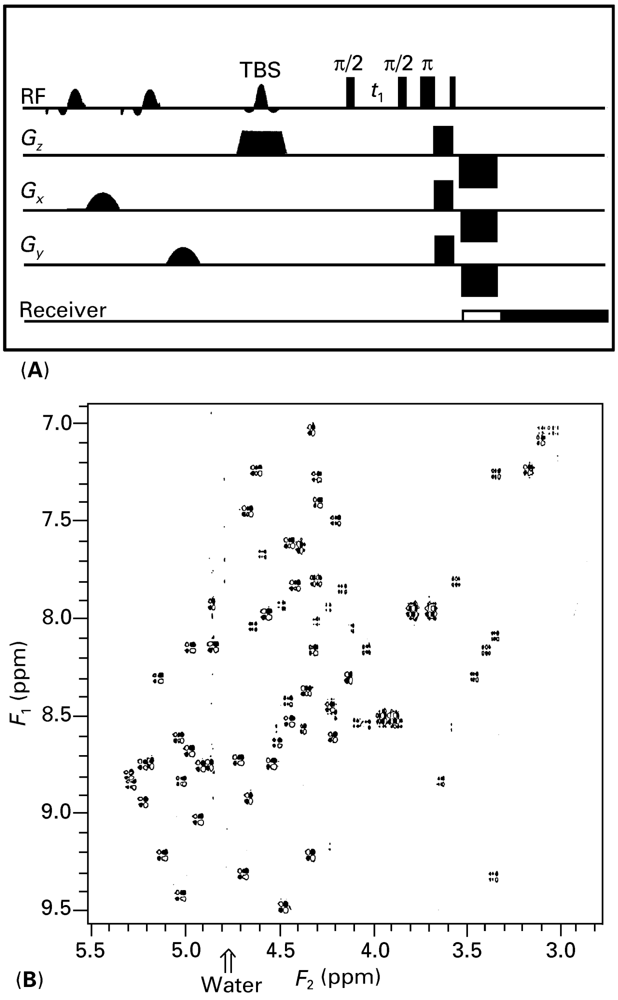



**Figure 8** (A) Pulse sequence for phase-sensitive version of gradient-enhanced double-quantum correlation method incorporating $T_1$ delayed CHESS sequence and TBS. (B) Phase-sensitive contour plot of data for 1 mM ubiquitin in 90% $H_2O$–10% $D_2O$ collected using this method. A water inversion null time of 200 ms was used to allow C$\alpha$H protons at 4.8 ppm to recover fully as expansion near $F_2 = 4.8$ (water) illustrates. Reproduced from Hurd R, John B, Webb P and Plant P (1992) *Journal of Magnetic Resonance* **99**: 632–637 with permission of Academic Press.

## Field maps and homogeneity adjustment

One relatively obvious application of three-axis gradients is to image and correct for any $B_0$ field inhomogeneity. Three-dimensional phase or frequency maps can be obtained and used to image the inhomogeneity of the sample, and with previously obtained maps for fixed offsets of the known shims a best-field solution can rapidly be made. Normally this approach works best with a strong solvent signal such as water, but in a limited way it can be accomplished using the deuterium-lock solvent signal.

## Summary

Gradients are useful as an integral part of multiple-pulse NMR methods. High-resolution NMR systems and probes continue to incorporate these devices and accordingly the use of these tools continues to become more common. In many ways, gradients are a perfect partner for the limitations of the native high–resolution NMR $B_1$ fields, and also work complement only to crafted RF pulse methods. When used appropriately, gradients have the ability to enhance the quality of most multiple-pulse NMR results.

## List of symbols

$B_0$ = applied magnetic field; $B_1$ = RF magnetic field strength; $D$ = diffusion coefficient; $F_1$ = evolution frequency; $g$ = strength of gradient pulse; $G$ = gradient amplitude; $I_z = z$ magnetization; $p$ = coherence order; $r$ = distance from gradient isocentre;

$t$ = gradient pulse duration; $\gamma$ = magnetogyric ratio; $\Delta$ = time between gradient pulses; $\tau$ = duration of gradient pulse; $\phi$ = magnetization phase.

*See also:* **Diffusion Studied Using NMR Spectroscopy; NMR Pulse Sequences; Product Operator Formalism in NMR; Solvent Suppression Methods in NMR Spectroscopy; Two-Dimensional NMR, Methods.**

## Further reading

Freeman D and Hurd RE (1992) Metabolite specific methods using double quantum coherence transfer spectroscopy. In Diehl P, Fluck E, Günther H, Kosfeld R and Seelig J (eds) *NMR: Basic Principles and Progress*, Vol 27, pp 200–222. Berlin: Springer-Verlag.

Hurd RE (1995) Field gradients and their application. In Grant DM and Harris K (eds) *Encyclopedia of NMR*. Chichester: Wiley.

Hurd RE and Freeman D (1991) Proton editing and imaging of lactate. *NMR in Biomedicine* **4**: 73–80.

Keeler J, Clowes RT, Davis AL and Laue ED (1994) Pulsed-field gradients: theory and practice. *Methods in Enzymology* **239**: 145–207.

Zhu J-M and Smith ICP (1995) Selection of coherence transfer pathways by pulsed field gradients in NMR spectroscopy. *Concepts in NMR* **7**: 281–288.

# Magnetic Resonance, Historical Perspective

**J W Emsley**, University of Southampton, UK
**J Feeney**, National Institute for Medical Research, London, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| Historical Overview |

## Introduction

NMR dates from 1938 when Rabi and co-workers first observed the phenomenon in molecular beams. This was followed in 1946 by the NMR work in the laboratories of Bloch and Purcell on condensed-phase samples. In the intervening 53 years there has been a wonderful revelation of how rich this spectroscopy can be, and only a flavour can be given here of the many significant developments. A very detailed account of the history is given in Volume 1 of the *Encyclopaedia of NMR*, which also includes biographies of many of those who created the subject as it is today. A shorter, but still very detailed, history can be found in five articles published in *Progress in NMR Spectroscopy*. Here we present a summary of the main developments under five headings: Establishing the principles; Solid-state and liquid crystal NMR; Liquid-state NMR; Biological applications of NMR; Magnetic resonance imaging. We also present three tables that give some of the important milestones in the development of NMR.

## Establishing the principles

NMR arises because some nuclei may have an intrinsic spin angular momentum, which has the consequence that they also have a magnetic dipole moment. The existence of a magnetic dipole moment for hydrogen nuclei was established in 1933 by Gerlach and Stern, who observed the effect of an applied magnetic field gradient on a beam of hydrogen molecules. The trajectories of the molecules are changed if their nuclei have magnetic moments. Inducing transitions between nuclear spin states by the application of electromagnetic radiation at the appropriate resonance frequency was introduced by Rabi and co-workers, also using molecular beams. In this experiment the beam passed first through a field gradient, which deflected the atoms in a direction dependent on the value of $m$, the magnetic quantum number, then through a homogeneous field, where they were subjected to the electromagnetic radiation, and finally through another field gradient whose sign was opposite to that in the first region. If the nuclei in the atoms do not absorb the radiation, then the effect of the two field gradients cancels, and the beam is undeflected. Absorption or emission of radiation leads to a net deflection of the beam. This simple experiment therefore provided a foretaste of the use of gradients to create or destroy signals.

The first successful NMR experiments on condensed-phase samples were done in 1945, and published in 1946, separately by the group at Stanford led by Bloch, who observed the protons in water, and a group at Harvard led by Purcell, who

also observed protons, but in solid paraffin. Unlike the beam experiments, in these the detection was of a net nuclear magnetization arising from the imbalance between states with different values of *m*, and it was crucial for their success that nuclear spin relaxation was occurring at a favourable rate. The first systematic experimental measurements of spin–lattice and spin–spin relaxation rates were published in 1948 by Bloembergen, Purcell and Pound, who also gave an interpretation of their magnitudes in terms of the dynamics of the molecules containing the nuclei.

The magnetic shielding of a nucleus from the applied field by the surrounding electrons was recognized to occur in atoms by Lamb, who published a method for calculating the effect in 1941. The aim of these calculations was to correct for the effect of the shielding on the resonance frequencies observed in molecular beam experiments and hence to obtain the true nuclear magnetic moment. Ramsey extended these calculations to nuclei in molecules in 1949–52, and in this same period the phenomenon was observed in the NMR of condensed-phase samples, first in the resonances of metals and metal salts by Knight in 1949, and in the following year by Proctor and Yu, who observed different resonances for $^{14}N$ in ammonium nitrate, and by Dickinson, who reported the same phenomenon for $^{19}F$ in various compounds (e.g $BeF_2$, HF, $BF_3$, KF, NaF, $C_2F_3CCl_3$). The physicists working on these problems thought these ‘chemical shift’ effects uninteresting and a nuisance, since they impeded the important task of measuring nuclear gyromagnetic ratios accurately! We can now pinpoint the years 1949–50 as the period when NMR ceased to be predominantly a technique of the ‘physicists’ and when the ‘chemists’ began to realize the potential usefulness of the ‘chemical shift’.

It was also at this time that the effects produced by spin–spin coupling were first observed. Experimental results now preceded theory. Proctor and Yu observed a multiplet for the $^{121}Sb$ resonance in a solution containing the ion $SbF_6^-$. They observed only five lines of the seven-line multiplet, and so were sidetracked into attempting to explain the splitting as incomplete averaging of the internuclear dipolar coupling. Dipolar coupling had been observed in molecular beam experiments, and its origin was well understood. The problem facing Proctor and Yu was that this interaction, being entirely anisotropic, should vanish if the molecules are rotating rapidly and isotropically, as in an isotropic liquid sample. Gutowsky and McCall also observed spin–spin splittings, but this time in the $^{31}P$ and $^{19}F$ resonances in the compounds $POCl_2F$, $POClF_2$ and $CH_3OPF_3$. They were able to deduce that the number of lines is determined by the product of the *m* values of the coupled nuclei. Ramsey and Purcell published an explanation of the splitting as arising from a rotationally invariant interaction between nuclear spins that proceeds via the electrons in the molecule. Spin–spin splitting was also observed at the same time by Hahn and Maxwell as a modulation on a spin-echo signal, Hahn having discovered the spin-echo phenomenon in 1949.

By 1952 all the basic, important interactions that affect NMR spectra had been demonstrated, and their relationship with molecular structure had been explained (see **Table 1**). The challenge then, as now, was how to exploit the value of NMR for samples of varying degrees of complexity, and this proved to be an exciting and rewarding quest. There were still many new effects to be discovered, and these began to appear quickly as the early pioneers started to explore this new spectroscopy. In 1953 Overhauser predicted that it should be possible to transfer spin polarization from electrons to nuclei. He delivered this prediction to an initially sceptical audience at a meeting of the American Physical Society in Washington, DC. Overhauser was a postdoctoral worker in Illinois when he made this prediction, and he had interested Slichter in the possibility of enhancing NMR signals in this way. Slichter and Carver succeeded in demonstrating the enhancement in lithium metal and all doubts about the nuclear Overhauser effect (NOE) were put to rest.

The chemical shift and spin–spin coupling phenomena were clearly destined to be discovered as soon as magnets became sufficiently homogeneous. They might be classed as inevitable discoveries. The Overhauser effect is different, and it is conceivable that it would have lain undiscovered for many years without the perception of one individual. We might call this a noninevitable discovery. One of the remarkable features of NMR development has been the number of such noninevitable discoveries, some of which have been fully exploited only many years after their discovery. Another such example is the invention by Redfield in 1955 of spin locking, a technique that produces a retardation of spin–spin relaxation in the presence of a radiofrequency field. This not only led to a method of studying slow molecular motions, but also provided a method for transferring polarization between two nuclei that are simultaneously spin-locked, as ingeniously demonstrated by Hahn and Hartmann in 1962.

There were many developments going on in the period 1955–65, some of which we will discuss later. The successes of the early pioneers encouraged the development of commercial spectrometers, and this provided increased access to NMR for a wider

**Table 1** Milestones in the development of NMR basic principles and solid–state and liquid-crystal NMR

| *Date* | *Milestone* | *Literature citation* |
|---|---|---|
| 1924–1939 | Early work characterizing nuclear magnetic moments and using beam methods | Frisch and Stern, *Z. Phys.* **85**: 4; Esterman and Stern *Phys. Rev.* **45**: 761, Rabi *et al.*, *Phys. Rev.* **55**: 526 |
| 1938 | First NMR experiment using molecular beam method | Rabi *et al.*, *Phys. Rev.* **53**: 318 |
| 1941 | Theory of magnetic shielding of nuclei in atoms | Lamb, *Phys. Rev.* **60**: 817 |
| 1945 | Detection of NMR signals in bulk materials | Bloch *et al.*, *Phys. Rev.* **69**:127; Purcell *et al.*, *Phys. Rev.* **69**: 37 |
| 1948 | Bloembergen, Purcell and Pound (BPP) paper on relaxation | Bloembergen *et al.*, *Phys. Rev.* **73**: 679 |
| 1949 | Hahn spin echoes | Hahn, *Bull. Am. Phys. Soc.* **24**: 13 |
| 1949 | Knight shift in metals | Knight, *Phys. Rev.* **76**: 1259 |
| 1950 | Discovery of the chemical shift | Proctor and Yu, *Phys. Rev.* **77**: 717; Dickinson *Phys. Rev.* **77**: 736 |
| 1951 | Discovery of spin–spin coupling | Proctor and Yu, *Phys. Rev.* **81**: 20; Gutowsky and McCall, *Phys. Rev.* **82**: 748; Hahn and Maxwell, *Phys. Rev.* **84**:1246 |
| 1952 | First commercial NMR spectrometer (30 MHz) | Varian |
| 1953 | Bloch equations for NMR relaxation | Bloch *et al.*, *Phys. Rev.* **69**: 127; Bloch, *Phys. Rev.* **94**: 496 |
| 1953 | Overhauser effect | Overhauser, *Phys. Rev.* **91**: 476; Carver and Slichter, *Phys. Rev.* **102**: 975 |
| 1953 | Theory for exchange effects in NMR spectra | Gutowsky *et al.*, *J. Chem. Phys.* **21**: 279 |
| 1953 | Proton spectrum of a liquid crystal | Spence *et al.*, *J. Chem. Phys.* **21**: 380 |
| 1954 | Carr–Purcell spin echoes | Carr and Purcell, *Phys. Rev.* **94**: 630 |
| 1955 | Solomon equations for NMR relaxation | Solomon, *Phys. Rev.* **99**: 559 |
| 1955 | Relaxation in the rotating frame | Redfield, *Phys. Rev.* **98**: 1787 |
| 1957 | Redfield theory of relaxation | Redfield, *IBM J. Res. Dev.* **1**: 19 |
| 1958 | Magic angle spinning for high-resolution studies of solids | Andrew *et al.*, *Nature* **182**: 1659; Lowe, *Phys. Rev. Lett.* **2**: 285 |
| 1963 | Liquid crystal solvents used in NMR | Saupe and Engelert, *Phys. Rev. Lett.* **11**: 462 |
| 1964 | Deuterium spectrum of a liquid crystal | Rowell *et al.*, *J. Chem. Phys.* **43**: 3442 |
| 1966 | NMR spectrum shown to be Fourier transform (FT) of free induction decay (FID) | Ernst and Anderson, *Rev. Sci. Instrum.* **37**: 93 |
| 1971 | Deuterium spectrum of a membrane | Oldfield *et al.*, *FEBS Lett.* **16**: 102 |
| 1976 | Cross-polarization magic angle spinning for solids | Schaeffer and Stesjkal *J. Am. Chem. Soc.* **98:** 1030 |

scientific community. The first commercial spectrometer (30 MHz for $^1H$) was marketed by Varian Associates in 1952, and many of the new early developments stemmed from Varian's research and development department. Sample spinning and field-frequency locking are just two examples that led to dramatic improvements in the quality of high-resolution spectra of liquids. However, the most significant development was the pulse Fourier transform (FT) method of acquiring spectra, which Anderson and Ernst realized at Varian, the first account of which appeared in 1966. At that time their spectrometer did not have an on-line, or even a close at hand computer on which to do the transform, and the exploitation of the method in a commercial spectrometer had to wait for the development of the on-line computer. In fact, the first commercial pulse FT spectrometer was marketed by Bruker in 1969.

Varian introduced superconducting magnets into NMR with a 200 MHz proton spectrometer, first produced in 1962, and whose field strength was soon increased so as to give proton resonance at 220 MHz.

By 1971 NMR was beginning to look like a mature spectroscopy with all the major developments in place. However, in that year Jeener suggested the idea of multidimensional spectroscopy, and in 1973 Lauterbur published his method for imaging of objects by applying magnetic field gradients. These two events stimulated Ernst and his collaborators to develop the first two-dimensional experiments, and a new age of rapid development in NMR began,

leading to the marvellous portfolio of experiments in NMR spectroscopy and imaging that are available today.

## Solid-state and liquid crystalline samples

The rapid and isotropic motion in normal liquids averages the anisotropic interactions to zero. The rapid motion also produces a long spin–spin relaxation time ($T_2$), and hence a very narrow NMR line. In most solids there is little or no motion and the NMR, lines may be split by very large anisotropic interactions, and will usually have a very short $T_2$, and hence broad lines. In fact, in the early days of NMR, studies of solids and liquids were seen to be quite different activities. Commercial spectrometers were produced mainly for liquid-state studies, since it was appreciated that the applications of NMR for mixture analysis and structure determination by chemists would be the major market. Spectrometers were usually designed either to obtain high-resolution spectra – that is, to resolve the small chemical shifts and spin–spin couplings exhibited by liquids – or for solid samples, where magnet homogeneity was not so important but special techniques were necessary in order to record the very broad line spectra. The NMR community was divided mainly into those working with liquids and those looking at solids.

We will restrict our description of the historical development of the NMR of solids and liquid crystals to showing how the gap between these two communities has narrowed, and indeed now overlaps. The first steps along this path were taken by Andrew, Bradbury and Eades in 1958, and by Lowe in 1959, who showed that rotation of a solid sample about an angle of 54.7° to the magnetic field can remove the second-rank, anisotropic contributions to NMR interactions for spin-$\frac{1}{2}$ nuclei. This means that, in principle, the dipolar interaction, which is entirely a second-rank, anisotropic interaction, and the anisotropic contribution to the chemical shift can be removed by using this ‘magic angle’ spinning (MAS) technique. The spectra obtained show spinning sidebands at the frequency of the rotation speed, and have intensities that depend on the relative magnitudes of the rotation speed and the magnitude of the interaction being averaged. The early experiments demonstrated that the spectral lines can be narrowed to reveal chemical shift differences, and even in some cases spin–spin couplings, but the samples that could be studied in this way were limited, and the method did not find wide application. The MAS experiment had to wait until 1976 before it was used to provide high-quality, high-resolution $^{13}C$ spectra from solid samples. Carbon-13 is a special case in being isotopically dilute at natural abundance and so the spectra can easily be simplified by proton decoupling. This produces spectra from a liquid sample that have a single line for each chemically equivalent group of carbons. The low isotopic abundance, however, also leads to a low signal-to-noise ratio, and in liquids it was not until the advent of the pulse FT method that $^{13}C$ spectra with a good signal-to-noise ratios could be obtained by time averaging. For a solid sample the time averaging is often inefficient because the ratio of the relaxation times $T_1/T_2$ is high. To overcome this problem, Schaefer and Stejskal used an idea proposed and demonstrated by Hahn and Hartmann in 1962 in which the $^{13}C$ and $^{1}H$ nuclei can be made to transfer polarization by subjecting each of them simultaneously to spin-locking radiofrequency fields.

In liquid crystalline samples the molecules move rapidly, so that the NMR interactions are averaged, but they do not move randomly, and this results in nonzero averaging of the dipolar couplings, the chemical shift anisotropies and the quadrupolar interactions. The first reported observation of a spectrum from a liquid crystalline sample was by Spence, Moses and Jain in 1953. It was the proton spectrum of a nematic sample, and consisted of a very broad triplet structure and had a low information content. Ten years later, Englert and Saupe recorded the $^{1}H$ spectrum of benzene dissolved in a nematic solvent and this consisted of a large number of sharp lines; its analysis gave three, partially averaged dipolar couplings whose values could be related to the relative positions of the protons and the orientational order of the sixfold symmetry axis of the benzene molecule. The study of liquid crystals themselves received a boost with the publication in 1965 by Rowell, Melby, Panar and Phillips of the spectrum given by the deuterons in a specifically deuterated nematogen. They obtained partially averaged quadrupolar splittings, which can be used to characterize the orientational order of the deuterated molecular fragments. The realization that NMR could give useful information about membranes and model membranes dates from the late 1960s and early 1970s, and the particularly valuable role of deuterium NMR in membrane studies stems from the publication in 1971 of a study by Oldfield, Chapman and Derbyshire.

## Liquid-state NMR

Following the detection of the NMR phenomenon and the subsequent discovery of the chemical shift

and spin–spin coupling, NMR emerged as one of the most powerful physical techniques for determining molecular structures in solution and for analysing complex mixtures of molecules. The potential of the method as a structural tool was almost immediately recognized. High-resolution NMR spectrometers were constructed in several laboratories (such as those of HS Gutowsky, RE Richards and JD Roberts, and JN Shoolery at Varian Associates) and the pioneering efforts of these scientists and others began to demonstrate the scope of applications of the technique in chemistry. The success of the method for chemists derives from the well-defined correlations between molecular structure and the measured chemical shifts and spin coupling constants. In retrospect, the achievements of the early workers were truly remarkable considering that they were working at such low magnetic fields (30/40 MHz for $^1$H) so that spectral dispersion was poor and the sensitivity was three orders of magnitude less than in present-day instruments. The ingenious adaptations of their instruments to increase the stability and resolution (for example, field-frequency locking, homogeneity shim coils and sample spinning) were absolutely essential to allow them to make progress in their structural determinations. As time progressed, the sensitivity was boosted initially by increasing the field strengths and improving the radiofrequency (RF) circuitry and probe designs, and subsequently by using spectral accumulation and Fourier transform methods.

Most of the important milestones in the development of the NMR technique for studies of solution state NMR are given in **Table 2**. By 1957 NMR was emerging as a powerful nondestructive analytical technique capable of providing structural information about the environment of more than 100 known nuclear isotopes. Initially the technique was held back by its relatively low sensitivity and the complexity of the $^1$H spectra of larger molecules. In the late 1950s, although many problems were identified for NMR study in areas such as polymer chemistry, organometallic chemistry and even biochemistry, the method was proving to be grossly inadequate for tackling them. For example, polymer scientists, acquired some of the early instruments hoping to determine stereotacticities and cross-linking in synthetic polymers; in fact it was not until several years later that improved instrumentation allowed such problems to be tackled successfully. Meanwhile, the method was enjoying considerable success in helping to solve molecular structures of moderately sized molecules ($M_r < 400$): it was particularly useful in natural product chemistry where it became possible to differentiate between several structures that satisfied the compositional data. It was also proving to be a very powerful method for defining stereochemical details of various structures, for example alkaloids and steroids. Not surprisingly, organic chemists were immediately attracted to this technique, which could reveal unresolved structural details about some of the molecules they had been studying for decades.

More challenging applications to larger molecules became possible only with the eventual improvements in sensitivity and spectral simplification. Although the manufacturers made steady progress in providing higher and higher field strengths, it was not until 1966 that a significant impact was made on the sensitivity problem with the arrival of Fourier transform methods and the use of dedicated computers for data acquisition. These methods also facilitated studies of less-sensitive nuclei and from 1966 to 1975 $^{13}$C studies at natural abundance became routine not only for structural studies but also for investigating rapid molecular motions (obtaining correlation times from $^{13}$C relaxation studies). During this period, structural determinations of fairly large molecules ($M_r$ ~3500) became commonplace and measurements of nuclear Overhauser effects were frequently used to identify protons that were near to each other. Fortunately, while this rapid expansion in applications work was underway, a few research groups continued to concentrate on understanding the basic spin physics. Some of the novel multipulse techniques developed at this time (such as INEPT and HMQC for indirect detection of insensitive heteronuclei via proton signals) were to prove of far-reaching value in eventually simplifying complex NMR spectra from large macromolecules.

## Biological applications of NMR

Biochemists became interested in the NMR technique long before it could provide them with the detailed information they were seeking. For example, the first $^1$H spectrum of a protein was recorded in 1957 and proved to be almost featureless. From these unpromising beginnings, who would have predicted that 40 years later the technique would be used to fully assign the resonances of proteins as large as 30 kda and to determine their three-dimensional structures? Early workers such as M Cohn, O Jardetzky and RG Shulman had sufficient vision to recognize the eventual potential of the method when they began their pioneering studies on nucleotides, amino acids, peptides, proteins, paramagnetic ion effects and metabolic applications. In the early days, brave attempts were made to solve the problem of signal overlap by studying partially deuterated,

**Table 2** Milestones in the development of solution-state NMR

| *Date* | *Milestone* | *Literature citation* |
|---|---|---|
| 1949–1950 | Discovery of the chemical shift | Knight, *Phys. Rev.* **76**: 1259; Proctor and Yu, *Phys. Rev.* **77**: 777; Dickinson *Phys. Rev.* **77**: 736 |
| | Discovery of spin–spin coupling | Proctor and Yu, *Phys. Rev.* **81**: 20; Gutowsky and McCall, *Phys. Rev.* **82**: 748; Ramsey and Purcell, *Phys. Rev.* **85**: 143, Hahn and Maxwell, *Phys. Rev.* **84**: 1246 |
| 1951 | Discovery of $^{1}$H chemical shifts | Arnold *et al.*, *J. Chem. Phys.* **19**: 507 |
| 1952 | First commercial NMR spectrometer (30 MHz) | Varian |
| 1953 | Overhauser effect | Overhauser, *Phys. Rev.* **91**: 476 |
| 1953 | Theory for exchange effects in NMR spectra | Gutowsky *et al.*, *J. Chem. Phys.* **21**: 279 |
| 1953–58 | Sample spinning used for resolution improvement | Bloch, *Phys. Rev.* **94**: 496 |
| | Field gradient shimming with electric currents | Golay, *Rev. Sci. Instrum.* **29**: 313 |
| | Magnetic flux stabilization (Varian) | Shoolery, *Prog. NMR Spectrosc.* **28**: 37 |
| | Variable temperature operation | Shoolery, *Prog. NMR Spectrosc.* **28**: 37 |
| | Spin decoupling | Bloom and Shoolery, *Phys. Rev.* **97**: 1261 |
| 1957 | Analysis of second-order spectra | Gutowsky *et al.*, *J. Am. Chem. Soc.* **79**: 4596; Bernstein *et al.*, *Can. J. Chem.* **35**: 65; Arnold, *Phys. Rev.* **102**: 136; Anderson, *Phys. Rev.* **102**: 151 |
| 1959 | Blood flow measurements *in vivo* | Singer, *Science* **130**: 1652 |
| 1959 | Vicinal coupling constant dependence on dihedral angle | Karplus, *J. Chem. Phys.* **30**: 11; **64**: 1793 |
| 1961 | First commercial 60 MHz field/frequency locked spectrometer (Varian A 60) | Varian |
| 1962 | First superconducting magnet NMR spectrometer (Varian 220 MHz) | Varian |
| 1962 | Indirect detection of nuclei by heteronuclear double resonance (INDOR) | Baker, *J. Chem. Phys.* **37**: 911 |
| 1964 | Spectrum accumulation for signal averaging | Ernst, *Rev. Sci. Instrum.* **36**: 1689 |
| 1965 | Nuclear Overhauser enhancements (NOEs) used in conformational studies | Anet and Bourn, *J. Am. Chem. Soc.* **87**: 5250 |
| 1966 | Fourier Transform (FT) techniques introduced | Ernst, *Rev. Sci. Instrum.* **36**: 1689; Ernst and Anderson, *Rev. Sci. Instrum.* **37**: 93 |
| 1969 | First commercial FT NMR spectrometer (90 MHz) | Bruker |
| 1969 | Lanthanide shift reagents used in NMR | Sievers, *NMR Shift Reagents,* Academic Press |
| 1970–75 | $^{13}$C studies at natural abundance become routine | |
| 1970 | First commercial FT spectrometer with superconducting magnet (270 MHz) | Bruker |
| 1971 | Pulse sequences for solvent signal suppression | Platt and Sykes, *J. Chem. Phys.* **54**: 1148 |
| 1971 | Two-dimensional (2D) NMR concept suggested | Jeener |
| 1972 | $^{13}$C studies of cellular metabolism | Matwiyoff and Needham, *Biochem. Biophys. Res. Commun.* **49**: 1158 |
| 1973 | $^{31}$P detection of intracellular phosphates | Moon and Richards, *J. Biol. Chem.* **248**: 7276 |
| 1973 | NMR analysis of body fluids and tissues | Moon and Richards, *J. Biol. Chem.* **248**: 7276; Hoult *et al.* *Nature* **252**: 285 |
| 1973 | 360 MHz superconducting NMR spectrometer | Bruker |
| 1974 | 2D NMR techniques developed | Aue *et al.*, *J. Chem. Phys.* **64**: 229 |
| 1976 | Early NMR studies on body fluids and tissues | Moon and Richards, *J. Biol. Chem.* **248**: 7276; Hoult *et al.*, *Nature* **252**: 285 |
| 1976–79 | $^{31}$P studies of muscle metabolism | Burt *et al.*, *J. Biol. Chem.* **251**: 2584; Burt *et al.*, *Science* **195**: 145; Garlick *et al. Biochem. Biophys. Res. Commun.* **74**: 1256; Jacobus *et al.*, *Nature* **265**: 756; Hollis and Nunnally, *Biochem. Biophys. Res. Commun.* **75**: 1086; Yoshizaki, *J. Biochem.* **84**: 11; Cohen and Burt, *Proc. Natl. Acad. Sci.* **74**: 4271; Sehr and Radda, *Biochem. Biophys. Res. Commun.* **77**: 195; Burt *et al.*, *Annu. Rev. Biophys. Bioeng.* **8**: 1 |
| 1977 | First 600 MHz spectrometer | Carnegie Mellon University |
| 1979 | Detection of insensitive nuclei enhanced by polarization transfer (INEPT) | Morris and Freeman, *J. Am. Chem. Soc.* **101**: 760 |

**Table 2** *Continued*

| *Date* | *Milestone* | *Literature citation* |
|---|---|---|
| 1979 | Detection of heteronuclear multiple quantum coherence (HMQC) | Mueller, *J. Am. Chem. Soc.* **101**, 4481; Burum and Ernst, *J. Magn. Reson.* **39**: 163 |
| 1979 | 500 MHz superconducting spectrometer | Bruker |
| 1980 | Surface coils used for *in vivo* NMR studies | Ackerman *et al.*, *Nature* **283**: 167 |
| 1980 | Pulsed-field gradients used for coherence selection | Bax *et al.*, *Chem. Phys. Lett.* **69**: 567 |
| 1981 | NMR used to diagnose a medical condition | Ross *et al.*, *N. Engl. J. Med.* **304**: 1338 |
| 1981–83 | Perfusion methods used for NMR studies of cell metabolism | Ugurbil *et al.*, *Proc. Natl. Acad. Sci*, **78**: 4843; Foxall and Cohen, *J. Magn. Reson.* **52**: 346 |
| 1982 | Full assignments obtained for small protein | Wagner and Wüthrich, *J. Mol. Biol.* **155**: 347 |
| 1983 | First 3D-structures of proteins from NMR data | Williamson *et al.*, *J. Mol. Biol.* **182**: 195; Braun *et al., J. Mol. Biol.* **169**: 921 |
| 1987 | 600 MHz superconducting spectrometer | Bruker; Varian; Oxford Instruments |
| 1988 | 2D-NMR combined with isotopically labelling for full assignments of proteins | Torchia *et al.*, *Biochemistry* **27**: 5135 |
| 1988 | Whole-body imaging and spectroscopy at 4.0 T | Barfuss *et al.*, *Radiology* **169**: 811 |
| 1989 | 3D-NMR on isotopically labelled proteins | Marion *et al.*, *Biochemistry* **28**: 6150 |
| 1990 | 4D-NMR on isotopically labelled proteins | Kay *et al.*, *Science* **249**: 411 |
| 1990 | Pulsed-field gradients routinely incorporated into pulse sequences | Bax *et al.*, *Chem. Phys. Lett.* **69**: 567; Hurd, *J. Magn. Reson.* **87**: 422 |
| 1992 | 750 MHz spectrometers | Bruker; Varian; Oxford Instruments |
| 1995 | 800 MHz spectrometer | Bruker |

large biological macromolecules at ever-increasing field strengths. However, a general solution to the signal overlap problem became available only with the arrival of multidimensional NMR methods.

The most important breakthrough came in 1975 with the development of the first two-dimensional (2D) NMR experiments, which had the capability of both simplifying complex spectra and also establishing correlations between nuclei connected either by scalar spin coupling through covalent bonds (COSY spectra) or by dipole–dipole relaxation pathways through space (NOESY spectra). These 2D experiments allowed the assignment of complex NMR spectra and provided distance information for use in structural calculations. The eventual demonstration of the full potential of these methods was made by Wüthrich and co-workers, which eventually led to the first determination of a complete structure for a globular protein in solution.

The extension of the multidimensional NMR approach to larger proteins was subsequently made possible by the development of 3D- and 4D-NMR techniques incorporating INEPT and HMQC pulse sequences that were applied to $^{13}$C- and $^{15}$N-labelled proteins. These latter developments were made at NIH by Bax and Clore and their co-workers. These multidimensional NMR methods provide the spectral simplification required to completely assign the spectra of proteins of up to 30 kDa and to determine their structures to a resolution similar to the 0.20 nm resolution X-ray structure (see the relevant milestone experiments in **Table 2**). Using the modern techniques, detailed structural and dynamic information can now be routinely obtained for complexes of proteins formed with nucleic acids and other ligands with overall molecular masses of ~30 kDa.

In the early 1970s a completely new area of NMR was opened by reports (by Moon and Richard and by Hoult and co-workers) showing that it was possible to record high-resolution $^{31}$P NMR spectra on cells and intact organs. This led to an exciting area of research into metabolic processes that allows the chemistry within living cells to be monitored directly. These methods have reached the stage where they can be used to diagnose disease, to monitor biochemical responses to exercise and stress, and even to follow the effects of drug therapy by using repeated non-invasive examinations. The possibility of combining this approach with spatial localization techniques in whole-body magnetic resonance imaging (MRI) presents enormous opportunities for future work.

## Magnetic resonance imaging (MRI)

Many of us can recall the great intellectual excitement that accompanied the publication of the early NMR experiments in 1973 showing how spatial information can be encoded into NMR signals. In particular, the simple approach adopted by Paul Lauterbur of using field gradients to produce the

spatial resolution required to give a two-dimensional image of water in glass tubes was a brilliant example of lateral thinking that provided a completely new way of viewing the NMR experiment. Even in the very early days, the pioneering workers in MRI (the word ‘nuclear’ having been dropped because it was thought that it would suggest to the patients that radioactivity was involved) realized that the technique would make its largest contribution in the area of noninvasive clinical imaging. By 1977 the first images of the human body were being reported, one of the earliest being that of a wrist showing features as small as 0.5 cm. At first the method was greeted with much scepticism because its sensitivity performance compared unfavourably with the well-established X-ray CT scanning methods: however, rapid instrumental advances soon allowed the MRI technique to show its full potential, particularly in the ability to provide high-contrast images for soft tissues and tissues in areas surrounded by dense bone structures. The development of the echo planar imaging (EPI) method by Mansfield and his co-workers allowed well-resolved images to be obtained from a single pulse and this opened up many new applications requiring short examination times, such as in heart, abdomen and chest imaging. Other important milestones in the development of the MRI technique are summarized in **Table 3**. There are now many applications where MRI is the favoured imaging method (such as brain scanning for detecting encephalitis or multiple sclerosis (MS) and for monitoring therapy treatment of MS). Most of the images examined are based on detecting $^{1}H$ nuclei. However, recent high-quality images of the airways in human lungs have been provided by helium or xenon images obtained after inhalation of the polarized inert gases by the patient. Another recent and exciting application, called functional MRI, attempts to study the working of the human brain; by stimulating the brain either through the

**Table 3** Milestones in the development of magnetic resonance imaging (MRI)

| *Date* | *Milestone* | *Literature citation* |
|---|---|---|
| 1973 | Spin-imaging methods proposed | Lauterbur, *Nature* **242**: 190; Mansfield and Grannell, *J. Phys. C* **6**: L422; Damadian, *NMR in Medicine*, Springer-Verlag |
| 1973 | NMR diffraction used for NMR imaging | Mansfield and Grannell, *J. Phys. C* **6**: L422 |
| 1973 | Zeugmatography; first two-dimensional NMR image | Lauterbur, *Nature* **242**: 195 |
| 1974 | Sensitive point imaging method | Hinshaw, *Phys. Lett.* **48**: 87 |
| 1974 | 2D NMR techniques developed | Aue *et al.*, *J. Chem. Phys.* **64**: 229 |
| 1975 | Slice selection in imaging by selective excitation | Garroway *et al.*, *J. Phys.* C **7**: L457; Sutherland and Hutchinson, *J. Phys. E* **11**: 79; Hoult, *J. Magn. Reson.* **35**: 69 |
| 1975 | Fourier zeugmatography | Kumar *et al.*, *J. Magn. Reson.* **18**: 69 |
| 1977–80 | Spin-imaging of human limbs and organs | Wehrli, *Prog. NMR Spectrosc.* **28**: 87 |
| 1977 | Echo-planar imaging | Mansfield and Pykett, *J. Magn. Reson.* **29**: 355 |
| 1977–78 | Whole-body scanning | |
| 1979 | Chemical shift imaging | Cox and Styles, *J. Magn. Reson.* **40**: 209 Brown *et al.*, *Proc. Natl. Acad. Sci.* **79**: 3523 Maudsley *et al.*, *J. Magn. Reson.* **51**: 147 Mauldsley *et al.*, *Siemens Forsch. Entwickl-Ber.* **8**: 326 |
| 1980 | Spin-warp imaging | Edelstein *et al.*, *Phys. Med. Biol.* **25**: 751 |
| 1980 | 3D-projection reconstruction | Lai and Lauterbur, *J. Phys. E* **13**: 747 |
| 1983 | Whole-body imaging at 1.5 T | Hart *et al.*, *Am. J. Roentgenol.* **141**: 1195 |
| 1984–87 | Gradient methods used for spatial localization | Bottomley, *US Patent* 480/228; Ordidge *et al.*, *J. Magn. Reson.* **60**: 283; Frahm *et al.*, *J. Magn. Reson.* **72**: 502 |
| 1984 | Combined imaging and spectroscopy on human brain | Bottomley *et al.*, *Radiology* **150**: 441 |
| 1985 | FLASH imaging | Haase *et al.*, *J. Magn. Reson.* **67**: 258 |
| 1985 | Magnetic resonance (MR) angiographic images | Wedeen *et al.*, *Science* **230**: 946 |
| 1986 | NMR microscopy imaging on live cell | Aguayo *et al.*, *Nature* **322**: 190 |
| 1987 | Echo-planar imaging at 2.0 T | Pykett and Rzedian, *Magn. Res. Med.* **5**: 563 |
| 1988 | Whole-body imaging and spectroscopy at 4.0 T | Barfuss *et al.*, *Radiology* **169**: 811 |
| 1991 | Functional MR-detection of cognitive responses | Belliveau *et al.*, *Science* **254**: 716; Prichard *et al.*, *Proc. Natl. Acad. Sci.* **88**: 5829 |
| 1993 | NMR microscopy using superconducting receiver coil | Black *et al.*, *Science* **259**: 793 |
| 1994 | Use of polarized rare gases in spin-imaging | Albert *et al.*, *Nature* **370**: 199 |

senses or by thought processes, it is possible to detect changes in MRI images of the brain. These are related to changes in oxygen levels in the blood induced in specific locations of the brain. This type of experiment opens up exciting possibilities for studying the human brain in action.

MRI scanners are now increasingly being used not only in research hospitals but also in the general hospital environment. The high-profile use of MRI as a major health-care tool has certainly increased the public awareness of NMR and drawn proper attention to the versatility of this exceptional phenomenon.

## List of symbols

$m$ = magnetic quantum number; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time.

*See also:* **Cells Studied By NMR; *In Vivo* NMR, Applications, Other Nuclei; *In Vivo* NMR, Applications, $^{31}$P; *In Vivo* NMR, Methods; Labelling Studies in Biochemistry Using NMR; Liquid Crystals and Liquid Crystal Solutions Studied By NMR; Macromolecule–Ligand Interactions Studied By NMR; Membranes Studied By NMR Spectroscopy; MRI Applications, Biological; MRI Applications, Clinical; MRI Instrumentation; MRI Theory; NMR in Anisotropic Systems, Theory; NMR of Solids; NMR Spectrometers; NMR Pulse Sequences; Nuclear Overhauser Effect; Nucleic Acids Studied Using NMR; Perfused Organs Studied Using NMR Spectroscopy; Proteins Studied using NMR Spectroscopy; Solid State NMR, Methods; Two-Dimensional NMR Methods.**

## Further reading

Grant DM and Harris RK (eds) (1996) *Encyclopedia of NMR*. Chichester: Wiley.

Emsley JW and Feeney J (1995) *Progress in Nuclear Magnetic Resonance Spectroscopy* **28:** 1.

# Manganese NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# Mass Spectrometry in Food Science

*See* **Food Science, Applications of Mass Spectrometry.**

# Mass Spectrometry, Historical Perspective

**Allan Maccoll**†, Claygate, Surrey, UK

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| **Historical Overview** |

## Introduction

Mass spectrometry has made many notable contributions to chemistry from the chemical physics of small molecules to the structures of large biomolecules. It is an instrument in which ions in a beam are separated according to their mass/charge ratio ($m/z$). Its humble beginnings lay in the works of physicists at the turn of the century. Up to the Second World War mass spectrometry was the province of the physicists along with a small band of physical chemists. However, the demands for accurate evaluation of the composition of aircraft fuel during the Second World War led to its extensive application to hydrocarbon analysis. Heartened by the success in this area, operators were encouraged to put 'dirty' organic chemicals in their instruments and so organic mass spectrometry was born. These developments were largely owing to the manufacturers responding to the demands for instruments to meet the needs of the chemists. The introduction of high-resolution instruments led to the developments of ion chemistry. This took place in the decades 1950–1980. By this time the mass spectrometric study of large organic molecules had been achieved and the prevailing interest switched to biomolecules – a good source of financial support in view of their medical relevance.

One of the important aspects of the development of mass spectrometry was the camaraderie (occassionally blighted by periods of frustration) that existed between the users and the manufacturers. This was nurtured by the introduction of user's meetings by Associated Electrical Industries (an offshoot of Metrovick). The users would foregather with the engineers responsible for instrumental development to explain *their* problems and requirements for instrument development. The author remembers well the confidence he gained from learning that his problems were not unique – other users had them too! Instrumental development was stimulated by the demands of the users and if the suggested instrumentation could be satisfactorily produced it soon became available. This made it an exciting period to live through.

## The beginnings

### Thomson

The origins of mass spectrometry lie in the work done in the Cavendish Laboratory in Cambridge by JJ Thomson and his colleagues at the start of the twentieth century on electrical discharges in gases. The first relevant work was the discovery of the electron, using a cathode ray tube. The rays from the cathode pass through a slit in the anode (**Figure 1**) and after passing through another slit pass between two metal plates and on to the wall of the tube. This wall had been treated with a phosphorescent material which glows where the beam strikes it. The beam can be diverted by applying a potential difference between the plates and also by superimposing a magnetic field. By adjusting the two fields so that there is no displacement of the beam Thomson was able to show that the particles carried a negative charge of around $10^{11}$ C kg$^{-1}$. Goldstein in 1886, using a perforated cathode, was able to show that there was always a beam travelling in the opposite direction to that of the electrons – the so called kanalstrahlen. Later, Wien showed that these were positively charged particles and concluded that they were positive ions. Thomson decided to investigate these particles. His positive ray apparatus (1912) is shown in **Figure 1**. A is a discharge tube producing positive ions which then pass through the cathode B and after collimation in the narrow tube BN are subjected to superimposed electric and magnetic fields (M,M′ P,P′). The displaced beams then travel to the fluorescent screen G where their effect is observed.

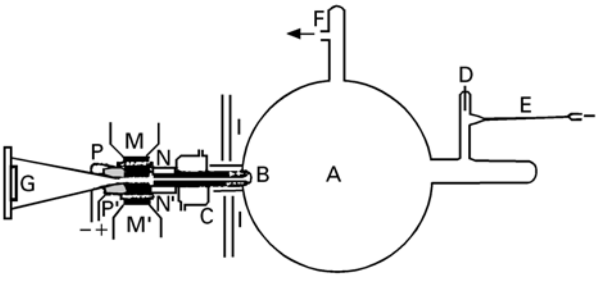



**Figure 1** Thomson's cathode ray tube.

**Figure 2** The parabolae.

In **Figure 2** the parabola formed by the top and bottom branches on the left-hand side are due to neon. Under better resolution they show the presence of isotopes at masses 20 and 22. Isotopes had previously been observed in studies of radioactivity. Thomson encouraged a research student in the Cavendish Laboratory, FW Aston, to build a mass spectrograph for further studies of stable isotopes. The research was interrupted by the war of 1914–1918 and so the work was not published until 1923.

## Aston

His spectrograph is shown diagrammatically in **Figure 3**. A beam of ions passes through the collimating slits $S_1$, $S_2$ into an electric field $P_1$, $P_2$. It then enters a magnetic field centred upon M and the divergent beam is brought to focus on a photographic plate P. The geometry ensures that, irrespective of the velocity of the ions, they are brought to a sharp focus on the photographic plate. This is known as velocity focusing. By 1923 Aston had realized that deviations from integral values of the relative molecular masses (Prout's Rule) were of considerable importance for the study of nuclear structure. However, as will be seen later, they were of inestimable importance in the development of organic mass spectrometry. Current values for some atoms are shown in **Table 1** ($^{12}C$ = 12.000 000).

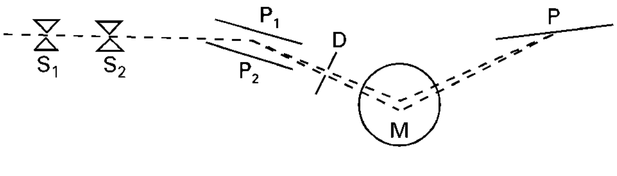



**Figure 3** Aston's mass spectrograph.

**Table 1** Accurate masses of some common atoms

| *Atom* | *Relative atomic mass* |
|---|---|
| Hydrogen | 1.007 825 |
| Carbon | 12.000 000 |
| Nitrogen | 14.003 074 |
| Oxygen | 15.994 915 |
| Fluorine | 18.888 405 |
| Sulfur | 31.972 074 |

## Dempster

In 1918, a Canadian working in the University of Chicago (AJ Dempster) developed a different type of apparatus for investigating positive rays (**Figure 4**). It involved a 180° magnetic field. Ions produced by the filament in G are accelerated into the magnetic field through $S_1$ and pass through $S_2$ and hence to the collector E. Such a geometry gives rise to direction focusing – ions will arrive at the collector irrespective of the direction they enter the magnetic field. The experimental arrangement is described by the fundamental equation of sector mass spectrometry (Eqn [1]), namely

$$m/z = 2V/B^2R^2 \qquad [1]$$

where $m$ is the mass of the ion, $z$ its charge, $B$ the magnetic field strength and $R$ the radius of the magnetic field. A fundamental difference between

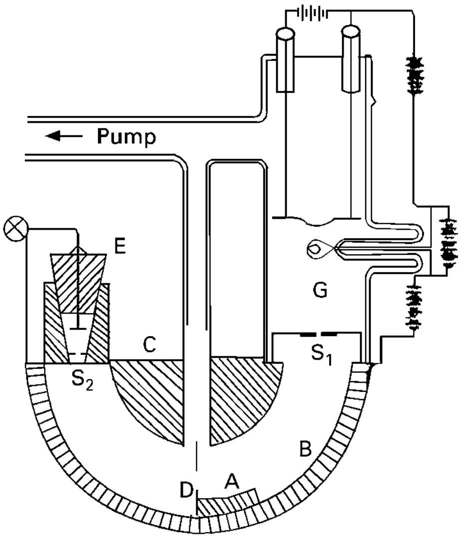



**Figure 4** Dempster's mass spectrometer.

Aston's instrument and that of Dempster is that Aston's spectra are obtained instantaneously whereas Dempster's have to be scanned. This can be done simply in two ways, either by scanning the electric field at constant magnetic field or by scanning the magnetic field at constant electric field (more sophisticated methods of scanning have been developed, leading to a better understanding of mass spectrometric processes). Most sector mass spectrometers use the latter method.

## Instrumental development

### The basic mass spectrometer

In the mass spectrometer shown in **Figure 5** the sample is held in the reservoir and led into the ionization chamber via a leak. On ionization the ions are accelerated into the magnetic sector and eventually arrive at the collector. The current is amplified and recorded.

JJ Thomson was very percipient in predicting organic mass spectrometry in 1913. He wrote in his book *Rays of Positive Electricity and their Application to Chemical Analysis*. "I have described at some lengths the applications of positive rays to chemical analysis: one of the main reasons for writing this book was the hope that it might induce others and especially chemists, to try this method of analysis. I feel sure that there are many problems in chemistry which could be solved with much greater ease by this than by any other method. This method is surprisingly sensitive – more so even than that of spectrum analysis, requires an infinitesimal amount of material and does not require this to be especially purified; the technique is not difficult if appliances for producing high vacua are available . . .". It is a reflection upon the chemists of the period that it took thirty years for Thomson's predications to be verified.

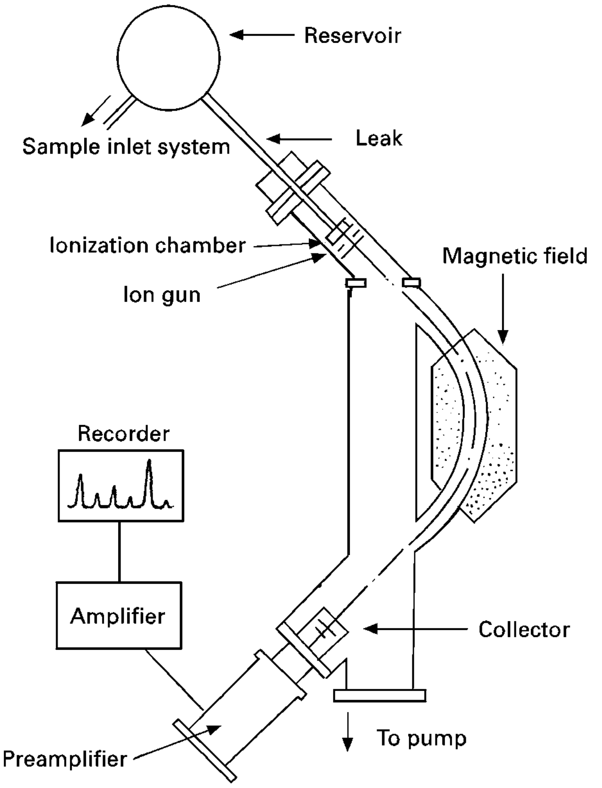



**Figure 5** A single focusing mass spectrometer (MS2).

One difficulty was that the apparatus, simple to a physicist, appeared very complex to a chemist. The application of mass spectrometry of chemistry had to await the commercial production of instruments. The impetus came in the 1940s when the war effort demanded rapid and accurate hydrocarbon analysis in connection with aviation fuels. The next big step came in the 1950s when it was realized that in addition to quantitative analysis the technique could be used for the qualitative (structural) analysis of organic compounds. A certain resistance had to be overcome to induce mass spectrometrists to put 'dirty' compounds into their instruments rather than 'clean' hydrocarbons. This gave mass spectrometer manufacturers a further impetus to develop more and more advanced instruments and led to a new discipline – organic mass spectrometry.

Before the advent of the mass spectrometer the determination of the relative molecular mass ($M_r$) of an organic compound was performed by quantitative analysis (empirical formula) and a rough $M_r$ was used to decide the number of empirical formula to make up the molecular formula. With the mass spectrometer the $M_r$ could be determined directly: however, there was more to come. It was noted earlier that the relative atomic masses of atoms were slightly different from integer values. If the $M_r$ of a compound could be accurately determined then there would be only one formula that would be consistent with it. So it was up to the manufacturers to produce instruments with sufficiently high resolution to be able to separate these values. A word about resolution or resolving power is appropriate here. Although there is no generally accepted definition, one that is widely used is the 10% valley definition. If two peaks of equal height are separated by $\Delta m$ and the valley between them is 10% of the peak height then the resolving power is said to be $m/\Delta m$. If one considers the doublet at $m/z$ 28 corresponding to $C_2H_4$ and $N_2$, $\Delta m$ is (28.031 299 –28.006 158) = 0.025 141. Thus $m/\Delta m = 1114$ and a resolving power of about 1000 would be required to separate

the two peaks. The search for higher and higher resolution led to the introduction of a double focusing mass spectrometer. It has been seen that while Aston's mass spectrograph gives velocity focusing, Dempster's mass spectrometer gave direction focusing. Nier and Roberts developed a geometry which ensured both velocity and direction focusing. This geometry formed the basis of the MS9 (Associated Electrical Instruments, AEI) (**Figure 6**) which for many years was the workhorse of the organic mass spectrometrists. Initially it had resolving power of 10000 but with modifications this value was raised tenfold. It became apparent that it would be advantageous if the ion beam could be selected before its subsequent analysis. This gave rise to the ZAB series of mass spectrometers (Vacuum Generators). These instruments also had the advantage of the ion beam being in the horizontal plane (the AEI instruments had the ion beam in the vertical plane) which made it much easier to add additional sectors when required.

## Representation of mass spectra

### The bar diagram

Mass spectra are usually represented by bar diagrams on which the relative intensity of peak or the relative abundance of an ion is plotted against the *m*/*z* value (**Figure 7**). The molecular peak $[M]^{\bullet +}$ is the one corresponding to the $M_r$ of the compound and the base peak is the most intense one in the spectrum. A further alternative is the use of the fraction of the total ion current carried by the ion in question. In the early days of mass spectrometry the operator had to laboriously develop the trace recorded on photographic paper or equally laboriously plot the ion current against the *m*/*z* ratio. More recently the spectra are electronically recorded and can be plotted out according to the whim of the operator.

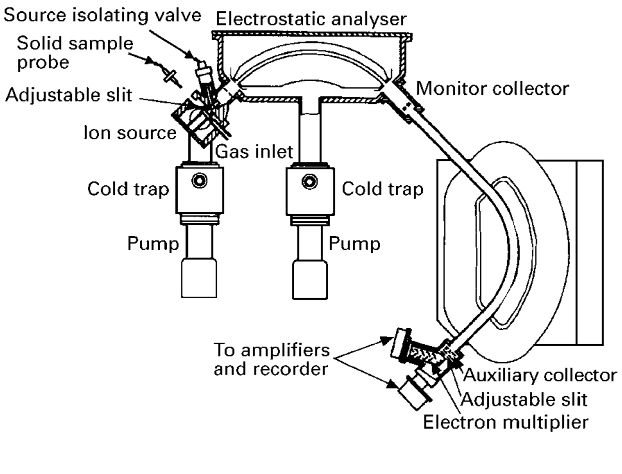



**Figure 6** A double focusing mass spectrometer (MS9).

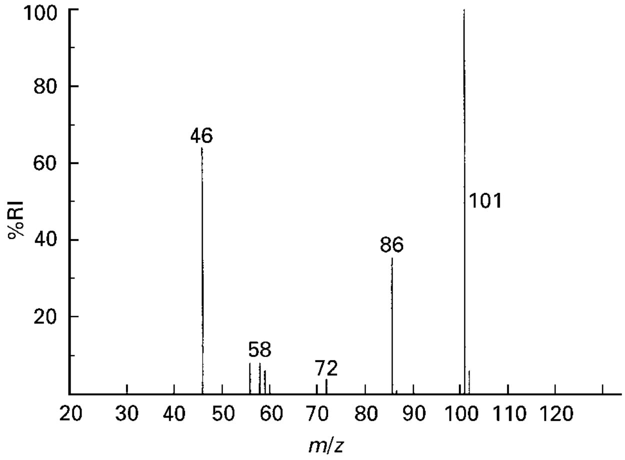



**Figure 7** Mass spectrum of $[HCONHC(CH_3)_3]$.

## The anatomy of a mass spectrometer

### The components of a mass spectrometer

The mass spectrometer consists essentially of a source, which produces a beam of ions, an analyser which separates the beam according to the *m*/*z* ratio and a collector which determines the fraction of the total ion current carried by each of the ions.

### Sector instruments

**The source** Probably the most widespread method of ion production is by electron impact. The other fundamental, though little used method, is that of photoionization. In recent years a number of other methods have been developed, such as fast atom bombardment (FAB) and electrospray (ES) both of which are known as 'soft' methods of ionization in that they transfer relatively little energy to the ion. A bonus with ESMS lies in the fact that multiply charged ions are produced, thus extending the mass range. Thus, for an ion $m^{20+}$ the effective mass range will be 20 times that of a singly charged ion. These two techniques have had considerable application in biological and medical mass spectrometry. An alternative soft ionization method is to use low-energy electrons in impact ionization. If the measurements are also carried out using a cooled source the process produces what are known as LELT (low energy, low temperature) spectra.

In the electron impact source a beam of electrons (usually 70 eV) impacts the gaseous substrate under investigation and removes an electron from it, thus producing an ion. This is a drastic method since

ionization energies are usually of the order of 10 eV and chemical energies of the order of a few volts. The processes occurring are:

$$e^- + M \rightarrow M^{\bullet+} + 2e^- \quad [2]$$

$$e^- + M \rightarrow F^+ + F^\bullet + 2e^- \quad [3]$$

$$e^- + M \rightarrow F^{\bullet+} + F + 2e^- \quad [4]$$

Process [2] represents ionization to form the molecular ion while [3] represents fragmentation to form an even electron ion and [4] represents fragmentation to form an odd electron ion. In writing equations for fragmentation it is essential that 'electron bookkeeping' be maintained. What is shown here is primary fragmentation – the ions $F^+$ can further fragment to give secondary fragments and so on.

**The analyser** It has been seen that a transverse magnetic field can separate an unresolved beam of ions according to their *m*/*z* values (**Figure 5**). Such a system gives direction focusing. An electric field (see **Figure 6**) can give direction focusing and so lead to a double focusing mass spectrometer capable of high-resolution measurements.

**The collector** The usual collector is an electron multiplier which can give gains of $10^7$ or more. The output is sent to a recorder or data system. An earlier form is the Faraday cup which collects the electrons – the current then being amplified and recorded.

### The quadrupole

Originally the tool of physicists and physical chemists, now with improved electronics the quadrupole mass spectrometer has become an essential instrument for biological and biomedical research. Originally described as a mass filter, it operates by using a combination of a quadrupole static electric field and a radiofrequency field which combine to focus an ion beam on a collector.

### The ion trap

This device is related to the quadrupole, being a three-dimensional quadrupole. The ion trap consists of a hyperbolic ring electrode (doughnut) and two hyperbolic end electrodes. To obtain a spectrum a variable amplitude radiofrequency is applied to the doughnut whilst the end plates are grounded. As the amplitude is increased ions of increasing *m*/*z* are collected.

### The time-of-flight mass spectrometer

In this instrument ions produced in the source are accelerated to a given velocity. The unresolved beam is then injected into a field-free region and the ions drift towards the collector. The velocities will be inversely proportional to the square roots of the masses. This means that a pulse of ions will split up according to the ionic masses. The unresolved beam thus becomes resolved in time. Provided that the response time of the electronics is sufficiently fast a spectrum can be recorded. Obviously an average over many such pulses is necessary to provide a reliable signal. Once again the electronics lie at heart of this problem, which demands very fast amplifiers. Initially the time-of-flight mass spectrometer (TOF) was the province of physicists and later of chemists but, with the tremendous advance in electronics, instruments are now produced that are capable of routine operation by relatively untrained operators.

### The ion cyclotron resonance mass spectrometer

An ion cyclotron resonance (ICR) spectrometer creates a pulse of ions in a magnetic field. These are brought into resonance by scanning the applied radiofrequency. From the cyclotron resonance frequency and the magnetic field strength the *m*/*z* ratio can be calculated. The use of a fast Fourier transform (FT-ICR) refines the method.

## The energetics of ionization and fragmentation

### The thermochemistry of ions

Just as the thermochemistry of neutral molecules has led to an understanding of the structure, stability and kinetics of chemical species, the thermochemistry of ions has led to a corresponding understanding of ionic species in the gas phase. Thus the enthalpy of formation ($\Delta_f H^\circ(M^{\bullet+})$ of the molecular ion is given by Equation [5].

$$\Delta_f H^\circ(M^{\bullet+}) = IP(M) - \Delta_f H^\circ(M) \quad [5]$$

where IE(M) is the ionization energy of the molecule and $\Delta_f H^\circ$ is the enthalpy of formation of the neutral molecules. Holmes and co-workers have published a very useful algorithm for estimating the enthalpies of

**Table 2** The enthalpies of formation of *n*-alkane molecular ions

| *Molecule* | $\Delta_f H°$ ($M^{\cdot+}$) (kJ mol$^{-1}$)[a] | $\Delta_f H°$($M^{\cdot+}$) (kJ mol$^{-1}$)[b] |
|---|---|---|
| $CH_4$ | 1142 | 1142 |
| $C_2H_6$ | 1025 | 1021 |
| $C_3H_8$ | 954 | 950 |
| $C_4H_{10}$ | 891 | 895 |
| $C_5H_{12}$ | 854 | 854 |
| $C_6H_{14}$ | 816 | 816 |
| $C_7H_{16}$ | 778 | 778 |

[a] Experimental value.
[b] Theoretical value.

formation of odd electron ions. Some typical values for hydrocarbons are shown in **Table 2**. The agreement between experimental and theoretical values is excellent. Often the enthalpies of formation of the substrate molecule are not known and so recourse has to be made to empirical methods such as that of Benson for estimation of the value.

In the case of the even electron ions one has, mainly, to have recourse to experimentally determined values. The enthalpies of formation of the even electron ions are given by Equation [6] where the appearance energy is represented by AE($F^+$), with $\Delta_f H°(F^{\cdot+})$, $\Delta_f H°(F^{\cdot})$and $\Delta_f H°$(M) being the enthalpies of formation of the ion, the radical and the molecule.

$$\Delta_f H°(F^{\bullet+}) \leq AE(F^+) + \Delta_f H°(F^{\bullet}) + \Delta_f H°(M) \quad [6]$$

In Equation [6] the inequality may be replaced by the equality in most instances. Some values for the primary carbonium ions are shown in **Table 3**. Values such as these can then be used in calculating ionization and appearance energies. These are, respectively, the lowest energy at which the molecular ion appears and the lowest energy at which a fragment ion appears. Thus the ionization energy is given by:

$$IE = \Delta_f H°(M^{\bullet+}) - \Delta_f H°(M) \quad [5a]$$

on rearranging Equation [5]. Similarly, the appearance energy is obtained by rearranging Equation [6].

$$AE(F^+) = \Delta_f H°(F^{\bullet+}) + \Delta_f H°(F^{\bullet}) - \Delta_f H°(M) \quad [6a]$$

Holmes and Lossing have developed an ingenious method of measuring the enthalpies of formation of neutrals by a further rearrangement of Equation [6]. This is extremely useful where the enthalpy of formation of the neutral has not been measured. The method depends on measuring the appearance energy of a fragment ion produced from different sources

$$\Delta_f H°(M) = \Delta_f H°(F^{\bullet+}) + \Delta_f H°(F^{\bullet}) - AE(F^+) \quad [6b]$$

and using the average value in Equation [6b].

**Table 3** Some values of the enthalpies of formation of carbonium ions

| *Molecule* | $\Delta_f H°$ *($F^{\cdot+}$)* (kJ mol$^{-1}$) |
|---|---|
| $CH_3$ | 1092 |
| $C_2H_5$ | 916 |
| $C_3H_7$ | 870 |
| $C_4H_9$ | 841 |
| $C_5H_{11}$ | 812[a] |
| $C_6H_{13}$ | 791[a] |
| $C_7H_{15}$ | 766[a] |

[a] Estimated value.

### Metastable ions

It will be seen in **Figure 6** that there are two important field-free regions (FFR) in the double focusing mass spectrometer, namely between the source and the electric analyser (FFRI) and between the electric and magnetic analysers (FFR2). It may so happen that in flight an ion decomposes in FFR2 in which case a diffuse peak appears in the mass spectrum at the position *m*/*z* given by Equation [7]

$$m^*/z = {m_2}^2/m_1 \quad [7]$$

for process [8]. A typical metastable peak is shown in **Figure 8** for the process

$${m_1}^+ = {m_2}^+ + (m_1 - m_2) \quad [8]$$

The appearance of a metastable peak is confirmation of a fragmentation route, but absence of the peak does not indicate the absence of a fragmentation. The reason is that metastable ions are relatively long lived. If the fragmentation is rapid no metastable will be seen. A special scan, keeping *B*/*E* constant, will record all the daughter peaks resulting from a given parent ion. Equally, a scan keeping $B^2/E$ constant will give all the progenitors of a given peak.

**Figure 8** A metastable in the mass spectrum of anisole.

These scans are very useful in mapping out the fragmentation patterns of a given ion.

### Collision induced dissociation

Another means of producing fragmentation involves collision processes – bimolecular as compared with the unimolecular processes previously discussed. In this method a beam of energetic ions is brought into collision with neutral molecules and fragmentation results – collision induced dissociation (CID). The spectra thus obtained were complex since they derived from an unresolved beam of ions. It was realized that it would be advantageous if the ions for collision were separated from the unresolved beam. This led to the development of a reversed geometry instrument – the ZAB, produced by Vacuum Generators. Finally there was the introduction of multisector instruments which gave rise to the technique of mass spectrometry–mass spectrometry (MS-MS). CID has proved very useful in assigning structures to fragment ions.

### The future

Further developments in fundamental mass spectrometry will have to await for universities to return to their basic task – the pursuit of fundamental research. At the present time many workers in the field have to design their research to attract funds. This often leads to hack research – not always in the best interest of the subject or the scientists. It is to be hoped that the new millennium will see the universities of the world returning to their proper research areas, namely fundamental research. Only in this way will mass spectrometry develop in its fundamental aspects which in turn will lead to new and more powerful techniques.

## The literature of mass spectrometry

1968 saw the first of the journals devoted to mass spectrometry. *Organic Mass Spectrometry* (OMS) and the *International Journal of Mass Spectrometry and Ion Physics* (IJMSIP). Later OMS spawned *Biomedical Mass Spectrometry* (BMS). IJMSIP has since changed its name to *The International Journal of Mass Spectrometry and Ion Processes* and latterly to the *International Journal of Mass Spectrometry*, while OMS and BMS have been incorporated in the *Journal of Mass Spectrometry*. The American Society for Mass Spectrometry has produced a Journal – *Journal of the American Society for Mass Spectrometry*. To facilitate rapid publication, *Rapid Communications in Mass Spectrometry* was born – the authors nominate their own referees.

## List of symbols

$B$ = magnetic field strength; $m$ = mass of an ion; $R$ = radius of the magnetic field; $V$ = electric field strength; $z$ = charge on an ion,; $\Delta_f H°$ = enthalpy of formation.

*See also:* **Chemical Ionization in Mass Spectrometry; Fast Atom Bombardment Ionization in Mass Spectrometry; Fragmentation in Mass Spectrometry; Ion Structures in Mass Spectrometry; Ion Trap Mass Spectrometers; Ionization Theory; Ion Energetics in Mass Spectrometry; Ion Collision Theory; Metastable Ions; Quadrupoles, Use of in Mass Spectrometry; Sector Mass Spectrometers; Statistical Theory of Mass Spectra; Time of Flight Mass Spectrometers.**

## Further reading

Aston FW (1924) *Isotopes*, 2nd edn. London: Edward Arnold.

Beynon JH and Morgan RP (1978) The development of mass spectrometry: an historical account. *International Journal of Mass Spectrometry and Ion Physics* **27**: 1–30.
Thomson JJ (1898) *The Discharge of Electricity through Gases.* London: Archibald Constable.
Thomson JJ (1913) *Rays of Positive Electricity and their Application to Chemical Analyses*, p 56. London: Longmans and Green.

# Mass Transport Studied Using NMR Spectroscopy

*See* **Diffusion Studied Using NMR Spectroscopy.**

# Materials Science Applications of X-Ray Diffraction

**Åke Kvick**, European Synchrotron Radiation Facility, Grenoble, France

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| **Applications** |

The X-ray diffraction technique is widely used in structural characterization of materials and serves as an important complement to electron microscopy, neutron diffraction, optical methods and Rutherford backscattering.

The early uses were mainly in establishing the crystal structures and the phase composition of materials but it has in recent years more and more been used to study stress and strain relationships, to characterize semiconductors, to study interfaces and multilayer devices, to mention a few major application areas.

One of the important advantages of X-ray diffraction is that it is a nondestructive method with penetration from the surfaces into the bulk of the materials. This article will outline some of the most important areas including some rapidly developing fields such as time-dependent phenomena and perturbation studies.

## X-ray sources

X-rays are electromagnetic in nature and atoms have moderate absorption cross-sections for X-ray radiation resulting in moderate energy exchange with the materials studied, making diffraction a nondestructive method, in most cases.

Traditionally X-rays are produced by bombarding anode materials with electrons accelerated by a >30 kV potential. The collision of the accelerated electrons produces a line spectrum superimposed on a continuous spectrum called *bremsstrahlung.*

The line spectrum is characteristic of the bombarded anode material and has photon intensities much higher than the continuous spectrum. The characteristic lines are generated by the relaxation of excited electrons from the electron shells and are labelled K, L, M, etc. and signify the relaxation L to K, M to K, etc.

A table of available laboratory wavelengths is given in **Table 1**.

The increased importance of X-ray diffraction in materials science is coupled to the recent emergence of a new source of X-rays based on *synchrotron radiation* storage rings. The synchrotron radiation is produced by the bending of the path of relativistic charged particles, electrons or positrons, by magnets causing an emission of intense electromagnetic radiation in the forward direction of the particles. The

**Table 1** Radiation from common anode materials

| *Radiation* | *Wavelength (Å)* | *Energy (keV)* |
|---|---|---|
| Ag $K_\alpha$ | 0.5608 | 22.103 |
| Pd $K_\alpha$ | 0.5869 | 21.125 |
| Rh $K_\alpha$ | 0.6147 | 20.169 |
| Mo $K_\alpha$ | 0.7107 | 17.444 |
| Zn $K_\alpha$ | 1.4364 | 8.631 |
| Cu $K_\alpha$ | 1.5418 | 8.041 |
| Ni $K_\alpha$ | 1.6591 | 7.742 |
| Co $K_\alpha$ | 1.7905 | 6.925 |
| Fe $K_\alpha$ | 1.9373 | 6.400 |
| Mn $K_\alpha$ | 2.1031 | 5.895 |
| Cr $K_\alpha$ | 2.2909 | 5.412 |
| Ti $K_\alpha$ | 2.7496 | 4.509 |
| Synchrotron | ~0.05–3 | 4.300 |

The value $\alpha$ is a mean of the $K_{\alpha 1}$ and $K_{\alpha 2}$ emissions. The synchrotron radiation is continuous and the range is the most commonly used. The range may be extended on both sides.

photons are generated over a wide energy range from very long wavelengths in the visible to hard X-rays up to several hundred keV.

The radiation is very intense and exceeds the available normal laboratory sources by up to 6 or 7 orders of magnitude. The synchrotron storage rings used for the radiation production, however, are large and expensive, with facilities characterized by storage rings with a circumference up to more than one thousand metres.

The main advantages of synchrotron radiation are:

1. continuous radiation up to very high energies (>100 keV);
2. high intensity and brightness;
3. pulsed time structure down to picoseconds;
4. high degree of polarization.

Figure 1 illustrates a modern synchrotron facility with many experimental facilities in a variety of scientific areas from atomic physics to medicine.

**Figure 1** Beam lines at the European Synchrotron Radiation Facility in Grenoble, France.

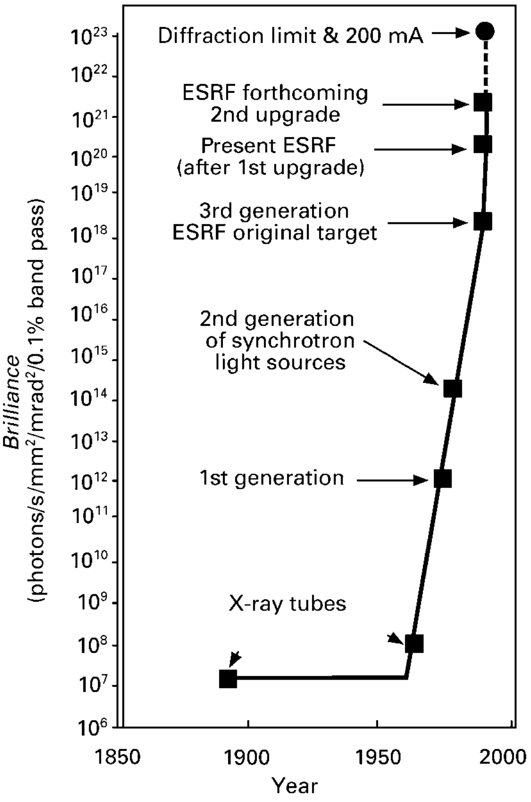



**Figure 2** The brightness defined as photons/s/mm²/mrad²/0.1% energy band pass for conventional and synchrotron X-ray sources. ESRF denotes the European Synchrotron Radiation source in Grenoble, France.

Figure 2 compares the brightness of the available X-ray sources.

## X-ray diffraction

The diffraction method utilizes the interference of the radiation scattered by atoms in an ordered structure and is therefore limited to studies of materials with long-range order.

The incoming X-ray beam can be characterized as a plane wave of radiation interacting with the electrons of the material under study. The interaction is both in the form of absorption and scattering. The scattering can be thought of as spheres of radiation emerging from the scattering atoms. If the atoms have long-range order the separate ‘spheres’ interfere constructively and destructively producing distinct spots, Bragg reflections, in certain directions.

The specific scattering angles, $\theta_{hkl}$, carry information on the long-range ordering dimensions and the intensity gives information on the location of the electrons within that order.

The basis for all material science studies using X-ray diffraction is Bragg's law:

$$\lambda = 2d_{hkl}\sin(\theta_{hkl}) \qquad [1]$$

where $\lambda$ is the wavelength of the incoming radiation, $d_{hkl}$ is the spacing of the (hkl) atomic plane and $\theta$ is the angle of the diffracting plane where constructive interference occurs. (see Figure 3).

Differentiation of Bragg's law gives the expression:

$$\Delta(\theta) = -\Delta(d)/d\tan(\theta) \qquad [2]$$

which is an important formula relating the observed changes in scattering angles to structural changes in the material.

The penetration depth of the probing radiation is an important parameter in designing a diffraction experiment. The penetration depth is associated with the absorption of the radiation, which is a function of the absorption cross-section of the material under study. The absorption can be calculated by the formula:

$$I/I_o = \exp(-\mu t) \qquad [3]$$

where $I_0$ is the intensity of the incident beam, I is the intensity of a beam having passed through t (cm) of material with an absorption coefficient of $\mu$ ($cm^{-1}$).

The absorption coefficient $\mu$ can be calculated as an additive sum over the different atomic species in the unit cell:

$$\mu = 1/V_c\,\Sigma(\sigma_n) \qquad [4]$$

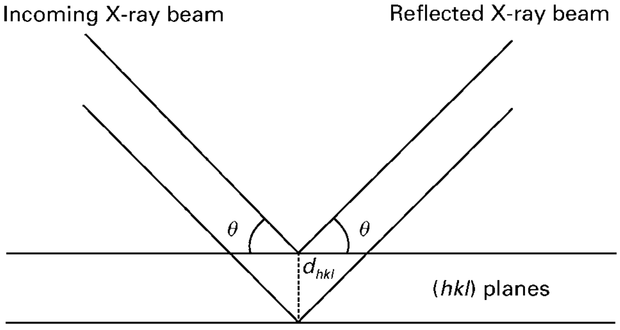



**Figure 3** Reflection from the planes (*hkl*) with interplanar spacing $d_{hkl}$.

where $V_c$ is the volume of the unit cell and $\sigma_n$ is the absorption cross-section for component n. The absorption cross-sections vary as a function of the wavelength and can be calculated using the Victoreen expression:

$$\mu/\rho = C\lambda^3 - D\lambda^4 + \sigma_{K-N}NZ/A \qquad [5]$$

where $\rho$ is the density of the material with atomic number Z and the atomic weight A. The constants C, D and $\sigma_{K\text{-}N}N$ vary with the wavelength. Tabulations for various materials can be found in International Tables for Crystallography, Vol III, pp 161 ff.

It can be noted that the absorption drops off with decreasing wavelength and the penetration depth can thus be changed with a change in wavelength. A quantity called penetration distance, $\tau$, is usually quoted for penetration depths and is defined as the distance where $I/I_0$ is reduced to 1/e. Penetration distances for a few elements are listed in Table 2, together with a comparison with other methods.

## Structure determinations

Historically, and even today, the structure determination of crystalline materials is the most important application of X-ray diffraction in materials science. The relative intensities of Bragg reflections carry information on the location of the electrons in the solids and thus give precise information on the relative positions and thermal motion of the atoms. Even information on the bonding electrons may be obtained.

The scattered intensities from different planes (hkl) in a crystal are measured using precise diffractometers that orient the sample with respect to the incident X-ray beam for all the possible diffraction planes in the crystal. Intensities are measured using scintillation, semiconductor CCD or imaging plate detectors. The measured intensities are converted, after various geometric corrections, to the amplitude

**Table 2** Penetration depth $\tau$ (1/e) in Al, Fe and Cu for various techniques in millimetres

| | *Al* | *Fe* | *Cu* |
|---|---|---|---|
| Scanning electron microscope | <0.001 | <0.001 | <0.001 |
| X-ray diffraction (Cu $K_\alpha$) | 0.14 | 0.007 | 0.005 |
| Synchrotron X-rays (80 keV) | 18 | 2.1 | 1.4 |
| Synchrotron X-rays (300 keV) | 36 | 11.5 | 10 |
| Neutrons (cold ) | 97 | 8.3 | 10 |

structure factors $|F_{hkl}|$ which are proportional to the scattered intensities.

The factors can be expressed as:

$$F_{hkl} = \sum^{j} f_j \exp\{-2\pi i(hu_j + k\nu_j + lw_j)\} \quad [6]$$

where the $f_j$ are the atomic scattering factors for atoms j and $u_j$, $v_j$ and $w_j$ are the fractional coordinates for atom j.

The electron density in the crystal unit cell can be determined from the structure factors $F_{hkl} = |F_{hkl}| \times e^{i\varphi hkl}$. The structure amplitudes ($|F_{hkl}|$) are obtained from the experiments and the phases $\varphi$ can be obtained from a number of phasing procedures. The structure factors can then be converted to a mapping of the electron density distribution $\rho$(xyz), which completely determines the crystal structure.

$$\rho(xyz) = 1/V_c\, \Sigma^{hkl} |F_{hkl}| e^{i\varphi hkl} \exp(-2\pi(hx + ky + lz)) \quad [7]$$

$V_c$ is the volume of the unit cell and x,y,z are the positional coordinates in the unit cell.

The most precise structure determinations are performed using a single crystal sample where hundred to thousands of separate Bragg spots can be used to determine atomic positions to a precision of better than 0.001 Å. In materials science, however, quite frequently the material is in a polycrystalline form and powder diffraction methods have to be employed.

In the powder diffraction method the Bragg reflections from many thousands of microcrystals with different orientation overlap giving rise to 'powder' diffraction rings rather than distinct Bragg spots. Since the powder patterns are in the form of rings rather than distinct spots (see Figure 4) the observed intensities from planes with identical or closely similar scattering angles are analysed in terms of a deconvolution function:

$$Y_i = B_i + \Sigma^{hkl} I(hkl)\, Ng \quad [8]$$

where $B_i$ is the background in point i, g is the powder peak shape function and N is a normalization factor. Once a structure model is obtained a refinement procedure proposed by Rietveld can be used to minimize the quantity

$$M = \Sigma^i\, W(Y_i(\text{obs}) - Y_i(\text{calc}))^2 \quad [9]$$

W is a weight factor based on counting statistics. Y(calc) may contain positional, thermal parameters, unit cells as well as peak shapes and these parameters are refined during the minimization.

The method is used for structure determinations and is well suited to handling multiphase systems. At present the method gives bond distances to within ≥0.005 Å precision, and ab initio structure determinations may handle systems with up to 200 parameters. Figure 5 gives an example of a powder pattern collected at the European Synchrotron Radiation Facility of a two-phase powder sample.

The broadening of a diffraction line may give important information on changes occurring during the processing of the material. The line width is affected by instrumental resolution, source size, divergence of the radiation, particle size, microstrain components and stacking faults. After appropriate correction qualitative information on the micrograin size can be obtained by using the Scherrer formula. The grain size affects the line broadening as a function of wavelength and scattering angle as:

$$\Delta 2\theta = \lambda/(D \cos\theta) \quad [10]$$

where D is the particle size.

In order to separate particle broadening from strain effects two or more reflections may be used. The strain varies as function of tan $\theta$ and the different effects in line broadening may thus be separated by determining the broadening from different (hkl) diffraction lines.

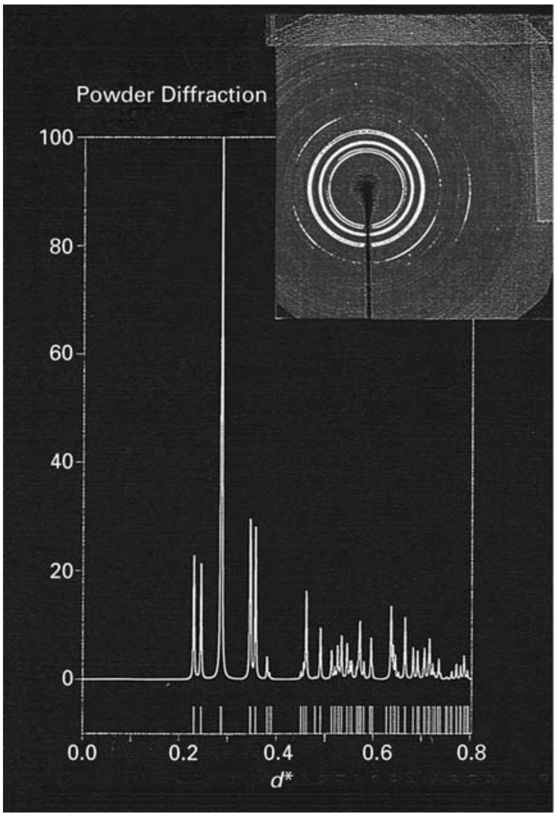
## Anomalous scattering

The absorption and scattering can change considerably around the absorption edges for the atoms. The absorption edge region is the atom specific energy region where the inner electrons are sufficiently excited by the X-rays to leave the atom or to be excited to an upper electronic shell. In this region the atomic scattering factor $f_j$ is modified and is no longer a real property but has, in addition to an extra real term $f'_j$, an imaginary term $f''_j$.

The term in Equation [6] changes to:

$$f_j = f_j^0 + f'_j + i f''_j \quad [11]$$

These changes can be used to alter the absorption and consequently the penetration depth by changing the wavelength around the absorption edges. The changes may be considerable and can amount to the equivalence of tens of electrons.



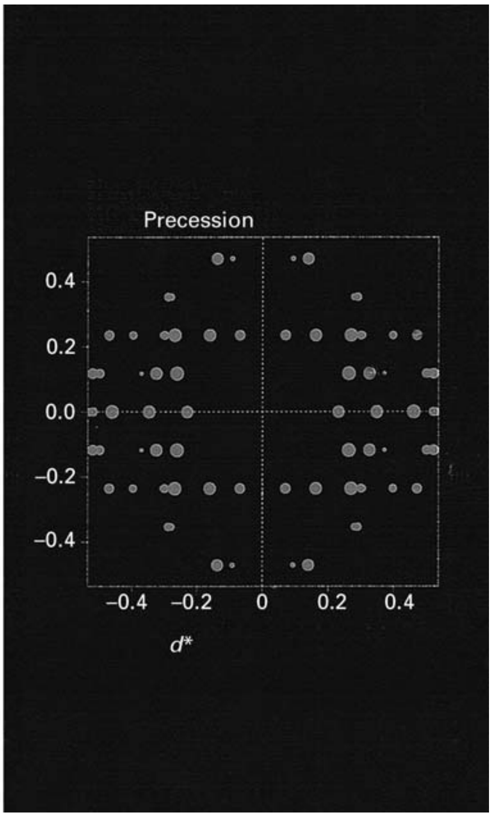




**Figure 4** The left-hand picture illustrates a powder diffraction pattern of $S_2N_2$ at a wavelength of 0.325 Å. The insert is the actual pattern which can be integrated and displayed as a function of $d^*$ [$d^* = (2 \sin \theta)/\lambda$]. The right-hand picture shows the result if one single $S_2N_2$ crystal is rotated in the X-ray beam. The size of the spots illustrates the difference in diffracted intensities from the separated Bragg reflections. Courtesy of Svensson and Kvick, 1998. (see Colour Plate 36).

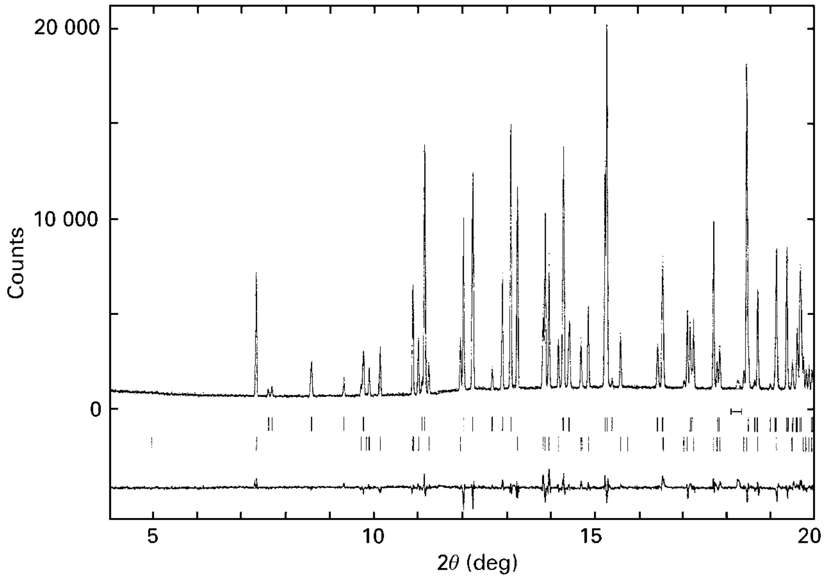



**Figure 5** The powder pattern from a two-phase mixture of orthorhombic $(CH_3)_2SBr_{2.5}$ and monoclinic $(CH_3)_2SBr_4$ recorded at a wavelength of 0.94 718 Å at the BM16 beamline at the ESRF. The two lines of vertical bars show the location of diffraction lines for the two different phases. The bottom pattern shows the difference between the observed and refined intensities from the Rietveld refinement. Courtesy of Vaughan, Mora, Fitch, Gates and Muir, 1998.

In addition to changes in absorption there are fundamental differences in the scattering which may be used in materials science. Examples include the determination of absolute optical configuration in polar compounds.

Normally, the Friedel pairs of structure factors, i.e. F(hkl) = F(-h-k-l), are identical. However, around the absorption edges this law breaks down since the imaginary term adds differently in the phase relation and the correct optical isomer can be determined.

The alternation of the scattering for a specific atom may also be used to resolve the partial contribution of that atom at a specific site by multiple wavelength differences studies.

The scattering contrast between almost isoelectronic elements such as Fe and Co may be enhanced sufficiently to allow precise determinations. Small chemical shifts in the energy of the edges depending on the valence state of the atom may also be used to differentiate between the valence states of the atoms in a compound.

The effect is particularly useful for structure determinations of macromolecules where multiple anomalous diffraction (MAD) experiments are combined to yield the phase factors necessary to obtain the electron density mapping.

The use of anomalous diffraction has gained in importance as synchrotron radiation sources have become available. The synchrotron sources provide easily tuneable radiation covering most of the atomic absorption edges.

## Magnetic scattering

During the last few years a minute scattering component coming from the magnetic moments of atoms has been exploited. The scattering from the magnetic moments are smaller than the scattering from the electrons by orders of magnitude. However, successful studies of magnetic structures and even magnetic atomic overlayers are now possible by using the high brightness synchrotron sources.

The synchrotron radiation also makes it possible to tune the wavelength to the absorption edges of the magnetic atoms where the magnetic scattering is strongly enhanced. For a more detailed account of anomalous diffraction and magnetism the reader is referred to the book edited by Materlik and co-workers (1994).

## Stress and strain relationships

Elastic X-ray stress analysis is based on recordings of the interplanar distances $d_{hkl}$. When stress is applied to a crystal the distances change from $d_0$ to $d_0 + \Delta d$ and the scattering in Bragg's law (Eqn [1]) will change accordingly. Using Equation [2] one obtains the relationship:

$$\Delta\theta = (\theta - \theta_0) = -\Delta d/d_0 \tan(\theta_0) = -\tan(\theta_0)\varepsilon \quad [12]$$

where $\theta_0$ is the scattering angle from a strain free state, and $\varepsilon$ is the elastic strain coefficient. $\theta$ is obtained from the diffraction experiment.

In strain analysis we can use the formalism given by Noyen and Cohen (1976) where the strain in the direction of a scattering plane (hkl) is measured as:

$$\varepsilon'(33)_{\phi\varphi} = (d_{\phi\varphi} - d_0)/d_0 \quad [13]$$

The plane spacing $d_0$ is measured from a strain free sample.

To obtain the strain in the sample system S one transforms the values from the laboratory coordinate system L.

The orthogonal coordinate systems $S_1$,$S_2$,$S_3$ and $L_1$,$L_2$,$L_3$, are defined as follows. $S_3$ is perpendicular to the sample surface and $S_2$ and $S_3$ are parallel to the sample surface. $L_3$ is the normal to the scattering plane hkl and makes the angle $\varphi$ with $S_3$. The $\phi$ angle is the angle between the projection of $L_3$ onto the sample surface and vector $S_1$. The measured quantity $\varepsilon_{\phi\varphi}$ can then be converted to the sample coordinate system by the transformation:

$$\begin{aligned}\varepsilon'(33)_{\phi\varphi} &= \varepsilon_{11}\cos^2\phi\sin^2\varphi + \varepsilon_{12}\sin 2\phi\sin^2\varphi \\ &+ \varepsilon_{22}\sin^2\phi\sin^2\varphi\,\varepsilon_{33}\cos^2\varphi + \varepsilon_{13}\cos\phi\sin^2\varphi \\ &+ \varepsilon_{23}\sin\phi\sin^2\varphi \quad [14]\end{aligned}$$

This important equation is linear in $\varepsilon_{11}$, $\varepsilon_{12}$, $\varepsilon_{22}$, $\varepsilon_{33}$, $\varepsilon_{13}$, $\varepsilon_{23}$ and these coefficients can thus be determined from six measurements at different angular values. When this is known the stress can be determined.

The stress ($\sigma$) and strains ($\varepsilon$) in the sample coordinate system are determined using Hooke's law:

$$\begin{aligned}\sigma_{ij} &= c_{ijkl}\,\varepsilon_{kl} \\ \varepsilon_{ij} &= s_{ijkl}\,\sigma_{kl} \quad [15]\end{aligned}$$

where the $c_{ijkl}$ are the elastic stiffness constants and $s_{ijkl}$ are the elastic compliances.

In the elastic case the observations may be expressed in terms of Young's modulus E and Poisson's ratio $\nu$:

$$(d_{\phi\varphi} - d_0)/d_0 = (1 + \nu/E)\sigma_{ij} - (\delta_{ij}\nu/E)\sigma_{kk} \quad [16]$$

where $\delta_{ij}$ is the Kronecker delta.

This formalism is the basis for the commonly used $\sin^2\varphi$ method, which is usually limited to the surface region of the sample.

The availability of more elaborate diffractometers and the highly penetrating synchrotron radiation also gives possibilities of deeper penetration into the sample by rotation of the sample around the scattering vector direction, $L_3$, and by wavelength tuning or change in the scattering angle. Transmission studies permit the study of stresses in specific 'gauge' volumes in the bulk.

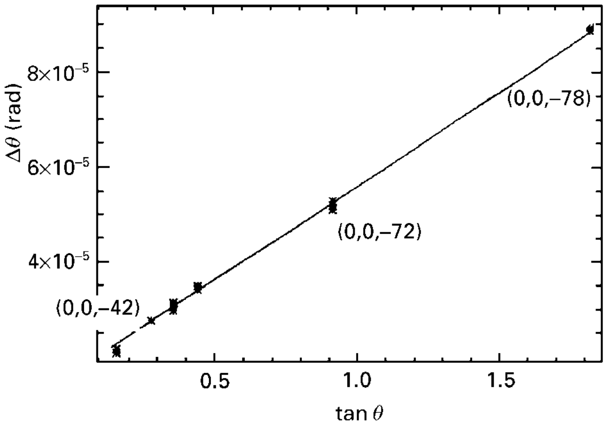



**Figure 6** The shifts in $\Delta\theta$ for (00*l*) reflections from a $LiNbO_3$ single crystal of 0.2 mm thickness subjected to an electric field of 50 kV $cm^{-1}$ along the crystallographic *c* direction as measured using synchrotron radiation of a wavelength of 0.307 Å at the beamline 1D11 at the ESRF. The *d*(33) element of the piezoelectric tensor can be evaluated to be $7.5(2) \times 10^{-12}$ $CN^{-1}$. Courtesy of Heunen, Graafsma, Kvick, 1997.

## Externally perturbed systems

Deformations due to perturbations may also be followed by X-ray diffraction.

These perturbations may be caused by electric field or changes in temperature and thus piezoelectric effects and thermal expansion coefficients may readily be determined by using Equation [2].

In the case of changes occurring from the application of an external electric field, the converse piezoelectric effect, the induced strain coefficient $\varepsilon_{ij}$ is related to the field by equation:

$$\varepsilon_{ij} = \sum_{k=1}^{3} d_{kij}E_k \quad [17]$$

where $E_k$ is kth component of the electric field and $d_{kij}$ is the kijth element of the third rank piezoelectric tensor.

Using Equation [2] Barsch has shown that the diffraction observable $\Delta\theta_r$ may be used to determine the piezoelectric tensor according to the formula (Coppens, 1992):

$$\Delta\theta_r = -E\tan\theta_r \sum_{k=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} e_k\, h_{r,i}\, h_{r,j}\, d_{kij} \quad [18]$$

where r refers to certain reflections (hkl) and $e_k$, $h_{r,i}$ and $h_{r,j}$ are, respectively, the components of unit vectors parallel to the electric field and the scattering vector for (hkl).

Figure 6 gives an illustration where the piezoelectric tensor element $d_{33}$ has been determined from measurements of a series of 00l reflections with the electric field aligned along the (00l) direction.

## Ion-implantation effects

Ion implantation in layers of semiconductor material is an important process where Equation [2] may be very useful for assessing the effect on the implantation.

If dopants are introduced by ion implantation, for instance in silicon, the tetrahedral radius of the dopant ($r_d$) differs from that of the substrate and changes in the $\theta$ angle will be observed.

For cubic silicon material the corresponding change in the unit cell dimension can give information on the doping level according to Vegard's law:

$$\Delta a/a = (1 - r_d/r_s)KC_dN^{-1} \quad [19]$$

where $r_d$ and $r_s$ are the tetrahedral radii of the dopant and substrate respectively, $C_d$ is the concentration of dopant and N is the number of substrate of atoms per unit volume.

The factor K takes account of the fact that only the lattice strain normal to the wafer is nonzero, whereas the in-plane strain is zero. For typical substrates with cubic structure the K values can be evaluated from the elastic stiffness constants.

## Superlattice structures

Physical properties of materials may be changed by creating ‘artificial’ structural periodicity by depositing alternate thin layers of different materials. Structures of this type are commonly used in optics, electrooptical applications and coatings for corrosion protection or thermal barriers.

The production and performance of these structures can be characterized by X-ray diffraction. The diffraction patterns from the compound structures are characterized by main diffraction peaks interspersed by satellite peaks. The periodicity of superlattices, $\boldsymbol{\Lambda}$, is given by the relationship:

$$\boldsymbol{\Lambda} = (j-k)\ \lambda/2(\sin\theta_j - \sin\theta_k) \qquad [20]$$

where j and k represent the satellite order number and $\theta_j$ and $\theta_k$ are the observed scattering angles at wavelength $\lambda$. The periods may vary depending on the production and the variation can be monitored by the various $\boldsymbol{\Lambda}$ values obtained from different satellite pairs. These period variation may be interpreted as surface roughness.

## Topography

Diffraction topography is an imaging technique based on Bragg’s law (Eqn [1]) and ranges from a rather qualitative inspection method of the scattering power over a crystal to a much more complicated method employing dynamical scattering theory to elucidate microscopic strains in the crystal. Many phenomena of importance to a materials scientist, such as stacking faults, growth bands, strain around grain boundaries, twins, or even dynamical phenomena such as acoustic waves or magnetic domain formation, have been studied by this method.

The studies can be divided into two main areas:

(a) orientation contrast;
(b) extinction contrast.

The former studies (a) map the variation of the scattering power across the X-ray beam and detects misalignment of certain portions of the crystal where the misalignment is larger than the divergence of the monochromatic incident beam. The misalignment can be due to rotations or dilations of the lattice. When monochromatic radiation is used these effects are seen as distinct bands or patterns of loss of scattered radiation.

A richer pattern of intensity variations is seen if continuous wavelength radiation is used. In this case, due to divergence or convergence of the diffracted beams at the boundaries, losses and gains in intensity are observed.

The second phenomenon (b) is observed when the scattering power in the crystal is alternated by strain

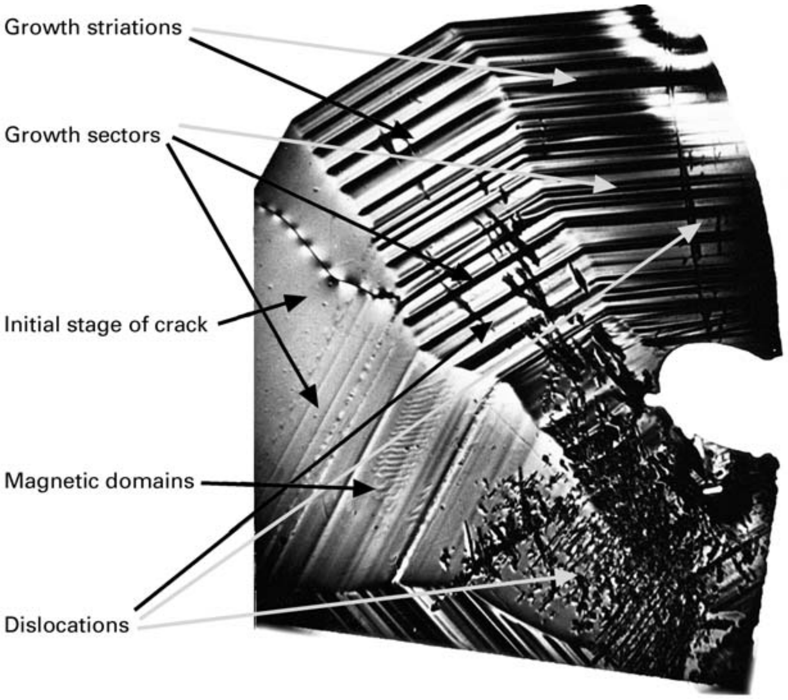



**Figure 7** Transmission topograph of a flux grown Ga-YIG ($Y_3Fe_{5-x}Ga_xO_{12}$, $x \approx 1$) crystal plate, Mo $K_{\alpha1}$-radiation ($\lambda$ = 0.709 Å). Courtesy of J Baruchel, 1998.

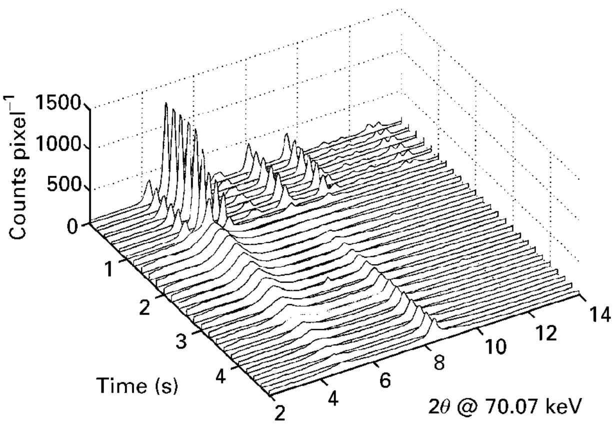



**Figure 8** The time-evolution of the powder pattern of an exothermic reaction of a mixture of Al and Ni powders. The powder patterns were recorded at the 1D11 beam line at the ESRF every 250 ms during the self-propagating high-temperature synthesis. The wavelength 0.177 Å was chosen to allow penetration of the sample. It can be noticed that the original peak from Al and Ni disappear or change during the reaction and the formation of new phases and the nucleation can be followed. Courtesy of Kvick, Vaughan, Turillas, Rodriguez, Garcia, 1998.

fields around defects without major realignment in the crystal.

The observed contrasts are rather complicated to quantify and an understanding of the effects of the X-ray wavefields is necessary to interpret the observations in detail. For a detailed account of the theory the reader is referred to Tanner (1976).

Topography uses two different methods of observation; reflection topography, which maps the surface region of the samples, and transmission topography, which also samples the bulk of the material. In the latter case the radiation must be chosen so that the absorption is small enough to allow bulk penetration.

Different volumes in the crystal may be sampled if the collimation of the incident and diffracted beams is carried out carefully. This method is called section topography.

Figure 7 gives an example of the information one may obtain from a transmission topograph.

## Time-resolved studies

Time-resolved X-ray diffraction has been used for a long time to study solid-state reactions. With the emergence of the new radiation sources it is now possible to follow solid-state reactions, phase transitions and physical changes caused by perturbations on much shorter time-scales.

Topography has already proved to be a suitable technique for studying magnetic domain formation or acoustic deformation down to time-scales of milliseconds.

The use of well collimated and high intensity synchrotron radiation beams is essential to reach the necessary time intervals without losing the statistical significance in the observed diffracted intensities. The white beam Laue technique has already been proven to facilitate studies down to the picosecond time regime for studies such as recombination of CO in myoglobin after flash photolysis. Nanosecond resolution has been obtained in a study of laser-annealing of defects in a silicon crystal.

The dynamics of reactions are either reversible or irreversible. Sufficient counting statistics can be obtained in the reversible processes through stroboscopic measurement where repeated measurements using short radiation time slices are performed. In the irreversible experiments the processes have to be followed by rapid consecutive exposures.

In many of the cases the reactions are often destructive to large single crystals and the powder method has to be employed. However, it has been proven that reactions such as polymerization and exothermic solid-state reactions may be followed down to the millisecond time-regime even when the reactions are irreversible.

Figure 8 exemplifies the time-evolution of an exothermic reaction between Al and Ni powders. In this time-resolved powder experiment one can follow the initial melting of Al and the formation of intermetallic phases as well as the crystallization process.

The rapid development of fast read-out CCD cameras in combination with the high brightness synchrotron sources promises to give X-ray diffraction a prominent role in the studies of dynamics of processes relevant to materials scientists.

## List of symbols

A = atomic weight; $B_i$ = background in point i; $C_d$ = concentration of dopant; C, D, $\sigma_{N-K}$, N = constants; $C_{ijkl}$ = elastic stiffness constants; d = atomic spacing; D = particle size; E = Young's modulus; $E_k$ = kth component of electric field; $f_j$ = atomic scattering factors; $F_{hkl}$ = amplitude structure factors; g = powder peak shape function; I = intensity of beam; $I_0$ = incident intensity of beam; j,k = satellite order number; K = constant; N = normalization factor; N = number of atoms; R = tetrahedral radius; $s_{ijkl}$ = elastic compliances; t = thickness of sample; $u_j$, $v_j$, $w_j$ = fractional coordinates for atom j; $\nu$ = Poisson's ratio; $V_c$ = volume of unit cell; W = weight factor; $Y_i$ = deconvolution function; Z = atomic number; $\delta_{ij}$ = Kronecker delta; $\Delta\theta$ = change of angle of scattering; $\varepsilon$ = elastic strain coefficient; $\theta$ = angle of scattering; $\lambda$ = wavelength; $\Lambda$ = periodicity of superlattices; $\mu$ = absorption

coefficient; $\rho$ = density; $\rho(xyz)$ = electron density distribution; $\sigma_n$ = absorption cross-section for component n; $\sigma$ = stress; $\tau$ = penetration distance; $\varphi$ = angle; $\phi$ = phase.

*See also:* **Fibres and Films Studied Using X-Ray Diffraction**; **Powder X-Ray Diffraction, Applications**; **Scattering and Particle Sizing, Applications; Scattering Theory.**

## Further reading

Authier A, Lagomarsino S and Tanner B (1996) *X-ray and Neutron Dynamical Diffraction*. New York and London: Plenum Press.

Coppens P (1992) *Synchrotron Radiation Crystallography*. London, San Diego: Academic Press.

*International Tables for Crystallography*, Vol I–III (1989) Dordrecht, Boston, London: Kluwer Academic Publishers.

*International Tables for Crystallography*, Vol A–C (1992) Dordrecht, Boston, London: Kluwer Academic Publishers.

Materlik G, Sparks CJ and Fischer K (1994) Amsterdam, London, New York, Tokyo: North-Holland.

Noyan IC and Cohen JB (1987) *Residual Stress*. New York, Berlin, Heidelberg, London, Paris, Tokyo: Springer-Verlag.

Tanner BK (1976). *X-ray Diffraction Topography*. Oxford: Pergamon Press.

# Matrix Isolation Studies By IR and Raman Spectroscopies

**Lester Andrews**, University of Virginia, Charlottesville, VA, USA

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

Matrix isolation is a technique for maintaining molecules in an inert medium at very low temperature for spectroscopic study. This method is particularly well suited for preserving reactive species in a solid environment. Elusive molecular fragments, such as free radicals that may be important intermediates for chemical transformations used in industrial reactions, molecules that are in equilibrium with solids at very high temperatures, weak molecular complexes that may be stable at low temperatures, and molecular ions that are produced in plasma discharges or by high-energy radiation can all be observed and characterized using infrared absorption and laser-excitation spectroscopies. The matrix isolation technique enables spectroscopic data to be obtained for reactive molecular fragments, many of which cannot be studied in the gas phase.

## Experimental apparatus

The experimental apparatus for matrix isolation experiments is designed with the methods of generating the molecular transient and performing the spectroscopy in mind. **Figure 1** is a schematic diagram of the laser-ablation matrix isolation apparatus for infrared absorption spectroscopy. Caesium iodide windows are typically employed. The rotatable cold window is cooled to 4–20 K by closed-cycle refrigeration or liquid helium. The matrix sample is introduced through the spray-on line at rates of 1–5 millimoles per hour; argon is the most widely used matrix gas, although neon, krypton, xenon and nitrogen are also used. In **Figure 1**, the Nd–YAG fundamental at 1064 nm in 5–50 mJ pulses of 10 ns duration is focused (spot size approximately 0.1 mm) onto a rotating metal target. Ablated metal atoms intersect the gas sample during co-deposition on the cold CsI window where collisions and reactions occur. The reactive species can be generated in a number of other ways: mercury arc photolysis of a trapped precursor molecule through the quartz window, evaporation from a Knudsen cell heated inside the chamber, chemical reaction of atoms evaporated from the Knudsen cell with molecules deposited through the spray-on line, and vacuum-ultraviolet photolysis of molecules deposited from the spray-on line by radiation from discharge-excited atoms. For laser excitation studies, the sample is deposited on a tilted copper wedge which is grazed by the laser beam, and light emitted or scattered at approximately 90° is examined by a spectrograph.

## Free radicals, anions and cations

The first free radical stabilized in sufficient concentration for matrix infrared detection was formyl (HCO). Hydrogen iodide (HI) was deposited in a

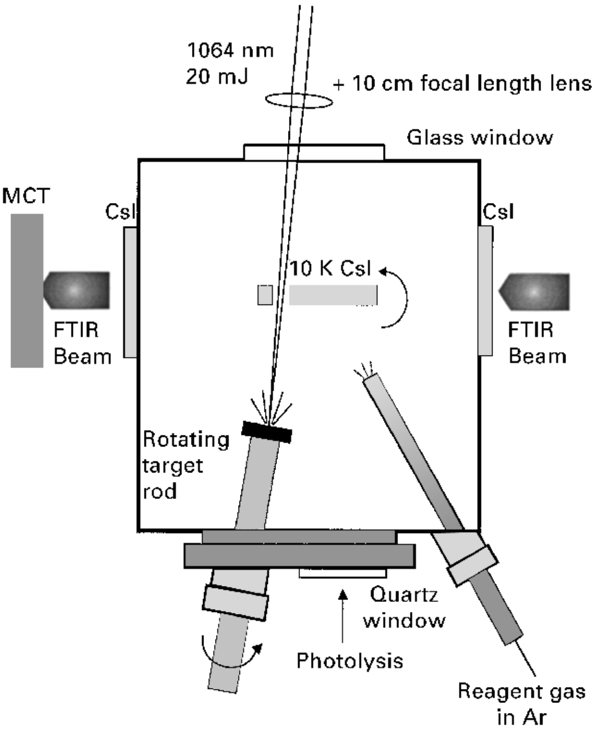



**Figure 1** Vacuum chamber for infrared matrix isolation studies using laser ablation.

carbon monoxide (CO) matrix and photolysed with a mercury arc; hydrogen atoms produced by dissociation of HI reacted with CO in the cold solid to produce HCO. The infrared spectrum of HCO provided vibrational fundamentals and information about the chemical bonding in this reactive species. Free radicals have been produced in matrices using a variety of techniques such as ultraviolet photolysis (Eqn [1]); vacuum ultraviolet photolysis (Eqn [2]); lithium atom (Eqn [3]); hydrogen atom (Eqn [4]); and lithium atom (and electron) (reaction [5]).

$$OF_2 + \text{ultraviolet light} \rightarrow OF^{\bullet} + F^{\bullet} \quad [1]$$

$$CHF_3 + \text{vacuum ultraviolet light} \rightarrow CF_3{}^{\bullet} + H^{\bullet} \quad [2]$$

$$CCl_4 + Li^{\bullet} \rightarrow CCl_3{}^{\bullet} + LiCl \quad [3]$$

$$H^{\bullet} + O_2 \rightarrow HO_2{}^{\bullet} \quad [4]$$

$$Li^{\bullet} + O_2 \rightarrow (Li^+)(O_2{}^-) \quad [5]$$

The first molecular ionic species characterized in matrices, $Li^+O_2{}^-$, was formed by the co-condensation reaction of lithium (Li) atoms and oxygen ($O_2$) molecules at high dilution in argon. The infrared spectrum exhibited a weak $(O \leftrightarrow O)^-$ stretching vibration and two strong $Li^+ \leftrightarrow O_2{}^-$ stretching vibrations, as shown at the top of **Figure 2** for the $^7Li$ and $^{16,18}O_2$ isotopic reaction. The 1–2–1 relative-intensity oxygen isotopic triplets in this experiment showed that the oxygen atomic positions in the molecule are equivalent and indicated an isosceles triangular structure. The ionic model for the bonding in $Li^+O_2{}^-$ was confirmed by contrasting intensities between the laser-Raman and infrared absorption spectra shown in **Figure 2**. Infrared intensities depend upon a change in molecular dipole moment and Raman intensities depend on a change in molecular polarizability during the vibration. Hence, modes between ions ($Li^+ \leftrightarrow O_2{}^-$) will be strong in the infrared and modes within an ion ($O \leftrightarrow C^-$) will be strong in the Raman spectrum. **Table 1** lists the observed frequencies. The contrasting frequencies when $Li^+$ is replaced by $Cs^+$ verify the ionic model for the bonding in the $(M^+)(O_2{}^-)$ molecules.

High-temperature molecules like lithium fluoride (LiF) can be trapped in matrices by evaporating the molecule from the crystalline solid in a Knudsen cell at high temperature or by reacting lithium atoms with fluorine molecules during condensation. The latter method has been used to synthesize the calcium oxide (CaO) molecule from the calcium atom–ozone reaction, which gives the calcium ozonide ion pair $Ca^+O_3{}^-$ and CaO diatomic products.

Continuous exposure of a condensing sample to argon-resonance radiation during sample condensation has been used to produce molecular cations for spectroscopic study. In the case of fluoroform

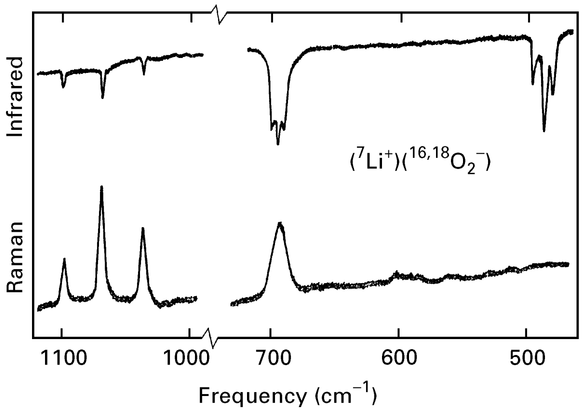



**Figure 2** Infrared and Raman spectra of lithium superoxide, $Li^+O_2{}^-$, using lithium-7 and 30% $^{18}O_2$, 50% $^{16}O^{18}O$, and 20% $^{16}O_2$ at concentrations of $Ar/O_2 = 100$. The Raman spectrum was recorded using 200 mW of 488-nm excitation at the sample.

**Table 1** Fundamental frequencies ($cm^{-1}$) assigned to the $\nu_1$ intraionic and $\nu_2$ and $\nu_3$ interionic modes of the $C_{2v}$ alkali metal superoxide molecules in solid argon

| *Molecule* | $\nu_1$ | $\nu_2$ | $\nu_3$ |
|---|---|---|---|
| $^{6}LiO_2$ | 1097.4 | 743.8 | 507.3 |
| $^{7}LiO_2$ | 1096.9 | 698.8 | 492.4 |
| $NaO_2$ | 1094 | 390.7 | 332.8 |
| $KO_2$ | 1108 | 307.5 | – |
| $RbO_2$ | 1111.3 | 255.0 | 292.5 |
| $CsO_2$ | 1115.6 | 236.5 | 268.6 |

($CHF_3$), photolysis produced the $CF_3^{\bullet}$ radical, which may be photoionized by a second 11.6-eV photon to give $CF_3^+$. The infrared spectrum of $CF_3^+$ revealed a very high C–F vibrational fundamental, which indicates substantial pi bonding in the planar carbocation.

## Complexes

Molecular complexes involving hydrogen fluoride (HF) serve as useful prototypes for the understanding of the important phenomenon of hydrogen bonding. An interesting chemical case is ammonia ($NH_3$) and HF, which on the macroscopic scale produce the salt ammonium fluoride but on the microscopic scale give the $NH_3$–HF complex. The infrared spectrum of this complex reveals vibrations for $NH_3$ and HF perturbed by their association in the complex. Fourier-transformed infrared spectroscopy is particularly advantageous in these studies because of the high vibrational frequencies for HF species and low sample transmission in this region. The ammonia symmetric bending motion is considerably blueshifted, and the hydrogen fluoride stretching fundamental is markedly redshifted. These shifts attest to a strong intermolecular interaction within the complex.

$$NH_3 + HF \rightarrow H_3N\text{- -}HF \qquad [6]$$

Ozone ($O_3$) is a very important molecule in the upper atmosphere, as it absorbs harmful ultraviolet radiation. In the laboratory, this ultraviolet dissociation of ozone provides oxygen atoms for chemical reactions, but in complexes red light can induce an O-atom transfer. Photochemical reactions of elemental phosphorus ($P_4$), an extremely reactive molecule well suited for matrix isolation studies, and ozone produce a new low oxide of phosphorus, $P_4O$, which is involved as a reactive intermediate in the striking of a match. The infrared spectrum of $P_4O$ shows a strong terminal PO bond stretching fundamental and characterizes the $P_4O$ structure as tetrahedral $P_4$ with an O atom cap. Further photochemical reactions of phosphorus and oxygen produce and trap the reactive molecule $^{\bullet}PO_2$, isoelectronic with the pollutant molecule $^{\bullet}NO_2$, which can also be formed by laser evaporation of red phosphorus followed by reaction with a stream of oxygen gas.

$$P_4 + O_3 \rightarrow (P_4)(O_3)\,\text{complex} \xrightarrow{h\nu} P_4O + O_2 \qquad [7]$$

## High temperature molecules

Titanium dioxide is a white solid used as a pigment in paints, but at 2200–2400 K some TiO and OTiO molecules exist in equilibrium with the solid. The OTiO molecule is bent (113±3°) with two strong double bonds. The TiO and OTiO molecules have been prepared by reacting laser-ablated Ti atoms with $O_2$ and trapped in solid argon for infrared spectroscopic analysis. **Figure 3** shows the spectrum of a sample prepared by reacting laser-ablated Ti with $O_2$ (0.5%) in excess argon followed by condensation at 7 K. The weak 1039.5 $cm^{-1}$ band is due to ozone which is made by combination of $O_2$ and O atoms produced in the ablation process. The weak 953.7 $cm^{-1}$ band is due to $O_4^-$ formed by combination of $O_2$ and $O_2^-$, the latter from capture of electrons by dioxygen. The strongest bands at 946.7 and 917.0 ($TiO_2$) and two weaker bands at 1012.8 ($TiO^+$) and 987.8 $cm^{-1}$ (TiO) are important products. The reaction was repeated for $^{18}O_2$ and mixtures of $^{16}O_2$, $^{16}O^{18}O$, $^{18}O_2$. **Table 2** lists the isotopic for frequencies for these molecules. The 16/18 isotopic ratio is characteristic of the normal vibrational mode. Note that the 917.0 $cm^{-1}$ band is strong enough to exhibit satellite adsorptions at 923.1, 919.9, 914.1 and 911.3 $cm^{-1}$ that are due to the minor $^{46}Ti$, $^{47}Ti$, $^{49}Ti$ and $^{50}Ti$ isotopes in natural abundance around the major $^{48}TiO_2$ isotopic band.

The 946.7 and 917.0 $cm^{-1}$ bands form triplet patterns in the statistical mixed isotopic experiment, which verifies the participation of two equivalent oxygen atoms, whereas the 1012.9 and 987.9 $cm^{-1}$ bands form doublets, which shows that a single oxygen atom is involved.

The 946.7 and 917.0 $cm^{-1}$ bands are due to the symmetric ($\nu_1$) and antisymmetric ($\nu_3$) Ti–O stretching fundamentals of OTiO. Titanium isotopic substitution provides basis for calculation of a lower limit (111±3°) to the OTiO valence angle whereas oxygen isotopic replacement gives a 115±3° upper limit to this angle owing to different anharmonicities. The true angle (113±3°) is the median of these limits.

The 16/18 oxygen isotopic ratios for the 987.8 and 1012.9 $cm^{-1}$ bands are different from values for the

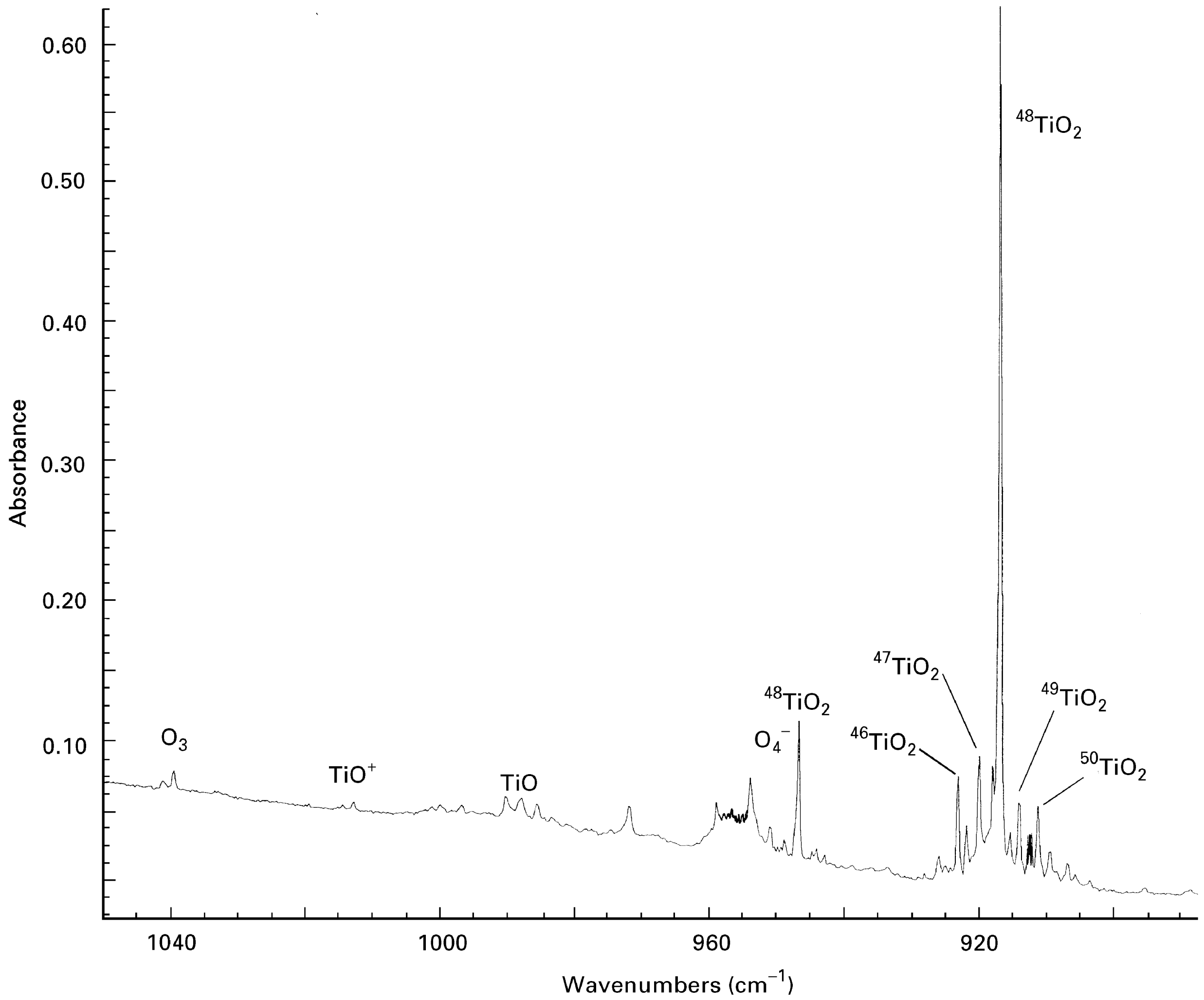



**Figure 3** Infrared spectra for laser-ablated Ti atoms and electrons co-deposited with $O_2$ (0.5%) in excess argon for 45 min at 7 K.

above frequencies. These ratios (1.0442 and 1.0443) approach the harmonic value (1.0446) for diatomic TiO. The 987.8 cm$^{-1}$ band is due to TiO in solid argon, which is redshifted from the 1000.0 cm$^{-1}$ gas phase value, and the 1012.9 cm$^{-1}$ band is due to $TiO^+$, which has not yet been observed by gas phase optical spectroscopy.

The use of pulsed-laser ablation to produce new chemical species for spectroscopic study is further

**Table 2** Major infrared absorptions (cm$^{-1}$) observed for laser-ablated titanium and dioxygen reaction products isolated in solid argon

| $^{16}O_2$ | $^{18}O_2$ | *Ratio* $^{16}O/^{18}O$ | Identification |
|---|---|---|---|
| 1039.6 | 982.3 | 1.0583 | $O_3(\nu_3)$ |
| 1012.9 | 969.9 | 1.0443 | $TiO^+$ |
| 987.8 | 946.0 | 1.0442 | TiO |
| 953.7 | 901.6 | 1.0578 | $O_4^-$ |
| 946.9 | 904.5 | 1.0469 | $^{48}TiO_2(\nu_1)$ |
| 917.1 | 881.7 | 1.0401 | $^{48}TiO_2(\nu_3)$ |

illustrated for the vanadium atom reaction with $N_2$. Laser-ablated metal atoms are sufficiently energetic to dissociate molecular $N_2$ into N atoms for reaction to form metal nitrides. A sample of 2% $^{14}N_2$ + 2% $^{15}N_2$ in argon was reacted with laser-ablated V atoms, and spectra from this experiment are shown in **Figure 4**. Four very weak bands ($A$ = absorbance = 0.002 to 0.001) were observed at 1026.2, 1010.3, 1014.4 and 1010.6 cm$^{-1}$ from nitrogen-14 with an identical set at 999.4, 993.6, 987.9 and 984.2 cm$^{-1}$ from nitrogen-15 on sample co-deposition at 10 K [**Figure 4A**]. Annealing to 25 and 30 K to allow diffusion and further reaction of trapped species slightly decreased the first band and increased the second, third and fourth bands in each set, and produced new weak bands at 997.8 and 971.4 cm$^{-1}$ [**Figure 4B, C**]. Broadband photolysis decreased the lowest two bands and increased the second band in each set [**Figure 4D**]. Further annealing to 40 K decreased the first three bands, increased a

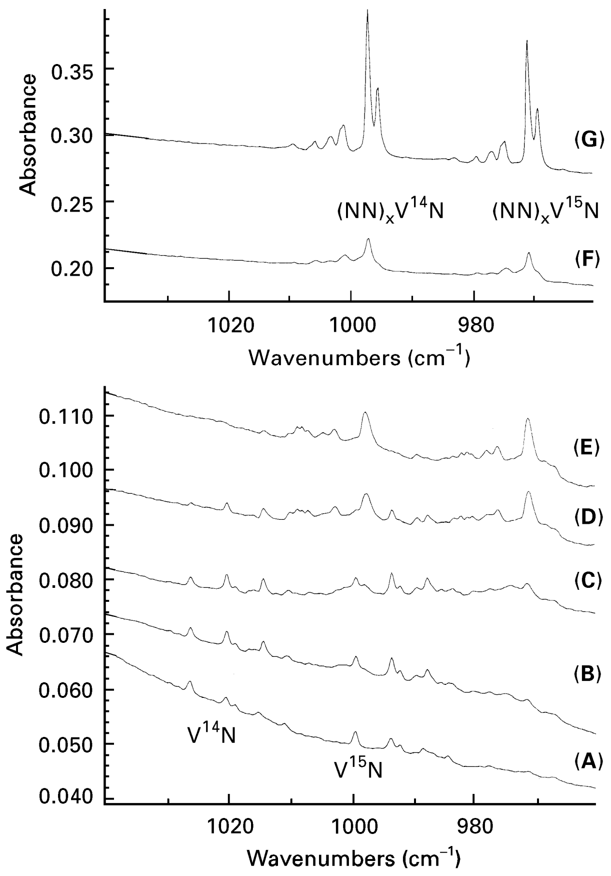



**Figure 4** Infrared spectra in the 1040–960 cm$^{-1}$ region for laser-ablated V atoms co-deposited with nitrogen. (A) 2% $^{14}N_2$ + 2% $^{15}N_2$ in argon co-deposited at 7 K for 1 h, (B) after annealing to 25 K, (C) after annealing to 30 K, (D) after annealing to 40 K, (E) after annealing to 43 K, (F) pure $^{14}N_2$ + $^{15}N_2$ co-deposited at 10–11 K for 1 h, and (G) after annealing to 30 K.

1002.8 cm$^{-1}$ band (and 976.5 cm$^{-1}$ counterpart) and increased the final 997.8 and 971.4 cm$^{-1}$ bands [**Figure 4E**]. A final annealing to 43 K destroyed the first two bands and increased the last two bands in each set [**Figure 4F**]. One experiment was done with a pure dinitrogen $^{14}N_2$ + $^{15}N_2$ mixture and the spectra are shown at the top of **Figure 4** for deposition at 10 K and annealing to 30 K; note growth of the strong 997.0 and 970.6 cm$^{-1}$ bands and satellites.

The VN example illustrates how the isolated VN molecule becomes successively ligated by dinitrogen to form $(NN)_x$VN on annealing in the solid argon matrix or on deposition in pure dinitrogen matrix. The sharp weak new band at 1026.2 cm$^{-1}$ in solid argon decreases on stepwise annealing while sharp bands increase at 1020.3, 1014.4 and ultimately 997.8 cm$^{-1}$. These bands exhibit sharp nitrogen-15 counterparts at 999.4, 993.6, 987.9 and 971.4 cm$^{-1}$, which define nitrogen 14/15 frequency ratios 1.0268, 1.0269, 1.0268 and 1.0272, respectively. These values are in excellent agreement with the harmonic VN diatomic ratio (1.0272) and attest to the pure V–N stretching character of these adsorptions. These bands showed no intermediate components with discharged (statistical) mixed isotopic dinitrogen, so a single nitrogen atom is involved in these vibrations. In solid dinitrogen, a sharp isotopic doublet at 997.0 and 970.6 cm$^{-1}$ also increased on annealing and revealed the diatomic ratio (1.0272) and is only 0.8 cm$^{-1}$ lower than the dominant nitrogen isotopic doublet surviving past 40 K annealing in solid argon (**Figure 4**).

The sharp 1026.2 cm$^{-1}$ band is due to the VN diatomic molecule isolated in argon. The evolution of bands in **Figure 4** from 1026.2 cm$^{-1}$ to 997.8 cm$^{-1}$ on annealing in solid argon and the 997.0 cm$^{-1}$ band in solid nitrogen provides a convincing picture for the attachment of dinitrogen ligands to the VN center. We note the appearance of five distinct bands for different ligated species. These are consistent with the eighteen electron rule in that these bands are due to $(NN)_x$VN with $x = 1, 2, 3, 4, 5$ where the maximum ligated species has eighteen electrons in the valence shell about vanadium.

*See also:* **IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; IR Spectrometers; IR Spectroscopy, Theory; Nonlinear Optical Properties; Raman Spectrometers.**

## Further reading

Andrews L (1984) Fourier transform infrared spectra of HF complexes in solid argon. *Journal of Physical Chemistry* **88**: 2940–2949.

Andrews L and Moskovits M (eds) (1989) *Chemistry and Physics of Matrix Isolated Species*. Amsterdam: Elsevier.

Andrews L and Smardzewski RR (1973) Argon matrix Raman spectrum of $LiO_2$. *Journal of Chemical Physics* **58**: 2258–2261.

Andrews L, Bare WD and Chertihin GV (1997) Reactions of laser ablated V, Cr and Mn atoms with nitrogen atoms and molecules. *Journal of Physical Chemistry A* **101**: 8417–8427.

Bondybey VE, Smith AM and Agreiter J (1996) New developments in matrix isolation spectroscopy. *Chemical Reviews* **96**: 2113–2134.

Chertihin GV and Andrews L (1995) Reactions of laser ablated Ti, Zr and Hf atoms with $O_2$ molecules in condensing argon. *Journal of Physical Chemistry* **99**: 6356–6366.

Zhou MF and Andrews L (1998) Matrix infrared spectra and density functional calculations of $Ni(CO)_x^-$, $x = 1$, 2, 3. *Journal of the American Chemical Society* **120**: 11499–11503.

Zhou MF and Andrews L (1999) Infrared spectra of the $C_2O_4^+$ cation and $C_2O_4^-$ anion in solid neon. *Journal of Chemical Physics* **110**: 6820–6826.

# Medical Applications of Mass Spectrometry

**Orval A Mamer**, McGill University, Montréal, Québec, Canada

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Applications |

Mass spectrometry has a breadth of application in medicine that is unequalled by any other spectroscopic technique. This is largely by virtue of having the uncommon ability to combine great sensitivity and specificity with near-perfect generality. Among the most common applications of mass spectrometry in medicine are the diagnosis and confirmation of known acquired and inherited metabolic disorders, characterization and investigation of those previously unknown, and the identification of intoxicants, whether inadvertently or deliberately administered. While this article will focus on the area of metabolic disease, the techniques described in brief here can be, and are, applied in a myriad other routine clinical situations. Some of these are measurement of serum homocysteine and methylmalonic acid as an indicator of increased risk for cardiovascular disease, and diagnosis of peptic ulcer, gastritis and gastric cancer caused by *Helicobacter pylori* by measurement of $^{13}C$-labelled carbon dioxide released by *H. pylori* from an oral dose of $^{13}C$-labelled urea. Mass spectrometry is the 'gold standard' in the clinical laboratory, and is at the heart of many reference methodologies. Serum cholesterol determinations are highly dependent on the natures of the wet chemistries used, and mass spectrometry with gas or liquid chromatographic or capillary electrophoresic inlets provide the only direct and unambiguous measurements available today. Metabolic disease investigation by mass spectrometry, then, is selected only as an example or model of how it may be applied routinely elsewhere in the clinical setting.

## Metabolic diseases

The identification of accumulating metabolites characteristic of a metabolic disease employs mass spectrometry in a qualitative manner. Compounds are identified in a fingerprint sense by computer-comparison of the mass spectra of unknowns with a library of reference spectra either purchased with the instrument or accumulated in-house from reference compounds. Mass spectrometry may also be used in a quantitative sense for the purpose of monitoring or evaluating patients, usually as part of a treatment protocol.

The most significant metabolic disorders are life threatening in the neonatal period if left untreated and are usually the result of the inherited inability to catabolize normal substrates derived from diet or from normal tissue degradation and recycling. The typical catabolic pathway can be represented by Equation [1],

$$A \xrightarrow{E_A} B \xrightarrow{E_B} C \xrightarrow{E_C} \cdot\cdot\cdot\cdot \longrightarrow N \qquad [1]$$

where A, B, ..., N are precursors, intermediates and final products, and subscripted Es are the enzymes or enzyme complexes that effect the changes associated with each metabolic step. A metabolic disease is said to occur when one (or more) of these enzyme-mediated steps fails to produce the required transformation, and the blood and tissue concentrations of the substrate or precursor for that step (or earlier steps) increase to levels that are toxic, inhibit other enzyme systems or significantly alter the pH or other characteristics of blood or tissues and adversely affect the patient's well-being. Other disorders may be the result of the inability to synthesize a required substance or the inability to transport something essential across a membrane. Identification of accumulating or missing substrates is therefore crucial to understanding the nature of the disorder and the eventual identification of the faulty enzyme or enzyme system.

The list of inherited metabolic diseases continues to grow; at this writing, there are 517 distinct disorders that have been at least partially characterized, with the sequences of the responsible defective proteins known for many of them. Most of these disorders are very rare; the incidence of a given disorder may vary from 1 live birth in 500 000 to 1 in 750, and is frequently dependent upon the degree to which an isolated population is inbred. A paediatrician may spend a lifetime in practice and not encounter any but the most common of these rare diseases.

Diagnosis in the very early postnatal period is critically important if the newborn is to survive and not suffer toxic accumulations of metabolites that frequently lead to retardation of physical and intellectual development. Through correct early diagnosis

of many diseases that result from errors of catabolism of dietary components, such as amino acids, diets restricted in these precursors may be instituted to prevent these accumulations.

Diagnosis *in utero* is often achieved by the mass spectral analysis of a small sample of the amniotic fluid for elevations of metabolites excreted by a possibly affected fetus. This invasive procedure is usually only applied when there is a reason to suspect that a fetus may be affected, such as a defect occurring in a prior birth. This provides a basis for an informed decision to be made either to terminate the pregnancy or to prepare for supportive intervention at birth.

The failure of an enzyme may be partial or complete, permanent or transitory. A genetic mutation that encodes for an enzyme that is partially or completely inactive will result in a permanent deficit. Some enzyme deficits that are due to mutations that lead to inefficient binding of a cofactor can be stimulated to a useful capacity by administration of large doses of the cofactor on a life-long basis.

## Gas chromatography–mass spectrometry

The use of mass spectrometry in medicine was greatly facilitated by the development of coupled gas chromatography–mass spectrometry (GC-MS). The first commercially successful GC-MS, the LKB-9000, became available in the mid 1960s, and it is estimated that fully half of the inherited metabolic disorders initially described in the 1960s and 1970s were discovered and investigated with various versions of this instrument. Direct coupling of fused-silica capillary columns to a proliferation of small and inexpensive bench-top quadrupole and ion trap instruments has now made practical the screening for metabolic disease in every live birth in developed countries.

An example of an inherited disorder that is the result of a permanent enzyme deficit is methylmalonic acidaemia, in which methylmalonic acid (MMA), derived in large part from valine and isoleucine, accumulates. In one of the common forms of this disorder, methylmalonyl-coenzyme A mutase, the enzyme responsible for conversion of methylmalonyl-coenzyme A (CoA) to succinyl-CoA is partially or completely inactive, and the clinical result is a severe and life-threatening keto acidosis, accompanied by high concentrations of MMA in the blood and in tissues and the urinary excretion of up to several grams of MMA per day. In other forms, mutase activity can be increased by administration of large doses of vitamin $B_{12}$ or one of its related cobalamins, which are necessary cofactors for the mutase. For diagnostic purposes, a sample of urine is collected from the patient and an extract is prepared containing the carboxylic acids present in the urine. The dried extract is treated with a chemical reagent that produces the trimethylsilyl (TMS) esters of the carboxylic acids, and this mixture is then analysed by capillary GC-MS. **Figure 1** is an organic acid profile obtained in this manner, and is an example of the use of GC-MS in the diagnosis of a mutase deficiency. Clearly evident is a very large GC peak due to

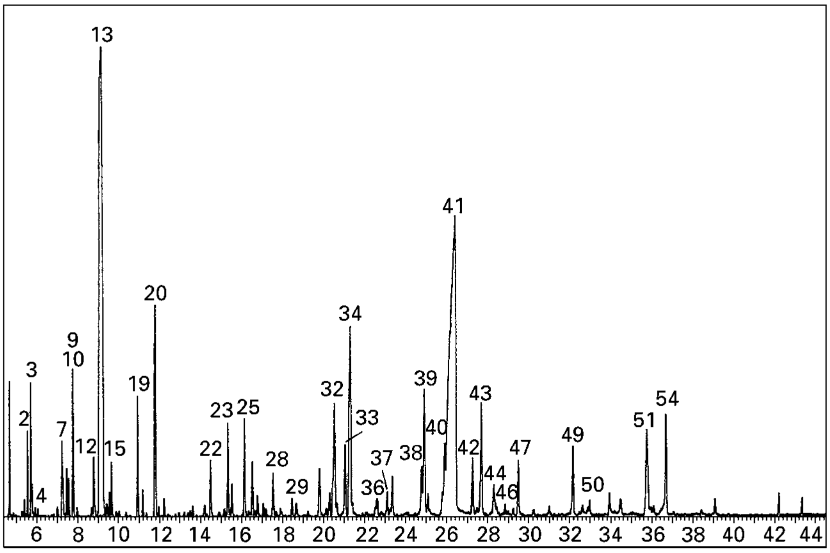



**Figure 1** The electron ionization total ion current chromatogram for the trimethylsilyl derivatives of the organic acids extracted from the urine of a patient with an inherited error of methylmalonyl-CoA mutase. The major acidic component is methylmalonic acid. The annotated peaks are identified in **Table 1**.

bis(trimethylsilyl) methylmalonate. Other acids present and identified are normally found in human urine. **Table 1** lists the identities of the eluting peaks in this figure and in the other profiles presented below.

Integrated ion-current peak areas can be related to the quantities of metabolites present in the extract. While relative sensitivities for ionization must be known and taken into account for precise determinations, frequently one is forced to assume equality for all metabolites identified when these are unknown. The result is usually accurate to within a factor of 2 or 3 and is sufficient for screening of large numbers of samples.

When more precise measurements are required, stable-isotope dilution techniques are employed. In this approach to quantitation, a measured amount of a stable-isotope-labelled analogue of the compound of interest (internal standard) is added to the fluid to be analysed. The sample is then prepared for analysis and the mass spectrometer is used to report the ratio of the unlabelled to labelled analogues. With relatively few and minor corrections applied, the concentration of the unlabelled metabolite is easily determined. The technique is insensitive to losses in sample separations and variable derivatization yields as these will affect both analogues in the same proportion. As an example, **Figure 2** illustrates the use of this technique in the accurate determination of the concentration of MMA in human serum.

The internal standard is 20 μg of [*Me*-$^2H_3$]MMA, which was added to 1 mL of serum. The acidic extract was converted to the TMS derivatives and analysed in a GC-MS mode termed 'selected-ion monitoring', in which the quadrupole analyser is stepped discontinuously between selected ions and reports only their intensities. In this case, the ions selected has masses 218.2 and 221.2 Da, which are moderately intense, characteristic and distinguishing fragment ions in the spectra of the TMS derivatives of MMA and [*Me*-$^2H_3$]MMA, respectively. The data are obtained as plots of the intensities of these two ions as functions of GC retention time, and resemble independent gas chromatograms for these two fragment ions. The peak areas are integrated and a correction of the 218.2 area is made for isotopic impurity in the internal standard. The 221.2 area is also corrected for natural-abundance heavy isotope inclusion in the light ion that is present as an $M$+3 signal at 221.2. From these calculations one may then conclude that the serum sample was 20.06 μM in MMA, about 10-fold greater than the normal upper limit.

Another inherited error of valine, isoleucine and leucine metabolism is branched-chain α-keto aciduria, also known as maple syrup urine disease because the odour of urine of these patients resembles that of maple syrup. Branched-chain $\alpha$-keto acid dehydrogenase, the enzyme complex that is used to oxidatively decarboxylate 2-ketoisovaleric, 2-keto-3-methylvaleric and 2-ketoisocaproic acids to the corresponding branched-chain CoA esters is defective, and very high concentrations of these keto acids accumulate in the blood and urine of these

**Table 1** Identities of acid peaks annotated in **Figures 1, 3, 4 and 5**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | Phenol | 16 | 2-Hydroxyisocaproic | 30 | 3-Phenyllactic | 44 | Vanillylmandelic |
| 2 | Lactic | 17 | (2*S*)-Hydroxy-3(*R*)- | 31 | Pimelic | 45 | Sebacic |
| 3 | 2-Hydroxyisobutyric | | methylvaleric | 32 | Hippuric (secondary | 46 | 4-Hydroxyphenyllactic |
| 4 | Glycolic | 18 | (2*S*)-Hydroxy-(3*S*)- | | derivative) | 47 | 3-Indoleacetic |
| 5 | Oxalic | | methylvaleric | 33 | 4-Hydroxybenzoic | 48 | 4-Hydroxyphenyl- |
| 6 | Glyoxylic oxime | 19 | Ethylmalonic | 34 | 4-Hydroxyphenylacetic | | pyruvic oxime |
| 7 | 4-Cresol | 20 | Succinic | 35 | 4-Hydroxybenzaldoxime | 49 | Palmitic |
| 8 | Pyruvic oxime | 21 | 4-Hydroxybenzaldehyde | 36 | Phthalic | 50 | 3-Hydroxysebacic |
| 9 | 3-Hydroxyisobutyric | 22 | Glutaric | 37 | Suberic | 51 | 4-Hydroxyhippuric |
| 10 | 3-Hydroxybutyric | 23 | Methylmalonic (3-TMS) | 38 | Vanillic | 52 | *N*-Acetyltyrosine |
| 11 | 2-Hydroxyisovaleric | 24 | 2-Methoxybenzoic (IS)[b] | 39 | Homovanillic | 53 | 5-Hydroxyindoleacetic |
| 12 | 2-Methyl-3-hydroxybutyric | 25 | Capric (IS)[b] | 40 | Azelaic | 54 | Stearic |
| 13 | Methylmalonic (2-TMS)[a] | 26 | 3-Hydroxyoctanoic | 41 | Hippuric | 55 | *N*-Acetyltryptophan |
| 14 | Benzoic | 27 | Mandelic | 42 | Citric | | |
| 15 | 2-Ethyl-3-hydroxy- | 28 | Adipic | 43 | 3-(3-Hydroxyphenyl)- | | |
| | propionic | 29 | 3-Methyladipic | | 3-hydroxypropionic | | |

[a] TMS=trimethylsilyl.
[b] IS=internal standard.

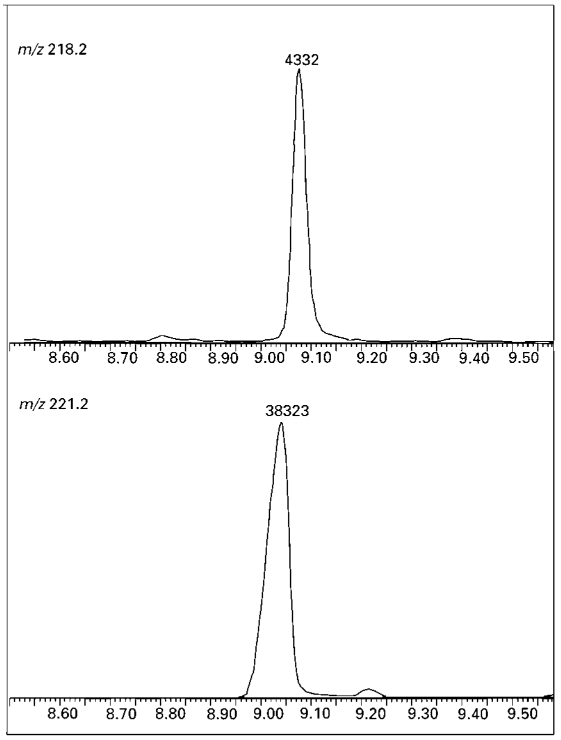
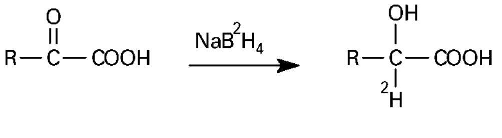



**Figure 2** Selected ion monitoring analysis of serum for methylmalonic acid. The *m/z* 218.2 ion represents the $[M-CO_2]^{\bullet+}$ McLafferty fragment of the TMS derivative of endogenous unlabelled methylmalonic acid. The *m/z* 221.2 ion is the fragment derived by a similar process from the internal standard $[Me\text{-}^2H_3]$ methylmalonic acid. The areas integrated for the ions are related to the relative concentration of the derivatives (see text). The deuterium-labelled analogue has a slightly shorter retention time than the unlabelled analogue, a phenomenon perhaps unexpected but commonly observed in these circumstances.

patients. They also have in their fluids large elevations of the corresponding 2-hydroxy acids made by reduction of the keto acids, probably by lactate dehydrogenase. In sample preparation, addition of sodium borohydride to the urine will reduce the three 2-keto acids to the corresponding three 2-hydroxy acids and eliminate (*E*) and (*Z*) isomerism in the TMS derivatives of the enols of the 2-keto acids, thereby simplifying the gas chromatogram. If sodium borodeuteride is added instead, the 2-hydroxy acids that are produced will bear single deuterium substitutions on carbon-2 (Eqn [2]) and the labelled/unlabelled ratios may be determined easily for each of the 2-hydroxy acids.



[2]

These three ratios then represent the ratios of the 2-keto acids originally in the sample to their corresponding 2-hydroxy acids. **Figure 3** illustrates the urinary organic acid profile of one of these patients after reduction of the urinary 2-keto acids by sodium borodeuteride to the labelled 2-hydroxy acids. The keto/hydroxy ratios are clinically significant as they are related to the $NAD^+$/NADH status of these patients. ($NAD^+$/NADH = the oxidized/reduced forms of nicotinamide-adenine dinucleotide.)

A transitory enzyme deficit is one that arises commonly as the result of the ingestion of a toxic material that is or can be metabolized to a suicide substrate for that enzyme, which if depleted must be replaced by *de novo* synthesis. Another common temporary deficit is the result of an immature enzyme system in the newborn that spontaneously resolves itself within a few days.

As an example of the former, Jamaican vomiting sickness presents with severe hypoglycaemia and acidosis that resembles known types of inherited acyl-CoA dehydrogenase deficiencies. The disorder is precipitated by consuming the unripe fruit of the akee plant that grows in the Caribbean area. The protoxin is an amino acid, hypoglycine A ($\alpha$-amino-$\beta$-(2-methylenecyclopropyl)propionic acid), which is metabolically converted to methylenecyclopropylacetic acid that irreversibly binds covalently to the flavin moiety of several acyl-CoA dehydrogenases and acts as a suicide substrate to permanently disable their active sites. Treatment is usually limited to supportive care during the period in which the inactivated enzyme system is replaced by new synthesis. The organic acid profile is dominated by very large concentrations of butyric, isovaleric and 2-methylbutyric acids, whose CoA esters require active dehydrogenases for further catabolism to crotonyl-CoA, 3-methylcrotonyl-CoA and 2-methylcrotonyl-CoA esters. Often these acids and other short-chain acids are also esterified to carnitine and excreted in the urine in easily detected amounts.

An example of a transient deficit that is the result of an immature enzyme system is a disorder in which the identification of increased concentrations of tyrosine and 4-hydroxyphenylpyruvic acid (4-HPPA) in urine of a premature newborn would suggest reduced activity of 4-HPPA oxidase. Administration of pharmacological doses of ascorbic acid, the cofactor required for activity of this enzyme, may overcome the temporary oxidase deficit and stimulate catabolism of the accumulating 4-HPPA. **Figure 4** is an illustration of this disorder, which is termed transient neonatal tyrosinaemia.

In addition to increased excretion of 4-HPPA, present here as the oxime, elevations of

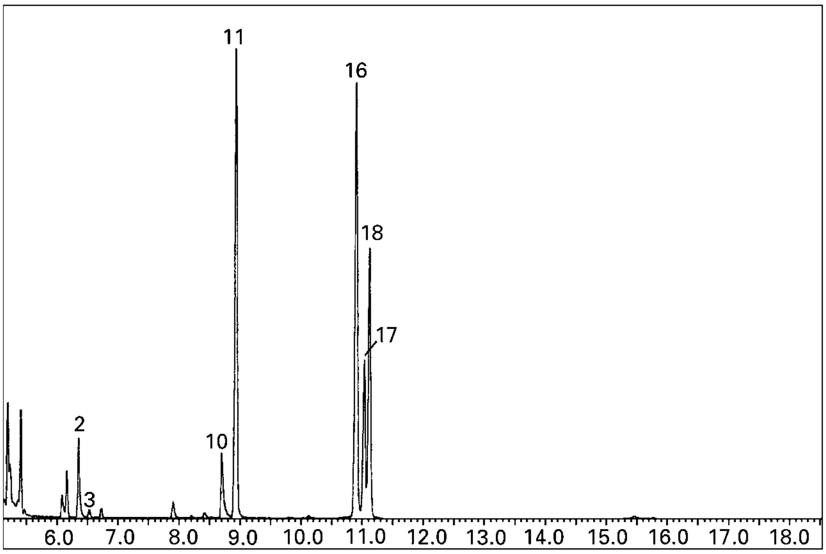



**Figure 3** The total ion current chromatogram obtained for the organic acids isolated from the sodium borodeuteride-treated urine of a patient with maple syrup urine disease. The ratios of the $^2$H-labelled to unlabelled analogues for 2-hydroxyisovaleric acid, 2-hydroxyisocaproic acid and the two 2-hydroxy-3-methylvaleric acid diastereomers are respectively 0.24, 34.5, 30.0 and 3.70.

4-hydroxyphenyllactic and 4-hydroxyphenylacetic acids are noted, and these are metabolites that are not important in normal catabolism and are derived from 4-HPPA by other enzymes. The oximes are made by addition of hydroxylamine hydrochloride to the urine during sample preparation for the purpose of detecting succinylacetone (see below)

Tyrosinaemia has an inherited form that is permanent and is due to inactive fumarylacetoacetate (FAA) hydrolase. As a result, FAA accumulates and is enzymatically converted into succinylacetoacetate (SAA), very low levels of which in turn severely inhibit 4-HPPA-oxidase. Many of the same accumulations seen in the transient neonatal form are measured as a result (**Figure 5**).

*N*-Acetyltyrosine is often noted in these organic acid profiles obtained for the hereditary form of tyrosinaemia, the result of *N*-acetylation occurring when blood levels of tyrosine rise to very high values. As a further complication, succinylacetone (SA), a

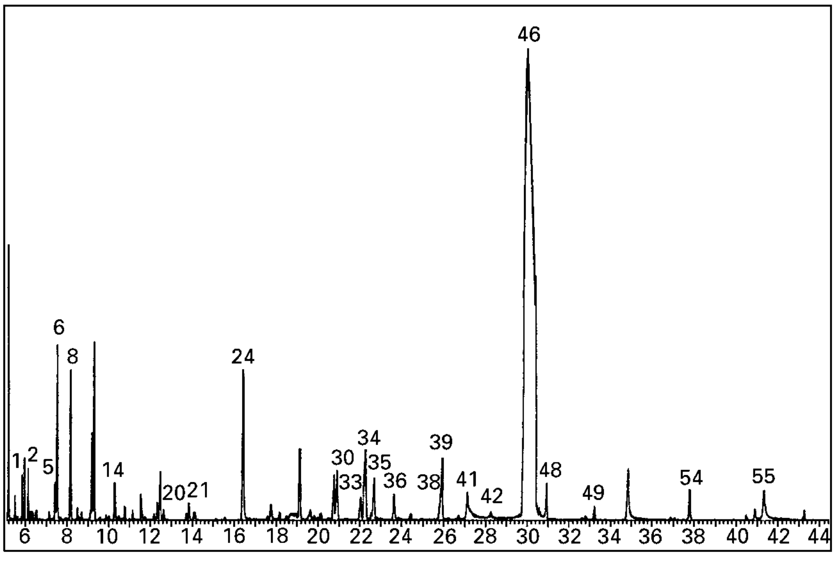



**Figure 4** The urinary organic acids excreted by a premature infant with immature 4-hydroxyphenylpyruvic acid oxidase. This disorder, also known as transient neonatal tyrosinaemia, is characterized by excretion of elevated amounts of 4-hydroxyphenylacetic, 4-hydroxyphenyllactic and 4-hydroxyphenylpyruvic acids. The urine was pretreated with hydroxylamine hydrochloride, which converts the keto acids into the respective oximes.

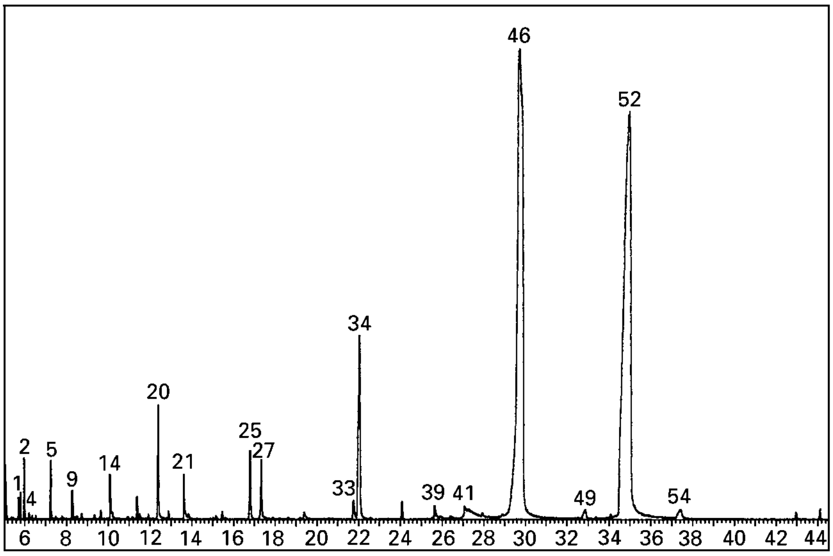


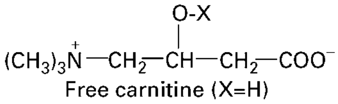

**Figure 5** The organic acid profile obtained for a patient with inherited tyrosinaemia. Very large concentrations of 4-hydroxyphenyllactic acid and *N*-acetyltyrosine are in evidence.

spontaneous decomposition product of SAA, markedly inhibits porphobilinogen synthetase and other enzymes, which leads to the large number of clinical presentations seen in this disorder. To distinguish the transient and inherited forms, one needs only to measure the serum concentration of SA. SA is unstable to the usual sample preparation methods and, to avoid losses, the oxime is made by addition of hydroxylamine hydrochloride to the serum or urine sample. Since all of the hydrogen atoms in SA are easily exchanged with aqueous protons, no deuterium-labelled analogue of SA resistant to back-exchange can be made for use as an internal standard, and synthesis of a suitably $^{13}C$-substituted analogue would be very difficult. This problem can be sidestepped by using 2-methoxybenzoic acid as the internal standard, as this does not occur in human metabolism and it elutes without interference by an endogenous metabolite. Ions selected for monitoring are *m/z* 209.1 and 212.1, the $[M-CH_3]^+$ fragments of the TMS derivatives of 2-methoxybenzoic acid and the positional isomers 3-[5-(3-methylisoxazolyl)] propionic acid and 3-[3-(5-methylisoxazolyl)] propionic acid. The latter two compounds are produced in the oximation of SA, and while they are separable on gas chromatography they are summed for quantitation. **Figure 6** illustrates an example of such an analysis of a urine sample.

Hence, this example represents the use of an unrelated compound as the internal standard, 2-methoxybenzoic acid in this case. It is necessary in these circumstances to carefully prepare calibration curves with samples of known composition to take into account the relatively different extraction and ionization efficiencies and appearance potentials for the ions selected.

## Fast atom bombardment and electrospray ionization

Several diseases are known that result in elevations in tissues and fluids of various esters of carnitine and reduce the availability of free carnitine, which is normally synthesized by humans and is necessary for the transport of long-chain fatty acids into mitochondria for oxidation. In several disorders arising from acyl-CoA dehydrogenase deficiencies, the accumulation of the acyl-CoA substrate frequently sequesters coenzyme A and reduces its availability for other unrelated but important and otherwise competent pathways. Carnitine administration can displace and make available much of the coenzyme A that had been isolated, and stimulate the excretion of the accumulating acidic metabolites now esterified to carnitine. Detection of reduced levels of serum or urinary free carnitine and elevations of esterified carnitine is therefore useful for diagnosis of a variety of metabolic disorders, among them congenital inability to synthesize carnitine. In this disorder, carnitine must be supplied by a carnitine-enriched diet as it is, in effect, a vitamin.



[1]

Free carnitine (X=H)
Carnitine esters (X=acyl)

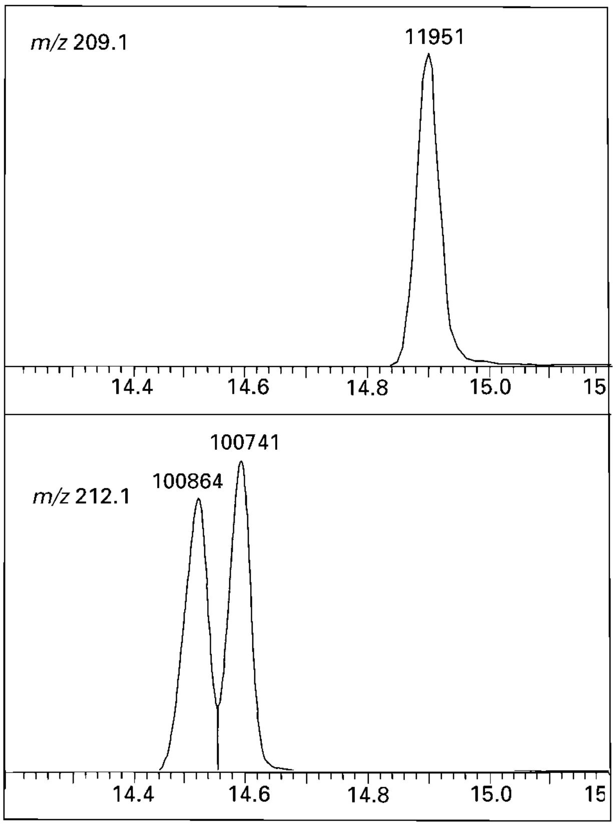



**Figure 6** Selected-ion monitoring analysis of a urine sample from a patient with the inherited form of tyrosinaemia. Succinylacetone, which is distinctly elevated here, is produced from fumarylacetoacetic acid and is measured as an indication of fumarylacetoacetate hydrolase inactivity, the precipitating cause of this form of tyrosinaemia. The ions monitored are the $[M - CH_3]^+$ fragments of the TMS derivatives of the internal standard 2-methoxybenzoic acid (*m/z* 209.1) and the two positional isomers of the oxime derivative of succinylacetone (*m/z* 212.1).

Carnitine and its esters (see [1]) cannot be introduced to the mass spectrometer by gas chromatography, as they incorporate quaternary amine functions and will decompose in the attempt. Fast atom bombardment (FAB) and electrospray ionization (ESI) can use the formal charge on the quaternary amine function to advantage, as carnitine and its esters are very easily desorbed from glycerol on the FAB probe and from aerosol sprays in ESI. **Figure 7A** illustrates the use of FAB in the quantitation of carnitine and its esters excreted in the urine of a patient presenting with a severe dicarboxylic aciduria associated with medium-chain acyl-CoA dehydrogenase (MCAD) deficiency.

D,L-[$Me_1$,$N$-$^2H_3$]carnitine, D,L-[$Ac$-$^2H_3$]acetylcarnitine, and D,L-decanoyl[$Me_1$,$N$-$^2H_3$]carnitine are added as internal standards, and the sample is prepared as the trideuteromethyl ester for FAB analysis. Esterification of the carnitine carboxylate function with perdeuteromethanol removes the zwitterionic character of carnitine and its esters, and leaves them with a full positive charge. It also raises their ion mass by 17 Da to avoid confusion between them and their nonesterified homologues. High-resolution (10 000) spectra confirm the presence of the ions of interest free of interference from unrelated ions of the same nominal masses. Quantitations are then simply a matter of applying the usual calculations associated with a stable-isotope dilution assay. The identity of the annotated peaks is given in **Table 2**, along with the concentrations determined. Large concentrations of acetyl-, propionyl-, nonanoyl-, suberyl- and sebacylcarnitines are found that reflect elevations of the free acids in the urine. As a second illustration of the technique, urine from a patient with propionic acidaemia (**Figure 7B**) is shown to have a large concentration of propionylcarnitine (**Table 2**).

ESI has made possible many recent advances in the application of mass spectrometry to diagnostic medicine. Its great sensitivity and applicability to minuscule sample sizes together with its ability to analyse aqueous solutions form the basis of its utility. Because it is also a desorption technique, it is especially useful and sensitive for compounds that incorporate formal charges or chemical groups that are easily ionized.

Trimethylaminuria is an inherited disease related to the inability to convert trimethylamine (TMA) metabolized from dietary sources into trimethylamine $N$-oxide. The result is a disorder that is not clinically acute but has the unpleasant effect of producing a body odour resembling that of rotting fish in those affected. The social consequences of this are severely debilitating. While it is relatively easy to diagnose this disorder, objective means are needed to evaluate the efficacy of treatment protocols. TMA can be quaternized with [$^2H_3$]methyl iodide and, with $^{15}$N-labelled TMA as the internal standard, an ESI method can be used to determine TMA concentrations in urine and blood. **Figure 8A** represents an example of the analysis of a normal urine for TMA. **Figure 8B** is obtained for a similar analysis of another normal urine spiked with 40 ng mL$^{-1}$ of unlabelled TMA.

A known amount of [$^{15}$N]TMA hydrochloride is added to the urine, the pH of the solution is increased to 12 in an ice bath, an ether extract is made, and an excess of [$^2H_3$]methyl iodide is added. The resulting ether solution is added to water, which is warmed to drive off the ether and excess methyl iodide, and the remaining solution is infused by syringe through the ESI probe. Selected-ion monitoring with averaging or short scan modes

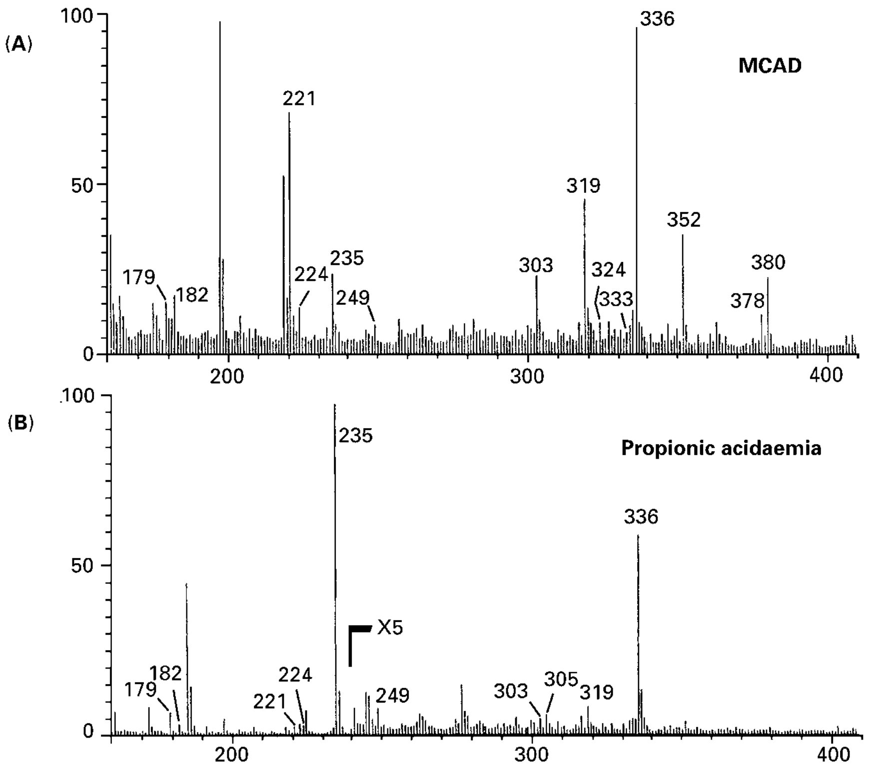



**Figure 7** Fast atom bombardment mass spectra obtained in glycerol matrix for quantitation of free carnitine and carnitine esters in the urines of a patient with medium-chain acyl-CoA dehydrogenase deficiency (A) and a patient with propionic acidaemia (B). The identities and concentrations of the annotated ions are listed in **Table 2**. The composition of the ions has been confirmed by measurements at 10 000 resolution.

**Table 2** Identities and concentrations of carnitine metabolites annotated in **Figure 7**

| *Carnitine* | *Mass* | *Ion composition* | *Concentration* (μmol/24 h) Fig. 7A | Fig. 7B |
|---|---|---|---|---|
| Free carnitine | 179.1475 | $C_8H_{15}{}^2H_3O_3N$ | 1688 | 2035 |
| Carnitine IS[a] | 182.1663 | $C_8H_{12}{}^2H_6O_3N$ | | |
| Acetylcarnitine | 221.1581 | $C_{10}H_{17}{}^2H_3O_4N$ | 1625 | 210 |
| Acetylcarnitine IS[a] | 224.1769 | $C_{10}H_{14}{}^2H_6O_4N$ | 25 | 4000 |
| Propionylcarnitine | 235.1737 | $C_{11}H_{19}{}^2H_3O_4N$ | | |
| Butyrlcarnitine | 249.1894 | $C_{12}H_{21}{}^2H_3O_4N$ | 29 | 19 |
| Octenoylcarnitine | 303.2363 | $C_{16}H_{27}{}^2H_3O_4N$ | 41 | 8 |
| Octanolylcarnitine | 305.2520 | $C_{16}H_{29}{}^2H_3O_4N$ | 563 | 11 |
| Nonanoylcarnitine | 319.2676 | $C_{17}H_{31}{}^2H_3O_4N$ | 67 | 16 |
| Adipylcarnitine | 324.2293 | $C_{15}H_{26}{}^2H_6O_6N$ | 28 | ND[b] |
| Decanoylcarnitine | 333.2833 | $C_{18}H_{33}{}^2H_3O_4N$ | 6 | ND[b] |
| Decanolylcarnitine IS[a] | 336.3021 | $C_{18}H_{30}{}^2H_6O_4N$ | | |
| Suberylcarnitine | 352.2606 | $C_{17}H_{26}{}^2H_6O_6N$ | 50 | ND[b] |
| Decenedioylcarnitine | 378.2763 | $C_{19}H_{28}{}^2H_6O_6N$ | 34 | ND[b] |
| Sebacylcarnitine | 380.2919 | $C_{19}H_{20}{}^2H_6O_6N$ | 68 | ND[b] |

[a] Internal standard.
[b] Not detected.

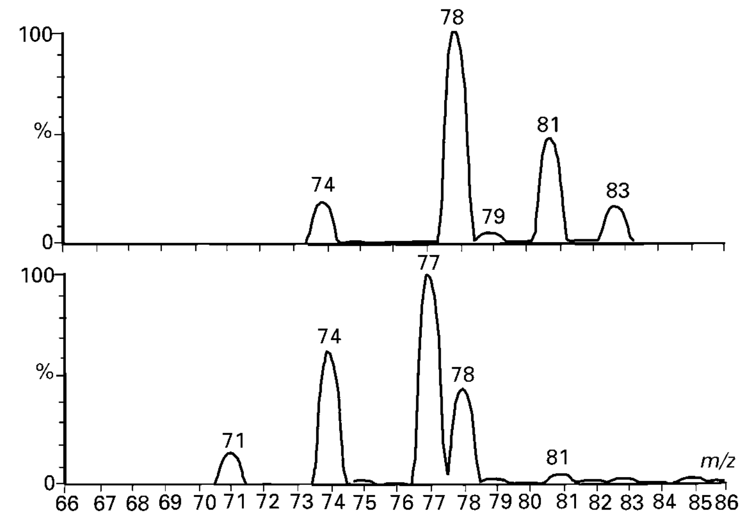



**Figure 8** Positive ion electrospray spectrum of tetramethylammonium ions produced by quaternization of trimethylamine (TMA) excreted by a patient with trimethylaminuria. When the amine fraction is quaternized with [$^2H_3$]methyl iodide, endogenous TMA is measured at *m*/*z* 77, while the [$^{15}$N]TMA internal standard is measured at *m*/*z* 78. The ion signal at *m*/*z* 74 is unlabelled tetramethylammonium produced endogenously, which would interfere with the analysis if unlabelled methyl iodide were used for quaternization.

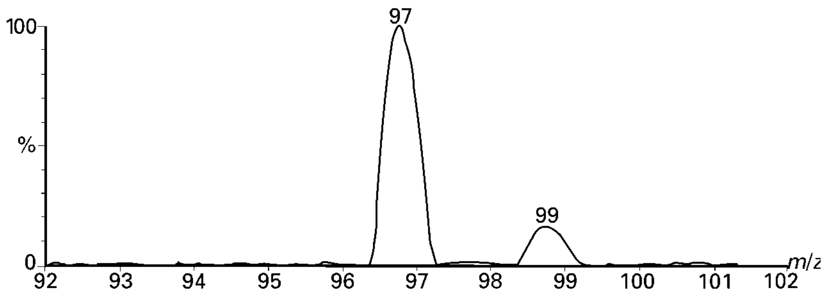



**Figure 9** Negative ion electrospray collision-induced neutral loss of 17 Da (hydroxyl radical) analysis of plasma inorganic sulfate as $H^{32}SO_4^-$ (*m*/*z* 97) and $H^{34}SO_4^-$ (*m*/*z* 99). Phosphate, which occurs in plasma in much larger concentrations, also presents an intense ion with mass 97 Da ($H_2PO_4^-$), but can be distinguished from $H^{32}SO_4^-$ and excluded from the analysis as it eliminates 18 Da (water) under similar conditions. Heavy isotopes of hydrogen, oxygen and sulfur included in $HSO_4^-$ in their natural abundances contribute an ion intensity at *m*/*z* 99 of about 5% of the intensity at *m*/*z* 97. Here the analysis shows that about 12% of the sulfate in the sample is derived from an oral load of $^{32}$S-labelled sulfate.

in multichannel array detection may be used. Calculations usually associated with stable-isotope dilution analyses are then applied. [$^2H_3$]Methyl iodide is used as the quaternizing reagent to avoid interference by endogenous tetramethylammonium ions (mass 74 Da) usually present in urine and blood.

A second use of ESI is in the measurement of inorganic sulfate in blood and urine. Sulfate is extruded actively from cells and resides virtually exclusively in the extracellular fluid compartment. Sulfur has four stable isotopes, $^{32}$S, $^{33}$S, $^{34}$S and $^{36}$S (95.02%, 0.75%, 4.21% and 0.02%, respectively, although these values may vary somewhat with the nature of the source), and as sulfate it is an end-metabolite of sulfur-containing amino acids and other physiologically significant organosulfur compounds. Sulfate does not respond well in positive-ion FAB or ESI, but in negative-ion ESI the ion $HSO_4^-$ is desorbed very efficiently from aqueous solutions. A stable-isotope dilution assay for sulfate based upon this fact uses highly enriched sodium [$^{34}$S]sulfate as the internal standard that is added to the biological fluid. The ratio of the $H^{32}SO_4^-$ analyte and $H^{34}SO_4^-$ reference ions could then be measured directly at *m*/*z* 97 and 99 if phosphate were not also present. $H_2PO_4^-$

also has mass 97 Da, and is present in concentrations much larger than that of sulfate. Collision-induced dissociation of $HSO_4^-$ yields $SO_3^{\bullet-}$ (80 Da) while $H_2PO_4^-$ yields $PO_3^-$(79 Da), which distinguishes them in the tandem mass spectrometer. It is then a simple matter to measure the relative abundances of $H^{32}SO_4^-$ and $H^{34}SO_4^-$ by neutral loss of an OH radical (17 Da), producing ions at *m/z* 80 and 82, respectively, as $H_2PO_4^-$ dissociates exclusively with the loss of water (18 Da) and does not interfere with the sulfate measurement. An adaptation of this technique can be used to monitor sulfate produced from $^{34}$S-containing amino acids in suspected errors of their catabolism. Radioactively-labelled sulfur-containing substrates have the drawback that $^{35}$S (the longest-lived radioisotope) has a half-life of only 87 days, and they must be prepared shortly before each use. [$^{34}$S]sulfate may also be used in tracer studies to follow the excretion of label in the urine following an oral dose of the labelled sulfate. These experiments are usually conducted with radioactive [$^{35}$S]sulfate for the purpose of determining a patient's extracellular fluid volume. An example of this last use is presented in **Figure 9**.

While this article presents and concentrates on examples of the common and routine use of mass spectrometry in the diagnosis of metabolic disease, many other applications are important in the identification and quantitation of metabolites in other areas of clinical medicine.

*See also:* **Biochemical Applications of Mass Spectrometry; Biomedical Applications of Atomic Spectroscopy; Chromatography-MS, Methods; Fast Atom Bombardment Ionization in Mass Spectrometry; Isotopic Labelling in Mass Spectrometry.**

## Further reading

Blasi F (ed) (1986) *Human Genes and Diseases*, Vol 8 in *Horizons in Biochemistry and Biophysics*. Chichester: Wiley.

Borum PR (ed) (1986) *Clinical Aspects of Human Carnitine Deficiency*. New York: Pergamon Press.

Burlingame AL and Carr SA (eds) (1996) *Mass Spectrometry in the Biological Sciences*. Totowa, NJ: Humana Press.

Chapman TE, Berger R, Reijngoud DJ and Okken A (eds) (1990) *Stable Isotopes in Paediatric Nutritional and Metabolic Research*. Andover, UK: Intercept.

Desiderio DM (ed) (1992) *Mass Spectrometry: Clinical and Biomedical Applications*, Vol 1. New York: Plenum Press.

Goodman SI and Markey SP (1981) *Diagnosis of Organic Acidemias by Gas Chromatography–Mass Spectrometry*. New York: Alan R. Liss.

Mamer OA (1994) Metabolic profiling: a dilemma for mass spectrometry. *Biological Mass Spectrometry*, **23**: 535–539.

Matsumoto I (ed) (1992) *Advances in Chemical Diagnosis and Treatment of Metabolic Disorders*, Vol 1. Chichester: Wiley.

Matsumoto I, Kuhara T, Mamer OA, Sweetman L and Calderhead RG (eds) (1994) *Advances in Chemical Diagnosis and Treatment of Metabolic Disorders*, Vol 2. Kanazawa: Kanazawa Medical University Press.

Matsuo T (ed) (1992) *Biological Mass Spectrometry*. Kyoto: Sanei Publishing.

Weaver DD (1989) *Catalog of Prenatally Diagnosed Conditions*. Baltimore: Johns Hopkins University Press.

Wolfe RR (1984) *Tracers in Metabolic Research*. New York: Alan R. Liss.

# Medical Science Applications of IR

**Michael Jackson** and **Henry H Mantsch**,
National Research Council Canada, Winnipeg, Manitoba, Canada

Copyright © 1999 Academic Press

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
|---|
| Applications |

Traditionally a technique associated with astronomy and organic chemistry, IR spectroscopy has emerged as a powerful technique for the analysis and classification of human tissues and fluids. Such an application is possible because of the sensitivity of IR spectroscopy to changes in tissue biochemistry. In general, IR spectroscopy is sensitive to the structure and concentration of the macromolecules (nucleic acids, proteins, lipids) present in cells and tissues and relatively insensitive to low molecular mass metabolites (such as glucose, lactate and individual amino acids). Disease states that affect tissue

ultrastructure are therefore potential targets for IR spectroscopic analysis. For example, cancer is often associated with the appearance of gross nuclear abnormalities, such as the presence of multiple nuclei in cells. As nucleic acids provide some of the most characteristic IR absorptions in biological materials, the changes in DNA associated with cancer result in changes to the spectral signature of malignant cells, and these changes are diagnostic. Changes in absorptions from other macromolecules have been reported to be diagnostic for a wide range of disorders affecting a variety of tissues.

A major advantage of IR spectroscopy is that a single instrument can in principle be used to characterize tissues affected by a wide range of disorders without the need for major reconfigurations of the instrument or the addition of probes (such as stains or other contrast-enhancing reagents) to the sample. In addition, valuable information concerning the molecular nature of the disease process is obtained. A general overview of available sampling methodologies and data analysis techniques will be presented, followed by illustrative examples.

## Methodological aspects

### General comments

Infrared absorption bands from biological materials are generally broad, largely owing to the presence of a heterogeneous population of chromophores. This heterogeneity has two sources. Firstly, functional groups are typically found in a variety of environments (e.g. polar/nonpolar, hydrogen-bonded/non-hydrogen-bonded), giving rise to slightly different but overlapping absorption profiles. Secondly, the complex biochemical nature of tissues, cells and biological fluids means that in many spectral regions overlapping absorptions from different functional groups are found (e.g. phosphate bands from nucleic acids overlap with collagen bands).

From a methodological viewpoint, this means that high-resolution measurements are not required. Typically, a spectral resolution of 2 or 4 $cm^{-1}$ is sufficient. From an interpretational viewpoint, the presence of such broad absorption bands often means that application of band-narrowing techniques such as Fourier self-deconvolution or derivation is required before information can be reliably extracted from spectra. Application of second-derivative methods has the additional advantage that variations in baseline slope and offset are eliminated. While reducing the width of broad overlapping absorption bands, these techniques also disproportionately enhance the weak but narrow water vapour bands in the 1500–1800 $cm^{-1}$ region, leading to many erroneous assignments in the literature. It is therefore essential that the spectrometer and any sampling accessories be well purged.

### Sampling methods

A wide variety of techniques exists for obtaining spectra from tissues and fluids; the simplest of these are transmission methods. For biological fluids, small volumes (typically 5–10 μL) of sample are placed between a pair of $CaF_2$ or $BaF_2$ windows. In the mid-IR the pathlength must be limited to 10 μm or less owing to the presence of intense absorptions from water. Even under these circumstances, absorptions from water dominate the spectrum, and identification of dissolved/suspended materials is difficult. Subtraction of a water spectrum is possible, but it should be borne in mind that the structure of water is altered by dissolved solutes. This alters the spectrum of water, such that water in biological fluids gives rise to spectra that are subtly different from that of pure water. Problems associated with water can be removed by drying the sample to form a thin film. Drying of samples may result in the loss of volatile components from biological fluids, and also results in loss of information relating to hydration of materials. An alternative approach that minimizes sample preparation and prevents loss of information is to acquire spectra in the near-IR spectral region. The absorption bands due to combination and overtone vibrations seen in the near-IR spectral region are significantly reduced in intensity compared to those of the fundamental vibrations. In practice, this means that longer pathlengths may be utilized before water absorptions become a severe problem. However, it should be remembered that only overtones and combinations of X–H vibrations, (X = N, C, O) show significant absorption intensity, reducing the information content of the near-IR spectral region.

For tissue samples, transmission measurements are made in essentially the same way. Samples may either be small pieces of tissue (1 $mm^3$) dissected from a larger sample or microtomed thin sections. In either case, the sample is compressed between calcium fluoride windows. For some tissues (e.g. skin, muscle) such an approach is not possible owing to the high mechanical strength of the sample. For such tissues, specialized sampling techniques and accessories are recommended, such as the attenuated total reflection (ATR) measurements using a ‘split pea’ accessory. Using this approach, tissue is compressed with a reproducible force against a small hemispherical ATR element and spectra are acquired.

A major drawback of these approaches is that the physical integrity of the tissue is destroyed,

preventing subsequent histological analysis. The investigator is therefore often uncertain of the exact nature of the sample being studied. This problem may be avoided with the use of an IR microscope. Using this technique, IR light is focused onto a small region of microtomed tissue and a spectrum is acquired. Spectra are sequentially acquired from the entire surface of the tissue by raster scanning with a computer-controlled stage. In this way a complete spectroscopic picture of the tissue section can be obtained with a high spatial resolution. The sample can then be stained and analysed by standard histological methods. This allows a direct correlation between spectra and sample histology, which aids in spectral interpretation.

### Analysis of preserved samples

In all studies involving biological materials, questions arise of sample degradation over time. With cells and tissues, it is possible to prevent this degradation by fixation with preservatives such as 70% ethanol or formalin. However, these fixatives exhibit strong IR absorptions and are thus a source of potential artefacts. Furthermore, these fixatives preserve tissue by inactivation of degradative enzymes, which may also introduce artefacts into spectra if this inactivation is reflected in changes in protein conformation. Suspension of cultured tumour cells in 70% ethanol significantly alters the absorptions arising from cellular proteins compared to suspension in isotonic saline. Most noticeably, the observation of two protein absorption bands at 1625 and 1680 $cm^{-1}$ is highly indicative of the formation of aggregates of protein molecules stabilized by intermolecular hydrogen bonding. This is generally associated with large-scale protein denaturation. The intensity of these absorptions is time dependent, increasing with prolonged suspension in the alcohol, owing to continued penetration of the alcohol into the cells. Fixation with ethanol also reduces the intensity of the ester C=O stretching vibration, suggesting a decreased lipid content in ethanol-fixed cells. This reduction in lipid content reflects solubilization of membrane lipids by ethanol.

Spectra of cells dried from formalin do not show the characteristic absorption associated with protein aggregates, indicating that formalin fixation does not induce protein denaturation. However, drying from formalin does result in the appearance of a series of sharp absorptions between 1000 and 1500 $cm^{-1}$. These narrow absorption bands arise from formaldehyde that is retained in salt crystals in the film. These distinctive absorptions are not present if the cell suspension is washed with isotonic saline before drying. Formalin should therefore be the fixative of choice for the vibrational spectroscopist, and cells and tissues should be rinsed with isotonic saline before spectral acquisition.

## Spectral interpretation and data analysis methods

### The group frequency approach

IR spectra have traditionally been interpreted by assigning absorptions that fall in particular frequency ranges to specific functional groups within molecules, in what is termed the functional group approach to spectral interpretation. Thus, absorption bands in the range 2800–3050 $cm^{-1}$ are attributed carbon–hydrogen stretching vibrations of ethylene, methylene and methyl groups, while absorptions at 1700–1750 and 1600–1700 $cm^{-1}$ are attributed to stretching vibrations of ester and amide C=O groups, respectively. A brief summary of the assignment of major absorptions in biological systems is presented in **Table 1**. Using such classical assignments, the biochemical and histological properties of tissues and cells may be inferred.

### Functional group mapping

In an extension of the functional group approach to spectral analysis, data acquired by mapping tissues using an IR microscope may be analysed using an approach termed functional group mapping. The principle behind functional group mapping is illustrated in **Figure 1**. A grid first defines the area of the sample to be analysed. A spectrum is then sequentially acquired from each pixel in the grid, and the intensity or peak frequency of an absorption band arising from a functional group of interest is calculated for each spectrum. The intensity or frequency of the absorption band is then plotted as a function of position within the grid to produce a map of intensity or frequency distribution. In essence, therefore, this functional group mapping produces an image of the distribution of materials throughout the sample. For complex samples such as tissue sections, the distribution of important constituents such as collagen, proteins, acylglycerides and nucleic acids throughout the sample can be measured. These 'chemical images' can then be directly compared with the stained tissue to allow correlation of spectral and histological features.

### Statistical and multivariate analysis

Spectroscopic differences between tissues or cell types are often subtle, leading to difficulties in

**Table 1** Frequencies of biologically important IR absorptions. All absorptions cover a range of frequencies, and only representative frequencies are given here

| *Frequency* ($cm^{-1}$) | *Assignment* |
|---|---|
| 3500 | OH stretch; water, carbohydrates |
| 3290 | Amide A (N–H stretch): proteins |
| 3050 | Amide B (N–H bend first overtone): proteins |
| 3010 | Olefinic =CH stretch: lipids, cholesterol esters |
| 2955 | $CH_3$ asymmetric stretch: lipids, proteins, carbohydrates, nucleic acids |
| 2925 | $CH_2$ asymmetric stretch: lipids, proteins, carbohydrates, nucleic acids |
| 2875 | $CH_3$ symmetric stretch: lipids, proteins, carbohydrates, nucleic acids |
| 2855 | $CH_2$ symmetric stretch: lipids, proteins, carbohydrates, nucleic acids |
| 2600 | S–H stretch: proteins |
| 2340 | Solution phase and enclathrated $CO_2$ |
| 2058 | SCN stretch: thiocyanate |
| 1740 | Ester C=O stretch: lipids, cholesterol esters |
| 1655 | Amide I (amide C=O stretch): proteins, α-helices |
| 1635 | Amide I (amide C=O stretch): proteins, β-sheet |
| 1545 | Amide II (amide N–H bend); proteins |
| 1515 | 'Breathing' vibration of tyrosine ring (C–C/C=C stretching) |
| 1467 | $CH_2$ bend: lipids, proteins, cholesterol esters |
| 1455 | $CH_3$ asymmetric bend: lipids, proteins, cholesterol esters |
| 1400 | $COO^-$ symmetric stretch: amino acid side chains, fatty acids |
| 1380 | $CH_3$ symmetric bend: lipids, proteins, cholesterol esters |
| 1280 | Amide III (C–N stretch) of collagen |
| 1240 | $PO_2^-$ asymmetric stretch: phospholipids, nucleic acids; amide III (C–N stretch) of collagen |
| 1204 | Amide III (C–N stretch) of collagen |
| 1170 | CO–O–C asymmetric stretch: phospholipids, cholesterol esters |
| 1150 | C–O stretch, carbohydrates (glycogen) |
| 1080 | $PO_2^-$ symmetric stretch: lipids, nucleic acids. C–O stretch, glycoproteins |
| 1060 | CO–O–C symmetric stretch: phospholipids, cholesterol esters |
| 1035 | C–O stretch, glycoproteins, carbohydrates |

determining the significance of any observed differences. This is illustrated by a comparison of baseline-corrected mean spectra of cultured melanoma, lung adenocarcinoma and breast tumour cells (**Figure 2**). Spectra of the three cell types appear essentially identical, dominated by absorptions from proteins and nucleic acids. However, subtle differences are apparent throughout the spectra, including the region exhibiting absorptions from nucleic acid phosphate vibrations. Given the subtle nature of these spectral differences, the question arises as to the significance of the differences. Obviously, significance must be assessed using statistical techniques. The simplest way to assess whether significant differences exist between sets of spectra is to perform Student's *t*-test. The results of *t*-tests performed to compare the second-derivative spectra of the three groups of cells at each wavelength between 1000 and 1750 $cm^{-1}$ are shown in **Figure 3**. For each comparison, the probability that differences between spectra at each wavenumber arose by chance is plotted. Thus, significant differences exist between spectra of breast and lung adenocarcinoma cells at a number of wavelengths. Similarly, significant differences exist between breast and melanoma cells and between melanoma cells and lung adenocarcinoma cells throughout the spectrum.

Such an analysis allows important conclusions to be drawn. For example, the greatest number of significant differences are found between spectra of breast tumour cells and melanoma cells, indicating that these two cell lines have the most different cellular biochemistry. In contrast, lung adenocarcinoma cells and melanoma cells appear to be the most similar as indicated by the high proportion of wavelengths at which no significant difference is found between spectra. Significant differences are found between all cell lines in the region of the nucleic acid phosphate vibrations, implying that the structure of the DNA in the three cell lines is different.

Analysis of differences between spectra by application of Student's *t*-test directly complements the traditional group frequency approach to spectral interpretation. Unfortunately, spectral regions in which significant differences exist between cell or tissue types are not necessarily the regions that allow optimal classification of spectra into groups based upon cell or tissue type. For example, significant differences exist between spectra of breast and melanoma cells between 1220 and 1240 $cm^{-1}$. However, lung adenocarcinoma cells also differ significantly from breast cells in this region. Thus, the only conclusion that can be drawn is that these cells are not breast cells. More definitive identification requires a comparison of a number of spectral regions, or even the entire spectral range.

Such definitive classification may be achieved with the aid of multivariate pattern recognition techniques such as hierarchical clustering, linear discriminant analysis (LDA) and artificial neural network analysis. Hierarchical clustering techniques compare sets of data (e.g. individually acquired spectra or spectra acquired by mapping of tissue) and group the data according to some measure of similarity. For mapping data, the application of cluster analysis

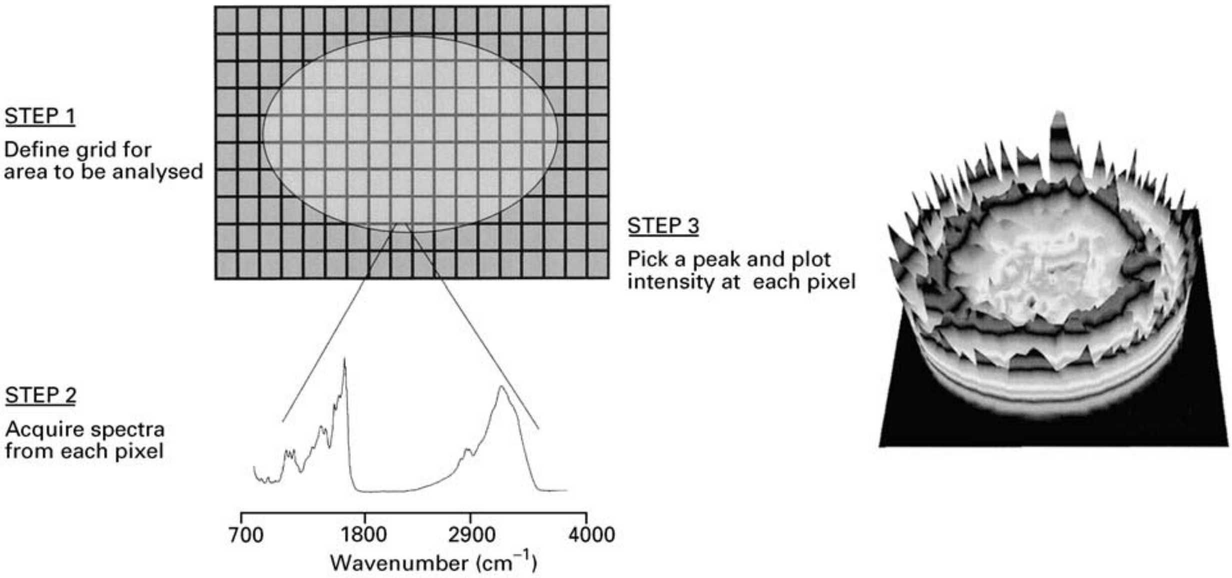



**Figure 1** Pictorial description of the procedure for functional group mapping.

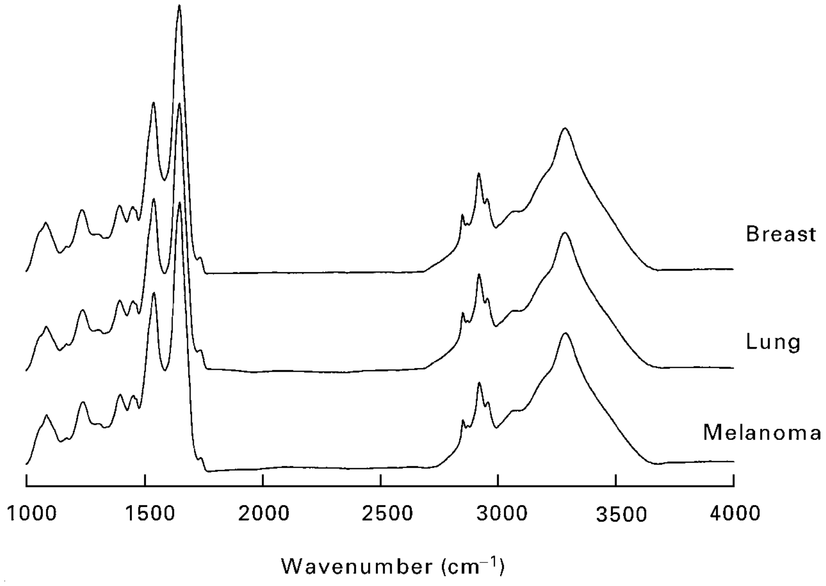



**Figure 2** Infrared spectra of cultured human breast and lung tumour cells and murine melanoma tumour cells.

methods allows the identification of regions of the tissue that give rise to similar spectra and, by inference, have similar biochemistry. The advantage of such methods is that they perform a direct comparison of the spectral data (i.e. they are unsupervised, requiring no input from the investigator), and do not require assignment of spectral features.

More powerful multivariate methods such as LDA make use of the fact that the investigator is often in possession of class assignments (e.g. cell type). In such cases, so-called supervised methods such as LDA can be trained to detect patterns in the spectral data unique to each class. Unknown spectra can then be analysed using the trained algorithm to determine

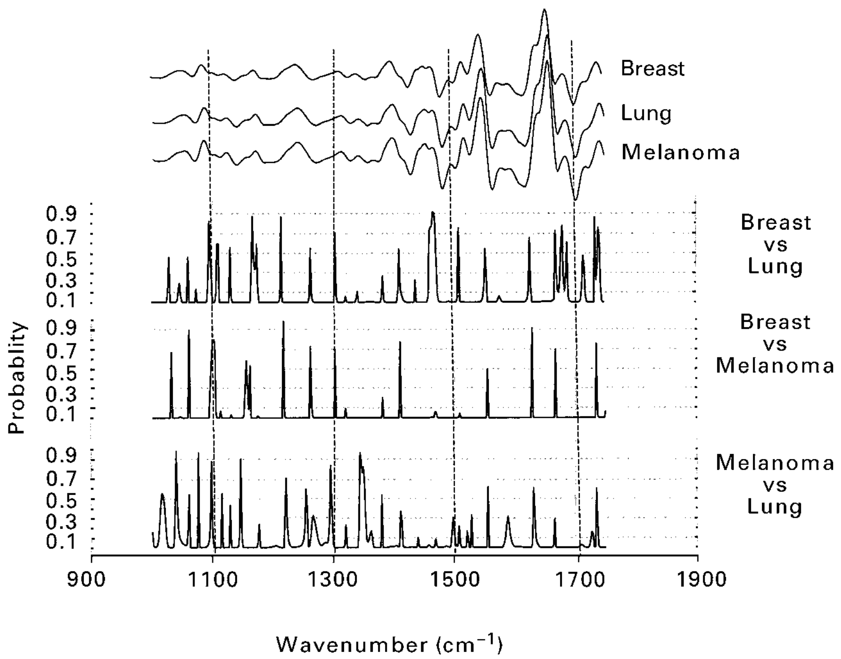



**Figure 3** Results of Student's *t*-test performed to assess the significance of differences between spectra of cultured human breast and lung tumour cells and murine melanoma cells.

the appropriate class assignment based upon the spectral patterns present.

## Applications of IR spectroscopy in medicine: Examples

With due regard to sampling, measurement and data analysis considerations, the application of IR spectroscopy to tissue, fluids and cells yields a remarkable amount of information. This information may relate to the concentration of particular analytes in biological fluids, the distribution of analytes within tissue, or the nature of biochemical changes associated with the disease process. In addition, information concerning metabolic processes in tissues and cells may be obtained. Examples will be presented to illustrate the general applicability of the technique.

### Analyte determination

A major advantage of IR spectroscopy is that an IR spectrum is the summation of the spectra of all of the IR-active species present in the sample, weighted with respect to concentration. This means that the concentration of each of the major IR active analytes in a biological fluid can be predicted from a single IR spectrum. In practice this is achieved using such techniques as partial least squares (PLS) regression analysis, in which IR spectra are regressed against laboratory values for the concentration of analytes of interest to calibrate the method. This calibrated regression method can then be used to predict the concentration of the analytes from the spectrum of a fluid of unknown composition. Application of PLS methods to IR spectra can result in predictive accuracy as good as that of standard laboratory tests for many clinically relevant analytes such as glucose, urea, triglycerides and total protein (**Figure 4**).

### Monitoring of metabolism

PLS techniques are required to decode the concentration information for most analytes present in biological fluids owing to the high degree of overlap between absorptions from most analytes. However, carbon dioxide exhibits a highly characteristic absorption in aqueous solution in a spectral region generally devoid of other absorptions (2340 cm$^{-1}$). This makes it relatively trivial to monitor $CO_2$ concentration. As $CO_2$ is a major by-product of cellular respiration, monitoring the intensity of this absorption in living systems should provide an indication of the rate of cellular metabolism. Obviously this cannot be done *in vivo* with current technology. However, monitoring the rate of respiration within suspensions of cells is possible. **Figure 5** shows the intensity of the absorption at 2340 cm$^{-1}$ as a function of time for a suspension of yeast cells incubated with glucose. An increase in intensity with time as the glucose is metabolized is readily apparent. A

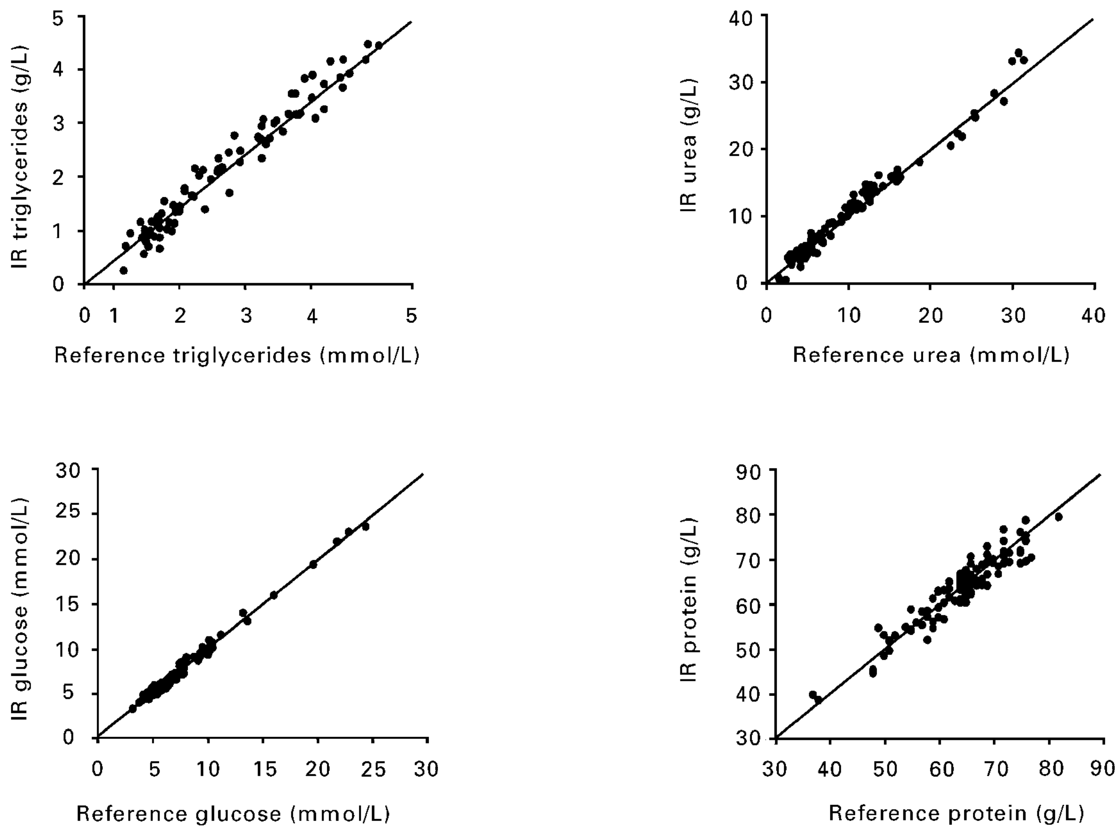



**Figure 4** Concentration of triglycerides, urea, glucose and protein predicted by partial least-squares analysis of infrared spectra of serum plotted against reference values.

second, weaker absorption is apparent at 2275 $cm^{-1}$, attributed to the naturally occurring $^{13}C$ isotope of $CO_2$. The increase in intensity of this second absorption with time in the presence of $^{13}C$-labelled glucose confirms that cellular respiration is indeed responsible for the increase in the intensity of the absorptions at 2275 and 2340 $cm^{-1}$ with time. More importantly, the ability to detect metabolically produced $^{12}CO_2$ and $^{13}CO_2$ should allow IR spectroscopy to be used as a method for the elucidation of metabolic pathways in living cells using specifically labelled materials.

### Grading of breast tumours

As discussed above, differences between spectra of cultured tumour cell lines may be subtle and difficult to discriminate from normal biological variability. Searching for diagnostic features in this relatively simple system may be compared to searching for the proverbial needle in a haystack, and multivariate pattern recognition methods are required. With this in mind, the task of identifying spectroscopic differences between sections of intact tissue becomes a daunting one. Not only must one find the subtle spectroscopic differences that exist between normal and diseased cells, but this information must now be extracted from the large signal arising from the matrix in which the cells sit. As an additional problem, the signal arising from the matrix may itself vary, either as a function of the disease or independently of the disease. The problem now becomes one of finding a needle in a field of haystacks.

The contribution of the extracellular matrix to spectra of breast tissue is shown in **Figure 6**. The spectra of low-grade breast tumours are vastly different from the spectra of cultured breast tumour cells. Breast tissue is composed predominantly of epithelial cells, connective tissue and adipose tissue. If we assume that cultured breast tumour cells and *in situ* breast tumour cells give rise to essentially similar spectra, then the differences seen between the spectra of tumours and cultured cells must reflect the presence of adipose and connective tissues. More specifically, a series of absorptions may be attributed to collagen and triglycerides. This example serves to illustrate not only the extent to which the extracellular matrix influences spectra, but also how spectroscopic studies of cultured cells can be used to identify such matrix effects.

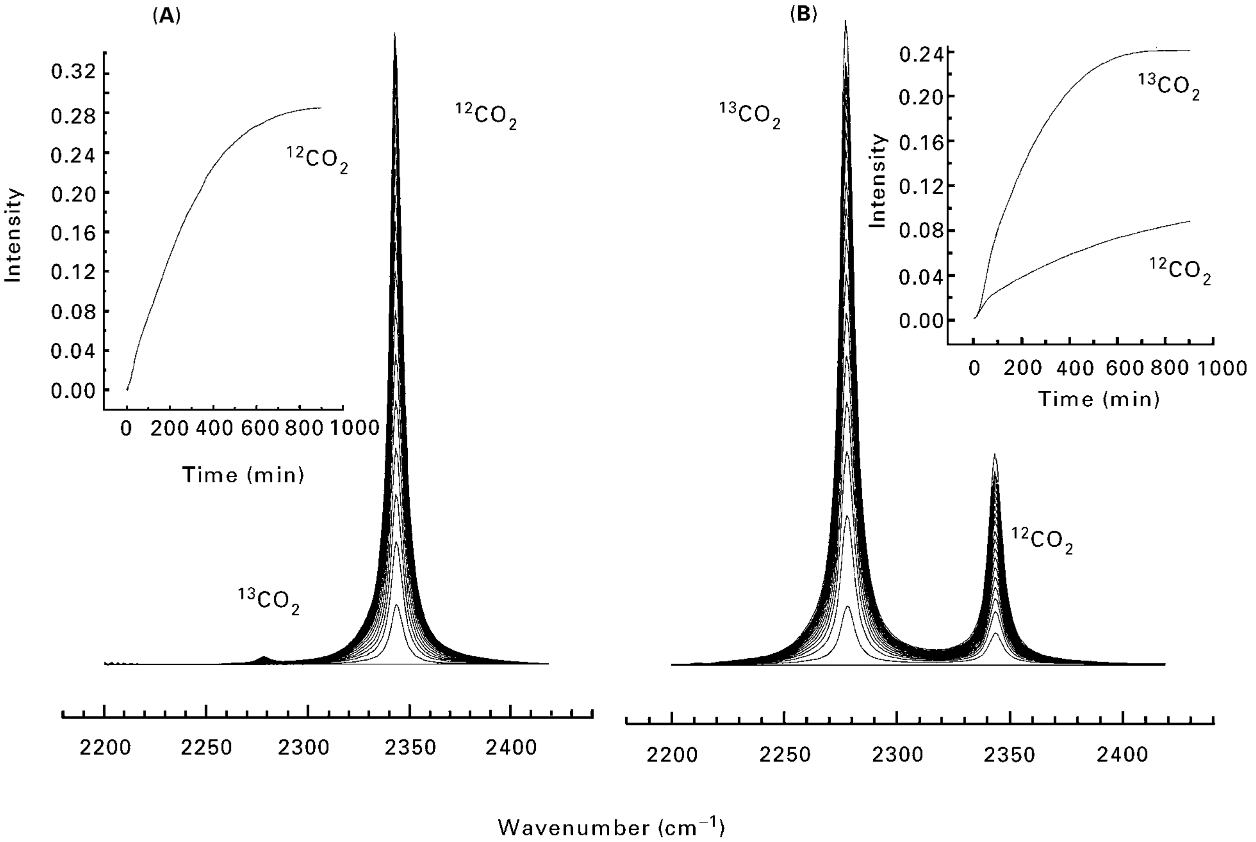



**Figure 5** The intensity of the $^{12}CO_2$ and $^{13}CO_2$ absorption bands as a function of time in suspensions of yeast incubated with (A) $^{12}C$ and (B) $^{13}C$ labelled glucose.

To overcome the problem of matrix interferences and allow grading of tumours, spectra must be analysed by multivariate pattern recognition methods. Much of the spectral data is not related to changes associated with disease progression but results from normal spatial variations in breast histology. This information is therefore redundant, and only serves to increase computation time. Application of a genetic algorithm can be used to determine which regions of the spectrum carry the most useful diagnostic information. These regions can then be used as input for a pattern recognition strategy. Using this approach, tumours can be correctly classified as low, intermediate or high grade with an accuracy approaching 90%.

## Microscopic studies of tumours

The problem of tissue homogeneity can also be addressed using IR microscopy and functional group imaging. Functional group maps showing collagen and adipose tissue distribution in a thin section of a subcutaneous murine tumour are shown in **Figure 7**. These maps are generated by plotting the intensity of characteristic absorptions from collagen and adipose tissue at each point within the map. The flat regions surrounding the tissue periphery correspond to the calcium fluoride substrate on which the tissue is supported. A low lipid content and a high collagen content characterize the periphery of the tissue, suggesting that this region of tissue is the epidermis and dermis. Underlying the epidermis/dermis, a large deposit of adipose tissue can be identified. A discrete band of collagen can then be seen forming a boundary layer around a central mass. This central mass is rich in protein and nucleic acids (as determined from functional group maps of the amide I and phosphate absorptions) but is low in collagen, identifying it as the main body of the tumour. The collagen found around the periphery of the tumour can therefore be recognized as part of a capsule surrounding the tumour. Functional group mapping thus allows identification of discrete histological structures within tissues.

## *In vivo* spectroscopy and imaging

IR spectroscopy can be used to probe tissue biochemistry *in vivo*, using fibre optic accessories to bring the IR light to the patient. Such *in vivo* studies are performed almost exclusively in the near-IR region of the spectrum, owing to the enhanced penetration of

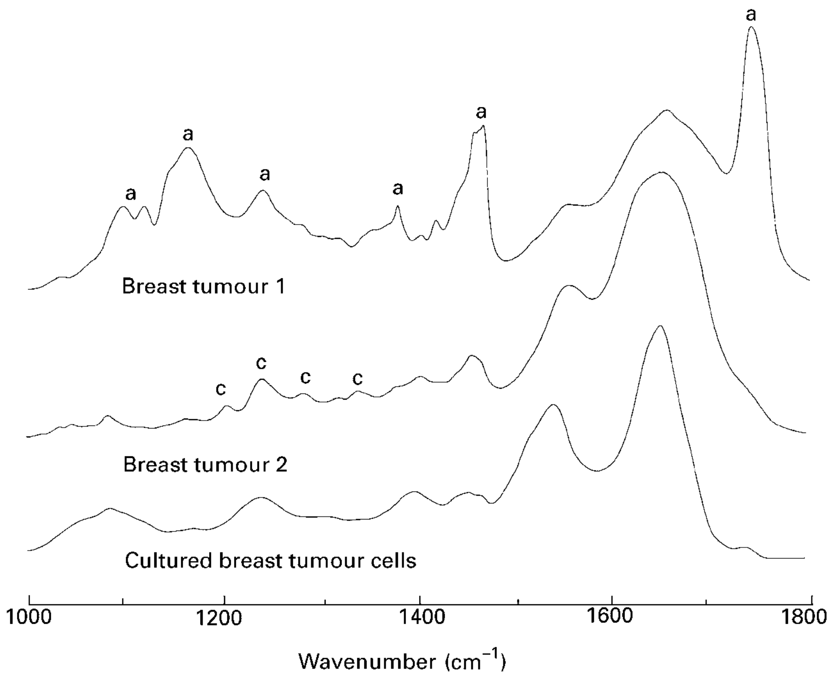



**Figure 6** Infrared spectra of cultured human breast tumour cells and two low-grade human breast tumours. Absorptions marked (a) and (c) arise from adipose tissue and collagen, respectively.

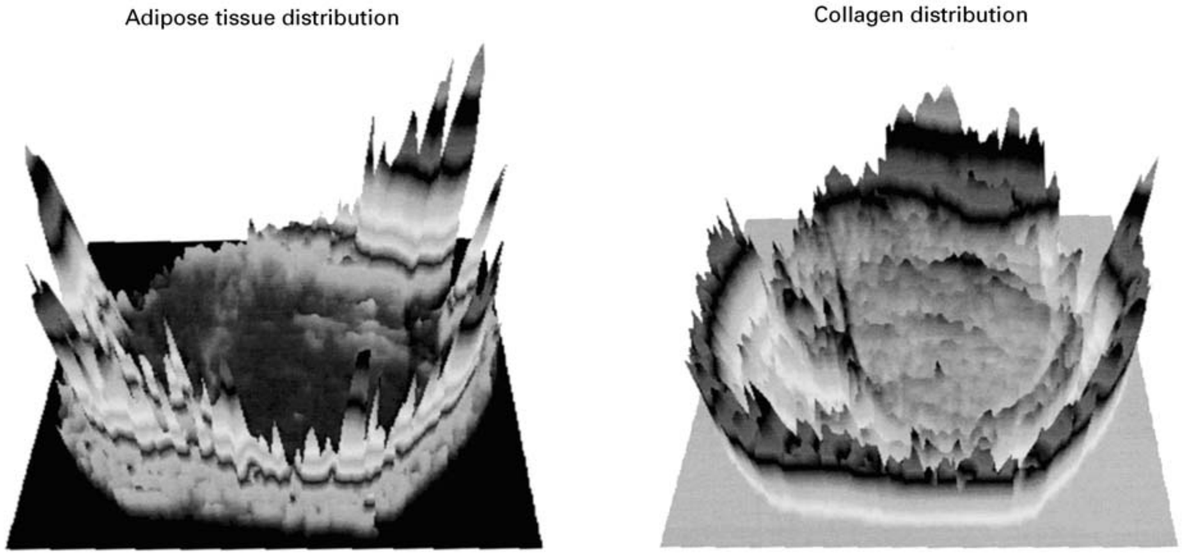



**Figure 7** Functional group maps of a thin section of a subcutaneous tumour from a nude mouse showing adipose tissue and collagen distribution.

near-IR light into tissues. However, applications have been limited by the low information content of near-IR spectra. To date, most near-IR spectroscopic studies have focused on two main areas: blood glucose analysis and haemodynamic monitoring.

Attempts to develop methods for the noninvasive determination of blood glucose have met with little success, despite three decades of research. Major problems include: (i) the complex nature of skin; (ii) variations in the thickness and composition of skin within and between individuals and the consequent difficulties with variations in light scattering; (iii) the relatively low concentration of glucose in body fluids; (iv) the low intensity of the glucose

signal in the near-IR; and (v) overlap of glucose absorption bands with absorption bands from water and other tissue components. Subtle effects such as changes in skin temperature and hydration further complicate the problem.

Haemodynamic monitoring has proved more successful. Haemodynamic monitoring using near-IR spectroscopy is possible owing to the occurrence of electronic transitions in the metal ions of haemoglobin and cytochrome species that are stimulated by absorption of near-IR light. Oxyhaemoglobin, deoxyhaemoglobin, oxidized cytochrome $aa_3$ and reduced cytochrome $aa_3$ exhibit absorptions in the near-IR region of the spectrum that are readily observed. Analysis of the relative intensities of these absorptions can therefore be used to assess tissue oxygenation and oxygen utilization within tissues. In addition, by combining an analysis of the haemoglobin and cytochrome absorptions with an analysis of tissue water absorptions, it is possible to extract information relating to haematocrit, tissue blood volume and blood flow.

The majority of such studies have been conducted using fibre optic-based systems. Spectra can be obtained from a small area of tissue by keeping to a minimum the distance between the fibre optic cables that bring the light to the tissue and those that return the light to the spectrometer. This approach provides information on a highly localized area, but does not provide gross haemodynamic information. Alternatively, the distance between the fibres may be increased and a large volume of tissue can then be probed. This approach provides information relating to the average haemodynamic parameters of the large volume of tissue. However, this approach does not allow localized haemodynamic information to be obtained. Recently, the availability of high-sensitivity charge-coupled-device (CCD) cameras that operate in the near-IR spectral region has been exploited to generate images that provide information concerning the haemodynamic parameters of large areas of tissue with a high spatial resolution. For example, an area of skin 10 cm × 10 cm can easily be imaged using a CDD camera with a 512 × 512 array detector. Spectra are obtained from each detector in the array, and each spectrum then corresponds to a region of the tissue approximately 200 μm × 200 μm. Haemodynamic parameters can be obtained from each spectrum, and these parameters can be used to produce an image. Thus, the global haemodynamic parameters of the 10 cm × 10 cm area of skin can be assessed, but information can also be obtained related to haemodynamic parameters with a spatial resolution limited only by diffraction.

The large amount of information generated in this manner (250 000 spectra) is ideally suited to multivariate analytical methods such as cluster analysis. Combining near-IR imaging and multivariate analysis allows one to determine nonsubjectively regions of tissue having similar haemodynamic properties (**Figure 8**). The applications of such techniques in areas such as wound healing are obvious.

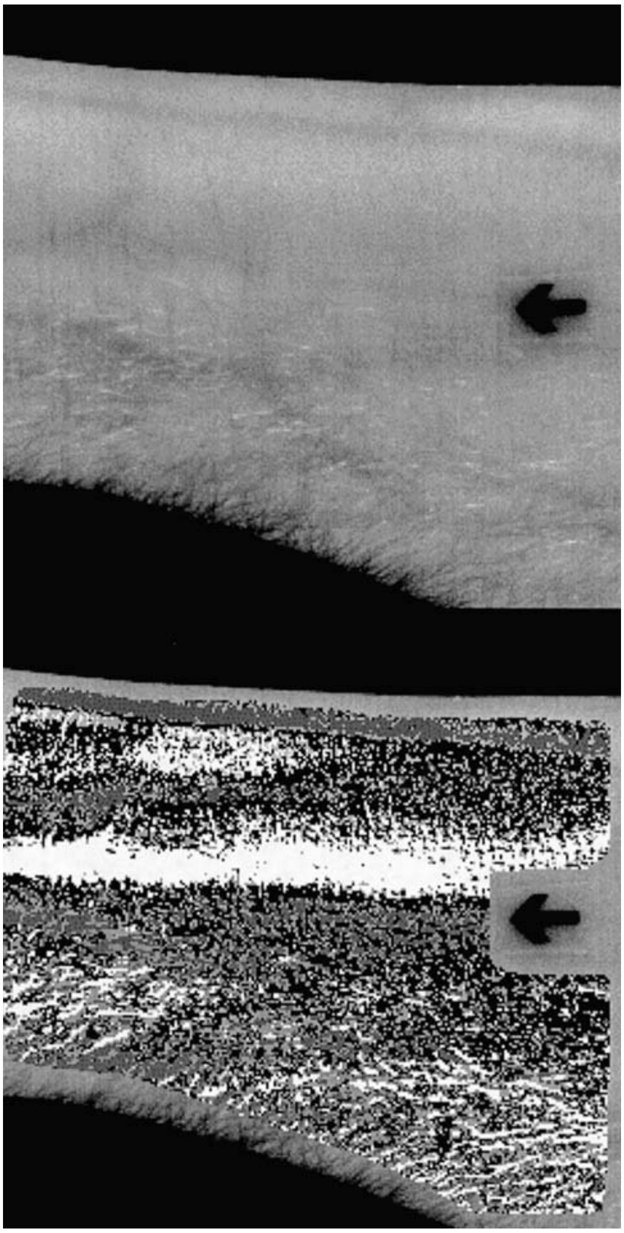


**Figure 8** Noninvasive near-infrared imaging of the human forearm using a 512 × 512 array of CCD detectors. Visible image (upper panel) and results of cluster analysis on each pixel for the deoxyhaemoglobin absorption at 760 nm (lower panel). Areas with similar shading have similar deoxyhaemoglobin levels.

*See also:* **Biochemical Applications of Raman Spectroscopy; FT-Raman Spectroscopy, Applications; Multivariate Statistical Methods.**

## Further reading

*Fourier Transform Infrared (FT-IR) Microspectroscopy. A New Molecular Dimension for Tissue or Cellular Imaging and in situ Chemical Analysis. Cellular and Molecular Biology* Special Issue (1998) Vol. 44, Issue no. 1.

Heise HM (1996) 'Near infrared spectroscopy for non-invasive monitoring of metabolites: state of the art'. *Hormone and Metabolic Research* **28**: 527–534.

Jackson M and Mantsch HH (1996) 'Biomedical IR spectroscopy'. In: Chapman D and Mantsch HH (eds) *FTIR Spectroscopy*, pp 311–340. New York: Wiley.

Jackson M and Mantsch HH (1996) 'FTIR spectroscopy in the clinical sciences'. In: Clark RJH and Hester RE (eds) *Biomedical Applications of Spectroscopy*, pp 185–215, Vol 25 in Advances in Spectroscopy. New York: Wiley.

Jackson M, Sowa M and Mantsch HH (1997) 'Infrared spectroscopy: a new frontier in medicine'. *Biophysical Chemistry* **68**: 109–125.

Manstch HH and Jackson M (eds) (1998) '*Infrared Spectroscopy: A New Tool in Medicine*'. Proc. SPIE 3257.

# Membranes Studied By NMR Spectroscopy

**A Watts**, University of Oxford, UK
**SJ Opella**, University of Pennsylvania, Philadelphia, PA, USA

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

Molecular motions in biomembranes are highly anisotropic. Since essentially all aspects of NMR spectroscopy are affected by molecular motions, the details of these motions strongly influence the spectroscopic approaches that can be applied to membrane samples. Although the motions of individual molecules or molecular groups may be very rapid (the molecular correlation time, $\tau_c$, is in the ns range for lipid hydrocarbon chains and amino acid side chains of proteins), the overall tumbling of the lipid and polypeptide molecules in membrane bilayers is very slow and can be in the ms or s range. For this reason, conventional solution-state NMR, which relies on rapid overall molecular reorientation to narrow the resonance lines, has found little success in membrane research; even a lipid molecule with a relative molecular mass ($M_r$) of ~1000 or a small peptide ($M_r$ ~ 3000) has, when in membrane bilayers, by solution-state NMR criteria a tumbling rate equivalent to a very large ($M_r$ >> 30 kDa) macromolecule in free solution. Only by sonicating the sample to produce small (diameter 20–50 nm) bilayer vesicles can conventional solution-state NMR methods be used. However, not only are there very serious concerns about the status of the lipids and polypeptides embedded in highly distorted lipid bilayers, but also those portions of proteins embedded in the bilayers may still reorient too slowly to yield resolvable resonances. Solution-state NMR spectroscopy has been applied to peptides and proteins in organic solvents, detergent micelles and bicelles (**Figure 1**), and structures have been determined in these model membrane environments. While micelles certainly provide a more relevant environment for membrane proteins than do mixtures of organic solvents, fully hydrated lipid bilayers remain the definitive environment for structural and functional studies of membrane proteins.

As an alternative approach to solution-state NMR methods, which are ineffective with lipid bilayer samples, solid-state NMR methods have been refined sufficiently to permit structural details to be obtained for membrane-embedded peptides and proteins. This usually requires isotopic enrichment, either through chemical synthesis or biosynthetic incorporation in expressed peptides and proteins. In the absence of routine X-ray crystallographic structural studies for these molecules, solid-state NMR spectroscopy has the potential to be a powerful and unique approach to determining the structures and describing the dynamics and functions of membranes and membrane-bound proteins. In addition, solid-state NMR spectroscopy has been widely used to describe lipid structure, dynamics and phase properties. Thus, solid-state NMR experiments can be

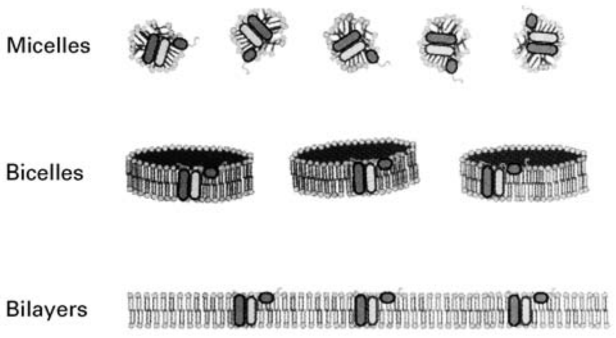



**Figure 1** Representation of the types of sample preparations used for studying membrane lipids and proteins. Micelles and bicelles are usually small (diameters ~ 20–50 nm) structures and bilayers can be sonicated into small (20–50 nm diameter) vesicles, or produced as extended (diameters >> 100 nm) multi-bilayered or single bilayered closed or open structures, depending upon the method of preparation. Natural membranes are usually as large bilayer fragments or closed structures containing a complex and heterogeneous mixture of lipids and proteins and possibly carbohydrates.

applied to both the polypeptide and lipid components of membrane bilayers, without the need to disrupt the sample through sonication or the addition of organic solvents or detergents.

## Nuclei used in membrane studies

With the exception of *J*-couplings, the major magnetic interactions (chemical shift, dipolar and quadrupolar couplings) for the nuclei exploited in biological NMR can be averaged with respect to the applied fields ($B_0$ ~ MHz and $B_1$ ~ kHz) by isotropic molecular motion of small molecules (**Table 1**). However, for biomembranes, any of these interactions may yield resonances with very broad lines and dominate the spectra, masking the resolution required for high resolution studies. Where these interactions can be exploited, their anisotropy (usually chemical shift, dipolar or quadrupolar) can give molecular orientational information from static samples, either oriented or as random dispersions (see below). Alternatively, magic angle spinning (MAS) of the sample can be used at spinning speeds ($\omega_r$) which are either fast enough to average the interaction completely ($\omega_r >> \sigma$, D, Q) to give high resolution-like solid-state NMR spectra, or may be moderated either to recouple a dipolar interaction, such as in rotational resonance or REDOR, or provide orientionally dependent spinning spectral side-bands for nuclei which display chemical shift anisotropy (e.g. $^{31}$P, $^{15}$N).

Naturally occurring $^{13}$C (natural abundance and with selective enrichment) and $^{31}$P nuclei have been extensively exploited in membrane NMR studies

**Table 1** Comparison of the strengths of the magnetic interactions in NMR and the ways in which they can be averaged to yield molecular information for lipids, peptides and proteins in membranes in liquid state and solid state approaches

| *Interaction* | *Liquids* (Hz) | *Solids* (Hz) | *Methods* |
|---|---|---|---|
| σ | 10–$10^4$ | 10–$10^4$ | MAS |
| J | $10^2$ | $10^2$ | Decoupling |
| D | 0 | $10^4$ | Decoupling, MAS |
| Q | 0 | $10^5$–$10^6$ | MAS |

Adapted with permission from Smith SO, Ascheim K and Groesbeck M (1996) Magic angle spinning NMR spectroscopy of membrane proteins. *Quarterly Review of Biophysics* **29**: 395–449. *σ*, Chemical shift anisotropy; J, J-coupling; D, dipolar coupling; Q, quadrupolar coupling.

(**Table 2**). However, replacement of $^1$H by $^2$H or $^{19}$F, and $^{14}$N by $^{15}$N, has also found widespread application, although to date $^{17}$O has not found application in these systems. Typical spectra for the more commonly exploited nuclei for lipids in bilayers are shown in **Figure 2**.

The need to average the strong dipolar coupling (~100 kHz) for $^1$H to obtain high resolution spectra has, until now, excluded widespread observation of this nucleus in membranes. Extensive protein deuteration, to leave a minor $^1$H density at a site of interest for observation in micellar suspensions, has been achieved. The realization that reorientation around the long molecular axis rotation of lipids and proteins in membrane bilayers in the liquid crystalline phase is sufficiently fast (at about $10^9$ Hz for lipids and $10^6$ Hz for proteins with radius ≤ 4 nm in fluid membranes) to average even homonuclear $^1$H dipolar couplings, has opened a new avenue for membrane studies for most observable nuclei, including $^1$H without the need for isotopic replacements. In addition, it is possible to perform magic angle oriented sample spinning (MAOSS) experiments to reap the benefits of both sample orientation and magic angle sample spinning in this situation.

## Nature of the sample

Depending upon the kind of information desired, membrane bilayer samples can be prepared either oriented with respect to the applied field, or as random dispersions. For most studies, full hydration (>30 wt% of water) is desired, especially for protein studies where denaturation may occur and biological function be lost without sufficient amounts of water present.

### Oriented membranes

Both natural and synthetic membranes can be effectively oriented and studied using NMR. In

**Table 2** Properties, advantages and disadvantages of the commonly used nuclei in studies of membranes

| *Nucleus* | *Relative sensitivity* | *Measured parameters* | *Advantages* | *Disadvantages* | *Common applications* |
|---|---|---|---|---|---|
| $^{1}H$ | 1 000 | High resolution spectra<br>Chemical shift,<br>$T_1$, $T_2$ | High sensitivity<br>Natural abundance | Reasonable spectra with small vesicles, micelles, high speed MAS or MAS of oriented bilayers<br>Several relaxation mechanisms<br>Overlapping resonances | Dynamic properties<br>Lipid diffusion |
| $^{2}H$ | 9 | Powder spectra<br>Quadrupole splitting<br>$T_1, T_2$ | Direct determination of order parameters and bond vectors<br>Measurable in cells and dispersed lipids<br>$T_1$ dominated by fast (ns) motions<br>$T_2$ dominated by slow ($\mu$s– ms) motions<br>Low natural abundance | Need for selective deutera tion<br>Low sensitivity | Ordering properties of phospholipids<br>Dynamic properties of phospholipids |
| $^{13}C$ | 16 | High resolution spectra<br>Chemical shift<br>$T_1$<br>Dipolar couplings | Natural abundance<br>$T_1$ dominated by one mechanism | Need MAS NMR to resolve spectra<br>Without selective enrichment, overlapping resonances | Dynamic properties of phospholipids<br>Lipid asymmetry<br>Ligand–protein interactions<br>Distance measurements |
| $^{31}P$ | 66 | High resolution and powder spectra<br>Chemical shift<br>$\sigma$<br>$T_1$<br>NOE | Natural abundance<br>Chemical shift anisotropy is sensitive to headgroup environment and phase properties of the bulk lipids<br>Measurable in cells and in dispersed lipids | Individual lipid classes cannot be resolved in mixed bilayer systems unless sonicated or MAS NMR is used | Quantitation of lipid composition<br>Lipid asymmetry<br>Phase properties |
| $^{15}N$ | 1.04 | High resolution spectra<br>Chemical shift<br>$\sigma$<br>$T_1, T_2$ | Cost of labelling is low<br>Can be incorporated in growth media<br>Chemical shift sensitive to conformation | Low natural abundance, means of labelling required<br>Overlapping resonance | Labelling of proteins and peptides<br>Structural and dynamic studies |
| $^{19}F$ | 830 | High resolution and powder spectra<br>Chemical shift<br>$\sigma$<br>$T_1$ | Chemical shift is sensitive to positional isomers<br>Order parameters can be obtained<br>High sensitivity<br>Measurable in cells and in dispersed lipids | Need for selective fluorination<br>Two factors contribute to the line shape, complicating the analysis<br>High power proton decoupling is difficult<br>May induce chemical perturbation compared to $^{1}H$ | Ordering properties of phospholipids |

$\sigma$, Chemical shift anisotropy; $T_1$, spin–lattice relaxation time; $T_2$, spin–spin relaxation time; NOE, nuclear Overhauser effect; MAS, magic angle spinning.

general, reducing the hydration level of biomembranes supported and oriented on a substratum (glass or mica plates) improves their orientation, but if less than limiting levels of hydration are used (~<30 wt%), alterations may occur in the lipid phase from bilayer to isotropic or hexagonal phases. Fortunately, for all orientational studies of membranes, a good internal check for orientation can be made from the $^{31}P$ NMR spectrum of the bilayer phospholipids, since the line positions in the spectra recorded at 0° and 90° are separated by 40–50 ppm (**Figure** 3). The average mosaic spread can then be estimated from the spectral line-width.

Synthetic, model membranes, made from pure lipids, are best oriented on glass plates by drying down an organic solvent (usually $CHCl_3$/MeOH) solution of the lipid at 1–5 mg $mL^{-1}$. Overloading the lipids onto a substratum generally produces less good

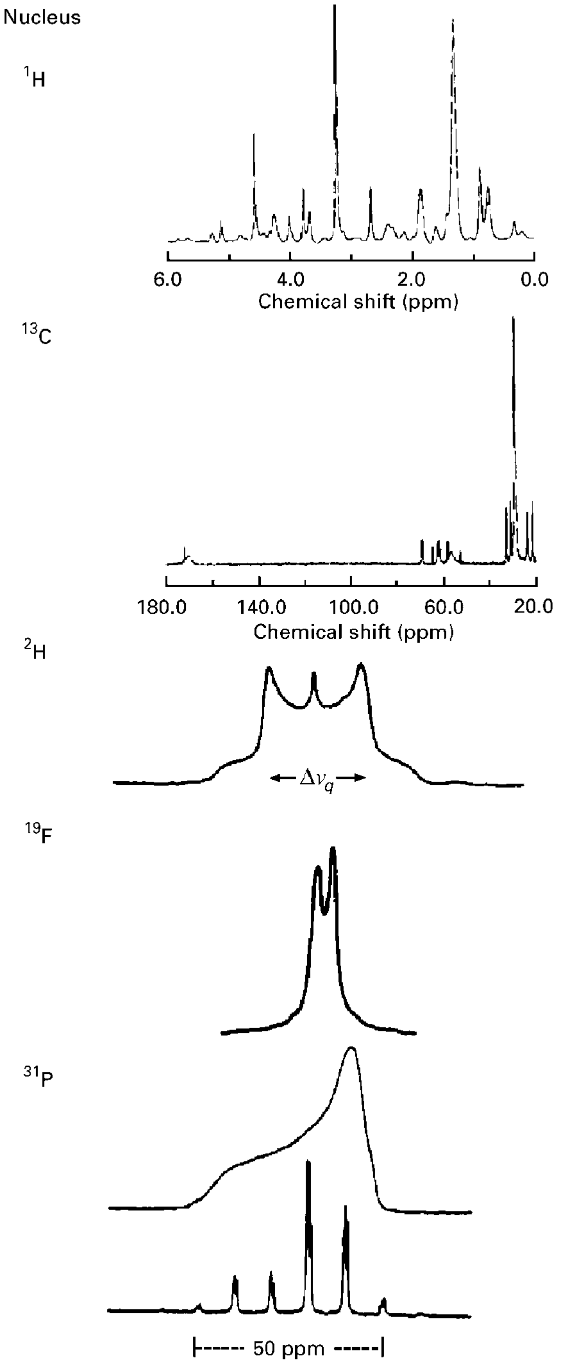



**Figure 2** Typical spectra for various nuclei exploited in studies of lipids in bilayers. $^{1}H$, a spectrum of phosphatidylcholine multi-bilayers, recorded under magic angle oriented sample spinning (MAOSS) conditions to give rise to narrow (~9 Hz) spectral lines; $^{13}C$, proton-decoupled spectrum of sonicated, small (diameters <50 nm) phospholipid vesicles; $^{2}H$, a typical deuterium NMR spherically averaged powder pattern for phospholipid bilayers deuterated in the choline head group, showing the way in which the quadrupole splitting ($\Delta\nu_q$) is determined; $^{19}F$, spectrum of sonicated vesicles of lipids specifically $^{19}F$-labelled in the 12′ position of both acyl chains; $^{31}P$, a static, spherically averaged powder pattern from mixed, cardiolipin, phosphatidylcholine and phosphatidylethanolamine bilayers showing the lack of spectral resolution of the three lipid types (upper spectrum), and under MAS conditions to resolve the three individual phospholipids (lower spectrum).

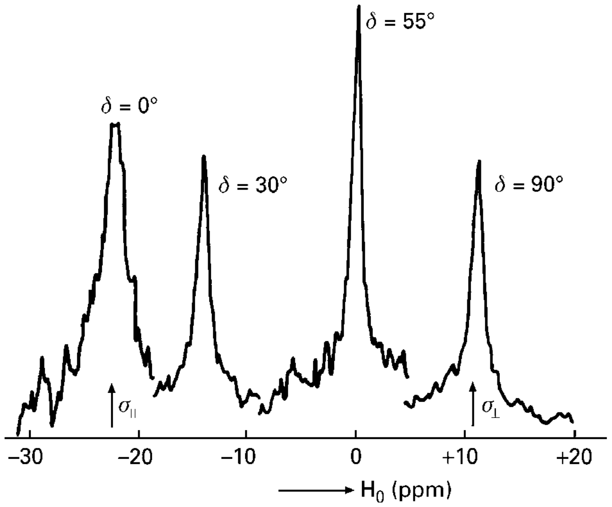



**Figure 3** Orientational dependence of the $^{31}P$ NMR spectra (at 36.4 MHz) of planar multi-bilayers of phosphatidylcholine, where $\delta$ is the angle of the applied field with respect to the membrane normal. $T$ = 77°C. Reproduced with permission from Seelig J and Gally HU (1976) Investigation of phosphatidylethanolamine bilayers by deuterium and phosphorus-31 nuclear magnetic resonance. *Biochemistry* **15**: 5199–5204.

orientation, and some lipids (notably zwitterionic ones such as phosphatidylcholines and phosphatidylethanolamines, both of which are major natural membrane components) orient better than others (anionic ones in particular), with cholesterol aiding orientation. Removal of solvent under high vacuum ($< 10^{-4}$ torr; 1 torr = 133.322 Pa) is followed by hydration, either by dropping buffer or water onto the film, or by incubating in a controlled atmosphere. It is also possible to orient hydrated random dispersions of lipids and proteins by applying pressure to the glass plates. In contrast, natural membranes are most readily oriented on glass plates using the isopotential spin dry ultracentrifugation (ISDU) method which involves centrifugation of a membrane dispersion followed by partial dehydration.

Membrane proteins prepared by solid-phase peptide synthesis or expression in bacteria can be reconstituted into lipids from detergents or organic solvents. The samples are placed on glass plates whose size and shape are determined by the radiofrequency coil with flat or square geometries for optimal filling factors in stationary experiments, or by the rotor for magic angle oriented sample spinning (MAOSS) NMR experiments. While most stationary experiments utilize samples arranged so that the bilayer normal is perpendicular to the direction of the applied magnetic field, it is possible to perform orientational dependent studies with the addition of a goniometer, which is usually rotated from the probe base without probe removal from the

magnet. For MAOSS NMR experiments, the smallest diameter circular thin glass plates which can be handled are usually 4 mm, but for greater sensitivity, larger (up to 14 mm) rotor probes can be used to accommodate larger plates.

### Random dispersions of membranes

Extensively sonicated dispersions of pure lipids usually form small single bilayer vesicles. These samples have been used extensively for $^{13}C$ NMR relaxation studies, which describe the motional properties of the various lipid groups, but have not been widely applied to proteins in membranes.

Natural membrane fragments or vesicles, large (diameter > 50 nm) liposomes, and hydrated unsonicated lipid dispersions are all indistinguishable (because of their slow tumbling) from the NMR perspective. They all give rise to anisotropically broadened spectral envelopes without the application of solid-state NMR techniques. For these systems, a spherically averaged powder pattern is observed which may be narrowed by molecular motion. Magic angle sample spinning (at a rate of $\omega_r$) narrows many of the spectral features, but some potential orientational information content may be lost at high spinning rates ($\omega_r >> $ CSA, D or Q; **Table 1**) where there are few spinning side bands to analyse.

Experimentally, the amount of sample required is determined by the sensitivity of the nucleus to be observed. Typically, mM of the sample (5–10 mg of lipid; 0.5–1 mg of a small peptide; 1–4 mg of a large protein) in a volume of about 0.7 mL or on 10–50 glass plates, may be needed for less sensitive ($^{2}H$, $^{15}N$) nuclei. For $^{1}H$, $^{19}F$, $^{31}P$ and $^{13}C$ in MAS NMR, somewhat less material is required. Cross-polarization from the abundant proton magnetization may improve sensitivity, and decoupling (5–10 kHz for $^{31}P$, 80–100 kHz for $^{1}H$) is routinely used to improve spectral shape and reduce spectral widths.

### Micelles and bicelles

Micelles (SDS is often used) and bicelles (made of long and short chain phosphatidylcholines) tumble isotropically in solution and can accommodate peptides and proteins (**Figure 1**) to provide better mimetics for membrane proteins and peptides than organic solvents, which themselves can induce secondary structures in peptides and proteins. Indeed, good protein function and high resolution NMR spectra can be obtained from micellar-protein complexes, with sufficient resolution to permit structural analysis. Bicelles align in the applied field with the membrane normal perpendicular to the applied field, and by doping the system with lanthanides (Tm, Yb, Er or Eu) the orientation can be turned through 90°. Using paramagnetic chelates, direct protein–lanthanide interactions can be abolished leading to better spectral resolution although some hysteresis effects due to molecular reorganization may complicate their use.

## Information content

### Macroscopic structures

Phospholipids can, in aqueous dispersion, form a range of macroscopic structures including bilayers (predominantly for long chain derivatives, $C_{12}$–$C_{24}$), hexagonal, cubic and rhombic structures (**Figure 4**). $^{31}P$ NMR of static samples is a good diagnostic way of identifying such structures. Spherically averaged powder patterns (with a CSA of ~40–50 ppm) from bilayers are reversed and reduced in spectral widths (by 0.5) for hexagonal structures because of the added degree of freedom of molecular motion along the hexagonal cylinder. Isotropic spectra arise from small (diameter ≥ 50 nm) vesicles, cubic, rhombic and micellar molecular arrangements, as a result of molecular reorientation of the structure with respect to the applied field which is fast enough to average the frequency-dependent $^{31}P$ chemical shift anisotropy ($\tau_c$ < 3–10 kHz ~ = 40–50 ppm on 200–600 MHz instruments). Similar spectral changes are observed with $^{2}H$ NMR for deuterated lipids in similar structures.

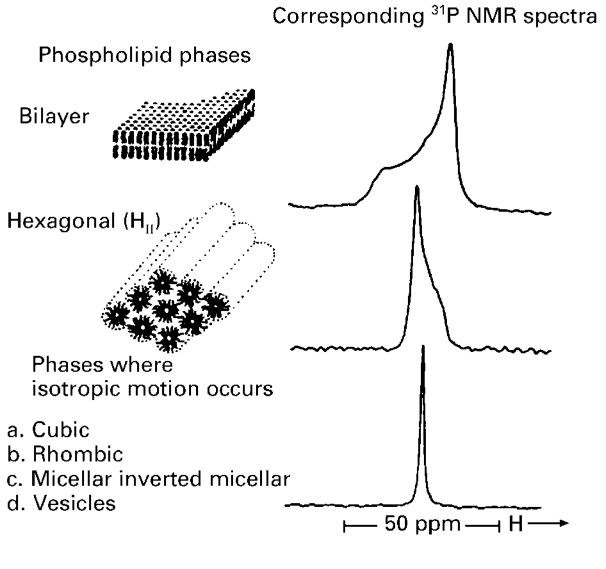



**Figure 4** Three different types of phospholipid phases and their corresponding $^{31}P$ NMR spectra. From Cullis PR and Kruijff B (1979) Lipid polymorphism and the functional roles of lipids in biological membranes. *Biochimica et Biophysica Acta* **559**: 399–420.

A two-component bilayer and isotropic $^{31}$P NMR spectrum from a membrane (usually a natural membrane) has been interpreted in terms of a major bilayer structure encompassing a much smaller (≤5%) population of inverted micelles. Although not the only explanation, the interpretation has many functionally attractive features (enhanced permeability, sites of membrane fusion, flip-flop regions, etc). Such isotropic spectral components are often produced by proteins interacting with the surface of lipid bilayers. The identity of the lipid type, in a mixed lipid membrane, which exists in this isotropic environment, can be determined using MAS $^{31}$P NMR methods.

## Molecular structure

Lipid bilayers are two-dimensional, axially symmetric structures with the normal to the plane of the bilayer being a reference axis. Lipids usually rotate quickly ($\tau_r$ ~ ns) around their long axis in fluid bilayers, but peptides and proteins may lack this motion which is determined by the association (controlled oligomerization or irreversible aggregation) behaviour of the various components, as well as the lipid dynamics.

**Lipid orientational information** By exploiting the anisotropic properties of the nuclear spin-interactions of $^{2}$H, $^{13}$C and $^{15}$N (**Table 1**), placed at specific positions in peptides and ligand, rather precise bond vectors have been determined in oriented membranes.

Deuterium has low natural abundance and a low γ, but the quadrupole coupling can give rise to large splittings in the $^{2}$H NMR spectrum ($\Delta\nu_q$(max) ~127 kHz), which is averaged by rotation of a methyl group around a $C_3$ axis (to 127/3 ~ 40 kHz). This high degree of orientational sensitivity has been exploited to determine the structure of specific residues (valines) involved in dimeric association for the gramicidin ion channel, and the structure of retinal within its binding site of bacteriorhodopsin and rhodopsin. Spectral simulations are necessary, as is some estimate of line broadening due to macroscopic mosaic spread (from $^{31}$P spectral widths), but as a general method, the information gained is *ab initio* and, as such, model-independent. Most amino acids are available in a $^{2}$H-labelled form for incorporation into peptides in solid phase synthesis, and a wide range of $^{2}$H-precursors are available for specific labelling of ligands and prosthetic groups.

**Peptide and protein orientation** Many results have been obtained with specific, selective and uniform labelling of polypeptide sites with the spin $S = \frac{1}{2}$ nuclei $^{13}$C and $^{15}$N. These results have included the structure determination of gramicidin in bilayers, the geometrical arrangements of a variety of helical peptides in bilayers, the three-dimensional structures of peptides in bilayers, and the orientations of bound ligands and prosthetic groups. Uniform labelling of proteins with $^{15}$N offers many advantages, and, indeed, was first developed for solid-state NMR spectroscopy before being applied to solution-state NMR where it has achieved nearly universal use. The nitrogen sites in the protein backbone are separated by two carbon atoms, leaving only minimal homonuclear $^{15}$N dipolar effects, which is the essence of dilute-spin solid-state NMR spectroscopy. Each residue has one amide nitrogen in the backbone (with the exception of proline) and some have distinctive side-chain nitrogen sites. Further, uniform labelling allows the use of expressed proteins, and shifts the burden from sample preparation to spectroscopy, where complete spectral resolution is the essential starting point for structure determination.

Multidimensional solid-state NMR experiments have been shown to yield completely resolved spectra of uniformly $^{15}$N labelled proteins in oriented lipid bilayers. In three-dimensional spectra, each amide resonance is characterized by three frequencies ($^{1}$H chemical shift, $^{15}$N chemical shift and $^{1}$H–$^{15}$N heteronuclear dipolar coupling), which provide the source of resolution among the various sites as well as the basic input for structure determination based on orientational constraints. The data shown in **Figure 5** are from a 50-residue protein in oriented lipid bilayers. More importantly, since the polypeptides are immobilized by the lipids on the relevant NMR time-scales, there can be no further degradation of line widths or other spectroscopic properties as the size of the polypeptide increases. Although larger proteins will have more complex spectra resulting from the increased number of resonances, there is no fundamental size limitation to solid-state NMR studies of membrane proteins.

In principle, a single three-dimensional correlation spectrum of an oriented sample of a uniformly $^{15}$N labelled protein provides sufficient information in the form of orientationally dependent frequencies for each amide site to determine the complete structure of the polypeptide backbone. The two-dimensional $^{1}$H–$^{15}$N heteronuclear dipolar–$^{15}$N chemical shift PISEMA spectrum in **Figure 5A** was obtained from a uniformly $^{15}$N labelled sample of the 50-residue fd coat protein in oriented bilayers; it contains resonances from all of the amide backbone (and side-chain) nitrogen sites. The resonances in the box in the upper left are largely from residues in the

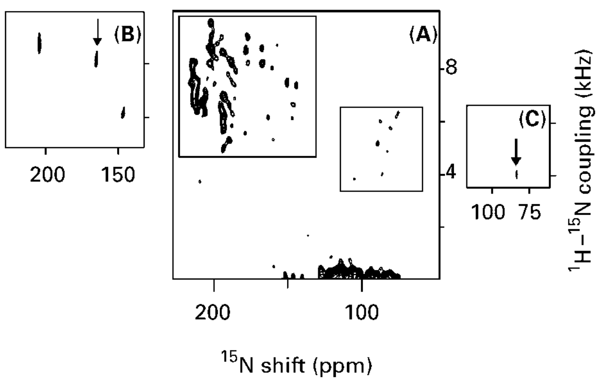



**Figure 5** Multidimensional solid-state NMR spectra of uniformly $^{15}N$ labelled fd coat protein in oriented lipid bilayers. Panel (A) is the complete two-dimensional $^{1}H$–$^{15}N$ heteronuclear dipolar–$^{15}N$ chemical shift spectrum. Panels (B) and (C) are spectral planes extracted from a three-dimensional correlation spectrum at specific $^{1}H$ chemical shift frequencies. The spectral regions in the planes correspond to the boxed regions of the complete two-dimensional spectrum with which they are aligned. Panel (B) contains resonances with $^{1}H$ chemical shift of 11.0 ppm. Panel (C) contains resonances with $^{1}H$ chemical shift of 11.6 ppm. The three orientationally dependent frequencies that can be measured for all of the resonances in the three-dimensional data set provide the input for structure determination. The arrow in panel (B) points to the resonance assigned to the amide of Leu 41 in the trans-membrane hydrophobic helix and that in panel (C) to the amide of Leu 14 in the in-plane amphipathic helix.

trans-membrane hydrophobic α-helix, and the resonances in the box on the right are from residues in the in-plane amphipathic α-helix. The two-dimensional PISEMA spectrum is generally well resolved, although there is some overlap among the resonances from residues in the trans-membrane α-helix because their peptide bonds have similar orientations, approximately parallel to the magnetic field. Two-dimensional planes extracted from a three-dimensional correlation spectrum of the same sample are shown in **Figures 5B** and **C**. The spectral regions in the planes correspond to the boxed regions of the complete two-dimensional spectrum with which they are aligned. There are very few resonances in the spectral regions in **Figures 5B** and **C** because only those resonances with a specific $^{1}H$ chemical shift, the third frequency in the three-dimensional spectrum, appear in each plane. These data illustrate how the three-dimensional correlation experiment contributes to these studies. First, it provides a substantial increase in resolution by separating resonances based on their $^{1}H$ chemical shift frequencies. Second, it enables the direct measurement of three orientationally dependent frequencies for each resolved resonances, the $^{1}H$ chemical shift, $^{1}H$–$^{15}N$ heteronuclear dipolar coupling and the $^{15}N$ chemical shift. Because the orientation of the protein-containing bilayers is fixed by the method of sample preparation, these frequencies can be used to determine the orientation of each peptide plane with respect to the applied magnetic field, since the magnitudes and orientations of the spin-interactions tensors for the amide sites in the molecular frame are known. Structures are then determined using the frequencies associated with individual resonances that have been assigned to specific residues as angular constraints. Using this approach the three-dimensional structure of the M2 trans-membrane segment from the acetylcholine receptor in oriented bilayers was determined. The $^{13}C$ chemical shift and $^{2}H$ quadrupolar coupling frequencies measured from selectively or specifically labelled samples in separate experiments also provide valuable structural information, and have been used together with the $^{15}N$ chemical shift and the $^{1}H$–$^{15}N$ heteronuclear dipolar coupling to determine the structure of the gramicidin channel at high resolution.

**Amplitude of motion–order parameters** Partial but fast ($\tau_c < \Delta\nu_{Q(\max)}$; $\Delta$ $CSA_{(\max)}$) motional averaging of spectral anisotropy permits an order parameter, and hence an angle of motional amplitude with respect to a fixed axis (the membrane normal), to be determined. Direct measurements of order parameters from the powder patterns of random bilayers dispersions of deuterated lipids are conveniently made from the measured quadrupole splittings ($\Delta\nu_q$) (**Figure 2**), determined from the spectral maxima, corresponding to the 90°-orientational spectral components. Thus:

$$\Delta\nu_q = \left(\frac{e^2Qq}{h}\right)\left(3\overline{\cos^2\theta} - 1\right)$$

where $(e^2Qq/h) = 127$ kHz and thus

$$|S| = \frac{(\Delta\nu_q)_{\text{obs}}}{(\Delta\nu_q)_{\max}}$$

and so $1 > |S| > 0$.

The order parameter profile ($S$ at various positions of measurement) usually observed for lipid acyl chains displays a plateau for lipid methylenes in the upper part of the bilayer ($C_2$–$C_{10}$; $S \sim 0.4$) as a result of water penetration and ordering of the upper part of the bilayer and then a decrease to the centre of the membrane ($C_{10}$–$C_{16}$ gives $S$ values of 0.4–0.1) (**Figure 6**). The discontinuity in $S$-profile at $C_{10}$–$C_{12}$ corresponds to the main position of unsaturation in

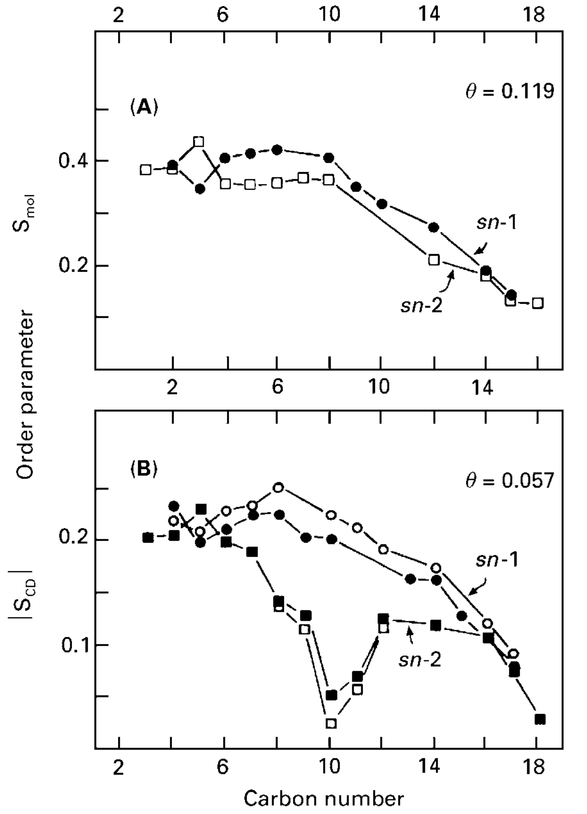



**Figure 6** The measured deuterium order parameters, $S_{CD}$ for the *sn*-1 (saturated) and *sn*-2 (unsaturated, C9=C10) bilayers of phospholipids specifically deuterated in their acyl chains, showing the consequence of the presence of double bonds causing the staggered conformation of the chains thus affecting the quadrupolar averaging (B), even though the molecular order parameter, $S_{mol}$, remains for each chain (A). Reproduced with permission from Seelig J and Seelig A (1977) Effect of single cis double bond on the structure of a phospholipid bilayer. *Biochemistry* **16**: 45–50.

natural membranes. At this position, the static kink in the chain contributes geometrically to the angle about which motional averaging occurs, thereby reducing the molecular order parameter ($S_{mol}$ determined from $\Delta\nu_{q(obs)}$), since:

$$|S_{mol}| = |S(\text{orientation})| \times |S(\text{amplitude})|$$

Labelled $^2$H lipid head groups give rise to small ($\Delta\nu_q$ < 10 kHz) quadrupole splitting, since their amplitude of motion with respect to the membrane normal is large, and their axis of averaging is not the membrane normal. In this case, the angle of molecular averaging cannot be uniquely determined, but the experimentally determined $\Delta\nu_Q$ values can be useful in studying ion binding (**Figure** 7), peptide and protein interactions, and electrostatic interactions, since the phospholipid head group acts as a molecular voltmeter and sensor of interfacial pH at the bilayer surface.

**Distance measurements within membranes** In magic angle spinning (MAS) NMR, sample spinning averages out the dipolar couplings which are required for distance determinations. However, internuclear distance measurements can be made using rotational resonance ($R^2$) for homonuclear spins and REDOR for heteronuclear spins, in which the dipolar couplings are reintroduced into the spectrum under special spinning conditions. By spinning (at a speed of $\omega_r$) the sample at multiples ($n$, where $n$ = 1,2,3,4……) of the chemical shift difference ($\Delta$ in Hz) between a specific spin pair such that $\omega_r = n\Delta$, then transfer of magnetization occurs between the spin pair, and the dipolar interactions are recoupled. Now, the dependence of the NMR spectral intensity with mixing time shows a dependence on the distance between the spin pairs, and hence the internuclear distance ($r$) can be determined. This approach has been used to determine $^{13}C$ spin pair distances to sub-nm resolution in membranes between neighbouring lipids, between lipids and proteins (**Figure** 8), within a ligand at informative sites to give the structure of ligand at its site of action in a membrane bound target and of a prosthetic group (retinal) in its binding site of a receptor.

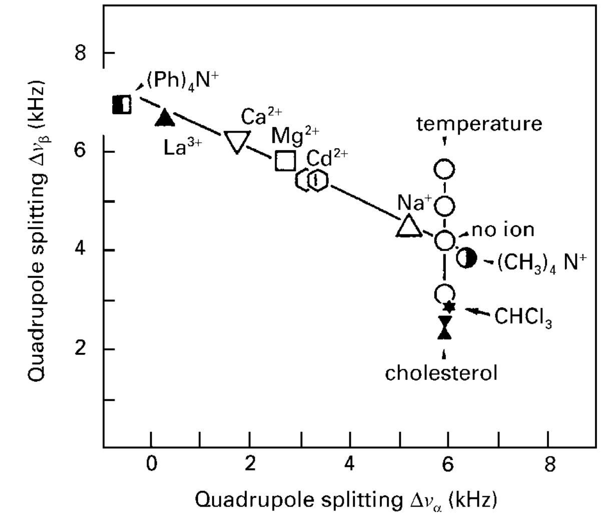



**Figure 7** Deuterium NMR quadrupole splittings ($\Delta\nu_q$) of the $\beta$-$CD_2$ group plotted against the corresponding splitting for the $\alpha$-$CD_2$ group of dipalmitoylphosphatidylcholine -$P(O_4)^-$–$\alpha CD_2$–$\beta CD_2$–$N^+(CH_3)_3$ in bilayers for a range of ions at constant ionic strength, $I$ = 1.05 M and at 59°C, showing the sensitivity of lipid head group orientation to surface charge. Reproduced with permission from Akutsu H and Seelig J (1981) Interaction of metal ions with phosphotidylcholine bilayer membranes. *Biochemistry* **20**: 7366–7370.

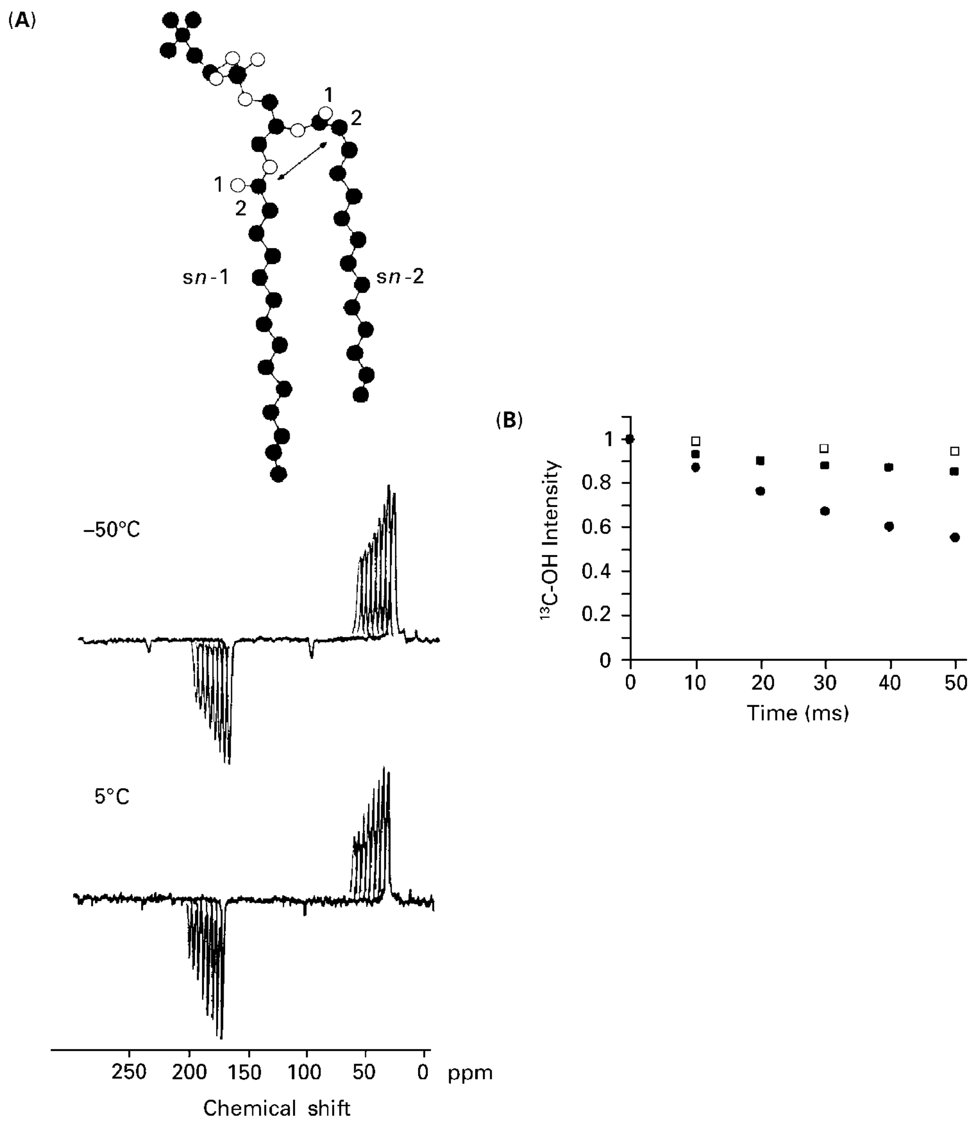



**Figure 8** Rotational resonance $^{13}C$ MAS NMR has been used to determine both an intramolecular distance within a lipid and an intermolecular distance between a lipid and a protein in bilayers. A train of spectra are shown at different mixing times for the $n = 2$ resonance condition at a spinning speed of 6248 ± 5 Hz, at two different temperatures, giving a distance (from an analysis of the intensity changes with mixing time) from the C1 on the *sn*-1 chain to the C2 on the *sn*-2 chain of 4.0–5.0 Å at –50°C in dipalmitoylphosphatidylcholine bilayers (A). To determine the intermolecular peptide–lipid distance, the spectral intensity changes due to magnetization transfer were determined under conditions of no magnetization transfer, that is, at off-resonance, (□), with unlabelled lipid to show the contribution from natural abundance (1.1%) $^{13}C$, (■), and with $^{13}C$-1,2-[2-$^{13}C$] labelled lipid in bilayers containing glycophorin labelled at residue, $^{13}C$–OH tyrosine-93 (●), to give a distance of 4.0–5.0 Å. Reproduced with permission from Smith SO, Hamilton J, Salmon A and Bormann BJ (1994) Rotational resonance NMR determination of intra- and intermolecular distance constraints in dipalmitoylphosphatidylcholine bilayers. *Biochemistry* **33**: 6327–6333).

Paramagnetic spectral broadening for a MAS NMR observed phosphate induced by a nitroxide spin-label, both at known positions in the primary sequence of membrane-bound bovine rhodopsin, a photoreceptor, has permitted some estimate of helix-loop distances to be made in the protein, supplementing other indirect structural details, even though the protein crystal structure is not available.

## Membrane dynamics

Both spin–lattice ($T_1$) and spin–spin ($T_2$) relaxation times give motional information about specific nuclei in a membrane system. Faster ($\tau_c \sim 10^7$–$10^9$ s) motions associated with acyl chain *trans–gauche* rotational isomerisms and protein residue motions affect $T_1$ values whereas slower ($\tau_c \sim 10^3$–$10^4$ s)

peptide backbone and membrane director fluctuations are inherently detected by $T_2$ values. Conventionally, high resolution NMR methods for measuring these relaxation times are used, the only difference being that broad spectral envelopes may show anisotropy in relaxation characteristics.

Membrane protein side chains (for example, Lys-$^{\varepsilon}CH_2$; Val-$^{\gamma}CH_2$) have been shown to possess fast (μs) motions from $T_1$ measurements of specifically deuterated residues in bacteriorhodopsin, even though the membrane environment was relatively rigid and crystalline. Similar approaches showed that the α-$CH_3$ group rotation is fast ($\tau_c$ << μs) for membrane-bound, specifically deuterated retinal in the same protein, even at –60°C. Somewhat slower motions (in the ms range) occur in peptide backbones of membrane proteins.

An enhancement of the $T_1$ for cardiolipin phosphates through direct contact of the haem group in cytochrome *c*, a peripherally bound protein, suggests that this protein undergoes considerable conformational distortions (molten globule) when at the bilayer surface, on the $10^{-6}$ s time-scale, a feature of peripherally bound proteins which may be general.

Slow (ms) director fluctuations (wobbling around the membrane normal) are induced in bilayer membranes by proteins, as shown by measurements of $T_2$. These motions are strongly coupled through the protein and membrane and may be significant in maintaining the protein in a dynamic equilibrium ready for function.

### Ligand structure

Small molecules activate a range of cellular responses following binding to membrane-bound receptors. Solid-state NMR methods permit the structure and binding kinetics of such small ligands, using isotropic labelling to aid assignment and enhance sensitivity. The β-ionone ring orientation (6-S *trans* or 6-S *cis*) of $^{13}C$-retinal in bacteriorhodopsin has been determined from direct measurements using rotational resonance MAS NMR, by comparison with the known crystal structure of retinal. Deuterated retinal in bacteriorhodopsin and mammalian rhodopsin has been observed and the structural changes induced upon light incidence determined to good precision using an *ab initio* approach. $^{13}C$-labelled retinal has proved particularly successful for use in describing sugar binding to sugar transporters, drug binding to P-type ATPases ($H^+/K^+$-ATPase and $Na^+/K^+$-ATPase) and acetylcholine binding to the membrane-bound receptor. In the absence of crystals of this class of proteins at the present time, detailed structural information from solid-state NMR approaches and observation of a range of isotopes are helping to elucidate functional descriptions by this approach.

*See also:* **Cells Studied By NMR; $^{13}C$ NMR, Methods; $^{13}C$ NMR, Parameter Survey; Diffusion Studied Using NMR Spectroscopy; Liquid Crystals and Liquid Crystal Solutions Studied By NMR; Macromolecule–Ligand Interactions Studied By NMR; NMR in Anisotropic Systems, Theory; $^{31}P$ NMR; Proteins Studied Using NMR Spectroscopy; Solid State NMR, Methods; Solid State NMR, Rotational Resonance; Solid State NMR Using Quadrupolar Nuclei.**

## Further reading

Cross TA and Opella SJ (1994) Solid-state NMR structural studies of peptides and proteins in membranes. *Current Opinion in Structural Biology* **4**: 574–581.

Glaubitz C and Watts A (1998) Magic angle-oriented sample spinning (MAOSS): a new approach toward biomembrane studies. *Journal of Magnetic Resonance* **130**: 305–316.

Gröbner G, Taylor A, Williamson, PTF *et al.* (1997) Macroscopic orientation of natural and model membranes for structural studies. *Analytical Biochemistry* **254**: 132–136.

Marassi FM, Ramamoorthy A and Opella SJ (1997) Complete resolution of the solid-sate NMR spectrum of a uniformly $^{15}N$-labeled membrane protein in phospholipid bilayers. *Proceedings of the National Academy of Sciences* **94**: 8551–8556.

Opella SJ (1997) NMR and membrane proteins. *Nature Structural Biology* **4**(suppl.): 845–848.

Opella SJ, Stewart PL and Valentine KG (1987) Protein structure by solid-state NMR spectroscopy. *Quarterly Review of Biophysics* **19**: 7–49.

Opella SJ, Marassi FM, Gesell JJ *et al.* (1999) Three-dimensional structure of the membrane embedded M2 channel-lining segment from nicotinic acetylcholine receptors and NMDA receptors by NMR spectroscopy. *Nature Structural Biology* **6**: 374–379.

Pines A, Gibby MG and Waugh JS (1973) Proton-enhanced nmr of dilute spins in solids. *Journal of Chemical Physics*, **59**: 569–590.

Pinheiro TJT and Watts A (1994) Resolution of individual lipids in mixed phospholipid membranes and specific lipid-cytochrome *c* interactions by magic angle spinning solid-state phosphorus-31 NMR. *Biochemistry* **33**: 2459–2467.

Ramamoorthy A, Marassi FM and Opella SJ (1996) Applications of multidimensional solid-state NMR spectroscopy to membrane proteins. In: Jardetzky O and Lefevre J (eds) *Dynamics and the Problem of Recognition in Biological Macromolecules*, pp 237–255. New York: Plenum.

Reid DG (ed) (1997) Protein NMR techniques. In: *Methods in Molecular Biology*. Totawa, New Jersey: Humana Press.

Sanders CR II and Landis GC (1995) Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies. *Biochemistry* **34**: 4030–4040.

Smith SO, Ascheim K and Groesbeck M (1996) Magic angle spinning NMR spectroscopy of membrane proteins. *Quarterly Review of Biophysics* **29**: 395–449.

Vold RR, Prosser RS and Deese AJ (1997) Isotropic solutions of phospholipid bicelles: A new membrane mimetic for high-resolution NMR studies of polypeptides. *Journal of Biomolecular NMR* **9**: 329–335.

Watts A, Ulrich AS and Middleton DA (1995) Membrane protein structure: the contribution and potential of novel solid state NMR approaches. *Molecular Membrane Biology* **12**: 233–246.

Watts A (1993) Magnetic resonance studies of phospholipid–protein interactions in bilayers. In: Cevc G (ed) *Phospholipids Handbook*, pp 687–740. New York: Marcel Dekker.

Watts A (1998) Solid state NMR approaches for studying the interaction of peptides and proteins with membranes. *Biochimica et Biophysica Acta* **1376**: 297–318.

Watts A (1999) NMR of drugs and ligands bound to membrane receptors. *Current Opinion in Biotechnology* **10**: 48–53.

# Mercury NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# Metastable Ions

**John L Holmes**, University of Ottawa, Ontario, Canada

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Theory |

## Metastable ions, their origins and observation

Metastable ions are those that dissociate en route from an ion source, through a mass analyser to an ion detection device. That ions can dissociate in flight was recognized by Hipple and Condon in 1946 when they provided the explanation for the presence of the diffuse peaks that they had observed in normal, electron ionization (EI) mass spectra obtained with a single-focusing magnetic sector mass spectrometer. It should be emphasized that the term metastable refers only to ions that are able to fragment in flight by virtue of internal energy that they acquired within the ion source, not after acceleration therefrom, nor by collisions with a target gas, nor by radiative excitation as they traverse the apparatus. They are thus unimolecular dissociations.

It was the striking and varied phenomenology produced by metastable ions that made them of particular interest in the 1960s and 1970s, when many laboratories had magnetic sector (B) instruments or double-focusing mass spectrometers consisting of an electric sector (E) and a magnetic sector arranged BE or EB (tandem mass spectrometers). Much of this article will describe observations made with tandem mass spectrometers.

First, however, consider a simple magnetic sector instrument. Singly charged molecular or fragment ions leaving the ion source will all have the same translational energy,

$$eV_{\text{acc}} = \tfrac{1}{2}\,M_1{v_1}^2 = \tfrac{1}{2}M_2{v_2}^2 = \tfrac{1}{2}M_n{v_n}^2$$

independent of their mass $M_n$. The ion acceleration

potential is $V_{acc}$ and $v_n$ are the ion velocities. As the magnetic analyser field strength is changed, ions of different mass will be transmitted to the detector giving rise to the normal mass spectrum, a series of sharp, well-defined signals for each *m*/*z* value for the ions present. If, however, an ion $M_1^{\bullet+}$ is metastable and fragments in flight

$$M_1^{\bullet +} \rightarrow M_2{}^+ + M_3^{\bullet}$$

the fragment ion $M_2^+$ will have less translational energy than an ion $M_2^+$ produced in the ion source. The fragment ion's translational energy is $\frac{1}{2}M_2v_1^2$ and is equal to $\frac{1}{2}(M_2^2)v_2^2$ given that for source ions $\frac{1}{2}M_1v_1^2 = \frac{1}{2}M_2v_2^2$. The fragment ion $M_2^+$ from metastable $M_1^{\bullet+}$ will thus appear in the normal mass spectrum at an apparent mass $(M_2^2/M_1)$.

For sector mass spectrometers, $V_{acc}$ is typically 2–10 kV and the timescale for ions to pass from ion source to detector is typically in the range 5–50 µs for ions of *m*/*z* ~ 100.

The signals at $(M_2^2/M_1)$ were observed in the majority of cases to be much wider than those of normal fragment ions and were typically of Gaussian profile.

Note that the presence of these diffuse peaks at nonintegral *m*/*z* values provides the observer with direct proof that a given fragmentation has indeed taken place, e.g.

$$\begin{array}{ll} C_5H_5{}^+ \rightarrow C_3H_3{}^+ + C_2H_2 \\ m/z\,65 \quad m/z\,39 \end{array}$$

This metastable dissociation of $C_5H_5^+$ ions will produce a peak at apparent mass 23.4 Da. Clearly these phenomena can (and do) assist in the interpretation of mass spectra by unequivocally linking a fragment ion with a specific precursor.

The reason for the diffuse shape, i.e. Gaussian profile, of metastable peaks is that upon dissociation a fraction of the internal energy of $M_1^{\bullet+}$ is converted into translational degrees of freedom of the products $M_2^+$ and $M_3^{\bullet}$. This kinetic energy is released isotropically and results in the broadening of the metastable peak relative to the width of a beam of undissociating ions from the source. In a tandem mass spectrometer of BE geometry, under conditions of good energy resolution, the shape of a metastable peak may be examined in detail and physicochemical information derived therefrom. In such a tandem mass spectrometer (shown in **Figure 1**) the electromagnet (B) is used to select an ion $M_1^{\bullet+}$ of interest. Under normal operating conditions the electric sector potential ($E_s$) is set to transmit all the ions from the ion source. If however, $M_1^{\bullet+}$ is metastable during its flight through the (second) field-free region between the two analysers, the fragment ion $M_n^+$ (or ions) will only be transmitted through the electric sector if its potential is reduced in the ratio $M_n(E_s)/M_1$. Scanning this potential from zero to $E_s$ will transmit all fragments from metastable $M_1^{\bullet+}$ ions at appropriate $M_n(E_s)/M_1$ values. Such a spectrum is known as a mass-selected-ion kinetic energy spectrum (MIKE).

## Metastable peak shapes and their significance: kinetic energy release

As stated above, the majority of metastable peaks have a Gaussian-type profile (**Figure 2A**). Indeed many peaks can be described by a simple variant of the Gaussian formula, $h = \exp(-\alpha w^2)$ (where $h$ is the fraction of the peak height at which the width, $w$, is

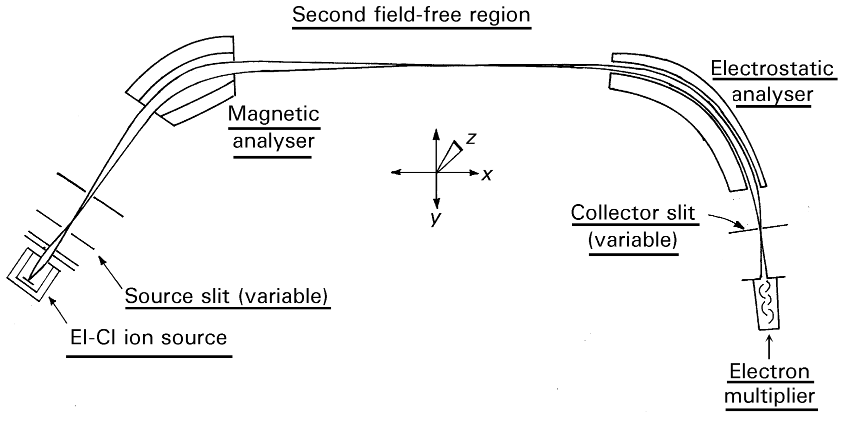



**Figure 1** Schematic drawing of a tandem mass spectrometer of BE geometry.

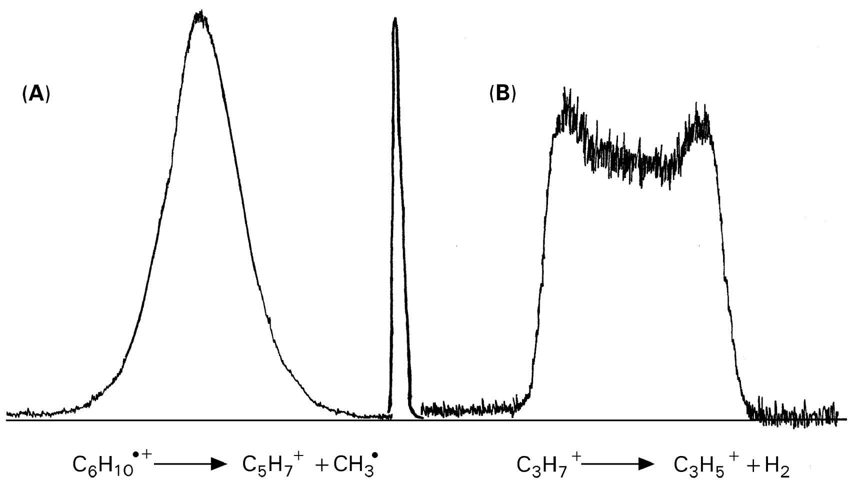



**Figure 2** Simple Gaussian (A) and 'dished' (B) metastable peaks. For comparison a typical profile for an undissociating ion beam is shown between them.

measured and $\alpha = \ln 2$) where the index 2 is replaced by $n$, taking on a range of values from about 1.4 to 2.1. Note that the distribution of released energies derived from the exact Gaussian function is Maxwell–Boltzmann in form.

A second class of peak shapes are the so-called flat-topped or 'dished' metastable peaks (**Figure 2B**). The dish in such peaks is instrument-dependent and arises from the magnitude of the kinetic energy release (KER) and its distribution (see below). As the peak is recorded by sweeping the electric sector voltage over the appropriate range, all fragment ions having large components of translational energy released in the $x$- and $y$-axial directions (see **Figure 1**) will be recorded. Those having large $z$-axial components may, however, be unable to pass the collector slit because of its finite height. These 'lost' ions result in the observed dish. The effect is only important for dissociations in which the magnitude of the KER is large (see below).

The relationship between the metastable peak characteristics and the magnitude of the KER ($T_h$) is given by

$$T_h = \frac{M_1{}^2}{16\, M_2 M_3 E_s}\left(W_h{}^2 - W_m{}^2\right)$$

where $W_h$ is the width of the energy-resolved metastable peak at fractional height $h$, measured in electric sector volts; $W_m$ is the corresponding width of the $M_1^{\bullet+}$ ion beam (also in sector volts). It has become customary to report KER values corresponding to half-height peak widths, $T_{0.5}$, they are usually given in meV. Under good energy-resolution conditions

$$\frac{W_{m(0.5)}}{W_{0.5}} < 0.1$$

making the main beam correction small to negligible.

For Gaussian-type peaks, $T_{0.5}$ values lie in the range 0–50 meV, while for flat-topped peaks, $T_{0.5}$ is from ~200meV to several volts.

It is believed that the great majority of Gaussian-type metastable peaks arise from ion fragmentations with no reverse energy barrier, $E_{rev}$ (**Figure 3**),

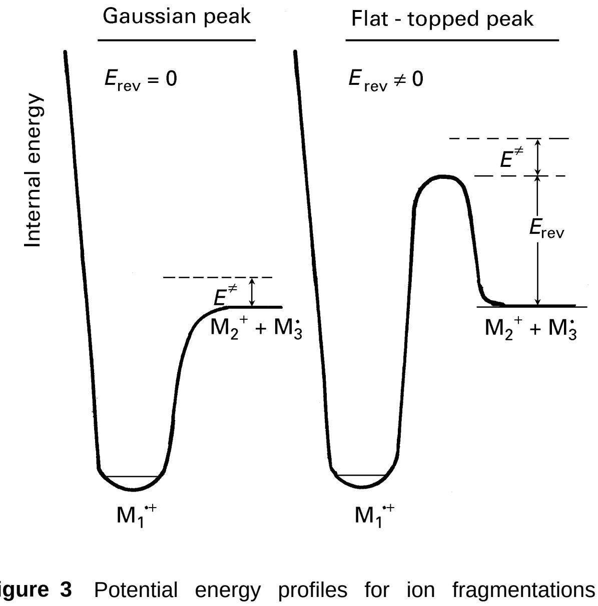



**Figure 3** Potential energy profiles for ion fragmentations giving rise to Gaussian and flat-topped metastable peaks.

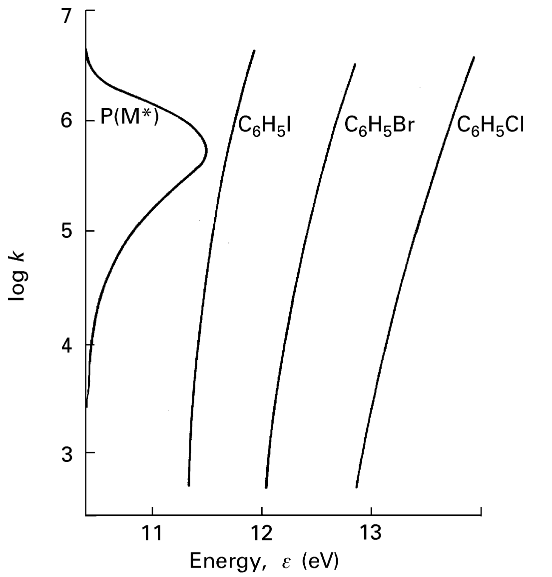



**Figure 4** Log $k$ vs. $\varepsilon$ curves for the dissociation of the phenyl halides $C_6H_5X^{\bullet+} \rightarrow C_6H_5^+ + X^\bullet$. The metastable ion window is shown as P(M*). This curve represents the probability that an ion having the corresponding dissociation rate constant, $k$, will fragment in the metastable ion time-frame, here ~1–3 μs. At the maximum, about 11% of the beam is dissociating.

whereas flat-topped peaks are associated with a finite or large $E_{rev}$. This is also illustrated in **Figure 3**. The origin of the KER for the Gaussian peak is the excess energy, $E^\ddagger$, above the threshold for the fragmentation necessary to achieve the rate constant ($k$) for the process; typically $k$ is ~$10^5$ $s^{-1}$. For the dissociation with the reverse energy barrier, $E_{rev}$ (a fixed energy) is augmented by $E^\ddagger$ (sometimes referred to as the non-fixed energy).

The range of rate constants ($k$) accessed in the metastable time frame is easily calculated using the first-order rate equation. Direct experimental measurements of the way in which a fragmentation's rate varies with internal energy ($E$) are made by 'photoelectron photoion coincidence' methods. A typical plot of ln $k$ vs. $\varepsilon$ curves with the metastable ion window included is shown in **Figure 4**. In general, the fraction of $E^\ddagger$ that is released as translational energy is small, about 3–10%. For dished peaks, much larger fractions of $E_{rev}$ may be released.

Composite metastable peaks are also known and these can arise from the following situations:

(a) A single ion structure $M_1^{\bullet+}$, dissociates metastably via two competing reactions to give isomeric fragment ions.
(b) A given ion $M_1^{\bullet+}$, has two stable isomeric forms (in an ion source) which metastably fragment via different transition states either to give fragment ions of different structure or to give fragment ions having only one structure.

Two examples are shown in **Figure 5**.

## Collision induced dissociations

The early experiments on collision induced dissociations (CID) were performed using sector instruments such as that shown in **Figure 1**. A small constricted region or cell within the second field-free region close to the ion beam focus, with a diffusion pump placed thereunder, has a target gas admitted to it. The differential pumping of the region is such that

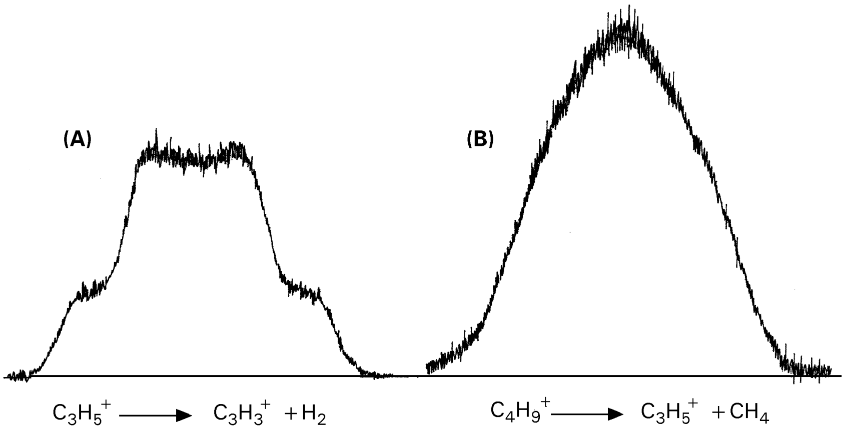



**Figure 5** Composite metastable peaks. (A) Peak for $C_3H_5^+$ ($m/z$ 41) ions dissociating to give isomeric fragment ions $C_3H_3^+$ ($m/z$ 39). The broad signal corresponds to the production of the cyclopropenium cation, while the narrower component is for the generation of the propargyl isomer $[HCCCH_2]^+$. (B) Signal for isomeric $C_4H_9^+$ ions, fragmenting via different transition states, to produce by methane loss the 2-propenyl cation, $[CH_3^+CCH_2]$ (not the thermochemically more stable allyl isomer $[CH_2CHCH_2]^+$).

very little gas diffuses into the high-vacuum region outside the cell. The cell, 2–3 cm long, is only a small zone within the second field-free region, ~1 m long. When sufficient gas has been admitted to reduce the intensity of the mass-selected ion beam by 10% (a condition corresponding to predominantly single collisions) the ion kinetic energy spectrum is recorded. Unlike the metastable-ion (MI) mass spectrum, the CID mass spectrum contains many signals and is similar in appearance to the normal mass spectrum of $M_1^{\bullet+}$ (with $M_1$ being a stable neutral molecule).

The peaks in this CID mass spectrum are all broadened by kinetic energy release, making the spectrum rather poorly resolved. Nevertheless, as described elsewhere, the CID mass spectrum can be used as an invaluable tool in ion structure assignments. If a charge of say 500 V is placed on the collision cell (necessitating its electrical isolation from the apparatus), then MI and CID peaks may be separated. MI peaks remain undisplaced, whereas all CID peaks are shifted to higher energy (positive charge) or lower energy (negative charge).

Note that poor background pressures in the field-free regions can lead to unwanted CID contributions to MI mass spectra. In general, more than four peaks in a MI mass spectrum may be regarded as showing possible high vacuum failure. A consideration of the likely energy requirements for the observed fragmentations may also help to resolve this problem.

## Metastable ions and ion structures

In the mid 1960s the first semiquantitative relationship between metastable-peak intensities and the structure of a fragmenting ion was proposed. This so-called ‘metastable-peak abundance-ratio test’ was based on the premise that when two (or more) competing fragmentations from the same ion give reasonably intense metastable peaks, then the ratio of their abundances may be used as a criterion for ion structure. Thus, if dissociating ions of a given molecular formula and *m/z* ratio, derived from a variety of precursor species, all display the same metastable peaks having similar relative abundances, then it can reasonably (but tentatively) be concluded that the fragmenting ions have the same structure. It is also possible that the fragmenting species consist of mixtures of common composition but having different structures.

The chief problem with this simple criterion was to decide what constitutes a ‘significant’ variation in metastable peak abundances. Indeed it was quickly realized that the ratio will be susceptible to the distribution of internal energies among the fragmenting ions. Notwithstanding these difficulties, it was observed that many even-electron hydrocarbon cations derived from a wide range of precursor molecules had closely similar metastable-peak abundance ratios, showing the ease with which isomeric hydrocarbon cations can rearrange prior to their metastable fragmentation. A good example are the $C_6H_{10}$ isomers, 30 of which have been examined; they all, irrespective of their initial structure, produce the cyclopentenyl, $C_5H_7^+$, cation when their molecular ions lose a $CH_3^{\bullet}$ radical. The metastable ion behaviours of the $C_5H_7^+$ fragment ions are identical.

A more stringent test for ion structure comes from the examination of energy-resolved metastable peaks. **Figure 6** shows the signals for loss of a hydrogen atom from the isomeric $C_2H_4O^{\bullet+}$ ions, $CH_3CHO^{\bullet+}$, cy-$CH_2OCH_2^{\bullet+}$ and $CH_2CHOH^{\bullet+}$. The shapes are clearly structure-characteristic and so could, for example, serve as identifiers for $C_2H_4O^{\bullet+}$ *m/z* 44 fragment ions generated from different precursor molecules. Thus, for example, it has been shown that ionized 2-propanol, $CH_3CH(OH)CH_3^{\bullet+}$, loses $CH_4$ by two competing eliminations, one yielding $CH_3CHO^{\bullet+}$, and the other $CH_2CHOH^{\bullet+}$. The metastable peak for H loss from *m/z* 44 being composed of the two right-hand signals in **Figure 6**, the narrow component sitting atop the broad.

It is worth noting however, that the above observations are phenomena that only allow the above $C_2H_4O^{\bullet+}$ ions to be distinguished. The details of the physical chemistry of the H-loss processes are known and, for example, the structure from which $CH_2CHOH^{\bullet+}$ ions actually lose their H atom is the (uncommon) $C_2H_4O^{\bullet+}$ isomer, the carbene ion $CH_3^{\bullet+}COH$.

## Isotopic labelling and metastable peaks

In studying ion fragmentation mechanisms, particularly molecular eliminations, it is important to discover which atoms are lost in the process. The use of isotopic labels to solve such problems is a long-established practice in organic chemistry. In general, the effect of label atoms on a normal EI mass spectrum does not provide definitive evidence for a reaction mechanism; this results from the possible multiple origins of fragment ions, especially those of lower *m/z* ratio. However, as described above, a metastable peak exactly defines an ion dissociation pathway and so here isotopic labelling has great utility. In particular, if MIKE mass spectra are examined, the incorporation of label in the sample molecule need not be complete; mass selection

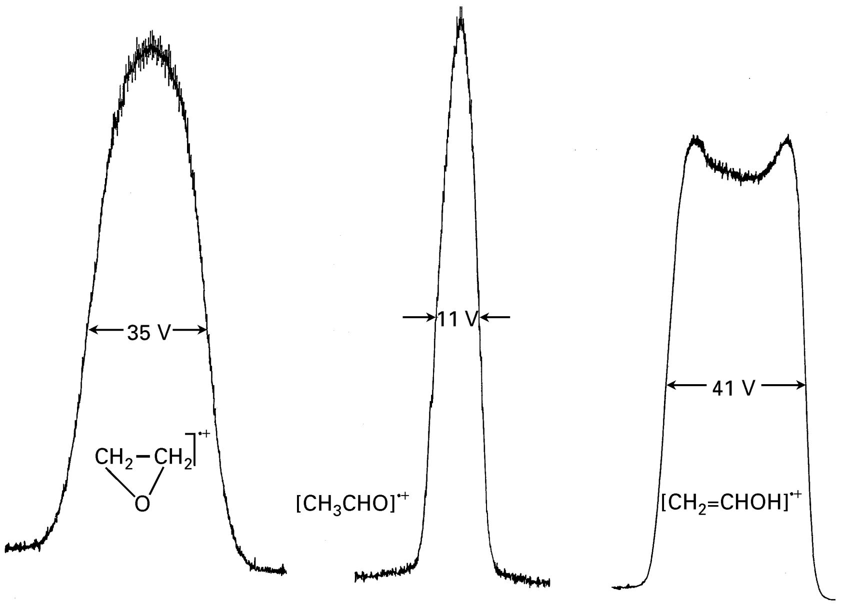



**Figure 6** The energy-resolved metastable peaks for the loss of H• from three isomeric $C_2H_4O^{\bullet+}$ ions. Note that the peak widths do not have a common scale. The half-height widths for the three signals are 35, 11 and 41 V from left to right.

allows the isolation of an ion containing the desired isotope(s). This is useful when for example, H,D exchange is performed in the mass spectrometer inlet system using $D_2O$ to exchange hydroxyl, amino or even α-to-keto hydrogen atoms.

An important result of labelling experiments on metastable ions was the demonstration that in many species, especially hydrocarbon cations, complete loss of the positional identity of H and C atoms preceded fragmentation. This, of course, obscures all reaction pathways and so structures of fragmenting ions and their products have to be determined by other routes. Such statistical loss of labelled atoms is referred to as 'atom scrambling'. The butyl cations, $C_4H_9^+$, can serve as an example. $CD_3(CH_3)_2C^+$ loses methane as $CD_3H$, $CD_2H_2$, $CDH_3$ and $CH_4$ in the statistical random ratio of 6!/5! to 3!6!/2!4!2! to 3!6!/2!3!3! to 6!/4!2! or 2:15:20:5 and $(CH_3)_3{}^{13}C^+$ loses $^{13}CH_4$ to $^{12}CH_4$ in the ratio of 1:3. All $C_4H_9^+$ isomers behave similarly, showing that at energies below the dissociation limit for $CH_4$ loss, many isomeric structures are accessed. Isotopic labelling and metastable ion studies can also reveal unexpected mechanisms, hidden in the normal mass spectrum. Two examples should suffice. The mass spectrum of benzoic acid, $C_6H_5COOH$, is very simple with the only major fragments corresponding to simple bond cleavages; e.g. $C_6H_5COOH^{\bullet+}$, $C_6H_5{}^+CO$ and $C_6H_5{}^+$, are the dominant signals. Metastable $C_6H_5COOH^{\bullet+}$ ions lose $OH^\bullet$ as expected, but astonishingly, metastable $C_6H_5COOD^{\bullet+}$ ions lose $OD^\bullet$ *and* $OH^\bullet$, in the ratio of 1:2. That a hidden exchange between carboxyl and the two *ortho* H atoms has taken place was shown by metastable *ortho*-dideuterio benzoic acid ions losing $OD^\bullet$ and $OH^\bullet$ in exactly the inverse ratio.

The unexpected complexity of ion rearrangement processes is also well exemplified by the metastable loss of HCN from ionized benzonitrile $C_6H_5CN^{\bullet+}$. Here, only some 5–7% of the nitrile carbon is present in the HCN eliminated. The participation of linear isomers of the aromatic nitrile ion has been demonstrated.

## Neutral species lost in metastable ion fragmentations

When metastable ions fragment, a neutral radical or molecule is generated. In most circumstances the identity of this neutral is not problematic and for example, is commonly a small molecule, $H_2O$, $CH_4$, CO, $C_2H_4$ or free radical $CH_3^\bullet$, $OH^\bullet$, $C_2H_5^\bullet$, $HC^\bullet O$, etc. The distinction between isobaric species (e.g. CO, $C_2H_4$, $C_2H_5^\bullet$, $HC^\bullet O$) may be made by the simple experiment described below.

If an ion-beam deflector electrode is placed in front of the collision cell in the second field-free region of a sector mass spectrometer (see **Figure 1**), then all mass-selected ions and their charged fragments may be deflected from the beam path, and only the neutral products from ion fragmentation

will enter the collision cell. Neutral fragments having kV translational energies are readily ionized by collision with a target gas (e.g. He or $O_2$), producing the characteristic mass spectrum of the neutral species.

This technique has, for example, shown that HNC rather than HCN is lost from metastable aniline molecular ions and that ionized methyl acetate loses $^{\bullet}CH_2OH$ radicals rather than $^{\bullet}OCH_3$ as the lowest-energy fragmentation route.

## List of symbols

$e$ = electron charge; $E$ = internal energy; $E_s$ = electric sector potential; $E_{rev}$ = reverse energy barrier; $E^{\ddagger}$ = excess energy; $h$ = fraction of the peak height at which $w$ is measured; $T_{0.5}$ = kinetic energy release calculated using the metastable peak width at half-height; $v_n$ = ion velocities; $V_{acc}$ = ion acceleration potential; $w$ = peak width; $W_h$ = width of energy-resolved metastable peak at fractional height $h$; $W_m$ = width of the $M_1^{\bullet+}$ ion beam; $\alpha = \ln 2$.

*See also:* **Fragmentation in Mass Spectrometry; Ion Energetics in Mass Spectrometry; Ion Structures in Mass Spectrometry; Isotopic Labelling in Mass Spectrometry; Neutralization-Reionization in Mass Spectrometry; Photoelectron-Photoion Coincidence Methods in Mass Spectrometry (PEPICO); Sector Mass Spectrometers.**

## Further reading

Burgers PC, Holmes JL, Mommers AA, Szulejko JE and Terlouw JK (1984) Collisionally induced dissociative ionization of the neutral products from unimolecular ion fragmentations. *Organic Mass Spectrometry* **19**: 442–447.

Cooks RG, Beynon JH, Caprioli RM and Lester GR (1973) *Metastable Ions*. Amsterdam: Elsevier, 1–296.

Derrick PJ and Donchi KF (1983) Mass spectrometry. In: Bamford CH and Tipper CFH (eds) *Comprehensive Chemical Kinetics* Vol 24, pp 53–247. Amsterdam: Elsevier.

Hipple JA, Fox RE and Condon EU (1946) Diffuse signals in mass spectra. *Physical Review* **69**: 347–351.

Holmes JL (1985) Assigning structure to ions in the gas phase. *Organic Mass Spectrometry* **20**: 169–183.

Holmes JL and Benoit FM (1972) Metastable ions in mass spectrometry. In: Maccoll A (ed) *MTP International Review of Science, Physical Chemistry, Series One*, Vol 5, pp 259–300. London: Butterworth.

Holmes JL and Terlouw JK (1980) The scope of metastable peak studies. *Organic Mass Spectrometry* **15**: 383–397.

Molenaar-Langeveld TA, Fokkens RH and Nibbering NMM (1986) An unusual pathway for the elimination of HCN from ionized benzonitrile. *Organic Mass Spectrometry* **21**: 15–22.

# Microwave and Radiowave Spectroscopy, Applications

**G Wlodarczak**, Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, France

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

Microwave spectroscopy covers, typically, the frequency range from 1–2 GHz to several THz. In this spectral region the rotational spectra of molecules which possess a permanent (or induced) dipole moment can be observed. The analysis of these spectra gives information on the geometrical structure of the molecule and the centrifugal distortion effects. Other information such as dipole moment, quadrupole coupling constants, spin rotation constants can also be determined. In the case of molecules that present large amplitude motions, like the internal rotation of a methyl group or ring deformation for cyclic compounds, the potential barrier is also accessible from the observation of line splittings. The large selectivity of microwave spectroscopy allows the diagnostic of interstellar medium, planetary atmospheres and minor components of the earth's atmosphere.

The number of molecules which have been studied by this method is increasing every year. Their complexity ranges from diatomic molecules to moderate size molecules (up to 30 atoms). Unstable species such as molecular ions, radicals, reactive molecules, molecular complexes, etc. can also be studied. The resulting molecular constants are regularly the subject of extensive compilations.

## Spectrum measurements

Microwave spectroscopy uses tunable coherent sources of radiation such as microwave synthetizers, solid state oscillators (Gunn diodes) or electronic tubes (klystrons). These oscillators can be operated in their fundamental mode (up to 120 GHz) but harmonic generation is commonly realized with frequency multipliers up to 500 GHz, and has been used to reach 1 THz on occasions. Backward wave oscillators are available up to 1.2 THz in their fundamental mode. **Figures 1** and **2** show typical rotational spectra recorded with this type of sources. Different techniques can be used to work in the THz region:

- far-infrared laser tunable sideband generation,
- mixing of two infrared radiations ($CO_2$ lasers) and a microwave tunable radiation,
- photomixing of two diode lasers,
- far-infrared Fourier transform spectroscopy.

The detectors are generally Schottky diodes or InSb helium-cooled bolometers in the millimetre and submillimetre wave region.

Stark modulation has been widely used at microwave frequencies, while source modulation is the most common technique for millimetre and submillimetre spectroscopy. Recently a new technique, microwave Fourier transform spectroscopy, has been developed to obtain increased frequency resolution and sensitivity. A macroscopic polarization is created in the sample by a microwave pulse of appropriate strength and duration. The sample emits then a signal which decreases in time owing to relaxation

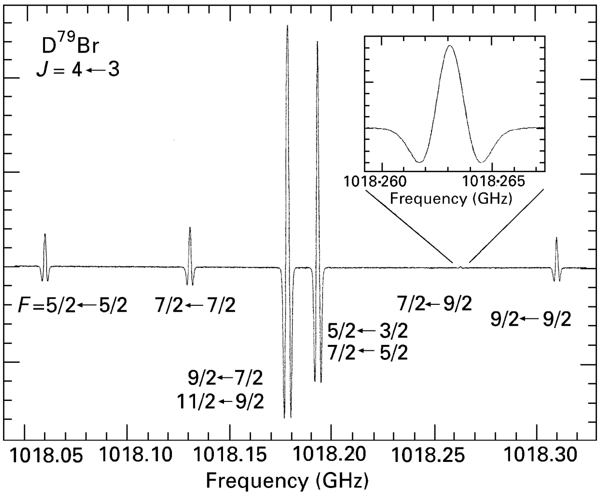



**Figure 2** A 32-MHz scan showing the $^{79}$Br hyperfine structure of the $J = 4 \leftarrow 3$ transition of D$^{79}$Br at 1.018 THz. Reproduced with permission from Saleck AH, Klaus T, Belov S and Winnewisser G (1996) THz rotational spectra of HBr isotopomers in their v = 0,1 states. *Z. Naturforsch* **51A**: 898.

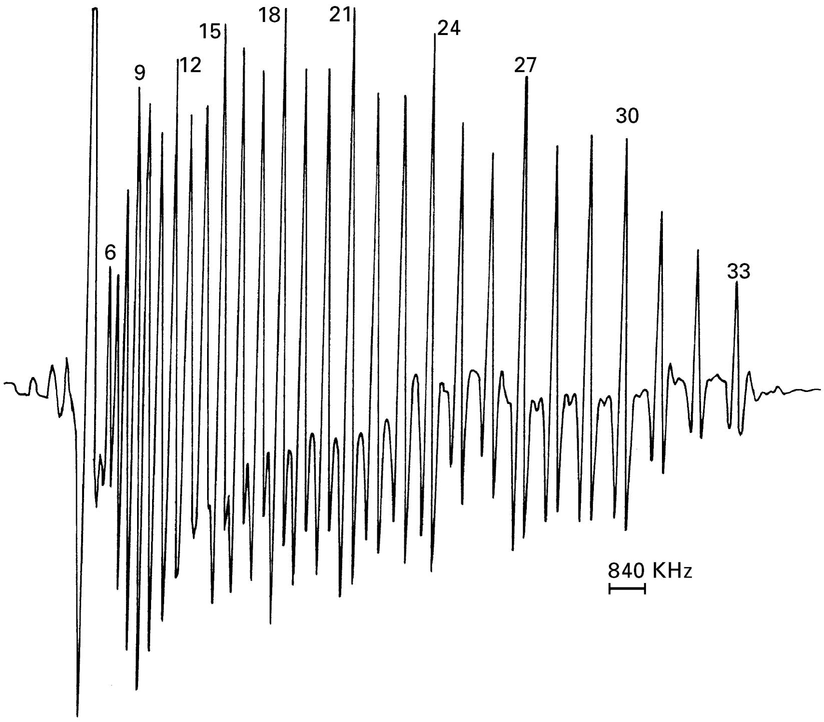



**Figure 1** The $J = 34 \leftarrow 33$ transitions of trioxane [$(H_2CO)_3$] at 358 GHz. The $K = 3n$ transitions are easy to recognize and their quantum number $K$ is indicated. Reproduced with permission from Gadhi J, Wlodarczak G, Boucher D and Demaison J (1989) The submillimeter-wave spectrum of trioxane. *Journal of Molecular Spectroscopy* **133**: 406.

effects. The Fourier spectrum of this decay signal contains the frequencies corresponding to rotational transitions of the sample. This technique is available between 1 and 40 GHz.

The sample is prepared in the gas phase, at low pressures (typically 10–30 mtorr) for stable or not too unstable species. Reactive species are prepared inside a free space cell, usually made of a pyrex glass tube, by different techniques: microwave discharge, a.c. or d.c. glow discharge, thermal decomposition or reaction, photochemical decomposition. This last process needs high power excimer or $CO_2$ lasers, but is much cleaner than electrical discharges. **Figure 3** shows spectral lines of linear $C_3H$, produced in a glow discharge of $C_2H_2$, CO and $H_2$ in a cell cooled to liquid nitrogen temperature to obtain stronger signals. Molecular ions can also be observed within electrical discharges: the sensitivity of the spectrometer is then increased by using the velocity modulation technique. This technique is based on the Doppler shift of ion spectra caused by the discharge electric field. An additive magnetic field usually produces an enhancement of the spectrum. For van der Waals or hydrogen-bonded complexes a supersonic expansion is generally needed. The monomers are diluted (several %) in rare gases (usually argon). The collisionless regime in molecular beams prevent the rapid destruction of the molecular complexes. **Figure 4** represents part of the spectrum of the phenol–water complex.

The isotopic species can be studied in natural abundance for $^{13}C$, $^{15}N$ and most of the elements. In the case of deuterated isotopomers an enriched sample is generally needed. The line frequency measurements are made with a very high accuracy: 1 kHz for microwave Fourier transform measurements between 2 and 20 GHz to 50–300 kHz for far-infrared measurements up to 2–3 THz. The line shape is usually a Doppler or a Voigt profile if the collisional broadening becomes important. The Doppler half-maximum halfwidth is given by $\Delta\nu_D$ (MHz) = $3.581\times10^{-7}$ $(T/M)^{1/2}$ $\nu_0$, where $T$ is the temperature in K, $M$ the molecular weight in atomic mass units, $\nu_0$ the frequency of the transition in MHz. For example the $J$ = 3–2 transition of CO at 345.8 GHz has a Doppler width of 400 kHz.

The collisional broadening increases linearly with pressure: $\Delta\nu_L = \gamma_L P$, where $\Delta\nu_L$ is the collisional half-maximum halfwidth. Typical values for collisional

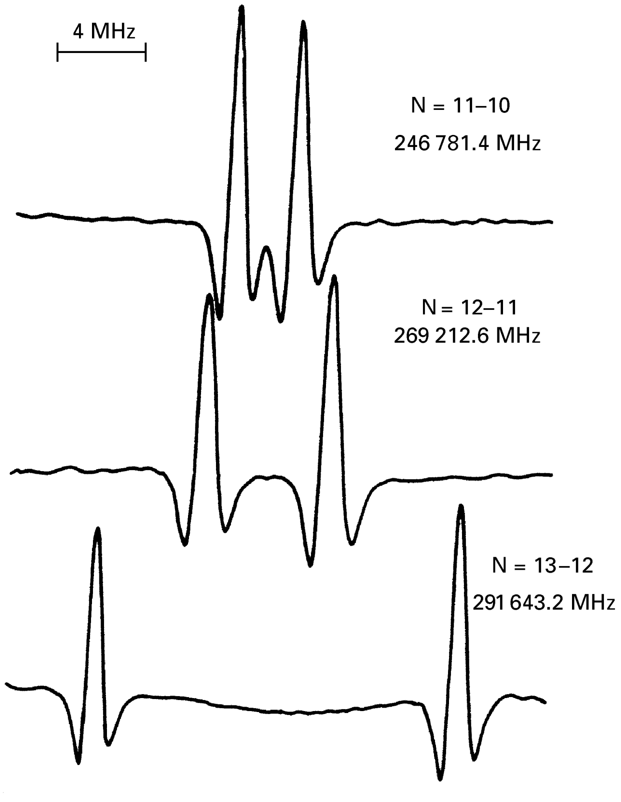



**Figure 3** Typical traces of the spectral lines for the $\nu_4$ ($^2\Sigma\mu$) state of $C_3H$ (CCH bending mode). The centre frequencies of the doublet lines are given in the figure. The spin splitting becomes large as *N* increases. Reproduced with permission from Yamamuto S, Saito S, Suzuki H *et al* (1990) Laboratory microwave spectroscopy of the linear $C_3H$ and $C_3D$ radicals and related astronomical observation. *Astrophysical Journal* **348**: 363.

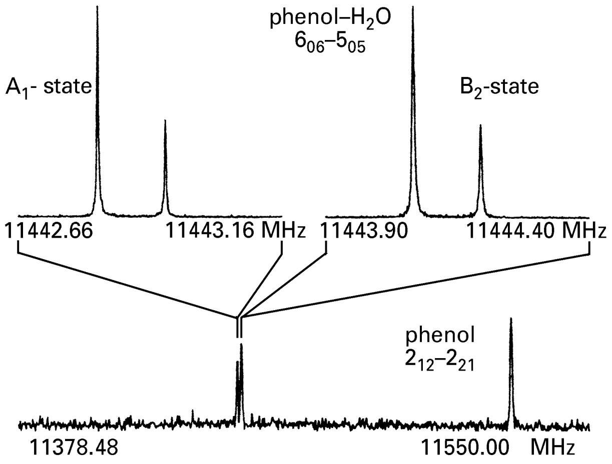



**Figure 4** A 172 MHz broad-band scan (lower trace) and high-resolution spectrum (upper trace) of phenol and water in helium at a stagnation pressure of 100 kPa. Three lines can be recognized in the low-resolution spectrum: a doublet consisting of a strong and a weak component of the $6_{06} \leftarrow 5_{05}$ transition of the phenol–$H_2O$ complex and the low frequency component of the $2_{12} \leftarrow 2_{21}$ internal rotation doublet of the phenol monomer. In the high-resolution spectrum, the lines appear as Doppler doublets. Reproduced with permission from Gerhards M, Schmitt M, Kleinemans K and Stahl W (1996). The structure of phenol–water obtained by microwave spectroscopy. *Journal of Chemical Physics* **104**: 967.

broadening parameters $\gamma_L$ are 1–10 MHz torr$^{-1}$, as shown in **Figure 5**; the value of $\gamma_L$ generally increases when the temperature decreases. At low pressures the line broadening is mostly due to the Doppler effect and is the general cause which limits the resolution of the spectrometer. The resolution can be increased by using sub-Doppler techniques such as saturated absorption spectroscopy (or Lamb dip spectroscopy) or by generating a microwave radiation perpendicular to the motions of the molecules when using a molecular beam.

Most often, molecules are studied in their ground electronic and vibrational state but rotational spectra in vibrationally excited states (up to 1000–1500 cm$^{-1}$) are commonly observed. **Figure 6** shows a stick diagram, reproducing the relative intensities of the lines, of the rotational spectrum of $C_3S$ in the ground and various excited states: the vibrational energies for $v_5 = 1$ and $v_4 = 1$ are, respectively, 150 and 490 cm$^{-1}$. The observation of rotational spectra in excited electronic states is more difficult because of their short lifetime.

## Rotational constants and geometrical structure

Most of the spectra recorded in the microwave region are rotational spectra. Among the molecular parameters used for the modelling of these spectra the most important are rotational constants. These constants are related to the principal moments of inertia by the relation $A = h/(8\pi^2 I_a)$. For light molecules such as hydrides, the spectrum lies in the millimetre or submillimetre wave region: for HCl the $J = 1$–0 transition, which is the lowest frequency transition, lies near 626 GHz. This is one of the reasons for the development of spectroscopy in the

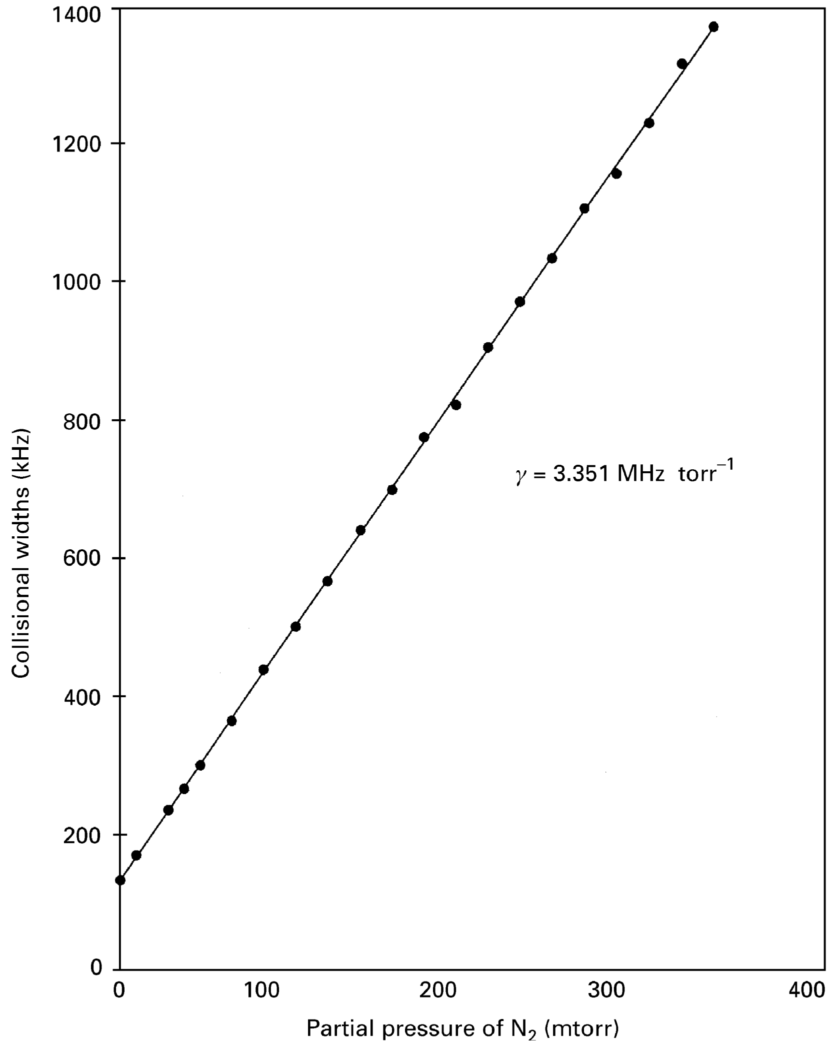



**Figure 5** Nitrogen-broadened widths of the $J = 8 \leftarrow 7$ line of $N_2O$ at $T = 295$ K; pressure of $N_2O$: 30 mtorr.

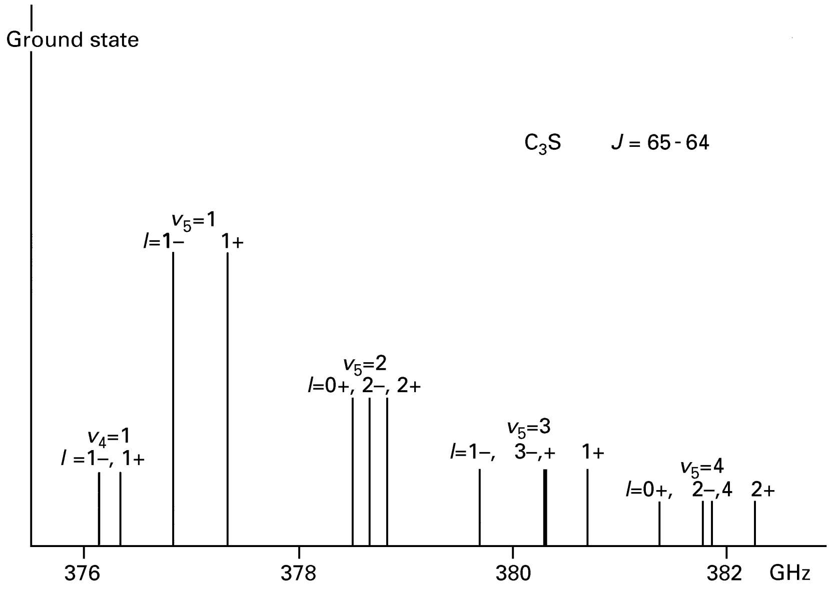



**Figure 6** Spectral pattern and relative intensities of $C_3S$ transition lines. Reproduced with permission from Tang J and Saito S (1995) Microwave spectrum of the $C_3S$ molecule in the vibrationally excited states of bending modes $v_4$ and $v_5$. *Journal of Molecular Spectroscopy* **169**: 92.

terahertz domain. For heavy molecules the spectrum lies mostly in the microwave region. Nevertheless high $J$ transitions are measured at millimetre and submillimetre wavelengths. This allows the analysis of the centrifugal distortion. The rotational Hamiltonian is generally expressed as a series expansion of the even powers of the angular momentum operator. The first term contains the rotational constants, the following terms contain, respectively, the quartic, sextic, octic, etc. centrifugal distorsion constants. The effects of these terms are important at high rotational excitation. This expansion is usually satisfactory except for floppy molecules, i.e. molecules which exhibit large amplitude motions ($H_2O$, $CH_3OH$, etc.). Quartic and sextic centrifugal distortion constants are usually determined during the analysis of the spectrum on a broad frequency range. They can be related to the force field and be estimated from *ab initio* calculations. Some empirical correlations were also found between these constants and the rotational constants: this allows also an estimation which can be useful at the beginning of the identification of a new spectrum. **Figure** 7 shows an example of correlation found between a sextic centrifugal distortion constant and the rotational constant $B$ for a class of symmetric top molecules.

The rotational constants are a source of information on the geometrical structure of the molecule. By combining the rotational constants of a parent molecule (usually the main species) and isotopically substituted daughter molecules it is now possible to determine a complete experimental structure which lies very close the equilibrium structure. These results mainly concern moderately sized molecules, containing between 3 and 8 atoms. A pure experimental equilibrium structure is difficult to obtain because the rotational constants in all fundamental

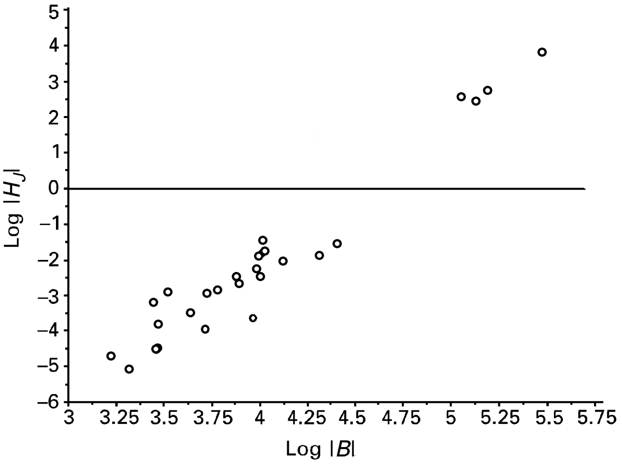



**Figure 7** Plot of log | $H_J$ | against the function of log $B$ for various $C_{3v}$ symmetric tops: $B$ is the rotational constant, $H_J$ is one sextic centrifugal distortion constant. Reproduced with permission from Demaison J, Bocquet R, Chen WD, Papousek D, Boucher D and Bürger H (1994) The far-infrared spectrum of methylchloride: determination and order of magnitude of the sextic centrifugal distortion constants in symmetric tops. *Journal of Molecular Spectroscopy* **166**: 147.

vibrational states are needed to determine the equilibrium rotational constants. For the higher excited states the data are usually obtained from high resolution IR spectroscopy. In many cases a lot of interactions occur between these vibrational states (Fermi resonances, Coriolis interactions, etc.) which make it difficult to obtain unperturbed rotational constants.

Another type of problem which arises in structure determination is the ill-conditioning of the system of equations that link the moments of inertia and the structural parameters. To obtain reliable parameters we have to incorporate additional data obtained from other techniques: electron diffraction, *ab initio* calculations, empirical relations, etc. Bond distances with an accuracy of 0.001 Å and bond angles with an accuracy of 0.2° can then be determined.

For molecules that contain a greater number of atoms, different conformers are present and generally observed: their spectra are rather well separated and the comparison between their intensities gives an order of magnitude of their relative energy difference.

## Determination of molecular parameters

Rotational constants and centrifugal distortion constants are not the only parameters accessible by microwave spectroscopy. In the case of linear and symmetric tops, doubly degenerate vibrationally excited states (bending modes for examples) are present and their spectra are slightly more difficult to analyse. More parameters are needed: l-type doubling constants, rotation–vibration interactions constants, etc. It should be noted that for symmetric tops the axial rotational constant is not accessible by pure rotational spectroscopy. Its determination remains a difficult problem, and needs the study of IR spectra (including hot bands and combination bands). Combined high-resolution microwave and IR studies have become the best way to determine a coherent set of molecular parameters for linear and symmetric tops. Owing to the growth in high quality experimental data, theoretical developments on the rovibrational Hamiltonian of symmetric tops have been made. The equivalence between different reductions of this Hamiltonian has been checked experimentally.

Large amplitude motions produces specific features in rotational spectra. Internal rotation of one methyl group induces a splitting of most of the lines into two components. This splitting depends on the barrier height and the geometry of the molecule. For high barriers the splitting is usually small but easily resolved by microwave Fourier transform spectroscopy. The splittings are larger in the excited torsional states but a complete theoretical description of the spectra has not yet been achieved. A prototype molecule in this field is acetaldehyde $CH_3CHO$.

Cyclic molecules containing an heteroatom in the cycle (O, S, etc.) also present large amplitude motions due to ring deformations, such as ring-puckering. The observation of rotational spectra in excited states is not a problem because these states are low in energy.

The behaviour of the rotational and centrifugal distortion constants with the vibrational quantum number can be related to the potential of the ring deformation (see **Figure 8**). Microwave measurements are then complementary to far-IR observations of the vibrational transitions (usually obtained at low resolution).

Hyperfine structure in rotational spectra is due to nuclei with a spin $I > \frac{1}{2}$ whose electric quadrupoles interact with the electric field gradient. The corresponding splittings depend on the nucleus involved in this interaction and its spin value. The spectra of numerous molecules containing $^{35}Cl$, $^{37}Cl$, $^{79}Br$, $^{81}Br$, $^{127}I$, $^{14}N$, $^{17}O$, $^{33}S$ have been studied and the corresponding quadrupole coupling constants determined. The deuterium coupling constants have been studied more recently because the splittings are smaller (several tenths of kHz) and were observed only by very high resolving spectrometers (molecular beam maser, microwave Fourier transform spectrometers). These

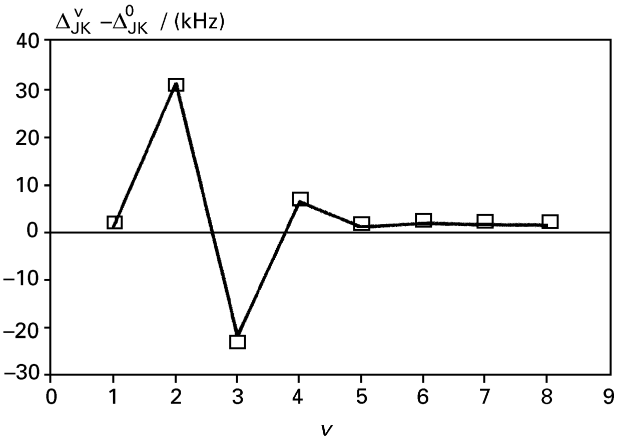



**Figure 8** Comparison of observed (□) and calculated (—) variations of the quartic centrifugal distortion constant $\Delta_{JK}$ with the ring-puckering quantum number *v* for methylene cyclobutane. Reproduced with permission from Charro ME, Lopez JC, Alonso JL, Wlodarczak G and Demaison J (1993) The rotational spectrum of methylene cyclobutane. *Journal of Molecular Spectroscopy* **162**: 67.

spectrometers also allow the more or less complete resolution of the hyperfine structure due to two or more nuclei. Spin rotation and spin–spin coupling constants are also accessible by measuring the transitions involving the lowest values of the rotational quantum numbers.

Dipole moments are determined by applying an external electric field (Stark effect). The accuracy of the experimental dipole moments is about 0.001 D under good conditions. It is mainly limited by the homogeneity of the electric field. The calibration is generally done by using the OCS dipole moment as a reference. The vibrational dependence of the dipole moment can also be studied. In some cases (allene for example) the molecules possesses a vibrationally induced dipole moment and no permanent dipole moment in the ground state. In some spherical tops ($CH_4$, $SiH_4$, etc.) a very small dipole moment induced by centrifugal distortion has been measured (~$10^{-5}$ D).

## Atmospheric applications

The atmospheric transmission between 0 and 1 THz, at the ground level, is dominated by the absorption lines of water vapour and, to a less extent, by some absorption lines due to molecular oxygen (magnetic dipolar transitions), as shown in **Figure 9**. These strong, broad absorption lines are a limiting factor for the observations of other signals, i.e. absorptions due to minor components of terrestrial atmosphere or interstellar emissions. Nevertheless microwave sensors plays an important role in atmospheric measurements either in ground-based facilities or air- and spaceborne ones. The advantages of microwave sensor are the following:

- accurate measurements over the altitude range 0–100 km, mostly independent of clouds and aerosols,
- high frequency resolution and good sensitivity using superheterodyne receivers,
- accurate measurements of ozone profile and trace constituents of importance in catalytic ozone destruction cycles (ClO etc.)

In any event the data collected have to be analysed together with data obtained in the UV, visible and IR part of the electromagnetic spectrum for a reliable interpretation.

The frequency of the centre of the absorption line is not the only parameter which is necessary. The line shape is dominated by molecular collisions up to

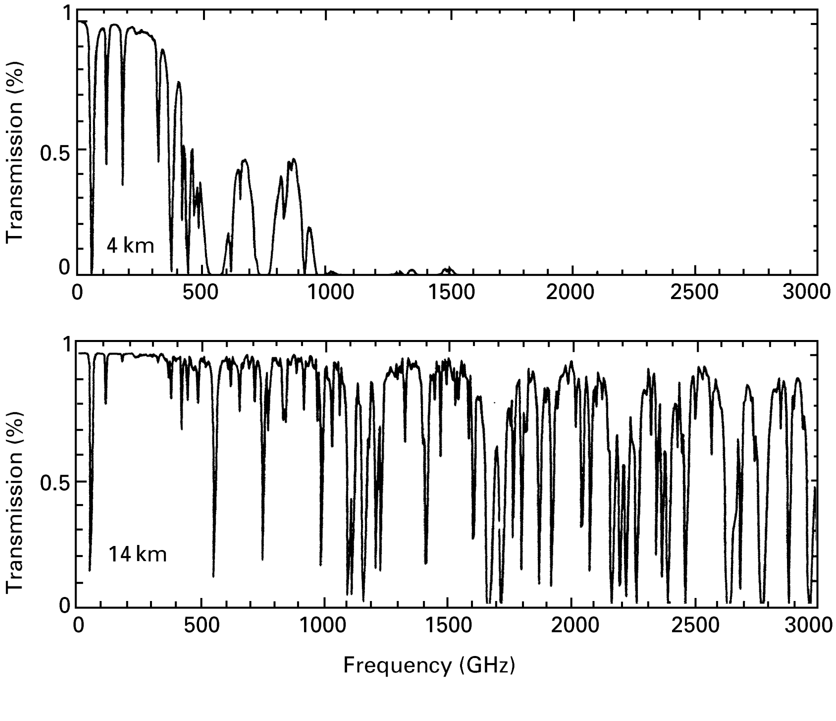



**Figure 9** Atmospheric transmission in the submillimetre and far-IR from (top) a very good high-altitude ground-based site (Mauna Kea at 4.2 km) and from (bottom) an airborne observatory (e.g. KAO at 12 km). The blocked regions are mostly caused by molecular absorption. Reproduced with permission from Phillips TG and Keene J (1992) Submillimetre astronomy. *Proceedings of IEEE* **80**: 1662.

an altitude of 80 km. The collisional broadening parameters with $N_2$ and $O_2$, and their temperature dependence, are determined in the laboratory: they are of a crucial importance for data inversion. Experimental laboratory data with an accuracy of 2–3% are now obtained for the collisional broadening coefficients; the temperature dependance is usually determined with a greater uncertainty but this does not influence the data inversion too much. These laboratory data are also useful benchmarks for theoretical calculations and model testing.

Millimetre-wave sensors represent the only ground-based technique for the observation of stratospheric ClO, the abundance of which is fully correlated to ozone depletion. Moreover this technique allows a continuous observation of ClO, and the analysis of its diurnal cycle, as showing in **Figure 10**. The most frequently observed line is the $J = \frac{15}{2} - \frac{13}{2}$ transition at 278.632 GHz, which is the most intense one. This line is also one of the less blended lines, (interferences with ozone lines located in the neighbourhood are not too strong). This line is also broadened by the hyperfine components. The total line shape contains the contributions of the successive atmospheric layers, and its inversion leads to the vertical concentration profile of ClO.

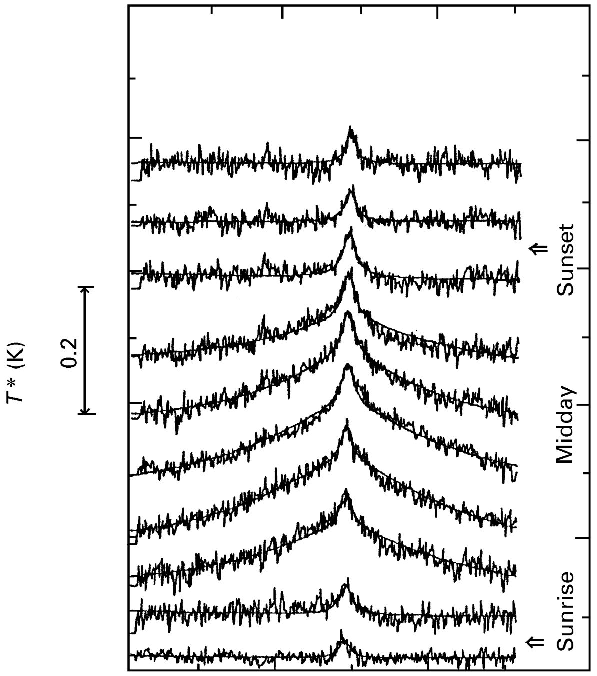



**Figure 10** Diurnal variations of the stratospheric ClO lines shape over McMurdo Station, Antarctica, averaged over the period 20–24 September 1987. de Zafra RL, Jaramillo M, Barrett J, Emmons LK, Solomon P and Parrish A (1989) New observations of a large concentration of ClO in the springtime lower stratosphere over Antarctica and its implications for ozone-depleting chemistry. *Journal of Geophysical Research* **94**: 11423.

Another application of microwave spectroscopy is the analysis of pollutants. Recently, microwave Fourier transform spectrometers have been used to analyse polluted air samples, in the frequency range 10–26 GHz. The air sample is supersonically expanded in a Fabry–Perot resonator, the technique being the same as the one used for the study of molecular complexes. The difference is in the carrier gas which is now air instead of argon. Laboratory studies show that the sensitivity decreases by a factor of 30 when argon is replaced by air. Nevertheless, the sensitivity is still high enough to allow the detection of most of the polar constituents of the sample. Another advantage, already mentioned above, is the very high frequency resolution, which permits the unambiguous identification of a great number of pollutants.

## Radioastronomy

One of the most fruitful application of laboratory microwave spectroscopy over the last twenty years is the analysis of the molecular content of interstellar clouds. These clouds contain gas (99% in mass) which has been mostly studied by radioastronomy, and dust, whose content has been analysed mostly by IR astronomy. The clouds rich in molecular content are dense or dark clouds (they present a large visual extinction), with a gas density of $10^3$–$10^6$ molecules $cm^{-3}$, and temperatures of $T < 50K$. At these low temperatures only the low-lying quantum states of molecules can be thermally (or collisionally) excited, i.e. rotational levels. Spontaneous emission from these excited states occurs at microwave wavelengths. In some warm regions of dense clouds (star formation cores) the absorption of IR radiation produces rotational emission in excited vibrational states. Other rich chemical sources are the molecular clouds surrounding evolved old stars, such as IRC+10216, and called circumstellar clouds.

In the 1980s and 1990s a lot of radiotelescopes were built, with large antennas (diameter = 10–30 m) and sensitive receivers in the millimetre and submillimetre range. More than 100 different molecular species were found in the interstellar medium (see **Table 1**) and, for some of them, various isotopic species were also detected. The identification of interstellar species is not easy because of the high density of lines in the spectra of some interstellar clouds. A millimetre wave spectrum of the Orion nebula is shown in **Figure 11**. This is owing to the richness of the chemistry in these clouds and also to the improved sensitivity of the latest generation

**Table 1** Interstellar molecules

| *Number of atoms* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *13* |
| $H_2$ | $H_2O$ | $NH_3$ | $SiH_4$ | | | | | | | |
| OH | $H_2S$ | $H_3O^+$ | | | | | | | | |
| SO | | | | | | | | | | |
| $SO^+$ | $N_2H^+$ | | | | | | | | | |
| NO | $SO_2$ | | | | | | | | | |
| SiO | HNO | | | | | | | | | |
| SiS | $SiH_2$? | | | | | | | | | |
| SiN | $H_2D^+$ | | | | | | | | | |
| NS | $NH_2$ | | | | | | | | | |
| HCl | | | | | | | | | | |
| HF | | | | | | | | | | |
| NaCl | | | | | | | | | | |
| KCl | | | | | | | | | | |
| AlCl | | | | | | | | | | |
| AlF | | | | | | | | | | |
| PN | | | | | | | | | | |
| NH | | | | | | | | | | |
| $CH^+$ | HCN | $H_2CO$ | $HC_3N$ | $CH_3OH$ | $HC_5N$ | $HCOOCH_3$ | $HC_7N$ | $CH_3C_5N$ | $HC_9N$ | $HC_{11}N$ |
| CH | HNC | HNCO | $C_4H$ | $CH_3CN$ | $CH_3CCH$ | $CH_3C_3N$ | $(CH_3)_2O$ | $(CH_3)_2CO$ | | |
| CC | $C_2H$ | $H_2CS$ | $H_2CNH$ | $CH_3NC$ | $CH_3NH_2$ | $CH_3COOH$ | $CH_3CH_2OH$ | | | |
| CN | $C_2S$ | HNCS | $H_2C_2O$ | $CH_3SH$ | $CH_3CHO$ | $C_6H_2$ | $CH_3CH_2CN$ | | | |
| CO | $SiC_2$ | $C_3N$ | $NH_2CN$ | $NH_2CHO$ | $CH_2CHCN$ | $C_7H$ | $CH_3C_4H$ | | | |
| CSi | HCO | l-$C_3H$ | HCOOH | $C_2H_4$ | $C_6H$ | | $C_8H$ | | | |
| CS | $HCO^+$ | *c*-$C_3H$ | $CH_4$ | $C_5H$ | $CH_2OCH_2$ | | | | | |
| CP | $HOC^+$ | $C_3O$ | *c*-$C_3H_2$ | $HC_2COH$ | | | | | | |
| $CO^+$ | OCS | $C_3S$ | l-$C_3H_2$ | l-$H_2C_4$ | | | | | | |
| | $HCS^+$ | $HOCO^+$ | $CH_2CN$ | $HC_3NH^+$ | | | | | | |
| | $CO_2$ | HCCH | $C_4Si$ | $C_4H_2$ | | | | | | |
| | CCO | $HCNH^+$ | HCCNC | | | | | | | |
| | MgNC | HCCN | HNCCC | | | | | | | |
| | MgCN | $CH_2D^+$ | $H_2COH^+$ | | | | | | | |
| | CaNC | $H_2CN$ | $C_5$ | | | | | | | |
| | $C_3$ | $SiC_3$ | | | | | | | | |
| | NaCN | | | | | | | | | |
| | $CH_2$ | | | | | | | | | |

of radiotelescopes. The characterization of the molecules present in these dense cloud requires a knowledge of the laboratory spectra. In some cases ($C_3H_2$, $HC_9N$, etc.) the identification was first made in the interstellar medium, before laboratory evidence. Nevertheless in the case of $HC_{11}N$, the highest membrane of the cyanopolyine series, interstellar detection was claimed at the beginning of the 1980s. This molecule was recently produced in the laboratory and its rotational spectrum does not fit the interstellar line. A search for $HC_{11}N$ with the new experimental data was at first unsuccessful but, finally, a deeper search confirmed the presence of $HC_{11}N$ in the interstellar medium. A lot of laboratory studies have been devoted to this family of molecules: the rotational spectrum of $HC_{17}N$ has been observed, and numerous hydrocarbons of the type $C_nH_m$, with $n > m$, have been produced in discharges and their spectra analysed.

The detection of isotopomers in interstellar medium is a source of information on the elemental isotopic ratio. Molecules containing the following atoms have been detected: D, $^{13}C$, $^{15}N$, $^{17}O$, $^{18}O$, $^{33}S$, $^{34}S$ and $^{36}S$. The deuterated species are of particular interest because their abundances bring useful information on the chemical processes which take place in the peculiar conditions of the interstellar medium (isotopic fractionation).

Molecular hydrogen is the dominant molecule; the second most abundant molecule, CO, is four orders of magnitude less abundant. But $H_2$ has no strong transitions in the microwave regions, CO is mainly used to map interstellar clouds in our galaxy and others, and also in quasars. The observation of

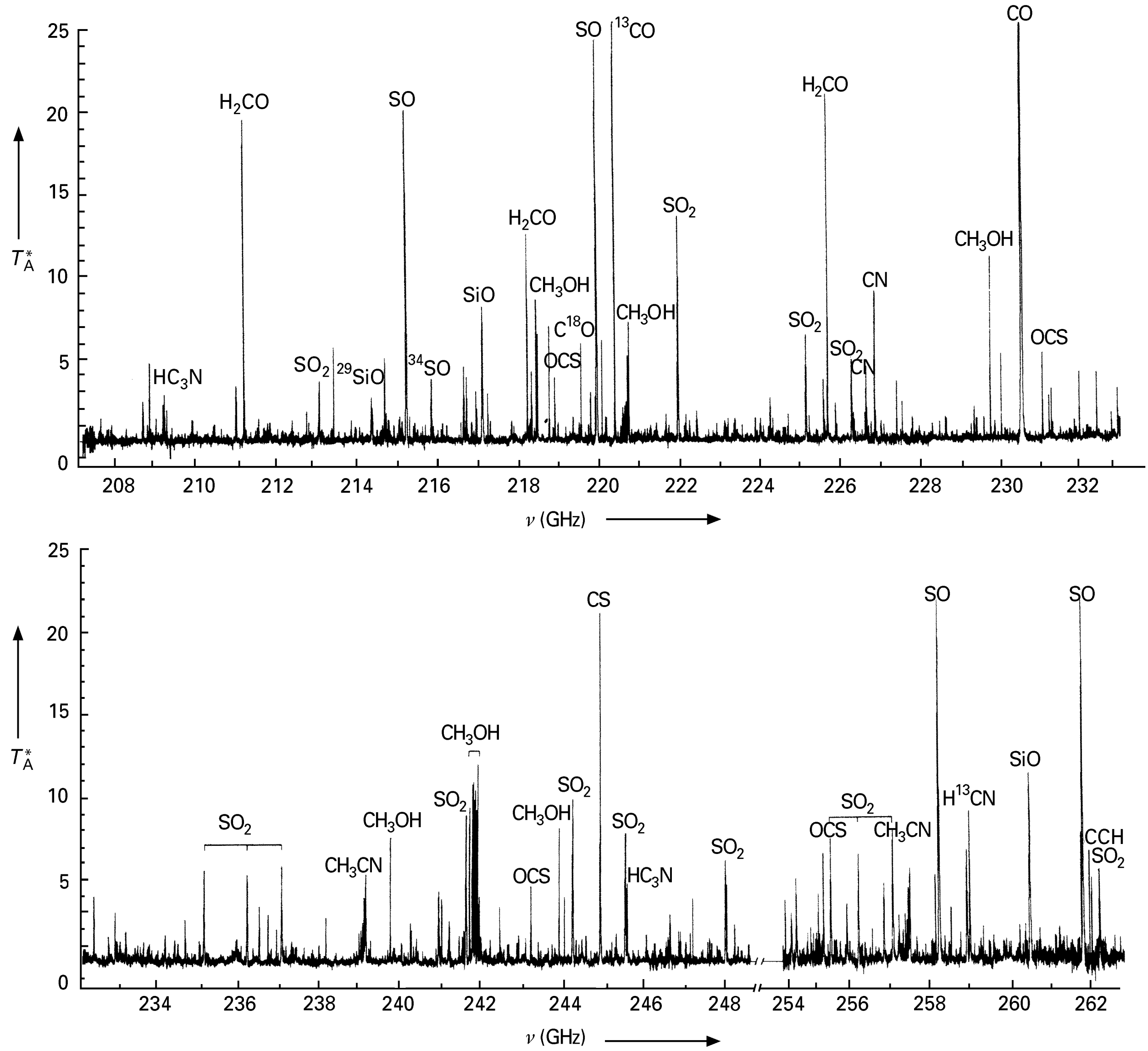



**Figure 11** Millimetre wave spectrum of the Orion nebula in the direction of the so-called Kleinmann–Low area. Rotational spectra from many molecules are seen; ν = frequency and $T_A$* = antenna temperature, a measure of emission intensity. Reproduced with permission from Blake GA, Sutton EC, Masson CR and Phillips TG (1987) Molecular abundances in OMC-1: the chemical composition of interstellar molecular clouds and the influence of massive star formation. *Astrophysical Journal* **315**: 621.

several lines of the same species gives information on the physical conditions in the interstellar cloud: temperature, molecular density. In the case of OH radical, the splitting of the observed microwave lines by the local magnetic field (Zeeman effect) is a way to evaluate its order of magnitude. Several molecular ions have been studied in the laboratory ($H_2D^+$, $H_3O^+$, $CH_2D^+$, etc.) because of their importance in interstellar chemistry, which consists mostly in gas phase ion–molecule reactions. But in many cases their reactivity prevents their interstellar detection. Radioastronomy has also been applied to the analysis of planetary atmospheres, together with infrared observations. Both CO and $H_2O$ were detected in Mars and Venus, $SO_2$ in Io (a satellite of Jupiter), CO and HCN in Neptune. In Titan, a satellite of Saturn, HCN, $HC_3N$ and $CH_3CN$ were detected, indicating a complex photochemistry. More detailed mappings were undertaken more recently with interferometers working in the millimetre-wave region.

Millimetre astronomy has also been found to be a powerful tool for the physicochemistry of comets. This was fully demonstrated by the observations of two exceptional comets: Hyakutake (1996) and

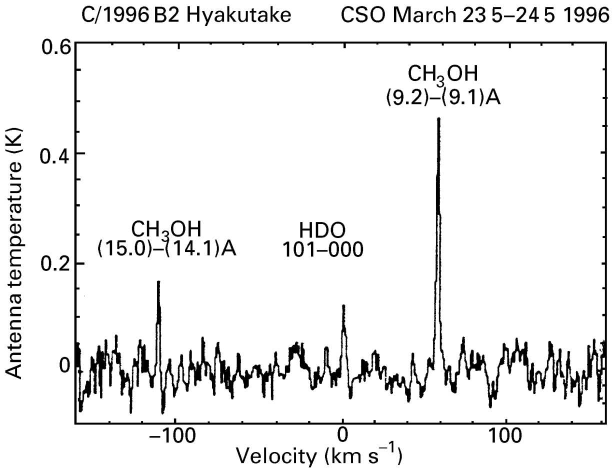



**Figure 12** The $1_{10} \leftarrow 0_{00}$ HDO line at 465 GHz, observed at the Caltech Submillimetre Observatory, in comet Hyakutake. Two lines of methanol are present in the same spectrum. Reproduced by permission from Crovisier J and Bockelée-Morvan D (1997) Comets at the submillimetric wavelength in *ESA Symposium*, Grenoble, France.

Hale–Bopp (1996–1997). The newly detected molecules in these two comets are: CS, $NH_3$, HNC, HDO, $CH_3CN$, OCS, HNCO, $HC_3N$, SO, $SO_2$, HCCS, HCOOH, $NH_2CHO$, CN, $CO^+$, $HCO^+$ $H_3O^+$. This number is considerably bigger than the total number of molecules previously in comets. **Figure 12** shows the detection of HDO and methanol in the comet Hyakutake.

Increasing amounts of data are being obtained at higher frequencies, i.e. in the submillimetre region. A recent survey of Orion was made between 607 and 725 GHz, and another one between 780 and 900 GHz started. These spectral regions are well suited for the detection of light hydrides. They are limited by the atmospheric windows. A continuous coverage will be available with the future satellite observatories, which are planned for the beginning of the third millennium.

## List of symbols

$m$ = molecular weight; $T$ = temperature (K); $\gamma_L$ = collisional broadening parameter; $\Delta\nu_D$ = Doppler half-maximum halfwidth; $\nu_0$ = transition frequency.

*See also:* **Atmospheric Pressure Ionization in Mass Spectrometry; Cosmochemical Applications Using Mass Spectrometry; Environmental Applications of Electronic Spectroscopy; Interstellar Molecules, Spectroscopy of; Microwave Spectrometers; Rotational Spectroscopy, Theory; Solid State NMR, Rotational Resonance; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading


Demaison J, Hüttner W, Tiemann E, Vogt J and Wlodarczak G (1992) *Molecular Constants mostly from Microwave, Molecular Beam, and Sub-Doppler Laser Spectroscopy*, Landolt–Börnstein, Numerical Data and Functional Relationships in Science and Technology (New Series) Group II, Vol 19. Berlin: Springer.

Encrenaz PJ, Laurent C, Gulkis S, Kollberg E and Winnewisser G (eds) (1991) *Coherent Detection at Millimetre Wavelengths and their Applications*. Les Houches Series. New York: Nova Science Publishers.

Gordy W and Cook CL (1984) *Microwave Molecular Spectra*. New York: Wiley.

Graner G, Hirota E, Iijima T, Kuchitsu K, Ramsay DA, Vogt J and Vogt N (1995) *Structure Data of Free Polyatomic Molecules*, Landolt–Börnstein, Numerical Data and Functional Relationships in Science and Technology (New Series) Group II, Vol 23. Berlin: Springer.

Kroto HW (1975) *Molecular Rotational Spectra*. London: Wiley.

Townes CH and Schawlow AL (1955) *Microwave Spectroscopy*. New York: McGraw-Hill.

# Microwave Spectrometers

**Marlin D Harmony**, University of Kansas, Lawrence, KS, USA

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Methods & Instrumentation**

Microwave radiation, defined roughly as electromagnetic radiation with a frequency in the range of 3000 to 300 000 MHz (wavelengths from 10 to 0.1 cm), finds extensive use in chemistry and physics chiefly for two spectroscopic applications. The first of these involves the study of certain magnetic materials, especially paramagnetic substances, and is generally known as electron spin resonance spectroscopy. The second involves the spectroscopic study of the rotational energy states of freely rotating molecules in the gas phase. This latter field of investigation, properly known as rotational spectroscopy but universally and synonymously identified as microwave spectroscopy, is the subject matter of this article. Any instrument used to detect, measure and record the discrete and characteristic absorption of microwave radiation by gaseous molecular samples is thus commonly known as a microwave spectrometer.

## General description

According to the well-known principles of quantum mechanics, the rotational energies of a rotating molecule, considered approximately as a rigid framework of atoms, are limited to certain discrete, quantized values $E_i$. Upon irradiation of a gaseous molecular sample by microwave radiation, an absorption of radiation is possible only if the frequency $\nu$ of the radiation satisfies the Bohr frequency relation

$$\frac{E_2 - E_1}{h} = \nu$$

where $E_1$ and $E_2$ are the initial and final rotational energies and $h$ is Planck's constant ($6.626 \times 10^{-34}$ J s). When a molecule in the quantum state 1 absorbs radiation and is excited to the quantum state 2 we say a spectral transition has occurred. The spectral transitions permitted by the Bohr relation are further limited by other quantum mechanical rules known as selection rules. The net result is that a particular molecule will exhibit typically tens, hundreds or even thousands of relatively sharp, discrete, rotational absorption *lines* in the microwave spectral region. For gas samples at pressures of less than approximately 100 mtorr the frequency widths of the absorption lines are very narrow (typically 0.1–1 MHz) so the resolving power of a microwave spectrometer is very high.

Quantum mechanical and electromagnetic theory provide an additional extremely important restriction upon the occurrence of rotational transitions, namely, to a first and generally adequate approximation they can occur only for molecules having non-zero electric dipole moments. Thus, microwave spectra occur for the polar molecules of water, carbon monoxide and acetone but not for the non-polar moleculess of methane, carbon dioxide and benzene. It is worth stressing also that rotational spectra are produced only by gaseous molecules, not by liquids or solids. While this seems at first a serious limitation it should be noted that it is possible to vaporize even very refractory materials at elevated temperatures. Thus, the microwave spectrum of gaseous sodium chloride (NaCl) molecules is perfectly well known.

On the other hand, microwave spectroscopy is generally not useful for heavy molecules, i.e. those with molecular weights in excess of a few hundred atomic mass units. The reasons for this will be discussed later, but the result is that microwave spectroscopy tends to be far less generally applicable than other spectroscopic techniques such as IR or NMR spectroscopy.

Some detailed applications and theoretical aspects will be described later, but it is worthwhile noting in this general discussion that microwave spectroscopy clearly distinguishes molecular isotopic composition. Thus the microwave spectrum of carbonyl sulfide (OCS) exhibits distinct and easily identifiable spectral lines for various isotopomers such as $^{16}O^{12}C^{32}S$, $^{16}O^{13}C^{32}S$, $^{16}O^{12}C^{34}S$, $^{17}O^{12}C^{32}S$, and $^{18}O^{12}C^{32}S$ in natural abundance. This means that microwave spectra provide very specific information about the individual isotopomers rather than some molecule imagined to be composed of the elements with their average atomic masses or weights.

## Experimental considerations

Microwave spectroscopic experimentation blossomed at the conclusion of World War II because of

the military developments in microwave technology, especially the development of practical microwave generators such as the klystron (vacuum tube) oscillator, and of microwave detectors such as the silicon point contact mixer diode. Later developments led to the backward wave oscillator (BWO) and still more recent work in solid state electronics has led to the availability of a variety of entirely solid state microwave generators such as the Gunn diode. In accord with Maxwell's equations of electromagnetism, the wavelength of microwave radiation is perfectly adaptable for transmission in conducting metal tubing known as a waveguide or (depending upon the frequency) in specially designed coaxial cables. Microwave devices for attenuation, power splitting, impedance matching, frequency measurement and directional coupling are available. The Further reading section should be consulted for details of these and other rather specialized microwave components.

**Figure 1** presents a block diagram of a typical continuous wave (CW) microwave spectrometer. The gas sample is contained typically in a one to three metre length of standard rectangular waveguide fitted at each end with vacuum-tight windows that are transparent to the microwave radiation. The microwave generator is a klystron, BWO or solid state device, and has provision as shown for electronic apparatus for frequency stabilization and measurement. Modern microwave generators can have stabilities and accuracies as high as $1{:}10^7$ or $1{:}10^8$, which translates to 1 kHz or better. The microwave generator has provision through some electronic means for scanning the frequency over some appropriate range at selectable speeds. After passing through the sample cell, the microwave radiation is detected and further processed by a system of signal amplifiers. A critical aspect for obtaining high sensitivity is the use of square-wave electric field-modulation. This modulation, typically at a frequency of 5–100 kHz, is applied to a central electrode insulated from, and running the length of, the cell walls. The square-wave electric field, through the phenomenon of the Stark effect (see below), modulates the absorption signal at the square-wave frequency, thus permitting narrow band amplification and lock-in detection. Finally the resulting spectrum is commonly observed on either an oscilloscope synchronized with the sweep speed of the microwave generator or a strip chart recorder. Most modern instruments are computer interfaced, allowing powerful spectral manipulations and analyses. In this case the computer normally handles other tasks such as frequency range and sweep speed selection and control.

There are numerous variations to the basic design. In particular, the common rectangular waveguide gas-cell is often replaced with other structures for specialized experiments. For example, microwaves can be propagated through free space utilizing special microwave horns or antennae, so the metal surfaces can be largely eliminated for the study of reactive molecules. In this free-space design, the waveguide cell is thus replaced with a relatively large volume glass cylindrical enclosure fitted at its ends with transmitting and receiving horns. In still another design the microwaves are resonantly enclosed in a cavity whose physical size satisfies the boundary conditions for an electromagnetic standing wave according to Maxwell's equations. A particularly useful design for experiments requiring continuous high-speed pumping of unstable molecules is the microwave Fabry–Perot cavity. This design consists of two appropriately designed metal reflectors, typically circular discs with spherically machined reflecting surfaces. Microwave radiation is coupled into and out of the Fabry–Perot with appropriately designed coupling irises and the entire cavity is then enclosed in a large vacuum chamber attached to a high-speed pumping system. The unstable gas molecules of interest are produced by some means external to the cavity and are then rapidly injected into and pumped out of the cavity continuously.

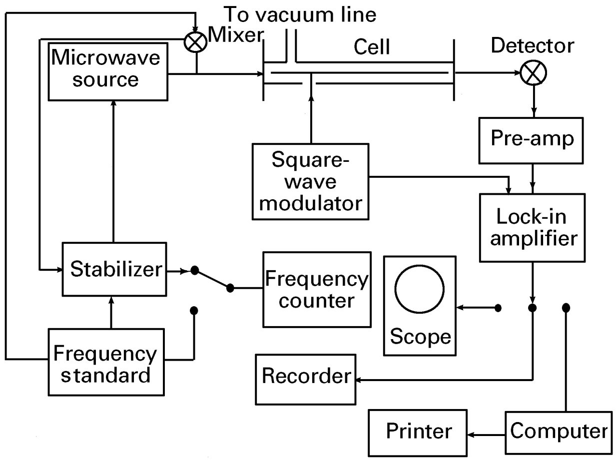



**Figure 1** Block diagram of a conventional Stark-modulated microwave spectrometer.

The CW microwave spectrometer just described is a typical *frequency*-domain instrument. In the late 1970s it was demonstrated that *pulsed time*-domain microwave spectroscopy could be practically performed in analogy to the techniques already well known in other fields such as NMR spectroscopy. **Figure 2** depicts a block diagram of a modern version of a pulsed Fourier-transform microwave spectrometer. The particular instrument shown utilizes a Fabry–Perot cavity and a pulsed-gas nozzle, and is especially useful for detecting microwave

spectra of molecular clusters in an expanding supersonic freejet.

Ignoring some of the details, which can be obtained from the Further reading section, the basic idea of the instrument is that a short pulse of monochromatic microwave radiation (approximately 1 μs in length) irradiates the gas sample in the cavity. If an appropriate transition exists within the bandwidth of the cavity (typically a few MHz), the radiation pulse produces a non-equilibrium ensemble of excited molecules which then immediately begin emitting radiation as they return to equilibrium after the pulse has dissipated. The resulting microwave emission is processed by a succession of coherent mixing processes which eventually yields a low-frequency signal for computer processing. Normally the experiment is repeated hundreds or thousands of times (at typically a 10 Hz repetition rate) to accumulate an observable signal. In accord with the theory of the coherent emission, Fourier transformation of the signal is found to produce the ordinary absorption spectral line. To scan a complete spectrum it is necessary to move the cavity resonance and microwave frequency along in small overlapping steps, repeating the entire signal accumulation process at each frequency.

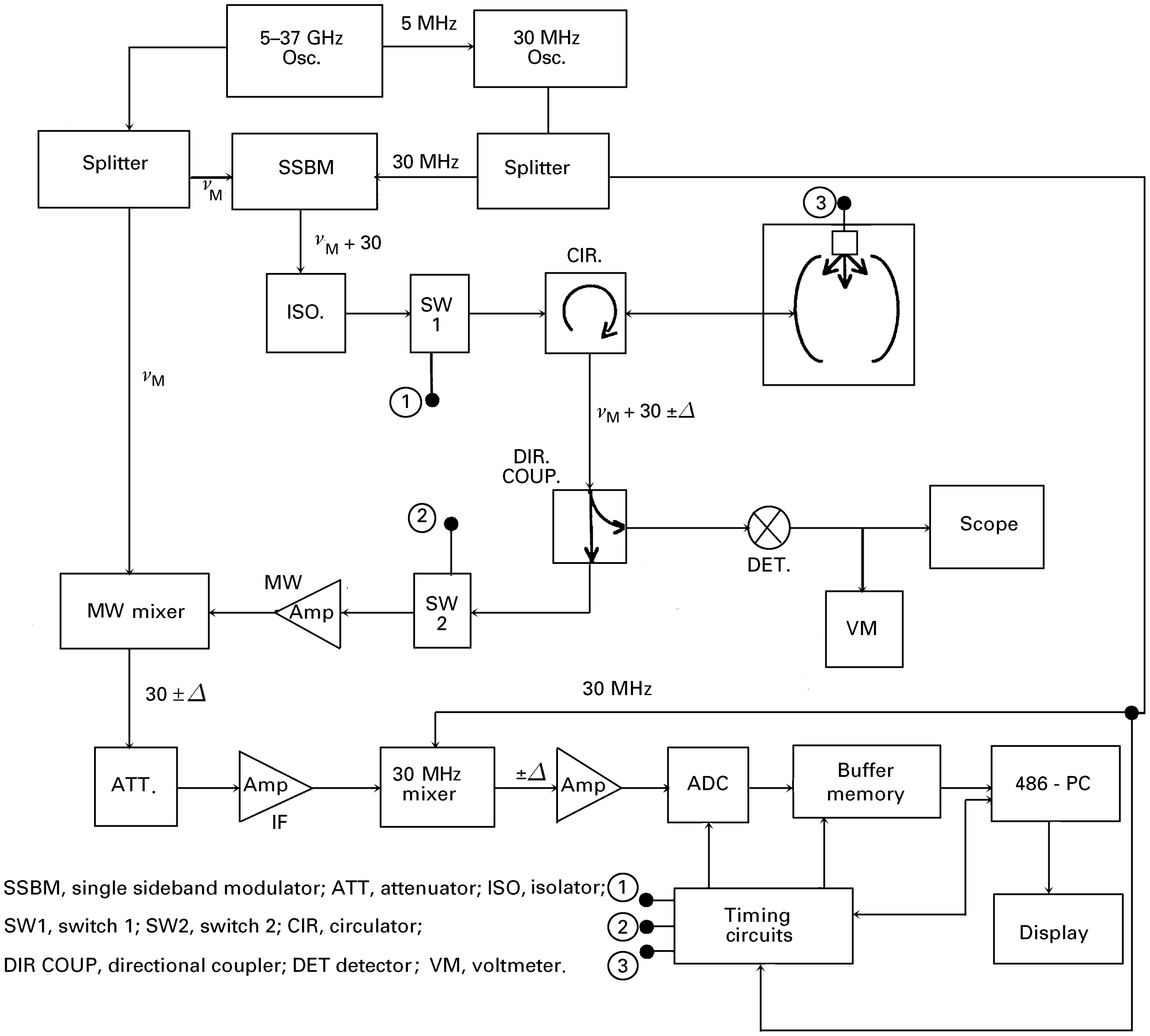



**Figure 2** Block diagram of pulsed Fourier-transform microwave spectrometer. Reproduced with permission of the American Institute of Physics from Harmony MD, Beran KA, Angst DM and Ratzlaff KL (1995). A compact hot-nozzle Fourier transform microwave spectrometer. *Review of Scientific Instruments* **66**: 5196–5202. Copyright 1995, American Institute of Physics.

The result is that the FT-microwave spectrometer (FTMWS) produces the 'same' spectrum as the CW-spectrometer in a much more complex fashion. What are its advantages? As with all spectroscopic experiments carried out in the time domain, the data collection is inherently more efficient, so that the ultimate sensitivity of the FT-spectrometer is substantially higher (perhaps by a factor of 10–100 in practice). In addition, the FT instrument yields much narrower line widths than achievable in typical CW experiments, so the spectral resolution is even higher than for ordinary CW experiments.

## Theoretical aspects of rotational spectra

The rotational quantum states of molecules are characterized by quantum numbers which specify the angular momentum of the rotating molecules. For all molecules, regardless of geometry, the quantum number $J$, with values, 0, 1, 2, . . ., specifies the total rotational angular momentum of the allowed energy states. (Note: we exclude from our discussion molecules having spin angular momentum, in which case a more careful specification of quantum numbers is necessary.) For all linear molecules this quantum number suffices to describe the rotation energy levels (aside from some special effects arising from vibrational motions) in the absence of additional applied fields. The permitted spectral transitions are limited by the selection rule $\Delta J = \pm 1$, i.e. transitions can occur only with a change of one unit of angular momentum. Thus, a typical observed microwave transition for $^{19}F^{12}C^{12}CH$ (in conventional notation) is the $J = 2 \leftarrow 1$, occurring at $\nu = 38824.64$ MHz. The notation means the molecule is excited from the lower $J = 1$ state to the higher $J = 2$ state. For the linear molecule, the rotational energy states in the simplest approximation are given by the expression

$$E = hBJ(J+1)$$

where $B = h/8\pi^2 I$ and $I$ is the classical moment of inertia of the molecule. The term $B$ is known as the 'rotational constant'.

For non-linear molecules, additional quantum numbers (or labels) are necessary, and moments of inertia must be defined for three axes, conventionally labelled $a$, $b$, $c$. Molecules such as $CH_3Cl$ or $NH_3$ can be shown to have $I_a < I_b = I_c$ and are known as prolate symmetric rotors. By convention the $a$-axis is chosen to lie along the molecular threefold (or higher) axis of symmetry. Then the theory for the rotating symmetric rotor shows that the energies depend now not only upon $J$ but also upon the quantum number $K$ which specifies the component of total angular momentum $J$ lying along the $a$-axis. The value of $K$ is limited to $-J, -J+1, \ldots 0 \ldots J-1, J$. The energy levels are then (to the first approximation again) expressed by

$$E = hBJ(J+1) + h(A-B)K^2$$

with definitions of the rotational constants as before, i.e. $B = h/8\pi^2 I_b$ and $A = h/8\pi^2 I_a$. The spectrum of the symmetric rotor is now determined by the selection rules $\Delta J = 0, \pm 1$ and $\Delta K = 0$. Note that the $\Delta K = 0$ rule leads to the result that the spectrum does not depend upon $A$ at all. Moreover, $\Delta J = 0$, which is a formal rule according to theory, leads to no observable microwave transition. The net result is that the symmetric rotor microwave spectrum is essentially of the same structure as that of the linear molecule.

Non-linear or general polyatomic molecules (known as asymmetric rotors) with no threefold or higher axes of symmetry will generally have $I_a \neq I_b \neq I_c$. The rotational energy levels for this case have a complex pattern, depending upon the rotational constants $A$, $B$ and $C$, the rotational angular momentum quantum number $J$, and two other pseudo-quantum numbers or labels related to $K$ for the symmetric rotor. The spectrum is specified by the rules $\Delta J = 0, \pm 1$ again, and some additional symmetry rules involving the pseudo-quantum numbers and the dipole moment components $\mu_a$, $\mu_b$ and $\mu_c$. Some typical observed transitions for bicyclobutane ($C_4H_6$) are the $J = 1_{1,0} \leftarrow 0_{0,0}$ at $\nu = 26625.55$ MHz and the $J = 2_{2,1} \leftarrow 2_{1,1}$ at $\nu = 23995.38$ MHz. The transitions with $\Delta J = +1$ are known as R-branch lines while the $\Delta J = 0$ transitions are known as Q-branch lines.

The previous description has been based upon the so-called rigid-rotor approximation. In fact, molecules deform as they rotate, leading to the phenomenon known as centrifugal distortion. This produces small corrections to the previously described energy expressions, usually amounting to changes of less than 0.1%. Because of the very high precision of microwave measurements, such changes are, however, easily detectable and can be accounted for by appropriate theory. A number of other factors contribute to the finer details of microwave spectra. Some of these will be described in the next section and additional information can be obtained by consulting the Further reading section.

In addition to understanding the frequency axis ($x$-axis) of microwave spectra, it is important to have some knowledge about the intensity (or $y$-) axis. The theory describing the absorption of microwave radiation is complex, but it is worthwhile looking at some of the key factors. In a useful approximate theory for an asymmetric rotor, the intensity (a quantity proportional to the fraction of absorbed radiation) is given for a microwave transition by

$$\text{Intensity} \propto \sqrt{ABC}\ \mu_g^2 \nu^2$$

where the rotational constants have been defined previously, $\mu_g$ is the dipole moment along one of the axes $g = a$, $b$, $c$, and $\nu$ is the frequency of the transition. The expression leads to several key conclusions:

(1) Microwave intensities vanish (i.e. no radiation is absorbed) if $\mu_g = 0$, that is if the molecule is nonpolar as mentioned earlier. Conversely, the squared dependence of $\mu$ strongly favours very polar molecules. Thus, all other factors being equal, the spectral intensities of nitriles (such as $C_2H_5CN$) with $\mu$ values of typically 4 debye, will be approximately $(4/0.08)^2$, i.e. 2500, times greater than those of simple alkanes such as propane ($\mu \approx 0.085$ debye).
(2) Intensities are generally greater at high frequencies, according to the $\nu^2$ dependence. Heavy molecules, with large moments of inertia and corresponding small rotational constants exhibit their transitions generally at low frequencies while the converse is true for light molecules. Thus heavy molecules tend to have ‘weak’ spectra while light molecules have ‘strong’ spectra.
(3) The factor $\sqrt{ABC}$ can be seen to emphasize the dependence upon molecular size and mass, or more precisely, upon moments of inertia. Small, light molecules are favoured because of their large rotational constants, while large, heavy molecules are discriminated against.

## Applications of microwave spectroscopy

### Structure determination

Microwave spectroscopy is the premier physical method for determining accurate and precise molecular structures, i.e. values of interatomic distances (bond distances) and angles (bond angles). This capability arises because the moments of inertia are directly related to the coordinates of the atoms as follows:

$$I_a = \sum_i m_i(b_i^2 + c_i^2)$$

with similar expressions for $I_b$ and $I_c$. In this expression, $m_i$ is the mass of the $i$th atom while $b_i$ and $c_i$ are the $b$- and $c$-axis coordinates of the atom. Assignment, measurement and analysis of microwave spectra yield precise values of rotational constants $A$, $B$ and $C$ and hence values of $I_a$, $I_b$ and $I_c$. Thus the latter quantities provide equations which permit the evaluation of atomic coordinates, $a_i$, $b_i$ and $c_i$. Once the coordinates are known, distances and angles are also known. Thus, the bond distance between atoms $i$ and $j$ is given by

$$R_{ij} = \sqrt{(a_i - a_j)^2 + (b_i - b_j)^2 + (c_i - c_j)^2}$$

A number of problems dealing with molecular non-rigidity must be considered if accurate and meaningful bond distances are to be obtained. Ideally, one would like to determine the coordinates (and hence structure) for the hypothetical vibrationless molecule. Methods for achieving this ideal (to various approximations) have been developed, so that numerous accurate structures have been determined from microwave spectral data. The Further reading section provides examples of such molecular structure determinations.

### Molecular electric dipole moments

It has been mentioned that microwave intensities are determined by the size of the electric dipole moment, so one might suppose that accurate measurements of intensities might provide values of $\mu$. This turns out not to be practical for various reasons. However, another very accurate procedure can be used. If an electric field is applied to a rotating molecule, a well-understood phenomenon known as the Stark effect splits the rotational transitions into a number of components. Precise measurements of these small splittings (typically several MHz) lead to very precise values of the electric dipole moment. Values of $\mu$ determined by this method refer to particular quantum states and are thus much more meaningful theoretically than those determined by classical bulk-gas relative permittivity (dielectric constant) measurements.

### Hyperfine structure

Molecules containing nuclei whose nuclear spin values satisfy $I \geq 1$ exhibit splittings of the rotational

transitions known as hyperfine structure. The predominant cause of these splittings (which for most common quadrupolar nuclei is typically several MHz or less) is the nuclear electric quadrupole interaction. Measurements of the splittings and application of appropriate theory lead to values of a quantity known as the quadrupole coupling constant, usually symbolized as $eQq$. In this expression $Q$ is the nuclear quadrupole moment (a fundamental nuclear constant), $e$ is the charge on the electron, and $q$ is the electric field gradient at the nucleus produced by the surrounding electron and nuclear charges. Coupling constants have been extensively measured for nuclei such as $^{35}$Cl ($I = \frac{3}{2}$), $^{14}$N ($I = 1$) and D ($I = 1$) in a variety of molecules. The resulting values provide important information about the chemical bonding of the atom in question. Note that several very common nuclei, such as $^{1}$H, $^{12}$C and $^{16}$O, have $I < 1$ and consequently produce no quadrupolar hyperfine splittings.

### Internal rotation

Molecules such as propane, methanol or acetone have methyl groups which undergo large amplitude torsional oscillations or internal rotation. This internal rotation is hindered in general by a potential barrier, and the well-known quantum mechanical theory for the effect often leads to observable splittings (typically a doubling) of microwave spectral lines. In general, for high barriers (>1000 cm$^{-1}$) the splittings are small (typically several MHz or less) while for low barriers (~300 cm$^{-1}$ or less) the splittings can be very large (100 MHz or greater). Because of these easily observed splittings microwave spectral measurements have led to a wealth of data on molecular internal rotation barriers. Several related phenomena, involving the puckering or inversion of four- or five-membered ring compounds, or the inversion about pyramidal nitrogen (as for $NH_3$), have also been extensively studied by microwave methods.

### Interstellar microwave spectra

One of the most exciting applications since the 1970s has been the observation of microwave (rotational) spectra of interstellar molecules. Common species such as formaldehyde, ammonia and methylamine and more exotic species such as HCO and H–CC–C≡C–CN have been detected in various interstellar media. The experimental technique differs substantially from that outlined in **Figures 1** and **2**. In this case the interstellar molecular spectra are detected by collecting microwave emissions from interstellar space with large radio telescopes equipped with sensitive microwave receivers. An interesting feature of the interstellar spectra is that the spectral lines are generally Doppler-shifted from their laboratory 'rest' frequencies because the absorbing medium is moving rapidly relative to the background radiation source.

### Multiple irradiation experiments

Microwave spectroscopy is often coupled with a second electromagnetic radiation field to perform specialized experiments. Thus microwave-optical double resonance (MODR) uses optical (say 400–800 nm) radiation simultaneously. The optical radiation transfers molecules to excited electronic states which are then probed by the microwave radiation before the excited molecules return to the normal ground state. Similar experiments utilizing infrared radiation (IRMDR) permit probing of excited vibrational states. Analogous experiments using two microwave fields (MMDR) and a microwave and radiofrequency field (RFMDR) are very commonly used to produce spectral simplification and to aid in spectral interpretation. The double resonance experiments have also been important for obtaining information about collisional energy transfer rates and mechanisms.

### Studies of weakly-bound complexes

Since about 1980 there has been great interest in performing microwave studies of weakly bound species such as $(H_2O)_2$, ArHCl and (HC≡CH)HCl. These species are studied with the unique instrument shown earlier in **Figure 2**, known as a pulsed-nozzle Fourier-transform microwave spectrometer. The weakly-bound species are formed by pulsing a gas-mixture through a small nozzle such that it undergoes a supersonic free-jet expansion. Complexes are formed rather abundantly in such expansions and are stabilized by the low temperatures (<5 K) achieved in the expansion. Pulsed FTMWS (synchronized with the pulsed nozzle) is then used to sensitively observe and study the rotational spectrum. Such investigations will surely continue to be of great future interest because they provide information on van der Waals and hydrogen-bonding forces, both of which are of critical importance to understanding intermolecular potentials.

### Analytical applications

The very high resolution and selectivity of microwave spectroscopy make it an excellent tool for qualitative analysis of gas-phase samples. Indeed, a substantial amount of effort has been placed by microwave spectroscopists in using the method to identify and characterize new chemical species,

especially those which are unstable and hence difficult to study by more conventional techniques. Because microwave spectroscopy is readily adaptable to continuously flowing gas samples (with special cell designs as mentioned earlier) it is an ideal method for investigating the products of combustion, pyrolysis, photolysis or electric discharges. Examples of such studies include OH, CS, $CH_2$=NH, $CF_2$=C=C=O, $HCO^+$, $HNN^+$ and many others. Of course, the last section described the unique application of microwave spectroscopy to unstable molecular clusters and earlier the high selectivity for isotopic analyses was mentioned.

The chief disadvantage of microwave spectroscopy for gas-phase analytical applications is that its sensitivity is not as high as for some other methods (such as laser fluorescence or mass spectrometry). For low molecular weight polar species such as $SO_2$, $NH_3$ and $NO_2$, analytical detection sensitivities using FT-MWS instruments certainly extend into the parts per billion (ppb) range. However, as the molecular size and mass increase or the polarity decreases the sensitivities may fall more typically into the ppm range. Naturally, as with all spectroscopic methods, appropriate preconcentration or preselection schemes may lead to effectively improved detection limits.

From the above it is clear that quantitative measurements at high sensitivities are most useful for a variety of small polar molecules which are of concern from the atmospheric environmental pollution point of view. Thus a substantial amount of effort has been and continues to be placed upon the development of field operable, portable microwave spectrometers for trace gas monitoring using both CW and FT instrumentation. Although there are likely to be continued applications of microwave spectroscopy to pure analysis problems in the future, it seems likely that the microwave spectrometer will continue to find its most exciting applications in the chemistry and physics research laboratory.

## List of symbols

$A$, $B$, $C$ = rotational constants; $b_i$, $c_i$ = $b$- and $c$-axis coordinates of the $i$th atom; $e$ = charge on an electron; $E$ = rotational energy; $h$ = Planck's constant; $I$ = moment of inertia, and the nuclear spin angular momentum quantum number; $J$ = angular momentum quantum number; $K$ = quantum number specifying component of $J$ lying along the $a$-axis; $m_i$ = mass of the $i$th atom; $q$ = electric field gradient; $Q$ = nuclear quadrupole moment; $R_{ij}$ = bond distance between atoms $i$ and $j$; $\mu$ = dipole moment; $\nu$ = frequency.

*See also:* **EPR Spectroscopy, Theory; Gas Phase Applications of NMR Spectroscopy; Microwave and Radiowave Spectroscopy, Applications; Rotational Spectroscopy, Theory; Solid State NMR, Rotational Resonance; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading

Balle TJ and Flygare WH (1981) Fabry–Perot cavity pulsed Fourier transform microwave spectrometer with a pulsed nozzle particle source. *Review of Scientific Instruments* **52**: 33–45.

Gordy W and Cook RL (1984) *Microwave Molecular Spectra*. New York: Wiley-Interscience.

Harmony MD (1981). In: Anderson, HL (ed) *AIP Physics Vade Mecum*, Chapter 15. New York: American Institute of Physics.

Harmony MD, Beran KA, Angst DM and Ratzlaff KL (1995) A compact hot-nozzle Fourier transform microwave spectrometer. *Reviews of Scientific Instruments*. **66**: 5196–5202.

Harmony MD, Laurie *et al* (1979) Molecular structures of the gas-phase polyatomic molecules determined by spectroscopic methods. *Journal of Physical Chemistry Reference Data* **8**: 619–721.

Harmony MD and Murray AM (1987). In: Rossiter BW and Hamilton JF (eds) *Physical Methods of Chemistry: Vol. IIIA – Determination of Chemical Composition and Molecular Structure*, Chapter 2. New York: Wiley.

Legon AC (1983) Pulsed-nozzle, Fourier-transform microwave spectroscopy of weakly bound dimers. *Annual Review of Physical Chemistry* **34**: 275–300.

Steinfeld JI and Houston PL (1978) In: Steinfeld JI (ed) *Laser and Coherence Spectroscopy*, Chapter 1. New York: Plenum.

Townes CH and Schawlow AL (1955) *Microwave Spectroscopy*. New York: McGraw-Hill.

Varma R and Hrubesh LW (1979) *Chemical Analysis by Microwave Rotational Spectroscopy*. New York: Wiley-Interscience.

# Mineralogy Applications of Atomic Spectroscopy

*See* **Geology and Mineralogy, Applications of Atomic Spectroscopy.**

# Molybdenum NMR, Applications

*See* **Heteronuclear NMR Applications (Y–Cd).**

# Mössbauer Spectrometers

**Guennadi N Belozerski**, St.-Petersburg State University, Russia

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Methods & Instrumentation**

To obtain the Mössbauer spectrum the radiation from a Mössbauer source should be directed onto the sample under study. In Mössbauer experiments it is not the absolute energy of the γ-quanta which is determined but the energy shift of the nuclear levels. The energy scanning is carried out by the use of the Doppler effect and the energy parameters ($\Gamma$, $\delta$) are expressed in velocity units, $v(E\ v/c)$. The Mössbauer spectrum is a measure of the dependence of the total intensity of radiation $I(v)$ registered by a detector in a definite energy region on the relative velocity $v$ of the source.

A schematic diagram of a Mössbauer experiment and the spectrum is shown in **Figure 1**. If both the source and the absorber are characterized by single lines of natural width $\Gamma_{nat}$, $\delta$ being zero, the spectrum will show maximum absorption at $v = 0$. In this situation the intensity, $I(0)$, registered by the detector is minimized (**Figure 1C**). When the source moves at a certain velocity $v$, the emission line $J_M(E)$ is displaced relative to the absorption line $J^a(E)$. The overlap then decreases and the intensity increases. Finally, at a velocity that may be considered to be infinitely large ($v = \infty$), the spectrum overlap becomes so small that any further increase in velocity will not result in a significant increase in relative intensity. This value of intensity may be described as $I(\infty)$. The fact that the line shapes of the source and absorber are described by Lorentzians causes the experimentally observed line for a thin absorber to be Lorentzian, and its half-height width is the sum of the line widths of the source and the absorber.

A typical device for accumulating the Mössbauer spectrum is the multichannel analyser, where the count rate is a function of a definite value of the Doppler velocity. The count rate is normalized relative to the off-resonance count rate. Hence, for transmission-mode Mössbauer spectroscopy relative intensities are always less than unity (or 100%). In Mössbauer scattering experiments relative intensities always exceed 100% and can reach several hundred percent in the case of electron detection from samples with a high abundance of the resonant isotope. It is most often that the $-v_{max}$ value corresponds to the first channel and the $+v_{max}$ value to the last channel. The quality of a Mössbauer spectrometer is determined by how accurately the modulation of the γ-quanta energy follows the chosen mode of movement.

## Typical Mössbauer spectrometers

The Mössbauer experiment may be in transmission mode, where γ-quanta are detected. The detector

registers not only the γ-rays of the Mössbauer transition, but also the background noise. The main process competing with resonance interactions in the transmission mode experiments is the photoelectric effect. In transmission experiments there are three sources of background: (i) γ- and X-rays of higher energies which may be Compton-scattered; Bremsstrahlung produced outside the detector may contribute to this too; (ii) high-energy γ- and X-rays having lost only a part of their energy in the detector; (iii) X-rays that are not distinguished by the detector from the Mössbauer quanta.

In scattering Mössbauer spectroscopy the processes competing with Mössbauer scattering are the Compton effect, Rayleigh scattering and classical resonant scattering of γ-rays. The Compton effect is to be specially taken into account when the source emits high-energy γ-rays in addition to the Mössbauer radiation. The typical experimental arrangements are presented in **Figure 2**.

In Mössbauer spectroscopy the shape of the spectrum and its area are the 'signals' conveying quantitative information on a phase. When the shape is known to be Lorentzian, for example, the amplitudes and the line positions are often used as para-meters of the signal $-I(v_i)$ value at $v_i$ ($i$ = 0, 1, ...).

Mössbauer scattering spectra obtained by detection of the γ-quanta or X-rays emitted out of the bulk of a material, convey information on the layer with a depth which is determined by the total linear absorption coefficient $\mu_a(E)$. The values of $\mu_a(E)$ for γ-rays and X-rays are generally different; therefore the Mössbauer spectra correspond to the layers, which are different in depth (from one to several μm). Backscattering Mössbauer spectroscopy is the most promising technique for applied research and industrial applications (see **Figure 2C**). The backscattering geometry is simple, efficient and suitable for any type of radiation. In such an experiment one can detect any radiation in different scattering channels. However, to detect γ-quanta, a special detector is needed. It has been shown by many experimentalists that the signal/noise ratio of the detection of γ-rays in the experimental geometry of **Figure 2B** is better than for detection of X-rays. At the same time the flat proportional counter has never been used to detect γ-quanta in the backscattering geometry of **Figure 2C**. Indeed, the direct Mössbauer radiation of an intensity which is 100 times as high

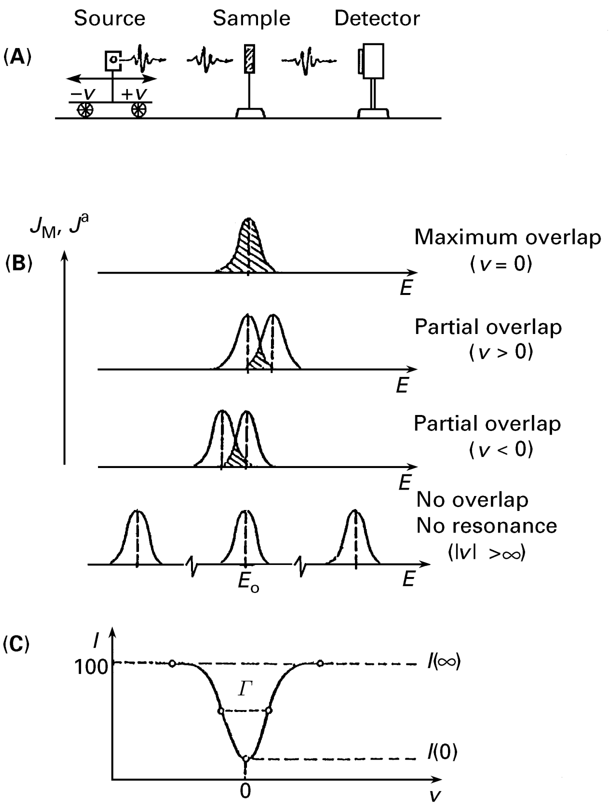



**Figure 1** Schematic illustration of the experimental arrangement (A) used to obtain a Mössbauer spectrum (C) for a single Lorentzian line both in the source and in the sample (B).

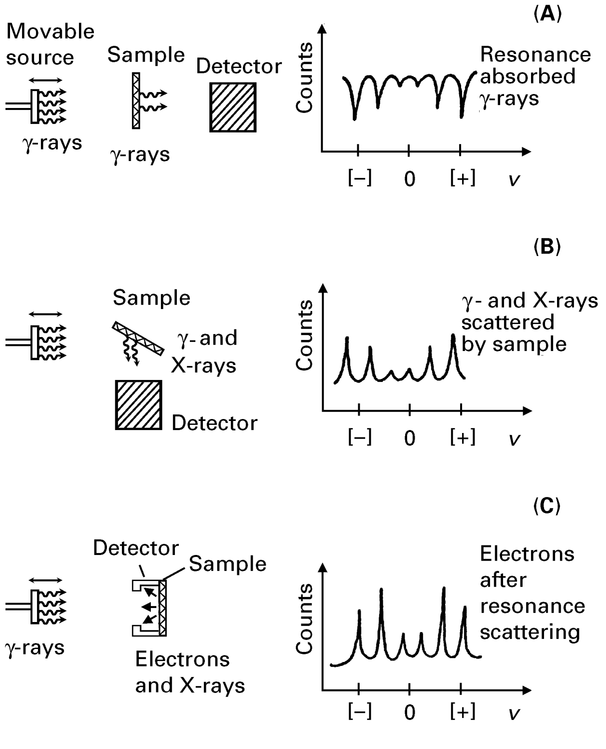



**Figure 2** Experimental arrangements and Mössbauer spectra for a $^{57}Co$ (Cr) source and a sample of α-Fe: (A) transmission geometry, (B) scattering geometry with the detection of γ- or X-rays, (C) backscattering geometry with the detection of X-rays and electrons. The source moves at a velocity *v*.

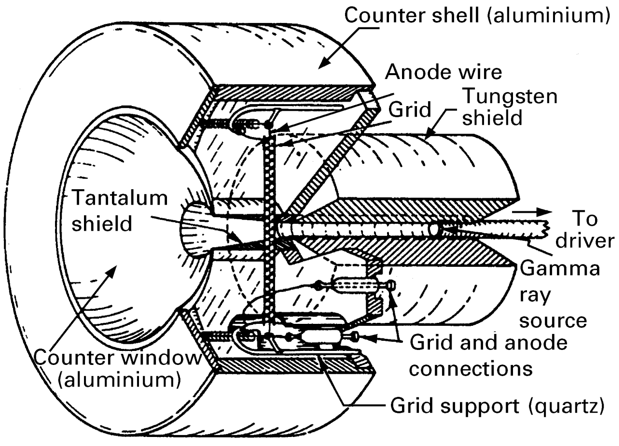



**Figure 3** Spectrometer based on the toroidal detector.

as the scattered intensity also passes through the detector such that the effect would be very small. The need to detect the resonantly scattered γ-quanta in a solid angle close to 2π stimulated the search for a detector capable of sensing scattered photons with an energy of 10–20 keV and which would be insensitive to the direct primary beam of γ-quanta.

The requirements have been met by the use of toroidal detectors. The main problem has involved the necessity to create the inner electric field with circular equipotential lines around the anode. For this purpose, one uses cylindrical grid wires surrounding the anode. Electrons produced within the counter volume travel to the grid and through it to the anode wire. After filling with a krypton–methane mixture the resolution for the 14 keV line for such a counter is ~15%. A section of a Mössbauer spectrometer using the counter is shown in **Figure 3**. This toroidal proportional detector is easy to handle and can be usefully applied to surface studies with high efficiency.

## Conversion electron Mössbauer spectroscopy

Mössbauer transitions are usually highly converted and are followed by the emission of characteristic X-rays and Auger electrons. (The total internal conversion coefficient is high. For most cases de-excitation of the nucleus is via the emission of conversion electrons followed by rearrangement of the excited atomic shell by X-ray emission and Auger processes. More than one electron is produced per resonant scattering event.) The detection of electrons has proved in many cases to be the most efficient means of observing the Mössbauer effect. The principal feature of Mössbauer spectroscopy based on the detection of electrons is that the average energy of an electron beam reaching the$_{xx}$ detector, and also the shape of the energy spectrum, depends on the depth $x$ of a layer d$x$ from which the beam has been generated. This provides interesting possibilities for layer-by-layer phase analysis. Various modifications of Mössbauer spectroscopy based on the detection of electrons have been developed including a technique which allows the Mössbauer signal from a very thin surface layer (~3 nm) of a homogeneous bulk sample to be distinguished. The techniques in this field of Mössbauer spectroscopy are classified as either integral or depth-selective.

The integral technique is called conversion electron Mössbauer spectroscopy (CEMS). In CEMS, of prime interest is the probability that electrons originating from a layer d$x$ at a depth $x$ with energy $E_0$ leave in a random direction from the surface with any energy and at any angle and will be registered by a detector. The electrons may be divided into several groups: conversion electrons, Auger electrons, low-energy electrons resulting from shake-off events and secondary electrons resulting from the re-emitted Mössbauer quanta and the characteristic X-rays. The energies and relative intensities of the first two groups of electrons for $^{57}$Fe and $^{119}$Sn are given in **Table 1**.

The development of CEMS as an independent analytical method came as a result of the development of gas-filled proportional counters for the detection of electrons. **Figure 4** illustrates the operating principle of such a CEM spectrometer. The proportional counter in CEMS detects all the electrons in the energy interval from about 1 keV up to the Mössbauer transition energy. In addition to the high efficiency, the proportional counters have an energy resolution allowing, if we need it, a certain depth selectivity to be obtained.

Phase analysis of multiphase mixtures, fine particles and disordered substances, as well as surface studies, require Mössbauer spectra to be recorded over a wide range of temperatures. The problem of

**Table 1** Main radiation characteristics for the $^{57}$Fe and $^{119}$Sn de-excitation

| | *$^{57}$Fe* | | *$^{119}$Sn* | |
|---|---|---|---|---|
| *Type* | *Energy* (keV) | *Probability per de-excitation ($C_i$)* | *Energy* (keV) | *Probability per de-excitation ($C_i$)* |
| K-conversion | 7.3 | 0.796 | | |
| L-conversion | 13.6 | 0.09 | 19.6 | 0.83 |
| M-conversion | 14.3 | 0.01 | 23.0 | 0.13 |
| KLL Auger | 5.4 | 0.543 | | |
| LMM Auger | | | 2.8 | 0.74 |

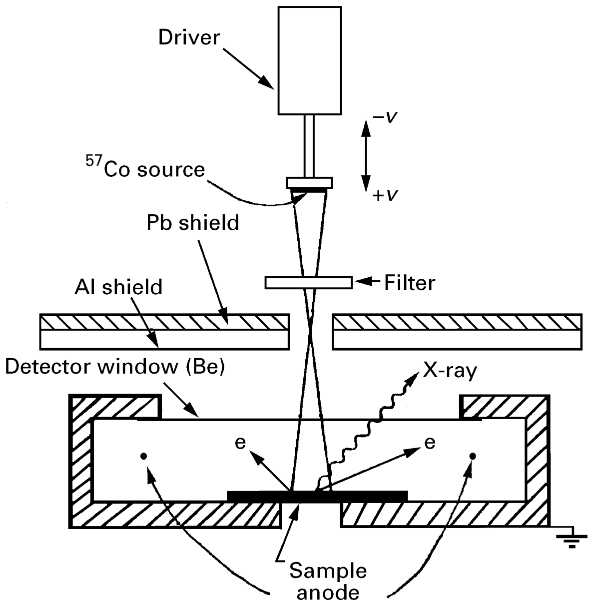



**Figure 4** Schematic picture of a spectrometer for backscattering studies.

the counter operation at temperatures other than ambient has received significant attention in Mössbauer spectroscopy. The counters can operate CEMS at low temperatures near 4.2 K and up to 1100 K.

Arrangements based on proportional counters which allow an independent and simultaneous recording of CEM spectra and X-ray Mössbauer spectra in backscattering geometry, and γ-ray absorption spectra in transmission, have been developed for industrial application purposes, see **Figure 5**. Due to the different escape or penetration ranges of the three radiations involved, the spectra give information on phases, depth and orientation. From a practical point of view the counters for γ-rays, X-rays and electrons must be separated and shielded to ensure independent detection.

In addition to the proportional counters, other types of gas-filled detectors are used in CEMS. First was the parallel-plate avalanche counter. In Mössbauer spectroscopy such detectors have been used as resonance detectors and at higher counting rates. These counters have found application in surface studies and are the effective tool for the registration of low-energy ($E < 1$ keV) electrons, which are practically impossible to detect with the proportional counter. Because of the high electron-detection efficiency this enables the measurement of reasonable spectra in a relatively short time for $^{57}$Fe, $^{119}$Sn, $^{151}$Eu, $^{161}$Dy and $^{169}$Tm. Second, scintillation detectors may be considered. Thin organic (crystal or plastic) scintillators are used for detecting electrons. Gas scintillation proportional counters with a good energy resolution (e.g. $R \approx 8\%$ at 6 keV) may also be constructed for CEMS as well as semiconductor detectors.

Channeltrons, microchannel plates and window less electron multipliers constitute a special group of detectors for CEMS. These have no entrance

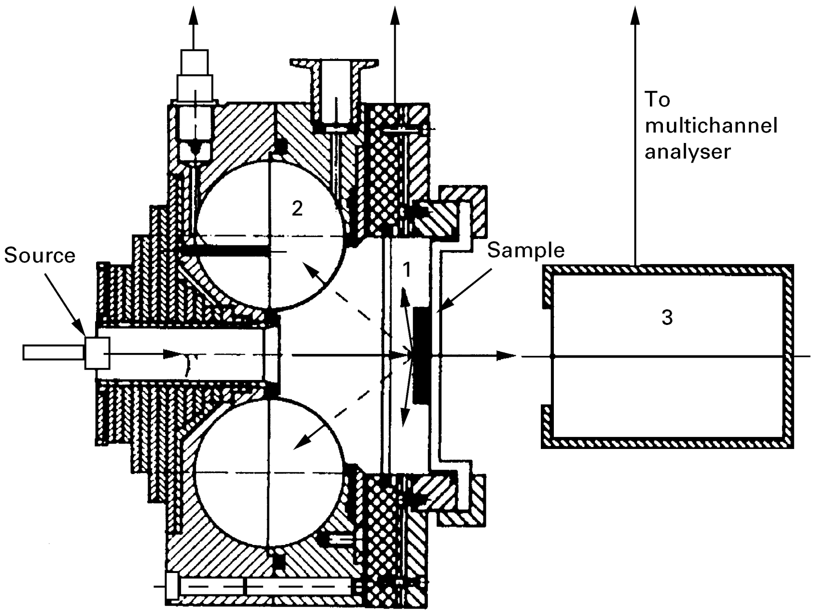



**Figure 5** Set-up for simultaneous recording of CEM spectra (1), X-ray Mössbauer spectra (2) and transmission spectra (3).

windows and are designed for vacuum operation which can be used to advantage in CEM spectrometers operating both at high and low temperatures. A new method of surface study has recently appeared, CEMS based on the detection of very low-energy electrons. Detectors in this group have no energy resolution. The pulse–height distribution at the out-put of these detectors is similar to the noise distribution.

Advantages of the best CEM spectrometers with a channeltron include their easy sample access, high cooling rate, capability of simultaneous transmission measurements and adaptability to on-line experiments. To increase the count rate, detection efficiency or the effect value, a bias potential is sometimes applied to the sample or to the input of the channeltron. The statistical quality of spectra is, as a rule, nearly as good as for the gas-filled ionization detectors. The effective technique of collecting secondary electrons by applying a bias potential between the sample surface and a channeltron has been used to develop a spectrometer for low-temperature measurements (see **Figure 6**). The beam of γ-quanta is incident at 45° to the sample surface. The sample is the first electrode in a system of electrodes used to attract the secondary electrons to the entrance of the channeltron and to accelerate them to an energy corresponding to the maximum detection efficiency.

#### Low energy electron Mössbauer spectroscopy (LEEMS)

Conversion electrons, KLL, KLM and KMM Auger electrons, photoelectrons and Compton-scattered electrons which are produced by γ-rays (with the energy above several hundred eV) in this context may be regarded as 'high energy' electrons emitted by the atom. Secondary electrons result from the interaction of the above electrons with matter. Also, there are electrons that are primarily produced with a very low energy. Two processes contribute to the intensity of the electrons. These are very low energy Auger electrons (LMM, MMM, MMN) and shake-off electrons.

Experimental data show a sharp peak in the number of electrons (related to Mössbauer events) at energies below 20 eV. These electrons supply information on a surface layer to a depth of ~5 nm. The detection of very low energy electrons offers the advantage of short data acquisition times (~77% of the electrons emitted from the Fe atom are low-energy Auger and shake-off electrons), and increases surface sensitivity compared to established procedures relying on the collection of electrons near 7.3 keV. CEMS detectors and techniques are summarized in **Table 2**.

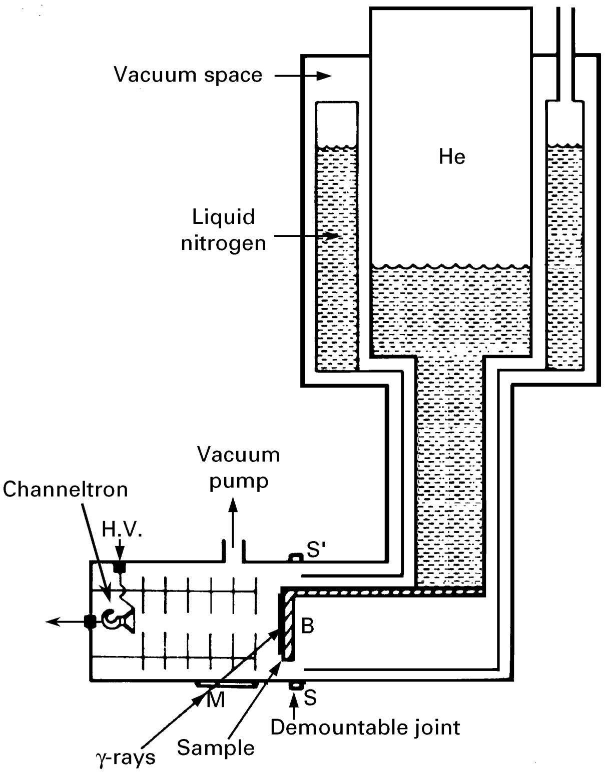



**Figure 6** CEMS spectrometer used to operate at 4.2 K. M, mylar window, B, cold finger. The detection assembly is screwed on to the dewar at SS′.

### Depth-selective conversion electron Mössbauer spectroscopy

The detection of electrons with energy $E$ by a β-spectrometer with high-energy resolution gives a Mössbauer spectrum corresponding to the phase at a

**Table 2** Detectors and electron detection techniques in CEMS

| *With energy resolution* | | *Without energy resolution* |
|---|---|---|
| Electron spectrometers | Magnetic | Parallel-plate avalanche counters |
| | Electrostatic | Channel electron multipliers |
| Ionization detectors | Proportional counters | Gas scintillation detectors |
| | Multiwire proportional counters | Microchannel plates |
| | Semiconductor detectors | Windowless multipliers |
| | Gas scintillation proportional counters | Organic scintillation detectors |
| | X-rays controlled proportional counters | Detection of light produced by microcharges |
| | | Geiger-Müller counters |

depth $x_1$ in the scatterer. If the known relationship between the energy of detected electrons and the depth of the layer through which they have passed is used, then depth-selective analysis of the surface layers can be performed. In depth-selective conversion electron Mössbauer spectroscopy (DCEMS) the group of electrons with a fixed energy leaving the surface at a certain angle within a small solid angle d$\omega$ are of interest.

There are a number of electrostatic and magnetic electron spectrometers that have been used, designed and developed for DCEMS. A schematic description of the DCEM spectrometer, based on the mirror analyser is depicted in **Figure 7**. Electrons, starting from inside the inner cylinder at angles close to 45° to the sample surface (1 $cm^2$), move out through slits in the inner grounded cylinder into a strong field region. The field bends the trajectories of the electrons back towards the inner cylinder. The group of electrons of interest passes through another slit, to be collected on the spectrometer axis. If a position-sensitive detector is placed on the axis, a series of Mössbauer spectra corresponding to different electron energies can be recorded simultaneously. Using three slits enables the simultaneous recording of K-, L- and M-conversion electron spectra.

An important methodology problem in DCEMS is the measuring time. For $^{57}Fe$ only the K-conversion electrons lead to a DCEMS spectrum. The thickness of the analysed layer in DCEMS is significantly less than in CEMS, being less than 80 nm for iron. It is dependent on parameters of the β-spectrometer. The resolution enhancement from 3% to 1% is significant for experiments involving the investigation of surface layers 0–5 nm thick. To investigate a thinner layer (0–2.5 nm thick), the β-spectrometer should detect separate groups of electrons in the 7.2–7.3 keV interval, and it is desirable to have R ≈ 0.5% and $\theta \approx 90°$. The maximum possible selectivity can probably be attained with electrostatic β-spectrometers whose accuracy of energy determination is about 1 eV and the half-width is –10 eV on the 7.3 keV line. To summarize, the experimentalist in DCEMS should try to use a detector with an efficiency close to 100%. There should be no window in the path of the electrons. The temperature of the sample may be varied in the range 4.2–1000 K.

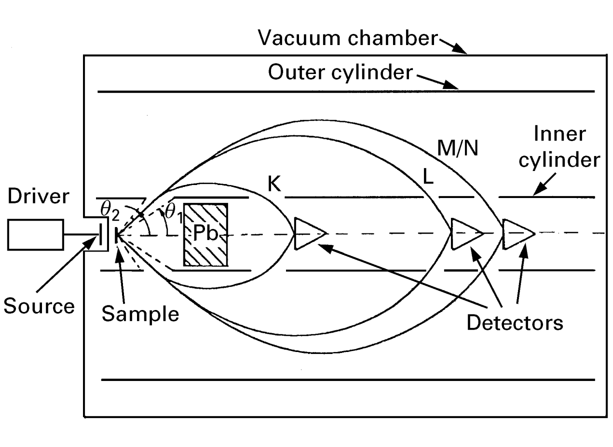



**Figure 7** Schematic diagram of a DCEM spectrometer based on the electrostatic cylindrical mirror analyser. Forward scattering geometry is used. $\theta_1$ and $\theta_2$, minimal and maximal angles for the input slit edge positions; Pb, lead shielding.

## Special Mössbauer spectrometers

There is a special situation where hyperfine interactions are present and a constant velocity $v_i$ is chosen, so that the incident radiation is on resonance with the scatterer's line number 2 ($v_i = v_2$). There is, in this situation, no unique relation between the energies of the incident and scattered γ-quanta. The scattered quantum may have the energy of the incident quantum, as well as the energy belonging to the line number 4. The same is true if relaxation processes or very complicated hyperfine interactions occur in the sample. To study the phenomenon the incident γ-quanta energy should be fixed and the energy spectrum of scattered γ-quanta will show directly the energy change of γ-quanta on scattering.

To obtain the energy distribution, a γ-ray detector is needed with an energy resolution of approximately $\Gamma_{nat}$. For this purpose a resonant filter is placed in front of the conventional detector (see **Figure 8**). This filter is a 'single line' Mössbauer absorber. Driving the filter ('analyser') in the constant acceleration mode and detecting the outgoing radiation allows the $I(v,v_i)$ spectrum to be produced (see **Figure 9**). The observed effect is determined now by the two elastic resonant scattering processes (by four $f$ factors). Two synchronized drive systems are necessary to observe the two scattering processes. This is known as selective-excitation double Mössbauer spectroscopy (SEDMS).

The method is demonstrated by considering the SEDM spectrum recorded for scattering at the energy corresponding to the $-\frac{1}{2} \rightarrow -\frac{1}{2}$ transition in a 9 μm thick $^{57}Fe$ foil (**Figure 9**). The Mössbauer spectrum consists of the second and fourth lines of the usual spectrum of $\alpha$-Fe, i.e. the lines corresponding to the $-\frac{1}{2} \rightarrow -\frac{1}{2}$ transition as well as to the $-\frac{1}{2} \rightarrow +\frac{1}{2}$ transition.

The main advantage of SEDMS is that the method offers a direct means by which the relaxation processes between sublevels of the excited nucleus can be observed. Indeed, the experimental spectrum $I(v,v_i)$ gives direct information on time-dependent hyperfine interactions which determine the nuclear level

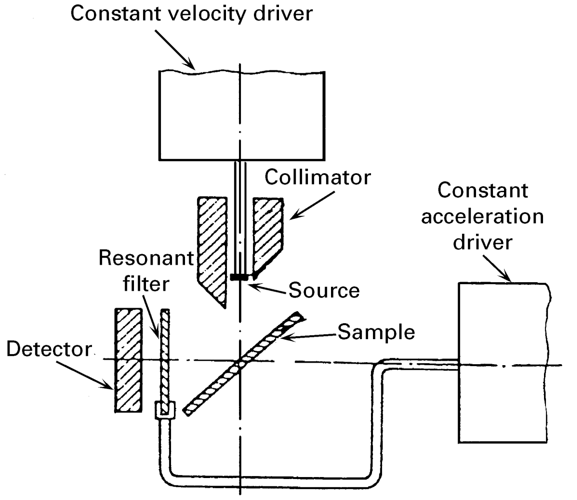



**Figure 8** A schematic experimental arrangement used for selective-excitation double Mössbauer spectroscopy.

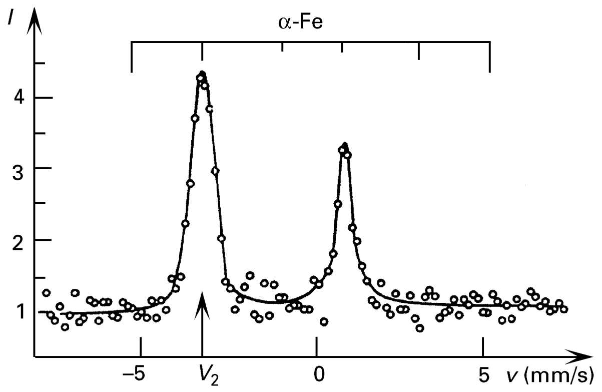



**Figure 9** SEDM spectrum of α-Fe.

splitting. The relaxation times in the region of $10^{-7}$–$10^{-10}$s are the most convenient to measure. Unfortunately, the necessity of having two successive resonant interaction processes results in a very low detected intensity. Indeed, the second part of a SEDMS experiment is a transmission experiment with the scatterer being the Mössbauer source.

Also, special Mössbauer spectrometers are used for total external reflection (TER) studies. On reflection at angles less than $\gamma_{cr}$ the electromagnetic field intensity falls off rapidly (for the metal iron mirror, $\gamma_{cr} = 3.8 \times 10^{-3}$ sr). The penetration depth for the radiation (i.e. the thickness $L$ of a layer under study) is taken to be equal to the depth at which the intensity is less by times e. If only the elastic scattering by electrons is considered, $L$ is evaluated to be 1.3 nm for an iron mirror.

An experimental set-up is given for studies of TER of Mössbauer quanta in **Figure 10**. The design of the Mössbauer spectrometer for TER studies ensures: (1) simple and reliable setting and measurement of the grazing angle $\gamma_{cr}$; (2) convenience in the adjustment of the angular beam divergence; (3) sample replacement without affecting the experimental geometry; (4) reproducibility of all source–collimator–sample distances; (5) sample rotation in the range 0–90°. The spectrometer consists of the analytical unit and electronic system for control, acquisition and processing of spectrometric data.

The analytical unit of the spectrometer comprises a vibration damping platform suspended on shock-absorbers. Mounted on the platform are guides of the 'wedge slide' type, which carry the driver, shielding screens, collimator to form narrow directed plane-parallel radiation beams, proportional counter and scintillation detector. A narrow plane-parallel γ-ray beam from the source rigidly attached to the driver is formed by the slit collimator and, through the entrance window of the dual detector, falls on the sample. The γ-radiation is reflected from the sample surface and passed through the exit window of the dual detector and slotted mask (screen), and detected by the scintillation detector D1.

Although the analysed layer is very thin, the technique has not been widely used due to the very low luminosity. Of no less importance is the fact that

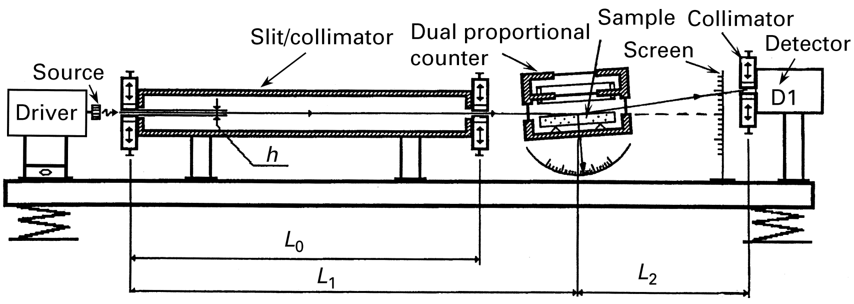



**Figure 10** An experimental set-up for studies of total external reflection of Mössbauer quanta. D1, scintillation detector. $L_0$ ~ 600 mm, $L_1$ ~ 700 mm, $L_2$ ~ 400 mm, $h$ = (1 ± 0.05) mm.

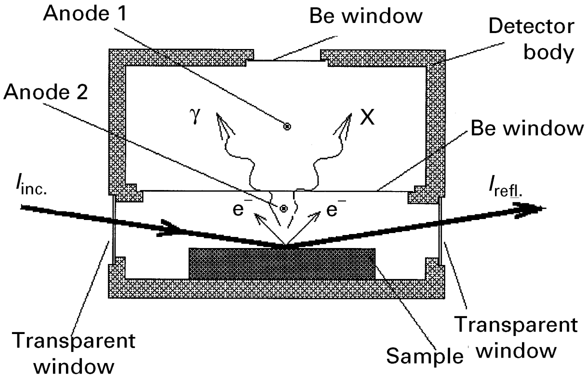



**Figure 11** Part of an experimental set-up (see **Figure 10**): the dual proportional counter.

interference effects complicate the interpretation of the experimental data. Substantial progress is achieved by detecting not only the mirror-reflected γ-quanta, but all secondary radiation leaving the surface when Mössbauer radiation is incident at an angle that is less than critical. The key part of an experimental set-up is the dual proportional counter (see **Figure 10**). A schematic picture of the dual-chamber gas proportional counter is shown in **Figure 11**. The sample under investigation is inside the electron chamber of the detector. The gas mixture in the chamber is He + 8% $CH_4$. The gas mixture for detection of γ- and X-rays is Ar + 8% $CH_4$. Thus during a single run (preset $\gamma$ value) one can obtain four Mössbauer spectra simultaneously: three from the combined detector and one from the scintillation detector (mirror-reflected γ-rays).

## Spectrum quality and quantitative information from Mössbauer spectra

The amplitude of Mössbauer lines in scattering experiments can often be greater than in a transmission geometry. However, the intensity loss of the scattered radiation of about two orders of magnitude makes it necessary to compare both the sensitivity of the two methods and the quality of the two spectra obtained. For a thin sample characterized by a single Lorentzian and the effective thickness $t_a$, the quality of the spectrum in relation to the quantity of information on the $t_a$ parameter (the information matrix element of interest), $[\hat{J}]_{t_a,t_a}$, is:

$$[\hat{J}]_{t_a,t_a} = \frac{\pi}{4}\varepsilon^2(0)I(\infty)\Gamma$$

where $\varepsilon(0)$ is the resonance effect magnitude. In order to increase $\varepsilon(0)$, the experimentalist needs to decrease the solid angle towards the detector and sample to prevent the source radiation from reaching the detector as a result of multiple nonresonant scattering in collimators and surrounding materials. This always gives a greater $\varepsilon(0)$ value, but the $I(\infty)$ value is decreased. The expression allows the evaluation of the limit when a further increase of $\varepsilon(0)$ values is no longer reasonable. After the optimal experimental conditions are chosen, the $\varepsilon^2(0)$ values are fixed for each sample under investigation. The quality of the spectrum is determined by the product $I(\infty)\Gamma$ and, as well as $I(\infty)$, it is also proportional to the measuring time.

In any spectroscopy, the intensity of the detected radiation may be written in the form:

$$I(\mathcal{E}) = C\int_{-\infty}^{\infty} L(E-\mathcal{E})\varphi(E)\mathrm{d}E + \xi(\mathcal{E})$$

where $\mathcal{E}$ is an energy parameter depending on the experimental setup, $L(E-\mathcal{E})$ is the instrumental line, $\varphi(E)$ is a function describing the response of the substance under investigation to monochromatic radiation, and $\xi(\mathcal{E})$ is the noise due to the stochastic processes.

In Mössbauer spectroscopy, any sample is characterized by $\mu_a(E)$. The simplest situation for recovering the $\mu_a(E)$ function from experimental data is in transmission spectroscopy, where $\varphi(E) = \exp[-\mu_a(E)d]$. There are two ways to find the $\mu_a(E)$ function. The first involves a hypothesis concerning the nature of this function. Analysis of an experimental spectrum amounts to the determination of the parameters characterizing $\mu_a(E)$ in accordance with the hypothesis.

The second way is connected with natural assumptions only on the nature of the $\mu_a(E)$ functions, for example, their smoothness. If there are no grounds for choosing a hypothesis, a certain initial assumption is made as to the nature of the required function. This often amounts to a search for an expression describing the response of the medium to monochromatic radiation, and sometimes 'an enhanced resolution of the method' is spoken of. The idea is that the best quality of the spectrum is attained using a source with a line shape described by the δ-function. Some methods of enhanced signal recovery have been developed for Mössbauer spectroscopy. As in sensitivity or resolution enhancement in other types of spectroscopy, a compromise has to be made between sensitivity and line width, as increasing the resolution always causes a decrease in sensitivity. Other types of data processing have been used to minimize distortion introduced by the measuring instruments.

## List of symbols

$E$ = energy; $\mathcal{E}$ = energy parameter depending on the experimental setup; $E_0$ = initial energy of electrons; $I(0)$ = intensity on resonance; $I(\infty)$ = intensity off resonance; $I(\mathcal{E})$ = intensity of the detected radiation; $I(v)$ = intensity at any velocity $v$; $I(v_i)$ = amplitude line at $v_i$ position; $I(v,v_i)$ = experimental spectrum SEDMS; $[\hat{J}]_{t_a,t_a}$ = information matrix element of interest; $J^a(E)$ = absorption line; $J_M(E)$ = emission line; $L(E–\mathcal{E})$ = instrumental line; $R$ = energy resolution; $t_a$ = effective thickness of the sample; $v$ = relative velocity; $\gamma_{cr}$ = angle of total reflection; $\Gamma$ = full width at half maximum; $\Gamma_{nat}$ = natural line width; $\delta$ = isomer (chemical) shift; $\varepsilon(0)$ = resonance effect magnitude; $\theta$ = direction electrons leaving the scatterer with an energy $E$; $\mu_a(E)$ = total linear absorption coefficient; $\xi(\mathcal{E})$ = noise due to the stochastic processes; $\varphi(E)$ = function describing the response of the substance under investigation to monochromatic radiation.

*See also:* **Calibration and Reference Systems (Regulatory Authorities); Mössbauer Spectroscopy, Applications; NMR Spectrometers; Quantitative Analysis.**

## Further reading

Andreeva MA, Belozerski GN, Grishin OV, Irkaev SM, Nikolaev VI and Semenov VG (1993) Mössbauer total external reflection: A new method for surface layers analysis. I. Design and developing of the Mössbauer spectrometer. *Nuclear Instruments and Methods* **B74**: 545–553.

Atkinson R and Cranshaw TE (1983) A Mössbauer backscatter electron counter for use at low temperature. *Nuclear Instruments and Methods* **204**: 577–579.

Balko B (1986) Investigation of electronic relaxation in a classic paramagnet by selective excitation double-Mössbauer techniques: Theory and experiment. *Physical Review B* **33**: 7421–7437

Bäverstam U, Bohm C, Ekdahl T and Liljequist D (1975) Method for depth selective ME-spectroscopy. In: Gruverman IJ and Seidel CW (eds) *Mössbauer Effect Methodology*, Vol 9, pp 259–276. New York: Plenum Press.

Belozerski GN (1993) *Mössbauer Studies of Surface Layers*. Amsterdam: Elsevier Science.

Flin PA (1975) Mössbauer backscattering spectrometer with full data processing capability. In: Gruverman IJ and Seidel CW (eds) *Mössbauer Effect Methodology*, Vol 9, pp 245–250. New York: Plenum Press.

Lippmaa M, Tittonen I, Linden J and Katila TE (1995) Mössbauer NMR double resonance. *Physical Review B: Condensed Matter* **52**(**14**): 10268–10277.

Meisel WP (1996) Surface and thin film analysis by Mössbauer spectroscopy and related techniques. In: Long GJ and Grandjean F (eds) *Mössbauer Spectroscopy Applied to Magnetism and Materials Science*, Vol 1, pp 1–30. New York: Plenum Press.

Nasu S (1996) High-pressure Mössbauer spectroscopy with nuclear forward scattering of synchrotron radiation. *High Pressure Research* **14**(**4–6**): 405–412.

Pasternak MP and Taylor RD (1996) High pressure Mössbauer spectroscopy: The second generation. In: Long GJ and Grandjean F (eds) *Mössbauer Spectroscopy Applied to Magnetism and Materials Science*, Vol 2, pp 167–205. New York: Plenum Press.

Schaaf P, Kramer A, Blaes L, Wagner G, Aubertin F and Gonser U (1991) *Nuclear Instruments and Methods in Physics Research* **B53** (**2**): 184–188.

Weyer G (1976) Applications of parallels-plate avalanche counters in Mössbauer spectroscopy. In: Gruverman IJ and Seidel CW (eds) *Mössbauer Effect Methodology*, Vol 10, pp 301–320. New York: Plenum Press.

# Mössbauer Spectroscopy, Applications

**Guennadi N Belozerski**, St.-Petersburg State University, Russia

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Applications**

The application of Mössbauer spectroscopy in diverse fields of qualitative and quantitative analysis is based on the ease with which hyperfine interactions can be observed. The information obtained from Mössbauer spectroscopy may be correlated with other methods by which HI can be examined such as NMR, EPR, ENDOR, PAC (perturbed angular correlations), nuclear orientation and neutron scattering. However, Mössbauer spectroscopy often proves to be experimentally simpler, more illustrative and an efficient method for studying applied problems. Mössbauer nuclei are ideal 'spies' supplying information on both the microscopic and macroscopic properties of solids.

Three factors may be identified as responsible for the widespread use of Mössbauer spectroscopy in both fundamental and applied research. First is the highest relative energy resolution $R \sim \Delta E/E$ and rather good absolute energy resolution $\Delta E \sim \Gamma_{nat}$ (the natural line width) (sometimes ~ $10^{-9}$ eV). Secondly, the absolute selectivity of Mössbauer spectroscopy means that in each experiment a response is registered from only one isotope of the element. Thirdly, Mössbauer spectroscopy has a high sensitivity that is determined by the minimum number of resonant atoms needed to produce a detectable response. In transmission Mössbauer spectroscopy for $^{57}$Fe, a response is given by a monolayer with an area of the order of 1 cm$^2$. Also important is the absence of any limitation on experimental conditions other than that the sample should be a solid.

## Applications in physics

Mössbauer spectroscopy offers a resolution sufficient to measure the effect of differing gravitational potentials on frequency or time as predicted by Einstein. The sign of the effect can be reversed by inverting the sense of travel over a fixed vertical path. Pound and Rebka measured the gravitational red shift in a 22-metre tower and observed $-5.1 \times 10^{-15}$ shift in the $\gamma$-ray energy of $^{57}$Fe. The source–detector setup was interchanged every few days to allow comparison of the results from a rising $\gamma$-ray beam to those from a falling one. When all of these measurements were combined, they yielded a result 0.9970 ± 0.0076 times the result predicted from the principle of equivalence.

There have also been some applications of Mössbauer spectroscopy in nuclear physics, to measure quadrupole moments of long-lived nuclear states by observing the orientation of a state at very low temperatures through the intensity ratios in a Mössbauer transition. The spectra for the case of the nuclear orientation in the $-\frac{11}{2}$ state of $^{119}$Sn by a quadrupole interaction are shown in **Figure 1**. In this case no macroscopic orientation of the hyperfine fields is needed. At very low temperatures the hyperfine splitting of the $-\frac{11}{2}$ state leads to an alignment in the $+\frac{3}{2}$ state yielding different intensities of the two Mössbauer lines. The increase of the intensity of line A located at higher velocities is clearly seen in the 16 mK spectrum. This uniquely indicates that the quadrupole moments of the $-\frac{11}{2}$ state and of the $+\frac{3}{2}$ state

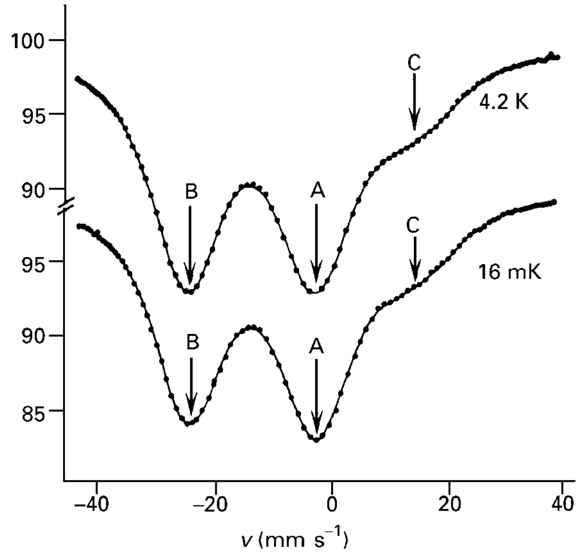



**Figure 1** Mössbauer spectra of the 23.9 keV transition in $^{119}$Sn with a source of $^{119}$Sn$^m$(OH)$_2$ (polycrystalline samples with one of the largest quadrupole splitting of ionic Sn$^{2+}$ compounds) at 4.2 K and 16 mK and a $^{119}$Sn:Pd (3 at% $^{119}$Sn) absorber chamber at a temperature of about 1.3 K are shown. The, $-\frac{11}{2}$ state decays by an *M*4 transition to the $+\frac{3}{2}$ state of $^{119}$Sn, which itself decays to the $-\frac{1}{2}$ ground state with the 23.9 keV *M*1 Mössbauer transition. The source was cooled inside the mixing chamber of the $^3$He/$^4$He dilution refrigerator specially designed for Mössbauer experiments. The weak line C at 1.5 mm s$^{-1}$ is attributed to Sn$^{4+}$ impurities in the source.

have the same sign. From a theoretical fit to the data it was deduced that $Q_{11/2} = -0.13 \pm 0.04$b. The same principle is used for measurements of temperature below 100 mK, i.e. by a $^{151}$Eu Mössbauer thermometer using an absorber of EuS.

More then 99.5% of all applications of Mössbauer spectroscopy are connected with hyperfine interaction parameters and structure factor determinations. An example of a sophisticated application is the study of the low temperature properties of magnetic impurities in metals that have an antiferromagnetic exchange interaction (Kondo effect). In order to study the very low temperature behaviour of a Kondo system, Mössbauer spectroscopy was used on two 'typical' Kondo systems, Fe:Cu and Fe:Au. The first system (Fe:Cu) showed expected Kondo-type properties – an extra polarization in the electron gas due to the correlations produced by the Kondo effect. The Fe:Au system, on the other hand, exhibited quite unexpected and striking results incompatible with those for Fe:Cu. In brief, in Fe:Au the temperature dependence of the susceptibility differed for $T > 10$ K from that for $T < 10$ K, yielding $\theta_1 = 10$ K and $\theta_2 = 0.5$ K for the Curie–Weiss $\theta$ of the two temperature regimes. Some data at low temperature are given in **Figure 2**. All data show that Fe:Au is clearly a system in which even at very low impurity concentrations the interaction effects are important at low temperatures and low external fields. Thus, it is difficult to extract parameters for the intrinsic Kondo behaviour of the isolated impurities from the experimental data on this system. The data show that the Kondo temperature for Fe:Au is most likely in the order of 10 K and not 0.5 K; the lower temperature seems to be connected with some kind of magnetic order.

Mössbauer spectroscopy is especially fruitful if the effective field, $H_{eff}$ is one of the parameters determining the shape of the spectrum. $H_{eff}$ is the resultant of the averaging of the hyperfine field $H_n$, but the way in which it is derived depends on the sample. The outcome is determined by comparing the frequency of the fluctuation of the $H_n$ direction with the Larmor precession frequency of the nucleus in the hyperfine field. In a ferromagnetic material, the temperature dependence of $H_{eff}$ is predicted to be the same as the temperature dependence of the expectation value of the atomic spin, and consequently of the magnetization. It thus provides a method for investigating magnetization without the application of external fields.

The study of the magnetic properties of materials and the dynamic aspects of magnetic interactions have been the most frequent applications of Mössbauer spectroscopy. The latter can be done

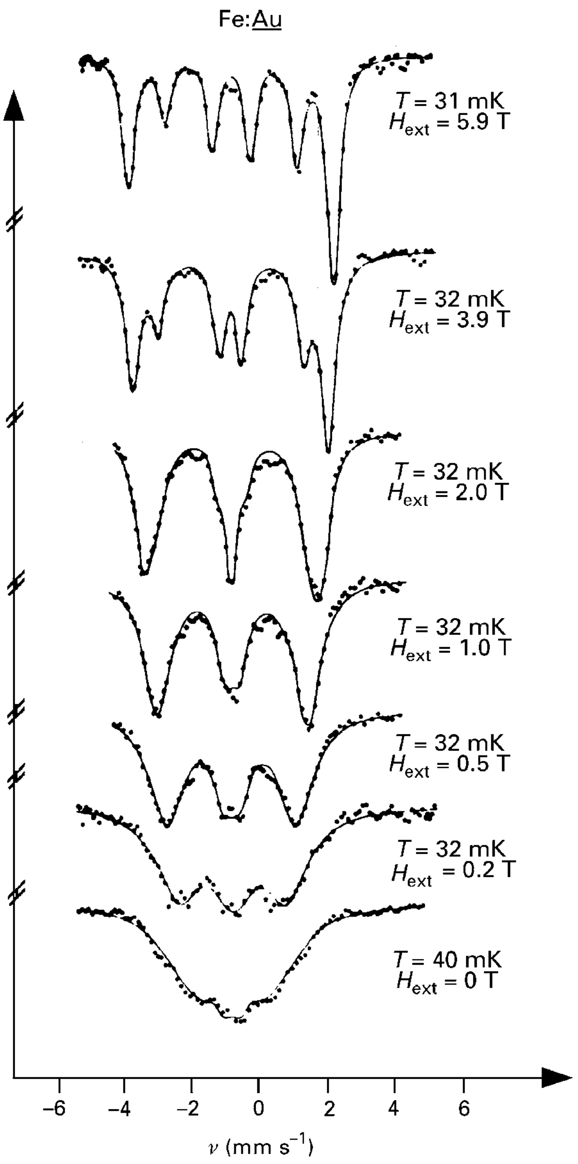



**Figure 2** Mössbauer spectra of the source of Fe:Au, prepared by diffusing $^{57}$Co into a foil of Au; the total impurity content (Co + Fe) was below 10 ppm. The source was attached to the mixing chamber of a $^3$He/$^4$He cryostat. The absorber was iron potassium hexacyanide at 1.3 K in the same external field ($H_{ext} \leq 6.0$ T). The solid lines are fits to the spectra including line broadening due to relaxation effects or a distribution of hyperfine field.

particularly conveniently in the case of single-domain particles. A reduction in the dimensions of single-domain particles increases the probability of thermal fluctuations of the magnetization orientations and this leads to superparamagnetic phenomena. One can suppress superparamagnetism either by cooling the sample or by applying an external magnetic field $H_{ext}$ (see **Figure 3**). Such studies permit the determination of the magnetic anisotropy, the alignment of magnetic domains, the particle size, the magnetic interaction among microcrystals, and surface effects.

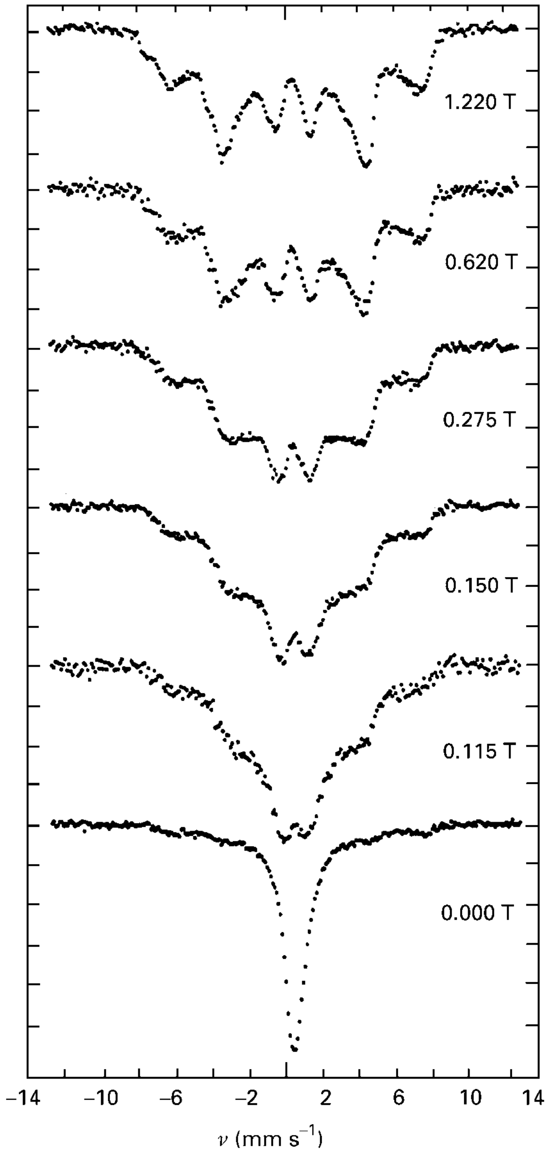



**Figure 3** Spectra of microcrystals of $Fe_3O_4$ obtained at 260 K in various applied magnetic fields. The iron atoms at the octahedral and the tetrahedral sites in $Fe_3O_4$ have different hyperfine field with opposite directions, and the corresponding Mössbauer lines are not completely resolved even in external field resolved $H_{ext}$ = 1.22 T.

A final illustration of the application of Mössbauer spectroscopy in physics is the study of spin texture in amorphous metals. Mössbauer spectroscopy provides information on both magnetic anisotropy and the texture of the principal electric field gradient (EFG) axes that reflects crystallographic features of the solid-state structure. It is known that spin texture in glassy metal ribbons, formed by high-speed quenching of fluid metals, is extremely sensitive to strains and stresses. This is partly consistent with the very low crystalline anisotropy in these materials. The results of a study of $Fe_{80}B_{20}$ are presented in **Figure 4**. The Mössbauer spectra for the contact with the quenching support and noncontact surfaces of the band are different. These differences in magnetic properties of the surfaces are due to the different cooling conditions during the band production. As the result of texture determination by Mössbauer spectroscopy, a distribution function $D(\theta_m, \Phi_m)$ (see **Figure 5**) was found for the contact and noncontact surfaces of the band, which describes the distribution of the relative volumes wherein $H_{eff}$ or the EFG are oriented along the $(\theta_m, \Phi_m)$ direction in a unit solid angle. Consideration of **Figure 4** shows that the magnetic anisotropy is along the *y*-axis. Hence, the magnetic dissipation fields on the alloy surface are significantly weaker. This is indicative of the high magnetic quality of the surface of the $Fe_{80}B_{20}$ alloy. This means that the local environment of Fe atoms is the same for the two sides of the band, i.e. there is no difference in the concentrations of Fe atoms and the metalloid atoms.

## Applications in chemistry

The three components of the hyperfine interaction – the isomer shift, the nuclear quadrupole splitting and the nuclear magnetic dipole splitting – have immediate chemical application. The isomer shift $\delta$ measures the charge density of the atomic electrons at the nucleus and is therefore directly related to chemical bonding and covalency. The ligand field splitting energy $\Delta$ depends on the gradient of the electric field produced by the other ions in the lattice, i.e. it is related to the point symmetry of the lattice information. $H_{eff}$ provides a sensitive tool for the detection of magnetically ordered states. Mössbauer spectroscopy is a major tool for studying the formation of new inorganic materials and probing structural and magnetic phase transformations in inorganic compounds. This can be demonstrated with some specific examples. As the first example, the participation of 3d and 4s electrons in the chemical bond of iron in a complex determines the total electron density at the Fe nucleus (see **Table 1**), and hence the isomer shift for the Fe complex. A change in the oxidation state influences the electron density at the nucleus in a major way, since such a change is usually associated with the removal (or the addition) of one or more valence electrons. Thus the two oxidation states differ in their outer valence electron number by an integer, and this results in a major difference in the isomer shift for the two oxidation states. This is illustrated in **Figure 6**, where the isomer shifts for monovalent, trivalent, and pentavalent gold

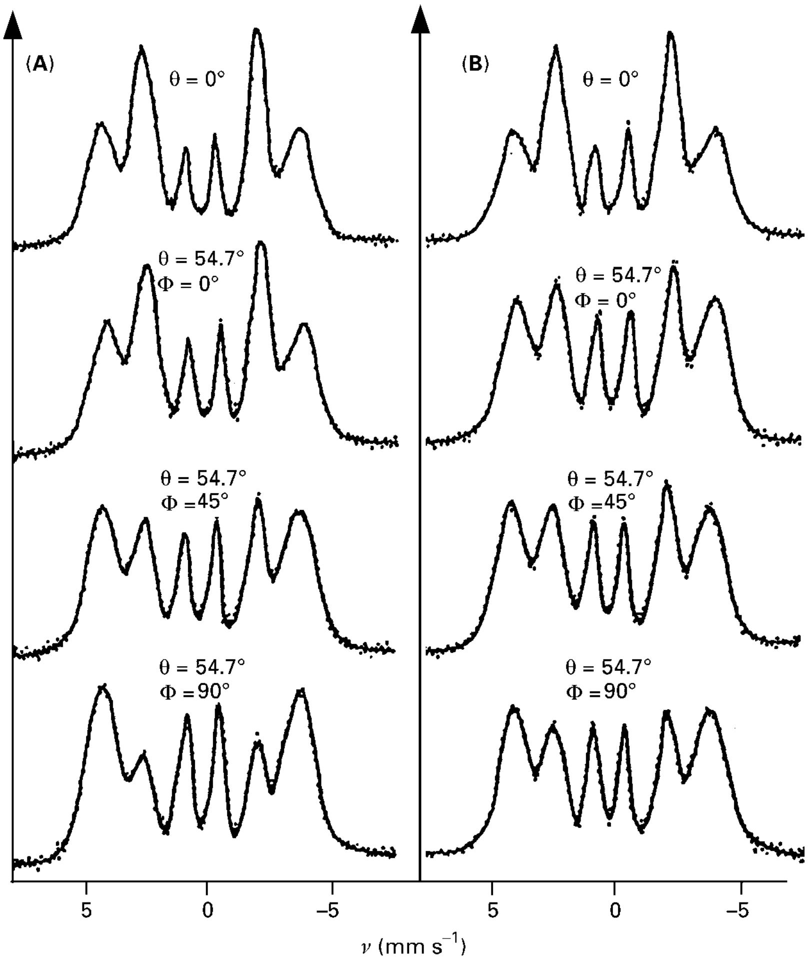



**Figure 4** Mössbauer spectra from a band of an amorphous alloy: (A) contact surface, (B) noncontact surface. In accordance with the foil preparation, the *z* principal axis is perpendicular to the foil surface, the *x* axis is along the band and the *y* axis is perpendicular to the rolling direction. The direction of the incident Mössbauer quanta is specified by the angles $\theta$ and $\Phi$.

complexes with various halide ligands are shown. The electron density $\rho(0)$ is the least in Au(I) compounds and the highest in Au(III) compounds. The spread in the isomer shift for the selected halide ligands is small within an oxidation state.

The ability to determine oxidation states is demonstrated in **Table 2** for the isomer shift data of fluorides of neptunium for various oxidation states represented by the $5f^n$ configuration.

**Table 1** Isomer shifts, electron configurations, and value of the electron density difference $\Delta\rho(0)$ for matrix-isolated Fe ions

| *Ion* | *Configuration outside Ar core* | *Isomer shift*[a] *(mm s$^{-1}$)* | *$\Delta\rho(0)$ (au)* |
|---|---|---|---|
| $Fe^0$ | $3d^64s^2$ | −0.75±0.03 | 16.0 |
| $Fe^+$ | $3d^7$ | +1.77±0.08 | 0.0 |
| $Fe^+$ | $3d^64s$ | +0.26±0.05 | 10.5 |

[a] Relative to Fe metal at 300 K.

The ice produced by the slow freezing of solutions that contain Mössbauer atoms, and the glass produced by their rapid freezing, are also suitable for investigation by Mössbauer spectroscopy. Systematic Mössbauer examinations of ices were started with investigations of the polymorphic transformation

**Table 2** Isomer shifts for the various fluorides of Np and the values of $\rho(0)$ for electron configurations, $5f^n$, nominally representing the oxidation states in fluorides

| *$5f^n$* | *Oxidation state* | *Electron density (au)* | *Isomer shift*[a] *(mm/s$^{-1}$)* |
|---|---|---|---|
| $5f^4$ | Np(III) | 6 965 162 | +34(1) |
| $5f^3$ | Np(IV) | 6 965 343 | −9(1) |
| $5f^2$ | Np(V) | 6 965 555 | −37(1) |
| $5f^1$ | Np(VI) | 6 965 793 | −63(2) |
| $5f^0$ | Np(VII) | 6 966 057 | −77(2)[b] |

[a] Relative to $NpAl_2$ at 4.2 K.
[b] For $Li_5NpO_6$.

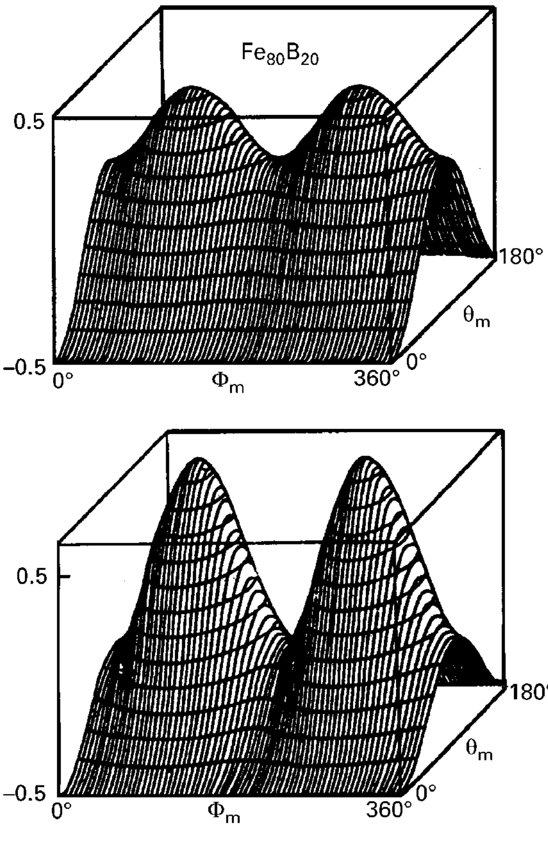



**Figure 5** Texture distribution functions $D(\theta_m, \Phi_m)$ for the band of amorphous alloy $Fe_{80}B_{20}$: (A) contact surface, (B) noncontact surface. The direction of the quantization axis is determined by the angles $\theta_m$ and $\Phi_m$.

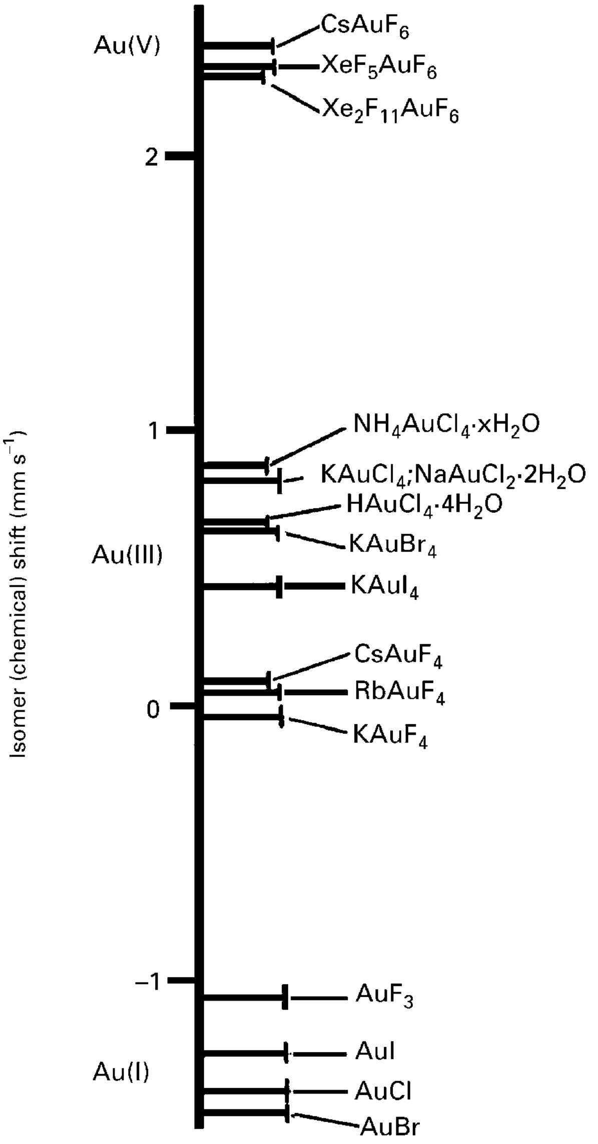



**Figure 6** Isomer shift scale for Au(I), Au(III), and Au(V) halides.

that occurs in the ice as a result of changes in temperature. Detailed investigations showed how the structure of the ice formed depends on the rate of freezing, and how this affects the Mössbauer parameters measured on the frozen solution. If rapid freezing is applied to equilibrium solution systems in which the equilibrium shifts rapidly (within 1–2 s) and measurably as a result of the change in temperature, Mössbauer spectroscopy provides reliable information only if the temperature of thermostating before the freezing is close to the solidification point.

Research into nonaqueous solvents and solutions is also worth note. When metal salts are dissolved in various solvents, it is possible to study the coordination of the solvent molecules to the cation and the extent of association between the cation and its anion.

The phenomenon of spin transition (spin crossover), is another example of a fruitful application of Mössbauer spectroscopy in chemistry. The phenomenon is observed in transition metal complexes with $d^4$, $d^5$, $d^6$, $d^7$ and $d^8$ electron configurations. Depending on the ligand field splitting energy, $\Delta$, relative to the mean spin pairing energy, $E_p$, the ground state of a particular transition metal complex shows either high-spin (HS) behaviour or low-spin (LS) behaviour. If $|\Delta - E_p| \sim k_BT$, both HS and LS spin states may coexist in thermal equilibrium: HS $\rightleftarrows$ LS.

Mössbauer spectroscopy reveals distinct resonance lines for the individual spin states provided their individual mean lifetimes are comparable to or longer than the lifetime $\tau$ of the nuclear excited state. If one determines the area fractions of the spin states involved (which in most cases come close to the actual molar fractions of HS and LS molecules), one may follow quantitatively the spin conversion as a function of temperature. The simplest way to observe the transition is to change the temperature. An unusual way to affect the spin transition behaviour is to

change the elastic properties of the crystal lattice, which are important for the spin transition behaviour. This was done by a partial isomorphous substitution of the central metal ions in mixed crystals of $[Fe_xMn_{1-x}(phen)_2(NCS)_2]$ by manganese ions, which differ in the volume and the electron configuration from the replaced metal ions (see **Figure** 7). The transition temperature shifts to lower temperatures upon dilution, and the residual paramagnetism at low temperatures increases systematically with decreasing iron concentrations $x$.

There are many other fields of application of Mössbauer spectroscopy in chemistry, including coordination chemistry, catalysis (heterogeneous), mixed-valence compounds, glasses, oxide and oxyhydroxide studies, and actinide chemistry.

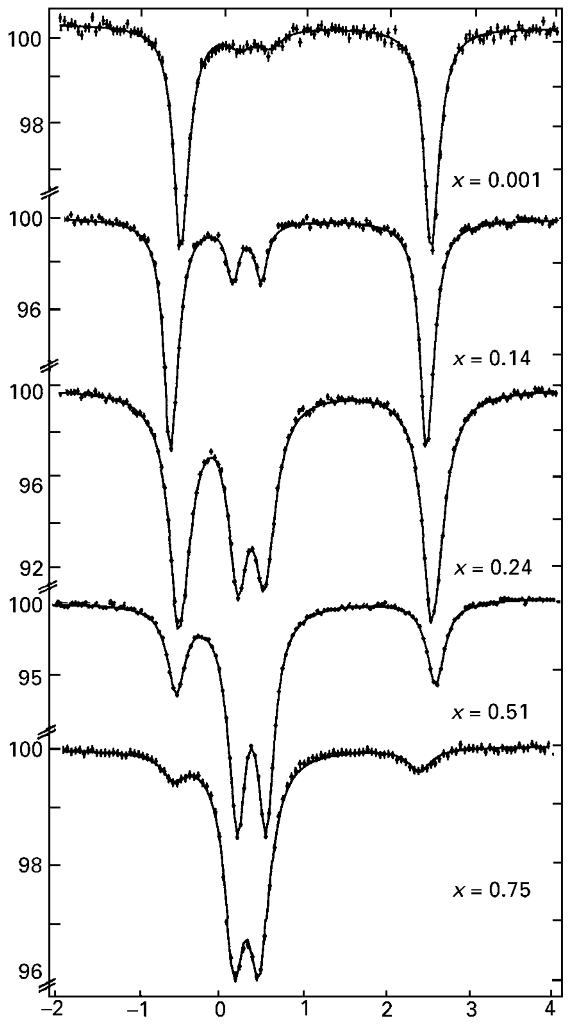



**Figure 7** Mössbauer spectra of $[Fe_xMn_{1-x}(phen)_2(NCS)_2]$ at 5 K and various iron concentrations $x$. The spectra demonstrate that, with increasing dilution of iron by manganese, the intensity of the quadrupole doublet of the high-spin state of iron(II) (outer two lines) increases steadily at the expense of the low-spin quadrupole doublet of iron(II) (inner two lines).

## Applications in biology and geology

Many biological molecules contain iron and Mössbauer spectroscopy is a useful tool for the study of proteins and enzymes. The measurement of the magnetic properties of transition metal elements in biological molecules by MS, NMR and EPR is an important way of characterizing the electronic state of the metal ion, and hence of providing a clue to the structure and function of the molecule. Mössbauer spectroscopy may be used to study their chemical state and bonding and to obtain qualitative data on the local structure and symmetry in their neighbourhood.

Haem proteins are the best-understood of these molecules, and the first systematic study of biological molecules using the Mössbauer effect was done on haemoglobin and its derivatives. Since then a great deal of work has been done on iron–sulphur proteins and on iron-storage proteins. Magnetic susceptibility data showed that the chemical state of the iron atom and its spin state are very sensitive to the nature of the sixth ligand. The Mössbauer spectra of these molecules have been valuable in confirming these earlier conclusions, and have yielded quantitative data on the way that the energy levels and wavefunctions of the iron atoms are affected by the ligand field and spin–orbit coupling in the protein. They have also been valuable in providing standard spectra for each of the four common states of iron, and a summary of the chemical shifts and quadrupole splittings measured at 195 K is shown in **Table 3**.

In disease, haemoglobin may become ferric (denoted Hi) and of low spin, e.g. in a haemoglobin cyanide HiCN, azide $HiN_3$ or hydroxide HiOH. The

**Table 3** Isomer shifts $\delta$ and quadrupole splittings $\Delta E_Q$ for haemoglobin at 195 K. The symbol Hi is used for ferric haemoglobin and Hb for ferrous haemoglobin

| *Material* | $\delta$ | $\Delta E_Q$ |
|---|---|---|
| HiF | (0.3) | (0.67) |
| $HiH_2O$ | 0.20 | 2.00 |
| HiCN | 0.17 | 1.39 |
| HiOH | 0.18 | 1.57 |
| Hb | 0.90 | 2.40 |
| $HiN_3$ | 0.15 | 2.30 |
| HiOH | 0.18 | 1.57 |
| Hb | 0.90 | 2.40 |
| HbNO[a] | – | – |
| $HbO_2$ | 0.20 | 1.89 |
| HbCO | 0.18 | 0.36 |

[a]Spectrum very broad.

most noticeable feature of the spectra is the appearance of magnetic hyperfine structure at 77 K. Spectra of HiCN at these temperatures are shown in **Figure 8**. At 4.2 K the magnetic hyperfine spectrum is well resolved, but is complex and asymmetrical owing to the anisotropy of the hyperfine interaction tensor.

Apart from haem proteins, there are studies of iron–sulfur proteins, iron transport and iron-storage compounds, iodine compounds and vitamin $B_{12}$. There are many publications connected with nitrogenase, oxygenase, hydrogenase and cytochrome P450–ferredoxin enzyme systems and medical and physiological applications.

Application of Mössbauer spectroscopy to crystallography, mineralogy and geology is very similar to chemical applications, with the main emphasis on phase analysis and structure determination. In view of the great importance of iron in the earth's crust and the widespread occurrence of this element in rock-forming minerals, earth scientists have naturally focused attention on applications of $^{57}$Fe Mössbauer spectroscopy. One of the most important groups of rock-forming minerals are the silicates, in which particular lattice position is often occupied by more than two atomic species. In these cases, accurate site occupancy numbers for each species cannot be obtained by diffraction alone. Mössbauer spectroscopy has been a useful tool for investigating the local properties of iron sites in complex crystal structures, particularly when employed on minerals with carefully defined (well-known) positional parameters.

Typical applications of Mössbauer spectroscopy to mineralogy and geology have been analysis of the oxidation state of iron at iron sites in minerals. The ferric to ferrous iron ratio reveals important information on the partial pressure of oxygen during crystallization, which is a parameter of geological significance. The study of area ratios of distinct hyperfine patterns has led to thermodynamic analyses of order–disorder phenomena in minerals. Application of Mössbauer spectroscopy to poorly crystallized materials of geological relevance may be fruitful. Studies have been made on hydrolysis and absorption in clay minerals and on coordination polyhedra of iron in glasses.

Mössbauer spectroscopy has made a modest contribution to the study of composite samples and separated mineral phases from the moon. The constituents of the lunar soil are of a highly complex nature. The primary aim has been to study the amount of iron in the soil, the distribution of iron over the soil constituents, and the particular valence states. The soils at the moon landing sites do not represent averages of the collected rocks. The spectrum of the soil (**Figure 9**) is the superposition of patterns of $^{57}$Fe in silicate minerals, silicate glasses, ilmenite, metallic iron, and troilite.

The identification of the magnetic patterns of metallic iron and FeS and the quadrupole-split pattern of paramagnetic $FeTiO_3$ is generally no problem in

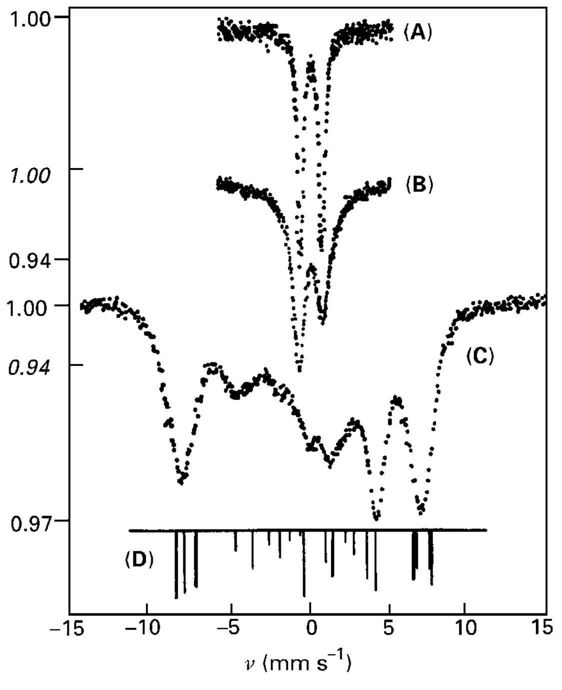



**Figure 8** Mössbauer spectra of a low-spin ferric haemoglobin HiCN – at (A) 195 K, (B) 77 K and (C) 4.2 K. The broadening at 77 K relative to 195 K is due to the slower electron spin relaxation rate. At 4.2 K the relaxation is so slow that the hyperfine pattern is resolved, but complex.

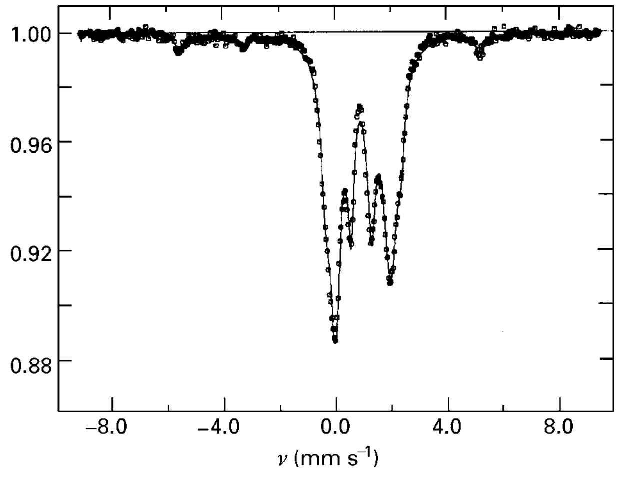



**Figure 9** Resonant absorption spectrum of $^{57}$Fe (295 K) in lunar soil from the Apollo 11 landing site at Mare Tranquillitatis. The absorption in the range between 0 and 0.3 mm $s^{-1}$ results primarily from iron in silicate glass, pyroxene and olivine; the peaks at lower and higher velocities are due to metallic iron.

spectra from soil absorbers kept at room temperature (**Figure 9**). However, the paramagnetic silicates in the soil cannot be identified easily because of the substantial overlap of their quadrupole-split patterns.

## Surface layer studies

Mössbauer spectroscopy has developed as one of the few methods available for investigation of solids differing in depth by several orders of magnitude. These can be layers with the depth of less than 1 nm (scattering at glancing angles) as well as layers that are about 10 μm deep (electron and X-ray detection). The technique is able to examine solids over a wide range of compositions, gaseous pressures, and temperatures and under conditions of practical interest. It should also be noted that Mössbauer spectroscopy is one of the best methods for *in situ* characterization of solid–solid and solid–solution interfaces. This lends itself to *in situ* studies of surfaces under various coatings and processes, surface magnetism and the effect of the gas phase on the properties of the surface layers and the structure and magnetic properties of epitaxially grown monolayers on the surface of oriented single crystals. The techniques of surface layer analysis find extensive applications in science and industry.

The thinnest layers to be studied in Mössbauer spectroscopy are those studied by GA DCEMS and (total external reflection) TER. For example, the TER method has been used to study the initial stage of oxidation of $\alpha$-Fe foil. The substantial progress in surface studies using glancing angles was achieved by simultaneous detection of electrons, $\gamma$-rays or X-rays as well as mirror reflected $\gamma$-rays, which allows reconstruction of the process of surface layer formation. The sample under investigation was inside the electron chamber of the detector. The results of the calculation and the fitting of the experimental spectra are very sensitive to the model of surface layer structure and to the density of the top layer.

For phase modification of the surface layer, the initial sample (sample I) was oxidized at 150°C in air for 4 hours (to give sample II). From experimental spectra (see **Figure 10**) it can be concluded qualitatively that in the top layers iron is oxidized and is characterized by different spectra with and without magnetic splitting. The sample II spectra in the right column of **Figure 10** clearly illustrate the phase transformation on the top of the α-Fe film. The spectra differ mostly at the lower grazing angles, where the penetration depth is minimal. For example, the contribution of the doublet is much larger for the spectrum of sample II at $\theta = 2.2$ mrad.

Thus even qualitative analysis clearly shows that the technique is more sensitive to the changes of chemical and magnetic states of ultra-thin surface layers than the usual CEMS techniques. Detailed information about distributions of different phases

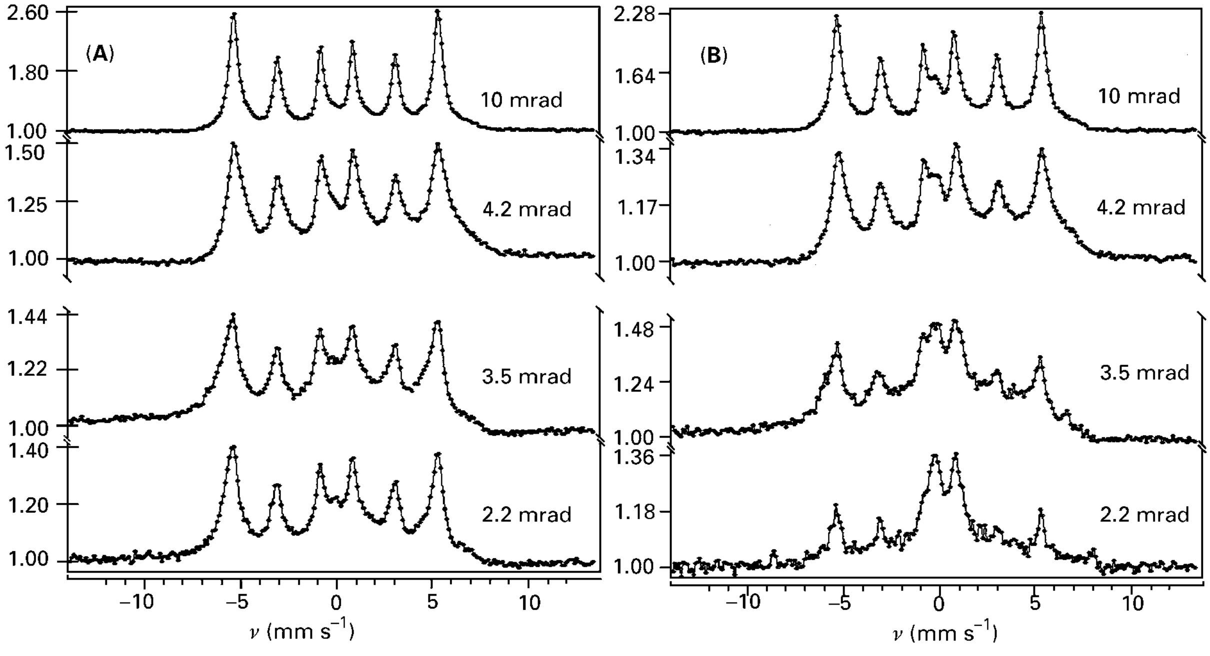



**Figure 10** Sets of TER-CEM spectra (A) for the initial sample $\alpha$-$^{57}$Fe (enriched up to 90% and ~ 50 nm thick), sputtered on a 10 mm thick beryllium disk (sample I) and (B) after oxidizing in air for 4 hours at 150°C (sample II). TER take place at grazing angle $\theta$ ~ 3 mrad.

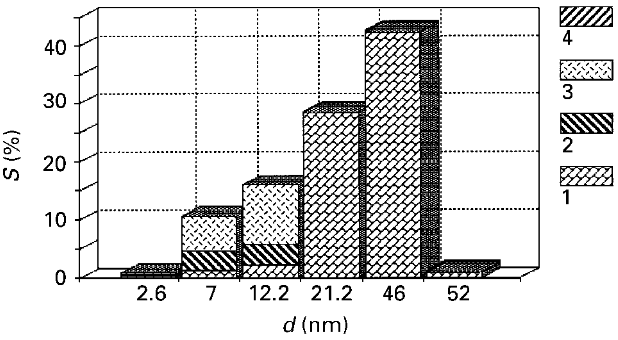



**Figure 11** The histograms of iron content $S$ in different phases of different sublayers (sample I) as function of the lower border of the sublayers. The sample is an iron film about ~ 50 nm thick, sputtered on the beryllium disk 10 mm thick. 1; sextet of $\alpha$-Fe; 2, broadened sextet with smaller magnetic splitting than in $\alpha$-Fe; 3, asymmetric strongly broadened sextet associated with hyperfine interaction parameters for $\alpha$-FeOOH; 4, the doublet.

with depth may only be obtained after quantitative analysis of sets of spectra. The results of calculations for the initial sample are given in **Figure 11**. It is seen that the $\alpha$-Fe phase is present in the top layer also. This points to the existence of an island phase structure of the top layer. The amount of FeOOH phase increases with depth but has a maximum value at ~10 nm. From analysis of histograms such as that shown in **Figure 11**, it can be seen explicitly how different oxide phases move inside the sample during oxidation processes.

## Industrial applications

The scope of problems addressed is extremely wide and includes metals research of industrial significance, steel and steel alloys, amorphous alloys and glasses, corrosion, surface magnetism and superparamagnetism, catalysis and solid-state reactions on solid surfaces, minerals, and mineral processing, superconductivity including high-temperature superconductivity, nanocrocrystals, thin films, ion implantation and laser treatment of metals, the coal industry, amorphous alloys and glasses, and numerous other miscellaneous applications. Areas of application of Mössbauer spectroscopy to corrosion studies are listed in **Table 4**.

Physical metallurgy is a rather wide field of applications of Mössbauer spectroscopy and it is possible to enumerate only the main topics: phase analysis, order–disorder alloys, surfaces, alloying, interstitial alloys, steel, ferromagnetic alloys, precipitation, diffusion, oxidation, lattice defects; etc. Alloys are well represented by the iron–carbon system, the mechanism of martensite transformation, high-manganese and iron–aluminium alloys, iron–silicon and Fe–Ni–X alloys.

**Table 4** Corrosion studies using Mössbauer spectroscopy

| *Processes studied* | *Corrosive environments, objectives, applications* |
|---|---|
| Passivation | Solutions and electrode surfaces in solutions at various potentials<br>Properties and structure of passive films in solutions, in gases or in vacuum |
| Corrosion in water and in aqueous solutions | Distilled water<br>The effect of oxygen<br>The effect of other admixtures in water<br>The temperature effect in aqueous media<br>Water vapour corrosion<br>Corrosion in power plants |
| Corrosion in gases | Pure oxygen or dry air<br>Atmospheric and water vapour corrosion<br>Aggressive gases<br>Combined action of both gas media and solutions |
| Corrosion in aggressive environments | Acids<br>Alkalis<br>Organic and natural media<br>Corrosion in tubing and autoclaves<br>Applications in agriculture |
| Specific corrosion processes | Stress corrosion<br>Corrosion beneath lake and polymeric coatings<br>Transformations of corrosion products to enhance the protective properties and to identify the corrosion products |
| Inhibition and passivation | The effect of special inhibiting or passivating admixtures on the on the composition and growth rate of protective films. |
| Corrosion of amorphous alloys | Materials science<br>A check on the theory used to describe corrosion of amorphous alloys |
| Internal oxidation | Materials science |

There have been a number of regional and international conferences devoted to industrial applications of Mössbauer spectroscopy or specific problems of materials science. The scope of problems addressed is extremely wide and ranges over all the above-mentioned problems. One can expect the contributions of Mössbauer spectroscopy in industry to divide into three areas: (1) as a research tool, (2) in quality control, and (3) for in-service evaluation. Unfortunately, there are still very few Mössbauer instruments in use for quality control, or for in-service evaluation of materials or surface. More than 99% of Mössbauer publications cover use of the technique as a research tool.

An example of the application of Mössbauer spectroscopy to metals research of industrial significance is the study of fractures and ruptures in steels (see **Figure 12**). The phase compositions of the metal microvolumes that are the sites for the propagation of fractures were determined. The spectra obtained from the fracture surfaces after various cyclic loadings are a superposition of a sextet with a single line in the centre ($\gamma$- and $\varepsilon$-phases). The central line is identical to that of the original sample, the sextet lines being broadened. Such line broadening may be attributed to different local environments of Fe atoms, or to significant local lattice distortions, or to both. It follows from **Table 5** that on increasing the loading rate the surface layer of the fracture is enriched in martensite. The appearance of the $\alpha$-phase at the site of destruction and its predominance there explains the contradiction between the intercrystallite character of the steel destruction and the high energy that is required. Martensite in the fracture layer is formed in front of a crack to be nucleated and propagated. The predominance of martensite in the fracture layers testifies to the large number of cracks in this area and to a significant stress relaxation preceding the formation and development of the cracks. This is the reason for the high energy of destruction of steels with the initial ($\gamma + \varepsilon$) structure. A comparison of the CEMS and X-ray detection data shows that the processes responsible for the destruction of steels with the $\gamma + \varepsilon$ structure are localized in a zone with a width less than 5 µm. On static loading tests, the width increases up to 10 µm.

## Concluding remarks

The technique has grown rapidly and is now an effective form of spectroscopy. Mössbauer spectroscopy using $^{57}$Fe is so simple and straightforward that it is now an integral part of many undergraduate laboratory curricula.

An explosive growth in the application of Mössbauer spectroscopy in a broad range of

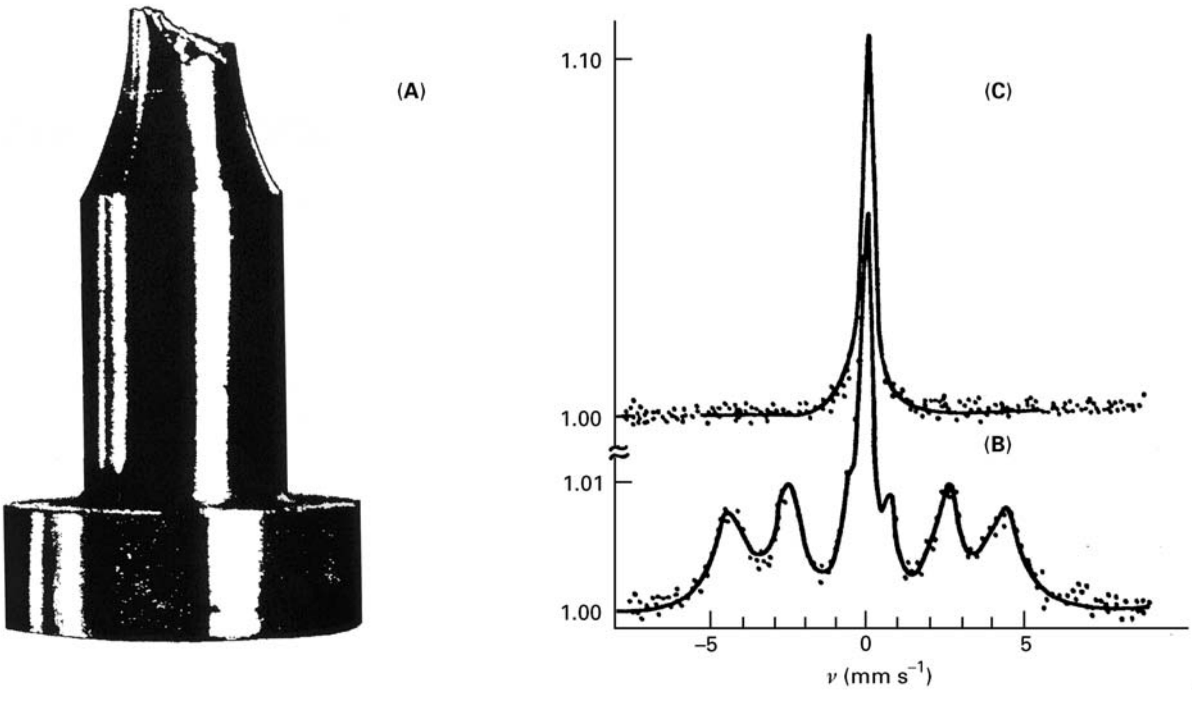



**Figure 12** (A) An 8 mm fracture of steel with the initial $\gamma + \varepsilon$ structure (19.2% Mn, 1.67% Si, 0.9% Ti, 0.07% C). (B) CEM spectrum of the steel from the fracture surface shown in (A). (C) CEM spectrum after annealing at 1050°C for 6 hours, followed by cooling in air.

**Table 5** Results of the phase analysis of iron-based alloy (19.2% Mn, 1.67% Si, 0.9% Ti, 0.07% C)

| *Loading* | *Detected radiation* | *Parameters of the $\gamma+\varepsilon$ phases* | | *Parameters of the magnetic phase* | | | *Phase composition (saturation effect and rescattering are accounted for)* | |
|---|---|---|---|---|---|---|---|---|
| | | $\delta_{\gamma+\varepsilon}$ | $\eta_{\gamma+\varepsilon}$ | $\langle H_{eff}\rangle$ | $\delta_\alpha$ | $H_\alpha$ | $C_{\gamma+\varepsilon}$ | $C_\alpha$ |
| Low-cycle fatigue | Electrons | −0.12 | 28 | 27.3 | −0.01 | 72 | 27.6 | 72.3 |
| | X-rays | −0.14 | 29 | 27.5 | −0.04 | 71 | 35 | 65 |
| Fatigue tests at 600 cycles min$^{-1}$ | Electrons | −0.12 | 10 | 27.5 | −0.01 | 90 | 8.8 | 91.2 |
| | X-rays | −0.14 | 26 | 27.5 | −0.01 | 74 | 31.9 | 68.1 |
| Impact tests | Electrons | −0.12 | 23 | 28.0 | −0.01 | 77 | 21.4 | 78.6 |
| | X-rays | −0.15 | 46 | 28.1 | −0.05 | 54 | 53.2 | 46.8 |

$\langle H_{eff}\rangle$ is the average effective magnetic field at a $^{57}Fe$ nucleus in the $\alpha$-phase; $\Delta\delta = \pm 0.01$ mm s$^{-1}$; $\eta_i$ are the relative areas (%) under the spectrum ($i = \alpha, \gamma+\varepsilon$); $C_i$ is the relative content of the $i$th phase (%), $\Delta\eta = \pm 5\%$, $C_\alpha = \eta_\alpha/(\eta_\alpha + 1.43\eta_\gamma)$. $H_{eff}$ values are given in tesla, isomer shifts are in mm s$^{-1}$

scientific areas has resulted in publications in a diverse number of journals. From January 1978, the Mössbauer Data Center in Virginia USA has published the *Mössbauer Effect Reference and Data Journal* monthly. This journal has become an invaluable resource for information on Mössbauer spectroscopy.

## List of symbols

$C_i$ = relative content of $i$th phase (%); $D(\theta_m, \Phi_m)$ = distribution function; $\Delta\rho(0)$ = value of the electron density difference; $E$ = energy; $E_P$ = mean spin pairing energy; $\Delta E$ = absolute energy resolution; $\Delta E_Q$ = quadrupole splittings; $\langle H_{eff}\rangle$ = average effective magnetic field; $H_{ext}$ = external field; $H_n$ = hyperfine field; $k_B$ = Boltzmann constant; $R$ = relative energy resolution; $T$ = temperature; $\Gamma_{nat}$ = natural line width; $\Delta$ = ligand field splitting energy; $\delta$ = isomer shift; $\Delta\delta$ = inaccuracy of the $\delta$ determination; $\eta_i$ = relative area (%) under the spectrum; $\Delta\eta$ = inaccuracy of the $\eta$ determination; $\theta$ = grazing angle; $\tau$ = lifetime of nuclear excited state; $\rho(0)$ = electron density.

*See also:* **Chemical Applications of EPR; Chemical Shift and Relaxation Reagents in NMR; Cosmochemical Applications Using Mass Spectrometry; Geology and Mineralogy, Applications of Atomic Spectroscopy; Industrial Applications of IR and Raman Spectroscopy; Interstellar Molecules, Spectroscopy of; Materials Science Applications of X-Ray Diffraction; Mössbauer Spectrometers; Mössbauer Spectroscopy, Theory; Solid State NMR, Methods; Stars, Spectroscopy of; Surface Studies By IR Spectroscopy; Zeeman and Stark Methods in Spectroscopy, Applications.**

## Further reading

Belozerski GN, Bohm C, Ekdahl T and Liljequist D (1982) A Mössbauer investigation of the surface of $\alpha$-iron. I. Corrosion and passivation in $H_2O_2$ solution. *Corrosion Science* **22**: 831–844.

Cadogan JM (1996) Mössbauer spectroscopy and rare-earth permanent magnets. *Journal of Physics D: Applied Physics* **29**: 2246–2254.

Clark SJ, Donaldson JD and Grimes SM (1996) Mössbauer spectroscopy. In: *Spectroscopic Properties of Inorganic and Organometallic Compounds*, Vol 29, pp 330–417. Cambridge: The Royal Society of Chemistry.

Cranshaw TE (1986) Metal research of industrial significance by Mössbauer spectroscopy. In: Long GJ and Steven JG (eds) *Industrial Applications of the Mössbauer Effect*, pp 7–24. New York: Plenum Press.

Gütlich P (1984) Spin transition in iron complexes. In: Long GJ (ed) *Mössbauer Spectroscopy Applied to Inorganic Chemistry*, Vol 1, 287. New York: Plenum Press.

Greneshe JM (1997) Nanocrystalline iron-based alloys nsinvestigated by Mössbauer spectroscopy. *Hyperfine Interactions* **110**: 81–91.

Hafner SS (1975) Mössbauer spectroscopy in lunar geology and mineralogy. In: Gonser U (ed) *Mössbauer Spectroscopy*, pp 167–199. New York: Springer-Verlag.

Jonson CE (1975) Mössbauer spectroscopy in biology In: Gonser U (ed) *Mössbauer Spectroscopy*, pp 139–166. New York: Springer-Verlag.

Morup S (1986) Industrial applications of Mössbauer spectroscopy to microcrystals. In: Long GJ and Stevens JG (eds) *Industrial Applications of the Mössbauer Effect*, pp 63–82. New York: Plenum Press.

Shenoy GK (1984) Mössbauer-effect isomer shift. In: Long GJ (ed) *Mössbauer Spectroscopy Applied to Inorganic Chemistry*, Vol 1, pp 57–78. New York: Plenum Press.

Steiner P, Belozerski GN, Gumprecht D, Zdrojewski W and Hufner S (1974) Impurity–impurity interaction in a very dilute Fe:Au system. *Solid State Communications* 157–160.

# Mössbauer Spectroscopy, Theory

**Guennadi N Belozerski**, St.-Petersburg State University, Russia

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| **Theory** |

Mössbauer spectroscopy is concerned with the scattering and emission of $\gamma$-radiation by atomic nuclei in the condensed phase. The phenomenon was discovered in 1958 by the German physicist R. L. Mössbauer. It makes use of the probability that the state of a system will remain unchanged when $\gamma$-quanta are absorbed or emitted with an energy that is exactly equal to the nuclear transition energy $E_0$.

Hence the $\gamma$-spectrum $J(E)$ of a Mössbauer source may be represented by the sum of a line $J_R(E)$ that is displaced due to recoil effects and broadened by the Doppler effect, and a line $J_M(E)$ with its centre at the energy that is exactly equal to the transition energy, the half-width being close to the natural one, $\Gamma_{nat}$. The $J_M(E)$ part of the spectrum is of particular interest and reveals itself most strikingly when the source and the sample under study (absorber) are in the solid state, see **Figure 1**. The following normalization conditions may be assumed:

$$\int_{-\infty}^{\infty} J_R(E)\,dE = 1 - f; \qquad \int_{-\infty}^{\infty} J_M(E)\,dE = f \quad [1]$$

A fraction $(1-f)$ of the disintegrations occurs with the energy transferred to the lattice, and the fraction $f$ is recoilless.

The particular dependence $J_R(E)$ is of no interest but it should be noted that the centre of gravity of this distribution is shifted by the amount $E_R$ relative to the transition energy $E_s$ in the source. $E_R = E_0^2/2Mc^2$ is the recoil energy imparted to an isolated nucleus of mass $M$ (where $c$ = vacuum speed of light). The energy distribution of the Mössbauer $\gamma$-quanta $J_M(E)$ may be considered as Lorentzian $L(E)$ with the full half-width $\Gamma$ such that $\Gamma = \Gamma_{nat}$. Owing to the normalization conditions, $J_M(E)$ may be written as

$$J_M(E) = \frac{f\Gamma_{nat}}{2}\frac{1}{(E - E_s)^2 + (\Gamma_{nat}/2)^2} = \frac{2f}{\pi\Gamma_{nat}}L(E) \quad [2]$$

The theoretical spectrum of the 129 keV $\gamma$-ray of $^{191}Ir$ absorbed by an atom in iridium metal is shown in **Figure 1**. The Mössbauer spectrometer is sensitive only to the narrow, recoil-free line at zero energy

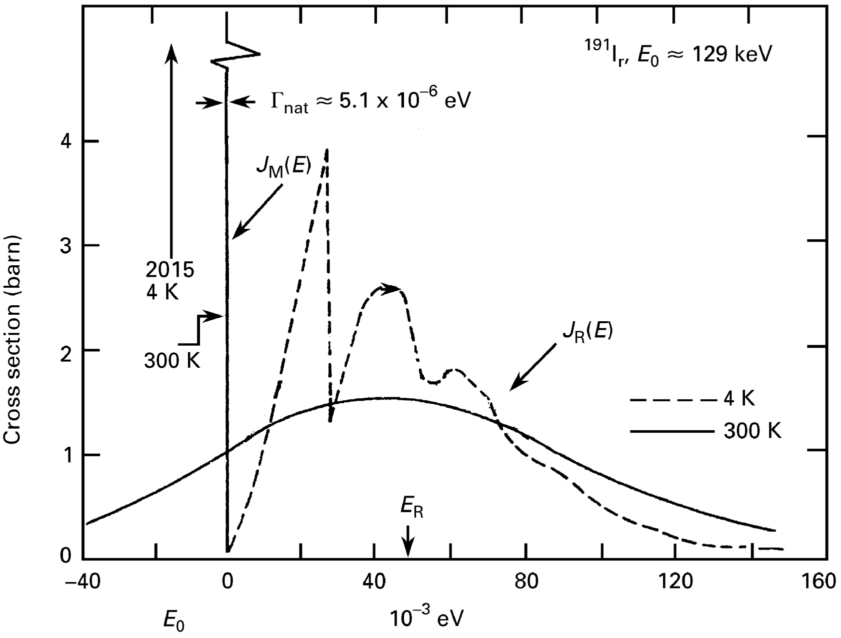



**Figure 1** Absorption cross section for $^{129}Ir$ at 4 K and 300 K. The Debye model was used to calculate the lattice vibration.

shift, which contains 5.7% of the total area under the curve at 4 K for $^{191}$Ir, $E_0 \approx 129$ keV.

There are two types of Mössbauer spectroscopic experiments based on scattering and transmission techniques. In transmission mode experiments, the resonant scattering leads to the sharp attenuation of the radiation intensity registered by a detector; it is therefore sometimes referred to as resonant absorption. The resonant absorption cross section is the total cross section of resonant scattering; the probability of detecting the scattered radiation in transmission spectroscopy may be neglected when the geometrical arrangement is appropriate. Scattering Mössbauer experiments involve the detection of either resonantly scattered $\gamma$-quanta or other radiation that is emitted in the process of resonant scattering or immediately after it.

The emission probability $f$ and the absorption probability $f'$ of recoilless $\gamma$-quanta and the temperature dependence of $f$ and $f'$ are determined by the $\gamma$-quantum energy, the mass of the nucleus, lattice vibrations and other properties of the sample. This is why the Mössbauer effect is not observed for all elements (see **Figure 2**).

## The Mössbauer effect probability

To obtain information on chemical bonds of atoms in solids from experimental data, an explicit theoretical relation is needed to associate experimental $f$ (or $f'$) values with the phonon spectrum and the force constants of the crystal. Unfortunately, this seemingly rather simple approach produces a number of problems that primarily result from the limited information that is available on the phonon spectra of solids of practical interest. Only for a cubic crystal where interatomic forces may be assumed to be harmonic is there a simple relation for the probability of $\gamma$-quantum resonant emission with a wave vector $\boldsymbol{k}$ when a nucleus undergoes a transition from an excited state to the ground state:

$$f = \exp\left\{-k^2\,\overline{x^2}\right\}$$

where $\overline{x^2}$ is the mean square displacement of the Mössbauer atom from its equilibrium position at temperature $T$.

The Debye model is most widely used and is assumed irrespective of the lattice symmetry and the number of atoms in the unit cell. Hence, from an observed $f$ value the effective characteristic temperature (Deby temperature) $\Theta$ may be evaluated and compared with the temperatures obtained for the substance by methods such as X-ray analysis or specific heat measurements. For very low and high temperatures (as compared with $\Theta$),

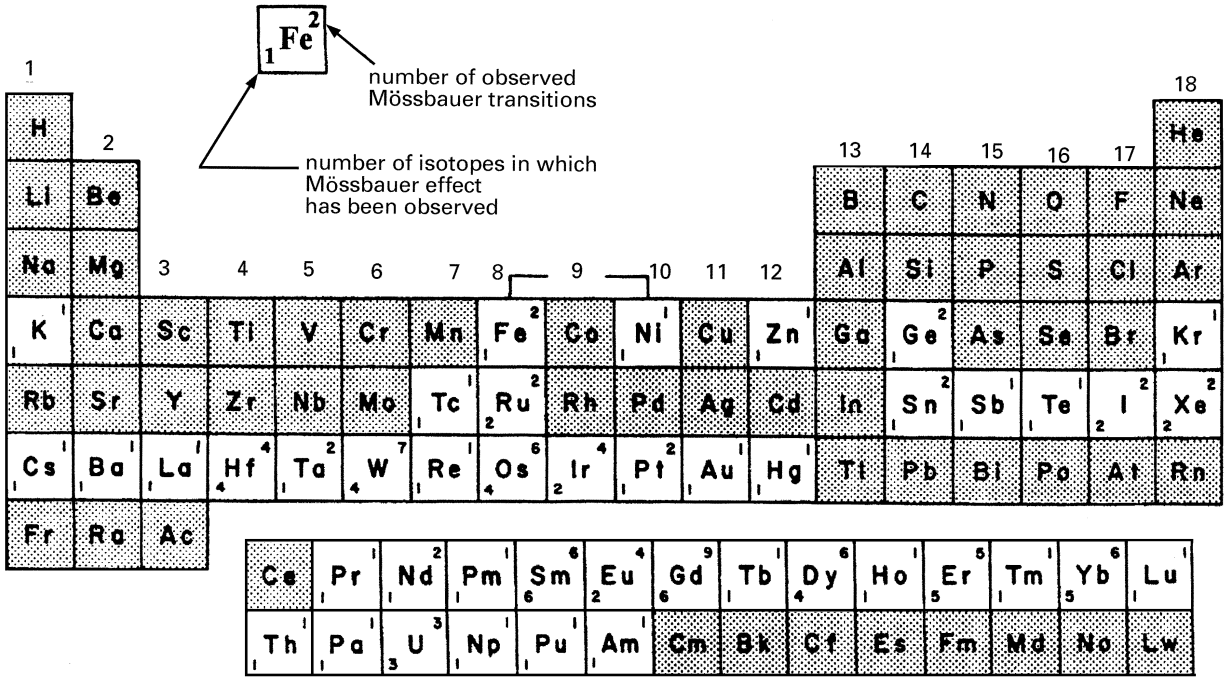



**Figure 2** Mössbauer periodic table.

$$f = \exp\left(-\frac{3E_R}{2k_B\Theta}\right) \qquad \text{for } T \ll \Theta$$
$$f = \exp\left(-\frac{6E_RT}{k_B\Theta^2}\right) \qquad \text{for } T > \Theta \qquad [3]$$

where $c$ is the velocity of light and $k_B$ is the Boltzmann constant. Using resonance atoms as admixtures in various compounds extends the number of solids in which the Mössbauer effect is observable and this expands the usefulness of the method. Unfortunately, the theoretical treatment is much more complicated than that of the ordinary Mössbauer effect.

In general one needs to take into account the dependence of phonon excitation on the direction of the recoil momentum of the nucleus with respect to the crystallographic axes. This must be considered in particular when strong anisotropy is observed for laminar crystals. The anisotropy factor $f'$ that is detected for some nontextured polycrystalline samples takes this into account in a phenomenon known as the Goldanskii–Karyagin effect.

## Hyperfine interactions and line positions in Mössbauer spectra

The energy of a nucleus, as well as of any system of charges and currents, changes upon interaction with an external electromagnetic field by an amount $E'$. Using classic electrodynamics, the energy may be described by the multipole moments series as

$$E' = q\varphi_0 - \boldsymbol{p}\boldsymbol{E}_0 - \boldsymbol{\mu}\boldsymbol{H}_0 - \frac{1}{6}\sum_{i,k=1}^{3} \boldsymbol{Q}_{ik}(\partial^2\varphi/\partial x_i\,\partial x_k)_0 - \cdots \qquad [4]$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electric and magnetic field strengths, respectively, $\varphi$ is the electrostatic potential, $q = eZ$ is the nuclear charge, $\boldsymbol{p}$, $\boldsymbol{\mu}$ are vectors of electric and magnetic dipole moments, $\boldsymbol{Q}_{ik}$ is the tensor of the electric quadrupole moment. The subscript zero indicates that the quantity is that at the centre of the nucleus. Moments of higher orders may be neglected. Since nuclei do not have electric dipole moments, the second term of Equation [4] is zero and the energy of a nucleus in an external electromagnetic field is determined by the product of the nuclear ($q$, $\boldsymbol{\mu}$, $\boldsymbol{Q}_{ik}$) and electron ($\varphi_0$, $\boldsymbol{H}_0$, $(\partial^2\varphi/\partial x_i\partial x_k)$) factors. In solid-state physics and in fields of application the first factors are supposed to be known.

As seen from Equation [4], the Hamiltonian $H$, describing the interaction of a nucleus with effective fields, may be represented as a sum of the two Hamiltonians, one for interactions of the nucleus with electrons ($H_Q$) and the other for interactions with the magnetic field ($H_M$):

$$H = H_Q + H_M \qquad [5]$$

The Hamiltonian of the electrostatic interaction is

$$H_Q = e\sum_{p=1}^{Z} \varphi(\boldsymbol{r}_p) \qquad [6]$$

where $\boldsymbol{r}_p$ is the radius vector of the $p$th proton, $\varphi(\boldsymbol{r}_p)$ is the electric potential in the vicinity of the $p$th proton, and the summation is over all protons, $p = 1,\ldots, Z$. The coordinate system is chosen such that the origin is at the centre of the nucleus and the axes $x^i (i = 1, 2, 3;\ x^1 \equiv x;\ x^2 \equiv y;\ x^3 \equiv z)$ are directed along the principal axis of the tensor of the electric field gradient (EFG) acting on the nucleus. $\varphi_p(0)$ is the electric potential at the centre of the nucleus due to the $p$th proton and $r_p^2 = \sum_{i=1} (x_p^i)^2$.

Using the fact that the electrostatic potential $\varphi$ satisfies the Poisson equation $\nabla^2\varphi = 4\pi\rho_e$, where $\rho_e = -e\ |\psi(0)|^2$ is the charge density at the nucleus ($r = 0$), and introducing the tensor of nuclear quadrupole moment $Q_{ik}$ and the EFG tensor $\varphi_{ik}$, we can now rewrite the interaction of the nucleus with the electric fields as a sum of two interactions

$$H_Q = H_\delta + H'_Q = -\frac{2\pi}{3}e^2\sum_{p=1}^{Z} r_p^2|\psi(0)|^2 + \frac{e}{6}\sum_{i,k=1}^{3}\varphi_{ik}Q \qquad [7]$$

The nucleus is here considered to be a sphere with a mean-square radius $\overline{r^2} = \sum_{p=1}^{Z} \overline{r_p^2}/Z$ for the ground state (g) and excited state (e). As a rule, $\overline{r_g^2} \neq \overline{r_e^2}$ and the nuclear charge is uniformly distributed inside the sphere. The external electric field acting on such a spherical nucleus does not split the levels but shifts them by the quantity

$$\delta E_{g,e} = \frac{2\pi}{3}e^2 Z r_{g,e}^2|\psi(0)|^2 \qquad [8]$$

The shift due to Coulomb interactions is of the order of $10^{-12}$ of the transition energy. The value of the shift for every nuclear level depends on the chemical state of the atom. This is characterized by the $|\psi(0)|^2_{as}$ parameter which is the electron density at the nucleus in the absorber (a) or in the source (s). In a Mössbauer spectrum this part of the full electrostatic interaction manifests itself as the isomer (chemical) shift $\delta$ between the centre of gravity of the

emission spectrum of the source and the centre of gravity of the absorption spectrum of the sample, which is called the absorber (**Figures 3A, B**). Thus the transition energy in the source $E_s$ is different from the energy $E_a$ in the absorber, both of them being different from the transition energy $E^0$ for $\overline{r^2} = 0$. It must be appreciated that in Mössbauer experiments it is not the absolute energy of the $\gamma$-quanta that is determined but the energy shift of the nuclear levels. The energy scanning is carried out by the use of the Doppler effect. Therefore, the energy parameters ($\Gamma$, $\delta$) are expressed in velocity units, $v$. For a pair of source and absorbed nuclei we may write

$$\delta = \frac{2\pi}{3} Ze^2(\overline{r_e^2} - \overline{r_g^2})(|\psi(0)|_a^2 - |\psi(0)|_s^2 \qquad [9]$$

The charge density at the nucleus is mainly determined by s electrons and only partially by p-electrons. The main effect of the p electrons and d electrons and any other electron shells that do not contribute directly to the electron density $|\psi(0)|^2$, is to shield the s electrons. The determination of the scale factor $(\overline{r_e^2}-\overline{r_g^2})$ in Equation [9] is called the isomer shift calibration. The interpretation of isomer shifts in Mössbauer spectra involves the correlation of a given $\Delta\rho_e(0)=|\psi(0)|_a^2-|\psi(0)|_s^2$ value (the electron density difference) with the known electronic structure of the Mössbauer atom or the change of the structure resulting from the examination of different samples. It should be noted that only one part in $10^{20}$ of the electrons in a solid directly participate in the isomer shift; the nuclear parameter $(\overline{r_e^2}-\overline{r_g^2})$ is of the order of $10^{-29}$ cm$^2$. The isomer shift is four orders of magnitude smaller than the Lamb shift caused by quantization of the electromagnetic field.

The second of the above-mentioned interactions is known as the electric quadrupole interaction, $H'_Q$. The value of $Q_{zz}$ when the nucleus is in the state $m = I$ is conventionally called the nuclear quadrupole moment $eQ = \langle I, I\ |Q_{zz}|\ I, I\rangle$. The EFG tensor in the principal axes, taking into account LaPlace's equation, is determined by two independent parameters: first, $\varphi_{zz}$, commonly called 'the electric field gradient' or 'the principal component of the electric field gradient tensor' and sometimes written as $\varphi_{zz} = eq$; second $\eta = (\varphi_{xx} - \varphi_{yy})/\varphi_{zz}$, called the 'asymmetry parameter', the axes being chosen such that $|\varphi_{zz}| > |\varphi_{xx}| > |\varphi_{yy}|$ with $0 < \eta < 1$.

In Mössbauer spectroscopy it is necessary to evaluate the eigenvalues of the $H'_Q$ Hamiltonian, that is the energies $E_Q^m$ for the ground state and for the excited state, the transition from which is followed by the emission of a Mössbauer $\gamma$-quantum (see **Figure 4A**). The line positions in Mössbauer spectra

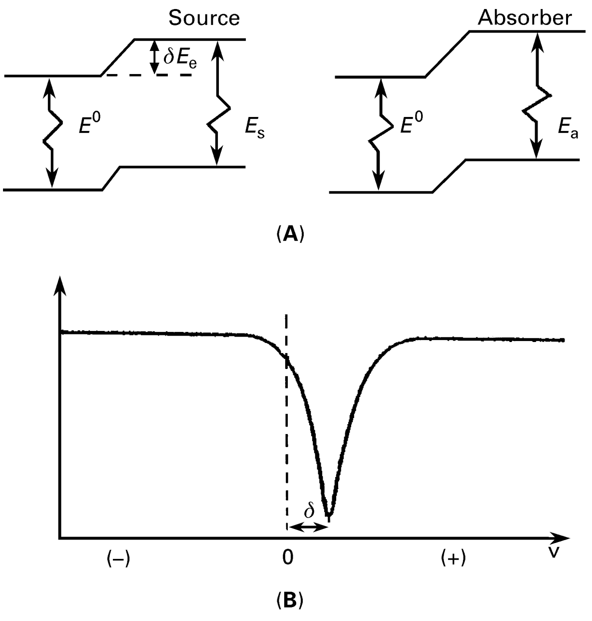



**Figure 3** (A) Energy level shifts for a $^{57}$Fe nucleus, resulting in the appearance of the isomer shift $\delta$. (B) The corresponding Mössbauer spectrum.

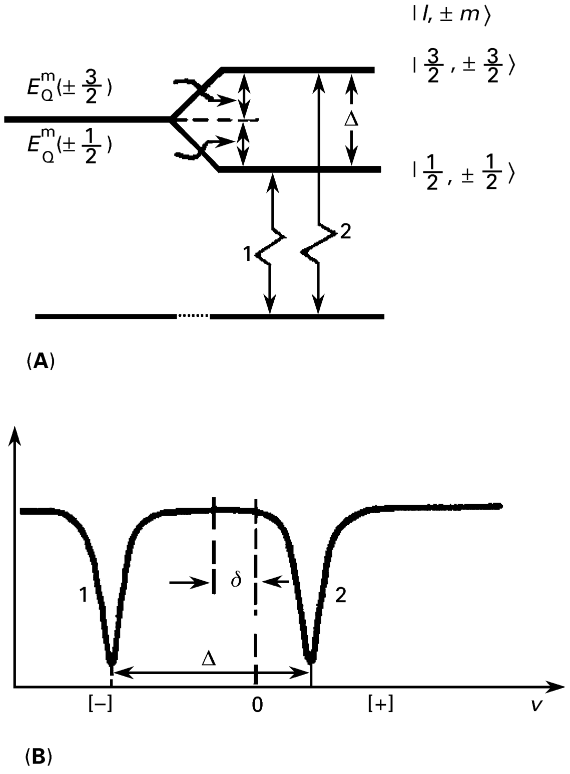



**Figure 4** (A) The splitting of the excited level of a $^{57}$Fe nucleus due to the electric quadrupole interaction $\Delta$. (B) The corresponding Mössbauer spectrum.

are determined by the eigenvalues of the sum Hamiltonian $H_Q$ for the nucleus in excited and ground states in the source and absorber, i.e. both $\delta$ value and $E_Q^m$. The intensities of the lines that provide valuable information on the structure of the surface layers are determined by the eigenvectors of the Hamiltonian $H'_Q$.

For the axially symmetric EFG tensor ($\eta$ = 0) the degeneracy of the nuclear energy levels is not completely split, and the energy depends only on the absolute value of the spin projection. The energy level displacement is given by the expression

$$E_Q^m = \frac{e\varphi_{zz}Q}{4I(2I-1)}3m^2 - I(I+1)] \qquad [10]$$

where $m$ is the value of the spin projection onto the quantization axis. If the nuclear spin is half integral, the quadrupole interaction will cause the levels to be at least twofold degenerate. If the spin values are integral, the level degeneracy for $\eta \neq 0$ may be completely lifted. For $\eta \neq 0$, the $E_Q^m$ values may be found by solving a secular equation that has no general analytical solution for $I > 2$. Of special interest in Mössbauer spectroscopy are the transitions between states with spin quantum numbers $I = \frac{1}{2}$ and $I = \frac{3}{2}$. This is the case for $^{57}$Fe, $^{119}$Sn, $^{125}$Te and many other nuclides. The quadrupole splitting, the distance between two lines, is equal to

$$\Delta = \frac{|e\varphi_{zz}Q|}{2}\left(\frac{1+\eta^2}{3}\right)^{1/2} \qquad [11]$$

According to Sternheimer, two primary sources of the EFG may be identified. First, charges on ions surrounding the nucleus (provided the symmetry of the surroundings is lower than cubic), and secondly, the unfilled valence shells (since filled shells possess a spherically symmetric charge distribution). The actual EFG at the nucleus is determined by the extent to which the electronic structure of the Mössbauer atom is distorted by electrostatic interactions with external charges. This leads to the so-called 'antishielding' effect, which is described by $1 - \gamma_\infty$.

The Hamiltonian for the interaction of the magnetic dipole moment of a nucleus with the effective magnetic field $\boldsymbol{H}_{\text{eff}}$ acting on it may be written

$$E_M = -g_I\mu_n \boldsymbol{I}\boldsymbol{H}_{\text{eff}} \qquad [12]$$

where $\mu_n$ is the nuclear magneton, $g_I$ is the gyromagnetic ratio, $\boldsymbol{I}$ is the nuclear spin operator (the quantization axis coincides here with the direction of $\boldsymbol{H}_{\text{eff}}$). The degeneracy of the nuclear levels is completely split. **Figure 5** depicts the splitting of the nuclear energy levels and the corresponding Mössbauer spectrum. The shift of the levels is determined by the expression

$$E_M^m = -g_I\mu_n m H_{\text{eff}} \qquad [13]$$

(where $m$ = spin projection onto the quantization axis). In $^{57}$Fe, where the transition multipolarity of interest is $M1$, $m_e - m_g$ = 0, ±1, and out of eight possible transitions in $H_{\text{eff}}$ only six are present (**Figure 5A**).

Often all three interactions, i.e. the electric monopole, magnetic dipole and electric quadrupole interactions, occur simultaneously. If the quadrupole interaction is small compared with the magnetic interaction ($E_Q^m \ll E_M^m$), a correction to the interaction energy may be applied using first-order perturbation

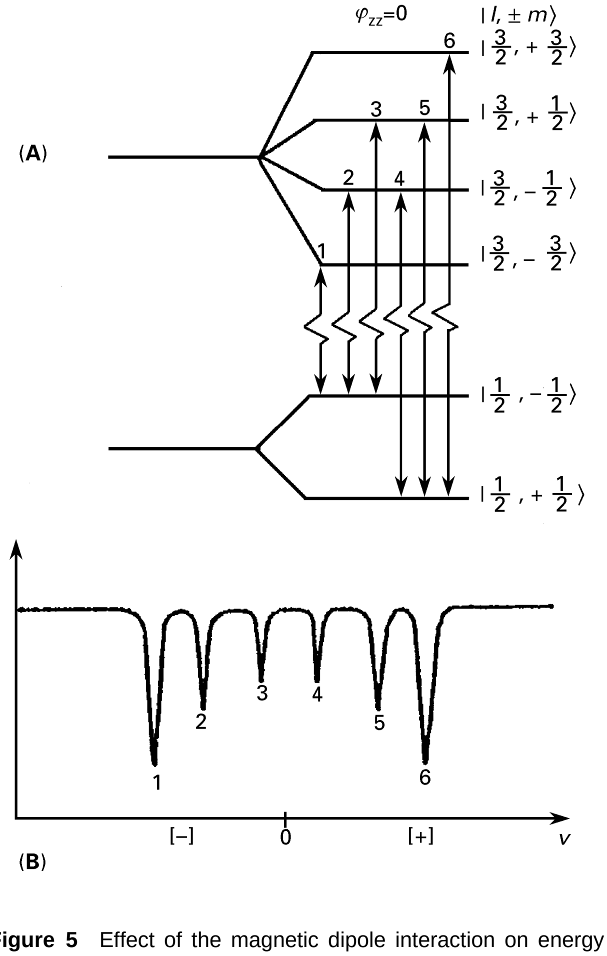



**Figure 5** Effect of the magnetic dipole interaction on energy level splitting in $^{57}$Fe. (A) Energy level diagram in the field $H_{\text{eff}} \neq 0$, $\varphi_{zz} = 0$. (B) The corresponding Mössbauer spectrum.

theory for a nondegenerate spectrum. For the case of an axially symmetric EFG tensor ($\eta = 0$) the level positions are given by

$$E_{\mathrm{M}}^{\mathrm{m}} = -g_I \mu_{\mathrm{n}} m H_{\mathrm{eff}} + \frac{eQ\varphi_{zz}}{4I(2I-1)} \left[3m^2 - I(I+1)\right] \frac{3\cos^2\theta - 1}{2} \qquad [14]$$

The splitting of the energy levels and the corresponding Mössbauer spectrum are shown in **Figure 6A** and **B**. If the $z$ axis of the axially symmetric EFG is parallel to the magnetic field $E_{\mathrm{Q}}^{\mathrm{m}} \sim E_{\mathrm{M}}^{\mathrm{m}}$, the hyperfine structure is also described by Equation [14]. The sublevels are not equidistant. This results in an asymmetric magnetically split Mössbauer spectrum as depicted in **Figure 6B**. For the more general case, there is a dependence of the sublevels shift on the angle $\theta$. If $E_{\mathrm{Q}}^{\mathrm{m}} \sim E_{\mathrm{M}}^{\mathrm{m}}$, $\eta = 0$ and $\theta \neq 0$, then the wavefunctions $\varphi_m$ describing a nuclear state with a definite spin projection $m$ onto the $z$ axis are not the eigenfunctions of that Hamiltonian. The wavefunctions of the nuclear state with energies given by the roots of the secular equation $\mathrm{Det}(H_{mm'} - \varepsilon\delta_{mm'}) = 0$ will be a superposition of $\varphi_m$ functions at different $m$ ($H_{mm'}$ is the matrix element of the Hamiltonian $H$).

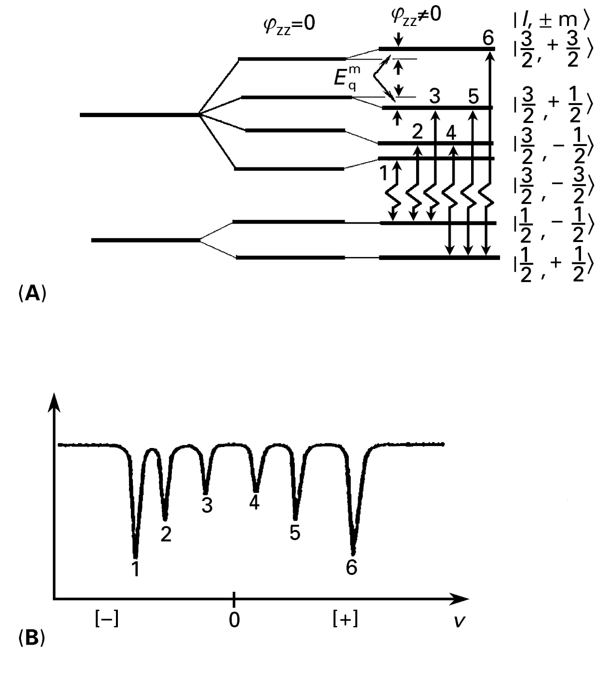



**Figure 6** (A) Energy level splitting diagram with combined hyperfine interactions ( $E_{\mathrm{Q}}^{\mathrm{m}} \ll E_{\mathrm{M}}^{\mathrm{m}}$ ) for $^{57}$Fe. (B) The corresponding Mössbauer spectrum.

The superposition will cause the relative line intensities of the Mössbauer spectrum to be different from those characterizing a pure magnetic interaction. This effect may also give rise to the appearance of additional lines in the Mössbauer spectrum.

## Relative intensities of spectral lines

In the absence of relaxation effects and saturation arising from finite sample thickness, the intensity of a spectral component is determined by the nuclear transition characteristics (see **Figures 3–6**). The most important of these are the spin and the parity of the excited and ground states of the Mössbauer nuclei, the multipolarity of the transition, and the direction of the wave vector $\boldsymbol{k}$ of the $\gamma$-quanta emitted with respect to a chosen direction which is specified, for example, by the magnetic field or by the electric field gradient that causes the nuclear level degeneracy to be lifted.

The probability $P$ of the occurrence of a nuclear transition of multipolarity $M1$ from a state $|I_{\mathrm{e}} m_{\mathrm{e}}\rangle$ to a state $|I_{\mathrm{g}} m_{\mathrm{g}}\rangle$, equals

$$P(I_{\mathrm{g}} m_{\mathrm{g}} 1M, I_{\mathrm{e}} m_{\mathrm{e}}; \theta, \varphi) = |G(m_{\mathrm{e}}, m_{\mathrm{g}})|^2 F_{\mathrm{L}}^{\mathrm{M}}(\theta, \varphi) |\langle I_{\mathrm{g}} \,||\, 1 \,||\, I_{\mathrm{e}}\rangle|^2 \qquad [15]$$

where $\theta$, $\varphi$ are the polar and azimuthal angles determining the direction of emitted $\gamma$-quanta in the coordinate system defined by the magnetic field direction, $M = m_{\mathrm{e}} - m_{\mathrm{g}}$; $G(m_{\mathrm{e}}, m_{\mathrm{g}}) = \langle I_{\mathrm{g}} m_{\mathrm{g}} LM \,|\, I_{\mathrm{e}} m_{\mathrm{e}}\rangle$, are the Clebsh–Gordan coefficients; $\langle I_{\mathrm{g}} \,|\,|\, 1 \,|\,|\, I_{\mathrm{e}}\rangle$ is the reduced matrix element which does not depend on the quantum numbers $m_{\mathrm{g}}$, $m_{\mathrm{e}}$. The angular function $F_L^{\mathrm{M}}(\theta, \varphi)$ is determined only by the transition multipolarity. The intensity of the Mössbauer line is proportional to the product of the Clebsh–Gordan coefficients and the $F_L^{\mathrm{M}}(\theta,\varphi)$ functions. Plots of angular dependence of the intensities of the spectral components are given in **Figure 7**. In the sample the purely magnetic hyperfine splitting of nuclear levels take place.

The effect of anisotropy of atomic vibrations in solids not only causes the Mössbauer effect probability $f$ to be anisotropic in single crystals, but may also lead to anisotropy in $f$ for nontextured polycrystalline samples consisting of randomly orientated crystallites. The relative line intensities of the Mössbauer spectrum (**Figures 3–6**) will be different for negative and positive velocities. Similar deviations may be caused by texture, that is by a preferred orientation of crystals in a polycrystalline sample.

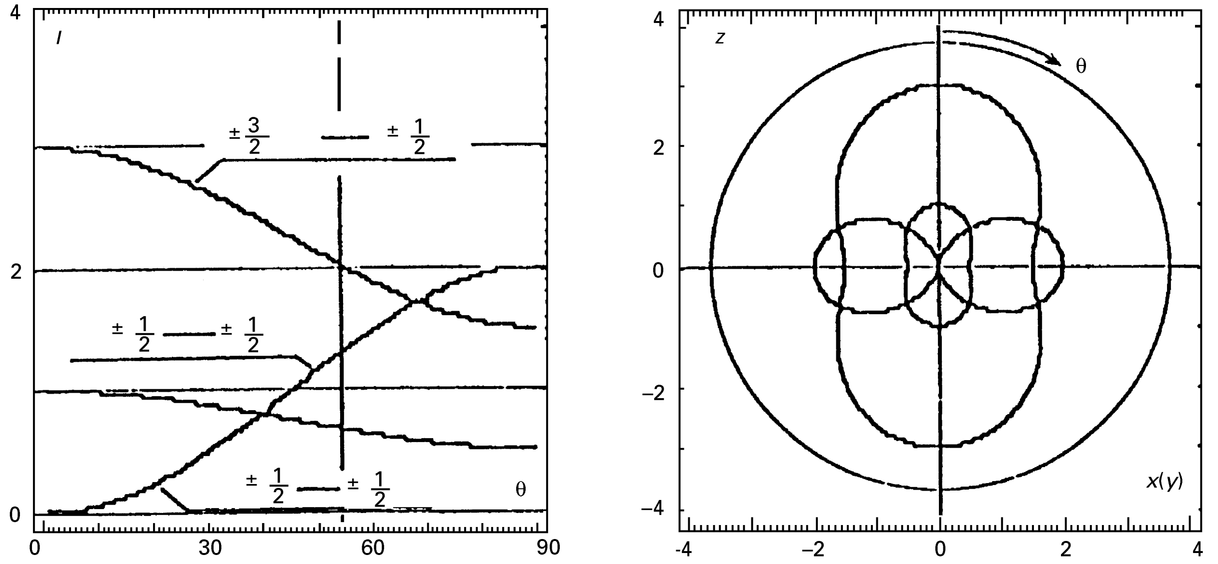



**Figure 7** Angular dependences of relative intensities of the hyperfine structure components for the $I_e = \frac{3}{2}$, $I_g = \frac{1}{2}$ transition in $^{57}Fe$, for magnetic dipole interaction. The polar angle $\theta$, defining the wave vector $\boldsymbol{k}$ of the emitted $\gamma$-quantum, is the angle between the radiation direction and the quantization axis. The quantization axis $z$ is parallel to $H_{eff}$.

## Resonance fluorescence and interference effects

The resonantly scattered radiation may interfere with the radiation scattered by electrons of the atom. The characteristic time – the lifetime of nuclear excited state $\tau \propto \Gamma^{-1}$ – is longer by several orders of magnitude than the lattice vibration periods. There is no correlation here between the initial and final positions of the atom. Despite this, the scattered wave remains coherent with the incident one.

The energy distribution of the scattered $\gamma$-radiation may differ substantially from that of the incident radiation and is determined by convolution of the emission and scattering spectra. The Rayleigh scattering spectrum intensity effectively coincides with the emission spectrum. The intensity of the resonantly scattered radiation follows the usual Lorentzian curve, while the contribution of the interference term to the total intensity of the scattered radiation takes the form of a dispersion curve. The use of Bragg reflections in a single-crystal scatterer permits a substantial reduction in the contribution from incoherent scattering. The interference pattern in this case may be unambiguously connected with the crystallographic and electronic structure.

The directions of Mössbauer diffraction, when the hyperfine splitting is absent, generally coincide with the directions of Rayleigh coherent scattering. However, the angular dependences of diffraction line intensities from nuclear scattering and Rayleigh scattering are different. Since the Debye–Waller factor decreases with the scattering angle, it is necessary to use large scattering angles to increase the contribution of nuclear diffraction to the total spectrum.

When the hyperfine splitting is present, the diffraction pattern caused by resonant scattering is much more complicated. Magnetic fields at the different Mössbauer atoms may be not parallel. Only one of the spin subsystems will participate in the coherent scattering of the quantum and there will be no cancellation of the scattering amplitudes. This leads also to the observation of pure nuclear diffraction maxima. When the hyperfine interaction energies are sufficiently different, it should be possible to tune the incident radiation to select a particular chemical environment, and then measure the diffraction pattern from only these atoms. Only Mössbauer effect diffraction can provide independent autocorrelation functions for atoms in different chemical environments. Two physical problems must be given special mention. First, the diffraction of Mössbauer radiation is dynamic in nature. Secondly, the suppression of inelastic scattering channels requires attention. The resulting effect is the nuclear resonant analogue of the Borman effect and is realized when a thick perfect crystal containing Mössbauer nuclei is set up at a diffraction angle and the transmittance of the crystal increases when the source velocity is such that the system is brought into resonance.

Considerable interest in pure nuclear back-reflections arises also from application to $\gamma$-optical devices, such as the filtering of Mössbauer radiation from the white spectrum of synchrotron radiation.

Extremely narrow band with ($10^{-6}$–$10^{-8}$ eV) and small angular width (0.4 arc second) have been obtained from the synchrotron radiation continuum. Progress in this technique has made it feasible to produce diffracted $\gamma$-quanta with intensities unattainable from conventional Mössbauer sources, thereby increasing interest in hyperfine spectroscopy. The standard experiment will be time-resolved observation of forward scattering from a polycrystalline target instead of the pure nuclear reflection from a single crystal that has been used to date. The use of synchrotron radiation may allow the Mössbauer effect to be observed in new isotopes. Such isotopes would need low-energy excited nuclear levels but need not have appropriate parent nuclei, and hence they are not given in **Figure 2**. The interference of the elastically scattered radiation gives rise to a mirror reflected wave. It is known that if electromagnetic radiation falls onto a mirror surface characterized by complex index of refraction $n = 1 - \sigma - i\beta$ at a glancing angle $\gamma \leq \gamma_{cr}$ the reflectivity $R$, i.e. the ratio of the reflected and incident intensities, becomes equal to unity. For real media there is always some absorption and the imaginary part of the index of refraction is not zero. However, if the $R$ value rises sharply when $\gamma$ becomes less than $\gamma_{cr}$, the situation is described as total external reflection (TER). The coherent amplification of the scattered wave under conditions of TER is analogous to diffraction on scattering from single crystals. The index of refraction depends only on the forward scattering amplitude and hence there is no phase shift between the waves scattered by various atoms and nuclei in the unit cell. In the presence of hyperfine splitting and of nonrandomly orientated quantization axes in the scatterer, polarization effects should also be taken into consideration. TER may be used for studies of very thin surface layers.

## Relaxation phenomena in Mössbauer spectroscopy

The term 'relaxation' is used to indicate that time-dependent effects occur in the system under study. The $\gamma$-quantum scattering leads, as a rule, only to change of the nuclear state, while the electronic system remains unchanged. Sometimes, i.e. in paramagnets, the interaction of the electronic shell with the environment may be comparable to or much weaker than the hyperfine coupling. The atom follows a random, stochastic 'path' through its allowed states owing to time-dependent, extra-atomic interactions, and as a result of the hyperfine interaction the Mössbauer spectrum will be affected.

There are two main types of relaxation in Mössbauer studies: paramagnetic and superparamagnetic relaxation. As a rule, the observed spectra are quite complicated. The simple relaxation processes for electronic spin $S = \frac{1}{2}$ can be analysed in terms of a fluctuating hyperfine field $H_n(t)$ which takes on the values $+H_n$ and $-H_n$. If off-diagonal terms in the hyperfine Hamiltonian are absent, then

$$H_M = A_z I_z S_z(t) = g_n \mu_n H_n(t) I_z \qquad [16]$$

where $A$ is the hyperfine coupling constant and $A_x = A_y = 0$. Under the influence of the electron–bath interaction, the electronic spin $S_z(t)$ fluctuates between the values $S_z = \pm\frac{1}{2}$ at some rate $\nu_R$. Using either stochastic arguments or a rate equation approach, one can arrive at a closed-form expression for the line shape. The simplicity of the result obtained has made this model very popular. If off-diagonal terms are present in $H_M$, such simplifications are not possible.

Relaxation theory, as it applies to Mössbauer spectroscopy, has two main approaches: perturbation calculations and stochastic models. The stochastic approach is easily visualized and adapted to various physical situations. The approach generally proceeds by considering the system divided into two parts: the radiating system and the 'bath'. Depending on the particular model, the bath can induce fluctuations in the radiating system by providing unspecified 'hits' or by being represented by a fluctuating effective magnetic field. Blume formulated the effective-field, nonadiabatic model in a particularly useful way by introducing the superoperator (or Liouville operator) formalism. The superoperator formalism is extended to the case where the nucleus, and the atomic electrons are treated as a fully quantum-mechanically coupled system. It is possible to carry out good calculations of all experimental relaxation spectra.

## List of symbols

$A$ = hyperfine coupling constant; $c$ = velocity of light; $E$ = energy; $E'$ = energy of interaction of a nucleus with an electromagnetic field; $E_0$ = energy of an excited state; $\boldsymbol{E}_0$ = electric field strength at the centre of the nucleus; $E^0$ = transition energy for $\overline{r^2} = 0$; $E_a$ = transition energy in absorber; $E_M^m$ = shift of levels due to magnetic interactions; $E_Q^m$ = eigenvalues of the $H'_Q$ Hamiltonian; $E_R$ = recoil energy; $E_s$ = transition energy in the source; $\delta E_{g,e}$ = value of the shift of a nuclear level; $e$ = electron charge; e = excited state (index); $G(m_e, m_g)$ = Clebsh–

Gordan coefficient; g = ground state (index); $g_l$ = gyromagnetic ratio; $f(f')$ = probability of recoilless emission (absorption), Lamb–Mössbauer factor; $F_L^M(\theta, \varphi)$ = angular functions; $\boldsymbol{H}$ = magnetic field strength; $\boldsymbol{H}_0$ = magnetic field strength at the centre of the nucleus; $\boldsymbol{H}_{\mathrm{eff}}$ = effective magnetic field acting on the nucleus; $H_{\mathrm{n}}$ = hyperfine field; $H$ = Hamiltonian; $H_{\mathrm{M}}$ = Hamiltonian for interaction with the magnetic field; $H_{\mathrm{Q}}$ = Hamiltonian for interaction with electrons; $H'_{\mathrm{Q}}$ = Hamiltonian; $H_\delta$ = Hamiltonian; $H_{mm'}$ = matrix element of the Hamiltonian $H$; $I$ = nuclear spin; $\boldsymbol{I}$ = nuclear spin operator; $J_{\mathrm{R}}(E)$, $J_{\mathrm{M}}(E)$ = energy distributions of Mössbauer $\gamma$-rays; $k_{\mathrm{B}}$ = Boltzmann constant; $\boldsymbol{k}$ = wave vector; $L(E)$ = Lorentzian line; $M$ = mass of nucleus; $M1$ = magnetic dipole transition; $m$ = spin projection onto the quantization axes; $n = 1 - \sigma - i\beta$ = the complex index of refraction; $\boldsymbol{p}$ = vector of electric dipole moment; $P$ = probability of a nuclear transition; $\boldsymbol{Q}_{ik}$ = tensor of the electric quadrupole; $q = eZ$ = nuclear charge; $R$ = reflectivity; $\boldsymbol{r}_p$ = radius-vector of the $p$th proton; $\overline{r^2}$ = mean-square radius; $S$ = electronic spin; $T$ = temperature; $v$ = velocity; $\overline{x^2}$ = mean square displacement of the Mössbauer atom; $Z$ = number of protons in the nucleus; $\Gamma$ = full width at half-maximum; $\Gamma_{\mathrm{nat}}$ = natural line width; $\gamma$ = glancing angle; $\gamma_{\mathrm{cr}}$ = angle of total reflection (the critical angle); $\gamma_\infty$ = antishielding factor; $\Delta$ = quadrupole splitting; $\delta$ = isomer (chemical) shift; $\varepsilon$ = resonance effect magnitude; $\delta_{mm'}$ = Kronecker symbol; $\eta = (\varphi_{xx} - \varphi_{yy})/\varphi_{zz}$ = asymmetry parameter; $\Theta$ = Debye temperature; $\theta$ = polar angle, specifying $\boldsymbol{H}_{\mathrm{eff}}$; $\mu_{\mathrm{n}}$ = nuclear magneton; $\boldsymbol{\mu}$ = vector of magnetic dipole moment; $\rho_{\mathrm{e}}$ = charge density at the centre of the nucleus; $\tau$ = lifetime of nuclear excited state; $\varphi$ = electrostatic potential; $\varphi$ = azimuthal angle, specifying $\boldsymbol{H}_{\mathrm{eff}}$; $\varphi_m$ = wavefunction; $\varphi(r_p)$ = electric potential in the vicinity of the $p$th proton; $\varphi_p(0)$ = electric potential at the centre of nucleus due to the $p$th proton; $\varphi_{xx,yy,zz}$ = $x,y,z$-component of the EFG tensor; $|\psi(0)|^2_{\mathrm{a,s}}$ = the electron density at the nucleus in the absorber (a) or in the source (s).

*See also:* **Electromagnetic Radiation**; **Mössbauer Spectrometers**; **Mössbauer Spectroscopy, Applications**; **NMR Principles**; **Scattering Theory**; **X-Ray Spectroscopy, Theory.**

## Further reading


Andreeva MA, Belozerski GN, Grishin OV, Irkaev SM and Semenov VG (1995) Mössbauer total external reflection. *Hyperfine interactions* **96**: 37–49.

Butz T, Ceolin M, Ganal P, Schmidt PC, Taylor MA and Troger W (1996) A new approach in nuclear quadrupole interaction data analysis: cross-correlation. *Physica Scripta* **54**: 234–239.

Deak L, Bottyan L, Nagy DL and Spiering H (1996) Coherent forward-scattering amplitude in transmission and grazing incidence Mössbauer spectroscopy. *Physical Reviews B: Condensed Matter* **53**: 6158–6164.

Gütlich P, Link R and Trautwein A (1978) *Mössbauer Spectroscopy and Transition Metal Chemistry* , p 280. Berlin: Springer-Verlag.

Hoy J (1997) Quantum mechanical model for nuclear resonant scattering of gamma-radiation. *Physics of Condensed Matter* **9**: 8749–8765.

Long GJ (ed) (1984–1989) *Mössbauer Spectroscopy Applied to Inorganic Chemistry*, Vols 1–3. New York: Plenum Press.

Long GJ and Grandjean F (eds) (1994) *Applications of the Mössbauer Effect*, International Conference on the Applications of Mössbauer Effect (ICAME-93), Vancouver, Vols I–IV. Amsterdam: Baltzer Science Publishers.

Mössbauer RL (1958) Kernresonanzfluoreszenz von Gammastrahlung in $^{191}$Ir. *Zeitschrift für Physik* **151**: N1, 124–137.

Shenoy GK and Wagner FE (eds) (1978) *Mössbauer Isomer Shifts*, p 780. Amsterdam: North-Holland.

Smirnov GV (1996) Nuclear resonant scattering of synchrotron radiation. *Hyperfine Interactions* **97/98**: 551–588.

Thosar BV and Srivastava IK (eds) (1983) *Advances in Mössbauer Spectroscopy Application to Physics, Chemistry and Biology*, p 924. Amsterdam: Elsevier.

Wertheim GK (1964) *Mössbauer Effect: Principles and Applications*, p 145. New York: Academic Press.


# MRI Applications in Food Science

*See* **Food Science, Applications of NMR Spectroscopy.**

# MRI Applications, Biological

**David G Reid**, **Paul D Hockings** and **Paul GM Mullins**, SmithKline Beecham Pharmaceuticals, Welwyn, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

Non-invasive MRI is at the forefront of clinical diagnostic imaging; its non-destructive nature also gives it great potential as a tool in biological research, involving animal models of disease. Because it is possible to scan the same animal as often, and over as long a period, as necessary, before and after experimental surgery and/or administration of test compounds, MRI is assuming increasing importance in longitudinal evaluation of novel pharmaceuticals and characterization of animal models of disease. Experiments can usually be designed so that each subject acts as its own control, increasing statistical power with smaller group sizes, and longitudinal studies are possible without killing groups of animals at each time point; two factors that, separately and in combination, offer dramatic sparing of laboratory animals. In general, measurement of anatomical features from MR images is much quicker than conventional invasive methodologies like tissue histology. It is often possible to acquire MRI data as three-dimensional images with isotropic resolution. These can be subsequently 'sliced' or rendered along arbitrary planes or surfaces to highlight irregular structures. The MR image is acquired *in situ*, so anatomy is undistorted by fixation, excision, sectioning and staining processes. Finally MRI methods developed to highlight features of animal disease models are often directly transferable to clinical trials and diagnoses.

MRI is so powerful because of the wide range of contrast mechanisms available to differentiate different organs, tissues and pathologies. The physicochemical basis of these contrast mechanisms, and the MR pulse sequences designed to exploit them, are treated in comprehensive standard works and other articles in this Encyclopedia. Important sources of MRI contrast are described below.

Differences in tissue water $T_1$ and $T_2$ relaxation times generally depend on differences in the extent to which water molecules interact with soluble macromolecules. Thus changes in the concentration of soluble proteins in oedema will usually cause changes in $T_1$ and $T_2$, so that MRI acquisition sequences weighted according to one (T1W or T2W) or both of these will distinguish oedematous from normal tissue. Water–macromolecule interactions are also the basis of magnetization transfer contrast (MTC), particularly effective at highlighting fibrous structures like cartilage. 'Pools' of water in which diffusion is more or less restricted by cell boundaries, or anisotropic environments, can be distinguished by diffusion weighted (DW) imaging. DWI is particularly effective at detecting cell swelling during ischaemic energy depletion, and in delineating the course of highly anisotropic microstructures like nerve cells. MR pulse sequences, which refocus magnetization using pulsed magnetic field gradients rather than spin echoes, produce images which are sensitive to differences in magnetic susceptibility between and within tissue, and are a function of the 'inhomogeneous $T_2'$', or $T_2$*. Because paramagnetic deoxyhaemoglobin and diamagnetic oxyhaemoglobin affect the magnetic susceptibility of neighbouring tissues in very different ways, $T_2$* weighted (T2*W) techniques can be used to define tissues where deoxy-haemoglobin has built up as a result of underperfusion, or where metabolic activation has increased oxygenated blood – Blood Oxygen Level Determination (BOLD). MR angiography (MRA) takes advantage of the different behaviours of moving and static nuclear spins, and can delineate vasculature and measure blood flow. Tissue perfusion can be measured using paramagnetic contrast reagents, usually stable chelates of gadolinium or manganese ions, or preparations of magnetic iron oxide particles, which reduce tissue relaxation times. T1W, T2W or T2*W images are obtained before and after administration (usually intravenously) of a contrast reagent; regions accessible to the reagent change in MRI intensity, and the time course of 'wash in' and 'wash out' gives a measure of perfusion status. Contrast reagents are widely used in animal models where the blood–brain barrier is compromised (such as demyelinating disorders and stroke), in studies of tumour perfusion, and as an alternative or adjunct to MRA.

## Practicalities

Although useful work is possible in vertical magnets designed for high resolution spectroscopy, most animal MRI is done in horizontal superconducting magnets with field strengths ranging from 2 to 7 T

(corresponding to $^{1}$H resonances from 86 to 300 MHz) and clear magnet bore diameters ranging from about 20 to 40 cm. Concentric shim, gradient and RF coils reduce the useable diameter of 20 and 40 cm systems to about 8 to 20 cm respectively. Subjects must usually be anaesthetized with a suitable inhalation (e.g. isoflurane, halothane) or injectable (e.g. alphaxalone/alphadalone, fentanyl/fluanisone and midazolam) anaesthetic compatible with the animal model under study. It is often necessary to coordinate, or 'gate', the acquisition of NMR data with heart beat and breathing, which can be done by monitoring the animal's electrocardiogram (ECG) and respiration, and triggering data acquisition in synchrony with one or both. Tracheal intubation and mechanical ventilation allow respiratory gating on the ventilation cycle. Whether triggering is necessary or not, ECG and respiratory monitoring are essential for ensuring animal well-being and effective anaesthesia in the magnet. Other vital parameters like rectal temperature and blood pressure are often also monitored, and animal temperature can be controlled with thermostatted heating blankets or air conditioning. Radio frequency probe and animal holder design is the province of the on-site engineer in many institutions, but increasingly manufacturers are offering these items ready made. At the conclusion of an experiment it is still usual to compare *in vivo* MRI measurements with more conventional histological or organ weight measurements. **Figure 1** shows an unconscious rat supported in an animal holder and connected to ECG and respiratory monitoring systems (right), and about to be inserted into a typical horizontal laboratory MR scanner (left).

## Applications

### Central nervous system

MRI has been fruitfully applied to a number of animal models of CNS conditions, such as demyelination (as in for instance experimental allergic encephalomyelitis, EAE), excitotoxicity and neurotoxicity, identification of the spread of neuronal depolarization in the cortical spreading depression phenomenon, identification of neuroanatomical

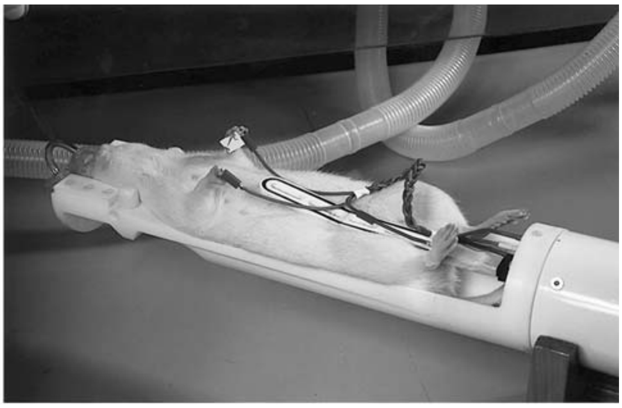


**Figure 1** Right: Unconscious laboratory rat mounted in a non-magnetic holder for MR scanning. Note the face mask for delivery of inhalation anesthetic, conducting sticky pad electrodes on fore and hind paws for ECG signal detection, and the lever (containing a fibre optic cable) placed over the abdomen for respiratory monitoring. Incisor and ear bars are also built into the assembly for stereotaxic positioning if necessary. Left: The entire animal holder about to be inserted into a 7 T laboratory scanner. Although the notional diameter of the horizontal superconducting magnet is 18.3 cm the addition of concentric shim, pulsed field gradient and resonator coils reduces the useable diameter to about 7 cm – adequate for most small laboratory rodents. The gradient coils produce linear variations in the magnetic field of up to 150 mT m$^{-1}$ (15 gauss/cm$^{-1}$) in each of three orthogonal directions; they are actively shielded to reduce induction of eddy currents in the magnet bore. ECG (electric) and respiratory (optical) signals are sent to monitors and triggering electronics outside the RF-impenetrable Faraday cage containing the magnet.

abnormalities in genetically modified animals, blood–brain barrier disruption using contrast reagents and localization of sites of action of psychoactive compounds using BOLD.

The versatility of MRI in this area is well illustrated by its application in models of stroke, where it has been widely used to study the evolution and properties of lesions produced by experimental cerebral ischaemia. Models investigated by MRI include permanent and transient versions of carotid artery, four vessel and middle cerebral artery occlusion (the latter commonly known as MCAO), in rats, mice, gerbils and larger animals like cats. The clarity of T2W images of ischaemic infarcts in some of these models makes this area extremely attractive for the development and implementation of 'high throughput' MRI screening strategies in testing neuroprotective treatments. MRI can exploit different sources of contrast to define the physiological events underlying ischemic injury. Thus **Figure 2** shows representative MR image slices through the brains of rats during an experiment to study the efficacy of an experimental neuroprotective treatment. The top row images are from a control subject, and the bottom row from a subject that received a neuroprotective treatment. Columns labelled (A) and (B) were acquired during a 100 minute period of MCAO using T2*W and DW respectively. In the ischaemic hemisphere (right hand side of each transverse brain image) buildup of paramagnetic deoxyhaemoglobin causes ischaemic regions with perfusion deficit to darken on the T2*W images due to T2* shortening. Additionally, cells swell and undergo cytoskeletal changes in response to energy depletion, which restricts the diffusion of tissue water. These regions show up bright relative to non-ischaemic tissue in DW imaging. Diffusibility changes are further emphasised if DW images are acquired using several different diffusion encoding gradient strengths, allowing a diffusion coefficient to be calculated for each pixel in the image, and the

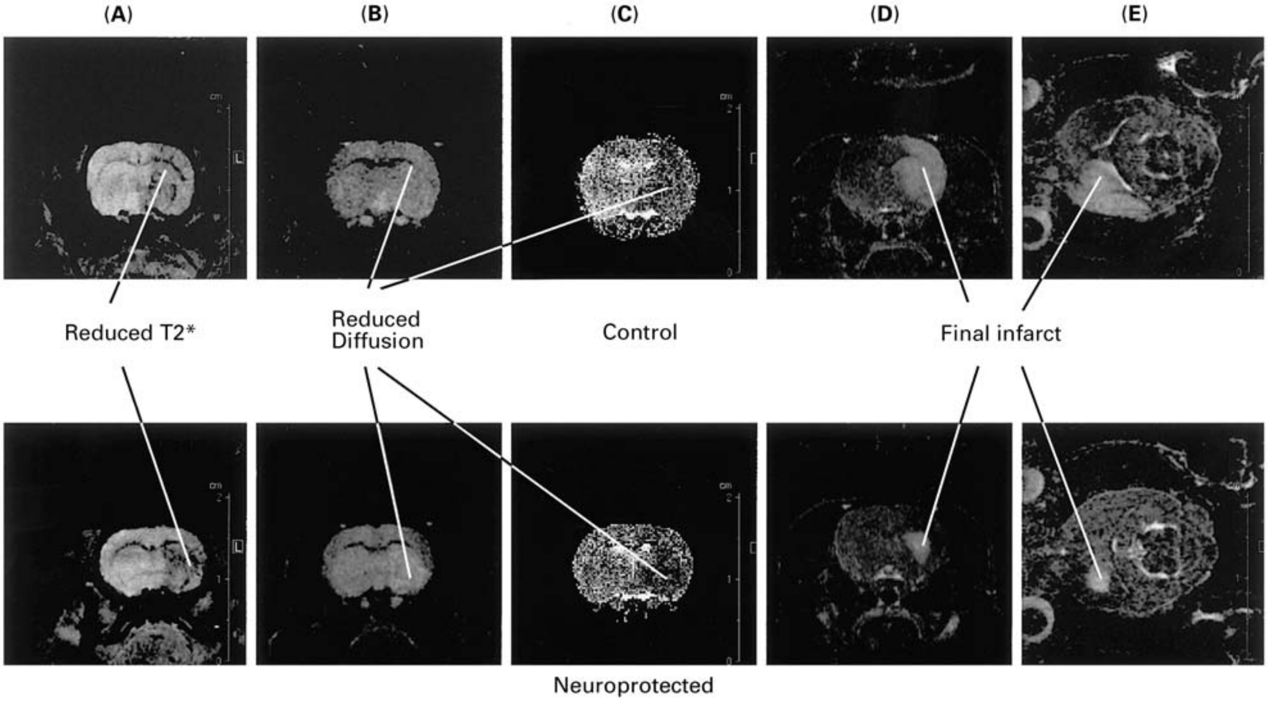



**Figure 2** 300 MHz MR images from the brains of rats subjected to temporary MCAO. The top row of images was acquired from a control animal, and the bottom row from an animal which received a prior neuroprotective treatment. Columns (A)–(D) show transverse images across the brain and column (E) shows a slice taken horizontally. The image columns show: (A) T2*W and (B) DW images acquired during the 100 min period of MCAO; areas of deoxyhaemoglobin buildup, and restricted diffusion, show up as dark and bright regions respectively in the affected (right) cerebral hemisphere; (C) Diffusion map plotting diffusion coefficients during the ischaemic period, calculated from images acquired with three different diffusion gradient strengths; areas of decreased diffusibility which show up bright in (B) manifest lower diffusion coefficients and hence appear dark in the map; Representative transverse (D) and horizontal (E) slices through 3D T2W images acquired 24 h after 100 min MCAO, in which oedema in infarcted regions appears bright. Pulse sequence conditions were: (A) Gradient echo technique, TE/TR = 13/1000 ms, flip angle $\alpha = 90^{\circ}$; (B) TE/TR = 64/1200 ms, diffusion sensitization applied in vertical direction, $b$ value = 11 370 s cm$^{-2}$; (C) Diffusion coefficient map calculated by exponential fitting the signal intensity decay to 3 $b$ values of 0, 2350 and 11370 s cm$^{-2}$; (D) and (E) Interecho delay = 6.5 ms, which a repetition (RARE) factor of 16 converts to a $TE_{effective}$ of 54 ms, TR = 1500 ms.

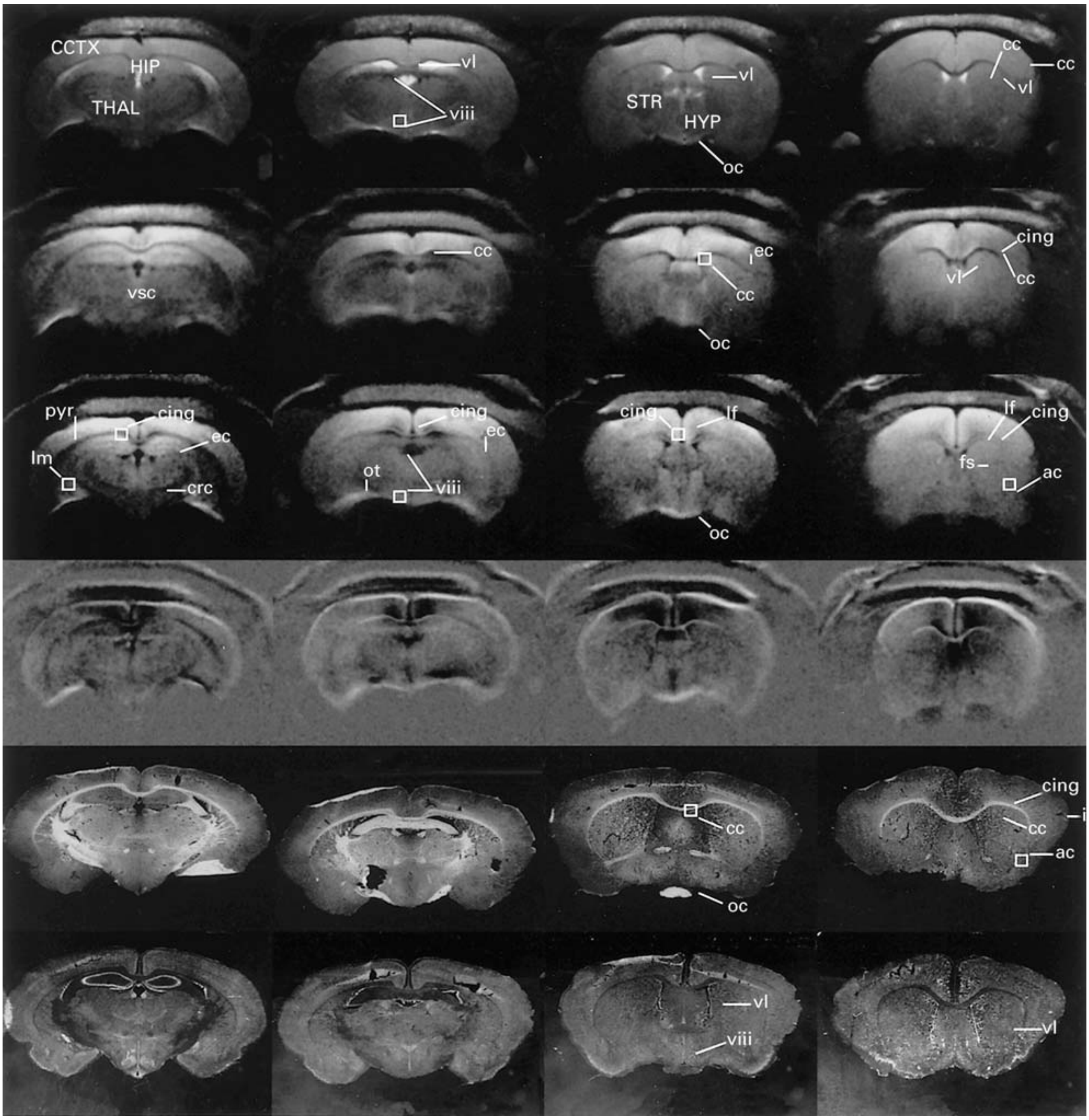



**Figure 3** Transverse diffusion weighted MR images of a rodent brain acquired at four different levels (images progress from caudal ('back') to rostral ('front') from left to right). Images were acquired with TE/TR = 82/2000 ms, field of view (FOV) = 2 cm, and diffusion sensitization $b = 0$ (top row), and $b = 29600\ \mathrm{s\,cm^{-2}}$ applied in a horizontal direction (2nd row) and orthogonal to the slice direction (3rd row). Further anatomical definition is apparent in difference images (4th row) calculated by digitally subtracting one diffusion sensitized image slice from another. Neuroanatomical structures delineated by the DWI method, and closely corresponding to structures identifiable using different histological stains (5th and 6th rows) are labelled as follows: CCTX – cerebral cortex; THAL – thalamus; HIP – hippocampus; cc – corpus callosum; STR – striatum; HYP – hypothalamus; ox – optic chiasm; ec – external capsule; 3v – third ventricle; LV – left ventricle; ot – optic tract; vsc – ventral spinocerebellar tract; ac – anterior commissure; cg – cingulum.

spatial dependence of the diffusion coefficient itself is displayed as a 'map' (column C). In contrast to DW and T2*W images, during and shortly after ischaemia no lesion is apparent using T2WI. However, 24 h after the transient MCAO oedema in the infarcted region manifests clearly as hyperintensity using T2W imaging; columns (D) and (E) show transverse and horizontal slices respectively through

the same 3D T2W datasets at this time point. Apart from the distribution of oedema, which can be easily quantified, the high isotropic (~ 140 μm) resolution facilitates observation of a number of other neuroanatomical structures, such as the fluid-filled ventricles. Concerted application of different MRI acquisition modalities enables one to measure areas undergoing energy depletion, perfusion deficit and oedema, and make inferences regarding areas at risk at early time points which are destined to evolve into infarcts, and those which may be salvageable by neuroprotective intervention.

The strong directionality of nerve fibres makes DW imaging a very useful method for delineating neuroanatomy, and for studying models of neurodegenerative disorders like demyelinating diseases. **Figure 3** shows slices through a rodent brain acquired with diffusion sensitization in different directions. Nerve fibres which run parallel to the diffusion gradient direction show up dark (due to relatively unrestricted diffusion along the fibre) while those running orthogonal to the gradient direction manifest as bright because diffusion across the axon is relatively restricted, so signal loss during diffusion sensitization is minimal.

## Cardiovascular system

Gating of the MRI acquisition to the cardiac and (preferably) respiratory cycles is essential in studies of cardiovascular anatomy and function. Images acquired at full systole, and diastole, enable one to measure the change in volume of the four chambers of the heart during a single contraction, and so calculate the ejection fraction. By preceding the MRI acquisition sequence with a selective presaturation method like DANTE, parts of the myocardium can be 'tagged' and their movement during the cardiac cycle mapped and correlated with cardiac dysfunctions.

Cardiac enlargement, or hypertrophy, is common in diseases like congestive heart failure, and can be a drug side effect. MRI is well suited to measuring changes in the cross-sectional area, or volume, of the chambers, and changes in wall thickness in response to hypertrophic stimuli. **Figure 4** shows transverse slices through the chests of three rats, orthogonal to the long axes of their hearts, acquired at diastole when the heart is fully distended. Image (A) is from a control animal, image (B) is from an animal which has received treatment which increases ventricle lumen size, and image (C) is from an animal after

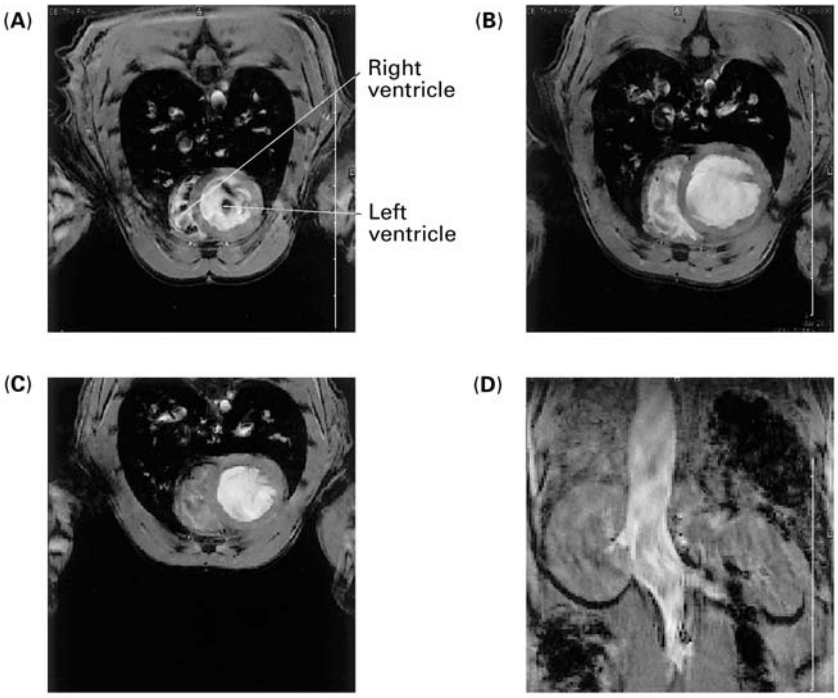



**Figure 4** Transverse 300 MHz gradient echo (TE/TR = 4/1000 ms, flip angle = 90°) images through the chests of (A) a control rat, and (B) and (C), animals subjected to experimental treatments which increase the heart ventricular lumen size, and wall thickness, respectively; image acquisition was triggered on the QRS complex of the ECG signal obtain images with the hearts in diastole and hence fully dilated. Panel (D), a coronal 'bright blood' image obtained from the same animal shown in (B), depicts the enlargement of the great vessels in the abdomen provoked by an aorto-caval shunt operation.

administration of an agent to increase wall-thickness, respectively; both effects are quantifiable from the images. The coronal image (D) shows the aorto-venacaval shunt (AVS) which caused the lumen increase seen in (B); the success of the operation is obvious from the distension of the descending aorta and inferior vena cava in this bright blood image. Without MRI the success of the AVS could only be confirmed *post mortem*.

Detection of atherosclerotic plaque would be extremely useful in evaluating therapy to reduce deposition in blood vessels. Atherosclerosis models usually involve feeding appropriate animals diets rich in fat and cholesterol to induce plaque; this process may take many months so that conventional longitudinal studies use large groups of animals killed at a number of time points. Plaque detection by MRI obviates this need for large multi-group studies. **Figure 5** shows slices from 3D T2W datasets acquired around peak cardiac systole, in a transgenic atherosclerosis-prone mouse, from the region above the heart containing the aortic arch and branch points of the vessels supplying the upper body, shown in the coronal slice (A). Panel (B) is a transverse slice through the branching vessels before induction of atherosclerosis, while (C) and (D) were obtained after a few months on a high fat diet. Strong flow and turbulence around the aortic arch region make it a primary site for plaque deposition, but cardiac and respiratory motion here make good image acquisition challenging. Nevertheless the buildup of atherosclerotic plaque is clearly visible and quantifiable in, for instance, the innominate artery.

MR angiography (MRA) can be used to define vascular anatomy. **Figure 6** shows 3D images from the brain, and the upper abdomen, of a rat acquired with a fast gradient echo technique. Static water in the field of view undergoes saturation on account of the high pulse repetition rate, but blood flowing into the field of view during acquisition gives a strong signal. The delineation of the portal vasculature achieved by this technique is further enhanced by administration of a suitable contrast reagent; the cerebral vasculature is well delineated without any enhancing agent.

## Liver

As the site of metabolism and toxicity of many xenobiotic compounds, non-invasive characterization of liver properties is of great interest. Many physiological and pharmacological interventions change liver size and morphology but its irregular shape can make quantification difficult; the use of pulse sequences giving adequate contrast between liver and

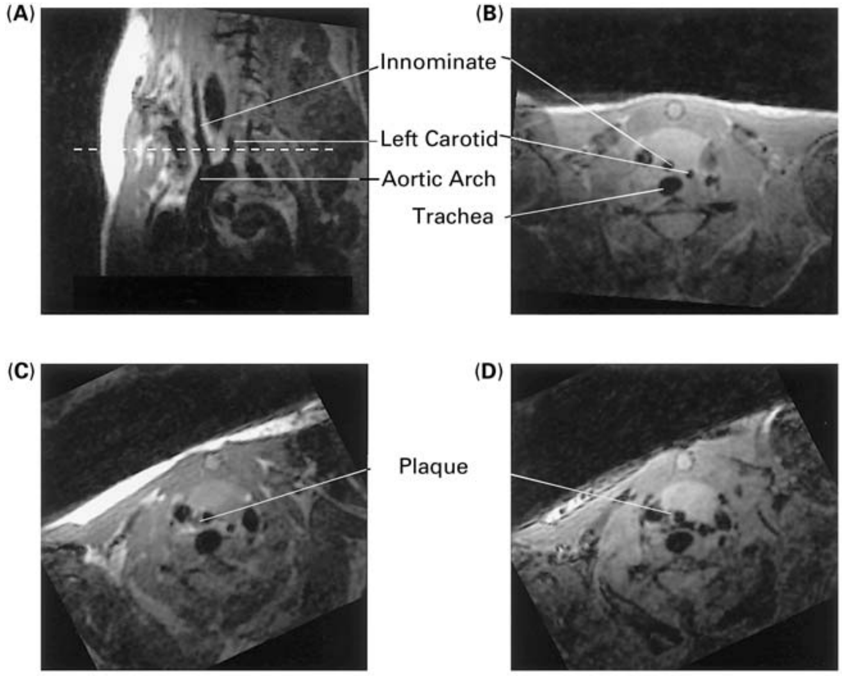



**Figure 5** Coronal (A) and transverse (B) – (D) image slices through the aortic arch region of an atherosclerosis-prone mouse acquired before (A) and (B) and 17 weeks after (C) and (D) commencement of a high fat diet, selected from 300 MHz 3D T2W datasets. Plaque is arrowed in (C) and (D); perivascular fat is removed from the latter by a fat suppression procedure. In these spin echo (TE/TR = 13/1000 ms) images triggered in full systole 65 ms after the QRS wave, rapidly moving blood gives no NMR signal and so appears black.

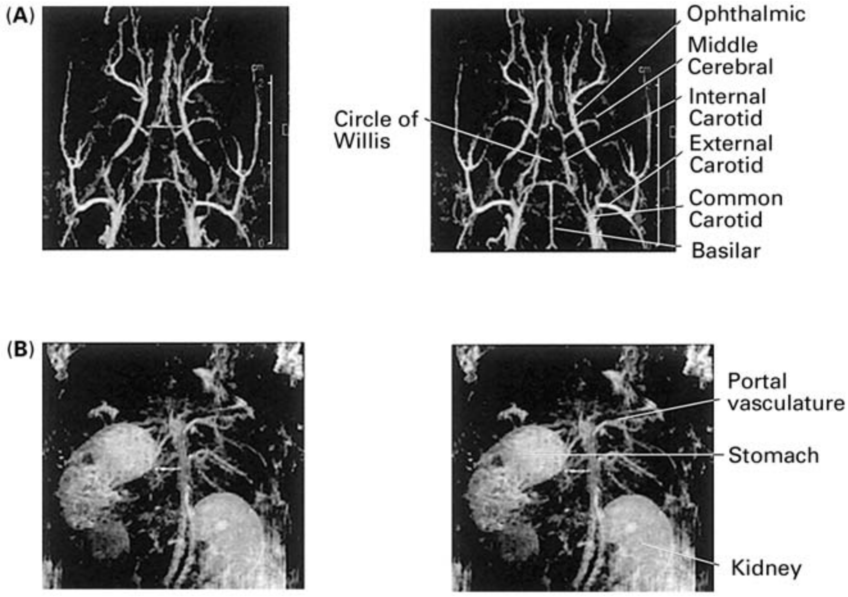



**Figure 6** 'Stereo pairs' of 'maximum intensity projection' bright blood MR angiograms acquired from rat brain (A) and abdomen (B). Contrast between flowing and static fluid was enhanced in (B) by administration of a colloidal magnetite contrast reagent which shortens $T_2$ and $T_2$* of blood relative to static tissue. The 3D effect can be best appreciated by viewing the images through stereo viewing glasses.

surrounding tissues is essential. **Figure 7** shows a series of coronal slices from a 3D image from a rat abdomen acquired with a T1W method. Such data allows accurate liver volume quantification and measurement of changes induced by natural diurnal variations and feeding, and hypertrophic stimuli. Many stresses also cause changes in liver ultrastructure; although *in vivo* MRI cannot resolve microscopic necroses, these often manifest in changes in the gross MRI properties of the tissue reflected in changes in relaxation or diffusion contrast, or altered susceptibility to contrast reagents. Note the excellent delineation of other abdominal organs, particularly the stomach, kidneys and adrenals, and the abdominal aorta.

## Kidney

Because it receives such a high proportion of the cardiac output, this organ is another important site of toxicity. T2W and proton density images delineate its anatomically and functionally distinct zones. **Figure 8** shows a slice along the median plane of a T2W 3D image of the kidney of a healthy rat. The divisions of the organ into outer and inner cortex, medulla and papilla, are obvious, as are neighbouring structures like the adrenal gland and fat pads. Treatment of the animal with regiospecific nephrotoxins produces characteristic changes in the MR images. Thus an inner cortical toxin brightens the corticomedullary boundary due to anomalous water buildup in this region. A papillary toxin evokes buildup of water in the inner zones of the organ leading to loss of medullary–papillary contrast, and swelling. Areas of anomalous MRI appearance correlate well with necrotic areas assessed by *post mortem* histology.

## Musculoskeletal system

Articular cartilage is readily visible by MRI. **Figure 9A** shows spin echo T2W image slices through the long dimension of a tibio-tarsal (ankle) joint of a rat subjected to an arthrogenic procedure. Degradation, remodelling, and swelling of the joint as the disease progresses can be clearly seen. **Figure 9B** displays images acquired from joints excised *post mortem* from a control and an arthritic rat; they were acquired on an instrument custom modified to operate with an autosampler – an example of high throughput biological MRI data acquisition.

## Oncology

MRI is a powerful technique for investigating the progression and properties of experimental tumours, as exemplified in **Figure 10**, which shows slices through a GH3 pituitary tumour implanted in a rat. Distinction of the tumour from surrounding tissue on the basis of relaxation time differences, and measurement of its volume, is straightforward. The left hand images were obtained with gradient (top) and spin echo (bottom) methodologies respectively while the rat breathed a normal air–anaesthetic mixture. The

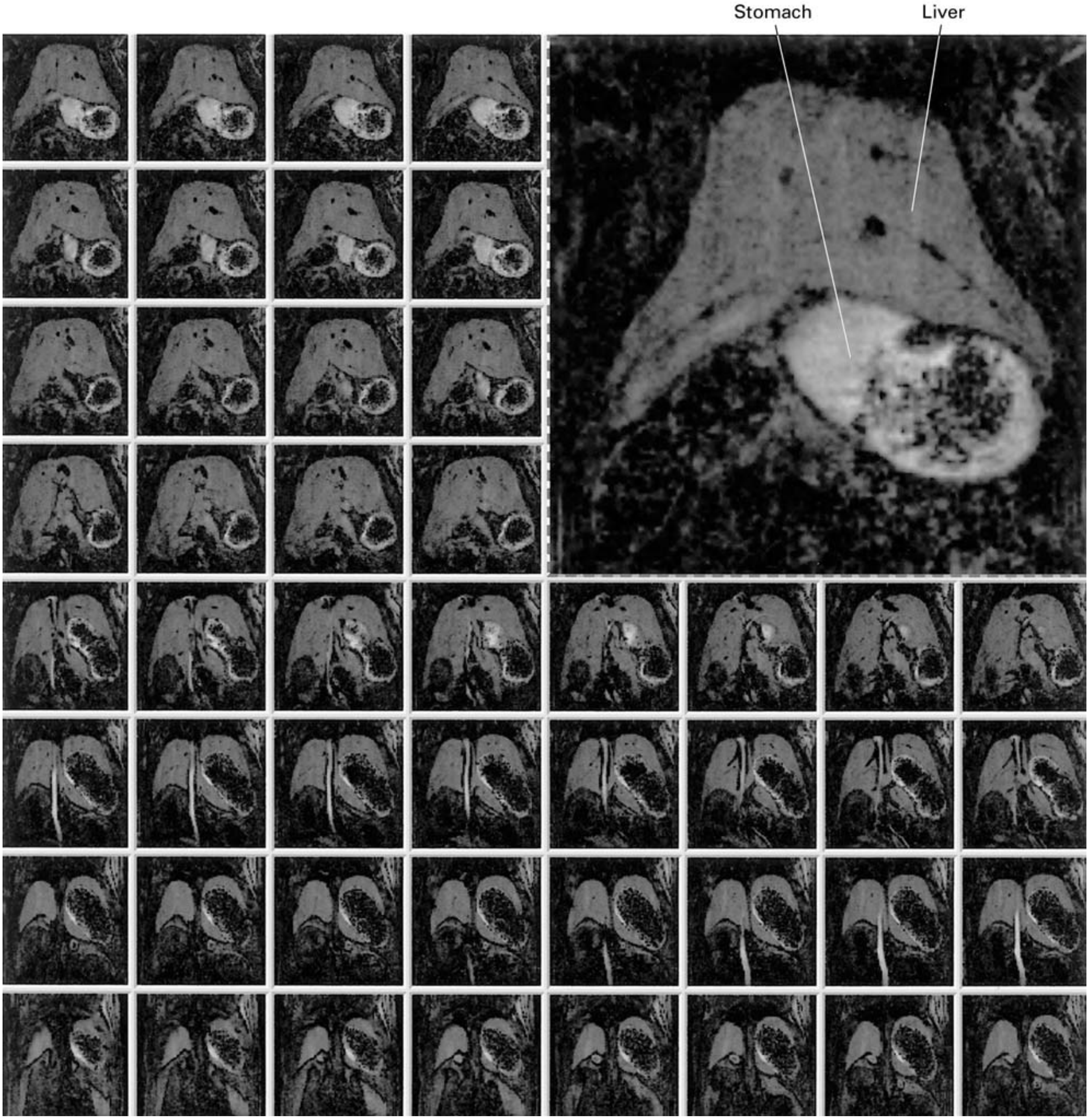



**Figure 7** Contiguous coronal sagittal slices through a 3D dataset (300 MHz) acquired from the upper abdominal area of the rat. The acquisition method, combining inversion recovery (950 ms) and segmented (16, TE = 3.3 ms) low flip angle (~ 30°) fast gradient echo readout, was designed to optimize contrast between liver and surrounding structures, but note also the excellent definition of the kidneys, stomach, and moving (bright) blood in the descending aorta. Abdominal fat was suppressed by selective saturation 3.25 parts per million (975 Hz) upfigeld of the water signal before readout.

right hand images were obtained after increasing the $CO_2$ content of the breathing mixture to 5% – a powerful vasodilatory stimulus. Oxygenated haemoglobin increases in the tumour reducing T2* relaxation, producing more signal in the gradient echo T2*W image. This is a dramatic example of the use of BOLD to study functional activation.

## Future developments

BOLD methodologies aided by fast techniques like echo planar imaging (EPI) in high field magnets promise the localization of the sites of action of neuroactive compounds. Cheaper actively shielded magnets will facilitate the use of MRI in biology,

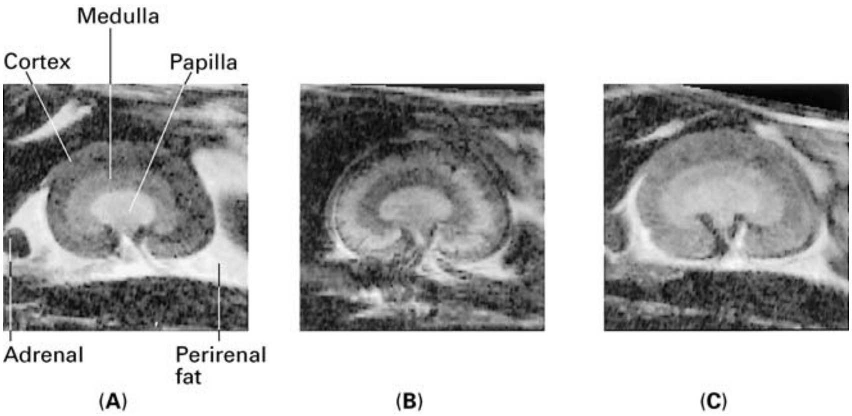



**Figure 8** Slices through the median planes of 3D 300 MHz T2W images (TE = 6.5 ms, TR = 1.5 s, multiecho segmentation, or RARE, factor = 32, $TE_{effective}$ = 104 ms) of kidneys from a control rat (A), and from rats treated with an inner cortical (B) and papillary (C) toxin. Note the clear differentiation in the control kidney between cortex, medulla and papilla, and also the good definition of perirenal fat and adrenal glands. Note also the evolution of a hyperintense band in the cortical toxin-treated kidney reflecting derangement of renal tubular function and water buildup in this region. The papillary toxin evokes a loss of papillary–medullary contrast. Marked swelling of both treated kidneys is also obvious and easily quantifiable.

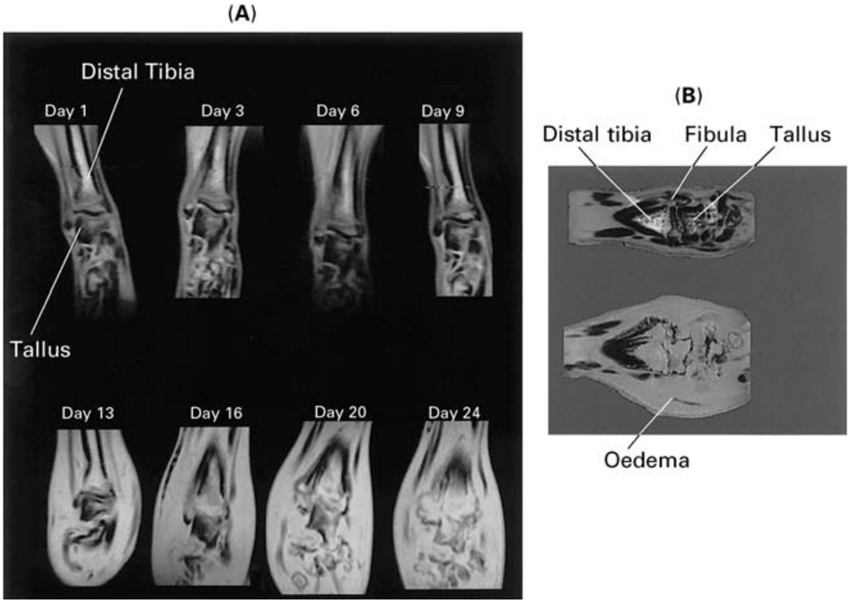



**Figure 9** (A) MR images from longitudinal assessment of degeneration of the posterior tibio-tarsal joint of a rat, rendered arthritic by intra-venous injection of a *Mycobacterium butyricum* suspension at Day 1 (200 MHz, TE/TR = 9/2500 ms, 100 × 100 μm in plane resolution, 1 mm trans-plane resolution). (B) 400 MHz images of excised tibio-tarsal joints from control and adjuvant-arthritic rats, acquired using autosampler technology (TE/TR = 8/1000 ms, 70 × 70 × 250 μm resolution). Reproduced by permission of Dr. Rasesh Kapadia, SB Pharmaceuticals, Upper Merion, PA.

pharmacology and toxicology as the systems become less demanding of laboratory space. Robust acquisition and processing software will remove the routine conduct of biological MRI from the hands of the NMR expert and place it in those of the biologist. Complete automation of *in vivo* experiments is unlikely, but fully automated imaging of fixed tissue is already possible; 3D images of fixed tissue, which can be subjected to a battery of image analysis procedures, will become valuable complements to conventional fixed tissue histology. Image analysis is the rate-limiting step in many experiments.

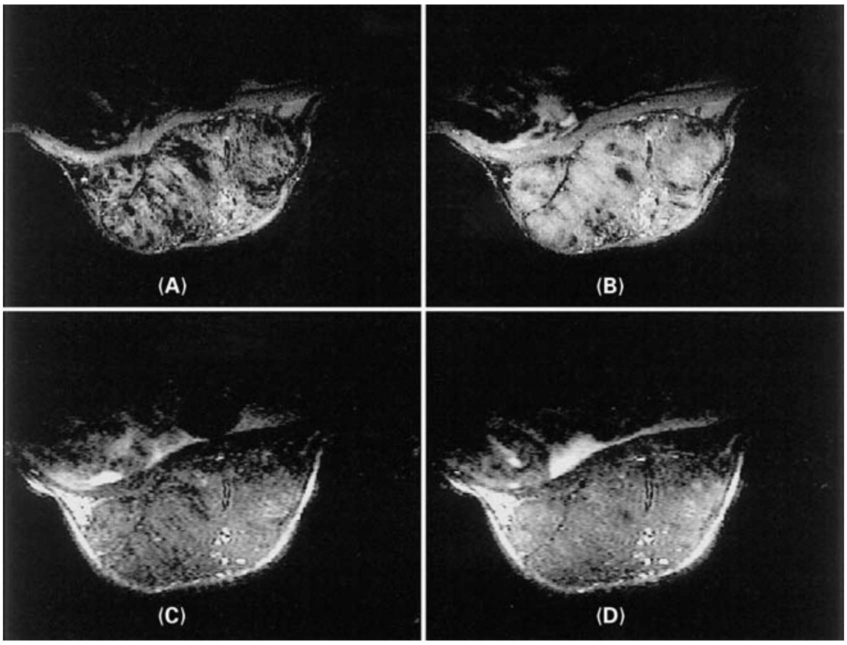



**Figure 10** 200 MHz MR images of a transplanted rat GH3 pituitary tumour. The top pair of images (A) and (B) were acquired with a T2*W gradient echo method (TE/TR = 20/80 ms, flip angle 45°) and the bottom pair (C) and (D) with T2W (TE/TR = 20/300 ms). The left hand images were acquired while the animal breathed normal air–anaesthetic gas mixture, while the right hand images were acquired shortly after switching the breathing mixture to carbogen (5% $CO_2$). Note the striking increase in intensity in the T2*W image as blood flow to the tumour increases due to vasodilation (B). This vasodilation is also reflected in the increase in T2W intensity in blood vessel cross-sections (D). Reproduced by permission of Dr Simon Robinson and Professor John Griffiths, St George's Hospital Medical School, London.

Perfection of automatic image coregistration and segmentation methods promise to break this logjam. MRI will be increasingly combined with *in vivo* spectroscopy, and other imaging methods like positron emission tomography (PET) to produce simultaneous anatomical, functional, metabolic and drug distributional information. Finally the interface between experimental and clinical MRI will strengthen as clinical trials are planned on the basis of laboratory protocols and *vice-versa*.

## List of symbols

AVS = aorto-venacaval shunt; BOLD = blood oxygen level determination; CNS = central nervous system; DW = diffusion weighted; EAE = experimental allergic encephalomyelitis; ECG = electrocardiograph; EPI = echo planar imaging; FOV = field of view; MCAO = middle cerebral artery occlusion; MRA = magnetic resonance angiography; MTC = magnetization transfer contrast; RARE = rapid acquisition with repeated echo; RF = Radio frequency; T1W = $T_1$ weighted; T2W = $T_2$ weighted; T2*W = $T_2$* weighted; TE = echo time; TR = repetition time.

*See also:* **Chemical Shift and Relaxation Reagents in NMR; Diffusion Studied Using NMR Spectroscopy; *In Vivo* NMR, Methods; *In Vivo* NMR, Applications – $^{31}$P; *In Vivo* NMR, Applications, Other Nuclei; MRI Applications, Clinical; MRI Applications, Clinical Flow Studies; MRI Instrumentation; MRI Theory; NMR Microscopy; NMR Relaxation Rates.**

## Further reading

Anderson CM, Edelman RR and Turski PA (1993) *Clinical Magnetic Resonance Angiography*. New York: Raven Press.

Bachelard H (1997) *Magnetic Resonance Spectroscopy and Imaging in Neurochemistry*, New York: Plenum.

Bushong SC (1996) *Magnetic Resonance Imaging: Physical and Biological Principles*, St Louis: Mosby-Year Book.

Callaghan PT (1991) *Principles of Nuclear Magnetic Resonance Microscopy* , Oxford: Clarendon.

Chen C-N and Hoult DI (1989) *Biomedical Magnetic Resonance Technology*. Bristol: Institute of Physics.

Elster, AD (1994) *Questions and Answers in Magnetic Resonance Imaging*. St Louis: Mosby-Year Book.

Flecknell, P (1996) *Laboratory Animal Anaesthesia*. London: Academic.

Gadian, DG (1995) *NMR and its Applications to Living Systems*. Oxford: Oxford University Press.
Underwood R and Firmin D (eds.) (1991) *Magnetic Resonance of the Cardiovascular System*. Oxford: Blackwell Scientific.
Yuh W, Brasch R and Herfkens R (eds) (1997) *Journal of Magnetic Resonance Imaging (Special Edition – MR Contrast Reagents)* 7: 1–262.

# MRI Applications, Clinical

**Martin O Leach**, The Institute of Cancer Research and The Royal Marsden Hospital, Sutton, Surrey, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

In 1973 both Lauterbur, and Mansfield and Grannell, proposed that a shift in resonance frequency, induced by a spatially varying magnetic field, could be used to encode the spatial location of nuclear magnetic resonance signals. Developments in the late 1970s demonstrated the feasibility of magnetic resonance imaging and led to the construction of clinical instruments, with the first whole-body image published in 1977 by Damadian and colleagues. Initially a range of different imaging techniques were employed, with commercial developments at first using filtered back-projection. While this is the standard method of reconstruction in X-ray computed tomography (CT), the limited homogeneity of early magnets, together with the inherent variations in human magnetic susceptibility, gave rise to considerable image artefacts. These approaches were superseded by spin-warp imaging, introduced by Edelstein and colleagues in 1980, an extension of Kumar and colleagues' Fourier zeugmatography technique. Spin-warp imaging remains the method used for most clinical magnetic resonance imaging. Imaging technique are discussed in more detail by Morris and Leach as given in the Further reading section.

The ensuing 20 years saw an unprecedented development in the scope and quality of magnetic resonance imaging, compared with the growth of previous medical imaging techniques. Although the advent of CT revolutionized diagnosis by providing high-quality cross-sectional images, its use has generally been limited to the detection and measurement of anatomical abnormality and it provides limited functional information. As CT was well-established when MRI was introduced, MRI initially supplied supplementary information, particularly in neurological examinations where the increased soft-tissue contrast of MRI and lack of bony artefacts allowed better depiction of the brain and spinal cord, together with improved visualization of physiological processes. Hardware has progressively developed, with the introduction of superconducting magnets leading to more stable and homogeneous magnets and allowing the introduction of higher-field magnets of up to 1.5 T to many hospitals. Magnet field strengths now range from 0.2 T, often with open configurations (based on electromagnets or resistive coils), aiding orthopaedic, paediatric and interventional applications, through 0.5 T (superconductive or permanent, with open designs again possible) used for a wide range of applications, to 1.0 T and 1.5 T superconducting designs, with manufacturers now developing 'short-bore' magnets with flared apertures to increase patient acceptability. High-field magnets are used where signal-to-noise is a principal concern, for angiography, functional MRI (brain activation), cardiology and real-time imaging. At 1.5 T, magnetic resonance spectroscopy is also possible, with many instruments being capable of proton spectroscopy, and some also having facilities for broad-band spectroscopy. Research sites have installed higher-field magnets, with many 3.0 T installations, some at 4.0–4.7 T and recent installations at 7 T and 8 T. These systems are primarily used for spectroscopy and for brain activation studies.

There has been a range of further developments in hardware. These include shielded gradient coils, facilitating high-speed imaging by reducing eddy currents, and large increases in the strength and switching speed of gradients, allowing clinical implementation of snapshot imaging, echo planar imaging and similar real-time techniques. Circularly polarized and phased-array coils have significantly increased the sensitivity of measurements, with modern systems having a wide range of coils. Automatic shimming techniques have improved fat signal suppression, as well as aiding spectroscopy. Self-shielded magnets have eased the

installation requirements for many clinical systems. These improvements have been accompanied by advances in pulse sequence design, versatility and reconstruction speed. Packages for specific clinical specializations are now available, providing pulse sequences and analysis techniques tailored to particular applications, e.g. functional neuroimaging and cardiac packages. In addition, a range of contrast agents with differing pharmaceutical properties have been developed that are leading to new clinical applications.

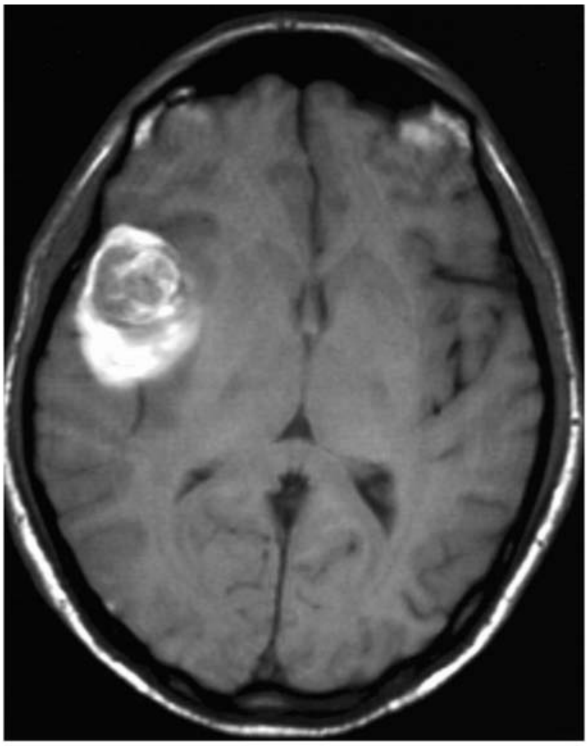
## Anatomical imaging

Clinical applications of MRI primarily make use of the high soft-tissue contrast, which can be readily manipulated by appropriate choice of pulse sequences, to demonstrate cross-sectional anatomy at any arbitrary orientation. One of the initial motivations for developing clinical MR imaging instruments was the observation that tumours had long $T_1$ relaxation times. Although this was shown not to provide a unique discriminator for cancer, the different $T_1$ and $T_2$ relaxation times, together with other intrinsic properties affecting the MR signal, allow contrast to be changed between tissues by selecting appropriate pulse repetition times and flip angles ($T_1$ weighting) and echo times ($T_2$ weighting). This allows abnormal or distorted anatomy to be seen, and aberrant tissues can often be identified by different relaxation properties. $T_1$ and $T_2$ relaxation times reflect the environment and ease of movement of water and fat molecules. The greater the water content and the greater the freedom of movement, the longer are $T_1$ and $T_2$. When water is tightly bound, magnetization transfer imaging techniques can be used to interrogate this bound compartment by exploiting the short $T_2$ and broad line shape. An off-resonance (several kHz) irradiation suppresses the bound component, without directly affecting unbound water. However, the signal of the unbound water is subsequently reduced by exchange with the partially saturated bound component. A difference image reveals the degree of magnetization transfer.

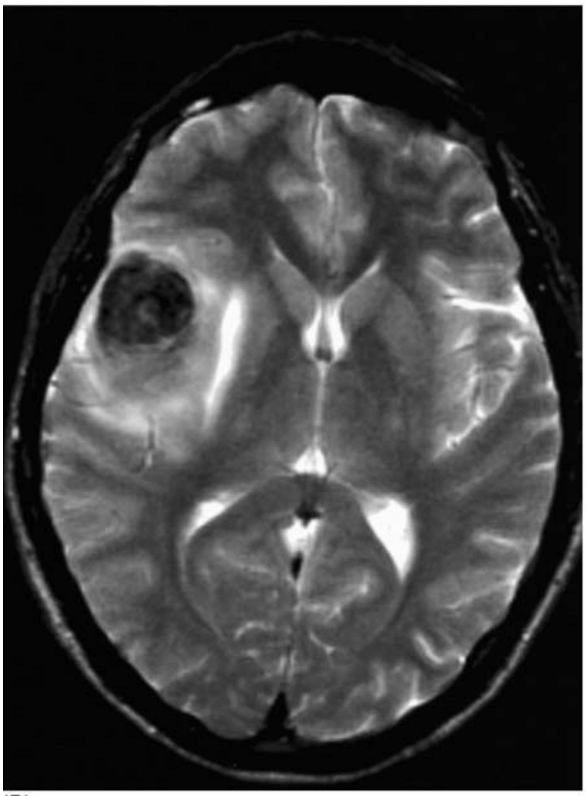
A basic clinical examination will employ both $T_1$- and $T_2$-weighted multislice imaging sequences, chosen in a particular plane. Typically a set of scout images (very rapid $T_1$-weighted images in several orientations) will be acquired to aid the prescription of these images (orientation and number of slices, etc.). A fast spoiled gradient-echo image (e.g. fast low angle single-shot (FLASH)) might be chosen with a 300 ms repetition time (TR), a 12 ms echo time (TE) and a 70° flip angle ($\alpha$), to provide $T_1$ weighting. A dual-echo spin-echo sequence with TR=2 s, TE=30 ms and 120 ms, and $\alpha$=90° would provide, respectively, proton density and $T_2$-weighted images. The gradient-echo image is subject to signal loss in areas of magnetic field inhomogeneity, or variations in magnetic susceptibility, for example in the brain adjacent to air-filled sinuses or near sites of previous haemorrhage. The effect can be minimized by selecting a very short TE, or using a $T_1$-weighted spin-echo sequence (see **Figure 1**). With these conventional sequences, straightforward anatomical examinations can be performed in most parts of the body that are free from movement. A number of additional gradient-echo sequences are available that exploit the principle of steady-state free precession. FISP (fast imaging with steady state precession) maintains the steady-state signal, and does not suffer signal loss from flowing blood, providing high signal from long-$T_2$ fluids, with a signal that does not depend strongly on TR. This is valuable for generating MR myelograms or for angiography. PSIF (a time reversed FISP sequence also called CE-FAST) provides strong $T_2$ weighting that is a function of TR.

A further basic sequence that is widely used is the inversion recovery sequence, in which the magnetization is initially inverted, and then sampled with a 90° pulse at an inversion time (TI) after the 180° pulse. This can provide a greater range of $T_1$-weighted contrast, and has the particular property that it can be used to null signal from a particular tissue on the basis of its $T_1$ relaxation time, by selecting the TI to sample signal from that tissue as it recovers through zero longitudinal magnetization. A widely used variant of the inversion recovery sequence is the STIR (short $\tau$ inversion recovery) sequence, which is used to null the signal from fat, which usually has a bright signal on $T_1$-weighted images and can obscure important anatomical detail. Similar sequences can be used to null the cerebrospinal fluid (CSF) signal in spine imaging, allowing the spinal cord to be clearly seen. A further variant is the FLAIR (fluid attenuated IR) sequence, which nulls CSF in the brain, enhancing visualization of brain tissue. Alternative methods are available to obtain fat, or water, images using selective excitation with, for example, binomial pulses, or conventional frequency-selective pulses, or by employing a multi-acquisition method sensitive to the phase difference between fat and water (the Dixon method).

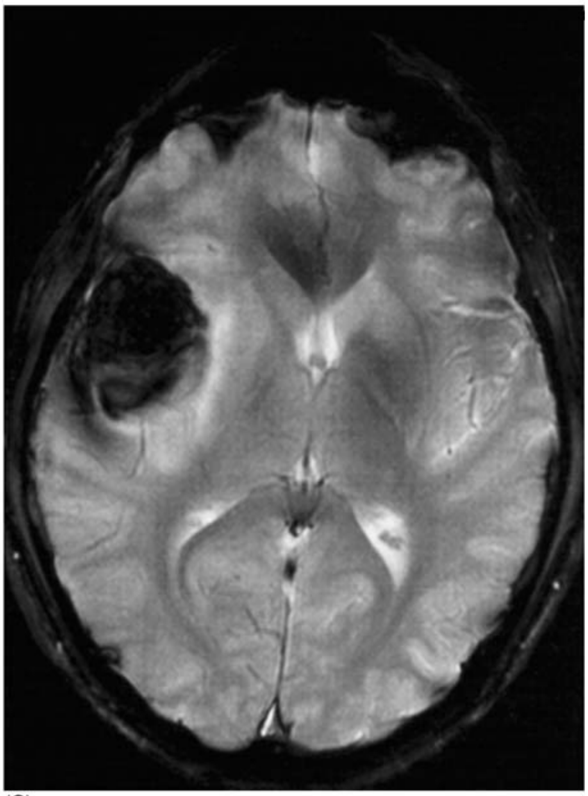
While providing excellent images in many parts of the body, acquisition times for these measurements are relatively long, reducing their value in moving tissues. In areas such as the abdomen, affected by respiratory and bowel movement, image quality can be improved by averaging, at the cost of longer measurement times. In the mediastinum, ECG triggering allows high-quality images at appropriate stages of the cardiac cycle to be obtained, despite the vigorous,



(A)



(B)



(C)

**Figure 1** Transaxial images through the brain of a patient with a haemorrhagic melanoma metastasis. (A) $T_1$-weighted spin-echo image (TR = 665 ms, TE = 14 ms, $\alpha = 80°$) showing bright signal in the regions of recent haemorrhage. (B) $T_2$-weighted turbo spin-echo image (TR = 4500 ms, effective TE = 90 ms, $\alpha = 90°$) showing bright signal from cerebrospinal fluid and low signal arising from $T_2$ shortening due to melanin deposits in the tumour. (C) $T_2^*$-weighted FLASH image (TR=1604 ms, TE = 35 ms, $\alpha = 30°$) showing increased $T_2^*$ signal loss within the tumour resulting from susceptibility changes due to melanin.

multidirectional motion. Image quality can be further improved by placing saturation slabs through moving high-signal regions, or by saturating in-flowing blood in adjacent planes. Respiratory gating, and methods of reordering phase encoding (ROPE) can also reduce motion effects. Although these techniques are still sometimes employed, major advances in imaging moving tissues, and in speeding up examinations, have been attained by a range of new rapid imaging techniques, made possible by recent advances in instrumentation. Where motion cannot be avoided, or where individual data sets building an image have to be acquired during movement, navigator echoes provide a way of accurately monitoring motion as well as providing the information necessary for correcting for the motion.

Turbo or magnetization prepared gradient-echo sequences have one or more preparation pulses, followed by a rapid succession of small flip angle pulses to interrogate the longitudinal magnetization, each encoding a different line in $k$-space, thus building up the image with only one preparation pulse. This sequence, and variants that further reduce the measurement time by reduced $k$-space sampling, provide rapid images that allow subsecond image acquisition, and a set of slices can be acquired within a breath-hold period. Preparation can include a large flip-angle pulse or an inversion pulse. Contrast and the relative weighting of spatial frequencies can be altered by changing the $k$-space sampling order. These techniques are often employed in 3D imaging sequences, to allow a 3D data set to be acquired in an acceptable time. A highly effective sequence providing rapid $T_2$-weighted measurements is the turbo spin-echo or RARE (rapid acquisition with relaxation enhancement) sequence. In this sequence, multiple echoes are acquired, each sampling a different line of $k$-space, thus speeding up acquisition of the image. A consequence of the many 180° pulses is a change in contrast in some tissues compared with spin-echo sequences, as well as increased power deposition. Contrast and resolution can also be varied by altering the $k$-space sampling scheme. Echo planar imaging uses a single-shot sequence to obtain a full image based on a single preparation or read-out pulse. This is one of the fastest imaging methods and places high demands on the gradient and acquisition system. It is now available on commercial systems, and is being applied to functional and physiological measurements, which are particularly sensitive to motion. A number of variants of the above techniques are in use, including GRASE (gradient and spin-echo), combining spin echo and gradient echo imaging and fast imaging with BURST RF excitation, which utilizes a sequence of RF pulses to generate images very rapidly.

Bone is not visible on MR images owing to the extremely short $T_2$ of hydrogen atoms in bone. The presence of bone can usually be inferred from the lack of signal, although estimation of bone volume is complicated by the relative shift in position of fat with respect to bone (the chemical shift artefact). In some areas of the body, signal voids from air spaces can also complicate interpretation. High resolution 3D imaging of joints can show excellent cross-sectional images of trabecular structure. The development of bone interferometry, based on the loss of signal in $T_2^*$-weighted images from susceptibility effects, has provided a means of measuring changes in trabecular bone mineral mass in diseases such as osteoporosis. $T_2^*$ includes the contribution of local magnetic susceptibility.

MRI is widely used in musculoskeletal and orthopaedic examinations. The use of site-specific surface coils, combined with 3D or narrow slice imaging sequences, allows the detailed structure of joints to be visualized (see **Figure 2**). Tendons can be seen as regions of low signal, and there is good contrast between cartilage, synovial fluid and the meniscus. Open magnet designs associated with fast imaging techniques facilitate kinetic imaging of joints and tissues. Absence of radiation and the ability to freeze motion have also extended the application of MRI to resolving problems in pregnancy and examining the fetus.

Contrast agents are now widely used to enhance the appearance of pathology, separating it from normal tissues based on differential uptake of a labelled pharmaceutical. These agents principally affect $T_1$, as they are usually paramagnetic compounds with several unpaired electrons. These cause increased intensity on $T_1$-weighted images in areas of high uptake because of the reduced $T_1$. They can also be used to affect $T_2$ using superparamagnetic or very small ferromagnetic particles, causing a loss of signal on $T_2$-weighted images. The most commonly used agent is gadolinium, usually chelated to diethylene triamine pentaacetic acid (DTPA) or similar compounds. The agent is injected intravenously and diffuses rapidly into the extracellular space. Its first use was to demonstrate breakdown of the normal blood–brain barrier (see **Figure 3**), but it is now widely used to delineate pathology outside of the brain, exploiting differences in blood vessel density and vascular permeability. Diagnosis may be based on the standard enhanced images, but often contrast is improved by subtracting post-contrast from pre-contrast images, or by performing fat-suppressed imaging.

More complex approaches exploit the dynamic behaviour of contrast agents to obtain physiological information, discussed further below. In some tissues,

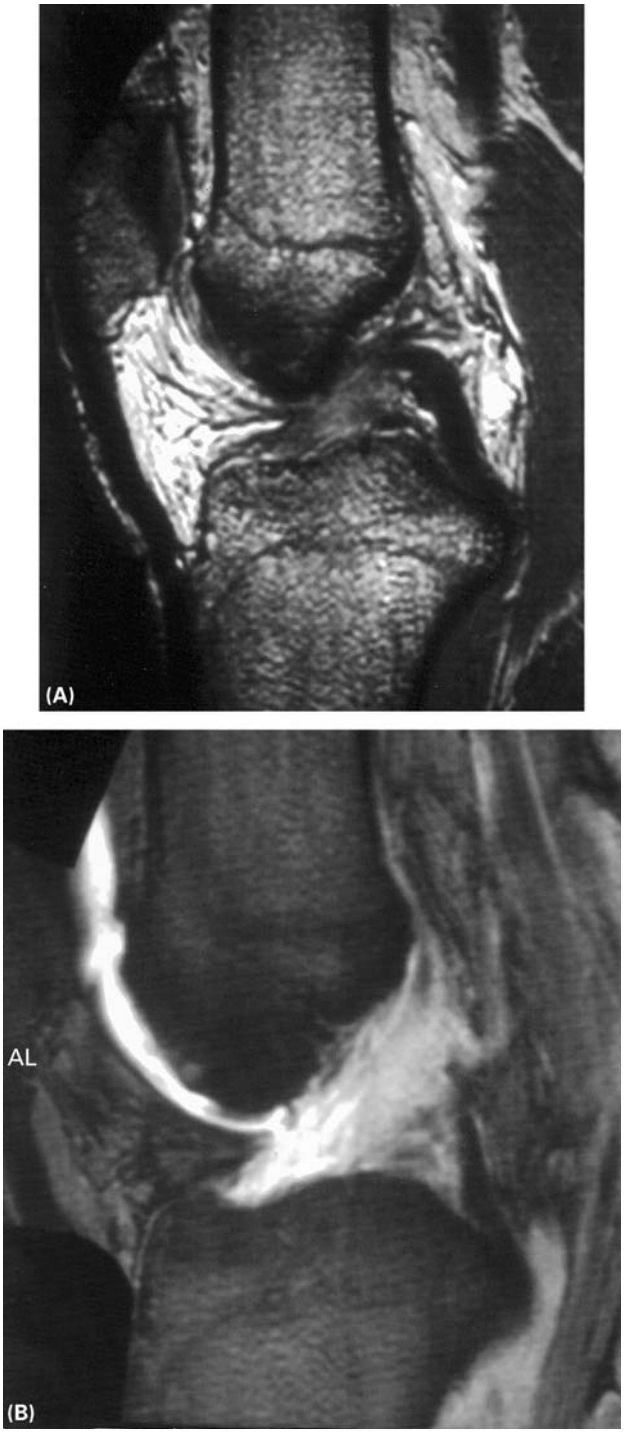



**Figure 2** Sagittal images through the knee of a patient with a ruptured anterior cruciate ligament. (A) A 3D FLASH image (TR = 25 ms, TE = 10 ms, $\alpha = 50°$) providing thin slice images (1.5 mm) showing trabecular bone structure. (B) A turbo spin-echo image (TR = 4500 ms, effective TE = 96 ms, $\alpha = 90°$) showing synovial effusion and oedema (4 mm slice thickness).

magnetization transfer techniques are employed to further improve contrast. While many agents are in development, the other major class of agent entering

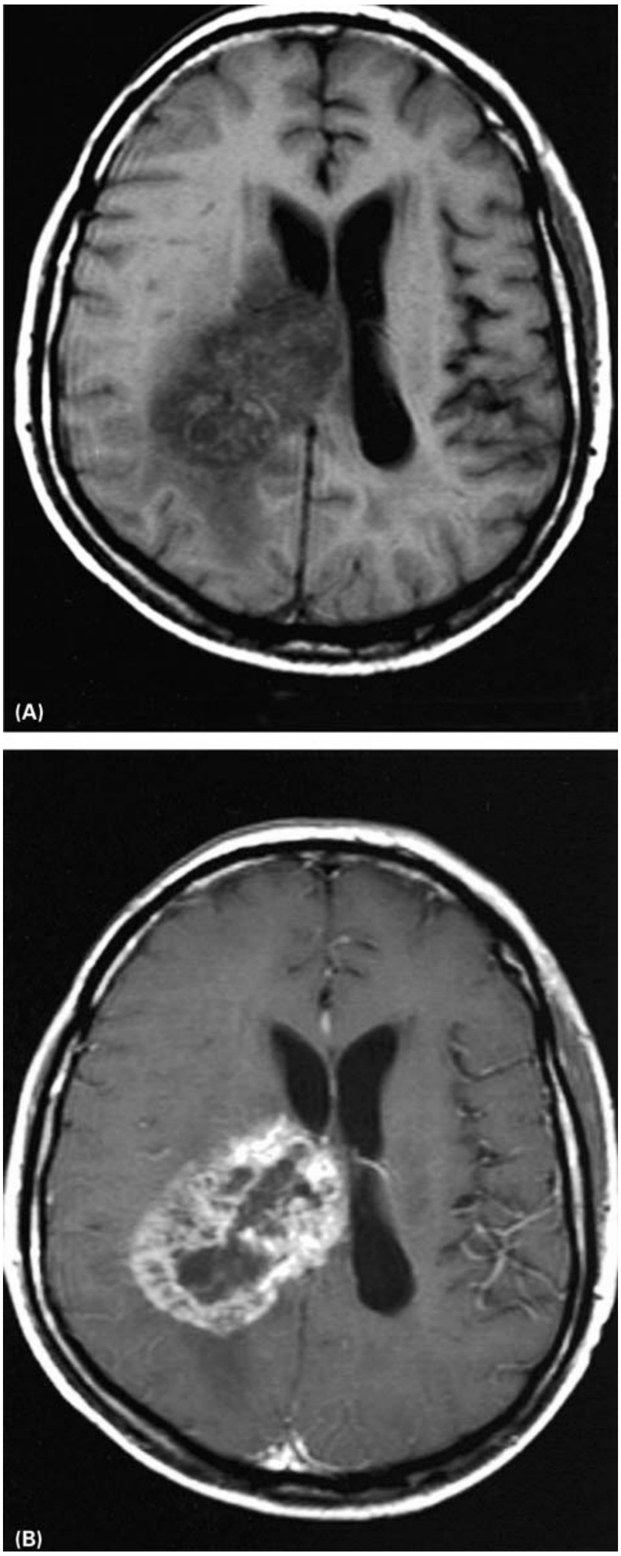
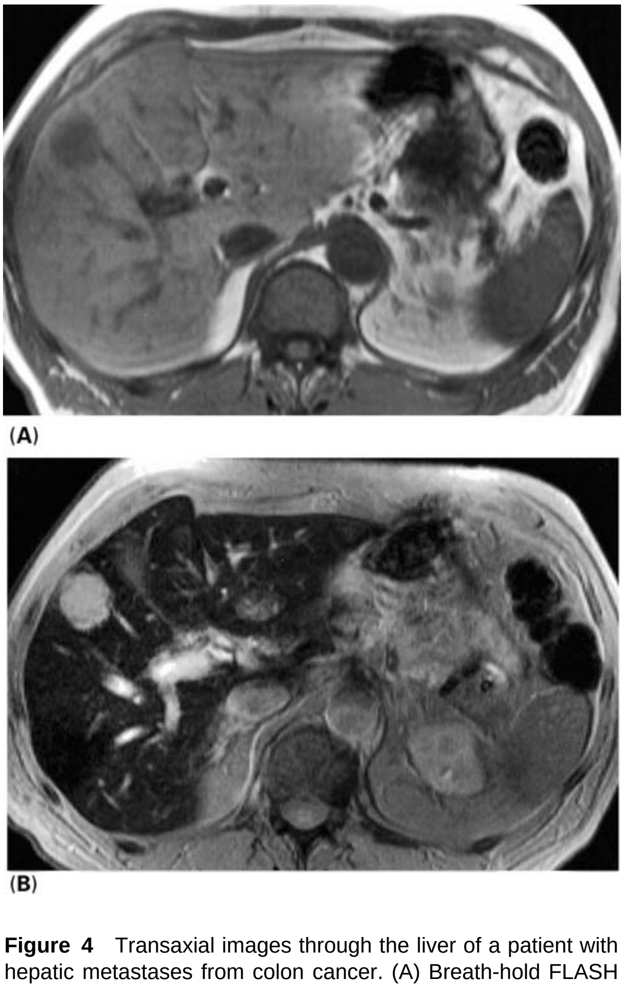



**Figure 3** Transaxial images through the brain of a patient with a glioma. (A) $T_1$-weighted spin-echo sequence showing a large tumour in the deep cerebral white matter. (B) The same slice following injection with 0.1 mmol kg$^{-1}$ of gadolinium contrast agent. In the latter slice, the sequence also had gradient moment rephasing to reduce artefacts from flowing blood, causing a slight change in white/grey matter contrast.

clinical practice is positive ($T_1$) liver agents such as gadoterate meglumine (Gd-DOTA) (taken up in tumour) and Mn-pyridoxal-5′-phosphate DPDP (taken up in normal liver cells) and negative ($T_2$*) liver agents such as superparamagnetic iron oxide particles (SPIO) which are taken up by the mononuclear phagacytosing system (Kupffer cells), reducing signal from normal liver (**Figure 4**).

## Measuring physiology and function

The nature of magnetic resonance measurements confers sensitivity to a range of properties of water molecules that can be exploited to measure functional aspects of tissues and fluids. Probably the most widely used feature is the sensitivity of MR measurements to motion. When imaging static tissues, motion of fluids or of other tissues presents as a problem to be minimized so as to reduce artefacts. The most common effect is misregistered signals at the same frequency (same position in the read-out direction) but displaced in the phase-encoding direction. This effect can be reduced by the strategies discussed above or by the use of gradient motion rephasing sequences, where the phase gain resulting from movement in the gradient is cancelled by reversed-polarity gradients. Subtraction of pairs of images with and without these additional gradient-lobes results in images of the moving material. The flow of fluids can be measured by bolus-tracking techniques, where a slice is saturated and inflow is observed, or where a distant slice is tagged (by inversion, for example) and the appearance of the tagged blood in the slice of interest is observed. Alternative approaches make use of the phase gain occurring in moving fluids, allowing the speed and direction of flow to be calculated from phase maps (**Figure 5**). Specific sequences can directly measure flow profiles in any arbitrary direction. These techniques are used to make direct measurements of flow velocity, cardiac valve performance, vessel patency and the effects of obstruction, but can also be used to produce flow images.



(A)

(B)

**Figure 4** Transaxial images through the liver of a patient with hepatic metastases from colon cancer. (A) Breath-hold FLASH $T_1$ image (TR = 80 ms, TE = 4.1 ms, $\alpha$ = 80°) showing limited lesion contrast. (B) Proton density-weighted FLASH image (TR = 127 ms, TE = 10 ms, $\alpha$ = 40°) showing darkening of the liver from application of 15 μmol kg$^{-1}$ of superparamagnetic iron oxide contrast agent, increasing the conspicuousness of the lesion.

Based on these flow-sensitive techniques, a major area of MRI development and application has been MR angiography. A range of time-of-flight and phase-contrast techniques are used to produce 3D data sets, or direct projection views, of vascular structure. 3D data sets are usually processed to produce a set of maximum intensity projections (MIPs), at different orientations, which can then be presented as a cine-loop display, giving apparent 3D visualization of vascular structures. The sensitivity of the measurement techniques has been improved with travelling saturation sequences and by the use of contrast agents and bolus-tracking approaches (**Figure 6**). A major advantage of MR angiography (MRA) is that registered high-resolution soft-tissue images can be obtained at the same time, aiding resolution of diagnostic problems. Initially the major area of interest was in carotid artery stenosis and in vascular abnormalities in the brain. Advances in technique now allow major vessels to be evaluated throughout the body, including the lung and peripheral vascular disease (**Figure 7**). It is now possible to use such approaches to replace expensive diagnostic angiography in application including screening for brain aneurysms and selection of donors for renal transplant.

While MRI is sensitive to bulk flow in vessels, it is also possible to assess the slower nutritive blood supply or perfusion of tissues, together with vascular permeability, and to measure the diffusion of water molecules within tissues. Diffusion is usually

measured by determining the loss of signal resulting from the additional dephasing of magnetization experienced by spins moving in a magnetic field gradient. Initially this was achieved by strong pairs of gradients on either side of the 180° pulse in a spin-echo sequence, resulting in moving spins receiving a net dephasing, whereas the phase change cancelled for static spins. The loss in signal is proportional to the diffusion coefficient, but is also affected by the dimensions of the structures in which the spin can move in the time available, leading to the term apparent (or restricted) diffusion coefficient (ADC). Early measurements with this approach showed that by sensitizing the gradients in different directions, it was possible to demonstrate the orientation of white-matter tracts in the brain. Molecules travelling along the tracts could travel a considerable distance, leading to a large loss of signal, whereas those travelling across the tracts could not move far, resulting in little loss of signal. This provided a powerful tool for analysing brain structure *in vivo*, and for better understanding of anatomical distortion due to disease. In early machines, the techniques were very susceptible to eddy currents induced in the magnet by the large gradient pulses, and by small bulk movements in tissues and fluids, which could give rise to much greater signal changes than the diffusion itself. These problems have been largely overcome by real-time imaging sequences and improved hardware. Diffusion measurements now commonly apply a set of six differently gradient-sensitized sequences to evaluate both the magnitude and spatial distribution of restricted diffusion, providing a diffusion tensor measurement. The method is now of considerable importance in the diagnosis of stroke and other ischaemic disease, where increased diffusion is an early and sensitive indicator of insult, providing the possibility of early and effective intervention before cell function is irreversibly lost.

Perfusion has also been measured using variants of the tagging or outflow techniques described above, where signal or apparent relaxation time changes occur as a result of the inflow or outflow of labelled spins. Contrast-enhanced studies provide a tracer, allowing the inflow or washout of the tracer, as seen on $T_1$ weighted images, to be used to derive perfusion. This approach is complicated for those positive contrast agents currently licensed for clinical use (gadolinium-labelled chelates) as they equilibrate rapidly with the extracellular space and also relax

**Figure 6** A maximum intensity projection of a set of MR time-of-flight angiography images, showing aneurysms on the circle of Willis (bright areas left and right of brain centreline).

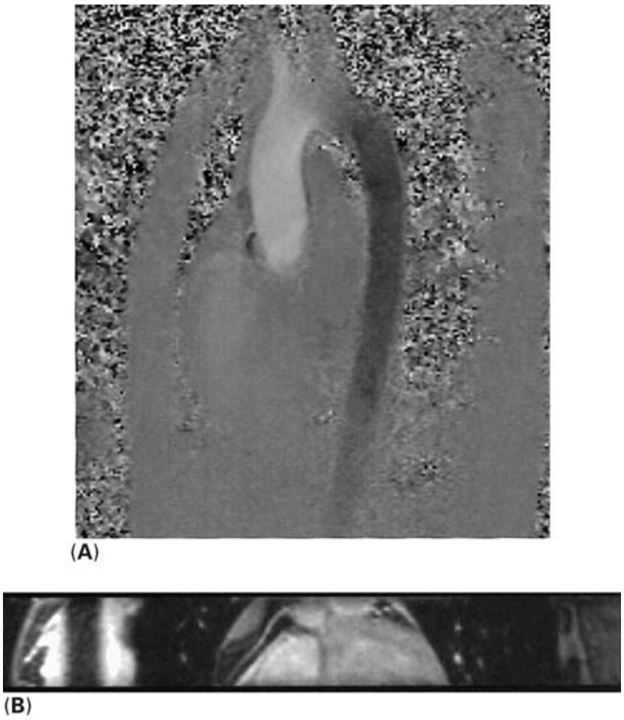



**Figure 5** Flow-sensitive images of blood flow. (A) An oblique coronal phase-contrast image through the ascending and descending aorta, where white shows flow out of the heart and up the ascending aorta, dark shows flow downwards, through the descending aorta. (B) A 3D FLASH image using navigator echo techniques to remove motion effects, showing the right coronary artery just above the aortic arch (the thin white vessel seen against a dark background, centre left of image). Both images were acquired with ECG triggering.

**Figure 7** A maximum intensity projection of a set of MR contrast-enhanced angiography images obtained from a 3D FISP sequence, following administration of a gadolinium contrast agent. The image shows the renal arteries and descending aorta (bright centre right with downward-angled renal arteries) and more faintly the upward-angled renal veins and kidneys, draining into the inferior venacava (centre left). The right kidney (to the left of the image) is reduced in size owing to involvement of a renal carcinoma (dark outline visible). At the top of the image the pulmonary veins can be seen clearly.

water molecules that distribute between the extracellular space, intracellular space and vascular space. Future generations of blood pool agents will be more effective in measuring perfusion as their distribution will be limited to the vascular space. An alternative, and more effective, approach is to make use of the local change in magnetic susceptibility that occurs as a bolus of high-concentration contrast agent passes through the vascular bed. Prior to the contrast agent equilibrating between the intra- and extravascular space, there is a large susceptibility gradient around each capillary, which will result in signal loss due to dephasing on gradient-echo sequences. Using $T_2$*-weighted fast-imaging sequences, this transient phenomenon can be detected. It is proportional to the blood volume in the image, and the timing is related to blood flow (**Figure 8**).

By using $T_1$-weighted images with positive extracellular contrast agents, combined with modelling techniques, it is also possible to calculate the volume of the extravascular extracellular space, and to calculate the permeability–surface area product governing the rate of transfer of the contrast agent out of the vascular system. This measure is of particular interest, as developing tumour vasculature is characteristically leaky. Development of this new vasculature by tumour-initiated growth factors is believed to be a necessary condition for tumour growth above the limit at which nutritional requirements can be supplied by simple diffusion, and is a target for new generations of anti-angiogenic therapies. Permeability and vascular volume can be calculated on a pixel-by-pixel base as colour-mapped functional images and superimposed on anatomical images. Such measurements require quantitative imaging sequences. Much useful information can be obtained by characterizing the behaviour of contrast uptake and washout, and studies have shown that this can be of value in identifying and characterizing tumours, and in monitoring response.

MRI also provides a number of approaches by which tissue motion can be measured. In principle, phase maps or tagging can be employed, although, owing to slice thicknesses larger than or comparable to the motion, this is rarely done. A more widely used approach in cardiac wall motion studies is the application of a one- or two-dimensional criss-cross pattern of parallel signal-suppressed lines on the object. After a defined period, short compared with $T_1$ relaxation, an image is read out and the movement of tissue relative to the original grid can be deduced. Appropriate software can provide for sophisticated wall motion studies (**Figure 9**). Associated with techniques for monitoring ventricular function based on flow, tissue perfusion studies and assessment of cardiac artery patency (see **Figure 5**), these provide a powerful range of techniques for cardiology.

A recent area of development has been the generation and application of hyperpolarized gases. Both $^3$He and $^{129}$Xe can be prepared at high nuclear polarizations (10–50%) compared with $^1$H (0.0006% at 1.5 T). This provides a very high signal, and initial measurements have shown the potential to image the lung air-spaces. This complements recent advances in fast very short echo-time sequences that have allowed the lung parenchyma to be imaged, as well as MRA approaches imaging the lung vasculature. Most measurements have been made with $^3$He, which has low solubility in tissues. $^{129}$Xe is of particular interest in measuring perfusion and other properties of tissue spaces, where it demonstrates a large tissue composition-dependent chemical shift. The potential for intravenous delivery using perfluorocarbon blood-substitutes and other suitable media is being evaluated. Studies using hyperpolarized gases require new imaging

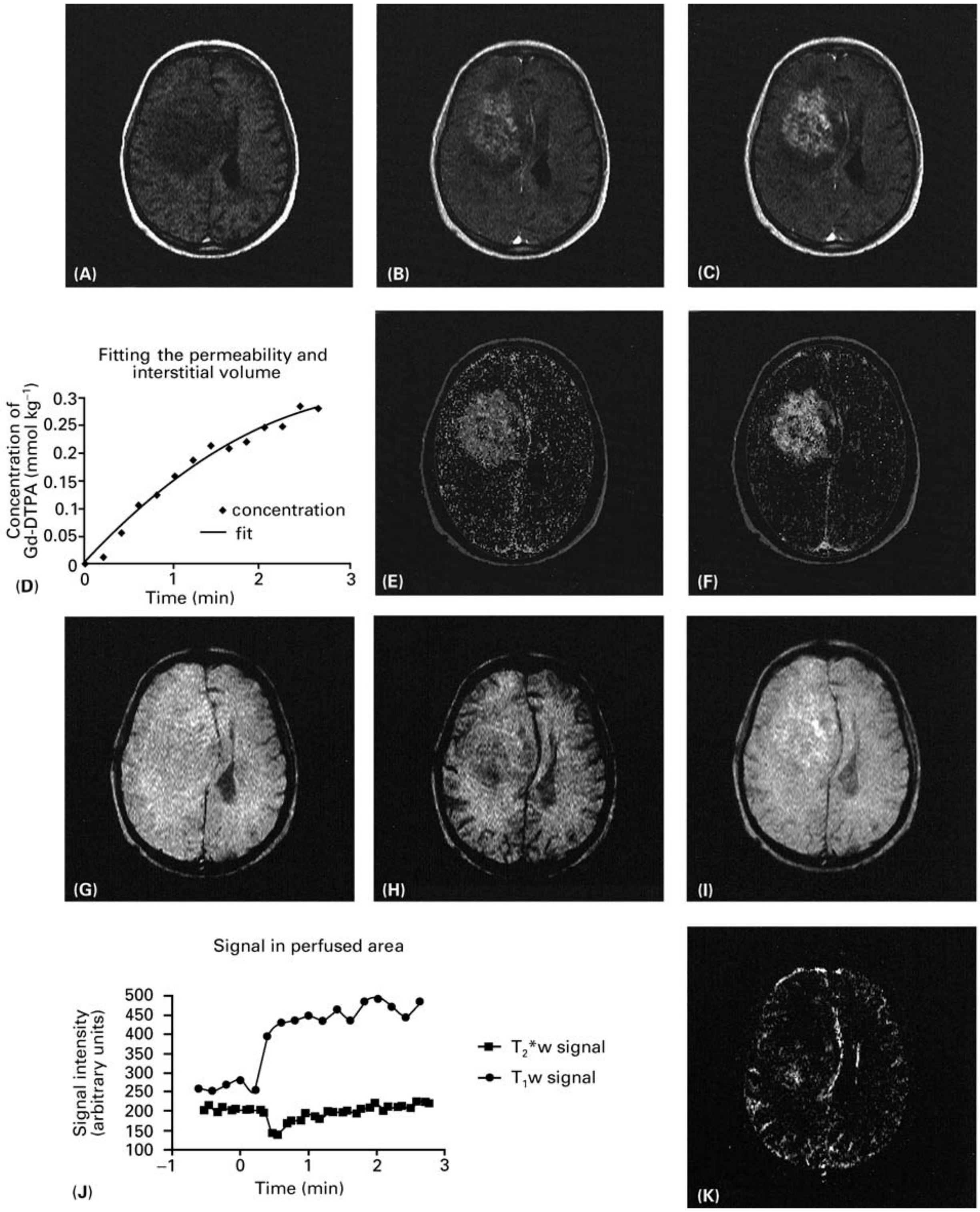



**Figure 8** Figures showing quantitative measurements of permeability and blood volume in trans-axial images through the brain of a patient with a recurrent glioma being treated with chemotherapy. (A–C) rapid $T_1$-weighted images showing uptake of the contrast agent (Gd-DTPA) in the tumour (pre-contrast, 0.8 min and 2.6 min). (D) A graph of the calculated concentration of Gd-DTPA in a volume of interest (points) compared with a constrained fit to a multicompartment model used to derive physiological features. (E) Pixel-by-pixel map of vascular permeability. (F) Pixel-by-pixel map of interstitial volume. (G–I) $T_2^*$ images obtained using the same sequences as for images (A–C) (pre-contrast, 0.28 min, 2.79 min), showing loss of signal due to the passage of contrast agent through the capillary bed; (J) Graph of signal intensity on $T_1$-weighted images, and on $T_2^*$-weighted sequences, where the integral of the signal drop on the latter curve is proportional to relative blood volume. (K) Pixel-by-pixel relative blood volume map. These images and calculated maps were obtained using sequences and methods developed by Ms I. Baustert and Dr G. Parker at the Royal Marsden Hospital/Institute of Cancer Research.

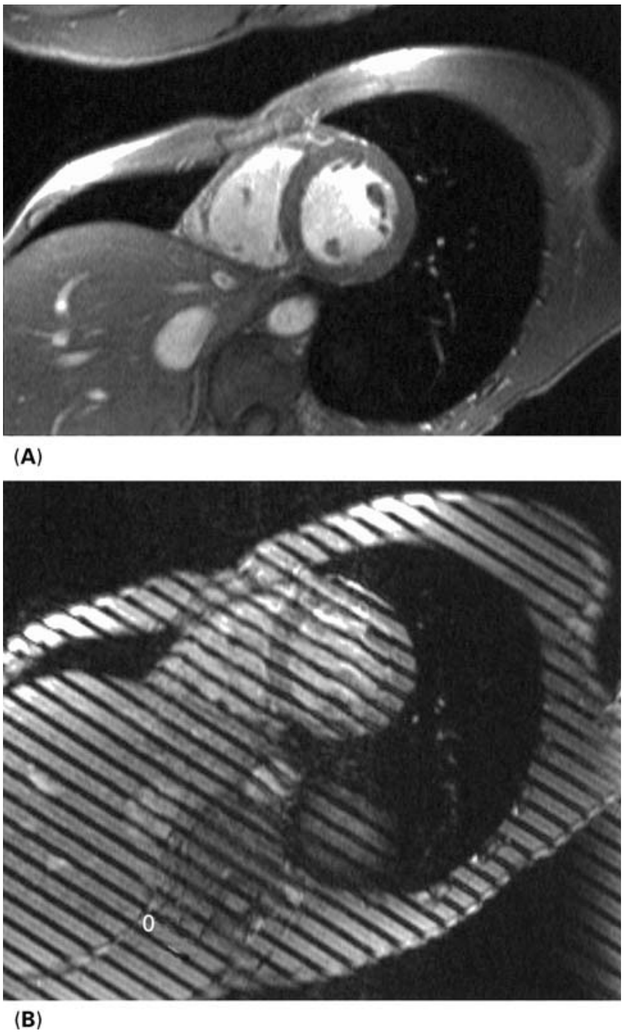


Figure 9 ECG-gated images through the heart showing bright blood and orientated to show left ventricle wall muscle. (A) Showing anatomy. (B) Tagged in one direction at early systole, to demonstrate myocardial wall motion.

approaches, as the polarization is exhausted by sampling and signal can only be restored by delivery of fresh hyperpolarized gas.

A major new area of functional MRI has been the discovery that brain activity associated with specific functional tasks causes a change in MR signal observable on $T_2$*-weighted imaging sequences. This is believed to result from brain activation causing increased local blood flow, which then provides an increased oxygen supply exceeding the increased demand. The blood thus contains proportionately less paramagnetic deoxyhaemoglobin, reducing the susceptibility between blood and surrounding tissues and thus reducing the susceptibility-induced signal loss. This approach provides higher-resolution images than the positron emission tomography techniques used previously, and allows functional activation measurements to be related to high-resolution images of local anatomy. Typically imaging is conducted with and without a stimulus, with subtraction or comparison of the two image sets to provide a difference image demonstrating the region of activation (**Figure 10**). Single-shot techniques are now being developed. The approach is being employed for basic neurological and psychiatric research, as well as in conditions affected by brain function. Signal-to-noise improves with field strength, and a number of centres are exploring the application of higher-field machines to improve the quality of these measurements. As with many of the more advanced techniques, motion and registration between measurements present problems, and sophisticated motion correction, image registration and mapping techniques are being developed.

MR spectroscopy (MRS) provides a complementary means of studying tissue function and metabolism. In the past, spectroscopic examinations have

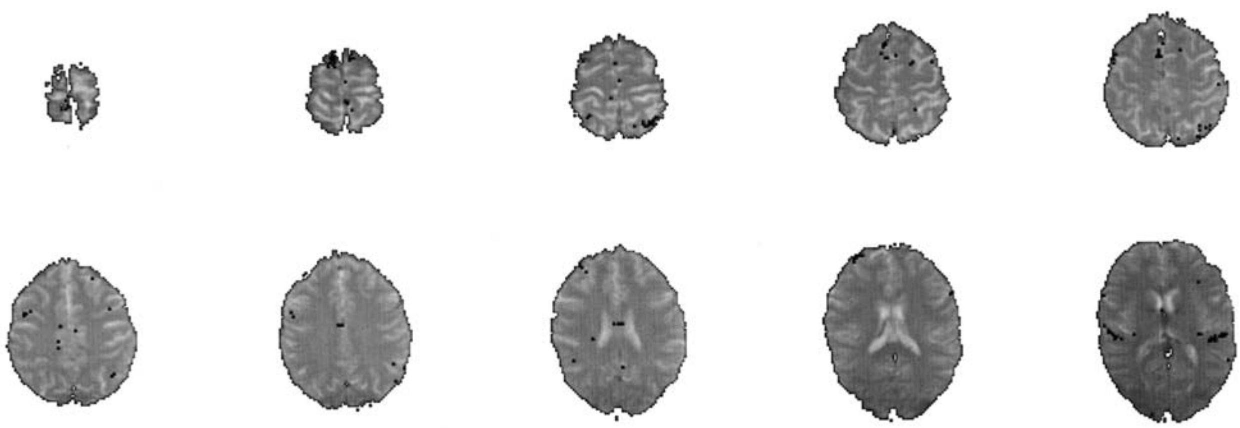


**Figure 10** A set of processed image planes through the head of a volunteer showing (black) areas of significant neural activation following exposure to a pure audio tone. Activation data were obtained at The Royal Marsden Hospital by Mr D. Collins using a real-time echo planar imaging (EPI) sequence, and processed at the Institute of Psychiatry by Dr J. Suckling.

often been distinct from imaging studies, but the increase in imaging speed, increased automation and more robust instrumentation have allowed spectroscopy to be integrated with imaging examinations. This trend will continue, allowing specific metabolic pathways, tissue metabolism via $^{1}H$ or $^{31}P$ spectroscopy, and drug distribution studies to be integrated with measurements of perfusion, diffusion or activation.

## Interventional techniques

The development of methods of guiding interventions or operations is a growing area of MRI. Following identification of suspicious lesions by MRI, it is often necessary to sample tissue to allow cytology or histopathology. Where MRI has provided better imaging, it is desirable to perform sampling using MRI, and eventually this might occur at the diagnostic visit. The design of most clinical MR systems using a cylindrical superconducting magnet design has limited access to the patient or biopsy site, presenting difficulties in performing biopsy or fine-needle aspirates in the magnet. A number of approaches are now being developed. MR-compatible biopsy tables designed for particular organs, often using specialist coils, are being designed for use with conventional systems. The breast is one such region, where MRI is demonstrating high sensitivity for the detection of breast cancer. Magnets have also been designed to provide open access, so that they can be used in the operating theatre or for more conventional image guided sampling. These systems employ either C-configuration magnets at about 0.2 T or a dual-doughnut superconducting design at 0.5 T, allowing access between the two superconducting rings. A particular objective of this latter design has been to enable interactive image guidance during neurosurgery. These approaches are requiring the development of a wide range of MR-compatible accessories, together with rapid imaging techniques and display technology.

Minimally invasive therapeutic approaches are also being piloted with MRI guidance and monitoring. These methods include high-intensity focused ultrasound, laser, electric current, RF hyperthermia and cryoablation. Areas of interest include breast, prostate and liver cancer. In principle, MR provides a valuable means of directly measuring temperature distributions in monitoring these treatments, although current techniques require a field strength of 1.5 T to provide adequate signal-to-noise ratio. An extension of these approaches is monitoring of intravascular or intra-gastrointenstinal tract using small surface coils providing high-resolution images local to the intervention.

## Acknowledgements

I am grateful to Dr Anwar Padhani, Mrs Janet McDonald and colleagues in the Diagnostic Radiology Department for providing many of the illustrations shown. Images and data used to illustrate this article were obtained as part of the Cancer Research Campaign supported research in the Magnetic Resonance Unit of the Royal Marsden Hospital and Institute of Cancer Research.

## List of symbols

$T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $T_2^*$ = transverse relaxation including susceptibility effects; TE = echo time; TI = inversion time; TR = repetition time; $\alpha$ = flip angle.

*See also:* ***In Vivo*** **NMR, Methods; Magnetic Field Gradients in High Resolution NMR; MRI Applications, Biological; MRI Applications, Clinical Flow Studies; MRI Instrumentation; MRI of Oil/Water in Rocks; MRI Theory; MRI Using Stray Fields; NMR Microscopy; NMR Pulse Sequences; Xenon NMR Spectroscopy.**

## Further reading

Edelman RR, Hesselink JR and Zlatkin MB (1996) *Clinical Magnetic Resonance Imaging*, 2nd edn. Philadelphia: WB Saunders.

Gadian DG (1995) *NMR and Its Application to Living Systems*. Oxford: Oxford University Press.

Glover GH and Herfkens RJ (1998) Future directions in MR imaging. *Radiology* **207**: 289–295.

Grant DM and Harris RK (eds) (1996) *Encyclopaedia of Nuclear Magnetic Resonance*. Chichester: Wiley.

Higgins CB, Hricak H and Helms CA (1992) *Magnetic Resonance Imaging of the Body*, 2nd edn. New York: Raven Press.

Leach MO (1988) Spatially localised NMR. In: Webb S (ed) *The Physics of Medical Imaging*, pp 389–487. Bristol: IOP Publishing.

Morris PG (1986) *Nuclear Magnetic Resonance Imaging in Medicine and Biology*. Oxford: Clarendon Press.

# MRI Applications, Clinical Flow Studies

**Y Berthezène**, Hôpital Cardiologique, Lyon, France

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| **Applications** |

Magnetic resonance imaging (MRI) is a useful, versatile diagnostic tool that can achieve contrast among different tissues by taking advantage of differences in $T_1$ relaxation times, $T_2$ relaxation times and proton densities. In recent years there has been considerable interest in the development of MRI techniques as a noninvasive method of measuring blood flow and tissue perfusion in certain clinical conditions. MRI flow measurements have been applied particularly to the vascular system and compared with other techniques, such as ultrasound. MRI provides a noninvasive method for quickly measuring velocity and volume flow rates *in vivo* using readily available methods and equipment. Flow quantification by means of MRI does not require the use of ionizing radiation and/or contrast agents, as X-ray techniques do. Unlike ultrasound, MRI measurements are not hindered by the presence of overlying bone and air.

The two principal methods of velocity measurement use either ‘time of flight’ or ‘phase contrast’ techniques. Time-of-flight methods are well suited for determining the presence and direction of flow, and phase-based methods are well suited for quantifying blood velocity and volume flow rate.

Furthermore, MRI offers the opportunity to quantitatively assess properties of tissue, such as perfusion and blood volume. Use of such quantification potentially allows tissue to be characterized in terms of pathophysiology and to be monitored over time, during the course of therapeutic interventions.

## Magnetic resonance flow measurements and MR angiography

### Time of flight

On cine gradient echo images, blood flow is bright. The high signal intensity (bright signal) of vessels on cine gradient echo images is achieved by the entry of unsaturated protons into the image, a phenomenon called time of flight or flow-related enhancement. By displaying flowing blood as high signal intensity, cine gradient echo images generally provide a better signal-to-noise ratio within the blood pool than spin-echo images, which demonstrate the arterial lumen as a dark region of signal void. Because the signal in cine gradient-echo imaging is based on through-plane movement of protons, this technique is occasionally less sensitive to slow flow or flow within the imaging plane (in-plane flow). In cases where blood flow is slow, the vessel is tortuous or flow is primarily in-plane, there may be diminished signal or even complete signal saturation on cine gradient echo images.

However, signal loss on cine gradient echo images can be used as a diagnostic aid in special circumstances. In cases of haemodynamically significant stenosis (as in aortic coarctation or aortic stenosis), a dark, fan-shaped flow jet can be seen on cine gradient echo images. This area of intravoxel dephasing results from the turbulent flow typically seen distal to a significant vascular narrowing. Aortic insufficiency may also manifest as a flow jet on cine gradient echo images. Although the relative size of the jet has been shown to correlate with the clinical severity of the stenosis, the appearance of the jet is highly variable and can be greatly affected by a variety of factors (e.g. imaging plane, pulse sequence, echo time). The jet may be small or even absent despite the presence of a high-grade (haemodynamically significant) vascular narrowing.

### Phase contrast (PC)

One good method for quantitatively measuring blood flow assesses the change in phase of the blood signal as the blood flows through a slice oriented perpendicular to the direction of flow. This method derives velocity from the phase of the MR signal, and calculates volume flow rate by multiplying the average velocity by the vessel’s area. To determine flow velocity, cine PC imaging takes advantage of the phase shifts experienced by moving protons (within blood) as they move along a magnetic field gradient. Bipolar flow-encoded gradients are applied to measure these phase shifts. This technique requires the operator to prescribe a velocity encoding that determines the flow-encoding gradient strength and sensitivity to flow direction(s) (anterior-to-posterior, anterior-to-posterior and left-to-right), which dictate the plane(s) of the gradient application.

The vascular information from cine PC acquisitions may be displayed as simple angiographic images (similar to cine gradient-echo images) in which all flow is bright or as phase map images in which the flow

directional information is coded as bright or dark and the flow velocity data are reflected in the signal intensity (relative brightness or darkness). With phase map cine PC imaging, blood flow can be quantified (millilitres per minute).

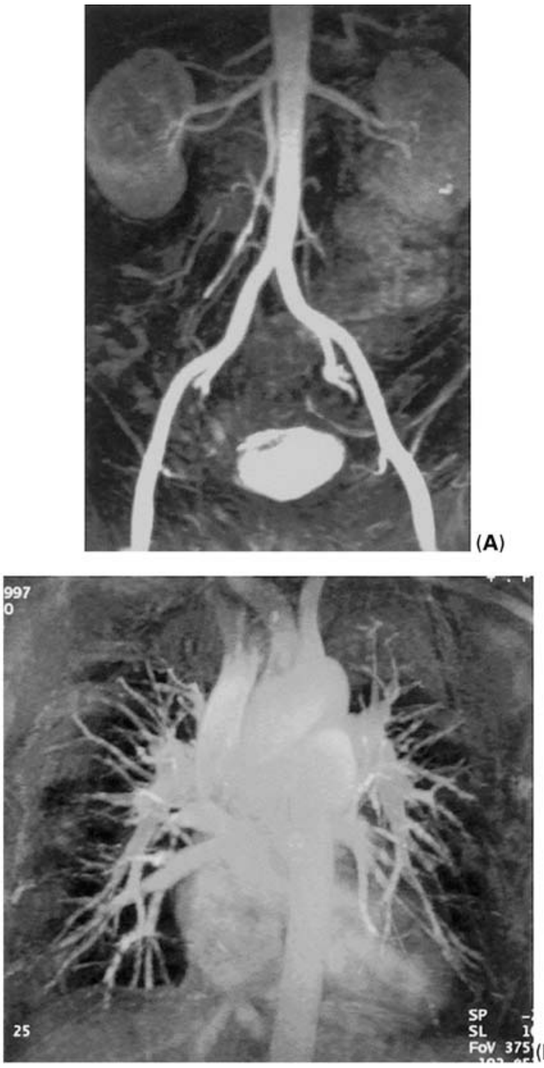
Ideally, if flow measurement is desired, one should choose an imaging plane perpendicular to the direction of the flow, select a velocity encoding at least as high as the fastest expected flow velocity, and prescribe the flow sensitivity to be in accordance with the direction of desired flow measurement. For measurement of normal flow within the ascending or descending aorta, for example, an axial cine PC prescription with a velocity encoding of 150 cm $s^{-1}$ and superior-to-inferior flow direction is appropriate. If slow flow is expected or the goal is to visualize flow in a false lumen, a lower velocity encoding such as 50 cm $s^{-1}$ may be more appropriate.

### MR angiography (MRA)

Angiography is the imaging of flowing blood in the arteries and veins of the body. In the past, angiography was only performed by introducing an X-ray opaque dye into the human body and making an X-ray image of the dye.

Many techniques have been developed for MRA of the great vessels, including gradient echo time-of-flight and phase-contrast techniques. Both time-of-flight and phase-contrast MRA methods can be implemented as either a sequential 2D or a true 3D acquisition. The encompassed MRA volume is analysed by post-processing with a maximum intensity projection (MIP) technique or with multiplanar reformatting (MPR). The MIP technique allows a rotational '3D' display of the vessel, viewed from different angles. MPR allows reconstruction of parallel thin slices in any orientation.

Three-dimensional gadolinium-enhanced MR angiography is a recently developed angiographic technique that can substantially improve the resolution, signal-to-noise ratio, speed and overall quality of vascular MRI. 3D gadolinium-enhanced MRA achieves its image contrast and hence its angiographic information from the $T_1$-shortening effect of gadolinium on blood. Because it is less dependent on inherent blood flow characteristics for the generation of vascular signal, 3D gadolinium-enhanced MRA is minimally degraded by flow-related artifacts. Three-dimensional gadolinium-enhanced MRA can be performed quickly (within a 20–40-s breath hold) on high-performance MR imagers. With a computer workstation, data from 3D gadolinium-enhanced MRA can be postprocessed to generate projection aortograms in any obliquity, that are similar to conventional angiograms (**Figure 1**).



(A) (B)

**Figure 1** 3D gadolinium-enhanced MR angiography of the abdominal aorta (A) and pulmonary vessels (B).

MRA is already used routinely in many centres for evaluation of the carotid arteries and intracerebral vasculature, aortography and assessment of the ileofemoral system. MRA of the coronary arteries is technically more difficult due to their relatively small size, their complex 3D anatomy and their constantly changing position within the thoracic cavity due to cardiac motion and respiration.

## Tissue perfusion

In the broadest sense, perfusion refers to one or more of various aspects of tissue blood flow. Parenchymal blood flow is the ratio of blood volume to the transit time of blood through the tissue. The different techniques of MR perfusion typically deal with blood

volume, transit times and blood flow as relative measures, although absolute quantification may also be possible. The two perfusion strategies are based either on induced changes in intravascular magnetic susceptibility ($T_2$* effect) or relaxivity ($T_1$ effect) and on tagging inflowing arterial spins.

### Dynamic contrast-enhanced MR imaging

Dynamic contrast-enhanced MRI is a method of physiologic imaging, based on fast or ultrafast imaging, with the possibility of following the early enhancement kinetics of a water-soluble contrast agent after intravenous bolus injection.

Bolus tracking techniques have been used to measure tissue perfusion, notably in the kidney, heart and brain. These methods are based on fundamental MR contrast mechanisms that promote either $T_1$ or $T_2$/$T_2$* enhancement. Gadolinium chelates administered in low doses lead to predominantly $T_1$-weighted signal increases, mediated by water proton–contrast agent dipolar relaxivity interactions. Alternatively, the $T_2$/$T_2$* dephasing of spins, due to a locally heterogeneous high magnetic susceptibility environment, has been exploited by using higher doses of gadolinium. If the curve of concentration versus time can be plotted as a known quantity of a tracer passes through an organ, organ perfusion can be calculated from the area under the curve. Alternatively, if the tracer is wholly extracted by the organ, the principles described by Sapirstein enable perfusion to be measured from the amount of tracer trapped by the organ.

In the normal brain, tight junctions of nonfenestrated capillaries effectively prevent Gd chelates from leaking into the interstitial space. Thus during bolus-tracking experiments, Gd-DTPA behaves like a true intravascular contrast agent as long as there is no brain abnormality that causes blood–brain barrier disruption, and regional cerebral blood volume may be determined by integrating the time-versus-concentration curve. As opposed to perfusion imaging of the brain, in other tissues such as the breast the technique is hampered by the fact there is nothing like a blood–brain barrier. Accordingly, Gd-DTPA will not be a true intravascular contrast agent. Nevertheless, by treating it mainly as an extracted tracer, it is possible to measure perfusion from the peak tissue enhancement. The model assumes a linear relation between tracer concentration and signal enhancement. Now that echo planar and ultrafast gradient-echo imaging can provide at least one image for each cardiac cycle during the passage of the tracer, measurement of myocardial perfusion with high resolution is possible.

### Arterial spin labelling

Blood flow imaging with MR by spin labelling, or spin tagging, of the water protons in the arterial source to a slice has the advantage that it is completely noninvasive, is a more direct assessment of blood flow, and may generate absolute blood flow quantification. Cerebral blood flow quantification has been accomplished by continuous adiabatic inversion of arterial spins and use of tracer kinetic models of cerebral blood flow determination. Qualitative cerebral blood flow mapping has also been described using echo planar sequences, a single inversion pulse to inflowing arterial spins, and subtraction of tagged and untagged echo planar images. In principle it is also quantifiable, to give absolute flow quantification.

## Clinical applications

### Brain

**Hyperacute stroke** Whereas conventional computed tomography and MRI are excellent modalities with which to detect and characterize central nervous system disease, they fail to depict acute ischaemia and infarction reliably at its earliest stages. Detection of cerebral infarction by dynamic MR contrast imaging is now possible. Some of the most promising work is being done with perfusion and diffusion imaging. Perfusion MRI characterizes how much brain tissue an occlusive blood clot has placed at risk (see **Figures 2** and **3**), whilst diffusion measurement shows how much tissue is already damaged or is possibly even dead.

**Flow-restrictive lesions** MR volume flow rate measurements have been used to evaluate the severity and haemodynamic significance of flow-restrictive lesions in the carotid, vertebral and intracranial vessels. A severe stenosis can result in a significant decrease in volume flow rate distal to the stenosis. Because brain perfusion relates directly to the volume of blood delivered, identifying an area of decreased volume flow rate distal to a stenosis may be of clinical importance.

Intracranial volume flow rate measurements are technically difficult using methods other than MRI, and for this reason the normal volume flow rates for intracranial vessels are not well established.

**Vascular flow reserve** Another evaluation process measures the change in volume flow rate in a given vessel before and after vascular challenge. In normal

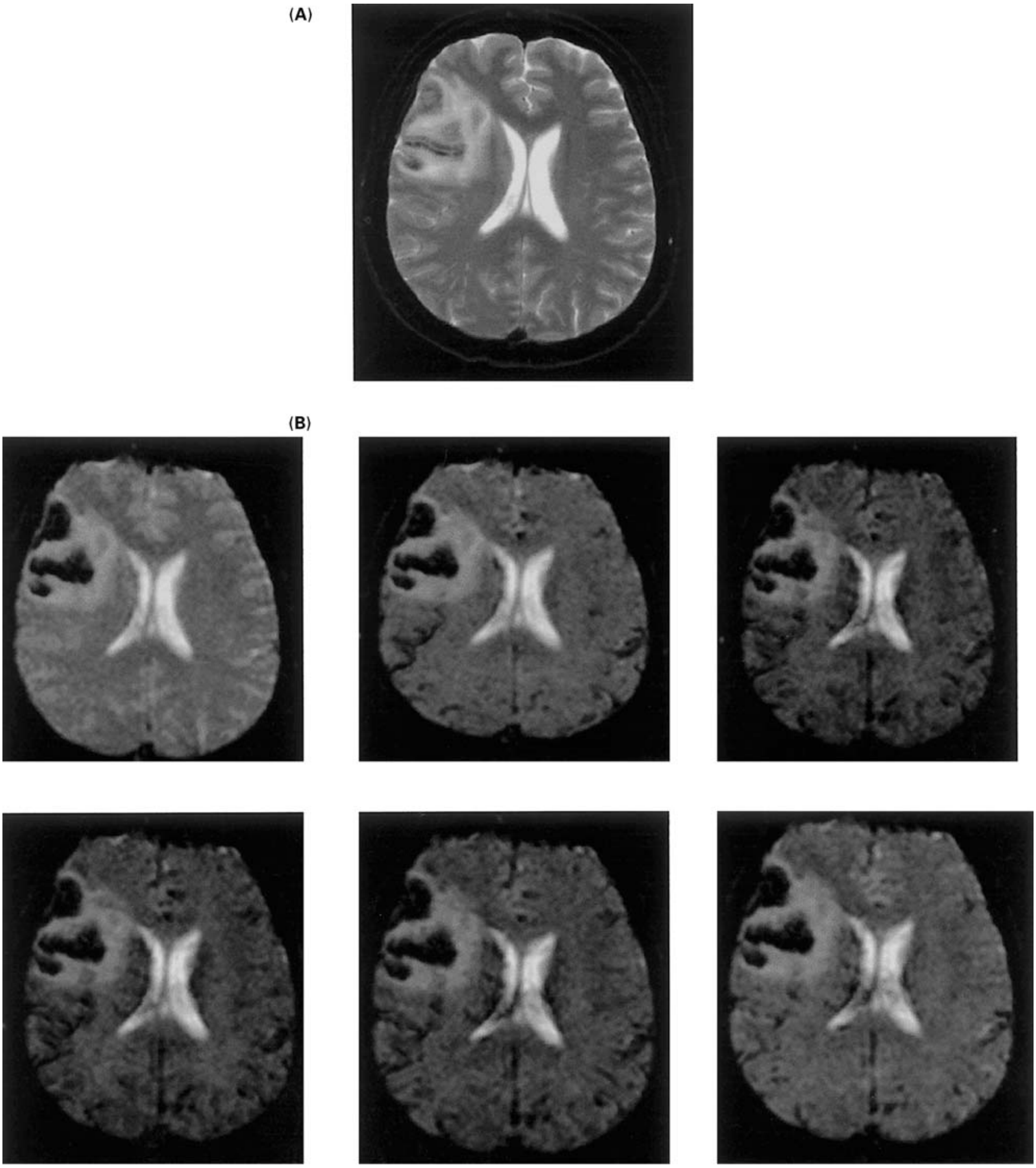



**Figure 2** Brain axial spin-echo $T_2$-weighted image (A) and sequential dynamic susceptibility-contrast in a patient with a right infarct (B).

situations, inhalation of $CO_2$ or intravenous injection of a vasodilator (acetazolamide), causes intracranial arteries to dilate, leading to an increase in flow velocity and volume flow rates in these vessels. The difference between the flow rate under routine conditions and maximal flow rate after chemically induced vasodilatation is designated as the flow reserve. In human subjects and specifically in patients with cerebrovascular disease the acetazolamide test is performed to evaluate the decrease in cerebral perfusion pressure through the investigation of the vasomotor reactivity (VMR), which is thought

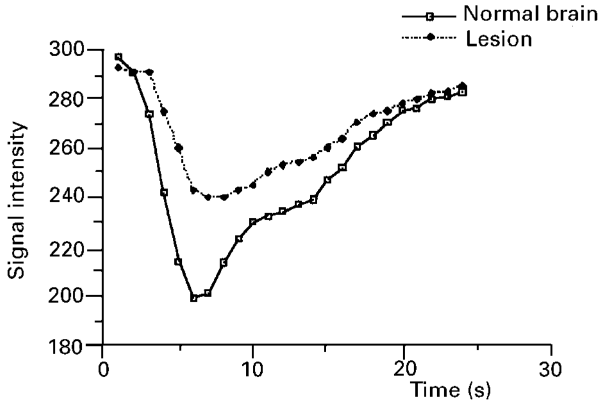



**Figure 3** Change in signal intensity during a rapid bolus contrast injection ($T_2$* effect) comparing normal brain and ischaemic regions. The lesion shows a less dynamic decrease in signal intensity than the contralateral normal region.

to reflect compensatory vasodilatation. In patients with occlusion or stenosis of more than 90% of the internal carotid artery, diminished VMR was reported to be significantly associated with low flow infarctions and higher rate of future ipsilateral stroke compared with patients with a normal or only slightly disturbed VMR. The quantification of the response of the blood vessels to the stimulus can be obtained by measuring cerebral blood flow, cerebral blood volume or blood flow velocity.

**Subclavian steal** In this syndrome due to occlusion of the subclavian artery proximal to the origin of the vertebral artery, the blood flow is reversed in the vertebral artery and redirected from the basilar artery into the arm. Phase contrast MRI can be used to determine the direction of vertebral artery flow. This information is valuable for monitoring the progression of disease, for assessing the magnitude of the steal, and in the postoperative setting, for determining the efficacy of vascular reconstructive surgery.

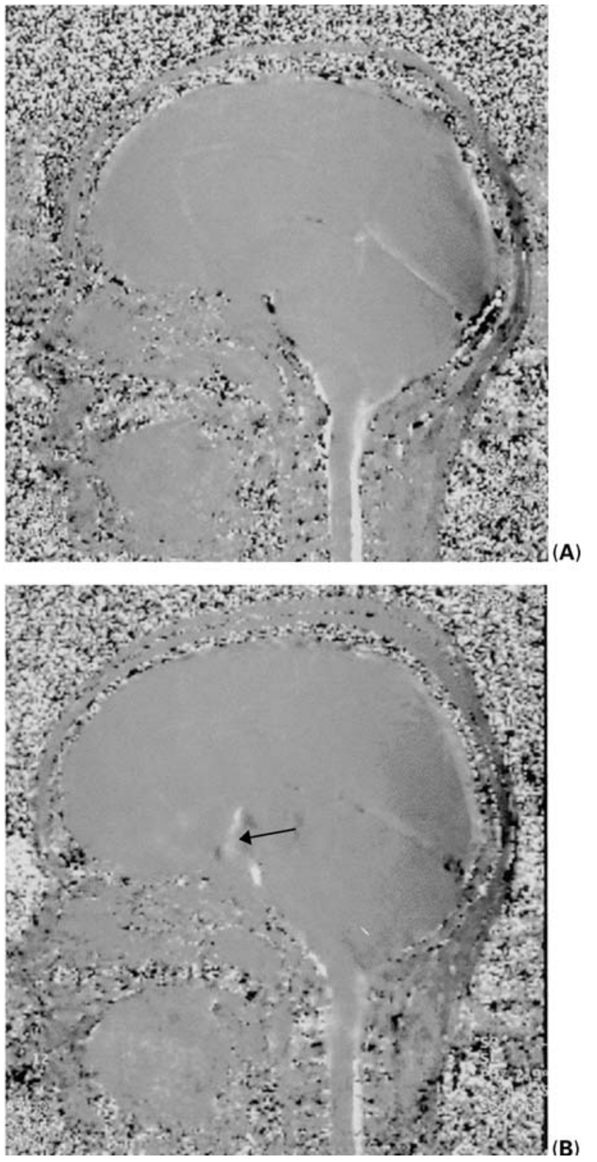



**Figure 4** Phase contrast image in a patient with a brain tumour before (A) and after (B) surgery. Before surgery no flow is seen in the third ventricle because of tumour compression. After surgery flow can be seen in the floor of the third ventricle (arrow).

**Cerebrospinal fluid flow** Phase contrast methods have been used to measure velocity and volume flow rates of cerebrospinal fluid in healthy volunteers and in patients with various diseases. This method can be used to measure the flow rate of cerebrospinal fluid through ventriculo-peritoneal shunts in patients with hydrocephalus (**Figure 4**).

## Thorax

**Valvular heart disease** The signal intensity of flowing blood during cine gradient echo imaging depends upon the nature of the flow. In general, flowing blood generates uniform high signal because of continuous replacement of magnetically saturated blood by fresh blood. Turbulence leads to loss of signal and so the turbulent jet of mitral regurgitation can be seen in the left atrium. The size of the signal void can be used as a semiquantitative measure of regurgitation but the signal void will vary with imaging parameters such as echo time. This is similar to colour flow Doppler where technical factors such as gain adjustment and filter setting are important. A more fundamental problem common to both is that

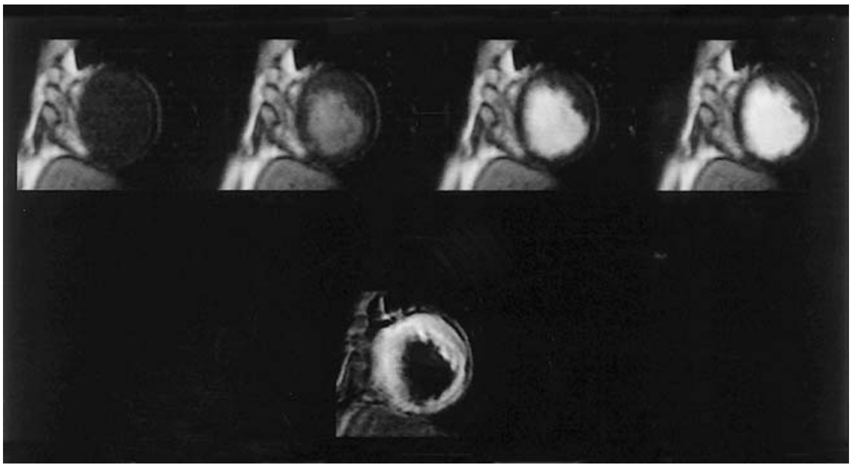


**Figure 5** Selected $T_1$-weighted images, of a single short-axis section, illustrating myocardial transit of the contrast agent in the left ventricle (top images). Myocardial perfusion is difficult to assess visually. However, postprocessing the image (factor image) demonstrates myocardial enhancement (bottom image).

the size of the regurgitant jet is influenced by many factors in addition to the severity of regurgitation, such as the shape and size of the regurgitant orifice and the size of the receiving chamber.

**Myocardial perfusion** MRI can be employed to evaluate myocardial perfusion at rest and during pharmacological testing. Ultrafast MRI sequences with image acquisition at every heart beat provide the opportunity to acquire dynamic information related to the passage of a paramagnetic contrast agent through the myocardial microcirculation and thus provides an indirect measure of myocardial perfusion (**Figure 5**). A myocardial region supplied by a severely stenosed coronary artery can be detected by a delayed increase in signal intensity and a decreased peak signal intensity. Recently, several tomographic images could be acquired during a unique bolus of a small amount of paramagnetic chelate allowing the study of almost the entire myocardial volume compared to the previous situation where only one slice was available.

### Great arteries

**Aorta** Next to congenital heart disease, the clinical utility of MRI has been most convincingly documented in patients with large vessel disease, and more specifically with acquired aortic disease. The wide field of view and the ability to freely adjust the orientation of imaging planes to the vessel direction do not only favour a clear depiction of the anatomy of the vessel lumen and vessel wall, but also facilitate the understanding of the relation to other anatomic structures within the chest and ensure highly accurate dimensional measurements.

Furthermore, it is relatively easy to combine the morphological information with functional aspects on blood flow, which can be assessed both qualitatively and quantitatively. The increased flow rate in arteries during systole, and in veins during both systole and diastole, enhances the contrast between intraluminal blood flow and vessel wall. Thus, a good image quality is usually obtained even without administration of intravenous MR contrast material.

Gradient echo techniques and phase velocity mapping are useful for demonstration and characterization of mural thrombus and for qualitative and quantitative assessment of aortic regurgitation associated with aneurysm of the ascending aorta.

There is substantial evidence demonstrating that from all the available modalities MRI has the highest sensitivity and specificity for detection of aortic dissection. MRI is not only well suited to identify an intimal flap, but can also detect aortic regurgitation and pericardial effusion with high accuracy.

The extent of aortic dissection is readily detected by NMR imaging and is displayed including involvement of other vessels. The entry and exit points are more difficult to localize, but there is no doubt that invasive investigation can be avoided with a combination of echocardiography and NMR imaging.

**Pulmonary arteries** The retrosternal position of central pulmonary arteries makes it difficult to assess pulmonary blood flow by Doppler echocardiography, especially in the presence of skeletal or lung abnormalities. NMR velocity imaging is not

technically constrained and is capable of accurate blood flow measurement in any plane. Flow can be accurately quantified in the left and right main pulmonary artery with use of phase velocity mapping.

MR velocity mapping is an accurate technique to measure volumetric pulmonary flow after repair of congenital heart disease. The consequences of pulmonary regurgitation on right and left ventricular function can be comprehensively evaluated by the combined use of MR velocity mapping and gradient-echo MRI of both ventricles. This unique information may have prognostic and therapeutic implications for the management of patients with (repaired) congenital heart disease.

The flow pattern in the main pulmonary artery differs between normals and patients with pulmonary hypertension. The latter have lower peak systolic velocity and greater retrograde flow during end systole.

Early studies have already indicated the possible role of MRI in detecting central pulmonary emboli with the use of conventional MRI techniques. MRA using fast 2D time-of-flight gradient-echo techniques combined with maximum intensity projections showed good sensitivity but only moderate specificity. Better results may be obtained with the use of phased-array coils or 3D MRA. Gadolinium-enhanced MRA of the pulmonary arteries, as compared with conventional pulmonary angiography, had high sensitivity and specificity for the diagnosis of pulmonary embolism. This new technique shows promise as a noninvasive method of diagnosing pulmonary embolism without the need for ionizing radiation or iodinated contrast material.

### Tumours

Dynamic contrast-enhanced MRI has been used as an additional imaging technique in various clinical applications, such as differentiation of benign from malignant lesions, tissue characterization by narrowing down the differential diagnosis, identification of areas of viable tumour before biopsy and detection of recurrent tumour tissue after therapy. This technique provides information on tissue vascularization, perfusion, capillary permeability and composition of the interstitial space.

Diagnosis in dynamic contrast-agent-enhanced breast MRI is primarily based on lesion contrast-agent-enhancement velocity, with breast cancers showing a faster and stronger signal intensity increase after contrast injection than benign lesions. The rapid enhancement seen in carcinomas is thought to be due to the angiogenic potential of malignant lesions.

While the dynamic technique proves very sensitive, specificity remains a problem: initial experiences with dynamic contrast-enhanced breast MRI suggest a clear-cut separation of benign and malignant lesions on the basis of their enhancement velocities. This concept has to be abandoned when more and more benign lesions have enhancement velocities comparable to or even higher than those of malignant tumours.

### Kidneys

MRI has advantages over both computed tomography and nuclear scintigraphy for assessing renal function, because it combines high spatial resolution with information on perfusion and function. Quantification of flow rate by phase contrast in the renal arteries and veins has the potential to provide estimation of renal blood flow, which could prove useful in a number of clinical situations, especially for studying renal vascular disorders and the effects of treatment, and for assessing renal transplants. Evaluation of renal perfusion with MRI has become feasible with the development of rapid data acquisition techniques, which provide adequate temporal resolution to monitor the rapid signal changes during the first passage of the contrast agents in the kidneys. More recently, magnetically labelled water protons in blood flowing into kidneys has been used to noninvasively quantify regional measurement of cortical and medullary perfusion. Dynamic MRI demonstrates renal morphology and reflects the functional status of renal vasculature. The measurement of renal perfusion by MRI could provide a noninvasive diagnostic method for monitoring the status of renal transplants and renal ischaemic lesions.

## Conclusion

With the above developments currently underway, the outlook for magnetic resonance flow measurements and contrast-enhanced MRI is bright. The opportunity to extract quantitative regional physiologic information in addition to anatomic information will definitely elevate MRI from 'anatomic imaging with soft tissue contrast' to 'a noninvasive technique for assessment of physiologic processes and tissue integrity with high spatial resolution', offering new power for diagnosis and treatment monitoring, and insights into the very mechanisms of disease physiopathology.

## List of symbols

$T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time.

*See also:* **Contrast Mechanisms in MRI; MRI Applications, Clinical; MRI Instrumentation; MRI Theory.**

## Further reading

Detre JA, Alsop DC, Vives LR, Maccotta L, Teener JW and Raps EC (1988). Noninvasive MRI evaluation of cerebral blood flow in cerebrovascular disease. *Neurology* **50**: 633–641.

Ho VB and Prince MR (1998) Thoracic MR aortography: imaging techniques and strategies. *Radiographics* **18**: 287–309.

Korosec FR and Turski PA (1997) Velocity and volume flow rate measurements using phase contrast magnetic resonance imaging. *International Journal of Neuroradiology* **3**: 293–318.

Mohiaddin RH and Longmore DB (1993) Functional aspects of cardiovascular nuclear magnetic resonance imaging. Techniques and application. *Circulation* **88**: 264–281.

Roberts TPL (1997) Physiologic measurements by contrast-enhanced MR imaging: expectations and limitations. *Journal of Magnetic Resonance Imaging* **7**: 82–90.

Sorensen AG, Tievsky AL, Ostergaard L, Weisskoff RM and Rosen BR (1997) Contrast agent in functional MR imaging. *Journal of Magnetic Resonance Imaging* **7**: 47–55.

# MRI Contrast Mechanisms

*See* **Constrast Mechanisms in MRI.**

# MRI Instrumentation

**Paul D Hockings**, **John F Hare** and **David G Reid**,
SmithKline Beecham Pharmaceuticals, Welwyn, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Methods & Instrumentation |

## Synopsis

Since 1973 when Paul Lauterbur published the first practical magnetic resonance imaging (MRI) method in Nature the one constant in this exciting area of science has been the rapid pace of change. Novel MRI methods forced the development of new technologies such as pulsed field gradients which have, again, opened the field to even more exciting pulse sequence developments. There has been a vast improvement in image quality over these years. Obviously many factors have contributed to this improvement and these will be discussed individually below. However, one factor stands pre-eminent and that is the improvement in pulsed magnetic field gradient technology. Improvements in gradient coil design have meant that gradients have become more linear and more sensitive, and the introduction of gradient shielding technology has reduced the problems of pre-emphasis and B0 correction to a thing of the past except for the most demanding methodologies. And, of course, there have been major innovations in gradient amplifier technology shortening rise times, increasing gradient strength and reducing gradient noise. Other major innovations of recent years that have significantly improved image quality have been the introduction of birdcage resonators and phased array coils. Oversampling of the receiver signal by the analogue-to-digital converter (ADC) has allowed the introduction of digital filtering techniques that prevent the folding of noise from outside the spectral width of interest back into the image. And, of course, there have been some technology improvements that have not contributed directly to improvements in image quality but have made the MRI technique easier to implement such as the introduction of self-shielded magnets and the enormous increase in computer power that has made 3D MRI techniques practical in terms of 3D Fourier transforms and image processing and display.

## Introduction

NMR spectrometers can be converted into MR scanners by the addition of gradient handling capacity and gradient amplifiers. In the crudest configuration the output of the gradient amplifiers can be fed into the room temperature shim set to create the linear magnetic field gradients necessary for imaging. Thus, modern NMR spectrometers with triple axis gradient sets can be used for micro imaging.

However, for biological and clinical MRI applications there are a number of additional hardware items that need to be considered. The basic components of the typical clinical superconducting MR scanner can be seen in **Figures 1** and **2**. The individual components are described in more detail in the text, but briefly, the magnet cryostat is kept at liquid helium temperature and houses the windings of the primary magnet and, if active shielding is used, a second set of superconducting coils outside the primary coils to reduce the fringe field effect. Inside the magnet bore, clinical scanners will usually have a passive shim assembly, active shim coils, gradient set, RF whole body coil and patient bed. Typically, the RF coils will be tuned to the proton frequency; however, the addition of RF coils tuned to other nuclei and the appropriate RF amplifiers will allow such nuclei to be imaged if the signal-to-noise ratio is sufficient.

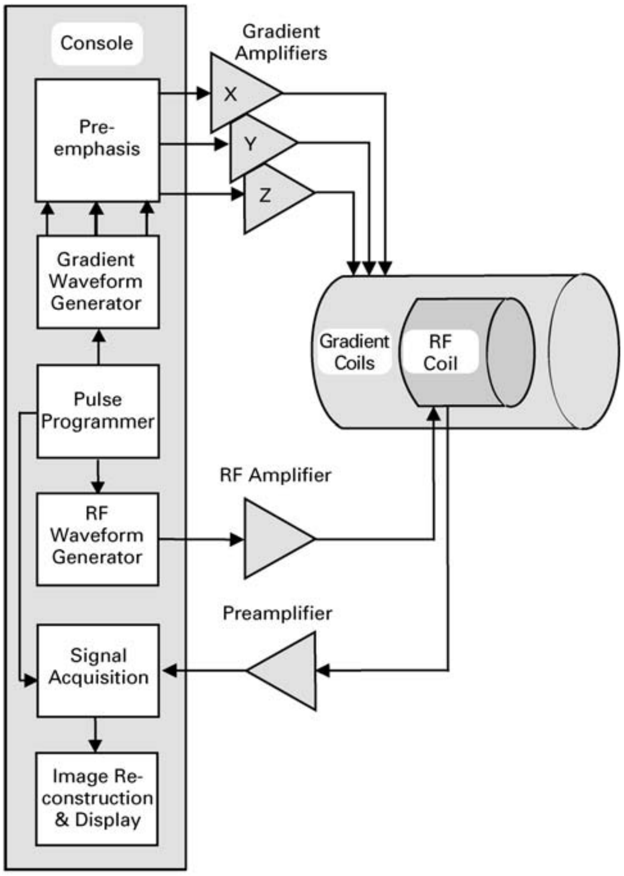



**Figure 2** The essential components of a typical clinical MRI system.

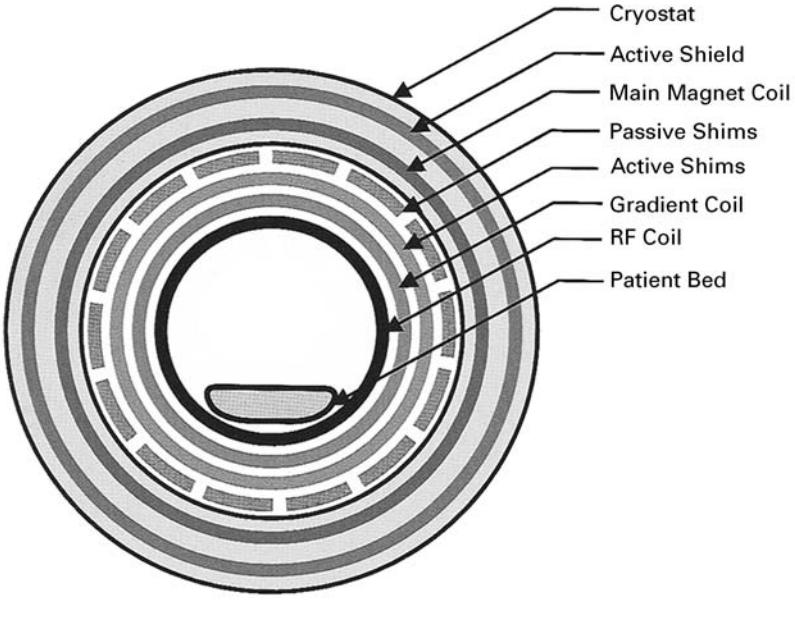



**Figure 1** Schematic cross-section through a typical superconducting clinical MR scanner. Within the cryostat (light blue) are the superconducting coils of the primary magnet (red) and active shield (green). In the bore of the magnet there are passive shim rods (grey), active shim coils (orange), gradient set (blue), whole body RF coil (black) and patient bed. The tractable diameter is generally half the magnet bore diameter. (See Colour Plate 37).

## Magnet

### Bore

First, one needs to decide the largest patient size that needs to be scanned, as this will govern the magnet bore size. There is a roughly two to one relationship between bore size and tractable patient diameter. Most clinical whole body scanners have a 1 m bore, but specialist magnets exist for imaging larger patients. Head only scanners will typically have a diameter of 60 cm. For smaller animals such as rats and rabbits a range of smaller bore magnets exists.

### Field

As with conventional high-resolution NMR, higher field strength produces more signal. However, for imaging applications this must be tempered by the consideration that differences in $T_1$ are generally greater at low field strengths and therefore images from lower field magnets intrinsically have more contrast. For micro-imaging applications where pixel size is approaching the distance water can diffuse during the application of the pulse sequence, practical experiments generally require field strengths in excess of 7 T. However, in the clinical realm superb images may be produced on systems with 0.5 T fields.

### Type

There are three types of magnets used for MRI.

**Superconducting** Higher field magnets are superconducting (**Figure 3**). The magnet coil sits in a pool of liquid helium at 4.2 K. In most animal imaging systems this is surrounded by a secondary liquid nitrogen temperature dewar (77 K) to reduce heat transfer to the liquid helium dewar (liquid helium is much more expensive than liquid nitrogen). Typically, the liquid helium would need to be topped up at intervals of 3 to 12 months. Liquid nitrogen is usually filled weekly. Modern clinical systems dispense with the secondary cryostat in favour of a helium refrigerator. These systems still need periodic refilling with liquid helium, generally at yearly intervals. The superconducting magnet offers high field strength, stability and homogeneity; however, the initial cost of the magnet can be an order of magnitude more than the electromagnets and permanent magnets and the extensive fringe field can make finding an appropriate installation site difficult.

**Electromagnets** Resistive magnets have less extensive fringe fields than superconducting magnets but require up to 60 kW to produce fields of 0.3 T and consume large quantities of cooling water (**Figure 4**). The open access design can be ideal for interventional MRI applications. Their main drawback, aside from the limited field strength available, is field instability due to fluctuations in the power supply and temperature.

**Permanent magnets** Like the electromagnet, the field strength of the current generation of permanent magnets is restricted to 0.3 T. For many applications this will be sufficient and, given the insignificant fringe magnetic field and open access design, will prove an ideal solution for some installations, particularly where the power supply and/or supply of cryogens is unreliable. However, permanent magnets require very careful temperature regulation to prevent drifts in field and they can be extremely heavy.

### Shielding

The stray fields emanating from superconducting magnets can pose a hazard to the surrounding environment. Unauthorized access within the 0.5 mT (5 gauss) field must be prevented to hinder entry of persons with cardiac pacemakers. In addition, fields as low as 0.1 mT can exert deleterious effects on colour computer monitors and analytical equipment such as scanning electron microscopes and mass spectrometers. When space is limited it may be necessary to shield the magnet to reduce its magnetic footprint. Passive shielding can be achieved by encasing either the magnet or the magnet room in ferromagnetic material. This iron shield can be both heavy and expensive. Alternatively, an active shield can be introduced by placing a second superconducting magnet outside the primary magnet and polarised in the opposite direction. The importance of this innovation to the whole body MRI market has been considerable, allowing magnets to be installed on sites throughout the world previously considered unsuitable or uneconomic and thereby contributing greatly to the overall market growth.

## Shim set

As in high-resolution NMR spectroscopy, it is not sufficient just to have a magnetic field of a certain value in the centre of the magnet. The field also needs to be homogeneous over the volume being sampled. The requirements for imaging are not nearly as stringent as for high-resolution spectroscopy but as the volumes being sampled are generally much larger the demands on the magnet design are equally exacting. Shimming is the process of optimization of the

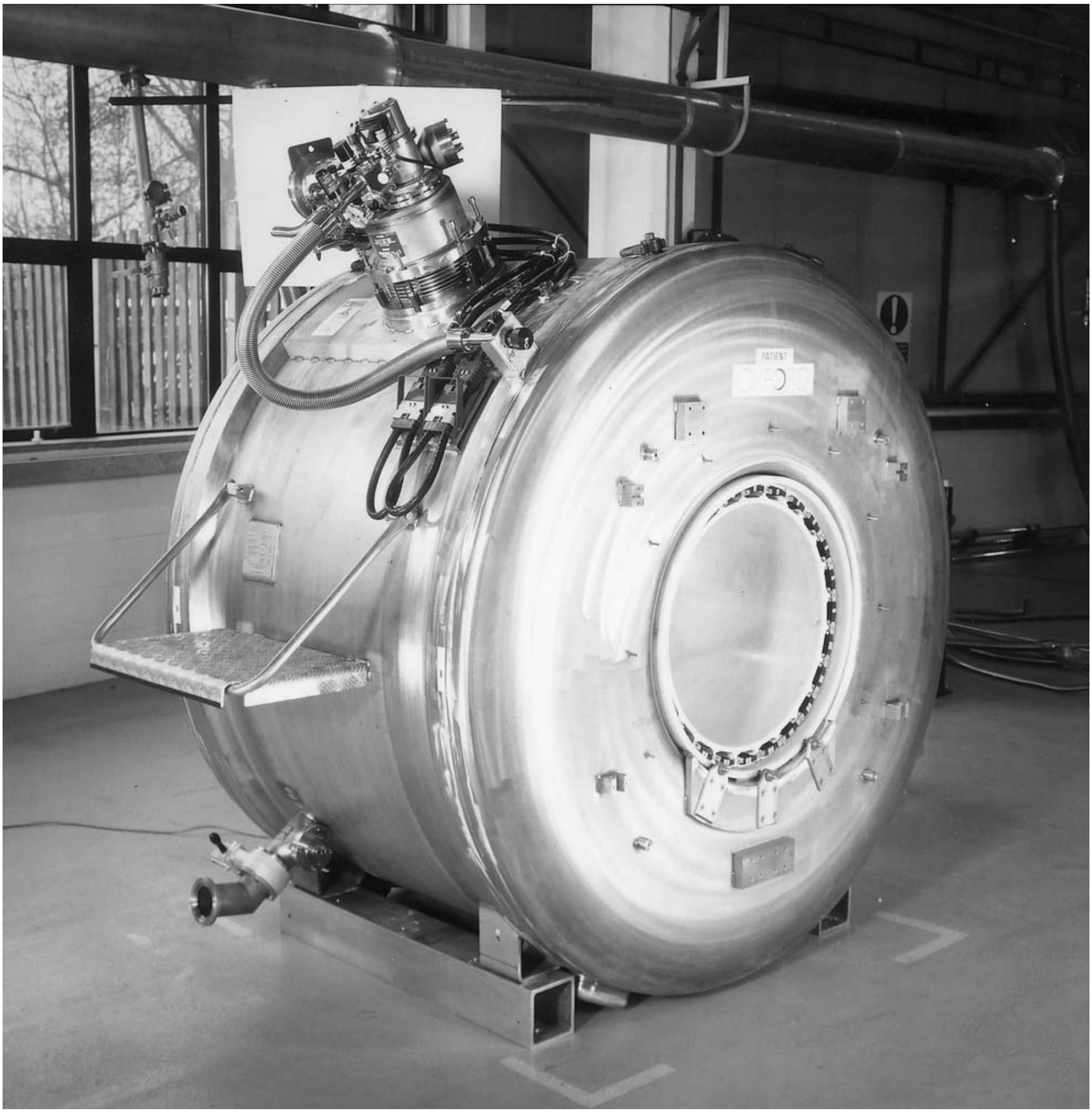


**Figure 3** Actively shielded 2.0 T whole body superconducting magnet. Reproduced by permission of Oxford Magnet Technology, Oxford, UK.

magnetic field homogeneity and is a two-stage procedure. In the first stage, the homogeneity of the primary magnet field is optimized in the absence of a sample. Magnets will either have several cryoshim coils with windings of different designs inside the cryostat or a series of iron rods placed around the room temperature bore of the magnet to balance imperfections in the field. Generally, the cryoshim currents or passive iron shims need only be adjusted on installation and can thereafter be left unless the magnetic environment changes through, for example, building work.

However, MRI subjects also introduce their own inhomogeneities into the magnetic field as tissue has a different magnetic susceptibility than air. These sample-induced field disturbances can be partially removed by the active shims. Small bore and clinical research instruments will typically include an active shim set with perhaps a dozen shim windings. Adjustment of the current in each coil to optimize

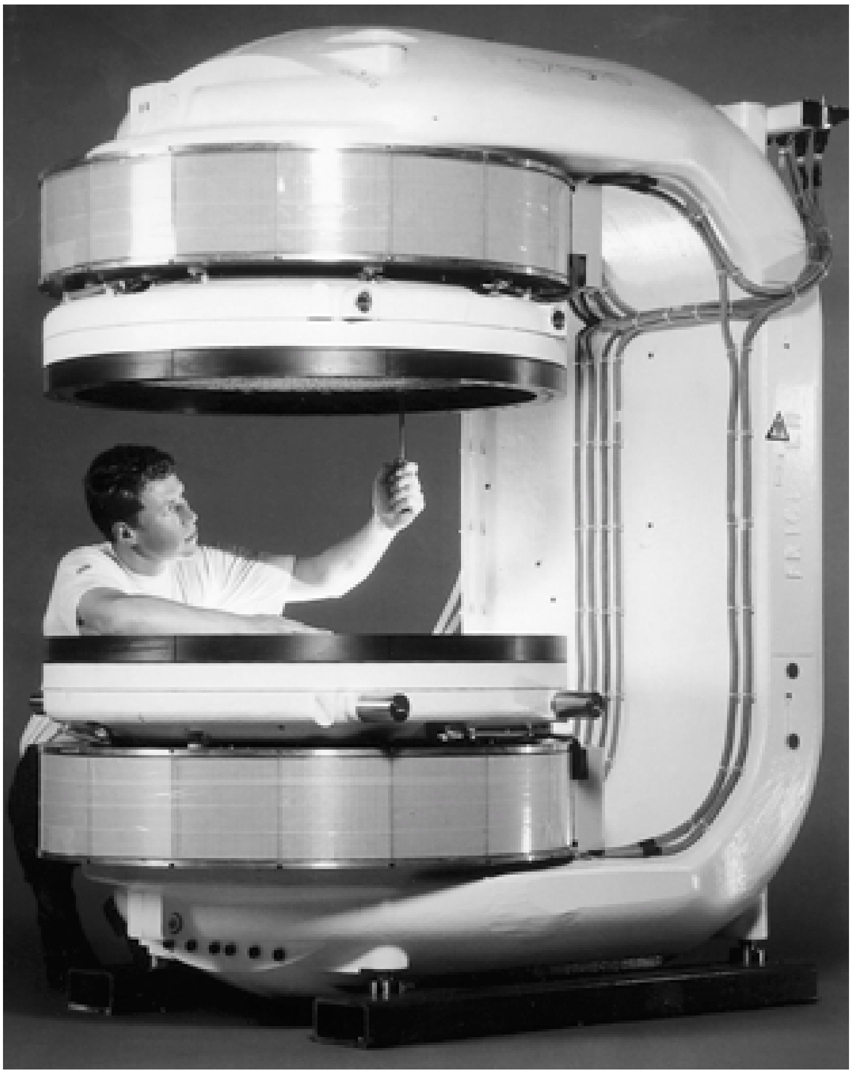


**Figure 4** Open access 0.24 T resistive electromagnet without cladding. Reproduced by permission of Oxford Magnet Technology, Oxford, UK.

the magnetic field homogeneity of the sample may be done by hand or by using a simplex minimization routine. Alternatively, one may first map the field inhomogeneities using an imaging method and then calculate the currents necessary to counteract the inhomogeneity in the sample. Many clinical scanners do not have active shim sets but rely solely on DC currents through the gradient set to shim in the *X*, *Y* and *Z* directions.

## Magnetic field gradients

Among the most critical components of an imaging system are the pulsed field gradients used to encode the images. Here, the characteristics that contribute to high quality images are the spatial linearity of the induced gradient pulses over the volume of interest and the decay characteristics of the gradient pulse. In the simplest system a linear gradient may be induced in the *Z*-axis by passing a direct current of opposite polarity through a Maxwell pair of coils wound on cylindrical formers. The greater the current the larger the linear field gradient imposed on top of the primary magnetic field.

### Gradient set

As described above it is possible to make a *Z* axis gradient set by winding a pair of circular coils onto a cylindrical former and passing a DC current through the coils such that the polarity is opposed. *X* and *Y* gradients can be formed using saddle coils. Today, most gradient sets are no longer wire coils wound onto formers but are streamline patterns

etched into copper sheet or cut into a copper plated cylinder. These have the advantage that the fabrication of complex current paths is easier and, generally, they are more compact. There are a number of conflicting parameters that must be considered when designing gradient coils. The sensitivity of the coil (in $T m^{-1} A^{-1}$), the region of acceptable linearity, the physical dimensions, the impedance and the shielding characteristics (more on this below) must all be weighed in the light of the proposed application. The coils will usually be embedded in epoxy resin to resist the torque generated when current is passed through the coils in the presence of the primary magnetic field. This torque would distort the shape of the coils and is the source of the drumming sound generated when the gradients are pulsed. Water cooling of the gradient set may be necessary for demanding applications with low field of view, thin slices and high duty cycle.

#### Amplifiers

The gradient pulse strength will be directly proportional to the current fed through the gradient coils. In modern clinical systems with echo planar imaging (EPI) capability the gradient amplifiers may need to produce 600 A. However, even for small animal systems in which the gradient amplifiers are more typically in the range of 50 A, current fed into the coil will take a finite time to reach the plateau value. That is to say that the gradient pulse will not be an ideal square function but will instead be trapezoidal. The duration of this rise time will depend on the inductance of the coil (hence low inductance coils are favoured for their short rise times) and on the voltage of the gradient amplifiers. Some systems are now provided with a 'booster' to raise the voltage and shorten the rise times. This 'booster' is basically a capacitor bank that discharges during the main amplifier switch on, increasing the voltage to drive the current through the coils. However, for fast imaging experiments such as EPI it is still important to minimize inductance in the design of the gradient coils. The other important criterion in selecting gradient amplifiers is low noise characteristics.

#### Preemphasis and active shielding

When the gradients are pulsed, residual fields called eddy currents are induced in the cryostat and other metallic structures. These fields decay with time constants typically in the order of tens of milliseconds, but for eddy currents in the cold cryostat vessel wall they may be hundreds of milliseconds long. Eddy currents can have a devastating effect upon image quality unless countered. Increasing the distance between the gradient coil and the magnet bore can reduce them, but as this will reduce the space available for the MRI subject it is often not an option. Eddy currents can be compensated for by overdriving the gradient waveform with a current that will itself counter the effect of the eddy currents. However, adjustment of this 'preemphasis' of the gradient pulse can be a tedious business as there are often eddy currents decaying with several different time constants.

Another approach to preventing eddy currents distorting the images is to shield the primary coil with a secondary coil placed outside the primary coil and connected to it in series. The secondary coil is designed to null the pulsed gradient field of the primary coil everywhere external to the coils but to have minimum effect in the centre of the coil. This approach has been almost universally adopted.

## Radiofrequency

As in high-resolution NMR, the nuclei in the MR imaging experiment, be they the water protons of the typical anatomical imaging experiment or other nuclei such as $^{19}F$, $^{31}P$ or $^{23}Na$, must first be excited. The requirements for amplitude and phase control of RF pulses are similar to those in high-resolution NMR spectrometers, though the addition of phase coherent frequency switching can be an advantage for multislice fast spin–echo experiments.

#### RF amplifiers

Clinical MR scanners used for fast imaging experiments may have up to 15 kW RF amplifiers. These high powers are necessary to reduce pulse duration in fast spin–echo imaging sequences. However, care must be taken that the amplifiers are linear otherwise the shaped pulses necessary for slice selection will be distorted and the slice profile degraded. Of course, many manufacturers are aware of this problem and compensate their pulse shapes for the known distortions induced by the RF amplifier so that the final pulse shape delivered to the RF coil is optimal. If slice profiles are inadequate it is always worth checking for non-linearity in the RF amplifiers.

#### RF probes

The alternating current generated by the RF amplifier is fed into a probe to create an alternating magnetic field at the Larmor frequency in the sample. There are a number of basic probe types each with their own advantages and disadvantages.

**Surface coils** The simplest type of RF coil is the surface coil. These usually consist of a single loop of wire and give high signal-to-noise ratio for surface structures due to the close coupling of the nuclei in the region of interest and the surface coil. They are used where high signal-to-noise ratio is of primary importance such as in localized spectroscopy experiments, functional imaging and experiments with nuclei other than the proton. The main disadvantage of the surface coil is the loss of signal intensity with distance from the coil, which results in signal intensity variation across the image and a limited field of view.

**Volume coils** Both the Alderman and Grant probe and the birdcage resonator use distributed capacitance to produce a relatively homogeneous RF field in the centre of the probe and hence uniformity of signal intensity across the image (**Figure 5**). Also, these coils lend themselves to operation in quadrature mode which brings a $\sqrt{2}$ increase in transmission efficiency and a corresponding $\sqrt{2}$ improvement in the signal-to-noise ratio upon reception. Volume and surface coil use can be combined so that the volume coil is used for transmission to produce uniform excitation across the MRI subject and the surface coil is used for reception to increase the signal-to-noise ratio. However, care must be taken that the two coils do not couple to each other, either by ensuring that their fields are orthogonal (geometric decoupling) or by employing active decoupling using additional electronic circuits.

**Figure 5** Clinical receive only volume coil. Reproduced by permission of Bruker Medical, Ettlingen, Germany.

**Phased array coils** In order to combine the signal-to-noise advantage of surface coils with the larger usable region obtained with volume coils, phased array coils can be used. These consist of an array of coils, each similar to a conventional surface coil, distributed over a surface. Each coil acts independently so that the required output signal can be obtained by combining the outputs from all or some of the elements. In order to reduce interaction between the adjacent coils, each one overlaps its immediate neighbours to minimize mutual inductance and is provided with its own preamplifier. The disadvantage of this approach is the relatively high price of the multiple amplifiers required.

## Faraday cage

The antennas used to detect NMR signals will pick up extraneous signals from the environment unless they are shielded in some way. In a high-resolution NMR instrument the bore of the magnet acts as a waveguide, effectively shielding the RF coil from the outside world. However, in imaging systems the dimensions of the magnet bore are often of the same order as the wavelength of the RF frequency of interest and then it is necessary to introduce additional shielding measures. The most common solution is to enclose the entire magnet in a continuous sheet or mesh of copper or aluminium. All services to this Faraday cage must be electrically filtered to ensure they do not act as gateways for environmental RF.

## Quality assurance

In addition to the physical hardware necessary to conduct an imaging experiment, every MR imaging lab will have a quality control process in place to identify spectrometer faults as they develop. In the clinical setting this will usually be included as part of the maintenance contract with the spectrometer manufacturer. Non-clinical labs will need to instigate their own procedures using standard phantoms. The parameters that need to be monitored are signal uniformity (RF coil homogeneity); signal-to-noise ratio; geometric linearity; spatial resolution; slice thickness; and relaxation time.

## Patient monitoring

A description of ancillary equipment for the holding and positioning of animal and human patients is

beyond the scope of this article. Similarly, monitoring equipment used for controlling animal or patient well-being such as pulse oximeters and blood pressure transducers will not be described. However, it is often necessary to monitor physiological parameters such as electrocardiogram (ECG) and/or respiration so that spectrometer acquisition can be synchronized with heart and/or abdominal motion. Many clinical systems have introduced optical transducers to convert the subject's ECG signal into optical signal for transfer via fibre optic lines to a monitoring device placed outside the magnet room. The advantage of the fibre optic line is that it cannot pick up extraneous RF and therefore does not need to be electrically filtered. Similarly, fibre optic and pneumatic devices are available for monitoring respiratory motion.

## Computing

The same computers can be used for MRI applications as for high-resolution NMR. However, MRI systems can quickly generate large datasets requiring 2D and, these days, 3D Fourier transformations and, if there are animals or patients in the magnet, the operator will not want to wait for long periods during data reconstruction. Therefore, thought should be given to installing an adequate computer workstation to operate the spectrometer console. In addition, other workstations will be needed for off-line processing of images. The demands of multi-planar reformatting of 3D data, image segmentation, surgery planning and so on, can also be quite intensive and so these additional machines also need to be high-end machines.

## Safety

### Magnetic field

The static magnetic field of any NMR instrument poses a hazard to persons with surgical implants. The large bore and horizontal geometry of most MR superconducting scanners means that the stray field can emanate for several metres and provision must be made to prevent members of the public being exposed to a potentially lethal threat. Normally, this will consist of appropriate warning signs and restricted access to areas where the field is above 0.5 mT (5 gauss). However, in addition to the well-known dangers of static magnetic fields there is a potential hazard to patients and volunteers from peripheral nerve stimulation due to switched magnetic field gradients. This occurs when strong gradients are switched on very rapidly and results in muscular twitching and, possibly, pain. Clinical systems that can achieve such fast gradient switching should have a gradient supervision unit to ensure that they meet the requirements of the appropriate regulatory agencies, e.g. the Medical Devices Agency in the UK and the Food and Drug Administration in the US. The switching of magnetic field gradients also generates acoustic noise, which is a potential risk to both patients and staff.

### RF

Pulse sequences that generate multiple 180° RF pulses can cause local tissue heating. Again the national regulatory agencies have laid down guidelines on RF power deposition in human subjects and a RF supervisor unit is necessary to ensure compliance.

### Cryogens

Superconducting magnets may contain hundreds of litres of liquid helium. In the event of either a spontaneous or emergency quench of the main magnetic field, possibly due to someone being trapped against the magnet by an uncontrolled ferrous object, the energy stored in the superconducting coils of the magnet dumps into the cryogenic liquid. The expansion factor for liquid helium is 760:1 so a large amount of cryogen gas is released into the surrounding space in a very short time. Magnet manufacturers have designed their magnets to fail safe under these conditions. However, there is still the risk of asphyxiation as an opaque fog of helium and perhaps nitrogen gas replaces the air in the magnet room. All clinical systems and large bore animal scanners should be fitted with a quench vent to allow these gases to escape safely. In addition, clinical scanners will require an oxygen detector set to alarm should the oxygen level in the magnet room fall below safe levels.

## Future trends

In the last few years MR functional imaging, in which activated regions of the brain can be visualized, and MR angiography, which visualizes flowing blood, have had a considerable impact on the specifications demanded of MR scanners. Both techniques benefit from high field strength and both rely on speed and hence gradient amplitude and switching speed. Combined with the inexorable drift to bigger and better magnets and magnetic gradient coils, there has been a move in the research MR field to follow clinical colleagues in demanding robust, easy

to use scanners. In the pharmaceutical industry and in university laboratories it is often necessary to train relatively MR illiterate scientists and technicians in the routine operation of the scanner. Automated tuning, shimming and resonance frequency adjustment makes this task easier. The introduction of actively shielded magnets to the small bore end of the market will mean these systems can be installed on crowded sites and thus greatly expand the potential market. In short, the future looks bright for continued improvement and expansion in the MR scanner market.

*See also:* **Contrast Mechanisms in MRI; Magnetic Field Gradients in High Resolution NMR; MRI Applications, Biological; MRI Applications, Clinical; MRI Theory; NMR Spectrometers; NMR Microscopy; NMR Relaxation Rates; Radio Frequency Field Gradients, Theory.**

## Further reading

Bushong SC (1996) *Magnetic Resonance Imaging: Physical and Biological Principles*. St Louis: Mosby.

Callaghan PT (1991) *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford: Clarendon.

Chen C-N and Hoult DI (1989) *Biomedical Magnetic Resonance Technology*. Bristol and Philadelphia: Institute of Physics.

Fukushima E and Roeder SBW (1981) *Experimental Pulse NMR: a Nuts and Bolts Approach*. Reading, MA: Addison-Wesley.

Gadian DG (1995) *NMR and its Applications to Living Systems*. Oxford: Oxford University Press.

Lerski RA and de Certaines JD (1993) Performance assessment and quality control in MRI by Eurospin test objects and protocols. *Magnetic Resonance Imaging* **11**: 817-833.

Shellock FG and Kanal E (1996) *Magnetic Resonance Bioeffects, Safety, and Patient Management*. Philadelphia: Lippincott-Raven.

# MRI of Oil/Water in Rocks

**Geneviève Guillot**, CNRS, Orsay, France

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

In recent years a large amount of basic and applied work on the application of NMR and MRI to the study of fluid distributions inside porous materials has appeared. With NMR one selectively observes one type of nucleus, by choosing the corresponding resonance frequency $\omega$ at a given static magnetic field intensity $B_0$ through the Larmor relationship

$$\omega = \gamma B_0 \qquad [1]$$

where $\gamma$ is the gyromagnetic ratio of the examined nucleus. The proton, which is abundantly available in both water and oil, is the nucleus most frequently observed. This means that in contrast to other noninvasive visualization techniques NMR directly probes the fluid (liquid or gas) phases within opaque porous matrices. At the same time, the unique feature of NMR is that the signal is sensitive to the physicochemical environment of the fluid. Thus, characterization of the porous material itself is also possible.

Apart from the NMR signal intensity, which is proportional to the transverse magnetization, the main quantities of interest are the relaxation times, $T_1$ (longitudinal) and $T_2$ (transverse). It is through the modification of the relaxation properties of the fluid inside the solid phase that one can obtain physicochemical information on the porous matrix such as pore size, permeability or surface chemistry. Moreover, diffusion and flow, or more exactly fluid particle displacements, can be measured and visualized by NMR techniques using pulsed gradient techniques. The susceptibility contrast between the fluid and the solid phases, however, is usually very strong in rocks, and consequently a significant line-broadening is observed. Thus, spin-echo methods must be used, and in some cases more specialized solid-state methods are necessary.

From these principles, new instruments for the characterization of oil wells by NMR have been designed and are now routinely used to obtain rock porosity, water, oil and gas saturations, and other quantities of interest to the oil engineer. Applications have also appeared in other fields such as civil engineering (water in cement, bricks or clays), polymer engineering (solvent in solid polymers or polymer–polymer mixtures) or fluid mechanics. Specific methods or hardware are being developed for nonmedical

applications of MRI and they may in turn find their way back into the medical field in the near future.

## Relaxation properties of fluids in rocks

### Surface effects

One usually observes faster relaxation rates for fluids inside a solid porous structure than for bulk fluids. This can be described as a surface relaxation effect, two or three fluid molecular layers having a specific relaxation rate much shorter than the bulk value. The origin of this shorter relaxation rate, for most mineral materials like rocks, is the presence of paramagnetic centres (usually iron). It is also considered that a reduction of molecular mobility or orientation could play a role. Whatever the origin of the surface relaxation, it can be shown that, under conditions of fast exchange, the relaxation rate measured for the fluid inside the pore space is proportional to the surface to volume ratio *S/V*, i.e. is inversely proportional to a characteristic pore size. The proportionality constant is the surface relaxation strength $\rho$ characteristic of the solid–liquid pair under consideration, its order of magnitude for water in sandstone is $8 \times 10^{-4}$ cm s$^{-1}$.

However, in many materials, pore sizes range over several orders of magnitude (from nanometres to hundreds of micrometres), and the experimental relaxation curves present a strong deviation from a monoexponential decay. A first approach is to use a stretched-exponential law to describe the relaxation curve, which has the advantage that a single relaxation parameter is obtained. Another approach is to calculate a relaxation time distribution from the relaxation curve by Laplace inversion; this is a mathematical task that presents some difficulties (the solution is not unique), but the inclusion of a regularization term, which is equivalent to favouring artificially smooth distributions, allows one to obtain reproducible results. With the assumption of fast exchange in each pore and slow exchange between different pores, the relaxation time distribution then gives the pore size distribution directly; this relationship is theoretically valid under the condition of uniformly distributed surface relaxation properties. The value of $\rho$ must be obtained independently, usually by the use of mercury porosimetry.

### Susceptibility contrast

The surface mechanism affects both $T_1$ and $T_2$. Another microscopic mechanism influences the apparent transverse relaxation and has important consequences in the methodology. This is the susceptibility difference between the fluid and the solid matrix. This difference creates an inhomogeneous magnetic field inside the fluid, and thus a broader line width. The main consequence is that it is almost always necessary to use spin-echo methods to observe fluids in a porous matrix.

Moreover, fluid molecular diffusion inside the field inhomogeneities causes decay of the transverse magnetization. The phenomenon cannot easily be described analytically, owing to the random character of molecular diffusion and the geometric complexity of porous media. Multi-echo sequences, such as CPMG (Carr Purcell Meiboon Gill), are employed to obtain the transverse relaxation curve, and it is considered that at interpulse spacing short enough (below 1 ms) and at low enough magnetic fields (below 0.2 T) the influence of field inhomogeneities is eliminated for many rock applications. One then recovers in liquids $T_2$ of the same order as $T_1$ within a factor 2 or so, that is to say equivalent physicochemical information. However, diffusion in gases being faster than in liquids, the apparent $T_2$ for gases in rocks can be shorter than for liquids. Because $T_2$ measurement times by CPMG are orders of magnitude faster than acquisition times for robust $T_1$ determinations, this method has become the standard protocol in the new logging instruments.

## Laboratory applications

### Methods

**Standard imaging sequence** The standard imaging sequence is the two-dimensional Fourier transform (2D FT) spin echo sequence, as described in **Figure 1**. It consists of a spin echo in coincidence with a gradient echo; the frequency encoding or read gradient pulses $G_{\mathrm{read}}$ and the phase encoding gradient pulses $G_{\mathrm{cod}}$ encode two orthogonal spatial directions; slice selection is obtained by the application of the gradient $G_{\mathrm{sel}}$ along the third orthogonal direction. The resolution within the image plane, or the voxel size $\delta r$, is fixed by the maximum applied gradient intensity $G$, and by the time duration of the gradient pulse $T$, through

$$\delta r = \frac{2\pi}{k} \qquad \text{with } k = \gamma GT \qquad [2]$$

The wave vector $k$ represents the maximum length explored in the reciprocal space of the image. With gradient intensities $G$ in the 20–100 mT m$^{-1}$ range, $k$ can be of order $10^4$–$10^5$ m$^{-1}$, or equivalently $\delta r$ can

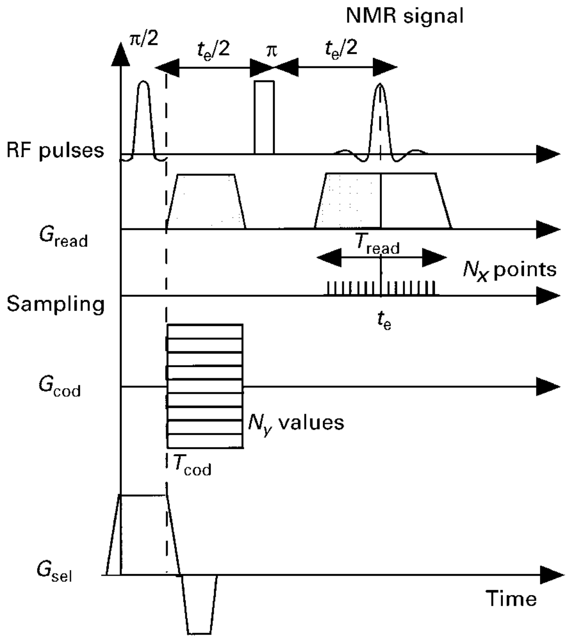



**Figure 1** The two-dimensional Fourier-transform spin echo NMR imaging sequence. A π/2 radiofrequency pulse flips the magnetization into the transverse plane, where it is refocused into a spin echo at the echo time $t_e$ by a π refocusing pulse applied $t_e/2$ after the first pulse. Spatial encoding is obtained (1) by using a shaped π/2 pulse, and simultaneously applying a selective gradient pulse $G_{sel}$ to define a slice within the object; (2) by applying two read gradient pulses so as to form a gradient echo in coincidence with the spin echo, and by sampling $N_{read}$ data points within the time $T_{read}$ in the presence of the gradient $G_{read}$; (3) by repeating the acquisition for $N_{cod}$ different gradient values applied during the time $T_{cod}$ with a maximum amplitude $G_{cod}$. One then computes the two-dimensional Fourier transform of the resulting $N_{read} \times N_{cod}$ data points in order to obtain a two-dimensional image. The products $G_{read} \times T_{read}$ and $G_{cod} \times T_{cod}$ are chosen to achieve the desired resolution within the image plane (see text).

be a few hundred μm. Use of longer pulse gradients has a limited efficacy for resolution improvement in the case of heterogeneous porous media with susceptibility broadening, corresponding to a short lifetime of the NMR signal. Thus, $\delta r$ is usually much larger than typical pore sizes in rocks. This also means that MRI will give images at a macroscopic scale of fluid distributions.

**More complex imaging sequences** The 2D FT sequence can be extended to three-dimensional imaging by using the phase encoding scheme instead of selection on the third axis. Chemical shift information can be extracted by adding a complementary chemical shift dimension; however, this procedure has rarely been used in practice for two main reasons: (1) four-dimensional data acquisition requires a prohibitive duration, and (2) the susceptibility effect spreads the spectra to the extent that the method is irrelevant in many practical situations. A more frequent approach is to extract 'water-only' and 'oil-only' images, with the simplification that the two chemical species are considered to give only single lines, and to use special protocols to eliminate local field inhomogeneities due to susceptibility as much as possible. This can be done only on rocks 'clean' enough and at magnetic field strengths above 1–2 T. Other relevant physicochemical information can be extracted by relaxation time imaging: the methods used are standard relaxation time measurement sequences combined with 2D FT or 3D FT imaging sequences.

**Resolution: choice of magnetic field, different methods** A usual rule in MRI is that a better resolution can be achieved at higher field intensity by an improvement of the signal-to-noise ratio. In the MRI of heterogeneous media, one must carefully examine the validity of that rule, since the spatial resolution is intrinsically limited by the line broadening due to the susceptibility contrast, which can be overcome only by increasing the gradient intensity. Since the susceptibility-induced field inhomogeneities are proportional to the magnetic field strength, the resolution achieved will be a compromise between the gradient intensity available from the instrument and the signal-to-noise ratio available. Orders of magnitude for the susceptibility internal gradients can be from 100 mT $m^{-1}$ (pores of 100 μm in a 1 T field) up to a few T $m^{-1}$ (pores of 1 μm in a 0.1 T field), comparable or much higher than the gradients available on large-scale imaging systems.

Other methods, which are usually considered as solid-state MRI methods because of their ability to obtain signals from samples with very short transverse relaxation times, are under development and offer very interesting possibilities for the exploration of small samples with very high controlled gradients. The first method uses fast oscillating gradients to obtain echoes at very short echo time; the second uses large static gradients and is called the STRAFI (Stray Field Imaging) method. The latter, which uses the very high gradients available in the stray fields of superconducting magnets (these can be as high as 10 to 100 T $m^{-1}$), probably offers the best possibility of going beyond the limit of large susceptibility gradients. However, these methods present the limitation that only objects of about 1 cm in size can be examined at the moment.

## Laboratory measurements

**Porosity, saturation** The NMR signal amplitude gives a fairly straightforward measurement of

porosity or fluid saturation when only one liquid (water or oil) is present in the porous sample, via simple calibration procedures such as reference measurements on bulk fluids in similar conditions. The relative accuracy achieved is usually of the order of 1% or better. Even so, extrapolation to zero time to eliminate relaxation weighting can be a difficult task in some iron-rich materials.

In diphasic cases (oil plus water), two simple techniques for the measurement of saturation have been suggested and used for laboratory applications. The first is to add to water a paramagnetic tracer, which effectively 'kills' the water signal by shortening its relaxation time below the observable limit; then only the oil signal is available. The second technique is to use NMR signals from other nuclei, such as deuterium ($D_2O$ replacing $H_2O$) or fluorine $^{19}F$ (a fluorinated oil replacing the normal oil); the latter nucleus presents the advantage of resonating at a frequency only 0.94 times lower than the proton frequency. Chemical shift $^1H$ imaging should be the first choice technique, of course, and examples of chemical shift-resolved images have been obtained in various laboratories, at fields of about 2 T, of sandstone or dolomite samples saturated with water and dodecane (**Figure 2**). However, as discussed above, this technique can work only at static magnetic fields high enough to produce resolved water and oil resonance spectra, and with reasonably 'clean' samples in which line broadening does not cause their overlap. In addition, as detailed below, the analysis of relaxation spectra has also proved to yield fruitful information.

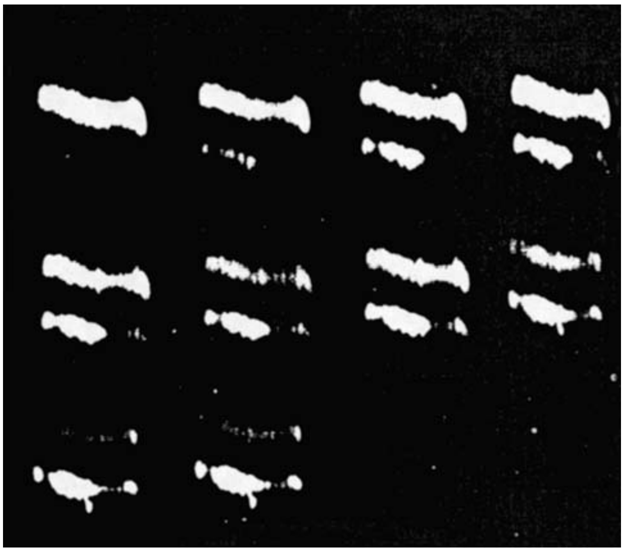


**Figure 2** Chemical-shift imaging (CSI) in laboratory MRI of oil/water in rocks: time course CSI images of oil (upper line) displacement by water (lower line) in a Baker dolomite core sample, over 30 h, obtained at a 1.89 T magnetic field strength. The absolute intensities are not normalized from one image to the other, thus the change in the ratio of oil and water intensities with time (from left to right) is the meaningful parameter. The oil signal is initially (upper left) more intense, but a uniform decrease in the oil signal and increase in the water signal with time is observed. Reproduced with permission from Majors PD, Smith JL, Kovarik FS and Fukushima E (1990) *Journal of Magnetic Resonance* **89**: 470–478.

**Information obtained from relaxation times** The different physicochemical phenomena that influence relaxation times should be taken into account with some care when examining NMR images. At the same time, they can be exploited as specific contrast mechanisms.

The general theoretical picture described above relating pore size distribution and relaxation time spectra works satisfactorily for solid materials of reasonably uniform surface chemistry, such as many model porous systems (glass bead or particle packs), and most sandstones. A number of laboratory studies have used it to deduce pore size distributions from longitudinal or transverse decay curves in saturated porous systems. The fact that NMR and mercury porosimetry, which measure respectively the accessible surface and the throat dimensions, give comparable pore size distributions can be explained by the regular geometry of these systems. One should also mention that an empirical correlation between hydraulic permeability and some representative relaxation time value have been observed to be more or less satisfactory. In other rocks with more irregular geometry, the relationship between throats and pore dimensions does not hold systematically.

Changes in fluid arrangement with saturation have been followed, for example, in drying or centrifugation experiments. The displaced water tends to occupy smaller and smaller pores as its saturation decreases, and the corresponding relaxation time spectrum is generally observed to be displaced to lower values. Light oils present lower surface relaxation strengths than water in many rocks, presumably because of the natural water-wet character of the rocks. As in the drying experiments, saturation changes in immiscible situations (water plus oil) are most apparent on the water part of the relaxation spectrum, which tends to be displaced to lower values as water is displaced out, while the oil part is generally less affected.

Surface wettability also has an influence on the surface relaxation process: hydrophobic treatments of originally water-wet surfaces, by grafting of organic chains or by coating of surfactant layers, are known to increase the water-proton relaxation times. Images weighted in wettability have thus been obtained from $T_1$-weighted images in water-

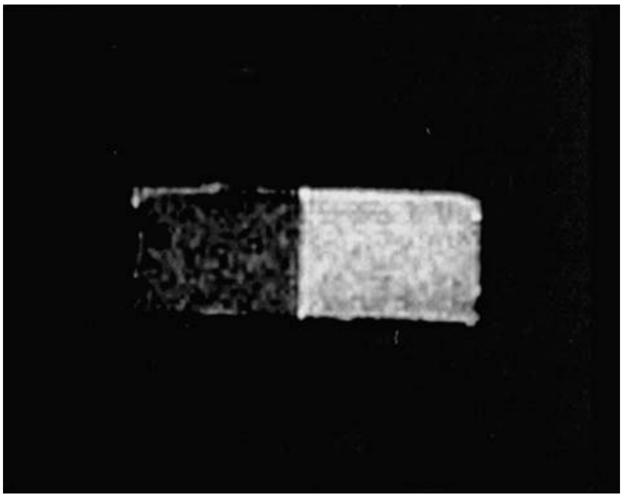


**Figure 3** Imaging of wettability contrast: images of water-saturated Fontainebleau sandstone samples of similar porosity (15%) and permeability, but with different surface treatments, obtained at a 0.1 T magnetic field strength. The right-hand sample is without treatment and naturally water-wet, and the left-hand sample was rendered oil-wet by chemical grafting of a silane chain. The image was acquired with a repetition time of 1 s, longer than the $T_1$ of the water-wet sample but shorter than the $T_1$ of the oil-wet sample; the resulting contrast due to different $T_1$-weighting by a factor 2 is much higher than the image signal-to-noise. Reproduced with permission from Guillot G, Chardaire-Rivière C, Bobroff S, Le Roux A, Roussel JC and Cuiec L (1994) *Magnetic Resonance Imaging* **12**: 365.

saturated rocks with different surface treatments (**Figure 3**). Moreover in mixed saturation (water plus oil) states the microscopic fluid arrangement depends on the surface wettability, and modifies the contribution to surface relaxation. Thus, NMR indices of wettability have been suggested from the shift of the water part of the relaxation spectrum at variable saturation. These indices are reasonably correlated to more traditional measurements of wettability properties. Fluid arrangement with respect to the solid surface has also been observed to influence the transverse relaxation for the wetting fluid via the susceptibility effect: indeed, the wetting fluid is in the vicinity of both a solid interface and the interface with the other fluid, while for the nonwetting fluid susceptibility effects play a role only on one fluid–fluid interface.

Another example of NMR relaxation weighting is the MRI study of mud filtration by rocks. Mud suspensions are used in oil-well drilling and their invasion into the surrounding rock is of importance for petroleum engineers. Water relaxation is faster in the presence of the mud particles, owing to their large surface area. Thus, the building of filtration cake has been followed quantitatively by MRI, as well as depth filtration of clay in natural rocks.

In many other potential application fields, the heterogeneous nature of the materials or their short transverse relaxation times cause similar difficulties in the collection of MRI images. Two strategies can be used. The first is to examine samples of realistic size (10–20 cm) at a moderate resolution, of the order of 1 mm, if $T_2$ values are long enough, typically longer than a few milliseconds: low-field equipment allows the collection of such images in many heterogeneous cases. When a finer resolution is necessary, other methods or specific equipment should be used.

For long enough transverse relaxation, images have been obtained by conventional liquid-state MRI sequences in different systems. From images of a solvent in a polymer matrix, quantitative measurements of solvent diffusion and possibly of matrix swelling have been performed in several systems, such as water–epoxy, water–nylon, methanol or chloroform–poly(methylmethacrylate). Elastomers are another example of samples with long enough $T_2$ and for which conventional MRI gives effective detailed information: the presence of voids in ill-cured elastomers is a spectacular source of contrast (corresponding to susceptibility defects), which can disappear with curing treatment. For building materials such as limestone and sandstone, the situation is comparable to that of oil-bearing rocks, and drying experiments have been monitored quantitatively. Similarly, the hardening of cement pastes is related quantitatively to the evolution of the water signal and of its longitudinal relaxation time (**Figure 4**).

For other samples, more solid-like or specific techniques should be used. Multipulse line-narrowing methods are well adapted to the case of solid polymers, such as adamantane, poly(methylmethacrylate) and polyacrylate. Fast gradient switching has been used to obtain one-dimensional images of water or solvent distribution in zeolite powders, with $T_2$ smaller than 1 ms. Moisture in building materials can cause spectacular damage and some groups have developed specific NMR instrumentation for moisture profile measurement at 1 mm resolution by point-to-point acquisition in bricks and mortars; these building materials are of very fine porosity and of an iron content (a few per cent) prohibitive for liquid NMR with conventional systems. But it is probably the STRAFI technique that will allow the finest resolution in solids to be achieved and that presents the highest efficiency for overcoming susceptibility broadening in heterogeneous materials.

## Flow and diffusion

**Methods** The simplest and most straightforward method for flow imaging inside porous materials at a

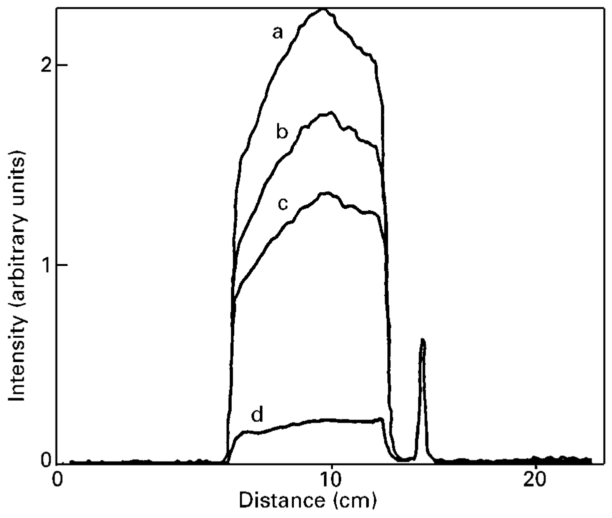



**Figure 4** Thickening of a white cement paste monitored by MRI: time evolution of 1D FT images of water in white cement obtained at 0.1 T over 4 h (curve a, 1 h; b, 2 h; c, 2.5 h; d, 4 h). The acquisition duration of each profile is a few seconds; the sharp peak on the right corresponds to a water reference sample; during cement thickening, this peak maintains the same intensity, while the signal from the cement paste decreases, corresponding to the progressive immobilization of water as solid hydrates and to the shortening of the remaining liquid water $T_2$. Reproduced with permission from Guillot G and Dupas A (1994) In: Colombet P and Grimmer AR (eds) *Applications of NMR Spectroscopy to Cement Science*, p 313. Amsterdam: Gordon and Breach.

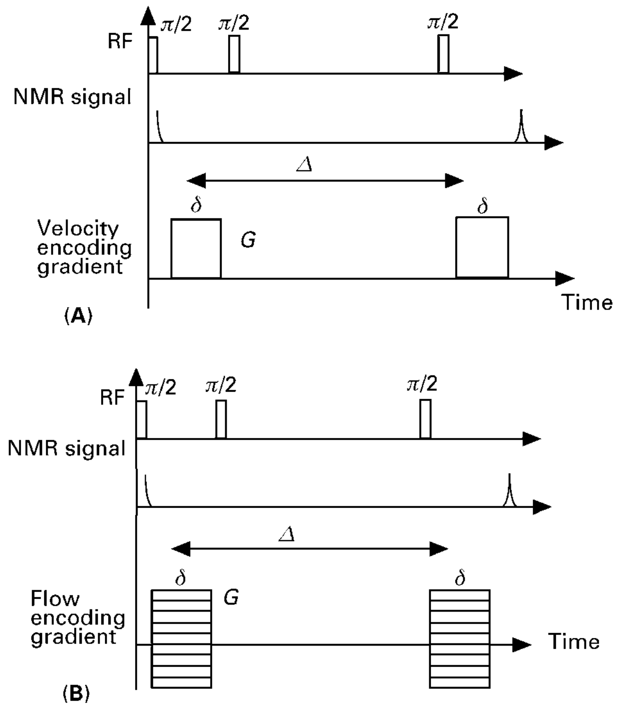



**Figure 5** Sequences using stimulated echoes for flow velocity imaging (A) and for the measurement of the displacement distribution function without imaging (B). The use of stimulated echoes often proves very convenient since the NMR signal can be observed for longer delays $\Delta$, limited only by the $T_1$ value. The pair of gradient pulses, when exactly matched, produce no phase shift for nonmoving spins, and for moving spins they induce a phase shift $\Delta\phi$ equal to the product of the wavevector $q = \gamma G\delta$, by the displacement $(r(\Delta) - r(0))$. For a uniform velocity field $v$, $(r(\Delta) - r(0)) = \Delta v$ everywhere in space and $v$ can be found from the phase-shift measurement at a given value of $q$ (A). In more complex flow situations, the full displacement distribution can be obtained from a Fourier transform analysis of data acquired with incremented $G$, and thus $q$, values (B).

macroscopic scale is to use paramagnetic solutions, which act as contrast agents just as in clinical applications of MRI. More refined techniques have been used and studies are currently in progress to study flow and diffusion. Their basis is generally the pulse field gradient-stimulated spin echo sequence. Susceptibility differences also create problems and can lead to an undervaluation of diffusion coefficients; multi-echo versions of this sequence derived from the CPMG echo train have been shown to compensate the susceptibility artefacts to a large extent.

Imaging of velocity is also possible at a macroscopic resolution. An appropriate gradient pulse pair causes a phase shift of the NMR signal. This phase shift is proportional to velocity (if all spins within each voxel move with the same velocity), via a controlled factor equal to the product of the wave vector $q$ and the time delay $\Delta$ between the gradient pulses (see **Figure 5A**). One can combine this velocity encoding gradient with the imaging gradient pulses, so as to compute a velocity image from the phase-shift image.

Another interesting and powerful approach for obtaining detailed information on the flow field inside porous materials (without imaging) is to study the displacement distribution function, which can be obtained by Fourier transformation of the NMR signal acquired for incremented values of the wave vector $q$ (**Figure 5B**). Of course it is also possible (but time-consuming) to make images of the displacement distribution. For these methods, the flow should be steady during the long data acquisition under the different gradient conditions, but this is a very realistic condition considering the low values of the Reynolds numbers normally encountered in the study of flow in porous media.

**Results** Some groups have mapped fluid velocity inside water-filled rocks, either sandstone or limestone samples, using the phase-shift method. The measured velocities have the expected order of magnitude and some reasonable correlation with rock porosity has been observed. However, one should be aware that in these studies the spatial

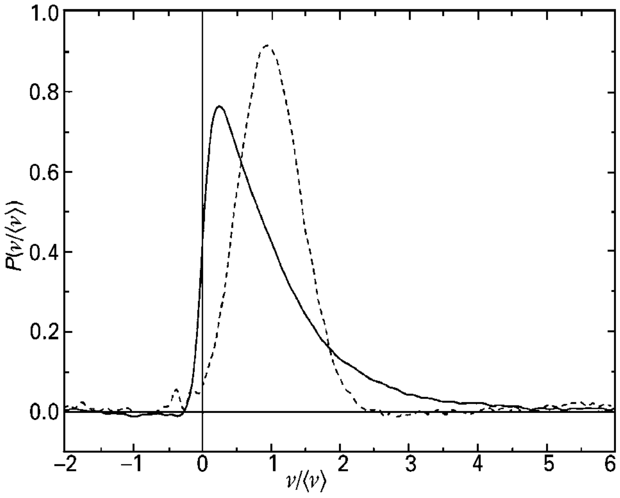



**Figure 6** Velocity probability distribution $P(v/\langle v\rangle)$ as a function of $v/\langle v\rangle$, where $\langle v\rangle$ is the average velocity, for water flowing in a glass bead pack. For short $\Delta$ (solid line), corresponding to a displacement $\Delta\langle v\rangle = 0.08d$, where $d$ is the bead diameter, the distribution can be described by an exponential decay, in agreement with the expected Stokes behaviour; here $d = 800$ μm, and $\Delta = 19$ ms. For long $\Delta$ (dashed line), corresponding to a displacement $\Delta\langle v\rangle = 7.3d$, the distribution can be described by a Gaussian law in agreement with the classical models of hydrodynamic dispersion in a porous media: here $d = 80$ μm, and $\Delta = 103$ ms. Courtesy of Lebon L, Leblond J and Hulin J-P PMMH, CNRS UMR 7636, ES-PCI, Paris, France.

resolution is usually larger than the pore sizes, i.e. larger than the scale of velocity variations, so the measured phase shift is related to some velocity value averaged in a complex way over the microscopic velocity distribution, in both space and time.

Other groups have measured, without imaging, the displacement distribution function for water in bead packs, and for water and oil in sandstone, and have examined its dependence on the time delay $\Delta$. At short $\Delta$, or at a mean displacement smaller than the pore size, the displacement distribution corresponds to the velocity distribution and has an exponential shape, in good agreement with numerical simulations of Stokes flow. At long $\Delta$, or for a mean displacement larger than a few pore sizes, the distributions have more a Gaussian shape that reflects the hydrodynamic dispersion of the fluid particles in the velocity field (**Figure 6**).

## Oil well logging

### Logging tools

A new generation of logging tools for measurement in the severe conditions encountered in oil formations has appeared in recent years. This has been made possible by the use of permanent magnets, such as samarium–cobalt alloys with Curie temperatures above 200°C. The sample of interest is the rock formation surrounding the tool, in contrast to the usual laboratory NMR situation where the sample of interest is inside the magnet and the RF probe, and different designs for the static magnetic field and for the RF probe, adapted to this specific geometry, have been developed. Working static magnetic fields of 10–100 mT can be obtained, with sensitive volumes of toroidal shape that are typically of 20–1000 $cm^3$, at a distance of a few centimetres away from the borehole wall. From their specific designs, the static magnetic field for most logging tools is in fact a field gradient of about 100 mT $m^{-1}$. Another severe constraint is that the tool must move continuously in the bore, at speeds of several cm $s^{-1}$. Under these conditions, chemical shift is not obsservable, and the only NMR pulse sequence fit for use is the CPMG echo train. Typical logs consist of one-dimensional images (along the bore axis) of the NMR signal intensity and of the relaxation time distribution extracted from the CPMG measurement, at a resolution of about 0.2–1 m; from these data, rock porosity and various information on recoverable oil can be computed. **Figure 7** shows as an example a prototype logging tool that was designed to attain a finer spatial resolution of 2 cm.

### Logging applications

The NMR signal intensity provides a measurement of the fluid-filled porosity. However, water in clay or shale has an apparent $T_2$ that is too short to be visible to the logging tools and the NMR signal comes mainly from water in larger pores and from oil. From the relaxation time distribution, an estimate of the movable fluid, called the free fluid index (FFI), is obtained by choosing a cutoff value, from *a priori* knowledge of the formation lithology. FFI is the proportion of fluid with relaxation times higher than this chosen value, or the proportion of fluid within pores larger than a given cutoff size. Laboratory measurements of centrifugeable water have shown a reasonable correlation with FFI measurements derived from NMR logs. An empirical estimate of permeability is also often calculated from FFI.

The differentiation of water from oil is based on the same general trends as presented above. The relaxation time spectrum can be separated into two parts: the water relaxation times are the shorter and change with the saturation state, whereas the light oils have the longer relaxation times, which are not strongly modified by confinement in the rock. It is also possible to detect the presence of gas, since at

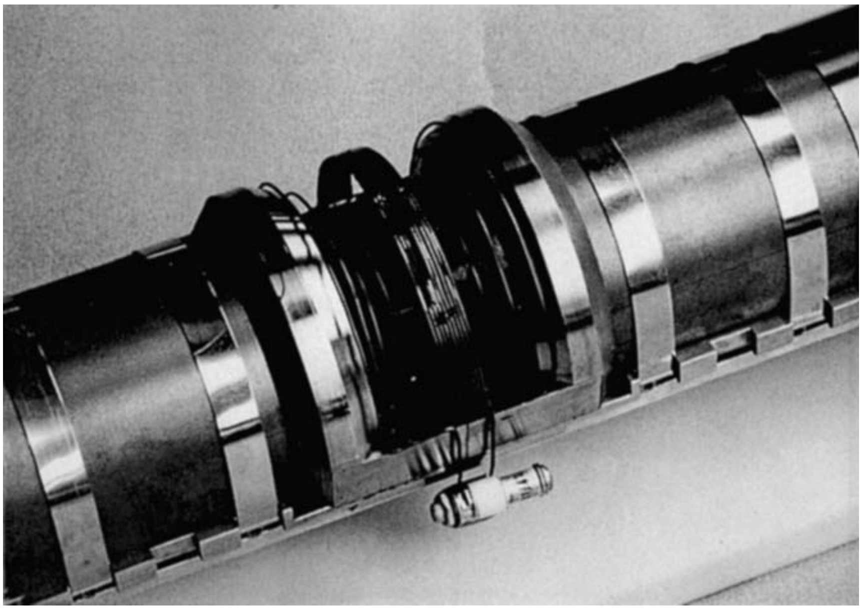


**Figure 7** Example of a NMR logging tool sensor prototype working at 4 MHz. The main magnetic field is produced by permanent magnets, plus V-shaped polar pieces to concentrate the magnet induction in the central plane, so as to define the measurement zone in the central area, with a spatial resolution along the tool axis of 2 cm. As the tool moves along the bore wall, the spins are prepolarized by the magnet induction before they arrive in the measurement zone; thus the standard logging speed can be as high as 15 cm $s^{-1}$. One can see the V-shaped main polar pieces between two cobalt–samarium permanent magnets, the RF antenna in the V space and the tuning capacitor. Courtesy of Locatelli M. LETI CEA-Technologies Avancées DSYS, Grenoble, France.

the high pressures in the reservoirs the corresponding density gives an NMR signal intensity only about 5 times lower than the signal from liquid oil or water. Gas can be distinguished from the other fluids through its specific relaxation behaviour: $T_1$ is a few seconds and $T_2$ is strongly influenced by diffusion effects in susceptibility-induced field inhomogeneities because of the higher diffusion coefficient of the gas. It has been shown that the amount of gas-filled porosity can be measured from $T_2$ acquisitions differently weighted in diffusion by changing the interpulse spacing in the CPMG sequence.

## List of symbols

$B_0$ = applied static magnetic field strength; $G$ = gradient pulse ($G_{read}$, frequency encoding; $G_{cod}$, phase encoding; $G_{sel}$, selective); $k$ = reciprocal-space wave vector; $q$ = wave vector; $T$ = pulse duration; $T_1$ = longitudinal (spin–lattice) relaxation time; $T_2$ = transverse (spin–spin relaxation time); $t_e$ = echo time; $\gamma$ = gyromagnetic ratio; $\delta$ = pulse duration; $\delta r$ = voxel size; $\Delta$ = delay between pulses; $\rho$ = surface relaxation strength (relaxation for the fluid owing to relaxation centres on the solid surface); $\omega$ = resonance frequency.

*See also:* **Contrast Mechanisms in MRI; Diffusion Studied Using NMR Spectroscopy; Geology and Mineralogy, Applications of Atomic Spectroscopy; MRI Applications, Clinical Flow Studies; MRI Instrumentation; MRI Theory; MRI Using Stray Fields; NMR Microscopy; NMR of Solids; NMR Principles; NMR Pulse Sequences; NMR Relaxation Rates; Relaxometers; Solid State NMR, Methods.**

## Further reading

Borgia GC (ed) (1991, 1994, 1996) *Proceedings of the International Meetings on Recent Advances in MR Applications to Porous Media:* Special Issues of *Magnetic Resonance Imaging* Vols **9**, **12**, **14**.

Brownstein KR and Tarr CE (1979) Importance of classical diffusion in NMR studies of water in biological cells. *Physical Review A* **19**: 2446–2453.

Callaghan PT (1993) *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford: Oxford University Press.

Edelstein WA, Vinegar HJ and Tutunjian PN (1988) NMR imaging for core analysis. *Society of Petroleum Engineers* 18272.

Kleinberg RL (1996) Well logging. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, Vol **8**, pp 4960–4969. Chichester: Wiley.

Special Issue of the *Log Analyst* on NMR Logging. November–December 1996.

Watson AT and Chang CT (1997) Characterizing porous media with NMR methods. *Progress in NMR Spectroscopy* **31**: 343–386.

# MRI of Rigid Solids

*See* **Rigid Solids Studied Using MRI.**

# MRI Theory

**Ian R Young**, Hammersmith Hospital, London, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Theory |

## Introduction

In essence the theory of nuclear magnetic resonance peculiar to magnetic resonance imaging (MRI) alone is very simple and can be simply summarized as being a specialist application of multidimensional Fourier transformation NMR, in which the various frequency axes are related to spatial ones by assuming that the gradient magnetic fields applied to encode space produce equivalent frequency variations. In practice, the situation is very much more complicated, and involves reviewing a number of different aspects of the data recovery process.

In reality, the most complex differences between small-scale high-resolution studies and those involving human subjects lie not so much in the spatial encoding process but in the interactions between the RF coils and the subject and in the desirable targets for human studies. These are not, to anything like the same extent as in high-resolution studies, dictated by the need for quantitative accuracy of the measurement of a multiplicity of chemical components but are, rather, driven by the requirement to highlight certain structures of the body with respect to others. Imaging in general, and MRI in particular, is driven predominantly by issues of contrast. Whole-body magnetic resonance spectroscopy (MRS) is, similarly, driven by rather different factors from those affecting normal work in small-bore high-field spectrometers. In many ways MRS is closer to MRI in terms of its strategies and problems and both are, in effect, considered in this article (the former by implication only).

Relative to normal spectroscopy, both MRI and MRS rely to a very much greater extent on comparisons of results from regions of tissue considered to be normal and those felt to include more or less severely diseased structures. Biological diversity ensures that the reproducibility of results from one subject to another will not be as good as in most spectroscopic studies. On the other hand, much of what is attributed to this cause is due to ill-considered research strategies and artefactual results, but there is an undeniable level of difference between individuals in animal species of all kinds so that the spread of results will always be greater than that obtained from small, passive samples. Sensate beings move in complex and more or less uncontrollable ways, have highly nonreproducible sizes and shapes and are hugely complex, so that practically all data obtained from them are contaminated by significant partial volume effects (which means that data sampled from a region of tissue contains components of multiple structures and a variety of different tissue types).

This article discusses the basic processes of spatial localization and the formation of images; it considers how the form of the signal-to-noise relationship is affected by the large, conducting load that the body represents, and how its movement affects the quality of the data obtained from studies. The issues surrounding the creation of contrast, which is the prime clinical desire of MRI strategies, are addressed in a separate article on contrast mechanisms (q.v.). This also discusses the development of a number of important artefacts, the formation of which is a derivation from the strategies outlined in this article. The reader may find it convenient to treat both as being closely related topics in *in vivo* NMR.

In all of what follows, it is assumed that the reader is familiar with conventional NMR theory as described elsewhere in this encyclopedia (see Further reading), as this article examines only the extensions needed to the theory and practice that are the basis of whole-body MR.

## Spatial localization

The basic concept of imaging is very simple. Gabillard first suggested the use of magnetic field gradients as a means of identifying positional data in NMR. However, it required the development of Fourier-transform NMR before these ideas could be usefully applied to imaging in its standard form. Mansfield and Grannell pioneered the application of gradients in FT-NMR, with the aim of achieving the analogue of optical diffraction, with resolution of lattice plane dimensions, while Lauterbur was the first to publish a two-dimensional image of an identifiable object.

In principle, assuming that $B_0$ is completely homogenous, application of a gradient $G_r$ (the component of a field varying along the $r$ axis parallel to the $Z$ axis (parallel to $B_0$)), results in a divergence of the spin resonant frequencies represented by

$$\omega_i = \gamma(B_0 + r_i G_r) \qquad [1]$$

where $r_i G_r$ is the magnitude of the gradient field at position $r_i$ along the gradient. If the object is multidimensional (as in the human body), the signal observed at frequency $\omega_i$ ($S_i$) is derived from all the spins in the plane orthogonal to the gradient axis through $r_i$.

If we have a series of planes normal to $r$ with respective uniform proton densities $P_i$ at distances $r_i$ (each generating a signal $S_i$) then, after demodulation, the signal from the whole object ($S_{\text{obj}}$) is given by

$$S_{\text{obj}} = \sum_{i=1}^{n} S_i = \sum_{i=1}^{n} A_i P_i \exp(-\mathrm{i}\gamma r_i G_r t) \qquad [2]$$

where $A_i$ measures the extent of the object in the plane at $r_i$. In the limit, this becomes

$$S_{\text{obj}} = \int_r P_r A_r \exp(-\mathrm{i}\gamma_r G_r)\mathrm{d}r \qquad [3]$$

ignoring all relaxation effects.

The above argument can be developed in two, or all three, dimensions to yield, for the latter, a relation of the form:

$$S = \int_x \int_y \int_z \rho_{ijk} \exp\left( -\frac{\mathrm{TE}}{T_{2ijk}} + \mathrm{i}\gamma\mathrm{TE} \right. \left. (G_x x_i + G_y y_j + G_z z_k) \right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \qquad [4]$$

where $\rho_{ijk}$ is the proton density of the ($ijk$) voxel, and $T_{2ijk}$ is its spin–spin relation time constant. This recognizes that there will be, at least, monoexponential transverse relaxation effects in the data, and that their magnitude will be those developed at a time TE (the echo time, the impact of which will be discussed later), which is the time at which it is assumed all spin dephasing due to the applied gradient is zero during the data acquisition.

In practice, the gradients require to be applied at different times to achieve the desired three-dimensional encoding, since the concurrent application of more than one gradient identifies a single axis, which is that determined by the vector sum of the applied fields.

## Spatial resolution

It is easier, at the beginning, to discuss how resolution is controlled by the spatial encoding process by considering a single-axis experiment (as in Equation [3]). In order to achieve a resolution of $n$ points along the direction of observation, the Nyquist criterion demands that there is a $n/2$ Hz spread of frequencies across the field of view being studied. Thus, the gradient $G_r$, the distance $r$ which spans the object, and time $t$ for which the signal is observed, are related by

$$n\pi = \gamma G_r r t \qquad (\gamma \text{ in Hz/T}) \qquad [5]$$

During the period $t$ at least $n$ data samples are required to identify the individual frequencies.

Conventionally, all imaging procedures acquire data in the presence of a gradient, depending on Equation [5]. In spatially localized spectroscopy, data are frequently measured in the absence of any gradient field and this approach could also, in theory, be useful in microscopy experiments, though in this case to minimize the impact of diffusion.

Essentially, there are only two fundamental imaging strategies, though with a growing number of variants of each. In one, the spins are excited and data are recovered along a direction determined by the vector sum of the applied gradients until enough has been collected, after which the magnetization is allowed to recover before it is excited again and data are acquired with a different gradient vector direction. The other strategy exploits the property that, after excitation spins retain the phase relationships into which they have been placed by the application of a short pulsed gradient field. In a good field, the relationships are held sufficiently well as the system

relaxes for another gradient to be applied for long enough, in what is usually an orthogonal direction, while data are recovered. Subsequently, after time for the magnetization to recover and be excited again, another, different, gradient pulse is applied and more data are obtained. The process is then repeated sufficiently often for a complete set of the information needed for a two-dimensional image to be obtained.

The former technique was that known as filtered back-projection, which, as originally applied, was a direct MR analogue of the original translate–rotate CT–X-ray scanner developed by Hounsfield. The first version of the other strategy, involving Fourier transformation in two or more directions, was proposed by Ernst and his colleagues in a form that ultimately proved much less useful than the spin warp method that has become the basis of the vast majority of clinical MRI.

The aim of every image recovery procedure is to obtain enough good information to fill all the locations of '*k*-space' (i.e. spatial frequency space) so that when the data processing has been completed there are no artefacts in the image due to missing or corrupt data.

**Figure 1** shows the sequence form (A,B) and coverage of *k*-space (C) developed during back-projection imaging. The various acquisitions form the spokes of a wheel and are produced in a plane (say the *X*/*Y* plane used as an example here) by the vector sums of the two gradients $x(=R\cos\theta)$ and $y(=R\sin\theta)$. Sufficient acquisitions are made to cover *k*-space at the density required. Suppose that the target of the image acquisition is an $n \times n$ matrix. The number of acquisitions needed to cover *k*-space is then $\pi n/2$. As will be discussed below, the efficiency of data acquisitions that fill *k*-space along orthogonal coordinates is better (as these require *n* acquisitions only) and this is one factor contributing to the unpopularity of back-projection acquisition.

It will be noted from **Figure 1** that if data were to be acquired as shown in **Figure 1A**, then regular sampling would result in oversampling at the start of the acquisition (as the spin frequency spectrum is spreading relatively slowly), though it may be just adequate later (after the gradients have flattened out and the full set of spin frequencies has developed). This form of acquisition was used in the early days of MRI by sampling the data nonlinearly (to account both for delays in amplifier response to command signals and for eddy currents arising from the changing gradients). Conjugate symmetry was also used to permit a significant reduction of the number of acquisitions as they only had to sample data though 180°.

Back-projection suffers from major problems in poor fields owing to those fields resulting, in effect, in the angular misplacement of points in *k*-space (which reconstruct to give streaks of intensity variation leaving the edges of structures). The centre of *k*-space is oversampled (and has an appropriately improved signal-to-noise ratio) as sampling can be continued as long as there is useful signal, though this results in angular undersampling at great distances from the centre. Nevertheless, the approach is not without merit. It is relatively impervious to the effects of motion (for reasons too complex to discuss in detail here) and it permits very high-resolution imaging of a local region in the body without the accompanying problem of aliasing that affects the spin warp method.

The spin warp technique and its development have, however, been the methods upon which practically all subsequent workers have based their work. The concept is shown in **Figure 2**. The block diagram of the data acquisition process is given in **Figure 2A** and the *k*-space strategy is indicated in **Figure 2B**. The data acquisition gradient is the same for each data recovery but the dotted lines indicate the varying nature of the phase-encoding pulses. The initial inverted gradient ($G_t$) (of the data acquisition) has the same value of

$$\int_0^{T_\mathrm{p}} G_t \,\mathrm{d}t$$

(where $T_\mathrm{p}$ is the width of the pulse) as the first half of the longer lower amplitude acquisition gradient ($G_a$) during the presence of which data are recovered.

Thus

$$\int_0^{T_\mathrm{ip}} G_\mathrm{it}\mathrm{d}t = \int_0^{T_\mathrm{ap}} G_{at}\mathrm{d}t \qquad [6]$$

where $T_\mathrm{ip}$ is the duration of the inverted gradient pulse and $T_\mathrm{ap}$ is half the duration of the acquisition gradient pulse.

During the warp gradient, spins are dephased, but they are then refocused at the centre of the data acquisition so that data sampling (which can be linear, as the gradient is constant throughout the data acquisition period) occurs through the echo peak. Thus, *k*-space is fully sampled for each line, and conjugate symmetry (as is needed for the technique in **Figure 1A**) is not necessary. The method is generally more robust than that in **Figure 1**, and in practice back-projection is also generally now

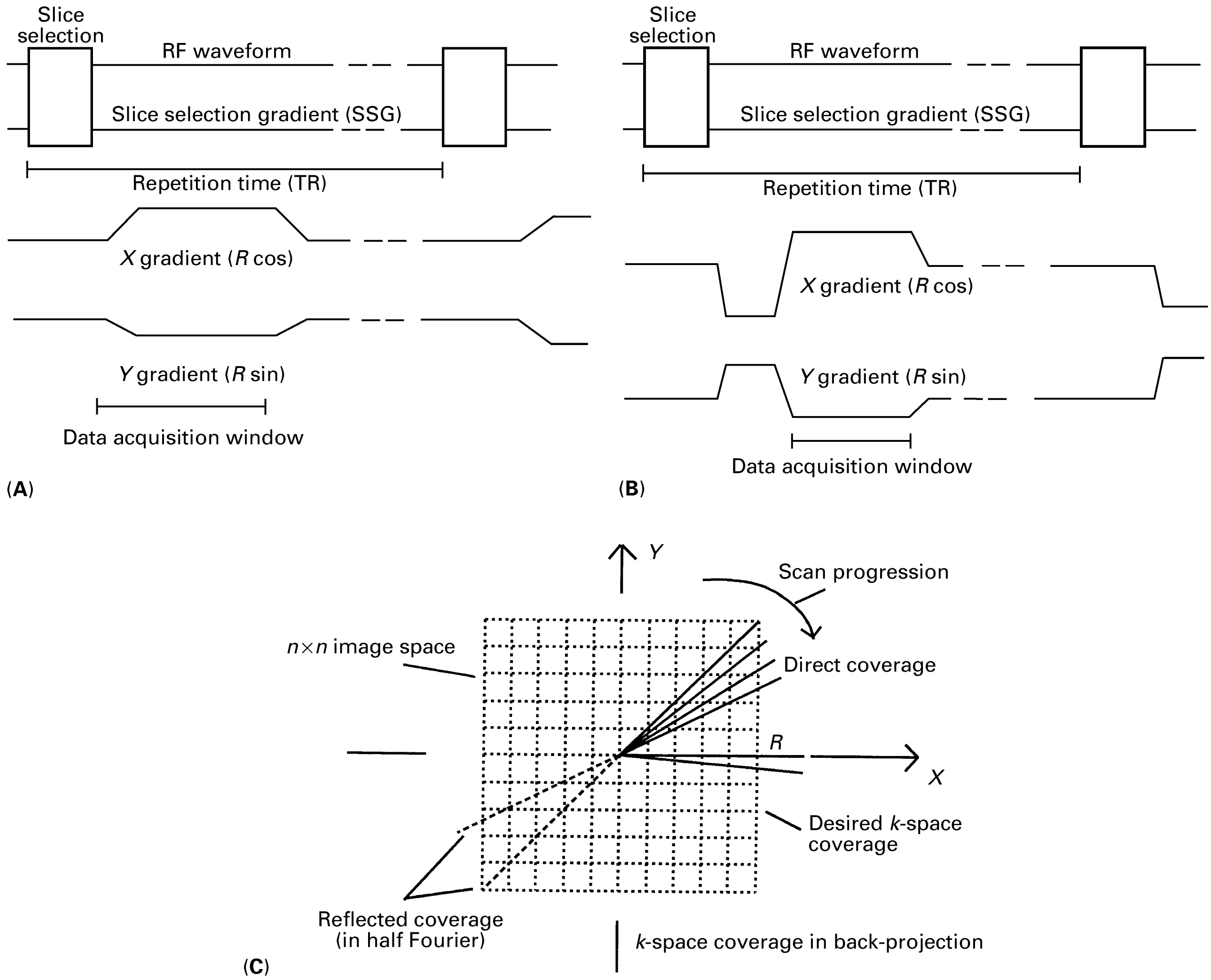



**Figure 1** Back-projection imaging. (A, B) Sequence form used for acquiring a back-projection data set for an image in the *X*-*Y* plane. At this time the slice selection procedure is simply shown as a block. It will be described later in the text. The *X* and *Y* gradients are constrained to define a series of vectors (given by $\theta = \tan^{-1}(Y/X)$ of constant magnitude $R$ $(= (x^2 + y^2)^{1/2})$: (A) shows the variant in the sequence used when conjugate symmetry is to be applied; (B) is the form of gradients used when an echo is to be formed during the flat regions of the gradients. (C) Coverage of *k*-space with back-projection. Workers reconstruct the data either using genuine back-projection algorithms (as in CT–X-ray) or by interpolating onto a two-dimensional grid, followed by a two-dimensional Fourier transformation.

implemented with echo formation (as in **Figure 2B**) except where it is desirable to acquire data as fast as is possible after excitation. The method in **Figure 2A** is a useful technique in the imaging of very short-$T_2$ proton moieties or nuclei such as sodium that also have short $T_2$ values.

In order to cover *k*-space completely, all its lines must be filled. Each line is obtained using a 'phase encoding' pulse (see **Figure 2A**), during which the spins precess at the frequencies dictated by the applied gradient. At the end of the gradient pulse, different groups of spins (isochromats) will have different phase relationships, which they retain in a perfect field along parallel lines even in the presence of another gradient applied orthogonal to the first. When enough phase-encoding pulses have been applied, each followed by a gradient in an orthogonal direction, the set of data generated will be given (ignoring relaxation and recovery effects) by

$$S_{x,y} = k\int_{0}^{m}\int_{-(n-1)}^{n} S_{p,t}\exp\left(\mathrm{i}\gamma(xG_x t + p\,\Delta\,G_y y\,T_\mathrm{p})\mathrm{d}p\,\mathrm{d}t\right) \qquad [7]$$

where $S_{xy}$ is the image dataset, assuming, in this case, that the readout direction is $x$, and the phase encode direction $y$. $\Delta G_y$ is the increment in the $y$ gradient between samples; $T_\mathrm{p}$ is the phase encode pulse

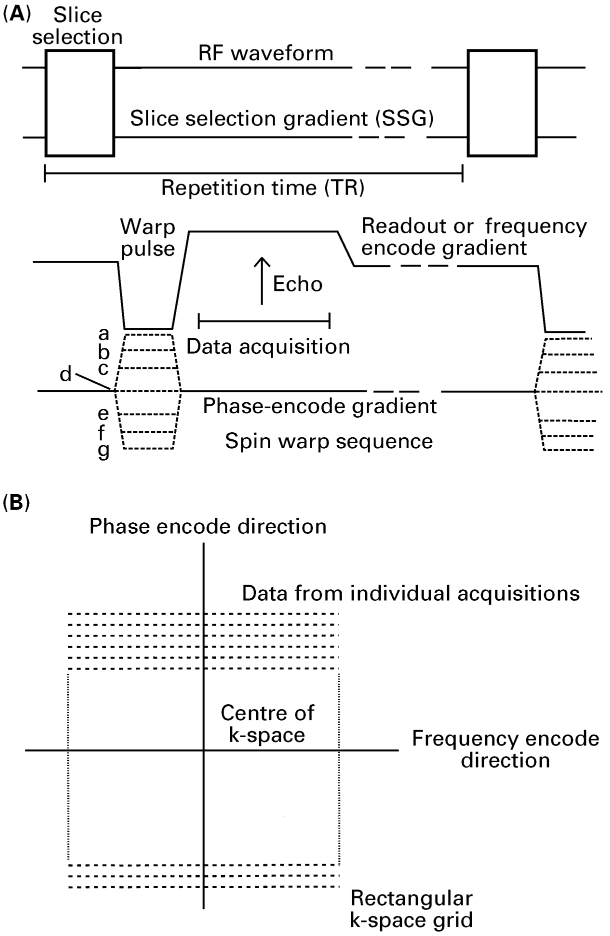



**Figure 2** Spin warp imaging. (A) Sequence structure used in spin-warp imaging (again the slice selection component is shown as a block). The data acquisition gradient is fixed throughout the procedure; the phase encoding gradient is stepped uniformly from one extreme to the other, hence the difference from one excitation to the next. (B) The *k*-space average resulting from the spin-warp sequence.

duration. $S_{p,t}$ is the signal recovered at time $t$ during recovery of the $p$th line of $k$-space.

A two-dimensional Fourier transform applied to the data results in a set of amplitudes associated with the set of positions $x$, $y$.

The process can be extended to three dimensions by adding another phase-encoding step in the third orthogonal direction. This is altered after the complete set of phase-encode steps in the second direction has been obtained. In order to obtain sufficient data to generate a volume data set with $n_i \times n_j \times n_k$ voxels, $n_j \times n_k$ acquisitions of the $n_i$ points in the readout direction are needed. If $n_j = n_k = 128$ (a relatively modest resolution target), 16384 acquisitions are required. Even if the acquisitions are repeated at 20 ms intervals, the recovery of the data takes around 5.4 minutes. Extra time, to allow for greater recovery of signal after excitation, can quickly result in very extended, and practically unacceptable, durations. Signal-to-noise ratio in such imaging procedures can be very good, but the risk of patient movement, or even refusal to proceed, becomes much greater.

## Slice selection

In most instances, single planes of data ('slices') are recovered during imaging. Slice selection is performed by selective excitation, in which an RF pulse is applied at the same time as a gradient. This selects a slice orthogonal to the direction of that gradient.

The process is illustrated in **Figure 3**. The RF waveform is modulated by a computer-generated pulse profile to give a burst of RF frequencies that are as uniform in amplitude as possible and that have minimal components outside the bandwidth wanted. The pulse profile ($B_1(t)$) results in a frequency spectrum $f_{B_1}(t)$ which, in turn, results in a range of magnetization flip angles $\alpha_t$ given by

$$\alpha_t = \gamma K_{B_1} f_{B_1}(t) t_{\mathrm{p}} \qquad [8]$$

where $K_{B_1}$ is a profile relating RF field intensity and the spectral content, and $t_{\mathrm{p}}$ is the pulse duration.

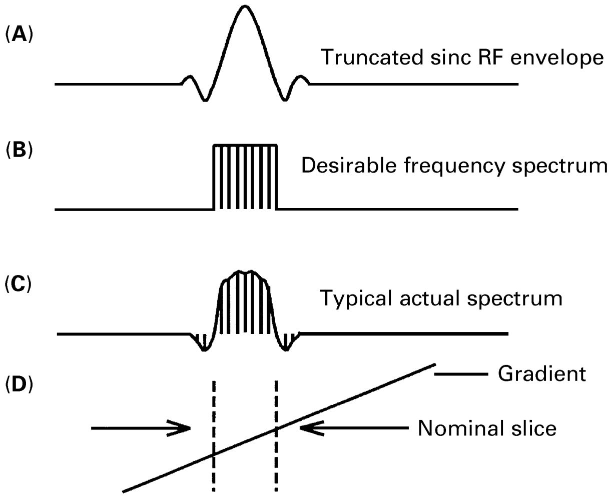



**Figure 3** Slice selection process. (A) Envelope of the RF pulse (typical of the simpler pulses used). (B) Desirable burst of frequencies. (C) More typically achieved burst of pulses. The frequencies shown as negative are, of course, of opposite phase to those in the main block of frequencies. (D) Applied gradient with slice selected marked on it relative to the components of (A) to (C). The slice is shown as perfect, but actual performance is generally significantly poorer.

Even as the $B_1$ irradiation is present, and before the gradient is completely removed, the excited spins start dephasing relative to each other and, at the completion of the process, little or no signal may be obtained. Another, inverted, gradient is then used to refocus the spins and reform the signal.

The width of the slice is determined by the gradient amplitude ($G$, conventionally measured in mT m$^{-1}$) and is given by

$$W = \frac{f_{B_1}(t)}{\gamma G} \qquad [9]$$

Initially RF pulses such as sinc pulses (with trapezoidal gradients) or high-order sinc pulses (with sinusoidal gradients) were used. Now much more sophisticated complex pulses (i.e. containing real and imaginary profile information) are used to obtain better and more exact slice selection.

Because tissue relaxation times are relatively long, scanning times are typically quite extended, with much apparently wasted time. Crooks and his colleagues followed an earlier suggestion and showed how multiple slices could be interleaved between each other, and acquired at the same time, by varying the operating frequency of the machine at successive acquisitions. Coincidentally, they also showed how to exploit the relatively long $T_2$ relaxation time of many tissues to recover more than one image from each slice excitation.

## Signal-to-noise ratio (SNR)

Although the basic formulation for signal-to-noise ratio is the same as in classic NMR, there are very significant differences in what actually happens. These arise from the fact that the body loading on the coils (which can generally be ignored in typical high-field (small-bore) systems), can easily dominate the noise from the resistance of the coils, and the input stages of the preamplifier, in all except very low main fields. The following discussion concentrates on factors that are important (or controllable) in the whole-body experiment. Thus we express Hoult and Richards' form for the signal-to-noise ratio, which is

$$\mathrm{SNR} = \frac{N_\gamma^2 \hbar^2 I(I+1) B_0^2 B_1(w) V s^{1/2}}{2(k_B T l \xi)^{1/2} (T \mu_0 \mu \omega p)^{1/4}} \qquad [10]$$

where $N$ is the Avogadro number, $\hbar$ is the Planck constant/$2\pi$, $I$ is the spin quantum number, $V$ the volume of the sample, $s$ the circumference and $l$ the length of the windings of the receiver coil, $k_B$ is the Boltzmann constant, $T$ is the coil absolute temperature, $\xi$ is the 'proximity factor' (dependent on things such as conductor spacing), $\rho$ is the resistivity of the material of which the coil is made, $\mu$ is the relative permeability, $B_1(w)$ is the field at the sample due to a unit current flowing in the coil ($\omega_0 = \gamma B_0$), as assuming only that field magnitude and coil design parameters are variables.

$$\mathrm{SNR} \propto B_0^{7/4} B_1(w) \frac{s^{1/2}}{(l\xi)^{1/2}} \qquad [11]$$

A unit volume is assumed as the target of the experiment. This derivation ignores any noise contributed by the object being studied, particularly one with large dimensions, and Hoult and Lauterbur later extended this formula to allow for the case where a large load is placed in the coil, to give the relationship

$$\mathrm{SNR} \propto \frac{B_0^2}{\left(k_d \frac{n^2}{\theta} (\xi \rho \mu \mu B_0)^{1/2} \left(k_b + k_c \frac{g}{a}\right) + k_d B_0^2 b^5 n^2 \frac{\sigma}{a^2}\right)^{1/2}} \qquad [12]$$

where the $k_c$ etc are numerical values. This describes the situation for a round 'headlike' object of radius $b$ and conductivity $\sigma$ in an $n$-turn saddle coil of included angle $\theta$, radius $a$ and length $g$.

The concept of the 'intrinsic SNR' was developed to demonstrate the sensitivity of system performance to changes in the main field in the situation where the coil is heavily loaded, giving it as

$$\mathrm{SNR} \propto B_0 \qquad [13]$$

in whole-body systems in which the body noise is the dominant factor (which it is in all except the lowest fields, or when the coils in use are very small or poorly coupled to the target tissue).

However, it was pointed out that in the form of data recovery generally used, in which multiple experiments are needed, relaxation effects cannot be ignored. Tissue $T_2$ is effectively constant over the range of fields used in current whole-body studies but $T_1$ shows a dependence on $B_0$ for which the

empirical relationship [14] was proposed,

$$T_{1A} = T_{1B} \left( \frac{B_{0A}}{B_{0B}} \right)^{\nu} \quad [14]$$

where $A$ and $B$ are two fields levels, and $\upsilon$ is found, empirically, to be in the range of 0.3–0.4 for the majority of relevant tissues.

The signal in the most basic form of study is

$$S = S_0 \left[ 1 - \exp\left( -\frac{\mathrm{TR}}{T_1} \right) \right] \exp\left( -\frac{\mathrm{TE}}{T_2} \right) \quad [15]$$

where TR is the time between successive excitations, and TE is the time between the middle of the excitation RF pulse, and the time at which the central point of $k$-space is acquired. (The American College of Radiology published a glossary of terms for whole-body NMR quite early in its clinical application. This ignored the traditional NMR notation quite cavalierly, but has become so well established in the biomedical community that it has been used in this article. Only in a few instances (some diffusion weighting and magnetization transfer terms) does it align with standard NMR practice.) Where TR is short relative to $T_1$ (as in volume scanning), equation [15] reduces to

$$S = S_0^* \frac{\mathrm{TR}}{T_1} = kS_0^* B_0^{-\upsilon} \quad [16]$$

where $S_0^*$ includes the signal dependency on $T_2$, and so on.

In addition, data have to be acquired over a finite period of time, as the acquisition gradient disperses the spin frequencies. Further, it is easier to recover data with a reduced bandwidth in lower fields as magnet homogeneities are conventionally expressed as fractions of the main field (and are a function of the unit design and relatively independent of field level), while dephasing effects due to field inhomogeneity are a function of actual field variations. For any one form of data acquisition, the noise voltage ($\propto f^{1/2}$), where $f$ is the bandwidth of the acquisition) is proportional to $B_0^{1/2}$.

Taking into consideration these latter two factors, the overall signal-to-noise of a practical imaging experiment is

$$\mathrm{SNR} \propto \left[ 1 - \exp\left( -\frac{\mathrm{TR}}{T_1} \right) \right] (pB_0)^{1/2} \quad [17]$$

where $p$ is the number of acquisitions needed to recover the data. Volume scans (where $p$ can be very large) can thus have excellent SNRs, even if the acquisition time is long. At the extreme where TR<<$T_1$, in a repetitive scanning situation,

$$\mathrm{SNR} \propto B_0^{0.2} \quad [18]$$

(assuming $\nu = 0.3$) and the benefit of increasing the field is marginal. Note, however, that as TR increases, the performance at higher fields improves with respect to that at lower ones. In practice, in the majority of whole-body situations, where noise from the body dominates that from the coil and electronics, the limits of performance are set by artefacts.

Fast scanning, one of the most significant issues in imaging because extended scanning times are unpopular with patients, and uneconomic for hospitals, has followed two very different routes. In the first, a method much used for MR angiography and in volume acquisitions, short repetition times are used with a reduced flip-angle, selected using the relationship derived by Ernst and Anderson that relates flip angle ($\alpha$), TR and $T_1$ for the optimum signal-to-noise ratio:

$$\alpha = \cos^{-1} \left[ \exp\left( -\frac{\mathrm{TR}}{T_1} \right) \right] \quad [19]$$

This is clearly only ideal for any one tissue in a complex system with multiple $T_1$ values, but, in spite of reservations about its contrast, it has been applied very successfully.

The alternative strategies depend on acquiring many lines of $k$-space consequent upon a single excitation. Mansfield developed echo planar imaging very early on, to acquire whole images in a single acquisition, while other popular forms of fast acquisition of $k$-space data imaging are more modest in their ambitions. All pay some price in signal-to-noise ratio, but all have their uses and particular capabilities.

Imaging of solids (or other systems with very short $T_2$ or $T_2^*$) involves special techniques to overcome the need to encode spatial data very quickly. The approaches are not dissimilar to those of medical MRI, but the demands on the machine are more extreme, and the topic has developed a specialized scientific community of its own.

## Contrast-to-noise ratio (CNR)

Clinicians who use MRI (and, indeed, in reality, all users who view MRI images as the basis of their studies) are actually less concerned about the signal-to-noise ratio of the data they have than they are about the contrast-to-noise ratio. In X-ray terms contrast is usually defined as

$$C = \frac{S_A - S_B}{S_A} \quad [20]$$

where $S_A$ and $S_B$ are the signals from two components A and B. In NMR this definition has proved inappropriate as $S_A$ can be zero or nearly so, as can easily occur, for example, with the inversion recovery sequence (see below), which actually delivers very high contrast. For MRI purposes, contrast-to-noise is better defined as

$$\mathrm{CNR} = \frac{S_A - S_B}{S_{0w} N} \quad [21]$$

where $S_A$ and $S_B$ are the signals as defined above, $S_{0w}$ is the fully recovered signal from water (or other suitable reference) and $N$ is the noise voltage. Much of medical MRI research has been devoted to studying how to maximize this ratio in a huge variety of diseases and circumstances.

## Spectral spatial selectivity

It is not appropriate here to enter into a discussion of spectral spatial selection methods or their problems. However, it is worth making the point that these are much closer to the concepts (and difficulties) of MRI than they are to those of traditional high-resolution spectroscopy. In particular, problems of apparent low signal-to-noise ratio may have much more to do with artefacts than traditional noise. MRI has a surprising amount, largely ignored, to contribute to high-resolution spectroscopy.

## Concluding remarks

The implementation of spatial localization in MRI (or MRS) is not hugely complex, nor is it difficult in practice. In many ways, it is deceptively easy in both instances, and it is necessary to be very aware of the range of artefacts that can rise from inadequacies of very many kinds in the design of hardware, in its implementation and in its use. It is hard to overemphasize the need for stability and accuracy in the equipment used, nor to underrate the complexity and subtlety of the effects that can arise from failure to achieve these. Unfortunately, providing even a simple discussion of these is beyond the scope of this article, though there is a brief reference to them in another article on contrast mechanisms in MRI (q.v.). Almost invariably, except in very low fields, results from whole-body experiments are more determined by artefact than they are by noise. Not infrequently, the presence of artefact is indicated only by an apparently excessive level of noise.

## List of symbols

$A_i$ = a measure of the extent of the object in the plane at $r_i$; $B_0$ = applied magnetic field strength; $B_1(t)$ = RF pulse profile; $B_1(w)$ = field at sample due to unit current in coil; $C$ = contrast; CNR = contrast-to-noise ratio; $f$ = bandwidth; $f(t)$ = frequency spectrum; $G$ = gradient field amplitude; $G_r$ = gradient field along $r$ axis (parallel to $z$ axis and to $B_0$); $\hbar$ = Planck constant/2π; $I$ = spin quantum number $k_B$ = Boltzmann constant; $K$ = profile relating RF intensity and spectral content; $l$ = receiver coil winding length; $n$ = number of observation points; $N$ = Avogradro number; $N$ = noise voltage; $P_i$ = proton density at position $r_i$; $r_i$ = position along $r$ axis; $s$ = receiver coil winding circumference; $S$ = signal strength; SNR = signal-to-noise ratio; $t$ = observation time; $t_p$ = RF pulse duration; $T$ = temperature; TE = echo time; TR = time between successive excitations; $T_1$ = spin-lattice relaxation constant; $T_{2ijk}$ = spin–spin relaxation constant for $\rho_{ijk}$; $T_p$ = duration of phase-encoding pulse; $T_{ip}$ = duration of inverted gradient pulse; $T_{ap}$ = half duration of acquisition gradient pulse; $V$ = ample volume; $\omega_0 = \gamma B_0$; $W$ = width of slice; $\alpha$ = flip angle; $\gamma$ = gyromagnetic ratio; $\mu$ = relative permeability; $\mu_0$ = permeability of free space; $\xi$ = proximity factor; $\rho$ = resistivity of coil material; $\rho_{ijk}$ = proton density of the $ijk$ voxel; $\omega$ = angular frequency.

*See also:* **Contrast Mechanisms in MRI; Fourier Transformation and Sampling Theory; Magnetic Field Gradients in High Resolution NMR; MRI Applications, Biological; MRI Applications, Clinical; MRI Applications, Clinical Flow Studies; MRI Instrumentation; NMR Principles; Two-Dimensional NMR Methods.**

## Further reading

Abragam A (1961) *Principles of Magnetic Resonance*. Oxford: Clarendon Press.

Budinger TF and Margulis AR (eds) (1986) *Medical Magnetic Resonance Imaging and Spectroscopy*. Berkely, CA: International Society for Magnetic Resonance in Medicine.
Edelstein WA, Hutchison JMS, Johnson G and Redpath T (1980) Spin warp NMR imaging and applications to human whole-body imaging. *Physics in Medicine and Biology* **25**: 751–756.
Foster MA and Hutchison JMS (eds) (1987) *Practical NMR Imaging*. Oxford: IRL Press.
Grant DM and Harris RK (eds) (1996) *Encyclopedia of NMR*. Chichester: Wiley.
Lauterbur PC (1973) Image formation by induced local interactions. Examples employing nuclear magnetic resonance. *Nature (London)* **242**: 191–192.
Mansfield P (1977) Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics* **10**: L55–58.
Mansfield P and Morris PG (1982) *NMR Imaging in Biomedicine*. New York: Academic Press.
Slichter CP (1980) *Principles of Magnetic Resonance*. Berlin: Springer-Verlag.
Young IR (1984) Signal and contrast in NMR imaging. *British Medical Bulletin* **40**(2): 139–147.

# MRI Using Stray Fields

**Edward W Randall**, Queen Mary and Westfield College, London and Instituto Superior Tecnico, Lisboa, Portugal

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

MRI is accomplished using gradients in the magnetic field flux density (normally referred to as the magnetic field). These gradients are generally produced in conventional imaging with the aid of sets of gradient coils. The suggestion to use the gradients that are present in the stray field of a superconducting magnet for imaging purposes was made first by Samoilenko and colleagues in 1988. The method has the advantage that these field gradients are large, of the order of 50 T $m^{-1}$. This has proved to be very useful not only for the imaging of solids but also additionally for the imaging of liquids in solids, neither of which can be imaged very satisfactorily or even at all with conventional MRI techniques. The method was generalized from one spatial dimension to three by Samoilenko and Zick, working at Bruker Spectrospin, but the technique was exemplified only by proton and fluorine work in diamagnetic solids such as organic polymers. The extension to multinuclear work, even quadrupolar nuclides, was accomplished by Randall and colleagues in one-dimensional studies. They also showed that the method gave good images of diamagnetic crystalline solids, and even of paramagnetic crystalline solids. Thus now virtually any nuclide, even those with electric quadrupole moments, can be imaged in any solid, with the possible exception of ferromagnetics. Because metals conduct electricity, they are not penetrated by radiofrequency fields and therefore can be imaged only when in powder form. Distortions of the images produced by magnetic susceptibility are greatly reduced in proportion to the gradient strength.

The stray field and its large gradient have been of use for spectroscopic studies also, mainly for the investigation of relaxation times and of diffusion processes. Additionally, the basic spin physics in the stray field has been of considerable interest in its own right. In the case of quadrupolar nuclides, for example, solid-state NMR spectroscopy in the stray fields gives a new method of determining the electric quadrupolar coupling constant. This may now be scanned spatially for the first time.

Initially the acronym STRAFI was used to denote stray field imaging, but more recently it has been used by Randall and colleagues more generally to denote simply the stray field itself. Studies are not confined simply to imaging. The three types of work in the stray field may thus be referred to as STRAFI imaging (or STRAFI-MRI), STRAFI spectroscopy and STRAFI diffusion studies. All of these are covered in this article, since the technique is relatively new.

It is possible to combine the techniques and not only study *localized* free induction decays, relaxation times and self-diffusion coefficients but also to constructs maps of these properties.

## The stray field

For superconducting magnets such as are used for NMR spectroscopy or for MRI studies, the central very homogenous field on the Z axis falls off rapidly as the distance from the centre of the magnet increases. Near the edge of the solenoid the gradient is particularly high. This is illustrated in **Figure 1**. The gradient can reach values of the order of 10–100 T $m^{-1}$ typically, depending on the value of the central magnetic flux density, the size of the bore of the magnet, and the details of primary coils. **Table 1** shows some typical values for proton resonance frequencies that have been reported for STRAFI work on various instruments: the positions in the gradient have not always been optimized.

Even higher fields are available on resistive magnets such as the 24 T magnet at the National High Magnetic Field Laboratory at the University of Florida. The maximum STRAFI gradient is approximately 138 T $m^{-1}$ at a flux density of about 17 T corresponding to a proton resonance frequency of 723 MHz. There is a 45 T system under construction.

High gradients give better spatial resolution in imaging experiments and allow the measurement of smaller diffusion constants. Additionally, they reduce the long-range effects of variations in magnetic susceptibility. Large fields are useful for the usual reasons, such as increased sensitivity, and for nuclides with low gyromagnetic ratios ($\gamma$) the higher resonance frequencies make probe ringing less of a problem. Thus the first solid-state STRAFI work on $^{14}N$ was accomplished at 32 MHz on the high-field, narrow-bore instrument shown in **Table 1**, in a stray field of 8.9 T.

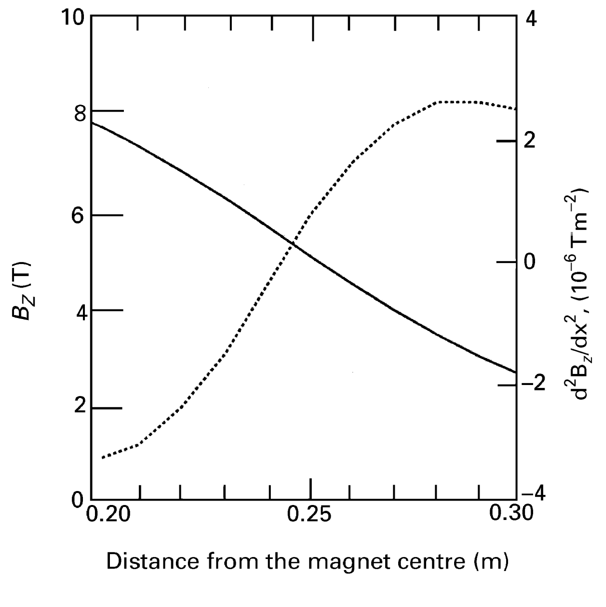



**Figure 1** Plots of the field strength $B_Z$ on the *Z* axis versus distance x from the magnet centre (solid line, left-hand abscissa) and of the second derivative of $B_Z$ with respect to *x* (dotted line, right-hand abscissa). The plots are for a typical superconducting magnet (Magnex Scientific Ltd) of bore 89 mm and a centre field of 9.4 T. Reproduced with permission from: McDonald PJ (1997) Stray field magnetic resonance imaging. *Progress in NMR Spectroscopy* **30**: 69–99.

**Table 1** Typical values for the centre and stray fields and their proton resonance frequencies

| *Centre field (T)* | *Centre frequency (MHz)* | *Bore (mm)* | *STRAFI (T)* | *Fringe frequency (MHz)* | *Gradient (T m$^{-1}$)* |
|---|---|---|---|---|---|
| 4.7 | 200 | 400 | 1.86 | 79 | 9.0 |
| 4.7 | 200 | 300 | 2.61 | 111 | 12.0 |
| 7.05 | 300 | 89 | 2.94 | 125 | 37.5 |
| 9.4 | 400 | 89 | 5.52 | 235 | 58.5 |
| 14.1 | 600 | 89 | 9.66 | 411 | 52.9 |
| 14.1 | 600 | 54 | 8.9 | 380 | 90.0 |
| 19.6 | 834 | 32 | 11.2 | 477 | 75.7 |

So far no magnet has been designed specifically for STRAFI work. This is a development that can be forecast confidently since it should be possible to obviate the expense associated with very high central field homogeneities. The exception is the portable system being developed in Aachen by Professor Blümich's group called the MOUSE, which uses an asymmetrical permanent magnet and employs the field outside the magnet. There are analogies with the systems developed for oil-well logging: although the fields are very low the *ratio* of the field strength to the gradient strength is similar to values in work on conventional superconducting magnets.

## The spin physics and STRAFI spectroscopy

With gradients of the order of 50 T $m^{-1}$ only a thin slice of the sample is excited for line widths of the order of a few kHz. The slice thickness for such narrow lines is governed (inversely) by the pulse duration, the value of $\gamma$, and the gradient strength. For 'isolated' protons in a 50 T $m^{-1}$ gradient and a pulse duration of 10 μs, the slice thickness is 80 μm, which increases to 360 μm for $^{23}Na$ (even if the quadrupole coupling is zero).

For broader lines the line width of the sample in the absence of the gradient can become the governing factor. The line has, of course, two components, one

due to homogenous broadening, governed by $T_2$, and a heterogeneous component, say for polycrystalline samples, that depends on the dominant interaction (apart from the gradient): dipolar interactions for dipolar solids and anisotropic liquids; the quadrupolar interaction for quadrupolar cases; or even the chemical shift anisotropy.

The STRAFI free induction signal for a single pulse has a decay that is governed by $T_2^*$ ($T_2^*$ is the apparent $T_2$ including inhomogeneity effects) and diffusion. Since the large gradient makes $T_2^*$ very short, it is usual to use the spin-echo technique of Hahn to overcome dead-time problems in the detection. The width of the echo is determined by $T_2^*$. The gradient term is negligible at the top of the echo where there is perfect refocusing. If diffusion is negligible then $T_2$ may be determined by arraying the spin echo delay ($\tau$) since the echo relaxation is then governed by $T_2$ only. The STRAFI echo is a true Hahn echo even for solids (because the strength of the gradient dominates other terms in the Hamiltonian), so that any two pulses will in general give an echo irrespective of their relative phase. In the absence of the large gradient, dipolar solids will give an echo only if the relative phase is different: this is the dipolar echo.

STRAFI Hahn echoes are very useful spectroscopically since they are obtained for both solids and liquids even in the same sample. Additionally, multiple primary Hahn echoes from just two pulses can be produced in the stray field even from solids.

It is usual to employ a train of pulses to produce a train of echoes. These are not to be confused with the multiple echoes mentioned above. Application of more than two pulses gives both primary and stimulated echoes. For a pulse angle of 90°, the sample response is as shown in **Figure 2** both for the EVEN sequence, $90_x - [\tau - 90_x - \tau - \text{echo} -]_n$, which has no phase change, and for the ODD sequence, which has $90_x - [\tau - 90_y - \tau - \text{echo} -]_n$. The two sequences produce the same intensity for the first echo, the relaxation properties of which are determined only by $T_2$, if diffusion is negligible. The echoes after the first one have more than one component, the number of which increases with the echo number. These contributions are all of the same sign for the ODD sequence, so that they augment each other – there is an analogy with constructive interference in diffraction – but not for the EVEN sequence, which produces echoes of opposite phase (destructive interference) and the up–down pattern shown in **Figure 2**. This has a (2 up, 2 down) pattern for 90° and an ($n$ up, $n$ down) pattern where $n$ is an integer given by $180/\alpha$ when $\alpha$, the pulse angle, is a submultiple of 180°. This is illustrated in **Figure 3**, which

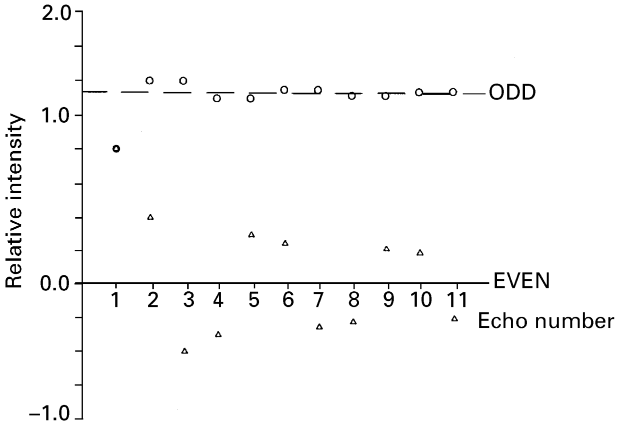



**Figure 2** Plots of the calculated normalized echo intensity in the absence of relaxation and diffusion for the ODD sequence (top, circles) and the EVEN sequence (bottom, triangles) versus echo number following a train of 90° pulses. The first echo intensity is the same for each sequence and is governed only by $T_2$ in the absence of diffusion. Subsequent echoes have contributions from $T_1$. The EVEN repeating pattern has two points up and two points down for 90°. Reproduced with permission from: Bain AD and Randall EW (1996) Spin echoes in static gradients following a series of 90 degree pulses. *Journal of Magnetic Resonance* **A123**: 49–55.

shows the responses for pulse angles of 60° (3 up and 3 down) and 45° (4 up and 4 down). The EVEN sequence is therefore very useful for the setting of the pulse angle. In fact, $\alpha$ varies across the slice but, remarkably, the observed value corresponds to the value at the middle of the slice, as confirmed by detailed calculations based on the Bloch approach or on the density matrix.

The odd sequence has the advantage as in normal spectroscopy that it produces line narrowing if $\tau$ is short enough to produce some spin locking, in which case the $T_2$ contribution to the STRAFI-echo decay can be replaced by $T_e$ (the effective decay time constant), which in the limit of spin locking may be approximated to $T_{1\rho}$. The line narrowing in the imaging experiments gives a greater spatial resolution. Even if there is no spin locking, extended echo trains may be produced since the last echoes have substantial contributions from $T_1$. Very long values of $T_1$ and $T_{1\rho}$, and hence $T_e$, can be obtained in many aprotic solids particularly if a low-$\gamma$ nuclide is being observed. A good example is $^{31}P$ in bone. In this case a very large number of pulses can be applied in a long pulse train which is of great advantage for signal accumulation (long echo train summation or LETS). The same behaviour can be seen even for quadrupolar nuclides; indeed the first such observation was for $^2H$ in heavy ice, for which 9000

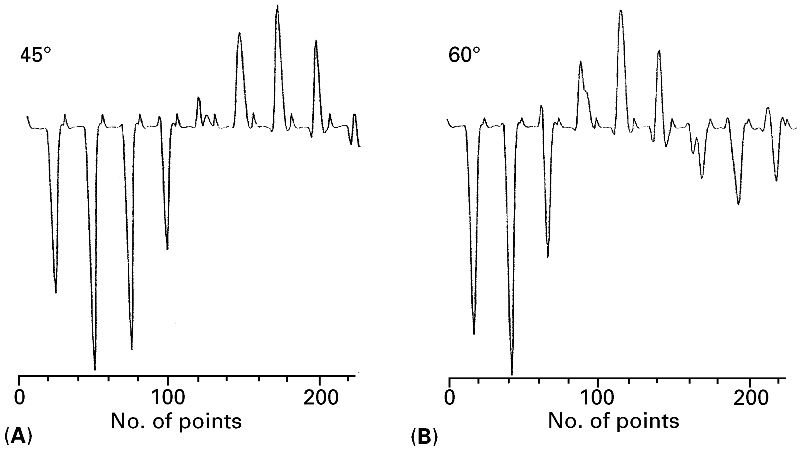



**Figure 3** Experimental data of the EVEN sequence and a pulse angle of 45°, which gives a 4-up and 4-down repeating pattern, and 60°, which gives a 3-up and 3-down pattern. The signals were for $^{1}H$ at 237.5 MHz in the 58 T $m^{-1}$ gradient of the Chemical Magnetics Infinity System at the University of Surrey for a silicone rubber sample. The $\tau$ value was 25 μs. Reproduced with permission from: Randall EW (1997) A convenient method for calibration of the pulse length in high field gradients using Hahn echoes. *Solid State NMR* **8**: 179–183.

echoes in each train have been produced. **Figure 4** shows the decay of the tops of these 9000 echoes. This decay is much longer than $T_2$ since echoes after the first have increasing degrees of $T_1$ governing their relaxation behaviour as the echo number increases. The decay is then said to be $T_1$-weighted. Of particular interest is the case of $^{14}N$, which for unprotonated nitrogens gives very long echo trains even when the value of the quadrupolar interaction is large, such as in sodium nitrite for which the literature value is about 4.9 MHz.

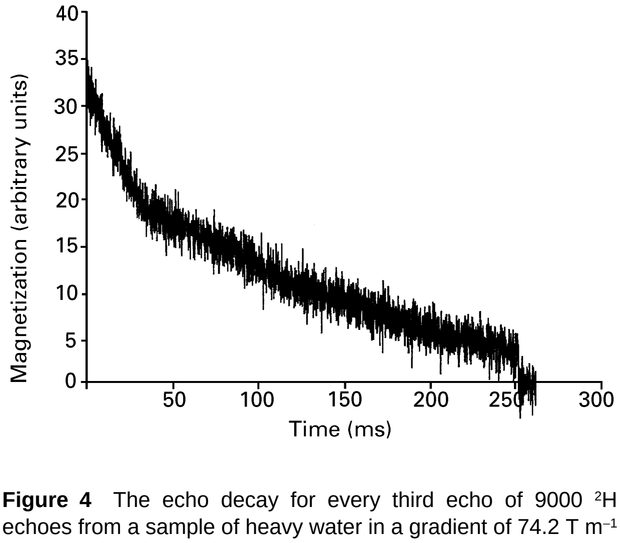



**Figure 4** The echo decay for every third echo of 9000 $^{2}H$ echoes from a sample of heavy water in a gradient of 74.2 T $m^{-1}$ and resonance frequency of 27.5 MHz. The tail of the decay is governed mainly by self-diffusion. The pulse duration was 4 μs (tip angle 90°). A total of 569 echo trains were accumulated with a repetition time of 0.5 s. Courtesy of Drs Teresa Nunes, Geneviève Guillot and the author.

STRAFI spectroscopy resembles normal spectroscopy except that the chemical shift and dipolar interactions are too small to perturb the signal appreciably. Studies are therefore confined to the study of relaxation times and diffusion. For heterogeneous solid samples, or liquids in solids, discrimination between different motional components is possible, the contributions for components with fast motions being dominant for the last echoes in a long sequence.

For quadrupolar nuclides in solid samples that have a large electric quadrupole interaction, $C_q$, the echoes exhibit splittings and modulations that depend on the magnitude of $C_q$, which may thus be measured.

## STRAFI-diffusion

Gradients have been used for the determination of diffusion coefficients since the earliest days of NMR. The advantages of pulsed gradients led to studies in fixed gradients being superseded. Now, however, the STRAFI technique, which is a fixed-gradient technique, has the advantage of employing large gradients that bring into measurable range smaller diffusion constants. Kimmich and colleagues were thus able to measure diffusion coefficients as small as $6 \times 10^{-11}$ $cm^2$ $s^{-1}$ in siloxane melts at 20.5°C in a gradient of 38.4 T $m^{-1}$. Even smaller diffusion coefficients should be accessible at gradients higher than this.

## STRAFI-MRI

STRAFI-MRI is a slice-selective method since even very short pulses excite only a narrow portion of the sample: Samoilenko's 'sensitive slice'.

The free induction decay of the excited slice contains spatial information, which can be revealed by Fourier transformation as in most other conventional NMR techniques. If the sample is a thin film with a thickness less than the width of the excited slice, the STRAFI image obtained in one spatial dimension is the image across the whole film. Larger samples, however, must be scanned to get the whole 1D projection. One way is to sweep the frequency of the pulse, but the older STRAFI method is to move the sample through the field. The first possibility is analogous to the method of frequency sweep in continuous-wave NMR spectroscopy. Field sweep in STRAFI-MRI has also been tried.

If the scanning is accomplished by sample movement, Fourier transformation is not necessary: the magnetization and the signal are directly proportional to the nuclear spin density. The 1D profile is generally obtained by translation of the sample along the $Z$ axis of the magnet. The motion can be continuous, if it is slow enough, or stepwise. The size of the step may become the major factor governing the spatial resolution if it is coarse. The slices may be interleaved so that there is no interference between the excitations of contiguous slices. Accumulation can be accomplished by repeating the motion from the starting position, since this allows relaxation to occur during the recycle time of the motion. An alternative for samples with a long $T_1$ is to use very long pulse trains as in the $^{31}$P example (the LETS protocol); see **Figure 5**.

The profiling can be repeated for different orientations around the same axis by rotation of the sample to produce a 2D image. Then the third dimension can be scanned by a second rotation orthogonal to the first. Back-projection algorithms can be used to produce the images. 3D work is very time-consuming by this method, since there are not the usual gains from the Fourier technique, but it is possible with axially symmetric samples to design the sample (normally called the phantom in imaging studies) so that useful but short experiments can be conducted in one dimension only. Examples are liquids diffusing into solids, such as solvents into organic polymers (**Figure 6**), or water diffusing into cements or concrete (**Figure 7**). In each case, relaxation weighting can be used to advantage. The system shown in **Figure 6** depicts the diffusion of hexafluorobenzene into the polymer from right to left. The sensitive planes for $^{1}$H and $^{19}$F are displaced one from the other because of the different values for the gyromagnetic ratios, and so are the images. The separation between the planes is about 4.9 mm in this example. The overall image therefore has a proton region (at the left) and a fluorine region at the right. In the sample these two regions overlap. There is a direct analogy with the chemical shift effect in conventional imaging.

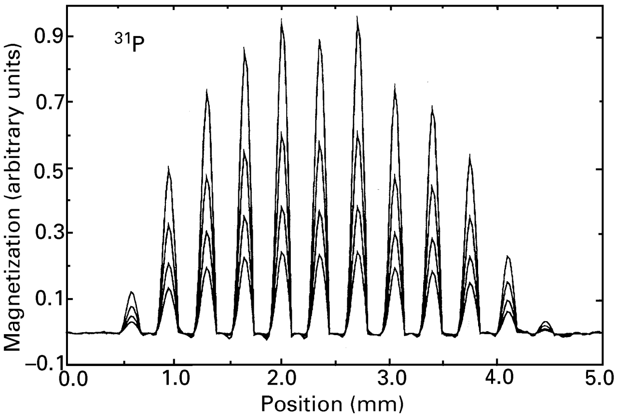



**Figure 5** $^{31}$P one-dimensional (semi-elliptical) profiles of a cylindrical sample containing ammonium hexafluorophosphate. At each position four summed echo trains are shown for long echo train summation (LETS). The highest intensity is for $n = 8192$ echoes. The gain in the signal-to-noise ratio is proportional to $n^{0.5}$ where $n$ is the number of echoes in the train. The variation in $n$ from the lowest to the highest intensity was 1024, 2048, 4096 and 8192. The odd sequence was used. Experimental details were resonance frequency 91 MHz; gradient strength 58 T m$^{-1}$; $\tau$ value 20 μs; pulse duration 3.6 μs; manual steps 0.35 mm; 64 averages of each echo train were used at 5.7 min per slice. Courtesy of Drs Duncan G. Gillies and Ben Newling.

The best position in the stray field is actually not where the gradient is maximum, because at this position the resonant slice is not planar since the magnetic field off-axis is not uniform. The optimum position is where the slice is planar. This occurs where the lines of equipotential lines are orthogonal to the $Z$ axis. Nearer the centre of the magnet the equipotential lines are curved one way, but further from the centre they curve another way. Exactly at the optimum position a flat disc extends someway off-axis, typically about 1 cm for superconducting magnets with bores of 8.9 cm.

The distortion of this disc away from the optimum position is not large for small samples, which are normally cylinders about 1–2 cm long and less than 1 cm in diameter, so that positions away from the optimum are frequently used. This is part of the reason for the apparent irregularities in **Table 1**. Thus, the inserts, RF channels and tuning boxes used for the study of $^{31}$P in the central field may be used for

**Figure 6** Profiles of a cylindrical phantom (5 mm thick, diameter 10 mm) of a polymer consisting of a 2:1 mixture of poly(methyl methacrylate) and poly(*n*-butyl methacrylate) following ingress of hexafluorobenzene (from the right): (A) projection produced from the first four echoes of 32 echoes; (B) projections from the last four echoes. The $^{19}$F image is spatially separated at the right from the $^{1}$H image at the left. The fluorine images exhibit spatially variable relaxation behaviour. The gradient strength was 37.5 T m$^{-1}$ and the resonance frequency was 123.4 MHz. Reproduced with permission from: Randall EW, Samoilenko AA and Nunes T (1996) Simultaneous $^{1}$H and $^{19}$F stray field imaging in solids and liquids. *Journal of Magnetic Resonance* **A117**: 317–319.

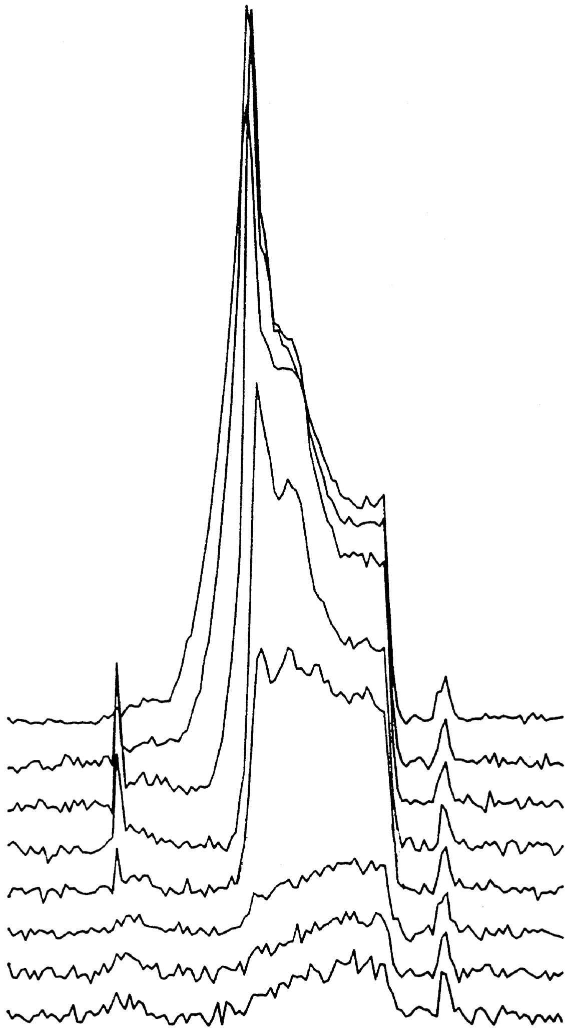


**Figure 7** Hydration and curing of a Portland (type I) cement contained in a glass tube and covered with water without initial mixing of the components. The $^{1}$H profiles were taken at various intervals thereafter: the top profile after 15 min; and the bottom one after 48 h. The peak at the right is from a plastic reference disc at the bottom of the vial, whereas the sharp peak at the left is from a parafilm top-cover and from water that initially condensed on it. The apparent loss of signal arises because of the growth of components with very short relaxation times as hardening occurs, which become invisible with the long echo times used ($\tau$ = 30 μs). The resonance frequency was 123.4 MHz and the gradient strength was 37.5 T m$^{-1}$. The ODD sequence was employed with trains of 64 echoes; 50 echo trains were accumulated for each point on each profile with a repetition time of 1 s. Courtesy of Dr Teresa Nunes.

protons in the stray field at a frequency that is rather lower than the optimum.

The gradient at the chosen position can be calibrated by measurement of an accurately known diffusion constant.

The spatial resolution has reached about 5 μm for $^{1}$H in the case of line narrowing by spin locking for thin samples by use of the Fourier technique in a gradient of about 38 T m$^{-1}$ in a 7 T magnet. For similar samples at the higher gradient strengths available on higher-field magnets, it should be possible to obtain a resolution of about 1 μm. Typically, however, for protons the resolution is about 50 μm or less.

For the $^{1}$H and $^{19}$F nuclides, the sensitive planes are spatially separated, with the fluorine-sensitive plane being at higher field close to the magnetic centre. The distance between these planes depends on both the field and the gradient but can be of the order of 5 mm. This means that both the $^{1}$H and $^{19}$F signal of a fluorohydrocarbon can be detected as the sample moves. For a small enough sample, for which the projection on the field axis is less than 5 mm, the two 1D profiles are completely separated, see **Figure 6**. This $\gamma$-displacement, similar to the chemical shift displacement in conventional imaging (which is too small generally to be seen with STRAFI work), is useful, otherwise the discrimination would have to be made with relaxation weighting alone. The $\gamma$-displacement itself can, however, be very small, so that overlap nearly always occurs for certain pairs of nuclides such as $^{23}$Na and $^{27}$Al, $^{63}$Cu and $^{65}$Cu and $^{75}$As and $^{115}$In (see **Table 2**).

For quadrupolar nuclides of half-integral spin there is a relatively narrow transition ($\frac{1}{2}$ to $-\frac{1}{2}$) and much wider transitions for the satellites. It can be that, as in spectroscopy, the latter are effectively not detected. In

**Table 2** STRAFI-frequencies of some nuclei in the 2.9 T stray field of a Bruker MSL 300 and the proportion of the signal in the central transition

| *Nucleus* | *I* | *Frequency* (MHz) | *Intensity of central transition (%)*[a] |
|---|---|---|---|
| $^{1}$H | $\frac{1}{2}$ | 123.4 | 100 |
| $^{19}$F | $\frac{1}{2}$ | 116.1 | 100 |
| $^{7}$Li | $\frac{3}{2}$ | 47.8 | 40 |
| $^{11}$B | $\frac{3}{2}$ | 39.6 | 40 |
| $^{23}$Na | $\frac{3}{2}$ | 32.2 | 40 |
| $^{65}$Cu | $\frac{3}{2}$ | 32.2 | 40 |
| $^{27}$Al | $\frac{3}{2}$ | 32.3 | 40 |
| $^{51}$V | $\frac{7}{2}$ | 32.5 | 19 |
| $^{59}$Co | $\frac{7}{2}$ | 29.2 | 19 |
| $^{115}$In | $\frac{9}{2}$ | 27.0 | 15 |

[a] The intensity of the central transition expressed as a percentage of the whole signal.

this case the image is reduced in intensity by an amount that depends on the spin quantum number as shown in **Table 2**. The resolution then is determined by the line width of the central transition, which to first order is given by $C_q^2/\nu$, where $\nu$ is the Larmor frequency. This can be of the order of kHz, so that the resolution is about the same as for dipolar-broadened $^{1}$H in polymers, say. Curiously, if the signal-to-noise ratio increases, so that the satellite transitions are detected, the spatial resolution decreases.

It is possible to distinguish the loss of signal from this source from the case of a lower spin density by a comparison of the echo trains produced by the odd and even pulse sequences.

In the case of nuclides with integral values of the spin quantum number, there is no narrow transition and the slice thickness will be very large if the line width is big because of a large value for $C_q$. Thus the observed spatial resolution may be a factor of 10 worse than the correct image: a feature of a millimetre in the phantom may appear to be a centimetre or more in the image. This has been illustrated for $^{14}$N in a series of samples in which the value of $C_q$ was changed from 0 to about 5 MHz.

It can be noted here that an observed profile, $O$, is the convolution of the actual physical profile, $P$, and the line shape function $L$:

$$O = L \otimes P \quad [1]$$

If $L$ is known, say for a powdered sample consisting of one component containing a known dipolar or quadrupolar interaction, from which the line shape can be computed, this source of heterogeneous broadening may be removed from the observed image by deconvolution. The resolution of the image is then determined partly by the homogenous broadening only, which is governed by $T_2$ or $T_{1\rho}$ if there is spin locking and line narrowing.

Samples that are paramagnetic have proved to be easy to image, at room temperature at least, rather surprisingly perhaps for high-resolution spectroscopists. The reason, as shown as long ago as 1950 by Bloembergen, is that the electron relaxation times are so short that the NMR line widths are not greatly affected. For $CuSO_4{\cdot}5H_2O$ the proton line widths are determined mainly by the proton–proton dipolar interaction, and therefore the imaging problem is no different from that for a diamagnetic hydrate. The same hold true for other ligands and for more paramagnetic samples. An advantage is that although $T_2$ is not greatly affected, $T_1$ is shortened considerably, as illustrated in **Table 3**, which is an aid to accumulation when short pulse-trains are employed.

The compounds in **Table 3** have produced 1D profiles for both the $^{1}$H and $^{19}$F nuclides.

A most striking result is the observation of STRAFI echoes for the quadrupolar $^{51}$V ($I = \frac{7}{2}$) nuclide in the paramagnetic vanadyl bis(acetonato) compound, which has two unpaired electrons on each vanadium atom. This result opens up the prospects of imaging for quadrupolar nuclides in paramagnetic situations in general.

## Conclusions: Prospects for the future

It seems clear that STRAFI-MRI will become the method of choice for the imaging of solids, and samples consisting of solids and liquids. Any type of nucleus may be used, including quadrupolar nuclides, in any type of solid. All types of materials may be addressed. The method will be seen to best advantage when simple 1D profiles are sufficient for the particular study. Its worth has been shown already for water in cements (**Figure 7**) and concrete and other building materials, and for organic liquids

**Table 3** Some values of proton relaxation times in paramagnetic compounds and values of the gram magnetic susceptibility $\chi_g$

| *Formula*[a] | $10^6\chi_g$ | $T_1$/ms | $T_2$/$\mu$s |
|---|---|---|---|
| Eu(fod)$^3$ | 3.91 | 277 | 25.3 |
| Dy(fod)$^3$ | 42.80 | 5.04 | 13.1 |
| Ho(fod)$^3$ | 35.06 | 5.40 | 14.0 |
| Yb(fod)$^3$ | 7.81 | 16.8 | 26.5 |

[a] fod is the 6,6-,7,7-,8,8,8-heptafluoro-2,2-dimethyl-3,5-octanedionato ligand.

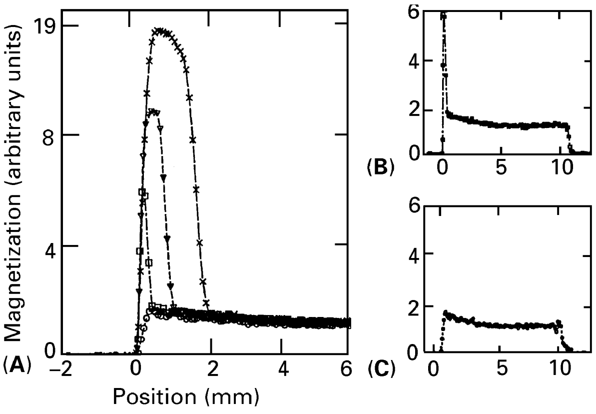



**Figure 8** One-dimensional ¹H profiles for poly(vinyl chloride) polymer exposed at the left to acetone vapour of different activities for 48 h at 20°C. In (A) the circles are for an activity of 0.13, squares for 0.35, triangles for 0.60; and crosses for 0.82. The image of the largely undisturbed solid polymer is to the right. At the left the image has overlapping images of the ingressing acetone and the softening polymer. Part (B) shows the points for an activity of 0.35 in more detail, and those for an activity of 0.13 are shown in (C) The macroscopic diffusion in (C) is of case I type. The onset of case II diffusion at the higher activities is shown clearly in (B) and in (A) – the acetone in this case has a very sharp front that is absent in the case 1 example. STRAFI-MRI has the advantage that all components, solids as well as the liquid, can be imaged simultaneously. It is then possible to employ relaxation weighting to separate the components of the image. Reproduced with permission from: Perry KL, McDonald PJ, Randall EW and Zick K (1994) Stray-field magnetic resonance imaging of the diffusion of acetone into poly(vinylchloride). *Polymer* **35**: 2744–2748.

in polymers, for which it has the advantage of imaging the undisturbed polymer, the softening polymer and the liquid phase in the softened polymer as illustrated in **Figure 8**. The macroscopic diffusion may be followed easily.

STRAFI-diffusion studies of microscopic diffusion can bring into range very slow self-diffusion coefficients.

STRAFI-spectroscopy in the excited slice can yield details of the relaxation characteristics in general, and even the values of $C_q$ for quadrupolar nuclides.

These various techniques may be combined for spatial mapping purposes, e.g. for relaxation times and quadrupole coupling constants.

There seems to be little doubt that STRAFI studies will be of great value in the general field of materials science.

As regards clinical work, *in vitro* studies have proved to be very valuable: for example, the study of materials and cements for dental work and hip-joint prostheses. One may predict that *in vivo* studies will emerge. Indeed, having the subject nearly outside the magnet has considerable advantages. For human patients, the question of the safety aspects of the large gradients will need to be addressed.

It is likely that special spectrometers will be designed exclusively for STRAFI studies in the near future.

## List of symbols

$C_q$ = quadrupolar interaction; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $T_2^*$ = effective $T_2$ in the presence of inhomogeneities; $T_e$ = effective decay constant for the STRAFI echo train; $T_{1\rho}$ = relaxation time in the rotating frame; $\alpha$ = tip angle; $\gamma$ = gyromagnetic ratio; $\tau$ = spin echo delay; $\nu$ = Larmor frequency; $\chi_g$ = magnetic susceptibility.

*See also:* **Magnetic Field Gradients in High Resolution NMR; MRI of Oil/Water in Rocks; MRI Theory; NMR in Anisotropic Systems, Theory; NMR Microscopy; NMR of Solids; NMR Principles; NMR Pulse Sequences; Solid State NMR, Methods.**

## Further reading


Bodart P, Nunes T and Randall EW (1997) Stray-field imaging of quadrupolar nuclei of half integer spin in solids. *Solid State NMR*, 8: 257–263.

Kimmich R and Fischer E (1994) One- and two-dimensional pulse sequences for diffusion experiments in the fringe field of superconducting magnets. *Journal of Magnetic Resonance* **A106**: 229–235.

McDonald PJ (1997) Stray field magnetic resonance imaging. *Progress in NMR Spectroscopy* 30: 69–99.

McDonald PJ and Newling B (1998) Stray field magnetic resonance imaging. *Reports on Progress in Physics* **61**: 1441–1493.

Randall EW (1997) ¹H and ¹⁹F MRI of solid paramagnetic compounds using large magnetic field-gradients and Hahn echoes. *Solid State NMR* 8: 173–178.

Randall EW (1997) A convenient method for calibration of the pulse length in high field gradients using Hahn echoes. *Solid State NMR* 8: 179–183.

# MS–MS

*See* **Hyphenated Techniques, Applications of in Mass Spectrometry; MS–MS and MS$^n$.**

# MS-MS and MS$^n$

**WMA Niessen**, hyphen MassSpec Consultancy, Leiden, The Netherlands

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

## Introduction

Tandem mass spectrometry (MS-MS) currently is an important technique, both in fundamental studies concerning the behaviour and structure of gas-phase ions, and in many analytical applications of MS. The history of MS-MS can be considered to go back to the observation and explanation of metastable ions by Hipple and Condon in 1945. Subsequently, the potential of metastable decompositions of ions was further investigated on magnetic sector instruments. MS-MS as a technique gained momentum in the mid-1970s. Reversed-geometry double focusing sector instruments were built in the laboratories of Cooks and of McLafferty and used for MS-MS studies, and the commercial ZAB range of sector mass spectrometers became widespread. The next major breakthrough was the introduction of triple-quadrupole instruments by the group of Yost in the early 1980s. Commercially available triple-quadrupole systems with an easy user interface stimulated the analytical use of MS-MS, which previously was primarily used in more fundamental studies. More recently, MS-MS applications have been implemented with other types of mass analysers, i.e. ion trap, time-of-flight (TOF) and Fourier-transform ion-cyclotron resonance (FT-ICR) mass spectrometers.

The principles, instrumentation and some (analytical) applications of MS-MS are reviewed in this article. More elaborate accounts on MS-MS may be found in the Further reading section.

## Principles of MS-MS

A basic instrument for MS-MS consists of a combination of two mass analysers with a reaction region between them. While a variety of instrument setups can be used in MS-MS, there is a single basic concept involved: the measurement of the *m*/*z* of ions before and after a reaction in the mass spectrometer; the reaction involves a change in mass and can be represented as

$$m_p^+ \rightarrow m_d^+ + m_n$$

where $m_p^+$ is the precursor (or parent) ion, $m_d^+$ is the product (or daughter) ion, and $m_n$ represents one (or more) neutral species. In terms of mass: $m_p = m_d + m_n$. The basic MS-MS experiment is the mass selection of the precursor ion in the first stage of analysis, the fragmentation of the precursor ion, e.g. metastable or by collision-induced dissociation (CID), and the mass analysis of the product ions in the second stage of analysis (product-ion scan).

The fragmentation of the precursor ion depends on the activation barrier of the reaction. The energy to overcome this barrier is due to the excess energy deposited in the precursor ion during ionization and transfer to the first mass analyser and, when applied, to the internal energy of the precursor ion gained by ion activation.

A metastable ion is an ion which during the ionization gained sufficient internal energy for fragmentation, but survived long enough to be extracted from the ion source. Such an ion may dissociate spontaneously during its flight from the ion source to the detector. Double focusing sector instruments can be used to detect the metastable ions.

However, in most cases, ion activation in a reaction region is applied to increase the internal energy of the ions transmitted from the ion source. Although ion activation by photodissociation and surface-induced collisions has been described, the

most widely applied method is collisional ion activation. The CID process results from the conversion of translational energy of the precursor ion into internal energy by collisions with a neutral target gas, e.g. helium or argon, admitted to the collision cell.

In CID, two collision-energy regimes should be considered: low-energy and high-energy collisions, depending on the initial translational energy of the precursor ion upon collision. High energy collisions (kV energy) are applicable in magnetic sector instruments as well as in certain applications of post-source decay in TOF instruments (see below), while low energy collisions are applied in most other systems (triple-quadrupole, ion trap and FT-ICR).

While in principle any ion generated in the ion source by any ionization technique can be subjected to MS-MS in the product-ion mode, MS-MS has especially found analytical applications in combination with soft ionization techniques, where without MS-MS only information on the intact molecule is obtained and no fragmentation is observed. MS-MS is then required to achieve structure informative fragmentation. However, there are some limitations as well. First of all, the fragmentation in CID may not always lead to sufficient fragmentation to allow unambiguous structure assignment. Furthermore, MS-MS is limited in practice to ions up to *m*/*z* of ~2000 because with larger ions the number of degrees of freedom is so large that the internal energy gained in CID is readily dissipated over a large number of bonds and no single bond acquires sufficient energy for cleavage on the observational time-scale of the instrument.

In addition to the product-ion scan, most MS-MS instruments allow other scan modes, e.g. neutral-loss and precursor-ion scan modes. The modes are not only useful in the elucidation of the fragmentation pattern of a particular compound, but also in the screening for a series of structurally-related compounds in complex samples. In the product-ion scan mode, the first mass analyser selects a particular precursor ion, while the product ions obtained by CID of this precursor are analysed in the second mass analyser. In the precursor-ion scan mode, this process is virtually reversed: the first mass analyser transmits all ions in a preset *m*/*z* window to the collision cell, while the second analyser selects only the ions of one particular *m*/*z*, e.g. a particular structure informative fragment for a series of ions or compounds. An example of the use of the precursor-ion scan mode is the monitoring of phthalate plasticisers by means of the common fragment ion at *m*/*z* 149 due to protonated phthalic anhydride. In the neutral-loss scan modes, both mass analysers are operated in the scanning mode, but at a fixed mass (*m*/*z*) difference, corresponding to a characteristic neutral loss. An example of the use of the neutral-loss mode is the monitoring of the $CO_2$ loss from deprotonated carboxylic acids. In addition, selective reaction monitoring (SRM) can be applied to monitor a specific reaction in the mass spectrometer. Both mass analysers are operated in the selection mode, i.e. selecting a particular precursor ion in the first and a particular product ion in the second mass analyser. SRM is extremely useful in the quantitative analysis of compounds in complex matrices, as a significant gain in selectivity may be achieved, leading to improvement of detection limits.

## Instrumentation for MS-MS

A wide variety of instruments for MS-MS are available.

### MS-MS in sector instruments

Initially, the study of metastable ions was performed using sector instruments. In a double-focusing magnetic sector instrument, a variety of MS-MS related experiments may be performed. Two reaction regions may be used in the field-free regions, i.e. between the ion source and the first sector and in between the two sectors (RR1 and RR2 in **Figure 1A**). A fragmentation reaction in the first field-free region of an instrument with either geometry (EB or BE) can be monitored using linked scan experiments. In a linked scan, both the magnetic field strength $\boldsymbol{B}$ and the electrostatic sector voltage are scanned but maintained in a fixed relationship throughout the scan, e.g. B–E linked scans for product-ion scans and $B^2$–E linked scans for precursor-ion scans. A disadvantage of the linked scan procedures is the limitation in resolution (typically ~1000 for the precursor ion and ~5000 for the product ion). The product ions generated in the second field-free reaction region of a reversed-geometry instrument (BE) may be monitored via a mass-analysed ion kinetic energy (MIKE) scan, where the electrostatic sector voltage is scanned while maintaining a constant magnetic field and accelerating voltage.

More advanced possibilities for MS-MS in sector instruments can be achieved in three- and four-sector instruments. The four-sector instruments enable MS-MS with high-resolution selection/analysis for both precursor and product ions. CID is performed with high-energy collisions, i.e. in most cases by single high-energy collisions with helium. The four-sector instruments obviously are complex to operate. In addition to the three- and four-sector instruments, a

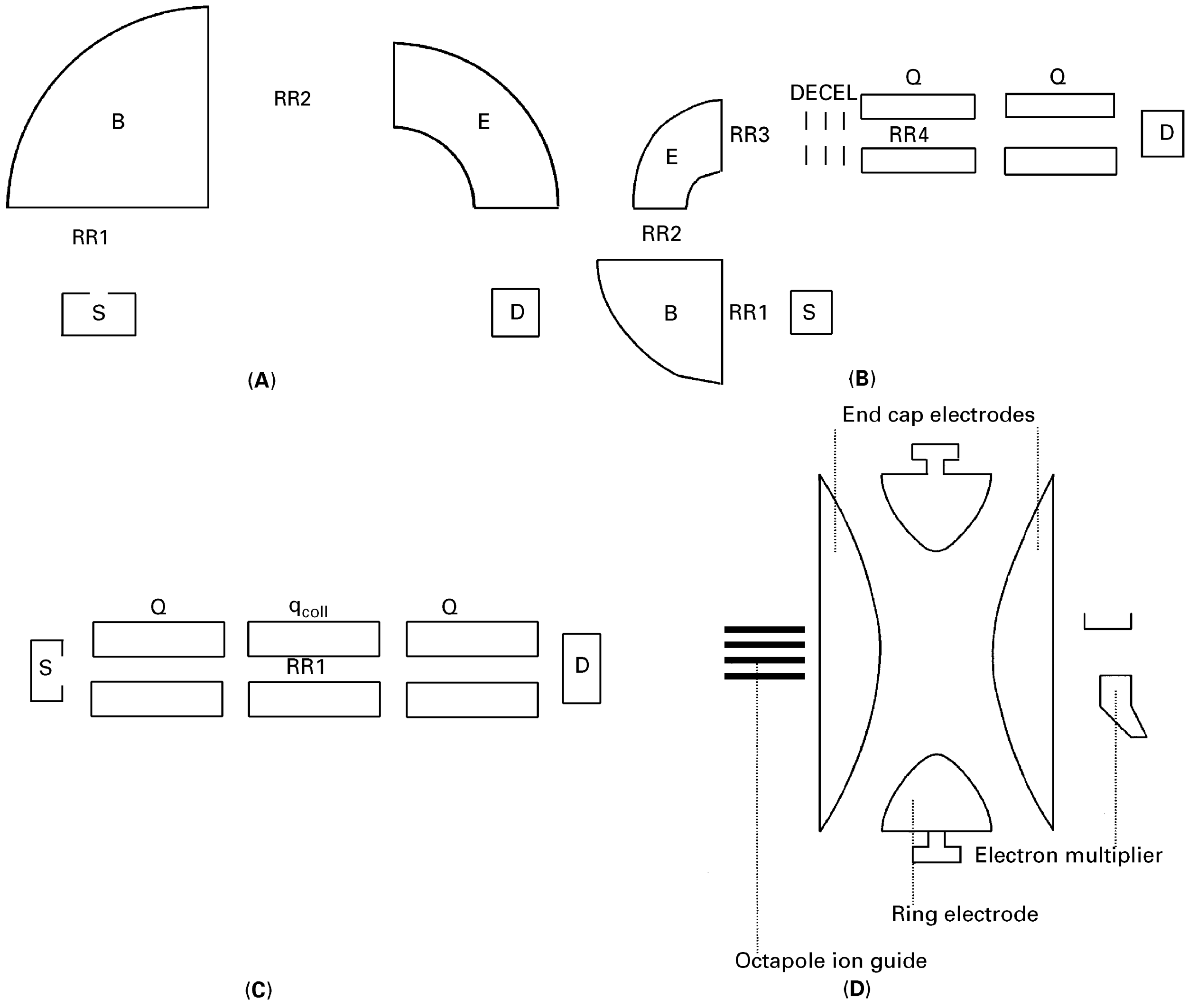



**Figure 1** Schematic diagrams of some MS-MS instrumentation: (A) double-focusing BE sector instrument, (B) hybrid BE–$q_{coll}$–Q instrument, (C) triple quadrupole Q–$q_{coll}$–Q instrument and (D) ion-trap instrument.

number of hybrid sector-based MS-MS instruments have been developed, i.e. combination of a double-focusing sector instrument as the front-end and a quadrupole, an ion trap or a TOF analyser as the back-end. These hybrid instruments, such as the BE–$q_{coll}$–Q hybrid schematically drawn in **Figure 1B**, allow high-resolution precursor-ion selection, while collisions can be achieved in either the high-energy or the low-energy regime, i.e. in the third field-free region or in the quadrupole collision cell (RR4 in **Figure 1B**), respectively.

### MS-MS in triple-quadrupole instruments

A triple-quadrupole instrument (Q–$q_{coll}$–Q) consists of two quadrupole mass analysers, connected by means of a collision cell, which is an RF-only quadrupole device (**Figure 1C**). The mass analysers can be used for rapid scanning or selection in various MS-MS scan modes. The collision cell transmits virtually all ions irrespective of their *m*/*z* without separation of different *m*/*z*, enabling efficient collection and refocusing of the product ions generated by low-energy, multiple collisions with argon. Good sensitivity is achieved with such an instrument. Two-fold improvements in sensitivity can be achieved by replacing the RF-only quadrupole with hexapole or octapole collision cells. In addition, some manufacturers position the three quadrupole elements in a banana shape to reduce the collisions of neutrals on the detector, thereby reducing the background noise. The triple-quadrupole instruments are currently the most widely used instruments for MS-MS, although most instruments are used for routine

quantitative bioanalysis in SRM mode after liquid chromatography (LC).

Recently, an extremely powerful hybrid Q–TOF system has been introduced, based on a quadrupole analyser as the front-end and a TOF analyser as the back-end (Q–$q_{coll}$–TOF geometry). As a result of the orthogonal geometry of the quadrupole and TOF, the TOF analyser enables high-resolution operation and thus accurate mass determination (up to 5 ppm), which can be extremely useful in structure elucidation by MS-MS, e.g. in peptide sequencing and impurity profiling.

## MS-MS and MS-MS$^n$ in ion-trap instruments

MS-MS in an ion-trap instrument is fundamentally different from MS-MS in sector and triple-quadrupole instruments. While in the latter the various stages of the process, i.e. precursor ion selection, CID and product-ion mass analysis are performed in different spatial regions of the instrument, in ion-trap instruments these stages are performed consecutively within the ion trap itself. It is 'tandem-in-time' rather than 'tandem-in-space' mass spectrometry. A simplified diagram of an ion-trap system is shown in **Figure 1D**.

The measurement process consists of a series of steps: ions are either generated inside the ion-trap or injected into the trap from an external source. A suitable storage voltage applied to the ring electrode enables trapping of the ions generated or injected. Next, the precursor ion is selected by applying an ion isolation RF waveform voltage at the endcaps. At the same time the voltage on the ring electrode must be ramped to a new value to store both the precursor ion and its product ions. Subsequently, CID is achieved by applying a resonance excitation RF voltage at the endcaps. This induces faster and more extensive ion trajectories in the trap, resulting in ion activation and subsequent fragmentation. The helium bath gas, present in the ion trap for stabilization of the ion trajectories, serves as the collision gas. In a normal product-ion MS-MS experiment, at this stage the product ions are scanned out of the ion trap towards the detector. However, the system also allows multiple stages of MS-MS, i.e. one of the product ions may be selected by means of an ion isolation RF waveform voltage at the endcaps, and subsequently activated and dissociated. This process can be performed up to ten times in current commercially available ion-trap systems, or as long as a sufficient number of ions remains; in practice, this typically allows four or five stages of MS-MS for most applications.

Compared with triple-quadrupole and especially with sector instruments, the ion-trap instrument provides more efficient conversion of precursor ion into product ions. However, the CID process via resonance excitation, although quite efficient in terms of conversion yield, generally results in only one (major) product ion in the product-ion mass spectrum. While with a triple-quadrupole instrument a series of product ions is observed, only one major fragment ion is observed with the ion trap. Other fragment ions are detected in the ion-trap system after multiple stages of MS-MS. The product-ion mass spectrum of the protonated triazine herbicide propazine described by Hogenboom and co-workers may serve as an example. In the spectrum from a triple-quadrupole instrument, the product ions are found at *m/z* 188, 146, 110 and 79. With the ion-trap instrument, three stages of MS-MS are required to observe all these ions, i.e. *m/z* 188 and a minor 146 in the first stage, *m/z* 146 in the second stage, and *m/z* 110, 104, 86 and 79 in the third stage. In this experiment, the most abundant product ion in each stage is selected as the precursor ion for the next stage. The fragmentation in the ion trap appears to be softer, or more controllable, than that in triple-quadrupole instruments. This stepwise fragmentation can be extremely useful in structure elucidation, but is a serious limitation in cases where confirmation of identity should be based on the detection of a number of particular product ions.

With the current commercial availability of various ion-trap MS-MS$^n$ systems, equipped with an external ion source and for use in both GC-MS and LC-MS, MS-MS in ion-trap systems can be expected to find more elaborate analytical applications.

## MS-MS in time-of-flight instruments

MS-MS in TOF instruments has only recently been reported by Kaufmann and Spengler. A reflectron-TOF should be used for MS-MS. While the ion activation may take place during ionization or even after ion acceleration, the actual fragmentation takes place after the ions have left the ion source (post-source decay). As a result, the precursor ions p and product ions d will have the same velocity, but different kinetic energy. Therefore, the MS-MS experiments can be performed in two steps, outlined in **Figure 2**: first, the instrument is operated in the linear mode, with the flight times of precursor and product ion being the same. Second, the instrument is operated in the reflectron mode. In this case, the precursor ion, having the higher kinetic energy, takes a longer path than the product ion. By comparison of the mass spectra from the linear and the reflectron

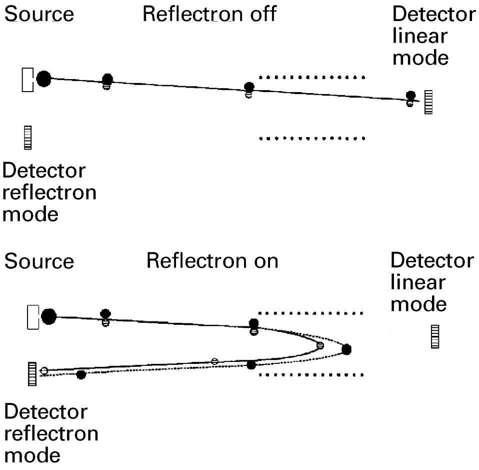



**Figure 2** Schematic diagram of the procedure of MS-MS with post-source decay in a reflectron TOF instrument. Reproduced with permission of Masson éditeur and John Wiley & Sons from de Hoffmann E, Charette J and Stroobant V (1996) *Mass Spectrometry, Principles and Applications*. Paris: Masson éditeur. Chichester: John Wiley & Sons.

mode, the fragment ions can be identified. Recently, it was also demonstrated that, via electrostatic ion manipulation of the ion beam, precursor-ion selection before post-source decay can be used. These approaches are extremely useful for the advanced mass analysis of ions generated by matrix-assisted laser desorption/ionization (MALDI).

An interesting approach to MS-MS is the combination of an ion-trap storage device and a reflectron time-of-flight instrument. While initially the ion trap was used only to store ions from the continuous ion source before the pulsed acceleration to the TOF analyser, it has recently been demonstrated by the group of Lubman that MS-MS in the ion trap before TOF mass analysis results in additional possibilities.

### MS-MS in FT-ICR instruments

In an FT-ICR instrument, which also is an ion-trapping device, MS-MS can be performed in a manner similar to MS-MS in ion-trap instruments. However, fragmentation by collisions is generally much less effective because of the significantly lower pressures in the FT-ICR cell; this is only partially compensated for by the longer ion residence times that are achievable.

In both the quadrupole ion traps and the FT-ICR instruments, MS-MS in the product-ion scan mode is feasible, but other scan modes such as neutral-loss and precursor-ion scans are not possible.

## Applications of MS-MS

Numerous applications of MS-MS have been reported in the literature, in which the MS-MS instrumentation is either used as a stand-alone instrument for sample introduction by a probe, or column-bypass injection in a liquid stream, or used in on-line combination with GC or LC. Especially in the latter area, where soft ionization strategies are frequently applied, MS-MS plays an important role. In this section, a number of applications of different MS-MS instruments is briefly reviewed. Most attention is focused on the analytical applications of MS-MS.

### Fundamental studies

The first studies with MS-MS concerned fundamental studies on the properties, structures and behaviour of gas-phase ions. This area continues to be important, as these fundamental studies lead to a better understanding of the processes involved in ion activation, fragmentation and CID. Fundamental studies can be subdivided into studies concerning the structures of ions in the gas phase, elucidation of reaction mechanisms and studies directed at obtaining thermochemical information from gas-phase species.

A currently important area of fundamental research in ion structures and reaction mechanisms, which also has significant analytical application, is the study of the fragmentation mechanisms of biomolecules in MS-MS. MS-MS is frequently applied to perform sequence analysis of biomolecules such as peptides, oligosaccharides and oligonucleotides. Fundamental studies are directed at elucidation of the formation of the various fragment ions observed, i.e. ions from backbone cleavages, side-chain specific ions and low mass ions such as immonium ions. In this respect, but also in other contexts, considerable research is directed at gaining a better understanding of the charge-remote fragmentations observed in the analysis of peptides, fatty acids and other compounds.

### Structure elucidation

MS-MS in the product-ion scan mode is generally quite successful in structure elucidation of unknowns. There are ample examples available in the literature. However, the interpretation of the production mass spectrum is not always straightforward. When a protonated molecule is selected as precursor ion, the fragmentation rules significantly differ from the well-known fragmentation rules valid for molecular ions generated by electron ionization. Upon fragmentation, protonated molecules have a

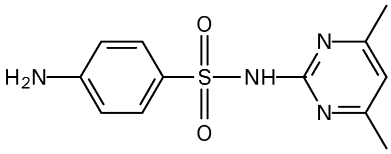
high tendency to rearrange and hydrogen shifts often take place. Knowledge of this type of fragmentation is less extensive and less systematically documented. Furthermore, the fragmentation in MS-MS may not result in a sufficient number of fragments to allow complete elucidation of the structure. In that case, additional strategies are required. In MS-MS studies with sample introduction via an electrospray interface for LC-MS, the combination of in-source CID and product-ion MS-MS of one of the fragment ions can be a useful tool to achieve further structure elucidation, as has been demonstrated, for example, by Bateman and co-workers for isomeric sulfonamides separated by capillary electrophoresis.

Because the experimental conditions for product-ion mass spectra are strongly instrument dependent, generally not very well standardized and difficult to exchange between instruments from different manufacturers, there are at present no generally applicable spectral libraries for MS-MS spectra that can assist in structure elucidation problems.

Two recent instrumental innovations can be applied to further assist in the interpretation of product-ion mass spectra. As discussed above, the MS-MS$^n$ capabilities of an ion-trap system allow the step-wise fragmentation of an analyte, facilitating the interpretation of the product-ion information and the fragmentation reactions involved. The use of a Q-TOF hybrid instrument allows accurate mass determination (at an accuracy of 5 ppm) of the product ions observed, which also facilitates the interpretation of the product-ion mass spectra.

### Screening for structurally-related compounds

The combination of MS-MS and a soft ionization method is frequently applied in screening food or environmental samples for contaminants. The strategies involved are briefly illustrated with an example related to the detection of the sulfonamide antibiotic sulfamethazine (SDM) and its possible conversion products in meat samples.




In the product-ion mass spectra of protonated SDM and one of its expected conversion products (deaminosulfamethazine, DAS), a common peak at *m/z* 124 is observed, corresponding to the protonated aminodimethylpyrimidine. This ion is complementary to the ion owing to the phenyl–$SO_2^+$ group, found at *m/z* 141 for DAS and at *m/z* 156 for SDM. To screen meat samples for the presence of SDM and its conversion products, a neutral-loss scan was applied with a common loss of 123 Da, corresponding to the loss of the aminodimethylpyrimidine. This provides a highly selective screening method for SDM-related compounds: only those compounds in the meat extract that show the neutral loss of 123 Da are detected. The result for a sausage extract, shown in **Figure 3**, indicates that both SDM and DAS are present in the meat extract, as well as a compound with *m/z* 280, most likely a hydroxy analogue, and a monochloro compound with *m/z* 298, where the chlorine attachment is due to the brine used in production of the sausage. In this particular case, the use of the precursor-ion scan mode with *m/z* 124 as the common product ion would have been successful as well. However, for many compounds the charge is preferentially retained at only one part of the molecule, and in CID the part of the molecule with the most-favourable ionization properties will be observed as an ion in the product-ion mass spectrum, while the complementary part is lost as a neutral and thus not detected. For example, this is the case with protonated nucleosides: after CID only the protonated nucleobase is observed, while no fragment ions due to the sugar ring are observed. The neutral-loss scan mode is now applicable to screen for modifications of the base, while possible modifications of the sugar ring remain undetected. The complementary precursor ion scan is not applicable in this case.

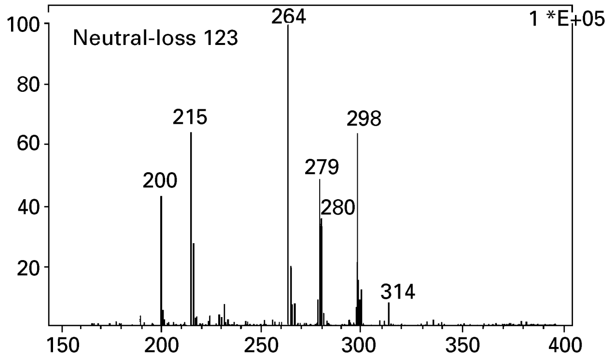



**Figure 3** Neutral-loss 123 Da mass spectrum obtained from the introduction of a sausage extract into a thermospray-MS-MS system. The peaks detected at *m/z* 200 and 215 were also present in the blank.

Similar screening procedures may be applied to perform Phase II metabolic profiling, i.e. based on the neutral-loss scan mode with a loss of 176 or 80 Da for glucuronide and sulfate conjugates, respectively.

### Quantitative bioanalysis

The use of MS-MS, especially in SRM mode, is an important application of MS-MS, especially in terms of instrument sales. Triple-quadrupole instruments are most widely used for this purpose, in most cases in an online LC-MS combination. In developing the method, the MS-MS parameters are optimized in such a way that only one or two intense fragment ions are preferentially observed. These fragment ions are subsequently used in the SRM procedure, selecting the precursor ion in the first mass analyser and the fragment ion in the second mass analyser. By means of SRM, a specific reaction in the gas phase is monitored. By careful selection of an appropriate reaction, excellent selectivity can be achieved, and the analyte of interest can be analysed at low levels in complex matrices. This approach is especially useful in high-throughput quantitative bioanalysis, required to acquire pharmacokinetic and pharmacodynamic data of new drugs and their metabolites during drug developmental studies. In such studies, an isotopically labelled internal standard is applied. This internal standard preferentially shows the same neutral loss as the analyte of interest, enabling the detection of the analyte and the internal standard at two different *m/z* values in the SRM procedure.

### Peptide sequencing

A special application of structure elucidation by product-ion MS-MS is the sequence determination of amino acids in a peptide, which has been reviewed by Papayannopoulos. The low-energy CID of a protonated peptide in a triple-quadrupole instrument leads to a series of fragment ions that originate from backbone cleavages. The cleavage of the peptide bond primarily leads to either an acylium ion, when the charge is retained on the N-terminal side of the peptide, or a protonated peptide when the charge is retained at the C-terminal side of the peptide. According to the Roepstorff–Fohlmann nomenclature, these ions are indicated as $b_n$ and $y''_n$ ions, respectively. The unknown amino acid sequence may be read from the two complementary series of ions in the product-ion mass spectrum. In high-energy CID, more advanced information can be obtained in peptide sequencing because, in addition to the backbone fragmentation, side-chain-specific fragment ions may be observed as well. This is especially important in the structure elucidation of isomeric amino acids, e.g. leucine and isoleucine. Given the extensive use of MALDI in peptide and protein mass analysis, the use of post-source decay processes to achieve peptide sequencing is of increasing importance.

*See also:* **Ion Trap Mass Spectrometers; Metastable Ions; Peptides and Proteins Studied Using Mass Spectrometry; Sector Mass Spectrometers; Surface Induced Dissociation in Mass Spectrometry; Time of Flight Mass Spectrometers.**

## Further reading

Bateman KP, Locke SJ and Volmer DA (1997) Characterization of isomeric sulfonamides using capillary zone electrophoresis coupled with nano-electrospray quasi-MS/MS/MS. *Journal of Mass Spectrometry* **31**: 297.

Busch KL, Glish GL and McLuckey SA (1988) *Mass Spectrometry/Mass Spectrometry*. New York: VCH Publishers.

Cooks RG, Beynon JH, Caprioli RM and Lester GR (1973) *Metastable Ions*. Amsterdam: Elsevier.

Hogenboom AC, Niessen WMA and Brinkman UATh (1998) Rapid target analysis of microcontaminants in water by on-line single-short-column liquid chromatography combined with atmospheric-pressure chemical ionization ion-trap mass spectrometry. *Journal of Chromatography* A **794**: 201.

Hunt DF, Shabanowitz J, Harvey TM and Coates ML (1983) Analysis of organics in the environment by functional group using a triple quadrupole mass spectrometer. *Journal of Chromatography* **271**: 93.

Johnson JV and Yost RA (1985) Tandem mass spectrometry for trace analysis. *Analytical Chemistry* **57**: 758A.

McLafferty FW (ed) (1983) *Tandem Mass Spectrometry*. New York: John Wiley & Sons.

Papayannopoulos IA (1995) The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrometry Reviews* **14**: 49.

Perchalski PJ, Yost RA and Wilder BJ (1982) Structural elucidation of drug metabolites by triple quadrupole mass spectrometry. *Analytical Chemistry* **54**: 1466.

Spengler B (1997) Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. *Journal of Mass Spectrometry* **32**: 1019.

Wu J-T, He L, Li MX, Parus S and Lubman DM (1997) On-line capillary separations/tandem mass spectrometry for protein digest analysis by using an ion trap storage/reflectron time-of-flight mass detector. *Journal of the American Society of Mass Spectrometry* **8**: 1237.

# Multiphoton Excitation in Mass Spectrometry

**Ulrich Boesl**, Technische Universität München, Germany

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Multiphoton excitation in mass spectrometry is determined by the characteristics of the applied light source (wavelength and intensity). One dominant feature therefore is spectroscopy. This concerns spectroscopy as a means of obtaining a species-selective ion source as well as mass selection for species-selective spectroscopy in a mixture. The latter leads to a rich manifold of techniques in many spectroscopic fields such as UV spectroscopy, cation and anion spectroscopy and photoelectron spectroscopy. Another dominant feature is photodissociation. Here multiphoton excitation allows variation from exceptionally soft ionization to very hard dissociation. It enables high yields of metastable decay products, which are particularly valuable for kinetic and energetic studies of molecular ions, a field of basic research in mass spectrometry. A third aspect is good compatibility with other techniques, e.g. chromatography or laser desorption, creating connections to different types of molecular systems (e.g. biomolecules) as well as to other analytical techniques. The last feature of multiphoton excitation dealt with in this article is options in analytical chemistry. The combination of speed, selectivity and sensitivity in a multiphoton ion source has great potential for environmental trace analysis or process integrated analysis of industrial production procedures, to name only two fields of application.

## Multiphoton excitation schemes

In **Figure 1** different types of multiphoton excitation are schematically presented.

**Figure 1A**. Resonance ionization spectroscopy is performed by monitoring molecular ions while tuning the laser wavelength. If the energy of one photon is in resonance with a neutral electronic excited state a second photon is able to be absorbed, giving rise to an ion current peak. Thus, the neutral UV-absorption spectrum is transferred to the ion current which can be recorded mass selectively (in opposition to the absorption). Thus UV spectroscopy and mass spectroscopy are combined as a two-dimensional technique.

**Figure 1B**. Instead of a fixed mass window, the laser wavelength may be kept constant and in resonance with a specific transition of one molecule in a mixture (e.g. molecule $M_A$) but out of resonance with other molecules (e.g. molecule $M_B$). Thus, a species selective ion source is achieved. This is unique in mass spectrometry and particularly useful for trace analysis. In addition, by selecting particular intermediate states, molecular cations may even be formed in a few vibrational or rotational levels of the ionic ground state. This is a very valuable feature for studying molecular structure, reaction kinetics or spectroscopy of molecular ions.

**Figure 1C**. Additional excitation of the molecular ions takes place at increased laser intensities, where photon absorption (in the ion and fragment ion manifold) and dissociation processes to fragment ions follow each other. Owing to this 'ladder switching' the degree of fragmentation is tuneable from soft ionization (mainly or solely molecular ions) to hard ionization (mainly small or even atomic fragment ions). At appropriate laser wavelengths and intensities metastable decay processes may dominate the mass spectrum. Multiphoton excitation and dissociation is also possible with delayed laser pulses and at different positions in space, giving rise to new forms of tandem mass spectrometry.

**Figure 1D**. If a tunable laser is used for secondary excitation of molecular ions, resonance dissociation spectroscopy of molecular cations may be performed. Here excited cationic levels are subject to laser spectroscopy. They serve as intermediate states for the process of resonance enhanced multiphoton dissociation. This is quite similar to resonance ionization spectroscopy of neutrals. The difference is that a dissociation instead of an ionization continuum is finally reached by multiphoton excitation. The advantage of this technique is that it is independent of high ion numbers (as necessary for absorption spectroscopy), fluorescence [necessary for laser-induced fluorescence (LIF)] or predissociation and therefore is fairly general. In addition, mass selectivity is intrinsic and one may benefit from state selective ion formation if resonance multiphoton ionization is used as an ion source.

**Figure 1E**. The degree of this state selective ion formation can be tested by analysing the kinetic energy of the emitted electrons while the laser

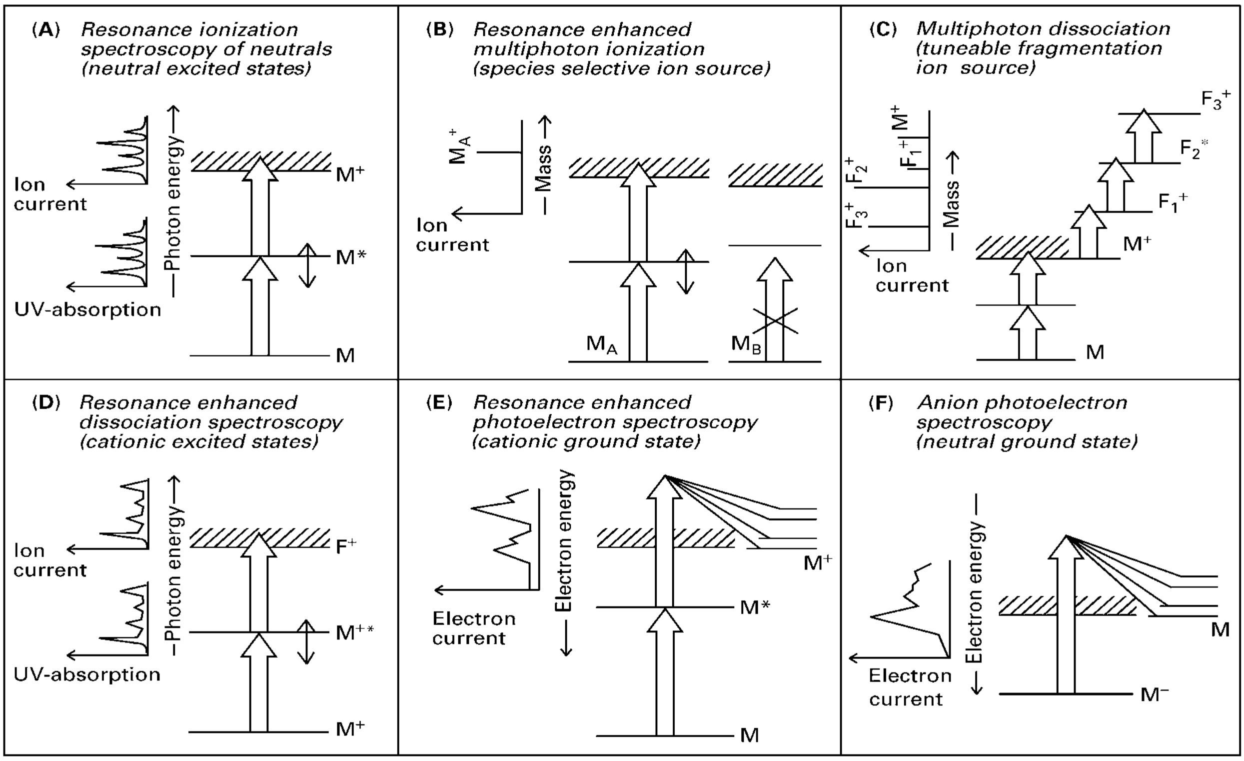



**Figure 1** Different multiphoton excitation schemes applicable to molecular systems in the gas phase in mass spectrometry.

wavelength is kept in resonance with the selected intermediate levels. This resonance enhanced photoelectron spectroscopy gives information about cationic ground states. It also helps in studies of the nature of the neutral intermediate states due to selection or propensity rules valid at the ionization and photoelectron emission process. A special, high-resolution version of photoelectron spectroscopy is zero kinetic energy photoelectron spectroscopy (ZEKE). If an intermediate excited state is involved, the first photon has to be kept in resonance with it while the second one is tuned over so-called ZEKE states very near to ionization thresholds. They end up in cationic states after field ionization. Instead of electrons, the cations may be recorded, allowing the option of mass selectivity.

**Figure 1F**. Finally, anion photoelectron spectroscopy (or photodetachment photoelectron spectroscopy) is discussed (although mostly a one-photon process). Owing to electron affinities being much smaller than ionization energies one photon absorption and detachment with lasers is possible for most anions. Either conventional photoelectron spectroscopy (**Figure 1F**) with a fixed laser wavelength or anion-ZEKE spectroscopy (similar but not equal to ZEKE) with tuned laser wavelengths is possible. The intriguing feature of anion photoelectron spectroscopy is that it starts with ionic species (mass selection before spectroscopy) and ends up with neutral species (access to short-lived species, radicals, weakly bound molecular systems, traces in complex mixtures, etc.).

## Experimental realization in a mass spectrometer

In **Figure 2** several experimental set-ups are summarized in a hypothetical instrument that consists of a central part (neutral source, ion sources, mass separator, ion detector) and several options for a variety of multiphoton excitation experiments (see **Figure 1**). The neutral inlet system or source may be narrow tubing, for an effusive molecular beam, or a pulsed nozzle for supersonic molecular beams. The advantage of the former is its very simple and inexpensive construction. By placing the end of the tubing between the electrodes of the ion extraction optics (ion source II) high gas densities may be achieved at the position of laser ionization without too high a load on the vacuum system. Supersonic beams, on the other hand, allow efficient cooling of the internal degrees of freedom of molecular motion. Well-structured spectra of even larger molecules are

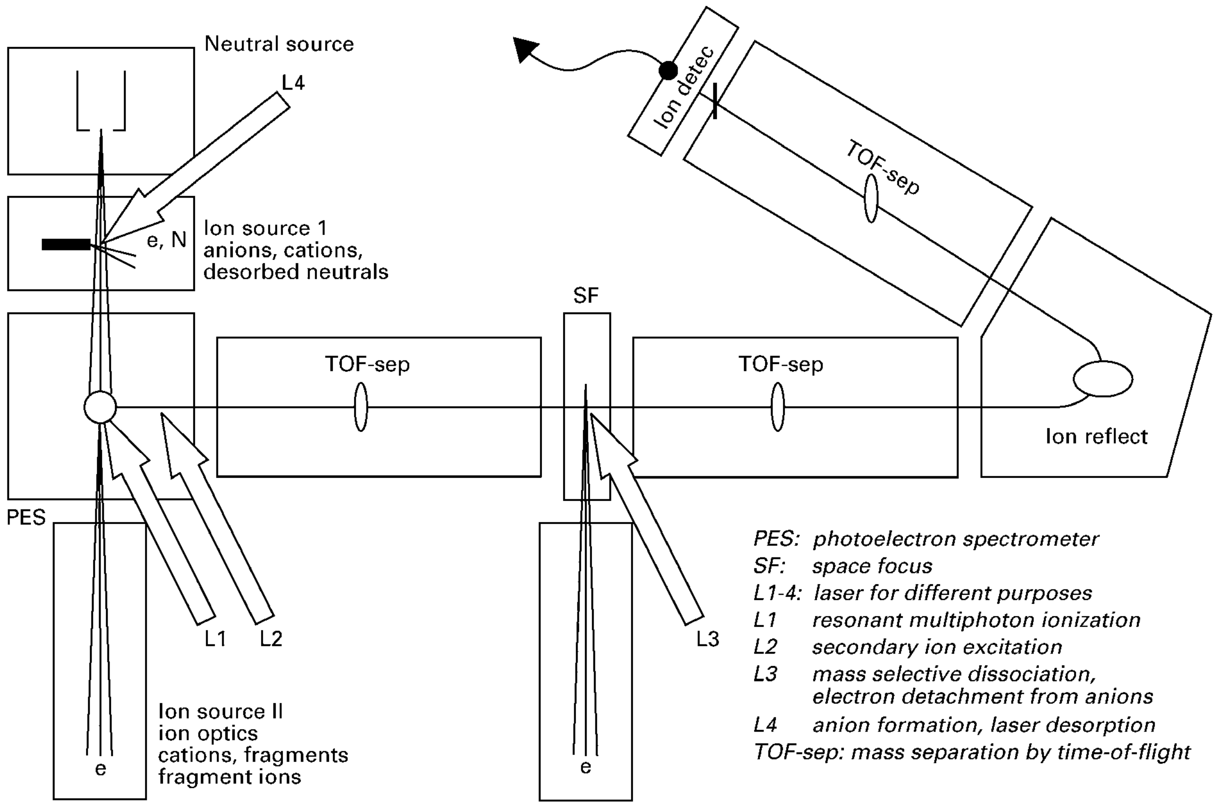



**Figure 2** A hypothetical instrument that summarizes experimental arrangements where multiphoton excitation is involved.

thus achievable, containing rich spectroscopic information and enabling highly selective excitation and multiphoton ionization. Large involatile neutral molecules may be transferred into the gas phase by laser desorption. The insertion of a gas chromatographic capillary column in the neutral source even makes possible species selection before the gas inlet.

There exist a number of laser-induced techniques for ionizing these neutrals, such as resonance-enhanced multiphoton ionization (in ion source I or II, **Figures 1A** and **1B**), laser-induced vacuum UV ionization (e.g. by 118 nm from a 3×355 nm/Nd:YAG laser), electron attachment (in ion source I) and electron ionization (in ion source II). Electrons may be supplied by laser-induced photoelectron emission from metal surfaces, preferably from thin wires made out of material having a low work function (e.g. hafnium). Ions formed in source I have to drift into the ion optics of the mass spectrometer (together with the neutral molecular beam) and can be extracted by a pulsed electric field while ionization in source II may be performed within a static electric field with instantaneous ion extraction.

Extracted ions will enter the field-free drift region of a time-of-flight (TOF) mass separator. Time-of-flight mass analysers have turned out to be the ideal mass selection tools for pulsed laser-induced multiphoton excitation and ionization. A first mass selective detection may be performed in the so-called space focus of the ion source. Even for very short field-free drift regions (e.g. some 15 cm) a mass resolution of 200 to 300 is possible in routine operations. A considerable enhancement of mass resolution is achieved by adding a special ion reflector with further field-free drift regions. In such a reflectron TOF analyser the ion cloud in the space focus (SF) is imaged onto the ion detector. This preserves the flight time distribution $\Delta t$, but extends the total flight time $t$ thus enhancing the mass resolution ($R_{50\%} = \frac{1}{2}t/\Delta t$).

Optional photoelectron spectrometers are included in **Figure 2**. These preferably are electric and magnetic field-free drift regions, allowing the analysis of electron kinetic energies by measuring electron flight times. They may be combined with ion source II for analysing photoelectrons emitted at resonance enhanced laser ionization (**Figure 1E**). A photoelectron TOF spectrometer can also be placed at the SF. Electrons due to photodetachment of anions may then be analysed (**Figure 1F**). Since in the space focus masses are already separated, this kind of

anion photoelectron spectroscopy is intrinsically mass selective. For neutral ← cation photoelectron spectroscopy at ion source II, mass selection is only possible by additional photoion/photoelectron coincidence techniques.

In **Figure 2**, several laser beams are supposed to be used in combination or separately. By laser beam L1 resonance enhanced multiphoton ionization and spectroscopy may be performed (**Figures 1A** and **1B**). At higher intensities of laser L1, multiphoton dissociation additionally takes place (**Figure 1C**). Secondary multiphoton dissociation is possible by laser beam L2. Mass selective analysis of the manifold of secondary fragments results in tandem mass spectrometry. Recording the intensity of a selected fragment ion as function of laser L2 wavelength allows cation UV-visible spectroscopy (**Figure 1D**). Both experiments (tandem mass spectrometry or cation UV-spectroscopy) are also possible with laser beam L3 coupled into the SF. By the latter experimental arrangement considerably higher secondary mass selectivity is achieved, while the former arrangement may give better sensitivity. In addition, with laser L3 mass selective neutral ← anion photoelectron (**Figure 1F**) and anion-ZEKE spectroscopy can be performed as mentioned above. Finally, laser beam L4 is used for photoelectron emission from a wire for anion formation by electron attachment. In addition, desorption of involatile neutral molecules into a supersonic molecular gas beam is performed with laser L4.

## Resonance multiphoton ionization: spectroscopy and selective ion source

Mass selective laser excitation was applied originally for the UV and VIS spectroscopy of neutral molecules. However, the benefits of both mass selected UV spectra and UV-resonance selective mass spectra were soon recognized.

In **Figure 3** (right-hand side) the resonance ionization spectra recorded mass selectively (at 106 and 107 amu) show the electronic origin of the $S_1 \leftarrow S_0$ transition of *p*-xylene (106 amu) and its $^{13}C_1{}^{12}C_7H_{10}$ natural isotopomer. An isotopic blueshift of +2.7 cm$^{-1}$ (corresponding to 0.02 nm) is found. On the left-hand side of **Figure 3** two UV-resonance selected mass spectra are shown, resulting from resonance enhanced multiphoton ionization at 272.17 nm and 272.14 nm. The latter is the band centre of the 107 amu selective laser spectrum and clearly gives rise to a strong relative enhancement of the heavier mass in the mass spectrum.

An extraordinary impact on molecular UV spectroscopy (in particular of large molecules) has been brought about by supersonic beam cooling of

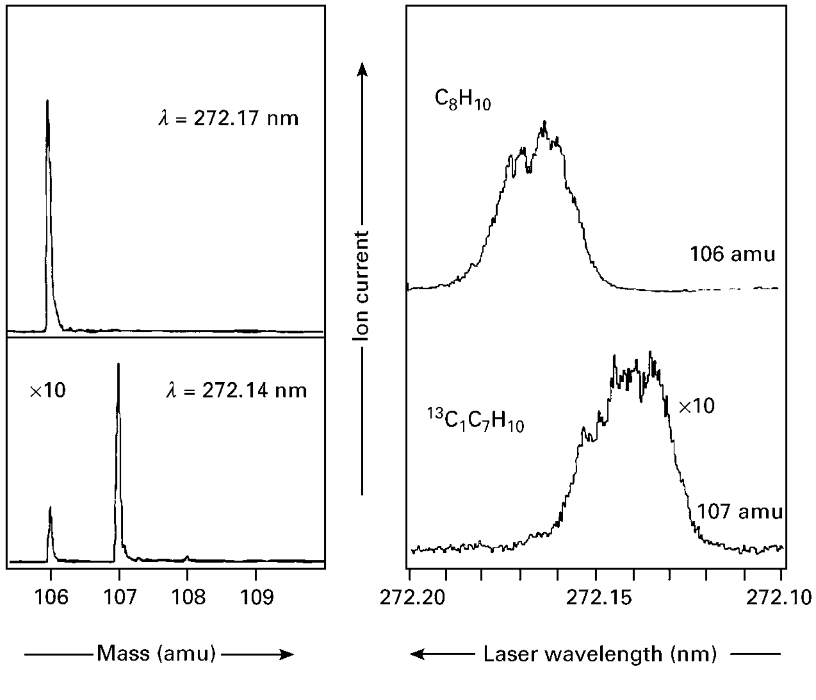



**Figure 3** Two options for resonance enhanced multiphoton ionization concerning ion source or spectroscopy: UV-resonance selected mass spectra (on the left) and mass selected molecular UV spectra (on the right) for *p*-xylene and its $^{13}C_1$-isotopomers.

internal molecular motions. The narrow line width in **Figure 3** is due to this effect and allows isotopomer-selective ionization as illustrated in **Figure 3**. However, even for large involatile molecules well-resolved spectra with vibrational fine structure can be observed. To get these molecules into the gas phase, high temperatures or laser desorption were necessary. **Figure 4** shows the spectra of heated gas samples containing dibenzo-dioxins seeded in argon gas and cooled in a supersonic molecular beam. For dibenzodioxins only their UV spectra in solution were known before the application of mass selective resonance ionization spectroscopy. On the left-hand side of **Figure 4**, the cold spectrum of unsubstituted dibenzodioxin is represented. Single vibronic bands of the $S_1 \leftarrow S_0$ transition appear, revealing low frequency 'butterfly' and torsional modes excited in $S_1$. This and further arguments suggest a slight non-planarity of dibenzodioxin in the $S_0$ and planarity of its $S_1$ and ionic ground state.

The combination of supersonic beams with mass selective spectroscopy has a further advantage. While mass selection allows discrimination against impurities, fragments and differently substituted (e.g. halogen, isotope) species, the high spectroscopic resolution enables the discrimination of structural isomers having the same mass. **Figure 4** clearly shows the significant difference of spectra of dibenzodioxins with two chlorine atoms substituted at symmetrically equivalent positions (i.e. the 2, 3, 7, and 8 positions). They differ only by their position relative to each other [e.g. neighbouring (bottom right) or not]. Wavelengths can easily be found (e.g. between 32 790 and 32 805 cm$^{-1}$) where these isomers may be ionized selectively. Spectroscopic studies of many different substituted aromatics have been performed. Effects such as increasing ionization energies with increasing degree of halogenation, decreasing energy of the $S_1$ state (e.g. below half the ionization energy) with increasing molecular size and strongly differing lifetimes of the intermediate states have been experimentally observed or theoretically determined, and of course, these effects have to be considered for efficient ionization schemes.

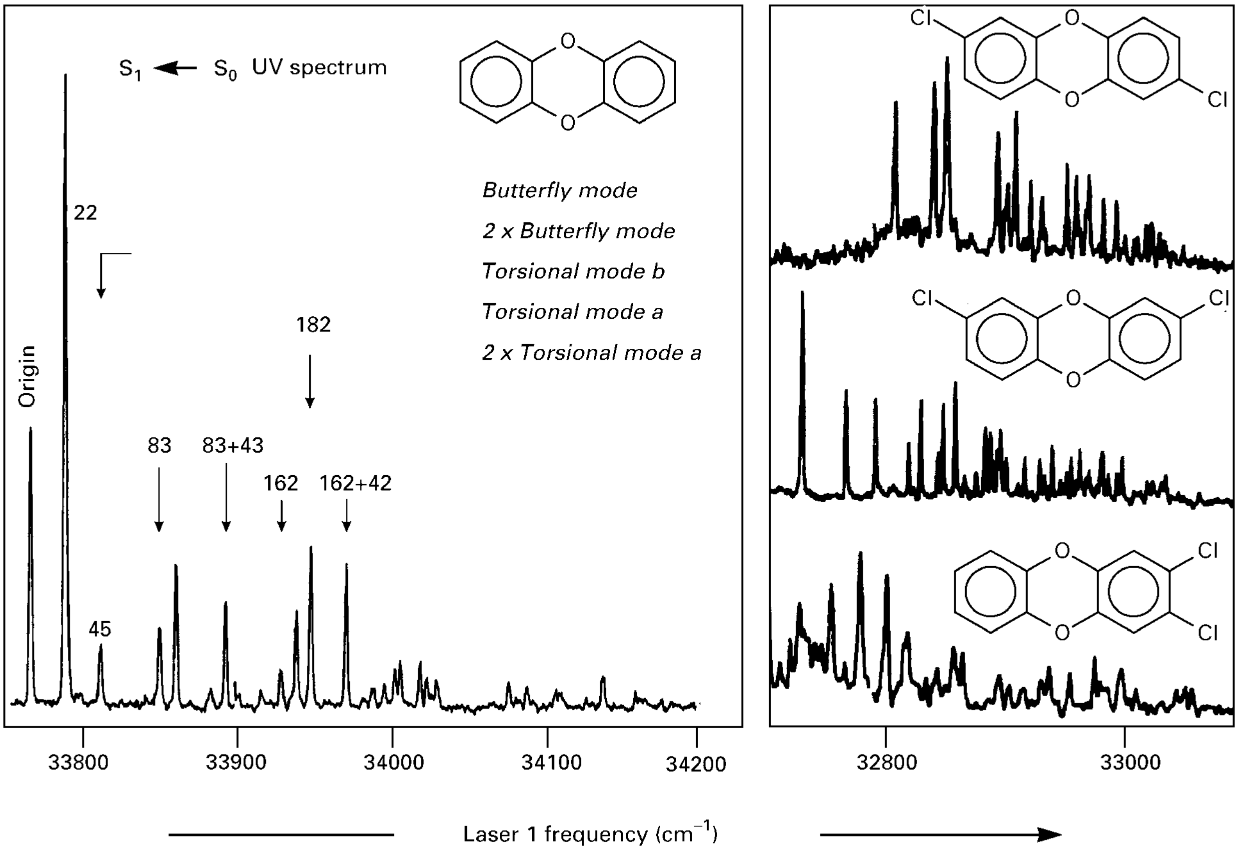



**Figure 4** On the left: the cold (supersonic beam cooled) UV spectrum of dibenzo-dioxin measured via multiphoton ionization. On the right: The dichloro-congeners (with chlorine at the 3, 8, the 2, 8, and the 2, 3 positions) show similar well-structured spectra, giving wavelengths for congener selective ionization.

## Multiphoton dissociation: tuneable fragmentation ion source

Species-selective ionization is one quality of multiphoton excitation in mass spectrometry. Another is the option to tune the degree of fragmentation from exceptionally low to very strong. The different possibilities of this feature are illustrated by the example of benzene in **Figure 5**. The variation of fragmentation is obtained by changing the laser intensity from $10^7$ W cm$^{-2}$ (corresponding to a UV-laser pulse with a 10 µJ pulse energy, a 10 ns pulse length and a 100 µm focus diameter) to $10^9$ W cm$^{-2}$. While the first mass spectrum is nearly free of fragments, the last mass spectrum consists mainly of carbon ions at mass 12. Both extremes cannot be achieved by conventional electron ionization or other ion sources and are unique for multiphoton excitation. It should be mentioned that these mass spectra have been taken with a short TOF mass spectrometer consisting only of three electrodes and one ion detector.

The reason for the astonishing feature of tuneable fragmentation is the specific excitation mechanism of multiphoton absorption in comparison with electron ionization which is a one-step process (the total energy is deposited in the neutral molecule in one step, causing ionization and fragmentation). Multiphoton ionization is due to consecutive absorption steps which are interrupted if fast decay channels are reached and are continued within the decay products (so-called 'ladder-switching'). Increasing the intensity increases the absorption probability (less time per absorption step), allowing the whole process to reach higher steps of the 'ladder of fragment ions' during the laser pulse. This process is indicated schematically in **Figure 5**. However, if very intense and very short laser pulses are used (i.e. femtosecond lasers), absorption is faster than decay and even ionization process and the total absorbed multiphoton energy is deposited in the molecule as a single event.

Since fragmentation is a rich source of information in mass spectrometry, this feature of multiphoton

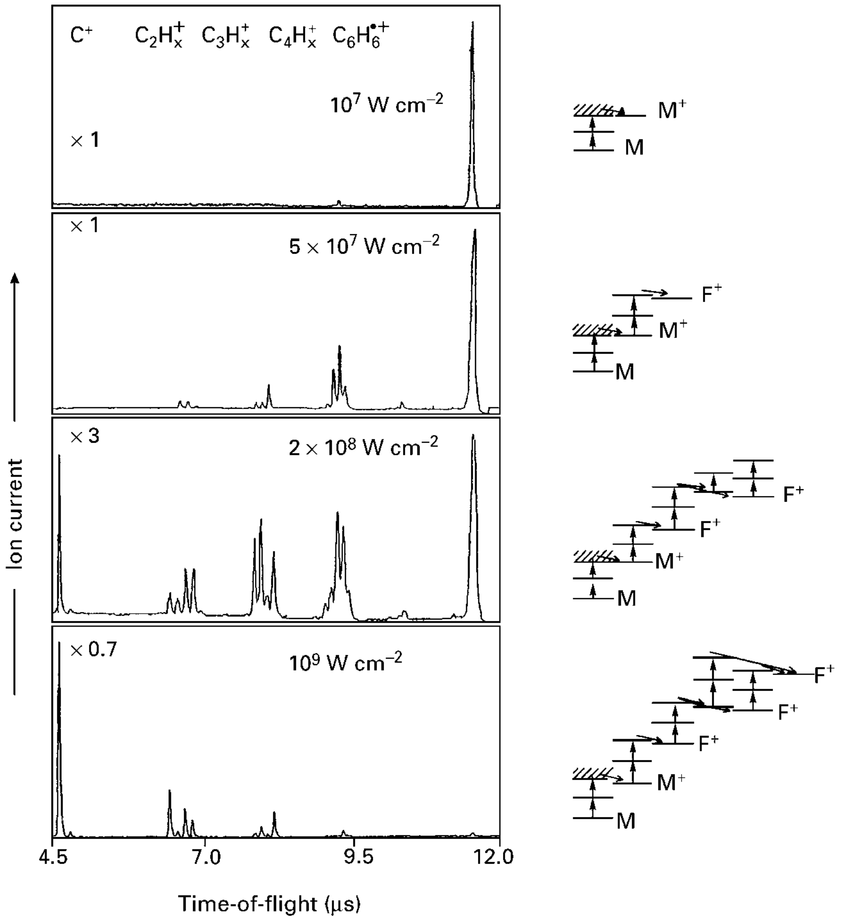



**Figure 5** Multiphoton dissociation of benzene at different laser intensities, inducing different degrees of fragmentation (from fragmentation-free at the top to dominating atomic fragment ions at the bottom of the figure). On the right of the spectra the model of 'ladder switching' is schematically displayed and explains the large variation of fragmentation by multiphoton absorption processes.

excitation may give a new impetus to this field of research. For example, increasing the laser intensity may allow one to tune for specific decay channels (e.g. breakage of bonds to functional groups, typical substituents or bonds within the molecular framework) and thus supply a low level but inexpensive type of tandem mass spectrometry. However, multiphoton excitation also allows access to the sensitive tool of metastable decay, the most informative source of fragmentation in mass spectrometry. While in conventional mass spectrometry considerable effort is necessary to study tandem mass spectra (mass separation before secondary collision-induced excitation) multiphoton excitation may create controlled ion decay dissociation (e.g. by enhancing specific metastable decay channels) very efficiently already in the ion source. This is illustrated and explained in **Figure 6**.

Here, part of a multiphoton ionization mass spectrum in the region around the benzene radical cation is represented. While the 'linear TOF' spectrum only reveals the molecular ion and its $^{13}C_1$-isotopomer, the 'reflectron TOF' spectrum additionally shows the fragment ions $C_6H_5^+$ and $C_6H_4^{\bullet+}$ although the ion source conditions were identical. (The difference in mass resolution will not be discussed here, see the Further reading section and other articles in this Encyclopedia.)

To understand this effect one should remember that product ions formed in the field-free drift region of a linear TOF instrument cannot be distinguished from their parent ion (same ion velocity). A reflectron TOF mass analyser, however, corrects for ion kinetic energy deviations and shows all fragment ions of the same mass independent of their ion kinetic energy (at least for an energy mismatch of ± 10%) at the same flight time. Thus **Figure 6** demonstrates the existence and high yield of metastable decay induced by multiphoton excitation (in **Figure 5** the total fragmentation yield was small, main peak was $C_6H_6^{\bullet+}$). An explanation for this is given at the top of **Figure 6**. By resonance enhanced multiphoton ionization, molecular ions $M^{\bullet+}$ with small internal energy and a narrow energy distribution $\Delta E$ are produced (soft ionization, no absorption in the ion). The absorption of a further photon in the molecular ion $M^{\bullet+}$ will then result in excited ions which still have a narrow energy distribution $\Delta E^*$. If the photon energy is in the range of the energy interval between the ion ground state and the threshold of a decay channel then the overlap of the metastable decay region (near $F^+$ decay threshold) and internal energy distribution $\Delta E^*$ will be large. In contrast to multiphoton excitation, electron ionization is usually performed with 70 eV electrons,

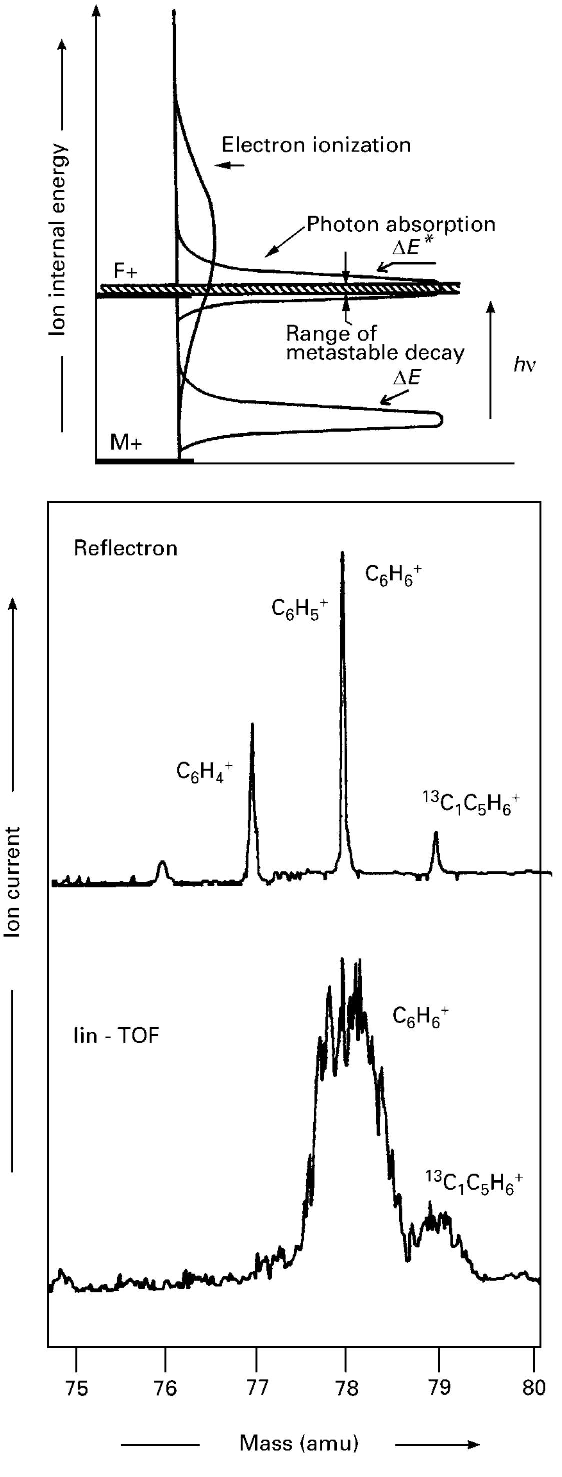



**Figure 6** Experimental demonstration (bottom) and schematic explanation (top) of the high yield of metastable decay products available with multiphoton dissociation.

resulting in a broad internal energy distribution whose overlap with the metastable decay range is relatively small (few metastable decay products). The special features of multiphoton excitation

concerning metastable decay have initiated a renaissance of ion kinetic decay studies.

Instead of a single colour excitation (laser L1, **Figure 2**) a second laser (L2 or L3) can be used. Secondary multiphoton excitation is then possible, which can be performed with different laser wavelengths, at different positions and different times from that of the ionizing laser L1. Thus, selective secondary excitation of a group of fragments (e.g. with L2 in ion source II) or even of single product ions with one well-defined mass (e.g. with L3 in the SF) is possible. This already is an approach to highly sophisticated tandem mass spectrometry with the benefit of well-defined internal ion energies and a high yield of secondary fragmentation. In **Figure 7**, a primary multiphoton mass spectrum of benzene and secondary mass spectra of four different fragment ions are shown. This experiment has been performed to elucidate multiphoton fragmentation and is shown here to illustrate its tandem mass spectrometry option. The primary mass spectrum is due to a laser wavelength of $\lambda_1 = 258$ nm (absorption maximum of benzene), while the secondary mass spectra of $C_4^{\bullet+}$, $C_4H^+$ and $C_4H_2^{\bullet+}$ have been performed with $\lambda_3 = 355$ nm. The $C_4H_3+$ fragment does not show fragmentation at $\lambda_3 = 355$ nm and has been excited

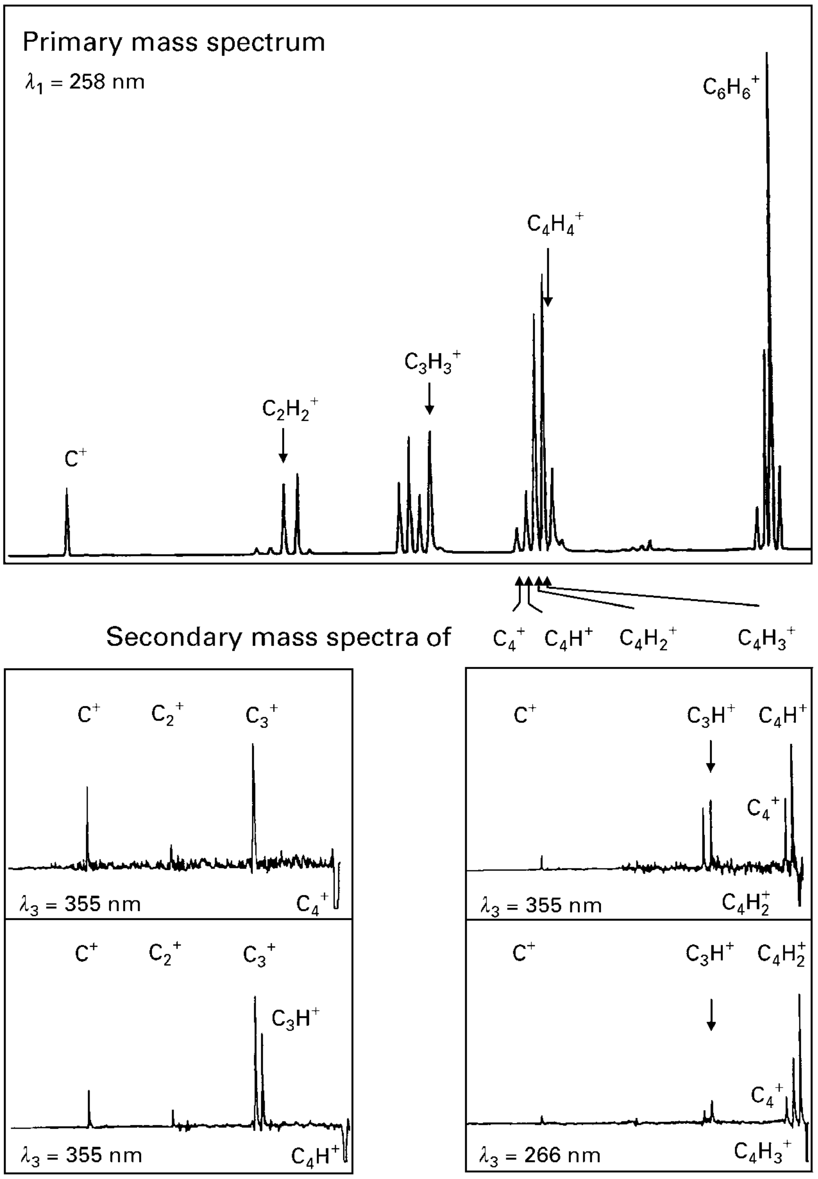



**Figure 7** Primary and secondary mass spectra of benzene, illustrating the options of multiphoton excitation for tandem mass spectrometry.

with $\lambda_3 = 266$ nm. Furthermore, $C_4H^+$ does not exhibit a secondary $C_4^{\bullet+}$ product ion as $C_4H_2^{\bullet+}$ and $C_4H_3^+$ although it dissociates at 355 nm. These effects already indicate a complex multiphoton fragmentation tree with several decay channels. It is not the task of this article to go into more detail of that specific process, rather this example should illustrate the available options of multiphoton excitation for tandem mass spectrometry.

## Multiphoton excitation for mass selective ion spectroscopy

There are several options for the use of multiphoton excitation for ion spectroscopy. One is resonance enhanced photoelectron spectroscopy. In contrast with conventional photoelectron spectroscopy, where single VUV photons are used, a neutral excited intermediate state is involved which allows the application of the low energy UV-photons available with lasers. Owing to small excess energies, the ionic ground state is most often studied by this technique. The involvement of different vibronic intermediate states results in a variation of the photoelectron spectrum and is a valuable means for obtaining additional information about the character of ionic and even of neutral intermediate states. Resonance enhanced photoelectron spectroscopy cannot be performed in a truly mass selective way. Either the neutral source consists only of one sort of molecule (which can be checked by synchronously observing the produced molecular ions) or photo-ion photoelectron coincidence spectroscopy is performed. But the high-resolution version of resonance enhanced photoelectron spectroscopy, ZEKE spectroscopy, can be performed in a true mass selective mode. Both methods are beyond the scope of this article (for further information see the Further reading section).

In addition to the study of the ionic ground state, resonance enhanced photoelectron spectroscopy allows one to characterize multiphoton ion sources with respect to the internal energy distribution of the produced ions. For favourable cases it even may supply information as to how to achieve state selective ion production. The example of fluorobenzene has been chosen in **Figure 8** to illustrate resonance enhanced photoelectron spectroscopy as well as further types of ion spectroscopy involving multiphoton excitation.

At the bottom of **Figure 8** the resonance ionization spectrum of fluorobenzene (cooled in a supersonic beam) is shown, exhibiting the electronic origin $0^0{}_0$ as well as other vibronic transitions (e.g. excitation of one quantum of the 6b-vibration). In

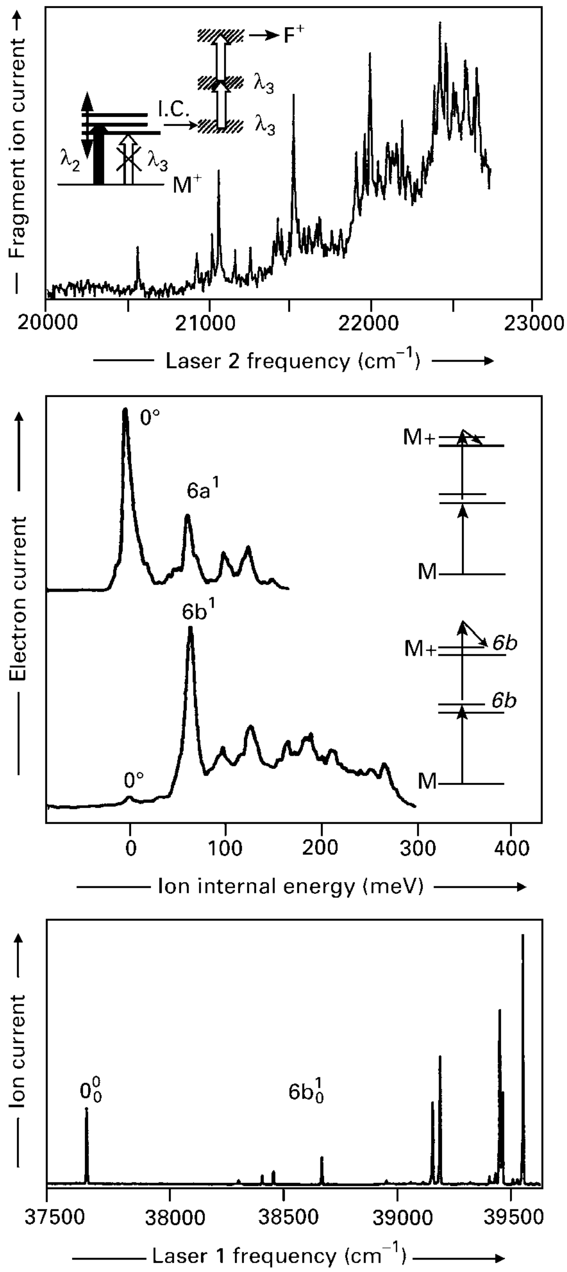



**Figure 8** Multiphoton excitation and ion spectroscopy: Bottom spectrum: cold, mass selected UV spectrum ($S_1 \leftarrow S_0$ transition) of fluorobenzene, providing wavelengths for efficient and selective ionization. Middle spectra: photoelectron spectra induced by UV-resonance enhanced two-photon absorption. Choosing different intermediate states [$S_1(0,0)$ or $S_1(6b^1)$] results in different populations of the final fluorobenzene radical cations. Top spectrum: spectroscopy of the excited ionic state of the fluorobenzene radical cation measured by multiphoton dissociation spectroscopy. The ions have been prepared via the neutral $0^0{}_0$ transition. A special excitation scheme has been used to optimize cation spectroscopy (for further details see text).

the middle of **Figure 8**, two photoelectron spectra are presented which involve different neutral intermediate states. A significant difference is observed,

indicating that the excitation of the vibrationless $S_1$ or of the 6b-mode in the $S_1$ induces different populations in the molecular ion. This supports assignments in the photoelectron spectrum, and allows production of state-selected fluorobenzene ions for further experiments (ion kinetics, ion spectroscopy).

Such a secondary experiment is displayed at the top of **Figure 8**. After production of fluorobenzene radical cations via resonance enhanced multiphoton ionization, preferentially in the ionic ground state, multiphoton dissociation spectroscopy has been applied. However, in contrast to **Figure 1D**, an even more elaborate excitation scheme has been used here, exploiting the experimental options of TOF mass spectrometry. Laser L2 (ion source II) is used solely for spectroscopy; laser L3 (SF) is responsible for detection by multiphoton dissociation. The benefit of this scheme is that laser L2 can be optimized for spectroscopy while dissociation by laser L3 is highly mass selective, efficiently suppressing background signals from other molecular or fragment ions. Using modern double-pulse lasers, the requirements for the laser system and mass analyser are still reasonable. The result is a very well-resolved ion spectrum. One should keep in mind that in the case of fluorobenzene, as for many other halogen-substituted and the non-substituted benzene cations, fluorescence spectroscopy is not possible (to say nothing about absorption spectroscopy) and only low-resolution conventional photoelectron spectra exist.

When treating ion spectroscopy one should not forget anions. Similar spectroscopic techniques may be used as for cation spectroscopy. For instance dissociation spectroscopy is also possible for molecular anions. Since excited anionic electronic states mostly do not exist, one uses infrared multiphoton dissociation to study vibrational levels of the ground state. Another interesting technique is the photoelectron spectroscopy of anions (photodetachment photoelectron spectroscopy), which exhibit a very specific feature. This technique differs from cation ← neutral photoelectron spectroscopy in two respects: (i) the final state is a neutral one; thus anion photoelectron spectroscopy delivers information about neutrals rather than ionic systems. (ii) The initial state is anionic; thus mass selection before spectroscopy is possible. As a result, mass selective spectroscopic information of neutral molecular systems is supplied which otherwise is not accessible. This is of particular interest for neutral systems which are only available in complex mixtures or are short-lived intermediate reaction products or radicals.

One example is shown in **Figure 9**. Metal–carbon–hydrogen complexes are of great importance as intermediates in heterogeneous catalytic reactions of hydrocarbons on metal surfaces, but are not available as pure samples. At the bottom of **Figure 9** an anion mass spectrum of such complexes is presented. These complexes have been formed by a gas phase reaction in the high density region of a supersonic molecular beam (acetylene seeded in argon). Iron atoms as well as low-energy electrons are supplied by laser ablation from an iron wire which is positioned as near as possible to the nozzle of the supersonic beam valve. Clearly, $FeC_iH_n^-$ complexes appear in the spectrum together with carbon and carbon hydride clusters. Mass selected anion photoelectron spectra which supply information about the neutral complexes $FeC_2$, $FeC_2H$ and $FeC_2H_2$ are presented at the top of **Figure 9**. Obviously, an increasing hydrogen content of the complex results (i) in a decreasing electron affinity (transition to state X), (ii) in a smaller change of molecular structure (Fe–C distance) between anion and neutral (Franck–Condon envelope of the Fe–C stretching frequency) and (iii) a strong decrease of energy gaps between higher electronic states (X, A, B, C) (a preliminary assignment in the case of $FeC_2H_2$). Although anion photoelectron spectroscopy is mostly owing to a one-photon process, it has been considered here to show the large variety of experimental possibilities when combining laser excitation and mass spectrometry.

## Combination of multiphoton excitation with other neutral sources: laser desorption

Since multiphoton excitation in mass spectrometry takes place in the more or less tight laser focus, which can easily be shifted in space and time or be subject to other variations, it can be combined with different ion optical or mechanical arrangements (e.g. sources of neutral molecular systems) without the need for much additional hardware. Thus, by combination with chromatography (particularly gas chromatography), species selection has successfully been realized. Another very promising combination, which has frequently been applied in the recent past for the study of involatile molecules (e.g. polycyclic aromatics, biomolecules), is that of laser desorption of neutral molecules and resonance enhanced multiphoton ionization. All the benefits of multiphoton mass spectrometry, such as soft ionization, selective ionization, controllable fragmentation or secondary excitation for tandem mass spectrometry, may be used in this field.

In **Figure 10**, the example of a biologically relevant molecule is displayed. Porphyrins containing central metal atoms (e.g. Mg or Fe) represent the

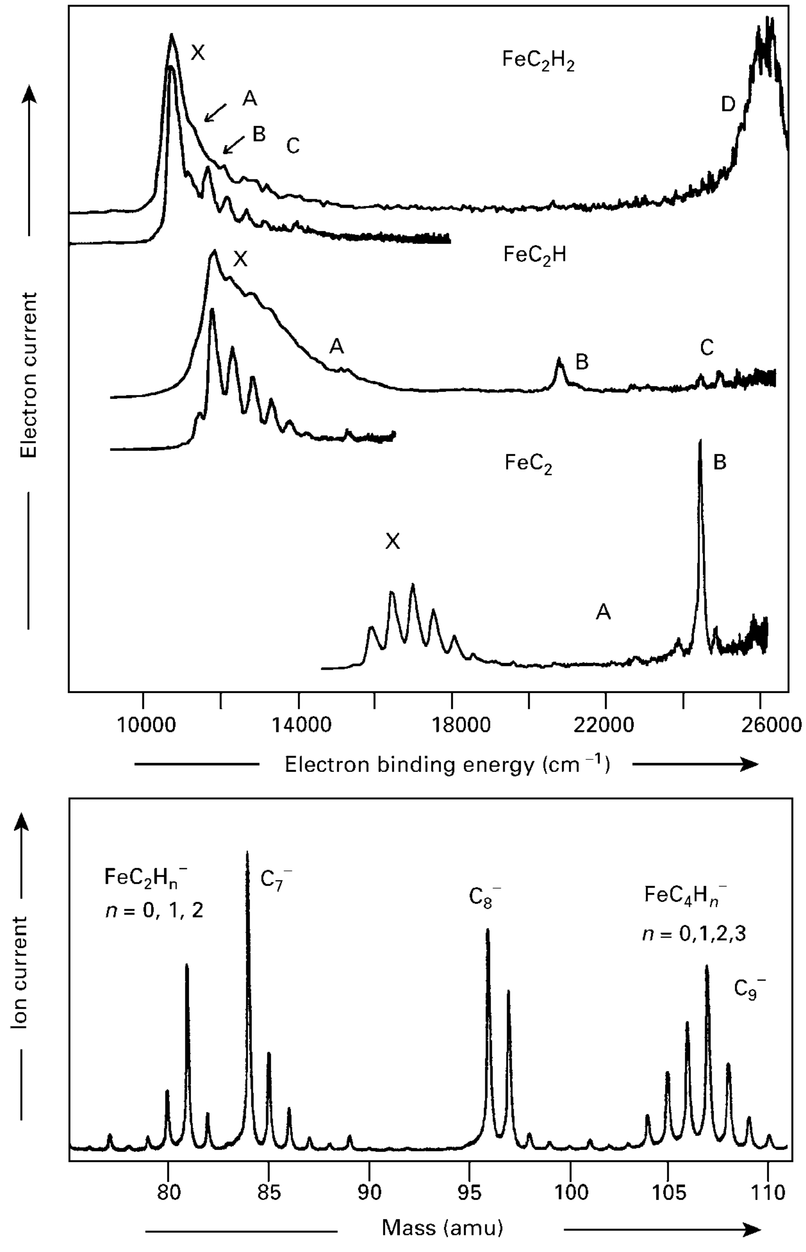



**Figure 9** Anion photoelectron spectroscopy. Its unique features are (i) intrinsic mass selectivity and (ii) neutrals as final states. Here, as an example the results for compounds of iron, carbon and hydrogen are shown which exist in catalytic processes, high-temperature terrestrial or low-temperature astrophysical chemistry. Bottom spectrum: a primary anion mass spectrum containing anions of the complexes of interest. Top spectra: anion photoelectron spectra obtained by electron kinetic energy analysis after laser-induced photodetachment. They reveal the change of molecular structure and electronic energies for increasing numbers of hydrogen atoms in the complex.

molecular frame of important biomolecules such as chlorophyll or haemoglobin which mainly differ in their central metal atom and specific side-chains. In **Figure 10**, the TOF mass spectrum of zinc meso-porphyrin-IX dimethyl ester is shown, revealing a small degree of fragmentation. Inset A displays the molecular ion on an expanded mass scale and nicely reveals the typical isotopic pattern of Zn convoluted with that of $^{13}C$-isotopomers (at the used wavelength of $\lambda_1 = 280$ nm and the internal molecular temperature of $\geq 300$ K, ionization is not isotope selective). The fragment peaks near 155 µs consist of a narrow peak ($M^+ - 64$) and a group of peaks ($>M^+ - 73$). The latter results from the loss of a side-chain (Zn-

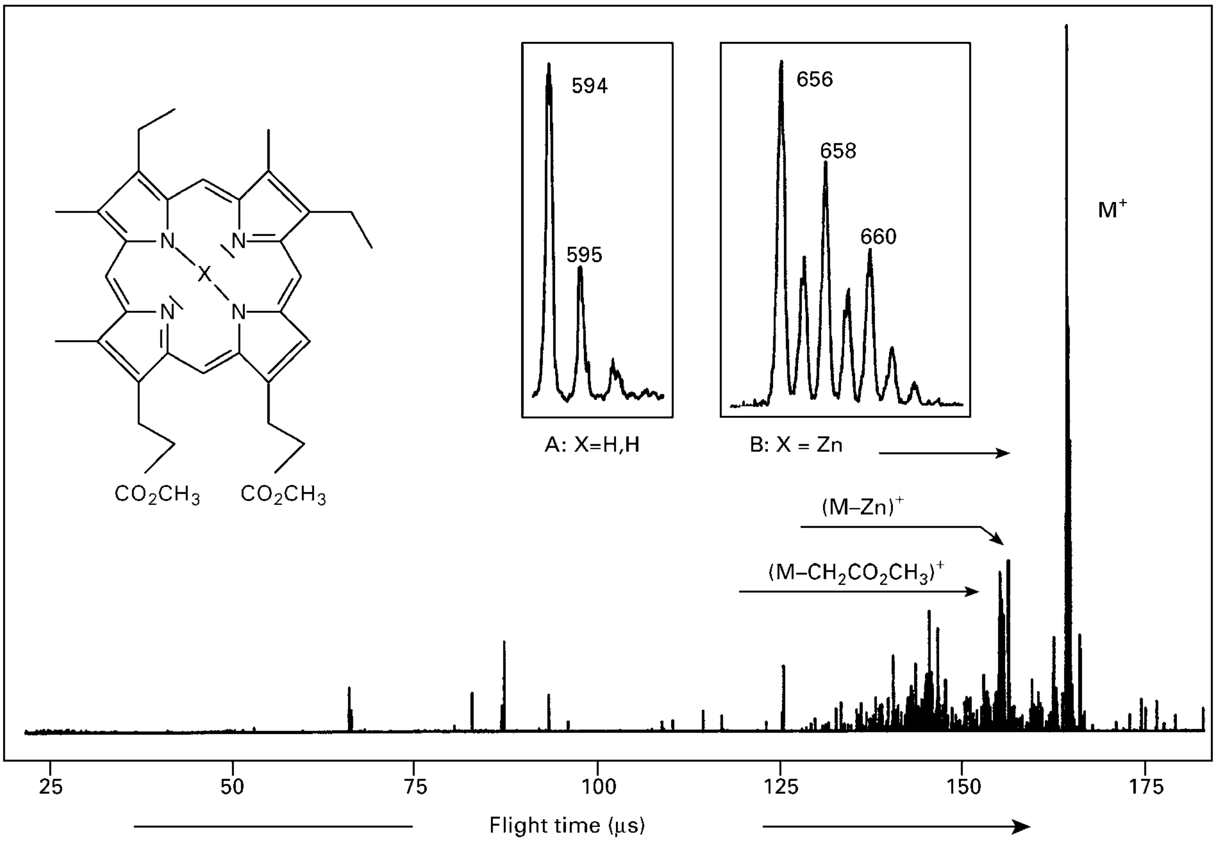



**Figure 10** The combination of neutral laser desorption and resonance enhanced multiphoton ionization. A TOF mass spectrum of zinc mesoporphyrin-IX dimethyl ester is presented. Inset *A*: the isotropic pattern (zinc, carbon) of the molecular ion. Inset *B*: the molecular ion of mesoporphyrin-IX dimethyl ester (TOF mass spectrum not shown) which displays the same isotope pattern as peak (M–64) due to loss of a zinc atom.

isotope pattern preserved) while the former is owing to the loss of the central Zn atom (no Zn-isotope pattern remains). It shows a nearly identical isotope pattern (owing to $^{12}C$–$^{13}C$ distribution) to the mesoporphyrin-IX dimethyl ester without a central atom, whose molecular ion is shown in inset B of **Figure 10**.

## Multiphoton mass spectrometry for chemical trace analysis

Both the high species selectivity and the soft ionization make resonance enhanced multiphoton excitation an excellent ion source for chemical trace analysis. In addition, the use of pulsed lasers combined with TOF mass spectrometry (also a pulsed technique) enables the recording of single mass spectra within a few milliseconds, depending on the repetition frequency of the laser pulses (typically 20 to 50 Hz). The unification of high speed, selectivity and sensitivity indicate that multiphoton mass spectrometry will become firmly established in the highly sophisticated analytical technology of the future. In **Figure 11**, the variety of applications is illustrated by examples from different fields, namely raw gas analysis of a waste incinerator, research for health care, production integrated analysis in the food industry and exhaust trace analysis of motor cars.

In **Figure 11A** (incinerator), traces of naphthalene (sub-ppb range) in the raw gas of a pilot waste incinerator have been recorded over 900 s. During this time a major failure of the combustion process (at 500 s) which gave rise to an increased concentration of polycyclic aromatics by many orders of magnitude has been studied. This event was preceded by small but significant fluctuations of the napthalene concentration, which were only detectable by a fast, but nevertheless selective and sensitive, method. Effects like these could be used in future for online control of incinerator plants. This could help to improve pollutant emission by reducing their formation rather than retaining them after the combustion process by expensive cleaning procedures. An example of

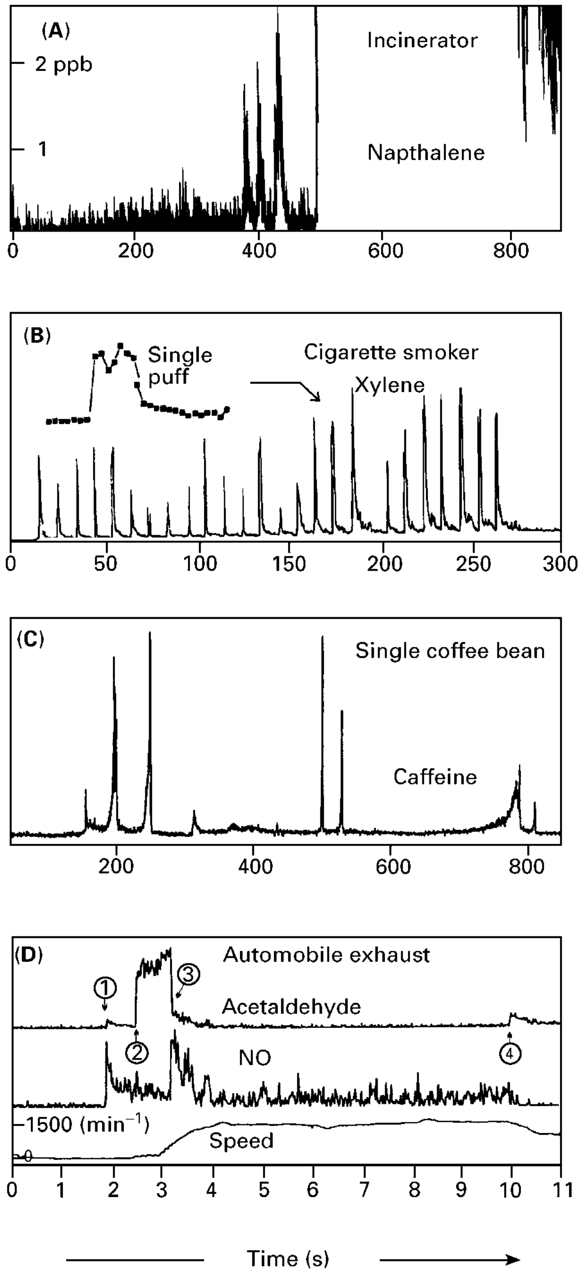



**Figure 11** Several examples of applications of multiphoton ionization mass spectrometry for chemical analysis. (A) Incinerator: traces of naphthalene have been recorded in the raw gas of a pilot waste incinerator. Here a major failure of the combustion and preceding fluctuations of the naphthalene concentration were studied. (B) Cigarette-smoker: the xylene concentration in the mouth space of a cigarette-smoker. (C) Single coffee bean: a single coffee bean has been heated to untypically high temperatures. Single short ejection events of caffeine have been observed due to distinct cracks of the plant skin of the coffee bean. (D) Automobile exhaust: acetaldehyde due to incomplete combustion and NO (complete combustion) during the starting of a cold engine measured in the exhaust. Different behaviour is observed when the ignition is starting (1), the fuel mixture becomes rich (2) and lean (3) and the speed of idle running is reduced to its regular frequency (4).

even more direct health care research is shown in **Figure 11B** (cigarette smoker). Here the xylene content in air recorded in the mouth space of a cigarette smoker is recorded. The 50-fold expansion of a single puff demonstrates the high temporal resolution of less than 0.1 s.

The next spectrum shows fluctuations of caffeine concentration in the headspace above one single green coffee bean which has been suddenly heated up to high temperatures (far above typical roasting temperatures). The very short spikes of caffeine emission are induced by $CO_2$ (pyrolysis product) causing distinct cracks of the plant skin. At the single ejection events solvated caffeine is swept along with the $CO_2$. Of course, other molecular components (e.g. those which are used as indicators for the progress of coffee roasting) can be monitored thus enabling online control of the roasting process.

In **Figure 11** (automobile), the behaviour of acetaldehyde (a dehydrogenation product of ethanol which is typical for incomplete combustion) and nitric oxide in the exhaust of a motorcar is recorded during the first seconds of a cold engine start. Unleaded gasoline containing 10% of ethanol has been used as fuel. Both pollutants strongly support the formation of ozone in the troposphere (e.g. smog in large cities). Their reduction, e.g. by an electronically improved regulation of the motor, is therefore of major interest. Thus, the highly dynamic processes in combustion engines must be studied, but appropriate fast analytical techniques are not yet commercially available. In **Figure 11D** five different situations initiated by distinct fast events of the motor regulation can be recognized. From 0 to 2 s, no combustion is active and none of the two pollutants are emitted. At time (1) ignition starts, causing a high peak of the combustion product NO which drops off to a level of some 100 ppm. At time (2), a fuel-rich mixture is injected for a better start of the engine. A sudden increase of acetaldehyde is observed, indicating incomplete combustion due to a shortage of oxygen. At time (3) the speed of idle running is reached (faster in a newly started engine for stabilization of the combustion) and the fuel is changed to a lean mixture. Acetaldehyde, the product of incomplete combustion drops instantly, while NO, the product of complete combustion, increases strongly again. However, heavy fluctuations of NO still indicate some instability of the whole process. After stabilization (NO level now at 80 to 100 ppm) the idle running frequency is reduced to the regular speed at time (4). The NO level drops significantly, but a slight increase and then a gradual decrease of acetaldehyde is observed. This is consistent with a reduced combustion temperature at lower

speed and desorption of fuel from the walls of the injection system for about half a second.

## List of symbols

$R_{50\%}$ = mass resolution; $t$ = time; $\Delta E$ = energy distribution; $\Delta E^*$ = internal energy distribution, $\lambda$ = wavelength.

*See also:* **Environmental Applications of Electronic Spectroscopy; Fragmentation in Mass Spectrometry; Ion Dissociation Kinetics, Mass Spectrometry; Ion Energetics in Mass Spectrometry; Laser Applications in Electronic Spectroscopy; Metastable Ions; Multiphoton Spectroscopy, Applications; Negative Ion Mass Spectrometry, Methods; Photoelectron–Photoion Coincidence Methods in Mass Spectrometry (PEPICO); Photoionization and Photodissociation Methods in Mass Spectrometry; Spectroscopy of Ions; Time of Flight Mass Spectrometers.**

## Further reading


Baer T (ed) (1996) Ion spectroscopy. *International Journal of Mass Spectrometry and Ion Processes (Special Edition)* **159**: 1–261.

Boesl U (1991) Multiphoton excitation and mass selective ion detection for neutral and ion spectroscopy. *Journal of Physical Chemistry* **95**: 2949–2962.

Boesl U, Weinkauf R and Schlag EW (1992) Reflectron time-of-flight mass spectrometry and laser excitation for the analysis of neutrals, ionized molecules and secondary fragments. *International Journal of Mass Spectrometry. Ion Processes (Special Edition)* **112**: 121–166.

Boesl U, Weinkauf R, Weickardt C and Schlag EW (1994) Laser ion sources for time-of-flight mass spectrometry. *International Journal of Mass Spectrometry. Ion Processes (Special Edition)* **131**: 87–124.

Boesl U, Heger HJ, Zimmermann R, Püffel PK and Nagel H (2000) Laser mass spectrometry in trace analysis. Submitted to *Encyclopedia of Chemical Analysis*. John Wiley & Sons. In press.

Gobeli DY, Yang JJ and El-Sayed MA (1985) Laser multiphoton ionization-dissociation mass spectrometry. *Chemical Reviews* **85**: 529–554.

Grotemeyer J, Boesl U, Walter K and Schlag EW (1986) A general soft ionization method for mass spectrometry: resonance-enhanced multiphoton ionization of biomolecules. *Organic Mass Spectrometry* **21**: 645–653.

Hayes JM (1987) Analytical spectroscopy in supersonic expansions. *Chemical Reviews* **87**: 745–760.

Kimura K (1987) Molecular dynamic photoelectron spectroscopy using resonant multiphoton ionization for photophysics and photochemistry. *International Reviews in Physical Chemistry* **6**: 195–226.

Letokhov VS (ed) (1985) *Laser Analytical Spectrochemistry*. Bristol: Adam Hilger.

Letokhov VS (1987) *Laser Photoionization Spectroscopy*. Orlando: Academic Press.

Lubman DM (ed) (1990) *Lasers and Mass Spectrometry*. New York: Oxford University Press.

Neusser HJ (1989) Lifetimes of energy and angular momentum selected ions. *Journal of Physical Chemistry* **93**: 3897–3907.

Vertes A, Gijbels R and Adams F (eds) (1993) *Laser Ionization Mass Analysis*. New York: John Wiley & Sons.

Zenobi R (1995) In situ analysis of surfaces and mixtures by laser desorption mass spectrometry. *International Journal of Mass Spectrometry Ion Processes* **145**: 51–77.

Zimmermann R, Lenoir D, Kettrup A, Nagel H and Boesl U (1996) On-line emission control of combustion processes by laser-induced resonance-enhanced multiphoton ionization mass spectrometry. *Twenty Sixth Symposium (International) on Combustion*, pp 2859–2868 Pittsburgh: The Combustion Institute.


# Multiphoton Spectroscopy, Applications


**Michael NR Ashfold** and **Colin M Western**,
University of Bristol, UK


Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Applications**

Multiphoton spectroscopy involves the excitation of an atom or molecule from one electronic state A to another B by absorption of two or more photons in contrast to more conventional spectroscopies that involve just a single photon. We identify two distinct variants of multiphoton spectroscopy. **Figure 1A** depicts a sequential multiphoton absorption, i.e. an incoherent multiphoton excitation proceeding via successive allowed one-photon transitions:

$$\mathrm{A} \xrightarrow{h\nu_1} \mathrm{M} \xrightarrow{h\nu_2} \mathrm{B} \qquad [1]$$

where the two photons may have the same (or, more generally, different) frequencies. Such schemes provide the basis for a wide range of double-resonance spectroscopies. By way of contrast, **Figure 1B** illustrates the simultaneous (i.e. coherent) absorption of two photons without the involvement of any resonant intermediate state M. The realization that an atom or molecule could undergo such coherent multiphoton excitation dates back to the early days of quantum mechanics, but the experimental demonstration of such multiphoton excitations had to await the advent of the laser to provide the necessary very high light intensities required to compensate for the inherently low multiphoton transition cross sections. With a conventional light source, a two-photon excitation might be $10^{20}$ times less probable than a one-photon absorption, but the $I^n$ intensity dependence combined with the $> 10^{12}$ increase in power density available with lasers makes such transitions relatively straightforward to observe with modern pulsed lasers. The fraction of photons absorbed will always be small, so identifying that a multiphoton excitation has occurred almost always involves monitoring some consequence of the multiphoton excitation rather than observing the absorption itself. Three of the more common consequences are depicted in **Figure 1B**. If the excited state B fluoresces, the multiphoton excitation spectrum can be obtained simply by monitoring the (multiphoton) laser-induced fluorescence (LIF) as a function of excitation wavelength. Given the high light intensities required to drive the multiphoton absorption step, however, it will generally be the case that some of the molecules excited to state B will absorb one (or more) additional photons and ionize. This is termed resonance-enhanced multiphoton ionization (REMPI), and is the most widely used method for detecting multiphoton absorption by gas-phase species. The third possible process for the excited state, dissociation (or any other loss mechanism), will reduce the photons or ions detected, and is a potential limitation that is discussed further below. The other important condition is that the B ← A two-photon (or multi-photon) excitation has a nonzero transition probability; the selection rules depend on the number of photons and the differences between one- and two-photon transitions have many analogues with those distinguishing infrared and Raman vibrational spectroscopy. If the selection rules are satisfied, then the spectrum obtained by measuring the ion yield (or the yield of the accompanying photoelectrons) as a function of excitation wavelength will provide a signature of the B ← A two-photon transition of the neutral molecule; analysis can provide structural (and, in some cases, dynamical) information about the excited state B.

## Multiphoton selection rules

As in one-photon spectroscopies, symmetry is crucial in determining multiphoton transition probabilities. A multiphoton transition between two states A and B is 'allowed' if the transition moment $\langle A | T_q^k(\hat{O}) | B \rangle$ is nonzero; i.e. if the product of the irreducible representations for the wavefunctions of state A and B and that of $T_q^k(\hat{O})$—the $q$th component of the spherical tensor of rank $k$ representing the multiphoton transition operator $\hat{O}$—contains the totally symmetric representation. Symmetry considerations ensure that only spherical tensors of either odd or even rank will contribute to any one-colour multiphoton excitation. Thus, for example, whereas one-photon electric dipole transitions must be carried by components of rank 1, only components of rank $k = 0$ and/or $k = 2$ can contribute to two-photon transitions brought about using photons of identical frequency and polarization. The $k = 0$ component (a scalar) can only contribute to a two-photon transition connecting states of the same symmetry. Identification of $k = 0$ components in two-photon excitation spectra is generally rather straightforward since they are forbidden (and 'thus disappear') when the spectrum is recorded using circularly polarized light. Sensitivity to the polarization state of the exciting radiation is one important feature distinguishing one-photon and multiphoton transitions. As **Tables 1** and **2** show, for all but the least symmetric molecules, at least some of the $k \neq 1$ components will span representations different from (or additional to) those of the one-photon electric dipole moment operator.

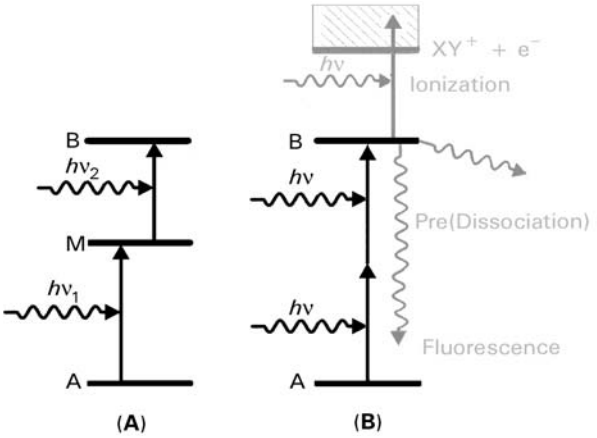



**Figure 1** Illustration of (A) sequential and (B) simultaneous two-photon excitation from state A to state B. Also shown in (B) are three possible fates of the excited state B: fluorescence, dissociation and further photon absorption that ionizes the molecule. This latter process it termed 2+1 resonance-enhanced multiphoton ionization (REMPI).

Multiphoton excitations can thus provide a means of populating excited states via transitions that are ‘forbidden’ in traditional one-photon absorption spectroscopy. Two-photon spectroscopy has proved to be particularly valuable in this regard, especially in the case of centrosymmetric molecules, e.g. $H_2$, $N_2$, $O_2$, the halogens, ethyne and benzene. All of these molecules have *gerade* ground states. Thus, in each case, one-photon absorption provides a route to populating the *ungerade* excited states but the *gerade* excited states are inaccessible unless the excitation is carried out using an even number of photons. Further inspection of **Table 1** hints at the increased complexity of coherent multiphoton excitation spectra. While governed by the same spin-conservation requirements and vibrational (i.e. Franck–Condon) restrictions as one-photon spectra, an *n*-photon excitation can support changes in rotational quantum number $\Delta J \leq n$.

## Experimental methods

The ‘basic’ multiphoton excitation experiment simply involves focusing tuneable laser radiation into a cell containing a low pressure (typically a few torr) of the atomic or molecular gas of interest and observing the resulting laser-induced fluorescence or, more commonly, the resulting ions or electrons (in the case of REMPI detection). In the latter experiment, the cell will typically be equipped with a pair of biased electrodes: an MPI spectrum is obtained simply by measuring the total ion, or the total photoelectron, yield as a function of excitation wavelength. The structure appearing in such a spectrum will reflect the resonance enhancements provided by the various rovibrational levels of the resonant intermediate electronic state(s) of the neutral, and may be analysed to provide spectroscopic (and thus structural) information about the excited neutral molecule. This ‘basic’ form of the REMPI experiment has limitations, and much of the recent experimental effort has been directed at improving both the selectivity and sensitivity of the technique.

One deficiency of this basic style of REMPI experiment is that all ions will be measured, irrespective of their masses, or that all photoelectrons will be counted, irrespective of their kinetic energies. A molecular MPI spectrum recorded in a static cell could therefore well include superimposed features associated with REMPI of the parent of interest, of neutral fragments arising from unintentional photodissociation of the parent molecule and of any other species present in the sample. Such potential ambiguities can usually be resolved by mass-resolving the resulting ions, and in most contemporary REMPI experiments this is achieved by time-of-flight (TOF) methods using either a linear TOF mass spectrometer, or a reflectron TOF mass spectrometer to provide enhanced resolution. A variety of fast charged particle detectors can be used, together with suitable time-gated signal-processing electronics, to monitor the REMPI spectrum associated with formation of any

**Table 1** Allowed changes in some of the more important quantum numbers and symmetry descriptors for atoms and molecules undergoing one-colour multiphoton transitions involving one ($k = 1$), two ($k = 0$ and 2) and three ($k = 1$ and 3) photons

| | *Rank of transition tensor, k (number of photons)* | | | |
|---|---|---|---|---|
| *Quantum number / property of interest* | *0 (2)* | *1 (1 or 3)* | *2 (2)* | *3 (3)* |
| (a) Atoms: | | | | |
| Orbital angular momentum, *l* of electron being excited | $\Delta l = 0$ | $\Delta l = \pm 1$ | $\Delta l = 0, \pm 2$ (but s $\nleftrightarrow$ s) | $\Delta l = \pm 1, \pm 3$ (but s $\nleftrightarrow$ p) |
| (b) Linear molecules (case (a)/(b)): | | | | |
| Axial projection of electronic orbital angular momentum, $\Lambda$ | $\Delta\Lambda = 0$ | $\Delta\Lambda = 0, \pm 1$ | $\Delta\Lambda = 0, \pm 1, \pm 2$ | $\Delta\Lambda = 0,..., \pm 3$ |
| linear molecules (case (c)): | | | | |
| Axial projection of total electronic angular momentum, $\Omega$ | $\Delta\Omega = 0$ | $\Delta\Omega = 0, \pm 1$ | $\Delta\Omega = 0, \pm 1, \pm 2$ | $\Delta\Omega = 0,..., \pm 3$ |
| (c) Centrosymmetric molecules: | | | | |
| Inversion symmetry, u/g | u $\leftrightarrow$ u<br>g $\leftrightarrow$ g | u $\leftrightarrow$ g | u $\leftrightarrow$ u<br>g $\leftrightarrow$ g | u $\leftrightarrow$ g |
| (d) Atoms and molecules: | | | | |
| Total angular momentum, *J* | $\Delta J = 0$ | $\Delta J = 0, \pm 1$ (but $J = 0 \nleftrightarrow J = 0$) | $\Delta J = 0, \pm 1, \pm 2$ (but $J = 0 \leftrightarrow J = 0,1$) | $\Delta J = 0,..., \pm 3$ (but $J = 0 \nleftrightarrow J = 0,1,2$; and $J = 1 \nleftrightarrow J = 1$) |
| Total parity, +/– | + $\leftrightarrow$ +<br>– $\leftrightarrow$ – | + $\leftrightarrow$ – | + $\leftrightarrow$ +<br>– $\leftrightarrow$ – | + $\leftrightarrow$ – |
| Electron spin, *S* | $\Delta S = 0$ | $\Delta S = 0$ | $\Delta S = 0$ | $\Delta S = 0$ |

**Table 2** Representations of the spherical tensor components $T_q^k$ ($\hat{O}$) of the one-colour, ($n$ = 1–3) transition operator

| *Number of photons, n* | *k* | *q* | $D_{\infty h}$[a] | $D_{6h}$ | $D_{3h}$[b] |
|---|---|---|---|---|---|
| 1 | 1 | 0 | $\Sigma_u^+$ | $A_{2u}$ | $A_2''$ |
| | 1 | ±1 | $\Pi_u$ | $E_{1u}$ | $E'$ |
| 2 | 0 | 0 | $\Sigma_g^+$ | $A_{1g}$ | $A_1'$ |
| | 2 | 0 | $\Sigma_g^+$ | $A_{1g}$ | $A_1'$ |
| | 2 | ±1 | $\Pi_g$ | $E_{1g}$ | $E''$ |
| | 2 | ±2 | $\Delta_g$ | $E_{2g}$ | $E'$ |
| 3 | 1 | 0 | $\Sigma_g^+$ | $A_{2u}$ | $A_2''$ |
| | 1 | ±1 | $\Pi_u$ | $E_{1u}$ | $E'$ |
| | 3 | 0 | $\Sigma_g^+$ | $A_{2u}$ | $A_2''$ |
| | 3 | ±1 | $\Pi_u$ | $E_{1u}$ | $E'$ |
| | 3 | ±2 | $\Delta_u$ | $E_{2u}$ | $E''$ |
| | 3 | ±3 | $\Phi_u$ | $B_{1u} + B_{2u}$ | $A_1'+A_2'$ |

[a] Assuming Hund's case (a) or (b) coupling. Ignore u/g labels for non-centrosymmetric linear molecules.
[b] $A_1'$ and $A_2''$ reduced to $A_1$, $A_2'$ becomes $A_2$, and $E'$ and $E''$ both transform as $E$ in $C_{3v}$ molecules.

single, user-selected, ion mass. In this way it is usually possible to distinguish spectral features associated with the parent from those arising from REMPI of neutral photofragments, or to distinguish different isotopomers of the same parent. Mass-resolved REMPI spectroscopy necessarily requires use of collision-free conditions; the precursor of interest in such experiments is thus introduced into the mass spectrometer source region as a molecular beam.

It often proves useful to measure the kinetic energies (KEs) of the resulting photoelectrons also. Such measurements also require use of a molecular beam so that their KEs (which are usually measured by TOF methods in a spectrometer designed to minimize stray electric and magnetic fields) can be recorded under collision-free conditions; they provide the basis for a number of variants of photoelectron spectroscopy discussed below.

## Applications

Less restrictive selection rules are just one of several benefits that can arise when using multiphoton excitation methods. Experimental convenience is another. A multiphoton excitation using visible or near-ultraviolet (UV) photons can often prove the easiest route to populating an excited state lying at energies that, in one-photon absorption, would fall in the technically much more demanding vacuum ultraviolet (VUV) spectral region. Other benefits derive from the fact that multiphoton excitations normally require the use of a focused pulsed laser because of the small excitation cross-section. The interaction is thus concentrated in a localized volume (the focal volume). The technique is therefore highly suitable for spatial concentration profiling, and well matched for use with supersonic molecular beams; many previously impenetrable molecular spectra have been interpreted successfully after application of multiphoton excitation methods to jet-cooled samples of the molecule of interest. This can be a huge benefit, especially in the case of REMPI where the resulting particles are charged and can be collected with far higher efficiency than could, for example, laser-induced fluorescence (LIF) from the excited state B. This benefit not only manifests itself in high sensitivity but, as we have seen, also offers additional species selectivity by allowing both mass analysis of the resulting ions and KE analysis of the accompanying photoelectrons.

### Spectroscopy, structure and dynamics of excited state species

REMPI spectroscopy is typically used to probe high-lying electronic states, for which dissociation is always likely, but it will discriminate in favour of the more long-lived states because of the competition between ionization and dissociation (as in **Figure 1B**). There is one class of excited states that are often relatively long-lived — Rydberg states, which thus tend to dominate REMPI spectra. Molecular Rydberg states are conveniently pictured as a positive ion core, consisting of the nuclei and all but

one of the valence electrons, with the remaining valence electron promoted to a state with a high principal quantum number, n. Such orbitals are large, spatially diffuse, and hence nonbonding, and are known as Rydberg orbitals. This is because the physical picture is very similar to that in the hydrogen atom, and the energy levels follow a modified Rydberg formula:

$$\tilde{\nu} = E_{\mathrm{i}} - \frac{R}{(n-\delta)^2} \qquad [2]$$

where $R$ is the Rydberg constant. As a written, $\tilde{\nu}$, $E_{\mathrm{i}}$ and $R$ must have the same units. $E_{\mathrm{i}}$ is the ionization limit of the ion core and $\delta$ is known as the quantum defect. It provides an indication of the extent to which the wavefunction of the Rydberg electron penetrates into the core region and its value is found empirically to be fairly constant for a given type of orbital. For molecules composed entirely of first-row atoms, typical values are $\delta = 1.0–1.5$ for s orbitals, $\delta = 0.4–0.8$ for p orbitals and $\delta \sim 0$ for all higher-l functions. Such qualitative ideas can be very useful for interpreting the patterns of excited states observed in many families of polyatomic molecules, though modifications due to configuration interaction (i.e. mixing between zero-order states sharing a common symmetry species but arising from different electronic configurations) can complicate such simple expectations. **Figure 2**, which shows a 2+1 REMPI spectrum of the NH radical serves to demonstrate several of these points. The spectrum is obtained by linearly polarized simultaneous two-photon absorption (at wavelengths ~271.2 nm) of NH radicals in their low-lying metastable excited $a^1\Delta$ state, followed by further one-photon excitation and detection of ions with $m/z$ 15. Rotational analysis confirms that the spectrum is carried by a two-photon transition, linking states of $^1\Delta$ (lower state) and $^1\Pi$ symmetry, while the observation of neighbouring vibrational brands (including hot bands originating from the $v = 1$ level of the $^1\Delta$ state) verifies that this is an electronic origin band. Changes in rotational quantum number $\Delta J \leq 2$ are clearly evident, as anticipated in **Table 2**. Knowing the ionization limit of the NH radical ($108\,804 \pm 5\,\mathrm{cm}^{-1}$, measured relative to the $X^3\Sigma^-$ ground state), we can deduce a value for the quantum defect of this state ($\delta = 0.79$) which, taken together with its known symmetry, suggests that this transition should be associated with electron promotion from the highest occupied doubly degenerate $1\pi$ orbital (the little-perturbed $2p_x$ and $2p_y$ orbitals of atomic N) to a $3p\sigma$ Rydberg orbital. The spectrum appears red-degraded, indicating a 13% reduction in the effective $B$ rotational constant upon electronic excitation. Multiphoton rotational line strengths (the multi-photon analogues of the Hönl–London line strength factors applicable to one-photon spectra) may be calculated, allowing derivation of the relative populations of the various initial quantum states contributing to the spectrum.

The simulated S branch contour shown in the top left part of **Figure 2** serves to illustrate another possible application of REMPI spectroscopy. Closer inspection of the experimental spectrum reveals that all transitions involving excited state levels with rotational quantum number $J' = 7$, 8 or 9 appear anomolously weakly. This is due to a very localized predissociation of the $v = 0$ level of the $f^1\Pi$ state of the NH radical. For these rotational levels in particular, the $f^1\Pi$ state predissociates at a rate that is comparable to, or greater than, the ionization rate; this competition leads to reduced ionization probability and a relative diminution of the eventual ion yield; multiphoton excitations proceeding via such predissociated levels thus appear with reduced relative intensity in the REMPI spectrum. In extreme cases the transitions involving such predissociated excited levels may show lifetime broadening as well. Clearly, in the case of more heavily predissociated excited states the REMPI signal is not only weaker (and thus harder to detect), but also less resolved, because of the increasing overlap of neighbouring lifetime-broadened spectral lines.

**Figure 3** shows an example involving the SO radical where the predissociation is so severe that an alternative detection scheme must be used. The necessary two-colour sequential double-resonance excitation scheme is indicated in the figure; the first step is designed to populate a single rotational level in the $A^3\Pi$ state. The fluorescence from this state is monitored, and a drop is seen when the second laser is tuned to a frequency appropriate for further excitation of these state-selected molecules to the $D^3\Pi$ state. **Figure 3** shows the A state fluorescence intensity as the second laser is scanned, and reveals a very broad Lorentzian peak (50 $\mathrm{cm}^{-1}$ full-width half-maximum). Use of the energy–time form of the uncertainty principle allows determination of the excited state lifetime (100 fs) from the width of the measured line shape. The D state is notionally a $4s\sigma$ Rydberg state, and its short lifetime is presumably indicative of significant mixing with a valence state, since the D state is the lowest-lying Rydberg state in SO.

This is just one of many instances where two-colour, double-resonance multiphoton spectroscopy can of great help in providing additional spectroscopic, structural and dynamical information about the

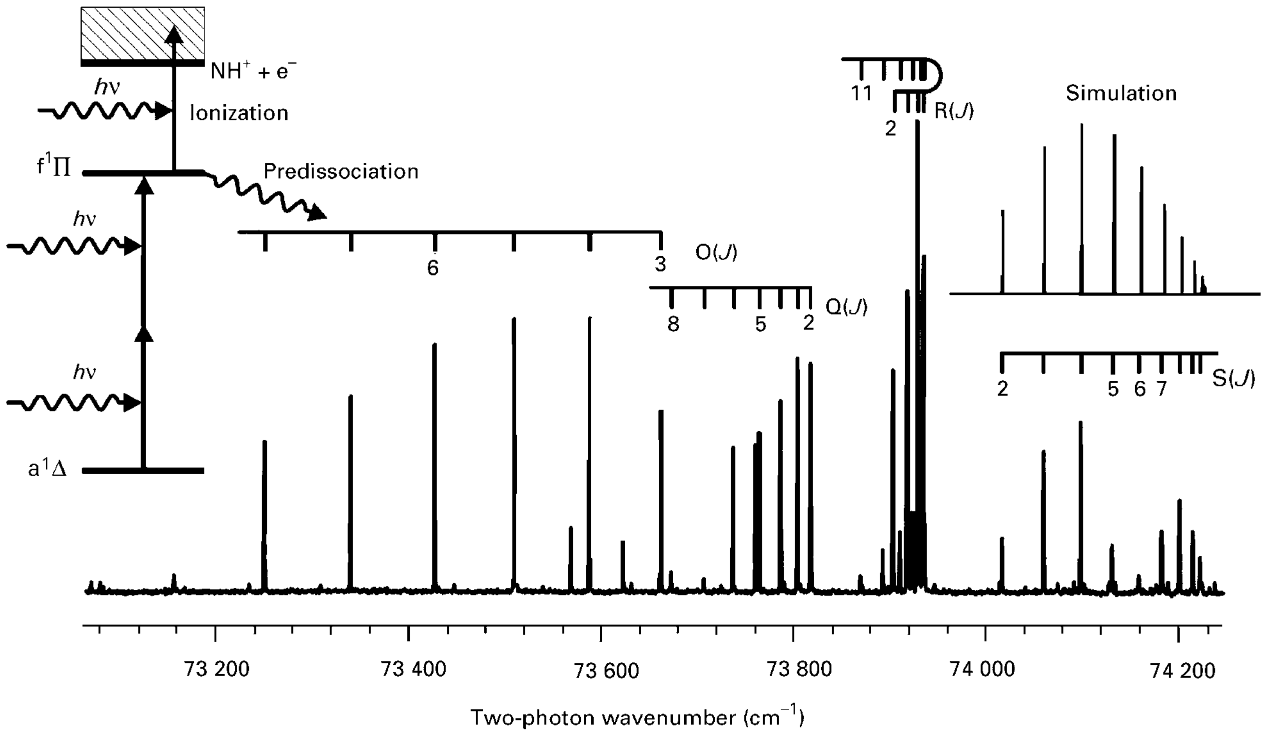



**Figure 2** Two-photon resonant MPI spectrum of the origin band of the $f^1\Pi \leftarrow a^1\Delta$ ($3p\sigma \leftarrow 1\pi$) transition of the NH radical obtained using the excitation scheme shown at the top left and monitoring the *m*/*z* 15 ion mass channel as a function of the laser wavelength. Individual line assignments are indicated via the combs superimposed above the spectrum. The simulation of the S branch (top right) highlights lines that appear in the experimental REMPI spectrum with reduced intensity because of competing predissociation.

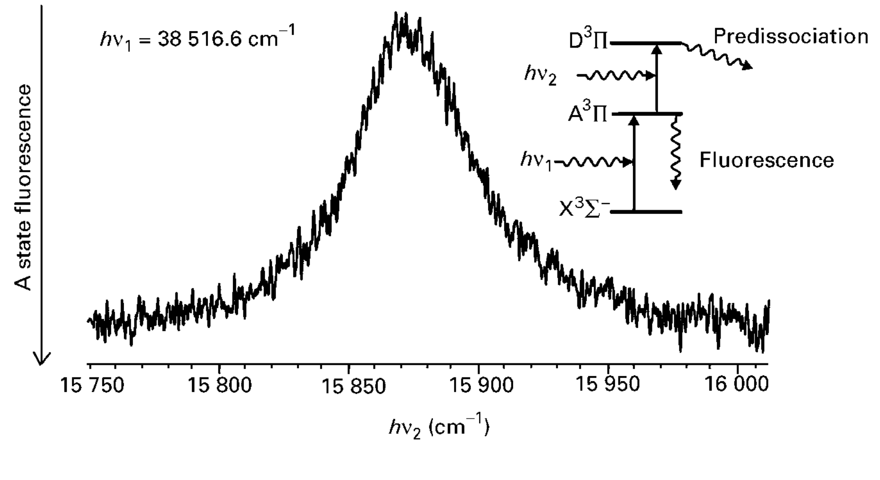



**Figure 3** A two-colour fluorescence depletion spectrum of one rovibronic line associated with the $D^3\Pi \leftarrow A^3\Pi$ transition in SO. The two-colour excitation scheme used (upper right) is required because of the very short lifetime (100 fs) of the $D^3P$ state. This results in the linewidth of 50 $cm^{-1}$ shown in the spectrum.

excited states of small and medium-sized gas-phase molecular species. Figure 4 shows the opposite extreme, where the final state is long lived (and ionization is used to detect that the multiphoton absorption has occurred), but double resonance is required to reach the states at all. The example involves states of the $S_2$ radical lying at energies around 75 000 $cm^{-1}$ where, without the simplification of jet

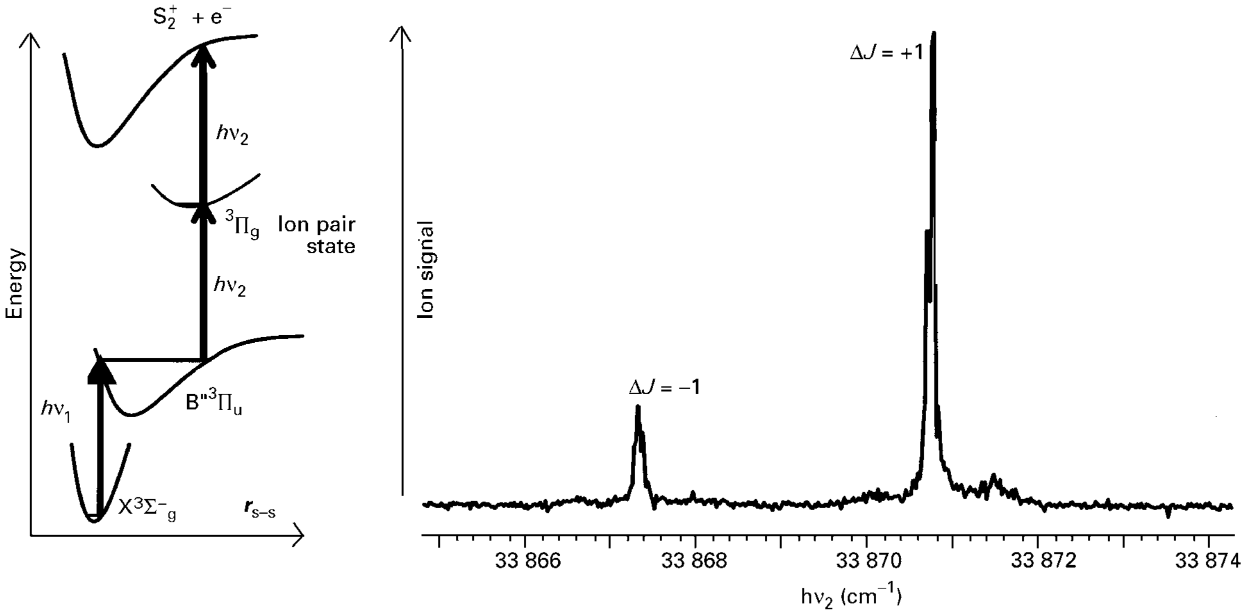



**Figure 4** Two-colour ionization spectrum exciting levels of a $^3\Pi_g$ ion pair state of $S_2$, using the scheme shown in the inset. Note that a two step process is required, both to give a net g ← g excitation and to overcome the poor Franck–Condon factors for the transition.

cooling and the additional state selectivity afforded by double resonance methods, the $S_2$ spectrum would be impenetrably complex. Further, the excited state of interest and the ground state of $S_2$ both have *gerade* parity. Recalling the selection rules listed in **Table 1**, we see that the excited state can only be reached by a spectroscopy that involves use of an even number of photons. The final state in this case is an ion pair state of $S_2$, a state that is best described as a pair of oppositely charged ions, $S^+S^-$, rather than a covalently bound S=S. As for Rydberg states, most if not all molecular species will have such ion pair states but, to date, their observation remains quite rare. This is because the equilibrium bond length in an ion pair state is generally much larger than in the ground state, with the result that the two states show little Franck–Condon overlap. Double-resonance spectroscopy can provide a means of accessing such states if, as here, and as illustrated, the first step is to the inner turning point of the wavefunction associated with a high vibrational level of an intermediate valence excited state while the second step is arranged to excite from near the outer turning point of this same vibrational level to the ion pair state of interest. In the case of $S_2$, the ground and ion pair states have B rotational constants of 0.30 cm$^{-1}$ and ~0.13 cm$^{-1}$, respectively, implying a 50% extension in equilibrium bond length when undergoing this double-resonant excitation. The versatility illustrated by these few examples serves to explain why multiphoton excitation methods in general, and REMPI in particular, continue to find widespread use as one of the most general and most sensitive species-selective methods of detecting atoms and small molecules (including radicals) in the gas phase.

## REMPI–photoelectron spectroscopy (PES)

We now consider the photoelectrons formed in the MPI process, and the information they may carry. The measurement of their kinetic energies has become an established technique, thereby providing a means of performing photoelectron spectroscopy on excited electronic states. Such measurements can therefore give important clues as to the electronic and vibrational make-up of the excited state. They also allow determination of such details as the number of photons involved in the overall ionization process and the source of any fragment ions. For example, a given daughter ion, $Y^+$, seen in the TOF spectrum of the ions resulting from REMPI of a parent, XY, can arise from photodissociation of the neutral parent followed by one- (or more) photon ionization of the fragment, i.e.

$$XY + nh\nu \rightarrow X + Y \quad [3a]$$

$$Y + mh\nu \rightarrow Y^+ \quad [3b]$$

or from MPI followed by photodissociation of the

resulting parent ion, i.e.

$$XY + nh\nu \rightarrow XY^+ \quad [4a]$$

$$XY^+ + mh\nu \rightarrow X + Y^+ \quad [4b]$$

or as a result of direct dissociative ionization of the parent, i.e.

$$XY + nh\nu \rightarrow (XY^+)^* \rightarrow X + Y^+ \quad [5]$$

Given the pulsed nature of the REMPI process, the electron KEs are almost always measured by TOF methods, either using a conventional (mu-metal-shielded) TOF spectrometer or a magnetic bottle photoelectron spectrometer. The latter offers the advantage of much higher collection efficiency, with comparatively little loss of ultimate KE resolution.

Recalling **Figure 1B** we note that when an MPI process is resonance enhanced by a bound excited state B, the vibrational structure in the resulting photoelectron spectrum will reflect the differences in the equilibrium geometries of state B and the parent ion, rather than between the ground state A and the ion as in traditional one-photon (e.g. He I) PES. Thus if the geometry and the vibrational level structure of the ion are already known, the vibronic structure evident in REMPI-PES can yield insight into the geometry of the resonance enhancing state B. If B is a pure Rydberg state, the electronic configuration of its core should be the same as that of the ionic state that lies at the convergence limit of the series to which it belongs. The Rydberg state and the ion will therefore be likely to have very similar geometries. Thus, by the Franck–Condon principle, we can anticipate that the final ionizing step in a REMPI process via such a Rydberg state will involve a $\Delta v = 0$ transition, leading to selective formation of ions with the same vibrational quantum number(s) as in state B. Since the photoionization is brought about using a (known) integer number of photons, the photoelectrons accompanying such state specific ion formation will have a narrow spread of KEs. In favourable cases the TOF spectrum of these photoelectrons can be resolved to the extent that individual rotational states of the ion are revealed, thus explaining the continuing appeal of REMPI-PES as a means of determining accurate ionization thresholds and of investigating photoionization dynamics in simple molecular systems.

Another form of PES has emerged that can provide a further order of magnitude improvement in energy resolution (i.e. $cm^{-1}$ resolution). This technique is now generally referred to as zero kinetic energy (ZEKE)-PES. In conventional photoelectron spectroscopy, and in REMPI-PES, we learn about the energy levels of the ion by measuring the photoelectron KEs as accurately as possible. ZEKE-PES also reveals the energy levels of the cation but is based on a different philosophy. The principle of the method is illustrated in **Figure 5**. In the particular double-resonant variant shown, one laser is tuned so as to populate a (known) excited state M, and a second laser pulse is then used to excite this population to the energetic threshold for forming one of the allowed quantum states of the ion. Any energetic electrons (e.g. those formed via an autoionization process) will quickly recoil from the interaction region. The ZEKE (threshold) electrons can be detected by application of a suitably delayed pulsed extraction field. A ZEKE-PES spectrum is obtained by measuring the excitation spectrum for forming photoelectrons with zero kinetic energy; precise energy eigenvalues are obtained because the spectral resolution is determined, ultimately, by the bandwidth of the exciting laser. It is now recognized that this description, while appealing, actually oversimplifies the physics. The 'ZEKE' electrons detected in an experiment as described actually derive from pulsed-field ionization of very high Rydberg states belonging to series converging to the threshold of interest. As a result, the ionization thresholds determined via this type of experiment will all be subject to a small, systematic shift to low energy. However, the magnitude of this shift scales with the applied extraction voltage, so the true thresholds can be recovered by recording such spectra using a number of different pulsed extraction voltages and extrapolating the observed line frequencies to zero applied field.

## Ion imaging

As REMPI is a very sensitive, selective and convenient means of detecting small localized concentrations of gas phase species, it is particularly suited to probing atomic or molecular products resulting from a gas-phase photodissociation or a crossed-beam reaction. Accurate knowledge of the energy disposal in such products, their recoil velocities and the angular distribution of these velocity vectors, the alignment of their rotational angular momenta, and the way all these quantities are correlated, can provide considerable insight into the detailed dissociation and/or reaction dynamics. Ion imaging is one way in which REMPI spectroscopy is being used to provide such information. The experiment is simple in concept. In the case of photodissociation, the precursor of interest, in a skimmed molecular beam, is photolysed to yield fragments, one of which is ionized selectively and in a

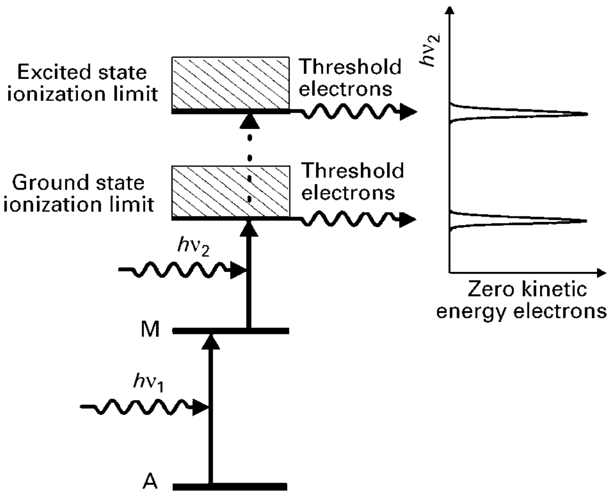



**Figure 5** Illustration of two-colour two-photon ZEKE excitation scheme, in which the first photon is fixed so as to be resonant with a known M ← A transition and the frequency of the second photon is tuned. As shown in the inset, the peaks in a ZEKE spectrum correspond to the onsets of new ionization thresholds.

quantum state-specific manner, by REMPI, as soon as it is created. The resulting cloud of ionized fragments continues to expand with a velocity and angular distribution characteristic of the original photolysis event, but is simultaneously accelerated out of the interaction region and arranged to impact on a position-sensitive detector, e.g. a microchannel plate behind which is mounted a phosphor screen that is viewed using a gated image-intensified CCD camera. The result in a squashed two-dimensional projection of its initial 3D recoil velocity distribution. This can be reconstructed mathematically to yield the speed and angular distributions of the tagged fragment, in the particular quantum state defined by the REMPI excitation wavelength. The structure in the image provides information about the speed and angular distribution of the tagged fragments and also, by energy and momentum conservation arguments, the quantum state population distribution in the partner fragments; such knowledge can provide a uniquely detailed view of the parent photofragmentation dynamics. By way of illustration, **Figure 6** shows ion images obtained by ionizing ground-state ($^2P_{3/2}$) Br atoms resulting from photolysis of $Br_2$ molecules at three different wavelengths. It is clear from these that the velocity and angular distribution of these Br atoms (defined relative to the electric vector, ε, of the photolysis laser – vertical in **Figure 6** as indicated by the double-headed arrow) depends on the photolysis laser wavelength. The radii of the partial rings apparent in each image give the speed of the atoms and hence, by energy conservation, the energy of the other fragment. The middle image reveals that $Br_2$ photolysis at 460 nm yields ground-state Br($^2P_{3/2}$) atoms (the tagged species) in conjunction with both another ground-state atom (outer ring) and a spin–orbit excited-state ($^2P_{1/2}$) partner, the inner ring. These product sets show different recoil anisotropies. The relative intensities of the two partial rings provides a measure of the branching ratios into these two product channels. Clearly, dissociation via a perpendicular transition ($\nu_{recoil}$ perpendicular to $\varepsilon$) yielding two ground state Br

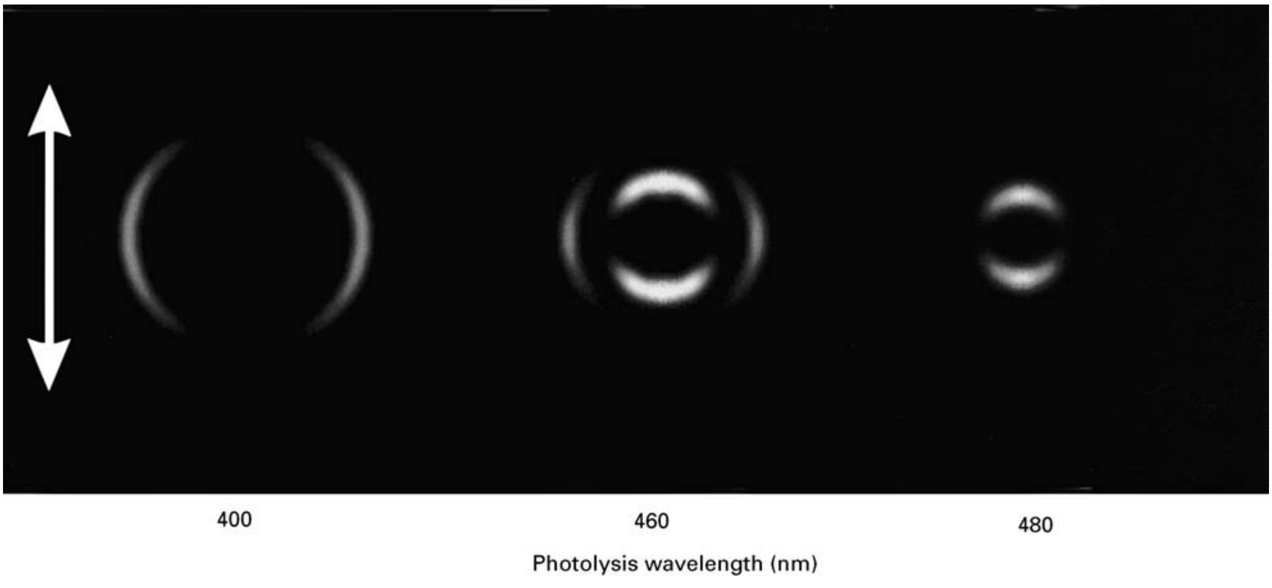



**Figure 6** Ion images of ground-state Br ($^2P_{3/2}$) atoms resulting from $Br_2$ photolysis at the specified wavelengths. The double-headed arrow indicates the plane of polarization of the photolysis laser radiation.

atoms is the dominant decay mechanism at 400 nm, whereas at longer excitation wavelength (e.g. 480 nm) the dominant fragmentation is to one ground-state ($^2P_{3/2}$) and one spin–orbit excited-state ($^2P_{1/2}$) Br atom, following a parallel excitation process. Analysis of images like these, and their dependence on photolysis wavelength, can provide much insight into both the mechanism and the timescale of the dissociation process.

## List of symbols

$B$ = rotational constant; $E_i$ = ionization limit; $h$ = planck constant; $J$ = rotational quantum number; $n$ = principal quantum number; $\hat{O}$ = multiphoton transition operator; $R$ = Rydberg constant; $T^k_q(\hat{O})$ = $q$th component of the tensor of rank $k$ representing $\hat{O}$; $\delta$ = quantum defect; $\varepsilon$ = electric vector; $\nu$ = frequency; $v$ = vibrational quantum number; $\tilde{\nu}$ = wavenumber.

*See also:* **Ion Imaging Using Mass Spectrometry; Laser Spectroscopy Theory; Multiphoton Excitation in Mass Spectrometry; Multiphoton Spectroscopy, Applications; Photoelectron Spectrometers; Photoelectron Spectroscopy; Photoelectron-Photoion Coincidence Methods in Mass Spectrometry (PEPICO); Photoionization and Photodissociation Methods in Mass Spectrometry; Time of Flight Mass Spectrometers.**

## Further reading

Ashfold MNR, Clement SG, Howe JD and Western CM (1993) Multiphoton ionization spectroscopy of free radical species. *Journal of the Chemical Society, Faraday Transactions* **89**: 1153–1172.

Ashfold MNR and Howe JD (1994) Multiphoton spectroscopy of molecular species. *Annual Review of Physical Chemistry* **45**: 57–82.

Heck AJR and Chandler DW (1995) Imaging techniques for the study of chemical reaction dynamics. *Annual Review of Physical Chemistry* **46**: 335–372.

Houston PL (1995) Snapshots of chemistry: product imaging of molecular reactions. *Accounts of Chemical Research* **28**: 453–460.

Kimura L and Achiba Y (1989) In Lin SH (ed) *Advances in Multiphoton Processes and Spectroscopy* Vol 5, pp 317–370. New Jersey: World Scientific.

Lin SH, Fujimara Y, Neusser HJ and Schlag EW (1984) *Multiphoton Spectroscopy of Molecules*. New York: Academic Press.

Müller-Dethlefs K and Schlag EW (1991) High resolution zero kinetic energy (ZEKE) photoelectron spectroscopy of molecular systems. *Annual Review of Physical Chemistry* **42**: 109–136.

# Multivariate Statistical Methods

**RL Somorjai**, Institute for Biodiagnostics, National Research Council, Winnipeg, Canada

Copyright © 1999 Academic Press

**FUNDAMENTALS OF SPECTROSCOPY**
**Methods & Instrumentation**

## Introduction, basic ideas, terminology

Spectroscopic methods are increasingly becoming the methods of choice for analysing a variety of experimental data, in chemistry, biology, food industry, medicine, etc. This popularity is well deserved. The spectral methods are generally faster, more accurate and frequently much cheaper than conventional analytical techniques. Furthermore, in biomedical applications they provide the means for noninvasive or minimally invasive diagnosis. However, these obvious advantages are somewhat offset by the indirect relationship between spectral features and the measurables or observables of interest (such as analyte concentrations or disease class assignments). Consequently, we have to model a generally complex and frequently nonlinear relationship. This relationship can be represented by

$$\mathbf{Y} = \mathbf{F}(\mathbf{X}) \qquad [1]$$

where $\mathbf{Y} = \{y, y_2, y_3, \ldots, y_K\}$ is the set of $K$ measurables (responses, observables, targets) (e.g. the concentrations of $K$ analytes or membership labels for $K$ classes), $\mathbf{X} = \{x_1, x_2, x_3, \ldots, x_N\}$ is the collection of $N$

samples (objects, patterns), with $\boldsymbol{x}_k$ the $k$th sample. Every sample is represented as an $L$-vector, with each of its $L$ elements corresponding to one of the $L$ spectral features (attributes, predictor variables, measurements) (such as wavelengths or frequencies), i.e. $x_k = \{x_k(1),\ x_k(2),\ x_k(L)\}$, and **F** is the model (function) that couples **Y** with **X** and whose parameters we are to optimize, implicitly or explicitly. Any practical scheme of optimization must necessarily be data-driven. Hence the larger $N$, the number of samples, the more reliable any prediction based on Equation [1]. In spectroscopy (whether infrared, IR, magnetic resonance, MR, or other), $K$ (the number of analytes or the number of classes) typically ranges from one to about a dozen, $L$ (the number of spectral wavenumbers or frequencies) is in the hundreds or thousands, and $N$ ought to be at least in the hundreds. Since in spectroscopic applications $L$, the dimension of the feature space, is invariably greater than one, and typically large, multivariate methods of analysis are necessary. Furthermore, $L$ is frequently larger than $N$, the sample space dimension. As this causes numerical problems, special approaches are needed and precautions have to be taken. These we can handle via some preprocessing procedure. Its most important goal is data reduction (compression). In practice, data compression is divided, somewhat artificially, into feature extraction or feature subset selection.

Exploratory data analysis and representation is usually the first step in probing the functional relationship between **Y** and **X**. This falls into the purview of unsupervised pattern recognition, with clustering methods the most common representatives. Hierarchical clustering is the older, having its roots and initial applications in taxonomy, psychology and the social sciences. Its limitations are due, paradoxically, to its flexibility: by changing the merging (or splitting) criterion and/or the dissimilarity measure, we can create almost any grouping of the samples. The final number of clusters depends on a user-selected threshold; hence it is subjective. The fact that the samples are partitioned into mutually exclusive (non-overlapping) groups must be regarded as unrealistic.

The other major variant of clustering minimizes some objective function of the intersample distances. The results are also dependent on user-selected parameters (e.g. the number of clusters and the type of the distance measure used). However, these results become more realistic if we allow overlap between clusters, i.e. if we accept that the samples can be fuzzy, having memberships in all clusters. Bezdek's fuzzy $c$-means clustering algorithm is the most popular of such clustering methods. Neither clustering variant can be expected to outperform methods based on supervised pattern recognition.

Although calibrating or classifying spectra are merely two different applications of the mathematical/statistical technique of regression, historically they have evolved independently, with different goals and requirements. Whenever possible attempts will be made to connect the two by drawing attention to common concepts and methods that only differ in their terminology and emphasis. Occasionally, for the sake of simplicity in presentation, discussion will cover certain aspects and peculiarities of the two disciplines separately.

Calibration *is* regression, but with a specific connotation: it involves the development of a quantitative, and generally linear relationship between, say, the concentrations of the analytes present and their spectral manifestations. Significantly, this relationship is continuous. In contrast, classification assigns class labels to the samples and thus establishes a categorical correspondence between the samples and their spectral features, a generally discrete relationship.

Both calibration and classification belong to the group of supervised pattern recognition methods. Supervision means that the model parameters in Equation [1] must be optimized using information provided by and extracted from the available data. To do this so that prediction is reliable and robust, leads naturally to the concept of partitioning the samples into training (design, learning) and validation (test, prediction) sets.

This article will focus on concepts and ideas, without delving into detailed description of individual methods. The interested reader is directed to Further reading for more in-depth accounts and discussions where there are also listed some of the most likely journal sources for the latest advances.

## Preprocessing

The characteristics of spectra demand that we use preprocessing to guarantee that the predictions from the optimized Equation [1] are reliable and robust. In fact, extensive experience suggests that if the preprocessing part of the analysis is done properly, then even the simplest calibration or classification methods will succeed.

At its most straightforward, Equation [1] describes some simple functional transformation of the original data. This might be as elementary as mean-centring and scaling ('whitening'), i.e. subtracting the overall sample mean from each individual sample and dividing by the sample variance. Other useful examples include smoothing (filtering), normalization (e.g. by the overall area under the spectra),

monotonic nonlinear operations (e.g. using the logarithms or powers of the spectral intensities), and analysing the (numerical) derivatives of the original spectra. For instance, the assignment of mid-IR spectra of biofluids into different disease classes is most successful when the classifier uses the first derivatives of the spectra.

Extracting some intrinsic structure from the data, while in the original $L$-dimensional feature space, is rarely successful because the majority of the $L$ spectral features are either redundant (typically owing to correlation) or represent irrelevant information ('noise'). Any of these will usually mask the discriminating features. We then interpret preprocessing as either a procedure that removes irrelevant features (a type of filtering) or one that finds the optimal subspace in which the data can be best analysed (a form of projection).

The number of spectral features $L$ is generally large and frequently larger than $N$, the number of samples. Furthermore, (adjacent) spectral features are often strongly correlated, i.e. they are not independent. If $L > N$, linear dependency or near-dependency (called multicollinearity by the calibration community) could lead to numerical instabilities, with consequent unreliable and non-reproducible results. For instance, when classifying spectra via linear or quadratic discriminant analysis (LDA or QDA), the occurrence of singular or near-singular covariance matrices causes matrix inversion problems. The overall conclusion, supported by extensive experience, is that we must somehow reduce the number of features. Although no general theoretical proof is available, in practice the ratio $N$:$L$ should be at least 10, but preferably 20, for reliable classification or calibration results. Instead, for spectra this ratio is more likely to be 0.1 or even less.

Data reduction can be achieved by feature extraction, which first transforms the original $L$ variables $\mathbf{X}$ into $L$ new variables $\mathbf{Z} = \mathbf{G}(\mathbf{X})$. $\mathbf{G}$ is nonlinear in general. However, principal component analysis (PCA), the most commonly used transformation, is linear: $\mathbf{Z} = \mathbf{AX}$, where $\mathbf{A}$ is a matrix. PCA is an unsupervised method, applied to the entire available data. It is carried out by diagonalizing the sample covariance matrix $\mathbf{S}$. If we sort the resulting eigenvalues in decreasing order, then the corresponding eigenvectors (principal components, PCs) point, in the original $L$-dimensional feature space, in directions of successively decreasing variance. The PCs are orthogonal and uncorrelated. Geometrically, PCA corresponds to a rotation of the original $L$ coordinate axes to $L$ new orthogonal axes formed by the $L$ PCs. We achieve the data reduction by retaining only the first $M << L$ PCs ($\mathbf{Z}$s) for further analysis. (Recall that the PCs are different linear combinations of all $L$ original features.) However, the first $M$ PCs, even if they account for a large fraction of the total variance in the data, are not necessarily the most discriminating directions along which classes may be separated. In fact, it is more useful to select those $M$ PCs that give rise to the lowest classification error. These are generally not the first $M$ in the ordered list. Similarly, in calibration, it makes more sense to select those $M$ PCs that show maximum correlation with the $\mathbf{Y}$. Such arguments lead naturally to the other data reduction method, feature subset selection. As a first application, we could select from the $L$ PCs the $M$ that best satisfy the above requirements. This is always beneficial. However, the general approach has several attractive characteristics which don't depend on a prior PCA, especially for features originating in spectra. The foremost of these is that the features of the optimal subset retain their spectral identity. This is, of course, important for interpretability, especially in MR spectroscopy, where the spectral peaks are relatively narrow and have distinct chemical identities. Another significant practical advantage is that once we identify a smaller subset of the complete set of features we never need to determine or measure the remaining features again.

Assume that we wish to select the $M << L$ best features, with 'best' defined appropriately for the given application. The naive approach of combining the individually best $M$ features rarely produces the best $M$-feature subset. Examining all possible $M$-feature subsets to find the best is usually prohibitive computationally. A possible solution is to restrict the search to a subspace of the complete space of all feature sets. The standard stepwise methods proceed along these lines. The forward selection (FS) version is computationally the least demanding. FS starts by selecting the $M$ individually best features. This is repeated, at each stage adding that feature which, when combined with those already chosen, gives the best result. The backward elimination (BE) algorithm starts with the complete set of features and sequentially removes the one that least degrades performance. The BE method is computationally prohibitive when the starting set of features is large, as is the case for spectra. Both have been generalized (e.g. more than one feature can be added or removed at a time). FS and BE can also be combined. Neither method is guaranteed to find the globally optimal $M$-feature subset. Branch-and-bound methods, which would guarantee the best possible subset, are rarely applicable to spectra. Most recently, optimal feature selection techniques based on genetic algorithms (OFS-GA) have been used with great success. The approach is natural

for preprocessing spectra or time series, for which the feature set originates from sequential discrete sampling from or measurement of an underlying continuum. Consequently, the dimensionality $L$ of the original feature space is large. The inherent order (along time or wavelength) implies that adjacent features will be correlated, a characteristic that causes redundancy. Unlike BE, OFS-GA can start with the original feature space. It maps this onto a bit string (zeroes and ones). GA operations, such as mutation and crossover are used on a population of bit strings. The performance criterion is the minimization of an objective (fitness) function $F$. If $F$ is the mean square error between the training set classification results and the a priori class indicator, then the best $F$ simultaneously maximizes the overall accuracy for the training set and the crispness of the class assignments. The input to the algorithm is $M$, the ultimate number of features (distinct subregions) required, and the type of feature space-reducing operation or transformation to be carried out in these subregions. Typical operations include (but are not confined to) replacing the individual features in a subregion by their average or variance. Selecting an $M << L$ and one of these operations lead jointly to substantial feature space reduction.

## Training versus test sets and crossvalidation

Suppose we knew exactly the multivariate distributions from which we drew our samples. Then, at least in principle, we could calculate the true error (called the Bayes error, $e_B$, which is the smallest achievable error given the exact distributions). In practice, we have a finite, limited number of $N$ samples. When the samples are random and $N$ is sufficiently large to be representative of the underlying distributions, the entire data set can be used to optimize the calibration or classifier parameters. Then predicting the concentrations of the $K$ analytes present in a new sample (calibration), or assigning this sample to one of the $K$ classes (classification), is analogous to interpolation. In contrast, for the typical case of small or moderate $N$, prediction, or class assignment corresponds to the much less reliable process of extrapolation. Thus, in most practical situations two contradictory requirements confront us. One the one hand, for reliability we would like to use all $N$ samples. When this is done, it is called the resubstitution (plug-in) method in the classification literature. However, the resubstitution or apparent error estimate $e_R$ is generally overoptimistic. The danger of overfitting threatens, without any guarantee that new samples, not used for training, would be correctly predicted or assigned. The consequence of overfitting is that the classifier or regressor will have poor generalizing power. On the other hand, let us partition the $N$ available samples into independent training and validation sets, with $N_T$ and $N_V$ ($N = N_T + N_V$) samples respectively. Then we may satisfy validation requirements but reliability suffers because $N_T$ is not sufficiently large.

Crossvalidation techniques may be used to compensate for the smallness of the sample set (unfortunately far too frequent in biomedical applications). Lachenbruch proposed the leave-one-out or delete-1 ($D_{[1]}$) method (a close relative of Tukey's jackknife). For classification problems this proceeds by successively deleting each of the $N$ samples from the data set, training a classifier on the remaining $N - 1$, and assigning the deleted sample to one of the classes. ($N_T = N - 1$ and $N_V = 1$). Thus we create and test $N$ ($N - 1$)-sample classifiers, each with performance only very slightly worse than that of the complete $N$-sample classifier. (The entire process is identical for calibration, with 'classification' and 'classifier' replaced by 'regression/calibration' and 'regressor'.) The result is a much-reduced bias for the classification (calibration) error, $e_{D[1]}$. Unfortunately, for small or even moderate $N$ the variance can be large. The delete-d ($D_{[d]}$) crossvalidation strategy can reduce the variance. The complete delete-d ($CD_{[d]}$) version produces $N!/d!(N - d)!$ classifiers, a huge number even for moderate $N$. (If $N = 50$ and $d = 10$, this number is ~$1 \times 10^{10}$). Thus, in practice we select uniformly a random subset of size $B$ out of these. As long as $N/B \to 0$, e.g. with $B = N^\delta$, $\delta > 1$, the $D_{[d]}$ strategy retains the efficiency of the $CD_{[d]}$. The expected classification error is the average of the $B$ individual errors, as measured on the $B$ sets of $d$ deleted, i.e. test samples.

The above are examples of nonparametric resampling (RS) methods, without replacement. The deservedly popular bootstrap method is also a resampling method, but with replacement. Because of replacement, some objects may appear several times in any of the $B$ bootstrap samples. This could cause difficulties for those classifiers that we optimize by inverting covariance matrices. A simple remedy is to perform the sample replacement only after we have selected for the current training set a particular subset of $N - d$ distinct samples. Again, we secure the end result by averaging the $B$ individual classifier outcomes.

Resampling, based on a random (Monte Carlo) subset selection is a very powerful crossvalidation approach. It is not constrained by the computational limitations of the $CD_{[d]}$, even for $d \approx N/2$. Further-

more, RS methods are equally effective in providing nonparametric confidence intervals/regions for any computed estimate.

## (Supervised) classification methods

These fall into two categories: parametric and nonparametric. The former assumes that the samples for the $K$ classes derive from some known distribution (usually multivariate normal), whereas the latter is distribution-free. There is no best method: the choice of the classifier depends both on the problem and on the property that we want to optimize. Most frequently, we minimize the classification error, in particular $e_V$ for the validation set, but this need not be the only choice. Our current concern is with the classification of spectra. In the following, we shall assume that the original, $L$-dimensional feature space was already reduced to an $M$-dimensional space, $M << L$ ($M$-space).

After 60 years, Fisher's LDA method is still the workhorse of the parametric methods. Of all the statistical classifiers, we understand LDA's theoretical properties best. We can derive it from ordinary (linear) least-squares (OLS) regression. LDA creates those linear combinations of the $M$ features that best separate the classes. For each class $k$ there is one of these discriminant functions $d_k$. The decision surfaces $g_{jk} = d_j - d_k = 0$ between any pair $(j,k)$ of classes are also linear, i.e. they are hyperplanes in $M$-space. The properties of LDA are optimal in the Bayes sense if all $K$ distributions are multivariate normal and have a common covariance matrix $\mathbf{S}$. In practice, these conditions are rarely satisfied. Nevertheless, LDA is surprisingly robust against even moderate deviations from normality and/or from the equality of the covariance matrices. LDA requires the optimization of $M + 1$ parameters for a $M$-dimensional feature space, the minimum number for a parametric statistical classifier.

If the covariance matrices $\mathbf{S}_k$ are different, then the classifier created is called quadratic discriminant analysis (QDA). The hypersurfaces separating the classes become quadratic (curved). Unfortunately, this additional flexibility comes at a cost: we have to optimize $O(M^2)$ parameters instead of the $O(M)$ in LDA. Usually $O(M^2) \geqslant N$, leading to overfitting, numerical instabilities, etc. Furthermore, QDA is less reliable, because it uses the $K$ $\mathbf{S}_k$s, each an estimate of the (generally unknown) population covariance matrix, whereas LDA uses the more reliable single, pooled sample covariance matrix.

Friedman suggested regularized discriminant analysis (RDA) to alleviate some of these problems. RDA introduces a two-parameter estimate of the sample covariance matrix. If $\mathbf{S}_k$ is the sample covariance matrix for class $k$, and $\mathbf{S}$ their pooled version, then we construct a two-parameter covariance matrix $\mathbf{S}_k(\lambda,\gamma)$, $0 \le \lambda \le 1$, $0 \le \gamma \le 1$. The four corners defining the extremes of the $\lambda,\gamma$ parameters represent well-known classifiers; ($\lambda = 0$, $\gamma = 0$) corresponds to QDA, ($\lambda = 1$, $\gamma = 0$) to LDA, whereas ($\lambda = 1$, $\gamma = 1$) gives the nearest-means classifier (assigning a sample to the class with the nearest training set mean, using Euclidean distance). When ($\lambda = 0$, $\gamma = 1$) a variance-weighted version of the nearest-means classifier is generated. The optimum $(\lambda, \gamma)$ pair is usually determined via crossvalidation. Excellent results have been reported even when the 'number of samples' to 'number of features' ratio $N{:}M$ was unfavourably low.

RDA modifies (shrinks) the eigenvalues of the original covariance matrices. Other methods exist with similar goals. Amongst these is DASCO (discriminant analysis with shrunken covariances). Although popular with the chemometric community, DASCO is quite complex, yet does not outperform RDA or the simpler ridge method. The latter also modifies the sample covariance matrix to be used with LDA from $\mathbf{S}$ to $\mathbf{S} + \alpha\,\mathbf{I}$, where $\alpha$ is a parameter to be optimized, $\mathbf{I}$ the unit matrix.

Among the nonlinear methods, neural net (NN) classifiers are relatively recent. Conceptually attractive, they are frequently used uncritically. A common danger is overfitting: because of the essentially unlimited flexibility of multilayer NNs, the training set can be classified perfectly (by creating an arbitrarily convoluted and complicated decision surface). Unfortunately, this high classification accuracy does not generally carry over to the validation set.

Recursive partitioning (decision tree) methods had found early use in numerical taxonomy, the social sciences and medicine. More recently, researchers in both artificial intelligence and statistics have substantially advanced these methods. There are two major types. Binary decision trees would be relevant in medical diagnosis, where the features are symptoms, with only yes/no answers for the decisions. If the features are continuous (spectral features qualify) then the decision trees are more complex, requiring the partitioning of each feature. The partitioning criteria and the sequence in which the features are probed are aspects of the tree design. The major advantage of the decision tree methods is ease of interpretability. To apply these methods to spectra, prior feature space reduction seems mandatory.

The nearest neighbour (nn) methods are nonlinear, nonparametric classification methods. They estimate probabilities at $x$ using the classes of adjacent training set points. The estimates will have smaller bias if only close neighbours are used, hence nearest

neighbours. In applying these methods a distance measure, e.g. Euclidean, is needed. Among the advantages are simplicity and an upper bound on their asymptotic misclassification error: $e_{nn} \leq 2e_B$. A major disadvantage in a high-dimensional feature space is the unacceptably large number of training samples required for accuracy.

## Calibration techniques

A major desideratum in most calibration work is linearity. If the relation between **Y** and **X** (Eqn [1]) appears nonlinear, then either the **X** or the **Y** (or both) are transformed until linearity is achieved. This is important for accurate, quantitative predictions of continuous responses such as analyte concentrations. This preoccupation with linearity guided the development of the most important calibration methods.

Because of their common origin (i.e. regression), calibration methods share many characteristics with classification methods. Nevertheless, we can introduce most calibration techniques more simply in terms of their original formulation. Furthermore, several of these methods treat data reduction and calibration (fitting) simultaneously. Wold's partial least squares (PLS) regression reduces the number of features on which the OLS regression is carried out. This is done by feature extraction, i.e. from the raw features the algorithm produces 'derived features' (DFs, the partial least-squares). The first DF is the linear combination of the $L$ original features that has maximum covariance (rather than correlation) with **Y**. Subsequent DFs are found in the same way, subject to being uncorrelated with previous ones. PLS can be described in terms of covariance regularization (maximization).

Principal component regression (PCR) is intimately related to PCA. If the spectral decomposition (diagonalization) of the cross-product matrix **X′X** is,

$$\mathbf{S} = \sum_{k=1}^{L} E_k \mathbf{u}_k \mathbf{u}_k'$$

then in PCR the matrix

$$\mathbf{S}^{-1} = \sum_{k=1}^{M} E_k^{-1} \mathbf{u}_k \mathbf{u}_k'$$

is used, where $M << L$ . $\mathbf{u}_k$ is the $k$th PC and the $M$ PCs are usually chosen to have the highest correlation with **Y**. There are other methods, e.g. shortest least-squares, but the above two are used most frequently.

## Discussion

Is there a best method for classification or regression? The answer must be: 'It depends on the data and the criterion to be optimized'. Nevertheless, some general recommendations can be made. The first is that preprocessing the data is probably the most important step to be taken for reliable, reproducible results. The second is that crossvalidation is essential if we want reliable predictions for new samples. Finally, as many samples should be collected as possible. If at least the first two are carried out properly, then even the simplest classifiers or regressors will work well. However, the most important message is: know your data!

Some important topics could not be dealt with because of space limitation. Outlier detection is amongst these. Every attempt should be made to find outliers before classification or calibration is carried out, otherwise the results will be distorted.

## List of symbols

$CD_{[d]}$ = complete delete-$d$; $d$ = number of samples deleted; $e_B$ = Bayes error; $e_R$ = resubstitution error; $e_{D(1)}$ = classification error; $e_{nn}$ = classification error; $L$ = number of variables; $N$ = number of samples; $N_T$ = number of samples in training sets; $N_V$ = number of validation sets; **S** = covariance matrix; **X** = collection of $N$ samples; **Y** = set of $K$ measurables; LDA = linear discriminant analysis; QDA = quadratic discriminant analysis; IR = infrared; MR = magnetic resonance; PC = principal component; PCA = principal component analysis; RDA = regularized discriminant analysis; OFS-GA = optical feature selection-genetic algorithm; FS = forward selection; BE = backward elimination; DASCO = discriminant analysis with shrunken covariances; NN = neutral net; nn = nearest neighbour; PLS = partial least squares; PCR = partial least squares.

*See also:* **Biofluids Studied By NMR; Calibration and Reference Systems (Regulatory Authorities); Computational Methods and Chemometrics in Near-IR Spectroscopy; Fourier Transformation and Sampling Theory; Laboratory Information Management Systems (LIMS).**

## Further reading

Bezdek JC (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press.

Bishop CM (1995) *Neural Networks for Pattern Recognition*. Oxford: Clarendon.

Breiman L, Friedman JH, Olshen RA and Stone CJ (1984) *Classification and Regression Trees* Belmont, CA: Wadsworth International Group.

Cook RD and Weisberg S (1982) *Residuals and Influence in Regression*. London: Chapman & Hall.
Deming SN and Morgan SL (1987) *Experimental Design: A Chemometric Approach*. New York: Elsevier.
Efron B and Tibshirani RJ (1993) *An Introduction to the Bootstrap*. New York: Chapman & Hall.
Fukunaga K (1990) *Introduction to Statistical Pattern Recognition*, 2nd edn. Boston: Academic Press.
Hand DJ (1997) *Construction and Assessment of Classification Rules*. Chichester: Wiley.
Jackson JE (1991) *A User's Guide to Principal Components*. New York: Wiley.
Kowalski BR (ed) (1984) *Chemometrics: Mathematics and Statistics in Chemistry* Dordrecht: Reidel.
Martens H and Naes T (1991) *Multivariate Calibration*. Chichester: Wiley.
McLachlan GJ (1992) *Discriminant Analysis and Statistical Pattern Recognition*. New York: Wiley.
Rencher AC (1995) *Methods of Multivariate Analysis*. New York: Wiley.
Ryan ThP (1997) *Modern Regression Methods*. New York: Wiley.
Sharaf MA, Illman DL and Kowalski BR (1986) *Chemometrics*. New York: Wiley.

**Relevant journals** *Analytical Chemistry Analytical Clinica Acta, Applied Statistics, Biometrics, Biospectroscopy, Chemometrics, Chemometrics and Intelligent Laboratory Systems, IEEE Transactions on Pattern Analysis and Machine Intelligence, International Journal of Pattern Recognition and Artificial Intelligence, Journal of Chemometrics, Journal of Classification, Journal of Multivariate Analysis, Journal of Royal Statistical Society, Neural Computation, Neural Networks, Neurocomputing, Pattern Recognition, Pattern Recognition Letters, Statistical Methods in Medical Research, Technometrics*

# Muon Spin Resonance Spectroscopy, Applications

**Ivan D Reid**, Paul Scherrer Institute, Villigen PSI, Switzerland
**Emil Roduner**, Universität Stuttgart, Germany

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

Spin-polarized positive muons, when injected into matter, can serve as magnetic probes for the investigation of various properties. The evolution of muon spin polarization rests on the same basis as conventional magnetic resonance techniques, but the methods involved in exploiting these probes are considerably different. The basic elements of muon spin resonance spectroscopy are described, with an emphasis on spectroscopic and kinetic applications in chemistry.

## The positive muon, a magnetic probe

Muon spin resonance spectroscopy is an offshoot of the experiment that proved parity nonconservation in pion and muon decay. The observation that the initial amplitudes and the relaxation rates of the muon precession signals depended on the nature of the stopping medium led to a new analytical method that is closely related to Fourier transform magnetic resonance.

The positive muon $\mu^+$ is an elementary particle that in the present context is best regarded as a light proton with a mass one-ninth that of the proton. Like the proton it is a spin-$\frac{1}{2}$ particle, but its magnetic moment is 3.18 times $\mu_p$. It is available for experiments as spin-polarized beams that can be implanted into most environments and then used to probe the local magnetic field, either as static spectators or by sampling a certain region as a mobile species. Because of its low mass and low linear energy transfer during the thermalization process, the muon causes relatively little damage near its stopping site. Nevertheless, it can probe radiation-chemical effects near the end point of its thermalization track.

In some environments the positive muon can capture an electron to form a hydrogen-like atom with the $\mu^+$ as a nucleus. Called muonium (Mu), this atom behaves chemically as a light hydrogen isotope. It is a paramagnetic species that can also serve as a static or diffusing magnetic probe. Some of the properties of muons and muonium are summarized in **Table 1**. Chemical reaction of Mu with unsaturated molecules leaves the muon as a polarized spin label in organic free radicals, for example

$$C_6H_6 + Mu \rightarrow C_6H_6Mu \quad [1]$$

Such species are routinely observed and can be characterized spectroscopically.

In comparisons of muons with protons and of muonium with hydrogen atoms, pronounced quantum effects occur whenever dynamics are involved. In this way, muons have been utilized to probe a large variety of properties and materials: insulators, semiconductors, metals, superconductors, insulators, gases, liquids, crystalline and amorphous solids, static and dynamic magnetic properties of all kinds, electron mobility, quantum diffusion, chemical reactivity and molecular structure and dynamics. The term adopted for the broad field of muon spin spectroscopy techniques, $\mu$SR, emphasizes the analogy with other types of magnetic resonance; for example EPR. '$\mu$S' represents 'muon spin', and 'R' in a more general sense stands simultaneously for 'rotation', 'relaxation' and 'resonance'.

## Muon production

Muons are produced artificially from pions formed by the collision of energetic protons with low-$Z$ targets. The pions decay with a lifetime of 26 ns:

$$\pi^+ \rightarrow \mu^+ + \nu(\mu) \quad [2]$$

The neutrino, ν, has negative helicity (angular momentum antiparallel to linear momentum), so from conservation laws the muon must also have negative helicity. By momentum selection, muon beams are obtained with a polarization close to 100%.

Two types of positive muon beams are possible: (1) surface muons, arising from pions decaying at rest close to the surface of the production target, with a momentum of 28 MeV/$c$ and a stopping range of 140 mg cm$^{-2}$ (corresponding to a water layer 1.4 mm thick); and (2) decay muons, from pions decaying in flight, having a selectable momentum normally in the range 70–130 MeV/$c$, and requiring a stopping range up to several centimetres of water. Surface muons are often preferred since they allow the use of thinner samples. They are also amenable to in-flight spin rotation where the crossed electric and magnetic fields of a Wien filter can be used to orientate their polarization with respect to their momentum.

**Table 1** Properties of $\mu^+$ and Mu

| | *Property* | *Symbol* | *Value*[a,b] |
|---|---|---|---|
| $\mu^+$ | Mass | $m_\mu$ | $1.883\,53\times10^{-28}$ kg = $0.1126\ m_p = 206.7864\ m_e$ |
| | Life time | $\tau_\mu$ | 2.197 134 μs |
| | Spin | $I$ | $\frac{1}{2}$ |
| | Magnetic moment | $\mu_\mu$ | $3.183\,344\ \mu_p$ |
| | Gyromagnetic ratio | $\gamma_\mu$ | 135.54 MHz T$^{-1}$ |
| Mu | Mass | $m_{Mu}$ | $0.0113\ 15 m_H$ |
| | Bohr radius | $a_0$ (Mu) | 0.053 17 nm = $1.0043 a_0$(H) |
| | Ionization potential | $I$ (Mu) | 13.539 eV = $0.9957 I$(H) |

[a] $\mu_p$ = magnetic moment of the proton.
[b] H = hydrogen.

Muon beams can be further classified on the basis of their time structure. Quasi-continuous muon sources allow high-precision measurement of the time interval between the arrival of the muon and its decay. However, the requirement in most experiments that each muon decay before the next is admitted to the sample limits the usable flux to about $10^5$ muons per second. Pulsed beams avoid this pile-up limit since a large number of muons are stopped simultaneously, followed by a long delay during which the decay of most muons can be accumulated, leading in principle to background-free histograms and a longer time window. Pulsed beams are also ideally suited for experiments that require coincidence with other pulses, e.g. laser or radiofrequency excitation.

## Detection of muon polarization

Conventional magnetic resonance uses radiofrequency techniques to detect the time evolution of spin polarization. This is impossible in muon spin resonance because of the extremely dilute concentration of the probes within the sample (often no more than one muon at a time). However, the decay of the muon

$$\mu^+ \rightarrow e^+ + \nu(e) + \bar{\nu}(\mu) \quad (\tau_\mu = 2.197\ \mu s) \quad [3]$$

does not conserve parity, leading to an anisotropic decay in which the positron ($e^+$) is emitted preferentially along the instantaneous direction of the muon spin at the moment of decay. Detection of the decay positron (and also of the muon's arrival in the case of time-differential experiments at a continuous-beam facility) involves standard particle-physics techniques, typically scintillating detectors coupled to photomultiplier tubes. Collection of an ensemble of correlated muon–positron events as a function of time and/or magnetic field in a given direction from the sample allows the evolution of muon polarization to be monitored.

Thus the basic methods of muon spin resonance differ considerably from those of conventional magnetic resonance spectroscopy. First, the probe is delivered to the sample with its spin polarization already prepared, and second the evolution of the

polarization is monitored statistically by collecting information on the decays of individual probes.

## Experimental techniques of muon-spin resonance

### Transverse-field muon spin rotation

The transverse-field $\mu$SR (TF-$\mu$SR) technique is shown schematically in **Figure 1A**. Muons enter the sample S either perpendicular to the magnetic field (as shown in **Figure 1A**) or (for a spin-rotated surface beam) along the field axis such that their polarization is perpendicular to the field. Each arrival is detected by a muon counter that starts a clock to measure the individual muon lifetime (at pulsed-beam facilities the start signal is synchronized to the muon pulse). The clock is recorded when the corresponding decay positron is detected in a positron counter, e.g. f or b, and the event is stored in a histogram $F(t)$ of decay positrons as a function of time spent by the muon in the sample. In the absence of any time evolution of the muon polarization, $F(t)$ is simply the radioactive decay curve (dashed line in **Figure 1A**) that corresponds to the muon decay. In a nonzero local magnetic field the muon will undergo a free induction decay (FID), similar to the response after a $\pi/2$ pulse in conventional magnetic resonance. Owing to the decay anisotropy (Eqn [3]) the FID appears in the histogram superimposed on the muon decay curve. $F(t)$ takes the shape indicated schematically in **Figure 1A**.

In many cases, different muons find themselves in different magnetic environments in the same sample, resulting in an FID with multiple frequencies. In this case it is convenient to work with the Fourier transform (FT) of the time-domain data, just as for conventional magnetic resonance. This is sometimes called FT-μSR.

Zeeman precession of the diamagnetic muon occurs at $\nu_\mu = \gamma_\mu B_i$ where the local magnetic field $B_i$ can deviate from the applied external field as a result of Knight shifts. Muonium, on the other hand, is observed mostly in low fields where two characteristic intra-triplet precession frequencies of the combined muon–electron systems occur. Below ~2 mT they become degenerate at $\nu_{Mu} = 13.94$ MHz mT$^{-1} \times B$. The two transitions are indicated in the Breit–Rabi diagram for Mu in **Figure 2** ($\nu_{Mu}$, solid arrows). Two further allowed transitions are of very high frequency

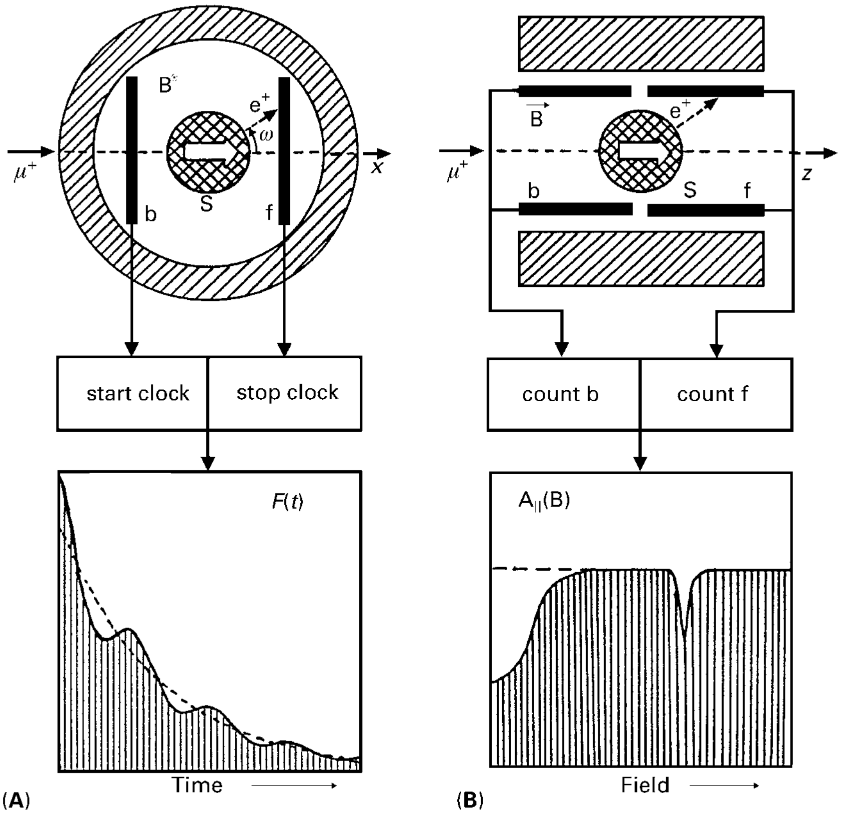



**Figure 1** Schematic drawing of the experimental apparatus used for muon spin resonance spectroscopy: (A) time-resolved transverse-, longitudinal- and zero-field studies; (B) time-integrated longitudinal fields (ALC).

and not normally observed (indicated by broken arrows). Muonated organic free radicals, however, are studied more easily in high fields where the muon transitions degenerate into two, independent of any other magnetic nuclei that may be present, corresponding to the first-order ENDOR frequencies,

$$\nu_{R\pm} = \left|\tfrac{1}{2}A_\mu \pm \nu_\mu\right| \quad [4]$$

These are indicated for Mu in **Figure 2**. The muon hyperfine coupling constant $A_\mu$ is obtained directly as the sum of the two observed frequencies (or the difference, in fields where $[\frac{1}{2}A_\mu - \nu_\mu]$ becomes negative). In general $A_\mu$ is in fact the component of a tensor $\tilde{\mathbf{A}}_\mu$ so its value depends on the orientation of the radical to the magnetic field $B$; however, most experiments are carried out with fluid samples where motion of the radical averages out the anisotropy, so just the isotropic value is observed.

### Zero-field and longitudinal-field muon spin relaxation (ZF-$\mu$SR and LF-$\mu$SR)

These experiments are performed in time-differential mode using the same setup as TF-μSR, except that the external field is absent or is applied parallel to the muon spin polarization in the beam (either by rotating the magnet or by removing spin rotation).

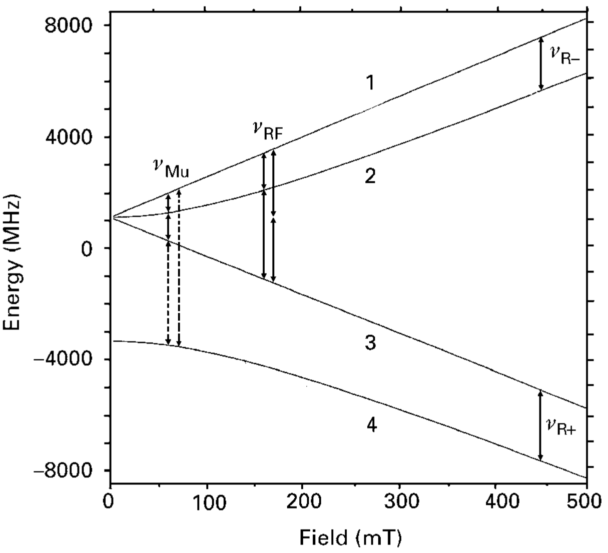



**Figure 2** Breit–Rabi diagram for muonium with transverse-field transitions indicated by arrows. The zero-field splitting corresponds to the muon hyperfine coupling constant ($A_\mu$ = 4463 MHz for Mu). In high fields, the four eigenstates become pure Zeeman states ($|1\rangle = |\alpha^\mu \alpha^e\rangle$, $|2\rangle = |\beta^\mu \alpha^e\rangle$, $|3\rangle = |\beta^\mu \beta^e\rangle$, $|4\rangle = |\alpha^\mu \beta^e\rangle$).

### Avoided-level-crossing resonance (ALC-$\mu$SR)

A basically different type of $\mu$SR takes advantage of the effect of avoided crossings of magnetic energy levels and is commonly performed in a time-integral mode as a function of longitudinal magnetic field (**Figure 1B**). Decay positrons are counted in the 'forward' ($N_f$) and in the 'backward' ($N_b$) direction with respect to the incoming muon beam. Of interest is the decay asymmetry

$$A_{\|}(B) = \frac{N_f - N_b}{N_f + N_b} \quad [5]$$

In sufficiently high fields, the eigenstates of any spin system are pure Zeeman states, and muons that are aligned with the external field are in an eigenstate where there is no time evolution of spin polarization, and $A_{\|}(B)$ is constant. Coupling of the muon to an unpaired electron, as in Mu or muonated radicals, leads to eigenstates that are mixtures of Zeeman states in low fields. The prepared non-eigenstate oscillates between eigenstates, resulting in partial depolarization of the muon and a strong field dependence of the detected decay asymmetry at low fields. The same effect operates at those high fields where there is an avoided crossing of two levels and eigenstates are mixtures of two Zeeman states. As **Figure 1B** indicates, sharp resonances are observed at these fields.

The resonances are distinguished by the selection rules $|\Delta M| = 0, 1, 2$, where $M$ is the quantum number for the $z$ component of the total spin. They appear to first order at resonance fields

$$B_r = \left|\frac{A_\mu + (\Delta M - 1)A_k}{2[\gamma_\mu + (\Delta M - 1)\gamma_k]}\right| \quad [6]$$

Thus, if $A_\mu$ is known, the nuclear hyperfine couplings $A_k$ are readily obtained, as are their relative signs, usually a difficult result in conventional magnetic resonance. The technique has been applied to semiconductors and to insulators, where dipolar and quadrupolar terms are determined in addition to the isotropic hyperfine couplings of Equation [6].

### Radiofrequency resonance techniques (RF-μSR)

RF-$\mu$SR resembles conventional CW (continuous wave) magnetic resonance, as it relies on resonance between an exciting RF field and a transition

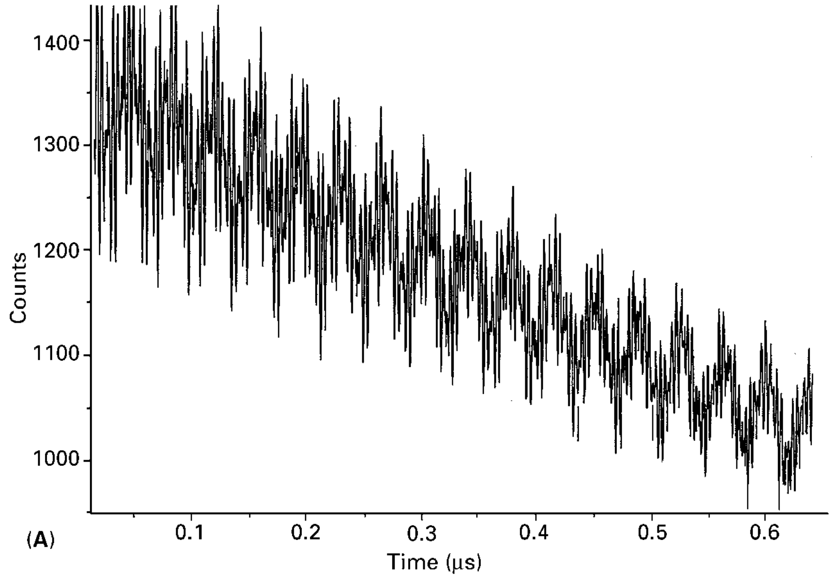
between two magnetic energy levels. The technique is of advantage, for example, for studying radicals in dilute solutions where the phase coherence of muon precession is lost in transverse fields during the formation process. The experiment is carried out in a longitudinal external field with a transverse RF field $B_1$. At high RF power a narrow double-quantum transition is observed in addition to the two power-broadened lines. The experiment can be conducted at constant field and variable RF frequency, or perhaps more conveniently at constant frequency and variable applied magnetic field as demonstrated nicely in a report on endohedral Mu and exohedral Mu adducts of fullerenes.

## Applications of muon spin resonance spectroscopy

To illustrate some of the experimental techniques outlined above, a selection of experiments using muon spin resonance spectroscopy are described.

### Spectroscopy of muonium-labelled organic free radicals

The prototypical sample used in $\mu$SR spectroscopy is benzene (see Eqn [1]). A raw histogram of the TF-$\mu$SR data obtained in a benzene sample at an applied field of 0.2 T is shown in **Figure 3A**. The precession



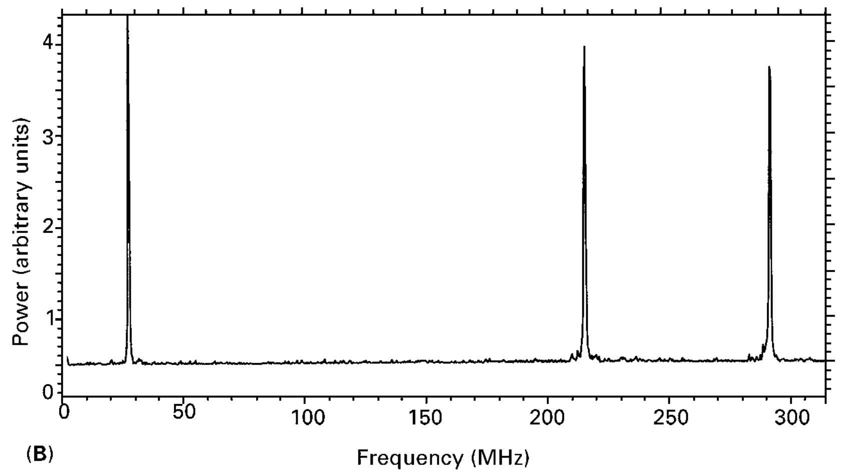




**Figure 3** TF-$\mu$SR spectra for positive muons implanted into benzene: (A) the raw time-differential histogram recorded for the experiment; (B) the Fourier transform of (A) showing the two transitions due to the $C_6H_6Mu$ radical (218.18 and 295.85 MHz, giving a hyperfine coupling constant of $A_\mu$ of 514.03 MHz) and the signal from muons in diamagnetic environments (27.1 MHz).

signal at 27 MHz arising from muons that have thermalized in a diamagnetic environment is easily seen, as is the overall decay curve due to the finite muon lifetime. **Figure 3B** shows the Fourier transform of the histogram, revealing not only the muon precession but also the ENDOR-like precession signals from muons which have reacted to form the free radical $C_6H_6Mu$.

Benzene is a simple example, since all addition sites are equivalent. However, consider 6,6-dimethylfulvene (**Figure 4**). Here there are four possible addition sites and thus potentially four radicals can be formed. **Figure 5A** shows the FT-$\mu$SR signal obtained in this compound at 0.3 T in ether solution. Two pairs of radical signals are seen, with hyperfine coupling constants of 105.2 and 203.5 MHz, respectively. To elucidate which radicals were observed, ALC spectra were also obtained for this compound (**Figure 5B**, lower trace). The resonances indicate where the muon's energy level has become degenerate with that of one of the H atoms in the radical (i.e. $\Delta M = 0$). Since the $A_\mu$ values are known from the FT-$\mu$SR, it is possible to devise a simulation to model the ALC spectrum, obtaining the various $A_p$ values within the

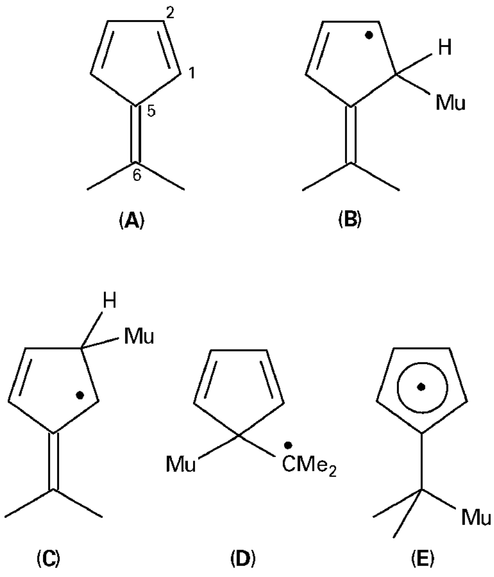



**Figure 4** The structure of 6,6-dimethylfulvene (A) and the four radicals (B–E) that may be derived from it by Mu addition. Reproduced with permission of the Royal Society of Chemistry from Rhodes CJ, Roduner E, Reid ID and Azuma T (1991) *Journal of the Chemical Society, Chemical Communications* 208–209.

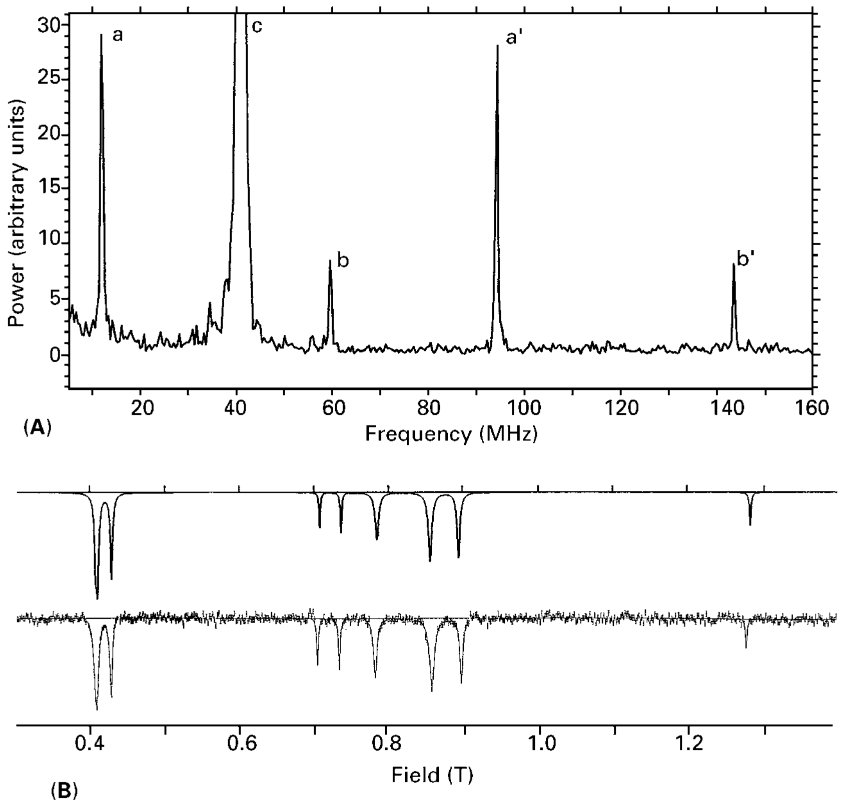



**Figure 5** Muon spin spectra from positive muons stopped in 6,6-dimethylfulvene: (A) FT-$\mu$SR spectrum showing the pairs of signals for each radical (a,a′; b,b′) and that from muons stopped in diamagnetic environments (c); (B) ALC spectra: lower trace, experimental data; upper trace, numerical simulation. Reproduced with permission of the Royal Society of Chemistry from Rhodes CJ, Roduner E, Reid ID and Azuma T (1991) *Journal of the Chemical Society, Chemical Communications* 208–209.

radicals, and to assign these values to specific sites. The results indicate that the two radicals are those formed by addition at the 1 and 2 positions in the ratio 1.4:1. The derived coupling constants and spin populations for these radicals are given in **Figure 6**.

Note that the ratio of the reduced muon coupling $A'_\mu$ ($A'_\mu = A_\mu (\mu_p/\mu_\mu)$) to the proton coupling at the same site is somewhat higher than unity. This is typical for muonated organic free radicals, the isotope effect normally lying between +5% and +15%. This can be understood as a simple consequence of the larger zero-point vibrational amplitude of the lighter isotope. Since the potential well for vibration is anharmonic, the average bond length is also larger. In other words, the muon and electron are moved towards the configuration of a free Mu atom, where $A_\mu$ is much higher than in the radical, and so their hyperfine coupling constant is correspondingly increased.

As noted earlier, the hyperfine coupling constant $A_\mu$ is in fact the component of a tensor $\tilde{\mathbf{A}}_\mu$ with a value dependent on the orientation of the radical to the magnetic field $B$. Whereas in solution the anisotropic part of the tensor is averaged out, in a crystalline solid the full anisotropic nature can be seen. **Figure 7** shows FT-$\mu$SR spectra from positive muons in naphthalene. The spectrum in (A) was obtained in acetone solution; the four lines are readily attributed to α and β radicals as indicated in the figure. The spectrum in (B) results from a large (~25 mm) single crystal – it is quite evident that the simple solution spectrum has been replaced by a much more complicated system. Indeed, the four lines have been split into 32 because there are two molecules per unit cell, because the crystal environment halves the symmetry of the parent molecules, and because additions from above and below the molecule are also not symmetric. None the less, the angular variation of the signals as the crystal was rotated about three orthogonal axes was mapped out and the hyperfine tensor was calculated for each of the 16 radicals. From symmetry and physical arguments, each was assigned to a specific addition site within the crystal. **Table 2** gives the results obtained for the tensor components.

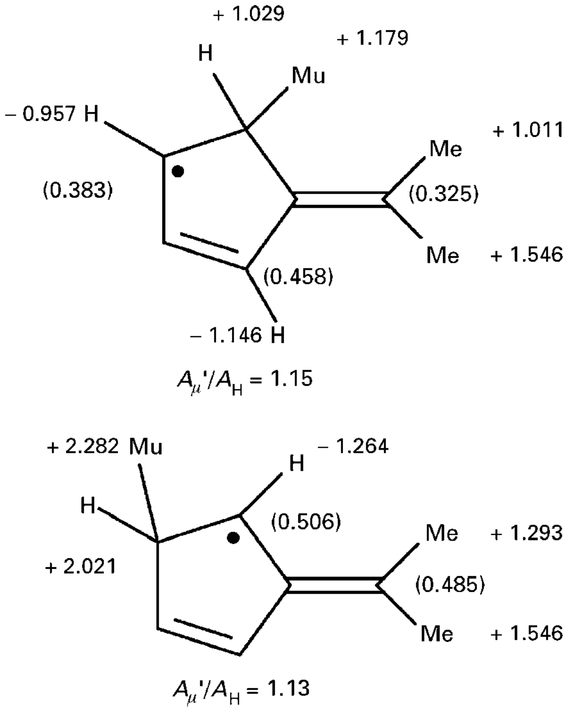



**Figure 6** Coupling constants in mT and (spin populations) for the two Mu adducts to 6,6-dimethylfulvene. Spin populations were obtained using $Q$ = –2.6 mT for α-protons and $Q$ = +2.925 mT for Me protons. Reproduced with permission of the Royal Society of Chemistry from Rhodes CJ, Roduner E, Reid ID and Azuma T (1991) *Journal of the Chemical Society, Chemical Communications* 208–209.

The orientation of the radicals within the crystal was deduced to be very similar to that of the parent radicals, but effects from the crystal field on the hyperfine coupling constants were quite evident, particularly for the β radicals; note particularly the differences in the isotropic components for addition from both sides of the molecule. Comparison with an ENDOR study of α-hydronaphthyl radicals showed that the earlier work had not detected all the possible radicals and that the two identified arose from addition at only one of the α sites; the ENDOR study did note see any β radicals. The reduced hyperfine coupling constants were ~10–20% larger than the proton values, although the comparison is not rigorous since the two studies were performed at greatly different temperatures.

As well as studies in condensed phases, $\mu$SR can also be carried out in the gas phase using either surface muons at moderate pressures or decay muons at higher pressures. Experiments with ethylene gas at room temperature and pressures of 25, 35 and 50 atm revealed FT-$\mu$SR signals that can readily be ascribed to the $CH_2Mu\dot{C}H_2$ radical. The resultant spectra are shown in **Figure 8**. The hyperfine coupling is 329.7 MHz, quite close to that found in liquid ethylene, and does not vary with pressure. The two radical signals are not of equal strength owing to the finite lifetime of the Mu precursor, which leads to dephasing of the muons and a corresponding loss of polarization. Calculations from the ratio of the polarizations of the two signals indicate a reaction rate coefficient of $16 \times 10^{-12}$ cm$^3$ molecule$^{-1}$ s$^{-1}$, compared to a value of $6 \times 10^{-12}$ cm$^3$ molecule$^{-1}$ s$^{-1}$ from relaxation measurements on the Mu signal in ethylene in a moderator gas at 295 K. The higher reaction

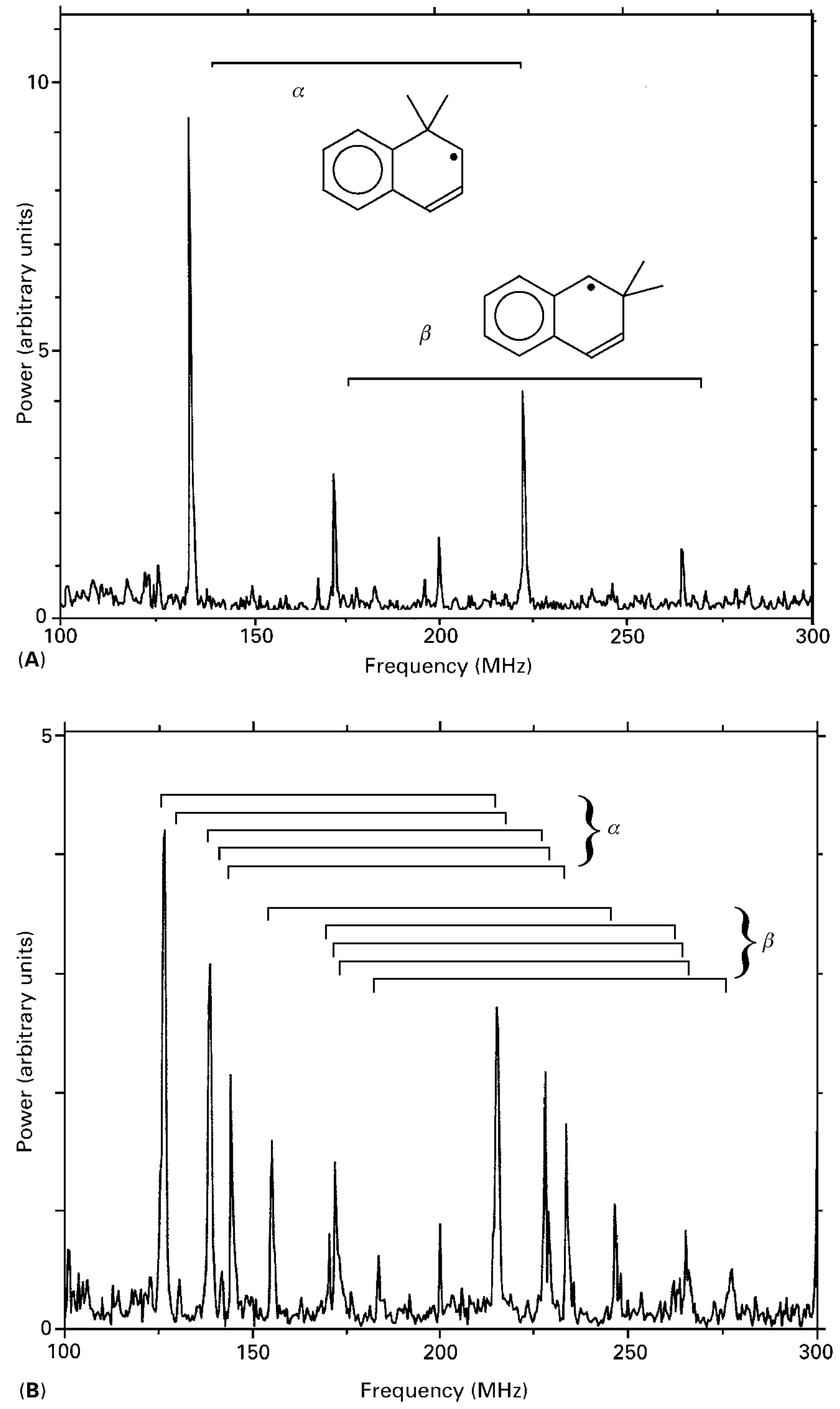



**Figure 7** FT-$\mu$SR spectra obtained for positive muons implanted in naphthalene samples: (A) in saturated acetone solution; (B) in a large single crystal. Reproduced with permission from Reid ID and Roduner E (1991) *Structural Chemistry* **2**: 419–431.

rate coefficient implies a Mu temperature of around 500 K, showing that it has not completely thermalized before the reaction with the ethylene. The broad nature of the radical signals has been explained in terms of the spin–rotational interaction within the radical as it undergoes collisions within the gas.

## Muonium and radical kinetics and dynamics

The mass ratio between Mu and H is unparalleled and pronounced isotope effects are to be expected. In reaction kinetics experiments, for example, considerable kinetic isotope effects (KIE) may be found.

Consider isotopic variations of the simple abstraction reaction

$$H + X_2 \rightarrow HX + X \qquad [7]$$

First, comparing the reaction with Mu instead of H,

**Table 2** Principal values of reduced hyperfine coupling tensors for Mu-substituted radicals in single-crystal naphthalene ($A'_{ii} = A'_{iso} + B'_{ii}$). Also given are the isotropic solution couplings and the tensors obtained for the $CH_2$ protons of analogous hydrogen radicals by ENDOR spectroscopy.

| *Nucleus*[a] | $A'_{iso}$ (MHz) | $B'_{11}$ (MHz) | $B'_{22}$ (MHz) | $B'_{33}$ (MHz) |
|---|---|---|---|---|
| CHMu muon $\alpha_a$ | 115.68 | –2.61 | –1.09 | 3.70 |
| CHMu muon $\alpha'_a$ | 108.75 | –2.64 | –1.06 | 3.70 |
| CHMu muon $\alpha_b$ | 117.98 | –2.94 | –0.95 | 3.89 |
| CHMu muon $\alpha'_b$ | 108.83 | –2.85 | –0.74 | 3.59 |
| CHMu muon $\alpha_{soln}$ | 112.17 | | | |
| CHMu muon $\beta_c$ | 128.71 | –2.82 | –1.69 | 4.51 |
| CHMu muon $\beta'_c$ | 143.54 | –3.11 | –1.27 | 4.38 |
| CHMu muon $\beta_d$ | 138.90 | –2.61 | –1.75 | 4.36 |
| CHMu muon $\beta'_d$ | 136.56 | –2.86 | –1.73 | 4.59 |
| CHMu muon $\beta_{soln}$ | 137.14 | | | |
| $CH_2$ proton 1 | 101.78 | –2.80 | –1.28 | 4.10 |
| $CH_2$ proton 1′ | 90.43 | –3.03 | –1.25 | 4.27 |

Reproduced with permission from Reid ID and Roduner E (1991) *Structural Chemistry* **2**: 419–431.

[a]The subscripts indicate the various inequivalent addition positions; a prime indicates addition from the opposite side of the molecular plane.

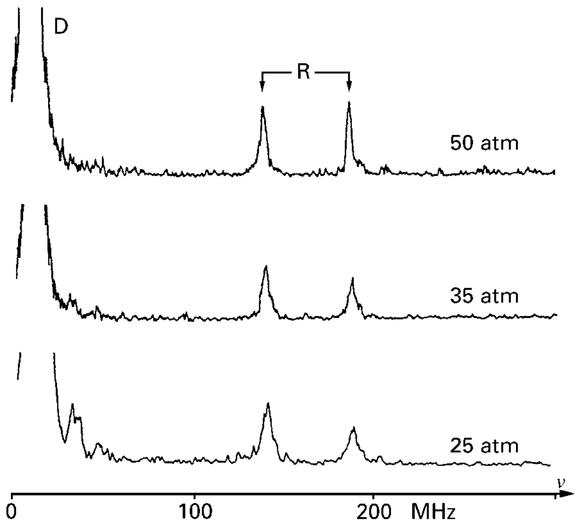



**Figure 8** FR-μSR spectra for ethylene gas at various pressures. D-muons in diamagnetic environments; R-muonated ethyl radical. Reproduced with permission of Baltzer Science Publishers from Roduner E and Garner DM (1986) *Hyperfine Interactions* **32**: 733–739.

one should expect a trivial KIE of about 3 owing to the lower mass and hence higher thermal velocity of Mu. However, the lower mass can also lead to greater zero-point vibrational energy effects in the activated complex and on the product side, causing a higher barrier for the Mu reaction so that reaction of the heavier isotope may be faster. Finally, quantum-mechanical tunnelling is often considered important in H-atom reactions, so that the lighter mass of Mu should favour it in crossing reaction barriers.

Compared to H-atom reaction studies, Mu reactivity experiments are reasonably simple. Positive muons are stopped, in a low (<0.8 mT) transverse field, in a moderator that produces a high yield of Mu, doped with varying concentrations $[X_2]$ of the reactant. Exponential relaxation of the coherent Mu precession signal is attributed to randomly occurring chemical reactions that transfer the muon to a different magnetic state where its precession frequency is significantly different. There is thus a linear relationship between the relaxation rate λ and $[X_2]$, leading directly to the thermally averaged bimolecular reaction rate constant $k$:

$$\lambda([X_2]) = \lambda(0) + k[X_2] \qquad [8]$$

where λ(0) is the relaxation rate for $[X_2] = 0$.

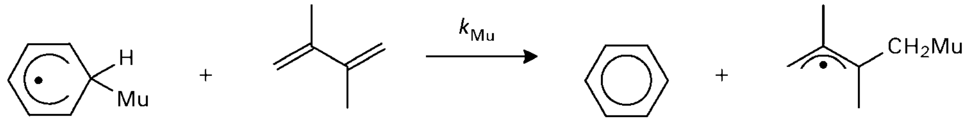
Many studies have been undertaken of such reactions, for example reaction of Mu with the halogen gases ($F_2$, $Cl_2$ and $Br_2$). For both $F_2$ and $Cl_2$ the data showed a significantly shallower slope in the Arrhenius plots of $k$ vs $T$ at temperatures below 200 K, clear evidence for tunnelling dominating the reaction in this region. This feature is absent from comparable H-atom reaction data. For $Br_2$, on the other hand, a negative slope was observed, indicating a negative activation energy $E_a$. In these highly exothermic reactions, the barrier is early on the reaction path and so the transition state approximates a slightly perturbed reactant molecule whose vibrations are only weakly dependent on the isotopic substitution. There may, however, still be a significant effect in the transmission coefficients for tunnelling, $\kappa_{H,Mu}$. **Table 3** summarizes some of the results obtained in this study and shows that there are indeed significant tunnelling effects even at the higher temperatures.

On the other hand, there are significant zero-point energy effects in the reaction of Mu with $H_2$ and $D_2$, since these reactions are highly endothermic ($\Delta H^0_R = 32$ and 38 kJ mol$^{-1}$, respectively) so the vibrationally adiabatic barrier is late. The potential energy surface of the $H_3$ system is known to a high accuracy, so comparison of the results for these

reactions with various theoretical treatments should help elucidate the extent to which the above effects are important and how well they can be predicted theoretically. Such measurements have been carried out over the temperature ranges 473–843 K and 598–843 K, respectively. Comparison with theoretical results showed that 3D quantum coupled states (CS) calculations of the reactions fit the data much better than improved canonical variation theory treatments, verifying the validity of both the potential energy surface and the CS treatment.

Studies of hydrogen KIE often involve reactions where a reagent attacks a H—R/D—R bond, transferring the isotope atom, whereas in Mu reactivity studies muonium is normally the attacking species. One example where Mu is thought to be transferred is the reaction of cyclohexadienyl radicals with 2,3-dimethyl-1,3-butadiene (DMBD).

In both pure benzene and pure DMBD, just one radical is observed via TF-$\mu$SR, with narrow lines ($\lambda_0$ <0.5 $\mu s^{-1}$). However, in binary mixtures of the compounds or in three-component mixtures with cyclohexane, the cyclohexadienyl lines broaden with increasing DMBD concentration while the line width of the allyl-type radical remains narrow. This obviously indicates a reaction of the cyclohexadienyl radical with DMBD. From the variation of line width with concentration (Eqn [8]) a reaction rate constant $k_{Mu}$ of $0.95 \times 10^6$ $M^{-1}$ $s^{-1}$ was obtained at 293 K. The protiated cyclohexadienyl radical has been observed in time-resolved EPR experiments, but no significant reaction with DMBD was seen, with an upper limit to $k_H$ of ~12 $M^{-1}$ $s^{-1}$. This yields an enormous KIE: $k_{Mu}/k_H \gtrsim 7.5 \times 10^4$ at room temperature. For such an affect Mu must be directly involved in the reaction and it is believed the scheme is:




Reaction of $C_6H_6Mu$ is favoured over $C_6H_7$ primarily because of its higher zero point vibrational energy, but tunnelling is also involved.

Other $\mu$SR methods can also be used to probe reactivity. For example, muon relaxation in longitudinal fields has recently been used to study the gas-phase reaction of the muonated ethyl radical with oxygen. Analysis of the results enabled a separation of the effects of chemical reaction and spin-exchange, leading to values of $k_{ch} = 8.4(3) \times 10^{-12}$ $cm^3$ $molecule^{-1}$ $s^{-1}$ and $k_{ex} = 2.8(2) \times 10^{-10}$ $cm^3$ $molecule^{-1}$ $s^{-1}$, respectively, at room temperature. It is striking that the chemical reaction rate is so much smaller than the spin-exchange rate, but this is explained by the anisotropy of the potential energy surface for chemical reaction leading to successful bond formation.

Dynamic effects in radicals also can be studied by $\mu$SR. Particularly useful is the $\Delta M = 1$ ALC-$\mu$SR resonance, which is only observable when there is anisotropy in the hyperfine coupling. Phenomenologically it can be thought of as occurring when the lower frequency $\nu_{R-}$ in Equation [4] passes through zero, i.e. a field where $\nu_\mu \sim \frac{1}{2}A\mu$ (Eqn [6]). Here a component of the hyperfine field is nullified by the external field and any anisotropy is seen by the muon as a transverse field, which causes it to precess and thus depolarize, leading to an observable resonance. Since the hyperfine tensor is anisotropic, this resonance should be seen unless the radicals are reorientating so fast as to average out the anisotropy. This was found to be the case for cyclohexadienyl radicals in a benzene monolayer on silica, for example, where the lack of the resonance down to 139 K was taken as evidence that the radicals remained mobile on the surface.

**Table 3** Reaction rate coefficients, kinetic isotope effects, and relative transmission coefficients for the reactions of H and Mu with the halogen gases.

| | $F_2$ | $Cl_2$ | $Br_2$ |
|---|---|---|---|
| $k_{298}$(Mu) ($10^{-11}$ $cm^3$ $molecule^{-1}$ $s^{-1}$ | 2.62±0.06 | 8.50±0.14 | 56.0±0.9 |
| $k_{298}$(H) ($10^{-11}$ $cm^3$ $molecule^{-1}$ $s^{-1}$ | 0.16±0.01 | 2.10±0.1 | 8.2±3.8 |
| $KIE_{298}$(Mu/H) | 16.4 | 4.0 | 6.8 |
| $\kappa_{Mu}/\kappa_H$ (298 K) | 5.7 | 1.4 | 2.3 |
| $\kappa_{Mu}/\kappa_H$ (250 K) | 8.0 | 2.1 | |
| $E_a$ (Mu) (kJ $mol^{-1}$) | 3.1± 0.3 | 2.7±0.2 | –0.4±0.8 |
| $E_a$ (H) (kJ $mol^{-1}$) | 9.2± 0.3 | 5.0±0.4 | |

Reproduced with permission from Gonzalez AC, Reid ID, Garner DM, Senba M, Fleming DG, Arseneau DJ and Kempton JR (1989) *Journal of Chemical Physics* **91**: 6164–6176.

A further study used this resonance to ascertain the anisotropic components of the hyperfine coupling in the radicals formed by irradiating bicyclo[2.2.1]hept-2-ene (norbornene) in its plastic phase. The tensor had already been derived from a single-crystal FT-$\mu$SR investigation but the ALC experiment was able to determine the (axial) anisotropy in a powder sample from the asymmetry of the shape of the resonance. The results agreed well with the FT-$\mu$SR data and confirmed effective partial averaging above $T_c$ (129 K) and a rapid approach to complete isotropy at $T_m$ (320 K).

## Potential and limitations of the muon as a probe

### Spin polarization and usable field range

Clearly, for a magnetic resonance type technique, the availability of a probe with a polarization close to 100% under all conditions is an immense advantage. Most techniques must work with Boltzmann populations, corresponding to polarizations of typically $10^{-5}$ for protons and $10^{-3}$ for electrons in commonly available magnetic fields and at room temperature. Muons have their full polarization independently of sample temperature and magnetic field and are therefore ideal probes, especially at low or zero field and at high temperatures.

Longitudinal fields up to about 7T have been used for ALC-$\mu$SR experiments. TF-$\mu$SR has been limited by the routinely available time resolution to ~3–4 T, corresponding to a Zeeman frequency of about 500 MHz. An important advantage over conventional magnetic resonance, however, is the availability of a fully polarized probe in *zero* field where additional information becomes accessible.

### Time resolution

Magnetic resonance techniques in time-resolved mode normally create transverse magnetization by application of a preparation pulse. This limits time resolution to typically 50 ns in EPR, and to several hundred nanoseconds in NMR, although a combination with optical techniques can break this limitation.

Since muons can be injected with their spins transverse to the external field, a preparation pulse is not necessary and time resolution is limited only to the accuracy with which $t_0$, the entrance time of the muon into the sample, can be measured. This is typically half a nanosecond when working with individual muons, and if needed it can be improved by a factor of about 5 at the cost of counting statistics.

### Frequency resolution

The muon lifetime of 2.2 μs sets a natural limit of 0.45 MHz on the widths of $\mu$SR lines. The nominal frequency resolution is given by the inverse length of the histogram. Using pulsed machines it is not uncommon to observe the FID over a time window of 20 μs, which leads to a nominal resolution of 0.05 MHz.

### Time window

The accessible time window for processes that can be observed directly extends between typically $10^{-9}$s and $10^{-5}$s. The limits are determined by the same parameters as time and frequency resolution. The windows accessible to NMR and EPR are thus extended considerably towards shorter times.

The time window can be extended to even shorter times if there is a muonium precursor state. Evolution of spin polarization in Mu occurs partly at frequencies near the Mu hyperfine frequency (4463 MHz in vacuum, broken arrows in **Figure 2**), which sets the timescale for loss of phase coherence during formation of the observed muonated species. The formation process can be studied indirectly in transverse fields by interpretation of shifts in the initial phase and concomitant loss of amplitude. Thus, processes occurring on a timescale down to 10 ps can be analysed, but the results rely on the validity of the underlying model.

### The muon as a local probe

In the same sense as nuclei in conventional magnetic resonance techniques, the muon is a local probe. While, for example, magnetization measurements give an integral response, the local magnetic probe may precess at different frequencies when it sits at different sites. The advantage of the muon is that it can be implanted in any material, including those that contain no other suitable NMR-active nuclei. For organic muonated radicals, the stopping site can be derived with confidence by analogy with the known chemistry of hydrogen atoms. In organic solids it is in general an interstitial site – this may be a nonbonding electron pair if oxygen or nitrogen are present. In semiconductors, some muons come to rest at bond centre sites. There are, however, many examples of materials where a firm determination of the stopping site is difficult.

It is important to consider to what extent the muon is also a perturbing probe. In comparison with the analogous proton or hydrogen defects, the differences are certainly small. The structure and properties of muonated radicals are very similar to those of their

hydrogen analogues and therefore well understood. On the other hand, a radical could be viewed as a strongly perturbed diamagnetic precursor molecule, which demonstrates how much the probe can change the electronic structure. In crystalline solids the properties measured are often not the same as those in the absence of the probe. It is clear, for example, that in a metal the positive charge of the muon causes charge polarization of the surroundings, leading to some structural relaxation.

### Sensitivity

Both the high spin polarization of the muon and the single-event detection technique lead to a high sensitivity of the muon spin resonance methods. The total number of muons needed for a routine $\mu$SR experiment in transverse fields is of the order of $10^7$, while X-band ESR at room temperature needs about $10^{11}$ unpaired electrons, and single-scan proton NMR needs on the order of $10^{17}$ spins to allow the observation of a narrow, unsplit signal. This sensitivity is particularly crucial for the observation of surface-adsorbed organic free radicals – it is difficult to maintain the minimum concentration for EPR detection under conditions where the radicals are mobile since they disappear via bimolecular termination reactions, whereas in TF-$\mu$SR one has a single muonated radical at a time in the entire sample. This allows adsorbed radicals to be observed up to temperatures where they desorb or undergo catalytic reactions. Termination reactions also represent a severe limitation to kinetic work in liquid solutions and in the gas phase, so the fact that the $\mu$SR technique guarantees ideal pseudo-first-order kinetics is clearly advantageous.

## List of symbols

$A_k$ = nuclear hyperfine coupling constant; $A_\mu$ = muon hyperfine coupling constant; $A'_\mu$ = reduced muon hyperfine coupling constant; $\tilde{\mathbf{A}}_\mu$ = muon hyperfine coupling tensor; $A_\parallel$ = decay asymmetry; $B$ = magnetic field strength; $B_{ii}$ = anisotropic components of the hyperfine coupling; $e^+$ = positron; $E_a$ = activation energy for reaction; $k$ = reaction rate constant; $k_{ch}$ = chemical reaction rate constant; $k_{ex}$ = spin-exchange reaction rate constant; $M$ = spin quantum number; $N_b$ = 'backward' decay positron count; $N_f$ = 'forward' decay positron count; $T_c$ = critical temperature for phase change; $T_m$ = melting point; $\Delta M$ = change in spin quantum number; $\Delta H_R^0$ = reaction enthalpy; $\gamma_\mu$ = muon gyromagnetic ratio; $\kappa$ = tunnelling transmission coefficient; $\lambda$ = relaxation rate; $\mu^+$ = positive muon; $\mu_p$ = proton magnetic moment; $\mu_\mu$ = muon magnetic moment; $\nu(e)$ = electron neutrino; $\nu(\mu)$ = muon neutrino; $\bar{\nu}(\mu)$ = muon antineutrino; $\nu_\mu$ = muon precession frequency; $\nu_{Mu}$ = muonium precession frequency; $\nu_R$ = radical precession frequency; $\tau_\mu$ = muon lifetime.

*See also:* **Chemical Applications of EPR; EPR, Methods; EPR Spectroscopy, Theory; Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy.**

## Further reading

Cox SFJ (1987) *Journal of Physics C: Solid State Physics* **20**: 3187–3319.

Davis EA and Cox SFJ (1996) (eds) *Protons and Muons in Materials Science*. London: Taylor & Francis.

Roduner E (1986) *Progress in Reaction Kinetics* **14**: 1–42.

Roduner E (1988) *The Positive Muon as a Probe in Free Radical Chemistry. Potential and Limitations of the μSR Techniques*, Vol. 49 in Lecture Notes in Chemistry. Heidelberg: Springer.

Roduner E (1993) *Chemical Society Reviews* **22**: 337–346.

Roduner E (ed) (1997) *Applied Magnetic Resonance* **13**.

Schenck A (1985) *Muon Spin Rotation Spectroscopy*. Bristol: Hilger.

Walker DC (1983) *Muon and Muonium Chemistry*. Cambridge: Cambridge University Press.

# N

# Near-IR Spectrometers

**R Anthony Shaw** and **Henry H Mantsch**, Institute for Biodiagnostics, National Research Council of Canada, Winnipeg, Manitoba, Canada

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**

**Methods & Instrumentation**

## Introduction

While near-infrared spectroscopy was for many years a sleeping giant, the diversity of instrumentation available today attests to its new-found and widespread acceptance as an analytical method. The aim of this article is to outline the measurement principles for each of the various types of instruments that are commercially available as of this writing, and to outline the strengths of each. Because analysis is by far the most common application, we also include accounts of specialized measurement techniques that have emerged for on-line, in-situ, and remote spectral measurements, as well as briefly outlining dedicated analysers founded upon near-infrared technology.

The near-infrared (near-IR) falls between the visible and mid-infrared regions, with corresponding vibrational frequencies in the terahertz range (**Figure 1**). It has become customary to report near-IR absorption positions in units of wavelength, either micrometres or nanometres, although it is becoming more common to see positions measured in wavenumbers ($cm^{-1}$). Simply the inverse of the wavelength (in cm), the wavenumber scale has become standard for mid-infrared spectroscopy and is becoming more commonly used for near-IR work. The advantages most often cited are that the scale is linear in energy and that it provides values of a convenient order of magnitude, particularly in the mid-infrared.

It seems odd at first glance that the near-IR region remained largely an afterthought for many years, given the relatively long and productive histories of both UV-visible and mid-infrared spectroscopy. However, the weak bands in the intervening near-IR region were long viewed as being superfluous; with broad, weak absorptions corresponding almost exclusively to overtones and combinations of vibrational modes, it was (and still is) generally considered that the structural information latent in the spectra is far more readily available from the mid-infrared spectra.

Times change. The 1980s and 1990s have witnessed revolutionary growth in the use of near-IR spectroscopy as an *analytical* technique. Near-IR spectrometers may now be found in applications ranging from agriculture to food processing and pharmaceuticals. Estimated at \$4 500 000 in 1977, near-IR instrumentation revenues exceeded the \$100 000 000 mark in 1997 and continue to grow at a compound rate of 20% annually. This demand has spurred the development of a variety of near-IR spectrometers ranging from dedicated analysers to research-grade systems of enormous versatility. Remote measurements may be carried out in conjunction with fibre optics, and, from a vantage point aboard the Hubble telescope, even more distant measurements are being carried out as the near-IR camera and multi-object spectrometer 'NICMOS' yields a new vision of the universe. Application of the same spectroscopic-imaging technology is yielding new insights here on Earth, and the next 20 years will be eventful ones as this technique begins to live up to its potential.

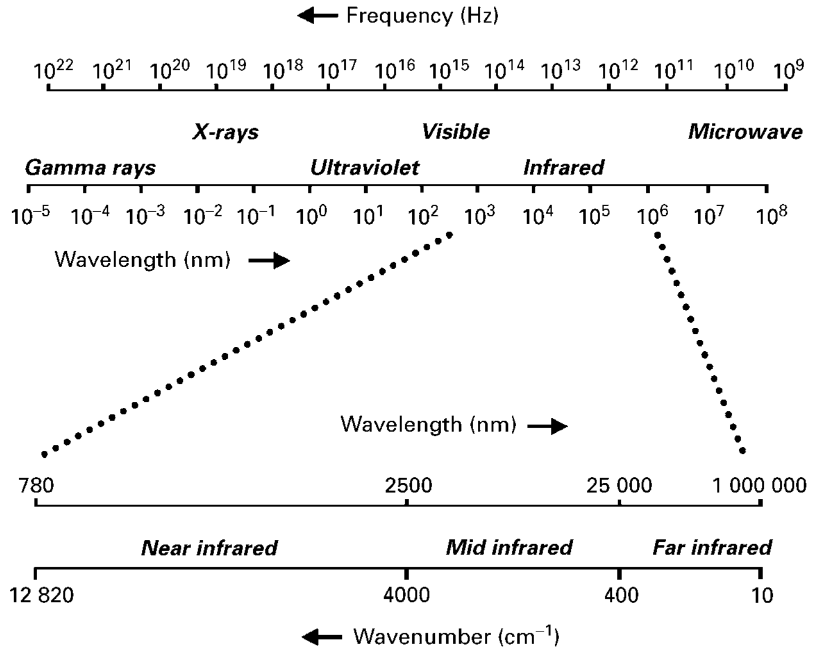



**Figure 1** The near-infrared lies between the visible and mid-infrared regions, ranging from 780 to 2500 nm.

## Instrumentation: an overview

### Historical development

In 1800, William Herschel assembled a spectrometer in which the Sun served as the source of radiation, a prism as the dispersing element, and a thermometer as the detector. Having first measured the distribution of radiant heat across the visible region, he then took the unprecedented step of moving the thermometer to a position beyond the red end of the spectrum. To his surprise, the thermometer not only registered heat, but a higher temperature than that found in the visible region. This signalled the discovery of a 'non-visible spectrum', and coincidentally of the first near-IR spectrometer.

The idea of a spectrometer dedicated to near-IR spectroscopy is relatively recent, and the early commercial near-IR instruments were simply UV-visible (or mid-infrared) spectrometers fitted with an additional detector and occasionally a second grating blazed for the near-IR. This equipment provided the basis for the pioneering work of Kermit Whetzel and Wilbur Kaye, who were largely responsible for laying the foundation of analytical near-IR spectroscopy. The development of modern near-IR instrumentation was spurred by research at the US Department of Agriculture; Karl Norris discovered that no commercial spectrometer of the time could provide diffuse reflectance measurements of the quality he required, and developed his own computerized near-IR spectrometer for meat analysis. The original aim of this work was to develop a convenient means to monitor the water content of agricultural products. The research was complicated by the observation that components other than water contributed to the absorption profiles; however, this observation was quickly turned to advantage when it was found that protein could be quantified accurately from the near-IR spectrum of wheat. Under the direction of Phil Williams, the near-IR method was adopted by the Canadian Grain Commission to replace routine Kjeldahl testing. It worked, and spurred the widespread adoption of near-IR spectroscopy as an analytical method, and with it the birth of dedicated near-IR instrumentation.

## What is a near-IR spectrometer?

It is easy to find a pair of near-IR spectrometers that would not be recognizable as belonging to the same species even placed side by side. This diversity has arisen largely as a result of the tremendous variety of applications, with each unique application satisfied by a unique set of design and performance characteristics.

Some of the spectral criteria that characterize and differentiate near-IR spectrometers are wavelength accuracy and precision; photometric accuracy, precision, and linearity; signal-to-noise; and resolution (or spectral bandwidth). Among these, the attributes

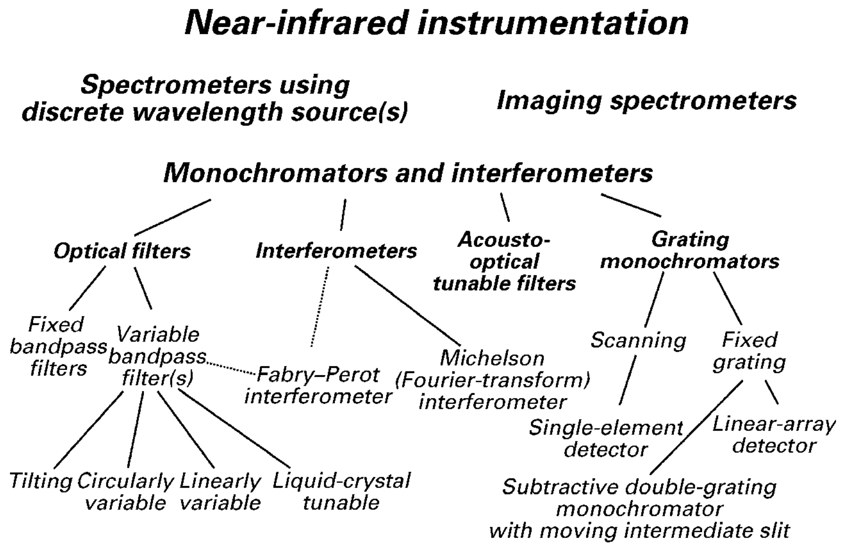



**Figure 2** Near-infrared instrumentation.

that are most important for analytical work are wavelength precision, photometric precision, and signal-to-noise. While the last may occasionally be sacrificed, it is generally the case that the signal-to-noise is the limiting factor governing the analytical accuracy; spectra are commonly measured with noise at single-digit micro-absorbance levels.

Practical considerations almost always play a role in choosing a near-IR spectrometer, reflecting specific demands that may be placed by a specific application. Among these considerations are size and/or portability; speed; susceptibility to extreme (e.g. industrial) environments; cost; measurement type; and flexibility (e.g. can the spectrometer be used only for transmittance measurements, or are reflectance and/or fibre optic measurements feasible?). Manufacturers consider all of these factors. For example, it is common to find commercial instrumentation housed in enclosures that are ruggedized to meet specific industry standards (National Electrical Manufacturers Association – NEMA).

For convenience, the spectrometers discussed in this article have been divided into the six design categories and variations illustrated in **Figure 2**. The instrumentation will be discussed after first introducing two essential components, the source and the detector.

## Near-IR sources

The most common broadband light source used for near-IR spectroscopy is the tungsten-halogen lamp, typically powered by a stabilized DC power supply. This source provides high energy throughout the near-IR, and has a long life with the halogen acting as a scavenger to deposit tungsten vapour back onto the filament.

## Detectors

No single detector covers the entire 780–2500 nm near-IR range. The detectors most commonly used for near-IR spectroscopy are lead sulfide (or lead selenide) photoconductors and silicon photodiodes, with PbS(Se) covering the region 1100 to 2500 nm and Si the visible to 1100 nm range. A huge advantage of these detectors is that both provide good signal-to-noise operation at room temperature. **Figure 3** illustrates the specific detectivities ($D^*$) as a function of wavelength for PbS, PbSe, Si, and other detectors. As illustrated, both PbS and PbSe become more sensitive at lower temperatures, and some spectrometer manufacturers do capitalize on this. The alternative detector most commonly used at wavelengths longer than 1100 nm is indium gallium arsenide (InGaAs), which extends to 1800 nm with a substantially better $D^*$ than either PbS or PbSe. At the expense of some sensitivity, the range of the InGaAs detector may be extended still further to 2200 nm, or even 2600 nm. Thermoelectric cooling is commonly used to increase $D^*$, but this improvement comes at the expense of long-wavelength response (**Figure 3**).

Fixed-grating spectrometers require a linear detector array for the simultaneous detection of the spectral elements dispersed along a line, and imaging applications require a two-dimensional detector array for spatial resolution. While both types of array are available for all of the detector materials included in **Figure 3**, those most commonly used are silicon, InSb and InGaAs.

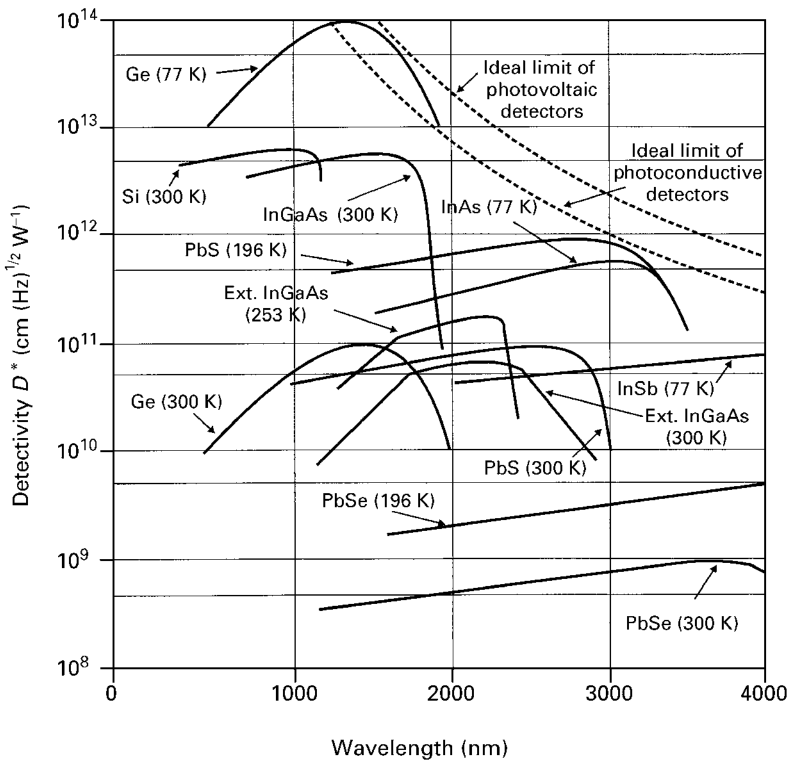



**Figure 3** Specific detectivities and wavelength ranges for near-infrared detectors.

## Near-IR instrumentation

The following sections are intended to illustrate the measurement principles for commercial near-IR spectrometers, and to provide an understanding of the strengths inherent in each design. Because near-IR spectroscopy and chemical analysis are inextricably linked, particular emphasis is placed on features relevant to this application.

### Grating instruments

**Scanning-grating spectrometers** In this category, it is important to distinguish spectrometers that are designed explicitly for quantitative analytical applications from those that are not. Several manufacturers offer conventional scanning single-monochromator spectrometers that extend from the ultraviolet region through the visible and near-IR. The main attraction of these instruments is the wide-wavelength coverage. A second monochromator may be included, which reduces the stray light and increases spectral resolution; these instruments have extraordinary photometric range, with linear response to 6 absorbance units and beyond. For analytical work however, these UV-visible near-IR spectrometers do not match the performance – in particular the combination of speed and signal-to-noise – that dedicated near-IR instrumentation provides.

On the other hand, commercial rapid-scanning monochromators that incorporate both PbS and Si detectors are capable of scanning the range 400–2500 nm in about 1 s. Typically 30 to 60 scans are co-added to provide a spectrum with very low noise levels (typically less than 30 microabsorbance units) within a minute or so. There are a wide variety of applications that require this combination of speed and accuracy, and the relatively high cost of these instruments is fully justified for these applications. These spectrometers are also found in research laboratories. The role of the 'high-end' research spectrometer is to test the feasibility of new analytical applications and, if successful, to provide a basis to judge how best to implement the method in the field. Is the full spectrum required, or will discrete wavelengths suffice? What spectral regions (and hence detectors) are essential? What is the minimum signal-to-noise that will still yield acceptable analytical results? It is often the case that the only way to answer these questions is to carry out a careful study using a full-range research-grade spectrometer.

**Fixed-grating spectrometers** The near-IR spectrum may be obtained by using a grating spectrometer with no moving parts at all. Since the grating disperses the spectrum of the source along an axis, the spectrum may be measured by an array of detector elements spaced along that axis (**Figure 4**). The most common detectors are silicon photodiode (PDA) and

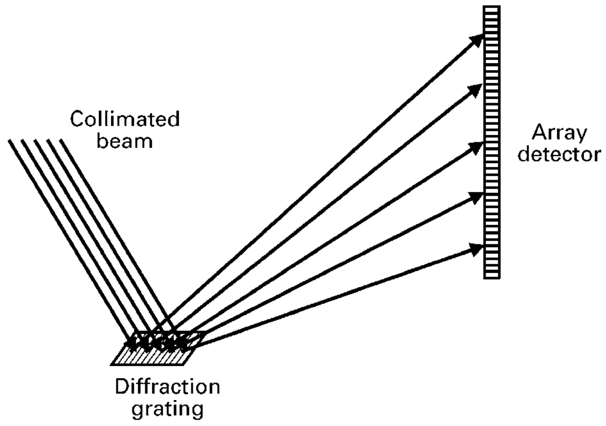



**Figure 4** The fixed-grating array detector. The grating spreads the spectrum across the array of detector elements, so that each element senses a different wavelength.

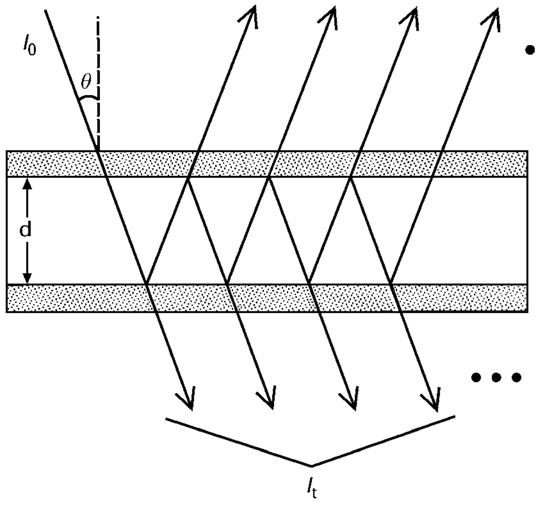



**Figure 5** The interference filter. The profile of the transmitted intensity $I_t$ is an Airy function whose bandwidth is governed by the reflectivity of the dielectric-coated reflective layers.

charge coupled device (CCD) arrays, with thermoelectrically cooled InGaAs arrays available for operation above 1100 nm. One advantage of this design is that it may be miniaturized to the point where the spectrometer is truly portable; these 'mini spectrometers' typically accept light via a single strand fibre optic cable of 10–1000 μm diameter, with resolution of 0.5 to 10 nm depending on the choice of grating and slit. They are inexpensive enough that several configurations, optimized for various resolutions and spectral bandwidths, may be linked in parallel by running several fibres from the same source.

It is possible to combine the mechanical stability of fixed gratings with a single-element detector. One manufacturer provides a subtractive double-monochromator system with the wavelength scanned by displacement of movable intermediate slits. Using slits cut in a rotating disk, this arrangement is capable of acquiring up to 1000 spectra per second. While the design was developed primarily for kinetic studies in the visible region, extension to the near IR is readily achieved.

### Interference filters and the Fabry–Perot interferometer

A number of commercial instruments make use of a set of interference filters to send discrete bands of radiation successively through the sample. In their simplest form, these filters appear as illustrated in **Figure 5**. Flat plates coated with dielectric reflecting layers are placed in parallel. Incident light undergoes multiple reflections in the region between the two plates, and the transmitted components interfere constructively or destructively as a function of (i) incident angle $\theta$, (ii) gap thickness '*d*', (iii) refractive index of the material between the reflective plates, and (iv) wavelength '$\lambda$'. If the light is incident normal to the surface of the filter the transmission profile has a maximum centred at a wavelength $\lambda_{max}$ equal to twice the gap thickness '*d*', with higher-order transmission maxima at $\lambda_{max}/2$, $\lambda_{max}/3$, $\lambda_{max}/4$, etc. The desired order is selected using appropriate high-pass and low-pass filters.

Quite elaborate filters may be built upon this principle, with bandwidths ranging from 0.25 up to 10% of the centre wavelength. A typical instrument includes several filters mounted on a rotating filter wheel (**Figure 6**). The wheel may include a set of filters with bandpasses tailored to fit the requirements of a specific application or, if the instrument is to be used for a variety of applications, a set of filters to survey the near-IR region of interest.

Because of the angular dependence of the transmission profile, the peak transmitted wavelength may be fine tuned by tilting the filter relative to the incident beam. In fact, this property has been capitalized on to scan small spectral regions using a single filter, and one commercial instrument used seven tilting filters of appropriate centre wavelengths to cover the 1400–2400 nm region. However with the advent of inexpensive holographic gratings, the tilting-filter instrument is no longer manufactured commercially.

The same interference principle may be applied to make a circularly-variable filter. The cavity thickness, and hence the centre wavelength, varies systematically around the circumference of the filter. Rotating the disk in a position between the source

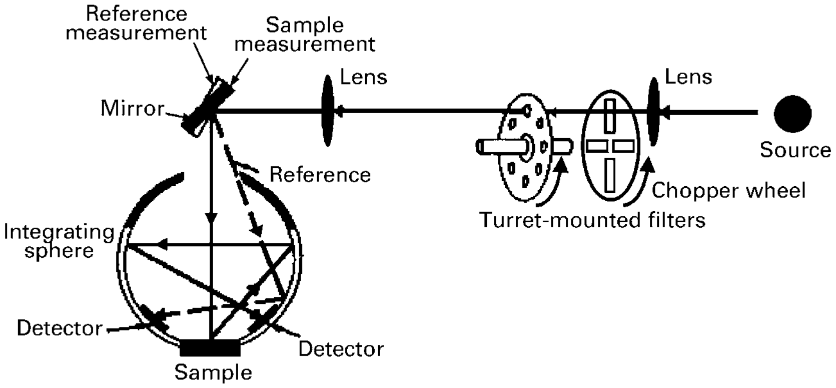



**Figure 6** A filter-based instrument operating in reflectance mode. The sample is probed with the mirror in the 'sample' position, and the beam is directed to the internal surface of the integrating sphere to measure the background. Courtesy of Bran+Luebbe Analyzing Technologies Ltd.

and the detector scans the wavelengths. Similarly, the cavity thickness may vary systematically along a line to produce a linearly variable filter suitable for use with an array detector. While neither of these has found wide application for analytical work, both find niches in applications where a broad bandpass is not an issue.

The Fabry–Perot spectrometer is an interference filter in which the centre transmitted wavelength is scanned either by varying the gap between the parallel plates, or by varying the refractive index of the medium separating them. Again, this design is not widely used for analytical work, primarily due to practical difficulties in manufacturing devices that work in first order across the near-IR region. Devices operating at higher order may have extraordinarily high resolution; however the free spectral range is correspondingly narrow.

The optical path of the liquid-crystal tunable filter comprises a series of polarizers and liquid crystal elements whose birefringent properties are electronically controlled. These devices can select wavelengths ranging from the visible to 1100 nm, with a bandwidth as low as 5 nm. The wide, circular aperture makes it particularly attractive as a means to select wavelength regions for transmission to a near-IR camera.

## The Michelson interferometer

The Michelson interferometer does not measure the infrared spectrum directly. Rather, an 'interferogram' is measured, and converted to a single-beam spectrum via Fourier transformation. Because of the critical role of this transformation, the method is generally referred to as Fourier-transform infrared spectroscopy, or 'FT-IR'. Instruments using this design have recently proven to be suitable for a number of near-IR applications.

The optical layout is illustrated in **Figure 7**. As the moving mirror is displaced, there is an increasing difference in the optical pathlength travelled by the two components incident on the detector. This results in interference between the two merging beams, and for a monochromatic source of wavelength $\lambda_1$ the resulting AC signal (a mirror is moving!) is a cosine wave. The Fourier transform of this interferogram is an infinitely narrow 'band' (delta function) at a position $1/\lambda_1$. (FT-IR spectra are typically reported in

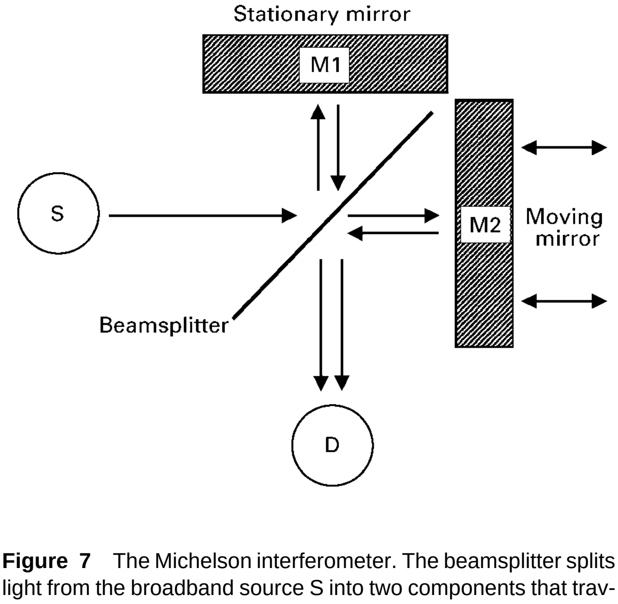



**Figure 7** The Michelson interferometer. The beamsplitter splits light from the broadband source S into two components that travel separate paths to mirrors M1 and M2. The recombined beams incident on the detector D move progressively further and further out of phase as the moving mirror is scanned. The interferogram is a plot of the AC component of the detector signal *vs* the path difference.

wavenumbers ($cm^{-1}$), the inverse of the wavelength (and, not coincidentally, the inverse of the unit of length measuring the optical-path-difference that defines the horizontal axis of the interferogram). The scale is linear in energy, and the resolution is a constant in these units.) For a polychromatic source the ideal interferogram is a superposition of cosine waves, one for each 'colour' in the spectrum of the source; the Fourier transform converts this to the emission spectrum of the source. The absorption spectrum is acquired by placing a sample before the detector in the optical path, repeating the measurement and ratioing the two single-beam spectra.

Although they have long been the design of choice for mid-IR spectroscopy, FT-IR instruments have only recently become popular in analytical near-IR applications. This emergence is due in large part to improvements in the interferometer design and materials, resulting in modern instruments that are very stable both in the short and the long term. Because the position of the moving mirror and the digitization interval are monitored very accurately by a parallel helium–neon laser interferometer, the FT-IR spectrometer is also characterized by superb wavelength (wavenumber) accuracy and precision.

The Michelson interferometer combines very good signal-to-noise with high spectral resolution. The resolution is governed by the maximum displacement of the moving mirror, and a resolution of 4 $cm^{-1}$ or better is readily achieved. For comparison, a bandpass of 10 nm is typical for grating near-IR spectrometers; the relatively wide slit is preferred in order to transmit enough energy to quickly record spectra of good signal-to-noise. **Figure** 8 graphically illustrates the resolution advantage of a Michelson interferometer operating at 4 $cm^{-1}$ or even at 16 $cm^{-1}$ nominal resolution. While the practical advantages of the higher resolution remain largely unexplored, there is little doubt that the enhanced ability to distinguish closely spaced absorptions will prove beneficial in some analytical applications.

Two final points specific to FT-IR spectrometers should be mentioned. First, the signal-to-noise can often be improved by using an optical bandpass filter to isolate the spectral region of interest. Second, this approach is particularly well suited for measurements where the light flux on the detector is low, for example for measurements using fibre optics where the light-gathering efficiency may be quite poor.

### Acousto-optical tunable filters

If an acoustic wave is produced in a suitable crystal, the result is that the refractive index varies periodically across the crystal. Intuition would then suggest that the crystal might act as a grating with diffraction properties governed by the wavelength of the sound wave.

In fact the crystal does act as a monochromator, but the mechanism underlying this phenomenon is more complex than the simple diffraction grating analogy would indicate. In fact the process is

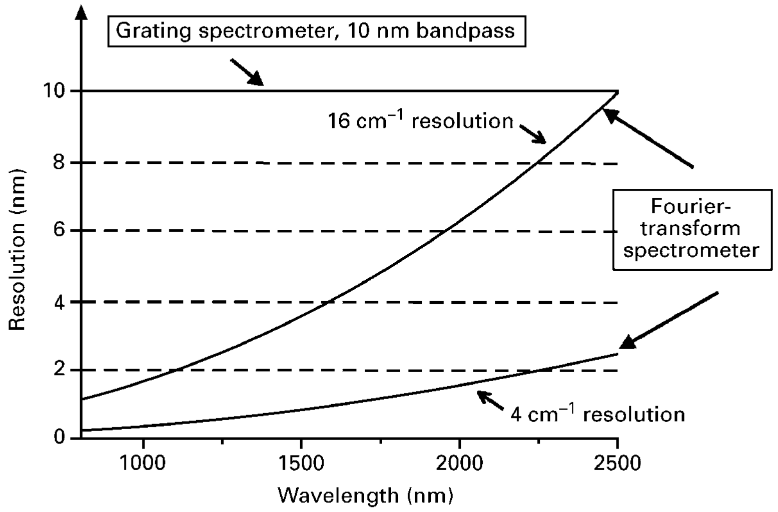



**Figure 8** Comparison of the resolution of a FT-IR (Michelson) spectrometer to that of a grating spectrometer. The solid lines represent the resolution of a grating spectrometer with a bandpass of 10 nm and the resolution (in nm) for a FT-IR spectrometer operating at a nominal resolution of either 4 $cm^{-1}$ or 16 $cm^{-1}$ (because noise levels rapidly increase for narrower slitwidths, a 10 nm bandpass is typical of analytical rapid-scanning grating spectrometers). For example, a FT-IR spectrometer operating at 4 $cm^{-1}$ resolution has an effective bandpass of 2.5 nm at 2500 nm, falling to less than 0.5 nm at 800 nm.

governed by a phonon–photon scattering mechanism the details of which lie beyond the scope of this article. The result, however, is that a tellurium oxide crystal with one or more piezoelectric transducers (PZTs) bonded to it acts as a monochromator (**Figure 9**). Wavelength selection is achieved by varying the driving frequency of the PZTs, and the intensity may also be varied by varying the driving power. The resolution is governed by the physical size of the crystal (or, more precisely, by the length of the path over which the light/crystal interaction takes place).

Two distinctive features of this design are the fact that it has no moving parts and that there is 'random access' to any pre-selected wavelength or set of wavelengths. Wavelength stability, precision, and accuracy are also very good, since the wavelength selection is controlled by the RF drive. But the most striking feature of the AOTF design is the speed of measurement.

### Spectrometers using discrete-wavelength sources

Light-emitting diodes (LEDs) are available with centre wavelengths throughout the NIR range, and are used as the basis of near-IR instruments with no moving parts. Because LEDs typically have bandwidths of 50 nm or larger, a small interference filter is required to select the desired centre wavelength and bandwidth. An appropriate set of filtered LEDs may then be assembled to span the wavelength range of interest, and the spectrum acquired by cycling through the set. Two advantages of this design are that the instrument can easily be miniaturized, and that the LED sources are stable for decades. Commercial instruments are in essence portable analysers that happen to be based upon near-IR spectroscopy, with built-in calibrations to give direct read-out of analyte levels.

### Imaging spectrometers

Every picture tells a story. This is as true in the near-IR as it is of the visible, and the combination of imagery and spectroscopy is a marriage with a short track record but enormous potential.

In its simplest form the measurement requires a near-IR camera, typically equipped with a two-dimensional silicon array detector, with some means to select wavelengths. Some of the advantages of spectroscopic imaging may be realized by simply using fixed bandpass filters to capture images at two or more discrete wavelengths; two judiciously chosen wavelengths often complement one another to provide information that is not available from either of the individual images.

The full potential of near-IR imaging spectroscopy is realized by replacing the fixed bandpass filters with a monochromator having a continuously variable bandpass. Several possibilities exist. The Fabry–Perot interferometer is suitable for scanning over small wavelength ranges at high resolution, and is used primarily for astronomical observations. Two elements most commonly used for imaging across wide spectral regions are the acousto-optic tunable filter and the liquid-crystal tunable filter. **Figure 10** illustrates the kind of data that can be measured

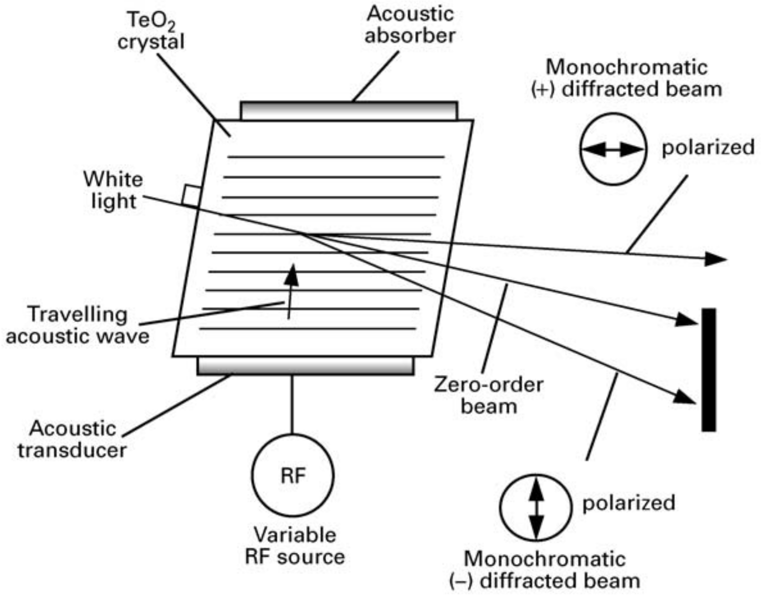



**Figure 9** Acousto-optical tunable filter. Courtesy of Brimrose Corporation.

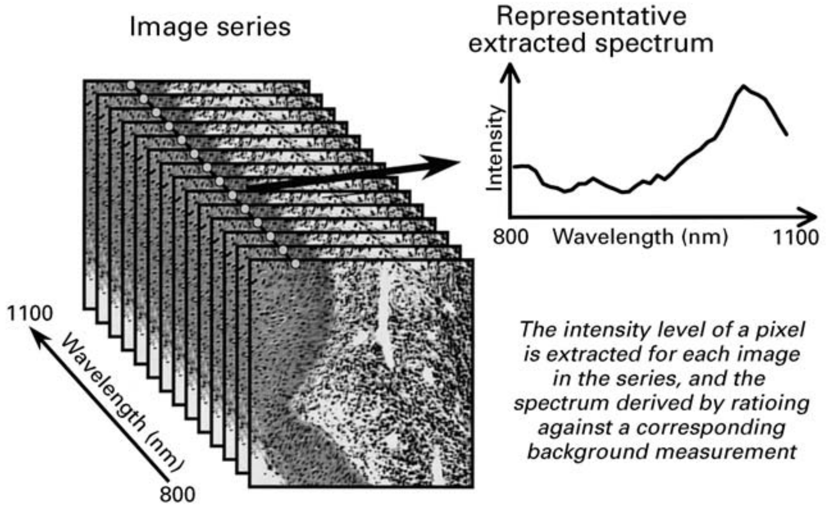



**Figure 10** Near-infrared spectroscopic imaging. A series of *N* images is acquired using a CCD-array camera, using optical filtering to step successively through the near infrared (800–1100 nm in this illustration). An *N*-point near-infrared spectrum may then be reconstructed for each pixel in the image. Courtesy of Jim Mansfield.

using this type of arrangement; a series of images is acquired at discrete wavelengths, and a spectrum then constructed for each pixel by extracting the corresponding intensity value for each image in the series. The analogous experiment has been carried out using a Fourier-transform spectrometer, using an infrared microscope in combination with an InSb focal-plane array detector. In that instance, an interferogram is collected at each pixel and transformed to yield the near-IR spectra.

## Sampling methods: from the laboratory to on-line

### Diffuse reflectance

Analytical applications of near-IR spectroscopy can only succeed if the sample can be presented to the spectrometer in a reproducible fashion. This is simple enough for liquids, but less straightforward for solid, physically inhomogeneous samples. The most common solution is to measure radiation that is diffusely reflected from the sample. **Figure 11** illustrates a typical configuration for this measurement, with two detector elements at 45° to the surface. Another way to collect the reflected radiation is to use an integrating sphere, as illustrated in **Figure 6**. In either event, the reflectance spectrum of a coarse, inhomogeneous sample is influenced to some degree by the surface terrain. This influence is typically minimized either by spinning the sample or translating the surface of the sample stepwise across the beam. It should be mentioned that the Michelson interferometer is not well suited for reflectance measurements of coarse moving samples. The moving sample modulates the intensity of the reflected light, giving rise to a noise component that superimposes on the interferogram. The result upon Fourier transformation is a noisy spectrum.

### Fibre optics, on-line and in-situ measurements

It is often desirable to measure near-IR spectra for samples that for reasons of convenience or necessity cannot be transported to the spectrometer for 'conventional' transmission or reflectance measurements. Examples range from processes such as fermentation and refining, where on-line monitoring is desirable,

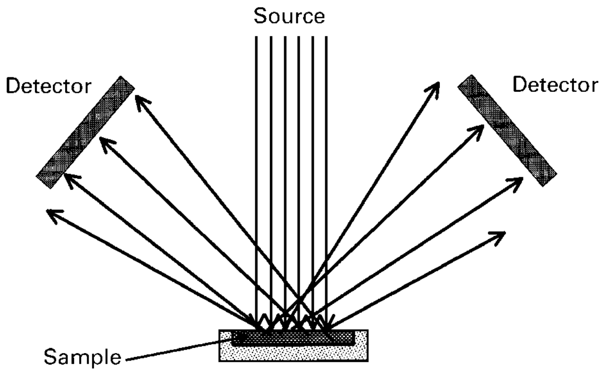



**Figure 11** Typical geometry used for diffuse-reflectance spectroscopy. The detectors may be the same materials, or two different materials (e.g. Si and PbS) for wider wavelength coverage.

to the analysis of pharmaceutical blends and incoming raw materials. A common solution is to use fibre optics or fibre-optic bundles to carry the spectral information from the sample to the spectrometer. Silica fibres are ideal for this purpose, as they are transparent from the visible through the near-IR.

One way to use fibre optics is for transmission measurements, with one bundle carrying near-IR radiation to the interface of a transmission cell, and a second bundle gathering the transmitted radiation and carrying it back to the spectrometer. Specialized fibre-optic probes exist for in-situ transmission measurements with the 'sending' and 'receiving' fibres in the same bundle. The immersion tip includes built-in optics to create a reproducible transmission path; light that has travelled through the sample is reflected back into the receiving fibres by a mirror built into the probe tip itself.

While transmission measurements are suitable for liquids, powders and other highly scattering samples require a different approach. One measurement simply uses a fibre-optic bundle with a measurement tip that exposes both sending and receiving fibres to the specimen. While many of the photons entering the sample are lost, a fraction of them are scattered in a path that leads them back into one of the receiving fibres. The background spectrum is typically reflectance from a ceramic standard. The 'interactance' spectrum is obtained by ratioing the single-beam powder spectrum against the single-beam ceramic reference spectrum, and reflects both the absorption properties of the sample and the wavelength dependence of the scattering efficiency. Probes based upon this general principle are suitable for checking the identity/purity of batch powder samples. The operator simply inserts the probe tip into the sample and waits for the analytical result to appear before moving onto the next measurement.

#### Remote reflectance

For many production-line applications, the most convenient way to measure the near-IR spectrum is by 'remote' diffuse reflection. The source and specialized collection optics are combined in a single device and simply aimed at the target, for example chemicals or food products moving along conveyor lines or webs and sheets in the production process. Just as is the case for laboratory measurements, these instruments vary in sophistication from single-wavelength devices to full grating spectrometers – the choice of the appropriate system is dictated by the difficulty of the analysis. Because water has extraordinarily strong absorptions in the near-IR, one of the more common on-line applications is monitoring moisture content; more than one manufacturer offers near-IR filter-based analysers exclusively for this purpose.

## Summary

From its original niche as a convenient method for grain analysis, near-IR spectroscopy has evolved to the point where new applications appear almost daily. The parallel evolution in hardware has led to modern instruments that could not have been contemplated by the founding fathers of near-IR analysis. Many devices faithfully produce analytical results running unattended around the clock, while others are used by operators who need not even be aware that near-IR spectroscopy is involved. In short, the technology has reached a certain maturity.

What does the future hold? In addition to ever-faster and more accurate instruments, one exciting prospect is the full realization of the marriage of spectroscopy with imaging. Practical roadblocks to the widespread adoption of this technique are rapidly disappearing as both imaging arrays and high-throughput tunable filters become more widely accessible, and the realm of possibilities seems limitless. This and the other extraordinary instruments that appear ever-more frequently are vital to our gathering and understanding the information that nature has to offer, and the next twenty years will surely be as eventful as the past twenty.

## List of symbols

$D^*$ = detectivity; $I$ = intensity; $\lambda$ = wavelength.

*See also:* **Fibre Optic Probes in Optical Spectroscopy, Clinical Applications; IR Spectrometers; IR Spectroscopy Sample Preparation Methods; IR Spectroscopy, Theory; Medical Science Applications of IR.**

## Further reading

Burns DA and Ciurczak EW (eds) (1992) *Handbook of Near-Infrared Analysis*. New York: Marcel Dekker.

Davies AMC and Williams P (eds) (1996) *Near Infrared Spectroscopy: The Future Waves*. West Sussex, UK: NIR Publications.

Davies AMC (ed) *Journal of Near-Infrared Spectroscopy*. (This journal regularly includes research articles reporting refinements and advances in near-IR instrumentation.)

Griffiths PR and de Haseth JA (1986) *Fourier Transform Infrared Spectroscopy*. Toronto: Wiley-Interscience.

McLure WF (1994) Near-Infrared spectroscopy: the giant is running strong. *Analytical Chemistry* **66** 43A–53A.

Murray I and Cowe IA (eds) (1992) *Making Light Work: Advances in Near Infrared Spectroscopy*. Weinheim: VCH.

Osborne BG, Fearn T and Hindle PH (1993) *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. Harlow, England: Longman Scientific & Technical.

Wetzel DL (1998) Contemporary near infrared instrumentation. In: Williams P and Norris K (eds) *Near-Infrared Technology in the Agriculture and Food Industries*, 2nd edn. St. Paul, MN: The American Association of Cereal Chemists.

# Negative Ion Mass Spectrometry, Methods

**Suresh Dua** and **John H Bowie**, The University of Adelaide, Australia

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

## Introduction

Investigations of both negative-ion and positive-ion mass spectrometry as aids to structure determination commenced in earnest in the middle of the twentieth century. The ease of formation of molecular cations, together with their characteristic fragmentations, led to the development of positive-ion mass spectrometry as one of the most important of all analytical techniques. In contrast, initial difficulties with the formation of suitable anions together with the early nonavailability of commercial mass spectrometers that could readily detect negative ions delayed the analytical development of negative-ion mass spectrometry for several decades. Modern instrumentation allows the measurement of the mass spectra of molecular anions ($M^{\bullet-}$), and of $(M–H)^-$ ions and multiply charged anions. The analytical applicability of the fragmentations of $(M–H)^-$ species will be outlined. Mechanistic information that can be obtained from gas phase studies with respect to rearrangement reactions of anions, and the collision-induced conversion of negative ions into neutrals and positive ions and the application of these techniques will also be described. Selected examples will be used to illustrate each section. References to reviews that provide further information concerning aspects of anion chemistry are listed at the end of the article.

## Analytical applications of negative ions

### The fragmentations of $(M–H)^-$ species

$(M–H)^-$ ions may be formed by deprotonation of organic molecules in the chemical ionization (NICI) source of a mass spectrometer using a strong base (e.g. $HO^-$ from $H_2O$ or $NH_2^-$ from $NH_3$), or using fast-atom bombardment (FAB), electrospray

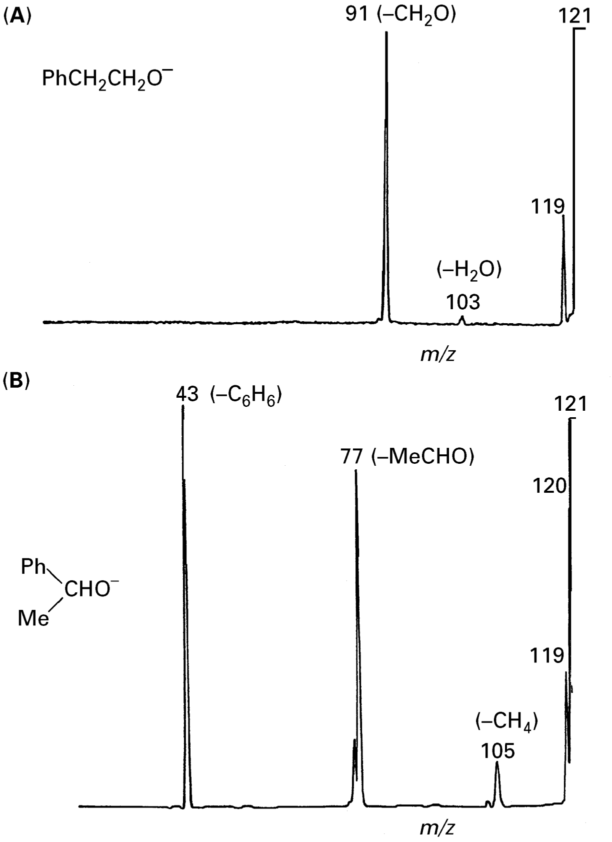



**Figure 1** Collision-induced $HO^-$ negative-ion chemical ionization tandem mass spectra (MS/MS) of (A) $PhCH_2CH_2O^-$ and (B) $PhCH(Me)O^-$. ZAB 2HF instrument. Argon collision gas: pressure of gas in first collision cell adjusted so that the reduction in the main beam is 10%.

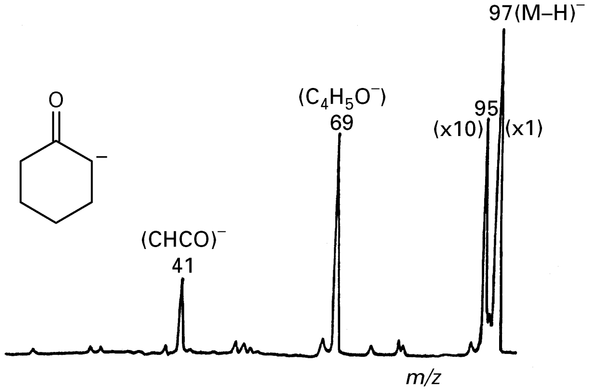
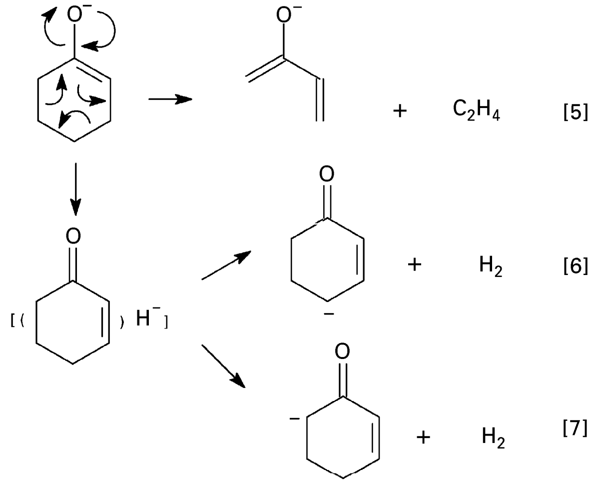



**Figure 2** Collision-induced $HO^-$ negative-ion chemical ionization tandem mass spectra (MS/MS) of deprotonated cyclohexanone. ZAB 2HF instrument. Details as for **Figure 1**.

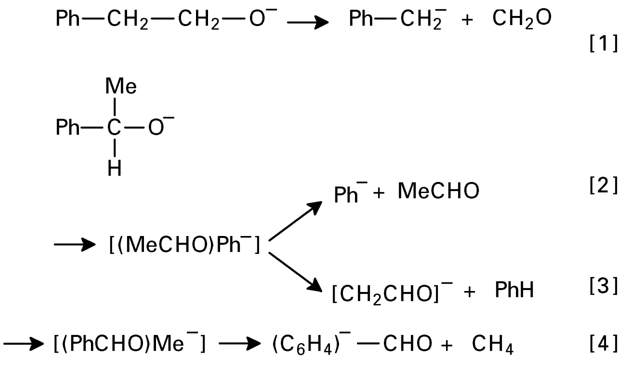



**Scheme 2**

ionization (ESI), atmospheric pressure ionization (API), matrix-assisted laser desorption ionization (MALDI) mass spectrometry, and a number of associated ionization techniques. The $(M–H)^-$ ions are generally formed with little excess energy and consequently undergo little decomposition. Molecular mass information can be obtained in this way for the majority of organic compounds, including those that do not form detectable molecular cations or where peaks from such cations are of small abundance in the positive-ion spectra (examples of the latter category include many long-chain compounds such as alcohols, ketones, acids and esters, and also some peptides and polysaccharides). Fragmentation data of $(M–H)^-$ ions may be obtained using collisional activation or some other method of energy activation of the parent anion. Ionization and energy activation techniques mentioned above are described in other articles.

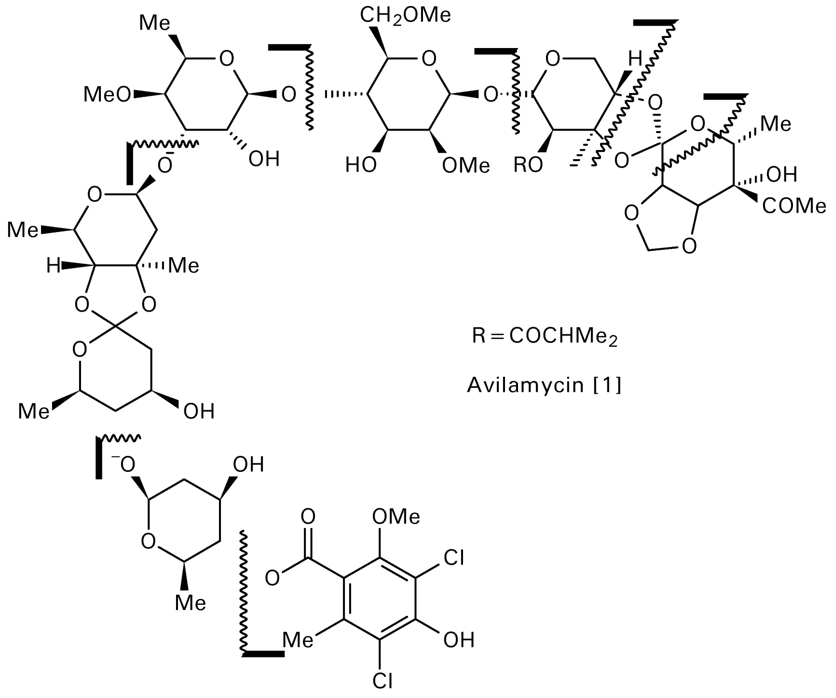
**Table 1** Characteristic negative-ion fragmentations of side chains of amino acid residues from $(M–H)^-$ ions of peptides

| *Residue* | *Loss ( or formation)* | *Mass* |
|---|---|---|
| Phe | $PhCH_2^-$ | 91 |
| Tyr | *p*-$HOC_6H_4CH_2^-$ | 107 |
| | $O{=}C_6H_4{=}CH_2$ | 106 |
| Trp | $C_9H_7N$ | 129 |
| Ser | $CH_2O$ | 30 |
| Thr | MeCHO | 44 |
| Cys | $H_2S$ | 34 |
| Met | MeSH | 48 |
| | MeSMe | 62 |
| | $^•CH_2CH_2SMe$ | 75 |
| Asp | $H_2O$ | 18 |
| Glu | $H_2O$ | 18 |
| Asn | $NH_3$ | 17 |
| Gln | $NH_3$ | 17 |



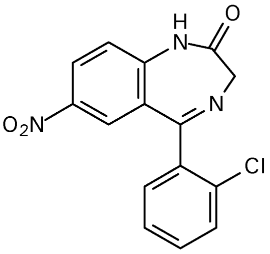

**Scheme 1**

The negative-ion fragmentations of $(M–H)^-$ ions derived from organic molecules are often simple and provide useful structural information. Fragmentations involving particular functional groups are often characteristic of those groups. Consider first the spectra of the two isomeric alkoxide anions shown in **Figure 1**. These spectra are collision-induced $HO^-$ NICI tandem mass (MS/MS) spectra that have been measured in a reverse sector mass spectrometer. The negative-ion fragmentations are rationalized in **Scheme 1**. They include simple cleavage (Eqns [1] and [2]), together with reactions of an anion within an anion neutral complex (Eqns [3] and [4]). A second example is shown for the $(M–H)^-$ parent ion of cyclohexanone in **Figure 2** cleavage mechanisms are outlined in **Scheme 2**. Fragmentations include a retro cleavage of the ring (Eqn [5]) together with two competitive losses of dihydrogen (Eqns [6] and [7]).



R = $COCHMe_2$

Avilamycin [1]



[2] Clonazepam

Many biologically important molecules form $(M-H)^-$ parent ions, and these often undergo characteristic fragmentations. For example, $(M-H)^-$ ions from molecules containing C–O and P–O bonds (Such as nucleosides, nucleotides and molecules containing saccharide linkages) undergo collision-induced cleavage reactions to yield stable $-O^-$ fragment anions. Negative-ion mass spectrometry is often the analytical method of choice for such molecules. A particular example is shown above for avilamycin [**1**], where C–O bond cleavages (the bold lines indicate the direction of charge retention) provide sequence information.

Positive-ion mass spectrometry has traditionally been the MS method of choice for sequencing peptides and proteins. However, the fragmentations of both $(M-H)^-$ and $(M-nH)^{n-}$ ions of peptides and proteins can also provide useful structural information. The negative-ion spectra of $(M-H)^-$ ions show two types of cleavages: (i) characteristic fragmentation of the side chains of some amino acid residues (these data are summarized in **Table 1**), and (ii) cleavages of the peptide backbone. As an illustration, MS/MS data from the $(M-H)^-$ parent ion of a 12-residue peptide are shown in **Figure 3**. The sequence of this peptide can be determined by the α and β negative-ion backbone cleavages shown in **Figure 3** (see **Scheme 3** for the nomenclature and mechanisms of α and β backbone cleavages). In addition, the ion formed by side-chain cleavage of $CH_3CHO$ from the Thr side chain (cf. **Table 1**) also produces backbone cleavage ions: these are indicated by arrows in **Figure 3**.

Further negative-ion backbone cleavages are initiated by the enolate anion of an Asp or Asn side chain. This is illustrated in the spectrum of the nonapeptide shown in **Figure 4**, with the characteristic Asp backbone cleavages rationalized in **Scheme 4**. The usual α and β cleavage ions are indicated on the formula shown in **Figure 4**.

## Radical anions

Parent radical anions may be formed following low-energy electron capture by a neutral if the electron affinity of the neutral is suitably positive. Particular examples are conjugated systems (such as α-diketones, α, β-unsaturated carbonyl systems, quinones and flavones, etc.) together with molecules containing specific functional groups that readily capture electrons (e.g. nitro, sulphonyl, phosphate esters, etc). For example, the limit of detection of the clonazepam [**2**] molecular anion is about 25 times lower than that of the corresponding molecular cation.

In some cases derivatization of a neutral with a functional group that readily accepts a low-energy electron can provide molecular mass information: for example, a long-chain alcohol can be converted into a perfluorobenzoate ester that readily yields a molecular anion.

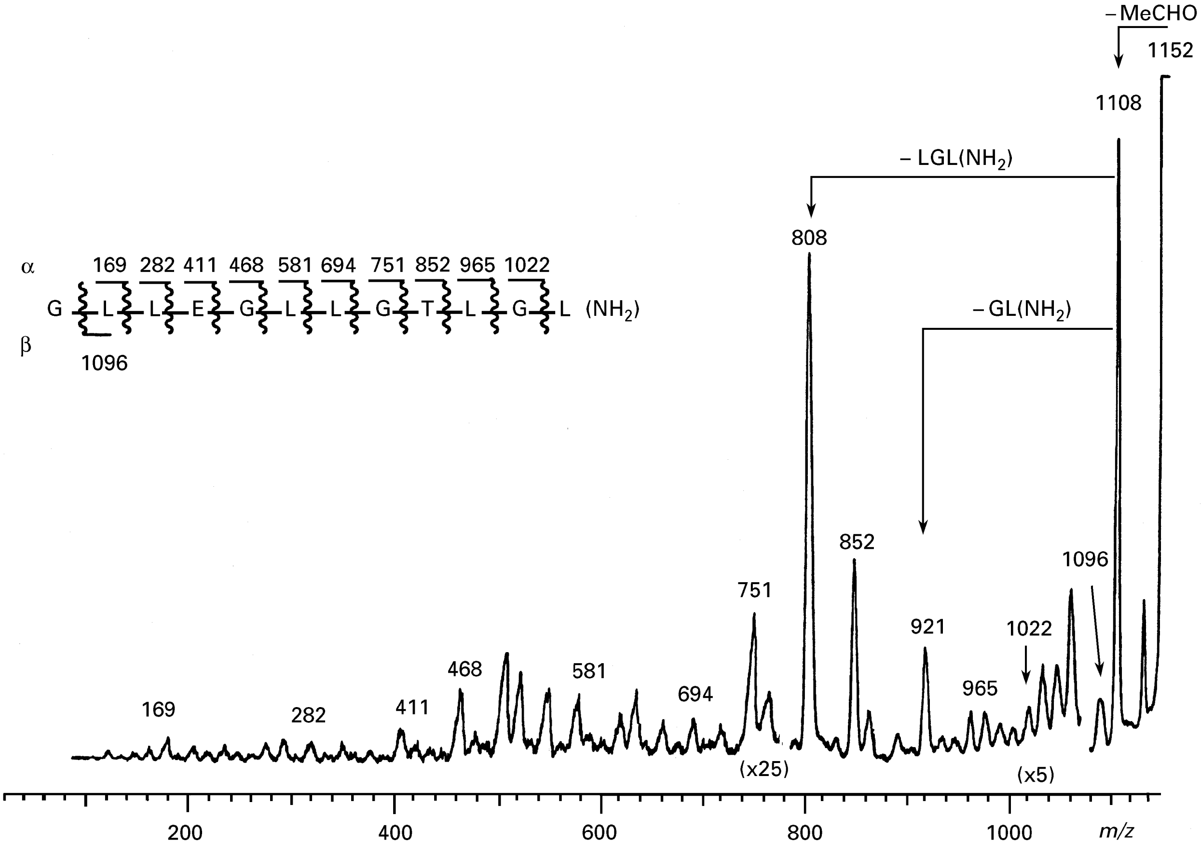
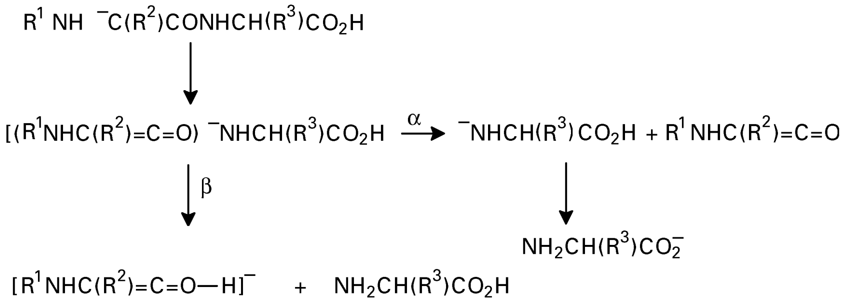
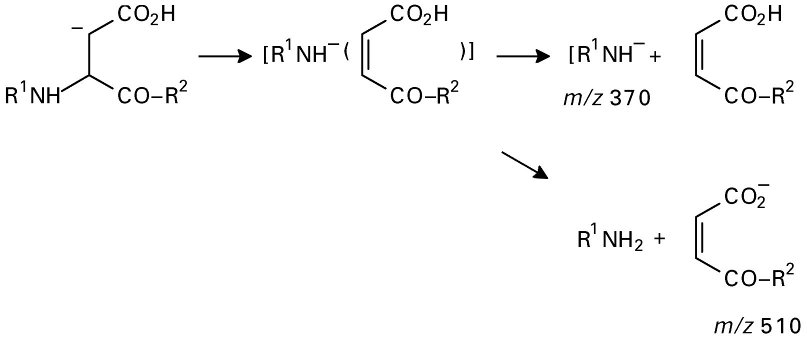



**Figure 3** Collision-induced negative-ion fast-atom bombardment tandem mass spectrum (MS/MS) of GLLEGLLGTLGL($NH_2$). ZAB 2HF instrument. Glycerol was used as matrix. Other details as for **Figure 1**. For the mechanisms of the backbone cleavages, see **Scheme 3**.




**Scheme 3**




**Scheme 4**

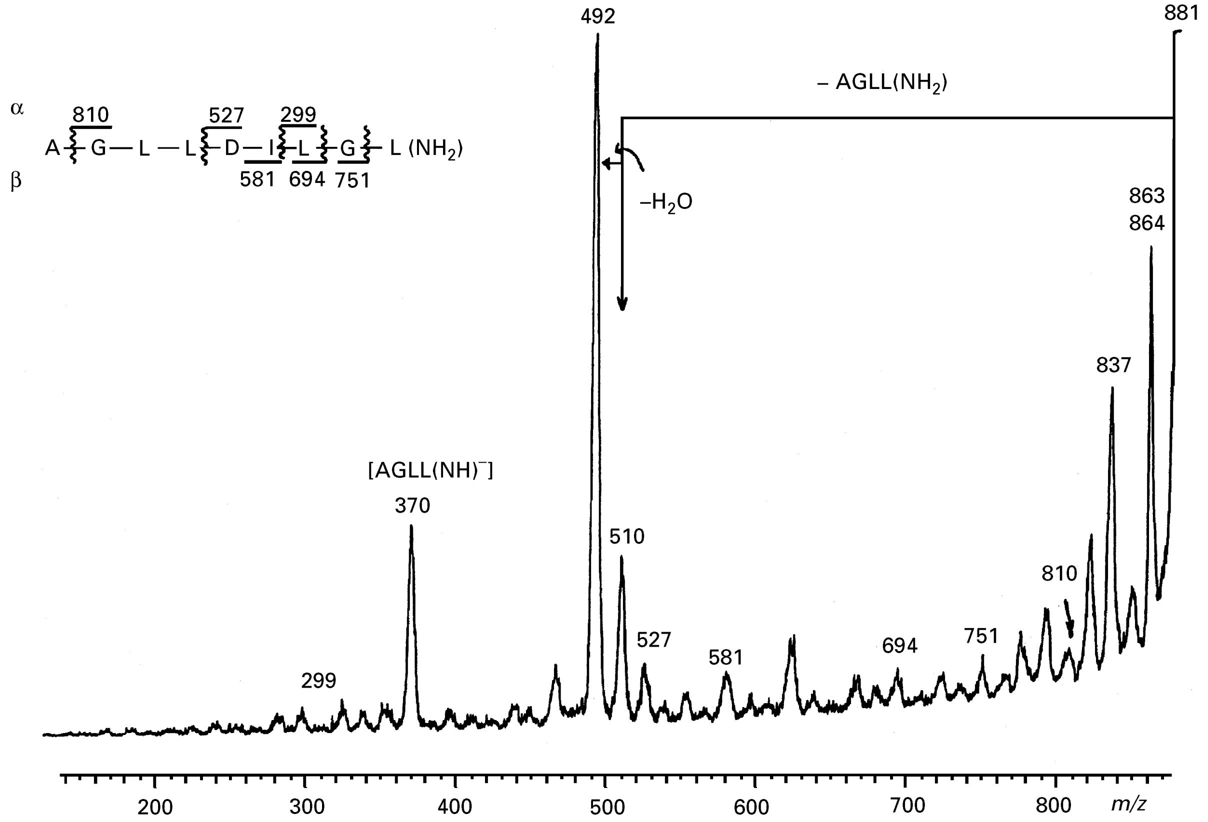



**Figure 4** Collision-induced negative-ion fast-atom bombardment tandem mass spectrum (MS/MS) of AGLLDILGL($NH_2$). ZAB 2HF mass spectrometer. Glycerol was used as matrix. Other details as for **Figure 1**. For the mechanisms of the backbone cleavages, see **Schemes 3** and **4**.

## Mechanistic studies of negative ions

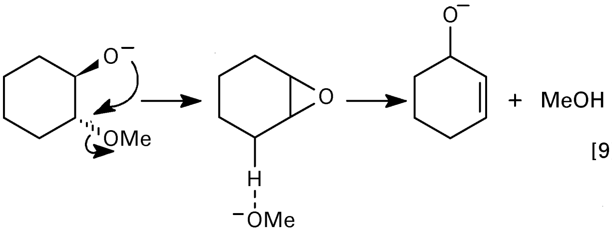
### The study of intramolecular rearrangements of negative ions

There are many rearrangement reactions of anions in the condensed phase in which the products and rates of reaction are dependent on the solvent and counterion used. The gas phase is thus the best medium for the investigation of the fundamental reactivity of the rearranging anion and the mechanism of the rearrangement. Many such rearrangements have been studied including well-known 1,2 anionic rearrangements such as the Wittig, Wolff, acyloin, negative-ion pinacol and Favorskii reactions.

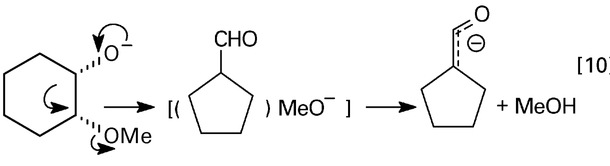
The pinacol–pinacolone rearrangement (Eqn [8]) is arguably the most famous of all acid-catalysed rearrangements and involves a simple Whitmore 1,2 methyl shift. Base-catalysed analogues of the pinacol rearrangement are not common, but the rearrangement does occur for deprotonated β-chlorohydrins. For example, base-catalysed rearrangement of *cis*-2-chlorocyclohexanol yields formylcyclopentane. The negative-ion pinacol rearrangement has been studied in the gas phase using $MeO^-$ as the leaving group. This is illustrated for the reactions shown in Equations [9] and [10]. The *trans* isomer (Eqn [9]) does not undergo the pinacol rearrangement, instead, an $S_Ni$ cyclization gives an epoxide, which ring-opens as shown to yield deprotonated cyclohex-2-en-1-ol and methanol. In contrast, the *cis* isomer (see Eqn [10]) undergoes the pinacol rearrangement as shown. The product anions of Equations [9] and [10] are identified from a comparison of their negative-ion mass spectra with those of independently synthesized anions.

The base-catalysed rearrangements of $\alpha$-halo ketones are classical examples of the reactions of ambident enolate anions in solution. The extent of each of the two reactions shown in Equations [11] and [12] is principally a function of the type of solvent used. A protic solvent solvates more strongly at the oxygen centre of the ambident anion and thus reaction proceeds through the carbanion centre to yield the Favorskii species as the major product (Eqn [11]). In marked contrast, the Favorskii rearrangement does not occur in the gas phase. Here,

$$Me_2C(OH)C(^+OH_2)Me_2 \longrightarrow Me_2C(OH)^+CMe_2 + H_2O$$
$$\longrightarrow Me^+C(OH)CMe_3$$
$$\longrightarrow MeCOCMe_3 + H^+ \qquad [8]$$




[9]




[10]

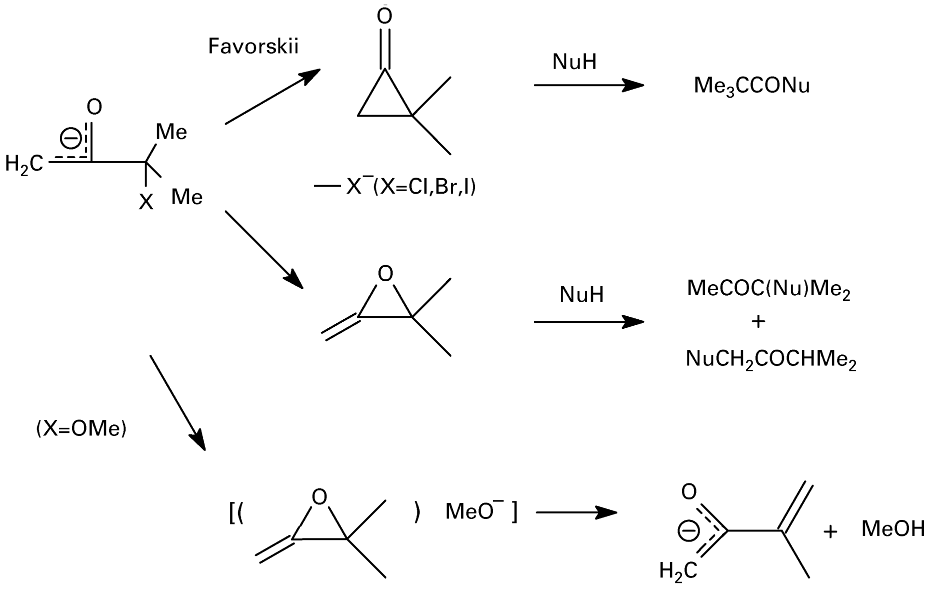
nucleophilic attack occurs exclusively via the more electron-rich alkoxide centre to form an allene oxide adduct that fragments as shown in Equation [13].

### Charge reversal and neutralization of negative ions

When a negative ion with high translational energy undergoes a soft collision with an inert gas atom (such as He or Ar) in a collision cell, some translational energy may be converted into internal energy. The energized ion rids itself of this internal energy in any one of a number of different ways; for example it may (i) radiate, (ii) cleave to give fragment negative ions (see, e.g. **Figures 1–4**), (iii) eject one electron to yield a neutral, or (iv) eject two electrons (either sequentially or synchronously) to yield a positive ion. These processes may occur for either parent or daughter anions, irrespective of whether they are radical anions or even-electron anions.

Charge reversal is that process by which an anion loses two electrons to form the corresponding cation. The process is best explained by reference to the schematic diagram shown in **Figure 5**. The particular polyatomic anion under study is selected using the analyser system of a sector mass spectrometer; it then proceeds into the first collision cell, which contains sufficient gas to effect single-collision conditions. The parent cations so produced have a range of internal energies and generally produce a mass spectrum containing peaks corresponding to the parent (cation) (the recovery signal: sometimes this is absent) together with a variety of fragment cations. The first application of charge reversal is that it is possible to make positive ions that cannot be formed by conventional ionization procedures. For example, *m*/*z* 31 from methanol in the positive mode is $CH_2{=}^+OH$: $CH_3O^+$ is not normally formed. However, $CH_3O^+$ can be formed by charge reversal of $CH_3O^-$. Similarly, $RCO_2^+$ is not a species formed readily by decomposition of any positive ion, but it can be formed by charge reversal of $RCO_2^-$.

The second application is analytical. If an $(M–H)^-$ ion undergoes little or no fragmentation following excitation, another option is to charge-reverse the $(M–H)^-$ to give a spectrum of positive ions that




[11]

[12]

[13]

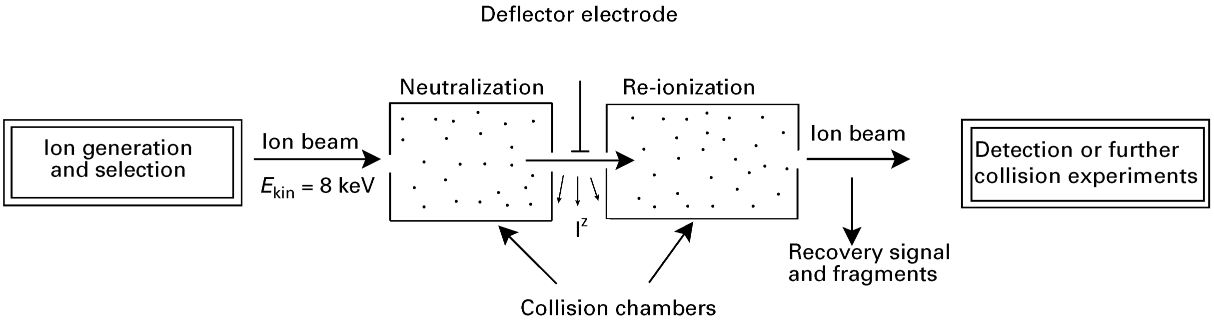



**Figure 5** Schematic representation of the two collision cells in a reversed sector mass spectrometer. Reproduced with permission of The American Chemical Society from Goldberg N and Schwarz H (1994) *Accounts of Chemical Research* **27**: 347–352.

should provide structural information. Alternatively, the structure of some fragment anion may need to be determined. Comparison of the charge-reversal spectrum of this anion with the charge-reversal spectra of anions of known structure can often identify the unknown. Suppose the major peak in a negative-ion spectrum corresponds to $C_2H_3O_2$ (*m/z* 59). The unknown is either $MeCO_2^-$ or $^-OCH_2CHO$. The charge-reversal spectra of these anions are reproduced in **Figure 6** each is characteristic and readily identifies the anion precursor.

Co-occurring with the charge-reversal process is that which produces an (even-electron) neutral by ejection of an electron from a radical anion, or a radical by loss of an electron from an even-electron anion. If a potential is applied after the first collision cell (see **Figure 5**) to deflect all ions, the only species proceeding into the second collision cell are neutrals formed in the first collision cell. If the second cell contains a gas (normally oxygen) that can ionize the neutrals, a composite positive-ion spectrum of all neutrals formed in the first collision cell can be obtained. This is called neutralization–reionization mass spectrometry (NRMS). The technique has a number of applications. First, the neutral formed in a negative-ion fragmentation may be identified. More importantly, interesting neutrals can be synthesized in the mass spectrometer, and their structures probed both experimentally (from their mass spectra) and theoretically (using computational methods). Fred McLafferty, a pioneer in the development and application of NRMS has stated that 'NRMS has made its largest contribution to research areas outside mass spectrometry in the characterization of unstable and radical species'. Several examples of the application of this technique are outlined below.

The radical HSO• is believed to be implicated in ozone depletion in the upper atmosphere by the reactions shown in Equations [14] and [15]. This species and its isomer HOS• have been made in the mass spectrometer by electron loss from the precursor anions $HSO^-$ and $HOS^-$. The $HOS^-$ precursor anion may be prepared by the collision-induced process summarized in Equation [16], while $HSO^-$ is

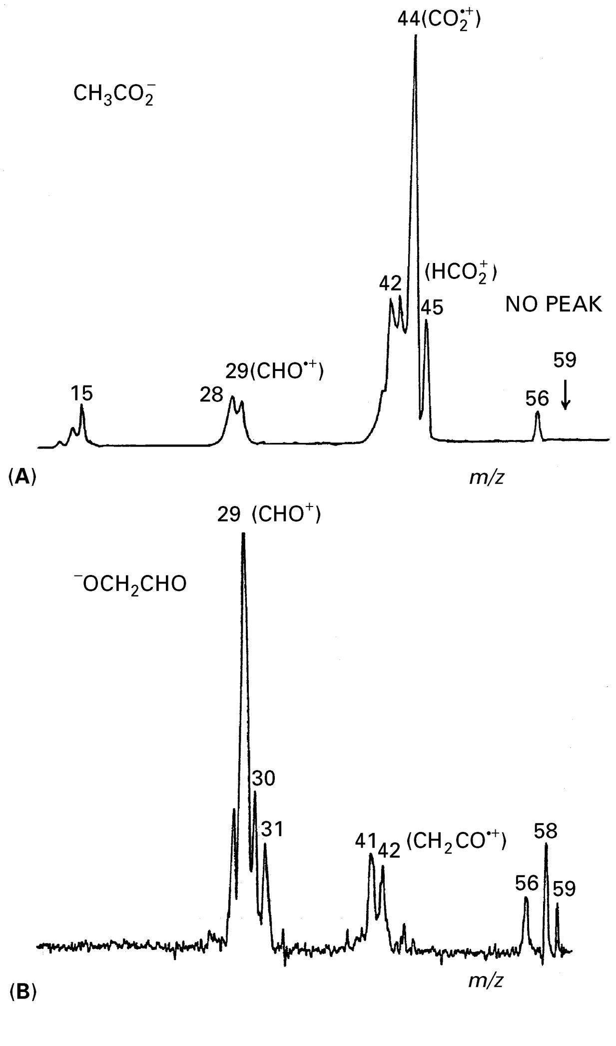



**Figure 6** Charge-reversal (positive-ion) tandem mass spectra of (A) $CH_3CO_2^-$ and (B) $^-OCH_2CHO$. ZAB 2HF mass spectrometer. Argon collision gas: pressure adjusted in the first collision cell so that the reduction in the main beam is 10%.

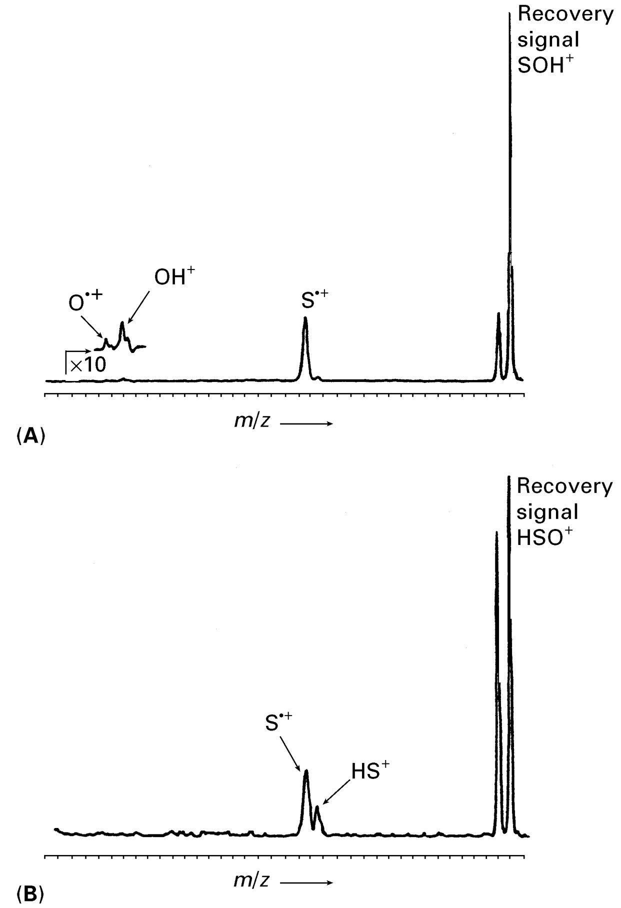



**Figure 7** Neutralization–reionization mass spectra of (A) HOS⁻ and (B) HSO⁻. Modified ZAB four-sector mass spectrometer of BEBE configuration (B = magnet sector; E = electric sector). Oxygen used as collision gas used in each collision cell (see **Figure 5**). The pressure of the gas in each cell was adjusted so the reduction in the main beam is 20% for each collision event. Reproduced with permission of ACS from Goldberg N and Schwarz H (1994) *Accounts of Chemical Research* **27**: 347–352.

prepared by admixture of $H_2S$ and $N_2O$ under conditions of negative-ion formation, perhaps by the reaction shown in Equation [17]. The reionization positive-ion spectra of the two non-interconverting radicals are shown in **Figure 7**.

$$HS^{\bullet} + O_3 \longrightarrow HSO^{\bullet} + O_2 \quad [14]$$

$$HSO^{\bullet} + O_3 \longrightarrow HS^{\bullet} + 2O_2 \quad [15]$$

$$[CH_3SOCH_2]^- \longrightarrow HOS^- + C_2H_4 \quad [16]$$

$$HS^- + N_2O \longrightarrow HSO^- + N_2 \quad [17]$$

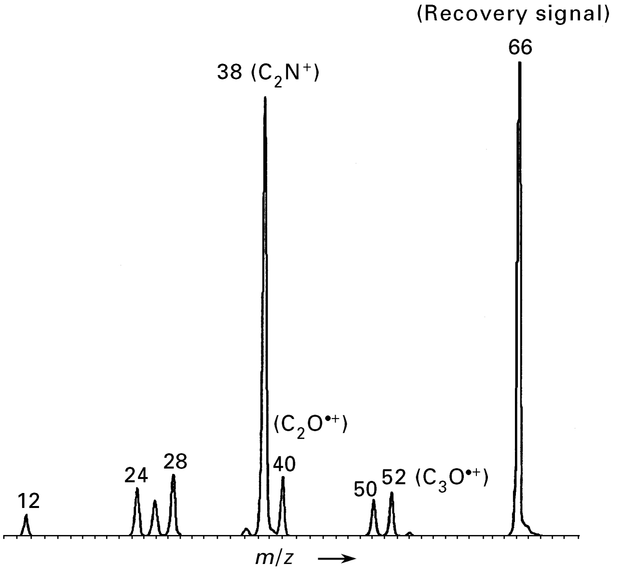



**Figure 8** Neutralization–reionization mass spectrum of N≡C–C≡C–O⁻. Details as in the legend to **Figure 7**. Reproduced with permission of Elsevier from Muedas CA, Sülzle D and Schwarz H (1992) *International Journal of Mass Spectrometry and Ion Processes* **113**: R17–R22.

There is much current interest in the formation of small cumulene and heterocumulene anions and their neutrals. Two examples are mentioned. The anion N≡CC≡C–O⁻ can be prepared by collisional activation of the precursor anion shown in Equation [18]. The corresponding radical is made by charge stripping. The positive-ion reionization spectrum of $NC_3O^-$ is shown in **Figure 8**.

$$NC{-}CH^-{-}CO_2Me \longrightarrow [(NC{-}CH{=}C{=}O)\ ^-OMe] \longrightarrow N{\equiv}C{-}C{\equiv}C{-}O^- + MeOH \quad [18]$$

Neutral $CH_2C_2$: is formed following electron loss from radical anion $CH_2C_2^{\bullet-}$. The anion is made by the reaction between allene and the monooxygen radical anion as shown in Equation [19].

$$O^{\bullet-} + CH_2{=}C{=}CH_2 \longrightarrow CH_2{=}C{=}\dot{C}{:}^- + H_2O \quad [19]$$

## Conclusions

Negative-ion mass spectrometry is a useful analytical technique that is complementary to more ubiquitious positive-ion method. There are particular classes of compounds where the spectra of $(M–H)^-$ ions provide more structural information than the corresponding positive-ion spectra.

Studies of intramolecular reactions of negative ions can provide useful mechanistic information.

Electron stripping of negative ions can be used to produce both neutrals and positive ions not available via other synthetic pathways.

*See also:* **Chemical Ionization in Mass Spectrometry; Fragmentation in Mass Spectrometry.**

## Further reading

Born M, Ingemann S and Nibbering NMM (1997) Formation and chemistry of radical anions in the gas phase. *Mass Spectrometry Reviews* **16**: 181–201.

Bowie JH (1984) The formation and fragmentation of negative ions derived from organic molecules (mainly radical anions). *Mass Spectrometry Reviews* **3**: 161–207.

Bowie JH (1990) The fragmentations of even-electron organic negative ions. *Mass Spectrometry Reviews* **9**: 349–379.

Bowie JH (1994) The fragmentations of $(M–H)^-$ ions derived from organic compounds. An aid to structure determination. In: Russell DH (ed) *Experimental Mass Spectrometry*, pp 1–39. New York: Plenum Press.

Eichinger PCH, Dua S and Bowie JH (1994) Review. A comparison of skeletal rearrangement reactions of even-electron anions in solution and in the gas phase. *International Journal of Mass Spectrometry and Ion Processes* **133**: 1–12.

Goldberg N and Schwarz H (1994) Neutralisation–reionisation mass spectrometry: a powerful "laboratory" to generate and probe elusive molecules. *Accounts of Chemical Research* **27**: 347–352.

# Neutralization–Reionization in Mass Spectrometry

**Chrys Wesdemiotis**, University of Akron, OH, USA

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Neutralization–reionization mass spectrometry (NRMS) involves the synthesis and characterization of neutral species inside a tandem mass spectrometer *via* gas-phase redox reactions. The neutrals are produced by neutralization of the corresponding mass-selected cations or anions (precursor ions) and are subsequently characterized by the mass spectra arising after reionization to positive or negative ions. The successive neutralization and reionization events are effected by collisions with gaseous targets at high kinetic energy. Equations [1] and [2] illustrate a neutralization–reionization (NR) sequence that starts with a positive precursor ion and reionizes the intermediate neutral to cations. These choices generally proceed with higher yields than alternative charge permutations and, therefore, have been employed in most NRMS studies reported so far; nonetheless, any charge combination of precursor ion and final product ions is possible and has been documented.

NRMS is uniquely suitable for the study of highly reactive neutrals that cannot be isolated and characterized in condensed phases, where interactions with neighbouring molecules would cause their immediate decomposition. By preparing the neutrals in the gas phase of the mass spectrometer, such destructive intermolecular collisions are avoided. Furthermore, the ionized forms of reactive neutrals are often stable, well known ions that can easily be generated in the mass spectrometer to serve as the starting material for the neutrals' synthesis. For example, carbene ion $H–C–NH_2^{\bullet+}$, which is available *via* electron ionization (EI) of cyclopropyl amine, can be used to synthesize the highly reactive prototype carbene H–C–$NH_2$ (a putative interstellar species); similarly, the distonic ion $^{\bullet}CH_2CH_2OCH_2^+$, which is formed by EI of 1,4-dioxane, can serve as the precursor for the hypovalent 1,4-diradical $^{\bullet}CH_2CH_2OCH_2^{\bullet}$ i.e. ring-opened oxetane (pyrolysis intermediate). NRMS has so far enabled the first experimental characterization of a large variety of neutral intermediates, in particular species with weak bonds, extra or missing valences, unpaired electrons, or a high tendency toward intermolecular isomerization (see below).

Knowledge of the intrinsic properties of such unusual neutrals is desirable for many reasons. They are postulated key intermediates in organic and biological reactions; they appear in the atmosphere, in interstellar clouds, and in other cosmic environments; they are involved in plasma and combustion chemistry and in photochemical processes and they play an important role in industrial and biological catalysis.

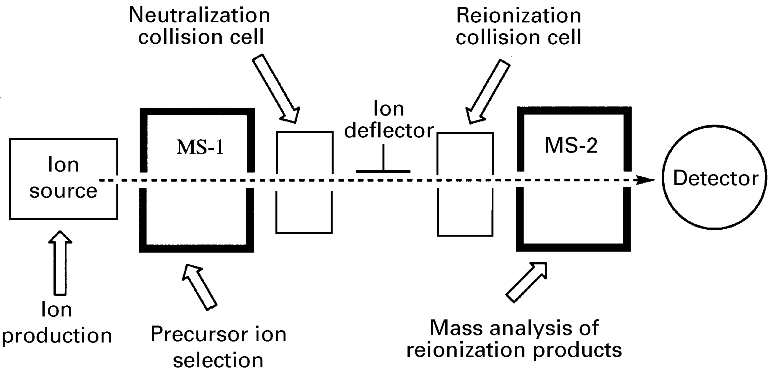



**Figure 1** Basic components of a mass spectrometer for NRMS.

## NRMS instrumentation

NRMS experiments are conducted with beam tandem mass spectrometers consisting of two mass-analyzing devices (MS-1 and MS-2) and equipped with at least two collision cells and an intermediate ion deflector in the field-free region between MS-1 and MS-2 (**Figure 1**). The desired precursor ion, *i.e.* the cationic or anionic form of the reactive neutral to be synthesized, is generated by the appropriate ionization method in the ion source, accelerated to a keV kinetic energy (3–10 keV), and separated from all other ions simultaneously produced in the ion source by MS-1. The mass-selected precursor ion beam is then partly neutralized by charge exchanging collisions with the neutralization target (Eqn [1]) in the collision cell following MS-1 (neutralization cell). The mixture of neutrals and ions exiting the cell passes the ion deflector, where the ions are removed from the beam path by electrostatic deflection. The remaining neutral beam subsequently enters the next collision cell (reionization cell) where it is ionized by collisions with the reionization target (Eqn [2]). The newly formed product ions are then dispersed according to their mass-to-charge ratio ($m/z$) by MS-2 to yield the NR spectrum of the precursor ion under study.

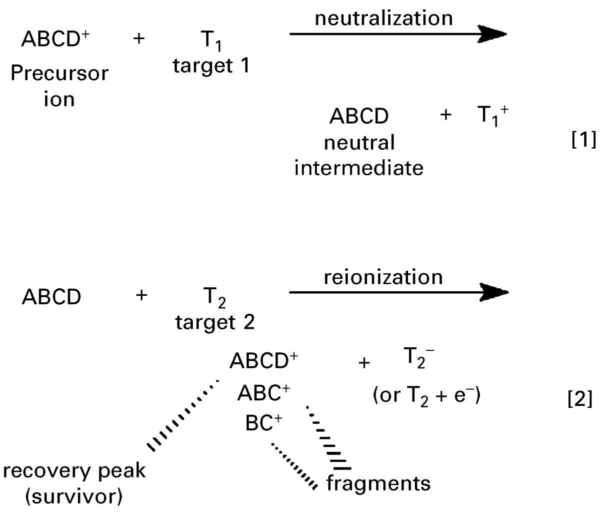
MS-1 and MS-2 may be simple, one-stage mass analyzers or tandem mass spectrometers. In the latter case, it is possible to form precursor ions by MS/MS of *specific* source-generated ions, and/or to analyze by MS/MS a *specific* product ion arising in the reionization event. Generally, the lifetimes of the ions subjected to NRMS experiments have been in the order of tens of microseconds. Depending on the distance between neutralization and reionization cells, the lifetimes of the neutrals generated in NRMS experiments range between a few tenths of a microsecond to a few microseconds. The field-free region between MS-1 and MS-2 may contain a third collision cell and a second intermediate ion deflector. Such a configuration allows one (a) to vary the lifetime of the neutral intermediates (by changing the neutralization and/or reionization cell) and (b) to probe the collisionally activated dissociation (CAD) of the neutrals *before* they are reionized.

## The neutralization and reionization events

Positive or negative precursor ions can be subjected to neutralization and the resulting neutral species may be reionized to cations or anions. Customarily, the charges of precursor and final ions are indicated by superscripts to the NR acronym. The most frequently performed experiment is $^+NR^+$; it involves neutralization of a cation followed by reionization of the neutral intermediate(s) to cations, *i.e.* a gas phase reduction–oxidation succession.

Cations with kinetic energies in the keV range can be neutralized with metal vapours (e.g. Na or Hg), atomic or molecular gases (e.g. Xe or $(CH_3)_3N$), and organic compound vapours (e.g. $CH_3SSCH_3$). Electron exchange takes place during the collision encounter of the fast moving precursor ion with the essentially stationary target atom/molecule. This encounter lasts only femtoseconds, a time much shorter than typical rovibrational periods. As a result, such neutralization is *vertical*, generating the incipient neutral in the geometry of its ionic precursor (Franck–Condon process, see **Figure 2**).

The neutralization yield and internal energy of the neutral species generated in the neutralization step (Eqn [1]) depend on several variables, including the internal energy of the precursor ion, the equilibrium geometries of precursor ion and neutral, and the target used for neutralization. One important factor determining the internal energy of the incipient neutral species emerging upon vertical neutralization is the thermochemistry of the charge exchange process.




The enthalpy change of this reaction, $\Delta H_N$, is given in Equation [3] where $IE_v(T_1)$ is the vertical ionization energy of the gaseous target oxidized and $RE_v(ABCD^+)$ the vertical recombination energy of the precursor ion reduced.

$$\Delta H_N = IE_v(T_1) - RE_v(ABCD^+) \qquad [3]$$

Exothermic neutralization ($\Delta H_N < 0$) usually leads to an excited species which may fragment if the energy deposited ($\Delta H_N$) exceeds the dissociation threshold. On the other hand, thermoneutral ($\Delta H_N = 0$) or endothermic ($\Delta H_N > 0$) neutralizations normally yield ground-state neutrals. With endothermic encounters, the energy deficit is supplied by the kinetic energy of the precursor ion. In general, thermoneutral or nearly thermoneutral reactions exhibit the highest yields.

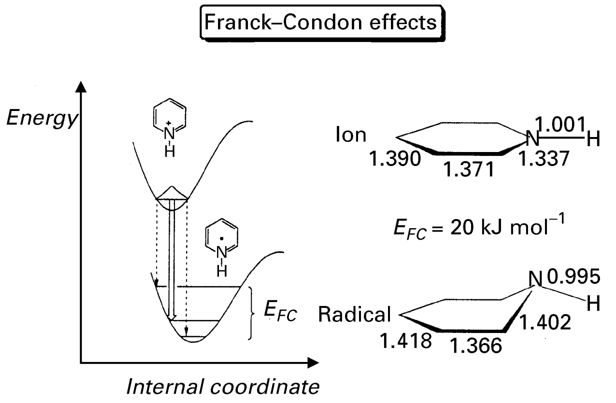



**Figure 2** Franck–Condon energies in vertical neutralization of pyridinium ion. Reprinted with permission of John Wiley and Sons from Turecek F (1998) Modelling nucleobase radicals in the mass spectrometer. *Journal of Mass Spectrometry* **33**: 779–795.

It is believed that neutralization involves glancing collisions with a relatively large projectile–target distance ('impact parameter'). An endothermic neutralizing collision may provide some extra internal energy to the newly formed neutral; the degree of activation in the nascent neutral increases with $\Delta H_N$, presumably because of the decreased impact parameter needed to convert some kinetic into internal energy, so that charge exchange becomes thermochemically possible. Additional excitation may be imparted by Franck–Condon effects; thus, if the equilibrium geometries of $ABCD^+$ and ABCD differ substantially, ABCD can emerge with rovibrational excitation, even if it is formed by a thermoneutral or endothermic process. For example, endothermic neutralization of pyridinium cations produces vibrationally excited pyridinium radicals due to the different ground-state structures of cation and radical (**Figure 2**). The average internal energy imparted to the radical by Franck–Condon effects ($E_{FC}$) is ~ 20 kJ mol$^{-1}$ in this case. This modest amount is not sufficient to cause N–H bond cleavage in the radical ($D$ = 108 kJ mol$^{-1}$), but increases the dissociation extent after reionization. In some systems, Franck–Condon effects can be very large; this is true for the neutralization of ground-state $(CH_3)_2Si^+OH$, which produces the radical $(CH_3)_2Si^{+\bullet}OH$ with enough internal energy for O–H or C–Si bond cleavage; therefore, such radicals do not survive intact until reionization, although $(CH_3)_2Si^{\bullet}OH$ is a bound species in its ground state. When molecular targets are used for neutralization, Franck–Condon effects play a more significant role than the thermochemistry of the neutralization reaction in determining the energy state of ABCD; further, in exothermic neutralizations, the excess $\Delta H_N$ can end up as excitation of the target ion ($T_1^+$ in Eqn [1]) and not in the incipient neutral ABCD. The energy state of ABCD also depends on the internal energy of the precursor ion $ABCD^+$; an activated precursor ion can give rise to an activated or vibrationally cool neutral, depending on the Franck–Condon overlap of Equation [1]. The former case is true for $NH_4^+$ neutralization, while the latter behaviour has been observed for neutralized $(CH_3)_2OH^+$, $C_2H_5OH_2^+$, and $(CH_3)_2Si^+OH$. Finally, it must be kept in mind that the neutralizing collision (Eqn [1]) may also cause some CAD of the precursor ion. The extent of concomitant CAD is minimized by avoiding strongly endothermic neutralizations whose small impact parameters may excite and dissociate the precursor ion (see above).

In the second $^+NR^+$ step (Eqn [2]), the neutral intermediate (ABCD) is collisionally reionized. This process, which resembles ionization by electron

impact, produces a molecular ion as well as fragment ions (see Eqn [2]). Customarily, the molecular ion ($ABCD^+$) has been called the 'survivor ion', as it originates from ABCD that survived intact. The survivor ion gives rise to the 'recovery peak' in the NR spectrum, *i.e.* the peak appearing at the same *m*/*z* value as the precursor ion. Regarding the reionization target ($T_2$ in Eqn [2]), $O_2$ and He are the most widely employed choices. $O_2$ yields more abundant recovery peaks and He more fragments, justifying their classification as 'softer' and 'harder' reionization gas, respectively. Other, less frequently used soft reionization gases are NO and $NO_2$. Reionization efficiencies depend significantly on the structure of the neutral and rise with the neutral's kinetic energy but decline with its internal energy. Furthermore, kinetic and internal energies of the neutral intermediate also affect how much internal energy is deposited on collisional ionization; here, the average internal energy gained increases with both.

With anionic precursors, the neutralization step entails electron removal, *i.e.* an oxidation reaction, as does reionization of a neutral to cations. Therefore, $O_2$ which is the best reionization target (see above) also is the target of choice for the neutralization of anions. Conversely, cation neutralization targets, which effect a gas-phase reduction, are most suitable for reionization to anions. Widely used targets for this purpose have been xenon and trimethylamine. Reionization to anions suffers from substantially poorer yields than reionization to cations, often by ca. 10 times or more; therefore, it has been employed to a much lesser extent. However, in certain cases it can provide superior structural information; for example, the unequivocal identification of oxygen-centered radicals, such as the carbonate radical $CO_3^{\bullet}$ and the diradical $^{\bullet}CH_2CH_2O^{\bullet}$ relied on their reionization to anions.

## The study of reactive neutrals *via* NRMS

The stability and unimolecular reactivity of the neutral generated in the neutralization step are characterized by the mass spectrum arising after reionization. The presence of a recovery peak in this spectrum provides evidence that the neutral intermediate has survived intact (*i.e.* undissociated) for microsecond(s). Whether it has retained the connectivity of the precursor ion is judged by comparison of the NR spectrum to the CAD spectrum of the precursor ion or the NR spectra of other, usually stable and known isomers. Both these strategies are presented below with representative examples.

The hypermetallic dilithium fluoride ($Li_2F$), a compound violating the octet rule, can be formed in the gas phase by neutralizing the known cation $[Li–F–Li]^+$, which is produced abundantly upon fast atom bombardment ionization of lithium trifluoroacetate. Neutralization with trimethylamine (TMA) and subsequent reionization with $O_2$ lead to the NR spectrum of **Figure 3A**. The CAD spectrum of $[Li–F–Li]^+$ (using $O_2$) is displayed in **Figure 3B**. The significant recovery peak in the NR spectrum convincingly shows that $Li_2F$ has survived intact. It is, thus, a stable species and bound in respect to Li + LiF. The experimentally documented stability is supported by theory which predicts that the reaction $Li_2F \rightarrow Li + LiF$ must overcome a barrier of 130–140 kJ mol$^{-1}$. The fragmentation patterns in NR and CAD spectra of **Figure 3** are similar. In both spectra, the most intense fragments are $LiF^{\bullet+}$(*m*/*z* 26) and $F^+$ (*m*/*z* 19), and the peak widths at half-height ($w_{0.5}$) follow the order $w_{0.5}$ ($LiF^{\bullet+}$, $Li^+$) < $w_{0.5}$ ($Li_2^{\bullet+}$) < $w_{0.5}$($F^+$). These common trends strongly suggest that the majority of NR fragments arise from decomposing survivor ions, not from dissociation of the inter-

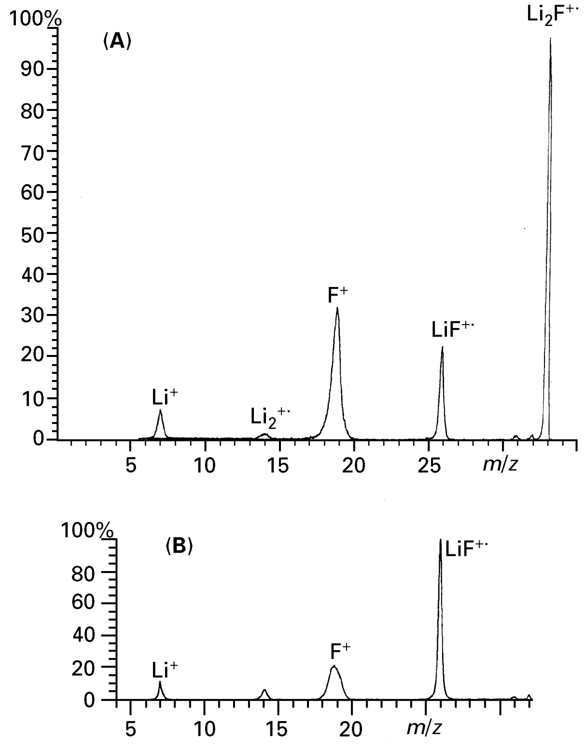
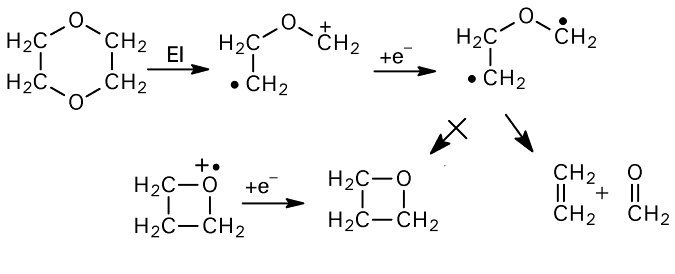



**Figure 3** (A) Neutralization–reionization ($^+NR^+$) and (B) collisionally activated dissociation (CAD) spectra of $Li_2F^+$. Reprinted with permission of Elsevier from Polce MJ and Wesdemiotis C (1999) Hypermetallic dilithium fluoride, $Li_2F$, and its cation and anion: a combined dissociation and charge permutation study. *International Journal of Mass Spectrometry* **182/183**: 45–52.




**Scheme 1**

mediate $Li_2F$, in keeping with the considerable thermodynamic stability of this hypermetallic compound.

A reactive neutral that dissociates extensively in the time span between neutralization and reionization is diradical $^\bullet CH_2CH_2OCH_2^\bullet$ (C–C ring-opened oxetane). This species is available through neutralization of the distonic ion $^\bullet CH_2CH_2OCH_2^+$ which is generated by electron ionization of 1,4-dioxane (**Scheme 1**). When Xe and $O_2$ are used for neutralization and reionization, respectively, the NR spectrum of **Figure 4A** is obtained; it is dominated by peaks at *m/z* ≤ 30, verifying that the incipient diradical emerging in the neutralization step decomposes to ethylene (28 Da) and formaldehyde (30 Da). A small fraction remains, however, undissociated, as indicated by the presence in **Figure 4A** of a survivor ion (*m/z* 58) and fragment ions of *m/z* > 30. The surviving diradicals could cyclize to the thermodynamically more stable oxetane molecule (cf. **Scheme 1**). Whether this happened is best determined by comparison of the NR spectra of $^\bullet CH_2CH_2OCH_2^+$ (**Figure 4A**) and ionized oxetane (**Figure 4B**). Consistent with the high stability of neutral oxetane, the corresponding NR spectrum (**Figure 4B**) contains a substantial recovery peak (*m/z* 58); further, the $C_3H_{0-3}^+$ fragments in this spectrum (*m/z* 36–39) are diagnostic for oxetane's cyclic structure which has three adjacent C atoms. These latter fragments are minuscule in the NR spectrum of $^\bullet CH_2CH_2OCH_2^+$ (**Figure 4A**), which shows that the diradical $^\bullet CH_2CH_2OCH_2^\bullet$ does not cyclize to oxetane within the time available between neutralization and reionization. Hence, the diradical is kinetically stable in respect to both isomerization and dissociation and exists as a bound, high-energy intermediate.

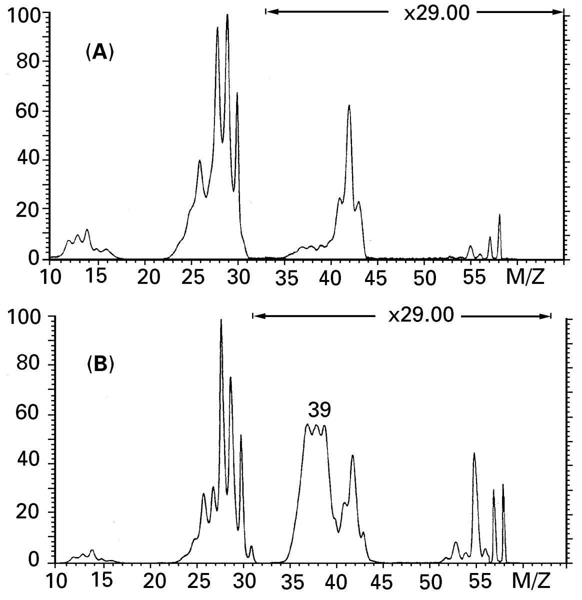



**Figure 4** Neutralization–reionization ($^+NR^+$) spectra of (A) $^\bullet CH_2CH_2OCH_2^+$ and (B) oxetane cation. Reprinted with permission of the American Chemical Society from Polce MJ and Wesdemiotis C (1993) The unimolecular chemistry of the 1,4-diradical $^\bullet CH_2CH_2OCH_2^\bullet$ in the gas phase. Comparison to the distonic radical ions $^\bullet CH_2CH_2OCH_2^+$ and $^\bullet CH_2CH_2OCH_2^-$ *Journal of the American Chemical Society* **115**: 10849–10856.

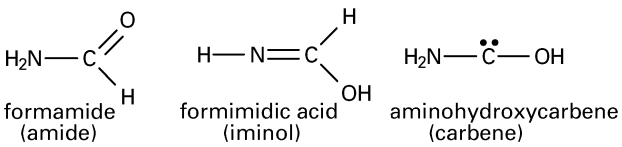
## Auxiliary NRMS methods

When the precursor ion is not isomerically pure or the intermediate neutrals rearrange partly or dissociate to products that are isobaric with those arising after reionization of the surviving neutrals, a straightforward characterization from the NR spectrum alone may be impossible. This problem has led to the development of an array of auxiliary NRMS methods that can be used to explore the stability and unimolecular reactivity of transient neutrals with complex and/or ambiguous NR spectra. A brief overview of these experimental approaches follows, along with examples demonstrating their usefulness.

### NRMS-CAD studies

The isomerization proclivity of a reactive intermediate can be interrogated by measuring the tandem mass spectrum of the survivor ion. This capability requires that the instrumentation available be equipped with an additional mass analyzing device after the one used for the mass separation of the NR products.

The NRMS-CAD method is illustrated with the $CH_3NO$ tautomers depicted in **Scheme 2**.




**Scheme 2**

Formimidic acid and aminohydroxycarbene, both tautomers of formamide, are reactive molecules that do not exist in condensed media. Using NRMS, they can be synthesized in the gas phase from the corresponding radical cations, which are formed *via* dissociative electron ionization of *N*-formylhydrazone (yields the iminol ion) and oxamide (yields the carbene ion). After reionization, all three tautomers give rise to abundant survivor ions, but the NR spectra are not distinctive enough to unequivocally determine whether the less stable iminol and carbene neutrals underwent isomerization to the most stable formamide molecule. This question can be answered by subjecting the three survivor ions to CAD and comparing the resulting spectra to the CAD spectra of authentic (i.e. source-generated) amide, iminol, and carbene ions (**Figure 5**). The ions are known not to interconvert.

The NR-CAD and CAD spectra of formamide and the carbene are essentially indistinguishable, providing strong evidence that the amide and carbene tautomers retain their structures upon NR and do not undergo any rearrangements. In sharp contrast, NR-CAD and CAD spectra of the iminol ion differ markedly from each other; the NR-CAD spectrum can be interpreted as a combination of the CAD spectra of genuine amide and iminol ions, in agreement with a partial tautomerization of the neutral iminol to the amide.

### CAD of the neutral intermediate (NCR studies)

Instruments with three successive collision cells and two intermediate ion deflectors allow for CAD of the neutral intermediate before reionization. With such an arrangement, the neutral is produced in the first cell, undergoes CAD in the second cell, and is reionized (along with its dissociation and/or isomerization products) in the third collision cell. The first deflector removes unneutralized precursor ions and the second one any ions formed during the neutral CAD step. This procedure gives rise to an NCR spectrum (neutralization – CAD – reionization) which, when compared to the corresponding NR spectrum, reveals insight about the favoured fragmentation and/or isomerization of the neutral species under study.

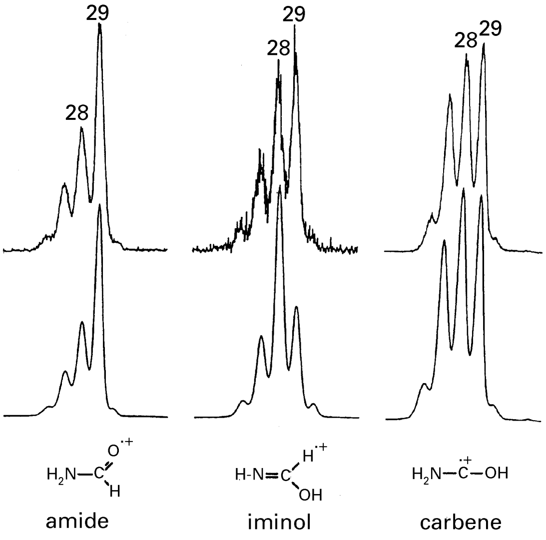



**Figure 5** Partial CAD spectra (using He) of $^+NR^+$ survivor ions (top) and CAD spectra of source-generated ions (bottom) of $H_2N$-$CH(=O^{\bullet+}$ (left), $HN=C(H)OH^{\bullet+}$ (centre), and $H_2N–C–OH^{\bullet+}$ (right). Reprinted with permission of Elsevier from McGibbon GA, Burgers PC and Terlouw JK (1994) The imidic acids H–N=C(H)–OH and $CH_3$–N=C(H)–OH and their tautomeric carbenes $H_2N$–C–OH and $CH_3$–N(H)–C–OH: stable species in the gas phase formed by one-electron reduction of their cations. *International Journal of Mass Spectrometry and Ion Processes* **136**: 191–208.

As an example, **Figure 6** shows the NR and NCR spectra of the distonic ion $^+CH_2CH_2O^\bullet$ whose neutralization creates the elusive diradical $^\bullet CH_2CH_2O^\bullet$ (ring-opened oxirane, cf. **Scheme 3**). The increased relative abundances of *m/z* 43, 29, 28, and 12–15 in the NCR spectrum support the occurrence of the top channel in **Scheme 3** (dissociation through acetaldehyde). Corroborative evidence that collisionally activated $^\bullet CH_2CH_2O^\bullet$ partially interconverts to $CH_3CH=O$, but does not ring-close to oxirane, is obtained by the NR-CAD spectrum of $^+CH_2CH_2O^\bullet$ (see previous section), which is consistent with a mixture of $^+CH_2CH_2O^\bullet$ and $CH_3CH=O^{\bullet+}$ and significantly different from the reference CAD spectrum of oxirane ion.

### Photodissociation/photoionization of the intermediate neutrals

Laser light can be used, in place of the reionization gas, to probe the reactivity and electronic state(s) of the neutrals produced in the neutralization step. Due to the low neutral currents generated in NRMS experiments (~ $10^6$ molecules per second), a coaxial arrangement of the neutral and laser beams and continuous laser light are necessary to ensure detectable photodissociation or photoionization yields. An

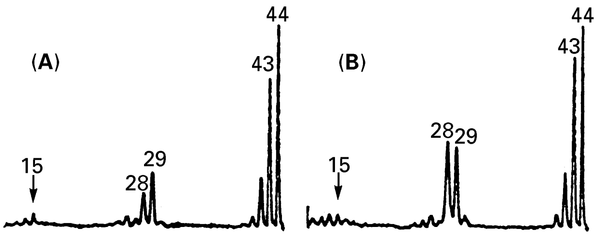



**Figure 6** (A) Neutralization-reionization ($^+$NR$^+$) and (B) neutralization–CAD–reionization ($^+$NCR$^+$) spectra of $=^+CH_2CH_2O^\bullet$. Hg was used for neutralization, $O_2$ for reionization, and He for CAD. Reprinted with permission of the American Chemical Society from Wesdemiotis C, Leyh B, Fura A and McLafferty FW (1990) The isomerization of oxirane. Stable $^\bullet CH_2OCH_2^{\bullet}$, $^\bullet CH_2CH_2O^\bullet$, and :$CHOCH_3$ and their counterpart ions. *Journal of the American Chemical Society* **112**: 8655–8660.

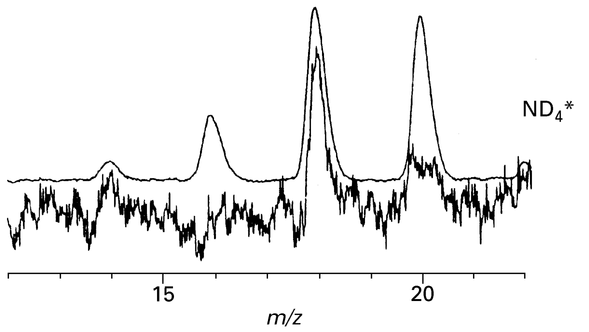


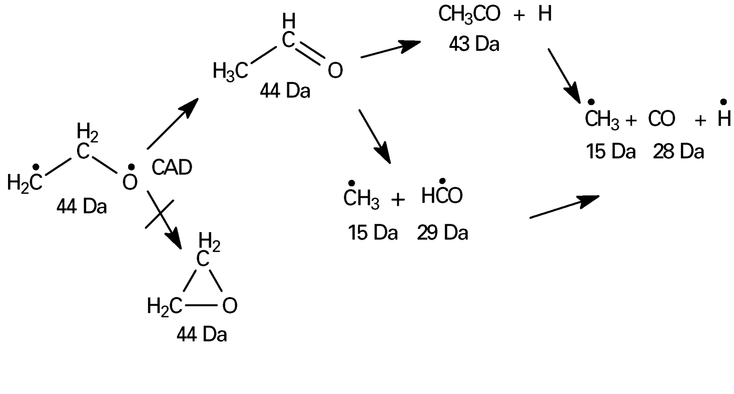

**Figure 7** Neutralization–photodissociation/photoreionization spectrum of $ND_4^\bullet$. Top: total reionization, bottom: after background subtraction. $ND_4^\bullet$ was formed by $ND_4^+$ neutralization with dimethyl disulfide. Reprinted with permission of Elsevier from Sadilek M and Turecek F (1996) Laser photolysis of $ND_4^\bullet$ and trimethylamine formed by collisional neutralization of their cations in the gas phase. *Chemical Physics Letters* **263**: 203–308.

argon laser, which emits intense lines at 488 nm (2.54 eV) and 514.5 nm (2.41 eV), is suitable for such experiments. The procedure is illustrated for the hypervalent radical $ND_4^\bullet$.

Collisional neutralization of $ND_4^+$ followed by laser irradiation of $ND_4^\bullet$ gives rise to the NR spectrum of **Figure** 7, top. A second NR spectrum is acquired with the laser off, so that the contribution of the background gases can be assessed. Subtraction of these sequential spectra gives the net change due to interaction with photons (**Figure** 7, bottom). The background-corrected spectrum includes a very weak positive contribution to the $ND_4^+$ survivor ion at *m/z* 22, indicating that the $ND_4^\bullet$ beam contains a minor fraction of long-lived excited electronic states with ionization energies ≤ 2.54 eV; for comparison, the IE of ground-state $ND_4^\bullet$ is 4.6 eV. The corrected NR spectrum also shows increased intensities of $ND_3^{\bullet+}$ (*m/z* 20) and $ND_2^+$(*m/z* 18), which are attributed to photodissociation of ground-state $ND_4^\bullet$ and subsequent collisional reionization with the background gases. The alternative scenario, namely collisional reionization to $ND_4^+$ followed by photodissociation of $ND_4^+$ to $ND_3^{\bullet+}$ and $ND_2^+$ is energetically impossible with the wavelengths used.

## Variable-time NRMS

A given NR product may arise from dissociation of the survivor ion or from dissociation of the neutral intermediate (see above). As explained above, NCR helps identify the predominant neutral fragmentations. An alternative approach for distinguishing overlapping neutral and ion dissociations is to systematically alter the time scales of these processes, using variable-time NRMS; here, the position of the reionization cell is varied, in order to change the *lifetimes* ('observation times') of the neutrals (ABCD) and reionized ions (ABCD$^+$ plus fragments). The cell need not be moved physically; instead, a segmented reionization cell can be used, in which the segments are individually floatable so that ions formed inside them are either prevented or permitted passage to MS-2, depending on the potentials applied. This




**Scheme 3**

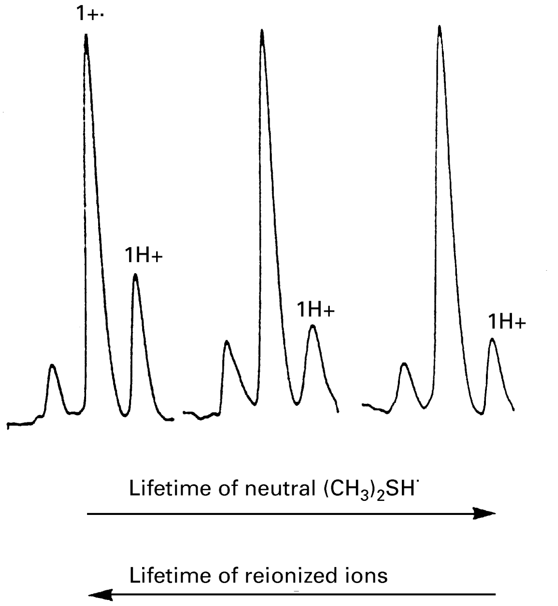



**Figure 8** Variable-time $^{+}NR^{+}$ spectra of $(CH_3)_2SH^+$ formed by chemical ionization of $(CH_3)_2S$ with $(CH_3)_2CH{=}OH^+$. Dimethyl disulfide and oxygen served as the neutralization and reionization targets, respectively. The observation times of the neutrals are 0.35 µs (left), 1.05 µs (centre), and 1.76 µs (right). The corresponding observation times of the ions are 3.44 µs (left), 2.73 µs (centre), and 2.02 µs (right). Reprinted with permission of Elsevier from Sadilek M and Turecek F (1999) Metastable states of dimethylsulfonium radical, $(CH_3)_2SH^\bullet$: a neutralization–reionization mass spectrometric and ab initio computational study. *International Journal of Mass Spectrometry* **185/186/187**: 639–649.

essentially corresponds to moving the entrance of the reionization cell; if the segments are floated such that the entrance moves towards (away from) MS-2, the observation time of the neutrals increases (decreases). Since the field-free region housing of the NR cells has a fixed length, increasing the observation time of the neutrals decreases that of the reionized ions (time elapsing between their formation and MS-2 entry) and *vice versa*.

**Figure 8** shows partial variable-time NR spectra of the dimethylsulfonium cation, $(CH_3)_2SH^+$, in which the lifetime of neutral $(CH_3)_2SH^\bullet$ (a hypervalent radical) gradually increases from 0.35 to 1.76 µs while that of the reionized ions decreases from 3.44 to 2.02 µs. The relative abundance of the survivor ion (designated $1H^+$) *decreases* upon *increasing* the observation time of the *neutrals* (equivalent with time available for *neutral dissociation*). This clearly points out that a fraction of radical $(CH_3)_2SH^\bullet$ dissociates in the microsecond time scale and that the peak labelled $1^{\bullet+}$ primarily originates through the neutral dissociation $(CH_3)_2SH^\bullet \rightarrow (CH_3)_2S + H^\bullet$ and subsequent reionization of $(CH_3)_2S$.

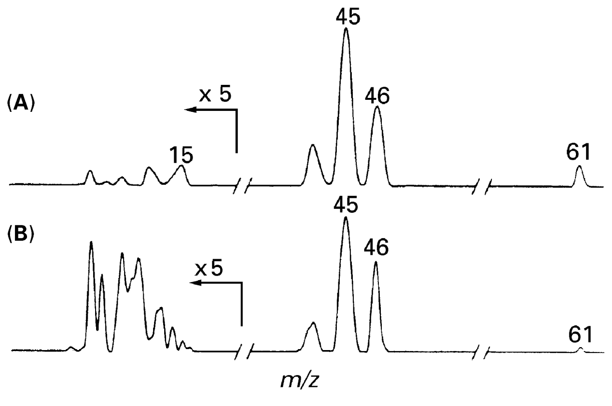



**Figure 9** Variable time $^{+}NR^{+}$ spectra of $CH_3SCH_2^+$ using $(CH_3)_3N$ for neutralization and $O_2$ for reionization. The observation times for $CH_3SCH_2^\bullet$ and reionized $CH_3SCH_2^+$ are (A) 2.44 and 1.70 µs and (B) 0.36 and 3.78 µs, respectively. Reprinted with permission of the American Chemical Society from Kuhns DW, Tran TB, Shaffer SA and Turecek F (1994) Methylthiomethyl radical. A variable-time neutralization-reionization and ab initio study. *Journal of Physical Chemistry* **98**: 4845–4853.

Another example is illustrated in **Figure 9**, which depicts variable-time NR spectra of the methylthiomethyl cation, $CH_3SCH_2^+$. In this case, the observation times for $CH_3SCH_2^\bullet$ and its reionization products are 2.44 and 1.70 µs, respectively, in the top and 0.36 and 3.78 µs, respectively, in the bottom spectrum. Now, the relative abundance of the survivor ion exhibits a 6-fold *decrease* upon *decreasing* the minimum neutral lifetime from 2.44 to 0.36 µs while increasing that of reionized $CH_3SCH_2^+$ from 1.70 to 3.78 µs. Thus, the depletion of the survivor ion in the lower spectrum must be due to an increased fraction of reionized $CH_3SCH_2^+$ that dissociates. This, in turn, strongly suggests that (a) the fraction of $CH_3SCH_2^\bullet$ dissociating within microseconds is negligible and (b) collisional reionization of $CH_3SCH_2^\bullet$ is accompanied by substantial energy deposition, thereby extending the unimolecular decomposition of the reionized ions into the microsecond time window.

## Neutral and ion decomposition difference (NIDD)

In this method, a neutral produced from an *anionic* precursor is reionized to *cations*, and one produced from a *cationic* precursor is reionized to *anions*, to thus obtain a $^{-}NR^{+}$ or $^{+}NR^{-}$ spectrum, respectively. This is then compared to the corresponding charge reversal spectrum ($^{-}CR^{+}$ or $^{+}CR^{-}$) of the precursor ion, where charge inversion is afforded by a two-electron transfer in a *single* collision (either one of the two collision cells in **Figure 1** can be used for this

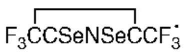
purpose). Under *single* collision conditions, the CR signals originate solely through ionic fragmentations of the charge-inverted ions. In contrast, the NR signals, which result from two spatially separated collisions involving neutral intermediates, may arise *either* from the neutral *and/or* the reionized ions. Normalization of the spectra to the sum of all fragments followed by subtraction of the CR from the NR intensities (Eqn [4]) results in the NIDD spectrum, in which processes due to the fragmentation of the neutrals have positive intensities, while ionic contributions (of the reionized ions) show up as negative peaks. Hence, the neutral and ionic processes of NR are readily deconvoluted.

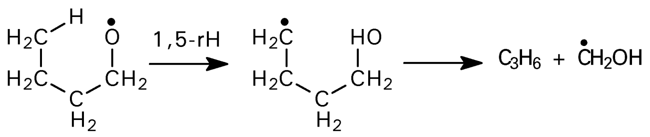
$$I_i(\mathrm{NIDD}) = [I_i(\mathrm{NR})/\Sigma_i I_i(\mathrm{NR})] - [I_i(\mathrm{CR})/\Sigma_i I_i(\mathrm{CR})] \quad [4]$$

where $I_i$ is the intensity of the *i*th ion.

NIDD is most suitable for the study of anionic precursors, which easily undergo two-electron loss in one step. The method has been applied extensively to alkoxide anions to decipher the chemistry of the corresponding radicals and cations, both of which are more reactive and less readily accessible than the anions. As an example, **Figure 10** shows the $^-$NR$^+$, $^-$CR$^+$, and $^-$NIDD$^+$ spectra of *n*-butoxide anion, $CH_3CH_2CH_2CH_2O^-$. Three prominent features are worth mentioning. (a) The recovery peaks are weak, consistent with the high reactivity expected for the oxenium cation $CH_3CH_2CH_2CH_2O^+$. (b) The favoured dissociations of this elusive cation are identified by the negative signals in **Figure 10C** and primarily lead to $C_3H_5^+$, $C_3H_3^+$, and $C_2H_5^+$ and $C_2H_3^+$; most likely, these products arise from elimination of $CH_2O$ or $C_2H_4O$ (inductive cleavages), followed by one or more dehydrogenations. (c) The fragments generated from the radical $CH_3CH_2CH_2CH_2O^{\bullet}$ are revealed by the positive signals in **Figure 10C.**

**Table 1** Examples of elusive neutral intermediates that have been synthesized and characterized by NRMS.

| *Intermediate type* | *Examples* |
|---|---|
| Ylides | $H_2N^+{=}N^-$(*iso*-diazine), $^-HC{=}N^+{=}NH$ (nitrileimine), $^-HC{=}N^+{=}CH_2$ (*N*-methylide of HCN), $C_3H_3NS$ (thiazole-2-ylide), $C_5H_5N$ (pyridine-2-ylide) |
| Carbenes | H–C–$NH_2$ (aminocarbene), H–C–$NO_2$ (nitrocarbene), $H_2N$–C–$NH_2$ (diaminocarbene), HO–C–OH (dihydroxycarbene), **:**C=C=O (carbonyl carbene), $CH_3O$–C–$OCH_3$, H–C–$OC_2H_5$ (ethoxycarbene), **:**C=C(H)CN (cyanovinylidene), $C_3H_4N_2$ (imidazol-2-ylidene and imidazol-4-ylidene), $C_3H_5NO$ (2-oxazolidinylidene), $H_2CCCC$:, $CH_3O$–C–$OCH_2CF_3$ |
| Cumulenes | O=C=C=O (ethylenedione), $NNCCN^{\bullet}$, $NCNCN^{\bullet}$, $NCCO^{\bullet}$, $CNCO^{\bullet}$, $CCNO^{\bullet}$, $OCCN^{\bullet}O$, $OCN^{\bullet}CO$, $NCCO_2^{\bullet}$, NCNCS, O=C=C=C=O (carbon suboxide), $NCCCO^{\bullet}$, O=C=C=CH–$CO_2H$ |
| Radicals | $HOSO_2^{\bullet}$ (bisulfite radical), $HOCO_2^{\bullet}$ (bicarbonate radical), $FC^{\bullet}(OH)_2$, $^{\bullet}CH_2NHNH_2$, $^{\bullet}CH_2CH_2Br$, $^{\bullet}CH_2N(H)$–COOH, $H_2NCH_2C^{\bullet}(OH)_2$, $(CH)_3)_2NCH_2^{\bullet}$, $CH_3CH_2C^{\bullet}(OH)_2$, $C_4H_6N^{\bullet}$ (neutral form of C–2 and C–3 protonated pyrrole), $C_4H_6N^{\bullet}$ (neutral form of N-3 protonated imidazole), $C_4H_5N_2^{\bullet}$ (neutral form of protonated pyrimidine), $C_4H_6N_3^{\bullet}$ (neutral form of protonated 2- or 4-pyrimidinamine) |
| Hypervalent radicals | $CH_2{=}CH(OH)_2^{\bullet}$; $C_2H_5OH_2^{\bullet}$; $(C_2H_5)_3O^{\bullet}$; *n*-$C_7H_{13}O^{\bullet}H(CH_3)$, $CH_3O(CH_2)_nOH^{\bullet}(CH_3)$ ($n = 2,4$), $(CH_3)_2N^{\bullet}H_2$ (several isotopomers), $H_3N^{\bullet}CH_2CO_2H$, $(CH_3)_3N^{\bullet}H$, $(CH_3)_4N^{\bullet}$, *n*–$C_6H_{11}N^{\bullet}H_3$, $C_6H_5CH_2N^{\bullet}H_3$, $(C_2H_5)_4N^{\bullet}$, *n*–$C_6H_{11}N^{\bullet}H(CH_3)_2$, $C_6H_5CH_2N^{\bullet}H_2(CH_3)$, *n*–$C_7H_{15}N^{\bullet}H(CH_3)_2$, $C_6H_5CH_2N^{\bullet}H(CH_3)_2$, $C_6H_5N^{\bullet}(CH_3)_3$, $C_6H_5CH_2N^{\bullet}(CH_3)_3$, $C_6H_5CH_2N^{\bullet}(C_2H_5)_3$, $HN(CH_2CH_2)_2N^{\bullet}H(CH_3)$, $N(CH_2CH_2)_3N^{\bullet}H$, $CH_3N(CH_2CH_2)_2N^{\bullet}H(CH_3)$, $(CH_3)_2N(CH_2)_nN^{\bullet}H(CH_3)_2$ ($n = 2,3,4,6$), $(NO_2)_2C_6H_4CH_2N^{\bullet}H(CH_3)_2$, $C_6F_5CH_2N^{\bullet}H(CH_3)_2$, $(C_2H_5)_2X^{\bullet}$ (X = Cl,Br,I), $H_3S^{\bullet}$ (various isotopomers) |
| Diradicals | $^{\bullet}CH_2OSi^{\bullet}$, $^{\bullet}CH_2NHCH_2^{\bullet}$, $^{\bullet}CH_2COO^{\bullet}$ (acetoxy), $^{\bullet}CH_2CH_2OCH_2^{\bullet}$, $^{\bullet}CH(CH_3)OCH_2^{\bullet}$, $^{\bullet}CH_2CH_2CH(OH)^{\bullet}$, $^{\bullet}CH_2CH_2SCH_2^{\bullet}$, $^{\bullet}CH_2CH_2CH_2CH(OH)^{\bullet}$ |
| Intermediates in atmospheric chemistry | $NH_2NO$ (nitrosamine), $HSO^{\bullet}$, $SOH^{\bullet}$, HSOH (hydrogen thioperoxide), $SOH_2$ (thiooxonium ylide), $O{=}S^{\bullet}OH$ (hydroxysulfinyl radical), $O_2S^{\bullet}H$ (hydrogensulfonyl radical), HOSOH (sulfinic acid), $CH_3SCH^{\bullet}$, $(CH_3)_2S^{\bullet}OH$ |
| Clusters | NSS, SNS (nitrogen disulfide), SSNS, $O_2SOSO$ (sulfur dioxide dimer), He@$C_{60}$ |
| Reactive closed-shell species | HOSSOH (dihydroxy disulfide), $FCO_2H$ (fluoroformic acid), RCONO (R = H, $CH_3$, $CF_3$), HC≡NS (thiofulminic acid), $CH_3O$–P=O, $CH_3S$–P=O, $CH_3PS_2$, $CH_3S$–P=S, $CH_3C$≡NS, $CH_3H_7NO$ (3-methyl-2-aziridinone), $CH_3CH{=}CH$–SH, $C_6H_5C$≡NS (benzonitrile–sulfide), $C_6H_5PS_2$, $C_6H_5S$–P=S |
| Organometallic species | $Fe(O_2)$ (peroxide), OFeO (dioxide), Fe(CO), $Fe(C_2H_4)$,$C_5H_5$–Fe–R (R = halogen, O, OH, $OCH_3$, $C_6H_5$, H), $C_5H_5FeC_5H_4{=}X$ (X=O, $CH_2$, CO), $SiNH_2^{\bullet}$, $Si_2O_2$, $Si_3N$, $D_2Si{=}CH_2$ (silaethene), $SiC_nH^{\bullet}$ ($n$=4,6) (silicon carbon hydrides), $(CH_3)_3SiC^{\bullet}{=}CH_2$, $(CH_3)_3SiCH{=}CH^{\bullet}$, $PrF_n$($n$ = 1-2), $(C_5H_5)_2Zr$ (zirconocene),  |




**Scheme 4**

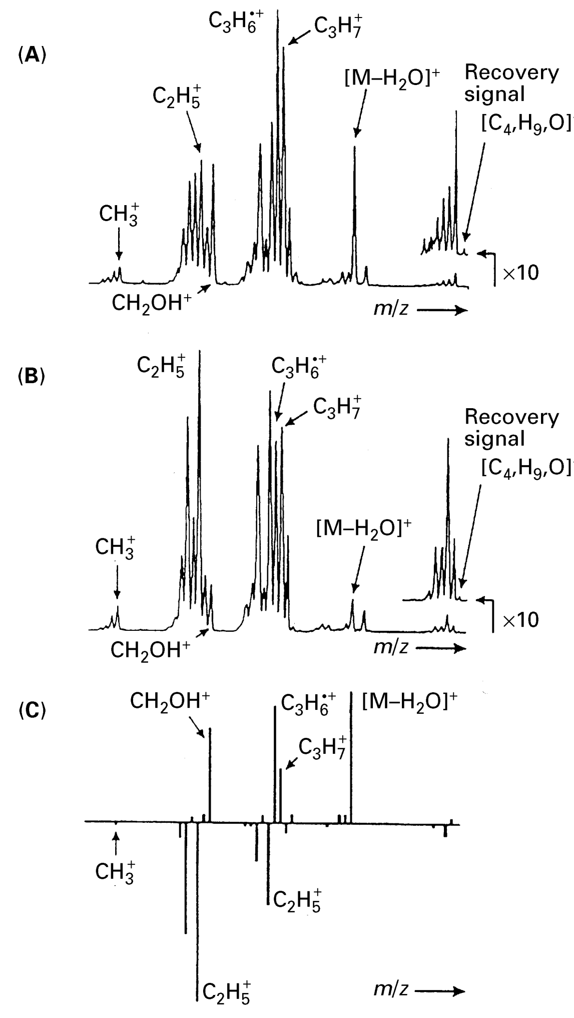



**Figure 10** (A) Neutralization–reionization ($^{-}$NR$^{+}$) spectrum of *n*-butoxide anions using $O_2$ for neutralization and reionization. (B) Charge reversal ($^{-}$CR$^{+}$) spectrum of *n*-butoxide anions using $O_2$. (C) $^{-}$NIDD$^{+}$ spectrum of *n*-butoxide anions. Reprinted with permission of Wiley-VCH from Hornung G, Schalley CA, Dieterle M, Schröder D and Schwarz H (1997) A study of the gas-phase reactivity of neutral alkoxy radicals by mass spectrometry: $\alpha$-cleavages and Barton-type hydrogen migrations. *Chemistry: a European Journal* **3**: 1866–1883.

Among them are the $\alpha$-cleavage products $C_3H_7^{\bullet}$+$CH_2O$ (weak); the intramolecular hydrogen transfer products $C_3H_6$+$^{\bullet}CH_2OH$; as well as the product of $H_2O$ loss which indicates the occurrence of a double hydrogen transfer to oxygen. Repeating these experiments with $CD_3CH_2CH_2CH_2O^{\bullet}$ reveals that the intramolecular H-rearrangement involves a 1,5-hydrogen migration (**Scheme 4**), similar to the Barton reaction observed in solution. On the other hand, the water elimination from $CD_3CH_2CH_2CH_2O^{\bullet}$ is found to proceed with partial H/D scrambling, leading to the losses of $H_2O$, HDO, and $D_2O$ in the ratio 1:10:5.

## Reactive neutrals investigated and outlook

The basic NRMS method and the auxiliary techniques outlined have enabled the discovery and characterization of a large variety of reactive neutrals that had eluded experimental studies due to their unusual structures and reactivities. The species studied so far are summarized in **Table 1** and include radicals, diradicals, carbenes, nitrenes, cumulenes, ylides, hypervalent species, and weak intermolecular complexes.

NRMS has primarily been a qualitative method. It determines whether a reactive neutral species is bound and to what it rearranges or dissociates, but it does not easily yield *quantitative* insight about bond dissociation energies and isomerization barriers or about the thermochemistry of the neutral under study; these data have generally been obtained by parallel theoretical calculations. NRMS also does not allow the assessment of the *bimolecular* reactivity of neutral intermediates, which is of paramount interest in organic and biological chemistry and in catalysis. However, in light of the new knowledge NRMS has yielded in less than two decades thus far (**Table 1**), the shortcomings mentioned provide the motivation for future efforts to overcome them.

*See also:* **Ion Dissociation Kinetics, Mass Spectrometry; Ion Energetics in Mass Spectrometry; Ion Imaging Using Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Ion Structures in Mass Spectrometry; Ion Trap Mass Spectrometers.**

## Further reading


Busch KL, Glish GL and McLuckey SA (1988) *Mass Spectrometry/Mass Spectrometry*. New York: VCH.

Goldberg N and Schwarz H (1994) Neutralization-reionization mass spectrometry: a powerful "laboratory" to generate and probe elusive neutral molecules. *Accounts of Chemical Research* **27**: 347–352.

Holmes JL (1989) The neutralization of organic cations. *Mass Spectrometry Reviews* **8**: 513–539.

McLafferty FW (1990) Studies of unusual simple molecules by neutralization-reionization mass spectrometry. *Science* **247**: 925–929.

McLafferty FW (1992) Neutralization-reionization mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **118/119**, 221–235.

Polce MJ, Beranova S, Nold MJ and Wesdemiotis C (1996) Characterization of neutral fragments in tandem mass spectrometry: a unique route to mechanistic and structural information. *Journal of Mass Spectrometry* **31**: 1073–1085.

Schalley CA, Hornung G, Schröder D and Schwarz H (1998) Mass spectrometric approaches to the reactivity of transient neutrals. *Chemical Society Reviews* **27**: 91–104.

Terlouw JK and Schwarz H (1987) The generation and characterization of molecules by neutralization-reionization mass spectrometry (NRMS). *Angewundte Chemie, International Edition; English* **26**: 805–815.

Turecek F (1992) The modern mass spectrometer: a chemical laboratory for unstable neutral species. *Organic Mass Spectrometry* **27**: 1087–1097.

Wesdemiotis C and McLafferty FW (1987) Neutralization-reionization mass spectrometry (NRMS). *Chemical Reviews* **87**: 485–500.

Zagorevskii DV and Holmes JL (1994) Neutralization-reionization mass spectrometry applied to organometallic and coordination chemistry. *Mass Spectrometry Reviews* **13**: 133–154.


# Neutron Diffraction, Instrumentation

**AC Hannon**, Rutherford Appleton Laboratory, Didcot, UK

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Methods & Instrumentation**

Neutron diffraction is a very powerful technique for investigating the structure of condensed matter. Crystal structures can be studied, either by diffraction from a polycrystalline powder or from a single crystal sample. Neutron diffraction is also a very important tool for the study of the structure of noncrystalline forms of matter, such as liquids and glasses. This article considers in detail the instrumentation for studying isotropic samples, either polycrystalline powders or non-crystalline materials. Neutrons may be obtained either from a nuclear reactor or from an accelerator-based source. They may be detected using either gas detectors or scintillator detectors. Typical neutron diffractometers at both reactor and accelerator-based sources are described, and a consideration of the resolution is given in each case.

## The neutron

The neutron was discovered by Chadwick in 1932. It is a neutral subatomic particle which has a finite mass, $m_n = 1.0087$ amu, similar to that of the proton. It has a spin of $\frac{1}{2}$ and a magnetic moment of −1.9132 nuclear magnetons. A neutron with speed $v$ has a de Broglie wavelength given by

$$\lambda = \frac{h}{m_n v} \qquad [1]$$

and thus the neutron exhibits wavelike behaviour including diffraction.

Normally neutrons only exist within the atomic nucleus and hence a nuclear reaction of some sort must be used in order to produce a beam of neutrons for neutron diffraction. Chadwick first produced neutrons by the interactions in beryllium of alpha particles from the decay of natural polonium. The first experimental demonstrations of the phenomenon of neutron diffraction were performed in 1936 and these also used radioactive sources. However, the neutron flux available from the decay of a radioactive source is very weak, and since neutron scattering is an intensity-limited technique, all neutron diffraction

experiments soon came to be performed using neutrons produced by either a nuclear reactor or an accelerator-based neutron source, both of which produce a much greater flux. The properties of the neutron source have important consequences both for the way in which a neutron diffraction experiment is performed, and for the results obtained, and hence it is worthwhile to discuss below the two important types of neutron source.

## Neutron sources

### Reactor sources of neutrons

Conventionally, from the 1940s onwards, neutron diffraction experiments were performed using a beam of neutrons derived from a nuclear reactor in which the neutrons are produced by the fission of $^{235}$U nuclei. The cross-section for neutron-induced fission of $^{235}$U is only high for slow neutrons with energies in the meV range, whereas the fast neutrons produced by fission have high energies in the MeV range. Hence, in order to sustain the fission process, a reactor includes a component, known as a moderator, which slows down the neutrons. The neutrons undergo inelastic collisions with the nuclei in the moderator so that they are in thermal equilibrium at the temperature of the moderator. The moderator normally contains large numbers of low mass nuclei (usually H or D) because the energy transferred in the inelastic collisions is maximized when the mass of the colliding nucleus is as close as possible to the neutron mass. The peak flux within the moderator is at a neutron speed $v_p$ given by

$$\frac{1}{2}m_n v_p^2 = k_B T \qquad [2]$$

where $T$ is the temperature of the moderator. For example, a temperature of 290 K corresponds to a neutron energy $E$ of 25 meV, a neutron wavelength $\lambda$ of 1.8 Å (0.18 nm), or a neutron speed $v$ of 2200 m s$^{-1}$. In order to perform a neutron diffraction experiment to study the atomic structure of condensed matter it is necessary to use neutrons whose wavelength is of a similar order of magnitude to the interatomic separations in materials. It is thus fortuitous that the process of moderation produces neutrons which, as well as being slowed down for maintaining the fission reaction, also have a wavelength suitable for performing neutron diffraction experiments.

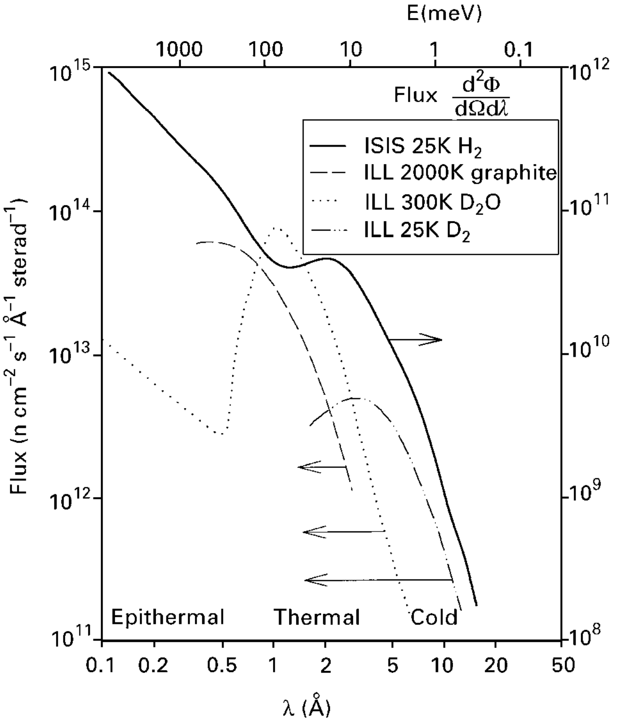



**Figure 1** The neutron flux distribution for three different moderators at the ILL reactor and for the liquid hydrogen moderator at the ISIS accelerator. The accelerator flux distribution is adjusted by a factor $10^3$ to represent the increased efficiency for time-of-flight experiments due to the pulse structure. Modified with permission from Price DL and Sköld K (1986) Introduction to neutron scattering in: *Neutron Scattering, Part A*, Sköld K and Price DL (eds). Orlando: Academic Press.

A neutron diffractometer uses a beam of neutrons which is obtained by viewing a moderator through a beam-tube or neutron guide which passes through the shielding around the neutron source. Note that, in practice, the moderator used as a source of neutrons for neutron diffraction experiments at a reactor may be separate from the moderator used to slow the neutrons in order to maintain the fission reaction. **Figure 1** shows the neutron flux for three different moderators at the world's preeminent reactor source of neutrons, the Institut Laue–Langevin (ILL) in Grenoble, France (see **Figure 2**). Reactor neutron sources produce a high flux of thermal neutrons ($E$~25 meV, $T$~290 K) and cold neutrons ($E$~1 meV, $T$~12 K), but they have little flux at higher epithermal energies ($E$~1 eV, $T$~12 000 K). This is a consequence of the fact that a reactor can only produce neutrons which are in thermal equilibrium with a moderator and there are practical limitations on the maximum temperature of the moderator.

The neutron flux produced by a normal nuclear reactor is unchanging with time and covers a wide

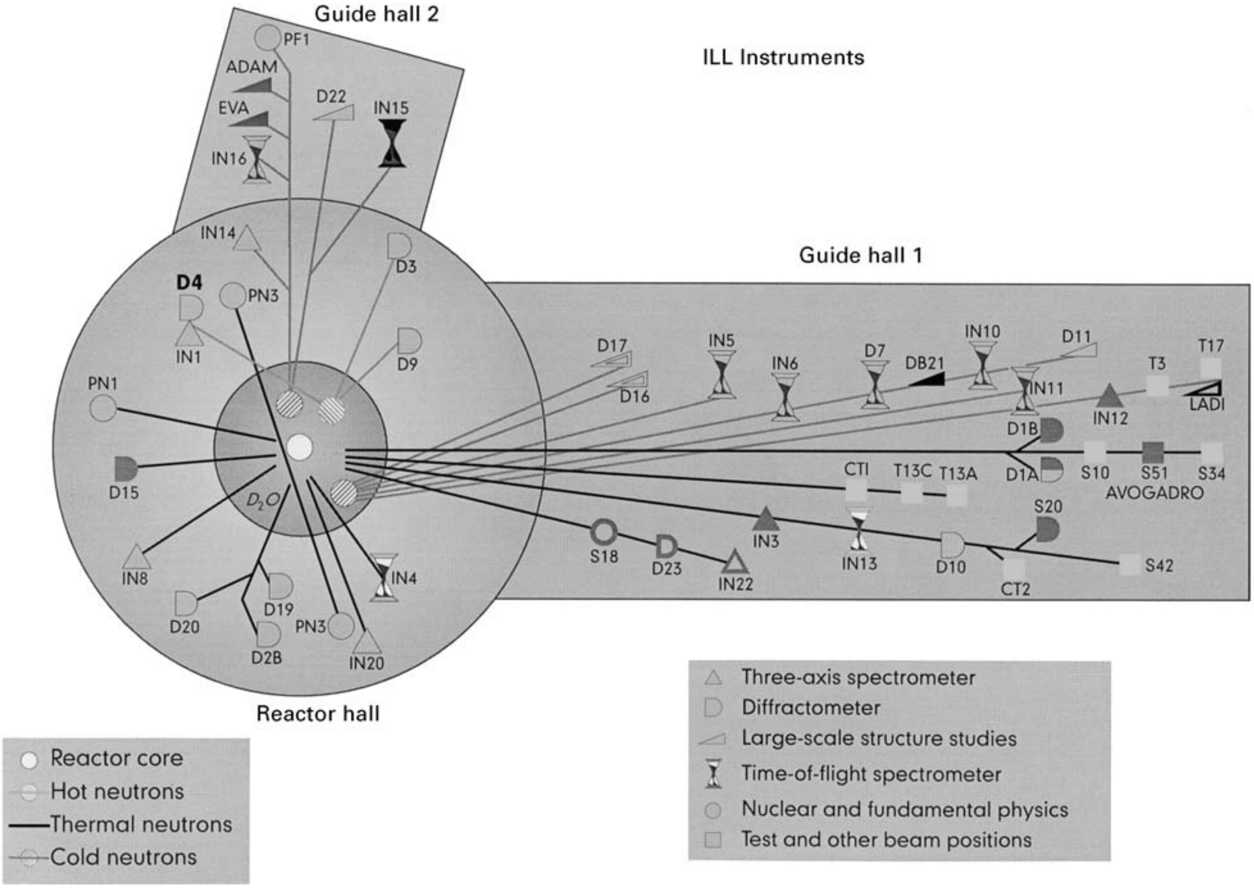



**Figure 2** The layout of the reactor and the neutron scattering instrumentation at the ILL reactor. Courtesy of H. Büttner.

range of neutron wavelengths. In order to perform a neutron diffraction experiment it is thus necessary to monochromate the neutron beam from a reactor so that it covers a narrow range of neutron wavelengths and the vast majority of the flux from the source is lost at this stage.

### Accelerator-based sources of neutrons

In more recent years neutron diffraction experiments have increasingly come to be performed using sources of neutrons which are based on a particle accelerator. A beam of charged particles is accelerated to a high energy and then fired at a target. Interactions between the particle beam and the nuclei in the target produce high energy neutrons which are then slowed down by a moderator.

The earlier accelerator-based neutron sources from the 1960s and 1970s use an electron linac to accelerate an electron beam to relativistic energies (~50 MeV). The electron beam is fired at a dense target made of a heavy element, usually uranium, and neutrons are produced by a two-stage process. First, the electrons are slowed down extremely rapidly due to the strong interaction with the electromagnetic field of the target nuclei, producing a cascade of bremsstrahlung photons. Secondly, some of these photons go on to produce neutrons by photoneutron reactions where the photon excites a target nucleus which subsequently decays with the emission of a neutron. About twenty electrons must be accelerated for each neutron produced.

More recent accelerator-based neutron sources use a linear accelerator, usually together with a synchrotron, to accelerate a beam of protons to a high energy (~800 MeV). The proton beam is fired at a heavy metal target (made, for example, of tantalum, uranium or tungsten) and neutrons are produced by the spallation process. Spallation is a violent interaction between the proton and the target nucleus which results primarily in the emission of neutrons, but also a variety of light nuclear fragments. In effect, the protons chip pieces off the target nuclei and each proton produces of the order of fifteen neutrons for a nonfissile target (or about twenty-five neutrons for a fissile target).

Accelerator-based sources are usually pulsed (typically with a pulse repetition rate of the order of

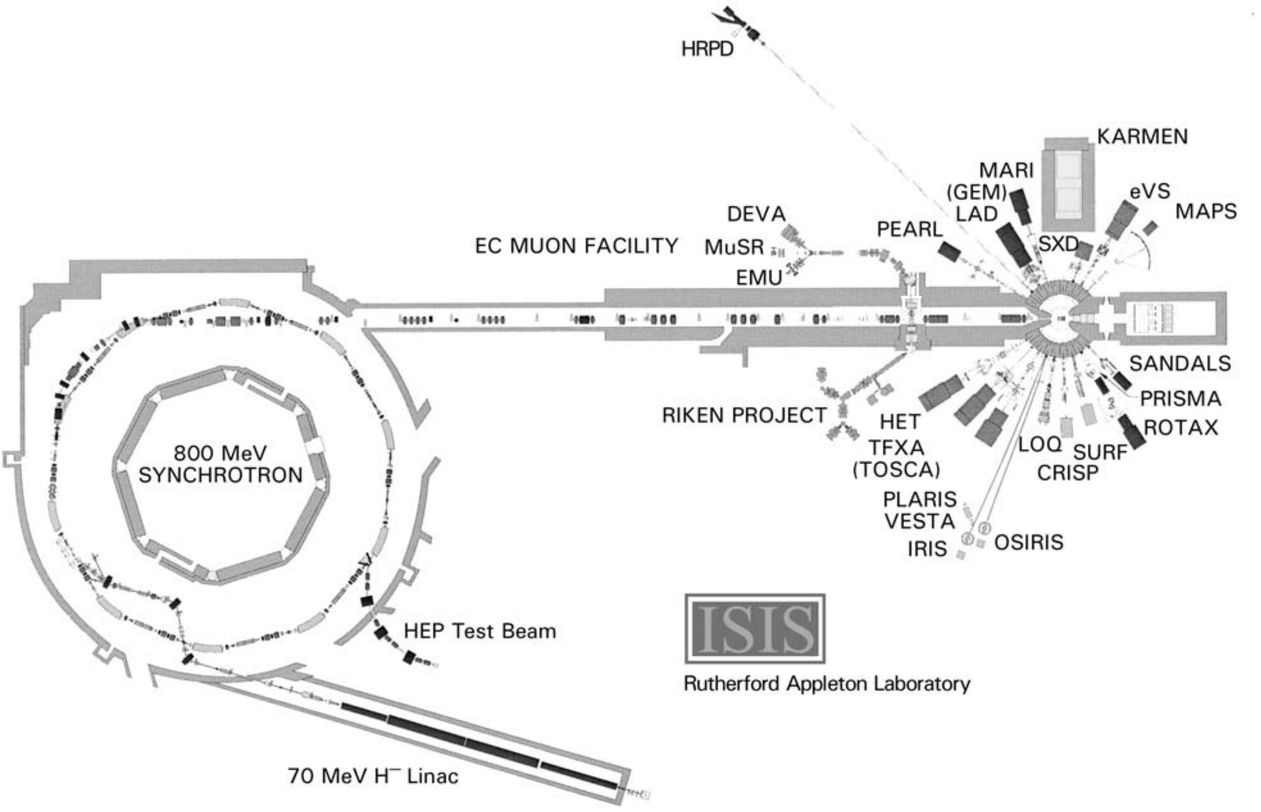



**Figure 3** The layout of the neutron source and the neutron scattering instrumentation at the ISIS spallation neutron facility.

50 Hz) and so they produce a pulsed neutron flux which is ideally suited to the time-of-flight neutron diffraction technique. (Accelerator-based neutron sources which are quasi-steady-state or intensity-modulated also exist. Furthermore, pulsed neutrons have also been produced in Russia by using a pulsed reactor.) This technique involves measuring the time-of-flight, $t$, taken for a neutron to travel the total flight path, $L$, from the moderator to the detector, via the sample. On the assumption of elastic scattering (i.e. initial and final neutron energies are the same, $E_i = E_f$) then

$$t = \frac{m_n}{h} L\lambda \qquad [3]$$

(or $t = 252.82L\lambda$ in convenient units of μs, metres and Ångstroms, respectively) and it is thus straightforward to determine the neutron wavelength. The use of the time-of-flight technique removes the need to monochromate the neutron beam and thus, even though the raw flux produced initially by an accelerator-based source is much less than that produced by a reactor source, the final flux available for neutron diffraction is of a comparable order of magnitude (see **Figure 1**).

The moderator at an accelerator-based neutron source is used to slow the neutrons down so that they have suitable wavelengths for neutron diffraction, in the same way as for a reactor neutron source. However, in order that the moderation process does not broaden the pulsed time structure of the neutron flux too much, the moderator must be relatively small. (Also note that, unlike a reactor, the process of moderation plays no role in the production of neutrons at an accelerator-based source.) This has the consequence that the neutrons produced by an accelerator-based source are under-moderated and there are many more epithermal neutrons than for a reactor source. **Figure 1** also shows the neutron flux for a moderator at the world's most intense pulsed neutron source, the ISIS spallation neutron source at the Rutherford Appleton Laboratory, UK (see **Figure 3**).

## Neutron detectors

The single most important component of a neutron diffractometer is the detector array. The lack of a

charge on the neutron means that it cannot directly create an electrical signal in a detector. Instead the neutron must be made to undergo a nuclear reaction which subsequently leads to an electrical signal.

The three most important reactions for neutron detection involve the $^3$He, $^{10}$B and $^6$Li nuclei:

$$
\begin{aligned}
&{}^3_2\mathrm{He} + {}^1_0\mathrm{n} \rightarrow {}^3_1\mathrm{H} + {}^1_1\mathrm{p} + 0.770\ \mathrm{MeV}\\
&{}^{10}_{5}\mathrm{B} + {}^1_0\mathrm{n} \rightarrow {}^7_3\mathrm{Li} + {}^4_2\mathrm{He} + 2.3\ \mathrm{MeV}\\
&{}^6_3\mathrm{Li} + {}^1_0\mathrm{n} \rightarrow {}^4_2\mathrm{He} + {}^3_2\mathrm{He} + 4.79\ \mathrm{MeV}
\end{aligned}
\qquad [4]
$$

Generally, the efficiency of detection of neutrons of wavelength, $\lambda$, by a detector of thickness $x$, containing a number density, $n_v$, of absorbing atoms with neutron absorption cross-section, $\sigma_{abs}(\lambda)$, may be expressed as

$$\text{Efficiency} = \varepsilon(1 - \exp(-n_v x \sigma_{abs}(\lambda))) \qquad [5]$$

where $\varepsilon$ is the fraction of absorption events which result in an output pulse from the detector. Usually the absorption cross-section used in detection is proportional to the neutron wavelength (a so-called $1/v$ cross-section):

$$\sigma_{abs}(\lambda) = \lambda \sigma^0_{abs} \qquad [6]$$

where $\sigma^0_{abs}$ is a constant. For good performance the ideal neutron detector should have a high neutron detection efficiency, a low intrinsic detector background, a low sensitivity to non-neutron events (particularly gamma rays), a good stability and a short dead-time.

## Gas detectors

A simple gas detector consists of a sealed metal tube, which contains either $^3$He or $BF_3$ gas under pressure (typically 10 bar; $10^6$ Pa), with a high voltage thin wire anode along its axis (see **Figure 4A**). When a neutron is absorbed (Eqn [4]) the energetic recoil particles produce ionization in the gas. The anode then collects electrons from the ionization and an electrical signal is produced. Although $BF_3$ is cheaper than $^3$He, it is little used now because it is less efficient and the gas is toxic. An advantage of using $^3$He gas is that it is self-regenerating because the triton produced when a neutron is absorbed undergoes beta decay with a half-life of 12.33 years to produce another $^3$He nucleus:

$$ {}^3_1\mathrm{H} \rightarrow {}^3_2\mathrm{He} + e^- + \bar{\nu}_e + 18.6\ \mathrm{keV} \qquad [7]$$

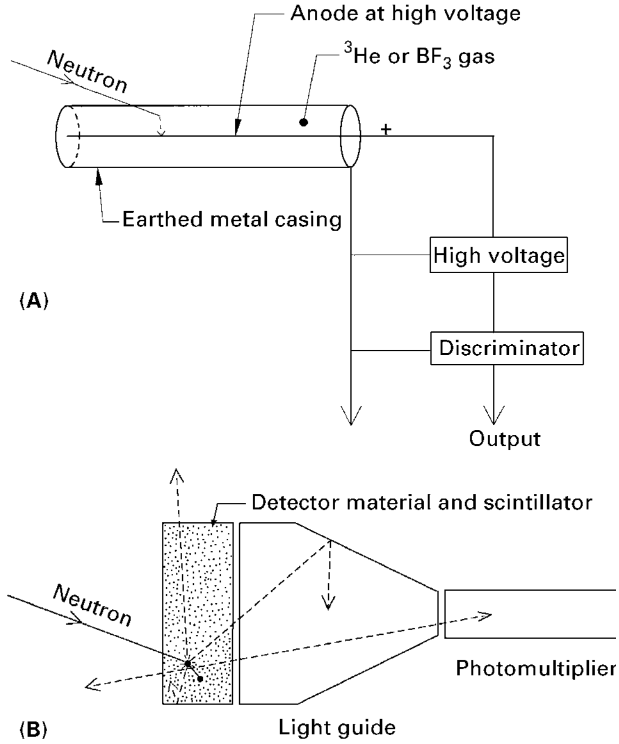



**Figure 4** Schematic of (A) a gas detector and (B) a scintillator detector (modified with permission from Windsor CG (1981) *Pulsed Neutron Scattering*. London: Taylor and Francis).

Gas detectors have a pulse height distribution in which the signal due to gamma rays is well separated from the neutron signal. Hence, the detector electronics can be made to discriminate very well against gamma rays, leading to a very low gamma ray sensitivity (~$10^{-8}$ at 1 MeV). Furthermore, gas detectors have a very low instrinsic detector background (at best ~3–5 counts per hour) and a good stability. In practice, modern neutron diffractometers often use multiwire detectors, where a large number of anodes are contained within a single envelope, in order to cover a larger solid angle and hence obtain a higher count rate. Also resistive-wire position-sensitive detectors are sometimes used for applications where it is useful to be able to determine the position along the length of the detector where a neutron is detected.

Gas detectors are relatively expensive and it is difficult to make the active element take up more complex geometries and shapes. Furthermore, it is difficult to make the active elements small, which is advantageous for good diffractometer resolution.

The most recent development in gas detector technology to be applied is the microstrip detector. In a microstrip detector the metal anode is replaced by a semiconducting glass plate on which very thin

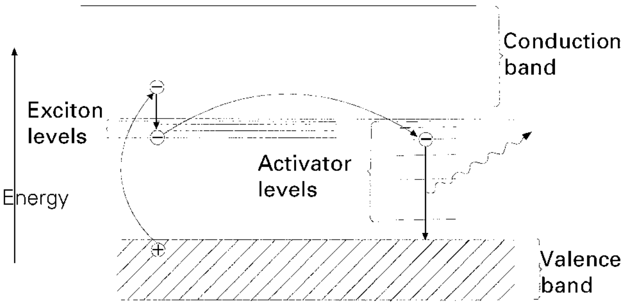



**Figure 5** The physical principles behind the operation of a scintillator material. Reproduced with permission from Bennington SM, Hannon AC and Forsyth JB (1993) *Rutherford Appleton Laboratory Report* RAL-93-083.

metallic strips are engraved by photolithographic means. In this way, the limitations on the size and shape of gas detector elements may be overcome in the future.

### Scintillator detectors

The aim with a scintillator detector is to combine the neutron absorbing element intimately with a scintillating phosphor so that the reaction products from the capture of a neutron strike the phosphor and produce a light flash, which is then detected by a photomultiplier tube (see **Figure 4B**).

There are two types of scintillator material in common use for neutron detection: $^6$Li/Ce glass scintillator and granular ZnS(Ag)/$^6$Li scintillator. **Figure 5** illustrates the physical principles behind the operation of an ideal scintillator material. The absorption of a neutron by a $^6$Li nucleus causes a large amount of energy to be deposited in the scintillator material, resulting in the excitation of electrons from the valence band into the conduction band. An excited electron can bind together with a hole to form an exciton which propagates through the lattice. When the exciton reaches an impurity activator atom (Ce for the glass scintillator, or Ag for the ZnS scintillator) it becomes trapped and then it recombines to emit a flash of light. Since the activator atoms have levels within the band gap of the scintillator material, the photon emitted has an energy less than the band gap and can pass through the material without absorption.

The earlier scintillator detectors used at ISIS were based on the lithium glass material. Such detectors have a high neutron detection efficiency (comparable to or even exceeding that of a gas detector), but also a very high gamma sensitivity (0.02 at 1 MeV), a large instrinsic detector background (~150 counts per minute) and a poor stability. Thus, neutron diffraction results obtained using lithium glass scintillator detectors can suffer from background and absolute normalization problems. More recent scintillator detectors are almost always based on ZnS. This gives a high neutron detection efficiency (comparable to, or even exceeding, that of a gas detector), a low gamma sensitivity (at best ~ $10^{-8}$ at 1 MeV), a low intrinsic detector background (at best ~ 12 counts per hour) and a much improved stability.

One of the advantages of scintillator detectors is that a high efficiency can be obtained in a small thickness of material because a higher density of absorbing atoms can be achieved than is possible with a gas detector (cf. Eqn [5]). This can be of particular importance for achieving a large efficiency for the detection of higher energy neutrons, for which the absorption cross-section is relatively small (cf. Eqn [6]). Another advantage of scintillator detectors is that a high degree of flexibility can be achieved in terms of the size and geometry of the active component. For example, this allows the detectors to be designed to be narrow and to follow the Debye–Scherrer cone so that the resolution width for the diffractometer is minimized.

## Neutron diffraction instrumentation

### Instrumentation – general principles

The purpose of a total neutron diffractometer is to measure the differential cross-section

$$\left(\frac{d\sigma}{d\Omega}\right)_{tot} = \frac{R_{tot}}{N\Phi d\Omega} \quad [8]$$

where $R_{tot}$ is the rate at which neutrons of wavelength $\lambda$ are scattered into the solid angle $d\Omega$ in the direction $(2\theta, \phi)$ (see **Figure 6**), regardless of whether or not they are scattered elastically (i.e. total diffraction). $N$ is the number of atoms in the sample and $\Phi$ is the flux of neutrons of wavelength $\lambda$ which is incident on the sample.

In general, the differential cross-section depends upon the scattering vector

$$\boldsymbol{Q} = \boldsymbol{k}_i - \boldsymbol{k}_f \quad [9]$$

where $\boldsymbol{k}_i$ and $\boldsymbol{k}_f$ are the neutron wave vector before and after scattering. The term momentum transfer (i.e. the momentum transferred to the sample) is commonly used for $Q$, although strictly speaking this term should be used for $\hbar Q$. Diffraction data

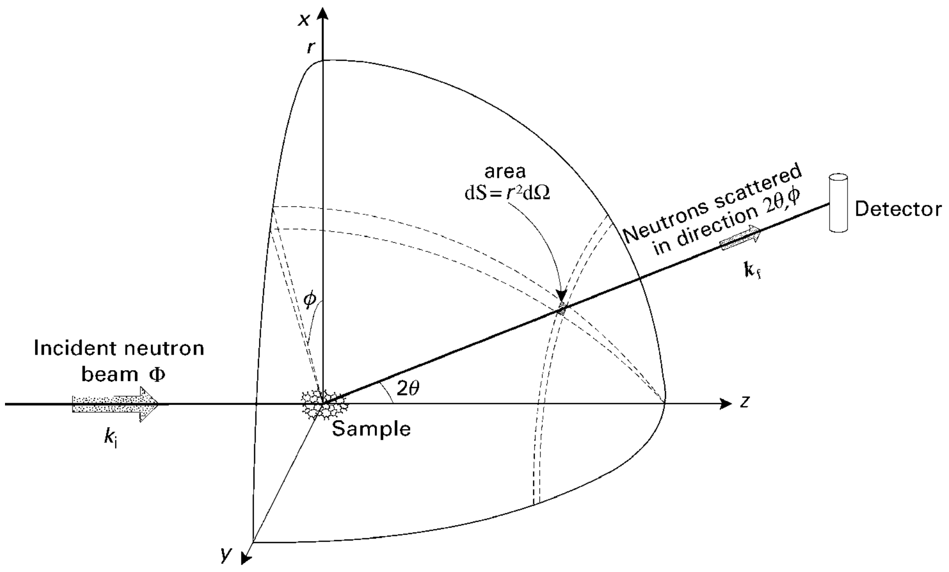



**Figure 6** The geometry for a neutron diffraction experiment.

may be treated by considering the scattering to be totally elastic so that the magnitudes of the initial and final neutron wave vectors are the same,

$$|\boldsymbol{k}_i| = |\boldsymbol{k}_f| \qquad [10]$$

The most frequently used type of neutron diffraction involves the study of the atomic structure of isotropic samples (polycrystalline powders, glasses, liquids, etc.) and hence this article gives greatest emphasis to the instrumentation for this type of experiment. For an isotropic sample, it is only the magnitude of the momentum transfer which is important (the direction is not important), and in this case the differential cross-section is a function of a single variable

$$Q = |\,\boldsymbol{Q}\,| = \frac{4\pi \sin\theta}{\lambda} \qquad [11]$$

For neutron diffraction experiments on polycrystalline powders it is usually more convenient to treat the differential cross-section as a function of $d$-spacing, $d(= 2\pi/Q)$, defined according to Bragg's law by

$$2d \sin\theta = \lambda \qquad [12]$$

Bragg peaks are then observed in the differential cross-section whenever the $d$-spacing satisfies

$$d = d_{hkl} \qquad [13]$$

where $d_{hkl}$ is a $d$-spacing between atomic planes in the crystal for which the structure factor is non-zero.

In order to produce high quality data, a neutron diffractometer must satisfy several requirements: The data must have a good statistical accuracy, which is obtained by having a high count rate. This is achieved by such factors as a bright source and a large total detector solid angle. The corrections which must be made to the data must be small. In particular, the background must be small, featureless and unchanging. The range in $Q$ must be as wide as possible, in order to provide high resolution in real-space. The reciprocal-space resolution must be as narrow as possible (see below).

## Reactor source instrumentation

**Figure** 7 shows a schematic of the layout of a typical neutron diffractometer at a reactor source. The neutron beam coming from the moderator at a reactor covers a wide range of wavelengths and is unchanging with time. It is thus necessary to use a single crystal monochromator so as to produce a monochromatic beam. The general principle of operation of a neutron diffractometer at a steady state source is the same as for a conventional X-ray diffractometer. The differential cross-section is measured as a function of $Q$ by moving the detector to different scattering angles, $2\theta$, and measuring the scattered count

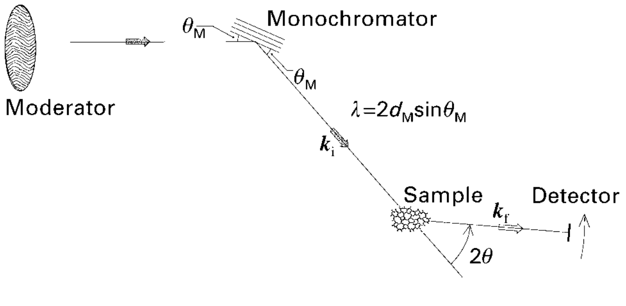



**Figure 7** Schematic of a neutron diffractometer for a steady-state (reactor) source.

rate. That is to say, $Q$ (or $d$) is scanned by varying $2\theta$ whilst keeping the neutron wavelength $\lambda$ constant (cf. Eqn [11]).

**Figure** 8 shows the D4 diffractometer at the ILL reactor which, for many years, has been the most successful reactor-based diffractometer for studying the structure of liquids and amorphous materials. The diffractometer uses neutrons from a hot graphite moderator at a temperature of 2400 K because this produces neutrons with short wavelengths and hence a high maximum $Q$ can be achieved, leading to good real-space resolution. A copper monochromator is used to produce neutrons with a wavelength of 0.7 Å, 0.5 Å or 0.35 Å, depending upon which copper reflection is selected. Most of the neutron flight path is evacuated in order to minimize background due to the scattering of neutrons by air. Two 64-wire $^3$He gas multidetectors are used in order to provide a large detector solid angle and hence a high count rate. The multidetectors are moved on compressed air across a marble floor in order to cover the full range of scattering angles. With a wavelength of 0.5 Å the $Q$-range of D4 extends from 0.3 to 24 Å$^{-1}$. (A higher maximum $Q$ may be attained at $\lambda = 0.35$ Å, but the flux available at this wavelength is too low for regular use.) A development program is currently underway to replace the multidetectors with microstrip detectors, increasing the count rate by an order of magnitude.

### Pulsed source instrumentation

**Figure** 9 shows a schematic of the layout of a typical neutron diffractometer at a pulsed neutron source. The time-of-flight technique (see Eqn [3]) is used to determine the wavelength of the detected neutrons and hence a monochromator is not needed. The differential cross-section is measured as a function of $Q$ with the detector at a fixed scattering angle, $2\theta$, and $Q$ (or $d$) is scanned by varying the neutron wavelength $\lambda$ (cf. Eqn [11]). The time-of-flight technique is thus a dispersive technique and a white beam covering a wide range of wavelengths is incident on the sample. For powder diffraction, pulsed source data have the simplifying property that time-of-flight

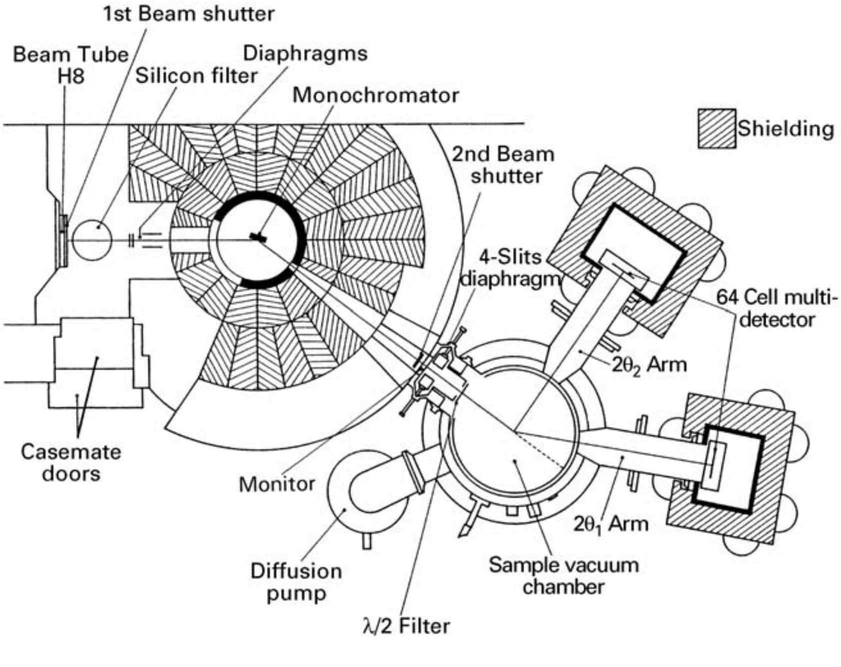



**Figure 8** The D4 neutron diffractometer at the ILL reactor. Reproduced with permission from American Institute of Physics, Clare AG, Etherington G, Wright AC, et al. (1989) A neutron-diffraction and molecular-dynamics investigation of the environment of $Dy^{3+}$ ions in a fluoroberyllate glass. *Journal of Chemical Physics* **91**: 6380–6392.

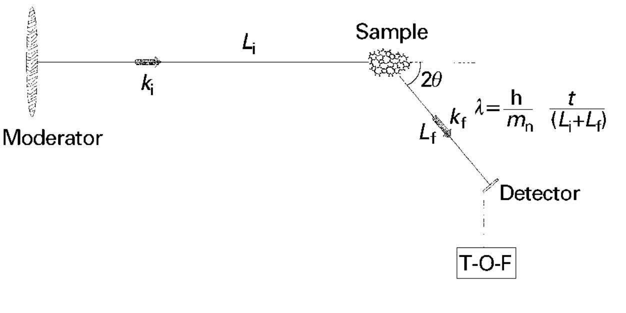



**Figure 9** Schematic of a neutron diffractometer for a pulsed (accelerator-based) source.

is proportional to $d$-spacing;

$$t = \frac{2m_n}{h} Ld \sin\theta \qquad [14]$$

(or $t = 505.64Ld \sin\theta$ in convenient units of μs, metres and Ångstroms, respectively).

A noteworthy advantage of the ability to measure a full diffraction pattern at a single fixed scattering angle is that complex sample environment equipment can be used (e.g. for high pressures) with two well-collimated flight paths for the incident and scattered beams so that background from the equipment is minimized.

In the analysis of data from a pulsed source diffractometer, it is necessary to normalize out the flux distribution $\Phi(\lambda)$ arising from the moderator. This is done by measuring the time-of-flight spectrum for a vanadium standard, as well as for the sample. Vanadium is used for this purpose because the scattering from vanadium is almost completely incoherent and the Bragg peaks are extremely small (nevertheless, it is necessary to remove the Bragg peaks from the data by some kind of smoothing process). In practice, the differential cross-section for the sample is determined by performing the following operation with the measured time-of-flight spectra:

$$\text{Normalized sample spectrum} = \left(\frac{\text{sample} - \text{empty container background}}{\text{vanadium} - \text{background}}\right) \qquad [15]$$

(A fully corrected dataset will also have been corrected for such effects as detector dead-time, attenuation and multiple scattering.) **Figure 10** illustrates the effect of flux normalization for pulsed diffraction data, using data from the LAD diffractometer described below. The vanadium time-of-flight spectrum is closely related to the flux distribution $\Phi(\lambda)$ arising from the moderator. The upturn in $\Phi(\lambda)$ at short times is due to high energy epithermal neutrons whilst the broad peak at

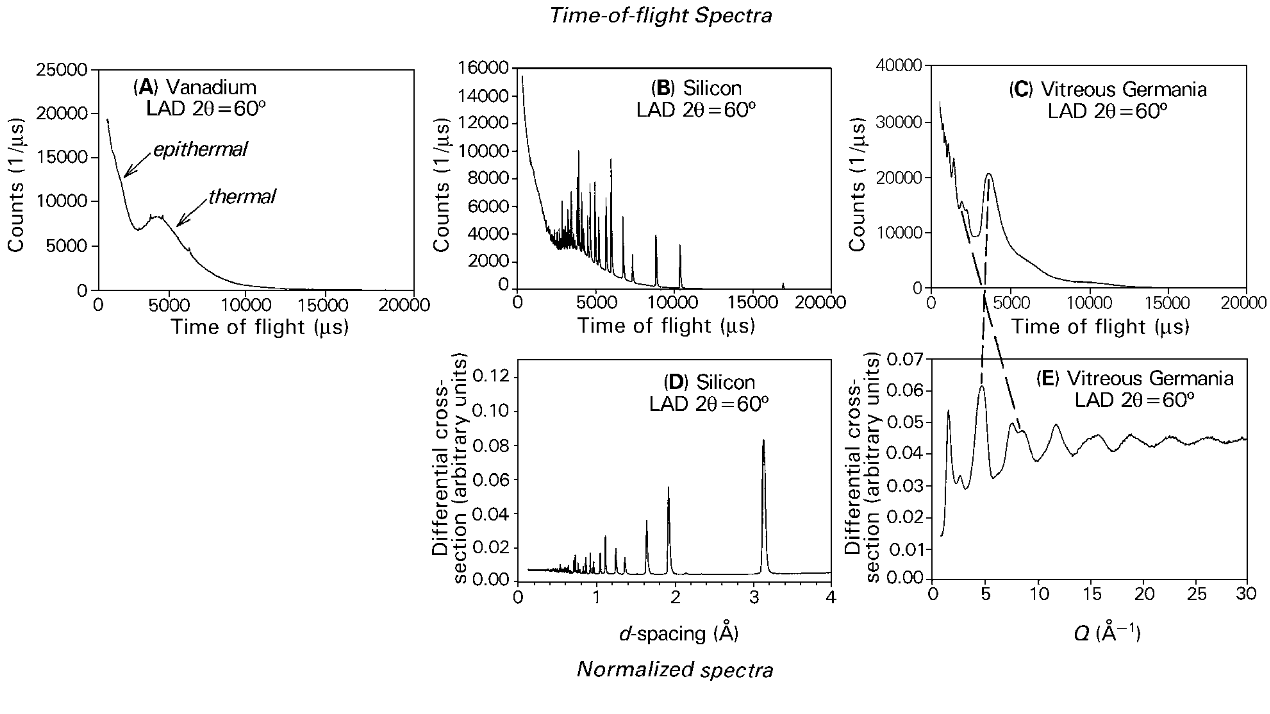



**Figure 10** Time-of-flight spectra for (A) vanadium, (B) polycrystalline silicon and (C) vitreous germania. Also shown are the normalized spectra for (D) polycrystalline silicon and (E) vitreous germania.

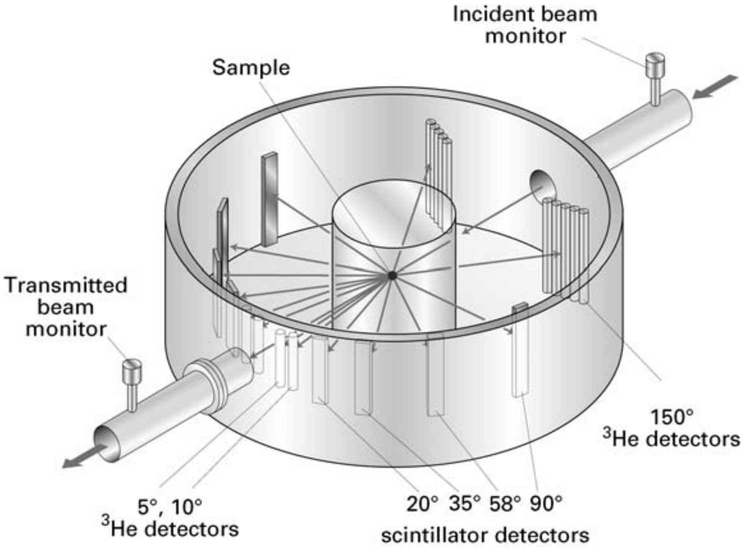



**Figure 11** The LAD diffractometer at the ISIS spallation neutron source.

intermediate times is the peak of a Maxwellian distribution whose position depends on the moderator temperature (cf. Eqn [2]). The spectra measured on a reactor diffractometer do not need to be normalized to a flux distribution, but a vanadium standard may still be used in an experiment in order to achieve an absolute normalization of the differential cross-section.

**Figure 11** shows the LAD (liquids and amorphous diffractometer) instrument at the ISIS spallation neutron source. LAD has been the main ISIS diffractometer for the study of disordered materials in the decade since the start of operations at ISIS and its design is typical for an early pulsed source diffractometer. The neutron beam comes from a liquid methane moderator at a temperature of 110 K. A cooled moderator is used in order to reduce the correction for inelasticity effects. The incident flight path, $L_i$, is 10.0 m long, leading to a moderately high reciprocal-space resolution. In practice, a pulsed source diffractometer has several different detector banks at different scattering angles so as to extend the $Q$-range of the data. LAD has detector banks at seven different scattering angles, some of which use $^3$He gas detectors, whilst the others use lithium glass scintillator detectors. The detectors are all in the horizontal plane. In order to minimize background there are large amounts of shielding between the detector banks. Thus the solid angle subtended by the relatively low detector area is small and hence the count rate of the instrument is not very high. The best resolution is obtained in the backward angle detectors ($2\theta = 150°$) with $\Delta Q/Q$ ~0.5–0.6%.

**Figure 12** shows the GEM (general materials) diffractometer at ISIS which is being constructed to replace the LAD diffractometer. This will be a state-of-the-art pulsed source diffractometer to be used for both non-crystalline diffraction and for powder crystallography. The incident flight path, $L_i$, will be relatively long at 17.0 m, giving a very small flight path contribution to the resolution. Thus the resolution in the backward angle detectors will be very good with $\Delta Q/Q$ ~0.2–0.3%. All of the detectors will be ZnS scintillators using narrow 5 mm active elements in order to minimize the angular contribution to the resolution. The detectors will cover a very large solid angle (area ~10 m$^2$, maximum azimuthal angle ~45°) so as to achieve a high effective count rate. A nimonic $t_0$ chopper at a distance 9.3 m from the moderator will be used to close off the beam at $t = 0$ and thus prevent very fast neutrons and prompt gamma rays from reaching the sample. This will prevent high energy neutrons from thermalizing in large pieces of sample environment equipment (e.g. high pressure) and then giving rise to a substantial background. In addition, two disc choppers will be used at flight paths of 6.5 m and 9.5 m to define a restricted wavelength range for the beam reaching the sample. This will be done so as to avoid frame overlap, which can be a significant problem for a diffractometer with a longer flight path. Frame overlap occurs when slower neutrons from a pulse of the source are overtaken by faster neutrons from the subsequent pulse. If the flux of the slower neutrons is significant, then a diffraction peak, which in reality is detected at the long time-of-flight, $t$, appears to be detected at the earlier time, $t - \tau_0$ (where $\tau_0$ is the period of the source), and this leads to spurious peaks in the data.

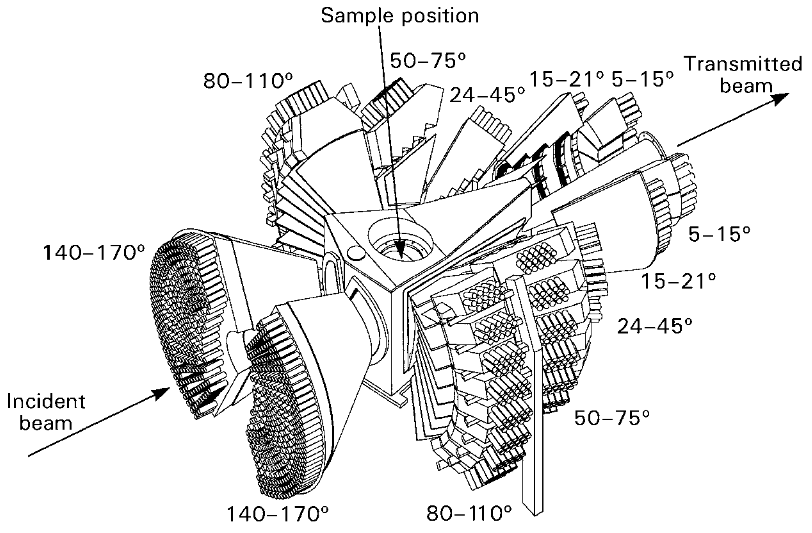



**Figure 12** The new GEM diffractometer at the ISIS spallation neutron source.

## Reciprocal-space resolution

A certain point in a diffraction pattern picks out the scattering at some momentum transfer $Q$. However, this $Q$ is never completely defined due to uncertainties in the various parameters $\alpha_i$ of the diffractometer used to determine $Q$, leading to an overall uncertainty $\Delta Q$. In practice the parameters $\alpha_i$ may be highly dependent, but for a simple discussion of $Q$-resolution it may be assumed that the parameters all act independently so that

$$\Delta Q = \left[\sum_i \left(\frac{\partial Q}{\partial \alpha_i}\Delta\alpha_i\right)^2\right]^{1/2} \qquad [16]$$

The parameters which are important in determining the resolution will depend on the details of the particular diffractometer being used and thus only an outline of some of the general principles and dependencies for diffraction from an isotropic sample can be given here.

### Resolution of a reactor diffractometer

For a reactor diffractometer the resolution width, $\Delta Q$, arises because the acceptance angle of the collimation around the monochromator, the size of the sample and the acceptance angle of the detector system are all greater than zero and also because the monochromator crystal has a mosaic spread. For a treatment of the $Q$-resolution in terms of independent parameters, these factors may be considered to lead to a geometrical uncertainty in angle, $\Delta\theta$, and an uncertainty in wavelength, $\Delta\lambda$. Differentiating the definition of $Q$ (Eqn [11]) thus gives

$$\frac{\Delta Q}{Q} = \frac{\Delta d}{d} = \left[(\cot\theta\Delta\theta)^2 + \left(\frac{\Delta\lambda}{\lambda}\right)^2\right]^{1/2} \qquad [17]$$

where the semi-angle, $\theta$, is in radians, not degrees.

**Figure 13** shows the experimental $Q$-resolution of the D4 reactor diffractometer (measured at $\lambda = 0.7$ Å for powder samples of YIG, Si, Ni and Fe with diameter 7 mm and height 45 mm, and with the detector at 1.455 m). For the range of angle shown, $Q\cot\theta$ changes little and hence the $Q$-resolution may be expressed approximately as

$$\Delta Q = (C_\theta + C_\lambda Q^2)^{1/2} \qquad [18]$$

where $C_\theta$ and $C_\lambda$ are constants. This simple independent-parameter expression predicts the general trend of **Figure 13**. However, the detailed behaviour, and in particular the upturn in $\Delta Q$ at low $Q$, is not predicted, and a full analysis of the $Q$-resolution must treat the parameters as being dependent.

The resolution of a reactor diffractometer may be improved by reducing the uncertainties in angle and wavelength. However, this can only be achieved at the expense of a corresponding decrease in the count rate of the diffractometer. For non-crystalline samples the $Q$-resolution is of less importance than for polycrystalline samples because the diffraction peaks are relatively broad (cf. **Figure 10**). The D4 diffractometer is used mainly to study non-crystalline samples and hence it has been

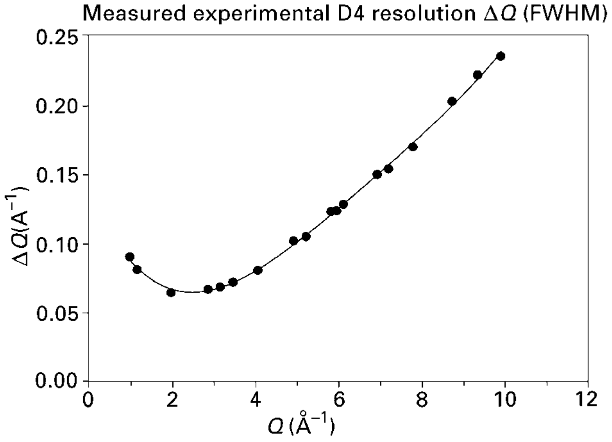



**Figure 13** The measured experimental resolution (full width at half maximum) of the D4 reactor diffractometer. Courtesy of HE Fischer.

optimized primarily to have a high count rate. Thus the $Q$-resolution of D4 is not the best that can be achieved at a reactor neutron source.

### Resolution of a pulsed source diffractometer

For a time-of-flight diffractometer the three major contributions to the $Q$-resolution are due to the geometrical uncertainty in angle, $\Delta\theta$, the uncertainty in flight path, $\Delta L$, and the uncertainty in time-of-flight, $\Delta t$, leading to the following expression for the total combined resolution width:

$$\frac{\Delta Q}{Q} = \frac{\Delta d}{d} = \left[(\cot\theta\Delta\theta)^2 + \left(\frac{\Delta L}{L}\right)^2 + \left(\frac{\Delta t}{t}\right)^2\right]^{1/2} \quad [19]$$

The angular uncertainty for a time-of-flight diffractometer, $\Delta\theta$, is similar to the reactor case in that it arises from the acceptance angle of the collimation in front of the moderator, the size of the sample and the acceptance angle of the detector. However, there is a major difference in its effect because of the way in which a single detector measures a complete diffraction pattern without moving to different scattering angles, $2\theta$, so that the angular contribution to $\Delta Q/Q$ is independent of $Q$. This contribution to the resolution is ideally suited to powder diffraction because it has a $\Delta d$ which is small for short $d$-spacing where Bragg peaks are very close together, and only becomes large at high $d$-spacing where Bragg peaks are well separated.

The flight path uncertainty, $\Delta L$, arises from the finite sizes of the moderator, sample and detector. Its contribution to the resolution is symmetric and may be approximated by a Gaussian, as is also the case for the angular contribution. The contribution to $\Delta d/d$ due to the flight path uncertainty may be minimized simply by using a very long flight path, $L$, although this may only be achieved at the expense of a corresponding decrease in the count rate.

At low and medium scattering angles, cot $\theta$ is relatively large and the angular contribution to the resolution tends to dominate. However, at backward angles ($2\theta \rightarrow 180°$) this contribution becomes very small so that a diffraction pattern can be measured with a very narrow angular resolution across its entire range. The time-of-flight uncertainty, $\Delta t$, is then of greater importance. Its main component is the time uncertainty which arises from the moderation process. This gives an exponential contribution to the resolution with a decay time, $\tau$, which is the mean time spent in the moderator for neutrons of a particular wavelength. **Figure 14** shows the resolution of the LAD time-of-flight diffractometer at backward

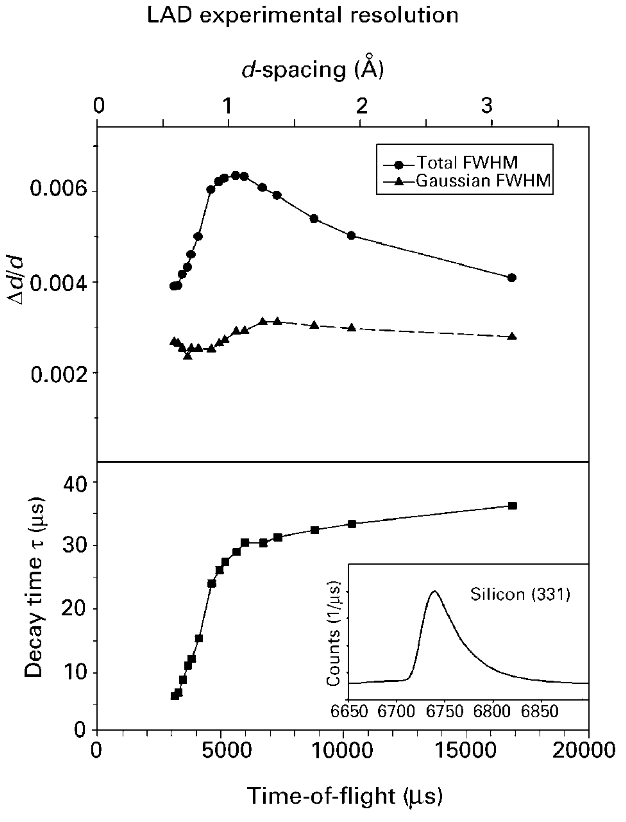



**Figure 14** The measured experimental resolution of the LAD pulsed neutron diffractometer at backward angle, $2\theta = 146°$. The total full width at half maximum (FWHM) and the FWHM for the Gaussian contribution are shown in the form of $\Delta d/d$. Also shown is the decay constant $\tau$ for the exponential contribution to the resolution. The inset shows the peak shape measured for a typical Bragg peak (the (331) reflection for polycrystalline silicon).

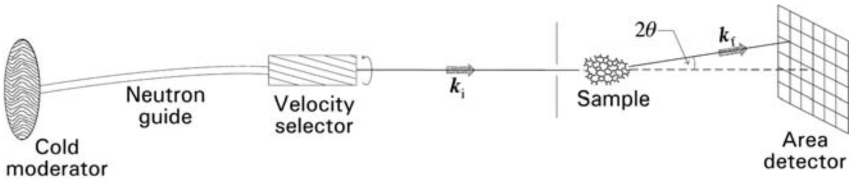



**Figure 15** Schematic of a small angle neutron scattering (SANS) diffractometer for a steady-state (reactor) source.

angle, obtained by fitting the convolution of a Gaussian and an exponential to Bragg peaks in experimental data. The Gaussian contribution to $\Delta d/d$ is almost independent of $d$-spacing, as predicted. The detailed behaviour of the decay time, $\tau$, depends upon the design of the moderator, but is generally small at short times (i.e. for high energy neutrons) and tends to a large value at long times (i.e. for low energy neutrons). The overall peak shape, Gaussian-exponential convolution (see inset to **Figure 14**), has a very sharp leading edge which can be advantageous in resolving overlapping peaks.

## Some other types of neutron diffraction instrumentation

In this article the greatest emphasis is given to wide angle neutron diffraction from isotropic samples. However, a range of other neutron diffraction techniques is also available and it is only possible to describe some of them briefly here. (One of the major advantages of neutrons is that polarized beams can be used and these are especially useful for the study of magnetic structures. However, a discussion of polarized neutron diffraction instrumentation is beyond the scope of this article.)

### Small angle neutron scattering at a reactor

A small angle neutron scattering (SANS) diffractometer is used to study structures whose dimensions range from 10 to 1000 Å (e.g. defects, voids, micelles, macromolecules, etc.) and as such their $Q$-ranges are in the region of 0.005 to 0.5 $Å^{-1}$. In order to access these low values of $Q$ the detector should be at low scattering angle, $2\theta$, and a cold moderator should be used so as to provide long wavelength neutrons (cf. Eqn [11]).

**Figure 15** shows a schematic of the layout of a typical small angle neutron scattering (SANS) diffractometer at a reactor source. A large two-dimensional multiwire area detector is placed at a low angle so as to cover as large a solid angle as possible and thus maximize the count rate. The detector is either placed symmetrically around the transmitted beam (as shown in **Figure 15**) to maximize the count rate for the lowest $Q$, or is placed off-axis to extend the $Q$-range upwards. The neutron beam is monochromated by a velocity selector which consists of a rotating cylinder with a set of helical grooves cut into its outer surface. The angular contribution to the resolution is unavoidably large at low angle because $\cot\theta$ becomes large (cf. Eqn [17]). Hence the monochromation produced by the velocity selector is relatively coarse ($\Delta\lambda/\lambda$~5–20%) so as to match the angular contribution and provide as large an incident flux as possible. The D11 diffractometer at the ILL a prime example of a reactor SANS instrument.

### Single crystal diffraction at a plused source

In an experiment on a single crystal sample the differential cross-section must be measured as a function of the magnitude and the direction of the momentum transfer vector, $\boldsymbol{Q}$. Bragg peaks are then observed in the differential cross-section whenever $\boldsymbol{Q}$ satisfies

$$\boldsymbol{Q} = \boldsymbol{\tau} \qquad [20]$$

where $\boldsymbol{\tau}$ is a reciprocal lattice vector of the crystal.

**Figure 16** shows a schematic of the layout of a typical pulsed source single crystal diffractometer. A goniometer is used to provide full control of the orientation of the sample in space so that any region of reciprocal-space of interest can be studied. An area

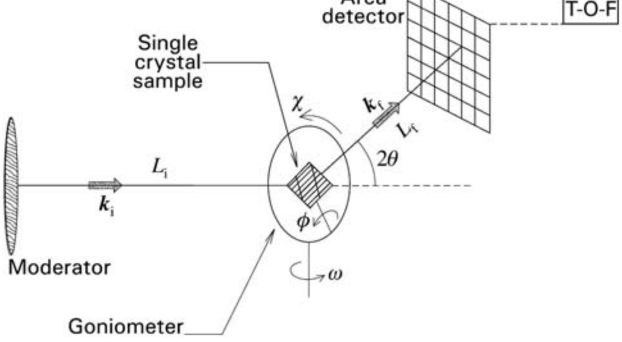



**Figure 16** Schematic of a single crystal diffractometer on a pulsed (accelerator-based) source.

detector combined with the use of multiwavelength data provides a three-dimensional sampling of reciprocal-space that may contain hundreds of Bragg reflections. This wide sampling for a pulsed source diffractometer is advantageous because it may reveal unexpected Bragg reflections or satellites and can show diffuse scattering between the Bragg peaks. The SXD diffractometer at ISIS is a prime example of a pulsed source single crystal diffractometer.

## List of symbols

$C_x$ = a constant; $d$ = $d$-spacing; $d_{hkl}$ = interlayer spacing between crystal planes; $(d\sigma/d\Omega)_{tot}$ = differential cross-section for total neutron diffraction; $E$ = neutron energy; $E_f$ = neutron energy after scattering (final); $E_i$ = neutron energy before scattering (initial); $h$ = Planck's constant; $\hbar$ = Planck's constant divided by $2\pi$; $k_B$ = Boltzmann's constant; $\boldsymbol{k}_f$ = neutron wavevector after scattering (final); $\boldsymbol{k}_i$ = neutron wavevector before scattering (initial); $L$ = total flight path of neutron; $L_f$ = flight path of neutron after scattering (final); $L_i$ = flight path of neutron before scattering (initial); $m_n$ = neutron mass; $N$ = number of atoms in sample; $n_v$ = atomic number density; $Q$ = magnitude of momentum transfer; $\boldsymbol{Q}$ = momentum transfer vector; $R_{tot}$ = rate at which neutrons are scattered into the specified solid angle; $t$ = time-of-flight of neutron; $T$ = temperature; $v$ = neutron speed; $x$ = detector thickness; $\alpha_i$ = diffractometer parameter; $\Delta d$ = resolution uncertainty for $d$-spacing, $d$; $\Delta Q$ = resolution uncertainty for momentum transfer, $Q$; $\Delta\theta$ = resolution uncertainty for semi-scattering angle, $\theta$; $\Delta\lambda$ = resolution uncertainty for neutron wavelength, $\lambda$; $\varepsilon$ = output pulse fraction for detector; $2\theta$ = scattering angle; $\lambda$ = neutron wavelength; $\sigma_{abs}(\lambda)$ = neutron absorption cross-section at neutron wavelength, $\lambda$; $\sigma^0_{abs}$ = neutron absorption cross-section at neutron wavelength of 1 Ångstrom; $\boldsymbol{\tau}$ = reciprocal lattice vector of crystal; $\tau$ = exponential decay time; $\tau_0$ = period of the pulsed source; $\phi$ = polar coordinate; $\Phi$ = neutron flux; $\Omega$ = solid angle.

*See also:* **Inelastic Neutron Scattering, Applications; Inelastic Neutron Scattering, Instrumentation; Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; Materials Science Applications of X-Ray Diffraction; Neutron Diffraction, Theory; Powder X-Ray Diffraction, Applications; Structure Refinement (Solid State Diffraction).**

## Further reading

Bacon GE (1969) *Neutron Physics*. London: Wykeham.

Bacon GE (1975) *Neutron Diffraction*, 3rd edn. Oxford: Oxford University Press.

Brown PJ and Forsyth JB (1973) *The Crystal Structure of Solids*. London: Edward Arnold.

Carpenter JM and Yelon WB (1986) Neutron Sources. In: (eds) Sköld K and Price DL. *Neutron Scattering, Part A*, pp 99–196. Orlando: Academic Press.

Egelstaff PA (ed.) (1965) *Thermal Neutron Scattering*. London: Academic Press.

Newport RJ, Rainford BD and Cywinski R (eds) (1988) *Neutron Scattering at a Pulsed Source*. Bristol: Adam Hilger.

Stirling GC (1973) Experimental techniques. In: Willis BTM (ed.) *Chemical Applications of Thermal Neutron Scattering*, pp 31–48. London: Oxford University Press.

Williams WG (1988) *Polarized Neutrons*. Oxford: Clarendon Press.

Willis BTM (1970) *Thermal Neutron Diffraction*. Oxford: Clarendon Press.

Windsor CG (1981) *Pulsed Neutron Scattering*. London: Taylor and Francis.

Windsor CG (1986) Experimental techniques. In: Sköld K and Price DL (eds) *Neutron Scattering, Part A*, pp 197–257. Orlando: Academic Press.

# Neutron Diffraction, Theory

**Alex C Hannon**, Rutherford Appleton Laboratory, Didcot, UK

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| Theory |

The cross-section measured in a neutron scattering experiment on a non-magnetic material depends in general upon a correlation function involving the positions in time and space of the atomic nuclei in the sample. For a total neutron diffraction experiment the diffraction pattern depends upon a correlation function involving the instantaneous interatomic vectors between the nuclei. However, for an elastic diffraction experiment the diffraction pattern depends upon a correlation function involving the time-averaged interatomic vectors between the positions of the nuclei. The correlation function is obtained by performing a suitable Fourier transform of the diffraction pattern and its features indicate which interatomic distances occur more or less frequently in the sample. The neutron correlation function is thus a powerful tool for the study of the atomic structure of glasses, liquids and disordered crystalline materials. For a crystalline sample the symmetry of the atomic structure gives rise to sharp Bragg peaks in the diffraction pattern involving only elastically scattered neutrons. The positions of the Bragg peaks and their systematic absences depend upon the symmetry of the crystal lattice. Meanwhile the intensities of the Bragg peaks depend upon the positions of the atoms in the unit cell. A single crystal neutron diffraction experiment yields both the symmetry of the lattice and the contents of the unit cell, whilst a powder neutron diffraction experiment is usually only used to reveal the unit cell contents for a crystal whose symmetry is already known.

## Fundamental theory

### The interaction between a neutron beam and a sample

If a sample is placed in a beam of neutrons then some of the neutrons will be removed from the transmitted beam due to scattering. There are two interactions which give rise to this scattering. The first of these is the nuclear force between a neutron and the nuclei of the sample. The second interaction is that between the magnetic moment of the neutron and the unpaired electrons of magnetic ions in the sample. The treatment of magnetic scattering (and polarized neutron scattering) is beyond the scope of this article. (There are two key differences between magnetic and nuclear scattering: magnetic scattering involves a form factor, and it also has a directional dependence (on magnetic moment) which does not arise for nuclear scattering). The basic properties of the neutron and their consequences for diffraction are summarized in **Table 1**.

**Table 1** The properties of the neutron–sample interaction and their consequences for diffraction

| *Property* | *Consequences* |
|---|---|
| No form factor for scattering | Data can be measured to high momentum transfer, leading to high real-space resolution |
| | Absolute normalization of results is relatively straightforward |
| Scattering length is different for different isotopes | Isotopic substitution can be used to extract element-specific information |
| Scattering length varies haphazardly across the periodic table | The positions of light atoms (especially hydrogen) can be determined in the presence of heavy atoms |
| | The positions of atoms of elements which are close in the periodic table can be distinguished |
| Interaction is relatively weak | Neutrons are highly penetrating and hence results are characteristic of the bulk of a sample, not the surface |
| | Bulky equipment can be used to subject the sample to a wide range of environments |
| Magnetic interaction between neutron and magnetic ions | Magnetic structures can be determined |

## Neutron cross sections

Consider a monochromatic beam of neutrons of energy $E_i$, wave vector $\boldsymbol{k}_i$ and flux $\Phi$ which is incident on a sample of $N$ atoms. The scattering of the neutrons may then be defined experimentally by the double differential cross section:

$$\frac{d^2\sigma}{d\Omega\, dE} = \frac{R_{inel}}{N\Phi\, d\Omega\, dE} \qquad [1]$$

where $R_{inel}$ is the number of neutrons scattered per unit time into the solid angle $d\Omega$ in the direction $(2\theta, \phi)$ (see **Figure 1**) with final energy $E_f$ in the range $E_i$–$E$ to $E_i$–$(E + dE)$. The energy $E$ $(= E_i – E_f)$ is that transferred by the neutron to the sample in a given scattering event. The double differential cross-section is the fundamental quantity for the scattering of neutrons and the cross-section for a given experimental arrangement (e.g. diffraction or transmission) is obtained from it by performing suitable integration(s).

In a total diffraction experiment all scattered neutrons are detected regardless of their final energy. In this case the relevant cross-section is the differential cross-section defined by

$$\left(\frac{d\sigma}{d\Omega}\right)_{tot} = \frac{R_{tot}}{N\Phi\, d\Omega} = \int_{-\infty}^{\infty} \left(\frac{d^2\sigma}{d\Omega\, dE}\right) dE \qquad [2]$$

where $R_{tot}$ is the number of neutrons scattered per unit time into the solid angle $d\Omega$ in the direction $(2\theta, \phi)$. This is to be contrasted with the case of elastic diffraction where only neutrons whose energy is unchanged (i.e. they are scattered elastically and have $E = 0$) are detected. In this case the relevant cross-section is

$$\left(\frac{d\sigma}{d\Omega}\right)_{el} = \frac{R_{el}}{N\Phi\, d\Omega} = \left.\frac{d^2\sigma}{d\Omega\, dE}\right|_{E=0} \qquad [3]$$

where $R_{el}$ is the number of neutrons scattered elastically per unit time into the solid angle $d\Omega$ in the direction $(2\theta, \phi)$. The elastic differential cross-

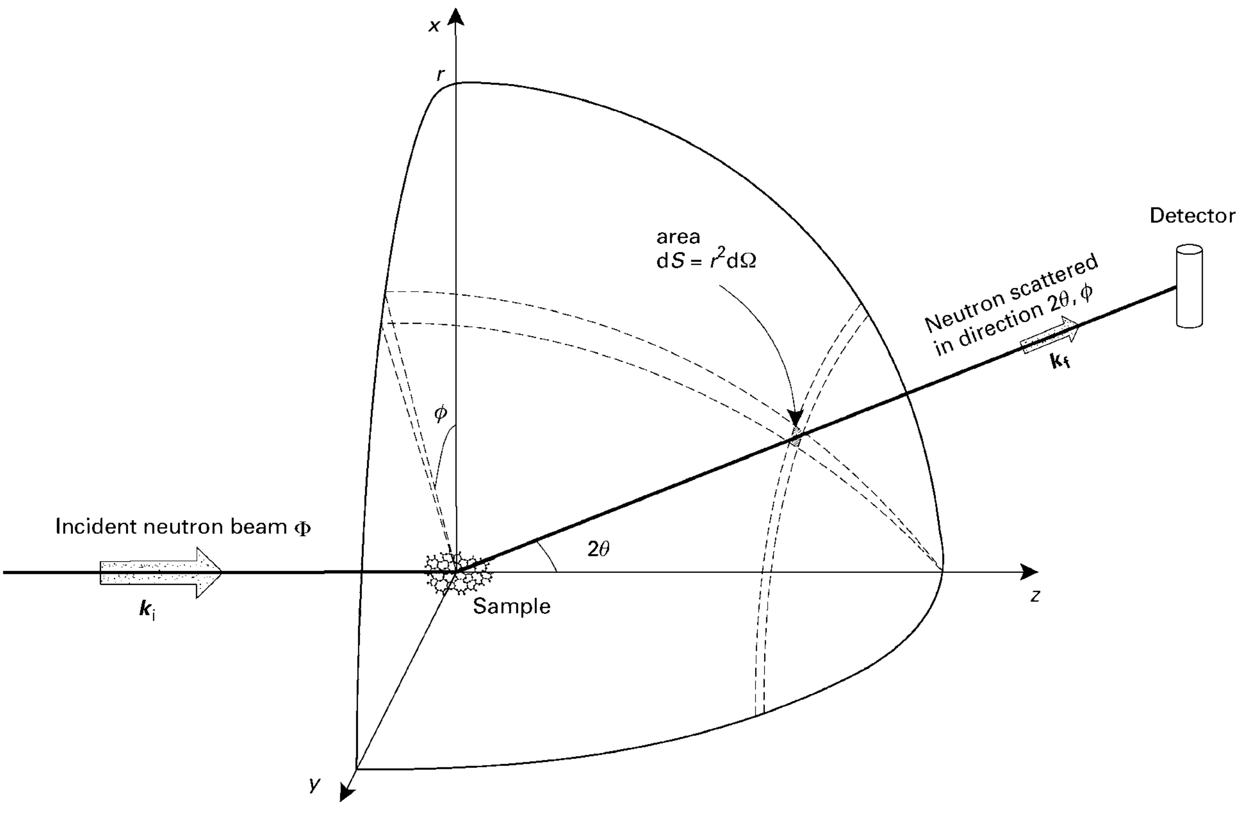



**Figure 1** Geometry for a neutron scattering experiment.

section is given by evaluating the double differential cross-section at $E = 0$.

### Neutron scattering length

To perform a neutron diffraction experiment, the neutron wavelength needs to be of similar magnitude to interatomic distances, say of the order of 1 Ångström (1 Å ≡ $10^{-10}$ m – Ångström units are almost universally used in the field of diffraction). This corresponds to a neutron energy of 25 meV, and thus the energies of neutrons used for diffraction are very much lower than typical X-ray energies used in diffraction. Nuclear forces are very short range (~$10^{-15}$ m), operating over much shorter distances than neutron wavelengths. Hence nuclei can be treated as point-like scattering centres which give rise to an isotropic scattered neutron wave (i.e. only s-wave scattering need be considered). The wavefunction of a neutron scattered by a single nucleus at the origin is thus expressed as

$$\Psi_{\mathrm{f}} = -\frac{b}{r}\exp(\mathrm{i}kr) \qquad [4]$$

where $k$ is the magnitude of the wave vector for the scattered wave and $r$ is a polar coordinate. $b$ is a constant, known as the (bound atom) scattering length, which has units of length. A positive scattering length corresponds to a phase change of $\pi$ between the scattered and incident waves. Most scattering lengths are positive, but there is a minority of cases where the scattering length is negative. The neutron scattering length corresponds to the form factor in X-ray diffraction, but with the important difference that it is a simple constant without any wave vector dependence.

### Fermi pseudo potential and the general expression for cross-section

The interaction between a neutron and a sample may be represented by the Fermi pseudo potential:

$$V(\boldsymbol{r}) = \frac{2\pi\hbar^2}{m_{\mathrm{n}}}\sum_{j=1}^{N} b_j\delta(\boldsymbol{r} - \boldsymbol{R}_j) \qquad [5]$$

where $m_{\mathrm{n}}$ is the neutron mass, and the summation is taken over the $N$ nuclei whose position vectors and scattering lengths are $\boldsymbol{R}_j$ and $b_j$ respectively. Use of the Fermi pseudo potential together with Fermi's golden rule (sometimes termed the Born approximation) eventually results in the following general expression from which all results for the nuclear scattering of neutrons may be derived:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}E} = \frac{1}{N}\frac{k_{\mathrm{f}}}{k_{\mathrm{i}}}\frac{1}{2\pi\hbar}\sum_{j,j'} b_j b_{j'}\int_{-\infty}^{\infty}\left\langle \exp\left\{-\mathrm{i}\boldsymbol{Q}\,.\left[\boldsymbol{R}_j(0) - \boldsymbol{R}_{j'}(t)\right]\right\}\right\rangle \exp(-\mathrm{i}Et/\hbar)\mathrm{d}t \qquad [6]$$

The summations $j$ and $j'$ are both taken over the $N$ nuclei of the sample and the angular brackets denote a thermal average at the temperature $T$ of the sample. $\boldsymbol{R}_j(t)$ is the position of the $j$th nucleus at time $t$ and $\boldsymbol{Q}$ is the scattering vector, defined by

$$\boldsymbol{Q} = \boldsymbol{k}_{\mathrm{i}} - \boldsymbol{k}_{\mathrm{f}} \qquad [7]$$

where $\boldsymbol{k}_{\mathrm{i}}$ and $\boldsymbol{k}_{\mathrm{f}}$ are the incident and scattered neutron wave vectors respectively. The term momentum transfer (i.e. the momentum transferred to the sample) is commonly used for $\boldsymbol{Q}$, although strictly speaking this term should be used for $\hbar\boldsymbol{Q}$.

## Total diffraction

### The static approximation

In the static approximation it is assumed that the incident neutron energy $E_{\mathrm{i}}$ is large compared with the excitation energies of the sample. This means that $k_{\mathrm{f}} \approx k_{\mathrm{i}}$ so that the scattering vector may be taken to have its elastic value:

$$Q = |\boldsymbol{Q}| = 2k\sin\theta = \frac{4\pi\sin\theta}{\lambda} \qquad [8]$$

where $\lambda$ is the wavelength of the neutrons. In the static approximation Equation [6] may be integrated according to Equation [2] to obtain Equation [9]. (In practice the assumptions inherent in the static approximation do not hold exactly and as a consequence an inelasticity correction must be made to experimental total diffraction data).

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{tot}} = I(\boldsymbol{Q}) = \frac{1}{N}\sum_{j,j'} b_j b_{j'} \times \left\langle \exp\left\{-\mathrm{i}\boldsymbol{Q}\,.\left[\boldsymbol{R}_j(0) - \boldsymbol{R}_{j'}(0)\right]\right\}\right\rangle \qquad [9]$$

Thus the total diffraction pattern $I(Q)$ depends upon the vectors $R_j - R_{j'}$ between the atoms, with a weighting according to scattering length. This is the basis for the use of total diffraction to measure the atomic structure of a sample. Total diffraction depends upon the positions of the atoms in the sample at an arbitrary time zero. Thus total diffraction yields an instantaneous 'snapshot' of the interatomic vectors.

### Coherent and incoherent scattering

For a consideration of coherent and incoherent scattering it is convenient to write Equation [9] in the form

$$I(\boldsymbol{Q}) = \sum_{j,j'} b_j b_{j'} \langle j, j' \rangle \qquad [10]$$

The value of $b_j$ is not the same for all the nuclei of a single element, owing to spin and isotopic incoherence, as will be explained below. Hence, to obtain a useful result, Equation [10] is averaged over all possible distributions of scattering length, making the assumption that there is no correlation between the values of $b_j$ for any two nuclei. Now the average value of $b_j b_{j'}$ differs, depending on whether $j = j'$ is satisfied;

$$\begin{aligned} \overline{b_j b_{j'}} &= \left(\bar{b}\right)^2, \; j \neq j' \\ \overline{b_j b_{j'}} &= \overline{b^2}, \quad j = j' \end{aligned} \qquad [11]$$

where $\bar{b}$ is the average scattering length (usually known as the coherent scattering length) for all nuclei of a particular element, whilst $\overline{b^2}$ is the average of the squared scattering length for the relevant element. The double summation of Equation [10] may thus be separated into $j \neq j'$ (distinct) terms and $j = j'$ (self) terms:

$$I(\boldsymbol{Q}) = i(\boldsymbol{Q}) + \sum_l c_l \overline{b_l^2} \qquad [12]$$

where $c_l = N_l$, $N$ is the atomic fraction for element $l$ and the distinct differential cross-section is

$$i(\boldsymbol{Q}) = \sum_{l,l'} \bar{b}_l \bar{b}_{l'} \sum_{\substack{j=1 \\ j \neq j'}}^{N_l} \sum_{j'=1}^{N_{l'}} \frac{1}{N} \times \left\langle \exp\left\{ -\mathrm{i}\boldsymbol{Q} . \left[ \boldsymbol{R}_j(0) - \boldsymbol{R}_{j'}(0) \right] \right\rangle \qquad [13]$$

The $l$ and $l'$ summations are over the elements in the sample (e.g. for $Li_2Si_2O_5$, $l =$ Li, Si, O). The $j$ (or $j'$) summations are then over all the $N_l$ (or $N_l'$) atoms of element $l$ (or $l'$), excluding terms where $j$ and $j'$ refer to the same atom. By the algebraic trick of adding and subtracting a $\Sigma_j \bar{b}_j^2 \langle j, j \rangle$ term the differential cross-section for total diffraction may be separated into its coherent and incoherent parts:

$$I(\boldsymbol{Q}) = \left\langle \bar{b}^2 \right\rangle_{\mathrm{av}} S(\boldsymbol{Q}) + \frac{\sigma_{\mathrm{inc}}}{4\pi} \qquad [14]$$

where the structure factor is given by

$$S(\boldsymbol{Q}) = \frac{1}{\left\langle \bar{b}^2 \right\rangle_{\mathrm{av}}} \sum_{l,l'} \bar{b}_l \bar{b}_{l'} \sum_{j,j'} \frac{1}{N} \left\langle \exp\left\{ -\mathrm{i}\boldsymbol{Q} . \left[ \boldsymbol{R}_j(0) - \boldsymbol{R}_{j'}(0) \right] \right\} \right\rangle = 1 + i(\boldsymbol{Q}) \Big/ \left\langle \bar{b}^2 \right\rangle_{\mathrm{av}} \qquad [15]$$

in which self terms ($j = j'$) are now included in the summation. The incoherent cross-section of the sample is

$$\sigma_{\mathrm{inc}} = 4\pi \left( \left\langle \overline{b^2} \right\rangle_{\mathrm{av}} - \left\langle \bar{b}^2 \right\rangle_{\mathrm{av}} \right) \qquad [16]$$

in which average values for the sample are defined as

$$\begin{aligned} \left\langle \bar{b}^2 \right\rangle_{\mathrm{av}} &= \sum_l c_l \bar{b}_l^2 \\ \left\langle \overline{b^2} \right\rangle_{\mathrm{av}} &= \sum_l c_l \overline{b_l^2} \end{aligned} \qquad [17]$$

($4\pi \langle \overline{b^2} \rangle_{\mathrm{av}}$ is the coherent cross-section for the sample). The interpretation of the coherent differential cross-section $\langle \overline{b^2} \rangle_{\mathrm{av}} S(Q)$ is that this is what would be measured from a sample for which all nuclei of element $l$ had a scattering length of $\bar{b}_l$. **Figure 2** shows the total diffraction pattern $I(Q)$ for a glass (for an isotropic sample, such as a glass, the direction of the $Q$ vector is not important). It is the coherent contribution to the differential cross-section that contains interference information relating to the positions of atoms in the sample. Also shown in **Figure 2** are the self and incoherent contributions – these are featureless and contain no interference information. For most elements the incoherent cross-section is

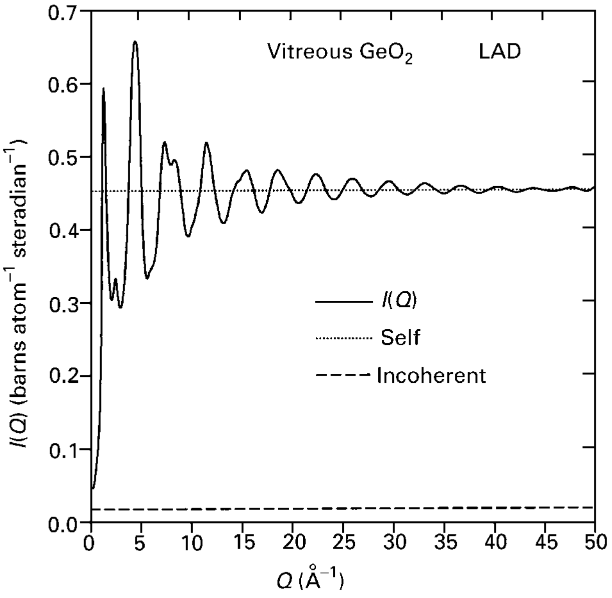



**Figure 2** The total diffraction pattern, $I(Q)$, for $GeO_2$ glass, shown together with the self and incoherent contributions.

relatively small, but a notable exception to this is hydrogen for which the incoherent cross-section is very large, so that the measured experimental diffraction pattern is dominated by the incoherent contribution.

### The sources of incoherence

The scattering length $b_j$ for an individual nucleus (i.e. the amplitude of the neutron wave which it scatters) is not the same for all nuclei of a particular element due to two factors. These are spin incoherence and isotopic incoherence.

Spin incoherence is due to the fact that a neutron and a nucleus of spin $I$ can form two different compound nuclei of spin $I \pm \frac{1}{2}$; the amplitude of the neutron wave scattered by the nucleus, and thus the scattering length, is generally different for the two different compound nuclei. Isotopic incoherence arises as a result of the presence of more than one isotope of a particular element.

## Total diffraction from a disordered sample

The total diffraction pattern measured in reciprocal-space may be interpreted in terms of a correlation function in real-space, as is discussed below. Although this approach to neutron diffraction was developed primarily for non-crystalline systems (liquids and glasses), it is increasingly being used for powder diffraction from disordered crystalline systems as well. For a disordered crystal, the correlation function method has the advantage that it reveals the local structure, rather than the structure averaged over all unit cells that is measured by Bragg diffraction (see below).

For an isotropic sample, the diffraction pattern depends only on the magnitude $Q$ of the momentum transfer $\mathbf{Q}$, and not its direction. In this case the distinct scattering, $i(Q)$ (Eqn [12]), is related to a neutron correlation function, $T(r)$, by a Fourier transform:

$$T(r) = T^0(r) + \frac{2}{\pi}\int_0^\infty Qi(Q)\sin(rQ)\,\mathrm{d}Q \qquad [18]$$

$T^0(r)$ is the average density contribution to the correlation function, given by

$$T^0(r) = 4\pi r g^0 \left(\sum_l c_l \bar{b}_l\right)^2 \qquad [19]$$

where $g^0$ is the average atomic number density in the sample, $c_l$ is the atomic fraction for element $l$, and the $l$ summation is over elements in the sample (e.g. for $Li_2Mo_2O_7$, $l$ = Li, Mo, O). For example, **Figure 3** shows the neutron correlation function, $T(r)$, for $GeO_2$ glass, obtained by Fourier transformation of the diffraction data shown in **Figure 2**. Also shown is the average density contribution $T^0(r)$.

The total correlation function is a weighted sum of partial correlation functions:

$$T(r) = \sum_l \sum_{l'} c_l \bar{b}_l \bar{b}_{l'} t_{ll'}(r) \qquad [20]$$

Each partial correlation $t_{ll'}(r)$ is related to a generalized van Hove distinct correlation function by

$$t_{ll'}(r) = 4\pi r G^{\mathrm{D}}_{ll'}(\boldsymbol{r}, 0) \qquad [21]$$

where

$$G^{\mathrm{D}}_{ll'}(\boldsymbol{r}, t) = \frac{1}{N_l}\sum_{\substack{j=1 \\ j\neq j'}}^{N_l}\sum_{j'=1}^{N_{l'}}\int\langle\delta(\boldsymbol{r}' - \boldsymbol{R}_j(0)) \times \delta(\boldsymbol{r}' + \boldsymbol{r} - \boldsymbol{R}_{j'}(t))\rangle \mathrm{d}\boldsymbol{r}' \qquad [22]$$

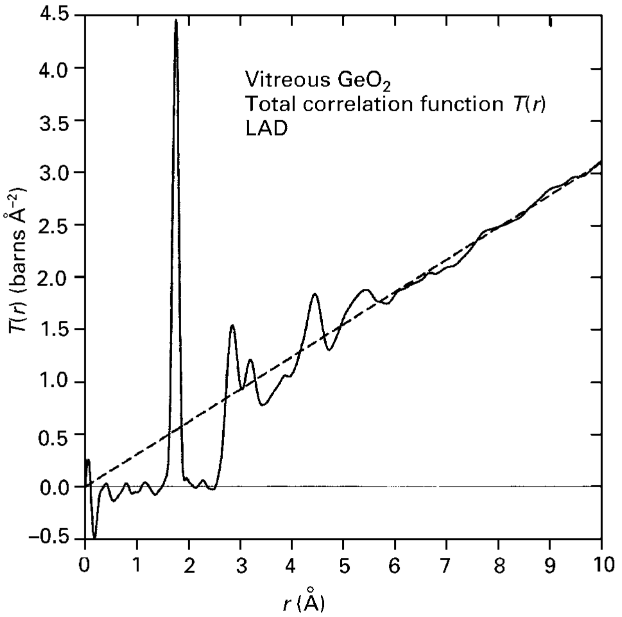



**Figure 3** The neutron correlation function, $T(r)$, for $GeO_2$ glass, obtained from the data in **Figure 2**. The dashed line indicates the average density contribution, $T^0(r)$.

Although the formal definition of the partial correlation function, $t_{ll'}(r)$, may appear complicated, its interpretation is simple: $rt_{ll'}(r)\mathrm{d}r$ is the average number of atoms of element $l'$ which are located in a spherical shell of radius $r$ to $r + \mathrm{d}r$, centred on an atom of element $l$. For example, **Figure 4** shows how the first two peaks in the correlation function, $T(r)$, arise for an $A_2X_3$ material with a network structure. For the experimental data for $GeO_2$ glass shown in **Figure 3**, the first peak in the correlation function arises from the nearest neighbour Ge–O bonds, whilst the second peak arises from the non-bonded O–O distance in the $GeO_4$ tetrahedra which connect together to form the random network structure.

In the harmonic approximation the contribution to $t_{ll'}(r)$ due to a single interatomic distance $r_{jk}$ with a root mean square thermal variation in distance of $\langle u_{jk}^2\rangle^{1/2}$ is

$$t_{jk}(r) = \frac{n_{jk}}{r_{jk}\left(2\pi\langle u_{jk}^2\rangle\right)^{1/2}} \exp\left[-\frac{\left(r - r_{jk}\right)^2}{2\langle u_{jk}^2\rangle}\right] \quad [23]$$

where $n_{jk}$ is the coordination number of $k$ atoms around the atom $j$. In reciprocal-space this corresponds to

$$i_{jk}(Q) = n_{jk}\,\bar{b}_j\bar{b}_k \frac{\sin\left(Qr_{jk}\right)}{Qr_{jk}} \exp\left(\frac{-\langle u_{jk}^2\rangle Q^2}{2}\right) \quad [24]$$

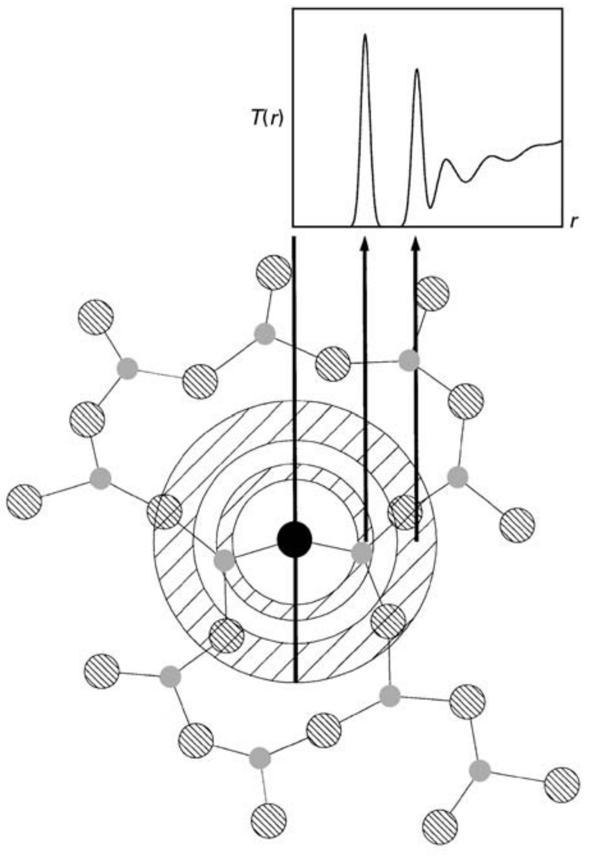



**Figure 4** A simulated neutron correlation function, $T(r)$, together with a fragment of an $A_2X_3$ network, showing how the peaks in the correlation function arise from the interatomic distances.

so that the total distinct scattering is given by

$$i(Q) = \sum_j \sum_k i_{jk}(Q) \quad [25]$$

## Small angle neutron scattering

Small-angle neutron scattering (SANS) involves the study of diffraction in the regime where the magnitude of the momentum transfer, $Q$, is small compared with the position of the first peak in the structure factor or the highest $d$-spacing Bragg peak. Thus SANS is used to study structures with dimensions of tens or hundreds of Ångströms which are large compared with interatomic distances in condensed matter. The individual scattering centres in a sample are not resolved and hence a continuous scattering length density may be introduced:

$$\rho_\mathrm{b}(\boldsymbol{r}) = \sum_l \bar{b}_l \sum_{j=1}^{N_l} \delta[\boldsymbol{r} - \boldsymbol{R}_j(0)] \quad [26]$$

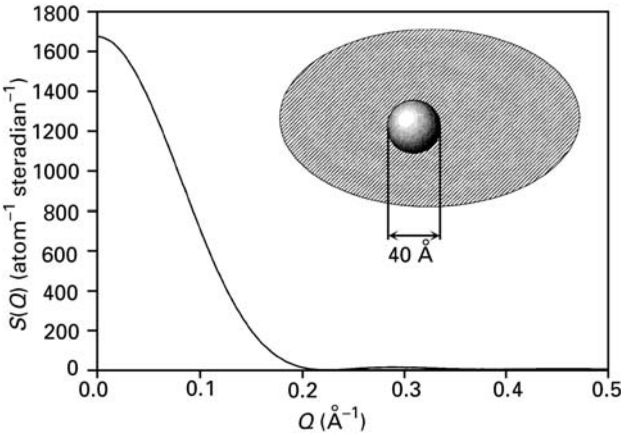



**Figure 5** The structure factor, $S(Q)$, calculated for a homogeneous sphere of radius 20 Å and uniform density 0.05 atoms Å$^{-3}$, surrounded by empty space.

The coherent differential cross-section observed in a SANS experiment is then

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{coh}} = \frac{1}{N}\left|\int\left[\rho_{\text{b}}(\boldsymbol{r}) - \rho_{\text{b}}^{0}\right]\exp(\text{i}\boldsymbol{Q}.\boldsymbol{r})\right|^{2} \quad [27]$$

where $\rho_b^0$ is the average scattering length density for the sample. A SANS diffraction experiment is thus sensitive to deviations of the scattering length density from the average value. To analyse data from a SANS experiment it is usually necessary to employ a model for the structures under investigation. For example, **Figure 5** shows the structure factor for a sphere with a homogenous scattering length density, surrounded by free space. As a consequence of the square modulus in Equation [27], SANS is only sensitive to the contrast in scattering length density. Thus the same structure factor as that shown in **Figure 5** would be obtained for an empty void of the same size embedded in a homogenous medium.

## Elastic diffraction from a crystalline sample

### Single crystal diffraction

An ideal crystalline material has an atomic structure which can be described in terms of infinite repetitions in three-dimensional space of a unit cell. If the sides of the unit cell are described by the three lattice vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, then repeating the unit cell ad infinitum by translations made up of all the possible (whole number) combinations of $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ will generate the ideal crystal structure.

For a single crystal the coherent elastic contribution to the differential cross-section is

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{coh el}} = \frac{(2\pi)^3}{\nu_0}\sum_{\tau}\delta(\boldsymbol{Q} - \boldsymbol{\tau}_{hkl})|F(\boldsymbol{Q})|^2 \quad [28]$$

where $\nu_0$ is the volume of the unit cell, given by

$$\nu_0 = \boldsymbol{a}.(\boldsymbol{b}\times\boldsymbol{c}) \quad [29]$$

$\boldsymbol{\tau}_{hkl}$ is a reciprocal lattice vector, given by

$$\tau_{hkl} = h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^* \quad [30]$$

where $h$, $k$ and $l$ are integers, and $\boldsymbol{a}^*$, $\boldsymbol{b}^*$ and $\boldsymbol{c}^*$ are the unit cell vectors of the reciprocal lattice, defined by

$$\boldsymbol{a}^* = \frac{2\pi}{\nu_0}\boldsymbol{b}\times\boldsymbol{c}; \quad \boldsymbol{b}^* = \frac{2\pi}{\nu_0}\boldsymbol{c}\times\boldsymbol{a}; \quad \boldsymbol{c}^* = \frac{2\pi}{\nu_0}\boldsymbol{a}\times\boldsymbol{b} \quad [31]$$

$F(\boldsymbol{Q})$ is known as the structure factor:

$$F(\boldsymbol{Q}) = \sum_{d}\bar{b}_d\exp\left(\text{i}\boldsymbol{Q}.\overline{\boldsymbol{R}}_d\right)\exp(-W_d) \quad [32]$$

where the $d$ summation is taken over the atoms in the unit cell and $\overline{\boldsymbol{R}}_d$ is the equilibrium position of the $d$th atom in the unit cell. The Debye–Waller factor, $W_d$, is a 'damping' term which depends upon the thermal atomic displacements. A complete treatment of the Debye–Waller factor is beyond the scope of this article, but its general behaviour is well illustrated by the simple case of a cubic Bravais lattice for which

$$2W = \frac{1}{3}Q^2\langle u^2\rangle \quad [33]$$

where $\langle u^2\rangle$ is the mean square thermal displacement of an atom from its equilibrium position.

The delta function of Equation [28], $\delta(\boldsymbol{Q}-\boldsymbol{\tau}_{hkl})$, shows that coherent elastic scattering only occurs when the following condition is satisfied

$$\boldsymbol{Q} = \boldsymbol{\tau}_{hkl} \quad [34]$$

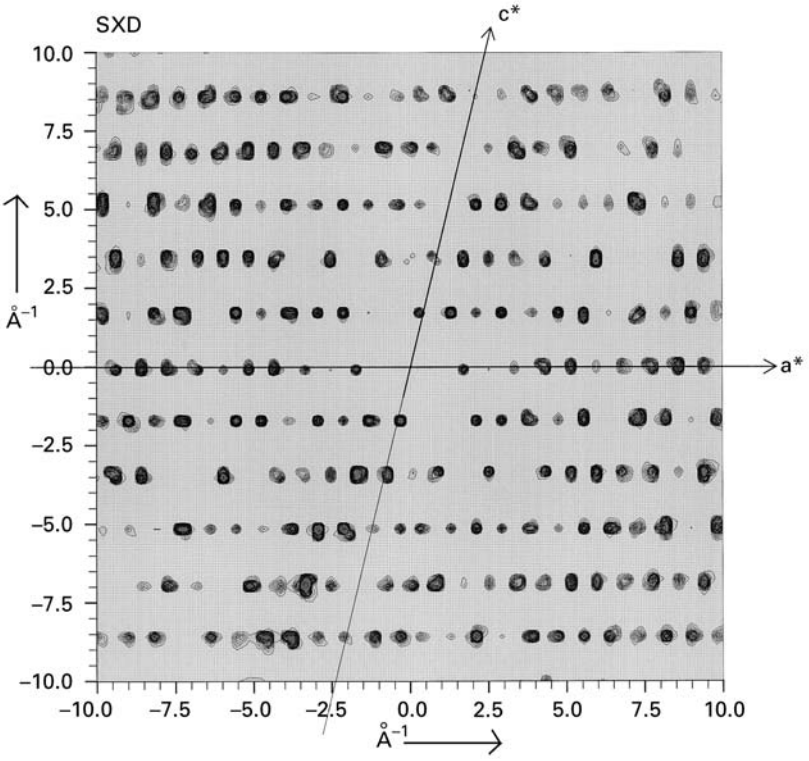



**Figure 6** A single crystal diffraction pattern for $KDCO_3$ ($a = 15.2$ Å, $b = 5.6$ Å, $c = 3.7$ Å, $\beta = 104.8°$). Reproduced courtesy of Fillaux F, Cousson A and Keen DA.

That is to say, peaks are observed in the diffraction data when the momentum transfer vector is equal to a vector of the reciprocal lattice. Such peaks are known as Bragg peaks and Equation [34] is Bragg's law, for diffraction from a single crystal. **Figure 6** shows a typical single crystal diffraction pattern; for every peak in the observed diffraction pattern the momentum transfer vector, $\boldsymbol{Q}$, satisfies Bragg's law. Each and every Bragg peak may be identified by a unique combination of indices ($h$, $k$, $l$).

The diffraction pattern of a single crystal contains a wealth of information in three-dimensional reciprocal-space and, if a sufficient number of Bragg peaks is observed, a complete description of the crystal structure may be determined. The symmetry of the Bragg peaks observed in the diffraction pattern depends upon the symmetry of the crystal structure. Thus the space group of the crystal can be determined from the symmetry observed in the diffraction pattern. The positions of the Bragg peaks in reciprocal-space depend upon the lattice parameters $a$, $b$, $c$, $\alpha$, $\beta$ and $\gamma$ (i.e. the magnitudes of the lattice vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, and the angles between them). Consequently the dimensions of the unit cell can be determined from the positions of the observed Bragg peaks. The intensity of each Bragg peak depends upon its structure factor, which itself depends on the positions of the atoms in the unit cell. Thus the contents of the unit cell can be determined from the integrated intensities of the observed Bragg peaks. In this way a complete description of the crystal structure may be developed from the single crystal diffraction pattern.

It should be noted that the structure factor, $F(\boldsymbol{Q})$, used in crystallography (Eqn [32]) has a different meaning and definition to the structure factor, $S(\boldsymbol{Q})$, used in studying diffraction from disordered samples (Eqn [15]). In addition, there is a fundamental difference between the Bragg scattering observed for crystals and the total diffraction cross-section in that the Bragg scattering is purely elastic, whereas total diffraction involves both elastically and inelastically scattered neutrons. This has the consequence that the correlation function which is addressed by the study of Bragg diffraction is the time-averaged correlation function $G^D_{ll'}(\boldsymbol{r}, \infty)$ (see Eqn [22]). Hence, for a crystal, the Bragg scattering provides information which is in principle complementary to total diffraction, which yields the instantaneous correlation function $G^D_{ll'}(\boldsymbol{r}, 0)$.

**Table 2** The planar $d$-spacings for each crystal system

| *Crystal system* | *d-spacing* |
|---|---|
| Cubic<br>$a=b=c$<br>$\alpha=\beta=\gamma=90°$ | $\frac{1}{d_{hkl}^2}=\frac{h^2+k^2+l^2}{a^2}$ |
| Tetragonal<br>$a=b$<br>$\alpha=\beta=\gamma=90°$ | $\frac{1}{d_{hkl}^2}=\frac{h^2+k^2}{a^2}+\frac{l^2}{c^2}$ |
| Orthorhombic<br>$\alpha=\beta=\gamma=90°$ | $\frac{1}{d_{hkl}^2}=\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}$ |
| Monoclinic<br>$\alpha=\gamma=90°$ | $\frac{1}{d_{hkl}^2}=\frac{1}{\sin^2\beta}\left(\frac{h^2}{a^2}+\frac{k^2\sin^2\beta}{b^2}+\frac{l^2}{c^2}-\frac{2hl\cos\beta}{ac}\right)$ |
| Hexagonal<br>$a=b$<br>$\alpha=\beta=90°$,<br>$\gamma=120°$ | $\frac{1}{d_{hkl}^2}=\frac{4}{3}\left(\frac{h^2+hk+k^2}{a^2}\right)+\frac{l^2}{c^2}$ |
| Rhombohedral<br>$a=b=c$<br>$\alpha=\beta=\gamma$ | $\frac{1}{d_{hkl}^2}=\frac{(h^2+k^2+l^2)\sin^2\alpha+2(hk+kl+lh)(\cos^2\alpha-\cos\alpha)}{a^2(1+2\cos^3\alpha-3\cos^2\alpha)}$ |
| Triclinic<br>$a\neq b\neq c$<br>$\alpha\neq\beta\neq\gamma$ | $\frac{1}{d_{hkl}^2}=\frac{\frac{h^2\sin^2\alpha}{a^2}+\frac{k^2\sin^2\beta}{b^2}+\frac{l^2\sin^2\gamma}{c^2}+\frac{2hk}{ab}(\cos\alpha\cos\beta-\cos\gamma)+\frac{2kl}{bc}(\cos\beta\cos\gamma-\cos\alpha)+\frac{2lh}{ac}(\cos\gamma\cos\alpha-\cos\beta)}{1+2\cos\alpha\cos\beta\cos\gamma-\cos^2\alpha-\cos^2\beta-\cos^2\gamma}$ |

## Powder diffraction

Although single crystal diffraction is the most powerful method for studying the structure of crystals, the powder diffraction method is also frequently used to reveal structural information for crystalline materials. In this case the sample is not a single crystal, but is instead a powder which contains numerous microcrystals, such that all orientations are equally likely. The observed diffraction pattern is then an average of Equation [28] over all possible relative orientations between the momentum transfer vector $Q$ and the crystal lattice. The measured diffraction pattern does not depend on the orientation between the sample and the neutron beam, and it is usual to identify each Bragg peak in the diffraction pattern by its $d$-spacing, $d_{hkl}$, defined by

$$|\tau_{hkl}|=\frac{2\pi}{d_{hkl}} \qquad [35]$$

Combining this definition with Equations [8] and [34] leads to Bragg's law for powder diffraction:

$$2d_{hkl}\sin\theta=\lambda \qquad [36]$$

Bragg's law indicates the combinations of $\theta$ and $\lambda$ for which Bragg peaks are observed in the powder diffraction pattern of a crystalline material. **Table 2** gives the formulae for the possible $d$-spacings of each of the crystal systems. Note that some of the $d$-spacings predicted by these formulae may not occur as peaks in the diffraction pattern, depending on the symmetry of the crystal structure; some Bragg peaks may have structure factors equal to zero, owing to the crystal symmetry, with the result that they are not allowed. The $d$-spacing of a Bragg peak may be interpreted as being the spacing for a set of planes of atoms in the crystal such that reflection from these planes gives rise to the Bragg peak (see **Figure 7**).

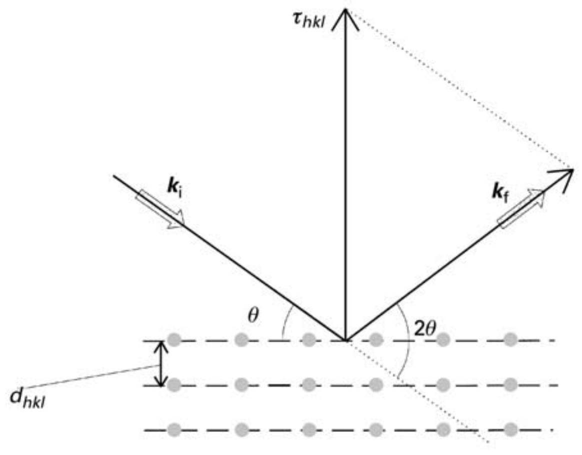



**Figure 7** Bragg planes with spacing $d_{hkl}$ which give rise to the *hkl* Bragg peak.

**Figure 8** shows the relatively simple powder diffraction pattern of polycrystalline silicon, measured by time-of-flight neutron diffraction. Each allowed Bragg peak for the crystal structure is observed as a sharp peak in the diffraction pattern. A fundamental difference between single crystal and powder diffraction is that the single crystal diffraction pattern has the full directional information described by Equation [28], whereas for the powder diffraction pattern this information has been collapsed down into a one-dimensional function. This difference has two important consequences: firstly, the loss of directional information makes it very much more difficult to deduce the symmetry of the crystal lattice from powder diffraction data. For this reason, powder diffraction is more commonly used to study the contents of the unit cell for crystals whose symmetry is already known; secondly, the Bragg peaks are much more likely to suffer from significant overlap in a powder diffraction pattern than in single crystal data. This means that it may not be possible to determine the structure factors of the reflections by integrating the intensities of individual peaks and instead the profile refinement method must be used to analyse the data. The profile refinement method involves fitting the whole of the diffraction pattern simultaneously in order to extract structural information. This method requires an explicit knowledge of the resolution function of the neutron diffractometer.

## List of symbols

$a$ = lattice parameter; $\boldsymbol{a}$ = lattice vector; $\boldsymbol{a}^*$ = reciprocal lattice vector; $b$ = lattice parameter; $\boldsymbol{b}$ = lattice vector; $\boldsymbol{b}^*$ = reciprocal lattice vector; $b_j$ = (bound atom) neutron scattering length for nucleus $j$; $\overline{b}_l$ = coherent scattering length for element $l$; $\overline{b_l^2}$ = mean square scattering length for element $l$; $\langle \overline{b}^2 \rangle_{av}$ = average value of $\overline{b}^2$ for a sample; $\langle \overline{b^2} \rangle_{av}$ = average value of $\overline{b^2}$ for a sample; $c$ = lattice parameter; $\boldsymbol{c}$ = lattice vector; $\boldsymbol{c}^*$ = reciprocal lattice vector; $c_l$ = atomic fraction for element $l$; $d_{hkl}$ = $d$-spacing; d$S$ = elemental area; $\mathrm{d}^2\sigma/\mathrm{d}E$ = double differential cross-section; $\mathrm{d}\sigma/\mathrm{d}\Omega$ = differential cross-

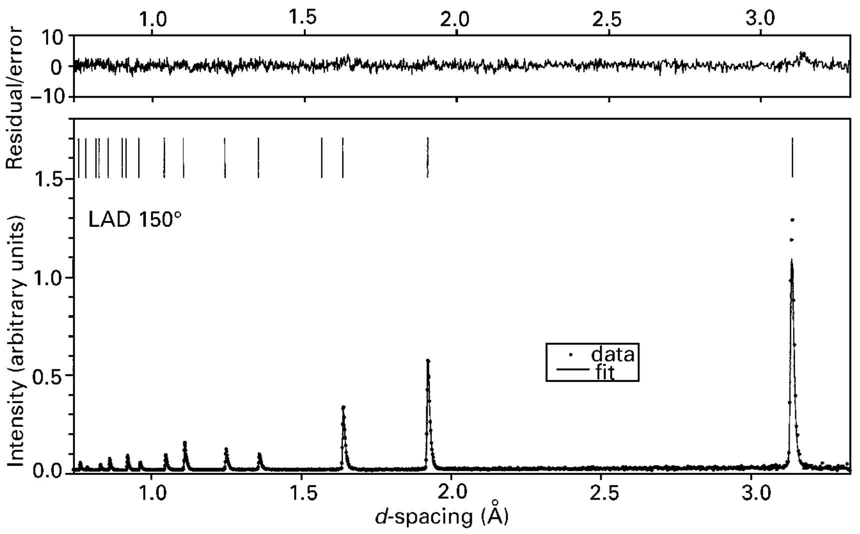



**Figure 8** The neutron diffraction pattern for polycrystalline silicon together with a fit obtained by profile refinement. Also shown is the residual divided by experimental error for the fit and vertical bars which indicate the positions of the allowed reflections under the space group $Fd\overline{3}m$.

section; $E_i$ = energy of neutrons in incident monochromatic beam; $E_f$ = energy of neutron after scattering; $E$ = energy transferred from neutron to sample; $F(Q)$ = structure factor in crystallography; $g^0$ = average atomic number density; $G^D_{ll'}(r,t)$ = generalized van Hove distinct correlation function; $h$ = index of Bragg reflection; $i(Q)$ = distinct scattering; $I$ = nuclear spin; $I(Q)$ = differential cross-section for total diffraction; $k$ = index of Bragg reflection; $\mathbf{k}$ = neutron wave vector; $\mathbf{k}_i$ = wave vector of incident neutron; $\mathbf{k}_f$ = wave vector of scattered neutron; $l$ = index of Bragg reflection; $m_n$ = neutron mass; $n_{jk}$ = coordination number; $N$ = number of atoms in sample; $N_l$ = number of atoms of element $l$ in sample; $\mathbf{Q}$ = momentum transfer; $r$ = polar coordinate; $R_x$ = the rate at which neutrons are scattered into a given final state; $\mathbf{R}_j(t)$ = position of $j$th nucleus at time $t$; $\overline{R}_d$ = equilibrium position of $d$th atom in unit cell; $S(Q)$ = structure factor for disordered materials; $t$ = time; $T$ = temperature of sample; $t_{ll'}(r)$ = partial neutron correlation function; $T(r)$ = neutron corelation function; $T^0(r)$ = average density contribution to neutron correlation function; $\langle u^2 \rangle$ = mean square thermal displacement of atom from equilibrium; $\langle u^2_{jk} \rangle$ = mean square thermal variation in interatomic distance; $V(r)$ = Fermi pseudo-potential; $W_d$ = Debye–Waller factor; $\alpha$ = lattice parameter; $\beta$ = lattice parameter; $\gamma$ = lattice parameter; $\delta(x)$ = Dirac delta function; $2\theta$ = scattering angle; $\lambda$ = neutron wavelength; $v_0$ = volume of unit cell; $\rho_b(r)$ = scattering length density; $\rho^0_b$ = average scattering length density; $\sigma_{inc}$ = incoherent cross-section; $\tau_{hkl}$ = reciprocal lattice vector; $\phi$ = polar coordinate; $\Phi$ = flux in incident neutron beam; $\Psi_f$ = wave-function of scattered neutron; $\Omega$ = solid angle.

*See also:* **Fourier Transformation and Sampling Theory; Inelastic Neutron Scattering, Applications; Inelastic Neutron Scattering, Instrumentation; Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; Material Science Applications of X-Ray Diffraction; Neutron Diffraction, Instrumentation; Powder X-Ray Diffraction, Applications; Structure Refinement (Solid State Diffraction).**

## Further reading

Bacon GE (1969) *Neutron Physics*. London: Wykeham.

Bacon GE (1975) *Neutron Diffraction*, 3rd edn. Oxford: Oxford University Press.

Brown PJ and Forsyth JB (1973) *The Crystal Structure of Solids* London: Edward Arnold.

Hannon AC, Howells WS and Soper AK (1990) ATLAS: A suite of programs for the analysis of time-of-flight neutron diffraction data from liquid and amorphous samples. *IOP Conference Series* **107**: 193–211.

Lovesey SW (1984) *Theory of Neutron Scattering from Condensed Matter. Volume 1: Nuclear Scattering*. Oxford: Oxford University Press.

Marshall W and Lovesey SW (1971) *Theory of Thermal Neutron Scattering*. London: Oxford University Press.

Price DL and Sköld K (1986) Introduction to neutron scattering. In: Sköld K and Price DL, (eds) *Neutron Scattering, Part A*, pp 1–97. Orlando: Academic Press.

Squires GL (1996) *Introduction to the Theory of Thermal Neutron Scattering*. Mineola: Dover Publications.

Wright AC (1974) The structure of amorphous solids by X-ray and neutron diffraction. *Advances in Structure Research by Diffraction Methods* **5**: 1–120.

Wright AC and Leadbetter AJ (1976) Diffraction studies of glass. *Physics and Chemistry of Glasses* **17**: 122–145.

# Nickel NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn)**.

# Niobium NMR, Applications

*See* **Heteronuclear NMR Applications (Y–Cd)**.

# Nitrogen NMR

**GA Webb**, University of Surrey, Guildford, UK

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

## Introduction

Nitrogen has two stable isotopes, $^{14}N$ and $^{15}N$, both of which are NMR active. In addition nitrogen is one of the most important atoms in organic, inorganic and biochemistry due to its occurrence in a variety of valence states with various types of bonding and stereochemistry. Thus the nuclear shieldings, spin–spin couplings and relaxation data obtained from nitrogen NMR studies are of widespread interest to molecular scientists. The relatively large range, in excess of 1000 ppm, found for the nitrogen chemical shifts of organic compounds indicates that the shifts are sensitive to relatively small electronic changes in the environment of the nitrogen atom. Consequently a change of solvent or substituent can have a very significant effect on nitrogen chemical shifts.

Solvent effects can be produced by specific interactions, such as protonation or hydrogen bonding and nonspecific interactions which may arise from solvent polarity effects. Both of these types of interaction are found in studies of solvent effects on nitrogen nuclear shieldings. The extent of substituent effects depends upon the position of substitution and the electronic nature of the substituent.

Most spin–spin couplings involving nitrogen nuclei are relatively small, and both positive and negative couplings are found. Both the sign and the magnitude of the couplings are influenced by the orientation of the nitrogen lone pair electrons with respect to the direction of the bond through which the coupling occurs.

$^{14}N$ nuclear relaxation is usually dominated by the quadrupolar mechanism which results in broad lines both in the nitrogen NMR spectrum and in the spectra of nuclei spin–spin coupled to nitrogen. The relaxation of $^{15}N$ is controlled by one, or more, of the less efficient processes.

## Nuclear properties

In natural abundance nitrogen exists in two isotopic forms, the most common is $^{14}N$ which is 99.635% abundant and has a spin of 1, the other is $^{15}N$ which has an abundance of 0.365% and a spin of $\frac{1}{2}$. Thus both isotopes are NMR active. The first report of a $^{14}N$ NMR spectrum appeared in 1950 when NMR was still very much in its infancy. However, both nitrogen isotopes have low NMR sensitivities relative to that of a proton in the same applied magnetic field; these are 0.00101 and 0.00104 for $^{14}N$ and $^{15}N$ respectively. Coupled with the relatively high cost of $^{15}N$ isotopic enrichment, the low sensitivities ensured that nitrogen NMR studies were not too widespread until the mid-1960s. Since then improvements in both NMR instrumentation and experimental techniques have generated a wide and growing interest in nitrogen NMR spectroscopy. Today the NMR spectra of both isotopes are normally studied at natural abundance due to the increased sensitivity which is currently available in high-field NMR spectrometers.

The fact that the $^{14}N$ nucleus has a spin of 1 implies that it has an electric quadrupole moment, which is able to interact with local electric field gradients in the molecule and produce rapid nuclear relaxation. A consequence of this is that $^{14}N$ NMR signals can be very broad, as much as a few kilohertz, and overlapping. Since the quadrupolar relaxation mechanism is independent of the magnitude of the magnetic field used in the NMR experiment, with a high-field spectrometer the $^{14}N$ signals appear to be relatively sharper and less overlapping than they do in a lower field instrument. Hence the widespread availability of high-field spectrometers has led to a renaissance of interest in $^{14}N$ NMR studies both from the point of view of increased sensitivity and from the diminishing in importance of the effects of quadrupolar relaxation in high-resolution NMR spectroscopy.

## Chemical shifts

The most widely used nitrogen chemical shift standard is neat nitromethane. This is used as an external reference by typically employing a set of coaxial tubes with the reference material in the inner one. Ideally an internal chemical shift standard is desirable since this removes the necessity of correcting for any magnetic bulk susceptibility difference between the sample and the standard. However, in nitrogen NMR this is usually not an option since the nitrogen shielding of a chemical shift standard may vary by up to 40 ppm due to molecular interactions in liquids and solutions. A possible candidate for the role of internal nitrogen chemical shift standard is

molecular nitrogen, $N_2$. This is present in nearly all solutions and gives a clear sharp signal in $^{14}$N NMR spectra. However, its shielding is not entirely immune from solvent effects; a range of about 2 ppm is found for solvent-induced variations. As shown in **Figure 1** the overall range of nitrogen chemical shifts for organic compounds is in excess of 1000 ppm.

Only approximate ranges are given as a general guide, and in practice most of them may be further divided into subsections to reflect the finer points of variation in molecular structure.The extent of the chemical shift range shown in **Figure 1** shows clearly that spectral assignments are usually very straightforward in nitrogen NMR. In general those nitrogen environments with σ-bonding only, and no nitrogen lone-pair electrons, are the most highly shielded ones, e.g. alkylamines. In contrast, when the nitrogen atom is involved in π-bonding and lone-pair electrons are available, then the shielding is decreased by a contribution from the paramagnetic term in the shielding expression, e.g. nitrites and nitroso compounds. In a molecular system where nitrogen is involved in multiple bonding an increase in shielding or decrease in chemical shift is observed if the nitrogen lone-pair is replaced by a covalent bond, e.g. in passing from pyridine to the pyridinium ion.

## Solvent effects on nitrogen shielding

Solvent effects on nitrogen shieldings can be remarkable and need to be carefully considered in the interpretation of nitrogen chemical shifts in terms of molecular structure. So far the largest variation reported in nitrogen chemical shift for a neutral molecular species, as a function of solvent, is about 50 ppm for pyridazine (1,2-diazine). This is shown in **Table 1** together with some other examples of typical solvent effects on nitrogen shieldings.

**Table 1** Examples of the range of solvent effects on nitrogen shieldings

| *Molecule* | *Range of solvent effects on nitrogen shielding (ppm)* |
|---|---|
| Pyridine | 38 |
| Pyridazine (1,2-diazine) | 49 |
| Pyrimidine (1,3-diazine) | 18 |
| Pyrazine (1,4-diazine) | 16 |
| 1,3,5-Triazine | 11 |
| Pyridine *N*-oxide | 30 |
| Indolizine and azaindolizine systems | |
| Bridgehead nitrogen | 1–3 |
| Pyridine-type nitrogen | 26–32 |
| Alkyl cyanides (nitriles) | 22–26 |
| Nitromethane | 11 |
| Methyl nitrate | 5 |
| *t*-Butyl nitrite | 26 |
| Methyl isothiocyanate | 10 |
| Azo bridge | 2 |
| Dinitrogen | 1.3 |

These data are quoted from systematic studies on dilute solutions where bulk susceptibility effects have been taken into account and the solvents used were chosen to encompass a broad range of properties.

In ionic species the largest solvent-induced variation of nitrogen shielding has been observed for the nitroso moiety of $[C(NO)(CN)_2]^-K^+$ in water and a number of alcohols. The range is about 200 ppm and is related to the p$K_a$ values of the solvents. The data reported in **Table 1** have been analysed in terms of the Kamlet–Taft system of solvent properties. This can be represented by the following expression:

$$\sigma(i,j) = \sigma_0 + a_i\alpha_j + b_i\beta_j + s_i(\pi_j^* + d_i\delta_j)$$

where $i$ denotes a particular nitrogen atom in the molecule examined, $j$ denotes the solvent, $\sigma$ is the relevant nitrogen shielding, $\sigma_0$ is the shielding observed in a solution using cyclohexane as a standard inert solvent, $\alpha_j$ is the hydrogen-bond donor strength of the solvent $j$, $\beta_j$ is the corresponding hydrogen-bond acceptor strength of the solvent, $\pi_j^*$ represents the solvent polarity/polarizability and $\delta_j$ is a correction for the 'superpolarizability' of aromatic and highly chlorinated solvents. The terms $a$, $b$, $s$ and $d$ are the relevant nitrogen shielding responses to the individual bulk solvent properties. If a solute nitrogen atom is protonated by the solvent then a large increase in shielding, often in excess of 100 ppm, occurs. Hydrogen-bonding from solvent to a solute nitrogen atom usually also produces a smaller increase in nitrogen shielding. In terms of ppm per scale unit of α the increase is about 21 for the nitrogen atoms of pyridine and pyridazine, about 17 for pyridine-type nitrogen atoms in the five-membered ring moieties of azaindolizines and about 10 for covalent cyanides. The magnitudes of these solvent-induced shifts reflect the relative strengths of the hydrogen-bonds concerned.

Since the $d$ term in the above equation is usually negligibly small in practice, nitrogen shielding responses to changes in solvent polarity are normally represented by the $s$ term. In terms of ppm per unit scale of $\pi^*$ the values found for $s$ may be of either sign. A positive sign indicates an increase in nitrogen shielding as the polarity of the medium increases. For pyridazine the value is about 13, whereas for other pyridine-type nitrogen atoms in azines and

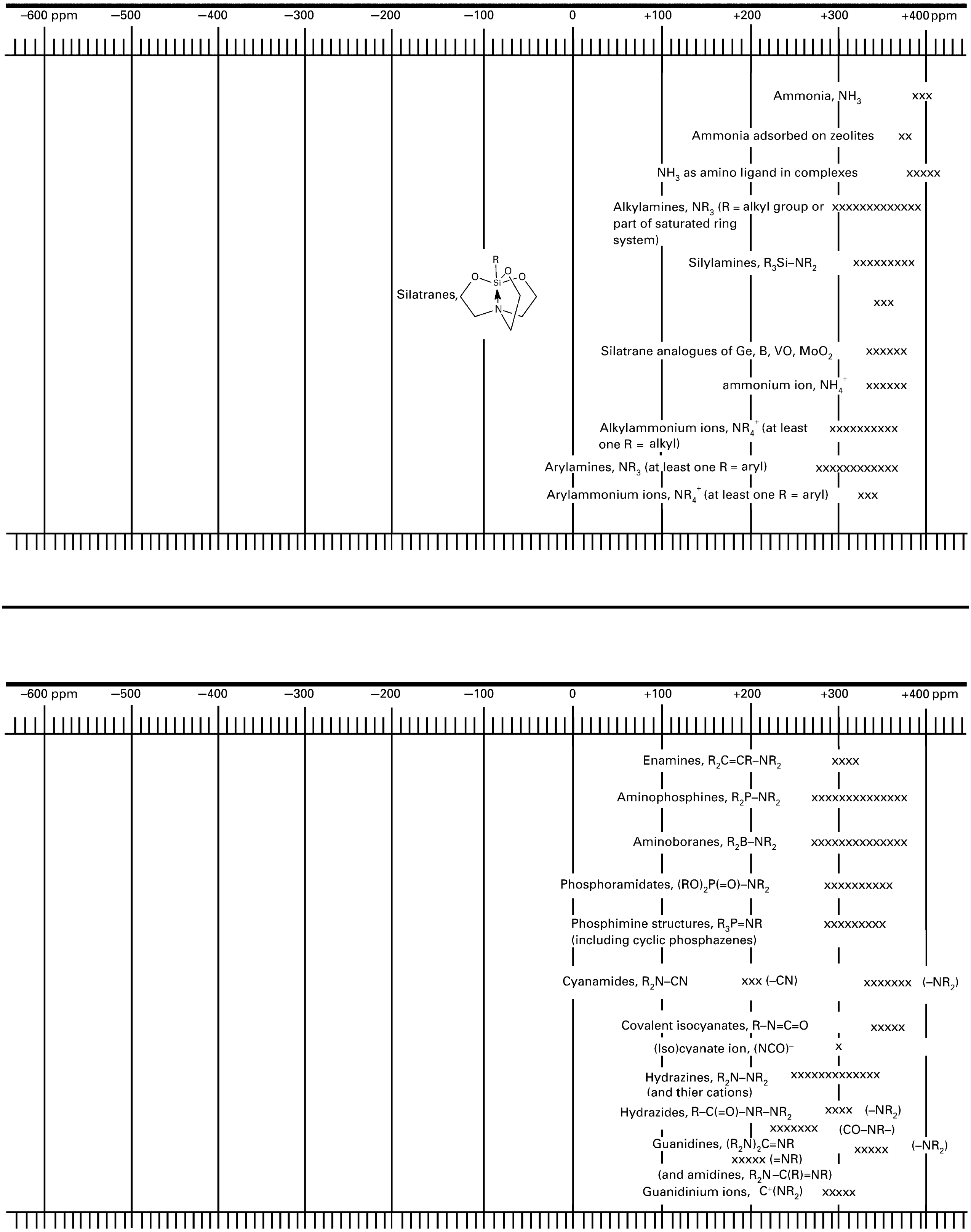



**Figure 1** *Continued*

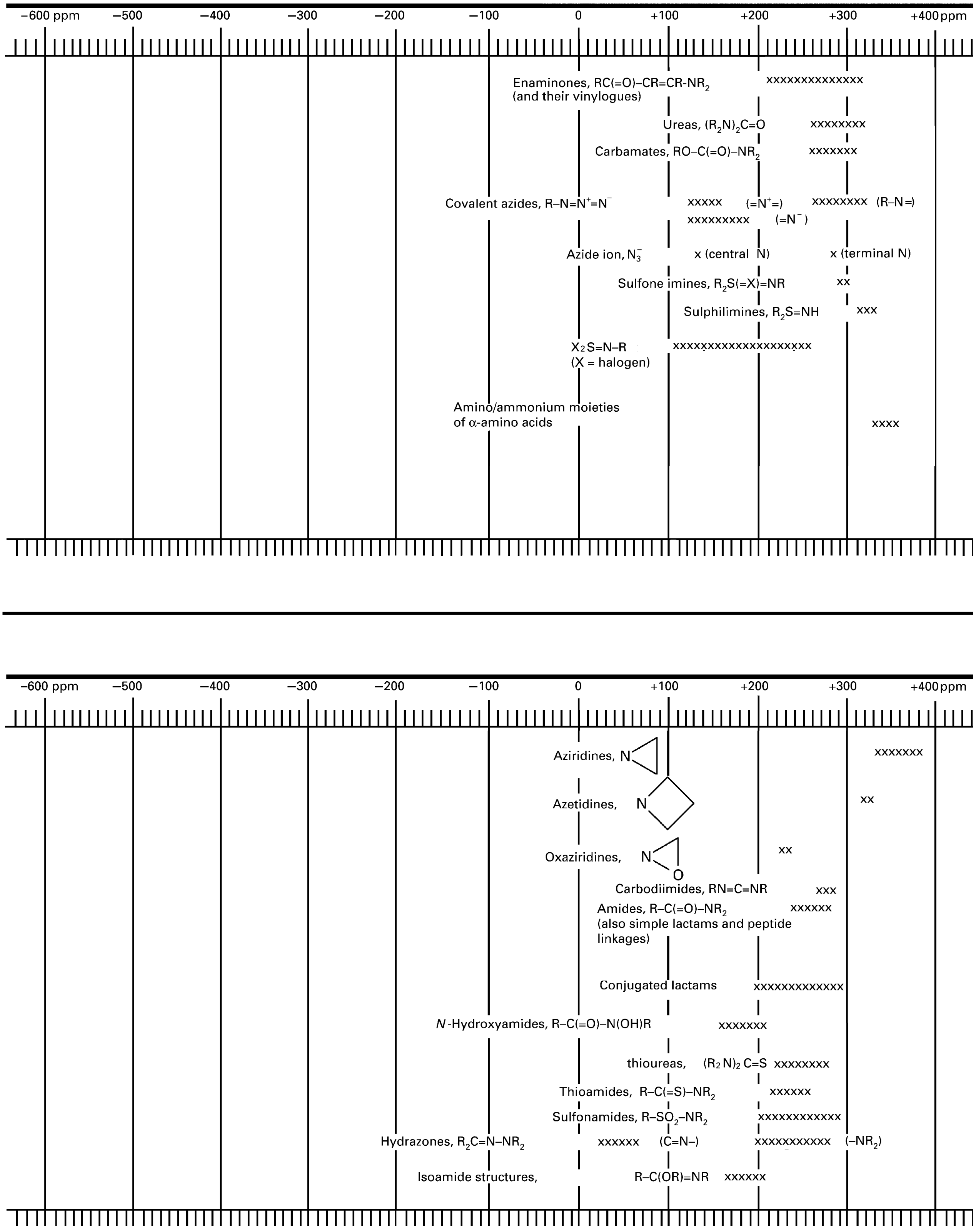
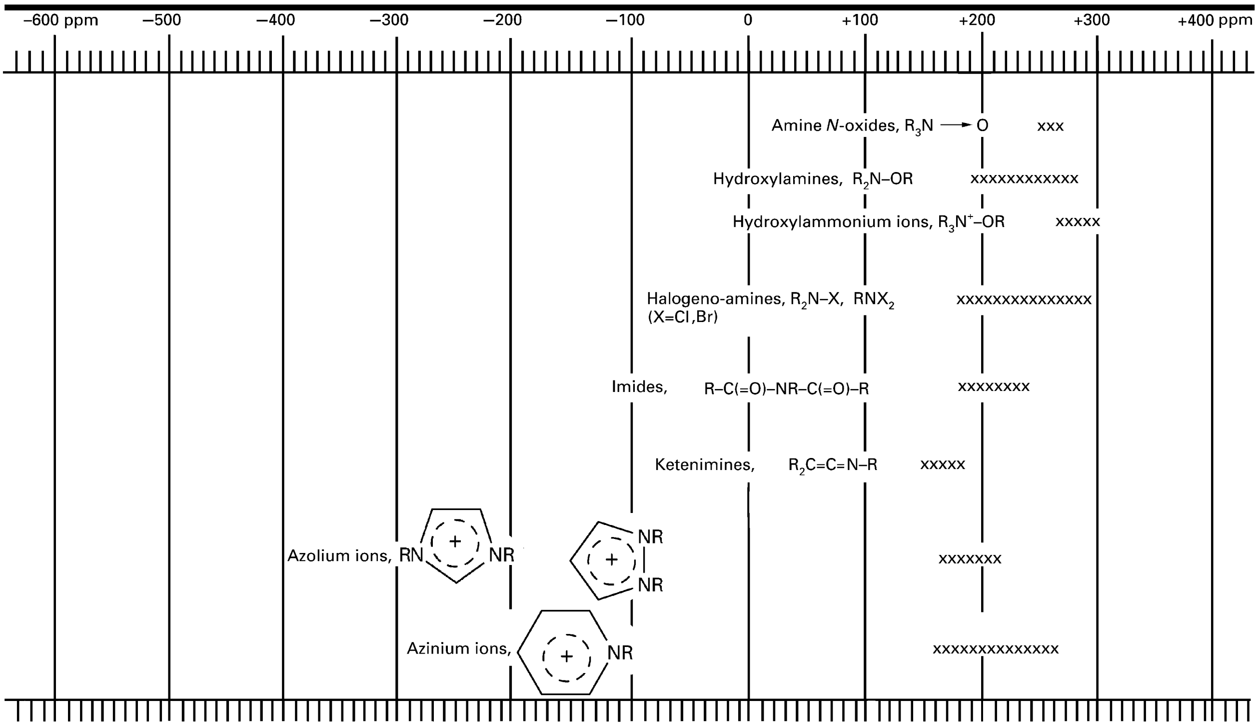



**Figure 1** *Continued*



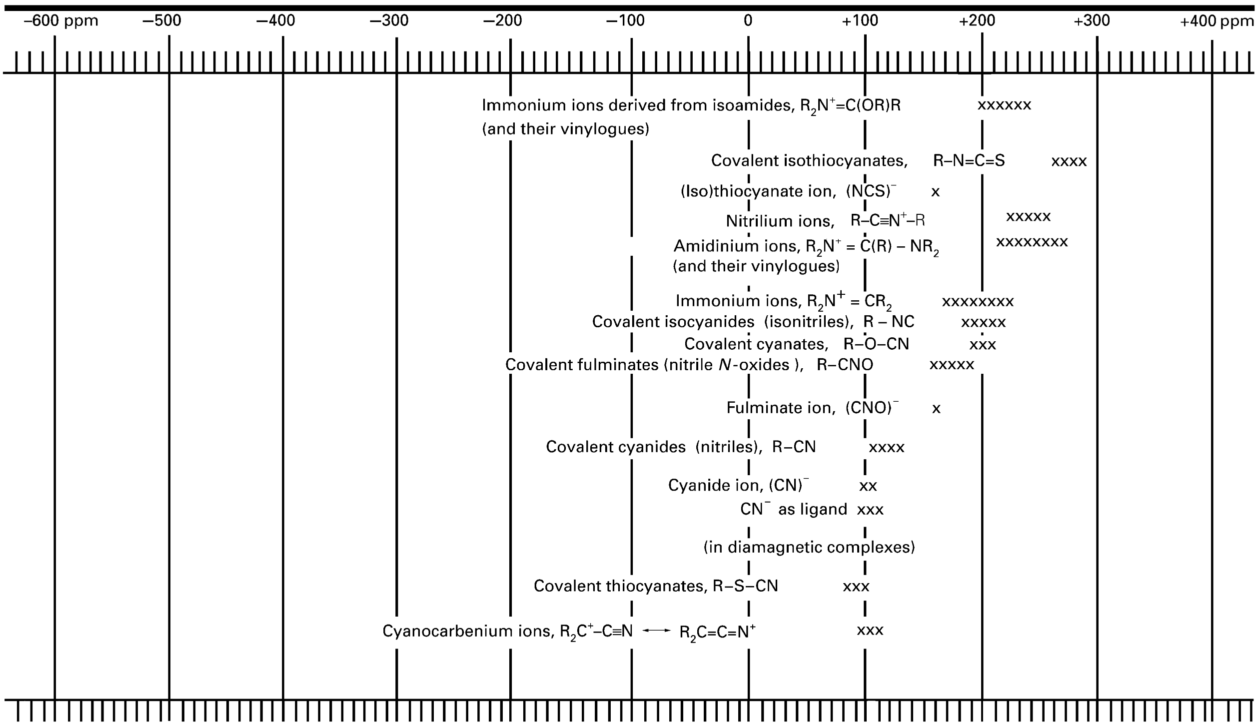
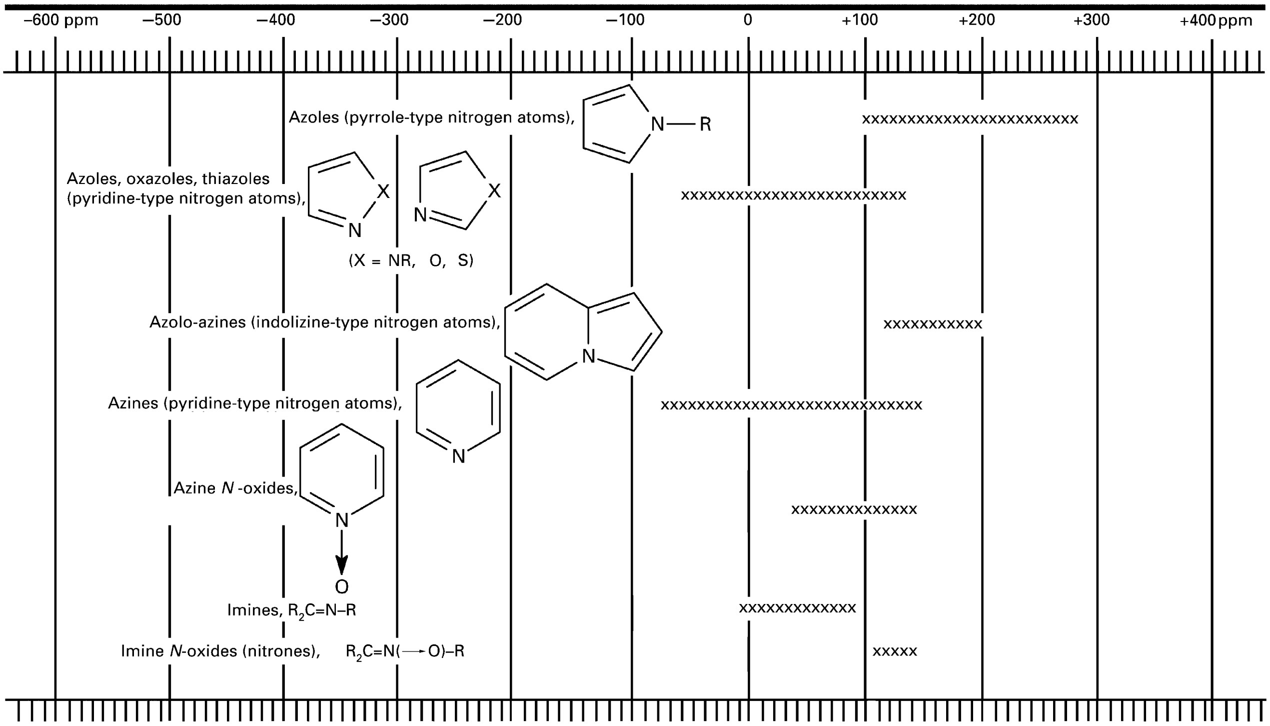




**Figure 1** *Continued*



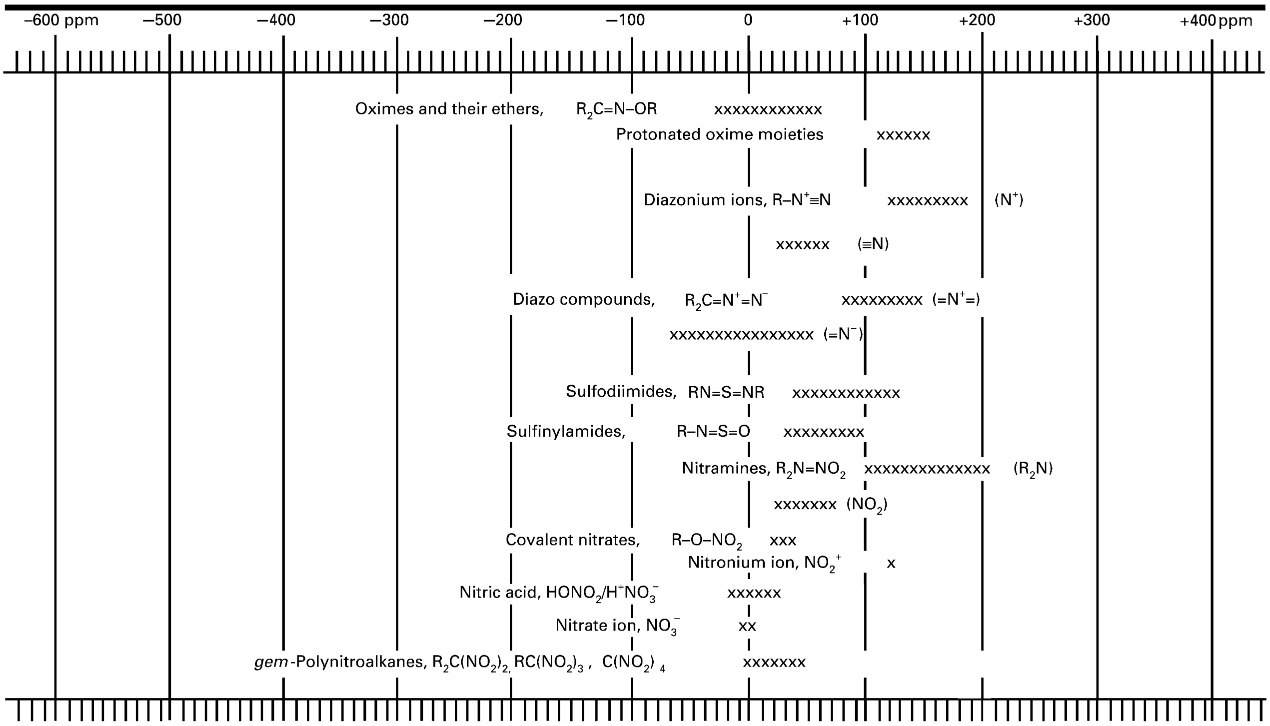




**Figure 1** *Continued*

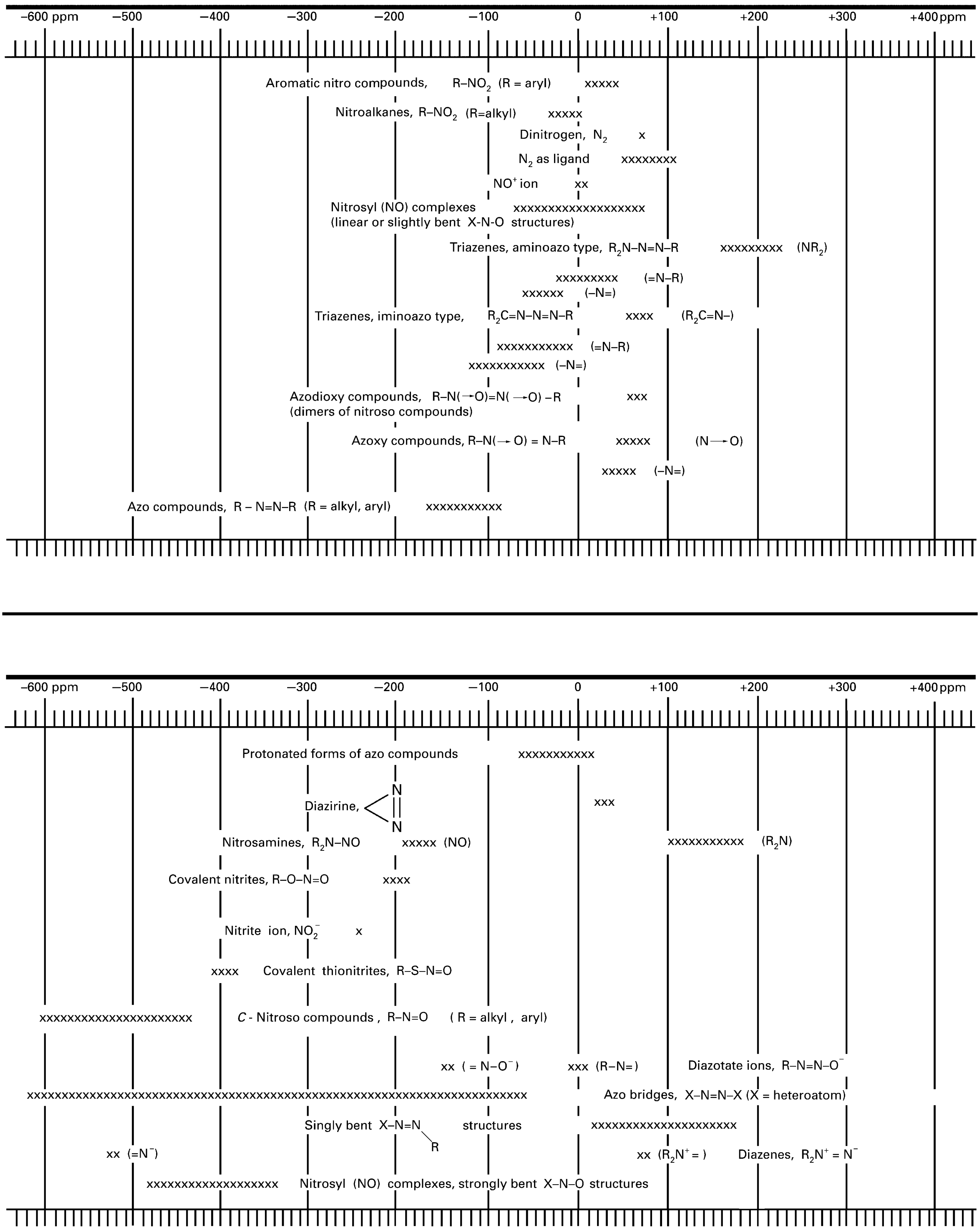



**Figure 1** Characteristic nitrogen shielding ranges for various classes of molecules and ions (referred to external neat nitromethane)

azaindolizines it is in the region of 5. The value of *s* is also found to be positive for the nitrogen atoms of covalent cyanides, whereas it is negative for covalent nitrites, isothiocyanates, C-nitro and O-nitro groups. The sign of *s* is related to the orientation of bond dipoles in the vicinity of the nitrogen atoms concerned. This point of view is supported by molecular orbital calculations incorporating the Solvaton model.

In general the *trans*-azo bridge nitrogen atoms reveal only a very small shielding variation as the solvent is changed. An exception occurs when trifluoroethanol is used as solvent. This is probably due to a hydrogen-bonding interaction with the $\pi$-electrons of the N=N bond. The nitrogen shieldings of nitroso groups show a diversity of solvent-induced effects. The observed range for $Bu^t$-NO is about 10 ppm. The major contribution to this range of solvent-induced shielding changes arises from variations in solvent polarity. This is probably due to the fact that the bulky *tert*-butyl group prevents any significant amount of solvent-to-solute hydrogen bonding in this case. In contrast, the O-nitroso groups in covalent nitrites show much larger solvent-induced nitrogen shielding ranges which include significant contributions from solvent-to-solute hydrogen bonding effects.

## Substituent effects on nitrogen shielding

A general trend in the values of the nitrogen shielding, which is observed for a large variety of structures, is for the shielding to decrease along the sequence, N–$CH_3$, N–$CH_2R$, N–$CHR_2$, N–$CR_3$, where R is an alkyl group. This is referred to as the $\beta$ effect and is useful for locating nitrogen-containing functional groups within a hydrocarbon chain. Other comparable effects may arise when alkyl groups replace hydrogen atoms at other positions in an aliphatic chain, as shown in **Figure 2**.

For aliphatic amines the nitrogen shielding decreases by 7.8 ppm when an alkyl group is introduced into the $\alpha$ position. It decreases by 18.6 ppm when the introduction occurs at the $\beta$ position. If the alkyl group appears at the $\gamma$ position the opposite effect is observed and the shielding increases by 2.4 ppm, whereas the introduction of the alkyl group at the $\delta$ position produces a further decrease of nitrogen shielding by 2.5 ppm. Local molecular interactions can also be monitored by means of changes in nitrogen chemical shifts. Interactions of protein fragments provide an example of this, as shown in **Figure 3**.

*Escherichia coli* thioredoxin has characteristic sites which are affected by the transformation between its reduced and oxidized forms. Variations in the nitrogen chemical shifts of the amino acid fragments have been used in a study of the binding of $Mg^{2+}$, $Ca^{2+}$ and $Ba^{2+}$ to *E. coli* ribonuclease HI.

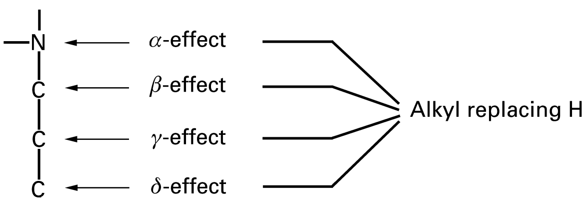



**Figure 2** Effect on nitrogen chemical shifts caused by replacing hydrogen atoms by alkyl groups in an aliphatic chain.

## Spin–spin couplings

Due to the normally rapid quadrupolar relaxation of $^{14}N$ nuclei, spin–spin coupling interactions between $^{14}N$ and other nuclei are only seldom observed. Thus most couplings involving nitrogen nuclei are usually measured from the NMR spectra of $^{15}N$-coupled nuclei. Spin–spin couplings between a nucleus X and $^{14}N$ and between X and $^{15}N$ are interconvertible by means of the following equation:

$$J(^{15}\text{N}{-}\text{X}) = -1.4027\ \text{J}\ (^{14}\text{N}{-}\text{X})$$

Spin–spin couplings involving $^{15}N$ play an important role in the application of nitrogen NMR to the structure determination of nitrogen-containing molecules, since their values are often characteristic of the type and number of intervening bonds between the coupled nuclei.

Values of $^1J(^{15}N–^1H)$ are negative in sign and considerably larger in magnitude than spin–spin couplings between this pair of nuclei when separated by more than one bond. If proton exchange occurs within a given molecule then the observed value of the $^{15}N–^1H$ coupling represents a weighted average

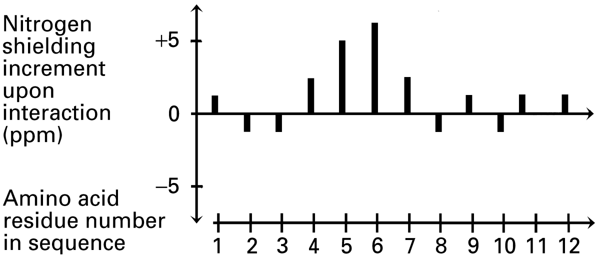



**Figure 3** Effect on nitrogen shielding due to interactions of protein fragments.

which includes coupling with the proton at all its sites of residence. Such a situation arises with some porphyrins, for example octaethylporphyrin in $CDCl_3$. At –53°C the $^{15}N$ NMR spectrum consists of a singlet and a doublet split by 98 Hz; this latter value is typical of $^1J(^{15}N–^1H)$ in pyrrole-type systems. At +28° C only a quintet is observed split by 24 Hz. This indicates that the NH protons are exchanging amongst the four nitrogen atoms at higher temperatures and that the long-range N–H couplings are close to zero.

If major structural differences between molecules are considered then values of $^1J(^{15}N–^1H)$ often show a reasonable correlation with the s-characters of the N–H bonds involved. However, notable exceptions are known and thus it may be unsound to try to estimate the s-character of a given N–H bond from the observed value of the $^1J(^{15}N–^1H)$ interaction. Any such correlation would have to rely on the dominance of the contact term to the spin–spin coupling interaction. In cases where this dominance occurs then a simple distinction may be made between a variety of structures by means of $^1J(^{15}N–^1H)$ values.

Since absolute values of $^1J(^{15}N–^1H)$ are normally of the order of 100 Hz, the collapse of multiplet patterns in the $^{15}N$ NMR spectra of NH moieties can be used to monitor proton exchange processes which occur at rates of around 100 $s^{-1}$. Some examples where this technique has been used are ureas, lactams, arginine and histidine. Values of $^2J(^{15}N–^1H)$) across a saturated carbon atom are normally small in magnitude and positive in sign. Such couplings reveal a dependence of the orientation of the lone-pair electrons on the nitrogen atom with respect to the C–H bond. The largest values are observed when the bond is *cis* to the lone pair. If the intervening carbon atom is tricoordinate, then the value of $^2J(^{15}N–^1H)$) is much larger, about 15 Hz. When the coupling occurs in an imino-type moiety the value of $^2J(^{15}N–^1H)$ depends critically upon the bond structure at the nitrogen atom in question. If the nitrogen atom has a lone pair of electrons, the coupling is usually large in magnitude and negative in sign. If protonation of the lone pair electrons occurs, then the absolute value of the coupling decreases significantly. An even more dramatic change occurs when an N-oxide is formed; the resulting coupling may then have a small positive value. Similar observations have been made for cyclic lactams of the uracil type; if the nitrogen has a lone pair of electrons the observed value of $^2J(^{15}N–^1H)$ is large, e.g. the anion of 3-methyluracil. The coupling is considerably reduced when a hydrogen atom is attached to the nitrogen.

Values of $^3J(^{15}N–^1H)$ in saturated systems are often larger than the corresponding two-bond couplings. However, the three-bond couplings depend upon the dihedral angle between the N–C and C–H bonds and attain maximum values for the *cis* and *trans* arrangements and a minimum value when the dihedral angle is around 90°. Thus for the *gauche* configuration, where the dihedral angle takes a value of about 60°, rather small absolute values of $^3J(^{15}N–^1H)$ are expected. For unsaturated systems, e.g. aza-aromatic structures, the values of $^3J(^{15}N–^1H)$ do not differ appreciably in absolute magnitude from those found in saturated systems. This is in contrast to the situation found for values of $^2J(^{15}N–^1H)$.

In general, few values have been found for $^{15}N–^1H$ couplings across more than three bonds. Those reported for pyridine, its cation and its N-oxide are small and positive in sign, whereas those for nitrobenzene are small and negative in sign.

Usually the values of $^1J(^{15}N–^{13}C)$ are negative in sign and less than ~35 Hz. However, the range of values reported is from +4.9 Hz for an oxaziridine derivative to –77.5 Hz for 2,4,6-trimethylbenzonitrile *N*-oxide. Some attempts have been made to relate $^1J(^{15}N–^{13}C)$ data to the amount of s-character of the N–C bonds. Such a correlation would depend upon the dominance of the contact interaction to the spin–spin coupling and this is not always the case in practice. In general, nitrogen–carbon couplings across one bond are larger than those across more bonds. However, there are exceptions. These usually arise when the nitrogen atom in question has a lone pair of electrons in an orbital with high s-character. Examples are provided by pyridine-type nitrogen atoms in azines and azoles and imino-type nitrogen atoms in imines and oximes. For pyridine the value of $^1J(^{15}N–^{13}C)$ is +0.62 Hz, that for $^2J(^{15}N–^{13}C)$ is +2.53 Hz and for $^3J(^{15}N–^{13}C)$ it is –3.85 Hz. In general it pays to be cautious in N–C spin–spin coupling data to determine the presence, or absence, of specific N–C bonding arrangements.

Values of $^2J(^{15}N–^{13}C)$ are usually smaller then those across one bond in saturated systems. If the coupling is across a carbonyl carbon atom then the absolute value of $^2J(^{15}N–^{13}C)$ increases significantly to about 4–12 Hz. Thus a means is provided for distinguishing between saturated and unsaturated molecular fragments. If the nitrogen atom involved in the coupling has a lone pair of electrons in an orbital with significant s-character then the sign and magnitude of $^2J(^{15}N–^{13}C)$ depend critically upon whether the coupled carbon nucleus is *cis* or *trans* to the lone-pair electrons.

Three-bond $^{15}N–^{13}C$ couplings are usually negative in sign and less than 5 Hz. For saturated systems attempts have been made to relate values of $^3J(^{15}N–^{13}C)$ to the dihedral angle between the C–C

and N–C bonds. Usually the rather small values observed for the couplings give rise to considerable errors in any such angle determination.

Couplings between $^{15}N$ and $^{13}C$ across more than three bonds are usually less than 1 Hz in absolute magnitude. A lone exception is a derivative of 1,2,4-triazine where a value of 3.9 Hz has been reported for the four-bond coupling between 1-N and a methyl group attached to a vinyl substituent.

The largest $^{1}J(^{15}N–^{15}N)$ values are found for *N*-nitrosoamines and the smallest for hydrazino moieties, nitramines and molecular $N_2$. Generally, values occur within the range of 10–20 Hz for N=N moieties. Exceptions are found when the moieties are charged, as they are in diazo compounds and azides, where the coupling is less than 10 Hz. The values of $^{2}J(^{15}N–^{15}N)$ across a nitrogen atom are close to zero in azides but may be as large as 11 Hz in some isomers of imino-type triazenes. Two-bond $^{15}N–^{15}N$ couplings across a carbon atom can also be significant, particularly when the carbon atom in question is tricoordinate.

Values of $^{1}J(^{31}P–^{15}N)$ cover a broad range and may be either positive or negative in sign. Hence, if only the absolute value is known their interpretation is not always straightforward. It seems that the value observed depends critically upon the dihedral angle between the nitrogen and phosphorus lone pairs of electrons. A significant algebraic decrease in the value of $^{1}J(^{31}P–^{15}N)$ occurs upon passing from tricoordinate to tetracoordinate phosphorus atoms.

A number of values of $^{2}J(^{31}P–^{15}N)$ have been reported for couplings across metal atoms in various complexes. For square-planar complexes a clear distinction is found in the values of the couplings depending upon whether the ligands in question are *cis* or *trans* to each other.

Values of $^{1}J(^{19}F–^{15}N)$ are positive in sign and usually between 150 and 460 Hz. Nitrogen–fluorine couplings across two or more bonds are also significant in magnitude.

$^{1}J(^{195}Pt–^{15}N)$ couplings have large absolute values, usually ranging from 100 to 580 Hz. They are found to be sensitive to the nature of the ligands, as well as to their arrangement in square-planar complexes. Normally the largest influence on the value of the coupling is exerted by the ligand that is *trans* to the nitrogen atom involved.

## Nitrogen nuclear relaxation

Since the $^{14}N$ nucleus has an electric quadrupole moment, this may interact with local electric field gradients in the molecule, to produce rapid nuclear relaxation and broad NMR transitions. The electric field gradient may vary due to molecular motion and this variation may be studied by means of $^{14}N$ NMR. A combination of the $^{14}N$ quadrupolar relaxation rate and $^{13}C–^{1}H$ dipolar relaxation data can be used to determine the rotational correlation times for motion about each principal molecular axis. This approach has been applied to pyrimidine, pyridazine and pyrazine. $^{14}N$ relaxation data for formamide are found to be very sensitive to the presence of both cations and anions. These results provide direct evidence for specific ion–amide interactions and a tentative model for the interaction of electrolytes in liquid formamide.

$^{14}N$ NMR line widths may be used as an aid to the assignment of nitrogen chemical shifts. In particular, a nitrogen atom bearing a partial positive charge tends to have a relatively small local electric field gradient associated with it, consequently its $^{14}N$ signal will be rather sharp. This approach has been extensively used in studies on meso-ionic compounds.

$^{15}N$ nuclear relaxation usually occurs through varying contributions from the dipole–dipole, spin-rotation, chemical shielding anisotropy and scalar coupling mechanisms. If the $^{15}N$ nucleus is directly bonded to a proton then the dipole–dipole interaction is normally dominant. In the case of *trans*-azobenzene, the $^{15}N$ relaxation occurs by a combination of the spin-rotation, dipole–dipole and chemical shielding anisotropy interactions. The relative proportions of these mechanisms contributing to the relaxation are found to vary considerably over the temperature range from 5°C to 80°C. The $^{15}N$ relaxation times of a number of aldoximes and ketoximes are between 25 s and 50 s. In general the dipole–dipole interaction is the major contributor but other mechanisms can account for more than half of the measured $^{15}N$ relaxation in these molecules.

*See also:* **Chemical Shift and Relaxation Reagents in NMR; NMR Relaxation Rates; Relaxometers; Solvent Suppression Methods in NMR Spectroscopy; Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy; Structural Chemistry using NMR Spectroscopy, Inorganic Molecules; Structural Chemistry using NMR Spectroscopy, Organic Molecules.**

## Further reading

Witanowski M and Webb GA (1972) *Annual Reports on NMR Spectroscopy* **5A**: 395–464.

Witanowski M and Webb GA (eds) (1973) *Nitrogen NMR*. London: Plenum Press.

Witanowski M, Stefaniak L and Webb GA (1977) *Annual Reports on NMR Spectroscopy* **7**: 117–244.

Witanowski M, Stefaniak L and Webb GA (1981) *Annual Reports on NMR Spectroscopy* **11B**: 1–502.

Witanowski M, Stefaniak L and Webb GA (1986) *Annual Reports on NMR Spectroscopy* **18**: 1–761.

Witanowski M, Stefaniak L and Webb GA (1993) *Annual Reports on NMR Spectroscopy* **25**: 1–480.

# NMR Data Processing

**Gareth A Morris**, University of Manchester, UK

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Methods & Instrumentation

Computers play a central part in modern NMR spectroscopy. Their use for the real-time control of pulsed NMR experiments has enabled the development of multiple pulse techniques such as two-dimensional NMR; this article deals with the part played by computers in the acquisition, processing and presentation of experimental NMR data.

All NMR experiments rely on the excitation of a nuclear spin response by a radiofrequency magnetic field, usually in the form of a short pulse or pulse sequence. This generates a rotating nuclear magnetic moment, which induces a small oscillating voltage in the probe coil: a ‘free induction decay’ (FID). The spectrometer receiver amplifies this voltage, shifts it down into the audio frequency range, and filters out high-frequency components before it is digitized by an analogue-to-digital converter (ADC). In modern instruments two receiver channels with phases 90° apart are used (‘quadrature detection’), allowing the relative signs of frequencies to be distinguished. From this point onwards the signal is handled digitally until it is presented to the experimenter as a printed spectrum or an interactive spectral display. The three main stages of processing are first, the acquisition of a filtered, averaged time-domain recording of the nuclear spin response; second, the generation of a frequency-domain spectrum, usually but not always by Fourier transformation; and third, postprocessing of the spectrum to aid its interpretation. The three stages are summarized briefly below, followed by more detailed discussions of the techniques used and some practical illustrations.

## Data acquisition

NMR experiments generally require the coaddition of a number of recordings of the nuclear spin response (‘transients’), often using different permutations of radiofrequency pulse and receiver phases (‘phase cycling’). The data are recorded as a set of complex numbers which sample the in-phase and quadrature NMR signals as a function of time. Early Fourier transform spectrometers had limited word length and required some scaling of the data before coaddition, but this is no longer necessary and successive transients are simply added to memory. Many recent spectrometers interpose a stage of digital signal processing before data summation: a wide receiver band width is used, and data are digitized very rapidly. These oversampled data are then filtered digitally before downsampling and addition to memory, giving better filtration of unwanted noise and signals, better effective ADC resolution, and less baseline distortion.

In two-dimensional (2D) NMR a series of FIDs is acquired using a pulse sequence containing a variable evolution period, which is incremented regularly to map out the behaviour of the nuclear signal as a function both of real time $t_2$ during the FID, and of the evolution time $t_1$. A typical 2D NMR experiment might acquire 512 FIDs of 1024 complex points, which would then be doubly Fourier transformed as a function of the two time variables. 3D and 4D NMR extend the principle to two and three evolution periods respectively.

## Spectrum generation

To make the raw experimental data interpretable they must be converted into a frequency-domain spectrum. The classical, and commonest, method is discrete Fourier transformation. This is a linear operation: the information content of the data is unchanged. Alternative methods such as maximum entropy reconstruction and linear prediction are nonlinear, changing the information content. In Fourier processing a weighting function is usually applied to the time-domain data before transformation; a DC correction based on the last portion of the data may also be performed. After weighting and any

zero-filling (see below), the fast Fourier transform (FFT) algorithm is used to produce the frequency spectrum. The resultant spectrum consists of a set of complex numbers sampling a defined frequency range at equal intervals. In general the real and imaginary parts of the spectrum will both be mixtures of absorption mode and dispersion mode signals; a spectrum suitable for display and analysis is obtained by taking linear combinations of the real and imaginary data ('phasing'), or, if that is not possible, by taking the modulus ('absolute value mode') or square modulus ('power mode').

In 2D NMR, weighting and zero filling are carried out on the FIDs as normal, but after Fourier transformation with respect to $t_2$ to give a series of spectra $S(t_1, F_2)$ the data matrix is transposed to give a set of 'interferograms' $S(F_2, t_1)$. A second Fourier transformation, with respect to $t_1$, yields the 2D spectrum $S(F_1, F_2)$. Because most coherence transfer processes cause amplitude rather than phase modulation as a function of $t_1$, information on the signs of $F_1$ frequencies is usually missing; it can be recovered by making two measurements using different phase cycles. If the two sets of measurements are combined before Fourier transformation, converting amplitude modulation into phase modulation, absolute value mode presentation is generally required because individual signals will show a 'phase-twist' line shape; this is also the case with simple experiments which use pulsed field gradients for coherence transfer pathway selection. If the two data sets are recombined appropriately after the first Fourier transformation, both phase cycled and pulsed field gradient 2D experiments can be made to yield pure double absorption mode line shapes. Two common recombination schemes are the 'hypercomplex' method of States, Haberkorn and Ruben, and the time-proportional phase incrementation (TPPI) method of Marion and Wüthrich.

## Postprocessing

Many chemical questions can be answered with a simple spectrum, but to extract all the useful information from an NMR experiment it can require a wide range of data processing methods. Integration (usually after baseline correction) and peak picking allow the relative numbers of spins responsible for different multiplets to be found, and chemical shifts and scalar couplings to be measured. Better measures of signal intensity and line shape can be found by least-squares fitting, if necessary in conjunction with line shape correction by reference deconvolution. Computers can also aid in the analysis of strongly coupled spectra and the spectra of dynamic systems.

## Zero-filling

Data acquisition produces a set of $N$ complex points, sampled at equal time intervals $\Delta t$, which describe the nuclear FID as $2N$ independent pieces of information. Discrete Fourier transformation of these data will produce a spectrum of $N$ complex points, at frequency intervals $1/(N\Delta t)$. A final absorption mode spectrum will contain just $N$ independent pieces of information, only half the amount originally acquired; the remaining $N$ points form the dispersion mode spectrum. Full use of the experimental data can be achieved if $N$ complex zeroes are appended to it before Fourier transformation ('zero-filling'). Transforming $N$ data points plus $N$ zeroes generates a spectrum of $2N$ complex points, at frequency intervals $1/(2N\Delta t)$ Hz. The $2N$ real points are independent, and contain the same information as the $2N$ imaginary points: the real and imaginary data are correlated, so the total information content of the spectrum is unchanged. Zero-filling thus improves the digital resolution of the frequency spectrum two-fold, as can be seen from **Figure 1**. Appending more than $N$ zeroes before transformation cannot increase the information content of the spectrum: the extra data points obtained simply interpolate between those produced by a single zero-filling.

## Time-domain weighting

All experimental NMR signals decay, sooner or later: most well-designed experiments sample the signal until it has decayed close to zero. The later stages of the recorded data contain less signal, but the noise remains more or less constant. Recording for too short a time will lose valuable data; recording for too long will emphasize the noise relative to the signal. The widths of spectral lines depend on the rate of decay of the NMR signal: resolution can be improved if the natural decay of the signal is counteracted. Both issues can be addressed by weighting the time-domain signal before Fourier transformation. The choice of weighting function determines the compromise between resolution and signal-to-noise ratio in the resultant spectrum.

The signal-to-noise ratio of a spectrum can be optimized by multiplying the experimental signal by a weighting function which matches the experimental decay envelope: 'matched filtration'. The Fourier transform of an exponential decay with time constant $T$ is a Lorentzian line shape with a full width at half height of $1/(\pi T)$ Hz. Thus to obtain the best signal-to-noise ratio for a Lorentzian line of width $W$ Hz the experimental data should be multiplied by a decaying exponential of time constant $1/(\pi W)$ s. This

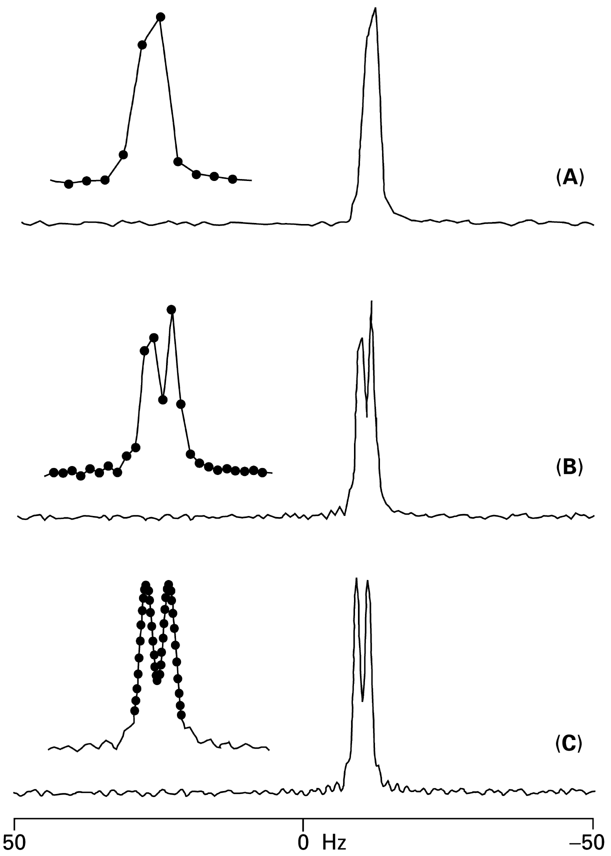



**Figure 1** Spectra of a doublet with splitting 2 Hz, centred at –10 Hz, calculated for a 64 complex point FID with (A) no zero filling; (B) one zero filling, to 128 complex points; and (C) four zero fillings, to 1024 complex points. The splitting becomes visible after one zero filling; further zero filling is equivalent to interpolation between the data points with a sin(*x*)/*x* function.

gives a spectrum with optimum signal-to-noise ratio, at the expense of a doubling of the line width to 2 *W* Hz. Where a spectrum with poor signal-to-noise ratio contains lines with a range of widths it can be helpful to try exponential weighting with several different time constants.

Time-domain weighting is equivalent to frequency-domain convolution. The convolution theorem states that the Fourier transform of the product of two functions $a(t)$ and $b(t)$ is the convolution of the two individual transforms $A(\nu)$ and $B(\nu)$:

$$\begin{aligned}\mathrm{FT}^{-}[a(t) \times b(t)] &= \int_{-\infty}^{\infty} a(t) \times b(t) \exp(-2\pi \mathrm{i} \nu t)\, \mathrm{d}t \\ &= \int_{-\infty}^{\infty} A(\nu') \times B(\nu - \nu')\, \mathrm{d}\nu' \\ &= \mathrm{FT}^{-}[a(t)] \otimes \mathrm{FT}^{-}[b(t)] \\ &= A(\nu) \otimes B(\nu) \qquad [1]\end{aligned}$$

where FT$^{-}$[ ] indicates Fourier transformation and ⊗ denotes convolution. Thus time-domain exponential weighting is equivalent to convolution, or smoothing, with a Lorentzian lineshape in the frequency domain. Matched filtration corresponds to smoothing the raw experimental spectrum with a function which matches the experimental line shape, as **Figure 2** illustrates.

Time-domain weighting is also extensively used for resolution enhancement. Since this emphasizes the later part of the experimental signal, the noise energy is increased with respect to the signal energy, and resolution enhancement reduces the signal-to-noise ratio of the resultant spectrum. The aim of resolution enhancement is to reduce line widths without degrading the signal-to-noise ratio unacceptably. The natural decay of individual NMR signals is normally exponential, but countering this decay by multiplication with a rising exponential would lead to steeply rising noise. To stop the exponential rise in noise a further weighting using a function with a steeper decline is required. A weighting function $W(t)$ composed of a rising exponential with time constant $t_e$ and a falling Gaussian with time constant $t_g$

$$W(t) = \exp(+t/t_e) \exp[-(t/t_g)^2] \qquad [2]$$

is generally the method of choice for resolution

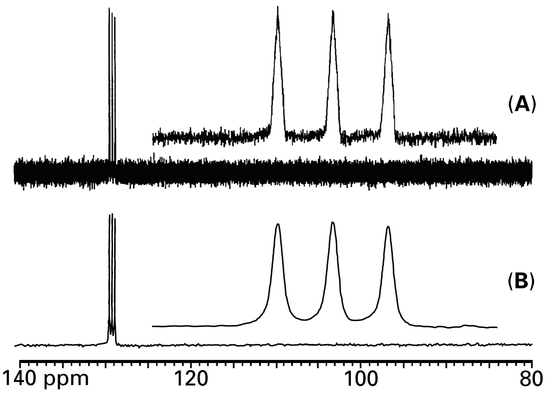



**Figure 2** 75.4 MHz spectra of the $^{13}$C triplet of deuteriobenzene in the ASTM (American Society for Testing and Materials) sensitivity test sample (60% deuteriobenzene/40% dioxane), with and without matched filtration. The unweighted spectrum (A) shows a signal-to-noise ratio of 18:1; the same data given an exponential multiplication with a time constant 1/(3π) s before Fourier transformation, corresponding to a 3 Hz Lorentzian line broadening, show (B) a signal-to-noise ratio of 92:1. An acquisition time of 5.462 s was used, with a spectral width of 12 000 Hz and one zero filling. The insets show expansions of the triplet signal, illustrating the broadening of the lines and the smoothing of the noise caused by the exponential multiplication.

enhancement. $W(t)$ can also be written as a time-shifted Gaussian

$$W(t) = \exp(+t_s^2/t_g^2)\exp[-(t-t_s)^2/t_g^2] \qquad [3]$$

where $t_s = t_g^2/(2t_e)$. If $t_e$ is equal to the decay constant of the experimental NMR signal, then multiplication by $W(t)$ before Fourier transformation converts a Lorentzian lineshape of width $1/(\pi t_e)$ Hz to a Gaussian of width $(2/\pi t_g)\sqrt{\log_e 2}$ Hz. Since spectra normally contain a range of line widths, it is usually necessary to experiment with $t_e$ and $t_g$ to find the best values for a given region of a spectrum. Because instrumental effects such as field inhomogeneity make experimental line shapes non-Lorentzian, resolution enhancement is best combined with reference deconvolution. **Figure 3** shows the application of Lorentz–Gauss resolution enhancement to a proton multiplet.

Even where neither sensitivity nor resolution enhancement is sought, time-domain weighting is desirable where some NMR signal survives at the end of the sampled data. Such a truncated dataset is equivalent to the full, untruncated signal multiplied by a window function; the convolution theorem shows that the resultant spectrum will contain the true line shapes convoluted by a 'sinc' [sin($x$)/$x$] function, giving rise to 'wiggles' on either sides of lines. Applying a weighting function $W(t)$, which brings the time-domain data smoothly to zero ('apodization'), can reduce or suppress such undesirable artefacts.

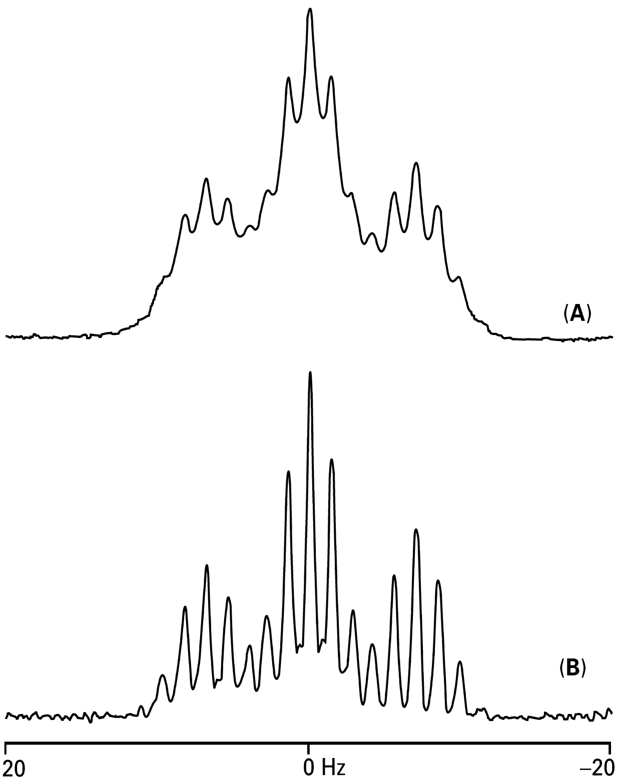



**Figure 3** Expansions of the multiplet at 5.1 ppm in the 400 MHz proton spectrum of geraniol in deuteriomethanol: (A) raw spectrum; and (B) spectrum after Lorentz–Gauss conversion using rising exponential weighting with a time constant of $1/\pi$ s and Gaussian weighting with a time constant of 1 s.

Most NMR spectra are presented in phase-sensitive mode, but this is not appropriate where the phases of signals vary rapidly or unpredictably with position, as in some magnetic resonance imaging and multidimensional NMR experiments. The modulus of a complex Lorentzian line shape shows a very broad base because of the contribution from the (imaginary) dispersion mode component. This can be suppressed if time-domain weighting is used to force the experimental signal into a form that is time-symmetric, for example using the function $W(t)$ above with a small $t_e$ ('pseudo-echo' weighting) or using a half sine-wave ('sine-bell' weighting). Although it is common to arrange for such weighting to leave the maximum of the weighted signal at the midpoint of the experimental data, this is neither necessary nor always desirable. Absolute value mode presentation is the norm where phase cycling or pulsed field gradients are used to produce signal phase modulated as a function of $t_1$ in 2D NMR; it is to be avoided where possible because overlapping peaks are distorted by interference between their dispersion mode parts.

## Fourier transformation

The classical frequency domain spectrum $S(\nu)$ is the Fourier transform of the FID $s(t)$:

$$S(\nu) = \int_0^{t_a} s(t)\exp(-2\pi i\nu t)\,dt \qquad [4]$$

where the integration limits reflect the fact that the FID starts at time zero and is recorded for a time $t_a$. Practical spectrometers use digital technology, so the FID $s(t)$ is digitized at regular intervals $\Delta t$ to give a time series of $M$ points $s_k = s[(k-1)\Delta t]$, where $(M-1)\Delta t = t_a$, and a discrete Fourier transform (DFT) is carried out using the Cooley–Tukey FFT algorithm. The DFT of a time series of $N$ complex points with spacing $\Delta t$ generates a frequency spectrum which is a series of $N$ complex points with spacing $1/(N\Delta t)$ Hz:

$$S_n = \frac{1}{\sqrt{N}}\sum_0^{N-1} s_k \exp(-2\pi ikn/N) \qquad [5]$$

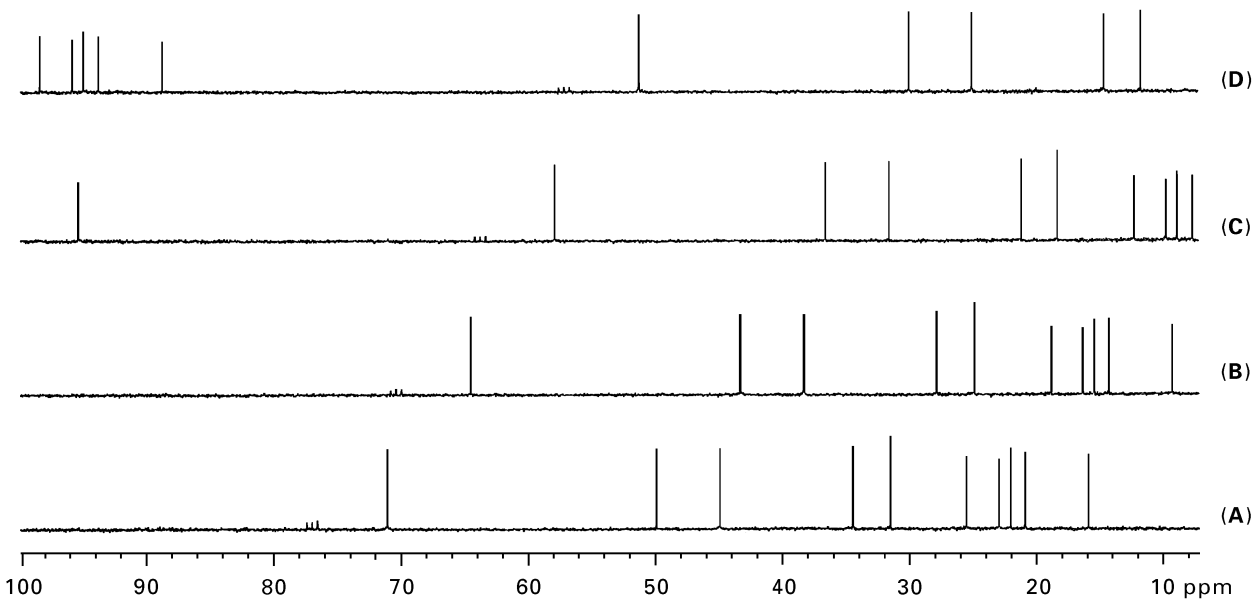



**Figure 4** 75.4 MHz proton-decoupled $^{13}C$ spectra of 30% menthol in deuteriochloroform, (A) recorded with all signals within the spectral window, and (B)–(D) with the transmitter displaced to high field in 500 Hz steps. Spectra (C) and (D) show the aliasing of the high-field signals to reappear at the low-field end of the spectrum.

where the frequency of the $n$th point is $(n-1)/(N\Delta t)$ Hz. Both the continuous and the discrete Fourier transform can use several different sign and normalization conventions; those given here are widely used, but others are equally valid. Spectrometers are almost invariably restricted by the FFT to Fourier transforming numbers of points $N$ which are powers of 2.

Discrete sampling in the time domain introduces an ambiguity into the frequency domain: signals at frequencies separated by multiples of $1/(\Delta t)$ Hz are indistinguishable, since their relative phases only change by multiples of $2\pi$ between sampling points. NMR signals which lie outside the range 0 to $(N-1)/(N\Delta t)$ will be 'aliased' by adding or subtracting the spectral width $1/\Delta t$ until they lie within this window, as seen in **Figure 4**. Since the signals that emerge from the spectrometer receiver may have positive or negative frequencies, the output of the DFT needs to be rotated by $N/2$ points $[1/(2\Delta t)$ Hz$]$ so that the digitized spectrum covers the range $-1/(2\Delta t)$ to $(N/2-1)/(N\Delta t)$ Hz. By convention, the resultant spectrum is plotted with frequency increasing from right to left, so that the most shielded nuclei (those with lowest chemical shift) lie at the right. Some older spectrometers sample the real and imaginary receiver channels alternately rather than simultaneously, using a real rather than a complex Fourier transform; signals outside the spectral width then 'fold' back into the spectrum by reflection about the frequency limits $\pm 1/(2\Delta t)$.

## Phasing

The measurement of NMR data must wait until the radiofrequency pulse and its after-effects have died away, which can take several tens of microseconds; analogue or digital filtration also delays the arrival of the NMR signal at the receiver output. The effect on an NMR signal of a delay time $\delta$ is to add a phase shift $\exp(2\pi i\nu\delta)$ to a signal of frequency $\nu$, causing the signal phase to vary linearly across the spectrum. In addition, the relative phases of the receiver reference signals and the transmitter pulse are arbitrary, so both a zero- and a first-order phase correction are needed to bring all signals into absorption mode. The phased spectrum $S_p(\nu)$ can be written

$$S_p(\nu) = S(\nu)\exp[-i(\phi_0 + \phi_1\nu\Delta t)] \qquad [6]$$

where the zeroth-order and first-order phase shifts $\phi_0$ and $\phi_1$ are normally determined either automatically, or by the spectrometer operator using an interactive display. **Figure 5** shows a typical spectrum before and after phasing.

First-order phase correction has one insidious effect, baseline distortion. A frequency-dependent phase shift cannot make up for the data that were lost during $\delta$; the baseline error is just the DFT of the missing data. However, provided $\delta$ is small compared to $\Delta t$, this baseline curvature can easily be corrected during postprocessing of the spectrum. In 2D

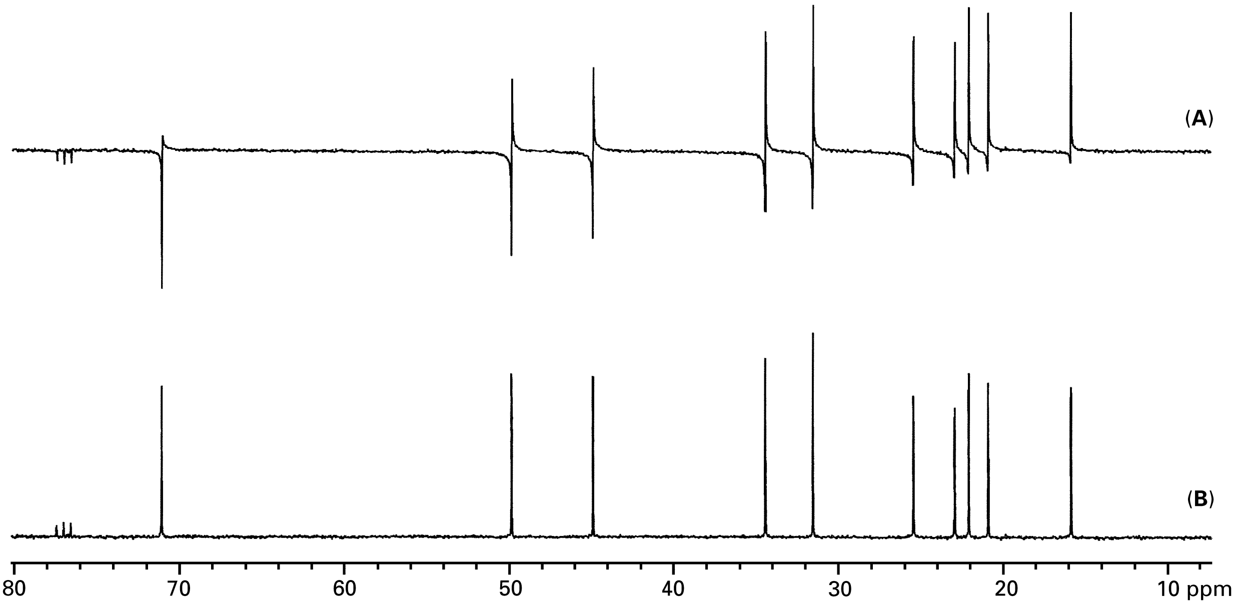



**Figure 5** 75.4 Mhz proton -decoupled 13c spectra of 30% menthol in deuteriochloroform, (A)b before and (B) after zero- and first-order phase correction.

NMR, there will be different time delays $\delta_1$ and $\delta_2$ in the two time dimensions; phasing again is normally carried out using an interactive display.

## Linear prediction

A FID and its Fourier transform contain exactly the same information, but sometimes this is insufficient to give a readily interpretable spectrum. Where there is adequate internal evidence within the FID, it may be possible to extrapolate the NMR signals forwards and/or backwards in time to synthesize missing data and hence create a time-domain signal that transforms to a clearer spectrum. The two commonest uses of such an extrapolation are backwards in time to replace data lost during the time $\delta$, and forwards in time to improve resolution. A typical digitized experimental FID contains a series of $n$ exponentially damped, complex signals, plus a background of random noise. The NMR signal can be written:

$$s_k = \sum_{j=1}^{n} \alpha_j \beta_j^k \quad [7]$$

where the complex number $\alpha_j = A_j \exp(i\phi_j)$ defines the phase $\phi_j$ and amplitude $A_j$ of the $j$th signal, and $\beta_j = \exp(2\pi i \Delta t \nu_j) \exp(-\Delta t / T_j)$ is determined by the frequency $\nu_j$ and decay constant $T_j$. The contribution made by component $j$ to point $s_k$ is just the contribution to $s_{k-1}$ multiplied by $\beta_j$. Linear prediction (LP) algorithms take a FID of $M$ points and fit this time series with a set of $m$ complex coefficients $a_j$

$$s_k = \sum_{j=1}^{m} a_j s_{k-j} \quad [8]$$

so that point $k$ is expressed as a linear combination of the previous $m$ points (forward prediction), or of the subsequent $m$ points (backward prediction)

$$s_k = \sum_{j=1}^{m} a_j s_{k+j} \quad [9]$$

A variety of algorithms exist for finding the coefficients $a_j$ and multipliers $\beta_j$, with which the experimental data can be extrapolated forwards or backwards; all share some common problems. Linear prediction has difficulty distinguishing between positive and negative decay constants $T_j$, and so is best suited to time series in which all the decay constants are either positive or negative, allowing spurious $\beta$ values to be rejected. The number $m$ of coefficients $a_j$ to be used has to be decided by the ex-

perimenter: too few, and peaks will be missed; too many and noise will be treated as signal. Non-Lorentzian line shapes make exponential damping a poor approximation, increasing the number of coefficients needed.

Linear prediction is (despite its name) a nonlinear method and can produce very misleading results, but with care it can greatly ease interpretation of poorly digitized spectra. It is particularly useful in 2D NMR, where signals are routinely truncated in the $t_1$ dimension.

## Maximum entropy reconstruction

Although linear prediction can be used to extract spectral data directly from a FID ('parametric LP'), it is commonly used to extrapolate the experimental time-domain data, which are then weighted and transformed as normal. Maximum entropy reconstruction, in contrast, seeks to fit the experimental FID with a model function that contains the minimum amount of information consistent with fitting experiment to within the estimated noise level. The criterion of minimum information corresponds to the maximum Shannon informational entropy $S(p)$, which for a probability distribution $p$ is defined as

$$S(p) = -\sum_n p_n \log_e p_n \qquad [10]$$

Maximum entropy methods have generated considerable controversy; they have been described (a little unkindly) as generating more heat than light. Their results can show spectacular improvements in signal-to-noise ratio, but this should not be confused with sensitivity of detection. Maximum entropy methods successfully pick out those signals that are above a defined threshold, but miss those below it; the signal amplitude estimates produced are comparable to those obtainable by simply fitting a model line shape to the Fourier transform spectrum. Thus for well-sampled experimental data the advantages of maximum entropy methods are largely cosmetic, and come at a high computational cost. Where such methods can be very valuable is with data sets that are damaged, incomplete, not sampled at uniform time intervals, or require deconvolution.

## Postprocessing techniques

The simplest and most widely used form of post-processing is integration of the signal intensity. The integral of a resonance is proportional to the number of spins contributing to it; thus the relative numbers of nuclei in different chemical groupings can be found by comparing the integrals of their signals under appropriate experimental conditions. Accurate integration almost always requires operator intervention to carry out baseline correction, varying according to need from simple offset and slope correction through to the subtraction of a baseline calculated by spline or polynomial fit to operator-defined regions of empty baseline.

A second common example of postprocessing is the listing of signal heights and positions ('peak picking'), for the measurement of chemical shifts and coupling constants, or as the first step in the extraction of parameters such as relaxation times, rate constants or diffusion coefficients. In principle, the integral of a signal should give a better estimate of signal amplitude than peak height, being independent of line shape; in practice, baseline errors and signal overlap mean that peak height measurements are usually preferred for intensity comparisons between corresponding signals in different spectra.

Where signals overlap, neither integration nor peak picking gives accurate signal intensities. Here iterative fitting can be used to decompose the experimental spectrum into contributions from individual lines, typically assumed to have Lorentzian or Gaussian shapes; the positions, amplitudes and widths of the theoretical line shapes are varied to minimize the sum of the squares of the differences between the experimental and calculated spectra. Such least-squares fitting is easily perturbed by instrumental distortion of the line shapes, for example as a result of static field inhomogeneity, so the best results require either great care with shimming or some form of compensation for instrumental line shape contributions.

One effective way to compensate for many instrumental sources of error is reference deconvolution. Since most instrumental errors (e.g. static field inhomogeneity and magnetic field instability) affect all signals equally, multiplying the experimental FID by the complex ratio of the theoretical and experimental signals for a reference resonance leads to a spectrum in which such errors have been corrected. This technique can be used to ensure that all lines are basically Lorentzian, and also to enforce strict comparability between different spectra in a series, for example to correct $t_1$ noise in multidimensional NMR.

Many other data processing techniques are used to extract useful information from experimental NMR spectra. Signal intensities compiled by peak picking may be fitted to an exponential or Gaussian function, as in the determination of relaxation times and

in pulsed field gradient spin-echo measurements of diffusion coefficients. The latter experiment can be extended to the construction of a pseudo-2D spectrum in which signals are dispersed according to chemical shift in one dimension and diffusion coefficient in the other ('diffusion-ordered spectroscopy'). In many spectra the extraction of chemical shift and coupling constant values is hindered by second-order effects ('strong coupling'); the analysis of strongly coupled spectra is most effectively carried out using quantum-mechanical simulation. This can be partially automated in favourable cases, the experimental spectrum being used as a target for least-squares fitting in which the variable parameters are the chemical shifts and coupling constants rather than simply the signal positions, amplitudes and widths. The extraction of kinetic parameters from exchange-broadened band shapes ('line shape analysis') can be similarly automated.

## List of symbols

$a$, $b$ = time-domain functions; $A$, $B$ = frequency-domain structures; $F$, $\nu$ = frequency; $i = \sqrt{-1}$; $m$ = number of complex coefficients; $M$, $N$ = number of complex points; $p$ = probability distribution; $s_k$ = $k$th point in free induction decay; $S_n$ = $n$th point in spectrum; $S(p)$ = Shannon informational entropy; $S_p(\nu)$ = phased spectrum; $s(t)$ = free induction decay; $t_1$ = evolution time; $t_2$ = real time; $t_e$ = exponential time constant; $t_g$ = Gaussian time constant; $t_s$ = time shift of Gaussian; $T$ = decay constants; $W$ = line width; $W(t)$ = weighting function; $\alpha_j$ = $j$th complex amplitude; $\beta_j$ = $j$th complex exponential; $\delta$ = delay time; $\phi$ = phase shift.

*See also:* **Fourier Transformation and Sampling Theory; Laboratory Information Management Systems (LIMS); NMR Principles; NMR Pulse Sequences; NMR Spectrometers; Two-Dimensional NMR, Methods.**

## Further reading

Freeman R (1997) *Spin Choreography*. Oxford, UK: Spectrum.

Freeman R (1987) *A Handbook of Nuclear Magnetic Resonance (2/e)*. Harlow, UK: Longman.

Hoch JC and Stern AS (1996) *NMR Data Processing*. New York: Wiley-Liss.

Lindon JC and Ferrige AG (1980) Digitisation and data processing in Fourier transform NMR. *Progress in NMR Spectroscopy* **14**: 27–66.

Morris GA, Barjat H and Horne TJ (1997) Reference deconvolution methods. *Progress in NMR Spectroscopy* **31**: 197–257.

Rutledge DN (ed) (1996) Signal treatment and signal analysis in NMR. *In*: Vol 18 of *Data Handling in Science and Technology*. Amsterdam: Elsevier.

# NMR in Anisotropic Systems, Theory

**JW Emsley**, University of Southampton, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Theory**

The simplest anisotropic system is a single crystal and in this case the spin interactions depend upon the orientation of the molecules with respect to the direction of the applied magnetic field, $B_0$, of the spectrometer. The anisotropies of the various interactions that affect the spectrum can be measured by changing the orientation of the crystal in the field. The spin interactions in polycrystalline or noncrystalline solids are still anisotropic, but now the spectra are invariant to sample orientation. Molecular motion can have a dramatic effect on the NMR spectra observed for anisotropic systems, and characterizing such motion is one of the principal uses of this spectroscopy.

The most complex anisotropic systems are the various liquid crystalline phases; now the molecular motion is similar to that in a normal, isotropic liquid in that there is rapid rotation and diffusion of the molecules. These phases differ from isotropic phases in that the motion is not random, and this has a profound affect on their NMR spectra.

The spectra produced by anisotropic systems are more complex than those from isotropic phases, but they contain more information. However, the spectral complexity is often too great for all the information to be obtained, and for solid samples it is usually preferable to dissolve the material in order to simplify the spectrum. Various other methods have

been developed that enable some, if not all, of the information to be extracted from the NMR spectrum of an anisotropic sample, and these are described elsewhere. The aim of this article is to show what can, in principle, be obtained by a successful analysis of the NMR spectra obtained from solid and liquid crystalline samples.

## General considerations

We will consider only experiments that use a strong magnetic field, that is those using mainstream spectrometers in which $B_0$ is greater than about one tesla. In this case, for nuclei with spin $I = \frac{1}{2}$, the largest magnetic interaction is that between $\mu$, the magnetic dipole moment of the nucleus, and the field, the Zeeman interaction. This leads to the field direction being, to a good approximation, a unique axis of quantization for the nuclear spins; and the experiments yield only the component, $T_{B,q}$, of the $q$th interaction along $\boldsymbol{B}_0$, and so these are the parameters that may be obtained from the spectra. For nuclei with $I > \frac{1}{2}$ the interaction of the nuclear electric quadrupole moment, $eQ_N$, of a nucleus of isotopic species denoted by $N$ with the electric field gradient, $\boldsymbol{V}_i$, at the site $i$ may be larger than the Zeeman interaction, and the spins will no longer be quantized along $\boldsymbol{B}_0$. We will not deal with this situation, and the interested reader is referred to the article on the theory of nuclear quadrupole resonance. The present discussion is therefore confined to spin $\frac{1}{2}$ nuclei, or to those with small quadrupole moments, or at sites of small field gradients, such that the Zeeman term still dominates the nuclear spin Hamiltonian.

For each of the spin interactions that affect the spectrum of an anisotropic sample, the value of $T_{B,q}$ may have both a part, $T_{B,q}$(isotropic), that is independent of the orientation with respect to $B_0$ of the molecule containing the spin, and a part, $T_{B,q}$(anisotropic), that is orientation dependent. Thus,

$$T_{B,q} = T_{B,q}(\text{isotropic}) + T_{B,q}(\text{anisotropic}) \qquad [1]$$

Four interactions may affect the spectrum of an anisotropic sample. The shielding (or chemical shift), $\sigma_i$, of the nucleus at a site $i$ always has a finite value of $T_{B,q}$(isotropic), $\sigma_B^0$, in an anisotropic sample. There will be a nonvanishing contribution $T_{B,q}(\text{anisotropic}) = \sigma_B^a$ provided that the symmetry of the site $i$ is lower than tetrahedral. The electron-mediated spin–spin coupling, $J_{ij}$, between a pair of nuclei in an anisotropic sample will also in general have two nonvanishing contributions, $J_{ij}^0$ and $J_{ij}^a$. Both terms are negligible when the two nuclei are in different molecules. For nuclei within one molecule, the anisotropic term ranges from being negligible to being very large depending on the nuclear isotopes involved. The dipolar spin–spin coupling, $D_{ij}$, has only an anisotropic term, as does the quadrupolar coupling, $eQ_NV_i$. We will return to the individual anisotropic interactions after discussing the features they have in common.

### Relationship between spectral parameters $T_{B,q}$ and molecular properties

To understand what information is provided on properties of the molecules, we need to know how to relate $T_{B,q}$ to components $T_{\alpha\beta,q}$ of the $q$th interaction in a reference frame fixed within a molecule. This relationship is a general one since all four interactions are second-rank tensors, and so:

$$T_{B,q} = \sum\nolimits_{\alpha,\beta} T_{\alpha\beta,q} \cos\theta_\alpha \cos\theta_\beta \qquad [2]$$

$T_{\alpha\beta,q}$ is a component in the ($xyz$) molecular axes, and $\theta_\alpha$ is the angle $\boldsymbol{B}_0$ makes with the axis $\alpha$. There are nine components of $T_{\alpha\beta,q}$:

$$\begin{matrix} T_{xx,q} & T_{xy,q} & T_{xz,q} \\ T_{yx,q} & T_{yy,q} & T_{yz,q} \\ T_{zx,q} & T_{zy,q} & T_{zz,q} \end{matrix}$$

The sum $T_{xx,q} + T_{yy,q} + T_{zz,q}$ is independent of the choice of the axes. For some systems $T_{\alpha\beta,q} \neq T_{\beta\alpha,q}$ (for a discussion see the paper of Smith, Palke and Gerig listed in Further reading), but the effect of this asymmetry can usually be neglected, and in this case there is a special set of axes ($abc$) such that only the diagonal elements are nonzero. These are known as the principal axes of the interaction tensor and the $T_{aa,q}$, $T_{bb,q}$ and $T_{cc,q}$ are principal components of $\boldsymbol{T}_q$. It is important to note that the different interactions will not usually share a common set of principal axes. In principal axes, Equation [2] is simplified to

$$T_{B,q} = T_{aa,q}\cos^2\theta_a + T_{bb,q}\cos^2\theta_b + T_{cc,q}\cos^2\theta_c \qquad [3]$$

There are obvious advantages in using principal axes, but their location in a molecule is not always known. This does not prevent us from working in principal axes when doing algebraic manipulations on Equation [2]. Thus, Equation [3] might seem inconsistent with Equation [1], but that this is not so

can easily be appreciated by replacing $\cos^2\theta_\alpha$ by $(1-\sin^2\theta_\alpha)$ to give

$$T_{B,q} = T_{aa,q} + T_{bb,q} + T_{cc,q} - T_{aa,q}\sin^2\theta_a - T_{bb,q}\sin^2\theta_b - T_{cc,q}\sin^2\theta_c \qquad [4]$$

There is, however, a more useful way of expressing $T_{B,q}$ as a scalar plus an anisotropic part, and this is derived by noting that when the molecules move rapidly and randomly, as in an isotropic liquid or gas, $T_{B,q}$ is averaged (denoted by the brackets $\langle\ \rangle$) to produce $\langle T_{B,q}\rangle_{\text{iso}}$:

$$\langle T_{B,q}\rangle_{\text{iso}} = \sum_{\alpha,\beta} T_{\alpha\beta,q}\langle\cos\theta_\alpha\cos\theta_\beta\rangle \qquad [5]$$

Using principal axes, and noting that $\langle\cos^2\theta_\alpha\rangle = \frac{1}{3}$

$$\langle T_{B,q}\rangle_{\text{iso}} = \tfrac{1}{3}(T_{aa,q} + T_{bb,q} + T_{cc,q}) = T_q^0 \qquad [6]$$

Remembering that the sum of the diagonal elements of $\boldsymbol{T}_q$ is independent of the axes, then we can see that Equation [6] is true for all choices of molecular axes even though we derived it in principal axes. Rearranging Equation [3] as

$$T_{B,q} = T_q^0 + \sum_\alpha T_{\alpha\alpha,q}(\cos^2\theta_\alpha - \tfrac{1}{3}) \qquad [7]$$

which leads to

$$T_{B,q} = T_q^0 + \tfrac{2}{3}\sum_\alpha T_{\alpha\alpha,q}(3\cos^2\theta_\alpha - 1)/2 \qquad [8]$$

Rearranging gives

$$T_{B,q} = T_q^0 + \tfrac{2}{3}[T_{aa,q} - \tfrac{1}{2}(T_{bb,q} + T_{cc,q})](3\cos^2\theta_a - 1)/2 + \tfrac{1}{2}(T_{bb,q} - T_{cc,q})(\cos^2\theta_b - \cos^2\theta_c) \qquad [9]$$

Equation [9] has the advantage of showing clearly what happens if the tensor $\boldsymbol{T}_q$ is axially symmetric in the molecular principal axes, that is, when $T_{bb,q} = T_{cc,q}$.

## Effect of molecular motion on the interactions in anisotropic systems

We have seen already that rapid, random motion averages $T_{B,q}$(anisotropic) to zero, but what is the effect of motion that is rapid but not random? The definition of rapid is that the rate of the motion is much greater than the magnitude of $T_{B,q}$(anisotropic). Slower motion will contribute to the relaxation rates and is not discussed here. The effect of motion is to produce $\langle T_{B,q}\rangle$, an averaged value of $T_{B,q}$, which can be thought of as a sum of contributions, $T_{B,q}(n)$ from $n$ configurations of the system, each with a normalized probability $P(n)$:

$$\langle T_{B,q}\rangle = \sum_n P(n)[T_q^0(n)(\text{isotropic}) + T_{B,q}(n)(\text{anisotropic})] \qquad [10]$$

For most systems the isotropic contribution will not have a strong dependence on $n$, and so we will concentrate on the averaging of the anisotropic term.

For solids it is better to consider the different kinds of motion that occur, and to see how they affect the averages.

**Rotation of whole, rigid molecules on lattice sites in crystals** Rotation about an $n$-fold axis, or hopping between $n$ equivalent sites, produces an averaged tensor, one of whose principal axes, $r$, lies along the rotation axis. The location of the other two principal axes, $s$ and $t$, will be in the plane orthogonal to $r$, and they are fixed with respect to the lattice and not the molecules. Thus,

$$\langle T_{B,q}\rangle = \langle T_q^0\rangle + \tfrac{2}{3}[T_{rr,q} - \tfrac{1}{2}(T_{ss,q} + T_{tt,q})](3\cos^2\theta_r - 1)/2 + \tfrac{1}{2}(T_{ss,q} - T_{tt,q})(\cos^2\theta_s - \cos^2\theta_t) \qquad [11]$$

The relationship between $T_{rr,q}$ and the components of $\boldsymbol{T}_q$ in the $(abc)$ principal frame is given by

$$T_{rr,q} = T_{aa,q}\cos^2\theta_{ar} + T_{bb,q}\cos^2\theta_{br} + T_{cc,q}\cos^2\theta_{cr} \qquad [12]$$

The angles $\theta_{ar}$, $\theta_{br}$ and $\theta_{cr}$ are between the axes $a$, $b$, $c$ and $r$. For two-site hopping, the location of $s$ and $t$ will depend upon the structure of the molecule.

When $n \geq 3$ the averaged tensor is axially symmetric about $r$, so that Equation [11] simplifies to

$$\langle T_{B,q}\rangle = \langle T_q^0\rangle + \tfrac{2}{3}[T_{rr,q} - \tfrac{1}{2}(T_{ss,q} + T_{tt,q})](3\cos^2\theta_r - 1)/2 \qquad [13]$$

and $T_{rr,q}$ and $\frac{1}{2}(T_{ss,q}+T_{tt,q})$ can be replaced by $T_{\|r,q}$ and $T_{\perp r,q}$.

These relationships apply to all the spin interactions when the molecule rotates as a whole because in this case the internuclear distances within a molecule do not change.

**Rotation of a group of nuclei within a molecule** The values of $\langle T_{B,q}\rangle$ for interactions involving nuclei only in the group that is moving relative to the crystal lattice (and hence relative to $\boldsymbol{B}_0$) are affected in the same way as described for whole-molecule rotation except that now the axes *s* and *t* are fixed in the nonrotating part of the molecule. The spin–spin interactions between nuclei in the rigid and moving parts of the molecules will be averaged in a way that depends on the structure of the molecules.

**Hopping (diffusion) of molecules between lattice sites** This kind of motion will usually occur in combination either with whole-molecule rotation on a lattice site or with internal motion in the molecules. The diffusion process will have a large effect on dipolar interactions between nuclei in different molecules. The exact magnitude of the effect will depend on the crystal structure, but it will always lead to a reduction in the intermolecular contribution to dipolar coupling.

**Molecular motion in liquid crystalline samples** In liquid crystalline samples the molecules rotate and diffuse rapidly but not randomly. The diffusive motion means that the distance between nuclei in different molecules is changing, with the result that the intermolecular contribution to the dipolar coupling is zero. The averaging produced by the rotational motion of whole, rigid molecules depends to some extent on the nature of the phase and the effect on it of the magnetic field of the spectrometer. We will consider here the simplest case only, which is when there is axial symmetry about $\boldsymbol{B}_0$. With uniaxial symmetry the result for a rigid molecule is to average the angular factors in Equation [9], which is usually expressed in slightly different way; thus

$$T_{B,q} = T_q^0 + \tfrac{2}{3}[T_{aa,q} - \tfrac{1}{2}(T_{bb,q} + T_{cc,q})]S_{aa}^B + \tfrac{1}{3}(T_{bb,q} - T_{cc,q})(S_{bb}^B - S_{cc}^B) \qquad [14]$$

where

$$S_{\alpha\alpha}^B = \langle 3\cos^2\theta_\alpha - 1)/2 \rangle \qquad [15]$$

$\theta_\alpha$ is the angle between $\boldsymbol{B}_0$ and axis $\alpha$. Note that now both the $T_{\alpha\alpha,q}$ and the $S_{\alpha\alpha}^B$ are unknown parameters that may be obtained from an analysis of the NMR spectra, and that their absolute values cannot be obtained separately. In practice, for chemical shifts and quadrupolar coupling it is usually possible to use values of the $T_{\alpha\alpha,q}$ obtained from the spectrum of a single-crystal sample to determine the $S_{\alpha\alpha}^B$. For dipolar coupling it is necessary to assume just one internuclear distance in order to obtain the separate values of the $S_{\alpha\alpha}^B$ and the $D_{ij}$.

The relationship between the values of the $S_{\alpha\alpha}^B$ and the orientational order of the molecules in the liquid crystalline phase can be derived by first introducing the concept of a director, $\boldsymbol{n}_i$. This is a unit vector at point *i* in the sample which defines the preferred orientation of the anisotropically shaped molecules. This can be quantified by a function, $P_{LC}(\beta,\gamma)$ which describes the probability that the director makes angles between $\beta$ and $\beta + d\beta$ and $\gamma$ and $\gamma + d\gamma$ in a frame (*abc*) fixed in the molecules, as shown in **Figure 1**. This probability density function has a maximum when $\beta$ is zero, whose magnitude grows towards a value of unity as the molecules become ordered by reducing temperature. The values of $\langle T_{B,q}\rangle$ can be thought of as arising from rapid motion of the molecules about the director giving an average, $\langle T_{n,q}\rangle$, of the $\boldsymbol{T}_q$ along $\boldsymbol{n}$, and then using the relationship

$$\langle T_{B,q}\rangle = \langle T_{n,q}\rangle (3\cos^2\theta_{Bn} - 1)/2 \qquad [16]$$

where $\theta_{Bn}$ is the angle between $\boldsymbol{n}_i$ and $\boldsymbol{B}$. If the directors are distributed relative to $\boldsymbol{B}_0$ then the spectrum will be a sum of spectra and will be like that from a polycrystalline powder. The averages $\langle T_{n,q}\rangle$ are given by

$$\langle T_{n,q}\rangle = T_q^0 + \tfrac{2}{3}[T_{aa,q} - \tfrac{1}{2}(T_{bb,q} + T_{cc,q})]S_{aa} + \tfrac{1}{3}(T_{bb,q} - T_{cc,q})(S_{bb} - S_{cc}) \qquad [17]$$

where the $S_{\alpha\alpha}$ are known as Saupe order parameters. They are defined as

$$S_{aa} = \int \tfrac{1}{2}(3\cos\beta - 1)P_{LC}(\beta,\gamma)\sin\beta\, d\beta \qquad [18]$$

$$S_{bb} - S_{cc} = \tfrac{3}{2}\int \sin^2\beta\cos 2\gamma P_{LC}(\beta,\gamma)\sin\beta\, d\beta\, d\gamma \qquad [19]$$

The important point to note is that they can act as tests for models of $P_{LC}(\beta,\gamma)$. This function can be

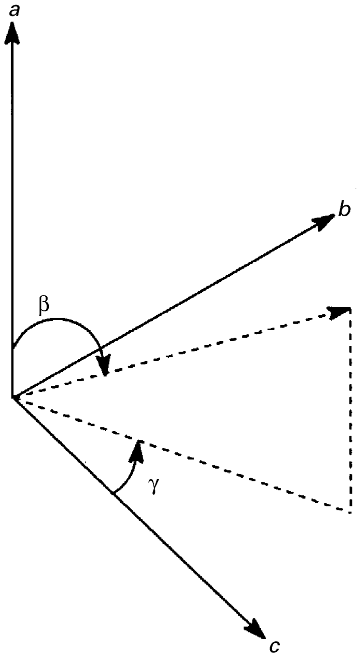



**Figure 1** Angles defining the orientation of the magnetic field, $\boldsymbol{B}_0$, in axes (*abc*) fixed in a molecule.

used to define an effective orienting potential, or potential of mean torque, $U_{\text{ext}}(\beta, \gamma)$, thus

$$P_{\text{LC}}(\beta, \gamma) = Z^{-1} \exp[-U_{\text{ext}}(\beta, \gamma)/k_{\text{B}}T] \quad [20]$$

with

$$Z = \int \exp[-U_{\text{ext}}(\beta, \gamma)/k_{\text{B}}T] \sin\beta \, \mathrm{d}\beta \, \mathrm{d}\gamma \quad [21]$$

where $T$ = temperature, $k_{\text{B}}$ = Boltzmann constant.

In the absence of a magnetic field, the directors vary in orientation relative to some space-fixed axes in a completely random way. The directors may, however, be aligned uniformly by a magnetic field. If the molecules comprising the liquid crystal phase have a positive anisotropy, $\Delta\chi$, in their magnetic susceptibility, then the directors align parallel to the field, so that $\theta_{Bn}$ is 0, and $\langle T_{B,q} \rangle = \langle T_{n,q} \rangle$. If $\Delta\chi$ is negative the directors align in a plane perpendicular to the field, making $\theta_{Bn} = 90°$ and $\langle T_{B,q} \rangle = -\frac{1}{2}\langle T_{n,q} \rangle$. In both cases the spectra are like those for a single crystal except the lines are much narrower because of the absence of broadening from intermolecular effects.

**The effect of internal motion on $\langle T_{n,q} \rangle$ for liquid crystalline samples** Fast internal motion produces an additional averaging of the $T_{n,q}$, in a similar way to the case of solid samples, but with the very important difference that there is a cooperative effect of changes in molecular shape and orientational order. For example, consider rotation about a bond through an angle $\phi$. The probability distribution becomes dependent on $\phi$ as well as $\beta$ and $\gamma$, and so too does the mean potential energy of the molecules in the phase. Thus, we revise Equations [20] and [21] to

$$P_{\text{LC}}(\beta, \gamma, \phi) = Z^{-1} \exp[-U(\beta, \gamma, \phi)/k_{\text{B}}T] \quad [22]$$

$$Z = \int \exp[-U(\beta, \gamma, \phi)/k_{\text{B}}T] \sin\beta \, \mathrm{d}\beta \, \mathrm{d}\gamma \, \mathrm{d}\phi \quad [23]$$

The mean potential may be divided into a purely anisotropic part, $U_{\text{ext}}(\beta, \gamma, \phi)$, which vanishes in the isotropic phase, and a part, $U_{\text{int}}(\phi)$, which does not:

$$U(\beta, \gamma, \phi) = U_{\text{ext}}(\beta, \gamma, \phi) + U_{\text{int}}(\phi) \quad [24]$$

Because the potential of mean torque has become dependent on the internal motion, then so too do the order parameters; that is, Equations [18] and [19] become

$$S_{aa}(\phi) = \int \tfrac{1}{2}(3\cos\beta - 1) P_{\text{LC}}(\beta, \gamma, \phi) \sin\beta \, \mathrm{d}\beta \quad [25]$$

$$S_{bb}(\phi) - S_{cc}(\phi) = \tfrac{3}{2} \int \sin^2\beta \cos 2\gamma \, P_{\text{LC}}(\beta, \gamma, \phi) \sin\beta \, \mathrm{d}\beta \, \mathrm{d}\gamma \quad [26]$$

Various strategies have been suggested for obtaining $P_{\text{LC}}(\beta, \gamma, \phi)$ from observed values of $\langle T_{B,q} \rangle$.

Having obtained $P_{\text{LC}}(\beta, \gamma, \phi)$, it is then possible to determine $P_{\text{LC}}(\phi)$, the probability that the molecule in the liquid crystalline phase is in a conformation defined by the angle $\phi$, as

$$P_{\text{LC}}(\phi) = \int P_{\text{LC}}(\beta, \gamma, \phi) \sin\beta \, \mathrm{d}\beta \, \mathrm{d}\gamma = \frac{\int \exp[-U(\beta, \gamma, \phi)/k_{\text{B}}T] \sin\beta \, \mathrm{d}\beta \, \mathrm{d}\gamma}{\int \exp[-U(\beta, \gamma, \phi)/k_{\text{B}}T] \sin\beta \, \mathrm{d}\beta \, \mathrm{d}\gamma \, \mathrm{d}\phi} \quad [27]$$

Note that in principle this differs from $P_{\text{iso}}(\phi)$, the

conformational probability distribution for the molecule in the isotropic phase, which is

$$P_{\text{iso}}(\phi) = \frac{\exp[-U_{\text{iso}}(\phi)/k_{\text{B}}T]}{\int \exp[-U_{\text{iso}}(\phi)/k_{\text{B}}T]\mathrm{d}\phi} \qquad [28]$$

## The anisotropic nuclear spin interactions

As noted earlier, we will confine our discussion to the case when the interaction of the spins with the magnetic field, the Zeeman interaction, is dominant so that the direction of the applied, static field determines the axis of quantization and the experiments yield the component $T_{B,q}$ of the various interactions in this direction. It is now convenient to define this as the $Z$ direction. A very detailed discussion of the spin interactions, and their contribution to the nuclear spin Hamiltonian has been given by Smith, Palke and Gerig.

### The Zeeman interaction

The contribution, $\mathcal{H}_{\text{Zeeman}}$, to the nuclear spin Hamiltonian, in units of hertz, has the form:

$$\mathcal{H}_{\text{Zeeman}} = -\frac{1}{2\pi}\sum_i \gamma_i(1 - \sigma_{ZZi})B_0 I_{Zi} \qquad [29]$$

where $\gamma_i$ is the magnetogyric ratio of the $i$th nucleus, and $I_{Zi}$ is the component along $Z$ of the nuclear spin angular momentum operator, $\boldsymbol{I}_i$. The important feature is that this Hamiltonian has an identical form in both isotropic and anisotropic systems, and hence the spectral consequences are exactly the same. That is, the resonance is shifted in energy depending on the electronic environment at the $i$th site. The difference between these two kinds of system lies in the contributions to $\sigma_{ZZ}$ in the two situations. Thus for a rigid molecule Equation [9] becomes

$$\sigma_{ZZ} = \sigma^0 + \tfrac{2}{3}[\sigma_{aai} - \tfrac{1}{2}(\sigma_{bbi} + \sigma_{cci})](3\cos^2\theta_{ia} - 1)/2 + \tfrac{1}{2}(\sigma_{bbi} - \sigma_{cci})(\cos^2\theta_{bi} - \cos^2\theta_{ci}) \qquad [30]$$

The isotropic term, $\sigma^0$, is much greater than the anisotropic contribution and so that there is only a small (~kHz) difference between the resonance frequency of a nucleus in an isotropic and anisotropic environment.

Note that the location of the principal axes ($abc$) depends on the site occupied in a molecule by the nucleus. The site symmetry in a rigid molecule determines whether $\sigma_{aai} = \sigma_{bbi} = \sigma_{cci}$ (isotropic site symmetry), $\sigma_{aai} \neq \sigma_{bbi} = \sigma_{cci}$ (axial site symmetry), or $\sigma_{aai} \neq \sigma_{bbi} \neq \sigma_{cci}$ (site asymmetry). The nature of the site symmetry is in fact revealed most clearly and easily by recording the spectra of polycrystalline samples, if possible. This is particularly true when the Zeeman term is the only one determining the observed spectrum. This situation can easily be realized, for example for $^{13}$C, $^{15}$N or $^{31}$P nuclei.

**Figure 2** shows the shape of spectra obtained for nuclei at sites of the three different symmetries.

### Electron-mediated spin–spin coupling

The contribution to the Hamiltonian, in hertz, is

$$\mathcal{H}_J = \sum_{i<j}\{J_{ij}^0\,[I_{Zi}I_{Zj} + \tfrac{1}{2}(I_{+i}I_{-j} + I_{-i}I_{+j})] + J_{ij}^a\,[I_{Zi}I_{Zj} - \tfrac{1}{4}(I_{+i}I_{-j} + I_{-i}I_{+j})]\} \qquad [31]$$

$J_{ij}^0$ ($\equiv T_{B,q}$(isotropic) in Equation [1] is the scalar contribution to coupling, which is the only contribution in isotropic environments. $J_{ij}^a$ is the anisotropic contribution, and its components in the principal, molecular frame can be obtained by substitution in Equation [9] for a rigid crystal, and Equation [14] for a rigid molecule in a liquid crystalline phase.

The operators $I_{+i}$ and $I_{-i}$ are defined as:

$$I_{+i} = I_{Xi} + \mathrm{i}I_{Yi} \quad \text{and} \quad I_{-i} = I_{Xi} - \mathrm{i}I_{Yi} \qquad [32]$$

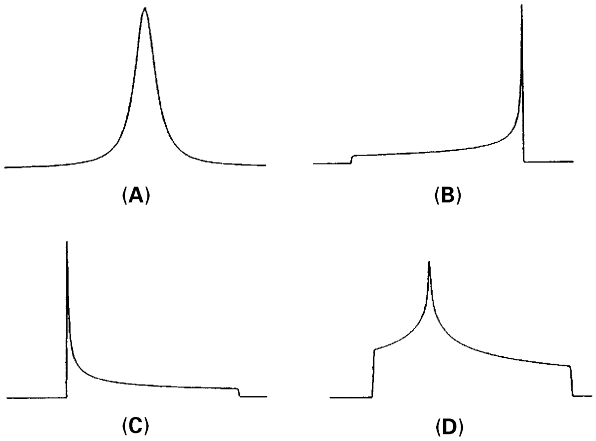


**Figure 2** Spectra for a polycrystalline sample for nuclei with different site symmetries and subject only to the Zeeman interaction: (A) $\sigma_{aa} = \sigma_{bb} = \sigma_{cc}$; (B) $\sigma_{aa} = \sigma_{bb} > \sigma_{cc}$; (C) $\sigma_{aa} = \sigma_{bb} < \sigma_{cc}$; (D) $\sigma_{aa} \neq \sigma_{bb} \neq \sigma_{cc}$.

The magnitude of $J^a_{ij}$ is predicted to be much smaller than either $J^0_{ij}$ or $D_{ij}$, the dipolar coupling between the same nuclei, when one of the nuclei is a proton. For other pairs of nuclei, $J^a_{ij}$ may not be negligible. Values have been obtained experimentally from the spectra of molecules dissolved in liquid crystalline phases.

### Dipolar coupling

This is the through-space coupling between a pair of nuclear magnetic dipoles. It is a purely anisotropic interaction and is also symmetric about $\boldsymbol{r}_{ij}$, the internuclear vector. It contributes a term to the Hamiltonian, in hertz, of

$$\mathcal{H}_D = \sum_{i<j} 2D_{ij}[I_{zi}I_{zj} - \tfrac{1}{4}(I_{+i}I_{-j} + I_{-i}I_{+j})] \qquad [33]$$

Note that this has identical spin operators to those involving $J^a_{ij}$ in Equation [31]. This means that the spectra of anisotropic systems depend on $(2D_{ij} + J^a_{ij})$, and that these two interactions cannot be determined separately from the spectra. For this reason, $J^a_{ij}$ is often referred to a pseudo-dipolar coupling.

Dipolar coupling is axially-symmetric about $\boldsymbol{r}_{ij}$, which is a principal axis, so that from Equation [9] there is just one contribution to $D_{ij}$:

$$D_{ij} = D_{aaij}(3\cos^2\theta_{ij} - 1)/2 \qquad [34]$$

$\theta_{ij}$ is the angle between $\boldsymbol{r}_{ij}$ and $\boldsymbol{B}_0$. $D_{aaij}$ for a fixed internuclear distance is given, in hertz, by

$$D_{aaij} = -[\mu_0 h\gamma_i\gamma_j/(16\pi^3)]/r_{ij}{}^3 \qquad [35]$$

where $\mu_0 = 4\pi\times 10^{-7}$ and is the permeability of free space. This makes it possible to determine molecular structure from dipolar couplings.

### The quadrupolar interaction

Remembering that we are restricting our discussion to the cases when the Zeeman term determines the axis of quantization of the nuclear spins, then the quadrupolar interaction contributes a term, in hertz, to the spin Hamiltonian of

$$\mathcal{H}_Q = \sum_i [eQ_N V_{ZZi}/(4hI_{Ni}(2I_{Ni} - 1)] \times [3I_{Zi}{}^2 - I_{Ni}(I_{Ni} + 1)] \qquad [36]$$

$I_{Ni}$ is the nuclear spin quantum number of spin $i$ of type $N$, which has quadrupolar moment $eQ_N$, and is at a site with an electric field gradient $\boldsymbol{V}_i$. It is a purely anisotropic interaction, and so for a rigid molecule in a solid crystalline sample,

$$V_{ZZi} = V_{aai}(3\cos^2\theta_a - 1)/2 + \tfrac{1}{2}(V_{bbi} - V_{cci})(\cos^2\theta_b - \cos^2\theta_c) \qquad [37]$$

## List of symbols

$a$, $b$, $c$ = molecule-fixed principal axes; $\boldsymbol{B}_0$ = applied magnetic field, $B_0 = |\boldsymbol{B}_0|$; $D_{ij}$ = dipolar spin–spin coupling; $h$ = Planck constant; $\mathcal{H}$ = Hamiltonian; $I$ = nuclear spin quantum number; $\boldsymbol{I}_i$ = spin angular momentum operator; $I_{Zi}$ = component of $\boldsymbol{I}_i$ along $Z$; $J_{ij}$ = electron-mediated spin–spin coupling; $k_B$ = Boltzmann constant; $\boldsymbol{n}_i$ = director (unit vector at point $i$); $P_{LC}(\beta,\gamma)$ = probability that director angle is in $\beta$+d$\beta$, $\gamma$+d$\gamma$; $P(n)$ = probability of configuration $n$; $Q_N$ = nuclear quadrupole moment; $r$, $s$, $t$ = lattice-fixed principal axes; $\boldsymbol{r}_{ij}$ = internuclear vector; $S_{\alpha\alpha}$ = Saupe order parameters; $T$ = temperature; $T_{B,q}$ = component of $q$th interaction along $\boldsymbol{B}_0$; $\mathbf{T}_q$ = interaction tensor; $T_{\alpha\beta,q}$ = component of $q$th interaction with respect to molecule fixed axes $\alpha$ and $\beta$; $U_{ext}$ = potential of mean torque; $\boldsymbol{V}_i$ = electric field gradient at site $i$; $x$, $y$, $z$ = molecular axes; $\gamma_i$ = magnetogyric ratio of $i$th nucleus; $\theta_i$ = angle of $\boldsymbol{B}_0$ with axis $i$; $\mu$ = nuclear magnetic dipole moment; $\mu_0$ = permeability of free space; $\sigma$ = shielding constant; $\phi$ = angle of rotation about bond; $\chi$ = magnetic susceptibility.

*See also:* **Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Diffusion Studied Using NMR Spectroscopy; Liquid Crystals and Liquid Crystal Solutions Studied By NMR; Solid State NMR, Methods; Solid State NMR Using Quadrupolar Nuclei; Solid State NMR, Rotational Resonance.**

## Further reading

Emsley JW (ed) (1985) *NMR of Liquid Crystals*, Dordrecht: Reidel.

Schmidt-Rohr K and Spiess HW (1994) *Multidimensional Solid State NMR and Polymers*. New York: Academic Press.

Smith SA, Palke WE and Gerig JT (1992) *Concepts in Magnetic Resonance* **4**: 107; (1992) **4**: 181; (1993) **5**: 151; (1994) **6**: 137.

# NMR Microscopy

**Paul T Callaghan**, Massey University, Palmerston North, New Zealand

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Methods & Instrumentation**

## Introduction

A microscope is generally regarded as an assembly of lenses used to give an image of small objects at high spatial resolution. By high spatial resolution is meant a resolution finer than can be resolved by the naked human eye, in other words better than 0.1 mm. The well-known ‘optical microscope’ uses either reflected or transmitted light to present an image and the resolution of this instrument is determined by the wavelength of the electromagnetic radiation, around 0.5 μm. There are many other forms of microscopy that use different radiations, such as the electron microscope, whose resolution is much finer because of the shorter wavelength of the electron beam, or the acoustic microscope which is very effective at probing near-surface properties using high-frequency sound waves. X-ray microscopes have been developed in recent years, although these are rather tricky to use because of the need to generate relatively monochromatic soft X-rays that can then be focused by Fresnel lenses. This article describes a very different microscope, based on the use of radio waves, the so-called ‘nuclear magnetic resonance microscope’.

If one were to present the concept of a radio wave microscope in the context of our usual perspectives on such devices, three factors would emerge. First, one would expect that, like the X-ray microscope, the device would have excellent powers of penetration, since radio waves pass easily through most matter, excluding good conductors such as metals. Second, one would expect that it would be almost ideally noninvasive, because of the very low energy of the radiofrequency photon. In this regard it could be contrasted with electron microscopy and X-ray microscopy, which are both capable of breaking covalent bonds between atoms. Finally, it might be guessed that such a device would have hopelessly poor spatial resolution because of the very long wavelengths (several metres!) of the radiation.

This last point would be certain to render radio wave microscopy useless were it not for the phenomenon of nuclear precession exhibited by atomic nuclei with nonzero spin when placed in a magnetic field. This property means that those nuclei (the ‘spins’) have associated with them a special radio frequency that depends precisely on the local magnetic field strength, and if that magnetic field strength is varied from place to place in the sample, then the measurement of the frequency of each nucleus will indicate where its parent atom or molecule is positioned. By this means one could build up an image of the atomic distribution, avoiding altogether the normal wavelength (or ‘diffraction’) limit to resolution. The technique used to measure the spatially dependent spin precession frequency is nuclear magnetic resonance (NMR).

The prospects for such a method seem almost too good to be true, but there is an Achilles heel. NMR relies on the measurement of the nuclear frequency by means of the exchange of radiofrequency photons whose energies are so weak that they must compete with the thermal noise that exists in the detection circuitry. It is this latter effect that determines the ultimate resolution of such an instrument, and, using the best possible stable nucleus (the proton) in thermal equilibrium at room temperature, and with a receiving antenna made from room-temperature metals, the limit is volume elements around $(10\ \mu m)^3$ for the image, an exceptionally poor resolution by comparison with all other microscopies. Indeed the NMR microscope is only just a microscope in the normal sense and one would hardly bother with it at all were it not for a few remarkably useful properties. It is exceptionally noninvasive as intimated earlier, and this makes it of special interest in *in vivo* biological applications. It is highly penetrating and can work perfectly well with optically opaque materials with minimal problems of ‘transparency’, and it is especially well suited to the study of liquid phases, a rather unusual property in the context of other microscopies. Most important of all, it enables the imaging of a number of molecular and atomic properties to which NMR is particularly suited. These include the local chemical composition, the local molecular order and rotational dynamics, and the local molecular translational motions. It is for these reasons that NMR microscopy, despite its rather poor spatial resolution, has found a number of important applications in science, technology and medicine.

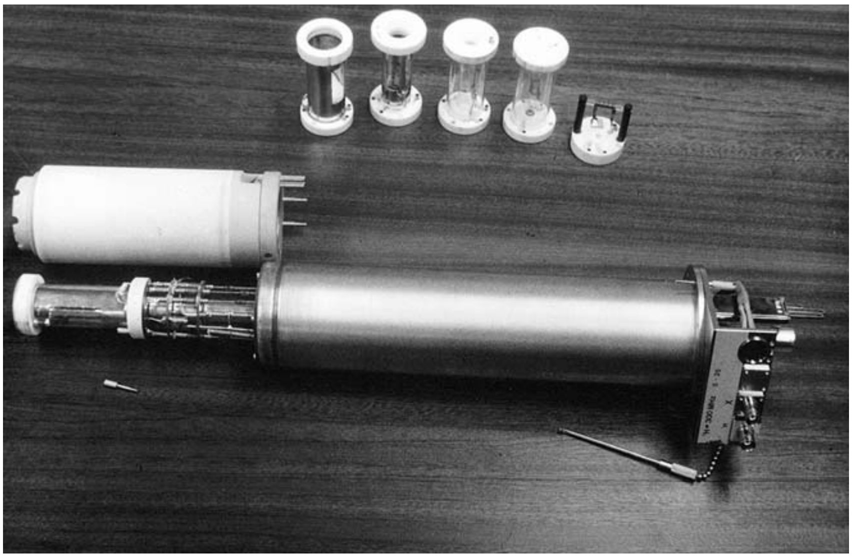


**Figure 1** Radiofrequency probe, gradient coils set and RF coil inserts used in NMR microscopy. The probe assembly is placed in the 89 mm diameter vertical bore of a superconducting magnet. The set of RF coils and resonators enable samples of different sizes to be inserted, and have diameters ranging from 25 mm diameters down to 2 mm. Photograph courtesy of Bruker Analytische Messtechnik, Karlsruhe, Germany.

## Nuclear magnetic resonance imaging

### Proton NMR

Nuclear magnetic resonance was first detected in 1945, although its extension to imaging applications was to wait until 1973. NMR is an enormous field of research and the subject of very many textbooks in its own right. This article will attempt to give only a cursory introduction to the phenomenon, and the reader should look to some of the articles elsewhere in this Encyclopedia for further information. An isolated proton, when placed in a magnetic field, $B_0$, occupies one of two quantum states with respect to the field direction and the quantum phases of those states rotate at the frequency $\omega = \gamma B_0$ (for example $\omega/2\pi$ = 300 MHz for a proton immersed in a 7 T field), where $\gamma$ is the gyromagnetic ratio of the proton, the factor that determines the ratio of its magnetic properties to its spin properties. This phase rotation arises from the combined magnetic and spin properties of the nucleus and is known as precession. The proton has a very high value of $\gamma$ (and hence radiofrequency photon energy) relative to other nuclei and it is highly abundant in most materials, as the nucleus of atomic hydrogen. It is thus the prime candidate for NMR microscopy.

In practice we deal with large ensembles of nuclei whose phases, in thermal equilibrium, are randomized so that the net magnetization presented by the sample is directed along $B_0$, and results from the slight preponderance of spin-up over spin-down (typically around 1 in $10^5$). Detection of the underlying precession frequency therefore requires the intervention of a small, transverse, oscillatory magnetic field, $B_1$ (the so-called resonant radiofrequency field, generated by a coil surrounding the sample). This field is applied as a pulse for a finite duration $t_p$ at the same frequency as the underlying precessional circulation, thus keeping its effect on the spins in step with their motion. As a result the spins gradually re-align themselves with respect to the combined effects of $B_0$ and $B_1$, the result being a reorientation of the net magnetization vector to an angle $\gamma B_1 t_p$ with respect to $B_0$. A '90°' RF pulse would be one that left the magnetization precessing in the transverse plane. The same sample coil is used to excite the spin system from equilibrium (relaxation time $T_1$). Thus the NMR spectrometer consists of a magnetic field, a coil as antenna, and a radio transceiver. In the case of a microimaging system, an additional requirement is a set of magnetic field gradient coils. Typical NMR microscope RF coils and gradient coils are shown in **Figure 1**.

### Position encoding using field gradients

The imaging principle is as follows. Suppose we now apply, in addition, a magnetic field gradient $\boldsymbol{G} = \nabla B_0$ (this can be achieved with specialized coil designs capable of generating quite uniform gradients along three orthogonal axes), then the precession frequency will depend on spin location and can be written

$$\omega = \gamma B_0 + \gamma \boldsymbol{G} \cdot \boldsymbol{r} \qquad [1]$$

Because the receiver uses heterodyne phase-sensitive detection with reference frequency $\gamma B_0$, the additional term, $\gamma \boldsymbol{G}\cdot\boldsymbol{r}$, shows up as a difference oscillation, generally in the audiofrequency part of the spectrum. As a consequence, the heterodyne signal detected via the NMR coil at time *t* after the spins have precessed in the gradient has the mathematical form

$$S(\boldsymbol{k}) = \int \rho(\boldsymbol{r}) \exp(\mathrm{i}2\pi \boldsymbol{k}\cdot\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \qquad [2]$$

where $\exp(\mathrm{i}2\pi\boldsymbol{k}\cdot\boldsymbol{r})$ is the local spin phase factor for the spins at position $\boldsymbol{r}$, $\boldsymbol{k} = (2\pi)^{-1}\gamma \boldsymbol{G}t$, and $\rho(\boldsymbol{r})$ is the density of spins at position $\boldsymbol{r}$ and is the quantity we seek to image. The remarkable fact about Equation [2] is that it is a simple Fourier transform in which $\boldsymbol{k}$ is the reciprocal-space wave vector conjugate to the spatial dimension $\boldsymbol{r}$. Thus the acquisition of a signal in a domain of '$\boldsymbol{k}$-space' leads to direct computation of $\rho(\boldsymbol{r})$ by means of Fourier inversion, a fact that is made possible by the ability of the NMR experiment to provide full phase information (i.e. the complex number $S(\boldsymbol{k})$ via the acquisition of both in-phase (real) and quadrature-phase (imaginary) signals.

Generally one acquires an image of the spin distribution in a 2-dimensional planar 'slice', with a frequency-selective RF pulse being used to excite only spins within a preselected plane, and the $\boldsymbol{k}$-space encoding being applied independently for two orthogonal directions within the plane, one direction via a fixed time duration variable-magnitude gradient pulse (phase-encoding) and the other direction via a fixed gradient pulse with the signal being sampled at successive time points during the evolution (read-encoding). A typical pulse sequence that performs these tasks along with the corresponding $\boldsymbol{k}$-space map, is shown in **Figure 2**. Note that this sequence employs the spin-echo method in which a separate 180° RF pulse is used to invert spin phases. This allows us to traverse both negative and positive regions of $\boldsymbol{k}$-space. The spin echo has important applications in the measurement of transverse relaxation, and of molecular flow and diffusion.

Note also that the finite time needed to encode the NMR magnetization in $\boldsymbol{k}$-space requires that the nuclear spin relaxation time, $T_2$, be sufficiently long. $T_2$ is determined predominantly by internuclear interactions, which in turn are motionally averaged by molecular tumbling. This means in effect that molecules in the liquid state have long proton $T_2$ values (10–1000 ms) while those in the solid state may have relaxation times as short as 10 μs. This results in a sharp discrimination of the signal in which the solid state component is entirely filtered out, unless special

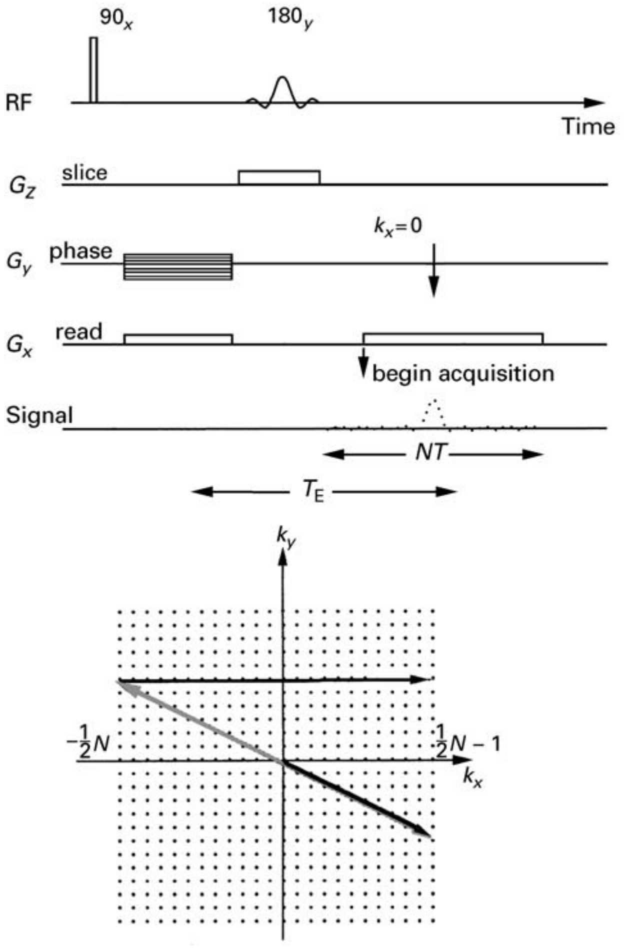



**Figure 2** RF and magnetic field gradient pulse sequence used in NMR microscopy along with the associated trajectory through $\boldsymbol{k}$-space. The frequency-selective 180° pulse is used to select a layer of spins (the slice plane) to participate in the spin echo and hence to contribute to the image. $NT$=acquisition time; $T_{\mathrm{E}}$=*i* echo time.

rapid-encoding methods are used. Most NMR microscopy images therefore arise from the liquid phase alone. Readers interested in the rarer solid-state options should consult references given under Further reading.

## Limits to resolution

The fundamental spatial resolution for NMR microscopy is limited by the three factors of intrinsic signal-to-noise, molecular self-diffusion and diamagnetic susceptibility effects. For protons at room temperature and with an RF coil at room temperature, the signal-to-noise limit is easily calculated under optimal experimental conditions and depends directly on the ratio of $T_1$ to $T_2$, the available experimental time for signal averaging, and the magnetic field strength. Typical resolution limits for

superconducting magnet systems for imaging times of the order of 30 min are (10 μm)$^3$ and (20 μm)$^3$ for $T_1/T_2$ values of unity and 100, respectively. These estimates assume 256 × 256 pixels with the sample exactly filling a solenoidal RF coil, a consequence of which is that the RF coil dimensions must progressively decrease with increasing resolution, thus limiting the highest resolution to very small samples. Note that resolution in one dimension can be traded against another so that it is possible to observe finer details (in-plane resolution) if a thicker slice can be tolerated. One of the highest-resolution examples yet reported, at (4.5 μm)$^2$ in-plane and 70 μm slice thickness, was obtained using a sample of onion cells contained with an RF coil of 1 mm diameter (see **Figure 3**).

In principle, the resolution could be further improved if the receiver coil temperature and/or its electrical resistance could be lowered. Some progress has been reported on the use of liquid nitrogen-cooled superconducting ceramics in which an effective signal-to-noise increase was obtained. The need for inductive coupling and the problem of poor filling factor means in effect that no major advantage has resulted to date.

The fact that the NMR imaging signal arises from nuclei whose parent molecules exist in a liquid state implies that imaging resolution will be strictly limited by self-diffusion. A rule of thumb is that diffusion limits will be important when the rms Brownian displacements over the ***k***-space encoding time are comparable with the pixel dimension. The usual consequence is severe signal attenuation, an effect that is apparent in **Figure 4**, which shows an image from water contained in a rectangular capillary of 100 μm wall spacing. Here only the layer of molecules near the wall, whose motions are impeded by the boundary, fails to suffer attenuation and therefore presents a bright edge effect. Such phenomena can provide potentially useful contrast. Another effect that also has its origin in boundaries associated with structural heterogeneity is illustrated in **Figure 5** which shows high-resolution NMR-microscope images of a section of geranium stem at two different read gradient strengths for which the acquisition times, $NT$, and echo time $T_E$ (see **Figure 2**) are different by a factor of 5. The image obtained at longer acquisition and echo time (**Figure 5A**) suffers from attenuation and distortion effects due to the differing diamagnetic susceptibility across cell wall boundaries. Note that the pixel dimension is 10 μm while the slice thickness is 500 μm. The subtle interplay between diffusion and susceptibility effects can also result in characteristic bright image features seen in **Figure 5A**. It is apparent that the susceptibility and diffusion effects can provide image features that are useful in discriminating boundaries between fluid regions. These complex phenomena and their elucidation are discussed in more detail in references listed at the end of the article.

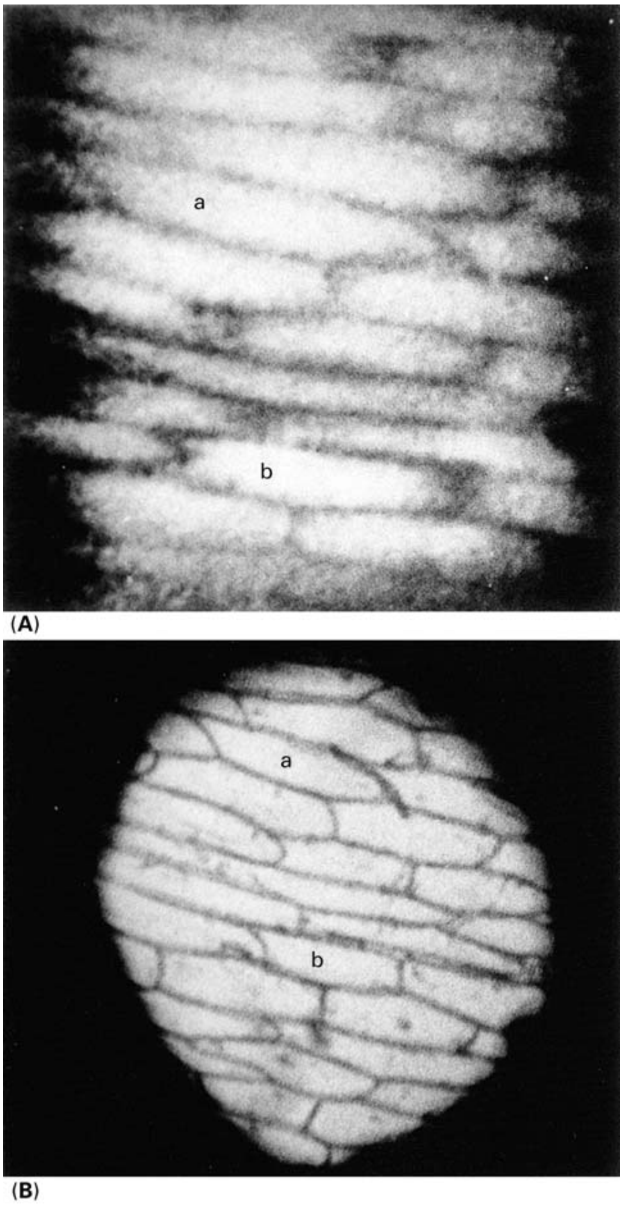



**Figure 3** NMR image (A) and corresponding optical micrograph (B) of the epidermal cells of *Alium cepa*. The pixel size is 5 μm with a slice thickness of 70 μm. Reprinted with permission of the International Society of Magnetic Resonance in Medicine. From Glover PM, Bowtell RW, Brown GD and Mansfield P (1994) *Magnetic Resonance in Medicine* **31**: 423–428.

## Image contrast

The particular utility of NMR microscopy lies in the contrasts that are available. These include, in

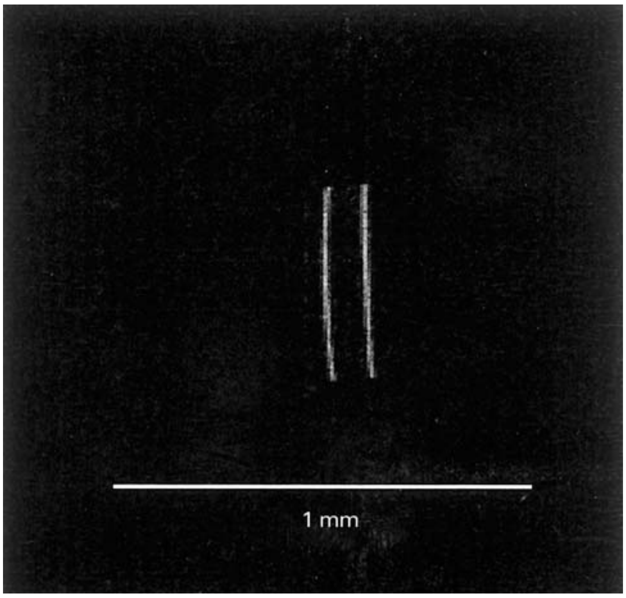



**Figure 4** NMR micrograph obtained from water in a rectangular glass capillary whose walls are spaced by 100 μm. The pixel dimension is 8 μm, comparable with the distance diffused by the water molecules over the echo time $T_E$, and hence the image intensity is severely attenuated, except at the walls where the molecular Brownian motion is restricted. Courtesy of S.L. Codd and the author.

addition to the usual spin density (or, by implication, molecular density) the susceptibility effects discussed above, the chemical shift (the small changes in frequency due to different electronic environments of the nuclei), the nuclear spin relaxation times (sensitive to the rotational dynamics of parent molecules and hence ideally suited to discriminating solids, liquids and semisolid phases), the molecular translational self-diffusion coefficient and the molecular translational flow rate. Each contrast, if it is to be appropriately quantified or enhanced, requires a particular modification to the basic imaging pulse sequence shown in **Figure 2**, details of which can be found in the book by Callaghan. For example, in the case of chemical shift imaging, the initial 90° RF pulse can be made frequency-selective so as to excite only those spins residing in a particular chemical site. For relaxation contrast the echo time can be altered by adjusting delay times or pulse repetition times. However, because the imaging pulse sequence is based on the use of magnetic field gradients in conjunction with spin echoes, this can make the image intensity especially sensitive to the effects of molecular self-diffusion when one is close to the diffusion limit. This is the reason for the bright edge effects seen in **Figure 4**.

A commonly used approach is to leave the parameters of the imaging spatial encoding (the part to the right of the grey line in **Figure 2**) unchanged and to add the pulse sequence needed for contrast encoding as a precursor. For example, in the case of

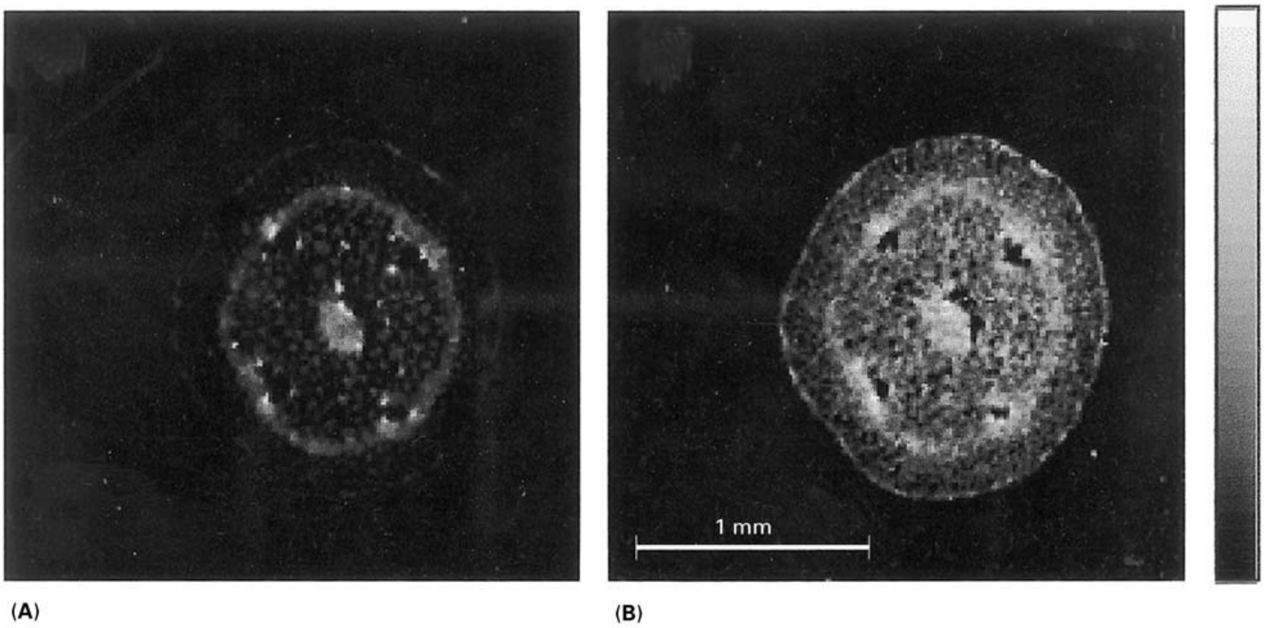



**Figure 5** Two images ($256^2$ pixels of $(10\ \mu m)^2$, slice thickness 500 μm) obtained with the spin-echo pulse sequence of **Figure 2** from a geranium stem in which different read gradients and echo times are employed: (A) 20 kHz with $T_E = 13.5$ ms; (B) 100 kHz with $T_E = 3.2$ ms. Reprinted with permission from Rofe CJ, Van Noort J, Back PJ and Callaghan PT (1995) *Journal of Magnetic Resonance B* **108**: 125–136.

$T_2$ contrast this would take the form of a prior $90_x$–$\tau$–$180_y$–$\tau$ spin-echo segment where the remaining echo amplitude used for spatial encoding depends on the relaxation that occurs over the delay time $\tau$. By acquiring separate images at different delay times $\tau$, one generates a three-dimensional data set (two for the image within the slice plane and one additional for $\tau$). Analysis in the third dimension enables one to display an image of relaxation rate. Another higher-dimensional encoding is that needed for molecular self-diffusion or flow analysis. This method is so important to NMR microscopy that it requires a separate discussion.

## Pulsed gradient spin-echo NMR

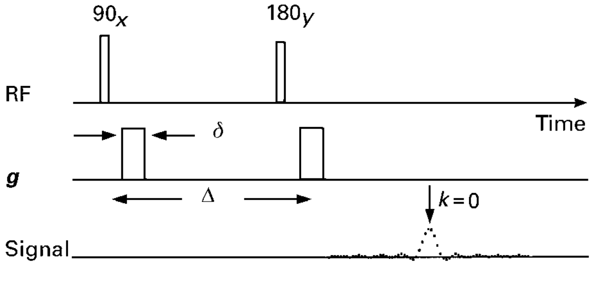
### Encoding for translational motion

The spin echo allows one to perform a useful trick in the encoding of spin positions. Consider the pulse sequence shown in **Figure 6**. At the echo maximum (the point where $k = 0$), the phase excursions of the spins return to zero, unless, that is, any of the parent molecules happen to have moved along the direction of the gradient. In this case the subtraction performed by the echo is imperfect and the residual phase shifts provide a signature for motion.

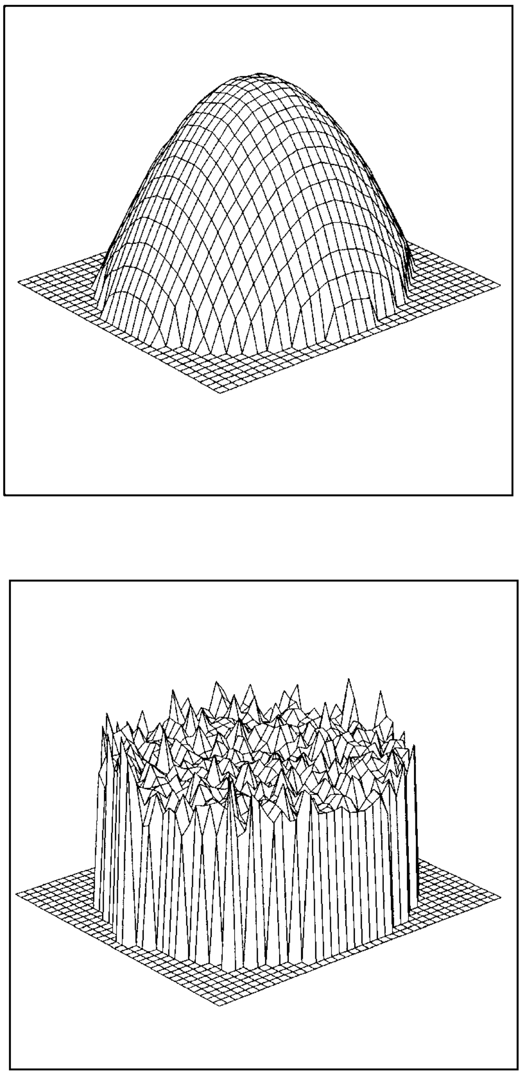
The pulsed gradient spin echo (PGSE) experiment of **Figure 6**, uses two narrow gradient pulses of amplitude $g$, duration $\delta$ and separation $\Delta$. These pulses effectively define the starting and finishing point of spin translational motion over the well-defined timescale, $\Delta$. A spin that moves by a distance $\boldsymbol{R}$ over time $\Delta$ will acquire a phase shift, $\gamma\delta\boldsymbol{g}\cdot\boldsymbol{R}$. Indeed the resulting phase encoding can be expressed in the same reciprocal space language that we have already seen; the difference now is that the wave vector, $q$ ($q = (2\pi)^{-1}\gamma g\delta$), is conjugate to molecular displacements over the time $\Delta$ between the gradient pulses, rather than in the case of imaging where it is conjugate to the actual spin positions. The analogue to Equation [2] is given by the echo amplitude,

$$E(\boldsymbol{q}, \Delta) = \int \overline{P_s}(\boldsymbol{R}, \Delta) \exp[\mathrm{i}2\pi\boldsymbol{q}\cdot\boldsymbol{R}]\, \mathrm{d}\boldsymbol{R} \qquad [3]$$

where the ‘average propagator’, $\overline{P_s}(\boldsymbol{R}, \Delta)$, gives the probability that a spin in the ensemble examined displaces by $\boldsymbol{R}$ over the encoding time $\Delta$. Just as inverse Fourier transformation of $S(\boldsymbol{k})$ in Equation [2] returns an image of the spin density $\rho(\boldsymbol{r})$, so inverse Fourier transformation of $E(q, \Delta)$ with respect to $q$ returns an image of the average propagator, $\overline{P_s}$.




(A)



(B)

**Figure 6** (A) Pulsed gradient spin echo sequence used to encode spin magnetization phase for molecular translational motion. (B) Velocity and diffusion maps for a water molecule flowing through a 2 mm diameter capillary. The images are shown as stackplots. The velocity profile is Poiseuille while the diffusion map is uniform. Courtesy of RW Mair, MM Britton and the author.

### Diffusion and flow

In the case of two simple examples of motion of importance in NMR microscopy, namely self-diffusion and flow, the propagator takes the form

$$\overline{P_s}(\boldsymbol{R}, \Delta) = (4\pi D_s\Delta)^{-3/2} \exp\left(-\frac{(\boldsymbol{R} - \boldsymbol{v}\Delta)^2}{4D_s\Delta}\right) \quad [4]$$

where $\boldsymbol{v}$ is the constant velocity and $D_s$ is the self-diffusion coefficient. When the PGSE encoding is used as a higher contrast dimension in conjunction with NMR imaging, the signal acquired is effectively modulated both in $\boldsymbol{k}$-space and $\boldsymbol{q}$-space. Generally the experiment is performed with two dimensions of $\boldsymbol{k}$ (the slice planes) and one dimension of $q$ (where $q = |\boldsymbol{q}|$), so that the signal is

$$\int(\boldsymbol{k}, q) = \int \rho(\boldsymbol{r}) E(q, \Delta) \exp(i2\pi \boldsymbol{k}\cdot\boldsymbol{r}) \, d\boldsymbol{r} \quad [5]$$

Inverse Fourier transformation of $S(\boldsymbol{k},q)$ with respect to $\boldsymbol{k}$ returns a set of 2-dimensional images modulated by $E(q, \Delta)$ while transformation with respect to $q$ returns, for every image pixel, $\overline{P_s}(Z, \Delta)$, the average propagator for displacements $Z$ along the gradient axis at position $\boldsymbol{r}$ within the image. By appropriate processing of these average propagators, details of the local motion can be calculated. For example, the width of $\overline{P_s}(Z, \Delta)$ is determined by the rms Brownian motion $(2D_s\Delta)^{1/2}$ whilst the displacement of $\overline{P_s}(Z, \Delta)$ along the $Z$ axis is determined by the flow displacement $v\Delta$ where $v$ is the local molecular velocity. In this manner maps of $D_s(\boldsymbol{r})$and $v(\boldsymbol{r})$ may be constructed, examples of which are shown in **Figure 6B**.

## Applications of NMR microscopy

Numerous factors militate against the widespread use of NMR microscopy: the resolution is poor by optical standards, the apparatus is expensive, the technique requires a high level of scientific expertise and the arrangements for sample loading are inconvenient and restrictive. Set against these are the uniquely noninvasive character of the method, its sensitivity to fluid phases, its unique ability to measure specific molecular properties and the especially powerful insights it can provide regarding fluid dynamics. Studies in which NMR microscopy has been able to provide unrivalled information include those concerned with membrane filtration, flow and dispersion in porous media, non-Newtonian flow in viscoelastic fluids, nonequilibrium phase transitions, electrophoresis and electroosmosis, interdiffusion of fluids, and the monitoring of chemical wave propagation. In biology the method has provided new insights into plant physiology *in vivo*, multicellular tumour development and angiogenesis in tumours, while the method enables detailed MRI investigations concerning rat and mouse physiology in the study of diseases and their treatment. It should be noted that while the majority of applications of the method have concerned $^1$H NMR, there are a number of important examples of its applications using other nuclei, including $^{19}$F, $^{31}$P, $^{13}$C, $^2$H and $^{17}$O. These rarer nuclei allow site- or molecular-specific labelling, and provide the opportunity to investigate new contrast schemes, for example the mapping of pH in the case of $^{31}$P.

A few examples selected from this range of applications are mentioned here. First, the use of NMR microscopy as a chemical mapping tool in plant physiology is illustrated in **Figure 7**, where two images of a castor bean stem cross section are illustrated alongside the total NMR spectrum from the stem. The spectral regions in the initial excitation pulse are indicated by the vertical line in each case and the images correspond to (A) water and (B) fructose. The sugar image is intense precisely at the phloem regions within the vascular bundles. There are many interesting applications of relaxation time weighting in order to obtain useful image contrast, some of which can be seen in chapter 5 of the book by Callaghan. One example from that chapter concerns profiles taken across several growth rings of a wood segment in which the water components associated with early wood, late wood and cell walls are separated by different spin relaxation times. This illustrates the potential of NMR microscopy to resolve differing aqueous components in porous media, in biological tissue and in food products. Another example concerns the sensitivity of $T_1$ and $T_2$ to the oxidation state of ions, an effect that is put to use in the imaging of chemical waves associated with the Belusov–Zhabotinsky reaction.

**Figures 8** and **9** show applications of NMR microscopy in the rheological investigation of complex viscoelastic fluids. In **Figure 8** comparative velocity and diffusion profiles are shown across the diameter of a 700 μm diameter capillary though which is pumped a solution of high-molecular-mass polymer undergoing laminar flow. The velocity profile is distinctly non-Poiseuille, consistent with shear thinning, while the polymer self-diffusion coefficients exhibit a dramatic enhancement once the shear rate (the velocity gradient) exceeds a characteristic value. This value corresponds to the slowest relaxation rate of the molecule $\tau_d^{-1}$, where $\tau_d$ is the so-called tube disengagement time. $\tau_d$ indicates the time

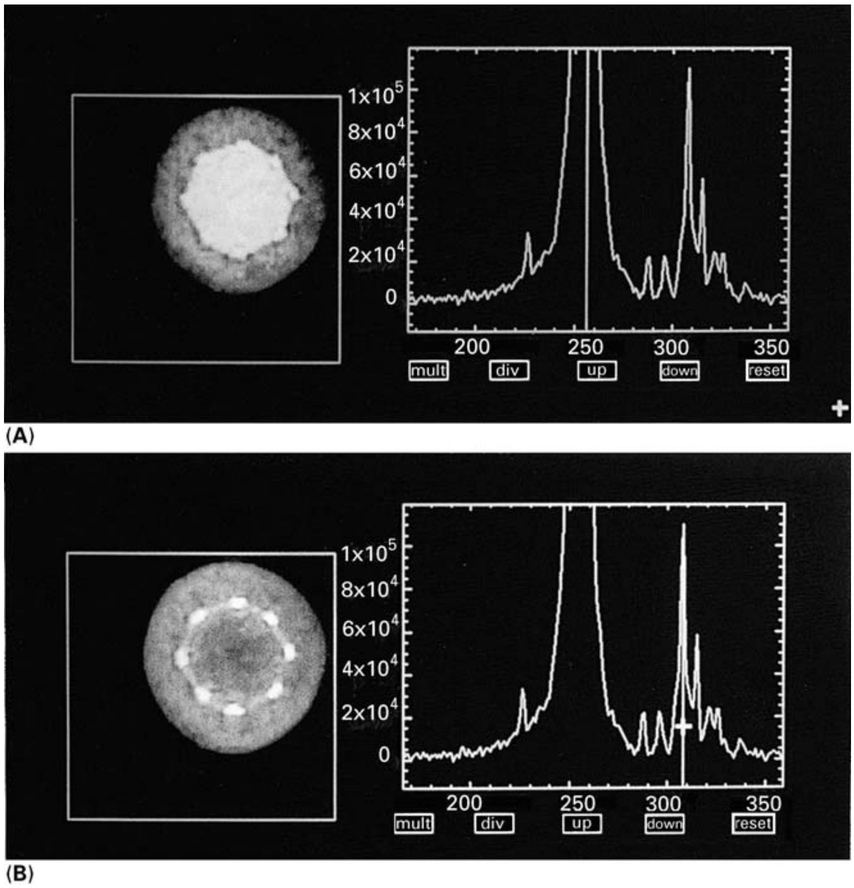



**Figure 7** Chemical shift-selective proton NMR images of castor bean stems along with the corresponding proton NMR spectra from the whole stem. The region of the spectrum used to generate the corresponding image is shown by the vertical line. The upper image corresponds to water distribution while the lower corresponds to fructose distribution. Note the confinement of the sugars to the site of the phloem. Courtesy of W. Koeckenberger.

taken for the polymer to completely reconfigure its conformation. This diffusion enhancement phenomenon is associated with the breakdown of chain entanglements under shear and illustrates the changes that occur when an externally imposed rate of strain competes with the natural Brownian dynamics of a molecular system. In the example shown here, the crossover in the competition can be observed because the diameter of the capillary is sufficiently small that high shear rates can be observed, thus emphasizing the importance of the microscopic length scale.

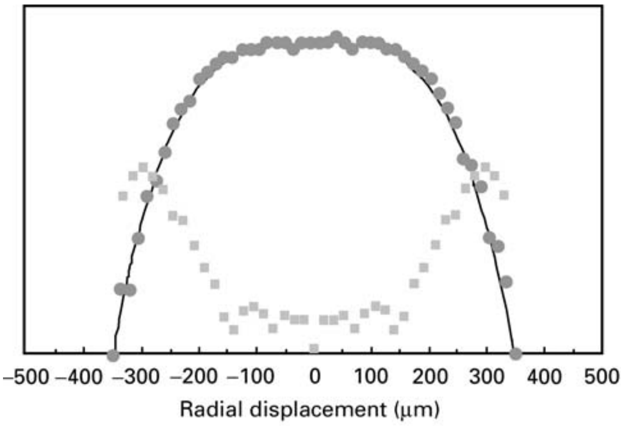



**Figure 8** Velocity (●) and diffusion (■) profiles taken across a diametral slice for a solution of 5% 1.6 MDa poly(ethylene oxide) in laminar flow through a 700 mm diameter capillary. Courtesy of Y Xia and the author.

The second example, shown in **Figure 9**, also illustrates some of the strange phenomena apparent when competitive crossover occurs in systems that exhibit flow instability or are close to a phase transition. The image shows the velocity profile obtained for a wormlike micelle solution sheared in the gap of a cone-and-plate rheometer, along with the resultant shear rate of the fluid. Flow instability phenomena have led to distinctive shear banding, in effect a first-order phase transition of the fluid maintained under non-equilibrium conditions. Microscopy has been able to provide unique insight in this emerging area of condensed-matter physics.

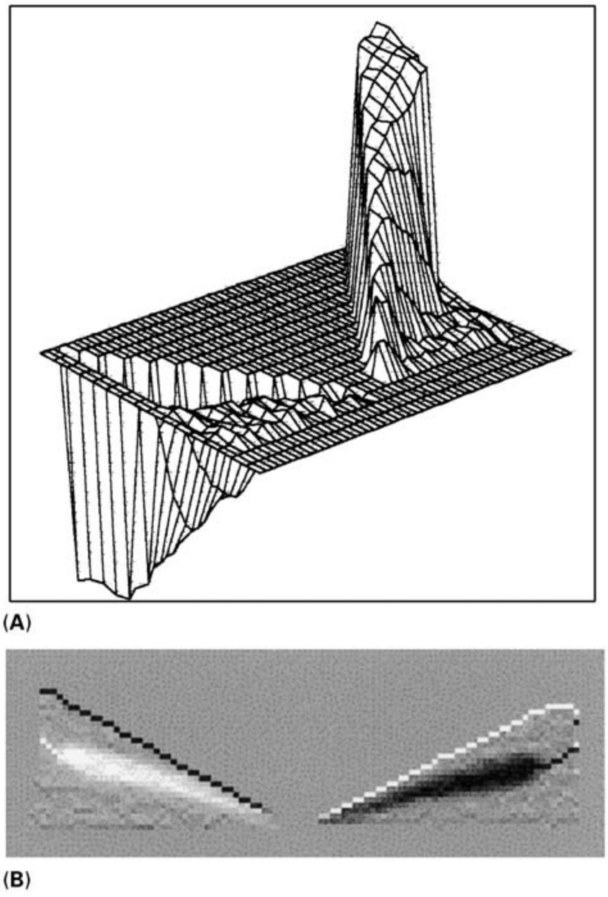



**Figure 9** Stacked profile map (A) of the fluid velocity across the gap of a 7° cone and plate system containing a semidilute worm-like micelle solution, and shear rate image (B) obtained by taking the derivation of the velocity across the gap. Dramatic shear banding effects are apparent. Note the use of an expanded field of view across the gap achieved by using different magnitudes of spatial encoding gradients. Courtesy of MM Britton and the author.

## Diffraction phenomena and higher spatial resolution

The diffractive physics at the heart of NMR imaging means in effect that any degree of structural homogeneity can be used to advantage in improving resolution. For example, in a periodic structure, the raw signal $S(\boldsymbol{k})$ will contain coherence peaks whose spacing bears an inverse relationship to the separations of 'scattering centres' in real space. Shorter-range spatial correlations can be investigated using Patterson function ($|S(\boldsymbol{k})|^2$) analysis. An entirely different but related diffraction phenomenon arises in the PGSE NMR experiment from the diffusion or dispersion of fluid within pores, where the boundary restrictions to molecular translation result in distinctive coherences in $E(q)$. All these approaches have the potential to greatly extend the resolution by which NMR field gradient methods can be used to probe liquid-phase morphology in materials science and biology.

## Conclusion

NMR microscopy is an expensive and sophisticated technique that requires specialist insight in order to gain maximum advantage. Every new application requires a significant degree of spectroscopic optimization (e.g. RF and gradient pulse sequence design) and hardware optimization (e.g. RF antenna and sample holding design) if the best results are to be achieved. The method has now been realistically assessed and is likely to find extensive use in chemical physics and in chemical engineering in the study of fluid dynamics, multiphase flow and dispersion, restricted diffusion and flow in porous media, the rheology of soft condensed matter, phase separation and chemical instability, pH and temperature mapping, permeation and interdiffusion, and in research concerning electrophoresis and current imaging. In biological applications the method has unrivalled capability in studying the distribution and transport of specific molecular species in plant and insect physiology *in vivo*. In medical research it is proving an important tool in rat and mouse studies *in vivo* and it holds considerable promise as a complementary method in histology.

While the intrinsic insensitivity of NMR confines both the spatial and temporal resolution possible in NMR microscopy, a number of important new technical developments may yet extend those limits. These include the use of superconducting RF coils, and the use of indirect optical pumping techniques to greatly enhance nuclear polarization. Furthermore, in systems exhibiting a degree of structural homogeneity, the limits to resolution may be significantly enhanced by taking advantage of diffraction methodology. These latter approaches, while scientifically challenging, provide a nice link between the field of NMR spectroscopy and the wealth of scattering techniques that are so widely used in the study of condensed matter.

## List of symbols

$B_0$ = applied magnetic field strength [flux density]; $B_1$ = applied transverse magnetic field strength [flux density]; $D_s$ = self-diffusion coefficient; $E$ = echo amplitude; $\boldsymbol{g}$ = gradient pulse amplitude; $\boldsymbol{G}$ = magnetic field gradient; $\boldsymbol{k}$ = reciprocal-space wave vector; $P_s$ = average propagator; $\boldsymbol{q}$ = wave vector conjugate to molecular displacement; $\boldsymbol{r}$ = position vector; $\boldsymbol{R}$ = distance moved by spin over time $\Delta$; $S$ = signal

amplitude; $t_p$ = pulse duration; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $\boldsymbol{v}$ = local molecular velocity; $\gamma$ = gyromagnetic ratio; $\delta$ = gradient pulse duration; $\Delta$ = gradient pulse separation; $\rho(\boldsymbol{r})$ = density of spins at $\boldsymbol{r}$; $\tau$ = delay time; $\omega$ = angular frequency of applied radiation/nuclear precession.

*See also:* **Contrast Mechanisms in MRI; Diffusion Studied Using NMR Spectroscopy; EPR Imaging; Fourier Transformation and Sampling Theory; Magnetic Field Gradients in High Resolution NMR; MRI Applications, Biological; MRI Instrumentation; MRI of Oil/Water in Rocks; MRI Theory; NMR Principles; NMR Pulse Sequences; NMR Relaxation Rates.**

## Further reading

Blümich B and Kuhn W (eds) (1992) *Magnetic Resonance Microscopy: Methods and Application in Materials Science, Agriculture, and Biomedicine*. Weinheim: VCH.

Callaghan PT (1991) *Principles of NMR Microscopy*. Oxford: Oxford University Press.

Callaghan PT (1996) NMR imaging, NMR diffraction and applications of pulsed gradient spin echoes in porous media. *Magnetic Resonance Imaging* **14**: 701–709.

Callaghan PT and Stepisnik J (1996) Generalised analysis of motion using magnetic field gradients. *Advances in Magnetic and Optical Resonance* **19**: 325–388.

Kimmich R (1997) *NMR Tomography, Relaxometry and Diffusiometry*. Berlin: Springer.

Mansfield P and Morris PG (1982) *NMR Imaging in Biomedicine*. New York: Academic Press.

# NMR of Solids

**Jacek Klinowski**, University of Cambridge, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| Applications |

## Introduction

NMR spectra cannot normally be measured in solids in the same way in which they are routinely obtained from liquids. For example, the width of the $^1$H NMR line in the spectrum of water is ~0.1 Hz, while the line from a static sample of ice is ~100 kHz wide. The reason for this is the existence of net anisotropic interactions which in the liquid are exactly averaged by the rapid thermal tumbling of molecules. A typical high-resolution spectrum of an organic compound in solution contains a wealth of information. The frequency of the radiation absorbed by the various non-equivalent nuclei in the molecule depends subtly on their chemical environments, giving rise to very sharp spectral lines. The parameters derived from such a spectrum (positions, widths, intensities and multiplicities of lines, relaxation mechanisms and rates) provide detailed information on the structure, conformation and molecular motion. This is not the case in a solid, where the nuclei are static and a conventional NMR spectrum is a broad hump which conceals most structural information. Although certain solids have sufficient molecular motion for NMR spectra to be obtainable without resorting to special techniques, we are concerned here with the general case, where there is no motion of nuclei and where conventional NMR, instead of sharp spectral lines, yields a broad hump which conceals information of interest to a chemist. Although the study of moments of such spectra and of various temperature-dependent parameters can still yield information on the degree of crystallinity, interatomic distances and molecular motion ('wide-line NMR'), we shall be primarily interested in ways of achieving high-resolution spectra, i.e. spectra which enable magnetically non-equivalent nuclei of the same spin species (e.g. $^{13}$C) to be resolved as individual lines.

The interactions to be considered in the solid state and their Hamiltonians are as follows:

(1) Zeeman interaction with the magnetic field, $\mathcal{H}_Z$;
(2) chemical shielding, $\mathcal{H}_{CS}$;
(3) dipolar interaction, $\mathcal{H}_D$;
(4) *J*-coupling, $\mathcal{H}_J$;
(5) quadrupolar interaction, $\mathcal{H}_Q$.

The total Hamiltonian is a sum of all these contributions:

$$\mathcal{H} = \mathcal{H}_Z + \mathcal{H}_{CS} + \mathcal{H}_D + \mathcal{H}_J + \mathcal{H}_Q \qquad [1]$$

with the quadrupolar term $\mathcal{H}_Q$ non-zero only for nuclei with $I > \frac{1}{2}$. In general, $\mathcal{H}_Z$, $\mathcal{H}_{CS}$, $\mathcal{H}_D$ and $\mathcal{H}_Q$ are

much larger than $\mathcal{H}_J$. *J*-Coupling is rarely observed in solids, so that $\mathcal{H}_J$ will henceforward be neglected.

The interaction Hamiltonians have the general form

$$\mathcal{H} = \boldsymbol{I} \cdot \mathbf{A} \cdot \boldsymbol{S} = [I_x\ I_y\ I_z] \begin{bmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yz} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{bmatrix} \begin{bmatrix} S_x \\ S_y \\ S_z \end{bmatrix} \qquad [2]$$

where $\boldsymbol{I}$ and $\boldsymbol{S}$ are vectors and $\mathbf{A}$ is a second-rank Cartesian tensor. We shall consider the various interactions in turn.

## The Zeeman interaction

The Zeeman Hamiltonian, which determines the resonance frequency of an NMR-active nucleus in the magnetic field $\boldsymbol{B}_o$ is

$$\mathcal{H}_Z = \boldsymbol{I} \cdot \mathbf{Z} \cdot \boldsymbol{B}_o \qquad [3]$$

where $\mathbf{Z} = -\gamma \hbar \mathbf{1}$, $\boldsymbol{I} = [I_x, I_y, I_z]$, $\boldsymbol{B}_o = [B_x, B_y, B_z]$ and $\mathbf{1}$ is a unit matrix. When the magnetic field is aligned with the *z*-axis of the laboratory frame of reference, $\boldsymbol{B}_o = [0, 0, B_o]$. The Zeeman interaction, which is directly proportional to the strength of the magnetic field, is thus entirely under the operator's control.

## Magnetic shielding

The effect known as the chemical shift, central to the application of NMR in chemistry, is caused by simultaneous interactions of a nucleus with surrounding electrons and of the electrons with the static magnetic field $\boldsymbol{B}_o$. The field induces a secondary local magnetic field which opposes $\boldsymbol{B}_o$ thereby 'shielding' the nucleus from its full effect. The shielding Hamiltonian is

$$\mathcal{H}_{CS} = -\gamma\hbar\, \boldsymbol{I} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{B}_o \qquad [4]$$

The shielding is anisotropic, which is quantified in terms of a second-rank tensor $\boldsymbol{\sigma}$ ('the chemical shielding tensor'):

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \qquad [5]$$

In strong magnetic fields $\boldsymbol{\sigma}$ is axially symmetric. When transformed into its principal reference system (PAS) by using rotation matrices, the tensor is described by three principal components $\sigma_{ii}$ ($i$ = 1, 2, 3):

$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \xrightarrow{\text{Rotation}} \begin{bmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix}$$

and three direction cosines, $\cos\theta_i$, between the axes of PAS and the laboratory frame.

The observed shielding constant, $\sigma_{zz}$, is a linear combination of the principal components:

$$\sigma_{zz} = \sum_{i=1}^{3} \sigma_{ii} \cos^2\theta_i = \frac{1}{3}\mathrm{Tr}\,\boldsymbol{\sigma} + \frac{1}{3}\sum_{i=1}^{3} (3\cos^2\theta_i - 1)\sigma_{ii} \qquad [6]$$

where Tr $\boldsymbol{\sigma}$ stands for the trace of the tensor. Since the average value of each $\cos^2\theta_i$ is $\frac{1}{3}$, the average value of $\sigma_{zz}$ in the NMR spectra of liquids (where there is random molecular tumbling) is the isotropic value:

$$\overline{\sigma_{zz}} = \frac{1}{3}\mathrm{Tr}\ \sigma = \sigma_{\mathrm{iso}} \qquad [7]$$

In solids the angle-dependent second term on the right of Equation [6] survives, giving rise to a spread of resonance frequencies, i.e. line broadening.

## Dipolar interactions

The Hamiltonian for the dipolar interaction between a pair of nuclei *i* and *j* separated by the internuclear vector $\boldsymbol{r}$ is given by

$$\mathcal{H}_D = R\boldsymbol{I}_i \cdot \mathbf{D} \cdot \boldsymbol{I}_j = R\,[I_{ix}\ I_{iy}\ I_{iz}] \times \begin{bmatrix} r^2 - 3x^2 & -3xy & -3xy \\ -3xy & r^2 - 3y^2 & -3yz \\ -3xz & -3yz & r^2 - 3z^2 \end{bmatrix} \begin{bmatrix} I_{jx} \\ I_{jy} \\ I_{jz} \end{bmatrix} \qquad [8]$$

where $R = \gamma_i \gamma_j \hbar\mu_o / 4\pi r^3$ is the dipolar coupling constant, $\gamma$ the nuclear gyromagnetic ratio and $\mathbf{D}$ the dipolar interaction tensor. In the PAS of the tensor, with the internuclear vector aligned along one

of the coordinate axes, we have $xy = yz = zx = 0$, $r^2 = x^2+y^2+z^2$ and the tensor becomes

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad [9]$$

It is clearly traceless (Tr $\mathbf{D}$ = 1−2+1 = 0).

The truncated dipolar interaction Hamiltonian may be written in the form

$$\mathcal{H}_\mathrm{D} = \frac{\gamma_i \gamma_j \hbar^2}{2r^3}[\boldsymbol{I}_i \cdot \boldsymbol{I}_j - 3I_{iz}\, I_{jz}](3\cos^2\theta - 1) \qquad [10]$$

where $\theta$ is the angle between $\boldsymbol{r}$ and the external magnetic field $\boldsymbol{B}_\mathrm{o}$. Since the average value $\overline{\cos^2\theta_{ij}} = \frac{1}{3}$, the isotropic average of the Hamiltonian is $\overline{\mathcal{H}_\mathrm{D}} = 0$, so that the dipolar interaction does not affect the NMR spectrum in solution. In the solid the interaction remains, greatly increasing the spectral line width.

## Quadrupolar interactions

Some 74% of all NMR-active nuclei have $I > \frac{1}{2}$, so that, in addition to magnetic moment, they possess an electric quadrupole moment brought about by non-spherical distribution of the nuclear charge. The quadrupole interaction broadens and shifts the NMR lines, and also affects their relative intensities.

When the quadrupolar Hamiltonian is considered as a perturbation on the Zeeman Hamiltonian, there is no general analytical solution for the eigenvalues of $\mathcal{H}_\mathrm{Z}$ in the (very rare) case when $\mathcal{H}_\mathrm{Z}$ and $\mathcal{H}_\mathrm{Q}$ are of comparable magnitude. When $\mathcal{H}_\mathrm{Q} >> \mathcal{H}_\mathrm{Z}$, the splitting of the nuclear states is very large and 'pure quadrupole resonance' (NQR) is observed even in the absence of a magnetic field. In the usual 'high field' case, $\mathcal{H}_\mathrm{Z} >> \mathcal{H}_\mathrm{Q}$, the quadrupole Hamiltonian in the PAS of the electric field gradient tensor is

$$\mathcal{H}_\mathrm{Q} = \frac{e^2 qQ}{4I(2I-1)}[3I_z^2 - I^2 + \eta(I_x^2 - I_y^2)] \qquad [11]$$

where $\eta$ is the asymmetry parameter which describes the symmetry of the electric field gradient. The definitions of $\eta$ and of the 'quadrupole frequency', $\nu_\mathrm{Q}$, which describes the magnitude of the interaction, are

$$\eta = \frac{V_{xx} - V_{yy}}{V_{zz}} \qquad \nu_\mathrm{Q} = \frac{3e^2 qQ}{2I(2I-1)h} \qquad [12]$$

Perturbation theory allows us to calculate the energy levels $E_m^{(0)}$, $E_m^{(1)}$ and $E_m^{(2)}$ (superscripts denote the order). Because of the first- and second-order shifts in energy levels, instead of a single (Larmor) resonance frequency $\nu_\mathrm{L} = [E_{m-1}^{(0)} - E_m^{(0)}]$, as with spin-$\frac{1}{2}$ nuclei, there are now several resonance frequencies:

$$\nu_m = \frac{E_{m-1} - E_m}{h} = \nu_\mathrm{L} + \nu_m^{(1)} + \nu_m^{(2)} \qquad [13]$$

Detailed calculations reveal that:

(1) The first-order frequency shift is zero for $m = \frac{1}{2}$, so that the central transition for non-integer spins (such as $^{27}$Al with $I = \frac{5}{2}$) is not affected by quadrupolar interactions to first order. It is thus advantageous to work with such nuclei, especially since the central transition is normally the only one which is observed: other transitions are so broadened and shifted as to be unobservable.

(2) The first-order shift is scaled by $\frac{1}{2}(3\cos^2\theta - 1)$.

(3) The second-order shift increases with $\nu_\mathrm{Q}^2$ and is inversely proportional to the magnetic field strength. Since the dispersion of the chemical shift, which is what we normally wish to measure, is proportional to $\boldsymbol{B}_\mathrm{o}$, it is advantageous to work at high fields, where the chemical shift effects make the maximum contribution to the spectrum. As the second-order frequency shift is always present for all transitions, the feasibility of obtaining useful spectra depends on the magnitude of $\nu_\mathrm{Q}$.

The very small quadrupole interactions of $^2$H and their sensitivity to molecular motion at a wide range of frequencies make this integer spin nucleus very useful for chemical studies. $^2$H NMR experiments normally use static samples, and dynamic information is extracted by comparing spectra measured at different temperatures with model computer simulations.

## Magic-angle spinning

Magic-angle spinning (MAS) is by far the most powerful tool in solid-state NMR. The technique averages anisotropic interactions by acting on the factor $(3\cos^2\theta - 1)$ in the Hamiltonians, which in solids is

not averaged to zero by rapid molecular motion. MAS was first introduced to deal with the dipolar interaction. It can be shown that when the sample is rapidly spun around an axis inclined at the angle $\beta$ to the direction of the magnetic field, the time-averaged value of the angle $\theta$, which an arbitrary internuclear vector makes with $\boldsymbol{B}_o$, is

$$\overline{3\cos^2\theta - 1} = \frac{1}{2}(3\cos^2\beta - 1)(3\cos^2\chi - 1) \qquad [14]$$

where $\chi$, is the angle between the internuclear vector and axis of rotation, is constant for each vector, because the solid is rigid. The result is that the term $\frac{1}{2}(3\cos^2\beta - 1)$ scales the spectral width, and that for $\beta = \cos^{-1}/\sqrt{3} = 54.74°$ (the 'magic angle'), $\overline{3\cos^2\theta - 1} = 0$. The dipolar Hamiltonian in Equation [10] is averaged to zero.

For MAS to be effective, the sample must be spun at a rate greater than the static spectral width expressed in Hz. As the homonuclear $^1$H–$^1$H interactions may lead to spectra which are as much as 50 kHz wide, it is not possible to spin the sample fast enough. Thus high-resolution solid-state $^1$H spectra of most organic compounds, where protons are generally close together, cannot be obtained with the use of MAS alone, but require the additional use of multiple-pulse techniques (see below). However, MAS is successful in removing homonuclear interactions for $^{13}$C, $^{31}$P and nuclei of small gyromagnetic ratios.

The chemical shift anisotropy is also reduced by MAS, because the tensor interactions controlling all anisotropic interactions in solids all have a common structure and may be expressed in terms of Wigner rotation matrices which are scaled by MAS.

## High-power decoupling

When dilute spins, such as $^{13}$C, interact via the dipolar interaction with $^1$H or other abundant nuclei, the large heteronuclear broadening of an already low-intensity spectrum is a considerable problem. High-power decoupling, used to remove heteronuclear coupling effects, applies a continuous, very-high-power pulse at the $^1$H resonance frequency in a direction perpendicular to $\boldsymbol{B}_o$. The $^{13}$C pulse is then applied, and the $^{13}$C free induction decay measured while continuing the $^1$H irradiation. The powerful decoupling pulse stimulates rapid $^1$H spin transitions, so rapid that the $^{13}$C spins experience only the time-average of the $^1$H magnetic moment, i.e. zero. Since the technique relies on selective excitation of the abundant and dilute nuclei, it can only remove heteronuclear interactions.

## Cross-polarization

Dilute nuclei, such as $^{13}$C and $^{15}$N, are more difficult to observe than abundant nuclei, such as $^1$H or $^{31}$P, particularly when they also have a low gyromagnetic ratio. However, the dilute and abundant nuclei are often in close proximity, and coupled via the dipolar interaction. Cross-polarization (CP) exploits this interaction to observe dilute nuclei, at the same time overcoming two serious problems often encountered in solid-state NMR: (i) because of a very small population difference in the polarized sample, NMR actually observes very few dilute spins and consequently the sensitivity of the experiment is low; (ii) spin–lattice relaxation times of spin-$\frac{1}{2}$ nuclei in solids are often very long, so that long delays are required between experiments and the spectral signal-to-noise ratio is poor.

The sequence of events during the $^{13}$C–$^1$H CP experiment is as follows. After the end of the 'preparation period', during which the sample polarizes in the magnetic field, a $\pi/2$ pulse is selectively applied to $^1$H along the $x$-axis of the rotating frame, aligning the $^1$H magnetization with the $y$-axis. A long pulse of amplitude $B_{1H}$ is then applied along the $y$-axis. Since the $^1$H magnetization is now aligned with the effective field in the rotating frame, it becomes 'spin locked' along this direction. At the same time, a long pulse of amplitude $B_{1C}$ is selectively applied to $^{13}$C along the $x$-axis. The amplitudes $B_{1H}$ and $B_{1C}$ are adjusted so as to satisfy the Hartmann–Hahn condition:

$$\gamma_H B_{1H} = \gamma_C B_{1C} \qquad [15]$$

The energies of $^1$H and $^{13}$C in the rotating frame are thus equal, and the two spin reservoirs can transfer magnetization in an energy-conserving manner during the 'contact time'. Finally, the $^{13}$C radiofrequency field is turned off and a free induction decay observed in the usual way. During the observation time the $^1$H field is still on, but serves as the high-power decoupling field to reduce the $^1$H–$^{13}$C dipolar broadening.

Detailed arguments show that the magnetization of $^{13}$C nuclei is theoretically increased by the factor of $\gamma_H/\gamma_C \approx 4$. After the $^{13}$C free induction decay signal has been measured, the magnetization of carbons is again almost zero, but the loss of proton magnetization is small. The CP experiment can be repeated without waiting for the carbons to relax. The only

limitations are the gradual loss of polarization by the $^{1}H$ spin reservoir, and the decay of the $^{1}H$ magnetization during spin locking. The latter process proceeds on a time-scale ('spin–lattice relaxation in the rotating frame') which is much shorter than the $^{13}C$ spin–lattice relaxation time.

## Multiple-pulse line narrowing

Although homonuclear dipolar couplings are in principle removable by MAS, with abundant nuclei they are often very strong. For example, the removal of the $^{1}H$–$^{1}H$ interaction in most organic compounds requires spinning rates far in excess of what is practically feasible. The alternative to MAS is to manipulate the nuclear spins themselves using 'multiple-pulse line narrowing' so as to average the dipolar interaction. The method uses specially designed sequences of pulses with carefully adjusted phase, duration and spacing. The result is that, when the signal is sampled at a certain moment during the sequence, the dipolar interaction is averaged to zero.

WAHUHA, the simplest multiple pulse sequence, is composed of four 90° pulses:

$$(\mathrm{P}_x - 2\tau - \mathrm{P}_{-x} - \tau - \mathrm{P}_y - 2\tau - \mathrm{P}_{-y} - \tau)\ n \text{ times} \qquad [16]$$

where $P_i$ represents rotation about the particular $i$-axis of the rotating frame and $\tau$ is the time interval between pulses. Over the sequence, the magnetic moments spend equal amounts of time along each of the three principal axes. The NMR signal is sampled in one of the $2\tau$ windows. Sequences have been developed involving from 4 to as many as 52 pulses. The entire sequence must be short relative to the relaxation time $T_2$, and the pulses themselves must also be very short.

Multiple pulse sequences average the dipolar Hamiltonian, but also affect other Hamiltonians to an extent which depends on the particular sequence. For example, the WAHUHA sequence scales chemical shift anisotropies by a factor of $1/\sqrt{3}$.

## Moments of an NMR line

Even when the dipolar $^{1}H$–$^{1}H$ interaction is not removed from the spectrum, the method of moments can provide important structural information. The $n$th moment of the line shape $f(\omega)$ about $\omega_0$ is defined as

$$M_n = \frac{\int_0^\infty (\omega - \omega_0)^n f(\omega)\,\mathrm{d}\omega}{\int_0^\infty f(\omega)\,\mathrm{d}\omega} \qquad [17]$$

where

$$M_0 = \int_0^\infty f(\omega)\,\mathrm{d}\omega$$

is the area under the line (the zeroth moment). For a normalized function $M_0 = 1$. The second moment is physically analogous to the moment of inertia of an object with the same shape as the line. If $f(\omega)$ is an even function of $\omega$, $M_n = 0$ for all odd values of $n$. It is convenient to calculate moments about the centre of gravity of the line shape, i.e. the value of $\omega_0$ for which the first moment is zero.

The second moment can be calculated from the interatomic distances in the solid containing pairs $i$, $j$ of dipolar-coupled nuclei. Van Vleck has shown that, for a polycrystalline powder composed of randomly oriented crystals in which we observe identical spin-$\frac{1}{2}$ nuclei, the second moment is

$$M_2^{\mathrm{homo}} = \frac{9}{16}\gamma^4\,\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2 \sum_j \frac{1}{r_{ij}^6} \qquad [18]$$

while for pairs of unlike nuclei the second moment is different:

$$M_2^{\mathrm{hetero}} = \frac{1}{4}\gamma_i^2\,\gamma_j^2\,\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2 \sum_j \frac{1}{r_{ij}^6} \qquad [19]$$

Thus, even when the interacting nuclei have very similar gyromagnetic ratios, the homonuclear second moment is larger by a factor of $\frac{9}{4}$ than the heteronuclear moment. This is because dipolar coupling between unlike spins cannot lead to an energy conserving mutual spin flip. The second moment is thus very sensitive to the kind of neighbour.

The method of moments has further advantages. First, since the second moment is inversely proportional to the sixth power of the internuclear distance, it is a very sensitive means of determining interatomic distances. Second, it can provide insights into the structure. For example, it was used to demonstrate the presence of groups of three equivalent protons in solid hydrates of strong acids, thus proving the

presence of oxonium ions, $H_3O^+$. Third, it is useful for the study of motion, because the moments are dramatically reduced when the dipolar interaction is partly or completely averaged out by an onset of a specific motion.

## DOR, DAS and MQ-MAS

We have seen that the second-order quadrupolar interaction, which affects all quadrupolar nuclei, is reduced, but not removed, by MAS. Its complete removal is the most important current problem in solid-state NMR. Three different techniques have been proposed to achieve this aim.

When the second-order quadrupole interaction is expanded as a function of Wigner rotation matrices, and we consider the case of a sample rapidly rotated about an angle $\beta$ with respect to $\boldsymbol{B}_o$, the average second-order quadrupolar shift of the central transition becomes

$$\nu_{\frac{1}{2}}^{(2)} = \frac{\nu_Q^{(2)}}{\nu_L}\left[I(I+1) - \frac{3}{4}\right][A_0 + B_2\,P_2(\cos\beta) + B_4\,P_4(\cos\beta)] \quad [20]$$

where $\nu_Q$ is the quadrupole frequency, $\nu_L$ is the Larmor frequency, $A_0$ and $B_0$ are constants and the $P_n(\cos\beta)$ terms are the Legendre polynomials

$$P_2(\cos\beta) = \frac{1}{2}(3\cos^2\beta - 1)$$
$$P_4(\cos\beta) = \frac{1}{8}(35\cos^4\beta - 30\cos^2\beta + 3) \quad [21]$$

There is no value of $\beta$ for which both the $P_2(\cos\beta)$ and the $P_4(\cos\beta)$ terms can be zero, so that the angle-dependent terms cannot be averaged by spinning about a single axis. Instead, in the ingenious 'double-rotation' (DOR) experiment the sample is spun *simultaneously* about two different axes $\beta_1$ and $\beta_2$, so that

$$P_2(\cos\beta_1) = 0 \qquad P_4(\cos\beta_2) = 0 \quad [22]$$

with solutions $\beta_1 = 54.74°$ (the conventional magic angle) and $\beta_2 = 30.56$ or $70.12$ °. As a result, only the $A_0$ term remains in Equation [20]. This is accomplished by a rotor-within-a-rotor probehead in which the centres of gravity of the two rotors, each spinning at a different angle with respect to $\boldsymbol{B}_o$, exactly coincide. Although the daunting engineering problems posed by the design of a DOR probehead have been overcome, it is very difficult to spin the two rotors simultaneously at sufficiently high spinning speeds, and the spinning rates are at present limited to ~6 and 1 kHz for the inner and outer rotors, respectively, compared with ~30 kHz achievable with MAS. This is an unfortunate limitation, since multiple spinning sidebands appear in the spectra if the rate of the rotation is lower than the strength of the quadrupolar interaction.

The technique known as 'dynamic-angle spinning' (DAS) adopts an alternative approach to DOR: the sample is rotated *sequentially* about two different axes, $\beta_1'$ and $\beta_2'$, which are chosen so that

$$P_2(\cos\beta_1') = -P_2(\cos\beta_2')$$
$$P_4(\cos\beta_1') = -P_4(\cos\beta_2') \quad [23]$$

with the solutions $\beta_1' = 37.38°$ and $\beta_2' = 79.19°$. The rotation axis is switched very rapidly, which poses technical problems, given that the minimum time required for changing the spinning angle must be shorter than the relaxation time of the nucleus being observed. As a result, DAS often cannot be applied to many nuclei, including $^{27}Al$, and is limited to the study of nuclei with long relaxation times (for example in amorphous samples, such as glasses).

Yet another solution to the problem, known as 'multiple-quantum magic-angle spinning' (MQ-MAS) relies on the fact that $B_2$ and $B_4$ are functions of $I$, $p$, $\eta$, $\alpha$ and $\beta$, where $p$ is the order of the multiquantum coherence and $\alpha$ and $\beta$ are the Euler angles corresponding to the orientation of each crystallite in the powder with respect to the rotor axis. Under fast MAS, the chemical shift anisotropy, heteronuclear dipolar interactions and the term proportional to $P_2$ in Equation [20] are removed, so that

$$\nu_{\frac{1}{2}}^{(2)} = \frac{\nu_Q^{(2)}}{\nu_L}\left[I(I+1) - \frac{3}{4}\right][A_0 + B_4\,P_4(\cos 54.74°)] \quad [24]$$

Although the second term, proportional to $P_4$, still causes substantial line broadening, it can be eliminated by using $p$-quantum transitions. A $p$-quantum transition (with $p = 3$ or $5$ for $^{27}Al$) is excited and the signal allowed to evolve during time $t_1$. As multiple quantum transitions are not directly observable by

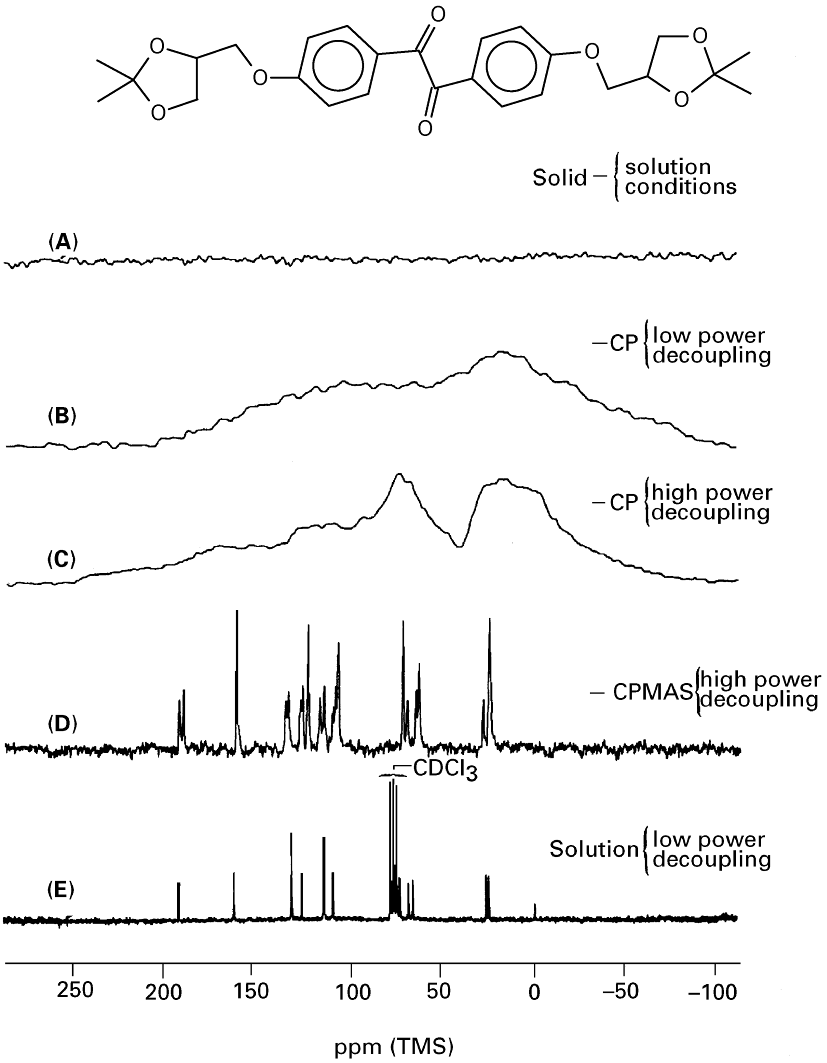



**Figure 1** $^{13}C$ NMR spectra of solid 4,4′-bis[(2,3-dihydroxypropyl)oxy]benzil. (A) Solution conditions using 60° $^{13}C$ pulses and 10 s recycle delays; (B) as in (A) but with $^{1}H$–$^{13}C$ cross-polarization, low-power proton decoupling and 1 s recycle delays; (C) as in (B) but with high-power proton decoupling; (D) as in (C) but with the addition of magic-angle spinning; (E) high-resolution spectrum of a solution in $CDCl_3$ with the same NMR parameters. Reproduced with permission of the American Chemical Society from Yannoni CS (1982) *Accounts of Chemical Research* **15**: 201–208. Copyright 1982 American Chemical Society.

NMR, a second pulse converts the signal into a single-quantum transition, which is observable. The technique enables a two-dimensional representation of the spectra, with a regular increment of $t_1$ providing a '$p$-quantum dimension', free of quadrupolar interactions. Although the optimal conditions for MQ-MAS are difficult to establish, the technique is being increasingly used for the study of quadrupolar nuclei of half-integer spin, such as $^{27}Al$, $^{85}Rb$, $^{23}Na$, $^{11}B$ and $^{93}Nb$.

Note that DOR, DAS and MQ-MAS do not remove the $A_0$ term in Equations [20] and [24]. Thus the position of the line in the spectrum, however narrow, does not correspond to the pure chemical shift, but includes the effect of the quadrupole interaction.

## Modern solid-state NMR

Magic-angle spinning has greatly enhanced our knowledge of a wide range of materials used in chemical, physical, biological and earth sciences and in the technology of glass and ceramics. It took nearly twenty years, since its discovery in 1958, for MAS to become a routine tool of structural investigation. The reasons were the difficulty of spinning the sample at the very high speeds required and the insufficiently high magnetic fields. However, the introduction of Fourier-transform NMR, cross-polarization and superconducting magnets during the 1960s and 1970s greatly improved the sensitivity of the spectra and enabled virtually all NMR-active

nuclei to be observed in solids. $^{1}H$ MAS NMR was used to examine polymers as early as 1972, and Schaefer and Stejskal were the first to combine CP and MAS in $^{13}C$ NMR studies of organics. Much important work, at first mostly with $^{13}C$ but later with other nuclei, has been done since. Since the early 1980s great progress has been made in the study of $^{29}Si$ and $^{27}Al$ in natural and synthetic molecular sieve catalysts and minerals, which is particularly significant since nearly a half of all known minerals are silicates or aluminosilicates.

High-resolution spectra of solids are now routinely obtained using a combination of CP and MAS (see **Figure 1**), and it is fair to say that CP-MAS has revolutionized materials science. The otherwise weak signals from dilute nuclei (such as $^{13}C$ or $^{29}Si$) are enhanced by cross-polarization, heteronuclear dipolar interactions are removed by high-power decoupling, chemical shift anisotropy and the weak dipolar interactions between dilute nuclei are averaged by fast MAS, and the signal-to-noise ratio is increased further thanks to the more frequent repetition of the experiment and the availability of high magnetic fields. Although the line widths is in such high-resolution spectra are still greater than these measured in liquids, the various non-equivalent nuclei can in most cases be separately resolved.

## List of symbols

$\boldsymbol{B}_0$ = magnetic flux density; **D** = dipolar interaction tensor; $\mathcal{H}$ = interaction Hamiltonian; $p$ = order of multiquantum coherence; $\mathrm{P}_i$ = rotation about the $i$-axis; $R$ = dipolar coupling constant; $T_1$, $T_2$ = relaxation times; $\gamma$ = nuclear gyromagnetic ratio; $\eta$ = asymmetry parameter; $\nu_Q$ = quadrupole frequency; $\nu_L$ = Larmor resonance frequency; $\sigma_{zz}$ = shielding constant; $\tau$ = time interval between pulses.

*See also:* **$^{13}$C NMR, Parameter Survey; $^{13}$C NMR, Methods; High Resolution Solid State NMR, $^{13}$C; High Resolution Solid State NMR, $^{1}$H, $^{19}$F; Magnetic Field Gradients in High Resolution NMR; NMR Principles; NMR Pulse Sequences; Solid State NMR, Methods; Solid State NMR, Rotational Resonance.**

## Further reading


Abragam A (1983) *The Principles of Nuclear Magnetism*. Oxford: Clarendon Press.

Andrew ER (1981) Magic angle spinning. *International Reviews of Physical Chemistry* **1**: 195–224.

Engelhardt G and Michel D (1987) *High-Resolution Solid-State NMR of Silicates and Zeolites*. Chichester: John Wiley.

Fukushima E and Roeder SBW (1981) *Experimental Pulse NMR – A Nuts and Bolts Approach*. Reading, MA: Addison-Wesley.

Fyfe CA (1983) *Solid State NMR for Chemists*. Ontario: CFC Press.

Mehring M (1983) *High-Resolution NMR Spectroscopy in Solids*, 2nd edn. New York: Springer-Verlag.

Slichter CP (1989) *Principles of Magnetic Resonance*, 3rd edn. New York: Springer-Verlag.

Stejskal EO and Memory JD (1994) *High Resolution NMR in the Solid State. Fundamentals of CP/MAS*. Oxford: Oxford University Press.

# NMR Principles

**PJ Hore**, Oxford University, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| **Theory** |

Nuclear magnetic resonance spectroscopy is an extraordinarily powerful source of information on the structure and dynamics of molecules. Almost every molecule one can think of has at least one magnetic nucleus already in place, exceedingly sensitive to its surroundings but interacting very weakly with them. As such, nuclear spins are ideal probes of molecular properties at the atomic level.

NMR spectra of molecules in liquids contain essentially five sources of information: the intensities of individual resonances (which depend on the number of nuclei responsible), chemical shifts (the interaction of nuclear spins with an applied magnetic field), spin–spin coupling (their interactions with one another), spin relaxation (the restoration of thermal equilibrium), and chemical exchange (the effects of conformational and chemical equilibria).

## Spin angular momentum and nuclear magnetism

Most atomic nuclei have an intrinsic angular momentum known as *spin*. Like the angular momentum of a gyroscope, nuclear spin is a vector quantity – it has both magnitude and direction. Unlike classical angular momentum, however, nuclear spin is quantized. Its magnitude is

$$\sqrt{I(I+1)}\hbar \qquad [1]$$

where $I$ is the *spin quantum number* of the nuclide in question and $\hbar$ is Planck's constant $h$ divided by $2\pi$. $I$ may be zero, or a positive integer or half-integer:

$$I = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \cdots \qquad [2]$$

**Table 1** gives the spin quantum numbers of some popular NMR nuclei.

The projection of the angular momentum vector $\boldsymbol{I}$ onto an arbitrary axis (labelled $z$) is also quantized:

$$I_z = m\hbar \qquad [3]$$

where the *magnetic quantum number*, $m$, can have values between $+I$ and $-I$ in integral steps:

$$m = +I, +I-1, \cdots, -I+1, -I \qquad [4]$$

The spin of a nucleus with $I = \frac{1}{2}$ (e.g. $^1$H) has magnitude ($\sqrt{3}/2$) $\hbar$ and $z$ component $I_z = \pm\frac{1}{2}\hbar$; for $I = 1$ (e.g. $^2$H), the spin angular momentum is $\sqrt{2}\hbar$, and $I_z = 0$ or $\pm\hbar$ (**Figures 1A** and **1B**). According to the uncertainty principle, the other two ($x$ and $y$) components of the angular momentum cannot be known once the magnitude and the $z$ component of $\boldsymbol{I}$ have been specified.

**Table 1** Nuclear spin quantum numbers of some popular NMR nuclides

| *I* | *Nuclide* | | | | | |
|---|---|---|---|---|---|---|
| 0 | $^{12}$C | $^{16}$O | | | | |
| $\frac{1}{2}$ | $^{1}$H | $^{13}$C | $^{15}$N | $^{19}$F | $^{29}$Si | $^{31}$P |
| 1 | $^{2}$H | $^{14}$N | | | | |
| $\frac{3}{2}$ | $^{11}$B | $^{23}$Na | $^{35}$Cl | $^{37}$Cl | | |
| $\frac{5}{2}$ | $^{17}$O | $^{27}$Al | | | | |
| 3 | $^{10}$B | | | | | |

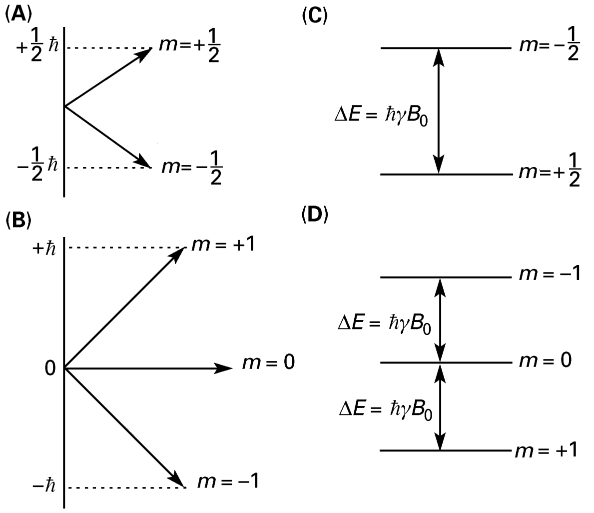



**Figure 1** Space quantization and energy levels of spin-$\frac{1}{2}$ and spin-1 nuclei. (A) and (C) spin-$\frac{1}{2}$; (B) and (D) spin-1. The energy level splittings produced by an applied magnetic field depend on the value of the gyromagnetic ratio, $\gamma$ (here taken as positive).

Closely associated with nuclear spin is a *magnetic moment* $\boldsymbol{\mu}$

$$\boldsymbol{\mu} = \gamma \boldsymbol{I} \qquad [5]$$

which is parallel or sometimes antiparallel to $\boldsymbol{I}$, with a proportionality constant $\gamma$ called the *gyromagnetic ratio*. As a consequence, both the magnitude and orientation of $\boldsymbol{\mu}$ are quantized. In the absence of a magnetic field, all $2I+1$ states of a spin-$I$ nucleus are degenerate, and the direction of the quantization axis is arbitrary.

In an applied magnetic field $\boldsymbol{B}_0$ with strength $B_0$, the spins are quantized along the field direction (the $z$-axis) and have an energy

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B}_0 = -\mu_z B_0 \qquad [6]$$

where $\boldsymbol{\mu} \cdot \boldsymbol{B}_0$ is the scalar product of the two vectors, and $\mu_z$ is the projection of $\boldsymbol{\mu}$ onto $\boldsymbol{B}_0$. Since $\mu_z = \gamma I_z$ and $I_z = m\hbar$, it follows that

$$E = -m\hbar\gamma B_0 \qquad [7]$$

That is, the $2I+1$ states are split apart in energy, with a uniform gap $\Delta E = \hbar\gamma B_0$ between adjacent levels (**Figures 1C** and **1D**).

The NMR experiment involves applying electromagnetic radiation of the correct frequency $\nu$ to 'flip' spins from one energy level to another, according to the selection rule $\Delta m = \pm 1$, i.e.

$$h\nu = \Delta E = \hbar\gamma B_0 \qquad [8]$$

which may be rearranged to give the *resonance condition*

$$\nu = \frac{\gamma B_0}{2\pi} \qquad [9]$$

The NMR frequency of a nucleus is proportional to its $\gamma$ and to the strength of the field; the $2I$ allowed transitions of a spin-$I$ nucleus have identical frequencies (e.g. **Figure 1D**). Typical magnetic fields used in modern NMR spectroscopy are in the range 4.7–20.0 T, giving proton ($^1$H) resonance frequencies of 200–850 MHz, falling in the *radiofrequency* region of the electromagnetic spectrum. **Table 2** gives the gyromagnetic ratios, resonance frequencies in a 9.4 T field, and natural isotopic abundances of some commonly studied NMR nuclei.

**Table 2** Gyromagnetic ratios, NMR frequencies (in a 9.4 T field), and natural isotopic abundances of selected nuclides

| | $\gamma$ ($10^7$ $T^{-1}$ $s^{-1}$) | $\nu$ (*MHz*) | *Natural abundance (%)* |
|---|---|---|---|
| $^1$H | 26.75 | 400.0 | 99.985 |
| $^2$H | 4.11 | 61.4 | 0.015 |
| $^{13}$C | 6.73 | 100.6 | 1.108 |
| $^{14}$N | 1.93 | 28.9 | 99.63 |
| $^{15}$N | −2.71 | 40.5 | 0.37 |
| $^{17}$O | −3.63 | 54.3 | 0.037 |
| $^{19}$F | 25.18 | 376.5 | 100.0 |
| $^{29}$Si | −5.32 | 79.6 | 4.70 |
| $^{31}$P | 10.84 | 162.1 | 100 |

The intensity of the observed NMR signal depends on the difference between the numbers of nuclei in the states involved in the transition. At thermal equilibrium the fractional difference in populations, of a spin-$\frac{1}{2}$ nucleus with positive $\gamma$, is given by the Boltzmann distribution:

$$\frac{n_\alpha - n_\beta}{n_\alpha + n_\beta} = \frac{e^{\Delta} - e^{-\Delta}}{e^{\Delta} + e^{-\Delta}} \cong \Delta = \frac{\hbar\gamma B_0}{2kT} = \frac{h\nu}{2kT} \qquad [10]$$

where $\alpha$ and $\beta$ denote the $m = +\frac{1}{2}$ and $m = -\frac{1}{2}$ levels, $k$ is the Boltzmann constant, and $T$ is the temperature in kelvin. The approximation made in Equation [10] is that the NMR energy gap $\hbar\gamma B_0$ is tiny by comparison with $kT$, which is the situation in essentially all NMR experiments. For protons ($^1$H) in a 9.4 T field, $\nu = 400$ MHz, so that $\Delta = 3.2 \times 10^{-5}$, giving a population difference of about one part in 31 000.

## Chemical shifts

Although the resonance frequency of a nucleus in a magnetic field is determined principally by $\gamma$, it also depends, slightly, on the immediate surroundings of the nucleus. This effect, the *chemical shift*, is of crucial importance for chemical applications of NMR because it allows one to distinguish nuclei in different environments. For example, the $^1$H spectrum of liquid ethanol (**Figure 2**) shows clearly that there are three types of protons (methyl, methylene and hydroxyl).

The chemical shift exists because the applied magnetic field $\boldsymbol{B}_0$ causes electrons in atoms and molecules to circulate around the nuclei. Somewhat like an electric current in a loop of wire, the swirling electrons generate a small local magnetic field that

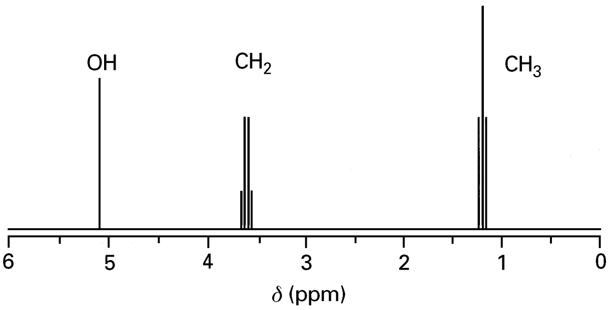



**Figure 2** Schematic [1]H NMR spectrum of liquid ethanol, $C_2H_5OH$. The three multiplets, at chemical shifts of 1.2, 3.6 and 5.1 ppm arise from the $CH_3$, $CH_2$, and OH protons. The multiplet structure (quartet for the $CH_2$, triplet for the $CH_3$) arises from the spin–spin coupling of the two sets of protons. Splittings are not normally seen from the coupling of the OH and $CH_2$ protons, because the hydroxyl proton undergoes rapid intermolecular exchange, catalysed by traces of acid or base.

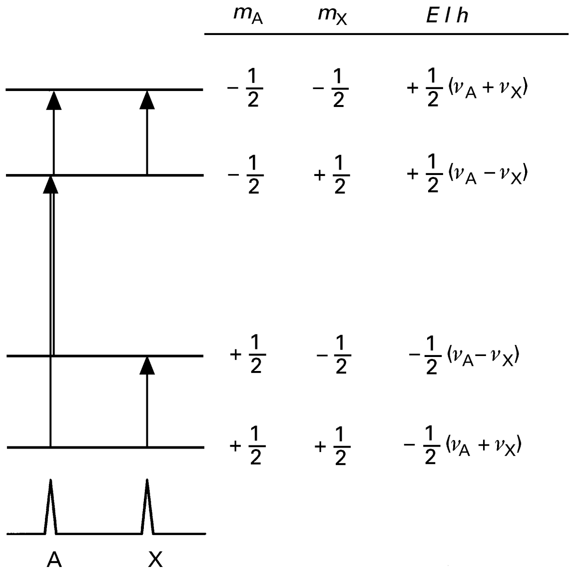



**Figure 3** Energy levels and NMR spectrum of a pair of spin-$\frac{1}{2}$ nuclei, A and X. $m_A$ and $m_X$ are the magnetic quantum numbers, $\nu_A$ and $\nu_X$ are the two resonance frequencies, and $E$ is the energy. The spin–spin coupling $J_{AX}$ is zero.

augments or opposes $\boldsymbol{B}_0$. This induced field $\boldsymbol{B}_{ind}$ is proportional in strength to $\boldsymbol{B}_0$ and, in atoms, is antiparallel to it. The net field $\boldsymbol{B}$ experienced by the nucleus is thus slightly different from $\boldsymbol{B}_0$:

$$B = B_0 - B_{ind} = B_0 - \sigma B_0 = B_0(1-\sigma) \qquad [11]$$

where the proportionality constant $\sigma$ is known as the *shielding* or *screening constant*. The resonance condition, Equation [9], thus becomes

$$\nu = \frac{\gamma B}{2\pi} = \frac{\gamma B_0}{2\pi}(1-\sigma) \qquad [12]$$

The shielding constant is determined by the electronic structure of the molecule in the vicinity of the nucleus: $\nu$ is thus characteristic of the chemical environment.

The relation between the energy levels of a pair of spin-$\frac{1}{2}$ nuclei A and X,

$$\frac{E}{h} = -\frac{m_A\gamma B_0(1-\sigma_A)}{2\pi} - \frac{m_X\gamma B_0(1-\sigma_X)}{2\pi} = -m_A\nu_A - m_X\nu_X \qquad [13]$$

and the NMR spectrum is shown in **Figure 3**.

The chemical shift is customarily quantified by means of a parameter $\delta$, defined in terms of the resonance frequencies of the nucleus of interest and of a reference compound:

$$\delta = 10^6 \times \left(\frac{\nu - \nu_{ref}}{\nu_{ref}}\right) \qquad [14]$$

$\delta$ is dimensionless and independent of $B_0$; values are usually quoted in parts per million (ppm). The most commonly used reference compound for [1]H and [13]C NMR is tetramethylsilane, $(CH_3)_4$ Si.

NMR spectra are displayed with $\delta$ increasing from right to left, with the reference compound at $\delta = 0$. As a consequence, nuclei with higher resonance frequencies (i.e. those that are less shielded) appear towards the left-hand side of the spectrum. Although spectra are now normally recorded at a fixed field strength, the old terms ‘upfield’ and ‘downfield’, meaning ‘more shielded’ and ‘less shielded’, dating from the days of field-swept NMR, are still in common use.

Chemical shifts are easily converted into frequency differences using Equation [14]. For example, the chemical shifts of the methyl and methylene signals of ethanol (**Figure 2**) are 1.2 and 3.6 ppm, respectively, giving a difference in resonance frequencies in a 9.4 T field of $(3.6 - 1.2) \times 10^{-6} \times 400$ MHz = 960 Hz.

The relative intensities of the signals in an NMR spectrum are proportional to the population differences (Eqn [10]), and therefore to the numbers of nuclei responsible for each signal. The $CH_3$, $CH_2$, and OH resonances of ethanol (**Figure 2**), for example, thus have integrated areas in the ratio 3:2:1.

## Spin–spin coupling

Magnetic nuclei interact not only with applied and induced magnetic fields, but also with one another. The result, for molecules in liquids, is a fine structure known as *spin–spin coupling*, *scalar coupling* or *J-coupling*, illustrated by the $^{1}H$ spectrum of ethanol in **Figure 2**.

The effect of spin–spin coupling on a pair of nuclear spins A and X is to shift their energy levels by amounts determined by the two magnetic quantum numbers and by the parameter that quantifies the strength of the interaction, the *spin–spin coupling constant*, $J_{AX}$. Thus, Equation [13] becomes

$$\frac{E}{h} = -m_A\nu_A - m_X\nu_X + J_{AX}m_Am_X \qquad [15]$$

For spin-$\frac{1}{2}$ nuclei, the energies are raised or lowered by $\frac{1}{4}J_{AX}$ according to whether the spins are parallel ($m_Am_X = +\frac{1}{4}$) or antiparallel ($m_Am_X = -\frac{1}{4}$). Equation [15] leads to the modified resonance condition for spin A:

$$\nu = \nu_A - J_{AX}m_X \qquad [16]$$

i.e. the resonance frequency of A is shifted from its chemical shift position by an amount that depends on the orientation of the X spin to which it is coupled. Since X has in general $2I + 1$ states, the A resonance is split into $2I + 1$ uniformly spaced lines, with equal intensities (because the different orientations of X are almost exactly equally likely). The effect that spin–spin coupling has on the energy levels of two spin-$\frac{1}{2}$ nuclei is shown in **Figure 4**. Each nucleus now has two NMR lines (a doublet).

The origin of spin–spin coupling is not the direct, through-space *dipolar interaction* of two magnetic moments: being purely anisotropic, this interaction is averaged to zero by the rapid end-over-end tumbling of molecules in liquids. Rather, the nuclei interact via the electrons in the chemical bonds that connect them. The interaction usually falls off rapidly as the number of intervening bonds increases beyond 3, so that the existence of a scalar coupling between two nuclei normally indicates that they are close neighbours in a molecular framework.

Equation [16] can easily be extended to describe more than two nuclei:

$$\nu = \nu_A - \sum_{i \neq A} J_{Ai}m_i \qquad [17]$$

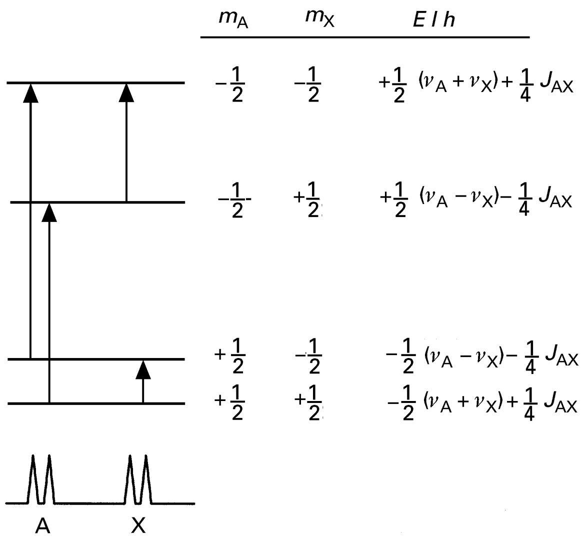



**Figure 4** Energy levels and NMR spectrum of a pair of spin-$\frac{1}{2}$ nuclei, A and X. $m_A$ and $m_X$ are the magnetic quantum numbers, $\nu_A$ and $\nu_X$ are the two resonance frequencies, $J_{AX}$ is the spin–spin coupling constant, and $E$ is the energy.

where the sum runs over all spins to which A has an appreciable coupling. If A is coupled to $N$ identical spin-$\frac{1}{2}$ nuclei (e.g. the three protons in a methyl group), it can be seen from Equation [17] that its resonance is split into $N + 1$ equally spaced lines with relative intensities given by the binomial coefficients

$$\binom{N}{i} = \frac{N!}{i!(N-i)!}, \quad i = 0, 1, 2, \cdots N \qquad [18]$$

Thus, the $CH_2$ and $CH_3$ resonances in ethanol (**Figure 2**) are respectively a 1:3:3:1 quartet and a 1:2:1 triplet.

This discussion of the *multiplet* (i.e. doublet, triplet, quartet, ...) structure arising from spin–spin coupling is valid in the *weak coupling* limit, i.e. when the difference in resonance frequencies of the coupled nuclei $|\nu_A - \nu_X|$ is much larger than their interaction $|J_{AX}|$. When this is not the case (*strong coupling*), the positions and intensities of the lines are modified, as illustrated in **Figure 5**. The origin of these effects lies in the NMR transition probabilities. As the coupling becomes stronger, the outer line of each doublet in **Figure 5** becomes weaker relative to the inner line. In the limit that the chemical shift difference is zero, the transitions leading to the two outer lines become completely forbidden, and the two inner lines coincide, so that only a single line is

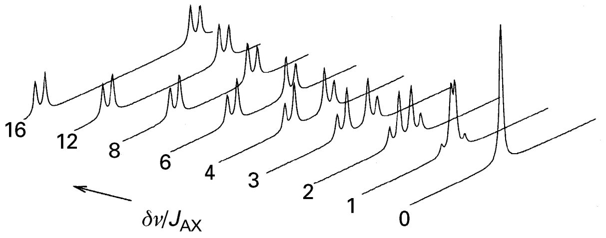



**Figure 5** Calculated NMR spectra of a pair of spin-½ nuclei for a range of $\delta\nu = \nu_A - \nu_X$ values between $16J_{AX}$ and zero.

observed. This is a general result: *spin–spin interactions between protons in identical environments do not lead to observable splittings.*

## Vector model of NMR

Considerable insight into the operation of simple NMR experiments may be derived from a straightforward *vector model*. It relies on the fact that while the individual nuclear magnetic moments behave quantum mechanically, the net magnetization of a large collection of nuclear spins obeys classical mechanics.

The motion of a classical magnetic moment $\boldsymbol{M}$, possessing angular momentum, in a magnetic field $\boldsymbol{B}$ is described by the differential equation

$$\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t} = -\gamma \boldsymbol{B} \times \boldsymbol{M} \qquad [19]$$

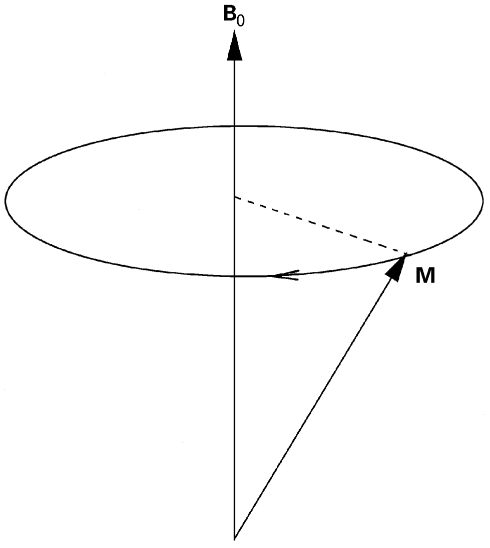



**Figure 6** The motion of a magnetization vector $\boldsymbol{M}$ in a magnetic field $\boldsymbol{B}_0$. $\boldsymbol{M}$ precesses around the field direction rather like the axis of a spinning gyroscope.

or equivalently,

$$\begin{aligned}\frac{\mathrm{d}M_{x'}}{\mathrm{d}t} &= -\gamma B_{y'} M_{z'} + \gamma B_{z'} M_{y'} \\ \frac{\mathrm{d}M_{y'}}{\mathrm{d}t} &= +\gamma B_{x'} M_{z'} - \gamma B_{z'} M_{x'} \\ \frac{\mathrm{d}M_{z'}}{\mathrm{d}t} &= -\gamma B_{x'} M_{y'} + \gamma B_{y'} M_{x'}\end{aligned} \qquad [20]$$

where $\boldsymbol{B} \times \boldsymbol{M}$ is the vector product of $\boldsymbol{B} = (B_{x'}, B_{y'}, B_{z'})$ and $\boldsymbol{M} = (M_{x'}, M_{y'}, M_{z'})$, and the $(x', y', z')$ coordinate system is called the *laboratory frame*.

These expressions describe the precession of $\boldsymbol{M}$ around $\boldsymbol{B}$ at *angular* frequency $\omega = \gamma B$, as may be seen by taking $\boldsymbol{B} = \boldsymbol{B}_0$, along the $z'$-axis:

$$\begin{aligned}\frac{\mathrm{d}M_{x'}}{\mathrm{d}t} &= +\gamma B_0 M_{y'} \\ \frac{\mathrm{d}M_{y'}}{\mathrm{d}t} &= -\gamma B_0 M_{x'} \\ \frac{\mathrm{d}M_{z'}}{\mathrm{d}t} &= 0\end{aligned} \qquad [21]$$

(see **Figure 6**). This motion is known as *Larmor precession*, and it occurs at the NMR frequency of the nuclear spins in the field $\boldsymbol{B}_0$:

$$\omega_0 = \gamma B_0 = 2\pi\nu \qquad [22]$$

An NMR experiment involves the application of a brief, intense burst of radiofrequency radiation, known as a 'pulse', along, say, the $x'$ axis in the laboratory frame. The frequency of this field, $\omega_{\mathrm{RF}}$ is very close to the Larmor frequency $\omega_0$. Regarding this linearly oscillating field as the sum of two counter-rotating fields, we may ignore the component that rotates in the opposite sense to the Larmor precession because, being $2\omega_{\mathrm{RF}}$ off-resonance, it has a negligible effect on the spins. The other component is

$$\boldsymbol{B}_1 = (B_1 \cos \omega_{\mathrm{RF}} t, -B_1 \sin \omega_{\mathrm{RF}} t, 0) \qquad [23]$$

The nuclear spins thus experience the sum of two magnetic fields: a strong static field $\boldsymbol{B}_0$ along the $z'$ axis, and a much weaker, time-dependent field $\boldsymbol{B}_1$

rotating in the $x'y'$ plane. $\boldsymbol{M}$ therefore precesses around the time-dependent vector sum of $\boldsymbol{B}_0$ and $\boldsymbol{B}_1$ (**Figure 7A**). To make this complicated motion easier to visualize, Equation [20] is transformed into the *rotating frame* $(x, y, z)$, a coordinate system rotating around the $z'$ axis at frequency $\omega_{RF}$, in which the radiofrequency field appears stationary. In this frame, the components of the bulk magnetization are

$$\begin{aligned} M_x &= M_{x'} \cos \omega_{RF}t - M_{y'} \sin \omega_{RF}t \\ M_y &= M_{x'} \sin \omega_{RF}t + M_{y'} \cos \omega_{RF}t \\ M_z &= M_{z'} \end{aligned} \qquad [24]$$

Differentiating Equations [24], and using Equations [20] with $\boldsymbol{B} = (B_1 \cos\omega_{RF}\, t, -B_1 \sin\omega_{RF}\, t, B_0)$ gives

$$\begin{aligned} \frac{dM_x}{dt} &= +\gamma\Delta BM_y \\ \frac{dM_y}{dt} &= +\gamma B_1 M_z - \gamma\Delta BM_x \\ \frac{dM_z}{dt} &= -\gamma B_1 M_y \end{aligned} \qquad [25]$$

or, more compactly,

$$\frac{d\boldsymbol{M}}{dt} = -\gamma \boldsymbol{B}_{eff} \times \boldsymbol{M} \qquad [26]$$

where $\boldsymbol{B}_{eff} = (B_1, 0, \Delta B)$, and $\Delta B = B_0 - \omega_{RF}/\gamma$. Equation [26] describes the precession of $\boldsymbol{M}$ about a static field $\boldsymbol{B}_{eff}$ at frequency $\gamma \boldsymbol{B}_{eff} = \gamma\sqrt{B_1^2 + \Delta B^2}$ in the rotating frame (**Figure 7B**). $\gamma\Delta B = \omega_0 - \omega_{RF} = \Omega$ is the offset of the radiofrequency field from resonance. To include chemical shifts, $\gamma B_0$ should be replaced by $\gamma B_0(1 - \sigma)$.

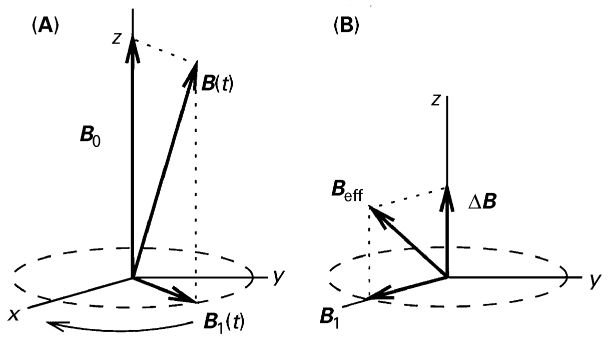



**Figure 7** The magnetic fields present in an NMR experiment in (A) the laboratory frame and (B) the rotating frame. $\boldsymbol{B}_0$ is the strong static field, $\boldsymbol{B}_1$ is the much weaker oscillating radiofrequency field, $\Delta\boldsymbol{B}$ and $\boldsymbol{B}_{eff}$ are respectively, the offset and effective fields in the rotating frame, and $\boldsymbol{B}(k)$ is the resultant of $\boldsymbol{B}_0$ and $\boldsymbol{B}_1$ in the laboratory frame.

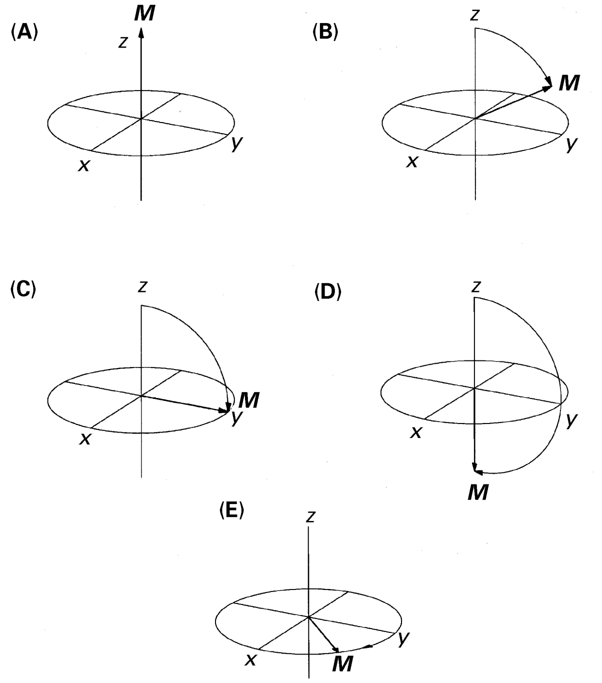



**Figure 8** The effect of radiofrequency pulses (in the rotating frame). (A) At thermal equilibrium, the net magnetization of the sample is parallel to the $\boldsymbol{B}_0$ direction. (B) A pulse along the *x* axis, whose strength $\boldsymbol{B}_1$ is much greater than the offset field $\Delta\boldsymbol{B}$, causes $\boldsymbol{M}$ to rotate in the *yz* plane at angular frequency $\gamma B_1$. (C) A 90° pulse, of duration $t_p$, $\gamma B_1 t_p = \pi/2$), rotates the magnetization from the 'north pole' (*z* axis) to the 'equator' (*y* axis) (D) A 180° pulse, of duration $t_p$ ($\gamma B_1 t_p = \pi$), rotates the magnetization from the 'north pole' to the 'south pole' ($-z$ axis). (E) Following a 90° pulse, the magnetization precesses around the 'equator' in the rotating frame at frequency $\Omega = \gamma\Delta B$. Relaxation is ignored throughout.

## Radiofrequency pulses

At equilibrium, in the absence of a radiofrequency field, the bulk magnetization of the sample $\boldsymbol{M}_0$ is parallel to the $\boldsymbol{B}_0$ direction ($z$ axis) with a magnitude proportional to the population difference ($n_\alpha - n_\beta$, for a spin-$\frac{1}{2}$ nucleus). If the radiofrequency field strength is much larger than the resonance offset ($B_1 >> \Delta B$) then $\boldsymbol{B}_{eff} \cong \boldsymbol{B}_1$ and the effective field lies along the $x$ axis in the rotating frame. The pulse therefore causes $\boldsymbol{M}_0$ to rotate in the $yz$ plane at frequency $\gamma B_1$ (**Figure 8A** and **8B**). In this way a short, intense monochromatic burst of radiofrequency

radiation can excite spins uniformly over a range of resonance frequencies, provided their offset frequencies Ω are much smaller than $\gamma B_1$. If the field is switched off after a time $t_p$, given by $\gamma B_1 t_p = \pi/2$, $\boldsymbol{M}$ is turned through 90° and is left along the $y$ axis (**Figure 8C**). A radiofrequency pulse with this property is known as a 90° pulse. If $t_p$ is twice this duration, the magnetization is *inverted* (a 180° pulse, **Figure 8D**); this is equivalent to exchanging the $n_\alpha$ and $n_\beta$ populations of a spin-$\frac{1}{2}$ nucleus.

### Free precession

Equation [26] may also be used to predict what happens *after* a 90° pulse. Setting $B_1 = 0$, the effective field is $\boldsymbol{B}_{\text{eff}} = (0, 0, \Delta B)$ and $\boldsymbol{M}$ precesses in the $xy$ plane at angular frequency $\gamma\,\Delta B = \Omega$, i.e. at the offset frequency determined by $\omega_{\text{RF}}$ and the chemical shift (**Figure 8E**):

$$\begin{aligned} M_x &= M_0 \sin \Omega t \\ M_y &= M_0 \cos \Omega t \\ M_z &= 0 \end{aligned} \qquad [27]$$

where $t$ is now the time after the end of the pulse. When several nuclei with different chemical shifts have been excited by the pulse, the $xy$ magnetization of the sample is the sum of several oscillating terms of the form of Equation [27].

### Free induction decay

Up to this point it has been assumed that the nonequilibrium state produced by the radiofrequency pulse does not relax back towards equilibrium. This is a reasonable approximation during the very short pulse. However, to describe the behaviour of the spins during the period of free precession that follows the pulse, relaxation must be included. This is traditionally done by allowing $M_x$ and $M_y$ to decay exponentially back to zero with a time constant $T_2$, while $M_z$ grows back to $M_0$ with a time constant $T_1$:

$$\begin{aligned} \frac{\mathrm{d}M_x}{\mathrm{d}t} &= +\gamma \Delta B\, M_y - \frac{M_x}{T_2} \\ \frac{\mathrm{d}M_y}{\mathrm{d}t} &= +\gamma B_1 M_z - \gamma \Delta B M_x - \frac{M_y}{T_2} \\ \frac{\mathrm{d}M_z}{\mathrm{d}t} &= -\gamma B_1 M_y - \frac{(M_z - M_0)}{T_1} \end{aligned} \qquad [28]$$

$T_1$ and $T_2$ are the *spin–lattice* and the *spin–spin relaxation times*. These expressions are known as the *Bloch equations*. With relaxation included, Equation [27] becomes

$$\begin{aligned} M_x(t) &= M_0 \sin \Omega t \exp(-t/T_2) \\ M_y(t) &= M_0 \cos \Omega t \exp(-t/T_2) \\ M_z(t) &= M_0[1 - \exp(-t/T_1)] \end{aligned} \qquad [29]$$

The two components $M_x$ and $M_y$ represent the detectable signal in an NMR experiment – the *free induction decay* (**Figure 9**). Fourier transformation of the free induction decay gives the NMR spectrum.

## Spin relaxation

Relaxation processes allow nuclear spins to return to equilibrium following a disturbance, e.g. a

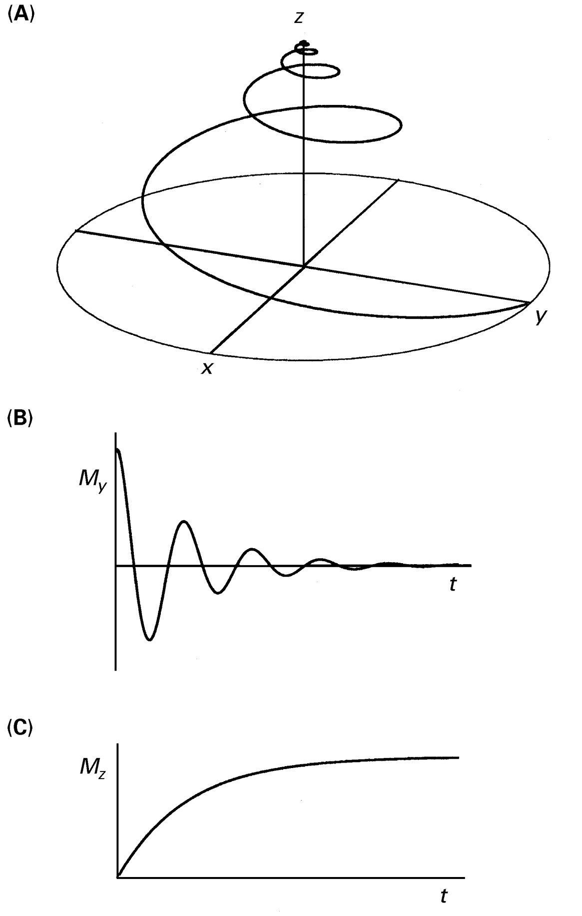



**Figure 9** Following a 90° pulse, the magnetization precesses around the $z$ axis and at the same time returns to its equilibrium position at the 'north pole' (A). The transverse components of $\boldsymbol{M}$ decay to zero with time constant $T_2$, the spin–spin relaxation time (B). The $z$ component of $\boldsymbol{M}$ grows back to $\boldsymbol{M}_0$ with time constant $T_1$, the spin–lattice relaxation time (C).

radiofrequency pulse. The relaxation times $T_1$ and $T_2$ characterize the relaxation of, respectively, the *longitudinal* and *transverse* components of the magnetization $\boldsymbol{M}$, respectively parallel and perpendicular to $\boldsymbol{B}_0$. Equivalently, $T_1$ is the time constant for the return to equilibrium of the populations of the spin states, while $T_2$ is the time constant for the dephasing of the *coherence* between spin states. In the absence of any significant spatial inhomogeneity of $\boldsymbol{B}_0$, or other sources of line broadening such as chemical exchange, the width of the NMR line (in hertz) is $1/\pi T_2$.

Spin–lattice relaxation is caused by randomly fluctuating local magnetic fields. A common source of such fields is the dipolar interaction between pairs of nuclei, modulated by molecular tumbling in a liquid. The component of these fields that oscillates at the resonance frequency can induce transitions between the spin states, so transferring energy between the spin system and the 'lattice' (i.e. everything else) and bringing the spins into equilibrium with their surroundings. In the simplest case, $T_1$ depends on the mean square strength of the local fields $\langle B_{\mathrm{loc}}^2\rangle$, and the intensity of the fluctuations at the resonance frequency $\omega_0$

$$\frac{1}{T_1} = \gamma^2\langle B_{\mathrm{loc}}^2\rangle J(\omega_0) \qquad [30]$$

where

$$J(\omega) = \frac{2\tau_{\mathrm{c}}}{1+\omega^2\tau_{\mathrm{c}}^2} \qquad [31]$$

is the *spectral density* function, and $\tau_{\mathrm{c}}$ is the *rotational correlation time* (roughly the average time the molecule takes to rotate through 90°).

Spin–spin relaxation has two contributions:

$$\frac{1}{T_2} = \tfrac{1}{2}\gamma^2\langle B_{\mathrm{loc}}^2\rangle J(\omega_0) + \tfrac{1}{2}\gamma^2\langle B_{\mathrm{loc}}^2\rangle J(0) \qquad [32]$$

The first is closely related to spin–lattice relaxation, and arises from the finite lifetime of the spin states, through the uncertainty principle. The second term is due to the loss of coherence caused by local fields of very low frequency (hence the $J(0)$ factor), which augment or oppose $\boldsymbol{B}_0$ and so give rise to a spread of resonance frequencies, and hence the dephasing of transverse magnetization. **Figure 10** shows the dependence of $T_1$ and $T_2$ on $\tau_{\mathrm{c}}$.

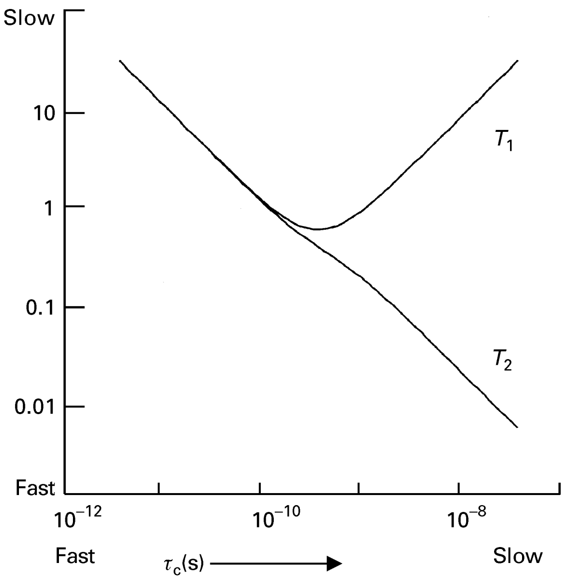



**Figure 10** The dependence of $T_1$ and $T_2$ on the rotational correlation time $\tau_{\mathrm{c}}$, using $\gamma^2\langle B_{\mathrm{loc}}^2\rangle = 4.5 \times 10^9\ \mathrm{s}^{-2}$ and $\omega_0/2\pi = 400$ MHz. The units for the vertical axis are seconds.

Relaxation times contain information on both $J(\omega)$ (i.e. on molecular motion) and $\langle B_{\mathrm{loc}}^2\rangle$, (i.e. on molecular structure via, for example, the $r^{-3}$ distance dependence of the dipolar interaction). A further relaxation phenomenon that provides important information on internuclear distances is the *nuclear Overhauser effect*.

## Chemical exchange

In addition to chemical shifts, spin–spin coupling and spin relaxation, NMR spectra are affected by, and may be used to study, chemical and conformational equilibria. Consider an equilibrium

$$\mathrm{A} \rightleftharpoons \mathrm{B} \qquad [33]$$

which exchanges the chemical shifts of two nuclei, with equal forward and backward rate constants, $k$. At low temperature, the NMR spectrum comprises two sharp resonances at frequencies $\nu_{\mathrm{A}}$ and $\nu_{\mathrm{B}}$. As the temperature is raised, the following sequence of events occurs: the two lines broaden and move towards one another until they coalesce into a broad flat-topped line which then narrows into a sharp single resonance at the average chemical shift $\frac{1}{2}(\nu_{\mathrm{A}} + \nu_{\mathrm{B}})$ (**Figure 11**).

The mid-point of this process, when the two lines just merge into one, occurs when

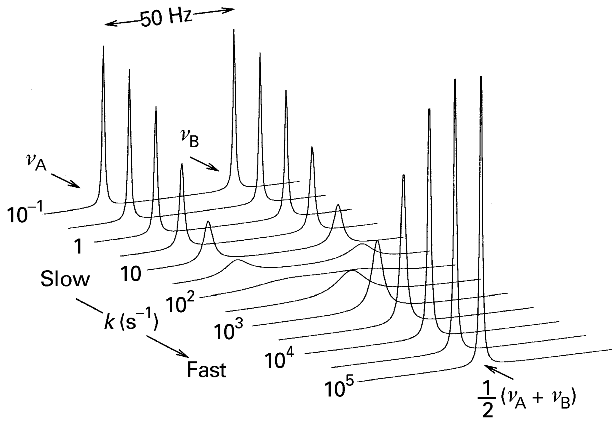



**Figure 11** Calculated NMR spectra for a pair of nuclei exchanging between two sites with equal populations. Spectra are shown for a range of values of the exchange rate $k$. The difference in resonance frequencies of the two sites, $\delta\nu$, is 50 Hz.

$$k = \frac{\pi(\nu_A - \nu_B)}{\sqrt{2}} = \frac{\pi\delta\nu}{\sqrt{2}} \quad [34]$$

For *slow exchange*, the exchange broadening of the two separate resonances is

$$\Delta\nu = \frac{k}{\pi} \quad [35]$$

while for *fast exchange*, the single line has an extra width

$$\Delta\nu = \frac{\frac{1}{2}\pi(\delta\nu)^2}{k} \quad [36]$$

Related but more complex expressions are found if the forward and backward rate constants differ, or if there are more than two exchanging species.

## List of symbols

$\boldsymbol{B}$ = magnetic field vector; $B$ = magnitude of $\boldsymbol{B}$; $\hbar$ = Planck constant $(h)/2\pi$; $\boldsymbol{I}$ = nuclear spin angular momentum vector; $I$ = nuclear spin quantum number; $J$ = spin–spin coupling constant; $m$ = nuclear magnetic quantum number; $\boldsymbol{M}$ = classical (macroscopic) magnetization vector; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $\gamma$ = gyromagnetic ratio; $\mu$ = nuclear magnetic moment; $\mu_z = z$ component; $\delta$ = chemical shift; $\sigma$ = shielding (screening) constant; $\tau_c$ = rotational correlation time; $\omega$ = angular frequency; $\omega_0$ = Larmor frequency.

*See also:* **Chemical Exchange Effects in NMR; Fourier Transformation and Sampling Theory; NMR Relaxation Rates; NMR Spectrometers; Nuclear Overhauser Effect; Parameters in NMR Spectroscopy, Theory of.**

## Further reading

Carrington A and McLachlan AD (1967) *Introduction to Magnetic Resonance.* New York: Harper and Row.

Ernst RR, Bodenhausen G and Wokaun A (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions.* Oxford: Clarendon Press.

Freeman R (1997) *A Handbook of Nuclear Magnetic Resonance*, 2nd ed. Harlow: Longman.

Günther H (1995) *NMR Spectroscopy*, 2nd edn. Chichester: Wiley.

Harris RK (1983) *Nuclear Magnetic Resonance Spectroscopy.* London: Pitman.

Hore PJ (1995) *Nuclear Magnetic Resonance.* Oxford: Oxford University Press.

McLauchlan, KA (1972) *Magnetic Resonance*, Oxford: Clarendon Press.

Sanders JKM and Hunter BK (1993), *Modern NMR Spectroscopy*, 2nd edn. Oxford: Oxford University Press.

# NMR Pulse Sequences

**William F Reynolds**, University of Toronto, Ontario, Canada

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Theory**

## Introduction

The single most important development in nuclear magnetic resonance (NMR) spectroscopy since the initial observation of the NMR phenomenon in bulk phases in 1945 was undoubtedly the introduction of pulse Fourier transform NMR by Anderson and Ernst. This technique provided greatly increased sensitivity per unit time, making it feasible to obtain spectra for low sensitivity/low abundance nuclei such as $^{13}C$. More importantly, it allowed the development of a wide variety of sophisticated and powerful multipulse experiments which have revolutionized the use of NMR spectroscopy in studies of molecular structure and dynamics. This article provides an overview of pulse sequence experiments. Many individual experiments are discussed in other articles.

## The classical vector model of NMR and the basic one-pulse Fourier transform experiment

Many NMR pulse sequences can be described either by a classical model describing the motions of magnetic vectors or by quantum mechanical models of different levels of sophistication. The attractive feature of the classical vector model is that it provides simple physical pictures of many of the basic pulse sequences. However, it does not work for many multipulse experiments that involve multiple quantum coherence. These experiments can only be described by quantum mechanical methods. Because of the insights which the vector model provides into many of the basic sequences, I will use this model wherever possible.

The fundamental magnetic properties of nuclei are well described elsewhere. I will begin with the bulk magnetization vector $\boldsymbol{M}$ for a series of nuclei of the same type. This is parallel to the external magnetic field $\boldsymbol{B}_0$ and is the resultant of individual magnetic moment vectors $\mu$, precessing about $\boldsymbol{B}_0$ with the Larmor angular velocity

$$\omega = -\gamma B_0 \qquad [1]$$

where $\gamma$ is the magnetogyric ratio of the nucleus (**Figure 1**). For a nucleus with spin quantum number $I = \frac{1}{2}$ $\boldsymbol{M}$ actually results from the slight excess of nuclei in the $\alpha$ spin state ($m_I = +\frac{1}{2}$) over those in the $\beta$ spin state ($m_I = -\frac{1}{2}$).

Now consider the effect of a 'pulse' of electromagnetic radiation of frequency corresponding to the Larmor frequency $\omega/2\pi$. This is applied so that the oscillating magnetic component of the pulse is in a plane at right angles to $\boldsymbol{B}_0$. This oscillating component can be resolved into two rotating components of angular velocity $\pm 2\pi\nu$. Only the component rotating in the same direction as the magnetic moments need be considered since the opposite component has no net effect on the nuclear magnetization. The former

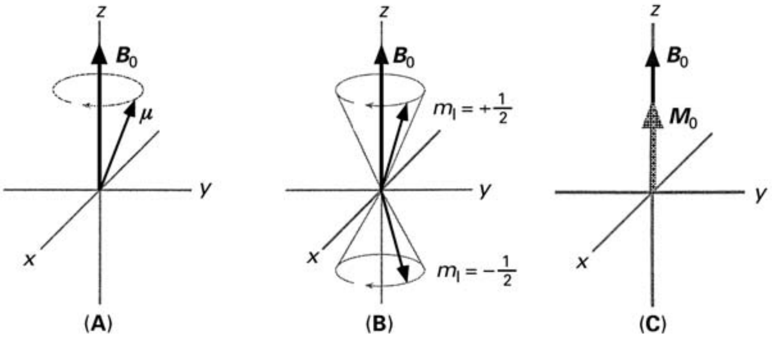



**Figure 1** (A) The precession of an individual magnetic moment $\mu$ about the external magnetic field $\boldsymbol{B}_0$. (B) The precession of magnetic moments in the $\alpha$ ($m_I = +\frac{1}{2}$) and $\beta$ ($m_I = -\frac{1}{2}$) spin states. (C) The resultant magnetic moment $\boldsymbol{M}$ of a large number of nuclei of the same kind, reflecting the small excess population of nuclei in the $\alpha$ spin state.

component is represented by a magnetization vector, $\boldsymbol{B}_1$, rotating in the $x$–$y$ plane at frequency $\nu$. However, to simplify the visualization, the Cartesian coordinate system is also assumed to be rotating at frequency $\nu$, called the rotating frame model (**Figure 2**). This allows us to concentrate on frequency differences rather than absolute frequencies in considering NMR experiments.

During the pulse, the individual magnetic moments and consequently the bulk magnetization vector precess about $\boldsymbol{B}_1$ with angular velocity $\gamma B_1$. If $\boldsymbol{B}_1$ is taken as defining the $x$-axis in the rotating frame, $\boldsymbol{M}$ rotates towards the $y$-axis through an angle:

$$\alpha(\text{radians}) = \gamma B_1 \tau \qquad [2]$$

where $\tau$ is the pulse duration in seconds. Thus if the duration of the pulse is just sufficient to rotate $\boldsymbol{M}$ through $\pi/2$ radians, it is called a 90° pulse. The resultant magnetization generated in the $x$–$y$ plane can then be detected by a receiver.

Now consider the one-pulse Fourier transform experiment. The pulse sequence is illustrated in **Figure 3**. The pulse does not excite a single frequency but rather a range of frequencies whose width (in Hz) is inversely proportional to the pulse duration (in s). The frequency excitation profile is provided by taking the Fourier transform of the time profile of the pulse (**Figure 3**). Quadrature detection distinguishes positive and negative frequencies, allowing one to position the transmitter frequency at the midpoint of the spectral window. Modern high resolution spectrometers typically have 90° pulses of duration 10 μs or less. While this allows excitation over a 200 kHz spectral window, it is important to have near uniform excitation over the entire spectral window. A 10 μs pulse provides near uniform excitation over ~25 kHz, which is adequate for most high resolution applications. However, solid state spectra have much wider spectral windows, requiring much shorter pulses.

After the pulse generates $x$–$y$ magnetization, the return of this magnetization to equilibrium is sampled as a function of time. This response is called the free induction decay (FID) signal. In the vector model, it can be regarded as the resultant of a series of individual magnetization vectors, each processing in the $x$–$y$ plane at some frequency $\Delta\nu$ relative to the transmitter frequency and decaying exponentially with a time constant $T_2$, characteristic of the return to equilibrium of $x$–$y$ magnetization. Each vector corresponds to a specific signal in the frequency spectrum, and thus the Fourier transform of the FID yields the frequency spectrum:

$$F(\nu) = \int_{-\infty}^{+\infty} f(t)\exp(2\pi i \nu t)\mathrm{d}t$$
$$= \int_{-\infty}^{+\infty} f(t)[\cos(2\pi\nu t) + \mathrm{i}\,\sin(2\pi\nu t)]\mathrm{d}t \qquad [3]$$

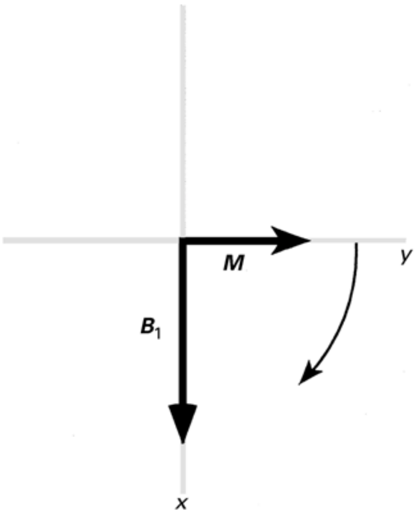



**Figure 2** The rotating frame coordinate system. The coordinate system is assumed to be rotating at the same frequency as $\boldsymbol{B}_1$ and consequently $\boldsymbol{B}_1$ appears to be stationary along the $x$-axis. The $x,y$ magnetization $\boldsymbol{M}$, generated after a pulse, will be stationary along the $y$-axis if the Larmor precession frequency is equal to the pulse frequency or rotating at a frequency $\Delta\nu$, corresponding to the frequency difference.

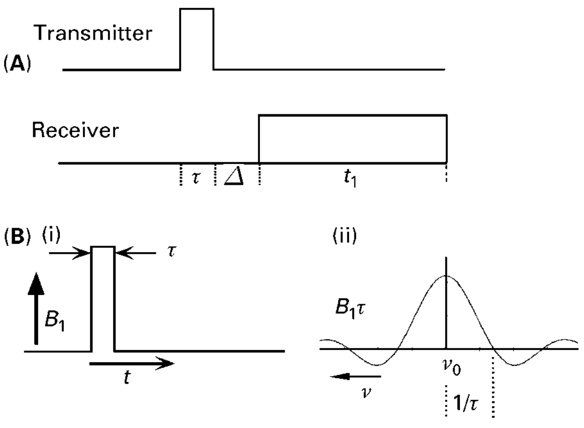



**Figure 3** (A) The basic pulse Fourier transform sequence; $\tau$ represents the pulse duration and $\Delta$ is a small delay, comparable to $\tau$, to ensure that the pulse is not detected by the receiver. Note that in this and subsequent pulse sequences, the duration of the pulse is exaggerated. The actual pulse duration is ~10 μs compared with an acquisition time, $t_1$, of ~1 s. (B) (i) The time profile of the pulse and (ii) the frequency excitation profile due to the pulse. The frequency profile is the Fourier transform of the time profile.

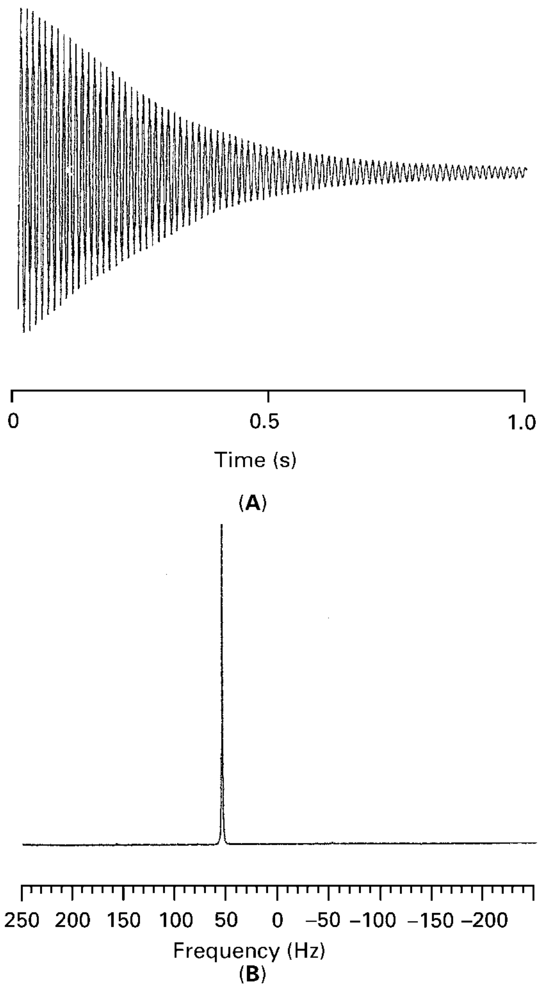
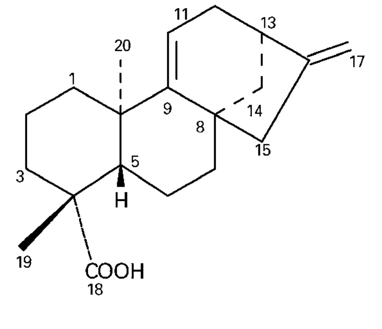



**Figure 4** The FID time signal (A) and resultant $^{1}H$ frequency spectrum (B) for a single off-resonance peak.

This is illustrated in **Figure 4** for a spectrum of a single off-resonance peak. **Figure 5** shows the FID and frequency spectrum for the aliphatic region of kauradienoic acid [1].




[1]

The two basic types of signals that can be detected in an NMR experiment (**Figure 6**) are an absorption signal (owing to detection of a signal at right angles to $\boldsymbol{B}_1$) and a dispersion signal (owing to a signal parallel to $\boldsymbol{B}_1$), corresponding to the cosine and sine terms in Equation [3]. The phase for an on-resonance (i.e. $\Delta\nu = 0$) peak can easily be adjusted to give an absorption signal via a zero frequency phase adjustment.

However, off-resonance peaks undergo additional phase shifts due to vector evolution during the finite pulse and the delay between the pulse and gating on the receiver (see **Figure 3**). For example, for a total time before acquisition of 20 µs, a peak at $\Delta\nu = 10\ 000$ Hz will rotate through an angle:

$$\begin{aligned}\alpha &= 2\pi(2.0 \times 10^{-5})(1.0 \times 10^{4}) \text{ radians} \\ &= 0.40\pi \text{ radians} = 72^\circ\end{aligned} \qquad [4]$$

introducing a significant dispersive component. Fortunately, this phase shift varies linearly with $\Delta\nu$ and thus can be corrected by applying a phase correction which varies linearly with frequency.

The acquisition time, $t_1$, is determined by the time required for $x$–$y$ magnetization to decay to near zero as well as by the desired data point resolution. Typical values for one-dimensional spectra range from 0.5 to 5 s. The ability to excite and acquire all signals for a given nucleus simultaneously provides a major sensitivity advantage over the older continuous wave (CW) method which involved slowly sweeping through the spectral window, exciting one signal at a time. Typically, one can acquire at least 100 FID signals in the time taken to acquire a CW spectrum. Since the signal-to-noise increases as the square root of the number of scans, this provides at least a 10-fold increase in sensitivity.

However, the acquisition of multiscan spectra introduces a new problem. Ideally, $\boldsymbol{M}$ should have returned to its equilibrium position along the $+z$-axis before the next pulse. Otherwise, the residual magnetization will fractionally decrease with each scan, a phenomenon known as saturation. Compounding the problem is the fact that the time constant for return to equilibrium along the $z$-axis, $T_1$, can be longer than $T_2$ (see below). One solution is to introduce an additional relaxation delay between the end of each acquisition and the next pulse. The second is to use a shorter pulse duration so that the rotation angle of $\boldsymbol{M}$, $\alpha$, is $< 90°$. Richard Ernst conclusively demonstrated that the second approach gives a superior signal-to-noise ratio. The ideal pulse flip angle, $\alpha_E$, called the Ernst angle, is given by:

$$\cos \alpha_E = \exp(-t_1/T_1) \qquad [5]$$

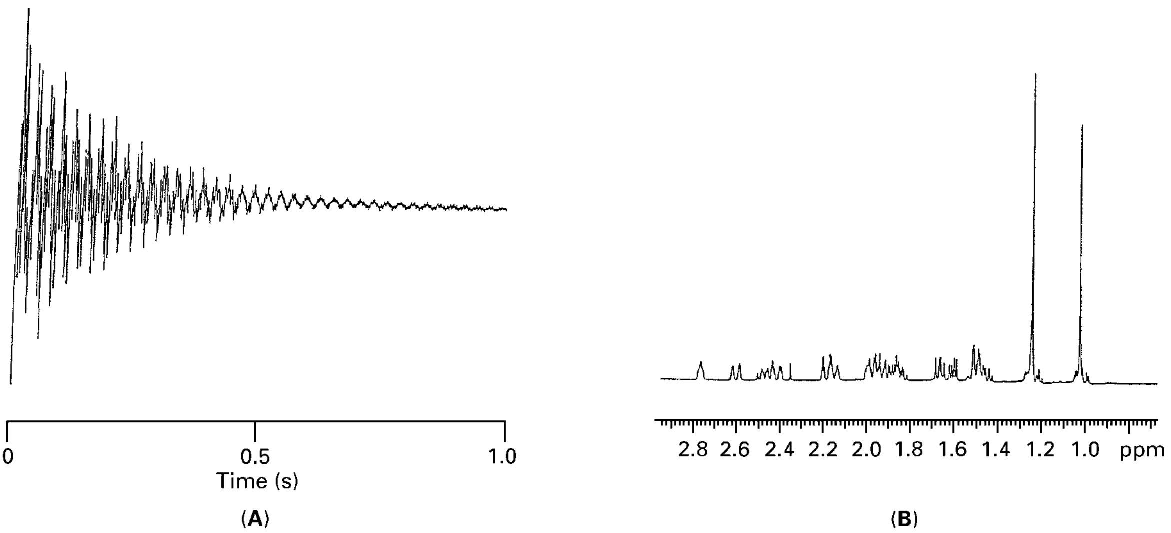



**Figure 5** The FID time signal (A) and [1]H frequency spectrum (B) for the aliphatic region of kauradienoic acid [1]. The scale along the bottom of the frequency spectrum is the δ scale [chemical shift in parts per million relative to $(CH_3)_4Si$].

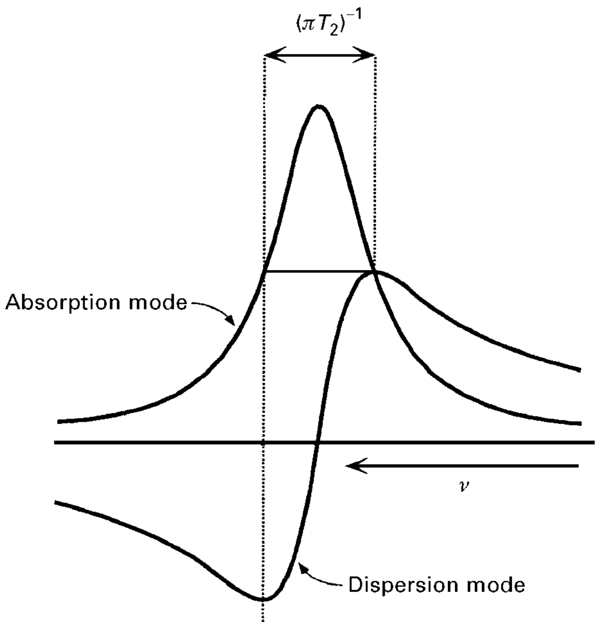



**Figure 6** Comparison of absorption (*v* mode) and dispersion (*u* mode) signal shapes. Spectra are usually phase corrected to give pure absorption mode peaks.

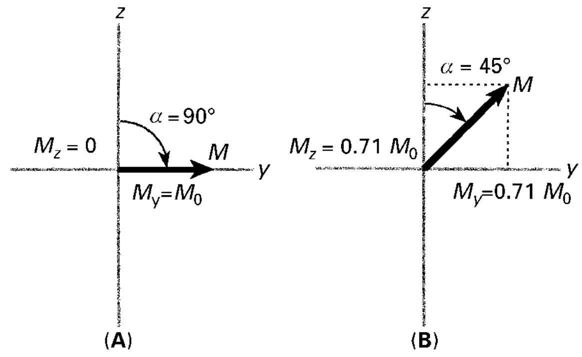



**Figure 7** Comparison of $z$ ($M_z$) and $y$ ($M_y$) magnetization immediately following (A) a $90^\circ_x$ pulse and (B) a $45^\circ_x$ pulse. The latter generates 71% of the amount of $M_y$ magnetization (and therefore 71% of the signal) while retaining 71% of equilibrium $z$ magnetization, compared with 0° $z$ magnetization after a 90° pulse. This allows most, if not all, of the equilibrium $z$ magnetization to be restored during the acquisition time, $t_1$, after a 45° pulse while a 90° pulse typically will require a lengthy delay after $t_1$ to restore equilibrium $z$ magnetization. If $T_1$ is very long compared with $t_1$, an even smaller pulse angle must be used (see Eqn [5]).

This is illustrated in **Figure** 7, using a simple trigonometric argument. However, $T_1$ may be significantly different for different peaks in a spectrum (e.g. a $^{13}C$ spectrum of a molecule containing protonated and non-protonated carbons). This requires a compromise choice of $\alpha$ and relative peak areas may no longer be quantitative.

Finally, the analogue voltage signal detected by the receiver must be digitized for computer storage and processing. This puts some constraints on data acquisition, as discussed in any of the texts listed in the Further reading section.

## Measurement of $T_1$ and $T_2$ relaxation times

The classical equations for the return of magnetization to equilibrium along the $z$- and $y$-axes are respectively:

$$\frac{dM_z}{dt} = -(M_z - M_0)/T_1 \text{ and } \frac{dM_y}{dt} = -M_y/T_2 \quad [6]$$

where $M_0$ is the magnitude of equilibrium $z$ magnetization. The spin–lattice or longitudinal relaxation time, $T_1$, reflects the effect of the component of the random fluctuating magnetic field (arising from the thermal motion of dipoles in the sample) at the Larmor frequency. However, the spin–spin or transverse relaxation time, $T_2$, is also affected by static magnetic field components, including any inhomogeneity of the magnetic field over the region of the sample. Consequently, transverse relaxation is in principle faster than longitudinal relaxation, i.e. $T_2$ can be smaller than $T_1$.

The inversion–recovery sequence can be used to measure $T_1$ (**Figure** 8). The 180° pulse inverts the equilibrium magnetization **M**. During the delay $t_1$, the magnetization begins to return to equilibrium. A $90°_x$ pulse then samples the magnetization remaining after $t_1$. Since the final pulse is a 90° pulse, it is necessary to include a relaxation delay, $\Delta$, to allow for return to equilibrium between scans. Ideally $\Delta \geq 5T_1$. However, it has been found that accurate values of $T_1$ can still be obtained using shorter values of $\Delta$, known as the fast inversion–recovery method. After a sufficient number of scans, $n$, have been collected to achieve adequate signal-to-noise, the experiment is repeated, systematically varying $t_1$ from small values out to $\sim 2T_1$. The intensity of each peak exponentially returns to equilibrium as $t_1$ increases (**Figure** 8). $T_1$ can be determined from a least-squares fit of the equation

$$\ln[S_\infty - S(t_1)] = \ln 2 + \ln S_\infty - t_1/T_1 \qquad [7]$$

This gives a linear plot of slope $-1/T_1$. However, this approach is particularly sensitive to errors in $S_\infty$ (obtained with $t_1 \geq 5T_1$). The alternative, more reliable, approach is to carry out an exponential fit to each relaxation curve. This does not require an accurate value of $S_\infty$ and is well suited to the fast inversion–recovery method.

The value of $T_2$ can be determined from the line width of a signal at half of its maximum height:

$$\Delta\nu_{1/2} = (\pi T_2)^{-1} \qquad [8]$$

However, this includes any contribution to the line width from magnetic field inhomogeneity. A 'true' $T_2$, independent of contributions from field inhomogeneity, can be obtained with the aid of a spin-echo or refocusing pulse sequence (**Figure** 9). This pulse sequence also forms a key component of many other multipulse experiments. Consider a single magnetization vector which is off resonance by $\Delta\nu$ Hz, either owing to chemical shift effects or field

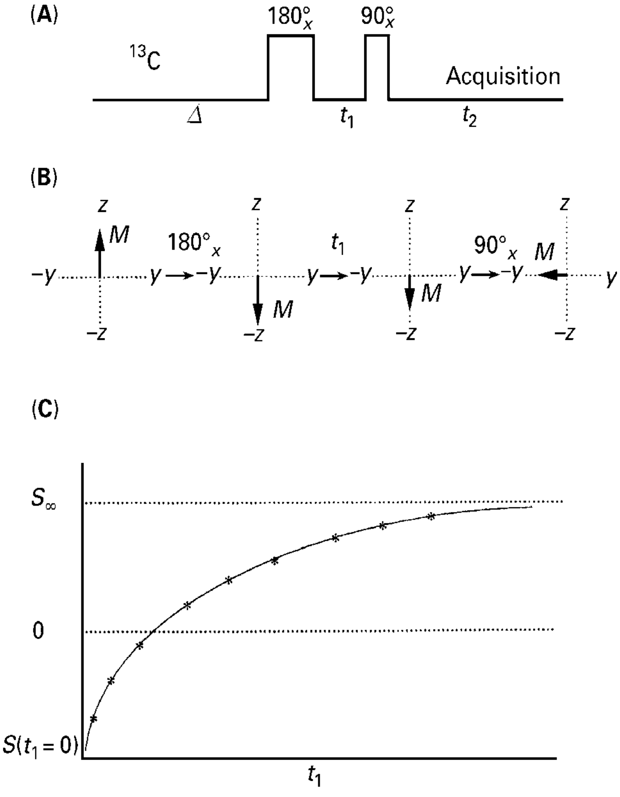



**Figure 8** (A) The inversion–recovery sequence used to measure $T_1$. The experiment is repeated with a number of different values of $t_1$. (B) The behaviour of a magnetization vector during the inversion–recovery pulse sequence. (C) A plot of signal intensity ($s$) versus $t_1$, illustrating the exponential return to the equilibrium value, $S_\infty$, as $t_1$ increases.

inhomogeneity. After the initial $90°_x$ pulse rotates it to the $y$-axis, it precesses during $t_1/2$ at angular velocity $2\pi\Delta\nu$ rad s$^{-1}$, rotating through an angle $\alpha$. The $180°_y$ pulse 'flips' it from the positive to the negative $x$ region (or vice versa), so that it is now at an angle $-\alpha$ with respect to the $y$-axis. During the second $t_1/2$ period it again rotates through $\alpha$, returning the vector to the $y$-axis, i.e. it is refocused (see **Figure** 9). The FID is then collected. The spin-echo sequence can be repeated a number of times, systematically varying $t_1$. Alternatively, one can generate an echo train in a single experiment, applying a $180°_y$ pulse at $t_1/2$, $3t_1/2$, $5t_1/2$, etc. and sampling the FID at the peaks of the echoes produced at $t_1$, $2t_1$, $4t_1$, etc. In either case, an exponential fitting process can be used to determine $T_2$ from the variation in signal intensity as a function of $t_1$.

## Spectral editing pulse sequences

These pulse sequences, which are important in $^{13}C$ NMR spectroscopy, allow assignment of individual

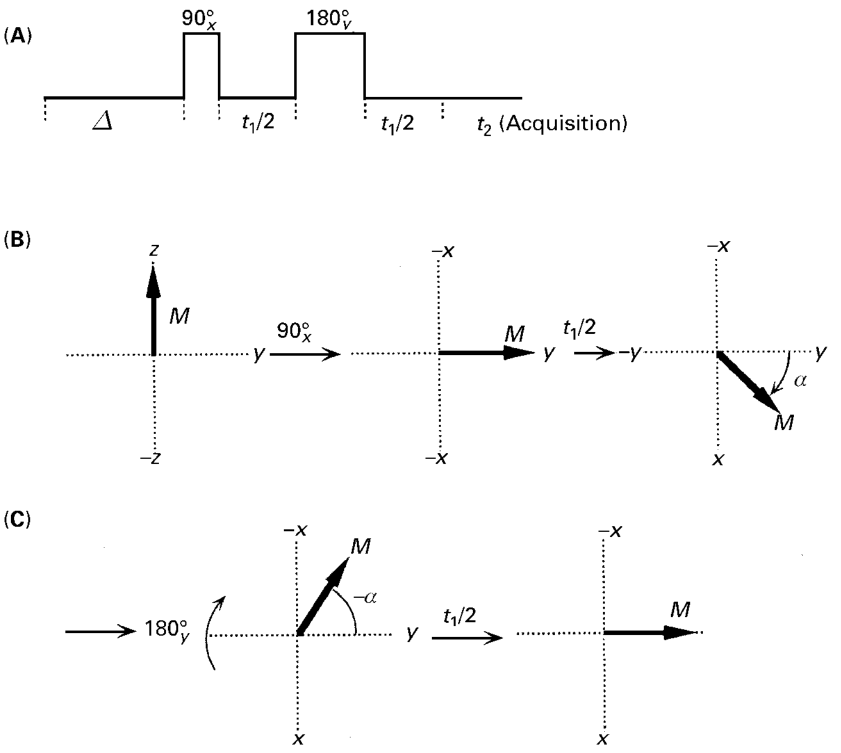



**Figure 9** (A) The spin-echo or refocusing sequence for measuring $T_2$. (B) Behaviour of a magnetization vector, corresponding to an off resonance ($\Delta\nu \neq 0$) signal, during the spin-echo sequence. The vector returns to the initial position after $t_1$, producing an 'echo'.

peaks in a heteronuclear spectrum in terms of the number of bonded hydrogens. There are three basic sequences which fall into two distinct classes. The first is the APT sequence which involves initial $^{13}$C excitation and $^{13}$C detection. **Figure 10** illustrates the multiplet patterns expected for $^{13}$C peaks coupled to 0, 1, 2 and 3 hydrogens. Now consider the effect of the APT pulse sequence (**Figure 11**) upon a $^{13}$C–$^{1}$H spin system. In the vector model, the $^{13}$C magnetization is excess $\alpha$ spin. This can be divided into two nearly equal magnetic vectors, corresponding to $^{13}$C α spin bonded to either $^{1}$H $\alpha$ or $\beta$ spins. Following the initial $90^\circ_x$ pulse, these two vectors begin to precess in the $x$–$y$ plane with angular velocities $2\pi(\Delta\nu \pm J/2)$ where $\Delta\nu$ is the chemical shift (in Hz), relative to the transmitter, and $J = {}^1J_{CH}$, the one-bond $^{13}$C–$^{1}$H coupling constant. After $t_1/2$, the $180^\circ_y$ $^{13}$C pulse flips the vectors about the $y$-axis. The simultaneous 180° $^{1}$H pulse inverts the equilibrium ($z$) $^{1}$H magnetization, interchanging $^{1}$H $\alpha$ and $\beta$ spin states. The two vectors continue to precess during the second $t_1/2$ period. At the end of this period, they

**Table 1** Vector evolution for various $CH_n$ ($n$ = 0–3) multiplets and resultant vectors at the end of $t_1$ for the APT sequence with various values of $t_1$

| | $t_1$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | | 1/4J[a] | | 1/2J | | 3/4J | | 1/J | |
| $n$[b] | $\alpha$(°)[c] | $\langle J\rangle$[d] | $\alpha$(°) | $\langle J\rangle$ | $\alpha$(°) | $\langle J\rangle$ | $\alpha$(°) | $\langle J\rangle$ | $\alpha$(°) | $\langle J\rangle$ |
| 0 | 0 | 1.00 | 0 | 1.00 | 0 | 1.00 | 0 | 1.00 | 0 | 1.00 |
| 1 | 0, 0 | 1.00 | $\pm\pi/4$ | 0.71 | $\pm\pi/2$ | 0.00 | $\pm3\pi/4$ | −0.71 | $\pm\pi$ | −1.00 |
| 2 | 0, 0, 0 | 1.00 | 0, $\pm\pi/2$ | 0.50 | 0, $\pm\pi$ | 0.00 | 0, $\pm3\pi/2$ | 0.50 | 0, $\pm2\pi$ | 1.00 |
| 3 | 0, 0, 0, 0 | 1.00 | $\pm\pi/4$, $\pm3\pi/4$ | 0.35 | $\pm\pi/2$, $\pm3\pi/2$ | 0.00 | $\pm3\pi/4$, $\pm9\pi/4$ | −0.35 | $\pm\pi$, $\pm3\pi$ | −1.00 |

[a] $J \equiv {}^1J_{CH}$, the one-bond $^{13}$C–$^{1}$H coupling constant for $CH_n$.
[b] $n$ = number of hydrogens directly bonded to carbon.
[c] Angles of rotation (relative to $y$-axis) of coupling vectors for the different peaks in each multiplet. These are calculated from the frequencies in **Figure 10** plus Equation [12].
[d] Vector average, relative to $y$-axis, of coupling vectors relative to an initial value of 1.00. Effects of $T_2$ relaxation during $t_1$ are not included.

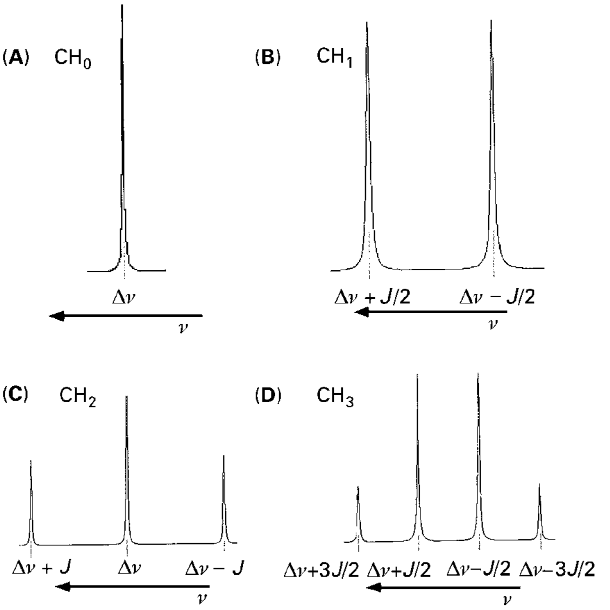



**Figure 10** Multiplet patterns arising from to ${}^1J_{CH}$ (one-bond ${}^{13}C$–${}^1H$ coupling constant) in the ${}^{13}C$ spectra of $CH_n$ groups ($n$ = 0–3).

are at angles with respect to the $y$-axis given by:

$$\alpha = \pm(2\pi J/2)t_1 = \pm\pi J t_1 \quad [9]$$

Thus, the ${}^{13}C$ chemical shift is refocused by the ${}^{13}C$ 180° pulse while the pair of 180° pulses allow ${}^1J_{CH}$ to evolve through $t_1$. Applying ${}^1H$ decoupling during acquisition rapidly scrambles ${}^1H$ spin states, producing a single, averaged vector which initially is along the $y$-axis but precesses at a frequency $\Delta\nu$ during FID acquisition. This results in a peak at $\Delta\nu$ in the frequency spectrum with an intensity determined by the vector average at the end of $t_1$. **Table 1** summarizes the results for $CH_0$, $CH_1$, $CH_2$ and $CH_3$ peaks for different values of $t_1$. Only a $CH_0$ peak is observed at $t_1 = 1/2J$ since the coupling vectors for the other multiplets average to zero. For $t_1 = 1/J$, $CH_0$ and $CH_2$ peaks are positive (upright) while $CH_1$ and $CH_3$ peaks are negative (inverted). This allows partial assignment of carbons in terms of numbers of attached protons. The main weakness of this approach is that it is sensitive to variations in ${}^1J_{CH}$ and thus may give unreliable results for compounds which have a wide range of ${}^1J_{CH}$ for carbons.

The other two spectral editing sequences are insensitive nuclei enhanced by polarization transfer (INEPT) and distortionless enhancements by polarization transfer (DEPT) (see **Figure 12**). Both of these sequences involve ${}^1H$ excitation, followed by

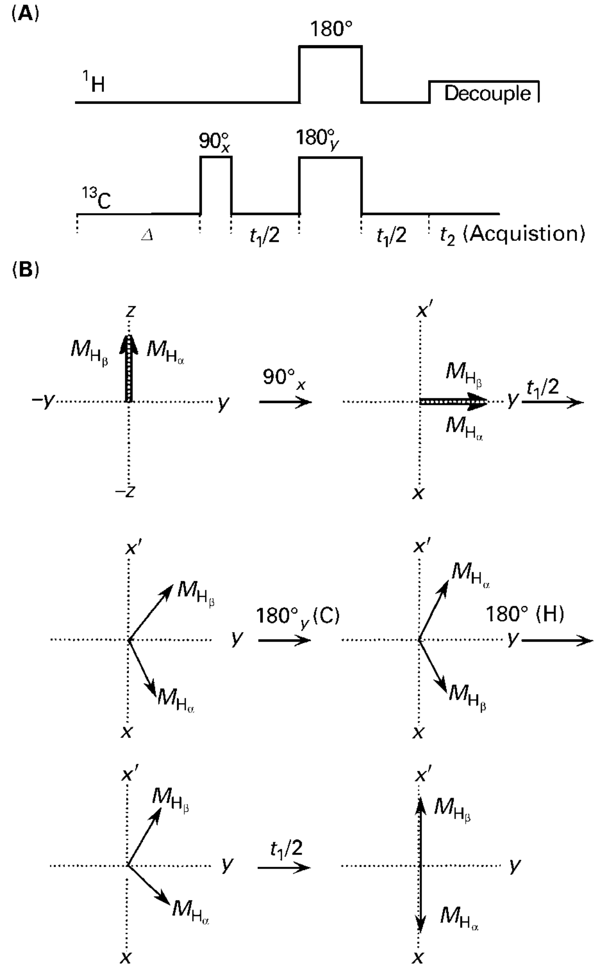



**Figure 11** (A) The APT (attached proton test) pulse sequence. (B) Behaviour of the ${}^{13}C$ magnetization due to a ${}^{13}C$–${}^1H$ spin pair during the APT sequence. The two components, corresponding to ${}^{13}C$ coupled to ${}^1H$ in $\alpha$ or $\beta$ spin states, precess at frequencies $\Delta\nu \pm J/2$ (where $J = {}^1J_{CH}$). The spin-echo sequence refocuses the chemical shift ($\Delta\nu$) but not $J$ (see Eqn [9]). This figure illustrates the result when $t_1 = J/2$ with vectors rotating through $\alpha = \pm\pi/2$.

magnetization transfer to the heteronucleus by a mechanism called polarization transfer. This provides heteronuclear signal enhancement by a factor of $\gamma_H/\gamma_X$, e.g. ~4 for ${}^{13}C$. In the case of INEPT, this can be adequately described by the vector model in terms of selective inversion of the populations of ${}^1H$ energy levels corresponding to carbons in the $\alpha$ spin state. However, although the sequences appear similar, DEPT can only be explained by a quantum mechanical model.

The two sequences use different forms of spectral editing. With INEPT, this is done by the choice of the final delay. If $\Delta_3 = J/2$, only CH carbons appear while $\Delta_3 = 3J/4$ produces a spectrum with CH and $CH_3$ up and $CH_2$ down. With DEPT, editing is carried out by varying the angle of the final ${}^1H$ pulse;

$\Theta = 90°$ yields only CH carbons while $\Theta = 135°$ yields CH and $CH_3$ up and $CH_2$ down. Because it relies on a pulse angle rather than a delay, DEPT is less sensitive than INEPT to variations in $^1J_{CH}$ and is thus the sequence of choice. Note that residual $^{13}C$ magnetization is suppressed by phase cycling in each case (see below) and thus non-protonated carbons are not observed with either sequence.

## Phase cycling for artifact suppression

Before the development of pulsed field gradient sequences (see below), most NMR pulse sequences included a phase cycle in which the phases of at least one pulse and the receiver were varied systematically. This was needed for one or more of several reasons, e.g. suppression of unwanted signals, suppression of artifacts due to hardware imperfections and/or incomplete return to equilibrium between scans and coherence pathway selection in multidimensional NMR. One example of each the first two kinds of phase cycle is briefly discussed below.

In the INEPT sequence (**Figure 12**), the final $90^\circ_y$ $^1H$ pulse sets up selective inversion of the populations of a pair of levels within the coupled AX ($^1H$–$^{13}C$) spin system. The $90^\circ_y$ $^{13}C$ pulse then generates two antiphase magnetization vectors of relative intensity +4 and –4 (relative to equilibrium $^{13}C$ magnetization) along the $\pm x$-axes owing to magnetization (polarization) transfer from $^1H$ arising from the selective population inversion. However, the $^{13}C$ $90^\circ_y$ pulse also generates a magnetization vector from the initial $^{13}C$ magnetization. It is desirable to eliminate the latter component to avoid complications with spectral editing. This is done by alternating the phase of the final $^1H$ pulse $90^\circ_y$, $90^\circ_{-y}$ while alternately adding and subtracting FID signals. With the $90^\circ_{-y}$ pulse, the antiphase $^{13}C$ vectors become –4, +4 but this is converted back into +4, –4 by subtracting this FID. However, the $^{13}C$ $90^\circ_{-y}$ pulse always generates a signal of the same phase owing to $^{13}C$ magnetization and thus is cancelled by the alternate addition and subtraction of FID signals.

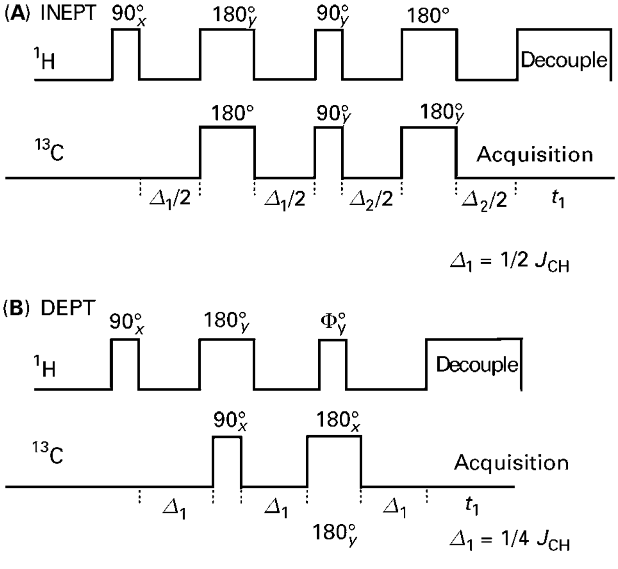



**Figure 12** (A) INEPT pulse sequence. (B) DEPT pulse sequence. The article on product operator formalism describes the behaviour of the DEPT sequence while the texts by Harris and Günther (see Further reading section) describe the behaviour of INEPT in terms of vector diagrams and energy levels.

Quadrature detection involves the use of two receivers. If the two receivers have different gains, 'quadrature image' peaks are generated at $-\Delta\nu$ for every true peak at $+\Delta\nu$. This can be eliminated by using a four-step CYCLOPS phase cycle in which the phase of a transmitter is cycled through relative phases $x$, $y$, $-x$, $-y$ along with the receiver. Derome (see Further reading) gives a very clear account of quadrature images and how they are suppressed by CYCLOPS phase cycling.

## Quantum mechanical methods for understanding pulse sequences

The ultimate approach for interpreting multipulse sequences and their resultant spectra is a full density matrix treatment. While this approach is ideal for simulating the spectrum generated by a multipulse experiment, the calculations are complex and do not provide obvious physical insights. A very useful and widely used simplified quantum mechanical approach involves product operator formalism. This focuses on the components of the density matrix which are directly relevant to the experiment. Product operator descriptions of several of the pulse sequences discussed here are given in another article.

Mastery of this approach is essential for anyone desiring to design new pulse sequence experiments and valuable for anyone wishing to understand modern NMR experiments.

## Multidimensional NMR experiments

Multidimensional NMR experiments have revolutionized the use of NMR spectroscopy for the structure determination of everything from small molecules to complex proteins. Since most of the 3D and 4D experiments are essentially combinations of two-dimensional (2D) experiments, this section will focus on 2D NMR. Only a basic overview will be given since many specific multidimensional experiments are discussed elsewhere.

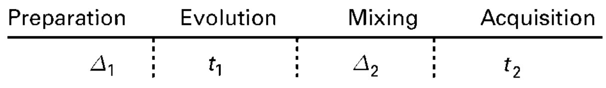



**Figure 13** A general two-dimensional NMR pulse sequence. Data are acquired during $t_2$ for a series of spectra in which $t_1$ is regularly incremented from 0 to some maximum value. Fourier transformation with respect to $t_2$ and then $t_1$ generates a spectrum with two difference frequency axes.

## Two-dimensional NMR spectroscopy

A generalized two-dimensional experiment is illustrated in **Figure 13**. The preparation time is usually a relaxation delay followed by one or more pulses to start the experiment. The evolution period establishes the second frequency dimension. A series of FID signals are collected with $t_1$ regularly incremented from 0 up to the desired maximum value, $t_1$ (max). The number of increments and $t_1$ (max) depend on the desired spectral width and data point resolution along the time-incremented axis. Depending on the experiment, the evolution period may contain one or more pulses, most commonly a spin-echo sequence. The mixing period, which is not required in some sequences, can be a 90° mixing pulse, a fixed delay, a more complex pulse such as a spin lock or isotropic mixing pulse or some combination of these. Finally, data is acquired during $t_2$, as in a 1D experiment. Double Fourier transformation, with respect to $t_2$ then $t_1$, yields a spectrum with two orthogonal frequency scales.

2D NMR experiments are designed to generate different kinds of frequency information along the two axes. The principle behind this frequency separation is most easily seen by considering modification of the APT sequence (**Figure 11**) to produce a 2D sequence, the heteronuclear *J*-resolved sequence. This involves replacing the constant $t_1$ period by an incremented $t_1$ period. Each $^{13}C$ signal in $f_2$ is then modulated by the evolution of $^1J_{CH}$ coupling vectors as $t_1$ is incremented (e.g. see **Table 1**). Fourier transformation with respect to $t_1$ at each $f_2$ frequency then produces a 2D spectrum where a cross section through each $^{13}C$ peak in $f_2$ will give a $^1J_{CH}$ multiplet pattern similar to one of those shown in **Figure 10**. Similarly, the INEPT sequence can be converted into a 2D sequence (the heteronuclear shift correlation sequence or HETCOR) by inserting a spin-echo sequence, $t_1/2$–180°(C)–$t_1/2$ immediately after the initial $^1H$ 90° pulse (see **Figure 12**). The extent of polarization transfer from $^1H$ to $^{13}C$ is then modulated by $^1H$ chemical shift evolution as $t_1$ is incremented, with the resultant 2D experiment having $^{13}C$ chemical shifts along $f_2$ and $^1H$ chemical shifts along $f_1$.

**Table 2** Characteristics of several commonly used 2D NMR pulse sequences

| *Sequence* | *Display mode*[a] | *$f_1$, $f_2$*[b] | *Transmission*[c] |
|---|---|---|---|
| COSY | D, OD | $\delta_H$, $\delta_H$ | $J_{HH}$ |
| DQCOSY[d] | D, OD | $\delta_H$, $\delta_H$ | $J_{HH}$ |
| TOCSY[e] | D, OD | $\delta_H$, $\delta_H$ | $J_{HH} \rightarrow J_{HH}$ |
| NOESY | D, OD | $\delta_H$, $\delta_H$ | H–H dipolar |
| ROESY[f] | D, OD | $\delta_H$, $\delta_H$ | relaxation |
| EXSY[g] | D, OD | $\delta_H$, $\delta_H$ | $H_A \rightleftharpoons H_B$ |
| HETCOR[h] | $f_1$, $f_2$ | $\delta_H$, $\delta_X$ | $^1J_{CH}$ |
| COLOC[h] | $f_1$, $f_2$ | $\delta_H$, $\delta_X$ | $^nJ_{CH}$ |
| FLOCK[h] | $f_1$, $f_2$ | $\delta_H$, $\delta_X$ | $^nJ_{CH}$ |
| HMQC[i] | $f_1$, $f_2$ | $\delta_X$, $\delta_H$ | $^1J_{CH}$ |
| HSQC[i] | $f_1$, $f_2$ | $\delta_X$, $\delta_H$ | $^1J_{CH}$ |
| HMBC[i] | $f_1$, $f_2$ | $\delta_X$, $\delta_H$ | $^nJ_{CH}$ |
| INADE-QUATE | DQ, SQ | $\delta_{C(1)}+\delta_{C(2)}$, $\delta_{C(1)}$ | $^1J_{CC}$ |

[a] D, OD: spectrum along diagonal with off-diagonal peaks between correlated protons, e.g. see COSY spectrum (**Figure 14**), $f_1$, $f_2$: different chemical shift scales along $f_1$, $f_2$, e.g. see HSQC spectrum (**Figure 15**), DQ, SQ: double quantum frequencies along $f_1$ (sum of frequencies of coupled $^{13}C$ peaks, relative to transmitter), regular (single quantum) $^{13}C$ spectrum along $f_2$.

[b] Chemical shift information appearing along each axis. {$\delta$ scale ≡ chemical shift in parts per million relative to internal reference [$\delta_1(CH_3)_4$ for $^1H$, $^{13}C$ and $^{29}Si$]}. Note that the $^1H$-axis normally also shows multiplet structure owing to $J_{HH}$.

[c] Parameter by which information is transmitted to establish correlations between the spectra on the two frequency axes: $J_{HH}$ = $^1H$–$^1H$ coupling constant, $H_A \rightleftharpoons H_B$ is chemical exchange between different sites, $J_{CH}$ = one-bond $^{13}C$–$^1H$ coupling constant, $^nJ_{CH}$ = *n*-bond ($n$ = 2 or 3) $^{13}C$–$^1H$ coupling constant, $^1J_{CC}$ = one-bond $^{13}C$–$^{13}C$ coupling constant.

[d] DQCOSY ≡ double quantum filtered COSY. This suppresses strong singlets (e.g. solvent peaks) and gives well-resolved off-diagonal peaks with up–down intensity patterns for coupled protons, i.e. those giving rise to the off-diagonal peak.

[e] TOCSY (also called HOHAHA) relays information among sequences of coupled protons. A cross section through the $f_2$ frequency of a specific proton shows $f_1$ peaks for all of the protons within the coupled sequence.

[f] ROESY ≡ NOESY in the rotating frame.

[g] The EXSY sequence is identical to the NOESY sequence but detects cross peaks between chemically exchanging hydrogens. Both EXSY and NOESY peaks may appear in the same spectrum.

[h] X nucleus (usually $^{13}C$) detected heteronuclear shift correlation sequences.

[i] $^1H$ detected heteronuclear shift correlation sequences. These are more sensitive than the earlier X-nucleus detected sequences [by $(\gamma_H/\gamma_X)^{3/2}$] but have more limited resolution along the X($f_1$) axis.

High-resolution 2D NMR pulse sequences can be based on information transfer via homonuclear or heteronuclear scalar coupling, dipolar relaxation or chemical exchange while solid-state 2D NMR experiments normally use dipolar coupling in place of scalar coupling. There are three basic modes of spectral

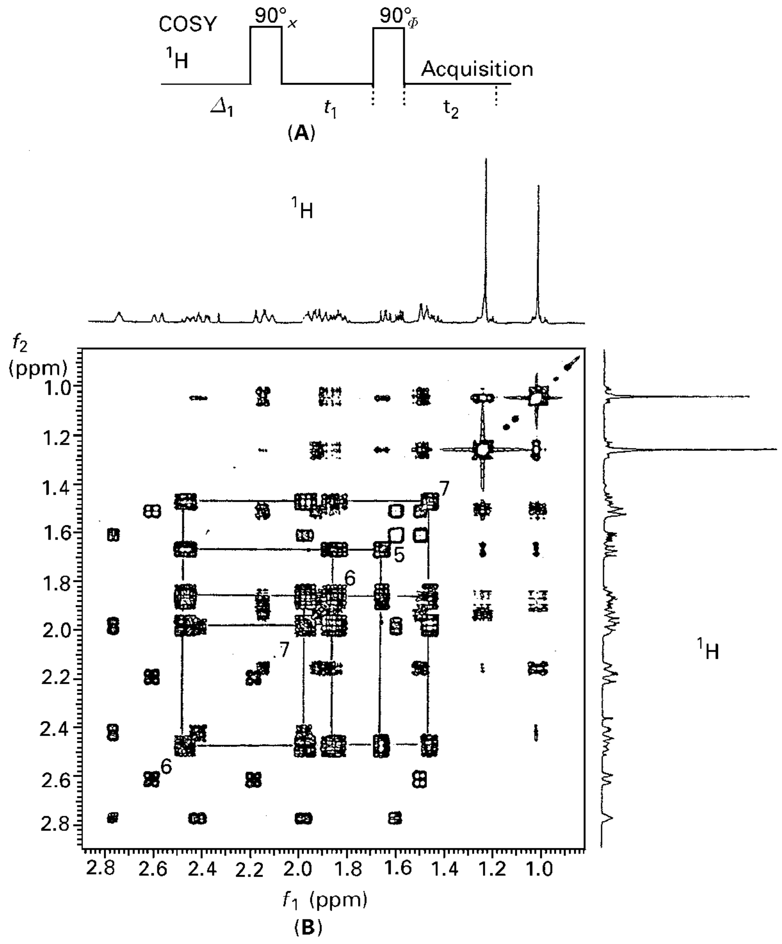



**Figure 14** (A) The COSY pulse sequence. (B) The COSY spectrum for the aliphatic region of [1], showing the $^1H$ spectrum along the diagonal and symmetric off-diagonal peaks between coupled protons. The connections for one molecular fragment [C(5)H–C(6)$H_2$–C(7)$H_2$)] are traced out.

display: with the normal spectrum along the diagonal and off-diagonal peaks for correlated signals (see COSY spectrum, **Figure 14**), different chemical shift information along the two axes (e.g. $^1H$ and $^{13}C$ or $^{15}N$, see **Figure 15**) and spectra with single quantum frequencies along $f_2$ and multiple quantum frequencies along $f_1$. The characteristics of many of the common high resolution 2D sequences are summarized in **Table 2**.

However, the real strength of multidimensional NMR is not in the information provided by a single experiment but rather in the synergy provided by carrying out several different experiments on the same molecule. This is illustrated below for kauradienoic acid [1], one of the very first molecules where combined 2D methods were used for spectral assignment. The $^1H$–$^1H$ COSY spectrum (**Figure 14**) and the $^1H$–$^{13}C$ shift correlation spectrum (**Figure 15**) for [1] allow assignment of molecular fragments involving sequences of protonated carbons. Further experiments allow completion of the structural and spectral assignments (see caption to **Figure 15**).

## Absolute value versus phase sensitive 2D spectra

Many of the original 2D sequences gave spectra which could not be phased since they involved different mixtures of absorption and dispersion modes for different peaks. To simplify displays, these spectra were plotted in absolute value mode, $(u^2 + v^2)^{1/2}$, where $u$ refers to dispersion mode and $v$ to absorption mode. While the individual absolute value mode peaks appear to be properly phased, they are distorted from Lorentzian shape with broad tails. Better resolution and sensitivity can be obtained if spectra are obtained in a manner which provides pure absorption mode peaks. There are now phase sensitive versions of most 2D sequences. There are two requirements for obtaining phase sensitive spectra. First, any fixed delay must include a spin-echo sequence to prevent chemical shift evolution. Second, one must acquire two separate data sets with

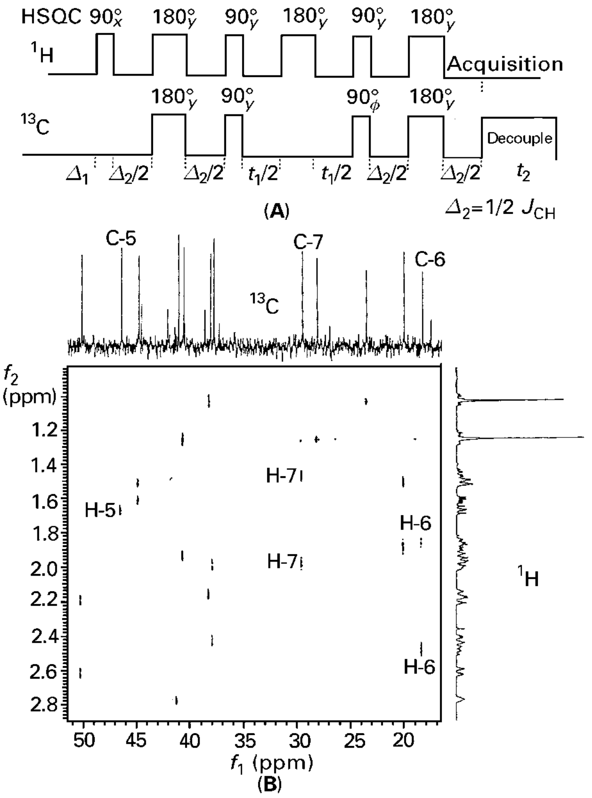



Figure 15 (A) The basic HSQC (heteronuclear single quantum coherence) pulse sequence. (B) The HSQC spectrum of the aliphatic region of [1] with the $^{13}C$ along $f_1$ and the $^1H$ spectrum along $f_2$. The $^{13}C$–$^1H$ connectivities are marked for the same molecular fragment as in **Figure 14**. Other sequences of protonated carbons can be determined from the same spectrum while an *n*-bond ($n = 2,3$) $^{13}C$–$^1H$ shift correlation spectrum such as HMBC, COLOC or FLOCK (see **Table 2**) can identify non-protonated carbons and tie together the molecular fragments into a complete structure.

one of the pulses having a 90° phase difference for the two spectra or one must increment the phase of one of the pulses by 90° with each time increment while doubling the number of time increments collected.

## Phase cycling for coherence pathway selection in 2D NMR

This important topic can only be understood in quantum mechanical terms. Owing to space limitations, only a very brief introduction can be given here.

The COSY sequence (**Figure 14**) will be used to illustrate the concept. A coherence level diagram for this sequence is given in **Figure 16**. Equilibrium $z$ magnetization is defined as having a coherence level of 0. The 90° pulse acts as a raising or lowering operator, i.e. it can change the spin quantum number of an individual nucleus by ±1, resulting in the generation of observable $x,y$ magnetization. This evolves during $t_1$ at frequencies determined by the chemical shift and homonuclear coupling. The second 90° pulse then can generate further coherence level changes, including level changes of 0 to ±2 associated with a pair of coupled nuclei. The receiver can only detect single quantum coherence and is chosen to be at coherence level +1. Thus, a pair of ideal pulses can generate a COSY signal by two paths with coherence level changes +1, 0 or −1, +2, respectively called P (or antiecho) and N (echo) signals. If both paths are detected, signals will occur at both $+\Delta\nu_1$ and $-\Delta\nu_1$ along the $f_1$ axis. However, phase cycling allows one to choose one path, rejecting the other. The change in coherence phase associated with a pulse is given by:

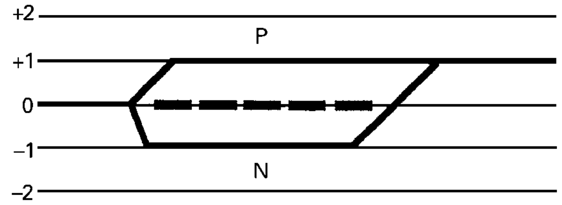



Figure 16 Coherence level diagram for the COSY sequence. The symbols N and P designate the N and P pathways.

$$\Delta\theta = \Delta p \cdot \Delta\phi \qquad [10]$$

where $\Delta\theta$ is the change in coherence phase (in increments of $\pi/2$), $\Delta p$ is the change in coherence level and $\Delta\phi$ is the change in pulse phase (in units of $\pi/2$). **Table 3** shows how this can be used to design two-step phase cycles for a coherence pathway section with the COSY sequence.

The dotted line indicates a third possible coherence pathway. If the initial 90° pulse is imperfect, there will be some residual $z$ magnetization which will be raised to coherence level +1 by the second 90° pulse. Suppression of this pathway requires two extra steps, yielding a four-step phase cycle. Finally, if one also wishes to incorporate a CYCLOPS cycle for $f_2$ quadrature image suppression, the total phase cycle is $4 \times 4 = 16$ steps. The number of scans in a 2D experiment should be some whole number multiple of the number of steps in the phase cycle. However, with COSY, the sensitivity is high enough that 16 scans are usually more than necessary to acquire good spectra and thus the phase cycle determines the minimum time for the experiment. Fortunately, pulsed field gradient sequences have overcome this problem.

**Table 3** Alternative two-step phase cycles for N-type pathway selection and P-pathway suppression for a COSY spectrum plus a four-step phase cycle which selects the N-pathway while suppressing both P and Z paths

| | *N* | | *P* | | | *N* | | *P* | |
|---|---|---|---|---|---|---|---|---|---|
| Scan | 1 | 2 | 1 | 2 | Scan | 1 | 2 | 1 | 2 |
| $\phi$(P1) | 0[a] | 0 | 0 | 0 | $\phi$(P1) | 0 | 3 | 0 | 3 |
| $\Delta\theta$(P1) | 0 | 0 | 0 | 0 | $\Delta\theta$(P1) | 0 | 1[b] | 0 | 3 |
| $\phi$(P2) | 0 | 1 | 0 | 0 | $\phi$(P2) | 0 | 0 | 0 | 0 |
| $\Delta\theta$(P2) | 0 | 2 | 0 | 0 | $\Delta\theta$(P2) | 0 | 0 | 0 | 0 |
| $\Delta\theta$(P1 + P2) | 0 | 2 | 0 | 0 | $\Delta\theta$(P1 + P2) | 0 | 1 | 0 | 3 |
| $\phi$(R)[c] | 0 | 2 | 0 | 2 | $\phi$(R) | 0 | 1 | 0 | 1 |

| | *N* | | | | *P* | | | | *Z* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scan | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| $\phi$(P1)[d, e] | 0 | 3 | 2 | 1 | 0 | 3 | 2 | 1 | 0 | 3 | 2 | 1 |
| $\Delta\theta$(P1) | 0 | 1 | 2 | 3 | 0 | 3 | 2 | 1 | 0 | 0 | 0 | 0 |
| $\phi$(P2) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\Delta\theta$(P2) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\Delta\theta$(P1 + P2) | 0 | 1 | 2 | 3 | 0 | 3 | 2 | 1[e] | 0 | 0 | 0 | 0[e] |
| $\phi$(R)[c] | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 3 |

The symbols N and P indicate the coherence level change from the first pulse and are respectively negative (–1) and positive (+1) for the two paths. The third path, from an imperfect initial 90° pulse, has zero coherence level change in the initial pulse and is thus given the symbol Z.

[a] Pulse and receiver phases *x, y, −x* and *−y* are, respectively, given as 0, 1, 2 and 3, corresponding to the number of $\pi/2$ phase increments relative to a $90°_x$ pulse.

[b] $\Delta p = -1$ for the coherence level change in the N pathway while $\Delta\theta$(P1) = 3. From Equation [10], $\Delta p \Delta\theta$ = (–1) (3) = –3. However, since a –270° phase shift corresponds to a −90° phase shift, $-3 \equiv 1$.

[c] $\phi$(R) ≡ receiver phase. When the sums of coherence phase changes in different scans match the receiver phase cycle, successive scans add, while when the relative phases change 0, 2 successive scans cancel. Thus in each case, the signals from the N path add while signals from the P path cancel.

[d] An alternative four-step phase cycle for N-path selection involves a $\phi$(P1) = 0, 1, 2, 3 and $\phi$(R) = 0, 3, 2, 1. For P-type selection $\phi$(P1) and $\phi$(R) should either both be 0, 1, 2, 3 or both be 0, 3, 2, 1. Another alternative for N-pathway selection is to expand the first two-step phase cycle to a four-step cycle with $\phi$(P2) = 0, 1, 2, 3 and $\phi$(R) = 0, 2, 0, 2.

[e] In this four-step phase cycle, the P-pathway is cancelled in steps 1 + 2 and in steps 3 + 4 while the Z pathway is cancelled in steps 1 + 3 and steps 2 + 4.

### Gradient pulse sequences

Many of the pulse sequences discussed above now have versions which incorporate magnetic field gradient pulses that can be used to replace phase cycling. They allow one to acquire a spectrum in greatly reduced time and/or with greatly reduced artifacts. For example, applying two identical field gradient pulses before and after the final 90° pulse in COSY selects the N (echo) path while suppressing the other paths.

## Pulse sequences which replace single pulses

A number of pulse 'sandwiches' have been developed to replace single pulses in specific cases. These include the following.

### Composite 180° pulses

The quality of the spectra obtained with many pulse sequences is strongly dependent on the precision of 180° pulses, particularly in the case of 180° inversion pulses for heteronuclei with broad spectral windows. Problems can arise due to mis-set pulses, inhomogeneity in pulses over the sample or incomplete excitation at large frequencies relative to the transmitter. These problems can be minimized by the use of composite 180° pulses, e.g. a $90^\circ_x$, $180^\circ_y$, $90^\circ_x$ composite pulse in place of a $180^\circ_x$ pulse (**Figure 17**).

### BIRD (bilinear rotating decoupling) pulses

BIRD pulses act as selective 180° $^1$H pulses either for protons directly bonded to $^{13}$C or not bonded to $^{13}$C, while simultaneously providing a $^{13}$C 180° pulse (**Figure 18**). Earlier uses of these pulses included partial $^1$H–$^1$H decoupling in HETCOR and optimization

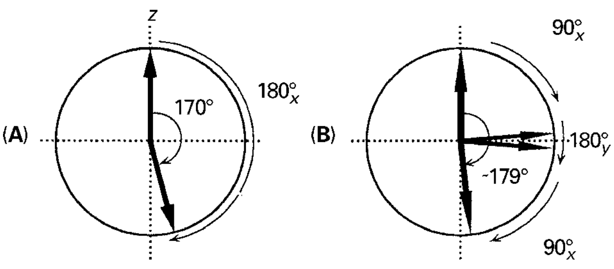
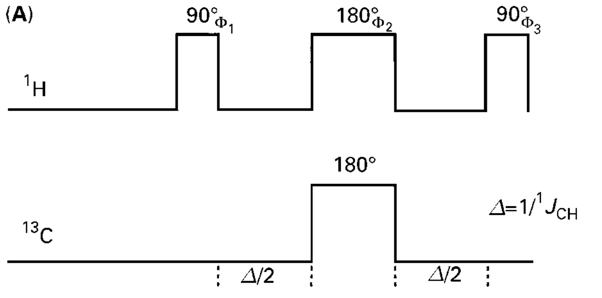



Figure 17 (A) Vector diagram illustrating the effect when a nominal 180° pulse is mis-set, resulting in only 170° rotation. (B) Illustration of how a composite 90°$_x$, 180°$_y$, 90°$_x$ compensates for the effect of a mis-set pulse. Compensation is less complete for off-resonance signals.

of performance of long-range $^{13}$C-detected $^{13}$C–$^1$H correlation sequences such as COLOC and FLOCK. The most common current use is for suppression of $^1$H–$^{12}$C magnetization in $^1$H-detected one-bond $^{13}$C–$^1$H correlation sequence, i.e. HMQC and HSQC. Although BIRD pulses can be explained by vector diagrams (**Figure 18**), a full understanding of these pulses requires a quantum mechanical treatment.

## Frequency selective pulses

The ability to selectively excite a narrow spectral region is important both for solvent suppression and because it often allows one to replace a full 2D experiment by a limited number of 1D experiments. A 'soft' (i.e. low power, long duration) pulse can be used for selective excitation but this does not generate




| $\Phi_1$ | $\Phi_2$ | $\Phi_3$ | $^1$H signal inverted |
|---|---|---|---|
| x | y | x | $^1$H - $^{12}$C |
| x | y | −x | $^1$H - $^{13}$C |
| x | x | −x | $^1$H - $^{12}$C |
| x | x | x | $^1$H - $^{13}$C |

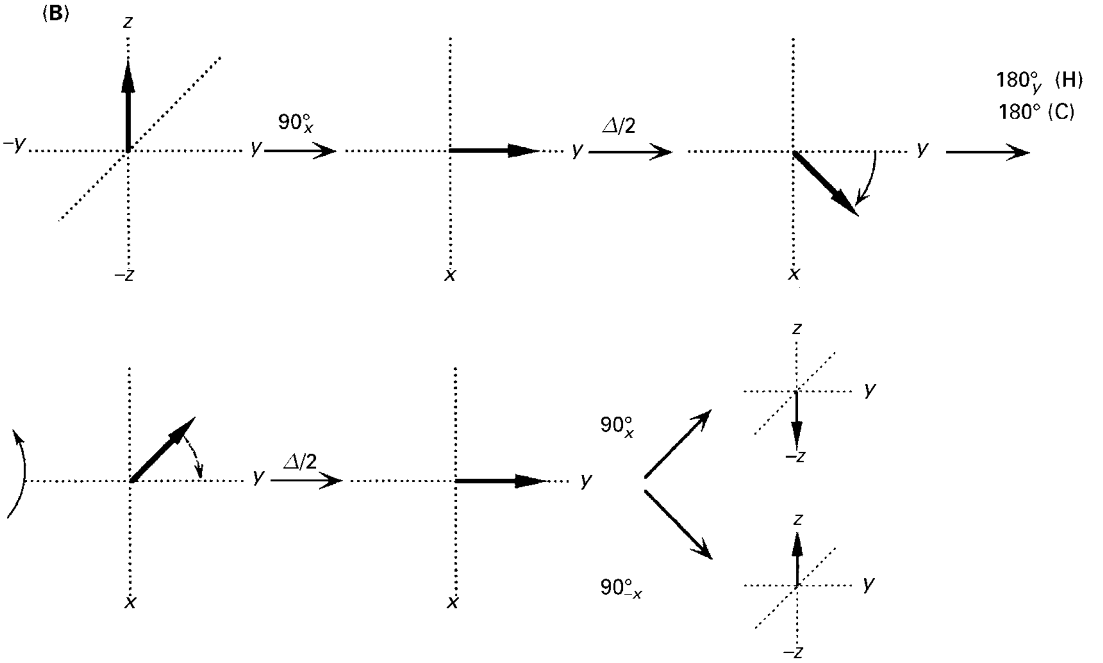



Figure 18 (A) The BIRD pulse sequence and effects of different combinations of phases within the BIRD pulse. (B) Vector diagram for a $^{12}$C–$^1$H spin system, illustrating how 90°$_x$, 180°$_y$, 90°$_x$ and 90°$_x$, 180°$_y$, 90°$_{-x}$ BIRD pulses, respectively, act as 180° and 0° $^1$H pulses. Since the BIRD pulse corresponds to the APT sequence (with $^1$H and $^{13}$C pulses interchanged) up to the point of the final 90° pulse, the effect of these two BIRD pulses on a $^1$H–$^{13}$C pair can be deduced from the data in **Table 1** for $n = 1$. With $\Delta = 1/J_{CH}$, the vectors associated with the $^1$H–$^{13}$C pair are refocused along the −y-axis and a 90°$_x$ pulse will rotate them back to the z-axis (0° pulse) while a 90°$_{-x}$ pulse rotates them to the −z-axis (180° pulse).

uniform excitation (see **Figure 3**). Better results are obtained from multiple pulse sequences. Modern spectrometers allow the generation of 'shaped' pulses whose time profiles are designed to produce the desired excitation profiles. A series of pulses of controlled amplitude, duration and phase (without intervening delays) are used which provide the desired profile. For example, generating a pulse with a time profile similar to the frequency profile in **Figure 3** will give a narrow square wave excitation profile. Both 90° and 180° pulses can be generated as well as pulses which simultaneously irradiate at two or more chosen frequencies. The bandwidth of each pulse can be adjusted to selectively irradiate a chosen signal or to cover a specific spectral region (e.g. irradiation of the amide $^{13}C{=}O$ region in a 3D or 4D protein spectrum).

## Broad-band decoupling pulse sequences

Multiple pulse sequences can also be used to provide effective broad-band decoupling. The original pulse sequence of this kind was the WAHUHA sequence of Waugh and co-workers which was designed to minimize broadening arising from homonuclear dipolar coupling in solid-state spectra.

For high-resolution NMR, the main interest has been in heteronuclear broad-band decoupling. Initially, the interest was in broad-band $^{1}H$ decoupling while acquiring heteronuclear (e.g. $^{13}C$) spectra. More recently, with the increasing use of $^{1}H$-detected 2D, 3D and 4D sequences involving $^{1}H$–X chemical shift correlation, the emphasis has been on decoupling of heteronuclei (e.g. $^{13}C$, $^{15}N$). This is much more demanding owing to the much wider heteronuclear chemical shift window. Increasingly effective decoupler pulse sequences have been developed with acronyms such as MLEV, WALTZ, GARP, DIPSI and WURST. Most are based on a composite 180° decoupler pulse which is subjected to a series of phase cycles. For example, WALTZ is based on a $90^{\circ}_{x}$, $180^{\circ}_{-x}$, $270^{\circ}_{x}$ composite pulse (which in shorthand form is designated $1\ \bar{2}\ 3$, justifying the name WALTZ).

## Summary

This article has given an overview of the many different multiple pulse experiments which have developed from the original pulse Fourier transform experiment. These experiments, along with major improvements in spectrometer instrumentation, have dramatically increased the range of structural and dynamic problems that can be studied by NMR spectroscopy.

## List of symbols

$\boldsymbol{B}_0$ = external magnetic field vector; $\boldsymbol{B}_1$ = rotating magnetic field vector arising from RF electromagnetic radiation; $u$ = dispersion mode; $\boldsymbol{M}$ = resultant of individual magnetic moment vectors; $v$ = absorption mode; α,β = spin states corresponding to allowed values of $m_I$; $\alpha$ = angle of rotation of $\boldsymbol{M}$ with respect to initial axis; $\gamma$ = magnetogyric ratio for nucleus, i.e. the ratio of magnetic moment/spin angular momentum; $\Delta$ = fixed delay; $\theta$ = phase of coherence; $\nu$ = frequency ($s^{-1}$); $\Delta\nu$ = frequency difference; $\phi$ = phase of pulse or receiver; $\omega$ = angular velocity (rad $s^{-1}$).

*See also:* **$^{13}C$ NMR, Methods; $^{13}C$ NMR, Parameter Survey; Fourier Transformation and Sampling Theory; High Resolution Solid State NMR, $^{1}H$, $^{19}F$; Magnetic Field Gradients in High Resolution NMR; NMR Principles; NMR Spectrometers; Product Operator Formalism in NMR; Proteins Studied Using NMR Spectroscopy; Solvent Suppression Methods in NMR Spectroscopy; Structural Chemistry using NMR Spectroscopy, Inorganic Molecules; Structural Chemistry using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Peptides; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals; Two-Dimensional NMR, Methods.**

## Further reading

Derome AE (1987) *Modern NMR Techniques for Chemistry Research*. Oxford: Pergamon Press.

Ernst RR (1992) Nuclear magnetic resonance Fourier transform spectroscopy (Nobel lecture). *Angewante Chemie* **31**: 805–823.

Freeman R (1997) *A Handbook of Nuclear Magnetic Resonance*, 2nd edn. Harlow: Addison Wesley Longman.

Freeman R (1998) Shaped radio frequency pulses in high resolution NMR. *Progress in NMR Spectroscopy* **32**: 59–106.

Günther H (1995) *NMR Spectroscopy*, 2nd edn. Chichester: Wiley.

Harris RK (1986) *Nuclear Magnetic Resonance Spectroscopy: A Physiochemical View*. Harlow: Longman.

Keeler J (1990) Phase cycling procedures in multiple pulse NMR spectroscopy of liquids. In: Granger P and Harris RK (eds) *Multinuclear Magnetic Resonance in Liquids and Solids*. Dordrecht: Kluwer.

Levitt M (1986) Composite pulses. *Progress in NMR Spectroscopy* **18**: 61–122.

Parella T (1998) Pulsed field gradients: a new tool for routine NMR. *Magnetic Resonance in Chemistry* **36**: 467–495.

Shaka AJ and Keeler J (1987) Broadband spin decoupling in isotropic liquids. *Progress in NMR Spectroscopy* **19**: 47–129.

# NMR Relaxation Rates

**Ronald Y Dong**, Brandon University, Manitoba, Canada

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Theory

How a nuclear spin system achieves thermal equilibrium by exchanging energy with its surrounding medium or the 'lattice' is governed by the NMR relaxation rates. The lattice consists of all degrees of freedom, except those of the nuclear spins, associated with the physical system of interest. Pulsed NMR provides a highly versatile and flexible tool to determine spin relaxation rates, which can probe the entire spectrum of molecular motions. These include molecular rotation, translational self-diffusion, 'coherent' rotational motion, and the internal motion in nonrigid molecules. Physical systems investigated by NMR range from condensed matter phases to dilute molecular gases. The theory of nuclear spin relaxation is now well understood, and the details are given in the classical treatise by Abragam. An elementary treatment of the same material can be found in the text by Farrar. In relating the measured spin relaxation rates to molecular behaviours, there are severe limitations and difficulties that a newcomer can often fail to appreciate. Several nuclear interactions may simultaneously all contribute to the relaxation of a spin system. These may include the magnetic dipole–dipole interaction, the quadrupole interaction, the spin–rotation interaction, the scalar coupling of the first and second kind, and the chemical-shift anisotropy interaction. Due to the need of estimating certain nuclear couplings and/or correlation times associated with molecular motions, considerable uncertainty may exist in identifying and separating these contributions. The semiclassical relaxation theory of Redfield is outlined in this short review to give expressions of spin relaxation rates in terms of spectral densities of motion. The treatment is semiclassical simply because it uses time correlation functions which are classical.

The most difficult problem in any relaxation theory is the calculation of correlation functions or spectral densities of motion. It is often possible to determine the mean square spin interaction $\langle H_q^2(t)\rangle$, where $H_q(t)$ is a component of the spin Hamiltonian which fluctuates randomly in time owing to molecular motions. The time dependence of the correlation function $\langle H_q(t)H_{q'}(t-\tau)\rangle$ can often be approximated by an exponential decay function of $\tau$, i.e.

$$\langle H_q(t)H_{q'}(t-\tau)\rangle = \langle H_q(t)H_{q'}(t)\rangle \ \exp(-\tau/\tau_c) \qquad [1]$$

where the angle brackets denote an ensemble average, and the correlation time $\tau_c$ for the motion can be determined with the help of experiments. There are many examples of exponentially decaying correlation functions. For instance, thermal motion of molecules in liquids was first treated in the classical BPP paper by Bloembergen. Spectral density calculations for liquids normally use a classical picture for the lattice. Quantum calculations of spectral densities are feasible for spin relaxation due to lattice vibrations or conduction electrons. Such calculations are, in general, impossible since the eigenstates of the lattice are often unknown. Molecular motions that are too fast ($\omega_0\tau_c << 1$) or too slow ($\omega_0\tau_c >> 1$) with respect to the inverse of the Larmor frequency $\omega_0$ are not amenable to nuclear spin–lattice relaxation ($T_1$) studies. Fortunately, measurements of spin–spin relaxation time ($T_2$) or spin–lattice relaxation time ($T_{1\rho}$) in the rotating frame can be used. Both $T_1$ and $T_2$ appear in the phenomenological Bloch equations, which describe the precession of nuclear magnetization in an external magnetic field. In NMR, the coupling between the lattice and the Zeeman reservoir of the nuclear spin system is magnetic in all cases except one. The exception is the quadrupole coupling between the nuclear quadrupole moment (for spin angular momentum $I > ½$) and the lattice via an electric field gradient, which is electrical in nature. When this coupling exists, it is generally more efficient than any magnetic coupling. Relaxation of a quadrupolar nucleus of spin $I = 1$ (i.e. $^2$H) will be explicitly addressed. The deuteron has a small quadrupole moment with a coupling constant $e^2qQ/h$ typically 150–250 kHz, large enough so that relaxation is dominated by the quadrupole interaction and small enough so that perturbation theory is applicable. In liquids, the couplings between nuclear spins are greatly reduced by rapid thermal motions of molecules. Since these couplings are weak and comparable to the coupling of the spins with the lattice, one can consider relaxation of individual spins or, at most,

groups of spins inside a molecule. For deuterated molecules in liquids, the dipole–dipole coupling between deuterons is much weaker than the quadrupole interaction. As a consequence, one can normally consider a collection of isolated deuteron spins in liquid samples.

## Theory

Suppose that an assembly of $N$ identical spin systems is considered. This allows a quantum statistical description of a spin system, for example the $k$th spin system in the ensemble. If the spin system is in a state with wavefunction or ket $|\psi_k\rangle$, the expectation value of a physical observable given by its operator $Q$ is

$$\langle Q \rangle_k = \langle \psi_k \mid Q \mid \psi_k \rangle \quad [2]$$

NMR spectroscopy deals with the observation of macroscopic observables rather than states of individual spin systems. Thus, one needs to perform an average over the members of the ensemble:

$$\langle Q \rangle = \sum_{k=1}^{N} \langle Q \rangle_k / N \quad [3]$$

In general, the ket $|\psi_k\rangle$ is time dependent and may be expanded using a complete orthonormal basis set of $m$ stationary kets $|\phi_\beta\rangle \equiv |\beta\rangle$:

$$|\psi_k\rangle = \sum_{\beta=1}^{m} C_\beta^k(t) |\beta\rangle \quad [4]$$

where the expansion coefficients $C_\beta^k$ are time dependent. This leads to

$$\langle Q \rangle = \sum_{\alpha,\beta} \langle \beta \mid Q \mid \alpha \rangle \sigma_{\alpha\beta}(t) \quad [5]$$

where the matrix elements of a density operator $\sigma$ are defined by

$$\sigma_{\alpha\beta} = \langle \alpha \mid \sigma \mid \beta \rangle = \overline{C_\alpha(t) C_\beta^*(t)} \quad [6]$$

and the bar denotes an ensemble average. Now the density operator is Hermitian and has real eigenvalues. In particular, its diagonal elements $\sigma_{\alpha\alpha}$ represent the probabilities of finding ket $|\alpha\rangle$ (or populations of $|\alpha\rangle$) in $|\psi\rangle$. The equation of motion for $\sigma$ can easily be obtained from the Schrödinger equation for $|\psi\rangle$,

$$\frac{\mathrm{d}}{\mathrm{d}t} |\psi\rangle = -\mathrm{i} H \, |\psi\rangle \quad [7]$$

where $H$ is an appropriate spin Hamiltonian (in angular frequency units) for the spin system. The result is the Liouville–von Neumann equation for the time dependence of the density operator $\sigma$:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = -\mathrm{i}[H, \sigma(t)] \quad [8]$$

A spin system with the Hamiltonian given by

$$H = H_0 + H'(t) \quad [9]$$

is now taken, where $H_0$ is the static Hamiltonian and $H'(t)$ represents time-dependent spin–lattice coupling. $H'$ is a random function of time with a vanishing time average [i.e. $\overline{H'(t)} = 0$], and $H_0$ includes the Zeeman interactions, static averages of dipolar and quadrupole couplings, and time-dependent radiofrequency (RF) interactions. Writing $\sigma$ and $H'$ as $\tilde{\sigma}$ and $\tilde{H}'$ in the *interaction* representation and using the second-order perturbation theory, the time evolution of the density operator can be shown to obey

$$\frac{\mathrm{d}\tilde{\sigma}}{\mathrm{d}t} = -\int_0^t \mathrm{d}t' \overline{[\tilde{H}'(t), [\tilde{H}'(t-t'), (\tilde{\sigma}(t) - \sigma_{\mathrm{eq}})]]} \quad [10]$$

where the bar is now used to indicate an average over all identical molecules in the sample. Using the eigenket basis of the static Hamiltonian $H_0$ (i.e. $H_0|\alpha\rangle = \alpha|\alpha\rangle$), Redfield has obtained a set of linear differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t} \tilde{\sigma}_{\alpha\alpha'} = \sum_{\beta\beta'} \exp[-\mathrm{i}(\omega_{\alpha'\alpha} - \omega_{\beta'\beta})t] \times R_{\alpha\alpha'\beta\beta'} [\tilde{\sigma}_{\beta\beta'}(t) - \tilde{\sigma}_{\beta\beta'}(\infty)] \quad [11]$$

where $\omega_{\alpha\beta} = \alpha - \beta$ ($\equiv E_\alpha - E_\beta$), $\tilde{\sigma}_{\beta\beta'}(\infty)$ corresponds to the matrix elements $\tilde{\sigma}_{\beta\beta'}$ at thermal equilibrium, and **R**, the Redfield relaxation supermatrix, is given by

$$R_{\alpha\alpha'\beta\beta'} = U_{\alpha\alpha'\beta\beta'} + U_{\beta'\beta\alpha'\alpha} - \delta_{\alpha'\beta'} \sum_\gamma U_{\gamma\gamma\beta\alpha} - \delta_{\alpha\beta} \sum_\gamma U_{\beta'\alpha'\gamma\gamma}$$

This treatment is closely related to the relaxation theory of Wangsness and Bloch. The $U$ functions are further simplified by examining, for example, $U_{\alpha\alpha'\beta\beta'}$,

$$U_{\alpha\alpha'\beta\beta'} = \int_0^{\Delta t} \mathrm{d}\tau \exp(-\mathrm{i}\omega_{\alpha'\beta'}\tau) G_{\alpha\beta\alpha'\beta'}(\tau) \times \left[\frac{\exp\{\mathrm{i}(\alpha-\beta-\alpha'+\beta')(\Delta t-\tau)\}-1}{\mathrm{i}(\alpha-\beta-\alpha'+\beta')\Delta t}\right] \quad [12]$$

where $G_{\alpha\alpha'\beta\beta'}(\tau)$ denote time correlation functions of a stationary random function $H'(t)$, which is by definition independent of the origin of time, and

$$G_{\alpha\beta\alpha'\beta'}(\tau) = \overline{H'^{*}_{\alpha'\beta'}(t) H'_{\alpha\beta}(t+\tau)} \quad [13]$$

where $H'_{\alpha\beta}(t) = \langle \alpha | H'(t) | \beta \rangle$. Note that the integrand is large only if $\tau << \tau_c$, the correlation time for $G_{\alpha\beta\alpha'\beta'}(\tau)$. Thus, the upper limit ($\Delta t$) of the integral can be set to infinity. The bracket in the integrand is a constant ($\Delta t >> \tau$) and equals 1. In this limit, $U_{\alpha\alpha'\beta\beta'}$ become the spectral densities $J_{\alpha\beta\alpha'\beta'}(\omega_{\beta'\alpha'})$ given by

$$J_{\alpha\beta\alpha'\beta'}(\omega_{\alpha'\beta'}) = \int_0^{\infty} \exp(-\mathrm{i}\omega_{\alpha'\beta'}\tau) G_{\alpha\beta\alpha'\beta'}(\tau) \mathrm{d}\tau \quad [14]$$

and the relaxation matrix elements are now given by

$$R_{\alpha\alpha'\beta\beta'} = J_{\alpha\beta\alpha'\beta'}(\omega_{\alpha'\beta'}) + J_{\alpha\beta\alpha'\beta'}(\omega_{\alpha\beta}) - \delta_{\alpha'\beta'} \sum_{\gamma} J_{\gamma\beta\gamma\alpha}(\omega_{\gamma\alpha}) - \delta_{\alpha\beta} \sum_{\gamma} J_{\gamma\alpha'\gamma\beta'}(\omega_{\gamma\alpha'}) \quad [15]$$

Because of the large heat capacity of the lattice relative to that of the nuclear spins, the lattice may be considered at all times to be in thermal equilibrium, while the time-varying spin states, in the absence of a RF field, evolve to thermal equilibrium because of the spin–lattice interactions. When the exponential argument [$(\omega_{\alpha'\alpha} - \omega_{\beta'\beta})$ in Equation [11]] is significantly larger than the spin relaxation rates, the exponential term oscillates rapidly in comparison with the slow variation in the density matrix due to relaxation. As a consequence, the impact of these terms becomes zero. The so-called secular approximation ($\omega_{\alpha'\alpha} = \omega_{\beta'\beta}$) effectively simplifies the equation of motion to

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\sigma}_{\alpha\alpha'} = \sum_{\beta\beta'}{}' R_{\alpha\alpha'\beta\beta'}[\tilde{\sigma}_{\beta\beta'}(t) - \tilde{\sigma}_{\beta\beta'}(\infty)] \quad [16]$$

where the prime on the summation indicates that only terms that satisfy $\omega_{\alpha'\alpha} = \omega_{\beta'\beta}$ are kept. Now the exponentials in front of those $R_{\alpha\alpha\beta\beta}$ terms in Equation [11] are clearly secular. These $R_{\alpha\alpha\beta\beta}$ parameters control the spin–lattice relaxation and are associated with the diagonal elements $\tilde{\sigma}_{\alpha\alpha}$, which specify the probabilities ($P_\alpha$) that spin states $|\alpha\rangle$ are occupied. The exponentials in front of $R_{\alpha\beta\alpha\beta}$ are also secular. These $R_{\alpha\beta\alpha\beta}$ parameters control the spin–spin relaxation. When only spin–lattice relaxation is considered, the important Redfield terms in the eigenbase representation are limited to the following two types:

$$R_{\alpha\alpha\beta\beta} = 2J_{\alpha\beta\alpha\beta}(\omega_{\alpha\beta}) - 2\delta_{\alpha\beta} \sum_{\gamma} J_{\gamma\beta\gamma\alpha}(\omega_{\gamma\beta})$$
$$R_{\alpha\alpha\alpha\alpha} = -\sum_{\gamma\neq\alpha} 2J_{\gamma\alpha\gamma\alpha}(\omega_{\gamma\alpha}) = -\sum_{\gamma\neq\alpha} R_{\gamma\gamma\alpha\alpha} \quad [17]$$

Now $H'(t)$ in Equation [9] determines what is called the spin relaxation mechanism. As an example, the dipole–dipole Hamiltonian or quadrupolar Hamiltonian with an axially symmetric ($\eta = 0$) electric field gradient tensor is given by

$$H'_{\lambda}(t) = \sum_{m_L} A_{2,m_L}\left\{D^2_{m_L,0}[\Omega(t)] - \overline{D^2_{m_L,0}}\right\} \quad [18]$$

where the time dependence arises via Euler angles $\Omega$ in the Wigner rotation matrices $D^2_{m,n}(\Omega)$, and $A_{2,m_L}$ is defined by

$$A_{2,m_L} = C_{\lambda}\rho_{2,0}T_{2,m_L}$$

with $C_{\lambda}\rho_{2,0} = \sqrt{3/8}\,(e^2qQ/\hbar)$ for the quadrupole ($I = 1$) terms, and for the dipolar Hamiltonian this is $-\sqrt{3/8}\,(\mu_0\gamma_i\gamma_j\hbar/\pi r_{ij}^3)$, where $\gamma$ is the gyromagnetic ratio of a nuclear spin, $r_{ij}$ is the internuclear distance between the spin pair, and $\mu_0$ is the magnetic vacuum permeability. $T_{2,m_L}$, the spin operators in the laboratory frame, are given for a deuteron by

$$T_{2,0} = \frac{1}{\sqrt{6}}(3I_z^2 - I^2)$$
$$T_{2,\pm1} = \mp\frac{1}{2}(I^{\pm}I_z + I_zI^{\pm})$$
$$T_{2,\pm2} = \frac{1}{2}(I^{\pm})^2$$

A similar set of equations can be written for the case of a pair of $I = \frac{1}{2}$ spins. When the cross-products between spin Hamiltonian matrix elements of different

$m_L$ values can be ignored (e.g. in liquids) where $m_L$ is the projection index of a rank $L$ (=2) interaction Hamiltonian, the spectral densities of Equation [14] become

$$J_{\alpha\beta\alpha'\beta'}(\omega_{\alpha\beta}) = \sum_{m_L} \langle \alpha \,|\, A_{2,m_L} \,|\, \beta \rangle \times \langle \alpha' \,|\, A_{2,m_L} \,|\, \beta' \rangle^* J_{m_L}(\omega_{\alpha\beta}) \qquad [19]$$

where

$$J_{m_L}(\omega) = \int_0^\infty G_{m_L}(\tau) \exp(-\mathrm{i}\omega\tau) \mathrm{d}\tau \qquad [20]$$

with

$$G_{m_L}(\tau) = \left\langle \left\{ D^2_{m_L,0}[\Omega(t)] - \overline{D^2_{m_L,0}} \right\} \times \left\{ D^{2*}_{m_L,0}[\Omega(t-\tau)] - \overline{D^{2*}_{m_L,0}} \right\} \right\rangle \qquad [21]$$

It should be noted that the $J_{m_L}(\omega)$ are quantities that are obtained from experiments without reference to any molecular dynamics model. Now, Equation [11] can be transformed back to the Schrödinger representation:

$$\frac{\mathrm{d}}{\mathrm{d}t} \sigma_{\alpha\alpha'}(t) = \mathrm{i}[\sigma, H_0]_{\alpha\alpha'} + \sum_{\beta\beta'} R_{\alpha\alpha'\beta\beta'}[\sigma_{\beta\beta'}(t) - \sigma_{\beta\beta'}(\infty)] \qquad [22]$$

The first term on the right-hand side describes spin precessions and is only important for spin–spin relaxation. According to Redfield, the above equation is valid provided that the relaxation elements are small in comparison to the inverse correlation time $\tau_c^{-1}$ of the thermal motion, i.e.

$$\frac{1}{R_{\alpha\alpha'\beta\beta'}} \gg \Delta t \gg \tau_c$$

where $\Delta t$ represents the time interval over which the density matrix of the spin system has not appreciably changed. Different nuclear spin relaxation mechanisms (i.e. quadrupole, dipole–dipole, spin–rotation, chemical-shift anisotropy, and scalar spin–spin relaxation) are surveyed below. The list of relaxation mechanisms is by no means inclusive, but contains the most commonly discussed mechanisms.

## Quadrupole relaxation

Let us apply the Redfield theory to a deuteron with its quadrupole moment experiencing a fluctuating electric field gradient arising from anisotropic molecular motions in liquids. When the static average of quadrupole interaction is nonzero, i.e. $\overline{H_Q} \neq 0$, it can be included in the static Hamiltonian $H_0$. The density operator matrix for a deuteron spin is of the dimension 3 × 3 and the corresponding Redfield relaxation supermatrix has the dimension $3^2 \times 3^2$. When only nuclear spin–lattice relaxation is considered, the spin precession term in Equation [22] is set to zero and the diagonal elements $\sigma_{\alpha\alpha}$ ($\alpha$ = 1, 2, 3) satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t} P_\alpha(t) = \sum_\beta R_{\alpha\beta}[P_\beta(t) - P_\beta(\infty)] \qquad [23]$$

where $P_1 \equiv P_1$, $P_2 \equiv P_0$ and $P_3 \equiv P_{-1}$ are the populations in spin states $|1\rangle$, $|0\rangle$ and $|-1\rangle$, respectively (see **Figure 1**), and $R_{\alpha\beta} \equiv R_{\alpha\alpha\beta\beta}$ given in Equation [17]. $R_{\alpha\beta}$ represents the transition probability per second from the spin state $\beta$ to the spin state $\alpha$ and $R_{\alpha\beta} = R_{\beta\alpha}$. Thus, nuclear spin–lattice relaxation involves transitions induced between nuclear states of different energies by the time-dependent part of quadrupolar interactions $H_Q(t) - \overline{H_Q}$. Solving Equation [23] in terms of linear combinations of the eigenstate populations $P_\alpha$ gives

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \exp(-t/T_{1Z}) & 0 \\ 0 & 0 & \exp(-t/T_{1Q}) \end{pmatrix} \times \begin{pmatrix} P_1(0) + P_0(0) + P_{-1}(0) \\ P_1(0) - P_{-1}(0) \\ -P_1(0) + 2P_0(0) - P_{-1}(0) \end{pmatrix} \qquad [24]$$

where the deuteron spin–lattice relaxation times $T_{1Z}$ and $T_{1Q}$ for relaxation of the Zeeman and quadrupolar orders, respectively, are

$$T_{1Z}^{-1} = K_Q[J_1(\omega_0) + 4J_2(2\omega_0)] \qquad [25]$$

$$T_{1Q}^{-1} = 3K_Q J_1(\omega_0) \qquad [26]$$

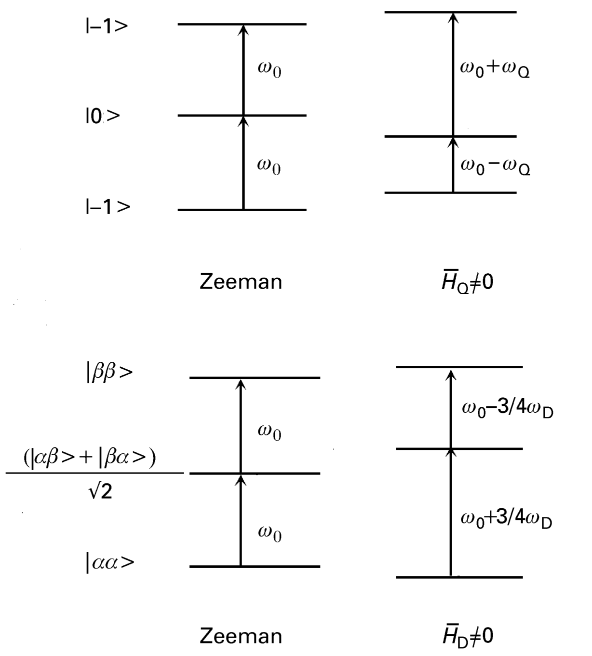



**Figure 1** Energy level diagram for a deuteron spin ($\eta = 0$) and for a pair of protons ($I = 1$ triplet; $I = 0$ is not shown) in an external magnetic field. $\omega_0/2\pi$ is the Larmor frequency.

where $K_Q = (3\pi^2/2)(e^2qQ/h)^2$. The asymmetry parameter $\eta$ of the quadrupolar coupling is assumed to be zero here. $T_{1Z}$ can be measured using an inversion–recovery pulse sequence, while $T_{1Q}$ can be obtained using the Jeener–Broekaert pulse sequence $90^\circ_x - \tau - 45^\circ_y - t - 45^\circ_y$.

When considering spin–spin relaxation, it is necessary to examine the off-diagonal elements $\sigma_{\alpha\beta}$ of the density operator matrix and there are three independent spin–spin relaxation times ($T_{2a}$, $T_{2b}$, and $T_{2D}$):

$$T_{2a}^{-1} = K_Q\left[\frac{3}{2}J_0(0) + \frac{5}{2}J_1(\omega_0) + J_2(2\omega_0)\right]$$
$$T_{2b}^{-1} = K_Q\left[\frac{3}{2}J_0(0) + \frac{1}{2}J_1(\omega_0) + J_2(2\omega_0)\right]$$
$$T_{2D}^{-1} = K_Q[J_1(\omega_0) + 2J_2(2\omega_0)] \quad [27]$$

The quadrupolar (solid) echo pulse sequence ($90^\circ_x - \tau - 90^\circ_y$) allows measurement of the spin–spin relaxation time $T_{2a}$. The double-quantum spin–spin relaxation rate $T_{2D}^{-1}$ can be determined using a double quantum spin–echo pulse sequence $90° - \tau - 90° - t_1/2 - 180° - t_1/2 - 90°$. The first two 90° pulses create the double-quantum coherence, which is refocused by a 180° pulse, and the spin–echo is detected by the last monitoring 90° pulse.

## Dipole–dipole relaxation

Treatment of spin–lattice relaxation of an isolated spin-$\frac{1}{2}$ pair by an intramolecular dipole-dipole interaction is identical to that for a spin-1 system (see **Figure 1**). Two like spin-$\frac{1}{2}$ nuclei separated by an internuclear distance $r$ are considered. The longitudinal or Zeeman spin–lattice relaxation time $T_{1Z}$ is given by Equation [25], but with $K_Q$ replaced by a different multiplicative constant $K_D = \frac{3}{2}(\mu_0\gamma^2\hbar/4\pi r^3)^2$ which determines the dipolar coupling strength. In solids with isolated spin-$\frac{1}{2}$ pairs, the Jeener–Broekaert sequence can be used to determine the dipolar spin–lattice relaxation time $T_{1D}$ which is the counterpart of $T_{1Q}$ described above. Now $T_{1Z}$ depends on the spectral density of the dipolar interaction fluctuations at the Larmor frequency and twice the Larmor frequency, whereas $T_{1D}$ depends in addition on the spectral density of dipolar fluctuations in the low-frequency region around the line width of dipolar couplings. Thus $T_{1D}$ can be quite sensitive to slow motions, which can significantly contribute to the spectral density at low frequencies. Similarly, the transverse or spin–spin relaxation rate for a spin-$\frac{1}{2}$ pair is, according to $T_{2a}^{-1}$ in Equation [27], given by

$$T_2^{-1} = K_D\left[\frac{3}{2}J_0(0) + \frac{5}{2}J_1(\omega_0) + J_2(2\omega_0)\right] \quad [28]$$

The spin–lattice relaxation rate ($T_{1\rho}^{-1}$) in the rotating frame is given by

$$T_{1\rho}^{-1} = K_D\left[\frac{3}{2}J_0(2\omega_1) + \frac{5}{2}J_1(\omega_0) + J_2(2\omega_0)\right] \quad [29]$$

where $\omega_1/\gamma$ is the spin-locking field $B_1$.

Now suppose the motional process (e.g. rotational Brownian motion in normal liquids) can be described by a single exponential correlation function of the form given in Equation [1]. The corresponding spectral density, which appears in the BPP theory, is a Lorentzian function:

$$J(\omega_0) = \frac{1}{5}\left(\frac{\tau_c}{1 + \omega_0^2\tau_c^2}\right) \quad [30]$$

where $\tau_c$ is a correlation time for the rotational motion. Hence,

$$T_1^{-1} = \frac{1}{5}K_D\left(\frac{\tau_c}{1 + \omega_0^2\tau_c^2} + \frac{4\tau_c}{1 + 4\omega_0^2\tau_c^2}\right) \quad [31]$$

$$T_2^{-1} = \frac{1}{5}K_D\left(\frac{3}{2}\tau_c + \left(\frac{5}{2}\right)\frac{\tau_c}{1+\omega_0^2\tau_c^2} + \frac{\tau_c}{1+4\omega_0^2\tau_c^2}\right) \quad [32]$$

In **Figure 2**, a sketch of these two equations as a function of $\tau_c$ is shown. As seen in this Figure, $T_1 = T_2$ in the extreme narrowing limit ($\omega_0\tau_c << 1$), which is observed in isotropic liquids. Note also that $T_1^{-1}$ goes through a maximum at $\omega_0\tau_c \sim 1$.

The above equations are developed for the rotational motion which modulates *intra*molecular dipole–dipole interactions in isotropic liquids. Translational motion can also cause spin relaxation by modulating *inter*molecular dipole–dipole interactions. Again for a pair of like spin-$\frac{1}{2}$ nuclei on two different molecules, the relaxation rate $T_{1T}^{-1}$ due to translation involves different correlation times and somewhat different spectral densities:

$$T_{1T}^{-1} = \frac{9}{8}(\mu_0\gamma^2\hbar/4\pi)^2[J_1(\omega_0) + 4J_2(2\omega_0)] \quad [33]$$

This relaxation mechanism is important when $^1$H NMR is used to study protonated samples and samples that contain paramagnetic impurities like oxygen. As the gyromagnetic ratio of an electron is about 1000 times greater than that of a proton, a small amount of paramagnetic impurity can drastically reduce the $T_1$ values in a sample.

For the case of an unlike spin-$\frac{1}{2}$ pair in a molecule (e.g. $^{13}$C–$^1$H), the Zeeman spin–lattice relaxation time for the resonant ($I$) spin depends on the spectral density at its Larmor frequency ($\omega_I$), at the sum ($\omega_I + \omega_S$) and difference ($\omega_I - \omega_S$) of the two Larmor frequencies for the spins $I$ and $S$:

$$T_1^{-1} = \frac{1}{3}K'_D[J_0(\omega_I - \omega_S) + 3J_1(\omega_I) + 6J_2(\omega_I + \omega_S)] \quad [34]$$

where $K'_D = \frac{3}{2}(\mu_0\gamma_I\gamma_S\hbar/4\pi r^3)^2$.

## Spin–rotation relaxation

A rotating molecule or a rotating group of atoms is a rotating charge system, thereby it can generate a magnetic field at the resonant nucleus. The fluctuating magnetic field depends on the magnitude and/or orientation of the angular momentum vector of the rotating molecule. Spin–rotation interactions can be an effective relaxation mechanism if the timescale of the fluctuating magnetic field at the nucleus is comparable to the inverse of the Larmor frequency. This is often encountered as the collisions among small molecules become infrequent. This mechanism is, therefore, important for gases, small molecules, and freely rotating groups ($-CH_3$, $-CF_3$). For molecules undergoing isotropic rotation in liquids, the longitudinal relaxation rate due to the spin–rotation interaction is

$$T_1^{-1} = \frac{2\pi kT}{\hbar^2}I_m C_{eff}^2 \tau_J \quad [35]$$

where $k$ is the Boltzmann constant, $T$ the absolute temperature, $I_m$ the moment of inertia of the molecule, $\tau_J$ the angular momentum correlation time, and $C_{eff}^2 = \frac{1}{3}(C_\parallel^2 + 2C_\perp^2)$ with $C_\parallel$ and $C_\perp$ being the principal elements of the spin–rotation interaction tensor **C** parallel and perpendicular to the symmetry axis of the molecule, respectively. While the correlation time $\tau_c$ for molecular rotation decreases with increasing temperature, $\tau_J$ becomes longer. They are related by the simple relation:

$$\tau_c\tau_J = \frac{I_m}{6kT} \quad [36]$$

Hence it is a characteristic feature for the spin–rotation relaxation to show an opposite temperature dependence to that observed for the other relaxation mechanisms.

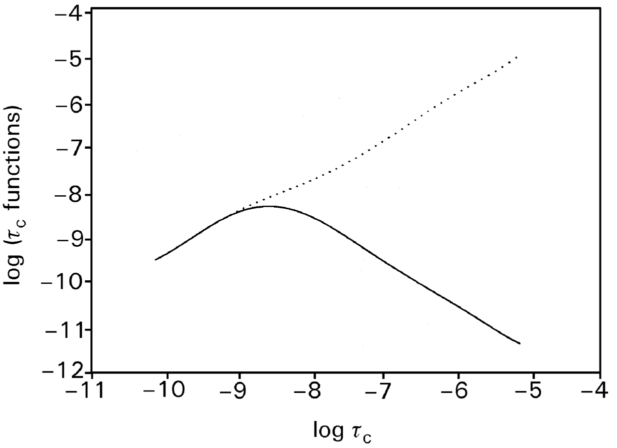



**Figure 2** Plot of the $\tau_c$ functions in Equation [31] (solid curve) and Equation [32] (dotted curve) as a function of the correlation time $\tau_c$ for $\omega_0/2\pi = 100$ MHz.

## Chemical-shift anisotropy relaxation

The magnetic field at a resonant nucleus depends on the applied $B_0$ field, as well as the screening of the applied field at the nucleus by the surrounding electrons. As the shielding (or chemical shift) tensor is anisotropic, its magnitude fluctuates with time due to modulations by molecular rotations. The fluctuating magnetic field at a nucleus from the chemical-shift anisotropy can cause spin relaxation. This relaxation mechanism depends on the square of the applied $B_0$ field and is therefore important at high fields. It also tends to dominate for nuclei exhibiting large chemical-shift ranges. For molecules with axial symmetry, one can obtain:

$$T_1^{-1} = \frac{2}{15}\gamma^2 B_0^2(\sigma_{\|} - \sigma_{\perp})^2 \frac{\tau_c}{1+\omega^2\tau_c^2} \quad [37]$$

$$T_2^{-1} = \frac{2}{15}\gamma^2 B_0^2(\sigma_{\|} - \sigma_{\perp})^2 \left[\frac{2\tau_c}{3} + \left(\frac{1}{2}\right)\frac{\tau_c}{1+\omega^2\tau_c^2}\right] \quad [38]$$

where $\sigma_{\|}$ and $\sigma_{\perp}$ are the principal components of the chemical-shift tensor parallel and perpendicular to the symmetry axis of the molecule, respectively. In contrast with the dipole–dipole relaxation, the longitudinal and transverse rates do not tend to the same value in the limit of short $\tau_c$, as seen in the above equations.

## Scalar relaxation

For spin–spin (scalar) coupling between spins $I$ and $S$, where $I = \frac{1}{2}$ and $S \geq \frac{1}{2}$, the interaction Hamiltonian involves the scalar coupling tensor **J**. The scalar relaxation of the nucleus $I$ can arise from either a time-dependent $S$ or a time-dependent **J** ('first kind'). If the relaxation time ($T_1(S)$) of the nucleus $S$ is short compared with $1/J$, where $J$ is the scalar coupling constant, the nucleus $I$ sees the average of the spin–spin interaction and does not show the expected multiplet, but a single line. The scalar relaxation correlation time $\tau_s$ is then equal to $T_1(S)$, and the scalar spin–spin contributions to the longitudinal and transverse relaxation times of the $I$ nucleus are given by

$$T_1^{-1} = \frac{8\pi^2 J^2}{3} S(S+1) \frac{\tau_s}{1+(\omega_I - \omega_S)^2\tau_s^2} \quad [39]$$

$$T_2^{-1} = \frac{4\pi^2 J^2}{3} S(S+1) \left[\tau_s + \frac{\tau_s}{1+(\omega_I - \omega_S)^2\tau_s^2}\right] \quad [40]$$

Scalar relaxation of the first kind involves collapse of spin multiplets. When $T_1(S)$ is primarily due to the quadrupole relaxation for $S > \frac{1}{2}$, this is often referred to as scalar coupling of the second kind. In the above equations, the denominators involving $\omega_I - \omega_S$ become very large when the Larmor frequencies $\omega_I$ and $\omega_S$ are very different. In this case, the scalar relaxation becomes unimportant for $T_1$, but still exists for $T_2$ due to the frequency-independent term in Equation [40].

## Spectral density of motion

As mentioned above, the evaluation of correlation functions or spectral densities is a daunting task for any relaxation theory. To further complicate the matter, different motional (or relaxation) processes can simultaneously occur in the material being studied by nuclear spin relaxation. However, the observed relaxation rate can often be given by

$$T_1^{-1} = T_{1a}^{-1} + T_{1b}^{-1} + T_{1c}^{-1} + \cdots \quad [41]$$

provided that different relaxation mechanisms labelled by the subscripts a, b, c occur at very different timescales. Otherwise, possible couplings between these processes may also exist and their contributions to relaxation must be properly treated.

In the above discussion of the quadrupole and dipole–dipole relaxation, the relaxation rates are written in terms of spectral densities for general applications. As an example, the reorientation correlation functions ($g_{mn}(t)$) for molecules rotating in an anisotropic medium are calculated using a rotational diffusion model. The rotational diffusion equation, which involves a rotational diffusion operator ($\Gamma$) and also contains the pseudopotential for reorienting molecules, must first be solved to get the conditional probability that a molecule has a certain orientation at time $t$ given it has a different orientation at time $t = 0$. This, together with the equilibrium probability for finding the molecule with a certain orientation, is required to work out $g_{mn}(\tau)$. In general, the orientational correlation functions can be written as a sum of decaying exponentials:

$$g_{mn}(t) = \sum_K (\beta_{mn}^2)_K \exp[(\alpha_{mn}^2)_K t] \quad [42]$$

where $m$ and $n$ represent the projection indices of a rank 2 tensor in the laboratory and molecular frames, respectively; $(\alpha_{mn}^2)_K/\rho$, the decay constants, are the eigenvalues of the rotational diffusion $\Gamma$ matrix and $(\beta_{mn}^2)_K$, the relative weights of the exponentials, are

the corresponding eigenvectors. In this model, the decay constants contain the model parameters $D_{\|}$ and $D\perp$ specifying rotational diffusions of the molecule about its long axis and perpendicular to the long axis. The spectral densities for a deuteron residing on the rigid part of a uniaxial molecule are the Fourier transform of the orientational correlation functions ($m = 0, 1$, or 2) to give

$$J_m(m\omega) = \sum_n [d^2_{n,0}(\beta_{M,Q})]^2 \sum_K \frac{(\beta^2_{mn})^2_K (\alpha^2_{mn})^2_K}{(\alpha^2_{mn})^2_K + m^2\omega^2} \quad [43]$$

These can now be substituted into Equations [25–27] to obtain deuteron relaxation rates.

By fitting the experimental spectral densities with the predictions from a certain motional model, its model parameters can then be derived. However, the derived motional parameters are model dependent. It is a price one normally has to pay when using NMR relaxation rates. Justification of NMR model parameters may be obtained by comparing them with those observed by other spectroscopic techniques.

## List of symbols

$B_1$ = RF field along an axis of the rotating frame; $\mathbf{C}$ = spin–rotation interaction tensor; $d^l_{mn}(\Omega)$ = reduced Wigner rotation matrix elements; $D^l_{mn}(\Omega)$ = Wigner rotation matrix elements; $eQ$ = nuclear electric quadrupole moment; $eq$ = electric field gradient at nucleus; $g_{mn}(t)$ = reduced time correlation function; $G_m(t)$ = time correlation function of spin coupling tensor; $H_0$ = static spin Hamiltonian; $H'(t)$ = zero-average, time-dependent spin Hamiltonian; $H_q(t)$ = $q$th component of time-dependent spin Hamiltonian; $I$ = nuclear spin angular momentum; $I_m$ = moment of inertia; $J$ = scalar coupling constant; $J_n(n\omega_0)$ = spectral density of the $n$th component of a fluctuating coupling tensor at frequency $n\omega_0$; $P_\alpha$ = population of the spin state $|\alpha\rangle$; $r_{ij}$ = internuclear distance; $\mathbf{R}$ = Redfield relaxation supermatrix; $R_{\alpha\alpha'\beta\beta'}$ = Redfield relaxation supermatrix elements; $T_{1Z}$ = longitudinal or Zeeman spin–lattice relaxation time; $T_{1D}$ = dipolar spin–lattice relaxation time; $T_{1Q}$ = quadrupole spin–lattice relaxation time; $T_{1T}$ = longitudinal relaxation time due to translation; $T_{1\rho}$ = rotating frame spin–lattice relaxation time; $T_{2,m}$ = spin operator tensor; $T_2$ = transverse or spin–spin relaxation time; $\gamma$ = nuclear gyromagnetic ratio; $\eta$ = asymmetry parameter of electric field gradient tensor; $\sigma$ = chemical-shift tensor; $\sigma$ = density operator; $\tilde{\sigma}$ = density operator in *interaction* representation; $\tau_c$ = correlation time; $\tau_J$ = angular momentum correlation time; $\tau_s$ = scalar relaxation correlation time; $\psi$ = wavefunction; $\omega_D$ = average dipolar coupling expressed as a frequency; $\omega_0$ = Larmor precession frequency; $\omega_Q$ = average quadrupole coupling expressed as a frequency; $\Omega$ ($\frac{1}{2}\alpha, \beta, \gamma$) = Euler angles; $90^\circ_x$, $45^\circ_y$ = RF pulses producing rotations of 90°, 45° about the $x$, $y$ axes of the rotating frame.

*See also:* **Chemical Shift and Relaxation Reagents in NMR; Liquid Crystals and Liquid Crystal Solutions Studied By NMR; NMR in Anisotropic Systems, Theory; NMR Principles; Nuclear Overhauser Effect.**

## Further reading

Abragam A (1961) *The Principles of Nuclear Magnetism*. Oxford: Clarendon.

Bloembergen N, Purcell EM and Pound RV (1948) *Physical Review* 73: 679.

Cowan B (1997) *Nuclear Magnetic Resonance and Relaxation*. Cambridge: Cambridge University Press.

Dong RY (1997) *Nuclear Magnetic Resonance of Liquid Crystals*, 2nd edn. New York: Springer.

Farrar TC (1989) *Introduction to Pulse NMR Spectroscopy*. Madison: Farragut Press.

Goldman M (1988) *Quantum Description of High-Resolution NMR in Liquids*. Oxford: Clarendon.

Jeener J and Broekaert P (1967) *Physical Review* **157**: 232–240.

Redfield AG (1965) *Advances in Magnetic Resonance* **1**: 1–32.

Slichter CP (1990) *Principles of Magnetic Resonance*, 3rd edn. New York: Springer.

# NMR Spectrometers

**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
**Methods & Instrumentation**

## Introduction

After the first observation of nuclear magnetic resonance in bulk phases in 1946 and the realization that it would be useful for chemical characterization, which first came with the discovery of the chemical shift in 1951, it was only a few years before commercial spectrometers were produced. By the end of the 1950s a considerable number of publications on the application of NMR to chemical structuring and analysis problems had appeared, and then during the 1960s and later it became clear that useful information could be obtained in biological systems. Since then, the applications and the consequential instrument developments have diversified and now NMR spectroscopy is one of the most widely used techniques in chemical and biological analysis. The very high specificity, the exploratory nature of the technique without the need to preselect analytes and its nondestructive nature have made it very useful despite its lower sensitivity compared to some spectroscopic methods.

A general description is given of the way in which a modern NMR spectrometer operates, of the various components that go into making a complete system and of the particular role that they play. A block diagram of the components of a high-resolution NMR spectrometer is given in **Figure 1**.

## Components and principles of operation of NMR spectrometers

### Continuous wave (CW) and Fourier transform (FT) operation

For many years, all commercial NMR spectrometers operated in continuous wave mode. This type of operation required a sweep of the NMR frequency or the magnetic field over a fixed range to bring each nucleus into resonance one at a time. These scans for $^1$H NMR spectroscopy would take typically 500 s to avoid signal distortion. Since most NMR spectra consist of a few sharp peaks interspersed with long regions of noise, this was a very inefficient process. A fundamental paper by Ernst and Anderson in 1966 pointed out the favourable gain in efficiency that could be obtained by simultaneously detecting all signals. This is achieved by the application of a short intense pulse of RF radiation to excite the nuclei, followed by the detection of the induced magnetization in the detector coil as the nuclei relax. The decaying, time-dependent signal, known as a free induction decay (FID) is then converted to the usual frequency domain spectrum by the process known as Fourier transformation (FT). For speed of implementation, in NMR computers this requires the data to have a number of values that is a power of 2, typically perhaps 16K points for modest spectral widths, up to 128K or even 256K points for wide spectral widths on high-field spectrometers (1K is 1024 or $2^{10}$ points). Acquisition of a $^1$H FID requires typically a few seconds and opens up the possibility of adding together multiple FID scans to improve the spectrum signal-to-noise ratio (S/N), since for perfectly registered spectra the signals will co-add but the noise will only increase in proportion to the square root of the number of scans. The S/N gain, therefore, is proportional to the square root of the number of scans. This, for the first time, made routine the efficient and feasible acquisition of NMR spectra of less sensitive or less abundant nuclei such as $^{13}$C.

### The magnet

The most fundamental component of an NMR spectrometer is the magnet. Originally, this would have been a permanent or electromagnet and these provided the usual configurations for field strengths up to 1.41 T (the unit of magnetic flux density is the tesla (T) equivalent to 10 000 gauss), corresponding to a $^1$H observation frequency of 60 MHz. Because the sensitivity of the NMR experiment is proportional to about the 3/2 power of the field strength, denoted $B_0$, there has been a drive to higher and higher magnetic fields. This led the commercial NMR manufacturers to develop stronger electromagnets for NMR spectroscopy that took the highest field strengths to 2.35 T, i.e. 100 MHz for $^1$H NMR observation. Materials suitable for electromagnets have a maximum saturation field strength at about this value and at this field the current used and the consequential water cooling required was a considerable running expense.

Because of the continued need for even higher strengths, NMR manufacturers have collaborated closely with magnet developers to produce high-resolution magnets based upon superconducting solenoids. The magnetic field is generated by a current circulating in a coil of superconducting wire immersed in a liquid helium dewar at 4.2 K. This bath is shielded from ambient temperature by layers of vacuum and a jacket of liquid nitrogen at 77 K, which is usually topped up at a weekly interval. A liquid helium refill is carried out at approximately 2-month intervals depending on the age and field strength of the magnet. The initial development of superconducting magnets was at 5.17 T, corresponding to 220 MHz for $^{1}$H and operated in continuous wave (CW) mode (q.v.)

Until about 1972, this represented the highest field strength, but then at regular intervals the available field strength gradually increased along with the emergence of wider-bore magnets, enabling the incorporation of larger samples. Thus, a 270 MHz spectrometer was produced along with a wide-bore 180 MHz machine, and subsequently the field was increased to allow $^{1}$H observation at 360 MHz, 400 MHz, 500 MHz, 600 MHz, 750 MHz; the observation frequency limit of any machine yet delivered to a customer is 800 MHz (mid 1999). The development of such magnets has required new technology in which part of the liquid helium bath is kept at about 2K by an adiabatic cooling unit, thereby allowing higher current to be used in the coils. This approach should lead to higher field strengths being available in

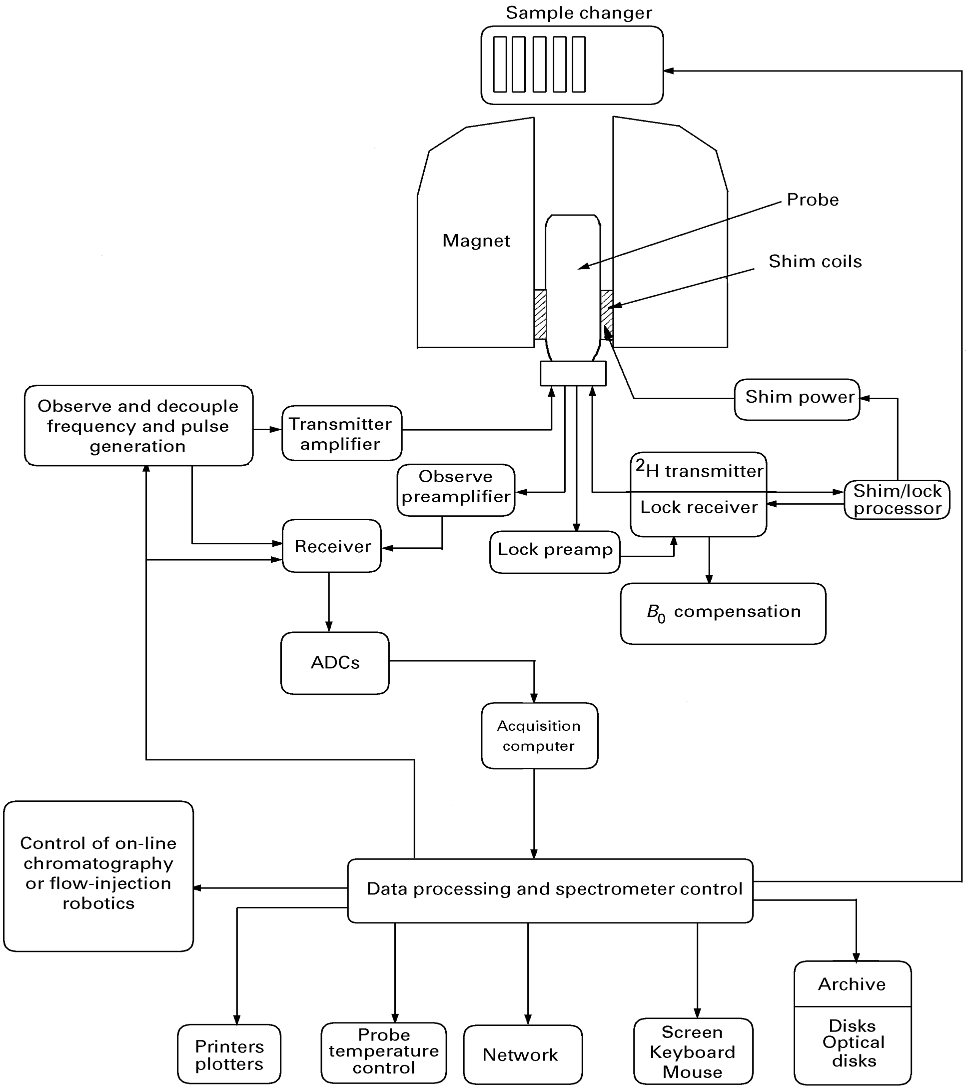



**Figure 1** A block diagram of the principal components of a modern NMR spectrometer.

the near future. A photograph of a superconducting magnet designed to give a field of 18.8 T and $^{1}$H NMR spectra at 800 MHz is shown in **Figure 2**, indicating the size that such magnets have reached. A modern, recently installed high-field NMR spectrometer using this type of superconducting magnet is shown in **Figure 3**. Nowadays, apart from very basic routine low-field spectrometers used, for example, for monitoring chemical reactions, all NMR spectrometers are based on superconducting magnets.

### NMR magnetic field optimization, signal detection and sample handling

Inserted into the magnet is the NMR detector system or probe. High-resolution NMR spectra are usually measured in the solution state in glass tubes of standard external diameters; 5 mm is the most common, but larger ones (10 mm) are used where improved sensitivity is required and sample is not limited. Also a range of narrow and specially designed tubes is available for limited sample studies, including 4 mm, 3 mm, 1.5 mm diameter and even smaller specially shaped cavities such as capillaries or spherical bulbs, plus tubes containing limited-volume cavities where the glass has a magnetic susceptibility tailored to be the same as that of a specific NMR solvent such as $D_2O$. The probe contains tunable RF coils for excitation of the nuclear spins and detection of the resultant signals as the induced magnetization decays away. A capability exists for measuring NMR spectra over a range of temperatures, typically 125–475 K.

Although modern high-resolution magnets have very high field stability and homogeneity, this is not sufficient for chemical analysis, in that it is necessary to resolve lines to about a width of 0.2 Hz; at 800 MHz this represents a stability of one part in $4 \times 10^9$. This performance is achieved in three ways: first by locking the magnetic field to the RF to ensure that successive scans are co-registered; second by improving the homogeneity of the magnetic field; and finally by sometimes spinning the sample tube.

Deuterated solvents are usually used for NMR spectroscopy to avoid the appearance of solvent peaks in the $^{1}$H spectrum. Deuterium is an NMR-active nucleus and the spectrometer will contain a $^{2}$H channel for exciting and detecting the solvent resonance. Circuitry is provided in the spectrometer for maintaining this $^{2}$H signal exactly on resonance at all times by detecting any drift from resonance caused by inherent magnet drift or room temperature fluctuations and for providing an error signal to bring the magnet field back on resonance by applying small voltages through subsidiary coils in the magnet bore. This is known as a 'field-frequency lock' and it means that successive scans in a signal accumulation run are registered exactly.

To improve the homogeneity of the magnet, an assembly of coils is inserted into the magnet bore (shim coils). These consists of about 20–40 coils specially designed so that adjustable current can be fed through them to provide corrections to the magnetic field in any combination of axes to remove the effects of field inhomogeneities. The criterion of the best

**Figure 2** A superconducting NMR magnet operating at 18.8 T for $^{1}$H NMR observation at 800 MHz demonstrating the size of these state-of-the-art magnets. Photograph courtesy of Bruker Instruments Inc., Billerica, MA, USA. (See Colour Plate 41a).

homogeneity is based upon the fact that when the $^2$H lock signal is sharpest (i.e. at the most homogeneous field) the signal will be at its highest. The currents in the shim coils are therefore usually adjusted to give the highest lock signal. Alternatively, it is possible, although less common, to 'shim' on the $^1$H NMR signal. It is possible to 'map' the field inhomogeneities using MRI methods involving magnetic field gradients prior to automatic compensation. This whole process is now largely computer-controlled in modern spectrometers.

NMR spectra are sometimes measured with the sample tube spinning at about 20 Hz to further improve the NMR resolution. This can introduce signal sidebands at the spinning speed and its harmonics, and on modern high-field machines with improved resolution, this is becoming less necessary and is undesirable in some cases.

In analytical laboratories where large numbers of samples have to be processed, automatic sample changers can play a large part in improving efficient use of the magnet time. These devices allow the measurement of up to about 120 samples in an unattended fashion with insertion and ejection of samples from the magnet under computer control. Automatic lock detection and optimization of sample spinning, NMR receiver gain and shimming are also standard. The data are acquired automatically and can be plotted and stored on backing devices. As an additional aid in routine work, it is possible to purchase an automated work bench that will produce the samples dissolved in the appropriate solvent in an NMR tube starting from a solid specimen in a screw-capped bottle and which will also dispose of samples safely and wash the NMR tube. It is possible to foresee the demise of the glass NMR tube in laboratories requiring high sample throughout. This can now be achieved using a flow probe type of NMR detector and automatic sample handling robots taking samples from 96-well plates. This is an extension of the technology used for direct coupling of chromatography, such as HPLC, to NMR spectroscopy.

**Figure 3** A modern high-resolution NMR spectrometer. A superconducting magnet is shown at the rear, in this case providing a field of 18.8 T corresponding to a $^1$H observation frequency of 800 MHz. Behind the operator is the single console containing the RF and other electronics and the temperature-control unit. The whole instrument is computer controlled by the workstation shown at the right. Photograph courtesy of Bruker Instruments Inc., Billerica, MA, USA. (See Colour Plate 41b).

**Excitation, detection and computer processing of NMR signals.**

The RF signal is derived ultimately from a digital frequency synthesizer that is gated and amplified to provide a short intense pulse. Pulses have to be of short duration because of the need to tip the macroscopic nuclear magnetization by 90° or 180° and at the same time to provide uniform excitation over the whole of the spectral range appropriate for the nucleus under study. Thus for $^{13}$C NMR, for example, where chemical shifts can cover more than 200 ppm, this requires 25 kHz spectral width on a spectrometer operating at 500 MHz for $^{1}$H, which corresponds to 125 MHz for $^{13}$C. To cover this range uniformly requires a 90° pulse to be $< 10$ μs in duration.

The RF pulse is fed to the NMR probe, which contains one or more coils that can be tuned and matched to the required frequency, this tuning changing from sample to sample because of the different properties of the samples such as the solution dielectric constant. The receiver is blanked off during the pulse and for a short period afterwards to allow the pulse amplifier to recover. The receiver is then turned on to accept the NMR signal that is induced in the coil as the nuclei precess about the field and decay through their relaxation processes. The detection coil is wound on a former as close as possible to the sample to avoid signal losses and is oriented with its axis perpendicular to the magnetic field. In a superconducting magnet the sample tube is aligned along the field, and this coil axis is therefore at right angles to the field and a simple solenoid, which would provide the best S/N, is not possible. Consequently most detector coils are of the saddle type.

The weak NMR signal is amplified using a preamplifier situated as close to the probe as possible, and then also in the main receiver unit where it is mixed with a reference frequency and demodulated in several stages to leave the FID as an oscillating voltage in the kHz range. This signal is then fed to an analogue-to digital converter (ADC) and at this point the analogue voltage from the probe is converted into a digital signal for data processing. ADCs are described in terms of their resolution, usually in terms of the number of bits of resolution: a typical high-field NMR FID is digitized to a resolution of 16 bits or one part in $2^{16}$ or 65 536. This digital signal can then be manipulated to improve the S/N ratio or the resolution by multiplying the FID by an appropriate weighting function before the calculation of the digital Fourier transform.

If only one ADC is used to collect the NMR FID, it is not possible to distinguish frequencies that are positive from those that are negative with respect to the pulse frequency. For this reason, the carrier frequency used to be set to one edge of the spectral region of interest to make sure that all of the NMR frequencies detected were of the same sign. This had the disadvantage of allowing all of the noise on the unwanted side of the carrier to be aliased onto the noise in the desired spectral region, hence reducing the final S/N by $\sqrt{2}$. To overcome this problem it is general practice now to collect two FIDs, separated in phase by 90°, either using two ADCs or multiplexing one ADC to two channels. This approach allows the distinction of positive and negative frequencies and means that the carrier can be set in the middle of the spectrum and the hardware filters can be correspondingly reduced in width by a factor of 2, giving an increase in S/N by $\sqrt{2}$. This process is termed quadrature detection.

In modern NMR spectrometers, the electronics are largely digital in nature, thus providing greater opportunities for computer control and manipulation of the signals. This includes the use of oversampling and digital filtering to improve the dynamic range of the signal acquisition.

Modern NMR spectrometers usually have two separate computer systems. One is dedicated to the acquisition of the NMR FID and operates in the background so that all necessary accurate timing requirements can be met. The FID is transferred, either at the end of the acquisition or periodically throughout it to enable inspection of the data, to the host computer for manipulation by the operator. These computers are based on modern operating systems such as UNIX. The computer software can be very complex, using multiple graphics windows on remote processors, and can, like any modern package, take advantage of networks, printers and plotters. Typical operations include manipulations of the signal-averaged FID by baseline correction to remove DC offset; multiplication by continuous functions to enhance S/N or resolution; Fourier transformation; phase correction; baseline correction of the frequency spectrum; calculation and output of peak lists; calculation and output of peak areas (integrals); and plotting or printing of spectra. It is common to have a separate computer workstation solely for data inspection and manipulation, networked to the host computer. This may be the same model as the host computer but is often an industry-standard model from a third-party supplier. NMR data processing software can also be purchased from a number of companies other than the instrument manufacturers, and these often have links to document production software or provide output of NMR parameters for input into other

packages such as those for molecular modelling. A number of approaches alternative to the use of FID weighting functions for improving the quality of the NMR data have been developed and are available from software suppliers. These include such methods as maximum entropy and linear prediction, and indeed it is now possible to purchase these as supplementary items from some NMR manufacturers.

## Multiple-pulse experiments and multidimensional NMR

Everything described so far applies to the basic one-dimensional NMR experiment in which the nuclear spin system is subjected to a 90° (or less) pulse and the FID is collected. A wide variety of experiments are reported in the literature and are routinely applied to measure NMR properties such as relaxation times $T_1$, $T_2$ and $T_{1\rho}$, which can be related in some cases to molecular dynamics. These experiments involve the use of several pulses separated by timed variable delays and are controlled by pulse programs written in a high-level language for ease of understanding and modification. The computer system will have software to interpret the data and calculate the relaxation times using least-squares fitting routines.

Such pulse programs are also used to enable other special one-dimensional experiments such as saturation or nonexcitation of a large solvent resonance (these are different in that the former method will also saturate NH or OH protons in the molecules under study through the mechanism of chemical exchange), or the measurement of nuclear Overhauser enhancement (NOE) effects which are often used to provide distinction between isomeric structures or to provide estimates of internuclear distances. Pulse programs are also used for measuring NMR spectra of nuclei other than $^1$H and sometimes in order to probe connectivity between protons and the heteronucleus. In this case, pulses or irradiation can be applied on both the heteronucleus and $^1$H channels in the same experiment. The commonest use is in $^{13}$C NMR where all spin–spin couplings between the $^{13}$C nuclei and $^1$H nuclei are removed by ‘decoupling’. This involves irradiation of all of the $^1$H frequencies while observing the $^{13}$C spectrum. In order to cover all of the $^1$H frequencies, the irradiation is provided as a band of frequencies covering the $^1$H spectral width; this is consequently termed noise decoupling or broad-band decoupling. Alternatively, it is possible to obtain the effect of broad-band decoupling more efficiently by applying a train of pulses to the $^1$H system, this being known as composite pulse decoupling.

Recently, a whole family of experiments have been developed that detect low-sensitivity nuclei such as $^{13}$C or $^{15}$N indirectly by their spin coupling connectivity to protons in the molecule. This involves a series of pulses on both $^1$H and the heteronucleus but allows detection at the much superior sensitivity of $^1$H NMR. Special probes have been developed for such ‘indirect detection’ experiments in which the $^1$H coil is placed close to the sample, and the heteronucleus coil is placed outside it, the opposite or ‘inverse geometry’ to a standard heteronuclear detection probe.

The one-dimensional NMR experiment is derived from measuring the FID as a function of time. If the pulse program also contains a second time period which is incremented, then a second frequency axis can be derived from a second Fourier transform. This is the basis for ‘two-dimensional’ NMR and its extension to three or even four dimensions. For example, a simple sequence such as

$$90°(^1\text{H})–t_1–90°(^1\text{H})–\text{collect FID for time } t_2$$

where $t_1$ is an incremented delay, results after double Fourier transformation with respect to $t_1$ and $t_2$ in a spectrum with two axes each corresponding to the $^1$H chemical shifts. This is usually viewed as a contour plot with the normal 1D spectrum appearing along the diagonal and any two protons that are spin coupled to each other giving rise to an off-diagonal contour peak at their chemical shift coordinates. This simple experiment is one of a large family of such correlation experiments involving either protons alone or heteronuclei. The extension to higher dimensions has already been exploited to decrease the amount of overlap by allowing spectral editing and the spreading of the peaks into more than one dimension. Hardware and software in modern NMR spectrometers allows this wide variety of experiments.

The increasingly complex pulse sequences used today rely on the ability of the equipment to produce exactly 90° or 180° pulses or pulses of any other angle. One way to do this is to provide trains of pulses that have the desired net effect of, for example, a 180° tip but which are compensated for any mis-setting. An example of such a ‘composite pulse’ is $90^\circ_x - 180^\circ_y - 90^\circ_x$, which provides a better inversion pulse than a single 180° pulse. Many complex schemes have been invented both for observation and for decoupling (especially for low-power approaches that avoid heating the sample). A universal approach to removing artefacts caused by electronic imperfections, and one which is also used to simplify spectra by editing out undesired components of magnetization, is the

use of ‘phase cycling’. This allows the operator to choose the phase of any RF pulse and of the receiver, and cycling these in a regular fashion gives control over the exact appearance of the final spectrum.

So far only pulses that excite the whole spectrum (hard pulses) have been described. For spectral editing purposes or to prove some NMR spin connectivity, it can be very convenient to perturb only part of spectrum, possibly only that corresponding to a given chemical shift or even one transition in a multiplet. This approach is achieved by using lower-power pulses applied for a longer period of time (e.g. a 10 ms 90° pulse will only cover 25 Hz). Such selective pulses are often not rectangular as are hard pulses but can be synthesized in a variety of shapes such as sine or Gaussian because of their desirable excitation frequency profiles. Modern research spectrometers can include such selective, shaped pulses in pulse programs.

## Instruments for special applications

### NMR of solids

Although $^{1}$H high-resolution NMR spectroscopy is possible in the solid, most applications have focused on heteronuclei such as $^{13}$C. High-resolution studies rely on very short pulses, so high-power amplifiers are necessary. Similarly, because of the need to decouple $^{1}$H from $^{13}$C and thereby to remove dipolar interactions not seen in the liquid state, high-power decoupling is required. However, the major difference between solution and solid-state high-resolution NMR studies lies in the use of ‘magic-angle spinning’ (MAS) in the latter case. This involves spinning the solid sample packed into a special rotor at an angle of 54°44′ to the magnetic field. This removes broadening due to any chemical shift anisotropies that are manifested in the solid-state spectrum and any residual $^{1}$H–$^{13}$C dipolar coupling not removed by high-power decoupling. Typical spinning speeds are 2–6 kHz or 120 000–720 000 rpm although higher speeds up to 25 kHz, at which the rotor rim is moving at supersonic velocity, are possible and necessary in some cases.

For nuclei with spin $> \frac{1}{2}$, MAS is insufficient to narrow the resonances and more complicated double angle spinning (DAS) or double orientation rotors (DOR) are necessary.

### NMR imaging

A whole new specialized subdivision of NMR has arisen in the allied disciplines of NMR imaging (magnetic resonance imaging or MRI) and NMR spectroscopy from localized regions of a larger object. MRI applications range from the analysis of water and oil in rock obtained from oil exploration drilling to medical and clinical studies, and spectroscopic applications include the possibility of measuring the $^{1}$H or $^{31}$P NMR spectrum from a particular volume element in the brain of a living human being and relation the levels of metabolites seen to a disease condition. Some experiments on smaller samples can be carried out in the usual vertical-bore superconducting magnets, but studies are more often performed in specially designed horizontal-bore magnets with a large, clear bore capable of taking samples up to the size of adult human beings. Because of their large bore, they operate at lower field strengths compared to analytical chemical applications, and typical configurations are 2.35 T with a 40 cm bore or 7.0 T with a 21 cm bore. Clinical imagers generally utilize magnetic fields up to 2 T with a 1 m bore. Imaging relies upon the application of magnetic field gradients to extra coils located inside the magnet bore in all three orthogonal axes including that of $B_0$ and excitation using selective RF pulses. Virtually all clinical applications of MRI use detection of the $^{1}$H NMR signal of water in the subject, with the image contrast coming from variation of the amount of water or its NMR relaxation or diffusion properties in the different organs or compartment being imaged. Very fast imaging techniques have been developed that allow movies to be constructed of the beating heart or studies of changes in brain activity as a result of visual or aural stimulation to be conducted.

### Benchtop analysis

Specialist tabletop machines can be purchased and these are used for routine analysis in the food and chemical industries. They operate automatically, typically at 20 MHz for $^{1}$H NMR, using internally programmed pulse sequences, and are designed to give automatic printouts of analytical results such as the proportion of fat to water in margarine or the oil content of seeds.

## Future trends

NMR spectroscopy has shown a ceaseless trend in improvements in S/N, field strength, new types of pulse experiments and computational aspects. This trend is not slowing down and, with the rapid advances in computers, it is probably accelerating; it is therefore difficult to predict NMR developments in the long term. However, some recent research developments mentioned below will certainly break through into commercial instruments.

**Higher magnetic field strengths** 800 MHz detection for $^{1}H$ NMR is the current (mid 1999) commercial limit and the first machines at this field have now been delivered and 900 MHz systems are being developed. Higher fields must be on the way and clearly an emotive figure would be the 1 GHz $^{1}H$ NMR spectrometer. This development will require the design of transmitter and detection technology working at or beyond the limit of RF methods and investigation of new superconducting materials for the magnets. Although the higher field strengths provide greater spectral dispersion and yield better sensitivity, it may be that some applications involving heavier nuclei are less suited to such fields because of the field dependence of certain mechanisms of nuclear spin relaxation, which could cause an increased line broadening and hence lower peak heights and delectability. It has been demonstrated that cooling the NMR detector to liquid helium temperature has the effect of improving the S/N by up to about 500%. This will have an even more dramatic effect on sensitivity than higher magnetic fields.

**New NMR pulse experiments** Four-dimensional experiments are reported in the literature and developments, through such approaches as selective excitation, allow the reduction of the enormous data matrices that result. This also means that new methods of detecting only the desired information in complex spectra are becoming possible through such approaches as the detection of $^{1}H$ NMR resonances only from molecules containing certain isotopes of other nuclei. Transfer of the use of pulsed magnetic field gradients has occurred from the MRI field to the high-resolution NMR area and this provides new ways of editing complex spectra with improved data quality and acquisition speed. This technology will find widespread application in the near future, for example in the measurement of diffusion coefficients and other forms of molecular mobility.

**Novel data processing** The advent of more widespread application of the maximum-entropy technique, where any prior knowledge about the system can be used to advantage, is imminent as the method becomes more widely available. It will probably gain more credence when careful bench-marking and comparisons have been completed. Undoubtedly, it will find application in all areas of NMR spectroscopy.

**Coupled techniques** The recent coupling of HPLC to NMR has been shown to be of great use in separating and structuring components of complex mixtures such as drug metabolites in body fluids. This technique has been extended to other chromatographic techniques such as supercritical fluid chromatography (SFC) and to the use of nuclei other than $^{1}H$ or $^{19}F$ which form the basis of most studies so far because of their high NMR sensitivity. The direct coupling of capillary electrophoresis (CE) and capillary electrochromatography (CEC) to NMR has also been developed and commercial systems based on these approaches will become available. The hyphenation of HPLC with both NMR spectroscopy and mass spectrometry has been achieved and the first commercial systems are now being produced. It is expected that a wealth of applications based on these technologies, such as the identification of drug metabolites, will be forthcoming. Finally, the technology that has led to the direct coupling of separation to NMR spectroscopy is leading to the demise of the glass NMR tube for high-throughput applications and its replacement by flow-injection robots.

## List of symbols

$B_0$ = magnetic field strength [flux density]; $T_1$ = spin–lattice relaxation times; $T_2$ = spin–spin relaxation time; $T_{1\rho}$ = spin–lattice relaxation time in the rotating frame.

*See also:* **Diffusion Studied Using NMR Spectroscopy; Fourier Transformation and Sampling Theory; Magnetic Field Gradients in High Resolution NMR; MRI Theory; NMR Data Processing; NMR Principles; NMR Pulse Sequences; NMR Relaxation Rates; Solid State NMR, Methods; Solvent Suppression Methods in NMR Spectroscopy; Two-Dimensional NMR, Methods.**

## Further reading

Ernst RR and Anderson WA (1996) Applications of Fourier transform spectroscopy to magnetic resonance. *Review of Scientific Instruments* **37**: 93–102.

Lindon JC and Ferrige AG (1980) Digitisation and data processing in Fourier transform NMR. *Progress in NMR Spectroscopy* **14**: 27–66

Sanders JKM and Hunter BK (1993) *Modern NMR Spectroscopy. A Guide for Chemists*, 2nd edn. Oxford: Oxford University Press.

# NMR Spectroscopy in Food Science

*See* **Food Science, Applications of NMR Spectroscopy.**

# NMR Spectroscopy of Alkali Metal Nuclei in Solution

**Frank G Riddell**, The University of St Andrews, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

The alkali metals, lithium, sodium, potassium, rubidium and caesium all possess NMR active nuclei, all of which are quadrupolar.

- Lithium has two NMR active isotopes $^{6}Li$ (7.4%) and $^{7}Li$ (92.6%), of which $^{7}Li$ is the isotope of choice due to its higher magnetogyric ratio and natural abundance. Both isotopes are available in isotopically enriched form making NMR tracer studies relatively easy.
- Sodium has only one NMR active nucleus, $^{23}Na$ (100%).
- Potassium has two NMR active isotopes $^{39}K$ (93.1%) and $^{41}K$ (6.9%), of which $^{39}K$ is the isotope of choice due to its much greater natural abundance and $^{41}K$ is observable only with the greatest difficulty.
- Rubidium has two NMR active isotopes $^{85}Rb$ (72.15%) and $^{87}Rb$ (27.85%), of which $^{87}Rb$ is the isotope of choice due to its much higher magnetogyric ratio despite its lower natural abundance.
- Caesium has only one NMR active nucleus, $^{133}Cs$ (100%).

Lithium is important as the treatment of choice for manic depressive psychosis and this has provoked a wide variety of NMR studies in an endeavour to probe its mode of action. Organolithium compounds are used extensively in synthetic organic chemistry and as industrial catalysts, especially in polymerization reactions.

Both sodium and potassium are essential for life. Potassium is the major intracellular cation in most living cells, with sodium having the second highest concentration. These concentrations are generally reversed in the extracellular fluids. The concentration differences across the cellular membrane are maintained by ion pumps, the most important of which is Na/K/ATPase. This enzyme pumps three sodium ions out of the cell and two potassium ions in for the consumption of one molecule of ATP. This enzyme consumes about one-third of the ATP produced in the human body, emphasizing the importance for life of maintaining the concentration gradients of these ions. In addition, large numbers of enzymes require the presence of sodium or potassium for them to function by mechanisms such as symport or antiport. The human need for sodium chloride as a part of the diet is recognized in many proverbs and sayings in common use, and in the word 'salary' which is a reminder that salt has in the past been used as a form of payment.

Although the chemistry of rubidium is close to that of potassium it cannot be used as a substitute for potassium in biological systems *in vivo*, although it has been used in studies of perfused organs and cellular systems. The same applies for similar reasons to caesium. These metals can be taken into biological systems where they generally replace potassium, but the ingestion of large amounts of the salts of either metal has severe physiological consequences leading in extreme cases to death.

Many reasons exist, therefore, for the development and implementation of NMR methods for the study of the alkali metals.

## Nuclear properties

The nuclear properties of the NMR active isotopes of the alkali metals are presented in **Table 1**.

**Table 1** Nuclear properties of the alkali metals

| *Isotope* | *Spin, I* | *Natural abundance (%)* | *Magnetogyric ratio,* $\gamma/10^7$ (rad $T^{-1}$ $s^{-1}$) | *Quadrupole moment* $Q/10^{-28}$($m^2$) | *NMR frequency,* $\Xi$ (MHz) | *Relative receptivity,* $D^c$ |
|---|---|---|---|---|---|---|
| $^6$Li | 1 | 7.42 | 3.937 | $-8 \times 10^{-4}$ | 14.716 | 3.58 |
| $^7$Li | 3/2 | 92.58 | 10.396 | $-4.5 \times 10^{-2}$ | 38.864 | $1.54 \times 10^3$ |
| $^{23}$Na | 3/2 | 100 | 7.076 | 0.12 | 26.451 | $5.25 \times 10^2$ |
| $^{39}$K | 3/2 | 93.1 | 1.248 | $5.5 \times 10^{-2}$ | 4.666 | 2.69 |
| $^{41}$K | 3/2 | 6.88 | 0.685 | $6.7 \times 10^{-2}$ | 2.561 | $3.28 \times 10^{-2}$ |
| $^{85}$Rb | 5/2 | 72.15 | 2.583 | 0.247 | 9.655 | 43.0 |
| $^{87}$Rb | 3/2 | 27.85 | 8.753 | 0.12 | 32.721 | $2.77 \times 10^2$ |
| $^{133}$Cs | 7/2 | 100 | 3.509 | $-3 \times 10^{-3}$ | 13.117 | $2.69 \times 10^2$ |

$\Xi$ is the observing frequency in a magnetic field in which $^1$H is at 100 MHz.
$D^c$ is the receptivity relative to $^{13}$C.
Quadrupole moments $Q$ are the least well determined parameters in this Table.
Data taken from: *NMR and the Periodic Table* (1978) Harris RK and Mann BE (eds) London: Academic Press.

## Quadrupolar relaxation and visibility

The NMR spectra of the alkali metals are dominated by the fact that all the isotopes are quadrupolar. Effective use of alkali metal NMR requires an understanding of the resulting quadrupolar interactions and the best ways to make use of them and to avoid their pitfalls. Many of the problems that arise and solutions adopted are similar to those involved with the halogens. Quadrupolar nuclei have an asymmetric distribution of charge which gives rise to an electric quadrupole moment. Apart from when the nucleus is in an environment with cubic or higher symmetry, the quadrupole moment interacts with the electric field gradient (EFG) experienced by the nucleus, giving rise among other things to quadrupolar relaxation. The strength of the quadrupolar interaction between the quadrupole moment ($eQ$) and the electric field gradient ($eq$) is given by the quadrupolar coupling $e^2qQ/h$. This can take from very small values to hundreds of MHz, depending on the magnitudes of $Q$ and $q$.

In solution, modulation of the EFG at the quadrupolar nucleus by isotropic and sufficiently rapid molecular motions (where $\omega\tau \ll 1$) leads to relaxation according to the expression:

$$\frac{1}{T_1} = \left(\frac{3}{40}\right)\left[(2I+3)/I^2(2I-1)\right] \times \left(1+\eta^2/3\right)\left(e^2qQ/\hbar\right)^2\tau_c$$

where $\eta$ is the asymmetry parameter associated with the EFG.

The alkali metal ions in solution are subject to relatively low quadrupolar interactions. This is particularly true for $^6$Li and $^7$Li and for $^{133}$Cs, which have inherently low quadrupole moments. Indeed $^6$Li and $^{133}$Cs have the two lowest known quadrupole moments and $^6$Li is often referred to as an 'honorary' spin $\frac{1}{2}$ nucleus. In aqueous solution the ions are solvated by charge dipole interactions with the water molecules. At any one instant the pattern of water molecules around the cation does not have spherical symmetry but is always close to it. Thus the quadrupolar couplings are low but are never zero. Typically, in aqueous solution and in the absence of extraneous influences, both isotopes of lithium show $Li^+$ line widths of < 1 Hz, $Na^+$ and $K^+$ show a line width of *ca.* 12 Hz, both rubidium isotopes show line widths of *ca.* 140–150 Hz and $^{133}$Cs shows a line width of *ca.* 1 Hz.

Because of the low values of the quadrupole moments for both isotopes of lithium, dipolar relaxation becomes important. In aqueous solution at ambient temperature dipolar relaxation accounts for over 75% of the relaxation of $^7$Li ($T_1 \sim 20$ s) and almost 100% of $^6$Li ($T_1 \sim 170$ s). In $D_2O$ solution with no $^1$H available for dipolar relaxation and virtually no quadrupolar mechanism available, the relaxation time of $^6$Li becomes very long ($T_1 \sim 830$ s). In contrast, it appears that despite the low quadrupole moment of $^{133}$Cs, dipolar relaxation does not contribute significantly.

In cases where molecular motion is restricted ($\omega\tau$ is not $\ll 1$) the situation is more complex. Such cases arise when the alkali metal is bound to the surface of a large molecule such as a protein or membrane surface and thereby has its motion restricted. The quadrupolar interaction with the nucleus shifts the energies of the Zeeman levels according to the square of the quantum number to a first approximation. Thus, the energy level splittings for a nucleus with $I = \frac{3}{2}$ (e.g. $^7$Li, $^{23}$Na, $^{39}$K and $^{87}$Rb) become as illustrated in **Figure 1**. With rapid isotropic motion, as described above, the multiple line pattern will

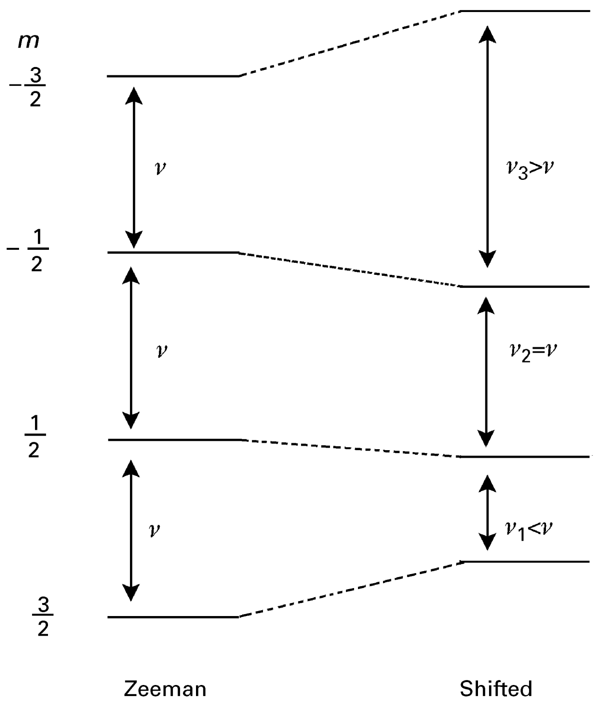



**Figure 1** Changes in the energy levels for a spin $\frac{3}{2}$ nucleus subject to a quadrupolar interaction. Note: The shifts in energy levels shown are exaggerated for clarity

collapse into a single line. In the absence of rapid isotropic motion the relaxation rate of the outer transitions, which combine to have 60% of the total intensity, is different from that of the inner transitions, which has 40% of the total intensity. The frequencies of the outer transitions also shift from the inner transition (dynamic frequency shift). There are three principal consequences of these changes for cases where the motion of the cation is restricted. First, line shapes become a double and not a single Lorentzian; secondly, two relaxation times are apparent; and finally, where the more rapid relaxation time becomes very short, partial or total invisibility of the signal from the outer transitions may occur. An excellent review of quadrupolar relaxation effects is given in the review by Springer given in the Further reading section.

## Biological applications

### Contrast reagents

One of the main problems in using NMR to study the alkali metals in biological systems is that the chemical shifts of the aqueous ions are essentially independent of the ion's surroundings in all cases except for $^{133}Cs^+$, making differentiation of intra- and extracellular cations difficult. This problem can be met by using a contrast reagent (either a shift or a relaxation agent) in one of the compartments, normally extracellular. A large number of aqueous shift reagents have been employed. They all work on the same principle, that a paramagnetic lanthanide, typically dysprosium, is enclosed in a complex by a ligand or ligands, and the resulting complex has several negative charges. With the overall charge on the complex being negative, the alkali metal cations are attracted to the negatively charged species and thereby brought into a region where the paramagnetic interaction induces a chemical shift change. Typical shift reagents for the alkali metals are given in **Table 2**. The resonances of the cations are also broadened by the process, but this broadening has been shown to be largely due the quadrupolar interaction of the cation with the reagent and not due to paramagnetic relaxation. For rubidium, which has a substantial line width that is comparable to the shifts capable of being induced by the best shift reagents, it is preferable to employ a relaxation agent to relax the signal from the extracellular $Rb^+$ into the baseline noise.

The first important application of alkali metal shift reagents was the use of dysprosium bistripolyphosphate $(DyP_3O_{10})_2^{7-}$ ($DyPPP_2$) to differentiate between intra- and extracellular $^{23}Na$ in human erythrocytes. This was soon followed by a similar experiment revealing the intracellular signal from $^{39}K$. In both cases it seems as if the intracellular metal ions in human erytrocytes are essentially 100% visible. The spectra obtained for the $^{39}K$ experiment are shown in **Figure 2**.

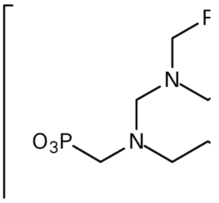
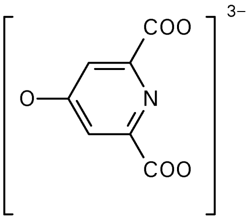
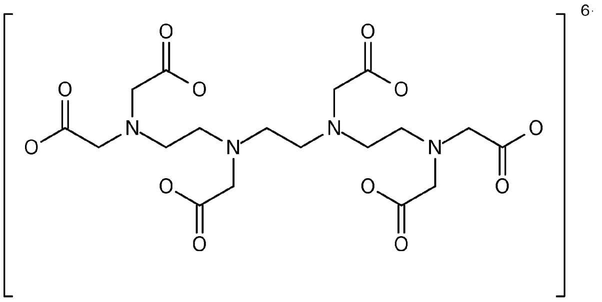
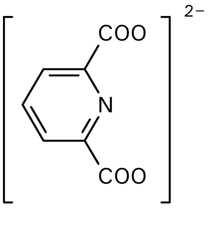
The maximum shift generated by the shift reagents varies with the alkali metal according to the number of shells of electrons shielding the nucleus from the paramagnetic centre. For example the maximum shifts available from $DyPPP_2$ are approximately: $^{7}Li^+$, 40 ppm: $^{23}Na^+$, 20 ppm; $^{39}K^+$, 10 ppm; $^{87}Rb^+$, 4 ppm; $^{133}Cs^+$, 2 ppm.

The shift reagent $DyPPP_2$ is commonly used for *in vitro* systems such as vesicles or with isolated erythrocytes. However, it displays considerable toxicity for *in vivo* systems, in which cases the shift reagent $TmDOTP^{5-}$ (see **Table 2**) is preferred. For example during *in vivo* studies of rat kidneys using $TmDOTP^{5-}$, three $^{23}Na^+$ signals were resolved, corresponding to intracellular $Na^+$, vascular $Na^+$ and intraluminal $Na^+$.

### Multiple quantum filtration

It has been shown that $^{23}Na$ double-quantum-filtered NMR spectroscopy can be used to detect anisotropic motion of bound sodium ions in biological systems. The technique is based on the formation of

**Table 2** Shift reagents for alkali metal cations

| *Shift reagent* | *Acid anion* |
|---|---|
| $[Dy(PPP)_2]^{7-}$ | Tripolyphosphate $(P_3O_{10})^{5-}$ |
| $[Dy(DPA)_3]^{3-}$ | Dipicolinate |
| |  |
| $[Dy(NTA)_3]^{3-}$ | Nitrilotriacetate $[N(CH_2COO)_3]^{3-}$ |
| $[Dy(CA)_3]^{6-}$ | Chelidamate |
| |  |
| $[Dy(THHA)]^{3-}$ | Triethylenetetraminehexaacetate |
| |  |
| $TmDOTP^{5-}$ | Thulium 1,4,7,10-tetraazacyclodecane-1,4,7,10-tetrakismethylenephosphonate |
| |  |

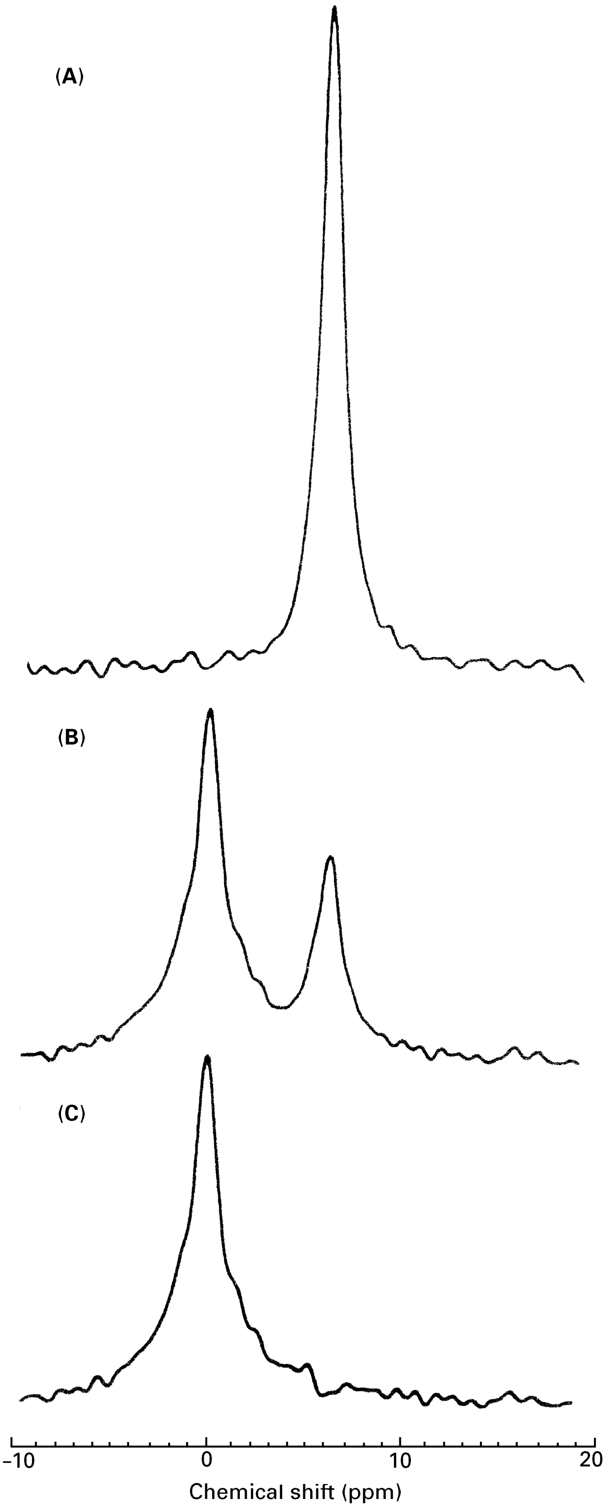



**Figure 2** $^{39}K$ NMR spectra recorded at 16.8 MHz of (A) resuspension medium containing 60 mM $K^+$, 6 mM $Dy^{3+}$ and 15 mM tripolyphosphate: (B) human erythrocytes in the same medium; and (C) difference spectrum after the subtration of 0.3 of the intensity of spectrum (A) from spectrum (B). For each spectrum 20 000 free induction decays were collected in approx. 20 min. Reproduced with permission of the Biochemical Society from Brophy PJ, Hayer MK and Riddell FG (1983) *Biochemical Journal* **210**: 961.

the second-rank tensor when the quadrupolar interaction is not averaged to zero. Isotropically tumbling $^{23}Na$, free in aqueous solution, is not seen by these methods. Such techniques allow, for example, the detection of $^{23}Na^+$ bound to macromolecules such as proteins or membranes or the detection of intracellular $^{23}Na^+$ if its motion is partially restricted. Triple-quantum-filtered spectra can also be used for similar purposes.

Multiple-quantum-filtered NMR offers the possibility of monitoring the intracellular $Na^+$ content in the absence of shift reagents provided that three criteria are met: (1) the contribution from intracellular $^{23}Na^+$ to the multiple-quantum-filtered spectrum is substantial, (2) that it responds to a change in intracellular $^{23}Na^+$ content and (3) that the amplitude of the extracellular multiple-quantum-filtered component remains constant during a change in intracellular $^{23}Na^+$ content.

### Lithium NMR

The use of $Li^+$ salts as the preferred treatment for manic depressive psychosis has spurred on the use of $^7Li$ NMR in biological systems, particularly work on cellular systems. The use of the shift reagent $DyPPP_2$ for $^{23}Na$ and $^{39}K$ to separate intra- and extracellular signals in human erythrocytes was rapidly followed by similar experiments with $^7Li^+$. The object of these experiments was to determine lithium transport rates across the erythrocyte membrane as a model for the blood–brain barrier. Comparisons were made of the transport rates of $^7Li^+$ into and out of the erythrocytes of manic depressive patients being treated with $Li^+$, with those of normal controls. These experiments have demonstrated that at extracellular concentrations ranging from 50 to 2 mM the efflux rate from the erythrocytes of the patients was significantly slower than for those of the controls. Moreover, the experiments have shown that the abnormal transport rate is a consequence of $Li^+$ treatment and is not a marker for the illness. Similar experiments have been carried out with other cellular systems including astrocytomas, neuroblastomas, rat hepatocytes and cultured Swiss Mouse 3T3 fibroblasts. The work on astrocytomas, an immortalized cell line from a human brain cancer, allowed visualization of $Li^+$ inside the cells which were supported on microcarrier beads (**Figure 3**) and showed that there is an active $Li^+$ extrusion pump present in these cells and, therefore, that there must also be a $Li^+$ pump present in astrocytomas in the brain.

It is widely believed that the enzyme interacting with $Li^+$ when it acts to control manic depressive psychosis is inositol monophosphatase. $^7Li^+$ NMR signals from $Li^+$ bound to the inositol monophosphatase enzyme have been observed. This suggests that $^7Li^+$ NMR could make an important contribution to the study of this and other lithium-sensitive enzymes.

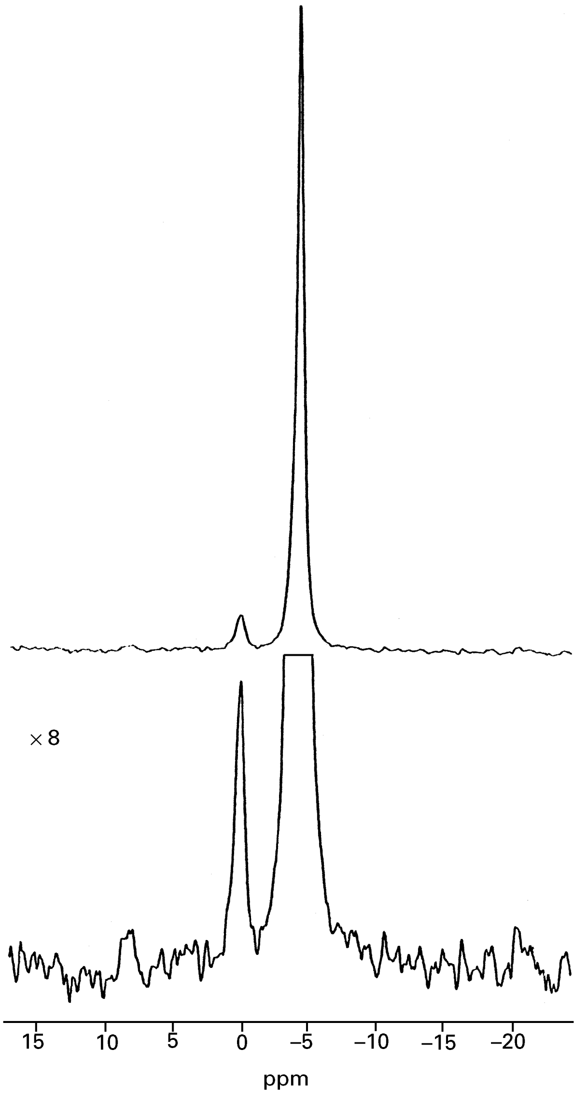



**Figure 3** $^7Li$ NMR spectra of astrocytomas on microcarrier beads in a buffer containing dysprosium tripolyphosphate shift reagent. Each spectrum is the sum of 48 acquisitions recorded at 194 MHz, $[Li^+]_{out}$ = 10 mM. Reprinted from Bramham J, Carter AN and Riddell FG *Journal of Inorganic Biochemistry* **61**: 273–284, copyright 1996, with permission from Elsevier Science.

NMR imaging techniques have been applied to study the concentrations and the pharmacokinetics of $Li^+$ in various parts of the human body. Such experiments have shown that the concentration of $Li^+$ in the brain and muscle is lower than that in the blood serum. Interestingly, they have also shown that after ingestion of $Li^+$ there is lag in the uptake of $Li^+$ into the brain. Maximum concentrations of $Li^+$ in the brain occur several hours after the concentration maximum in the serum has been reached. Such

experiments point towards better treatment regimes for patients.

### Sodium, potassium, rubidium and caesium NMR

A very substantial body of literature exists on the use of $^{23}Na$ and $^{39}K$ NMR to study cardiac function. Such experiments are typically performed on isolated perfused rat hearts, although hearts from other species including guinea pigs and dogs have also been used. Often the experiments involve shift reagents, although more recently multiple quantum filtration has been extensively used to differentiate between pools of ions. During ischaemia (low or zero blood flow which mimics a heart attack) there is an accumulation of intracellular sodium, cellular swelling and energy deficiency that participate in the transition to irreversible ischaemic injury. Such changes can be followed readily by a combination of $^{23}Na$ and $^{31}P$ NMR techniques. These experiments have provided valuable information on the behaviour of hearts under conditions of ischaemia and their recovery afterwards during reperfusion. They provide valuable information on methods for the resuscitation of ischaemic hearts and their protection against hypoxic injury. Although not present in normal biological systems in other than trace amounts, $Rb^+$ and $Cs^+$ have been used on several occasions as $K^+$ analogues in the above experiments, thus extending the range of nuclei available for study. Other similar experiments have been performed on hearts from genetically hypertensive rats. Similar experiments have been performed on other organs such as kidney and liver from small animals.

Studies of $^{23}Na^+$ in cellular systems have been performed on cells such as superfused isolated rat cardiomyocytes, *Methanobacterium thermoautotrophicum*, porcine vascular endothelial cells, the halotolerant bacterium *Brevibacterium* sp., *Escherichia coli*, murine TM3 Leydig and TM4 Sertoli cell lines, and mouse 3T3 fibroblasts. These experiments have been employed to determine the NMR visibility of $^{23}Na^+$, its intracellular concentration, membrane transport properties and dynamics and its ionic mobility inside the cells. $^{23}Na$ NMR studies have contributed to the study of Na/K/ATPase.

Since the principal cytoplasmic cation is $K^+$, NMR experiments on $^{39}K^+$ in cellular systems should give valuable information on the intracellular environment. That they have been used much less frequently than experiments on $^{23}Na^+$ is because of the lower receptivity of $^{39}K$. The utility of $^{39}K^+$ studies is shown by work on $^{39}K^+$ from *E. coli* after plasmolysis. The $^{39}K^+$ signals are 100% visible and show biexponential relaxation, with both components relaxing very rapidly. The result was attributed to a substantial interaction between the $^{39}K^+$ and the polynegatively charged surface of the ribosomes.

The uptake of $Rb^+$ into human erythrocytes has been studied by $^{87}Rb$ NMR using the relaxation agent $LaPPP_2$ to contrast the two pools of $Rb^+$. Uptake was linear over a 24 h period.

With $^{113}Cs^+$ NMR there is no need for a contrast reagent to separate the intra- and extracellular signals since the chemical shifts of the intra- and extracellular signals are well separated. Variations of the phosphate concentration in the suspension buffer are sufficient for this purpose. Uptake of $^{113}Cs^+$ into human erythrocytes was observed to be linear with a rate of 0.33 mM $h^{-1}$ at an extracellular $Cs^+$ concentration of 10 mM. When the cells were removed to a $Cs^+$-free buffer they retained the $Cs^+$, indicating that there is no transport mechanism available for the removal of $Cs^+$ from the cells. $Cs^+$ was shown to replace $K^+$ inside the cell. The favourable properties of $^{113}Cs^+$ as indicated above, primarily its chemical shift range without the use of shift reagents and its low quadrupolar interactions, have led to its use as an analogue of $K^+$ in several studies of its tissue compartmentation.

### Mediated membrane transport

A variety of NMR methods exists for the study of the mediated transport of alkali metal ions through model biological membranes. Substrates that mediate the transport include the ionophoric antibiotics such as monensin [1], channel forming peptides such as gramicidin and the peptaibols, other channel forming substrates such as amphotericin and the brevitoxins, or synthetic carriers, for example, those based around crown ether-like skeletons such as [2]. For such experiments large unilamellar vesicles (LUV) formed from phospholipid are prepared and a chemical shift difference between the intra- and extravesicular compartments is established by means of a shift reagent. For rapid exchange of ions across the membrane ($k > 10$ $s^{-1}$) dynamic line broadening provides information on the transport kinetics (**Figure 4**). For cases where the transport rate is comparable to the relaxation rate a magnetization transfer technique can be employed. In this experiment one of the two signals, normally the extracellular signal, is inverted by a simple pulse sequence. Chemical exchange then causes a time-dependent reduction in the signal of the other resonance. Analysis of signal intensities against time gives the transport kinetics. For relatively slow exchange ($k < 10^{-3}$ $s^{-1}$) isotope exchange is used, e.g. $^{6}Li/^{7}Li$ or $^{7}Li/^{23}Na$. In such experiments the concentration gradients of the cations form the driving force for the transport. Such experiments have provided extremely valuable

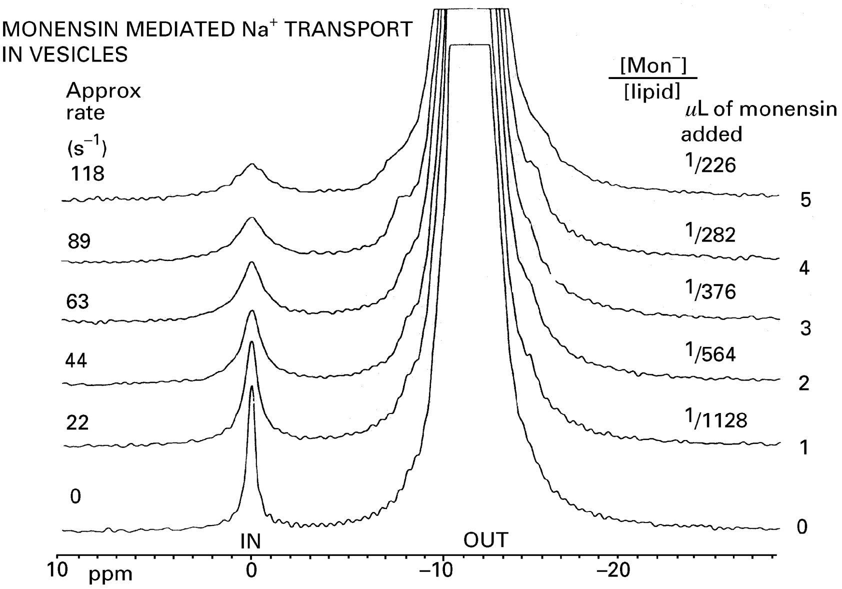



**Figure 4** Changes in the $^{23}$Na NMR spectra recorded at 21.16 MHz of LUV containing 120 mM NaCl on addition of increasing microlitre amounts of a dilute solution of monensin in methanol at 303 K. The surrounding solution contains 10 mM $Na_5P_3O_{10}$, 70 mM NaCl and 4.0 mM $DyCl_3$. Reprinted from Riddell FG and Hayer MK *Biochimica Biophysica Acta* **817**: 313–317, Copyright © 1985, with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

insights into the kinetics and mechanism of mediated transport. For the ionophoric antibiotics such as monensin they have shown that one ionophore transports one alkali metal and that the rate limiting step is almost invariably release of the alkali metal ion at the membrane surface. However, several synthetic ionophores, e.g. ([2], $n = 3$, $R_1 = R_2 = C_{10}H_{21}$), which transports $Na^+$ at a rate comparable to that of monensin, exhibit diffusion as the rate limiting step. For gramicidin for example, these experiments have confirmed that two molecules are required to come together to form a pore. For the peptaibols the transport has been shown to occur by 'barrel stave' assembly of peptide molecules across the membrane forming a pore inside the 'barrel' that is of sufficiently long duration to allow complete exchange of the intra- and extravesicular media. For the brevetoxins, selectivity for various ions was probed by changing the ions involved in the concentrations gradients. The dependence on cholesterol incorporation in the membrane was studied.

So-called 'bouquet molecules', based on a central crown ether or cyclodextrin unit equipped with pendant arms that are also capable of complexing cations and are long enough to traverse a lipid bilayer, have been studied in vesicles with a $Na^+$ / $Li^+$ gradient across the membrane using both $^7$Li and $^{23}$Na NMR. Such systems show a one-for-one exchange of $Na^+$ for $Li^+$ (antiport). These molecules were found to transport $Na^+$ at similar rates in fluid- and gel-state membranes; this suggests that ion passage occurs preferentially by a channel mechanism and not by the carrier mechanism. Monensin, known to operate as a carrier, was shown to transport at a slower rate in a gel-state membrane.

Another interesting aspect of these experiments is their ability to probe the effect of changes in the membrane composition on the transport kinetics. Thus, placing positive and negative charges on the membrane surface causes changes in ionophore mediated transport rates, and the introduction of pharmaceuticals such as chlorpromazine and imipramine cause changes in the nigericin mediated $Na^+$ transport rates.

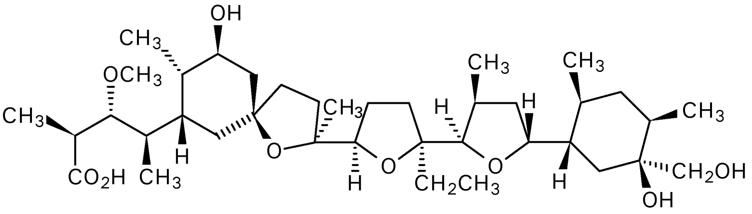
## Chemical applications

### Covalently bound lithium

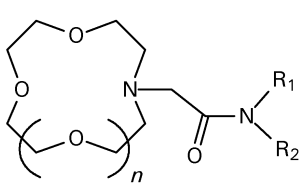
Lithium covalently bound to carbon may be observed by $^7$Li NMR in lithium alkyls. In such molecules $^7$Li has a small chemical shift range (~12 ppm). Tables of chemical shifts and coupling constants are to be found in the review by Günther.

### Metal ion complexation studies

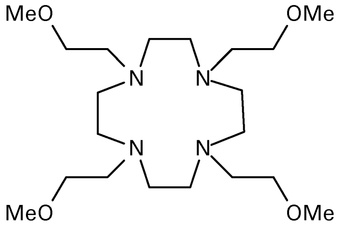
Although biological applications have been the major use of alkali metal ion NMR, it has also proved to be valuable for the study of complexation of the alkali metals in host–guest systems by suitably designed ligands, e.g. the crown ethers and cryptands. Two parameters are important in detecting complexation: chemical shift changes and decreases



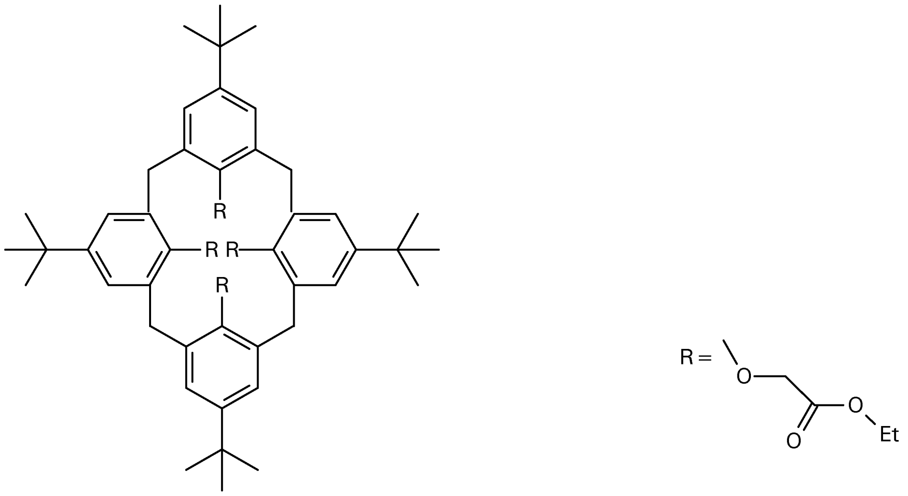
[1]



[2]



[3]



[4]

in relaxation times as a result of enhanced quadrupolar interactions.

Often dynamically broadened alkali metal NMR spectra can be seen as a result of exchange between the free and complexed cation. A good example is provided by the dynamic $^{7}Li$ and $^{23}Na$ spectra for the interaction of $Li^{+}$ and $Na^{+}$ with the pendant arm macrocycle 1,4,7,10-tetrakis(2-methoxyethyl)-1,4,7,10-tetraazacyclododecane [3]. The dynamic $^{7}Li$ spectra are shown in **Figure 5**. Evidence of a slowly exchanging 1:1 complex and of a 2:1 complex in rapid equilibrium with the 1:1 complex between calixarene [4] and $Na^{+}$ is provided by studies of this system by $^{23}Na$ and $^{1}H$ NMR.

Frequently, when the alkali metal ion is exchanging between the complex and the solution, the temperature variation of $T_1$ and/or $T_2$ for the metal can give information about the exchange kinetics. The complexed ion has a much shorter $T_1$ (and $T_2$) value due to strong quadrupolar interactions with the ligand. In the slow exchange limit the observed $T_1$ value approaches that of the ion free in solution, whilst in the rapid exchange limit the $T_1$ value is an average of the values for the complexed and free

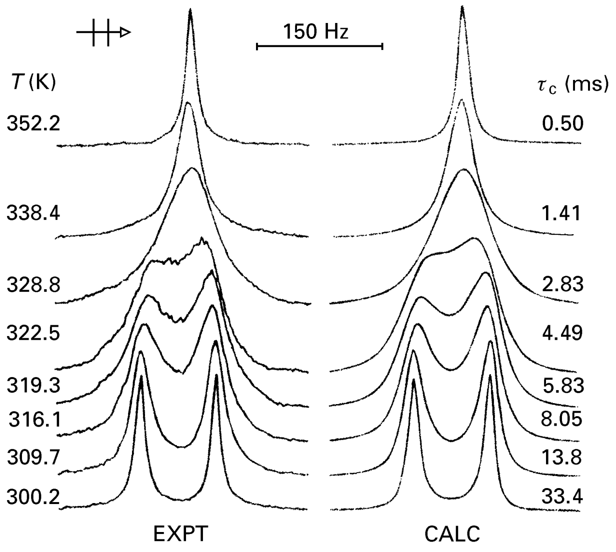



**Figure 5** Typical exchange-modified $^7$Li NMR spectra recorded at 116.59 MHz of a dimethylsulfoxide solution of solvated $Li^+$ and [3]. Experimental temperatures and spectra appear on the left of the figure and the best fit calculated line shape and lifetime values on the right. Complexed $^7Li^+$ appears as the left-hand side, high frequency, peak. Reprinted with permission from Stephens AKW, Dhillon RS, Madbak SE, Whitbread SL and Lincoln SF *Inorganic Chemistry* **35**: 2019–2024, copyright 1996, American Chemical Society.

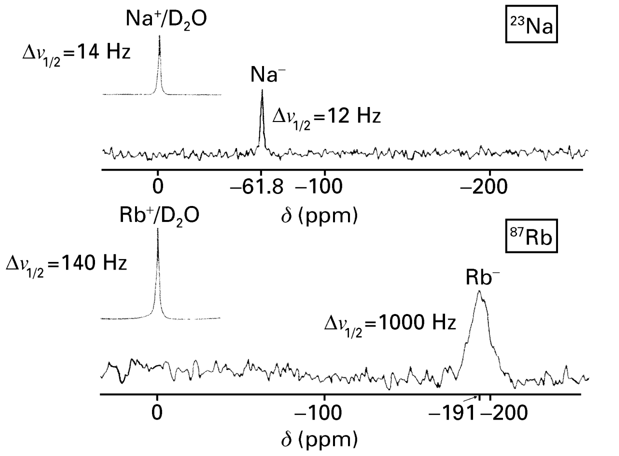



**Figure 6** $^{23}$Na and $^{87}$Rb NMR spectra of solutions of sodium and rubidium in 1,4,7,10-tetraoxacyclododecane (12-crown-4). Negative chemical shift values correspond to a decrease in resonance frequency and an increase in nuclear shielding. Reproduced with permission of The Royal Society of Chemistry from Holton DM, Edwards PP, Johnson DC, Page CJ, McFarlane W and Wood B (1984) *Journal of the Chemical Society, Chemical Communications*, 740–741.

ions. In between these extremes $T_1$ follows a sigmoid curve when plotted against $1/T$.

### Alkali metal anions in solution: alkalide ions

Under conditions of the most rigorous purity and using high vacuum techniques in dipolar aprotic and similar solvents, the alkali metals yield metal anions ($M^-$) known as the alkalides. Thus sodium in hexamethylphosphoric triamide solutions gives rise to sodide ($Na^-$) ions. Sodium and rubidium in 1,4,7,10-tetraoxacyclododecane (12-crown-4) give sodide and rubidide ions (**Figure 6**). These anions are most readily identified by their NMR spectra which occur substantially to low frequency of the chemical shift standards of the alkali metal salts in $D_2O$ solution. The alkalide ions are formed by the addition of one electron to the partially filled outermost S orbital. This is expected to lead to a substantial shielding increase, the observation of which is a good indication of the ion formation. Chemical shifts of the alkalide anions vary slightly with solvent and temperature but are near the following values: $Na^-$, −62 ppm, $K^-$, −103 ppm, $Rb^-$, −191 ppm, $Cs^-$, −280 to −300 ppm.

The spectra of the sodide ion and the potasside ion at low temperatures show relatively sharp line widths, indicative of low ion–solvent interactions suggesting that these ions in solution resemble those in the gas state. On the other hand, at room temperature the line width of the rubidide ion (~1000 Hz vs. 140 Hz for $Rb^+$ in $H_2O$) indicates that there is quadrupolar broadening and the observed shift falls short of that calculated for a gaseous-like ion, unlike the shifts of the sodide and potasside ions. This strongly suggests that there are interactions between the solvent and the rubidide ion. In the case of caesium dissolved in crown ethers the species $Cs^+e^-$ has also been observed.

## List of symbols

$eq$ = electric field gradient strength; $e^2qQ/h$ = quadrupolar coupling; $eQ$ = quadrupole moment strength; $I$ = spin quantum number; $T_1$ = longitudinal relaxation time; $T_2$ = transverse relaxation time; $\tau$ = correlation time for molecular motion (s); $\omega$ = Larmor frequency (rad $s^{-1}$).

*See also:* **Biofluids Studied By NMR; Cells Studied By NMR; Halogen NMR Spectroscopy (excluding $^{19}$F); *In Vivo* NMR, Applications, $^{31}$P; *In Vivo* NMR, Applications, Other Nuclei; Membranes Studied By NMR Spectroscopy; NMR Relaxation Rates; Perfused Organs Studied Using NMR Spectroscopy; Relaxometers.**

## Further reading

Bramham J and Riddell FG (1994) Cesium uptake studies on human erythrocytes, *Journal of Inorganic Biochemistry* 53: 169–176.

Brophy PJ, Hayer MK and Riddell FG (1983) Measurement of intracellular potassium ion concentration by NMR, *Biochemical Journal* **210**: 961–963.

Edwards PP, Ellaboudy AS and Holton DM (1985) NMR spectrum of the potassium anion $K^-$, *Nature (London)* **317**: 242–244.

Edwards PP, Ellaboudy AS, Holton DM and Pyper NC (1988) NMR studies of alkali anions in non-aqueous solvents. *Annual Reports on NMR Spectroscopy* **20**: 315–366.

Günther H (1996) Lithium NMR, In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, p. 2807, Chichester: Wiley.

Laszlo P (1996) Sodium-23 NMR, *In*: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, p. 4551. Chichester: Wiley.

Lindman B and Forsén S (1978) The Alkali Metals, *In*: Harris RK and Mann BE (eds) *NMR and the Periodic Table*, London: Academic Press.

Mota de Freitas D (1993) Alkali metal nuclear magnetic resonance, *Methods in Enzymology* **227**: 78–106.

Riddell FG (1998) Studying biological lithium using nuclear magnetic resonance techniques, *Journal of Trace and Microprobe Techniques* **16**: 99–110.

Sherry AD and Geraldes CFGC (1989) Shift reagents in NMR spectroscopy in lanthanide probes, *In*: Bünzli J-CG and Chopin GR (eds) *Life, Chemical and Earth Sciences, Theory and Practice*, Amsterdam: Elsevier.

Springer CS (1996) Biological systems, spin-3/2 nuclei. *In*: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, p. 940. Chichester: Wiley.

# NMR Spectroscopy, Applications

*See* **Diffusion Studied Using NMR Spectroscopy; Drug Metabolism Studied Using NMR Spectroscopy; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals; Biofluids Studied By NMR; Carbo-hydrates Studied By NMR; Cells Studied By NMR; Structural Chemistry Using NMR Spectroscopy, Peptides; Proteins Studied Using NMR Spectroscopy; Nucleic Acids Studied Using NMR; Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules; Structural Chemistry Using NMR Spectroscopy, Organic Molecules.**

# NOE

*See* **Nuclear Overhauser Effect.**

# Nonlinear Optical Properties

**Georges H Wagnière**, University of Zurich, Switzerland
**Stanisław Woźniak**, A. Mickiewicz University, Poland

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Theory**

Nonlinear optical phenomena manifest themselves as special forms of light scattering and refraction. Ordinary, linear, light scattering may be viewed as a quasi-simultaneous absorption and re-emission of a photon of same frequency $\omega$. The event occurs on a very short timescale; for radiation in the UV-visible region, of the order of $10^{-15}$ s, or 1 fs. Light scattering is not to be confounded with resonant absorption and emission. Atoms and molecules absorb and emit light at particular, selected frequencies; they scatter light within the whole spectrum of electromagnetic radiation. Resonant absorption and emission are connected with a change of energy state of the atom or molecule; light scattering is an elastic process, leaving the atom or molecule in the same state after as before the event. Under conditions that will be specified in more detail below, forms of elastic scattering may also occur in which more than one incident photon are involved. For instance, it may happen that two photons of frequency $\omega$ incident on a molecule merge to form a single re-emitted photon of frequency $2\omega$. This process is called second-harmonic generation and is a nonlinear optical phenomenon.

In the following, the most important nonlinear optical effects are reviewed. The emphasis is laid on the material systems in which they have been observed. Nonlinear optical properties of both organic and inorganic materials are presented in tabular form and commented on.

## Basic quantities

The semiclassical description of light scattering attributes the effect to a light-induced (electric) dipole moment $\boldsymbol{p}^{(1)}$ in the molecule that oscillates with the frequency of the incident radiation $\omega$ and becomes the source of quasi-immediate re-emission at the same frequency. Mathematically, this is expressed as

$$\boldsymbol{p}^{(1)}(\omega) = \alpha^{(1)}(-\omega;\omega)\cdot\boldsymbol{E}(\omega) \qquad [1a]$$

$\boldsymbol{E}(\omega)$ is the electric field strength of the incident light at the frequency $\omega$, and $\alpha^{(1)}(-\omega;\omega)$ the molecular *frequency-dependent polarizability* of first order. The descriptive parenthesis $(-\omega;\omega)$ indicates incidence at frequency $\omega$ and quasi-simultaneous re-emission at frequency $\omega$. For a macroscopic sample we have the corresponding equation:

$$\boldsymbol{P}^{(1)}(\omega) = \varepsilon_0\,\chi^{(1)}(-\omega;\omega)\cdot\boldsymbol{E}(\omega) \qquad [1b]$$

where $\boldsymbol{P}^{(1)}(\omega)$ is the volume polarization and $\chi^{(1)}(-\omega;\omega)$ the macroscopic susceptibility. $\chi$ is the appropriate sum of the molecular contributions $\alpha$.

If two separate radiation frequencies, $\omega_1$ and $\omega_2$, act on a molecule at the same time, we will mainly have separate and distinct scattering at these two basic frequencies. However, if the coherence and intensity of the incident radiation are sufficiently high, we may observe the generation of 'overtones' of frequency $\omega_1+\omega_2$:

$$\boldsymbol{p}^{(2)}(\omega_1+\omega_2) = \alpha^{(2)}(-\omega_1-\omega_2;\omega_1,\omega_2):{}^1\boldsymbol{E}(\omega_1)\;{}^2\boldsymbol{E}(\omega_2) \qquad [2]$$

We notice that the induced polarization $\boldsymbol{p}^{(2)}$ for this second-order effect is proportional to the product of the field strengths ${}^1\boldsymbol{E}(\omega_1)$ and ${}^2\boldsymbol{E}(\omega_2)$. The parenthesis $(-\omega_1-\omega_2;\omega_1,\omega_2)$ indicates incidence at frequencies $\omega_1$ and $\omega_2$ and scattering at frequency $\omega_1+\omega_2$.

For $\omega_1=\omega_2$, sum frequency generation becomes tantamount to second-harmonic generation (SHG), i.e. incidence at frequency $\omega$ and quasi-simultaneous re-emission at the doubled frequency $2\omega$:

$$\boldsymbol{p}^{(2)}(2\omega) = \alpha^{(2)}(-2\omega;\omega,\omega):\boldsymbol{E}^2(\omega) \qquad [3]$$

We encounter a situation for which the induced polarization is no longer linearly but now quadratically dependent on the field strength of the incident radiation. Correspondingly, the intensity of the radiation generated at frequency $2\omega$ is also quadratically dependent on the intensity $I(\omega)$. In contrast, the intensity of 'ordinary' light scattering at frequency $\omega$ is linearly dependent on $I(\omega)$.

As well as sum frequency generation, the nonlinear optical effect of difference frequency generation is also possible, $\omega_1 - \omega_2$ being generated from $\omega_1$ and $\omega_2$. For $\omega_1 = \omega_2$, we then have optical rectification, namely, the induction of a static electric polarization of frequency 0, by the radiation field:

$$\boldsymbol{p}^{(2)}(0) = \alpha^{(2)}(0; \omega, -\omega) :\ {}^1\boldsymbol{E}(\omega)\ {}^2\boldsymbol{E}(-\omega) \quad [4]$$

(Mathematically, $\boldsymbol{E}(-\omega)$ is expressed as the complex conjugate of $\boldsymbol{E}(\omega)$.)

**Tables 1** and **2** and **Figure 1** summarize a number of distinct nonlinear optical effects, in particular including those of third order arising from the combined influence of three frequencies, $\omega_1$, $\omega_2$, $\omega_3$. These effects are collectively called four-wave mixing, as three incident waves combine coherently to give a fourth resulting one of frequency $\omega_1 \pm \omega_2 \pm \omega_3$. The radiation-induced polarization depends to third order on the electric field strength of the incident radiation, namely, on the triple product ${}^1\boldsymbol{E}(\omega_1)\ {}^2\boldsymbol{E}(\pm\omega_2)\ {}^3\boldsymbol{E}(\pm\omega_3)$. In the particular case where $\omega_1 = \omega_2 = \omega_3$, we may have third-harmonic generation, proportional to $E^3(\omega)$. Correspondingly, the intensity of the third harmonic radiation $I(3\omega)$ depends on $I^3(\omega)$.

The variety of nonlinear optical phenomena is indeed vast, but with increasing order their observation becomes more and more difficult. Conventional, thermal light sources do not produce radiation of sufficient coherence and intensity to induce observable nonlinear effects. The field of nonlinear optics has been made accessible by the laser. Laser light is generated by induced emission and can be pictured as consisting of wave trains oscillating in phase. This allows the generation of tightly bundled, sharply focusable beams of high intensity.

The key quantity to understanding the nonlinear optical response of a given molecule is the corresponding generalized polarizability $\alpha^{(n)}(\ldots)$, sometimes also called molecular susceptibility or, for $n > 1$, hyperpolarizability, where $n$ denotes the order of the effect. The first-order polarizability $\alpha^{(1)}$ is a second-rank tensor. The tensor is symmetric and, therefore, in general has six independent elements, assumed to be defined in a symmetry-adapted molecular reference frame $x$, $y$, $z$. These six different tensor elements manifest themselves in point groups belonging to the triclinic symmetry system: $xx(\equiv \alpha^{(1)}_{xx})$, $yy$, $zz$, $xy = yx$, $yz = zy$, $zx = xz$. In the monoclinic system, there occur four independent elements. In higher systems, by symmetry the tensor becomes diagonal. In the cubic system, all three diagonal elements are the same. For an isotropic medium, we obtain a single scalar average.

The second-order polarizabilities $\alpha^{(2)}(\ldots)$ are third-rank tensors. Such tensors in general vanish in centrosymmetric media, as they are parity-odd. In the triclinic symmetry system, there are $3^3 = 27$ independent tensor elements. In a molecule of higher symmetry, some elements become zero, others may

**Table 1** Overall classification of nonlinear optical effects

| *Frequency of incident radiation* | *Frequency of scattered radiation* | *Order of effect (n)* | *Rank of susceptibility tensor* | *Name/description* |
|---|---|---|---|---|
| $\omega$ | $\omega$ | 1 | 2 | Rayleigh scattering, ordinary refraction |
| $\omega_1$, $\omega_2$ | $\omega_1 + \omega_2$<br>$\omega_1 - \omega_2$ | 2 | 3 | Sum-frequency generation<br>Difference-frequency generation |
| $\omega_1$, $\omega_2$, $\omega_3$ | $\omega_1 + \omega_2 + \omega_3$<br>$\omega_1 + \omega_2 - \omega_3$<br>$\omega_1 - \omega_2 + \omega_3$<br>$\omega_1 - \omega_2 - \omega_3$ | 3 | 4 | Four-wave mixing |
| $\omega_1$, $\omega_2$, $\omega_3$, $\omega_4$ | $\omega_1 \pm \omega_2 \pm \omega_3 \pm \omega_4$ | 4 | 5 | Five-wave mixing |

**Table 2** Some important nonlinear optical effects

| *Frequencies of interacting electric fields* | *Effect* | *First experiment* |
|---|---|---|
| $\omega + \omega \to 2\omega$ | Second-harmonic generation (SHG) | a |
| $\omega - \omega \to 0$ | Optical rectification (OR) | b |
| $2\omega - \omega \to \omega$ | Parametric amplification (PA) | c |
| $\omega + \omega + 0 \to 2\omega$ | Electric field-induced second- harmonic generation (EFISH) | d |
| $\omega + \omega + \omega \to 3\omega$ | Third-harmonic generation (THG) | d |
| $\omega + \omega + \omega - \omega \to 2\omega$ | Second-harmonic generation by five-wave mixing | e |

[a] Franken PA, Hill AE, Peters CW and Weinreich G (1961) *Physical Review Letters* **7**: 118.
[b] Bass M, Franken PA, Ward JF and Weinreich G (1962) *Physical Review Letters* **9**: 446.
[c] Giordmaine JA and Miller RC, (1965) *Physical Review Letters* **14**: 973.
[d] Terhune RW, Maker PD and Savage CM (1962) *Physical Review Letters* **8**: 404.
[e] Shkurinov AP, Dubrovskii AV and Koroteev NI (1993) *Physical Review Letters* **70**: 1085.

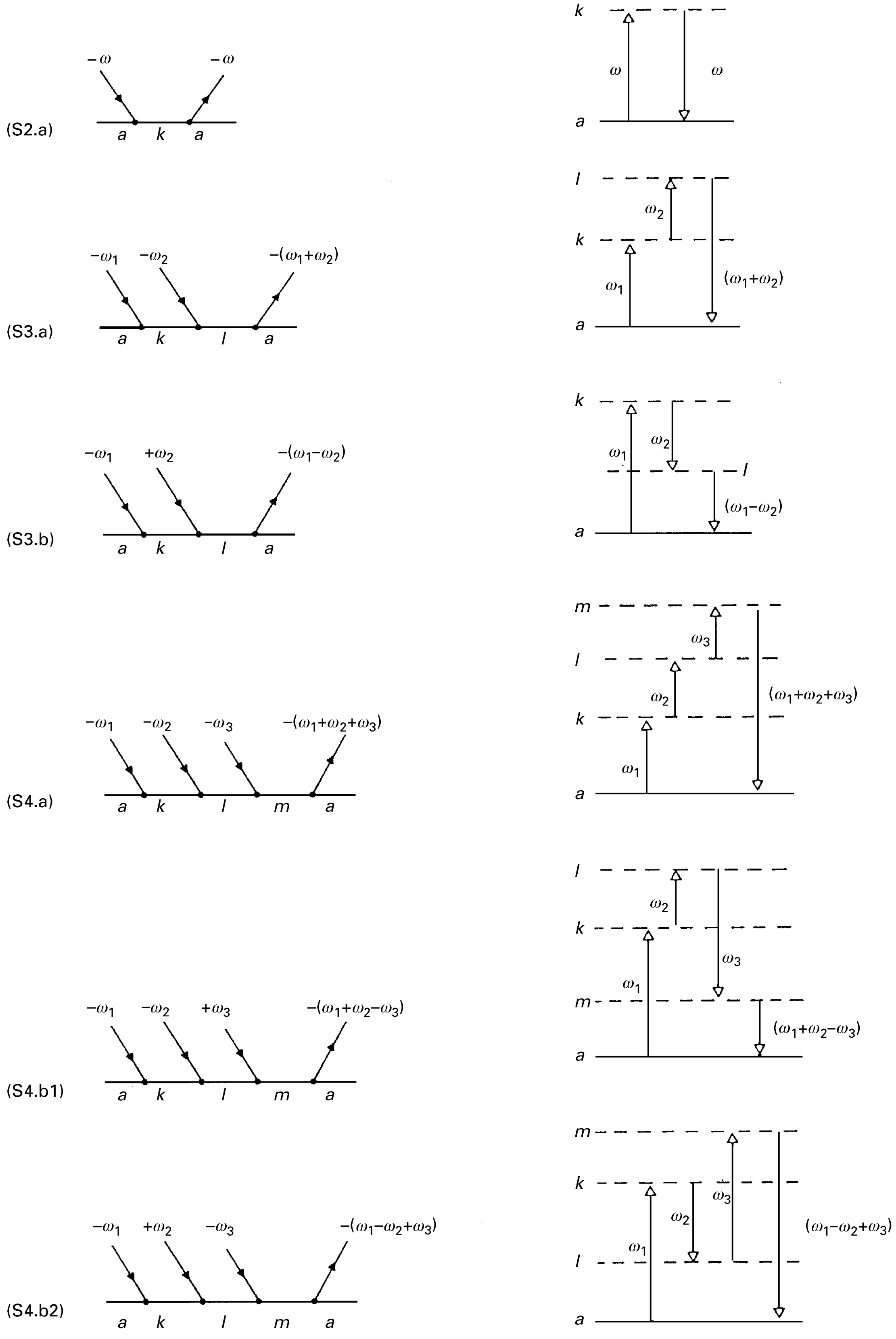



**Figure 1** Ward graphs (at left) and ladder graphs (at right) for linear (S2.a), second-order nonlinear (S3.a, b), and third-order nonlinear (S4.a, b1, b2) elastic scattering processes. The broken horizontal lines in the ladder graphs represent virtual, nonstationary states of the molecular system. Reproduced with permission from Wagnière GH (1993) *Linear and Nonlinear Optical Properties of Molecules.* Basel: Verlag Helvetica Chimica Acta.

**Table 3** Independent nonvanishing elements of $\chi^{(2)}(-\omega_1-\omega_2;\, \omega_1, \omega_2)$ for crystals of given symmetry classes

| *Crystal system* | *Crystal class* | *Nonvanishing tensor elements* |
|---|---|---|
| Triclinic | 1 | All elements are independent and nonzero |
| | $\bar{1}$ | Each element vanishes |
| Monoclinic | 2 | *xyz, xzy, xxy, xyx, yxx, yyy, yzz, yzx, yxz, zyz, zzy, zxy, zyx* (twofold axis parallel to $\hat{y}$ ) |
| | *m* | *xxx, xyy, xzz, xzx, xxz, yyz, yzy, yxy, yyx, zxx, zyy, zzz, zzx, zxz* (mirror plane perpendicular to $\hat{y}$) |
| | 2/*m* | Each element vanishes |
| Orthorhombic | 222 | *xyz, xzy, yzx, yxz, zxy, zyx* |
| | *mm*2 | *xzx, xxz, yyz, yzy, zxx, zyy, zzz* |
| | *mmm* | Each element vanishes |
| Tetragonal | 4 | *xyz* = −*yxz*, *xzy* = −*yzx*, *xzx* = *yzy*, *xxz* = *yyz*, *zxx* = *zyy*, *zzz*, *zxy* = −*zyx* |
| | $\bar{4}$ | *xyz* = *yxz*, *xzy* = *yzx*, *xzx* = −*yzy*, *xxz* = −*yyz*, *zxx* = −*zyy*, *zxy* = *zyx* |
| | 422 | *xyz* = −*yxz*, *xzy* = −*yzx*, *zxy* = −*zyx* |
| | 4*mm* | *xzx* = *yzy*, *xxz* = *yyz*, *zxx* = *zyy*, *zzz* |
| | $\bar{4}$2*m* | *xyz* = *yxz*, *xzy* = *yzx*, *zxy* = *zyx* |
| | 4/*m*, 4/*mmm* | Each element vanishes |
| Cubic | 432 | *xyz* = −*xzy* = *yzx* = −*yxz* = *zxy* = −*zyx* |
| | $\bar{4}$3*m* | *xyz* = *xzy* = *yzx* = *yxz* = *zxy* = *zyx* |
| | 23 | *xyz* = *yzx* = *zxy*, *xzy* = *yxz* = *zyx* |
| | *m*3, *m*3*m* | Each element vanishes |
| Trigonal | 3 | *xxx* = −*xyy* = −*yyz* = −*yxy*, *xyz* = −*yxz*, *xzy* = −*yzx* , *xzx* = *yzy*, *xxz* = *yyz*, *yyy* = −*yxx* = −*xxy* = −*xyx*, *zxx* = *zyy*, *zzz*, *zxy* = −*zyx* |
| | 32 | *xxx* = −*xyy* = −*yyx* = −*yxy*, *xyz* = −*yxz*, *xzy* = −*yzx*, *zxy* = −*zyx* |
| | 3*m* | *xzx* = *yzy*, *xxz* = *yyz*, *zxx* = *zyy*, *zzz*, *yyy* = −*yxx* = −*xxy* = −*xyx* (mirror plane perpendicular to $\hat{x}$) |
| | $\bar{3}$,$\bar{3}$*m* | Each element vanishes |
| Hexagonal | 6 | *xyz* = −*yxz*, *xzy* = −*yzx*, *xzx* = *yxy*, *xxz* = *yyz*, *zxx* = *zyy*, *zzz*, *zxy* = −*zyx* |
| | $\bar{6}$ | *xxx* = −*xyy* = −*yxy* = −*yyx*, *yyy* = −*yxx* = −*xyx* = −*xxy* |
| | 622 | *xyz* = −*yxz*, *xzy* = −*yxz*, *zxy* = −*zyx* |
| | 6*mm* | *xzx* = *yzy*, *xxz* = *yyz*, *zxx* = *zyy*, *zzz* |
| | $\bar{6}$*m*2 | *yyy* = −*yxx* = −*xxy* = −*xyx* |
| | 6/*m*, 6/*mmm* | Each element vanishes |

Reproduced with permission from Boyd RW (1992) *Nonlinear Optics.* Boston: Academic Press.

be equal to each other (see **Table 3**). For instance, in the chiral cubic point group O, the only nonvanishing elements are: $xyz = yzx = zxy = -xzy = -yxz = -zyx$; in the achiral point group $T_d$: $xyz = yzx = zxy = xzy = yxz = zyx$. For these cases, a second-order nonlinear response can only be detected if ${}^1\boldsymbol{E}$ and ${}^2\boldsymbol{E}$ are nonparallel to one and the same symmetry-adapted coordinate axis. As will be seen in more detail in the next section, in an isotropic medium, the averaged value of $\alpha^{(2)}$ only fails to vanish if the individual molecules of which the medium is composed are chiral, and if the frequencies $\omega_1$ and $\omega_2$ of the incident radiation are different.

The third-order polarizabilities $\alpha^{(3)}$ (...) are parity-even fourth-rank tensors. In the triclinic system, there occur $3^4 = 81$ independent and nonzero tensor elements. The presence of symmetry leads to corresponding simplifications.

The same symmetry considerations, which here are stated for individual molecules, may of course also be applied to macroscopic systems, in particular to crystals. The determining aspect is here the overall crystal symmetry, and instead of the molecular polarizabilities $\alpha^{(n)}$, we consider the bulk susceptibilities of the crystal $\chi^{(n)}$.

## Organic media

### Organic liquids and solutions

The nonlinear optical properties of solutions of organic molecules have been investigated extensively, although the selection rules for second-order nonlinear optical effects in isotropic liquids are quite restrictive. In order to be noncentrosymmetric, a fluid must consist of, or contain, chiral molecules. Such a chiral medium is 'optically active' and not superposable on its mirror image. Although sum and difference frequency generation are then possible, the important special cases of second-harmonic generation and optical rectification are still forbidden. The respective molecular polarizabilities $\alpha^{(2)}(-2\omega;\, \omega, \omega)$ and $\alpha^{(2)}(0; \omega, -\omega)$ vanish upon isotropic averaging.

Second-harmonic generation may be induced in any liquid medium (or gas) if an external static

electric field is applied to it, whereby the medium loses its centrosymmetry and the conditions for second-harmonic generation are fulfilled. The generalized polarizability leading to this effect may be expressed as $\alpha^{(3)}(-2\omega; \omega, \omega, 0)$ and is described by a fourth-rank tensor. With electric field strengths applicable in the laboratory, the effect is in general quite small. However, if the liquid is composed of polar molecules (not necessarily chiral), the applied electric field will also partially align them. This then leads to an additional, temperature-dependent, contribution to second-harmonic generation that can be stronger. It is proportional to the molecular dipole moment $\mu$ and to an average value of the tensor $\alpha^{(2)}(-2\omega; \omega, \omega)$, often denoted in the literature by $\beta$. Electric field-induced second-harmonic generation, in the literature sometimes abbreviated as EFISH, has been widely applied to study solutions of polar organic molecules in nonpolar solvents. To allow extraction of significant molecular data, the interaction between solute molecules should be negligible and the influence of the nonpolar solvent must be taken into account as an averaged correction.

Although second-harmonic generation attained by the EFISH effect is in general weak, the method has been applied widely and successfully. The molecular data so obtained serve as a point of departure for the interpretation of the nonlinear optical properties of molecular crystals and arrays and for the design of novel systems. One observes (see **Tables 4** and **5**) that particularly large quantities for $\beta$ are found in molecules containing one or more electron-donor substituent(s), such as $-NH_2$ (amino), one or more electron-acceptor substituent(s), such as $-NO_2$ (nitro), bound to a polarizable $\pi$ electron system (containing conjugated C=C double bonds).

The tensor elements of the molecular quantity $\alpha^{(2)}(-2\omega; \omega, \omega)$ and, therefrom, the averaged quantity $\beta$ may in principle be calculated quantum mechanically. $\alpha^{(2)}$ may be expressed in terms of the energy levels of the molecule and the electric dipole transition moments between the corresponding quantum states. Exact (*ab initio*) calculations are very cumbersome, but a number of simplified procedures (semiempirical calculations) have been applied to this problem and their results allow a reasonably successful interpretation of the measured results, in particular where strong charge-transfer effects come into play.

Third-order nonlinear optical effects, such as third-harmonic generation or other kinds of four-wave mixing phenomena, occur in all media, irrespective of their symmetry. This follows from the parity-even property of the corresponding tensors. Consequently, third-harmonic generation can be observed both in liquids and gases. Some results on organic molecules are given in **Tables 4** and **5**. In general $\gamma$, the dominant component of $\alpha^{(3)}(-3\omega; \omega, \omega, \omega)$, is a small quantity leading to correspondingly small effects.

### Organic layers and crystals

Any surface or interface breaks the inversion symmetry and is therefore a possible source of second-order effects. Owing to their surface sensitivity, second-harmonic generation measurements have developed into a very useful tool for probing the orientation of organic molecules in well-structured monolayers, such as those obtainable by the Langmuir–Blodgett technique (see **Table 6**). The surface susceptibility may in general be written as

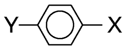
$$\chi_{\mathrm{S}}^{(2)} = \chi_{\mathrm{SM}}^{(2)} + \chi_{\mathrm{SB}}^{(2)} \qquad [5a]$$

where $\chi_{\mathrm{SM}}^{(2)}$ stands for the part arising from the adsorbed molecules and $\chi_{\mathrm{SB}}^{(2)}$ for the background contribution of the adjoining media. In order to obtain strong signals, the molecules in the layer must themselves be noncentrosymmetric. Often one chooses the adjoining bulk media to be centrosymmetric (air, water, glass; see **Figure 2**). Then

$$\chi_{\mathrm{S}}^{(2)} \approx \chi_{\mathrm{SM}}^{(2)} \qquad \chi_{\mathrm{SM}}^{(2)} \gg \chi_{\mathrm{SB}}^{(2)} \qquad [5b]$$

Among organic crystals, one of the most frequently used for second-harmonic generation is urea, composed of noncentrosymmetric molecules arranged in a noncentrosymmetric fashion, according to the tetragonal space group $P\bar{4}2_1m = \mathrm{D}_{2d}^3$ (see **Table 7**). Much attention has been devoted to the design and fabrication of even more efficient media, based on large $\beta$ values obtained from EFISH experiments. In some cases, such as that of *p*-nitroaniline, $\beta$ is large, but the molecules crystallize in a centrosymmetric space group, rendering the crystal useless. One strategy to overcome such difficulties consists in making the molecules chiral, thereby forcing them into a noncentrosymmetric crystal structure. From a theoretical point of view, one is interested in relating the bulk susceptibility of the crystal $\chi_{ijk}^{(2)}$ to the susceptibilities of the individual molecules $\alpha_{xyz}^{(2)}$ in their respective positions and orientations in the unit cell. Neglecting intermolecular interaction, this may be written as a sum

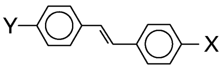
$$\chi_{ijk}^{(2)} = \frac{1}{V} L_{ijk} \sum_{s} \sum_{x_s y_s z_s} C_{ix_s} C_{jy_s} C_{kz_s} \alpha_{x_s y_s z_s}^{(2)} \qquad [6]$$

**Table 4** Properties of *para*-disubstituted benzenes: 

| *X* | *Y* | *Solvent* | $\lambda_{max}$ (nm)[a] | $\mu$ ($10^{-30}$ cm)[b] | $\alpha^{(1)}$ ($10^{-40}$J $m^2V^{-2}$) | $\alpha^{(2)} \equiv \beta$ ($10^{-50}$J $m^3V^{-3}$) | $\alpha^{(3)} \equiv \gamma$ ($10^{-60}$J $m^4V^{-4}$) |
|---|---|---|---|---|---|---|---|
| NO | $NMe_2$ | *p*–Dioxane | 407 | 20.7 | 23.3 | 4.44 | |
| $NO_2$ | Me | *p*–Dioxane | 272 | 14.0 | 17.8 | 0.78 | 0.99 |
| $NO_2$ | Br | *p*–Dioxane | 274 | 10.0 | 20.0 | 1.22 | |
| $NO_2$ | OH | *p*–Dioxane | 304 | 16.7 | 16.7 | 1.11 | 0.99 |
| $NO_2$ | OPh | *p*–Dioxane | 294 | 14.0 | 28.9 | 1.48 | 1.11 |
| $NO_2$ | OMe | *p*–Dioxane | 302 | 15.3 | 16.7 | 1.89 | 1.23 |
| $NO_2$ | SMe | *p*–Dioxane | 322 | 14.7 | 21.1 | 2.26 | 2.10 |
| $NO_2$ | $N_2H_3$ | *p*–Dioxane | 366 | 21.0 | 20.0 | 2.81 | 1.11 |
| $NO_2$ | $NH_2$ | Acetone | 365 | 20.7 | 18.9 | 3.41 | 1.85 |
| $NO_2$ | $NMe_2$ | Acetone | 376 | 21.3 | 24.4 | 4.44 | 3.46 |
| $NO_2$ | CN | *p*–Dioxane | | 3.0 | 18.9 | 0.22 | 0.86 |
| $NO_2$ | CHO | *p*–Dioxane | 376 | 8.3 | 18.9 | 0.07 | 0.86 |
| $CHC(CN)_2$ | OMe | *p*–Dioxane | 345 | 18.3 | 26.7 | 3.63 | 3.70 |
| $CHC(CN)_2$ | $NMe_2$ | $CHCl_3$ | 420 | 26.0 | 31.1 | 11.85 | |

[a] $\lambda_{max}$ denotes the wavelength of the lowest electronic transition; [b]$\mu$ denotes the ground-state dipole moment; the other quantities are explained in the text.

Data from Cheng L-T, Tam W, Stevenson SH, Meredith GR, Rikken G and Marder SR (1991) *Journal of Physical Chemistry* **95**: 10631; converted therefrom into SI units (see **Table 9**).

**Table 5** Properties of 4,4′-disubstituted stilbenes: 

| *X* | *Y* | *Solvent* | $\lambda_{max}$ (nm)[a] | $\mu$ ($10^{-30}$ cm)[b] | $\alpha^{(1)}$ ($10^{-40}$J $m^2V^{-2}$) | $\alpha^{(2)} \equiv \beta$ ($10^{-50}$J $m^3V^{-3}$) | $\alpha^{(3)} \equiv \gamma$ ($10^{-60}$J $m^4V^{-4}$) |
|---|---|---|---|---|---|---|---|
| CN | OH | *p*–Dioxane | 344 | 15.0 | 35.6 | 4.81 | 6.42 |
| CN | OMe | $CHCl_3$ | (340) | 12.7 | 37.8 | 7.04 | 6.67 |
| CN | $N(Me)_2$ | $CHCl_3$ | 382 | 19.0 | 43.3 | 13.33 | 15.43 |
| $NO_2$ | H | *p*–Dioxane | 345 | 14.0 | 32.2 | 4.07 | 7.53 |
| $NO_2$ | Me | *p*–Dioxane | 351 | 15.7 | 38.9 | 5.56 | 9.51 |
| $NO_2$ | Br | *p*–Dioxane | 344 | 10.7 | 42.2 | 5.19 | 12.10 |
| | | $CHCl_3$ | (356) | 11.3 | 36.7 | 6.67 | 5.56 |
| $NO_2$ | OH | *p*–Dioxane | 370 | 18.3 | 36.7 | 6.30 | 12.84 |
| $NO_2$ | OPh | *p*–Dioxane | 350 | 15.3 | 46.7 | 6.67 | 9.88 |
| $NO_2$ | OMe | *p*–Dioxane | 364 | 15.0 | 37.8 | 10.37 | 9.75 |
| | | $CHCl_3$ | (370) | 15.0 | 37.8 | 12.59 | 11.48 |
| $NO_2$ | SMe | *p*–Dioxane | 374 | 14.3 | 43.3 | 9.63 | 13.95 |
| | | $CHCl_3$ | (380) | 14.3 | 42.2 | 12.59 | 12.35 |
| $NO_2$ | $NH_2$ | $CHCl_3$ | 402 | 17.0 | 35.6 | 14.81 | 18.15 |
| $NO_2$ | $N(Me)_2$ | $CHCl_3$ | 427 | 22.0 | 37.8 | 27.04 | 27.78 |

For footnotes, see **Table 4**.

where $i$, $j$, $k$ denote the coordinate system of the crystal, $x_s$, $y_s$, $z_s$ that of the molecule $s$ in the unit cell. $L_{ijk}$ is a local-field correction, $V$ the volume of the unit cell. The trigonometric factors $C_{ix_s}$ relate the molecular coordinate systems to the crystal. This purely additive orientated gas model presents a useful first approximation for the interpretation of data on organic molecules. To refine it, intermolecular interaction in the crystal must be included in the calculation. For crystals of strongly polar molecules, methods based on the dipole–dipole approximation have been successful.

## From harmonic generation to parametric amplification

### Conservation of photon energy

The photons involved in a nonlinear optical process must fulfil the requirement of energy conservation. For a three-wave mixing effect in which the incident photons are of frequency $\omega_1$, $\omega_2$, leading to an outgoing photon of frequency $\omega_3$, this implies

$$\omega_1 + \omega_2 = \omega_3 \qquad [7]$$

For sum-frequency generation, where $\omega_3 = (\omega_1 + \omega_2)$, this is automatically fulfilled. For difference-frequency generation, where $\omega_3 = (\omega_1 - \omega_2)$, the above equation as such evidently cannot be satisfied; we must write

$$\omega_1 + \omega_2 = \omega_3 + 2\omega_2 \qquad [8]$$

This means that for each incident photon of frequency $\omega_2$ there are two outgoing photons of the same frequency. Simultaneously with the generation of a new wave of frequency $\omega_1 - \omega_2$, the incident wave of frequency $\omega_2$ is parametrically amplified. If the nonlinear medium is placed between two mirrors reflecting at the frequencies $\omega_2$ and (or) $\omega_3$, this parametric effect may be increased. One calls such a device a parametric oscillator (see **Figure 3**). From this point of view, $\omega_1 \equiv \omega_P$ corresponds to the so-called pump wave, $\omega_2 \equiv \omega_S$ to the (amplified) signal wave, and $\omega_3 \equiv (\omega_1 - \omega_2) \equiv \omega_I$ to the idler wave. Equation [8] may be simplified to

$$\omega_1 = \omega_2 + \omega_3, \qquad \text{or} \qquad \omega_P = \omega_S + \omega_I \qquad [9]$$

The fundamental process then appears to be the conversion of a photon of higher frequency $\omega_P$ into two photons of lower frequency $\omega_S$ and $\omega_I$. Interestingly, this process may go on in a parametric oscillator merely as a result of sending in a pump wave. The signal photons are first generated inside the cavity by spontaneous emission and then coherently amplified. Carried out in this manner, the intensity of

**Table 6** Surface susceptibility $\chi^{(2)}(-2\omega;\omega,\omega)$ and molecular second-order nonlinear polarizability $\alpha^{(2)}(-2\omega;\omega,\omega)$ for organic monomolecular layers on water

| *Molecule* | $\chi^{(2)}_{ijj}$ *(10⁻²⁰ m V⁻¹)* | $\alpha^{(2)}_{zzz}$ *(10⁻⁵⁰ J m³ V⁻³)* |
|---|---|---|
| $C_8H_{17}(C_6H_4)_2CN$ | 46[a] | 9.2 |
| $C_9H_{19}(C_6H_4)_2CN$ | 46 | 9.2 |
| $C_{10}H_{21}(C_6H_4)_2CN$ | 46 | 9.2 |
| $C_{12}H_{25}(C_6H_4)_2CN$ | 46 | 9.2 |
| $C_{14}H_{29}COOH$ | 0.21 | 0.030 |
| $C_{17}H_{35}COOH$ | 0.17 | 0.026 |
| $C_{22}H_{45}COOH$ | 0.17 | 0.026 |
| $C_{17}H_{35}CH_2OH$ | 0.25 | 0.041 |
| $C_{12}H_{25}(C_{10}H_6)SO_3Na$ | 0.75 | 0.28 |
| $C_8H_{17}(C_6H_4)_2COOH$ | 12[b] | 2.2 |
| $C_7H_{15}(C_4N_2H_2)C_6H_4CN$ | 8 | 3.0 |
| $C_5H_{11}(C_6H_4)_3CN$ | 15 | 2.8 |

Data from Rasing Th, Berkovic G, Shen YR, Grubb SG and Kim MW (1986) *Chemical Physics Letters* **130**: 1 and Berkovic G, Rasing Th and Shen YR (1987) *Journal of the Optical Society of America B* **4**: 945.

Fundamental wavelength $\lambda$ = 532 nm.

[a] For surface density $3.0 \times 10^{18}$ molecules $m^{-2}$.

[b] For surface density $2.5 \times 10^{18}$ molecules $m^{-2}$.

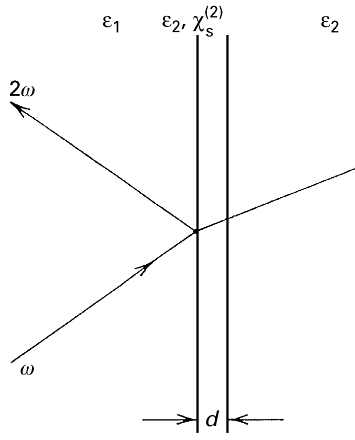



**Figure 2** Sketch of second-harmonic generation from an interface between two isotropic media. The interfacial layer of thickness $d$ is specified by a linear dielectric constant $\varepsilon_2$ and a second-order surface nonlinear susceptibility $\chi^{(2)}_s$. Reproduced with permission of John Wiley and Sons from Shen YR (1984). *The Principles of Nonlinear Optics.* New York: © 1984 John Wiley and Sons.

**Table 7** Experimental second-order nonlinear optical susceptibilities $d_{il}$ of organic crystals

| *Crystal* | *Symmetry* | $d_{il}$ ($10^{-12}$ m $V^{-1}$) | | *Reference* |
|---|---|---|---|---|
| MBBCH (2,6-bis(*p*-methylbenzylidene)-4-*t*-butylcyclohexanone) | Orthorhombic | $d_{31}$ | 15 | a |
| | $mm2 = C_{2v}$ | $d_{32}$ | 12 | |
| | | $d_{33}$ | 4 | |
| | | $d_{eff}$ | 12 (I) | |
| BBCP (2,5-bis(benzylidene)cyclopentanone) | $222 = D_2$ | $d_{14}$ | 7 | a |
| *m*-NA (*m*-nitroaniline) | $mm2 = C_{2v}$ | $d_{31}$ | 13.05 | b |
| | | $d_{32}$ | 1.09 | |
| | | $d_{33}$ | 13.72 | |
| | | $d_{eff}$ | 10.35 (I) | |
| 5NU ($P2_12_12_1$; 5-nitrouracil) | $222 = D_2$ | $d_{14}$ | 8.7 | c |
| POM (3-methyl-4-nitropyridine-1-oxide) | $222 = D_2$ | $d_{14}$ | 9.6 | d |
| MNA (2-methyl-4-nitroaniline) | Monoclinic | $d_{11}$ | 167.6 | e, j |
| | $m = C_s$ | $d_{12}$ | 25.1 | |
| | | $d_{33}$,$d_{13}$,$d_{31}$ | ~$10^{-3}$ $d_{11}$ | |
| | | $d_{eff}$ | 20.8 (I) | |
| L-PCA (*L*-pyrrolidone-2-carboxylic acid) | Orthorhombic | $d_{14}$ | 0.22 | f |
| | $222 = D_2$ | $d_{eff}$ | 0.20 (I) | |
| MAP (methyl-(2,4-dinitrophenyl)amino-2-propanoate) | Monoclinic | $d_{21}$ | 16.8 | g, j |
| | $2 = C_2$ | $d_{22}$ | 18.4 | |
| | | $d_{23}$ | 3.7 | |
| | | $d_{25}$ | –0.54 | |
| | | $d_{eff}$ | 16.3 (I) | |
| | | $d_{eff}$ | 8.8 (II) | |
| Urea ($CO(NH_2)_2$) | Tetragonal | $d_{14}$ | 1.4 | h, i |
| | $\bar{4}2m = D_{2d}$ | | | |

Fundamental wavelength $\lambda = 1.064$ μm. *Data for different frequencies available.
(I) For type I phase-matched SHG; (II) for type II phase-matched SHG.
[a] Kawamata J, Inoue K and Inabe T (1995) *Applied Physics Letters* **66**: 3102.
[b] Huang G-F, Lin JT, Su G, Jiang R and Xie S (1992) *Optical Communications* **89**: 205.
[c] Puccetti G, Perigaud A, Badan J, Ledoux I and Zyss J (1993) *Journal of the Optical Society of America B* **10**: 733.
[d] Zyss J, Chemla DS and Nicoud JF (1981) *Journal of Chemical Physics* **74**: 4800.
[e] Levine BF, Bethea CG, Thurmond CD, Lynch RT and Bernstein JL (1979) *Journal of Applied Physics* **50**: 2523.
[f] Kitazawa M, Higuchi R, Takahashi M, Wada T and Sasabe H (1995) *Journal of Applied Physics* **78**: 709.
[g] Oudar JL and Hierle R (1977) *Journal of Applied Physics* **48**: 2699.
[h] Catella GC, Bohn JH and Luken JR (1988) *IEEE Journal of Quantum Electronics* **24**: 1201.
[i] Halbout J-M, Blit S, Donaldson W and Tang CL (1979) *IEEE Journal of Quantum Electronics* **QE-15**: 1176.
[j] Nicoud JF and Twieg RJ (1987) In: Chemla DS and Zyss J (eds) *Nonlinear Optical Properties of Organic Molecules and Crystals*, Vol.1, pp 227–296. London: Academic Press.

the signal wave becomes linearly dependent on the intensity of the incident pump wave.

Evidently, a photon of frequency $\omega_P$ may break up into two photons of lower frequency in an infinity of ways, depending on the relative frequencies $\omega_S$ and $\omega_I$. In order to select which frequency $\omega_S$ should be amplified, the parametric oscillator must be correspondingly tuned. The most important and practical way to achieve this tuning is by phase matching in a crystal.

### Conservation of photon momentum: phase matching

To optimize the intensity of a coherent nonlinear optical effect, there must be conservation of photon momentum. For sum frequency generation this requirement is expressed as

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_3 \qquad [10a]$$

**Table 8** Experimental second-order nonlinear optical susceptibilities $d_{il}$ of inorganic crystals

| *Materials* | *Symmetry* | | $d_{il}$ ($10^{-12}$ m $V^{-1}$) | $d_{il}$ $\lambda$(μm) | *Reference* |
|---|---|---|---|---|---|
| Quartz ($\alpha$-$SiO_2$) | 32 = $D_3$ | $d_{11}$ | 0.46 | 1.06 | a |
| | | $d_{14}$ | 0.009 | 1.0582 | b |
| $LiIO_3$ | 6 = $C_6$ | $d_{31}$ | 6.43 | 2.12 | a |
| | | | 7.11 | 1.06 | |
| | | | 8.14 | 0.6943 | |
| | | $d_{33}$ | 6.41 | 2.12 | |
| | | | 6.75 | 1.318 | |
| | | | 7.02 | 1.06 | |
| $LiNbO_3$ | 3*m* = $C_{3v}$ | $d_{31}$ | 5.77 | 1.15 | a |
| | | | 5.95 | 1.06 | |
| | | $d_{33}$ | 29.1 | 2.12 | |
| | | | 31.8 | 1.318 | |
| | | | 34.4 | 1.06 | |
| | | $d_{22}$ | 3.07 | 1.0582 | b |
| $KNbO_3$ | *mm*2 = $C_{2v}$ | $d_{31}$ | – 15.8 | 1.064 | g |
| | | $d_{32}$ | – 18.3 | | |
| | | $d_{33}$ | – 27.4 | | |
| | | $d_{24}$ | – 17.1 | | |
| | | $d_{15}$ | – 16.5 | | |
| $Ba_2NaNb_5O_{15}$ | *mm*2 = $C_{2v}$ | $d_{31}$ | – 14.55 | 1.0642 | b |
| | | $d_{32}$ | – 14.55 | | |
| | | $d_{33}$ | – 20 | | |
| $BaTiO_3$ | 4*mm* = $C_{4v}$ | $d_{15}$ | –17.2 | 1.0582 | b |
| | | $d_{31}$ | – 18 | | |
| | | $d_{33}$ | – 6.6 | | |
| $NH_4H_2PO_4$(ADP) | $\bar{4}$2m = $D_{2d}$ | $d_{14}$ | 0.48 | 0.6943 | b |
| | | $d_{36}$ | 0.485 | | |
| $KH_2PO_4$(KDP) | $\bar{4}$2*m* = $D_{2d}$ | $d_{14}$ | 0.49 | 1.0582 | b |
| | | $d_{36}$ | 0.599 | 1.318 | a |
| | | | 0.630 | 1.06 | |
| | | | 0.712 | 0.6328 | |
| $KD_2PO_4$(KD*P) | $\bar{4}$2*m* = $D_{2d}$ | $d_{14}$ | 0.528 | | c |
| | | $d_{36}$ | 0.528 | | |
| GaP | $\bar{4}$3*m* = $T_d$ | $d_{14}$ | 35 | 3.39 | b |
| | | $d_{36}$ | 58.1 | 10.6 | a |
| | | | 77.5 | 2.12 | |
| | | | 99.7 | 1.06 | |
| GaAs | $\bar{4}$3 *m* = $T_d$ | $d_{14}$ | 188.5 | 10.6 | b |
| | | $d_{36}$ | 151 | 10.6 | a |
| | | | 173 | 2.12 | |
| $AgGaSe_2$ | $\bar{4}$2*m* = $D_{2d}$ | $d_{36}$ | 57.7 | 10.6 | a |
| | | | 67.7 | 2.12 | |
| $AgSbS_3$ | 3*m* = $C_{3v}$ | $d_{31}$ | 12.6 | | c |
| | | $d_{22}$ | 13.4 | | |
| $Ag_3AsS_3$ | 3*m* = $C_{3v}$ | $d_{31}$ | 15.1 | | c |
| | | $d_{22}$ | 28.5 | | |
| CdS | 6*mm* = $C_{6v}$ | $d_{33}$ | 36.0 | | c |
| | | $d_{31}$ | 37.7 | | |
| | | $d_{36}$ | 41.9 | | |
| CdSe | 6*mm* = $C_{6v}$ | $d_{15}$ | 31 | 10.6 | b |
| | | $d_{31}$ | 28.5 | | |
| | | $d_{33}$ | 55.3 | 10.6 | a |

**Table 8** *Continued*

| *Materials* | *Symmetry* | | $d_{il}$ ($10^{-12}$ m $V^{-1}$) | $d_{il}$ $\lambda$(μm) | *Reference* |
|---|---|---|---|---|---|
| | | | 65.4 | 2.12 | |
| Te | $32 = D_3$ | $d_{11}$ | $5 \times 10^3$ | 10.6 | b |
| $\beta$-$BaB_2O_4$(BBO) | $3m = C_{3v}$ | $d_{11}$ | 1.6 | 1.064 | d |
| | | $d_{22}, d_{31}$ | < 0.08 | | |
| $LaBGeO_5Nd^{3+}$ | | $d_{eff}$ | 0.296 | 1.064 | e |
| $KTiOPO_4$ | $mm2 = C_{2v}$ | $d_{15}$ | 1.91 | 1.064 | h |
| (KTP) | | $d_{24}$ | 3.64 | | |
| | | $d_{31}$ | 2.54 | | |
| | | $d_{32}$ | 4.35 | | |
| | | $d_{33}$ | 16.9 | | |
| $RbTiOPO_4$ | $mm2 = C_{2v}$ | $d_{15}$ | 6.1 | 1.064 | f |
| | | $d_{24}$ | 7.6 | | |
| | | $d_{31}$ | 6.5 | | |
| | | $d_{32}$ | 5.0 | | |
| | | $d_{33}$ | 13.7 | | |
| ZnO | $6mm = C_{6v}$ | $d_{31}$ | 2.1 | 1.0582 | b |
| | | $d_{15}$ | 4.3 | | |
| | | $d_{33}$ | −7.0 | | |

[a] Absolute values: Choy MM and Byer RL (1976) *Physical Review* **B14**: 1693
[b] Shen YR (1984) *The Principles of Nonlinear Optics*. New York: Wiley
[c] Boyd RW (1992) *Nonlinear Optics*. Boston: Academic Press
[d] Eimerl D, Davis L, Velsko S, Graham EK and Zalkin A (1987) *Journal of Applied Physics* **62**: 1968
[e] For type I phase-matched SHG; Capmany J and Garcia Sole J (1997) *Applied Physics Letters* **70**: 2517
[f] Zumsteg FC, Bierlein JD and Gier TE (1976) *Journal of Applied Physics* **47**: 4980
[g] Biaggio I, Kerkoc P, Wu L-S, Günter P and Zysset P (1992) *Journal of the Optical Society of America B* **9**: 507
[h] Vanherzeele H and Bierlein JD (1992) *Optics Letters* **17**: 982.

$\boldsymbol{k}_i$ denotes the wave vector of the corresponding beam. To avoid reduction of effective beam interaction length due to finite cross-sections, collinear phase matching is aimed at. One then may write equation [10a] in scalar form

$$k_1 + k_2 = k_3 \qquad [10b]$$

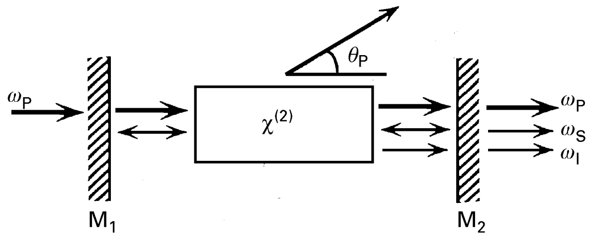



**Figure 3** Schematic representation of a singly-resonant optical parametric oscillator. Pump wave of frequency $\omega_P$, (reflected) signal wave of frequency $\omega_S$, idler wave of frequency $\omega_I$. The signal wave $\omega_S$ becomes amplified. $\theta_P$ denotes the angle of orientation of the direction of propagation with respect to the crystal optic axis. Adapted with permission from Tang CL and Cheng LK (1995) *Fundamentals of Optical Parametric Processes and Oscillators*. Amsterdam: Harwood Academic Publishers.

With $k_i = n_i 2\pi/\lambda_i$, where $n_i \equiv n(\omega_i)$ stands for the refractive index of the medium at frequency $\omega_i$ and $\lambda_i$ for the vacuum wavelength, this may be expressed as

$$\omega_1 n_1 + \omega_2 n_2 = \omega_3 n_3 \qquad [10c]$$

In a lossless medium, $n(\omega)$ in general increases monotonically with $\omega$ owing to normal dispersion. In an isotropic medium such as a liquid, $n(\omega)$ is independent of beam polarization. It can then easily be shown that for $\omega_1 \leq \omega_2 < \omega_3$, Equation [10c] cannot be satisfied.

In a uniaxial birefringent crystal, excluding propagation along the optic axis, an incident beam may, depending on its polarization, be made ordinary or extraordinary. The ordinary (o) and extraordinary (e) rays, perpendicularly polarized with respect to each other, will each experience a different index of refraction, $n^{(o)}(\omega) \neq n^{(e)}(\omega)$. According to the crystalline medium, the frequency $\omega$ and the angle of incidence with respect to the optic axis, situations may be found where the phase-matching condition is fulfilled.

For sum-frequency generation in a positive uniaxial crystal, in which $n^{(e)} > n^{(o)}$, the phase-matching condition may be satisfied in two different ways:

$$n_1^{(e)}\omega_1 + n_2^{(e)}\omega_2 = n_3^{(o)}\omega_3 \qquad \text{Type I} \qquad [11a]$$

$$n_1^{(o)}\omega_1 + n_2^{(e)}\omega_2 = n_3^{(o)}\omega_3 \qquad \text{Type II}$$

or

$$n_1^{(e)}\omega_1 + n_2^{(o)}\omega_2 = n_3^{(o)}\omega_3 \qquad [11b]$$

Similarly for the parametric effect in a negative uniaxial crystal, in which $n^{(e)} < n^{(o)}$ (see **Figure 4**):

$$n_P^{(e)}\omega_P = n_S^{(o)}\omega_S + n_I^{(o)}\omega_I \qquad \text{Type I} \qquad [12a]$$

$$n_P^{(e)}\omega_P = n_S^{(e)}\omega_S + n_I^{(o)}\omega_I \qquad \text{Type II}$$

or

$$n_P^{(e)}\omega_P = n_S^{(o)}\omega_S + n_I^{(e)}\omega_I \qquad [12b]$$

Crystals belonging to the cubic crystal system are isotropic, and therefore unsuited for phase-matching. Tetragonal and trigonal crystals are uniaxial; those of the orthorhombic, monoclinic and triclinic symmetry are biaxial. The description of phase-matching in biaxial crystals is somewhat more complicated than in uniaxial crystals, but it essentially rests on the same principles.

The search for birefringent crystals with good phase-matching properties is of great technical importance in nonlinear optics. Although phase-matching has been achieved in organic crystals (see **Table 7**), inorganic materials appear so far to offer a greater variety of possibilities.

## Inorganic media

### Noncentrosymmetric crystals

Inorganic crystals are widely applied for second-harmonic generation and for optical parametric processes. Some frequently used materials: for second-harmonic generation from the near-IR into the visible and beyond $KH_2PO_4$ (KDP), $KD_2PO_4$ (KD*P) and more recently $\beta$-$BaB_2O_4$(BBO). For optical parametric amplification into the mid-IR: $AgGaSe_2$, GaSe; the visible and near-IR: $LiNbO_3$, $KTiOPO_4$, $KNbO_3$; and into the visible and UV: $\beta$-$BaB_2O_4$ and $LiB_3O_5$.

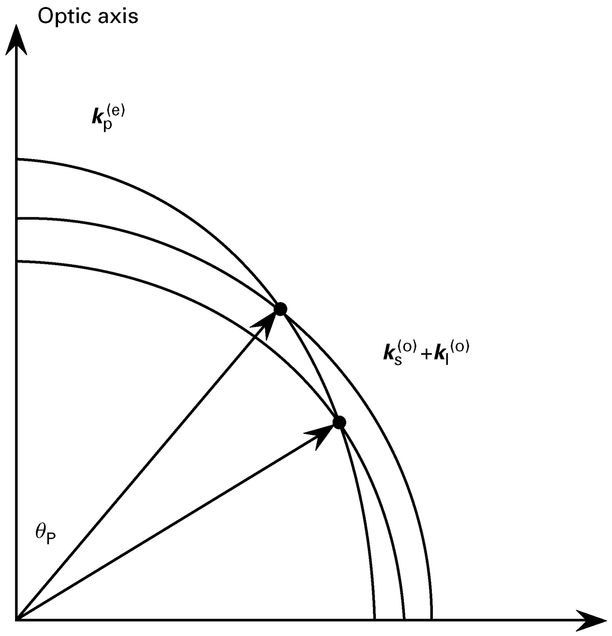



**Figure 4** Phase-matching in an optical parametric process to achieve photon momentum conservation is based on the use of birefringence to compensate for normal material dispersion. In an uniaxial crystal, the ordinary wave (o) is polarized perpendicularly to the plane defined by the direction of propagation and the optic axis. The corresponding value of $k^{(o)}$ (or $n^{(o)}$) is independent of $\theta_P$, the angle of orientation of the direction of propagation with respect to the optic axis. $k_S^{(o)}$ and $k_I^{(o)}$ therefore lie on a circle. The extraordinary wave (e) is polarized in the plane defined by the direction of propagation and the optic axis. The value of $k_P^{(e)}$ (or $n_P^{(e)}$) in its dependence on $\theta_P$ is described by an ellipse. In a negative uniaxial crystal, and for given values of $\omega_P = \omega_S + \omega_I$, the ellipse for $k_P^{(e)}$ may intersect the circle for $k_S^{(o)} + k_I^{(o)}$. At the corresponding angle $\theta_P$ there is phase matching. Rotation of the crystal relative to the direction of propagation of the waves correspondingly leads to tuning of the frequencies of the signal and idler waves. Adapted with permission from Tang CL and Cheng LK (1995). *Fundamentals of Optical Parametric Processes and Oscillators*. Amsterdam: Harwood Academic Publishers.

**Table 8** shows experimental second-order nonlinear optical susceptibilities for different tensor components $d_{il}$ and various fundamental wavelengths. The quantities $d_{il}$ are defined as follows:

$$\chi_{ijk}^{(2)}(2\omega) = 2d_{ijk}(2\omega) \qquad [13]$$

The second and third indices of $d_{ijk}$ are then replaced by a single symbol $l$ according to the piezoelectric

contraction:

| $jk:$ | 11 | 22 | 33 | 23, 32 | 31, 13 | 12, 21 |
|---|---|---|---|---|---|---|
| $\ell:$ | 1 | 2 | 3 | 4 | 5 | 6 |

The nonlinear susceptibility tensor can then be represented as a $3 \times 6$ matrix containing 18 elements. In the transparent region, i.e. outside of absorption bands, one may assume the validity of the Kleinman symmetry condition, which states that the indices $i$, $j$, $k$ may be freely permuted:

$$d_{ijk} = d_{kij} = d_{jki}, \quad \text{etc.} \qquad [14a]$$

One then finds, for instance,

$$d_{123} = d_{312}, \quad d_{14} = d_{36} \qquad [14b]$$

In this case there are only 10 independent elements for $d_{il}$.

**Table 8** shows that the values for $d_{il}$ may vary over several orders of magnitude, and that it is not necessarily the crystals with the highest values that are most commonly used. The technical applicability is partly also determined by other qualities, such as phase-matching properties, ease of crystal growth, mechanical strength, chemical inertness, temperature stability and light-damage threshold.

A quantity often used to characterize the optical properties of nonlinear optical materials is the Miller index:

$$\delta_{ijk} = \frac{d_{ijk}(2\omega)}{\chi_{ii}^{(1)}(2\omega)\, \chi_{jj}^{(1)}(\omega)\, \chi_{kk}^{(1)}(\omega) \cdot \varepsilon_0} \qquad [15]$$

$\chi^{(1)}(2\omega)$ represents the linear susceptibility for the doubled frequency $2\omega$, $\chi^{(1)}(\omega)$ that for the fundamental frequency $\omega$. One finds that for most materials $\delta$ is not far from a mean value of about $2 \times 10^{-2}$ m$^2$ C$^{-1}$, suggesting that in a given substance nonlinear and linear susceptibilities are closely related.

Noncentrosymmetric crystals show other properties in addition to frequency conversion, for instance the linear electro-optic or Pockels effect: the linear change of the refractive index induced by an applied DC electric field. Furthermore, the point groups $C_n$ and $C_{nv}$ allow for the existence of a permanent electric dipole moment. Indeed, crystals such as $LiNbO_3$ ($C_{3v}$) and $BaTiO_3$ ($C_{4v}$) are well-known for their ferroelectric properties. Crystals transforming according to point groups containing only rotations, such as $C_n$, $D_n$, T and O are chiral and therefore optically active. In **Table 8** we find quartz $\alpha$-$SiO_2$($D_3$), $LiIO_3$($C_6$) and Te($D_3$).

## Harmonic generation in metal vapours

Third-harmonic generation can in principle occur in all matter, as it is not tied to the condition of noncentrosymmetry. While the effect has been investigated in liquids and solids, the use of gases, in particular alkali metal vapours, has proved particularly interesting. In spite of the relatively low density of atoms, the third-harmonic generation efficiency can become quite high, up to 10%. The limiting laser intensity in gases is orders of magnitude higher than in condensed matter. Furthermore, the sharper transitions in gases allow strong enhancement of $\chi^{(3)}$ near *resonances*, especially three-photon resonances, which are electric dipole-allowed with respect to the atomic ground state. In sodium vapour this corresponds to transitions $3s \rightarrow 3p$, $3s \rightarrow 4p$, etc. Enhancement may in principle also occur via intermediate one-photon resonances, of same symmetry as three-photon resonances; or by two-photon resonances at transitions of symmetry $3s \rightarrow 4s$, $3s \rightarrow 5s$, or $3s \rightarrow 3d$, etc. The resonance enhancement of $\chi^{(3)}(3\omega)$ will evidently be diminished by concurrent multiphoton (or single-photon) absorption. In tuning $\omega$, a compromise must be sought, whereby the anomalous dispersion of $\chi^{(3)}$ is maximized in comparison to energy dissipation through absorption.

The anomalous dispersion of $\chi^{(3)}(3\omega)$ near resonances may also be used to achieve phase matching,

$$n(\omega) = n(3\omega) \qquad [16a]$$

which in a normally dispersive isotropic medium would be impossible. Considering an alkali atom A, and assuming $\omega$ to be below, and $3\omega$ to be above a strong $s \rightarrow p$ transition, we find

$$n_A(\omega) > n_A(3\omega) \qquad [16b]$$

Phase matching may be achieved by admixture of a buffer gas B. Such an inert gas must be transparent at frequency $3\omega$ and above; then

$$n_B(\omega) < n_B(3\omega) \qquad [16c]$$

The relative concentration of the inert gas is adjusted, so as to have for the mixture M,

$$n_{\mathrm{M}}(\omega) = n_{\mathrm{M}}(3\omega) \qquad [16d]$$

High conversion efficiencies have, for instance, been achieved with the mixtures Rb:Xe (10%) and Na:Mg (3.8%).

## Four-wave mixing

Beside third-harmonic generation, there exists a large variety of four-wave mixing effects. Depending on the combination of frequencies, on the occurrence of intermediate resonances and on the polarization of the light beams involved, the manifestation of these phenomena may be very different. We limit our considerations to a few selected examples.

### Coherent Raman spectroscopy

In coherent anti-Stokes Raman spectroscopy (CARS) two beams of frequency $\omega_1$ and $\omega_2$ are mixed in the sample to generate a new frequency $\omega_s = 2\omega_1 - \omega_2$. If there is a Raman resonance at $\omega_1 - \omega_2 = \Omega$, an amplified signal is detected at the anti-Stokes frequency $\omega_1 + \Omega$ (see **Figure 5**). The corresponding susceptibility $\chi^{(3)}(-\omega_4;\, \omega_1, \omega_2, \omega_3)$ may be written $\chi^{(3)}(-\omega_1 - \Omega;\, \omega_1, -\omega_1 + \Omega, \omega_1)$. The major experimental advantage of CARS and of other coherent Raman techniques is the large, highly directional signal produced, of the order of $10^4$ times more intense than would be obtained for conventional spontaneous Raman scattering. Usually, CARS experiments are performed with pulsed lasers delivering a peak power of the order of 10 – 100 kW. High frequency-resolution measurements with CW lasers are also possible. CARS experiments have been performed in gases, liquids and solids and on a variety of substances, ranging from diamond to aqueous solutions of biological macromolecules. Of particular interest is the use of CARS for combustion diagnostics. The coherent Raman signals can easily be separated from the luminescent background in flames.

Other, related coherent Raman effects are also represented in **Figure 5**, such as the case (C) where the signal beam is detected at the Stokes frequency. The Raman-induced Kerr effect (B) may be interpreted as the quadratic influence of an electric field of frequency $\omega_2$ on the elastic scattering of radiation at a frequency $\omega_1$, or vice versa. In this case the phase-matching (or wave-vector-matching) condition is fulfilled for any angle between beams 1 and 2, while in cases (A) and (C) it may only be met for certain angles of the beams with respect to each other.

### Degenerate four-wave mixing

The process governed by the third-order susceptibility $\chi^{(3)}(-\omega;\, \omega, -\omega, \omega)$ is called degenerate four-wave mixing. It may lead to a variety of highly interesting effects, one of them being that the index of refraction $n(\omega)$ becomes dependent on the incident light intensity $I(\omega)$:

$$n(\omega) = n_0(\omega) + n_2(\omega)I(\omega) \qquad [17]$$

For a single-mode laser beam with a Gaussian transverse intensity distribution, the index of refraction at the centre of the beam will then be larger than at its periphery, provided $n_2(\omega)$ is positive. Thereby the medium will act as a positive lens, tending to bring the incident beam to a focus at the centre on the beam. However, only if the intensity of the laser beam is sufficiently large will this self-focusing effect be able to counteract the beam spread due to ordinary diffraction.

An effect that may also occur with other nonlinear optical phenomena, but that has been extensively

**Table 9** Conversion from CGS-esu to SI units for *n*th order optical quantities

| *Conversion factor for n ≥ 1* | *[SI] ← [CGS-esu]* | *Dimension in SI units* |
|---|---|---|
| $\chi^{(n)}[\mathrm{SI}] = \dfrac{4\pi}{(3\times 10^4)^{n-1}}\,\chi^{(n)}[\mathrm{esu}]$ | | $\left(\dfrac{\mathrm{m}}{\mathrm{V}}\right)^{n-1}$ |
| $\alpha^{(n)}[\mathrm{SI}] = \dfrac{10^{-7}}{(3\times 10^4)^{n+1}}\,\alpha^{(n)}[\mathrm{esu}]^*$ | | $\mathrm{J}\left(\dfrac{\mathrm{m}}{\mathrm{V}}\right)^{n+1*}$ |
| * The case $n = 0$ corresponds to the conversion factor for a permanent electric dipole moment: | | |
| $\mu[\mathrm{SI}] = \dfrac{1}{3}\times 10^{-11}\mu\ [\mathrm{esu}]$ | | $\mathrm{J}\dfrac{\mathrm{m}}{\mathrm{V}} = \mathrm{cm}$ |

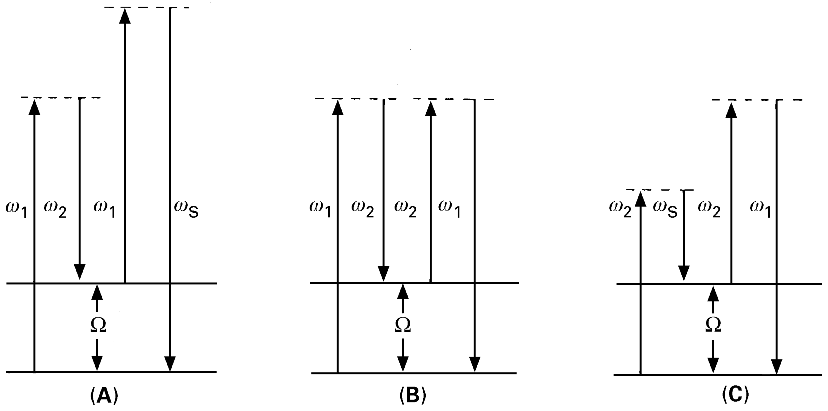



**Figure 5** Ladder graphs for four-wave mixing effects containing Raman processes. In all cases there is assumed an intermediate Raman-type resonance at the frequency Ω. (A) The coherent anti-Stokes Raman (CARS) process. (B) The process responsible for stimulated Raman spectroscopy (SRS) as well as the Raman-induced Kerr effect (TRIKE). (C) The coherent Stokes Raman spectroscopy (CSRS). Adapted with permission from Levenson MD (1982), *Introduction to Nonlinear Laser Spectroscopy.* New York: Academic Press.

studied in the frame of degenerate four-wave mixing, is phase conjugation. Here we consider not a single beam of frequency $\omega$, but four different beams: the collinear counterpropagating pump beams 1 and 2 interfere in the $\chi^{(3)}$-active medium to form an induced static grating. From this grating a signal wave 3, incident at a given angle with respect to 1 and 2, is scattered and reflected. The coherent reflected wave 4 is phase conjugate with respect to 3. For instance, if 3 is a forward-travelling plane wave

$$E_3(z,t) = A_3 \cos(\omega t - kz - \phi_0) \qquad [18a]$$

the corresponding phase-conjugate wave 4 will be

$$\begin{aligned} E_4(z,t) &= A_4 \cos(-\omega t - kz - \phi_0) \\ &= A_4 \cos(+\omega t + kz + \phi_0) \qquad [18b] \end{aligned}$$

It will travel backwards and behave as if the time $t$ had been replaced by $-t$. A nonlinear medium susceptible to degenerate four-wave mixing can thus be used as a phase-conjugate mirror. A left circularly polarized incident beam will be reflected as a left circularly polarized beam, and not as a right circularly polarized one as would be the case upon ordinary reflection. The phase conjugation process can be thought of as the generation of a time-reversed wavefront. If the input signal wave in passing through a medium before entering the phase-conjugate mirror suffers a wavefront distortion, the phase conjugate wave reflected back through the medium will remove this distortion. The phenomenon of phase conjugation can, for instance, be used to correct for aberrations induced by amplifying media.

## Particular aspects of nonlinear optics

### Higher order electromagnetic effects

The interaction energy of a molecular system with the radiation field may formally be expanded into a multipole series. The first term in this expansion contains the electric dipole–electric field interaction; in the second term appear the magnetic dipole–magnetic field interaction and the electric quadrupole interaction with the electric field gradient of the radiation, and so on. If the wavelength of light is large compared to the molecular dimensions, the higher multipole effects tend to be small and are often negligible from an experimental standpoint. The discussion until now has therefore considered only dominant electric dipole contributions to the molecular polarizability or bulk susceptibility. However, depending on molecular symmetry, there are situations where magnetic dipole and electric quadrupole interactions may become measurable. For instance, owing to these, weak second-harmonic generation may also be observed in some centrosymmetric crystals. Furthermore, the interplay of electric dipole, magnetic dipole and electric quadrupole interactions in chiral media leads to natural optical activity and to related higher-order nonlinear circular differential effects.

Particular nonlinear optical phenomena arise also when static electric or magnetic fields are applied. The molecular states and selection rules are thereby modified, leading, for instance, to higher-order, nonlinear-optical variants of the linear (Pockels) and quadratic (Kerr) electro-optical effect, or of the linear (Faraday) and quadratic (Cotton–Mouton) magneto-optical effect.

### Incoherent higher-harmonic scattering

We have seen that coherent second-harmonic generation is forbidden in liquids, even in chiral ones. This is due to the fact that the relevant molecular quantity $\alpha^{(2)}(-2\omega;\omega,\omega)$ vanishes when averaged over all possible molecular orientations:

$$\langle\alpha^{(2)}(-2\omega;\omega,\omega)\rangle = 0 \qquad [19a]$$

However, the inhomogeneity of the liquid at the molecular level and the fact that every molecule is an individual scatterer of radiation are not fully taken into account. The superposition of this molecular scattered radiation is partly incoherent. It consists mainly of 'ordinary' Rayleigh scattering at the basic frequency $\omega$; but if the molecules are noncentrosymmetric, some incoherent radiation of frequency $2\omega$ is also generated. This hyper-Rayleigh scattering, though weak, is clearly detectable with pulsed lasers of megawatt peak power. Its intensity is proportional to the square of $\alpha^{(2)}(-2\omega;\omega,\omega)$, which upon averaging over all spatial orientations in the liquid does not vanish:

$$\langle|\alpha^{(2)}(-2\omega;\omega,\omega)|^2\rangle \neq 0 \qquad [19b]$$

From the directional dependence and the depolarization ratios of the scattered radiation, information may be gained on particular tensor elements of $\alpha^{(2)}(-2\omega)$. The method has the advantage over EFISH measurements that it is also applicable to noncentrosymmetric molecules that do not posses a permanent dipole moment, in particular to, 'octopolar' molecules of symmetry $D_{3h}$ (such as tricyanomethanide $[C(CN)_3]^-$) or of symmetry $T_d$ (such as $CCl_4$). It is to be expected that progress in laser technology and light detection systems will further improve the applicability of the method.

## List of symbols

$C_{ix_s}$ etc. = trigonometric factors; (e) refers to the extraordinary ray; $\boldsymbol{E}$ = electric field strength of incident radiation; $I(\omega)$ = intensity of incident/scattered radiation; $i,j,k$ = coordinate system of crystal; $\boldsymbol{k}_i$ = wave vector of beam $i$; $L_{ijk}$ = local-field correction; $n$ = order of nonlinear effect; $n_i$ = refractive index of medium at $\omega_i$; (o) = refers to the ordinary ray; $p^{(n)}$ = molecular induced electric dipole moment ($n$th-order effect); $P^{(1)}$ = volume polarization; $V$ = volume of unit cell; $x_s$, $y_s$, $z_s$ = coordinate system of molecules; $\alpha^{(n)}$ = molecular polarizability of $n$th order; $\delta_{ijk}$ = Miller index (see equation [15]); $\varepsilon_0$ = permittivity of free space; $\lambda$ = wavelength; $\mu$ = static molecular dipole moment; $\phi$ = phase angle; $\chi^{(1)}$ = macroscopic susceptibility; $\chi^{(2)}_S$ = surface susceptibility; $\omega$ = photon frequency.

*See also:* **Electromagnetic Radiation; Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Linear Dichroism, Theory; Multiphoton Spectroscopy, Applications; Optical Frequency Conversion; Raman Optical Activity, Applications; Raman Optical Activity, Spectrometers; Raman Optical Activity, Theory; Raman Spectrometers; Rayleigh Scattering and Raman Spectroscopy, Theory; Symmetry in Spectroscopy, Effects of.**

## Further reading

Andrews DL (1993) Molecular theory of harmonic generation. Modern nonlinear optics, Part 2. *Advances in Chemical Physics* 85: 545–606.

Bloembergen N (1965) *Nonlinear Optics*. NewYork: WA Benjamin.

Boyd RW (1992) *Nonlinear Optics*. Boston: Academic Press.

Chemla DS and Zyss J (1987) *Nonlinear Optical Properties of Organic Molecules and Crystals*, Vols 1 and 2. London: Academic Press.

Clays K, Persoons A and De Maeyer L (1993) Hyper-Rayleigh scattering in solution. Modern nonlinear optics, part 3. *Advances in Chemical Physics* 85: 455–498.

Flytzanis C (1975) Theory of nonlinear susceptibilities. In: Rabin H and Tang CL (eds) *Quantum Electronics*, Vol. I, Nonlinear Optics, part A. New York: Academic Press.

Lalanne JR, Ducasse A and Kielich S (1996) *Laser–Molecule Interaction*. New York: Wiley.

Levenson MD (1982) *Introduction to Nonlinear Laser Spectroscopy*. New York: Academic Press.

Shen YR (1984) *The Principles of Nonlinear Optics*. New York: Wiley.

Tang CL and Cheng LK (1995) *Fundamentals of Optical Parametric Processes and Oscillators*. Amsterdam: Harwood Academic.

Wagnière GH (1993) *Linear and Nonlinear Optical Properties of Molecules*. Basel: Verlag HCA, VCH.

Yariv A (1975) *Quantum Electronics*. New York: Wiley.

Zel'dovich BY, Pilipetsky NF and Shkunov VV (1985) *Principles of Phase Conjugation*. Berlin: Springer-Verlag.

# Nonlinear Raman Spectroscopy, Applications

**W Kiefer**, Universität Wurzburg, Germany

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

Linear, spontaneous Raman spectroscopy is a powerful tool for structural analysis of materials in the gaseous, liquid or solid state. Its scattering cross-section can be increased considerably by resonance excitation, i.e. irradiation in spectral regions where there is strong absorption or by applying surface enhanced methods like SERS (surface enhanced Raman scattering). Also, the scattering volume, determined by the dimensions of a focused laser beam, can be as small as a few $\mu m^2$ if a microscope is incorporated in a Raman spectrometer. There are, however, cases where ordinary Raman spectroscopy has limitations in allowing the derivation of the desired information. For example, particular vibrational modes of specific symmetry are neither allowed in linear Raman scattering nor in infrared absorption, but their vibrational bands show up in what is called a hyper-Raman spectrum, because there is a nonvanishing contribution from the nonlinear part of the induced dipole moment. Also, fluorescence simultaneously excited with visible laser light, may obscure the Raman scattered light. This can often be overcome by near-infrared laser excitation. Another way is to apply nonlinear coherent Raman techniques like CARS (coherent anti-Stokes Raman spectroscopy). In general, nonlinear optical properties of materials can only be obtained using nonlinear optical methods. One of the major advantages of nonlinear coherent Raman spectroscopy is its possible high resolution of up to three orders of magnitude better than its linear counterpart. In addition, these methods allow spectral information to be obtained from scattering systems which produce a high light background like flames, combustion areas, etc. In recent years there has been a dramatic development in time-resolved linear and nonlinear Raman spectroscopy due to the availability of commercial pico- and femtosecond lasers which allows direct insight into the dynamics of molecules in their ground or excited electronic state. After a short description of the various nonlinear Raman techniques, typical applications will be given for these methods.

## A short description of nonlinear Raman techniques

Spontaneous nonlinear as well as coherent nonlinear Raman methods are considered here. These are based on the contributions of the nonlinear part of the induced dipole moment (spontaneous effects) or the induced polarization (coherent effects) to the intensity of the frequency shifted light. In the first case, the Raman signal is generated in a spontaneous, incoherent but nonlinear optical process, whereas in the second case the Raman information is contained in a coherent laser beam whereby the nonlinear polarization acts as a coherent light source.

## Hyper-Raman effect

Generally, the induced dipole moment $\boldsymbol{p}$ in a molecular system is written as

$$\boldsymbol{p} = \alpha \boldsymbol{E} + \frac{1}{2}\beta \boldsymbol{EE} + \frac{1}{6}\gamma \boldsymbol{EEE} + \ldots \qquad [1]$$

where $\alpha$ is the polarizability, $\beta$ the hyperpolarizability and $\gamma$ the second hyperpolarizability. $\boldsymbol{E}$ is the incident electric field. The nonlinear terms in Equation [1] are usually small compared to the linear term which gives rise to normal, linear Raman scattering. However, when the electric field is sufficiently large, as is the case when a high-powered laser is focused on the sample, contributions from the second term in Equation [1] are sufficiently intense to be detected. This scattering is at an angular frequency $2\omega_L \pm \omega_R$, where $\omega_L$ is the angular frequency of the exciting laser beam and $-\omega_R$ and $+\omega_R$ are the Stokes and anti-Stokes hyper-Raman displacements, respectively. Scattering at $2\omega_L \pm \omega_R$ is called hyper-Raman scattering.

The hyper-Raman effect is a three-photon process involving two virtual states of the scattering system. The level scheme for Stokes hyper-Raman scattering is presented in **Figure 1**.

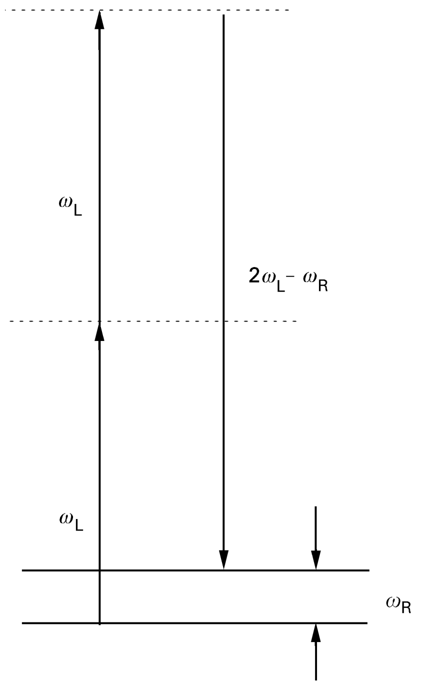



**Figure 1** Schematic level diagram for Stokes and hyper Raman scattering.

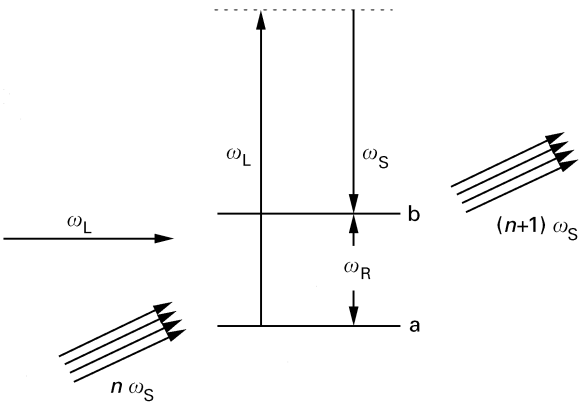



**Figure 2** Schematic diagram for stimulated Raman scattering as a quantum process.

The importance of the hyper-Raman effect as a spectroscopic tool results from its symmetry selection rules. It turns out that all infrared active modes of the scattering system are also hyper-Raman active. In addition, the hyper-Raman effect allows the observation of 'silent' modes, which are accessible neither by infrared nor by linear Raman spectroscopy.

## Stimulated Raman effect

The stimulated Raman process is schematically represented in **Figure 2**. A light wave at angular frequency $\omega_S$ is incident on the material system simultaneously with a light wave at angular frequency $\omega_L$. While the incident light beam loses a quantum ($\hbar\omega_L$) and the material system is excited by a quantum $\hbar\omega_R = \hbar(\omega_L - \omega_S)$, a quantum $\hbar\omega_S$ is added to the wave at angular frequency $\omega_S$, which consequently becomes amplified. It can be shown theoretically that a polarization at Stokes angular frequency $\omega_S$ is generated via the third-order nonlinear susceptibility $\chi^{(3)}$. Including a degeneracy factor, the polarization oscillating at angular frequency $\omega_S$ is given by Berger and co-workers (1992):

$$\boldsymbol{P}^{(3)}(\omega_S) = \frac{6}{4}\varepsilon_0\chi^{(3)}(-\omega_S;\ \omega_L, -\omega_L,\ \omega_S) \times |\boldsymbol{E}(\omega_L)|^2\,\boldsymbol{E}(\omega_S) \qquad [2]$$

where $\varepsilon_0$ is the permittivity constant of vacuum.

Optimum gain for this effect is found at the centre of the Raman line where $\omega_R = \omega_L - \omega_S$. There, the gain constant for stimulated Raman scattering at Stokes frequency is given by

$$g_S \propto \left(\frac{d\sigma}{d\Omega}\right)\cdot\frac{1}{\Gamma} \qquad [3]$$

where ($d\sigma/d\Omega$) is the differential Raman cross-section and $\Gamma$ represents the line width of the molecular transition ($\omega_R$). From Equation [3] we immediately recognize that in stimulated Raman scattering processes where only *one* input laser field with frequency $\omega_L$ is employed a coherent Stokes wave is generated for those Raman modes which have the highest ratio between differential Raman cross-section and line width $\Gamma$. The distinctive feature of stimulated Raman scattering is that an assemblage of coherently driven molecular vibrations provides the means of coupling the two light waves at angular frequencies $\omega_L$ and $\omega_S$ by modulating the nonlinear susceptibility.

## Nonlinear Raman spectroscopies based on third-order susceptibilities

From the discussion on stimulated Raman scattering it is clear that during this nonlinear process coherently driven molecular vibrations are generated. In what is usually called the stimulated Raman effect only one input field ($\omega_L$) is used for this type of excitation. We have seen that only particular Raman modes, i.e. those with highest gain factors, give rise to stimulated Stokes emission. Thus, for molecular spectroscopy in which we are interested in determining *all* Raman active modes, excitation with one strong laser field would not serve the purpose, although it would

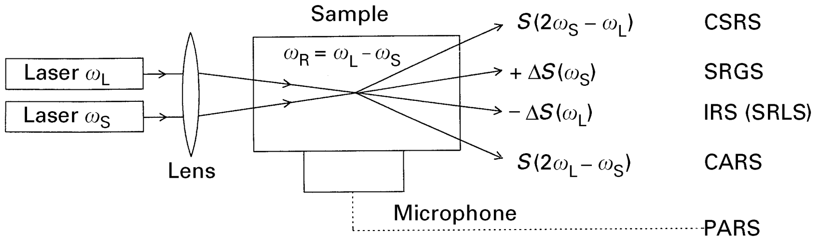



**Figure 3** Schematic diagram for a few techniques in nonlinear (coherent) Raman spectroscopy (CSRS: Coherent Stokes Raman Spectroscopy; SRGS: Stimulated Raman Gain Spectroscopy; IRS: Inverse Raman Spectroscopy (= SRLS: Stimulated Raman Loss Spectroscopy); CARS: Coherent anti-Stokes Raman Spectroscopy; PARS: Photoacoustic Raman Spectroscopy).

provide very high signals in the form of a coherent beam, but unfortunately, only at one particular vibrational frequency. However, the advantages of stimulated Raman scattering, being high signal strength and coherent radiation, can be fully exploited by a very simple modification of the type of excitation. The trick is simply to provide the molecular system with an intense external Stokes field by using a second laser beam at Stokes angular frequency $\omega_S$ instead of having initially the Stokes field produced in the molecular system by conversion of energy from the pump field. Thus, by keeping one of the two lasers, e.g. the laser beam at Stokes angular frequency $\omega_S$ tunable, one is now able to excite selectively coherent molecular vibrations at any desired angular frequency $\omega_R$ assuming the transitions are Raman allowed. A variety of nonlinear Raman techniques based on this idea have been developed, which combine the wide spectroscopic potentials of spontaneous Raman spectroscopy and the high efficiency of scattering, strong excitation and phasing of molecular vibrations in a macroscopic volume of substance, that are the features inherent to stimulated Raman scattering.

The following acronyms of some of these nonlinear coherent Raman techniques have been widely used: CARS, CSRS (coherent Stokes Raman spectroscopy), PARS (photoacoustic Raman spectroscopy), RIKE (Raman induced Kerr effect), SRGS (stimulated Raman gain spectroscopy), IRS (inverse Raman scattering) also called SRLS (stimulated Raman loss spectroscopy). A schematic diagram of these methods is illustrated in **Figure 3**. The common physical aspect is the excitation of Raman active molecular vibrations and/or rotations in the field of two laser beams with angular frequencies $\omega_L$ and $\omega_S$ in such a way that their difference corresponds to the angular frequency of the molecular vibration $\omega_R$ ($= \omega_L - \omega_S$). The strong coupling between the generated coherent molecular vibrations with the input laser fields via the third-order nonlinear susceptibility $\chi^{(3)}$ opens the possibility for various techniques.

The most powerful of these methods is CARS since a new coherent, laser-like signal is generated. Its direction is determined by the phase-matching condition

$$\boldsymbol{k}_{AS} = 2\boldsymbol{k}_L - \boldsymbol{k}_S \qquad [4]$$

where $\boldsymbol{k}_{AS}$, $\boldsymbol{k}_L$ and $\boldsymbol{k}_S$ are the wave vectors of the anti-Stokes signal, pump and Stokes laser, respectively. The laser-like anti-Stokes signal is therefore scattered in one direction, which lies in the plane given by the two laser directions $\boldsymbol{k}_L$ and $\boldsymbol{k}_S$ and which is determined by the momentum vector diagram shown in **Figure 4**. Therefore, CARS is simply performed by measuring the signal $S(2\omega_L - \omega_S) = S(\omega_L + \omega_R)$, which is a coherent beam emitted in a certain direction. These coherent signals with anti-Stokes frequencies are generated each time the frequency difference of the input laser fields matches the molecular frequency of a Raman active transition.

The mixing of the two laser fields can also produce radiation on the Stokes side of the $\omega_S$-laser. The direction of this coherent Stokes Raman scattering (CSRS) signal is again determined by a corresponding momentum conservation diagram, which leads to a different direction (see **Figure** 3), labelled by $S(2\omega_S - \omega_L)$. Since the CSRS signal is in principle weaker than the CARS signal, and because the former may be overlapped by fluorescence, the CARS technique is more frequently used.

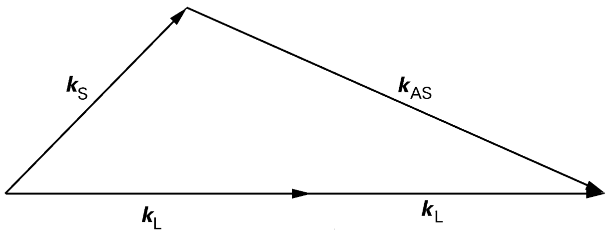



**Figure 4** Momentum conservation for CARS (representation of Equation 4).

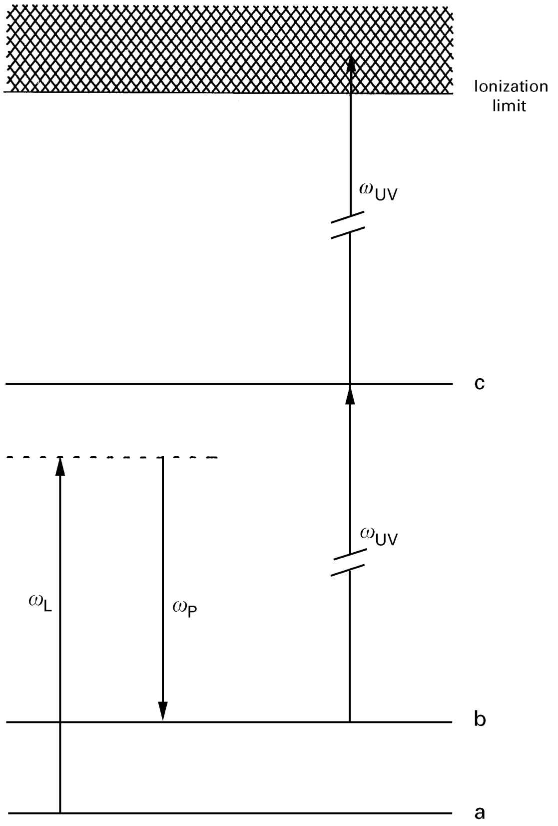



**Figure 5** Energy-level diagram illustrating the two excitation steps of Ionization Detected Stimulated Raman Spectroscopy (IDSRS).

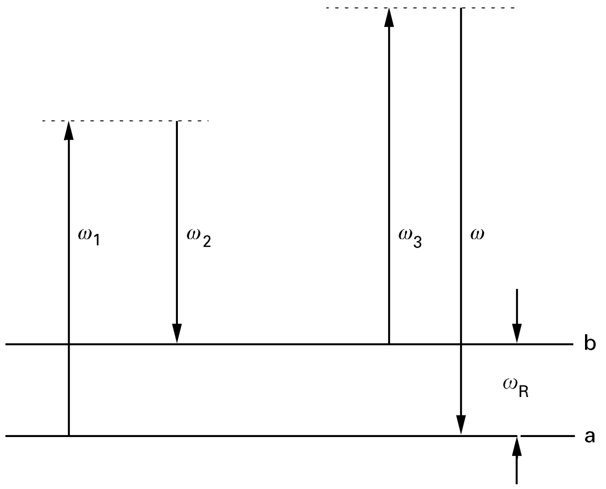



**Figure 6** Schematic diagram representing the four-wave mixing process: a polarization is generated at the frequency $\omega = \omega_1 - \omega_2 + \omega_3$.

The interaction of the electric fields of the two $\omega_L$ and $\omega_S$ lasers with the coherent molecular vibrations yields also a gain or a loss in the power of the lasers. The method where the gain at the Stokes frequency (labelled in **Figure 3** by $+\Delta S(\omega_S)$) is measured is generally referred to as 'stimulated Raman gain spectroscopy' whereas the 'inverse Raman scattering' (IRS) is the terminology commonly used to designate the induced loss at the pump laser frequency (**Figure 3**, $-\Delta S(\omega_L)$). IRS is also often called stimulated Raman loss spectroscopy (SRLS). In order to get full Raman information of the medium, it is necessary to tune the frequency difference $\omega_L - \omega_S$; then, successively all Raman-active vibrations (or rotations, or rotation–vibrations) will be excited and a complete nonlinear Raman spectrum is then obtained either by measuring newly generated signals (CARS, CSRS) or the gain (SRGS) or loss (SRLS) of the pump or the Stokes laser, respectively. In what is called broadband CARS, the Stokes ($\omega_S$) is spectrally broad, while the pump laser ($\omega_L$) is kept spectrally narrow, resulting in the simultaneous generation of a broad CARS spectrum. For the detection of the latter a spectrometer together with a CCD camera is needed.

In photoacoustic Raman spectroscopy (PARS), due to the interaction of the two input laser fields ($\omega_L$, $\omega_S$) a population of a particular energy level ($\omega_R$) of the sample is achieved. As the vibrationally (or rotationally) excited molecules relax by means of collisions, a pressure wave is generated in the sample and this acoustic signal is detected by a sensitive microphone.

A technique which combines the high sensitivity of resonant laser ionization methods with the advantages of nonlinear coherent Raman spectroscopy is called IDSRS (ionization detected stimulated Raman spectroscopy). The excitation process, illustrated in **Figure 5**, can be briefly described as a two-step photoexcitation process followed by ion/electron detection. In the first step two intense narrow-band lasers ($\omega_L$, $\omega_S$) are used to vibrationally excite the molecule via the stimulated Raman process. The excited molecules are then selectively ionized in a second step via a two- or multiphoton process. If there are intermediate resonant states involved (as state c in **Figure 5**), the method is called REMPI (resonance enhanced multi-photon ionization)-detected stimulated Raman spectroscopy. The technique allows an increase in sensitivity of over three orders of magnitude because ions can be detected with much higher sensitivity than photons.

The nonlinear Raman techniques discussed above are special cases of a general four-wave mixing process, which is schematically illustrated in **Figure 6**. Here, three independent fields with angular frequencies $\omega_1$, $\omega_2$ and $\omega_3$ may be incident upon the matter. A fourth field, which is phase coherent relative to the input fields, is then generated at angular frequency $\omega = \omega_1 - \omega_2 + \omega_3$. When the angular frequency

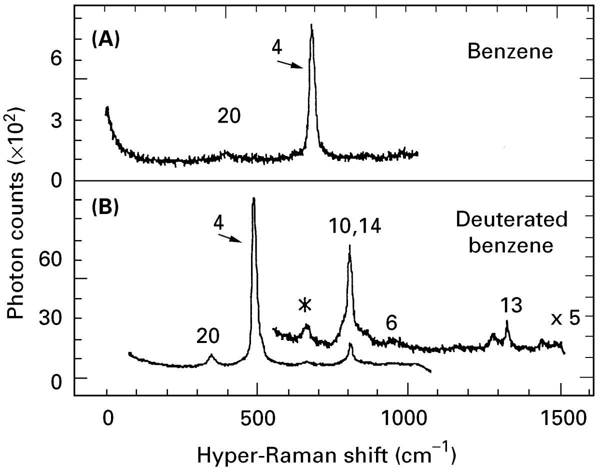



**Figure 7** Hyper-Raman spectra of $C_6H_6$ excited with a Nd:YAG laser ($\lambda_0$ = 1.064 nm) Q-switched at 1 kHz (A) and of $C_6D_6$ in the lower spectrum with the laser Q-switched at 6 kHz (B). Reproduced by permission of Elsevier Science from Acker WP, Leach DH and Chang RK (1989) Stokes and anti-Stokes hyper Raman scattering from benzene, deuterated benzene, and carbon tetrachloride. *Chemical Physics Letters* **155**: 491–495.

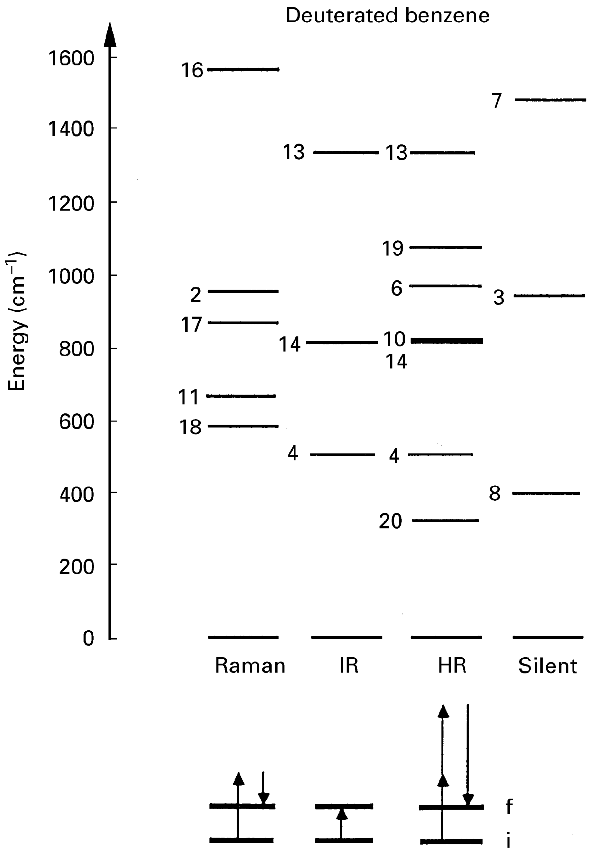



**Figure 8** Vibrational energy levels of $C_6D_6$ (energy < 1600 cm$^{-1}$) grouped by their activity from the ground state, i.e. Raman, IR, or hyper-Raman (HR). Modes which are not active in Raman, IR, or hyper-Raman are grouped. Reproduced by permission of Elsevier Science from Acker WP, Leach DH and Chang RK (1989) Stokes and anti-Stokes hyper Raman scattering from benzene, deuterated benzene, and carbon tetrachloride. *Chemical Physics Letters* **155**: 491–495.

difference $\omega_1 - \omega_2$ equals the Raman excitation angular frequency $\omega_R$, the signal wave at $\omega$ is enhanced, indicating a Raman resonance. For example, a CARS signal is Raman resonantly generated when $\omega_1 = \omega_3 = \omega_L$, $\omega_2 = \omega_S$ and $\omega_L - \omega_S = \omega_R$.

## Applications

### Applications of spontaneous nonlinear Raman spectroscopy (Hyper-Raman scattering)

Since its discovery in 1965, hyper-Raman spectra have been observed in all three states of aggregate. However, reasonable signal-to-noise ratios could only be obtained for a convenient measurement time after the development of fast pulsed, high power lasers and highly sensitive detectors (multichannel diode arrays or charge-coupled devices (CCDs)). Before that time only a few gases had been studied which included ethane, ethene and methane. Only vibrational spectra of modest resolution have been obtained in these studies. A number of group IV tetrahalides have been studied in the liquid phase. Other liquids whose Raman spectra have been reported include water and tetra-chloroethene. Probably most hyper-Raman work was performed in crystals: $NH_4Cl$, $NH_4Br$, calcite, $NaNO_2$, $NaNO_3$, $LiNbO_3$, $SrTiO_3$, caesium and rubidium halides, rutile, $PbI_2$, CuBr, diamond and quartz. Stimulated hyper-Raman scattering has been observed from sodium vapour, resonance hyper-Raman scattering from CdS and surface enhanced hyper-Raman scattering from $SO_3^{2-}$ ions adsorbed on silver powder.

Technological advances, i.e. CW pumped acoustooptically Q-switched Nd:YAG lasers with repetition rates of up to 5 kHz combined with multichannel detection systems have increased the ease of obtaining hyper-Raman signals. By making use of this advanced technology, hyper-Raman spectra of benzene and pyridine could be obtained. Spectra from benzene, deuterated benzene and carbon tetrachloride have been measured with high signal-to-noise ratios. As examples, we show in **Figure 7** the hyper-Raman spectra of benzene and deuterated benzene. The observed hyper-Raman bands are labelled by numbers (4, 6, 10, 13, 14, 20) and correspond to the $\nu_4$ ($A_{2u}$), $\nu_6$ ($B_{1u}$), $\nu_{10}$ ($B_{2u}$), $\nu_{13}$ ($E_{1u}$), $\nu_{14}$ ($E_{1u}$) and $\nu_{20}$ ($E_{2u}$) vibrations of $C_6D_6$, respectively. **Figure 8** shows the low-lying

vibrational energy levels for $C_6D_6$ grouped by their activity involving transition from the ground state, i.e. Raman, IR, hyper-Raman (HR) and none of the above which are grouped as silent. Note that in the third column four modes with energy below 1500 cm$^{-1}$ are only hyper-Raman active and three modes of symmetry $A_{1u}$ and $E_{1u}$ are both IR and hyper-Raman active. Except for the $\nu_{19}$ ($E_{2u}$) mode all hyper-Raman active modes can be found in the spectrum displayed in **Figure 7**. The modes of class $B_{2g}$ are active in the second hyper-Raman effect which is controlled by the fourth rank second hyperpolarizability tensor $\gamma$.

Hyper-Raman scattering under resonance conditions for molecules in the gas phase was observed in 1993. High quality rotational resonance hyper-Raman spectra of $NH_3$ were obtained using blue incident radiation at half the $\tilde{X} \rightarrow \tilde{A}$ transition energy. Also hyper-Raman scattering of methyl iodide for excitation with a laser line which has been tuned through the two-photon resonance with the absorption band of a predissociative Rydberg transition in the VUV (175–183 nm) was reported. Similarly to linear resonance Raman scattering, overtones or combination bands can also be observed for resonantly excited hyper-Raman sc attering. An example is given in **Figure 9** where several higher order modes of methyl iodide can be observed.

The use of CW pumped acoustooptically Q-switched Nd:YAG lasers (repetition rates of 5 kHz), synchronously gated photomultiplier tubes, and synchronously gated two-dimensional single-photon counting detectors has improved the signal-to-noise ratio of hyper-Raman spectra. Considerable further improvements have been obtained with mode-locked pulses (at 82 MHz) from a Nd:YAG laser to observe the surface-enhanced hyper-Raman signal from pyridine adsorbed on silver. In these studies, hyper-Raman signals were observed with intensities close to spontaneous Raman scattering. It was shown that surface enhanced hyper-Raman scattering (SEHRS) has become a useful spectroscopic technique. In view of the recent advances in laser and detector technology, significant improvement in SEHRS sensitivity will come rapidly from the use of an intensified CCD camera for hyper-Raman signal detection and the use of a continuously tunable mode-locked Ti:sapphire laser as the excitation source.

## Applications of coherent anti-Stokes Raman spectroscopy (CARS)

The advantages of CARS, i.e. high signal strength, very high spectral or temporal resolution, discrimination against fluorescence, etc., have opened new ways to study molecular structure. In the following

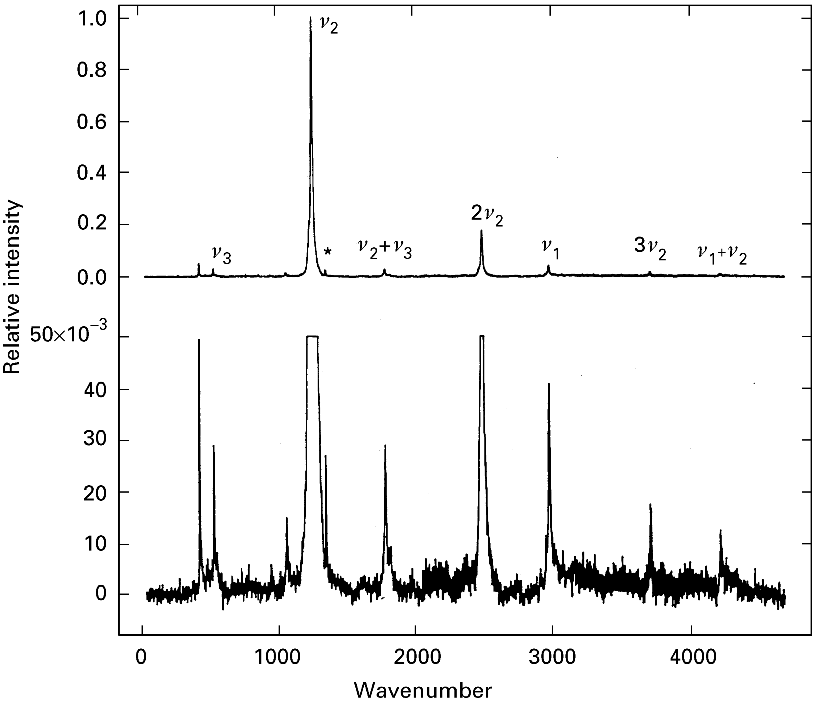



**Figure 9** Resonance hyper Raman spectrum of $CH_3I$ vapour excited at 365.95 nm. Reproduced by permission of Elsevier Science from Campbell DJ and Ziegler LD (1993) Resonance hyper-Raman scattering in the VUV. Femtosecond dynamics of the predissociated C state of methyl iodide. *Chemical Physics Letters* **201**: 159–165.

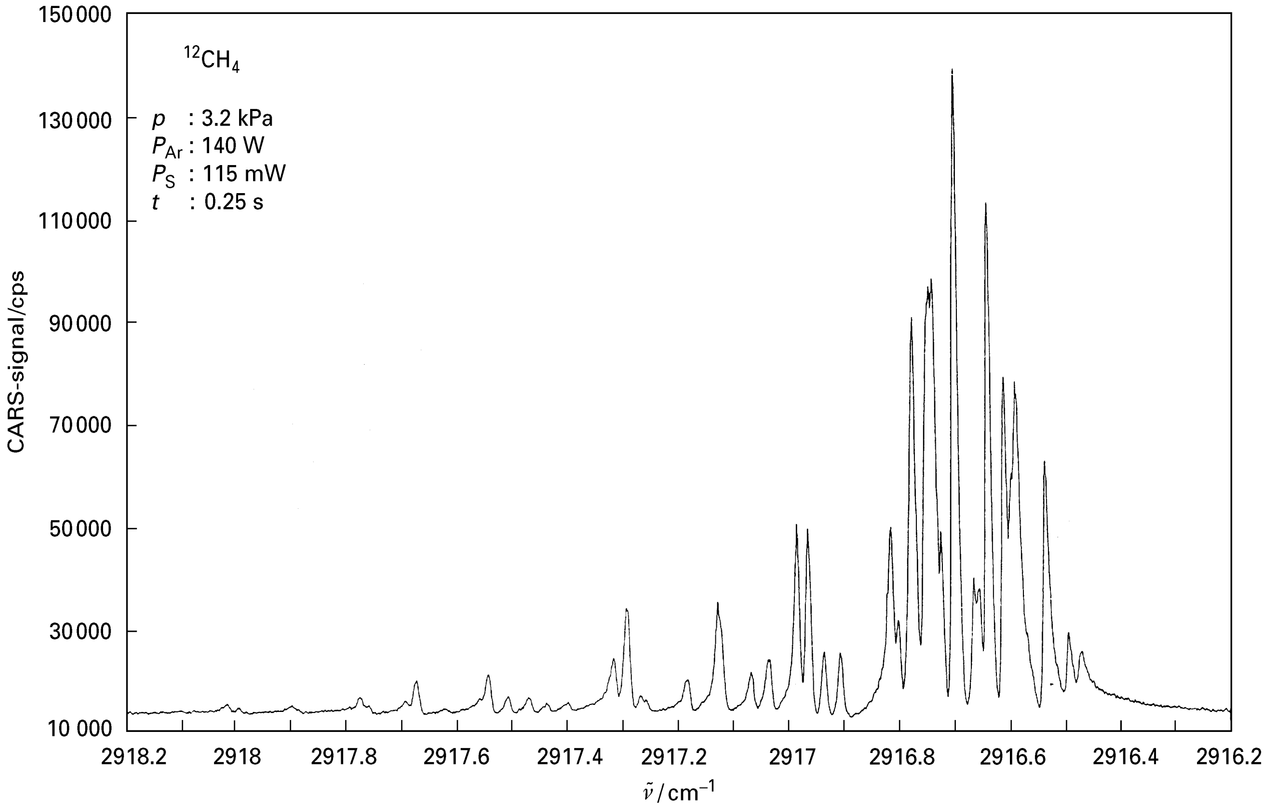



**Figure 10** High-resolution CARS spectrum of $\nu_1$ band of methane. Reproduced by permission of VCH Verlag from Schrötter HW (1995) Raman spectra of gases. In: Schrader B (ed.) *Infrared and Raman Spectroscopy*, pp 277–297. Weinheim: VCH Verlag.

some selected examples will be given to demonstrate the capability of this nonlinear coherent technique.

The 1980s and 1990s have seen a remarkable growth in the number of CARS applications to molecular and physical properties, particularly in the field of gas-phase systems. The latter are challenging because of low sample densities and the narrow transition line widths make them attractive for high resolution studies. Gas-phase CARS spectra have been obtained so far at pressures down to a few pascal, at temperatures ranging from a few K to 3600 K, and at a resolution better than about $10^{-3}$ cm$^{-1}$.

Mainly, the Q-branches of simple molecules, like di-, tri-, and four-atomic as well as spherical $XY_4$ top molecules have been studied. As an example **Figure 10** shows the Q-branch of methane. The complicated rotational structure seen there has been resolved by applying this powerful nonlinear Raman technique. This very high resolution of the order of $10^{-3}$ cm$^{-1}$ allows us to study in detail collisional effects, which is of particular importance as a basis for the determination of temperatures and pressures.

One very active area of the gas-phase CARS technique has been the remote sensing of temperature and species in hostile environments such as gas discharges, plasmas, flames, internal combustion engines, and the exhaust from jet engines. The high signal intensity and the excellent temporal and spectral resolution of CARS make it a favourite method for such studies. For example, CARS has been used to measure state populations and changes in discharges of $H_2$, $N_2$ and $O_2$ at pressure ranging from a few kPa down to 0.6 Pa. Also, gas-phase CARS can be employed to monitor $SiH_2$ intermediates in their investigations of silane plasmas commonly used in amorphous silicon deposition processes. Many laboratories are engaged in combustion research. Combustion studies in engines include thermometry in a diesel engine, in a production petrol engine, and thermometry and species measurements in a fully afterburning jet engine. Investigations on turbulent and sooting flames were performed.

Temperature information from CARS spectra derives from spectral shapes either of the Q-branches or of the pure rotational CARS spectra of the molecular constituents. In combustion research it is most common to perform thermometry from nitrogen since it is the dominant constituent and present everywhere in large concentration despite the extent

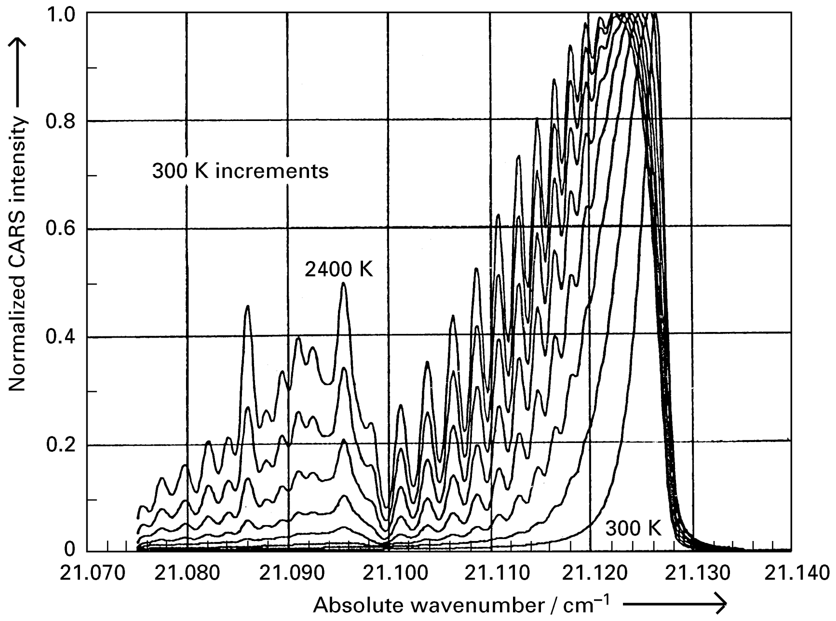



**Figure 11** Temperature dependence of $N_2$ CARS spectrum from 300 to 2400 K in 300 K increments (Hall and Eckbreth, 1984).

of chemical reaction. The Q-branch of nitrogen changes its shape due to the increased contribution of higher rotational levels which become more populated when the temperature increases. **Figure 11** displays a calculated temperature dependence of the $N_2$ CARS spectrum for experimental parameters typically used in CARS thermometry. Note that the wavenumber scale corresponds to the absolute wavenumber value for the ~2320 cm$^{-1}$ Q-branch of $N_2$ when excited with the freqency doubled Nd:YAG laser at 532 nm (≅ 18 796 cm$^{-1}$), i.e. $\tilde{\nu}_{AS}$ = 18 796 + 2320 = 21126 cm$^{-1}$. The bands lower than about 21 100 cm$^{-1}$ are due to the rotational structure of the first vibrational hot band.

For the case that there are not too many constituents in the gas under investigation, the use of the pure rotational CARS technique may be superior to vibrational CARS thermometry since the spectra are easily resolvable (for $N_2$ the adjacent rotational peaks have a spacing of approximately 8 cm$^{-1}$) compared with the congestion of the rotational lines in the vibrational bands of the Q-branch spectra (see **Figure 11**). An experimental comparison of rotational and vibrational CARS techniques, under similar conditions has been made that demonstrates that rotational CARS may be viable for flame-temperature measurements up to 2000 K. Of course, the pure rotational approach cannot be applied for spherical molecules which have no pure rotational CARS spectrum. An elegant method, using Fourier analysis based on the periodicity of pure rotational CARS spectra has been introduced recently.

In addition to temperature measurements the gas-phase CARS technique also provides information on the fluctuating properties occurring for instance in turbulent combustion systems. However, concentration measurements are more difficult to perform than temperature ones because the absolute intensity is required, while temperature measurements are only based on the shape of the spectrum. Simultaneous information on the relative concentrations between several species are easier to obtain.

Quantitative gas-phase CARS spectroscopy has also been applied to probing species in a laboratory chemical reactor and to temperature measurements inside incandescent lamps. Another interesting area is that of CARS applied to free expansion jets. The key benefits of this technique are the spectral simplification of cold molecules and the increased concentrations of small van der Waals complexes obtained under the non-equilibrium jet conditions.

CARS is also used for the study of samples in the condensed phase. The major experimental advantage of CARS (and most nonlinear coherent Raman techniques) is the large signal produced. In a typical CARS experiment in a liquid or a solid, the applied laser power of the pump and Stokes laser ($10^4-10^5$ W) generates an output power of up to 1 W, while conventional Raman scattering would give a collected signal power of ~$10^{-4}$ W with the same lasers. Since the CARS output is directional, the collection angle can be five orders of magnitude smaller than that needed in spontaneous scattering. Taken together, these two factors imply that CARS is nine orders of

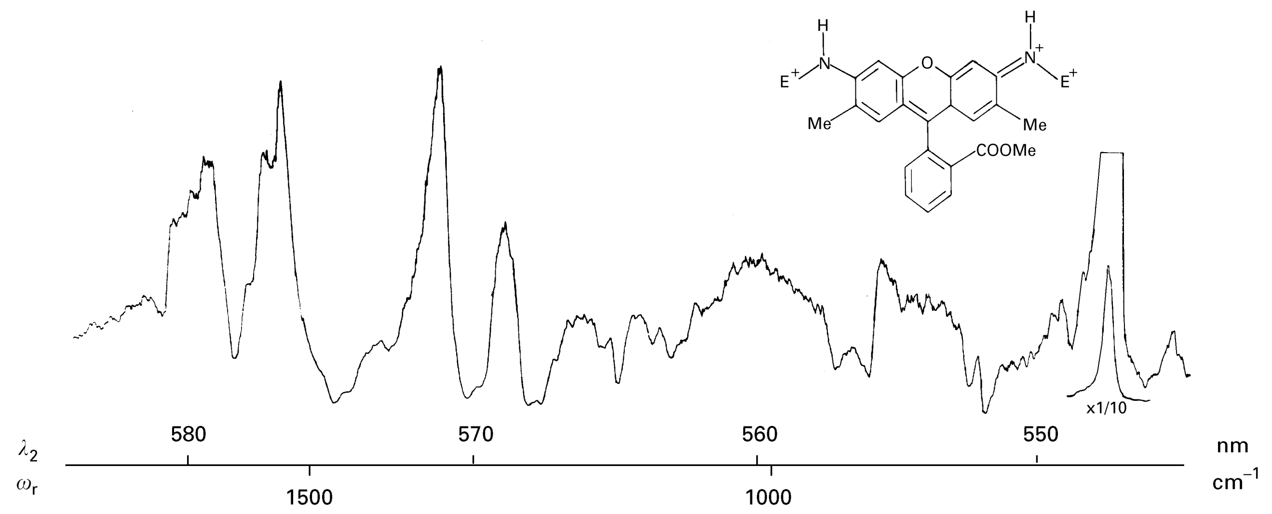



**Figure 12** CARS spectrum of rhodamine 6G in solution (Carreira and Horovitz, 1982).

magnitude less sensitive to sample fluorescence than spontaneous scattering. The advantage is actually even greater since the CARS signal is at higher frequency than any of the input laser frequencies.

While it is nearly impossible to obtain Raman spectra of highly luminescent materials, e.g. dye solution, it was the CARS technique which first overcame this problem because of the reasons mentioned above. As an example the CARS spectrum of a rhodamine 6 G (R6G) water solution is displayed in **Figure 12**. The vibrational modes of the strongly luminescent R6G molecule can be seen. It should be mentioned at this point that by long wavelength excitation, i.e. for example excitation with the 1.064 μm line of a CW Nd:YAG laser or by making use of the SERS effect luminescence-free linear Raman spectra can be obtained. Since the latter methods are in any case much easier to perform than CARS or other nonlinear Raman techniques, they are to be preferred. However, if one is interested in obtaining structural as well as electronic properties of absorbing materials through resonance excitation, there are many cases where linear resonance Raman spectroscopy is limited because of the mentioned strong luminescence.

On the other hand, many, particularly organic, substances show considerable third-order nonlinear susceptibilities $\chi^{(3)}$, as for example polyacetylenes, polydiacetylenes or chlorophyll. For such systems, resonance CARS spectroscopy is a suitable tool to obtain resonance Raman information via the anti-Stokes, coherent spectroscopic method. However, in performing resonance CARS spectroscopy in solids one must realize that this technique results in a fairly complicated arrangement between the sample and the coherent beams. First, the phase-matching conditions (Eqn [4], **Figure 4**) have to be obeyed, where the momentum vectors depend also on the refractive index of the solid media. Therefore a continuous adjustment of the crossing angle between the incident laser beams ($\boldsymbol{k}_L$, $\boldsymbol{k}_S$) as well as of the angle between the pump laser beam and the CARS beam ($\boldsymbol{k}_L$, $\boldsymbol{k}_{AS}$) is required during the scan of the CARS spectrum. Secondly, in order to excite particular phonons in the crystals, the difference between the pump and the Stokes beam wave vectors must coincide with the wave vector of the coherently excited phonon in the crystal ($\boldsymbol{k}_L - \boldsymbol{k}_S = \boldsymbol{k}_{phonon}$). Depending on the strength of absorption and sample thickness, CARS in solids is either performed in transmission or in reflection (backscattering CARS).

As an example of resonance CARS studies in solids, for which a linear resonance Raman study has been impossible to perform because of simultaneous strong luminescence, we considered here investigations on colour zones in substituted diacetylene crystals originating from partial polymerization.

For a long time it has been known that diacetylene monomer single crystals undergo, upon thermal annealing or exposure to high-energy radiation, topochemical solid-state polymerization. From this reaction, polymer chains are formed which have a substantial π-electron delocalization, forming a pseudo-one-dimensional electronic system. Colour zones occur in such crystals due to different chain lengths and CARS studies were performed on these zones in crystals with low polymer content, where the polymer chains were embedded in the monomer matrix. As mentioned, resonance Raman excitation within the strong absorption of the polymer chains, i.e. within the absorption of the colour zone, produced high luminescence levels which obscured the bands in linear Raman spectroscopy. In contrast,

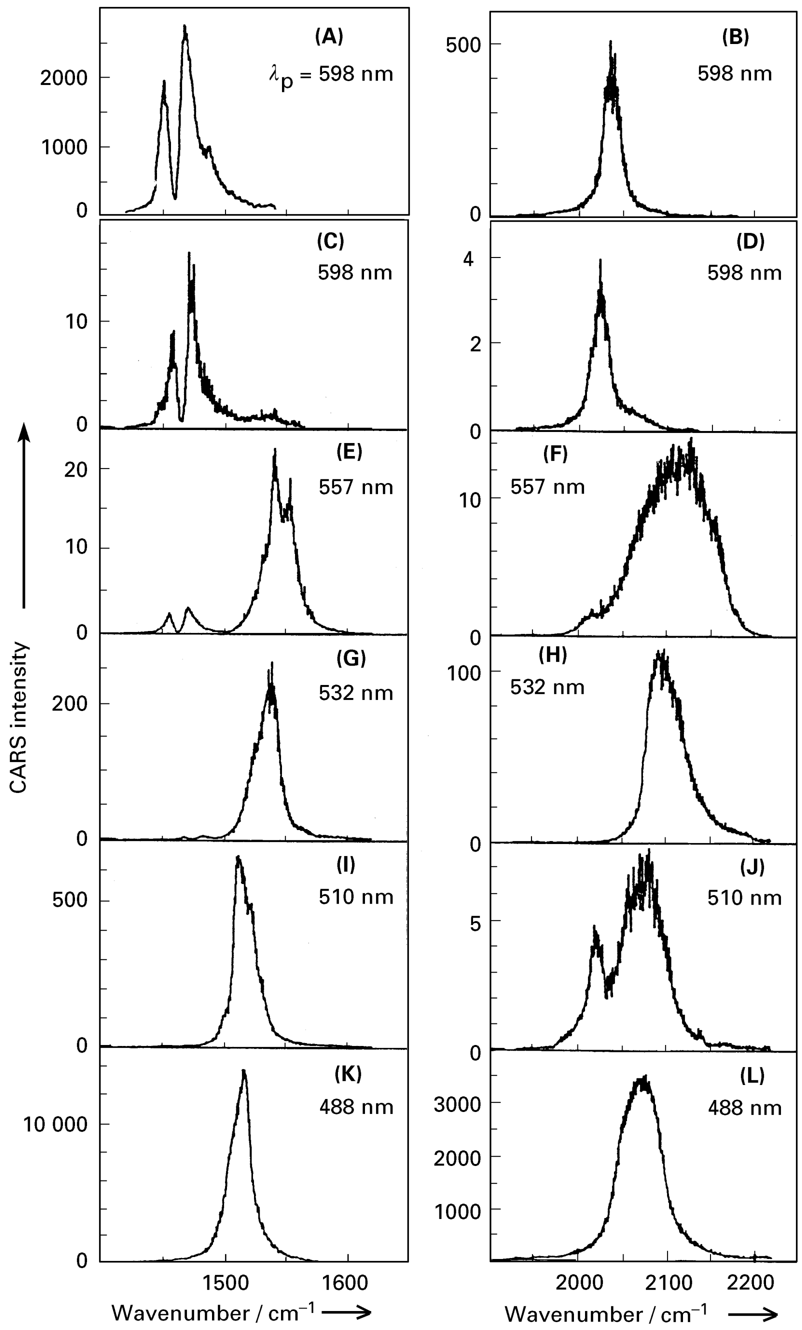



**Figure 13** Resonance CARS spectra of a substituted diacetylene single crystal (FBS-DA) at 10 K. The pump wavelength $\lambda_p$ used is labelled for each spectrum. (A) and (B) show CARS spectra of the P-colour zone, and (C)–(L) those for the Y-colour zone. Spectra on the left side correspond to the C=C stretching region, and those on the right side to the C≡C stretching region. For further details, see text. Reproduced by permission of John Wiley & Sons from Materny A and Kiefer W (1992) Resonance CARS spectroscopy on diacetylene single crystals. *Journal of Raman Spectroscopy* **23**: 99–106.

luminescence-free resonance CARS spectra can be obtained, as shown in **Figure 13** for the case of an FBS DA crystal at 10 K (FBS = 2,4-hexadiynylene-di-p-fluorobenzene sulfonate, DA = diacetylene). On the left and right panels of **Figure 13** CARS spectra are displayed for the region of the C=C and C≡C stretching region around 1500 cm$^{-1}$ and 2100 cm$^{-1}$, respectively. Spectra (A) and (B) are those of the P colour zone (P = principal) and (C) – (L) those of the Y-colour zone (Y = yellow). Note the very different CARS intensities as well as band shapes for the various excitation wavelengths of the pump laser ($\lambda_p$, which corresponds to $\omega_L$ of the CARS process as outlined above) which are due to different resonant

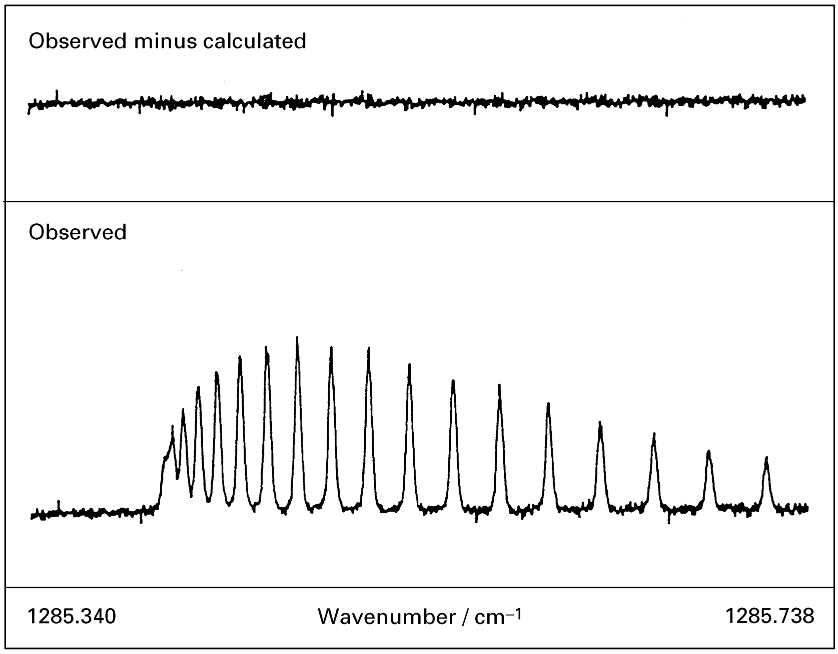



**Figure 14** High resolution multi-pass stimulated Raman gain spectrum (SRGS) of the Q-branch of the lower component of the Fermi resonance diad of $^{12}C^{16}O_2$ at a pressure of 200 Pa (1.5 torr). Reproduced by permission of John Wiley & Sons from Saint-Loup R, Lavorel B, Millot G, Wenger C and Berger H (1990) Enhancement of sensitivity in high-resolution stimulated Raman spectroscopy of gases. *Journal of Raman Spectroscopy* **21**: 77–83.

enhancements. Comparing spectrum (K) with (C), for example, shows in addition the very high dynamic range (at least four orders of magnitude) inherent in this type of spectroscopy. Analysing the CARS spectra together with the absorption spectra of several substituted DA crystals, one is able to derive important structural as well as electronic properties of this type of crystal.

It should be mentioned that there are some disadvantages of CARS: (i) an unavoidable electronic background nonlinearity that alters the line shape and can limit the detection sensitivity; (ii) a signal that scales as the square of the spontaneous scattering signal (and as the cube of the laser power), making the signals from weakly scattering samples difficult to detect; and (iii) the need to fulfil the phase matching requirements. While other techniques avoid these difficulties, CARS still remains the most popular coherent nonlinear technique.

### Applications of stimulated Raman gain and inverse Raman spectroscopy (SRGS, IRS)

The advantages of SRGS and IRS are that (in contrast to CARS) the signal is linearly proportional to the spontaneous Raman scattering cross-section (and to the product of the two laser intensities), and that the phase-matching condition is automatically fulfilled. The fact that the resolution of the nonlinear Raman techniques is limited only by the laser line widths gives the stimulated Raman techniques particular appeal under conditions where interference from background luminescence is problematic or in situations where very high resolution is required. The main disadvantage of these techniques, however, is that they are quite sensitive to laser noise. The latter requires high stability in laser power.

Due to complexity, only a few stimulated Raman gain and loss spectrometers with a main application in high resolution molecular spectroscopy have been built since the fundamental developments around 1978.

Here, we present an instructive example for each of the two techniques (SRGS, IRS) emphasizing the high resolution capability of these methods.

The Q-branches of numerous molecules, particularly of linear and spherical top molecules have been analysed by means of SRGS and IRS. As an example of a recent high resolution SRGS spectrum we show in **Figure 14** the spectrum of the Q-branch of the lower component of the Fermi resonance diad of $^{12}C^{16}O_2$ at 1285 cm$^{-1}$. The spectrum has been recorded at a pressure of 200 Pa (1.5 torr). The excellent agreement with a calculation assuming Voigt line profiles is demonstrated by the residual spectrum in the upper trace.

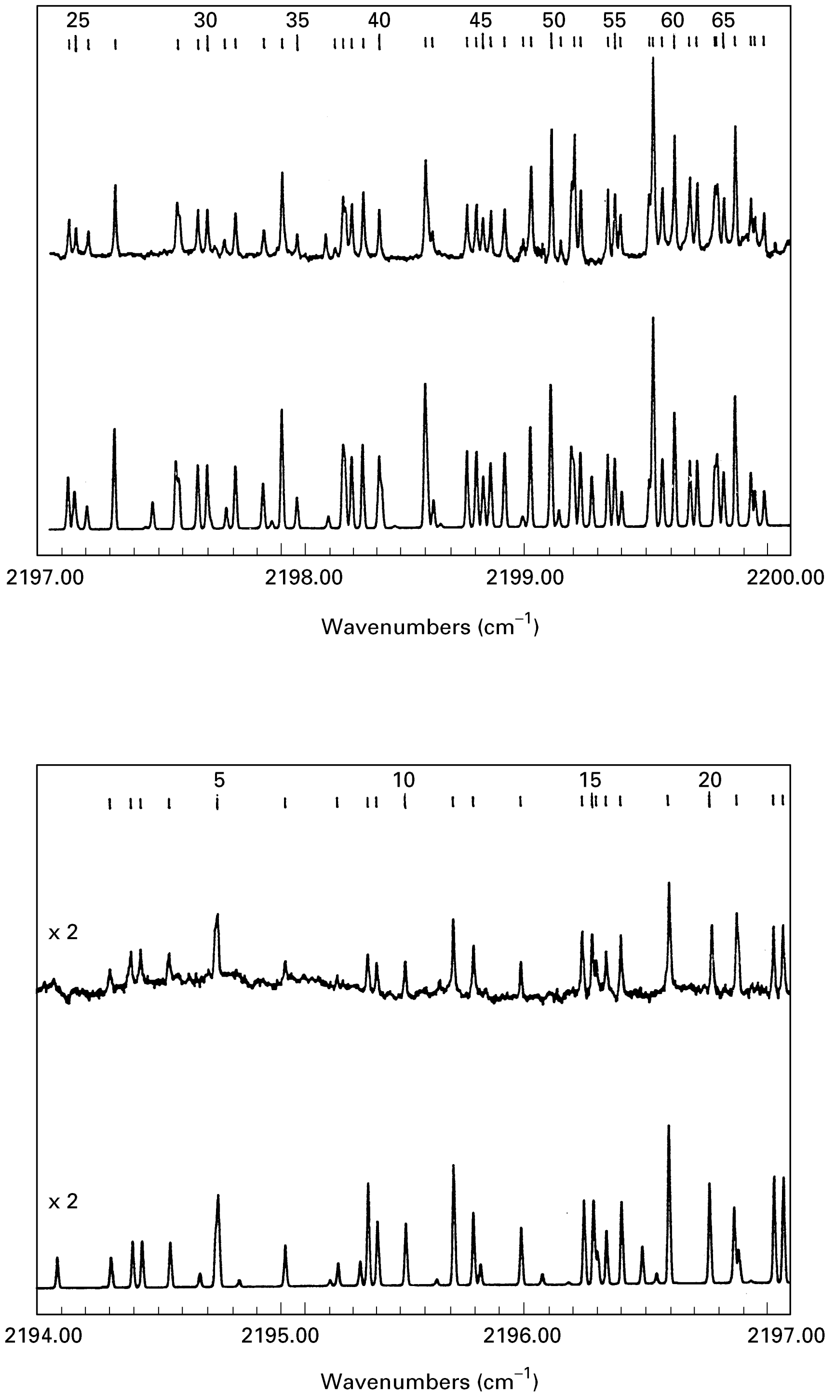



**Figure 15** High resolution inverse Raman Spectrum of the $\nu_2$ Q-branch of $CH_3D$ between 2194 and 2200 cm$^{-1}$. Upper traces : Observed, lower traces: calculated spectra. Reproduced by permission of John Wiley & Sons from Bermejo D, Santos J, Cancio P *et al* (1990) High-resolution quasicontinuous wave inverse Raman spectrometer. Spectrum of $CH_3D$ in the C-D stretching region. *Journal of Raman Spectroscopy* **21**: 197–201.

An example for high-resolution IRS is given in **Figure 15**, where the $\nu_2$ Q-branch of $CH_3D$ is displayed. This spectrum represents a Doppler-limited spectrum of the C–D stretching band. The authors were able to assign the observed transitions by performing a theoretical fit to the observed data which allowed them to refine some of the rotational–vibrational constants.

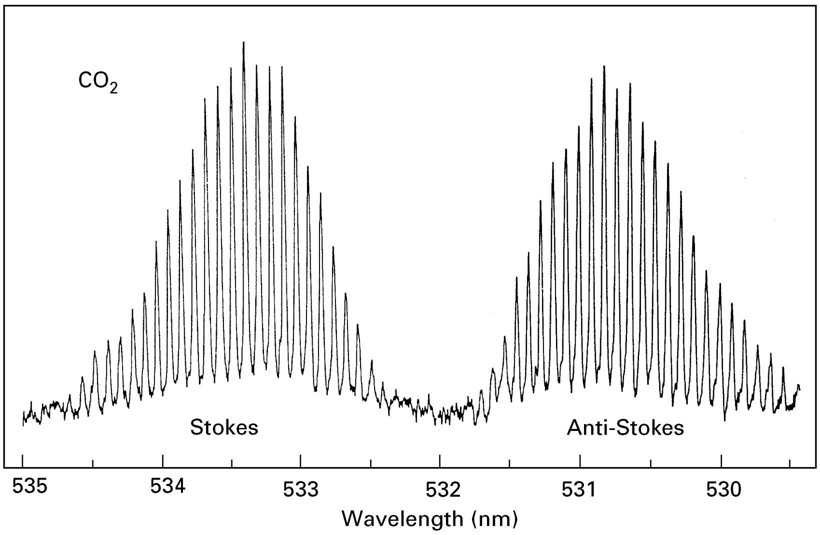



**Figure 16** The pure rotational photoacoustic Raman (PARS) spectrum of $CO_2$ gas at a pressure of 80 kPa (600 torr); pump laser wave length at 532 nm. Note the complete absence of any acoustical signal due to Rayleigh scattering (at 532 nm). Reproduced by permission of Academic Press from Barrett JJ (1981) Photoacoustic Raman Spectroscopy. In: Harvey AB (ed) *Chemical Applications of Nonlinear Raman Spectroscopy*, pp 89–169. New York: Academic Press.

## Applications of photoacoustic Raman spectroscopy (PARS)

As discussed above in photoacoustic Raman spectroscopy (PARS) the energy deposited in the sample by excitation of, for example, a vibration by the stimulated Raman process leads to pressure increases through relaxation to translational energy and can therefore be detected by a sensitive microphone.

When the pump ($\omega_L$) and Stokes ($\omega_S$) beams have only small frequency differences, as can be achieved, for example, by using a frequency-doubled Nd:YAG laser for $\omega_L$ and a dye laser with amplifier pumped by the third harmonic of the same Nd:YAG laser for $\omega_S$, the recording of pure rotational PARS spectra becomes possible. Such a spectrum at medium resolution is shown in **Figure 16**. The striking feature of this spectrum is the absence of a strong Rayleigh component at the pump wavelength (532 nm) because at that wavelength no energy is deposited in the sample.

The PARS technique has been extended to study vibrational–rotational transitions with high resolution (~0.005 cm$^{-1}$). For example, a high resolution PARS spectrum of the lower component of the Fermi resonance diad of $CO_2$ at a pressure of 1.6 kPa (= 11 torr) could be obtained with high signal-to-noise ratio.

In another PARS study it was shown that photoacoustic Raman spectroscopy is a sensitive technique for obtaining Raman spectra of hydrogen-bonded complexes in the gas phase. PARS spectra of the CN stretching $\nu_1$ region of HCN as a function of pressure revealed bands which could be assigned to HCN dimers and trimers.

## Applications of ionization detected stimulated Raman spectroscopy (IDSRS)

Above we have discussed how the sensitivity in determining Raman transitions can be enormously increased by employing nonlinear Raman schemes in which the shifts in vibrational state populations due to stimulated Raman transitions are probed by resonance-enhanced multiphoton ionization. As ions can be detected with much higher sensitivity than photons, the signal-to-noise ratio in the nonlinear Raman spectrum of, for example, NO could be improved by a factor of $10^3$ by this method. In fact, one can obtain sufficient sensitivity to characterize the Raman transitions of species even in molecular beams. The high sensitivity of IDSRS made it, for instance, possible to investigate the degenerate Fermi doublet of benzene in such a molecular beam experiment. The two Fermi subbands could be recorded separately by selectively tuning the UV laser into resonance with electronic transitions from one of the two states. When the Stokes laser is tuned, then the rovibrational structure of only one Raman transition is recorded. **Figure 17** shows in the upper part the lines belonging to $\nu_{16}$ and in the lower part those assigned to $\nu_2 + \nu_{18}$ in the same spectral region.

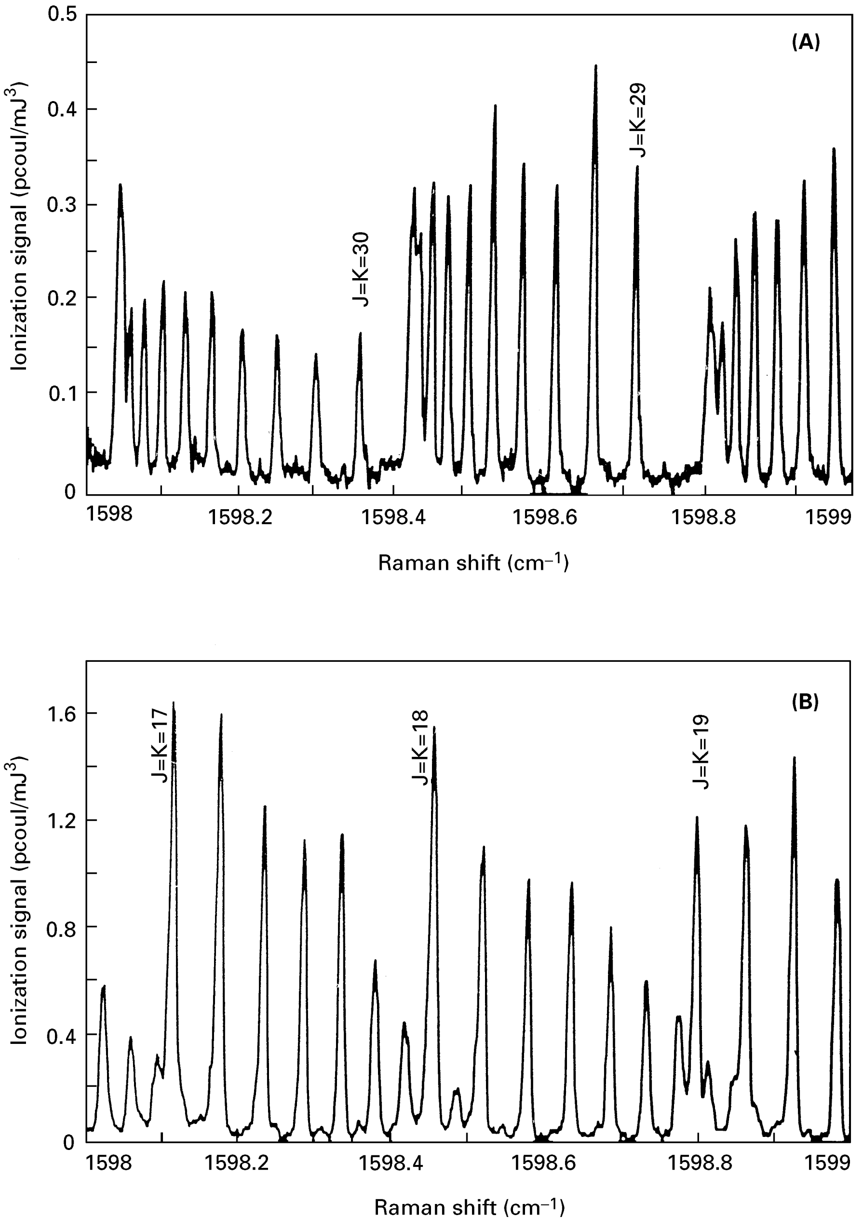



**Figure 17** Ionization detected stimulated (IDSRS) spectra of benzene in the region of overlap between O-branch transitions of $\nu_{16}$ and the S-branch transitions of $\nu_2 + \nu_{18}$. (A) UV laser tuned to 36 467 cm$^{-1}$; (B) UV laser tuned to 36 496 cm$^{-1}$. Reproduced with the permission of the American Institute of Physics from Esherick P, Owyoung A and Pliva J (1985) Ionization-detected Raman studies of the 1600 cm$^{-1}$ Fermi diad of benzene. *Journal of Chemical Physics* **83**: 3311–3317.

## List of symbols

$d\sigma/d\Omega$ = differential Raman cross-section; $E$ = electric field; $g_S$ = gain constant; $k$ = wave vector; $p$ = dipole moment; $\alpha$ = polarizability; $\beta$ = hyperpolarizability; $\gamma$ = 2nd hyperpolarizability; $\Gamma$ = line width; $\varepsilon_0$ = permittivity of vacuum; $\lambda$ = wavelength; $\chi^{(3)}$ = 3rd order nonlinear susceptibility; $\omega_L$ = angular frequency of exciting beam; $+\omega_P$ = anti-Stokes hyper-Raman displacement; $-\omega_R$ = Stokes hyper-Raman displacement.

*See also:* **Matrix Isolation Studies By IR and Raman Spectroscopies; Nonlinear Optical Properties; Nonlinear Raman Spectroscopy, Instruments; Nonlinear Raman Spectroscopy, Theory; Photoacoustic Spectroscopy, Theory; Raman Optical Activity, Applications; Raman Optical Activity, Theory; Rayleigh Scattering and Raman Spectroscopy, Theory; Surface-Enhanced Raman Scattering (SERS), Applications.**

## Further reading


Acker WP, Leach DH and Chang RK (1989) Stokes and anti-Stokes hyper Raman scattering from benzene, deuterated benzene, and carbon tetrachloride. *Chemical Physics Letters* **155**: 491–495.

Barrett JJ (1981) Photoacoustic Raman spectroscopy. In: Harvey AB (ed) *Chemical Applications of Nonlinear Raman Spectroscopy*, pp 89–169. New York: Academic Press.

Berger H, Lavorel B and Millot G (1992) In: Andrews DL (ed.), *Applied Laser Spectroscopy,* pp 267–318. Weinheim: VCH Veilagsgesellschaft.

Bermejo D, Santos J, Cancio P *et al* (1990) High-resolution quasicontinuous wave inverse Raman spectrometer. Spectrum of $CH_3D$ in the C-D stretching region. *Journal of Raman Spectroscopy* **21**: 197–201.

Campbell DJ and Ziegler LD (1993) Resonance hyper-Raman scattering in the VUV. Femtosecond dynamics of the predissociated C state of methyl iodide. *Chemical Physics Letters* **201**: 159–165.

Carreira LA and Horovitz ML (1982) Resonance coherent anti-Stokes Raman spectroscopy in condensed phases. In: Kiefer W and Long DA (eds) *Nonlinear Raman Spectroscopy and its Chemical Applications*, pp 429–443. Dordrecht: D. Reidel Publishing Company.

Esherick P, Owyoung A and Pliva J (1985) Ionization-detected Raman studies of the 1600 $cm^{-1}$ Fermi diad of benzene. *Journal of Chemical Physics* **83**: 3311–3317.

Hall RJ and Eckbreth A (1984) Coherent anti-Stokes Raman spectroscopy (CARS): Application to combustion diagnostics. In: Ready F and Erf RK (eds) *Laser Applications* **5**: 213–309. New York: Academic Press.

Harvey AB (1981) *Chemical Applications of Nonlinear Raman Spectroscopy*. New York: Academic Press.

Kiefer W and Long DA (1982) *Nonlinear Raman Spectroscopy and its Chemical Applications*. Dordrecht: D. Reidel.

Kiefer W (1995) Nonlinear Raman Spectroscopy. In: Schrader B (ed.) *Infrared and Raman Spectroscopy*, pp 162–188. Weinheim: VCH Verlag.

Kiefer W (1995) Applications of non-classical Raman spectroscopy: resonance Raman, surface enhanced Raman, and nonlinear coherent Raman spectroscopy. In: Schrader B (ed.) *Infrared and Raman Spectroscopy*, pp 465–517. Weinheim: VCH Verlag.

Materny A and Kiefer W (1992) Resonance CARS spectroscopy on diacetylene single crystals. *Journal of Raman Spectroscopy* **23**: 99–106.

Saint-Loup R, Lavorel B, Millot G, Wenger C and Berger H (1990) Enhancement of sensitivity in high-resolution stimulated Raman spectroscopy of gases. *Journal of Raman Spectroscopy* **21**: 77–83.

Schrötter HW (1995) Raman spectra of gases. In: Schrader B (ed.) *Infrared and Raman Spectroscopy,* pp 227–297. Weinheim: VCH Verlag.

# Nonlinear Raman Spectroscopy, Instruments

**Peter C Chen**, Spelman College, Atlanta, GA, USA

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES

Methods & Instrumentation

## Introduction

When exposed to large electric fields generated by intense sources of light (e.g. a laser), the charges in a material exhibit a nonlinear response. The resulting induced polarization of charge $P$ is described by the series expansion

$$P = \chi^{(1)}E + \chi^{(2)}EE + \chi^{(3)}EEE + \dots \quad [1]$$

where the second-, third-, and higher-order terms account for the nonlinear contribution. The coefficients $\chi^{(n)}$ are separate complex susceptibility tensor elements that describe the magnitude of the nonlinear contribution. The $E$s are the applied electric fields from the lasers with the form $E = A\exp[-i(\boldsymbol{k}x - \omega t)]$, where the $\boldsymbol{k}$ is the propagation or wave vector, $\omega$ is the angular frequency, and $x$ and $t$ indicate space and time, respectively. Since $\chi^{(1)} >> \chi^{(2)} >> \chi^{(3)}$, lasers with sufficiently large $E$s are required in order for the second and third terms to be significant. Most nonlinear Raman techniques rely on the third-order term to drive the induced polarization that generates an intense output beam.

Nonlinear spectra are produced by monitoring the intensity of the output beam while varying some parameters, such as the frequency $\omega$ of one or more of the laser beams. When the difference in frequency between two laser beams matches the frequency of a Raman-active mode, the resulting resonance enhances the nonlinear optical effect, causing a change in the intensity of the output beam. The result is a peak in the nonlinear Raman spectrum. Some nonlinear Raman techniques that use this approach are given in **Table 1**.

**Table 1** Comparison of some nonlinear Raman techniques

| *Technique* | *Comment* | *Variable* | *Output* | *Advantages* | *Disadvantages* |
|---|---|---|---|---|---|
| CARS | Most popular form of nonlinear Raman | $\omega_1$ or $\omega_2$ | Intensity of newly generated light at $\omega_4 = \omega_1 - \omega_2 + \omega_3$ | Fluorescence-free, intense signal at new wavelength | Phase matching required owing to dispersion, nonresonant background, complex line shape |
| CSRS | Nonparametric version of CARS | $\omega_1$ or $\omega_2$ | Intensity of newly generated light at $\omega_4 = \omega_1 - \omega_2 + \omega_3$ | Intense signal at new wavelength, can be used to observe dephasing effects | Susceptible to fluorescence, phase matching required due to dispersion, nonresonant background, complex line shape |
| SRG | Induced amplification of $\omega_2$ | Modulation of $\omega_1$ | Increase in intensity of $\omega_2$ when $\omega_1 - \omega_2 = \omega_{Raman}$ | No phase matching, no nonresonant background, linear with concentration | Sensitivity limited by stability of probe laser, difficult to multiplex |
| SRL | Induced reduction in intensity of $\omega_1$ | Modulation of $\omega_2$ | Decrease in intensity of $\omega_1$ when $\omega_1 - \omega_2 = \omega_{Raman}$ | No phase matching, no nonresonant background, linear with concentration | Sensitivity limited by stability of probe laser, difficult to multiplex |
| RIKES | Raman-induced birefringence | Modulation of $\omega_2$, $\omega_1$ is CW | Induced change in $\omega_1$ polarization for $\omega_1 - \omega_2 = \omega_{Raman}$ | Nonresonant background can be suppressed, no phase matching | Limited sensitivity, susceptible to turbulence and birefringence from windows, optics, sample |
| DFWM | Laser-induced grating | | Beam of light at $\omega_4 = \omega_1 = \omega_2 = \omega_3$ | Very sensitive, no phase matching | Multiple mechanisms (local and nonlocal), not a Raman technique |
| IRSFG | $\chi^{(2)}$ process, IR and Raman active | | Intensity of newly generated light at $\omega_3 = \omega_1 + \omega_2$ | Surface-specific | Requires tunable coherent IR source, relatively low signal intensity |

### Coherent anti-Stokes Raman spectroscopy (CARS)

Perhaps the best-known and most widely used form of nonlinear Raman spectroscopy is CARS. One of its attractive properties is that it generates an intense beam of light at a new frequency that is anti-Stokes (blue shifted) and spectrally separable from the input beams. Therefore, CARS is not susceptible to fluorescence (red shifted) and mechanisms (e.g. non-local effects) that can affect elastic scattering of the input beams.

CARS relies on the third-order term from Equation [1] which can be expanded as

$$\chi^{(3)} A_1 A_2 A_3 \exp[-i((\pm \boldsymbol{k}_1 \pm \boldsymbol{k}_2 \pm \boldsymbol{k}_3)\, x - (\pm\, \omega_1 \pm \omega_2 \pm \omega_3) t)] \qquad [2]$$

where the subscripts are labels for three input laser fields. This term can cause new light to be generated at frequency combinations corresponding to $\pm\,\omega_1 \pm \omega_2 \pm \omega_3$. CARS involves the generation of light at the specific output frequency $\omega_4 = \omega_1 - \omega_2 + \omega_3$. Raman-like peaks in the spectra are obtained when $\omega_1$–$\omega_2$ is tuned to the frequency of Raman-active vibrations or rotations. Judicious selection of the three input fields so that $\omega_1$ or $\omega_4$ matches the frequencies of coupled higher lying electronic levels can lead to the same type of enhancement observed in resonance Raman spectroscopy.

The intensity of the generated CARS beam can be written as

$$I_4 = \frac{144\pi^4 \omega_4 |\chi^{(3)}|^2 I_1 I_2 I_3 L^2}{c^4 n_1 n_2 n_3 n_4} \left\{ \frac{\sin(L\Delta k/2)}{L\Delta k/2} \right\}^2 \qquad [3]$$

where the $I$s correspond to the intensities of the beams, $n$ is the refractive index, and $c$ is the speed of light. This equation indicates that the output beam intensity varies as the product of the input laser intensities and the square of their overlap length $L$ in the sample. The squared sinc function on the right is equal to 1 when the phase of the input and output beams are matched (i.e. phase matched).

Peaks in the nonlinear Raman spectrum are produced when $\chi^{(3)}$ changes while varying the frequencies of the input beams. The intensity of the CARS process varies as the squared modulus of the nonlinear susceptibility:

$$\chi^{(3)} \propto N \sum \frac{\mu_{ac}\mu_{cb}\mu_{bd}\mu_{da}}{(\omega_c - \omega_1 - i\Gamma_{ac})(\omega_b - (\omega_1 - \omega_2) - i\Gamma_{ab})(\omega_d - (\omega_1 - \omega_2 + \omega_3) - i\Gamma_{ad})} \qquad [4]$$

which is a fourth-ranked tensor that is summed over all possible states. $N$ is the concentration, and the $\mu$s are transition dipole moments. The three products in the denominator are resonant terms that approach a minimum value of $i\Gamma$ (the dephasing linewidth) when a laser combination frequency matches the frequency of a level. The labels for the transition moments and the angular frequencies correspond to those shown in the CARS energy level diagram in **Figure 1**.

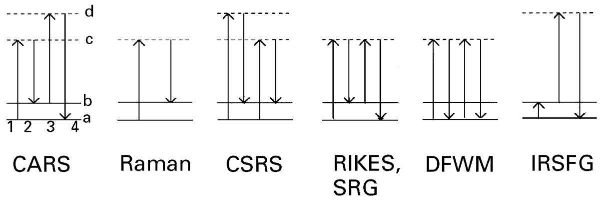



**Figure 1** Energy level diagram for CARS, Raman, CSRS, DFWM, RIKES/SRG, and IRSFG. The dotted horizontal lines represent virtual levels and the solid horizontal lines represent ground, rotational, or vibrational levels. The output frequency corresponds to the downward arrow furthest to the right in each diagram. Electronic enhancement may be achieved if the virtual levels are replaced by real levels.

Equation [4] can be used to compare spectra from incoherent Raman and CARS. First, while conventional Raman varies linearly with the sample concentration, the CARS signal varies as $|\chi^{(3)}|^2$ and is therefore proportional to the square of the $N$. Furthermore, $\chi^{(3)}$ is a summation of terms, including both resonant and nonresonant contributions ($\chi^{(3)} = \chi_{\mathrm{res}}{}^{(3)} + \chi_{\mathrm{nr}}{}^{(3)}$). Contributing terms near resonance are primarily imaginary ($i\Gamma$ dominates), while non-resonant terms are primarily real ($i\Gamma$ is negligible). The nonresonant contributions result in a nonzero background, which determines the detection limits of the technique to around 0.1% in the condensed phase and 10 ppm in the gas phase. Therefore, although the nonlinear signal is more intense, the sensitivity of CARS for trace analysis is not necessarily higher than that of more conventional techniques. Finally, since the observed signal goes as $|\chi^{(3)}|^2$, the cross-product between the $\chi_{\mathrm{res}}{}^{(3)}$ and $\chi_{\mathrm{nr}}{}^{(3)}$ can contribute dispersion-like character. Therefore, CARS peaks often have asymmetric line shapes, especially when the nonresonant background is large relative to the resonant peak.

### Other nonlinear Raman techniques

In addition to CARS, other closely related but less commonly used nonlinear Raman techniques have been developed. The energy level diagram for coherent Stokes Raman spectroscopy (CSRS) is shown in **Figure 1**. Unlike CARS, CSRS is nonparametric; the final state is not the same as the initial state. Therefore, CSRS spectra may exhibit extra peaks due to coherence dephasing. Furthermore, the CSRS output beam is generated to the Stokes (lower frequency) side of $\omega_3$. CSRS is therefore more susceptible to spectral interference from fluorescence and Rayleigh scattering of the input beams.

Other Raman-based forms of nonlinear spectroscopy include stimulated Raman gain (SRG) or stimulated Raman scattering, stimulated Raman loss (SRL) or inverse Raman spectroscopy, and Raman induced Kerr effect spectroscopy (RIKES). Some information on these techniques are provided in **Table 1**. Many of these other forms do not produce light at wavelengths that are different from the input lasers, do not involve phase matching, and may be susceptible to multiple effects that may interfere with the measurement. Consequently, these techniques have not been as widely used as CARS.

### Other nonlinear techniques

Several other forms of nonlinear spectroscopy have been developed that are not strictly based on Raman-active vibrations or rotations. Degenerate four-wave mixing (DFWM) is a $\chi^{(3)}$ technique where all input and output frequencies are identical. Because it does not involve the generation of light at new frequencies, it can rely on non-local mechanisms other than the local electronic polarizability (e.g. electrostriction). The selection rules for DFWM are closely related to those of one-photon techniques (e.g. absorption). DFWM using infrared beams is therefore used to probe infrared absorbing transitions instead of Raman-active transitions.

Finally, other nonlinear techniques can be used to obtain spectra that are both infrared and Raman active. Infrared sum frequency generation (IRSFG) is a surface-specific nonlinear technique that relies on $\chi^{(2)}$. The coherently generated output beam has a frequency of $\omega_3 = \omega_1 + \omega_2$, where $\omega_1$ is in the infrared region. The selection rules for IRSFG require that the medium be anisotropic and that the transition be both IR and Raman active.

Although the remainder of this article will focus primarily on CARS, the described advances in instrumentation and methods typically also benefit other forms of nonlinear spectroscopy.

## Experimental setup

Conventional Raman spectroscopy involves the collection and spectral analysis of light that is incoherently scattered in many directions. Nonlinear Raman spectroscopy requires careful alignment and overlap of multiple laser beams in order to produce a coherent output beam. Phase matching is also required for CARS and some other closely related nonlinear techniques (e.g. CSRS).

### Overlap

Nonlinear Raman spectroscopy requires spatial and temporal overlap of the input beams. All beams should be spatially overlapped, which can be achieved by ensuring that all beams are parallel or collinear as they enter the lens that focuses them into the sample. Spatial overlap at the sample position can then be verified by temporarily placing a knife edge or a small pinhole into the focal point overlap region. If the spatial properties of the beams (e.g. divergence and diameter) are poor or not well matched, spatial filters and additional lenses may be used to improve the quality of the overlap. Temporal overlap of the incoming beams at the sample is also essential, since the response times of some mechanisms (i.e. the local electronic polarizability) are on the order of femtoseconds. Therefore, most CARS systems use a single fixed-wavelength laser to pump all tunable lasers. Temporal overlap is then optimized using optical paths that delay any beams that would otherwise arrive at the sample prematurely. Temporal overlap may be confirmed by scattering light at the overlap region into a fast photodiode when working with nanosecond pulses. For shorter pulses, temporal overlap may involve the use of an autocorrelator. Finally, the frequency and polarization of input beams and the detection system should be adjusted as needed. Polarization optics may be inserted both in the pump beams and in the detection system.

### Phase matching

Phase matching is required for CARS experiments in normally dispersive media (i.e. condensed phase samples). The exponential terms from Equation [2] can be written as $\omega_1 + \omega_3 = \omega_2 + \omega_4$ (conservation of energy) and $\boldsymbol{k1} + \boldsymbol{k3} = \boldsymbol{k2} + \boldsymbol{k4}$ (conservation of momentum). The magnitude of each $k$ vector is $|\boldsymbol{k}| = n\omega/c$, and the direction corresponds to the direction of the beam as it propagates through the sample. In a dispersionless material (the refractive index is constant for all wavelengths of light) both conditions may be satisfied using collinear alignment of all beams (see **Figure 2A**). For most materials,

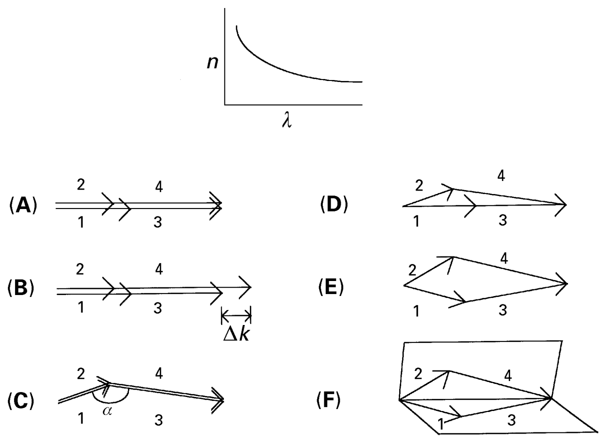



**Figure 2** The refractive index in a sample with normal dispersion increases with decreasing wavelength. The phase matching diagrams are as follows: (A) collinear phase matching in the gas phase where dispersion is negligible (e.g. gas phase); (B) phase mismatch $\Delta k$ encountered when using collinear geometry in a sample with normal dispersion; (C) possible arrangement in RIKES, SRG, and SRL where the angle $\alpha$ between beams is not critical and phase matching calculations are not needed; (D) conventional phase matching in condensed phase ($\Delta k = 0$); (E) BOXCARS phase matching; and (F) folded BOXCARS phase matching.

however, the refractive index increases with the frequency of light. Since ***k4*** has the highest frequency and therefore the largest refractive index, it is disproportionately long, causing $|\boldsymbol{k2}| + |\boldsymbol{k4}|$ to be greater than $|\boldsymbol{k1}| + |\boldsymbol{k3}|$ (see **Figure 2B**). The discrepancy in length indicates the presence of a phase mismatch $\Delta k$ between the beams. The result is a loss in the efficiency of the output beam, described by the squared sinc function in Equation [3] and shown in **Figure 3**.

**Figure 2D** shows how this problem can be fixed by introducing an angle between ***k2*** and ***k4*** to match the phases of the beams. The fact that ***k4*** is emitted along its own unique trajectory provides the ability to separate spatially the CARS output beam from the pump beams, other nonlinear processes, or other

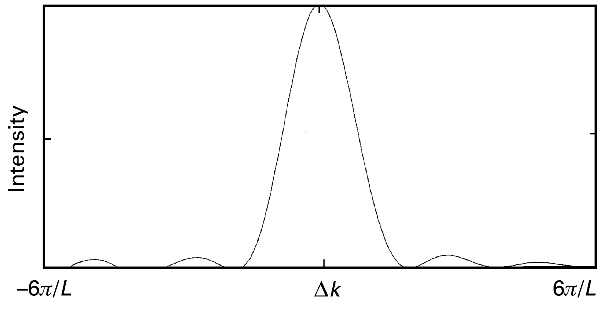



**Figure 3** Effect of the phase matching $\Delta k$ on the intensity ($I$) of the CARS output beam.

sources of spectral interference. Additional spatial discrimination may be achieved using BOXCARS phase matching, where an angle is introduced between ***k1*** and ***k3*** to increase further the angle between ***k4*** and ***k2*** (see **Figure 2E**).

In the gas phase, dispersion may be negligible, making collinear phase matching possible (see **Figure 2A**). However, the BOXCARS approach is often preferred because it allows spatial discrimination between the input and output beams. Additional spatial discrimination may be achieved using a three-dimensional form called folded BOXCARS (see **Figure 2F**).

Unfortunately, the angles required for phase matching often vary when the laser frequencies change. The magnitude of each $k$ vector depends upon both its frequency $\omega$ and the frequency-dependent refractive index $n$. Changing the frequency of any one of the four beams forces one other beam frequency to change. Therefore, the scanning of beam frequencies while producing spectra usually requires adjustment of the phase matching angles in order to avoid a phase mismatch. Without correction, the growing phase mismatch can be approximated by

$$\Delta k \approx \frac{n\Delta\omega}{c}\sqrt{2 - 2\cos\theta} \qquad [5]$$

where $n$ is the approximate refractive index, $\Delta\omega$ is the change in frequency, and $\theta$ is the angle between the two beams with changing frequencies. Therefore, this phase mismatch problem can also be minimized by reducing the angle $\theta$ between changing $k$ vectors.

## Instrumentation

In conventional Raman spectroscopy, the required instrumentation includes (1) a fixed-wavelength narrowband laser, (2) a filter, monochromator, or some other means for rejecting Rayleigh scattering, and (3) a detection system for spectrally analysing and measuring the intensity of the scattered light. Factors such as spectral resolution and scan range depend primarily upon the detection system. For CARS and other forms of nonlinear Raman spectroscopy, however, the scanning of wavelengths is often performed by the laser instead of the detection system. Therefore, the quality of the spectra depends primarily upon the lasers.

### Lasers

Most nonlinear Raman spectrometers include a fixed-wavelength laser that pumps one or more continuously tunable lasers. Some common pump lasers

include the Nd:YAG laser, excimer laser, nitrogen laser, and argon ion laser. The traditional source for broad tunability has been the dye laser, which is tunable over several tens of nanometres. For example, a common configuration involves the second harmonic of an Nd:YAG laser ($\lambda = 532$ nm) split into two beams, one for pumping a dye laser ($\omega_2$) and the other for both $\omega_1$ and $\omega_3$ (see **Figure 4**). Tuning of the dye laser frequency causes the frequency difference $\omega_1 - \omega_2$ to pass through Raman-active rotations or vibrations.

In recent years, however, dye lasers have been replaced or enhanced by sources that are more broadly tunable or that allow extension of wavelength into regions that are inaccessible by dyes. Difference frequency generation, sum frequency generation, and stimulated Raman scattering are nonlinear optical processes that can generate tunable light in the infrared and UV regions. Ti:sapphire lasers, tunable over a range of roughly 700–900 nm, are widely commercially available in both CW and pulsed (mode-locked) versions. Optical parametric devices such as the optical parametric oscillator (OPO) and the optical parametric amplifier (OPA) are nonlinear devices that are continuously tunable over wide regions of the spectrum. For example, optical parametric oscillators (OPOs) pumped by the third harmonic ($\lambda = 355$ nm) of an Nd:YAG laser can produce tunable signal and idler beams that cover a range of roughly 450–1800 nm.

The temporal behaviour of the laser source is also an important factor to consider. CW dye lasers can have low noise and extremely narrow bandwidths ($10^{-4}$ cm$^{-1}$) for high-resolution work. However, their peak powers are low (~watts), making their use with CARS possible for only the strongest Raman

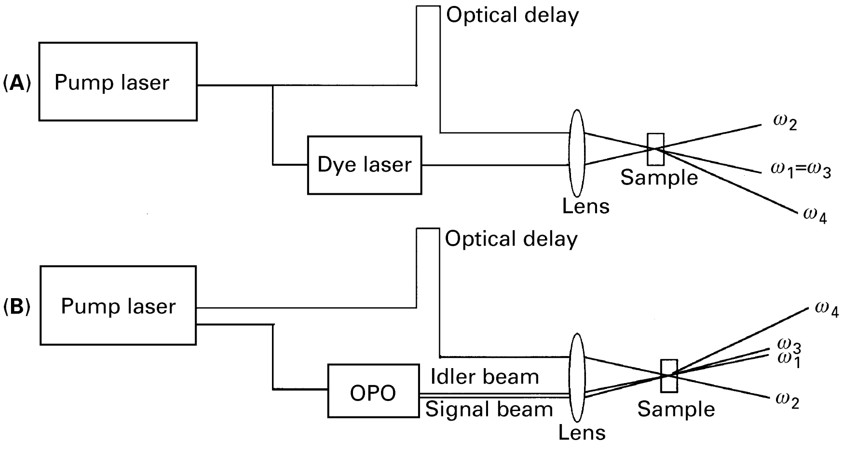



**Figure 4** Experimental setups, illustrating two possible CARS spectrometers using (A) a single dye laser where $\omega_1 = \omega_3$, and (B) an optical parametric oscillator for single-wavelength detection.

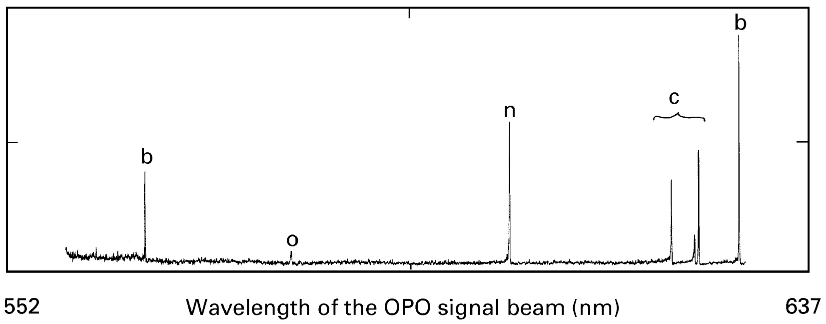



**Figure 5** A CARS vibrational spectrum produced by monitoring the output beam intensity (at $\omega_4$) while wavelength scanning an OPO (see Figure 4(B)). This spectrum shows Raman-active peaks from benzene (b), oxygen (o), nitrogen (n), and cyclohexane (c) covering a range from 681 cm$^{-1}$ ($\lambda_{OPO}$ = 552 nm) to 3098 cm$^{-1}$ ($\lambda_{OPO}$ = 637 nm). Zero frequency shift corresponds to $\lambda_{OPO}$ = 532 nm.

transitions. Unless extremely high resolution work is needed, Q-switched lasers are preferable because they can generate high peak power (MW to GW) nanosecond pulses with sufficiently narrow bandwidth (0.01–0.2 $cm^{-1}$) for most Raman applications. The noise due to these lasers may be problematic, given the relatively low repetition rate (<100 Hz) and shot-to-shot fluctuations of several per cent. For example, the mode-beating effects in Q-switched Nd:YAG lasers can cause huge shot-to-shot variations in the nonlinear optical effect. Fortunately, this mode-beating problem can be addressed by using an injection seeded Nd:YAG. Mode-locked lasers provide incredibly high peak powers (many gigawatts) in an extremely short pulse (femtoseconds to picoseconds). These lasers typically also have high repetition rates (kHz or MHz) which is convenient for signal averaging. Owing to the inverse relationship between bandwidth and pulse duration, however, mode-locked lasers generate spectrally broad pulses that may not permit adequate resolution for the system under study. Mode-locked lasers are often used for time-domain CARS.

Although narrow bandwidth is often preferred for high-resolution studies, some applications benefit from the use of a laser having a broad bandwidth in conjunction with multiwavelength detection. Spectra generated by wavelength scanning a narrowband laser can contain unwanted effects when working with samples or environments that change rapidly (e.g. turbulent combustion systems). On the other hand, broadband dye lasers with broad bandwidths (>100 $cm^{-1}$) can be used with multiwavelength detection to perform single-shot CARS spectroscopy.

The following are some key properties of a laser system:

- tuning range
- linewidth
- pulse length
- pulse energy (peak power)
- coherence
- polarization
- stability and reproducibility
- practical issues (e.g. cost, maintenance, convenience).

### Detection system

The primary functions of the detection system are to reject unwanted light, to measure the intensity of the output signal, and to analyse spectrally or temporally the output signal if needed. Collection optics (e.g. lenses, optical fibres), wavelength separation devices, detectors, and associated electronics are common components of the detection system.

Possible sources of unwanted light include fluorescence, Rayleigh scattering and ambient room light. A well-designed system will provide three means for rejecting this unwanted light and for minimizing potential damage to optics, slits, and detectors from intense beams of light. Spectral rejection may be accomplished using a combination of wavelength separation devices such as filters, prisms, gratings, or monochromators. Temporal rejection may be achieved using electronic gating, optical gating, or lock-in amplification if the signal is driven by pulsed or modulated lasers. Spatial filters that interrupt the input beams can be incorporated into the detection system if phase matching can cause the output beam to leave the sample at a different angle than that of the input beams.

Measurement of the intensity may be performed using a broad range of photoemissive or semiconductor detectors. Common issues to consider include wavelength sensitivity, damage or saturation threshold, linearity, and noise. Photoemissive detectors such as photomultiplier tubes are fast and sensitive in the UV and visible regions. They are, however, insensitive in the infrared region and easily damaged by high levels of light. Semiconductor photodiodes are less sensitive but more rugged, and may be used for more intense signals (>$10^7$ photons per pulse). Photoemissive and semiconductor detectors may also be used in multichannel form for multiwavelength detection. Examples include charge-coupled devices (CCDs) and photodiode arrays with or without microchannel plate intensifiers.

If needed, spectral analysis of light from the output beam can be achieved using a simple monochromator with a multiwavelength detector (CCD or diode array). The spectral resolution is determined by the size of the monochromator, the width of the entrance slit, the density and order of the grating, and the distance between individual elements in the detector array. Fast temporal analysis may be achieved using a fast detector such as a streak camera with picosecond or subpicosecond temporal resolution.

Most detection systems operate in one of four possible modes: single-wavelength detection, scanning detection, multiwavelength detection, and time-resolved detection. The simplest of these, single-wavelength detection, involves the detection of light at one fixed wavelength with rejection of light at all other wavelengths. The detection system may be as simple as a narrowband dielectric filter in front of a photodetector, although the use of a monochromator allows more flexibility for control of bandwidth and selection of wavelength. Scanning detection is needed for measurement of an output beam that is changing in wavelength. It typically involves wavelength

scanning of a monochromator with a photodetector, although broadband filters may be used if the change in wavelength of the output beam is small. Multi-wavelength detection is required when the output beam contains multiple frequency components that form a complete spectrum. The equipment for this mode is briefly discussed in the preceding paragraph. Time-resolved detection provides temporal information when the temporal behaviour of the output beam provides information such as the response of a system to an externally controlled stimulus.

The following are some figures of merit for detectors:

- time response and resolution
- wavelength response, discrimination, and resolution
- spatial discrimination
- sensitivity
- stability and noise
- practical issues – cost, convenience
- multichannel vs single channel
- saturation and damage threshold
- linear response and dynamic range.

## Techniques

### Acquisition of spectra

Spectral information may be acquired in three ways. The first method is the conventional approach where one or more of the laser fields frequencies are tuned to match Raman-active resonances. The second approach is to use a broadband source that allows the spectral information to be obtained in a single shot. The third is to use a time-resolved approach, where time between the pulses is varied and the sample response is measured as a function of the delay.

### Scanned CARS

In conventional frequency-domain CARS, either $\omega_1$ or $\omega_2$ is scanned so that $\omega_1 - \omega_2$ passes through Raman-active resonances. As the difference in frequency between these two beams is tuned to each resonance, a resonance enhancement of the nonlinear optical effect occurs, leading to a peak in the intensity of the output beam. Spectra are produced by plotting the intensity of the output beam as a function of $\omega_1 - \omega_2$.

The output beam is typically monitored using a scanning detection system because $\omega_4 = \omega_1 - \omega_2 + \omega_3$ varies as the input frequencies are varied. However, single-wavelength detection may be accomplished if $\omega_4$ is held constant by simultaneously tuning $\omega_3$ to compensate for the changes in $\omega_1 - \omega_2$. One way to accomplish this compensation is to let $\omega_1$ and $\omega_3$ be generated by an OPO idler and signal beam (see **Figure 4**). As the OPO beams are tuned, $\omega_1 - \omega_2$ changes, but $\omega_2$ and $\omega_4$ remain constant. This approach also reduces the phase mismatch during a scan because the angles between the scanned beams ($\theta$ in Equation [5]) may be reduced to zero.

Shot-to-shot noise in the laser system can degrade the quality of the spectra for scanned CARS. Since the signal depends on the product of three input intensities, relatively small noise in the pump laser can result in a much greater noise in the output beam intensity. This problem is especially problematic in Q-switched Nd:YAG lasers that are not injection seeded. Furthermore, shot-to-shot temporal jitter between pulses in a system that does not have a single pump laser can result in noisy spectra. Such noise problems may be corrected by simultaneously monitoring and dividing the signal by the individual pump beam intensities. Alternatively, parts of the input beams may be focused into a separate reference cell to simultaneously generate a non-resonant signal to correct for fluctuations.

### Single-shot CARS

Single-shot CARS may be accomplished by using one or more broadband lasers in addition to one or more narrowband lasers for the input beams. Each frequency element of the broadband laser(s) can independently mix with the narrowband frequency, contributing a separate frequency element to the output beam. This approach, called multi-colour CARS, multiplex CARS, or single-shot CARS, typically uses a broadband dye laser and multiwavelength detection in order to capture simultaneously a region of a few hundred wavenumbers of a rotational and/or vibrational spectrum. For example, dual broadband CARS involves the use of a single broadband dye laser for $\omega_1$ and $\omega_2$, and a fixed narrowband frequency beam for $\omega_3$ (e.g. the pump beam for the dye laser). The resulting technique provides a relatively simple way to obtain single-shot rotational spectra in the range 0–150 $cm^{-1}$. Unlike scanned CARS, the spectral resolution for this technique is often determined by the detection system.

This approach is especially useful in the analysis of gas-phase combustion and other systems where turbulence may be a problem. In the condensed phase, the range of coverage may be limited by phasematching.

### Time-resolved nonlinear Raman

Time-domain CARS involves the use of short picosecond or femtosecond pulses to generate the nonlinear Raman signal. Up to three separately timed

excitation pulses may be combined in the sample, resulting in the generation of a pulse of light called a photon echo. Measurement of size of the photon echo as a function of the delay time between pulses can be used to determine values of both the energy relaxation times $T_1$ and the phase relaxation times $T_2$. Time resolution of several femtoseconds is possible.

Another option for performing time-resolved nonlinear Raman spectroscopy is to use a fast detector such as a streak camera. By combining short picosecond or femtosecond pulses with longer nanosecond pulses, a generated signal can be produced that evolves over time. This approach can be used to obtain simultaneously both frequency and time domain information.

## List of symbols

$c$ = speed of light; $E$ = applied electric field; $I$ = beam intensity; $\mathbf{k}$ = propagation or wave vector; $L$ = overlap length; $N$ = concentration; $n$ = refractive index; $P$ = polarization of charge; $t$ = time; $x$ = space; $\theta$ = angle between beams; $\mu$ = transition dipole moment; $\omega$ = angular frequency; $\chi^{(n)}$ = complex susceptibility tensor element.

*See also:* **Laser Applications in Electronic Spectroscopy; Light Sources and Optics; Multiphoton Spectroscopy, Applications; Nonlinear Optical Properties; Nonlinear Raman Spectroscopy, Applications; Nonlinear Raman Spectroscopy, Theory; Optical Frequency Conversion; Raman Spectrometers.**

## Further reading

Bloembergen N (1992) *Nonlinear Optics*. Redwood City, CA: Addison-Wesley.

Boyd RW (1992) *Nonlinear Optics*. San Diego, CA: Academic Press.

Eckbreth AC (1996) *Laser Diagnostics for Combustion Temperature and Species*, 2nd edn. Amsterdam: Gordon and Breach.

Levenson MD and Kano SS (1988) *Introduction to Nonlinear Laser Spectroscopy*, revised edition. San Diego, CA: Academic Press.

Mukamel S (1995) *Principles of Nonlinear Optical Spectroscopy*. New York: Oxford University Press.

Shen YR (1984) *The Principles of Nonlinear Optics*. New York: Wiley.

Wright JC (1996) Nonlinear laser spectroscopy. *Analytical Chemistry* **68**: 600A–607A.

Wright JC (1982) Applications of lasers in analytical chemistry. In: Evans TR (ed) *Techniques of Chemistry*, Vol 17, pp 35–179. New York: Wiley.

Yariv A (1989) *Quantum Electronics*, 3rd edn. New York: Wiley.

Zinth W and Kaiser W (1993) Ultrafast coherent spectroscopy. In: Kaiser W (ed) *Topics in Applied Physics*, 2nd edn, Vol 60, pp 235–277. Berlin: Springer-Verlag.

# Nonlinear Raman Spectroscopy, Theory

**J Santos Gómez**, Instituto de Estructura de la Materia, CSIC, Madrid, Spain

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Theory**

## Introduction

In a typical spontaneous Raman experiment, an incident, nonresonant photon of energy $\hbar\omega_P$ interacts with the molecule and is scattered into a photon of energy $\hbar(\omega_P \pm \omega_R)$ where $\omega_R$ is the frequency of a vibrational mode. The molecule undergoes a transition that balances the gain or loss of field energy. The spectroscopic information is extracted by measuring the energy change of the scattered photon.

From a classical point of view, the molecule is polarized by its interaction with the input field at $\omega_P$ and an oscillating dipole at frequency $\omega_P$ is induced. As the molecular polarizability itself is modulated by internuclear motion at frequency $\omega_R$, lateral bands appear at combination frequencies. The molecular dipole oscillating at $\omega_P \pm \omega_R$ radiates a field at these frequencies. Even with a coherent input field, as provided by lasers, the output field is incoherent because the phases of individual scatterers are not correlated.

In a typical nonlinear Raman experiment the molecule interacts with two strong coherent fields at frequencies $\omega_P$ (pump) and $\omega_S$ (Stokes). As we will see below, for strong field the response of the system is nonlinear and the molecular polarizability at a given frequency is periodically modulated at the driving

frequencies and at new frequencies that are linear combinations of these $2\omega_P$, $2\omega_S$, $\omega_P \pm \omega_S, \ldots$ . If we focus on the oscillating polarizability component at $\omega_P - \omega_S$, the interaction with either of the input fields or with a third input field at $\omega_0$ gives rise as before to a component of the induced dipole at a combination frequency, such as $\omega_P + (\omega_P - \omega_S) = \omega_{AS}$, $\omega_P - (\omega_P - \omega_S) = \omega_S$, $\omega_0 + (\omega_P - \omega_S)$, whose amplitude depends on intrinsic molecular properties and on the product of three field amplitudes. The input fields being temporal and spatially coherent, there is a definite phase relationship between driven dipole oscillations of different molecules in the interaction volume, giving rise to a coherent macroscopic polarization in the medium.

The created polarization acts as a source and a new coherent field at frequency $\omega_\sigma$ grows to some extent or, if $\omega_\sigma = \omega_P$, $\omega_S$ or $\omega_0$, the corresponding amplitude increases or decreases depending upon the relative field-polarization phase.

The above behaviour is quite general and will be observed for any molecule and frequency combination, provided that some macroscopic symmetry requirements are met. The interaction just described is a four-wave process: a macroscopic polarization at a signal frequency $\omega_\sigma$ builds up through a nonlinear frequency mixing of three frequency components of the electromagnetic field. In a general $(n+1)$-wave process, $n$ frequency components are mixed to produce an oscillating polarization at $\omega_\sigma = \pm(\omega_1 \pm \omega_2 \pm \omega_3 \pm \cdots \pm \omega_n)$. The magnitude of the nonlinear polarization depends on the product of $n$ field amplitudes and an $n$th-order nonlinear susceptibility $\chi^{(n)}(\omega_\sigma; \omega_1, \omega_2, \ldots \omega_n)$, which is a material property.

The connection of this nonlinear optical effect with spectroscopy lies in the fact that $\chi^{(n)}$, and hence the signal strength, will be enhanced whenever a linear combination of a subset of $m \leq n$ frequency components approaches the energy difference of two molecular states of proper symmetry. This $m$-photon resonant enhancement — which can in principle be observed even in the absence of real transitions, as will be the case for states with no thermal population — has been exploited to develop a number of nonlinear spectroscopic techniques that differ in order $n$, order of the used resonance $m$, number of colours (i.e. different actual laser fields that provide the frequency components), spectral and temporal resolution and the actual method used to detect the resonances by monitoring either the power generated at $\omega_\sigma$ or the change in amplitude, polarization or phase at some of the input frequencies.

The simplest $m$-photon resonance is obtained for $m = 2$ when $\Omega_{fg} = (E_f - E_g)/\hbar \simeq \omega_1 \pm \omega_2$. If we consider $\omega_1$ and $\omega_2$ to be typical optical frequencies, the + sign correspond to two vibronic states $g$ and $f$, from different electronic states, in resonance with $2\omega_1$, $2\omega_2$ or $\omega_1 + \omega_2$: two-photon absorption-like resonance (TPA).

Most nonlinear Raman techniques correspond to the − sign above: $\Omega_{fg} \simeq \omega_P - \omega_S$ (Raman-like resonance), where frequency subscripts have been converted to the usual Raman convention. These resonances can contribute to enhance $n$th-order processes for $n \geq 2$. For isotropic media, all even-order susceptibilities vanish and Raman techniques involve 3rd, 5th, 7th,… order nonlinear susceptibilities. The ordering comes from a perturbative development of the field–molecule interaction and, as far as the perturbative approach can be applied, successive orders correspond to much smaller terms. Hence we can study the main spectroscopic features with only the lowest-order nonvanishing term, and most nonlinear Raman techniques are 3rd-order four-wave processes, which will constitute the main topic in this article. These techniques are indeed readily implemented with present laser technology in a variety of media, including gases at low pressure.

We can perform Raman spectroscopy by monitoring different properties of the macroscopic field as we scan $\omega_P$ or $\omega_S$ in such a way that $\omega_P - \omega_S \simeq \Omega_{fg} \equiv \omega_R$, leading to several techniques that have usually been identified by acronyms. The energy level diagrams for the main nonlinear Raman techniques are depicted in **Figure 1**.

With only two colours we can monitor the intensity of the generated beam at the anti-Stokes frequency $\omega_{AS} = \omega_P + \omega_R$ (coherent anti-Stokes Raman spectroscopy, CARS) or the Stokes frequency respect to $\omega_S$, $\omega_{CSRS} = \omega_S - \omega_R$ (coherent Stokes Raman spectroscopy, CSRS). Alternatively, we can monitor the intensity increase at $\omega_S$ (stimulated Raman gain spectroscopy, SRG) or the intensity decrease at $\omega_P$ (stimulated Raman loss, SRL, also known as inverse Raman spectroscopy, IRS). If we focus on the polarization of the created field or the change in polarization state of input fields, we arrive at different polarization variants of the above techniques, such as polarization CARS, CARS ellipsometry or Raman-induced Kerr effect (RIKES). The input frequencies can be tuned to match additional one-photon resonances that can enhance the signal by several orders of magnitude, leading to techniques such as resonance CARS. A third colour $\omega_0$ can be used for different purposes such as fine tuning of additional one-photon resonances or shifting the signal frequency $\omega_\sigma = \omega_0 + \omega_R$ to a more convenient region; this expands the possibilities, leading to Raman resonant four-wave mixing (FWM). The different choices of input and signal beams are collected

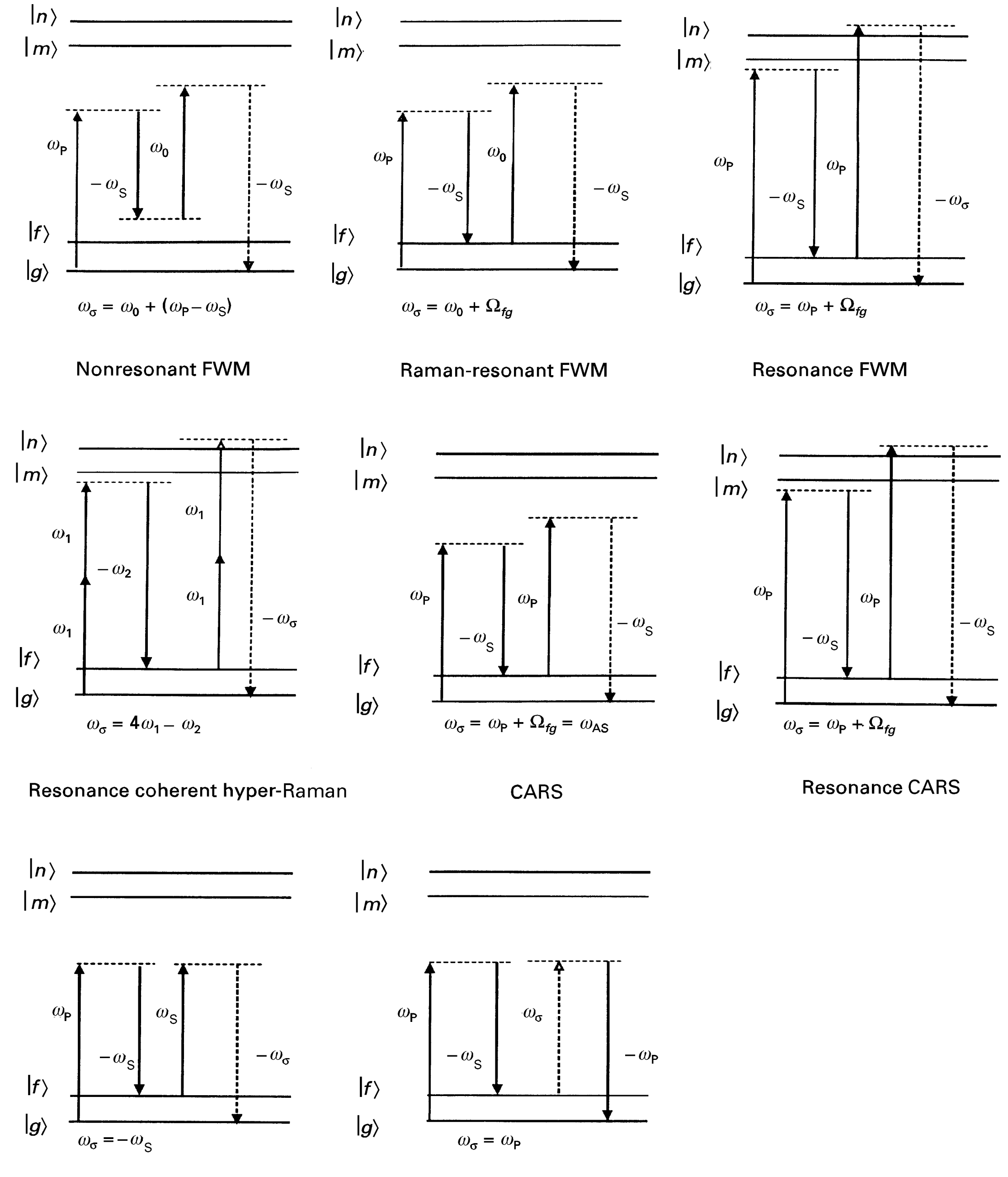



**Figure 1** Energy level diagrams for nonlinear Raman spectroscopic techniques. The input fields are shown as arrows pointing upwards for positive frequency (downwards for negative) in the arguments of the nonlinear susceptibility, and the output field as a dashed arrow, which must close the diagram. Schematic molecular energy levels are represented to show the main Raman resonance and additional one-photon resonances. Horizontal position does not imply time ordering. The energy balance involves the field and the material excitation in different ways for different techniques. In pure nonresonant processes, the energy of created photons balances that of those destroyed. In SRL and SRG, each scattering event leads to an excited molecule. In Raman-resonant FWM and CARS, the energy of destroyed photons is shared among material excitation and created photons depending on the relative magnitude of the imaginary and real parts of the nonlinear susceptibility.

in **Table 1**, along with the names of the associated nonlinear Raman spectroscopic techniques.

The spectroscopic use of three-photon resonances has been demonstrated, although these usually require additional enhancement through one-photon electronic resonance, or the use of condensed media. Closely related to Raman spectroscopy is coherent hyper-Raman scattering (see **Figure 1**), the stimulated analogue of the spontaneous hyper-Raman effect, for which a three-photon resonant

**Table 1** Nonlinear Raman spectroscopic techniques

| | *FWM* | *CARS* | *RIKES* |
|---|---|---|---|
| Name | Four-wave mixing | Coherent anti-Stokes Raman spectroscopy | Raman-induced Kerr effect |
| Monitored effect | Intensity at $\omega_\sigma = \omega_0 + \omega_R$ | Intensity at $\omega_{\text{anti-Stokes}} = \omega_P + \omega_R$ | Intensity at $\omega_S$ polarized $\perp$ |
| Input/output field | $\omega_0$, $\omega_P$, $\omega_S$ / $\omega_0 + \omega_R$ | $\omega_P$, $\omega_S$ / $\omega_A$ | $\omega_P$, $\omega_S$ / $\omega_S$ |
| Input/output polarization | All possibilities with paired Cartesian index | $E_{\omega_P x}$, $E_{\omega_s x}$ / $E_{\omega_A x}$<br>$E_{\omega_P x}$, $E_{\omega_s y}$ / $E_{\omega_A y}$ | $E_{\omega_P}$ circular, $E_{\omega_s x}$ / $E_{\omega_s y}$ |
| *Effective susceptibility* | $\chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma;\omega_0\omega_P-\omega_S)=\chi^{(3)NR}_{\mu\alpha\beta\gamma}+\frac{1}{6\epsilon_0\hbar}$<br>$\times\sum_{g,f}(N_g-N_f)\alpha_{\alpha\mu}\alpha_{\gamma\beta}\left(\frac{\Delta_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}+\mathrm{i}\frac{\Gamma_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}\right)$ | $\chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma;\omega_P\omega_P-\omega_S)=\chi^{(3)NR}_{\mu\alpha\beta\gamma}+\frac{1}{3\epsilon_0\hbar}$<br>$\times\sum_{g,f}(N_g-N_f)\alpha^2_{x\alpha}\left(\frac{\Delta_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}+\mathrm{i}\frac{\Gamma_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}\right)$ | $\chi^{(3)}_{\text{eff}}(-\omega_S;\omega_P-\omega_P\omega_S)$<br>$=\frac{\mathrm{i}}{2}\left(\chi^{(3)}_{yyxx}-\chi^{(3)}_{yxyx}\right)\approx\frac{\mathrm{i}}{2}\left(\chi^{(3)R}_{yyxx}-\chi^{(3)R}_{yxyx}\right)$ |
| Signal[a] | $I_{\omega_\sigma}=\frac{9}{4}\frac{\omega_\sigma^2}{\varepsilon_0^2c^4n_\sigma n_0n_Pn_S}\left\lvert\chi^{(3)}_{\mu\alpha\beta\gamma}\right\rvert^2I_{\omega_0}I_{\omega_P}I_{\omega_S}L^2\,\mathrm{sinc}^2(\Delta kL/2)$ | $I_{\omega_A}(L)=\frac{9}{16}\frac{\omega_A^2}{\varepsilon_0^2c^4n_An_P^2n_S}\left\lvert\chi^{(3)}_{\alpha xx\alpha}\right\rvert^2I^2_{\omega_P}(0)I_{\omega_S}(0)L^2$<br>$\times\,\mathrm{sinc}^2(\Delta kL/2)$ | $\propto\left\lvert\chi^{(3)RIKES}_{\text{eff}}(-\omega_S;\omega_P-\omega_P\omega_S)\right\rvert^2I_P^2I_S$ |

| | *OHD-RIKES* | *SRL* | *SRG* |
|---|---|---|---|
| Name | *Optically heterodyne detected RIKES* | Stimulated Raman loss or inverse Raman spectroscopy | Stimulated Raman gain |
| Monitored effect | Intensity at $\omega_S$ polarized $\perp$ | Intensity decrease at $\omega_P$ | Intensity at $\omega_S$ |
| Input/output fields | $\omega_P$, $\omega_S$ / $\omega_S$ | $\omega_P$, $\omega_S$ / $\omega_P$ | $\omega_P$, $\omega_S$ / $\omega_S$ |
| Input/output polarization | $E_{\omega_P}$ circular, $E_{\omega_s x}$ / $E_{\omega_s y}$ | $E_{\omega_P x}$, $E_{\omega_s x}$/$E_{\omega_P x}$ | $E_{\omega_P x}$, $E_{\omega_S x}$/ $E_{\omega_S x}$ |
| Effective susceptibilty | $\chi^{(3)}_{\text{eff}}(-\omega_S;\omega_P-\omega_P\omega_S)$<br>$=\frac{\mathrm{i}}{2}\left(\chi^{(3)}_{yyxx}-\chi^{(3)}_{yxyx}\right)\approx\frac{\mathrm{i}}{2}\left(\chi^{(3)R}_{yyxx}-\chi^{(3)R}_{yxyx}\right)$ | $\chi^{(3)}_{x\alpha x\alpha}(-\omega_P;\omega_S\omega_P-\omega_S)=\chi^{(3)NR}_{x\alpha x\alpha}+\frac{1}{6\epsilon_0\hbar}$<br>$\times\sum_{g,f}(N_g-N_f)\alpha^2_{\alpha x}\left(\frac{\Delta_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}+\mathrm{i}\frac{\Gamma_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}\right)$ | $\chi^{(3)}_{x\alpha x\alpha}(-\omega_S;\omega_P-\omega_P\omega_S)=\chi^{(3)NR}_{x\alpha\alpha x}+\frac{1}{6\epsilon_0\hbar}$<br>$\times\sum_{g,f}(N_g-N_f)\alpha^2_{x\alpha}\left(\frac{\Delta_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}-\mathrm{i}\frac{\Gamma_{fg}}{\Delta_{fg}^2+\Gamma_{fg}^2}\right)$ |
| Signal[a] | $\propto\mathrm{Im}\left(\chi^{(3)RIKES}_{\text{eff}}\right)I_PI_S$<br>or $\propto\mathrm{Re}\left(\chi^{(3)RIKES}_{\text{eff}}\right)I_PI_S$ | $I_{\omega_P}(L)=I_{\omega_P}(0)\left[1-\frac{3\omega_P}{\varepsilon_0c^2n_Sn_P}\mathrm{Im}\left(\chi^{(3)SRL}_{x\alpha x\alpha}\right)I_{\omega_S}(0)L\right]$ | $I_{\omega_S}(L)=I_{\omega_S}(0)\left[1-\frac{3\omega_S}{\varepsilon_0c^2n_Sn_P}\mathrm{Im}\left(\chi^{(3)SRG}_{x\alpha\alpha x}\right)I_{\omega_P}(0)L\right]$ |

[a] Signal expression for plane monochromatic input waves. $L$ is the interaction length.

enhancement $\Omega_{fg} \simeq 2\omega_1 - \omega_2$ can be detected in $\chi^{(5)}$ in a six-wave process, for example by monitoring the intensity of the generated field at $\omega_\sigma = 4\omega_1 - \omega_2 \simeq 2\omega_1 + \omega_R$. We will not explicitly consider these higher-order resonances in the following, but the theory involved closely follows that for the simplest two-photon resonant four-wave case.

From a fully quantum-mechanical view, the system matter+field undergoes a transition from an initial state $|n_P = n, n_S = 0, g\rangle$, for which the input field has $n$ photons, the output field is at the vacuum state and the molecule in an internal state $g$, to a final state $|n_P = n - 1, n_S = 1, f\rangle$. Owing to the low Raman scattering cross section, the probability of finding one scattered photon within the interaction region at the time of the next scattering event is vanishingly small at low input intensity. As a consequence, the initial state of the output mode is always the vacuum state and the number of scattered photons is linearly dependent on the number of photons in the input field alone. At high input intensity or in the presence of a second external field at $\omega_S$, $n_S \neq 0$ and the matrix element depends on the product of $n_P \times n_S$. Therefore, spontaneous and nonlinear Raman are closely connected. In fact, the spontaneous Raman effect cannot be considered among linear optical phenomena such as absorption, diffraction or stimulated emission because the scattered field has a different frequency.

The dependence of the spatial, spectral and temporal signal properties on those of the input beams, along with the broad range of lasers available, opens many possibilities ranging from high-resolution studies of gases at Doppler-limited resolution using narrow-bandwidth lasers to high-temporal-resolution studies of intramolecular and intermolecular dynamics and relaxation processes at the femtosecond scale in condensed media. Although the underlying theory is essentially the same, high-resolution studies are better described in the frequency domain, looking at the Raman resonances in the nonlinear susceptibility of the medium, allowing for interpretation of line position, intensity and (with the inclusion of relaxation terms) line width. Broadband ultrashort pulses lead to an impulse response, where the molecule evolves freely after the sudden excitation involving many states and is probed at later times. This situation is better described in the time domain, looking at the time evolution of the nonlinear response function, which shows relaxation and dynamic processes, directly and still contains information about the spectrum in the form of oscillations or beating among different excited modes. Midway between these limits, lasers with small bandwidth and short duration can be used to excite a selected set of molecular levels and study their evolution.

## Theory

The theory relating first principles to practical signal expressions for an arbitrary material system can be considered to be well understood. We present here a sketch of its main aspects, indicating the approximations needed and its scope. Our goal is to give a practical signal expression in terms of Raman scattering tensor components, field intensities and polarizations. Our model system is an isotropic medium composed of polarizable units, small compared with the wavelength, interacting only through their coupling to a common thermal bath, such as a gas or dilute solution, in the presence of a superposition of quasi-monochromatic ($\delta\omega_i \ll \omega_i$) laser pulses.

Nonlinear interaction can be studied on a full quantum basis by expressing the Hamiltonian in terms of quantum operators for the material and the radiation degrees of freedom, but the most fruitful treatment has been the semiclassical approach in which the electric field operator is replaced by its expectation value and considered as a function, though quantum operators are used for the material. Only electric dipole polarization will be addressed.

Owing to the coherence, we need to consider the macroscopic evolution of the field in a medium that shows a macroscopic polarization induced by the field–matter interaction. This will be done in three steps. First, the polarization induced by an arbitrary field will be calculated and expanded in power series in the field, the coefficients of the expansion being the material susceptibilities $\chi^{(n)}$ (frequency domain) or response function $\mathbf{R}^{(n)}$ (time domain) of $n$th-order. Nonlinear Raman effects appear at third order in this expansion. Second, the perturbation theory derivation of the third-order nonlinear susceptibility $\chi^{(3)}$ in terms of molecular eigenstates and transition moments will be outlined, leading to a connection with the spontaneous Raman scattering tensor components. Last, the interaction of the initial field distribution with the created polarization will be evaluated and the signal expression obtained for the relevant techniques of **Table 1**.

### Macroscopic nonlinear polarization

We consider a unit volume that contains many molecules but is small compared to the wavelength. The electric field can be considered uniform inside this volume. By expanding the polarization at a given time in power series in the field, we obtain for a given point:

$$\boldsymbol{P}(t) = \boldsymbol{P}^{(0)}(t) + \boldsymbol{P}^{(1)}(t) + \boldsymbol{P}^{(2)}(t) + \boldsymbol{P}^{(3)}(t) + \cdots \quad [1]$$

where $\boldsymbol{P}^{(0)}$ stands for the static polarization, $\boldsymbol{P}^{(1)}$ is linear in the field, $\boldsymbol{P}^{(2)}$ quadratic and so on (bold typefaces represent a vector, tensor or matrix magnitude).

The next approximation is to consider only the *local response*: the polarization at the point $\boldsymbol{r}$ depends only on the field at this point. In condensed media, the local field sensed by the molecule includes the depolarizing effect of the surrounding molecules, and it does depend on the field values at different points. In an anisotropic medium the material response shows *spatial dispersion* and is sensitive to both frequency and wavevector of the driving fields.

Nonlinear Raman processes are governed by the third-order polarization, which can be expressed in the time domain, for a generic electric field, in component form:

$$P_\mu^{(3)}(t) = \epsilon_0 \int_{-\infty}^{\infty} \mathrm{d}\tau_1 \int_{-\infty}^{\infty} \mathrm{d}\tau_2 \int_{-\infty}^{\infty} \mathrm{d}\tau_3 \; R_{\mu\alpha\beta\gamma}(\tau_1, \tau_2, \tau_3)\, E_\alpha(t-\tau_1)\, E_\beta(t-\tau_2)\, E_\gamma(t-\tau_3) \qquad [2]$$

where summation over repeated indices is assumed.

$\mathbf{R}^{(3)}(\tau_1, \tau_2, \tau_3)$ is a fourth-order tensor ($\mathbf{R}^{(n)}$ has rank $n$+1) of components $R^{(3)}_{\mu\alpha\beta\gamma}(\tau_1, \tau_2, \tau_3)$ where $\mu\alpha\beta\gamma$ represent Cartesian coordinates describing the field polarization. It is a real function of $n$ time variables (because $\boldsymbol{E}$ and $\boldsymbol{P}$ are physical observables) and vanish whenever any of its time arguments are negative because the polarization at time $t$ cannot depend on the field values at later times. Owing to the implicit summation, the above equation is invariant to any permutation of the pairs $(\alpha, \tau_1)$, $(\beta, \tau_2)$ etc. (intrinsic permutation symmetry).

To study narrow resonances, the frequency domain allows for a better description of the material response in terms of susceptibilities. To obtain the corresponding expressions we take the Fourier transform of $\boldsymbol{E}(t)$ and $\boldsymbol{P}(t)$.

$$\boldsymbol{E}(t) = \int_{-\infty}^{\infty} \mathrm{d}\omega\, \boldsymbol{E}(\omega) \exp(-\mathrm{i}\omega t)$$

$$\boldsymbol{E}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\tau\, \boldsymbol{E}(\tau) \exp(\mathrm{i}\omega \tau)$$

$$\boldsymbol{P}(t) = \int_{-\infty}^{\infty} \mathrm{d}\omega\, \boldsymbol{P}(\omega) \exp(-\mathrm{i}\omega t)$$

$$\boldsymbol{P}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\tau\, \boldsymbol{P}(\tau) \exp(\mathrm{i}\omega \tau) \qquad [3]$$

By substituting Equation [3] into Equation [2] and rearranging, we arrive at

$$P_\mu^{(3)}(\omega_\sigma) = \epsilon_0 \int_{-\infty}^{\infty} \mathrm{d}\omega_1 \int_{-\infty}^{\infty} \mathrm{d}\omega_2 \int_{-\infty}^{\infty} \mathrm{d}\omega_3 \; \chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma; \omega_1\omega_2\omega_3) E_\alpha(\omega_1) E_\beta(\omega_2) E_\gamma(\omega_3) \delta(\omega - \omega_\sigma) \qquad [4]$$

where $\omega_\sigma = \omega_1 + \omega_2 + \omega_3$, and the third-order nonlinear susceptibility $\chi^{(3)}_{\mu\alpha\beta\gamma}$ is defined by

$$\chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma; \omega_1\omega_2\omega_3) = \int_{-\infty}^{\infty} \mathrm{d}\tau_1 \int_{-\infty}^{\infty} \mathrm{d}\tau_2 \int_{-\infty}^{\infty} \mathrm{d}\tau_3 \; R^{(3)}_{\mu\alpha\beta\gamma}(\tau_1\tau_2\tau_3) \exp\left(\mathrm{i}\sum \omega_i \tau_i\right) \qquad [5]$$

$\boldsymbol{\chi}^{(3)}$ is a field- and time-independent fourth-rank tensor, and a function of three frequency arguments. Its general properties can be derived from that of the nonlinear response: $\boldsymbol{E}(\omega)$ and $\boldsymbol{P}^{(3)}(\omega)$ include both positive and negative frequencies, related by the conditions $\boldsymbol{E}(-\omega) = \boldsymbol{E}^*(\omega)$, $\boldsymbol{P}(-\omega) = \boldsymbol{P}^*(\omega)$. The reality condition of $\mathbf{R}^{(n)}$ implies that $[\boldsymbol{\chi}^{(3)}(-\omega_\sigma; \omega_1\omega_2\omega_3)]^* = \boldsymbol{\chi}^{(3)}(\omega_\sigma; -\omega_1 - \omega_2 - \omega_3)$, $\boldsymbol{\chi}^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma; \omega_1\omega_2\omega_3)$ is invariant under the permutation of the pairs $(\alpha, \omega_1)$, $(\beta, \omega_2)$.

The integration in Equation [4] considers the existence of several frequency combinations matching the condition $\omega_\sigma = \Sigma_j\, \omega_j$, within the bandwidth of the applied field. A single laser whose bandwidth is large enough to include frequency components that match $\omega_a - \omega_b = \omega_R$ can drive the nonlinear response. However, more frequently the laser bandwidth is smaller than $\omega_R$ and different beams with distinct centre frequencies at needed to match the Raman resonance. Let us now consider a field composed of a superposition of quasi-monochromatic beams, i.e. the amplitude at a carrier frequency is modulated by a complex envelope that defines its spectral shape. The central frequencies are chosen to match the Raman resonant four-wave mixing scheme (see **Table 1**), which can be later adapted to other colour choices. Three laser beams are present of frequencies $\omega_0$, $\omega_P$, $\omega_S$, with $\omega_P - \omega_S = \omega_R$ as the only resonant frequency combination.

The field can be expressed as

$$\boldsymbol{E}(t) = \sum_{\omega_j = \omega_0, \omega_P, \omega_S} \frac{1}{2}[\boldsymbol{E}_{\omega_j}(t) \exp(-\mathrm{i}\omega_j t) + \boldsymbol{E}_{-\omega_j}(t) \exp(\mathrm{i}\omega_j t)] \qquad [6]$$

The envelope $\boldsymbol{E}_{\omega_j}(t)$ is now time-dependent, and contains both amplitude and phase information, thus defining the spectrum (centred at $\omega_j$) of each field. It is implicit in the above considerations that $\boldsymbol{E}_{\omega_j}(t)$ changes smoothly in the timescale of $(\omega_j)^{-1}$. By substituting Equation [6] into Equation [4], we found that $\boldsymbol{P}^{(3)}(t)$ can be recast as a sum of quasi-monochromatic components at central frequency $\omega_\sigma$. The envelope $\boldsymbol{P}^{(3)}_{\omega_\sigma}(t)$ will vanish due to the fast oscillating exponential factors except when $\omega_\sigma$ is close to any linear combination of three frequencies chosen among the input frequencies, leading to

$$\boldsymbol{P}^{(3)}(t) = \sum_{\omega_\sigma} \frac{1}{2}\Big[\boldsymbol{P}^{(3)}_{\omega_\sigma}(t)\exp(-\mathrm{i}\omega_\sigma t) + \boldsymbol{P}^{(3)}_{-\omega_\sigma}(t)\exp(\mathrm{i}\omega_\sigma t)\Big] \quad [7]$$

Only combinations including $\pm(\omega_P - \omega_S)$ will show resonant enhancement. Among these, we are interested in the term that oscillates at a centre frequency $\omega_\sigma = \omega_0 + \omega_P - \omega_S$, which corresponds to a field $\omega_0$ being scattered into an output field $\omega_0 + \omega_R$ due to the driving effect of the resonant pair $\omega_P$, $\omega_S$. This term is

$$\begin{aligned}\boldsymbol{P}^{(3)}_{\omega_\sigma}(t) = \epsilon_0 \frac{D}{4} \int_{-\infty}^{\infty} \mathrm{d}\tau_1 \int_{-\infty}^{\infty} \mathrm{d}\tau_2 \int_{-\infty}^{\infty} \mathrm{d}\tau_3 \\ \times\ \boldsymbol{\Phi}^{(3)}_{-\omega_\sigma;\,\omega_0\,\omega_P-\omega_S}(t-\tau_1, t-\tau_2, t-\tau_3) \\ \times\ \boldsymbol{E}_{\omega_0}(\tau_1)\ \boldsymbol{E}_{\omega_P}(\tau_2)\boldsymbol{E}^*_{\omega_S}(\tau_3)\end{aligned} \quad [8]$$

where $\boldsymbol{\Phi}^{(3)}(t-\tau_1, t-\tau_2, t-\tau_3) = \mathbf{R}^{(3)}(t-\tau_1, t-\tau_2, t-\tau_3)\exp(i\sum_j \omega_j\tau_j)$ is an *envelope response function* that relates the envelope of the third-order polarization at frequency $\omega_\sigma$ with the envelopes of the fields when their central frequency is tuned to the values given by the subscripts of $\boldsymbol{\Phi}^{(3)}$. The relationship between $\boldsymbol{\chi}^{(3)}$ and $\boldsymbol{\Phi}^{(3)}$ is:

$$\boldsymbol{\chi}^{(3)}(-\omega_\sigma;\omega_0\omega_P-\omega_S) \int_{-\infty}^{\infty} \mathrm{d}\tau_1 \int_{-\infty}^{\infty} \mathrm{d}\tau_2 \int_{-\infty}^{\infty} \mathrm{d}\tau_3\ \boldsymbol{\Phi}^{(3)}_{-\omega_\sigma;\omega_0\omega_P-\omega_S}(\tau_1,\tau_2,\tau_3) \quad [9]$$

The $D$ factor above comes from the identification of terms in the summation that are identical owing to the intrinsic permutation symmetry. We must now compare the characteristic time scale of the material $T_1$ (population relaxation) and $T_2$ (phase relaxation) with that of the laser pulse $\tau_p$, taken as the period of the fastest amplitude change in the field (similar to the pulse duration for a smooth, narrow-bandwidth pulse, but close to the inverse of the full bandwidth for a multimode laser).

For $\tau_p \ll T_2$ each pulse appears to the system as a $\delta$ impulsion at fixed time $t_i$, $A\ \boldsymbol{E}_{\omega_a}(t) = A_{\omega_a}\delta(t-t_i)$, where $A_{\omega_a}$ is the complex pulse area. The polarization is

$$\boldsymbol{P}^{(3)}_{\omega_\sigma}(t) = \epsilon_0 \frac{D}{4} \boldsymbol{\Phi}^{(3)}_{-\omega_\sigma;\omega_0\omega_P-\omega_S}(t-t_0, t-t_P, t-t_S)\ \boldsymbol{A}_{\omega_0}\boldsymbol{A}_{\omega_P}\boldsymbol{A}^*_{\omega_S} \quad [10]$$

$\boldsymbol{\Phi}^{(3)}$ can now be directly measured by changing the times $t_i$ at which the pulses are applied. A typical implementation of this scheme is to excite a Raman mode by applying $\omega_P$ and $\omega_S$ simultaneously at $t = 0$ and provide some variable delay for the response to evolve before applying the field at $\omega_0$ that is scattered into the signal field $\omega_\sigma$.

On the other hand, for $\tau_p \gg T_2$ the macroscopic polarization relaxes very fast, the system loses memory of previous field values, and $\boldsymbol{P}^{(3)}_{\omega_\sigma}(t)$ follows adiabatically the field envelope, leading to

$$\boldsymbol{P}^{(3)}_{\omega_\sigma}(t) = \epsilon_0 \frac{D}{4} \boldsymbol{\chi}^{(3)}(-\omega_\sigma;\omega_0\omega_P-\omega_S)\ \boldsymbol{E}_{\omega_0}(t)\ \boldsymbol{E}_{\omega_P}(t)\ \boldsymbol{E}^*_{\omega_S}(t) \quad [11]$$

This expression is the starting point for the vast majority of nonlinear Raman experiments. One special case that enters into this category is CW experiments with monochromatic waves ($\tau_p \to \infty$ and the amplitudes of incoming fields and nonlinear polarization become constant). The corresponding expression is

$$P^{(3)}_{\omega_\sigma\mu} = 2^{1-n}\epsilon_0 D\chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma;\omega_0\omega_P-\omega_S)\ E_{\omega_0\alpha}E_{\omega_P\beta}E^*_{\omega_S\gamma} \quad [12]$$

where individual vector and tensor components are used. $D$ is the number of distinguishable permutations; thus $D = 6$ for processes involving three different input frequencies but $D = 3$ for two-colour processes (i.e. $\omega_0 = \omega_P$ or $\omega_0 = \omega_S$), which necessarily include two identical frequency arguments, whose permutation does not produce a distinguishable term. In this context $\omega_j$ and $-\omega_j$ must be considered different. The $2^{1-n}$ factor is linked to the $\frac{1}{2}$ factors appearing in Equations [6] and [7].

Although the order of frequencies in $\chi^{(3)}_{u\alpha\beta\gamma}(-\omega_\sigma;\ \omega_0\omega_P - \omega_S)$ has no special meaning itself, owing to the intrinsic permutation symmetry, we will keep it fixed (Raman convention), using $\omega_P$ and $-\omega_S$ as second and third arguments. Besides the intrinsic permutation symmetry, the medium macroscopic symmetry imposes further restrictions on the tensor index of nonvanishing, independent components $\chi^{(3)}_{u\alpha\beta\gamma}$. A very important result is that the $n$th-order susceptibility vanishes for even $n$ in media showing inversion symmetry and $\boldsymbol{\chi}^{(3)}$ contributes the lowest-order nonlinearity. For isotropic media, it can be shown that $\chi^{(3)}_{u\alpha\beta\gamma}$ vanishes if some Cartesian index appears an odd number of times in the subscript.

The symmetry properties of $\chi^{(3)}_{u\alpha\beta\gamma}$ in any medium can be established considering that it transforms as a fourth-rank polar tensor under the macroscopic symmetry operations. The symmetry of the microscopic polarizable units can be used to simplify the microscopic expression of $\chi^{(3)}_{\mu\alpha\beta\gamma}$.

### Microscopic origin of the nonlinear Raman susceptibility

We must now relate $\boldsymbol{\chi}^{(3)}$ to molecular eigenstates and transition moments. This can be accomplished through standard quantum-mechanical perturbation methods. Only a sketch of the steps involved will be given here.

The microscopic origin of the nonlinear response is the distortion induced in the molecular charge distribution due to the electrical field. The presence of a microscopic dipole produces a macroscopic polarization in the unit volume $\boldsymbol{P} = N\langle \varepsilon r\rangle$, where $N$ is the number density of polarizable units and $\langle \varepsilon r\rangle$ the expectation value of the dipole moment induced in each unit. In order to evaluate $\langle \varepsilon r\rangle$ we will use the density matrix formalism, because it is the easiest way to relate microscopic properties to macroscopic ones and to cope with macroscopic coherence effects. In the absence of fields, the medium is supposed to be described by an unperturbed Hamiltonian $\mathbf{H}_0$ and to be at equilibrium. When the fields are applied, the field–matter interaction contributes a time-dependent term $\boldsymbol{V}(t) = -\boldsymbol{E}(t)\boldsymbol{P}(t)$ to the global energy. The evolution of the system under this perturbation can be described through the equation of motion of the density operator:

$$\mathrm{i}\hbar\frac{\partial\boldsymbol{\rho}(t)}{\partial t} = [\mathbf{H}(t),\boldsymbol{\rho}(t)] - \mathrm{i}\hbar\boldsymbol{\Gamma} \qquad [13]$$

where $\mathbf{H}(t) = \mathbf{H}_0 + \boldsymbol{V}(t)$ and the term $\mathrm{i}\hbar\Gamma$ is introduced to include relaxation. The evolution of the matrix density elements is given by:

Diagonal elements:

$$\mathrm{i}\hbar\frac{\partial\rho_{ii}(t)}{\partial t} = [\boldsymbol{V}(t),\boldsymbol{\rho}(t)]_{ii} - \mathrm{i}\hbar\Gamma_{ii}\rho_{ii}(t)$$

Nondiagonal elements

$$\mathrm{i}\hbar\frac{\partial\rho_{ij}(t)}{\partial t} = [\boldsymbol{V}(t),\boldsymbol{\rho}(t)] + \hbar\rho_{ij}(t)(\omega_{ij} - \mathrm{i}\Gamma_{ij}) \qquad [14]$$

where $\omega_{ij} = (E_i - E_j)/\hbar$ and $\mathrm{G}_{ij} = \frac{1}{2}(\mathrm{G}_{ii}+\mathrm{G}_{jj}) + \mathrm{G}^{\mathrm{proper}}_{ij}$, $\mathrm{G}_{ij}$ being the homogeneous width of level $i$. $\mathrm{G}^{\mathrm{proper}}_{ij}$ includes the effect of dephasing processes that destroy the coherence between states $i$ and $j$ without altering their populations. Once we solve Equation [14] for $\rho_{ij}(t)$, the dipole moment can be calculated as $\langle \varepsilon r\rangle = e\mathrm{T}\boldsymbol{r}(\rho(t)\boldsymbol{r})$.

Diagonal elements $\rho_{ii}(t)$ represent the fractional population of state $i$. Equation [14] shows that the population factors change from their initial values owing to the perturbation, while the energy relaxation mechanisms tend to restore them to the equilibrium distribution with a rate constant $\Gamma_{ii} \sim 1/T_1$.

Nondiagonal terms in Equation [14] produce a macroscopic polarization; $\rho_{ij}(t)$ departs from zero at equilibrium and includes an oscillatory term at the frequency $\omega_{ij}$. The existence of a nondiagonal term $\rho_{ij}(t) \neq 0$ is linked to the creation of a coherent superposition of states $i$ and $j$ due to the perturbation. This coherence is broken through the interaction with the molecules of the bath, with a rate constant of the order of $\Gamma_{ij} = 1/T_2$. Usually this dephasing is much faster than the energy relaxation and $T_2 << T_1$.

Two distinct strategies exist for solving Equation [13]. The best-suited to our previous treatment starts by expanding the density matrix in a power series in the perturbation. By identifying terms we arrive at

$$\boldsymbol{P}^{(3)}(t) = Ne\ \mathrm{Tr}\ (\boldsymbol{\rho}^{(3)}(t)\ \boldsymbol{r}) \qquad [15]$$

from which we can evaluate $\boldsymbol{P}^{(3)}(t)$ once we have obtained $\boldsymbol{\rho}(t)$ iteratively, after expanding $\boldsymbol{\rho}$ in a basis of eigenstates of $\mathbf{H}_0$. In dense media local field factors must be included to relate the microscopic field at the molecule site to the macroscopic field of Equation [2].

This approach correctly describes the nonresonant nonlinear response and, with the inclusion of decay terms, allows treatment of Raman resonances as far as the density matrix remains close to its equilibrium value.

In the case of strong resonance, i.e. high-intensity fields closely tuned to narrow transitions, the perturbative series does not converge and this approach will fail. An alternative solution uses a simplified two-level or three-level atom model, along with effective multiphoton operators, and solves Equation [14] without using perturbation techniques. The result is a field-dependent *effective* susceptibility, that can describe population transfer and saturation effects. It can be related to $\chi^{(n)}$ by including correction factors that are a function of the amplitude, the transition moments and the detuning from exact resonance.

The final result for the perturbative approach is given without proof. Using a time-ordered, double diagrammatic technique, it can be shown that $\chi^{(3)}$ includes a sum of 48 terms, each corresponding to a diagram that specifies a time-ordered sequence of individual interactions. Quantum interference among these paths leads to a partial cancellation whenever proper dephasing is not relevant.

Among the remaining 24 terms, we select those with Raman-resonant denominators. The other terms can be considered a weakly dispersive nonresonant background for the Raman process. Typical resonant and nonresonant terms are shown in **Table 2**.

Let us now suppose that only two states $g$ and $f$ are connected through a Raman-type resonance ($\Omega_{fg} - (\omega_P - \omega_S) \simeq 0$), the rest of the denominators being well removed from zero. By inspection, it is possible to collect terms that start at either of these two levels and contribute to the spectral features of $\chi^{(3)}$. After referring all the frequencies to the lower $g$ level, we arrive at

$$\begin{aligned}\chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma;\omega_0\omega_P-\omega_S) &= \chi^{(3)NR}_{\mu\alpha\beta\gamma} + \frac{N}{6\epsilon_0\hbar^3}\sum_{g,f}\frac{1}{\Omega_{fg}-(\omega_P-\omega_S)}\\ &\times\sum_j\left(\frac{R^\mu_{gj}R^\alpha_{jf}}{\Omega_{jg}-\omega_\sigma}+\frac{R^\alpha_{gj}R^\mu_{jf}}{\Omega^*_{jg}+\omega_0}\right)\\ &\times\sum_k\left[\rho^{(0)}_{gg}\left(\frac{R^\beta_{gk}R^\gamma_{kf}}{\Omega_{kg}-\omega_P}+\frac{R^\gamma_{gk}R^\beta_{kf}}{\Omega_{kg}+\omega_S}\right)\right.\\ &\left.-\rho^{(0)}_{ff}\left(\frac{R^\beta_{gk}R^\gamma_{kf}}{\Omega^*_{kg}-\omega_P}+\frac{R^\gamma_{gk}R^\beta_{kf}}{\Omega^*_{kg}+\omega_S}\right)\right]\end{aligned} \qquad [16]$$

where $N$ is the total number density of the chromophore. The summation in $j$ and $k$ runs over all molecular levels, whereas that in $g$, $f$ runs over all pairs close to Raman resonance, which give rise to close-lying lines. When there is more than one chromophore, an additional summation over species should be included.

If an optical frequency $\omega_a$ is close to one-photon electronic resonance with an intermediate state $j$, the terms including the denominator $\Omega_{jg} - \omega_a \simeq 0$ will be dominant and we obtain a good approximation by retaining only these (resonant nonlinear Raman spectroscopy). In the absence of such additional electronic resonances, we can neglect the damping coefficient for all denominators whose absolute value is far from zero and write them in terms of real Raman scattering tensor components for the transition from

**Table 2** Some terms in the microscopic expression of the third-order non-linear susceptibility for Raman resonance

$$\begin{aligned}\chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma;\omega_0\omega_P-\omega_S) &= \frac{N}{6\varepsilon_0\hbar^3}\sum_{gkfj}\rho^{(0)}_{gg}\\ &\times\left\{\frac{1}{[\Omega_{fg}-(\omega_P-\omega_S)](\Omega_{kg}+\omega_S)}\right.\\ &\times\left(\frac{(R_{gk})_\gamma(R_{kf})_\beta(R_{fj})_\alpha(R_{jg})_\mu}{\Omega_{jg}-\omega_\sigma}+\frac{(R_{gk})_\gamma(R_{kf})_\beta(R_{fj})_\alpha(R_{jg})_\mu}{\Omega^*_{jg}+\omega_0}\right)\\ &+\frac{1}{[\Omega_{fg}-(\omega_P-\omega_S)](\Omega_{kg}-\omega_P)}\\ &\times\left(\frac{(R_{gk})_\beta(R_{kf})_\gamma(R_{fj})_\alpha(R_{jg})_\mu}{\Omega_{jg}-\omega_\sigma}+\frac{(R_{gk})_\beta(R_{kf})_\gamma(R_{fj})_\mu(R_{jg})_\alpha}{\Omega^*_{jg}+\omega_0}\right)\\ &+\text{ other resonant terms}^a\\ &+\frac{1}{[\Omega^*_{fg}+(\omega_0+\omega_P)](\Omega^*_{jg}+\omega_P)}\\ &\times\left(\frac{(R_{gk})_\mu(R_{kf})_\gamma(R_{fj})_\alpha(R_{jg})_\beta}{\Omega^*_{kg}+\omega_\sigma}+\frac{(R_{gk})_\gamma(R_{kf})_\mu(R_{fj})_\alpha(R_{jg})_\beta}{\Omega_{kg}+\omega_S}\right)\\ &+\frac{1}{[\Omega^*_{fg}+(\omega_0+\omega_P)](\Omega^*_{jg}+\omega_0)}\\ &\times\left(\frac{(R_{gk})_\mu(R_{kf})_\gamma(R_{fj})_\beta(R_{jg})_\alpha}{\Omega^*_{kg}+\omega_\sigma}+\frac{(R_{gk})_\gamma(R_{kf})_\mu(R_{fj})_\beta(R_{jg})_\alpha}{\Omega_{kg}+\omega_S}\right)\\ &\left.+\text{ other nonresonant terms}^b\right\}\end{aligned}$$

$\Omega_{ab} = (E_a - E_b)/\hbar - i\Gamma_{ab}$

$(R_{ab})_\mu$ = component along $\mu$ of the transition dipole matrix element $\langle b|\,er\,|a\rangle$

[a] For $\omega_P - \omega_S \approx \omega_R$ as the only frequency combination resonant with a molecular transition, Raman resonant terms dominate the nonlinear response and its change across the resonance defines the spectral shape. Additional one-photon resonance is possible through the denominators such as $(\Omega_{kg} - \omega_P)$ whenever input frequencies approach electronic transitions. The spectral behaviour in this case is dominated by a few terms in the summation.

[b] The sum of all nonresonant terms constitutes a weakly dispersive background. Some terms are nonresonant for a Raman transition but can give rise to two- or three-photon absorption for a different set of input frequencies.

$g$ to $f$ excited by the laser at frequency

$$\omega - (\alpha_{\delta\epsilon}(\omega))_{gf} = \frac{1}{\hbar}\sum_j \left(\frac{R^{\epsilon}_{gj}R^{\delta}_{jf}}{\omega_{jg}-\omega} + \frac{R^{\delta}_{gj}R^{\epsilon}_{jf}}{\omega_{jf}+\omega}\right)$$

We can further simplify by ignoring the weak dispersion of $(\alpha_{\delta\varepsilon}(\omega))_{gf}$ and dropping the $gf$ subscript. The final expression is shown on **Table 1**. We see that, in the presence of resonances, the third-order susceptibility is generally complex:

$$\chi^{(3)} = \chi^{(3)\mathrm{NR}} + \chi'^{(3)\mathrm{R}} + \mathrm{i}\chi''^{(3)\mathrm{R}} \qquad [17]$$

The non-resonant part $\chi^{(3)\mathrm{NR}}$ is real and includes the contribution of every species present. It will usually show a weak dispersion, being approximately constant for extended frequency ranges. In this case, the permutation of the frequency arguments alone does not alter much the numerical value of the susceptibility. This is the *Kleinman's symmetry conjecture* which, combined with the strict intrinsic permutation symmetry allows to freely permute the Cartesian subscripts in $\chi^{(3)\mathrm{NR}}$. This property, although not exact, is of great relevance for the polarization nonlinear Raman techniques.

$\chi^{(3)\mathrm{R}}$ has both real and imaginary parts, and is additive if different transitions from the same or other chromophore overlap. Disregarding multiplicative factors, the dispersion of $\chi^{(3)\mathrm{R}}$ has the form

$$\chi'^{(3)\mathrm{R}} \sim \sum_{g,f} \alpha^2(\rho^{(0)}_{gg} - \rho^{(0)}_{ff})\frac{\Delta_{fg}}{\Delta^2_{fg}+\Gamma^2_{fg}}$$
$$\chi''^{(3)\mathrm{R}} \sim \pm\sum_{g,f} \alpha^2(\rho^{(0)}_{gg} - \rho^{(0)}_{ff})\frac{\Delta_{fg}}{\Delta^2_{fg}+\Gamma^2_{fg}} \qquad [18]$$

The imaginary part has a Lorentzian profile of halfwidth $\Gamma_{fg}$, centred at exact resonance. This Lorentzian line shape comes from the assumed form for the damping terms, and more realistic models should be used in Equation [14] when studying line shapes. The spectral shape is the same as for spontaneous Raman lines. The real part of $\chi^{(3)\mathrm{R}}$ is multiplied by the detuning, and shows a dispersive line shape. $\chi^{(3)\mathrm{R}}$ depends on the population difference, instead of just the population of the initial state as does spontaneous Raman. This dependence can lead to saturation whenever appreciable population is transferred to the excited level.

### Generation, evolution and detection of the signal field

The last step is to analyse how the induced nonlinear polarization creates a new wave or interferes with the existing ones to generate the nonlinear Raman signal. The nonlinear wave propagation equation (taking $\boldsymbol{P}^{(2)} = 0$) is

$$\nabla \times \nabla \times \boldsymbol{E}(\omega) = \frac{\omega^2}{c^2}\boldsymbol{\varepsilon}(\omega)\boldsymbol{E}(\omega) + \omega^2\mu_0\boldsymbol{P}^{(3)}(\omega) \qquad [19]$$

where $\boldsymbol{\varepsilon}(\omega) = 1 + \chi^{(1)}(-\omega;\omega)$ is the dielectric tensor that accounts for the linear propagation. We look for solutions in the form of running plane waves propagating collinearly along the $z$ axis, orthogonal to the electric field vector, and suppose that the amplitude of the signal field will change slowly along $z$. We will restrict ourselves to the small-gain regime, in which neither appreciable depletion of the input beams nor population transfer occurs, leading to an amplitude rate change given by

$$\frac{\partial E_{\omega_\sigma\mu}(z)}{\partial z} = \frac{\mathrm{i}\omega_\sigma}{2cn_\sigma}\left(\frac{D}{4}\right)\chi^{(3)}_{\mu\alpha\beta\gamma}(-\omega_\sigma;\omega_0\omega_\mathrm{P}-\omega_\mathrm{S})$$
$$E_{\omega_0\alpha}(z)E_{\omega_\mathrm{P}\beta}(z)E^*_{\omega_\mathrm{S}\gamma}(z)\exp(\mathrm{i}\Delta kz) \qquad [20]$$

where $\Delta k = (k_0 + k_\mathrm{P} - k_\mathrm{S}) - k_\sigma$ is the phase mismatch, i.e. the difference between the wave vector of the 'polarization wave' at frequency $\omega_\sigma$ and that of the signal field at the same frequency. The former value is a combination of the refractive index at the input frequencies, whereas the latter depends on the refractive index at $\omega_\sigma$. Owing to linear dispersion, these two values are in principle different. If $\Delta k \neq 0$ the change rate of the signal amplitude shows an oscillatory behaviour, whereas for $\Delta k = 0$ (perfect phase match) a monotonic change is obtained (increase or decrease depending upon the argument of the complex susceptibility) leading to a higher signal. In the small-signal regime we can integrate Equation [20] and evaluate the intensity at the detector, given by $I\omega_j = \frac{1}{2}\varepsilon_0 cn_j|E_j|^2$.

We are at last able to give a plane wave signal expression for the different nonlinear Raman techniques, collected in **Table 1**.

In several techniques we monitor the intensity at a new frequency (normal CARS, Raman resonant four-wave mixing), a new polarization (RIKES) or both (polarization CARS). The generated wave does not interfere with the input waves and the output intensity depends on the product of three input

intensities, the square of the interaction length, within the coherence length of lasers, and the square of the effective susceptibility. If $\omega_\sigma = \omega_P$ or $\omega_S$, the phase matching is automatic; in other cases it must be set by controlling the refraction index or crossing the beams.

The quadratic dependence with $\chi^{(3)}$ leads to interference among neighbouring lines, from the same or different species, and with the nonresonant background in dense media, which can lead to strongly distorted line-shapes. The different tensorial properties of $\chi^{(3)R}$ and $\chi^{(3)NR}$ can be exploited to reduce background interference, by using input waves with circular polarization or linear along a direction chosen to cancel the nonresonant part.

In other techniques, we monitor the change in intensity for the pump or Stokes beam. We can solve Equation [20] considering the input field as variable, leading to an exponential amplitude change that can be approximated by a linear change in the small signal regime, or keeping it constant as before. For the latter method we must include the interference of signal and input fields at the detector, which shows a close connection with heterodyne detection techniques. Both methods lead to the same result in the small-gain regime, showing a decrease in the pump intensity (SRL) and an increase in the Stokes intensity (SRG), linear with the imaginary resonant part of the susceptibility $\chi^{(3)\prime R}$, i.e. proportional to the spontaneous Raman spectrum and free from interference effects, automatically phase-matched and dependent on the product of pump and Stokes intensities.

The signal expressions in **Table 1** allow for an analysis of the spectral properties of the relevant techniques. In actual experiments, beams are focused and the detailed geometry must be considered.

## List of symbols

$A_\omega$ = complex pulse area; $D$ = number of distinguishable permutations of the pairs $(\omega, \alpha)$; $\langle er \rangle$ = expectation value of the molecular dipole operator; $E_i$ = energy of level $i$; $\boldsymbol{E}(t)$ = electric field vector; $E_\mu(t)$ = component along $\mu$ of the electric field vector; $\boldsymbol{E}(\omega)$ = amplitude of the Fourier component of $\boldsymbol{E}$ at frequency $\omega$; $\boldsymbol{E}_\omega$ = amplitude of the monochromatic wave at frequency $\omega$; $E_{\omega\mu}$ = component along $\mu$ of $\boldsymbol{E}_\omega$; $\boldsymbol{E}_\omega(t)$ = envelope of the quasi-monochromatic field at central frequency $\omega$; $\mathrm{H}(t)$ = time-dependent total Hamiltonian; $\mathrm{H}_0$ = unperturbed Hamiltonian in the absence of electromagnetic fields; $I_\omega$ = intensity of the monochromatic wave at frequency $\omega$; $N$ = number density of molecules; $n_\sigma$ = refractive index at frequency $\omega_\sigma$; $\boldsymbol{P}(t)$ = macroscopic polarization vector; $\boldsymbol{P}^{(n)}$ = $n$-th order term in the series expansion of $\boldsymbol{P}$ in powers of the applied field; $P_\mu^{(3)}$ = component along $\mu$ of $\boldsymbol{P}^{(3)}$; $\boldsymbol{P}(\omega)$ = amplitude of the Fourier component of $\boldsymbol{P}$ at frequency $\omega$; $P_\mu^{(3)}(\omega)$ = amplitude of the monochromatic polarization wave at frequency $\omega$; $\boldsymbol{P}_\omega^{(3)}(t)$ = envelope of the quasi-monochromatic third-order polarization component at central frequency $\omega$; $P_{\omega\mu}^{(3)}$ = component along $\mu$ of $P_\omega^{(3)}$; $\mathbf{R}^{(n)}(\tau_1, \tau_2, \tau_3)$ = third-order nonlinear response function tensor; $R_{\mu\alpha\beta\gamma}^{(3)}(\tau_1, \tau_2, \tau_3)$ = component of $\mathbf{R}^{(3)}$ with Cartesian index $\mu\alpha\beta\gamma$; $R_{ab}^{\mu}$ = component along $\mu$ of the transition dipole matrix element for transitions from $a$ to $b$; $T_1$ = population relaxation characteristic time; $T_2$ = phase relaxation characteristic time; $V(t)$ = interaction operator; $(\boldsymbol{\alpha}(\omega))_{gf}$ = Raman scattering tensor for transitions from state $g$ to state $f$; $(\alpha_{\delta\varepsilon}(\omega))_{gf}$ = component $\delta\varepsilon$ of the Raman scattering tensor; $\boldsymbol{\Gamma}$ = relaxation matrix; $\Gamma_{ij}$ = matrix element of $\boldsymbol{\Gamma}$; $\Gamma_{ij}^{\mathrm{proper}}$ = pure dephasing term associated with element $\rho_{ij}$ of the density matrix; $\Delta\boldsymbol{k}$ = phase mismatch vector; $\boldsymbol{\rho}(t)$ = density matrix of the material system; $\rho_{ij}(t)$ = matrix element of $\boldsymbol{\rho}_k$; $\boldsymbol{\rho}^{(3)}(t)$ = third order term in the development of the matrix density in power series of the interaction; $\tau_\mathrm{p}$ = characteristic time of the radiation (it represents the period of the fastest amplitude change in a laser pulse); $\boldsymbol{\Phi}^{(3)}_{-\omega_\sigma;\,\omega_0\omega_\mathrm{p}-\omega_\mathrm{s}}(\tau_1, \tau_2, \tau_3)$ = third-order nonlinear envelope response function; $\boldsymbol{\chi}^{(n)}(-\omega_\sigma; \omega_1\omega_2\omega_3)$ = $n$th-order nonlinear susceptibility tensor; $\chi_{\mu\alpha\beta\gamma}^{(3)}(-\omega_\sigma; \omega_1\omega_2\omega_3)$ = component of $\boldsymbol{\chi}^{(3)}$ with Cartesian index $\mu\alpha\beta\gamma$; $\boldsymbol{\chi}^{(3)NR}$ = nonresonant part of the third-order nonlinear susceptibility; $\boldsymbol{\chi}'^{(3)R}$ = real resonant part of the third-order nonlinear susceptibility; $\boldsymbol{\chi}''^{(3)R}$ = imaginary resonant part of the third-order nonlinear susceptibility; $\omega$ = frequency of the radiation, $E = \hbar\omega$; $\omega_\mathrm{R}$ = frequency corresponding to a Raman transition; $\omega_{ij} = (E_i - E_j)/\hbar$ = frequency corresponding to the energy gap for levels $i$ and $j$; $\Omega_{ab} = \omega_{ab} - \mathrm{i}\Gamma_{ab}$ = compact complex notation for detuning and relaxation; $\hbar$ = Planck constant/$2\pi$.

*See also:* **Nonlinear Optical Properties; Nonlinear Raman Spectroscopy, Applications; Nonlinear Raman Spectroscopy, Instruments; Raman Optical Activity, Applications; Raman Optical Activity, Theory; Raman Spectrometers.**

## Further reading

(This work is mainly based on Santos (1996) and Butcher (1990.)

Bloembergen N (1965) *Nonlinear Optics*. Reading, MA: WA Benjamin.

Butcher PN and Cotter D (1990) *The Elements of Nonlinear Optics*. Cambridge: Cambridge University Press.

Clark RJH and Hester RE (eds) (1988) *Advances in Nonlinear Spectroscopy*. Chichester: Wiley.
Eesley GL (1981) *Coherent Raman Spectroscopy*. Oxford: Pergamon Press.
Lee D and Albrecht AC (1993) In: Prigogine I and Rice SA (eds) *Advances in Chemical Physics*, Volume LXXXIII, pp 43–87. New York: Wiley.
Levenson D and Kano SS (1988) *Introduction to Nonlinear Laser Spectroscopy*. San Diego: Academic Press.
Mukamel S (1995) *Principles of Nonlinear Optical Spectroscopy*. New York: Oxford University Press.
Prior Y (1984) A complete expression for the third-order susceptibility. Perturbative and diagrammatic approaches. *IEEE Journal of Quantum Electronics* **QE-20**: 37–42.
Santos Gómez J (1996) Coherent Raman spectroscopy. In: Laserna JJ (ed) *Modern Techniques in Raman Spectroscopy*, pp 305–340. Chichester: Wiley.

# NQR, Applications

*See* **Nuclear Quadrupole Resonance, Applications.**

# NQR, Spectrometers

*See* **Nuclear Quadrupole Resonance, Instrumentation.**

# NQR, Theory

*See* **Nuclear Quadrupole Resonance, Theory.**

# Nuclear Overhauser Effect

**Anil Kumar** and **R Christy Rani Grace**,
Indian Institute of Science, Bangalore, India

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Theory |

## Introduction

The Nuclear Overhauser Effect (NOE) has become one of the key effects for obtaining the structures of molecules, especially biomolecules, in solution by NMR spectroscopy. The effect was first enunciated by its discoverer, Overhauser, as a large polarization of nuclear spins in metals on saturation of electrons to which they are coupled. Neither its discoverer, nor the famous people in the audience at the conference where he first presented his calculation in 1953, nor Charles Slichter, who took upon himself the task of experimentally verifying Overhauser's assertions, could have imagined that the effect would become the backbone of modern NMR. The effect remained as a curiosity and was sparingly used until its transformation into the nuclear–nuclear Overhauser effect, in which one of the nuclear spins is saturated and the nearby nuclear spin shows changes in the intensity of its NMR signal. For small molecules, in the fast motion limit, the Overhauser effect is positive for nuclear spins with the same sign of $\gamma$ (the magnetogyric ratio) and negative for those with opposite sign.

The nuclear–nuclear Overhauser effect, in its early days, was used for enhancing the polarization of less-sensitive nuclei, such as natural abundant $^{13}C$ or $^{15}N$ and in double-resonance experiments for obtaining additional information on relaxation of molecules. It was utilized, in conjunction with the spin–lattice relaxation of various carbons, for the characterization of anisotropy of molecular reorientations. However, its largest utility continued to be for obtaining information on molecular conformations, by selectively saturating the magnetization of a specific proton in the molecule and monitoring the changes in the intensity of resonances of nearby protons.

In continuous wave (CW) NMR, one of the problems is the separation of the Overhauser effect from the coherent effects of radiofrequency (RF) irradiation. For example, even at very low RF field (~ 0.02 Hz) CW double resonance experiments, the observed intensities can only be explained by including coherent effects, as splitting or broadening of directly connected transitions. In higher power experiments, the lines are split and several coherent effects are simultaneously present. The development of one-dimensional (1D) Fourier transform NMR immediately freed the nuclear Overhauser effect from the shackles of such coherent effects. The irradiation can be time-shared, such that the saturating RF field is applied before the observation pulse. Since, during observation, the radiofrequency field and hence the coherent effects are absent, the spectrum contains information only on relaxation. The development of 1D Fourier transform NMR, with its added sensitivity advantage, opened the field of NMR, to the study of proton NMR of biomolecules. As soon as such spectra were recorded, the assignment of the large number of resonances became a major problem. Selective decoupling and NOEs were employed, respectively, to assign and to obtain information on the secondary structures of peptides and proteins. In small peptides, it became a matter of routine to pick-up a few characteristic NOEs and assign the secondary structure. For larger peptides, containing 20–50 amino acid residues, efforts were directed towards obtaining a large number of (50–100) selective decoupling and NOE experiments, to assign the spectra and to obtain the three-dimensional structure of the molecules. However, it was hard work to manually plot, measure and analyse these large number of 1D spectra. The development of two-dimensional (2D) NMR removed this obstacle. COSY NMR spectroscopy with its many improvements, is used for identifying coherently coupled spins, while the 2D NOESY is used for obtaining large-scale information on distances between nearby spins. The success of this algorithm prompted workers to look at even larger molecules (of relative molecular mass 10–15 kDa), crowding even the 2D spectra, and necessitating addition of a third or even fourth dimension to remove overlap. This was achieved by labelling the biomolecules with either $^{13}C$ and/or $^{15}N$ and spreading the 2D information into a third or fourth dimension, as a function of attached carbon or nitrogen chemical shifts. Selective polarization transfers (from a selected heteronucleus to a selected proton) followed by NOE are being developed to enhance the resolution and to reduce the burden of 2D or 3D NMR experiments.

## Theory

In the NOE, the non-equilibrium magnetization of a spin is shared by nearby spins through mutual dipole–dipole relaxation. Any other relaxation mechanism, if present, attenuates the effect. For example, dissolved paramagnetic oxygen has a deleterious effect on the NOE. It is therefore necessary to have clean solvents and often it becomes necessary to remove the dissolved oxygen for a quantitative estimate of NOEs.

The source spin is brought to a non-equilibrium state, for example by selectively saturating its magnetization by a low power, long pulse. During this irradiation, the transfer of magnetization to the other nearby spins also takes place and this can be monitored as a function of the irradiation time. The irradiation can be performed for a sufficiently long time, so that a steady-state is reached between the irradiation and the relaxation and this gives rise to what is known as 'steady-state NOE'. However, such experiments, particularly in case of slowly reorienting molecules, give NOEs between fairly distant spins in the molecule, through migration of magnetization over several intervening spins (the phenomenon is known as 'spin-diffusion'). This has the disadvantage that the NOE loses selectivity of information. A simple alternative is to reduce the irradiation time or to monitor the growth of the NOE, as a function of irradiation time. This has been called 'truncated-driven-NOE'. In both these experiments, the NOE transfer takes place in the presence of the RF field. In another experiment, the source spin is selectively inverted (by a selective $\pi$ pulse) and the NOE is monitored in the absence of the RF field. This is known as 'transient-NOE', and was first suggested by Solomon.

The 2D NOE (NOESY) experiment is formally equivalent to a large number of transient NOE experiments. Each cross-section of a 2D NOESY spectrum is equivalent to a 1D selective transient NOE experiment, in which the peak corresponding to the diagonal is selectively inverted. The NOE transfer in both NOESY and 1D transient NOE takes place in the absence of the RF field and, theoretically, the two experiments are identical (except for a factor of 2 in the intensities of 1D experiments which will be explained later). The theory of the NOE, given in the following subsections is divided into three parts, (i) steady-state NOE, (ii) transient NOE and (iii) 2D NOE (NOESY). Later sections describe the NOE in the rotating frame, heteronuclear NOE and transferred NOE.

In this article the NOE between various spins will be discussed, ignoring the spin–spin coupling between various spins and the associated splitting of lines into multiplet components. A spin will be assumed to be irradiated (saturated or inverted) as a whole and detected as a whole. This description leaves out many interesting effects, such as selective saturation of various transitions, direct pumping effects of irradiation or inversion, multiplet effects between various transitions of a spin arising from cross-correlations and strong coupling effects.

### Steady-state NOE

Consider a two spin-$\frac{1}{2}$ system, with spins $I$ and $S$ relaxation coupled to each other. Relaxation of these two spins in the presence of mutual cross-relaxation is described by coupled rate equations known as Solomon's equations:

$$\frac{\mathrm{d}I_z(t)}{\mathrm{d}t} = -\rho_I(I_z(t) - I_z^{\mathrm{eq}}) - \sigma_{IS}(S_z(t) - S_z^{\mathrm{eq}}) \qquad [1]$$

$$\frac{\mathrm{d}S_z(t)}{\mathrm{d}t} = -\sigma_{SI}(I_z(t) - I_z^{\mathrm{eq}}) - \rho_S(S_z(t) - S_z^{\mathrm{eq}}) \qquad [2]$$

where $I_z(t)$ and $S_z(t)$ describe the instantaneous values of the longitudinal magnetization of the spins $I$ and $S$ and $I_z^{\mathrm{eq}}$ and $S_z^{\mathrm{eq}}$ represent their equilibrium values respectively. The terms $\rho_I$ and $\rho_S$ are the self-relaxation rates of the two spins, respectively, and $\sigma_{IS} = \sigma_{SI}$ is the mutual cross-relaxation (NOE) rate between the spins. All these rates depend on the mechanics of relaxation operative in the spin system and the state of mobility of the molecule in which they are embedded.

On selective saturation by RF irradiation at the resonance frequency of one of the spins (say $S$), it is generally assumed that the magnetization of the irradiated spin is fully saturated at all times, that is $S_z(t) = 0$, for all values of $t$. The steady state solution of the Solomon's equations is then obtained by making the time derivatives on the left-hand-side of Equations [1] and [2] as equal to zero, yielding the steady state value of $I_z$ from Equation [1] as

$$I_z^{SS} - I_z^{\mathrm{eq}} = \frac{\sigma_{IS}}{\rho_I} S_z^{\mathrm{eq}} \qquad [3]$$

Since $I_z^{\mathrm{eq}}/S_z^{\mathrm{eq}} = \gamma_I/\gamma_S$, one obtains the well-known steady-state-NOE as

$$\eta = \frac{I_z^{SS} - I_z^{\mathrm{eq}}}{I_z^{\mathrm{eq}}} = \frac{\gamma_S}{\gamma_I}\frac{\sigma_{IS}}{\rho_I} \qquad [4]$$

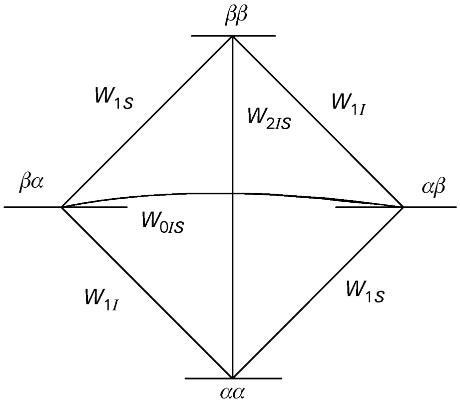



**Figure 1** Energy-level diagram of a two spin-½ system, showing the transition probabilities and the spin states. The designation αα means both the spins are in the +½ state, while *αβ* means the first spin (I) is in the +½ state and the second spin (*S*) is in −½ state. $W_{1I}$ gives the single quantum transition probability for the I spin when it changes its spin state. $W_0$ and $W_2$ are the zero and double-quantum transition probabilities, respectively, by which both the I and S spins simultaneously change their spin states.

For two relaxation coupled spins, each spin-$\frac{1}{2}$, the energy level diagram consists of four levels (**Figure 1**).

From this diagram it is seen that $W_{1I}$ and $W_{1S}$ are the single quantum transition probabilities for the spin *I* and *S* respectively. The term $W_{0IS}$ is the zero quantum transition probability for the flip-flop process, in which the two spins are in opposite spin states and exchange their spin states; $W_{2IS}$ is the double quantum transition probability, by which the two spins in the similar states simultaneously change their spin states. In homonuclear systems, these probabilities are proportional to the spectral densities at 0, $\omega$ and $2\omega$ frequencies, respectively, for $W_0$, $W_1$ and $W_2$, where $\omega$ is the Larmor frequency. Using population rate equations of the four levels of the two-spin system and using these transition probabilities connecting the four levels, it is straightforward to show that

$$\begin{aligned}\sigma_{IS} &= W_{2IS} - W_{0IS}\\ \rho_I &= W_{0IS} + 2W_{1I} + W_{2IS}\\ \rho_S &= W_{0IS} + 2W_{1S} + W_{2IS}\end{aligned} \quad [5]$$

Using Equation [5] in Equation [4] the NOE on spin *I*, on saturating spin *S*, is obtained as

$$\eta_I\{S\} = \left(\frac{\gamma_S}{\gamma_I}\right)\frac{W_{2IS} - W_{0IS}}{W_{0IS} + 2W_{1I} + W_{2IS}} \quad [6]$$

The origin of the NOE can now be explained in the following way. After the selected spins absorb a certain amount of energy from the RF field, they pass this energy to the other spins and to other degrees of freedom, known as the lattice. In other words, they start to relax. Mutual dipolar interaction between the nearby spins provides an effective mechanism for exchange of energy among the spins and to the lattice. However, if these spins are embedded in a rigid solid, the dipolar interaction is time-independent and does not couple the spins to the lattice, except in the case of like spins (same $\gamma$), the dipolar interaction in rigid solids does provide a mechanism of rapidly passing the energy from one spin to the next (via its energy conserving flip-flop, $W_0$, terms). This process, known as spin-diffusion in solids, is responsible for migration of magnetization across the whole sample and to relaxation sinks such as rapidly rotating methyl groups or rapidly relaxing paramagnetic centres. On the other hand, if the interacting spins are on molecules in solution, which are rapidly reorienting, the dipolar interaction becomes time dependent and, if there are molecular motions at the Larmor or twice the Larmor frequency, it couples the spins to the rotational degrees of freedom of the molecule and transfers the spin energy to the lattice. The lattice is assumed to have infinite heat capacity and remains at thermal equilibrium at room or sample temperature. The coupling of the spin to the lattice at the Larmor frequency is provided additionally by several other mechanisms, such as modulation of chemical shift anisotropy, quadrupolar interaction and paramagnetic relaxation centres. All these contribute to $W_1$ but not to $W_0$ and $W_2$, and attenuate the NOE, as seen from Equation [6]. It is only the dipolar interaction that has two spin terms, such as $I_+S_-$, $I_-S_+$ and $I_+S_+$, $I_-S_-$, which contribute respectively, to $W_0$ and $W_2$ and in turn to the NOE. The first two are flip-flop terms or zero-quantum terms, in which the two spins exchange energy among each other without requiring exchange of energy with the lattice (for $\gamma_I = \gamma_S$). Such processes dominate for slowly reorienting molecules such as biomolecules; and therefore the NOE is very significant in such cases, as hardly any energy is lost to the lattice. Here $W_0$ dominates over all the other terms, and as seen from Equation [6], the NOE is large and negative (−1). The third and the fourth terms, $I_+S_+$ and $I_-S_-$ are the two quantum terms and the exchange energy of the spins with the lattice at twice the Larmor frequency. In this case both the spins flip simultaneously either up or down. If molecular motions are fast enough to have spectral densities at $2\omega$, then these processes dominate. It can

be shown that in such circumstances the NOE is positive and small, with a maximum of +0.5.

The spectral densities for the various motional regimes are schematically sketched in **Figure 2**.

In the slow motion or long correlation time limit ($\tau_c >> 1/\omega_0$), the spectral densities are maximum near the zero frequency and minimum at $\omega_0$ and $2\omega_0$, while for fast reorienting molecules $\tau_c << 1/\omega_0$, the spectral density is almost equal for all the frequencies till $\omega_c = 1/\tau_c$, where $\tau_c$ is the correlation time for molecular reorientations. For the dipolar interaction between spins $I$ and $S$, the transition probabilities for isotropic molecular reorientations are obtained as

$$\begin{aligned} W_{0IS} &= \frac{1}{10}k^2 \frac{\tau_c}{1+(\omega_I-\omega_S)^2\tau_c^2} \\ W_{1I} &= \frac{3}{20}k^2 \frac{\tau_c}{1+\omega_I^2\tau_c^2} \\ W_{1S} &= \frac{3}{20}k^2 \frac{\tau_c}{1+\omega_S^2\tau_c^2} \\ W_{2IS} &= \frac{6}{10}k^2 \frac{\tau_c}{1+(\omega_I+\omega_S)^2\tau_c^2} \end{aligned} \quad [7]$$

where $k = (\mu_0/4\pi)\gamma_A\gamma_X\hbar r_{IS}^{-3}$.

**Homonuclear case** When the two spins are of the same species, i.e. $\gamma_I = \gamma_S$, then the steady state homonuclear NOE for selective saturation of one of the spins with the other spin having a well-resolved chemical shift is obtained by substituting in Equation [6] $\omega_I = \omega_S = \omega$,

$$\eta_{\max} = \frac{5+\omega^2\tau_c^2-4\omega^4\tau_c^4}{10+23\omega^2\tau_c^2+4\omega^4\tau_c^4} \quad [8]$$

This curve is plotted in **Figure 3**, and has the value +0.5 for fast reorientation with respect to $\omega$ (short correlation limit $\omega\tau_c << 1$) and –1 for slow reorientations (long correlation limit, $\omega\tau_c >> 1$).

These can also be seen for homonuclear spins, by direct substitution of $W$s in the two limits in Equation [6]. For example, for the short correlation time limit, $W_0 : W_1 : W_2 :: 2 : 3 : 12$, yielding $\eta = \frac{1}{2}$. For the long correlation time limit, $W_1$ and $W_2$ are negligible and $\eta = -1$. From Equation [8] and from **Figure 3** it is seen that $\eta = 0$, for $\omega\tau_c = 1.118$. This is often called the critical correlation time limit. In this limit, $W_0 = W_2$, $\sigma = 0$ and the laboratory frame NOE is zero. Experiments have been designed in which both the spins are spin-locked along an axis making an angle $\theta$ from the **z**-direction, yielding a non-zero NOE. This will be described in the section on 'rotating-frame NOE'. The short correlation time limit, $\omega\tau_c << 1$ is generally applicable to small molecules (relative molecular mass ≤ 1 kDa) at low spectrometer frequencies, as was often the case in the early days of NMR. For larger size molecules, which reorient slowly ($\omega\tau_c >> 1$) the NOE is negative, but very useful in obtaining information on the proximity of the spins.

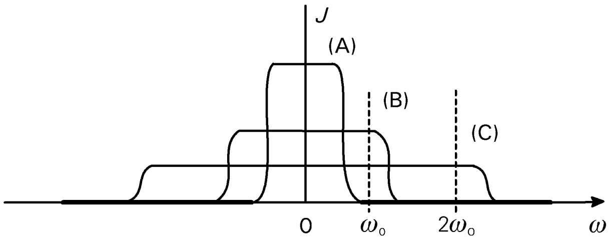



**Figure 2** Spectral density $J(\omega)$ for three values of correlation time, plotted as a function of frequency $\omega$. The spectral density has a cutoff frequency $\omega_c = 1/\tau_c$, where $\tau_c$ is the correlation time of molecular reorientations. As molecular reorientations become faster, $\tau_c$ decreases and the spectral density dispersion becomes flatter. The terms $T_1$, $T_2$ and NOE depend on the value of the spectral densities at 0, $\omega_0$ and $2\omega_0$, where $\omega_0$ is the Larmor frequency. (A) Spectral density for slowly reorienting molecules which have long correlation times ($\tau_c >> 1/\omega_0$). In such cases the spectral density has a negligible value at $\omega_0$ and $2\omega_0$, but large values at low frequencies. (B) Spectral density for intermediate values of correlation times, for which $\tau_c \approx 1/\omega_0$. (C) Spectral density for small molecules undergoing fast reorientation, which have short correlation times ($\tau_c << 1/\omega_0$) and the spectral density has nearly equal values from 0 to $2\omega_0$. Since the area under the curves is constant, the spectral density has different magnitudes at each frequency in the above three cases.

The observation of negative NOE among the spins with the same sign of $\gamma$ and in the short correlation time limit, however, gave rise to some excitement during the late 1960s. It was soon found that the negative NOE was owing to what was called a 'three-spin-effect'. The explanation is as follows. When the first spin is saturated, the second spin is enhanced in intensity. By logical extension this means that the third spin is reduced in intensity. This of course requires that the three spins are in almost a linear configuration, such that the direct positive NOE from the first spin to the third is less than the transmitted NOE via the second spin. The observation of the negative three-spin-effect for homonuclear spins in the short correlation time limit is thus a signature of the linearity or near-linearity of the three spins. The other observation of a negative NOE between two protons, without the intervention of a third spin in a polypeptide by Balaram and coworkers, was the first evidence of molecules

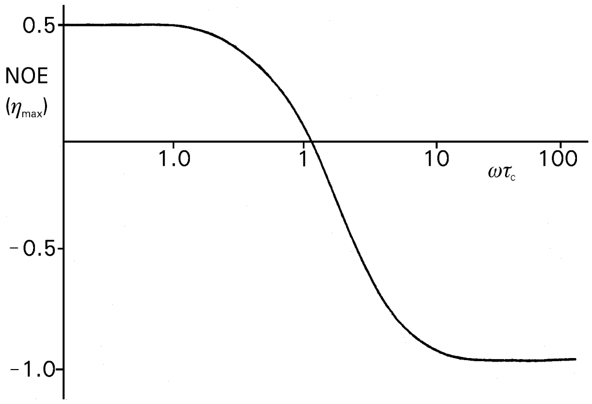



**Figure 3** Variation of homonuclear NOE enhancement, Equation [8], plotted as a function of $\omega\tau_c$. Note the logarithmic scale of $\omega\tau_c$. For small molecules with short $\tau_c$, the limiting value for $\eta_{max}$ is +0.5. In practice, since relaxation mechanisms other than dipolar are also efficient in this extreme narrowing limit, positive enhancements as large as this are rarely observed. For large molecules with long $\tau_c$, the limiting value of $\eta_{max}$ is −1. Biomolecules and small molecules in viscous solvents come into this category and generally give significant NOEs. In the central region where $\eta_{max}$ varies rapidly with $\omega\tau_c$, the NOE enhancements depend on the spectrometer frequency and the molecular tumbling rate.The value of $\eta_{max}$ passes through a null for $\omega\tau_c \approx 1$.

tumbling at rates slower than the Larmor frequency. Ever since then, larger molecules have been studied by NMR spectrometers operating at higher frequencies, and negative NOEs between protons have become the backbone of NMR research. There is an additional advantage of negative NOEs, which becomes apparent in the transient NOE experiment, described in the next section.

## Transient NOE

In the transient NOE experiment, the perturbed spin is selectively inverted rather than saturated. Since this can be done in times short compared with $T_1$ and $T_2$ of the spins, the NOE during the pulse is neglected and the migration of the magnetization is observed after the pulse, in the absence of the RF field. The time evolution of magnetization is obtained using Solomon's Equations [1] and [2]. Substituting the initial condition, $S_z(0) = -S_z^{eq}$, one obtains a biexponential time evolution for $I_z$ and $S_z$ magnetization, assuming $\rho_I = \rho_S = \rho_{IS}$, as

$$\frac{S_z(t) - S_z^{eq}}{S_z^{eq}} = -e^{-(\rho_{IS}+\sigma_{IS})t} - e^{-(\rho_{IS}-\sigma_{IS})t} \qquad [9]$$

$$\frac{I_z(t) - I_z^{eq}}{I_z^{eq}} = -e^{-(\rho_{IS}+\sigma_{IS})t} + e^{-(\rho_{IS}-\sigma_{IS})t} \qquad [10]$$

While the recovery of the inverted spin $S_z$ to its equilibrium value is biexponential, that of $I_z$ magnetization shows an initial growth and then a decay. Equation [10] can be rewritten as

$$\frac{I_z(t) - I_z^{eq}}{I_z^{eq}} = [1 - e^{-2\sigma_{IS}t}]e^{-(\rho_{IS}-\sigma_{IS})t} \qquad [11]$$

This gives the NOE on spin *I*, which is positive for positive s and negative for negative s. The NOE on spin *I* grows, reaches a maximum and then decays to zero. The initial rate of growth is obtained by differentiating Equation [11] with respect to time and taking the limit $t \to 0$, the so-called initial rate approximation, yielding

$$\frac{d}{dt}\left(\frac{I_z(t) - I_z^{eq}}{I_z^{eq}}\right)_{t\to 0} = 2\sigma_{IS} \qquad [12]$$

The initial rate of growth of NOE thus gives a direct measure of the cross-relaxation rate $\sigma_{IS}$ and by inference the distance $r_{IS}$ ($\sigma_{IS}$ is proportional to $r_{IS}^{-6}$)

The advantage of the transient NOE experiment is that the transport of magnetization takes place in the absence of RF irradiation and also the dynamics of Solomon's equations are identical to the 2D NOE experiment, described in the next section. The driven experiment has no 2D analogue and the solution given earlier in Equation [4] has the limitation that the details of saturation are not included. In fact if one uses Equation [2] instead of Equation [1], for the steady-state solution, by substituting $dS_z(t)/dt = 0$, $S_z(t) = 0$, one obtains a wrong result

$$\frac{I_z(t) - I_z^{eq}}{S_z^{eq}} = \frac{\rho_S}{\sigma_{IS}} \qquad [13]$$

Boulat and Bodenhausen earlier and recently Karthik have shown how this anomaly can be removed by describing the details of the saturation process of spin *S*.

## 2D NOE (NOESY)

Selective saturation or inversion of each transition out of a large number of closely spaced transitions of various protons of a protein is both tedious and difficult. The development of the two-dimensional nuclear Overhauser effect (2D NOE or NOESY) experiment was therefore a turning point in the application of NMR for the study of biomolecules. The 2D NOE experiment, **Figure 4A**, uses three 90°

pulses. The first pulse flips the magnetization of all the spins in the molecule to the transverse plane, which are then allowed to evolve during a frequency labelling period $t_1$. The second pulse flips the magnetization to the longitudinal direction. This non-equilibrium magnetization is allowed to relax and gives NOEs according to Equations [1] and [2] during the mixing period $\tau_m$. The state of the spin system is read by the third 90° pulse with the signal being recorded as a function of the time variable $t_2$. A complete set of data $s(t_1, t_2)$, after Fourier transformation with respect to both $t_1$ and $t_2$ yields a 2D spectrum (**Figure 4B**).

The magnetization components, which have the same frequencies in time domains $t_1$ and $t_2$, lie along the diagonal of the NOESY spectrum, while those magnetization components which have crossed over from one spin to another spin during the mixing time $\tau_m$, owing to the NOE, lie on both sides of the diagonal and are called the cross-peaks. Indeed, it has been shown that a cross-section parallel to $\omega_2$ is identical, except for a factor of 2, to a 1D transient difference NOE experiment in which the peak on the diagonal is selectively inverted. Suppression of transverse magnetization during $\tau_m$ and the growth of longitudinal magnetization during $\tau_m$ (giving rise to an axial peak in the 2D spectrum) can be achieved either by phase cycling or by the use of gradients. The factor of 2 difference arises between the 1D and 2D experiment owing to the fact that the axial magnetization does not contribute in the 2D, but it does in the 1D experiment. However, the rate of transfer and hence the information content is identical in the two experiments. The first 2D NOE spectrum of a small protein, basic pancreatic trypsin inhibitor (BPTI) is shown in **Figure 5**.

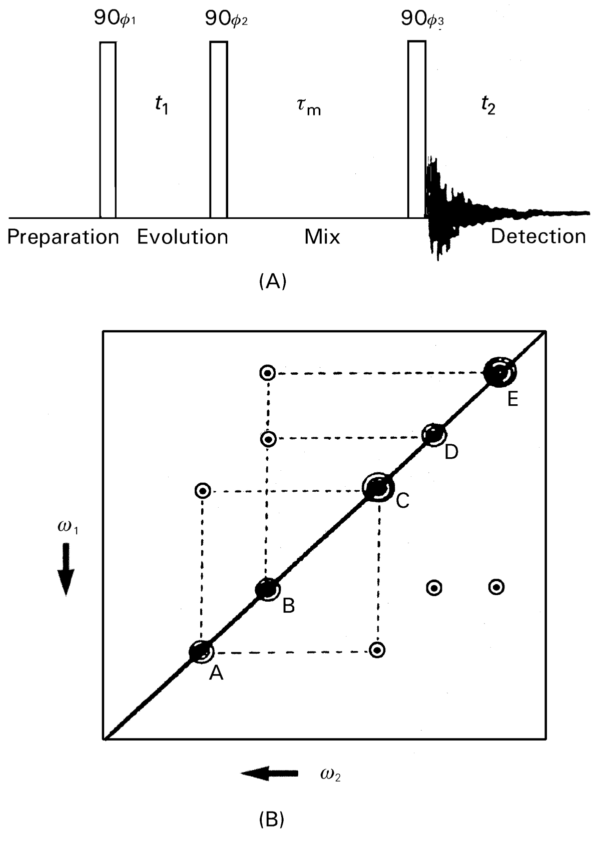



**Figure 4** (A) The 2D NOESY pulse sequence, which uses three 90° pulses. The times $t_1$ and $t_2$ are the evolution and detection periods, respectively; $\tau_m$ is the mixing time during which only longitudinal magnetization is retained, either by gradients or by cycling the phases ($\phi_1$, $\phi_2$, $\phi_3$) of the pulses. (B) Schematic NOESY spectrum, showing that in such spectra the NOEs are manifested as cross-peaks between the various spins, the resonances of which lie on the diagonal.

A large number of cross-peaks are observed, each indicating a NOE or exchange between the protons of the corresponding diagonals. Exchange also gives rise to cross-peaks identical to the negative NOE (same sign as the diagonal) and can only be distinguished from NOE by the use of the rotating frame NOE method, to be discussed later. The data in the NOESY experiment are analysed by measuring the peak volume of a cross-peak in a series of 2D experiments with various mixing times $\tau_m$. The initial rate of growth of the NOE is directly proportional to $1/r^6$. To obtain the proportionality constant, the rate is compared with some known distance. However, this procedure is strictly valid for only two relaxation coupled spins. Since there are in general, several spins simultaneously coupled by relaxation to each other, the two-spin problem is generalized in the following manner.

**Generalized Solomon's equations** If there are more than two spins, relaxation coupled to each other, then the two-spin Solomon's Equations [1] and [2] can be generalized, in the following manner

$$\frac{d(I_{zi}(t) - I_{zi}^{eq})}{dt} = -\rho_i(I_{zi} - I_{zi}^{eq}) - \sum_{j \neq i} \sigma_{ij}(I_{zj} - I_{zj}^{eq}) \quad [14]$$

which states that there are $n$ $(i = 1 \ldots n)$ coupled equations describing the self-relaxation of each spin via the term $\rho_i$ and the cross-relaxation with the other spins via $\sigma_{ij}$. This is a straightforward extension of the pairwise interaction and it neglects any cross-terms (cross-correlations) that may be present between the relaxation of various spins. It has been shown that the effect of cross-correlations on the total NOE (the average NOE, neglecting differences in the intensities of various transitions of a spin) are

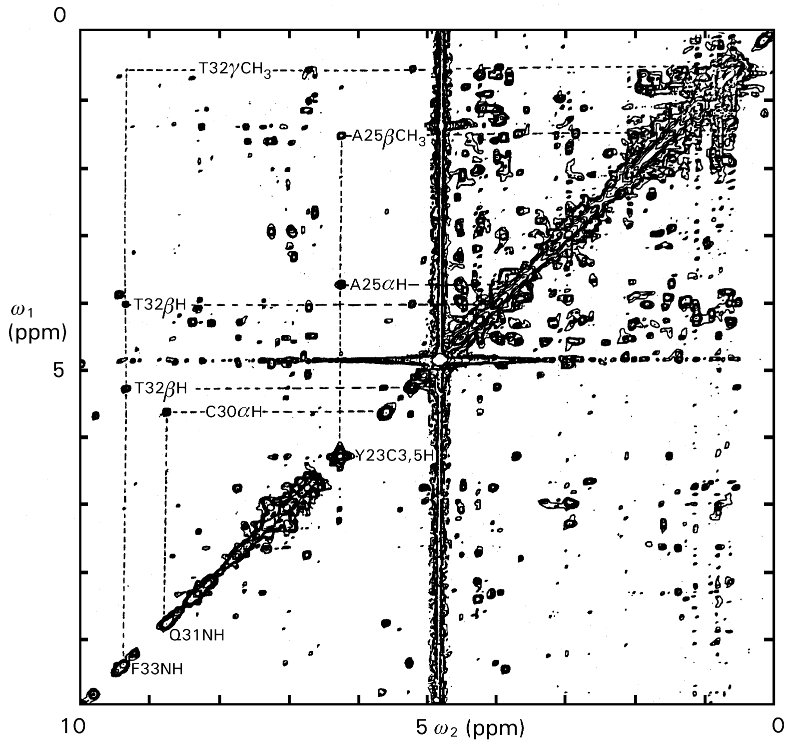



**Figure 5** Contour plot of the [1]H NOESY spectrum at 360 MHz of the basic pancreatic trypsin inhibitor. The protein concentration was 0.02 M, solvent $D_2O$, pD = 3.8, $T$ = 18°C. The spectral width was 4000 Hz; 512 data points were used in each dimension; 56 transients were accumulated for each value of $t_1$. The mixing time $\tau_m$ was 100 ms. The absolute value spectrum, obtained after digital filtering in both dimensions with a shifted sine bell, is shown. NOE connectivities for selected amino acid residues are indicated by the broken lines. Reproduced with permission of Academic Press from Kumar A, Ernst RR and Wüthrich K (1980) *Biochemistry and Biophysics Research Communications* **95**: 1.

generally small. In this review, the effect of cross-correlation on NOE will not be dealt with and the reader is referred to several articles on this field, including a recent review by the authors. The general solution of Equation [14] is a multiexponential time evolution of magnetizations which are coupled to each other. Once the geometry of the spins is known, it is possible to calculate the various rates of Equation [14] and compute the expected auto- and cross-peak intensities of the NOESY experiment. These computed intensities are then iteratively fitted to the observed intensities, to converge on possible structure(s) consistent with the observed intensities. Often there are differences between the computed intensities and the observed intensities that arise from internal motions, which in turn when built into the calculations give information on the internal motions. Anisotropy of reorientation of the molecules also plays a role and can also be built into the NOE calculations.

Three-dimensional structures of a large number of biomolecules (proteins, peptides, oligonucleotides and oligosaccharides) have been obtained using information derived from the NOESY experiments. The reader is referred to the 1986 book by Wüthrich and the *Encyclopedia of Nuclear Magnetic Resonance* for an exhaustive review up to 1996.

## ROESY

For intermediate size molecules for which $\omega\tau_c \approx 1$, the zero-quantum ($W_{0IS}$) and double-quantum ($W_{2IS}$) transition probabilities are nearly equal and the cross-relaxation rate $\sigma_{IS}$ approaches zero. In such cases there is no NOE. Bothner-By came up with the fascinating idea of doing cross-relaxation in the transverse plane by spin-locking the magnetization, using RF fields. He named the technique as CAMEL-SPIN (cross-relaxation appropriate for minimolecules emulated by locked spins), but it is now known as ROESY (rotating frame NOESY). Both 1D and 2D versions are known, and are shown schematically in **Figure 6**. The method will be explained using the 1D experiment; the 2D logic is identical. The first 90° pulse (**Figure 6A**) flips the magnetization to the

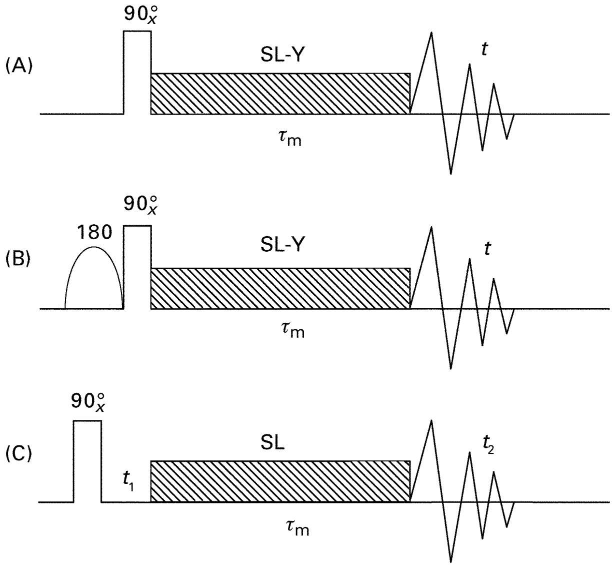



**Figure 6** Rotating frame NOE pulse sequences. The 1D experiment requires two sequences, represented by (A) and (B). (A) is the reference experiment in which a $90^\circ_x$ non-selective pulse is applied on all the spins, followed by a spin-lock along the *y*-direction for a time $\tau_m$ and the state of the spin system is detected. (B) The control experiment in which a selective 180° pulse, inverts the magnetization of the spin from which the NOE is to be observed before the $90^\circ_x$ pulse and the experiment is continued as (A). The 1D NOE spectrum is the difference between the spectra obtained with the sequence (A) and (B). (C) The 2D ROESY sequence. The times $t_1$ and $t_2$ are the evolution and detection periods and $\tau_m$ is the mixing time. SL refers to the low power spin-locking RF field.

transverse plane, followed by a spin-lock using a 90° phase-shift.

The spin-locked magnetization of the two spins, which differ in chemical shifts, decay and cross-relax according to the rate equations,

$$\frac{\mathrm{d}m_I(t)}{\mathrm{d}t} = \rho_I m_I(t) - \sigma_{IS} m_S(t)$$
$$\frac{\mathrm{d}m_S(t)}{\mathrm{d}t} = \sigma_{SI} m_S(t) - \rho_S m_S(t) \qquad [15]$$

where $m_I$ and $m_S$ are the transverse magnetization of spins $I$ and $S$ spin-locked along the RF field. For the 1D case, two experiments (reference and control) are performed. For the reference experiment (**Figure 6A**), the initial condition is $m_I(0) = m_S(0) = 1$, while for the control experiment (**Figure 6B**), in which the magnetization of spin $S$ is selectively inverted just before non-selective spin-lock, the initial condition is $m_I(0) = 1$, $m_S(0) = -1$. The solution of Equation [15] for these two cases can be written, respectively, (for $\rho_I = \rho_S = \rho$ and $\sigma_{IS} = \sigma_{SI} = \sigma$) as

$$m_I^r(t) = \mathrm{e}^{-(\rho+\sigma)t}$$
$$m_I^c(t) = \mathrm{e}^{-(\rho-\sigma)t} \qquad [16]$$

The NOE is the difference of the two experiments and is therefore given by

$$\eta = \mathrm{e}^{-(\rho-\sigma)t} - \mathrm{e}^{-(\rho+\sigma)t} \qquad [17]$$

$\eta_{\max}$ is obtained as

$$\eta_{\max} = \left(\frac{\rho+\sigma}{\rho-\sigma}\right)^{-\frac{(\rho-\sigma)}{2\sigma}} - \left(\frac{\rho+\sigma}{\rho-\sigma}\right)^{-\frac{(\rho+\sigma)}{2\sigma}}$$
$$t_{\max} = \frac{1}{2\sigma}\ln\left(\frac{\rho+\sigma}{\rho-\sigma}\right) \qquad [18]$$

$\rho$ and $\sigma$ in the above equations for homonuclear spins are obtained as

$$\rho = \frac{3}{4}\gamma^4\hbar^2\left[\frac{3}{8}J(0) + \frac{15}{4}J(\omega) + \frac{3}{8}J(2\omega)\right]$$
$$\sigma = \frac{3}{4}\gamma^4\hbar^2\left[\frac{1}{6}J(0) + \frac{3}{2}J(\omega)\right] \qquad [19]$$

For isotropic Brownian motion, the spectral densities are obtained as

$$J(0) = \frac{1}{r^6}\frac{24}{15}\tau_c,$$
$$J(\omega) = \frac{1}{r^6}\frac{4}{15}\left(\frac{\tau_c}{1+\omega^2\tau_c^2}\right)$$
$$J(2\omega) = \frac{1}{r^6}\frac{16}{15}\left(\frac{\tau_c}{1+4\omega^2\tau_c^2}\right) \qquad [20]$$

Using these values for the spectral densities, the expressions for $\rho$ and $\sigma$ in Equation [19] reduce to

$$\rho = \frac{1}{20}\left(\frac{\gamma^4\hbar^2}{r^6}\right)\left\{5\tau_c + \frac{9\tau_c}{1+\omega^2\tau_c^2} + \frac{6\tau_c}{1+4\omega^2\tau_c^2}\right\}$$
$$\sigma = \frac{1}{20}\left(\frac{\gamma^4\hbar^2}{r^6}\right)\left\{4\tau_c + \frac{6\tau_c}{1+\omega^2\tau_c^2}\right\} \qquad [21]$$

The maximum NOE for these conditions is plotted in **Figure 7**, which has a maximum value of 38.5% for $\omega\tau_c << 1$ and 67.5% for $\omega\tau_c >> 1$.

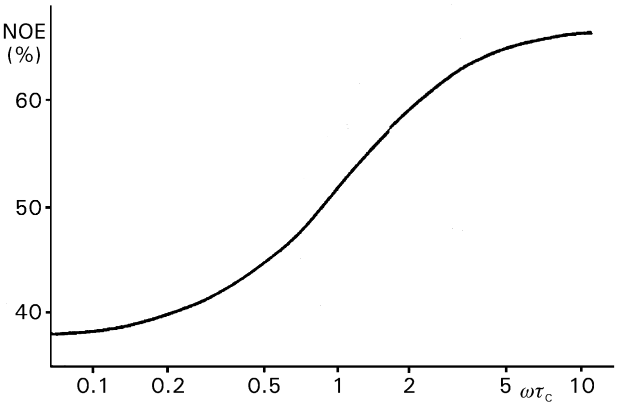



**Figure 7** Plot of maximum rotating frame NOE (from Equation (18)) for a homonuclear two spin-½ system as a function of $\omega\tau_c$. The rotating frame NOE is positive for all values of correlation time.

The NOE is positive and does not have a null for any correlation time. This result holds for the 2D experiment as well, with the diagonal and the cross-peaks having opposite signs (positive NOE). A typical 2D ROESY spectrum is shown in **Figure 8**.

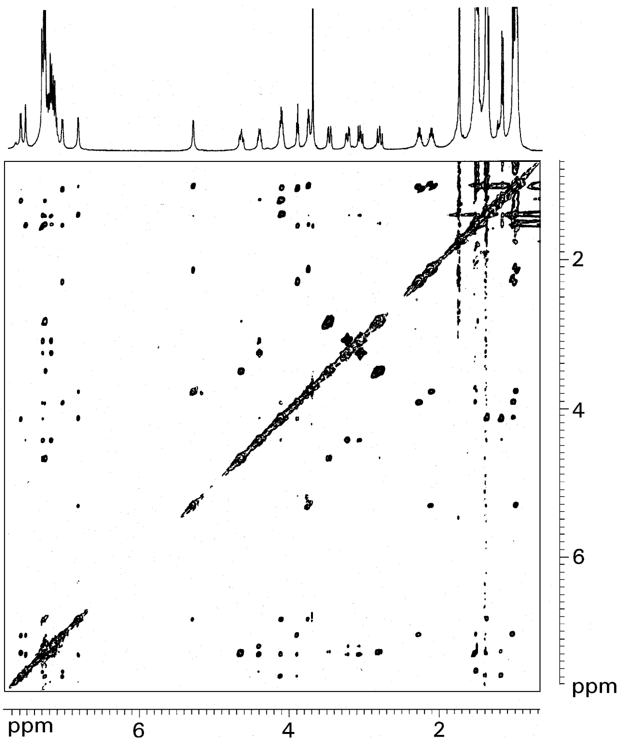



**Figure 8** Contour plot of the 2D $^1$H ROESY spectrum of a 0.5 M solution of Boc-Val-Ala-Phe-Aib-Val-Ala-Phe-Aib-OMe in $CDCl_3$, recorded at 400 MHz. A 2.25 kHz spin-lock field has been used during the 300 ms mixing period. 64 scans were performed for every $t_1$ value and 512 × 1k data were acquired. Zero filling was used to give a 1k × 1k size of the displayed absorptive part of the spectrum. The diagonal drawn is negative and the cross-peaks are positive. Unpublished results by Das C, Grace RCR and Balaram P.

ROESY has several advantages and disadvantages. The advantages are (i) the NOE is finite (non-zero) for all sizes of the molecules, (ii) a positive NOE means that there is leakage to the lattice present and hence the magnetization does not migrate over long distances. The limited spin-diffusion helps in selecting the nearest neighbours. (iii) The three-spin effect yields a negative NOE; this can be looked at as an advantage or a disadvantage. However, the disadvantages are (i) The ROESY intensities are sensitive to the magnitude of the spin-locking RF field and to their resonance offsets. (ii) There is also a coherence transfer due to *J*-coupling, known as TOCSY (total correlation spectroscopy). In fact an identical pulse scheme can also be used for obtaining a TOCSY spectrum, with which one identifies all the resonances which are *J*-coupled through bonds. For example, all the resonances of an amino acid residue could be identified by taking a cross-section at either the NH or $\alpha$H position. The salient features of ROESY, TOCSY and NOESY are listed in **Table 1**.

Some of these differences are therefore utilized for differentiating the TOCSY and ROESY peaks, in particular the strength of the spin-locking field and the mixing time, as well as the sign of the cross-peak. Experiments have also been designed for obtaining clean TOCSY as well as clean ROESY spectra.

**Table 1** Comparison of the salient features of ROESY, TOCSY and NOESY

| *Feature* | *ROESY* | *TOCSY* | *NOESY* |
|---|---|---|---|
| Net transfer | Yes | Yes | Yes |
| Pure absorptive | Yes | Yes (almost) | Yes |
| Sign with respect to the diagonal | Opposite (+ve NOE) | Same | Opposite for $\omega\tau_c << 1$ Same for $\omega\tau_c >> 1$ |
| Mixing time | Large (> 100 ms) | Small (<100 ms) | <500 ms |
| RF field amplitude needed | Small | Large | Not applicable |

## Heteronuclear NOE

So far we have discussed the NOE between nuclei having the same $\gamma$. The maximum use of these homonuclear NOEs has been for proton–proton NOEs. However, the heteronuclear NOE also has a long history. In the early days of NMR, the heteronuclear NOE was basically utilized for the enhancement of the intensities of low $\gamma$ nuclei. For example, the steady state NOE, on spin I on irradiation of spin S when spins $I$ and $S$ are mutually dipolar-relaxation coupled is given by Equations [4] and [6]. For $\omega\tau_c << 1$, $W_0{:}W_1{:}W_2{::}2{:}3{:}12$, $\sigma = \frac{1}{2}\rho$, yielding

$$\eta_I\{S\} = \frac{1}{2}\frac{\gamma_S}{\gamma_I} \qquad [22]$$

Thus for $^{13}C$ the NOE is 2, while for $^{15}N$ it is –5. The $^{13}C$ magnetization could thus be enhanced by 200 % by adding the NOE to carbon equilibrium magnetization. For $^{15}N$, the enhancement was –4, which is like a 400% increase in the signal. These enhancements have been and are routinely utilized in heteronuclear NMR experiments.

As $\omega\tau_c$ increases, the heteronuclear NOE decreases in magnitude (approaching zero for $\omega\tau_c >> 1$) since in the heteronuclear case the flip-flop process becomes non-energy conserving and requires spectral densities at a frequency equal to the difference in the Larmor frequencies of the two spins. Even the rotating frame NOE for heteronuclei is very small as $\sigma$ approaches zero. Therefore heteronuclear NOE has been used only for small molecules in the $\omega\tau_c << 1$ limit, and there is even a 2D version known as HOESY which uses identical logic as NOESY, except that the first two pulses are applied on the source spin and the third pulse on the target spin, with the detection on the target spin.

## Transferred NOE (TRNOE)

NOE measurements are also made on spin systems undergoing chemical exchange in which a small ligand molecule (peptide, drug molecule) exchanges between free and bound states with a large host molecule (protein). Under this condition, the rotational correlation time of the small molecule changes from a regime of $\omega\tau_c << 1$ in the free state to $\omega\tau_c >> 1$ in the bound state. While sharp resonances of the ligand are observed only from its free state, the bulk of the NOE takes place when the ligand is in the bound state, giving rise to negative NOE between the various protons of the ligand. This is known as the transferred NOE. The observed NOE is related to the various exchange rates, concentration ratios of the two molecules and the binding constants. The transfer NOE experiment is thus often exploited for obtaining information on the exchange kinetics of ligand–protein interactions.

## Conclusions

The nuclear Overhauser effect arises from mutual dipolar relaxation of nearby spins and is utilized for identifying spins which are close in proximity in reorienting molecules. The NOE thus provides information on the three-dimensional structures of the molecules in solution and has led to an alternate methodology of obtaining structure of biomolecules in solution as compared with those obtained by single crystal X-ray crystallography. Since the NOE is extremely sensitive to the modulation of the distance between the spins owing to internal motions, as well as to the anisotropy of molecular reorientations, this methodology has led to detailed information on internal motions, conformational flexibility, local order parameters, exchange between chemically inequivalent sites and anisotropic reorientations of the molecules, some of which are correlated to their biological functions. The field is growing rapidly with the help of modelling software, molecular mechanics calculations, development of newer experimental protocols such as isotopic labelling of spins ($^{13}C$ and $^{15}N$), measurement of $J$-coupling between heteronuclei, leading to conformational information, and the utilization of information contained in cross-correlations.

### Acknowledgement

RCR Grace acknowledges CSIR for financial support.

## List of symbols

$I_z(S_z)$ = longitudinal magnetization of spin $I$ (and $S$); $J$ = coupling constant, $m_I$ = transverse magnetization of spin $I$; $T_1$, $T_2$ = relaxation times; $W = T_1/T_2$ transition probability; $\gamma$ = magnetogyric ratio; $\eta$ = NOE; $\rho$ = self-relaxation rate; $\sigma_{IS}$ = mutual cross-relaxation rate; $\tau_c$ = correlation time; $\tau_m$ = mixing time; $\omega$ = Larmor frequency.

*See also:* **Chemical Exchange Effects in NMR; Macromolecule–Ligand Interactions Studied By NMR; Magnetic Resonance, Historical Perspective; NMR Pulse Sequences; NMR Relaxation Rates; Nucleic**

**Acids Studied Using NMR; Proteins Studied Using NMR Spectroscopy; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Peptides; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals; Two-Dimensional NMR Methods.**

## Further reading


Abragam A (1989) *Time Reversal, an Autobiography*, p 164. Oxford: Oxford University Press.

Balaram P, Bothner-By AA and Dadok JJ (1972). *Journal of the American Chemical Society* **94**: 4015.

Boulat B and Bodenhausen G (1992). *Journal of Chemical Physics* **97**: 6040.

Chiarpin E, Pelupessy P, Cutting B, Eykyn TR and Bodenhausen G (1998) *Ang Chemie* (in the press).

Dalvit C and Bodenhausen G (1990) *Advances in Magnetic Resonance* **14**: 1.

Ernst RR, Bodenhausen G and Wokaun A (1987) *Principles of Nuclear Magnetic Resonance in One and Two-dimensions*. Oxford: Clarendon Press.

Karthik G (1999) *Journal of Chemical Physics* **110**: 4992.

Kumar Anil, Grace RCR and Madhu PK (1998) *Progress in NMR Spectroscopy* (in the press).

Levy GC, Lichter R and Nelson GL (1980) *Carbon-13 Nuclear Magnetic Resonance for Organic Chemists*. NewYork: Wiley Interscience.

Nageswara Rao BDN (1970) Nuclear spin relaxation by double resonance. *Advances in Magnetic Resonance* **4**: 271.

Neuhaus D and Williamson M (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. NewYork: VCH.

Noggle JH and Schirmer RE (1971) *The Nuclear Overhauser Effect – Chemical Application*. London: Academic Press.

Overhauser AW (1996) In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, Vol 1, p 513. Chichester: Wiley.

Sanders JKM and Hunter BK (1987) *Modern NMR Spectroscopy*. Oxford: Oxford University Press.

Slichter CP (1978) *Principles of Magnetic Resonance*. NewYork: Springer-Verlag.

Wüthrich K (1976) *NMR in Biological Research: Peptides and Proteins*. Amsterdam: North Holland.

Wüthrich K (1986) *NMR of Proteins and Nucleic Acids*. NewYork: Wiley.


# Nuclear Quadrupole Resonance, Applications


**Oleg Kh Poleshchuk**, Tomsk Pedagogical University, Russia

**Jolanta N Latosińska**, Adam Mickiewicz University, Poznan, Poland


Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| **Applications** |

The NQR frequencies for the various nuclei vary from several kHz up to 1000 MHz. Their values depend on quadrupole moments of the nucleus, the valence electrons state and the type of chemical bond in which the studied atom participates. Using the NQR frequencies, the quadrupole coupling constant (QCC) and asymmetry parameter ($\eta$) can be calculated with either exact or approximate equations, according to the spin of the nuclei. For a polyvalent atom, the NQR frequencies depend on the coordination number and hybridization. The NQR frequency shifts for a single-valent atom in different environments can be classified as follows:

1. The greatest shifts of the NQR frequencies are determined by the valence electron state of the neighbouring atoms and can reach 1200–1500%. The changes of the NQR frequencies depend on ionic character, and for example for chlorine the lowest frequency corresponds to a chloride ion, whereas the highest corresponds to the chlorine atom in $ClF_3$.
2. Changes within the limits of one type of chemical bond (with the same kind of atom) can reach 40–50%. Thus the frequency changes of C–Cl bonds vary from 29 to 44 MHz.
3. The range of possible shifts of the NQR frequencies is reduced to 10–20% when only one class of compound is studied. In this case the shifts are determined by the surroundings and donor–acceptor properties of the substituents. For example, for halogen substituted benzenes, the changes of the frequencies for the C–Cl bonds are about 9% for C–Br bonds ~ 12% and for C–I bonds ~ 18%.
4. The changes of NQR frequencies caused by the occurrence of intramolecular and intermolecular interactions lie within the limits of 3–40%.

5. The shifts of NQR frequencies caused by crystal field effects in molecular crystals reach a maximum of 1.5–2%. However within a series of similar compounds this effect, as a rule, does not exceed 0.3%.

Such classification of NQR frequency shifts allows the determination of structural non-equivalencies from NQR spectra. Understanding how the frequency shifts depend on chemical non-equivalence and how they are determined by distinctions in distribution of electronic density in a free molecule is very important. Moreover, crystallographical non-equivalences are observed, and are a result of distinct additional contributions of the crystal field to the electric field gradients at a nucleus. It is obvious that the division of structural non-equivalencies into molecular and crystal types is not relevant in coordination and ionic crystals, in which there are no individual molecules. NQR, being highly sensitive to subtle changes in electron density distribution, provides diverse information on the structural and chemical properties of compounds.

## Molecular structure studies using NQR

When applied to structural investigations, NQR spectra may prove an effective tool for the preliminary study of crystal structure in the absence of detailed X-ray data. Such parameters as spectroscopic shifts, multiplicity, spectroscopic splitting, resonance line width, the temperature dependence of resonance frequencies and relaxation rates all afford useful structural information and provide insight into the factors determining the formation of certain structural types.

The violation of chemical equivalence of resonance atoms due to a change in chemical bonding, such as, for example, dimerization in group IIIA halogenides, leads to a significant splitting of the spectroscopic multiplet caused by a difference in the electronic structure of bridging and terminal atoms.

### Crystallographic structures

For crystallographically non-equivalent atoms the corresponding components of the electric field gradient (EFG) at the respective sites differ from each other in magnitude and direction due to the crystal field effect. This generally includes a contribution to the EFG of electrostatic forces between molecules, dispersion forces, intermolecular bonding and short-range repulsion forces. Physically non-equivalent sites differ from each other only in the direction of the EFG components, their magnitudes being identical. To distinguish between such sites, Zeeman analysis of the NQR spectrum is required.

**Table 1** The angles (°) between the directions of the C–Cl bonds and the crystal axes in 1,3,5-trichlorobenzene according to NQR and X-ray results

| *Chlorine position* | *Crystal axes* | | | | | |
|---|---|---|---|---|---|---|
| | *a* | | *b* | | *c* | |
| | *NQR* | *X-ray* | *NQR* | *X-ray* | *NQR* | *X-ray* |
| Cl-1 | 119.1 | 118.8 | 149.5 | 150 | 81.7 | 80.9 |
| Cl-3 | 64.1 | 65.0 | 31.3 | 30.2 | 73.7 | 74.3 |
| Cl-5 | 24.8 | 24.0 | 90.4 | 90.5 | 114.8 | 114.0 |

The intensities of spectroscopic lines are also important. They reflect the relative concentration of resonance nuclei at certain sites although one also has to take into account the transition probabilities and lifetimes of the energy states of the system investigated. The correspondence between the number and intensities of frequencies and the number of non-equivalent sites occupied by a resonant atom in a crystal lattice is very helpful in a preliminary structure study made with the use of NQR.

NQR single-crystal Zeeman analysis can provide information about special point positions occupied by the quadrupole atoms. This Zeeman analysis determines the orientation of the EFG components with respect to the crystal axes, which essentially facilitates the most difficult and time-consuming stage of X-ray analysis. **Table 1** gives a comparison of the angle values between the crystal axes and the EFG $z$-axis at the chlorine atoms in 1,3,5-trichlorobenzene (determined by the two methods).

As one can see from **Table 1** the angle values obtained by the two methods show rather good numerical agreement. To illustrate more completely the type of structural information that can be obtained by NQR spectroscopy we consider in more detail an NQR study of $BiCl_3$, whose structure is known. Two chlorine atoms are involved in bridging to two other Bi atoms while another chlorine atom is involved in bridging to only one other Bi atom. The $^{35}Cl$ and $^{209}Bi$ NQR parameters of $BiCl_3$ measured by means of the single-crystal Zeeman method are listed in **Table 2**.

The assignment of $^{35}Cl$ resonances is unambiguous owing to direct correspondence of the lower frequency line of double intensity to the two Cl atoms which are crystallographically nearly equivalent. The asymmetry parameter for the chlorine atoms has been determined, using the zero-splitting cone. According to the results the $BiCl_3$ crystal belongs to the Laue symmetry of $D_{2h}$ and therefore to the orthorhombic crystal class. The Cl(1) and Bi atoms

**Table 2** $^{35}Cl$ and $^{209}Bi$ NQR spectra of $BiCl_3$

| Isotope | T (K) | Transition frequencies (MHz) $\frac{1}{2}-\frac{3}{2}$ | $\frac{3}{2}-\frac{5}{2}$ | $\frac{5}{2}-\frac{7}{2}$ | $\frac{7}{2}-\frac{9}{2}$ | $e^2Qqh^{-1}$ (MHz) | η (%) | Assignment |
|---|---|---|---|---|---|---|---|---|
| $^{35}Cl$ | 291 | 15.952(2) | – | – | – | 30.960 | 43.1 | Cl(1) |
| | | 19.173(1) | | | | 38.145 | 17.8 | Cl(2) |
| $^{209}B$ | 294 | 31.865 | 25.132 | 37.362 | 51.776 | 318.900 | 55.5 | |

occupy special point positions in the lattice since only two non-equivalent directions of the corresponding EFG components have been detected for them, while four orientations have been found for the EFG components at the site of Cl(2,3) atoms. The number of observed $^{35}Cl$ resonances suggests that the molecule possesses a mirror plane and therefore the centrosymmetric group $P_{nma}$. The lower frequency resonance of double intensity is then assigned to the chlorines out of the mirror plane while the higher frequency line corresponds to one lying in the mirror plane. The shorter (i.e. more covalent) bond length of the latter corresponds to the higher resonance frequency.

Another example is provided by the $^{35}Cl$ results for phosphorus pentachloride. In the gas phase this is known to be a trigonal bipyramid but the usual solid $PCl_5$ is likewise known to be an ionic crystal, $(PCl_4)^+(PCl_6)^-$. In accordance with this, and the detailed crystal structure, there are four resonances at high frequency that correspond to the $(PCl_4)^+$ group and six at a lower frequency that correspond to the $(PCl_6)^-$ group (**Table 3**).

It has recently been observed that quenching of the vapour of $PCl_5$ gives rise to a new metastable crystalline phase which can be preserved essentially indefinitely at low temperatures. The $^{35}Cl$ spectrum of this has two low frequencies, corresponding to the axial chlorine atoms, and three identical higher frequencies, corresponding to the equatorial substituents, which is strong evidence that this new phase is the corresponding molecular solid. Another example is taken from the chemistry of antimony pentachloride. It is known that the molecule of antimony pentachloride at 210 K has a trigonal bipyramid structure, and at 77 K it is a dimer. In **Table 3** the experimental $^{35}Cl$ NQR frequencies and their assignment to equatorial and axial chlorine atoms are given. In the dimer molecule, bridging chlorine atoms have much lower NQR frequencies, as in dimer molecules of group IIIA compounds. From **Table 3** it is also clear that in the monomer the NQR frequencies of equatorial chlorine atoms are greater than the axial ones.

In the dimer, the ratio of $^{35}Cl$ NQR frequencies is in agreement with the results of *ab initio* calculations. In dimers of transition elements, such as $NbCl_5$ and $TaCl_5$, the $^{35}Cl$ NQR frequencies of equatorial chlorine atoms are higher than those of axial, whilst those of bridging chlorine atoms are higher, on average, than those of terminal chlorines. A similar inversion of NQR frequencies in dimers of transition and non-transition elements is explained by a significant multiplicity of the metal–halogen bonds and hence by electron transfer from p-valent orbitals of the halogen atoms to the vacant d-orbitals of the central atom.

The fact that the difference between chemically non-equivalent atomic positions is readily revealed by NQR spectroscopic splitting may be utilized to identify geometric isomers. Octahedral complexes of tin tetrachloride, $[SnCl_4{\cdot}L_2]$ exist as either *cis* or *trans* isomers. In the *cis* isomers the axial and equatorial non-equivalence produces considerable splitting in the NQR spectra. In the *trans* complexes, all four chlorine atoms are chemically equivalent with identical electron density distribution. Splitting in the NQR spectra of these isomers arises therefore from the crystallographic non-equivalence of the chlorine positions. Indeed, the observed NQR splitting of two complexes (**Table 4**) provides evidence

**Table 3** $^{35}Cl$ quadrupole resonance frequencies for phosphorus and antimony pentachlorides

| Compound | $\nu\,^{35}Cl$ (MHz) | Assignment |
|---|---|---|
| $(PCl_4)^+(PCl_6)^-$ | 28.395, 28.711, 29.027, | $(PCl_6)^-$ |
| | 30.060, 30.457, 30.572 | $(PCl_6)^-$ |
| | 32.279, 32.384, 32.420 | $(PCl_4)^+$ |
| | 32.602 | $(PCl_4)^+$ |
| $PCl_5$ | 29.242, 29.274 | Axial |
| | 33.751 | Equatorial |
| $SbCl_5$ | 30.18 | Equatorial |
| | 27.85 | Axial |
| $Sb_2Cl_{10}$ | 27.76 | Equatorial |
| | 30.18 | Axial |
| | 18.76 | Bridging |

**Table 4** $^{35}$Cl Quadrupole resonance frequencies and asymmetry parameters of some [$SnCl_4$·2L] and [$SbCl_5$·L] complexes

| *Complex* | *ν $^{35}$Cl (MHz)* | *η (%)* | *Assignment* |
|---|---|---|---|
| [$SnCl_4$·$(NC_5H_5)_2$] | 17.644 | 15.4 | – |
| | 17.760 | 15.4 | – |
| [$SnCl_4$·$(OPCl_3)_2$] | 19.030 | 11.7 | Equatorial |
| | 19.794 | 11.1 | Equatorial |
| | 21.132 | 2.3 | Axial |
| [$SbCl_5$·$OPCl_3$] | 24.399 | 11.3 | *Cis* |
| | 25.821 | 2.5 | *Trans* |
| | 26.119 | 4.7 | *Cis* |
| | 27.314 | 4.9 | *Cis* |
| [$SbCl_5$·$NCCCl_3$] | 26.008 | 6.2 | *Cis* |
| | 26.313 | 10.3 | *Cis* |
| | 26.409 | 2.2 | *Trans* |
| | 27.297 | 11.2 | *Cis* |

for the *cis* configuration of [$SnCl_4$·$(OPCl_3)_2$] and the *trans* configuration of [$SnCl_4$·$(NC_5H_5)_2$], which is confirmed by X-ray data.

However, it is not always possible to assign equatorial and axial chlorine atoms solely on the basis of the splitting of the $^{35}$Cl NQR frequencies. For *cis* isomers, the ratio of the NQR frequencies of equatorial and axial chlorine atoms is fixed by several factors that determine the optimum crystal structure, among which is the influence of donor molecules, L is not a major contribution. In practically all structural investigations of complexes that exhibit this type of equatorial Sn–Cl bond it is usually noted that axial chlorines have lower relative NQR frequencies compared with equatorial atoms (**Table 4**). However, interpretation in structure terms is difficult because of the large values of the asymmetry parameters of chlorine atoms, which (**Table 4**) differ considerably for axial and equatorial atoms. On the other hand, donor ligands, L, influence the change in electron density of equatorial and axial Cl atoms and cause a relative lowering of frequencies of axial atoms in comparison with those of equatorial chlorines. Such a relation of frequencies in *cis* isomers is explained from the point of view of the 'mutual ligand influence' concept in non-transition element complexes. In *cis* complexes $SnCl_4L_2$ the interaction of –Sn–L bonds will be stronger with Sn–Cl bonds which are in a *trans* position, than with *cis*-Sn–Cl bonds. For these complexes, the mutual ligand influence establishes the greater *trans* effect, and leads to a redistribution of the electronic density on the chlorine atoms. In $SbCl_5$·L complexes (**Table 4**) to assign axial and equatorial chlorine atoms signals, even allowing for a knowledge of frequencies, splittings and asymmetry parameters, it is necessary in a number of cases to use the temperature dependences of NQR frequencies. In this case the NQR frequency is described by a square-law dependence: $\nu(T) = A + BT + CT^2$. A positive sign of the $C$ coefficient indicates that a NQR frequency arises from an axial chlorine atom. Thus it appears that in some complexes the frequency of an axial chlorine atom lies above those of equatorial atoms. Apparently, in these complexes the spatial influence of the donor molecules on the NQR frequencies of the equatorial chlorine atoms is dominant.

The width of the NQR signal also provides structural information. In molecular crystals of high order and purity the line width is not much different from the value determined from the sum of the spin–lattice relaxation and spin–spin relaxation times. In the majority of inorganic compounds, the lines are, however, inhomogeneously broadened by lattice imperfections such as defects, vacancies, admixtures and dislocations, so that their widths are mainly determined by the crystal inhomogeneity. A systematic study of spectroscopic shifts and broadening produced by a continuous change of the relevant sources of that broadening is an effective approach to the investigation of problems concerned with the distribution of mixtures over a matrix, the nature of their interaction with the matrix, the mechanisms of disorder and the local order in vitreous compounds.

## Studies of bonding using NQR spectra

### The Townes–Dailey approximation

The most widely used approach to provide a meaningful account of bonding trends within a series of related compounds is that formulated by Townes and Dailey for the interpretation of nuclear quadrupole interactions (NQI). The electric field gradient at a quadrupole nucleus ($q_{zz}$) arises mainly from electrons of the same atom. To a first approximation, it is possible to consider that the internal electrons will form a closed environment with spherical symmetry and, consequently, do not contribute to the EFG. Actually, polarization of the internal electrons is taken into account through the Sternheimer antishielding factor ($\gamma_\infty$). However, if the comparison of NQI is only for the purpose of chemical interpretation and is not accompanied by discussion of their absolute meanings, the polarization of the internal electrons can be neglected. Among valent electrons, those that are on s orbitals with spherical symmetry do not

contribute to the EFG and the main contribution is caused by p electrons; the contribution of d and f electrons is much less significant because of their greater distance from the nucleus and their smaller participation in hybridization.

The quantitative consideration of the contributions to the EFG results in expressions of the following type, applied here to nuclei of chlorine in chloroorganic compounds:

$$q_{zz}{}^{\mathrm{Cl}} = [(1 - s^2 + d^2 - I_{\mathrm{B}} - \pi) + I_{\mathrm{B}}(s^2 + d^2)]q_{\mathrm{at}}{}^{\mathrm{Cl}} \quad [1]$$

where $s^2$ and $d^2$ are contributions from s and d orbitals to the hybridization of the chlorine atom bonding orbital, $I_B$ is the ionic character of the bond (the chlorine atom carries a partial negative charge) and $q_{at}{}^{Cl}$ is the gradient of p electron density on the chlorine atom. The valent orbitals can be represented by somewhat modified expressions in that some treatments include the three nuclear p orbital populations, and the axes $x$, $y$ and $z$ are usually defined so that the $z$-axis coincides with the bond direction of the considered halogen or with an axis of symmetry in the molecule. These approaches can be more convenient for discussing bonding and the contribution of lone pair electron orbitals.

On the basis of the above interpretation NQI can be unequivocally represented in terms of the population of orbitals. Actually, NQR spectroscopy allows the determination, at best, of two parameters ($e^2Qqh^{-1}$ and $\eta$) and in many cases ($I = \frac{3}{2}$) only $e^2Qqh^{-1}$ can be obtained. However, the above-mentioned Equation [1] contains four parameters ($s$, $d$, $I_B$, $\pi$) which cannot be determined from one or two experimental parameters. It is therefore necessary to include approximations, which neglect the d orbital participation in hybridization, and also, in some cases, p bonding and to consider that the s orbital hybridization is a small contribution and remains constant in a series of compounds. Thus changes of $e^2Qqh^{-1}$ are directly related to bond ionic character or p electron charge transfer in the case of a hydrogen bond.

In the case of nitrogen, whose nucleus has a spin $I = 1$, the situation is more favourable, as the experiment allows the determination of both the nuclear quadrupole coupling constant and the asymmetry parameter, and it is possible to make conclusions about sp hybridization, if the molecular geometry is known. However, the meaning of nuclear NQI for $^{14}N$ p electrons is not known with as much accuracy as for chlorine, bromine or iodine, but the estimation of a 9–10 MHz contribution can be considered reliable, as it is based on the analysis of a large number of experimental data. An even more difficult situation is the case of the antimony atom, for which very little reliable NQI data exist.

## Donor–acceptor interactions

In addition to interpreting experimental NQI values in terms of orbital populations, a role is also played by structural, dynamic and simple contributions, which change the NQI such that their experimental meanings differ from those expected for a hypothetical molecule or complex in an isolated condition at rest. In the case of molecular complexes there is an additional contribution which results in NQI changes caused by a change in hybridization owing to a change in the donor and the acceptor molecule geometry. A typical example is given by $MX_3$ complexes, where M is a group IIIA or VB element and X is a halogen or methyl group. With complex formation, the X–M–X angle and appropriate hybridization change, and these result in changes in NQI interpretation even in the absence of any charge transfer. This complicates unequivocal interpretation of experimental NQR data.

It is necessary to answer the question: are the shifts of the NQR frequencies caused by a transition from a pure, non-complexed mixture of initial substances to a complex, or because of a charge transfer or for other reasons? This rather important question depends on several factors. With a small charge transfer the NQR frequency shifts are defined mainly by crystal or solid-state effects, which are caused by distinct effects in the crystal environment of molecules as the result of the transition from individual components to the complex (crystal electrical field, intermolecular interactions, thermal movement); the complexation shifts can reach several hundred kHz (in the specific case of a resonance on a chlorine nucleus a shift of the order 200 kHz is typical). When the observable shifts have larger values (one to several MHz) they cannot be considered as caused by crystal effects and it is then possible with confidence to attribute them to electronic effects arising from a charge transfer. However, it is necessary to take into account other contributions, such as hybridization changes. The hybridization change accompanying deformation of a flat $AlMe_3$ molecule to a pyramidal form formally results in an NQI change on the aluminium atom, even if there is no electron population change or change in ionic character of the bonding orbitals; thus, the shift of the NQR frequency in this case can be determined as having both a charge and a hybridization contribution.

A more difficult situation exists when, in the free compounds, there are strong intermolecular interactions: the perturbation of these interactions by complex formation can result in an increased NQI in a complex which contradicts the usual, simple prediction of an NQI reduction upon transition to a complex. Such situations are met in complexes of mercury halogenides such as $HgBr_2$·dioxan and $HgI_2$·dioxan. More complete interpretation of the experimental results can be achieved if in a complex there is present a number of quadrupole nuclei; this allows a comparison of shifts for each of them. In addition, there is often useful structural data available from X-ray diffraction. Finally for an estimation of a relative role of the various contributions to observable NQR frequency shifts, one can resort to theoretical calculations.

### Intermolecular interactions

Another example arises in the situation where a quadrupole halogen atom makes a symmetrical bridge between two metal atoms, which is often the case in polymeric metal halides of composition $MX_n$ ($n$ = 3–5). Dimers of metal halides, e.g. $AlCl_3$, $GaBr_3$, $TaCl_5$, $SbCl_5$ or $[Al_2Br_7]^-$, have thus been attractive for NQR investigators, because of the differences between the resonance frequencies of bridging ($X_b$) and terminal ($X_t$) halogen atoms. The dimers have structures with either two bridging M–$X_b$–M bonds of about equal length, as in $GaBr_3$, or two bonds of significantly different length, as in complexes of oxygen donor ligands with mercuric halides, or one bridging bond as, e.g. in the aluminium heptabromide anion $[Al_2Br_7]^-$. **Figure 1** presents the structure of one type of symmetric bridging dimeric halide.

The spatial structure of the $MX_3$-type dimer is shown in **Figure 1A** whereas **Figure 1B** presents the same structure projected onto the plane of the bridging bonds. The halogen atoms involved in the bridges are, together with metal atoms (M), in one plane which is perpendicular to the plane of the other terminal halogen atoms ($X_t$) and the metal atoms.

Analysis of the Zeeman splitting of the NQR spectra of halides of non-transition metals from group IIIA of the periodic table permits a determination of the directions of the main axes of the EFG tensor on the bridging halogen atoms. These directions are marked in **Figure 1B**. The axis of the greatest gradient of the electric field on the bridging halogen atom is perpendicular to the plane that contains the metal atoms and the bridging halogen atoms. The same axis, but on the terminal halogen atoms, lies along the metal–halogen bond. The orientation of the main

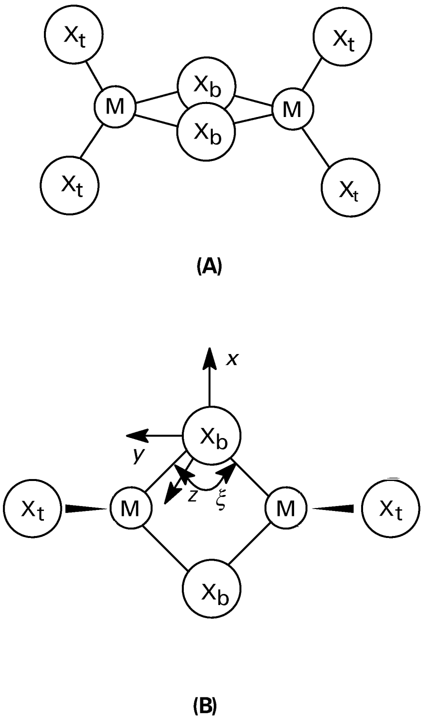



**Figure 1** (A) Spatial structure of $MeX_3$ type halides with two bridging bonds. (B) The same structure projected onto the plane of bridging bonds. The directions of the main axes of the EFG tensor on the bridging halogen atoms are marked, in the case when $\xi$<109°28′.

axes of the EFG tensor in the case of other types of bridging dimeric halides is similar.

We have indicated that the NQR frequencies of the bridging halogen atoms in non-transition metal compounds are lower than those of the terminal atoms and that this relationship is reversed in the case of transition metal compounds. **Table 5** includes the NQR frequencies of the bridging and terminal halogen atoms, the values of the EFG asymmetry parameter for the bridging and terminal halogen atoms, as well as the angles at the bridging atoms (see **Figure 1B**) for some non-transition and transition metal dimers.

X-ray structural and electron diffraction data indicate that, independent of the nature of the central metal atom, the length of the bridge is always greater than the distance between the terminal atoms. In non-transition metal compounds the effective negative charge on the bridging halogen atoms is smaller than that on the terminal atoms. The opposite situation is found in transition metal halides. In these compounds the effective negative charge on the bridging atoms is greater than that on the terminal ones. The explanation for this lies in the nature of the metal valence shells (p–d transfer). In the

**Table 5** NQR spectral parameters of some non-transition and transition metal dimer halogenides

| *Compound* | *ν(X<sub>b</sub>) (MHz)* | *η<sub>b</sub> (%)* | *ν(X<sub>t</sub>) (MHz)* | *η<sub>t</sub> (%)* | *ξ(°)* |
|---|---|---|---|---|---|
| $GaCl_3$ | 14.667 | 47.3 | 19.084 | 8.9 | 86.0 |
| | | | 20.225 | 3.4 | |
| $GaI_3$ | 133.687 | 23.7 | 173.650 | 0.9 | 94.5 |
| | | | 174.589 | 2.8 | |
| $AlBr_3$ | 97.945 | – | 113.790 | – | 82.0 |
| | | | 115.450 | | |
| $AlI_3$ | 111.017 | 18.1 | 129.327 | 0 | 93.5 |
| | | | 129.763 | | |
| $InI_3$ | 122.728 | 29.7 | 173.177 | 1.1 | 94.5 |
| | | | 173.633 | | |
| $NbCl_5$ | 13.290 | 58.8 | 7.330 | – | 101.3 |
| $NbBr_5$ | 105.850 | 58.8 | 59.500 | – | 101.3 |
| $WBr_5$ | 114.580 | 44.9 | 81.660 | – | 98.6 |

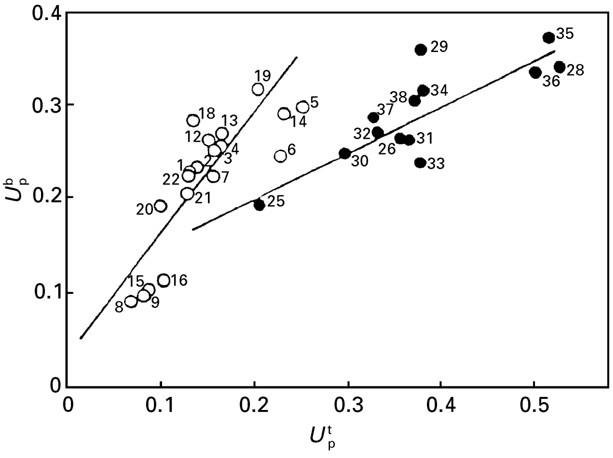



**Figure 2** $U_p^b$ versus $U_p^t$ for halogen atoms in halides of transition (o) and non-transition (•) metal elements.

framework of the Townes and Dailey theory this means an effective lowering of the resonance frequency owing to the decrease in the $p_x$ and $p_y$ orbital population of the terminal halogens.

**Figure 2** presents a correlation found by the linear regression method between the values of unpaired p-electron densities for the bridging ($U_p^b$) and terminal ($U_p^t$) halogen atoms for a large number of compounds. The existence of these correlations is connected with the influence of the p-electron distribution on NQR frequencies, and the bridging and terminal halogen atoms are connected through the central atom, which acts as a buffer. One can see from **Figure 2** that the experimental points can be divided into two groups, corresponding to non-transition and transition metal compounds. This is in good agreement with the difference in NQR frequencies for the bridging and terminal halogen atoms for all the compounds studied. As a consequence, the unpaired electron density differs for the compounds of both groups, although $U_p^b$ increases with increasing $U_p^t$ for all compounds studied. The observed correlations are approximate, which points to the influence of other factors, such as crystallographic uncertainty and the position of the molecules in the unit cell of the crystal.

### Charge distribution in biological molecules

NQR spectroscopy can be also used to provide detailed information on the structure and conformation of biologically active systems. It offers a unique possibility of determining the quadrupole coupling constants and, as a consequence, effective charges, and in this way allows a determination of the electronic structure of the molecule. NQR spectroscopy appears to offer a powerful tool for the investigation of various chemical effects in the solid phase of many nitrogen-containing compounds. Analysis of the quadrupole coupling constants of nitrogen atoms allows an estimation of the electron density distribution on the nitrogen nuclei and enables the analysis of charge distribution in chemical bonds involving nitrogen.

For example, the effect of substitution at the 1-H position of the 2-nitro-5-methylimidazole [2] ring can be analysed (**Table 6**), Hitherto, different imidazole derivatives have been studied by $^{14}N$ NQR, by the continuous wave method, by $^{13}C$ NMR and $^{35}Cl$ NQR (pulse methods). However, the problem of the effect of a substituent at the 1-H position on the

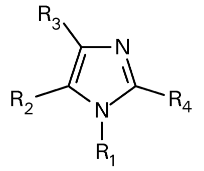
**Table 6** Chemical names and substituents of the compounds studied

| *No.* | *Compound* | *R¹* | *R²* | *R³* | *R⁴* |
|---|---|---|---|---|---|
| [1] | Imidazole | H | H | H | H |
| [2] | 2-Nitro–5-methylimidazole | H | $NO_2$ | H | $CH_3$ |
| [3] | 1-(2-Hydroxyethyl)-2-nitro-5methylimidazole (metronidazole) | $CH_2CH_2OH$ | $NO_2$ | H | $CH_3$ |
| [4] | 1-(2-Carboxymethyl)ethyl-2-nitro-5-[2-(P-ethoxy-phenyl)ethenyl] imidazole | $CH_2CH_2OCOCH_3$ | $NO_2$ | H | $CH{=}CHPhCH_3O$ |

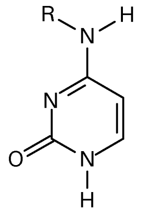
**Table 7** Quadrupole coupling constants and asymmetry parameters for imidazole derivatives

| | | *Nitrogen Nucleus* | | | | | |
|---|---|---|---|---|---|---|---|
| | | −N= | | −NR | | $-NO_2$ | |
| No. | *T* (K) | $e^2Qq_{zz}h^{-1}$ *(MHz)* | $\eta$ | $e^2Qq_{zz}h^{-1}$ *(MHz)* | $\eta$ | $e^2Qq_{zz}h^{-1}$ *(MHz)* | $\eta$ |
| [1] | 291 | 3.222 | 0.119 | 1.391 | 0.930 | – | – |
| | 77 | 3.253 | 0.135 | 1.418 | 0.997 | – | – |
| [2] | 296 | 3.243 | 0.250 | 1.569 | 0.821 | 1.225 | 0.356 |
| | 193 | 3.249 | 0.249 | 1.546 | 0.878 | 1.244 | 0.360 |
| [3] | 296 | 3.299 | 0.150 | 2.467 | 0.317 | 0.936 | 0.381 |
| | 193 | 3.320 | 0.156 | 2.479 | 0.320 | 0.950 | 0.381 |
| [4] | 296 | 3.755 | 0.038 | 2.566 | 0.238 | 0.921 | 0.239 |
| | 193 | 3.779 | 0.039 | 2.565 | 0.238 | 0.931 | 0.230 |

electron distribution in 2-nitro-5-methylimidazole [2] derivatives has not been considered. The results of NMR-NQR double resonance studies on a series of imidazoles (**Table 6**) are collected in **Table 7**.




[1-4]




[5]

As follows from the data in **Table 7**, the introduction of $NO_2$ and $CH_3$ at positions 2 and 5 of the imidazole ring, respectively, leads to an increase in nuclear quadrupole coupling constants on both nitrogens of the ring and a decrease in the value of this parameter on the nitrogen from the $NO_2$ group. Results of a bond population analysis carried out according to the Townes–Dailey method for imidazole and its derivatives are displayed in **Table 8**.

According to the assumed notation $n_{NC}$ and $n_{NR}$ stand for the population of N–C and N–R bonds, $n_a$ stand for the lone pair population of an N atom, while $\pi$ is the $\pi$-electron density. A comparison of the data collected in **Table 8** shows that in the case of a 3-substituted nitrogen atom, the lone pair electron density and NR bond population increase with increasing length of a substituent at 1-H position in the imidazole ring, both at 193 and 269 K. The substitution of an $NO_2$ group at the 2-position and a methyl group at the 5-position of the imidazole ring leads to an insignificant redistribution of electron density (of the order of 2%) relative to that in pure imidazole. Only the introduction of a substituent at the 1-H position of the ring causes considerable changes. Similarly for a 2-substituted nitrogen, the bond population changes in a characteristic way – the $\pi$-electron density and population of the N–C bond decrease with elongation of the substituent at 1-H position of the imidazole ring (**Table 8**). Interestingly, a qualitatively similar but much more effective phenomenon is the electron redistribution on the nitrogen atoms of the $NO_2$ groups. Therefore it can be concluded that, in the case of 2-nitro-5-methyl derivatives of imidazole, with increasing length of the chain of the aliphatic substituent at the 1-H position of the imidazole ring, the electron density of the $\pi$-orbital and $\sigma$-bond of the 2-substituted atom, as well as the $\pi$-orbital and $\sigma$-bond of the $NO_2$ group towards the 3-substituted nitrogen, undergoes redistribution. Of essential importance is the character of substituents, i.e. $CH_2CH_2OH$ and $CH_2CH_2OCOCH_3$ groups are electron density acceptors, while for 1H-imidazole and 1-acetylimidazole, the opposite tendency was observed, i.e. the $\pi$-electron density and the population of the $\sigma$-bond increased. The results also indicated that the effect of temperature on the bond populations in the imidazole derivatives studied is negligible.

The 4-N-derivatives of cytosine [5], R-substituent R = H, $CH_2Ph$, $CH_2CH_2Sh$, $CH_2CH_2Ph$ and naphtha-

**Table 8** Population of the −NH–, −N=,$-NO_2$ bonds for imidazole and its derivatives

| | *−NH−* | | | *−N=* | | *$-NO_2$* | | |
|---|---|---|---|---|---|---|---|---|
| *No.* | $n_a$ | $n_{NC}$ | $n_{NR}$ | *p* | $n_{NC}$ | *p* | $n_{NC}$ | *T(K)* |
| [1] | 1.330 | 1.120 | 1.330 | 1.640 | 1.260 | – | – | 77 |
| | 1.340 | 1.140 | 1.330 | 1.640 | 1.270 | – | – | 293 |
| [2] | 1.350 | 1.129 | 1.330 | 1.657 | 1.556 | 1.072 | 1.182 | 193 |
| | 1.361 | 1.139 | 1.330 | 1.652 | 1.556 | 1.070 | 1.178 | 296 |
| [3] | 1.672 | 1.250 | 1.368 | 1.449 | 1.390 | 1.042 | 1.128 | 193 |
| | 1.669 | 1.250 | 1.366 | 1.452 | 1.394 | 1.041 | 1.126 | 296 |
| [4] | 1.648 | 1.250 | 1.340 | 1.430 | 1.384 | 1.045 | 1.117 | 193 |
| | 1.648 | 1.250 | 1.340 | 1.430 | 1.384 | 1.044 | 1.116 | 296 |

lene, have aroused significant interest mainly because of their biological significance. 4-*N*-methyl and 4-*N*-acetylcytosine have been found in nuclei acids as rare bases. The results of the analysis of bond populations for cytosine and its derivatives performed according to the Townes–Dailey theory are described below. On the NH nitrogen, the changes in electron density distribution induced by substitution are insignificant. The symmetry of charge distribution at the NH nitrogen in 4-*N*-thioethylcytosine is the lowest while in 4-*N*-naphthalenecytosine it is the highest. On the other hand, the greatest value of the $z$-component of the EFG tensor was detected in the vicinity of the NH in pure cytosine while the smallest was in naphthalenecytosine, which implies that the electrons are drawn away from NH by the system of bonds depending on the substituent. Thus, it can be concluded that the introduction of a substituent at the amine group of cytosine at 4-N does not cause any changes when the substituent contains a chain such as $CH_2CH_2$ which separates the aromatic system from the cytosine ring. However, when the aromatic system is separated only by one $CH_2$ group, or is not separated at all, there occurs a strong inductive effect which is responsible for drawing $\pi$-electron density from the NH nitrogen. For a 2-substituted nitrogen the symmetry of charge density at –N= is 50% lower. This is if one compares that of 4-*N*-phenylmethylcytosine with that of 4-*N*-thioethylcytosine, and about 20% lower than that of 4-*N*-naphthalenecytosine. On the other hand, the quadrupole coupling constant and thus the $z$-component of the EFG tensor at the –N= nitrogen, is much higher in the derivatives whose substituents are not separated by the $CH_2CH_2$ chain. The difference between the population of the NC bond for pure and substituted cytosine is small –0.0001 for 4-*N*-thioethylcytosine but as great as 0.132 for 4-*N*-phenylmethylcytosine, which is still much less than the difference between the populations of $\sigma$ and $\pi$ bonds. Such a situation indicates that the changes in $\sigma$-bond population are dominant. In 4-*N*-naphthalenecytosine the change in $\pi$-electron density is dominant. This implies that –N= in phenylmethylcytosine and naphthalenecytosine plays the role of a buffer. In 4-*N*-phenylmethylcytosine and 4-*N*-naphthalenecytosine the electron density of a free electron pair at the nitrogen NHR is significantly delocalized. The electron density on the NH bond in 4-*N*-naphthalenecytosine, 4-*N*-phenylmethylcytosine and 4-*N*-thioethylocytosine relative to that in pure cytosine changes by 0.053, 0.038 and 0.005, respectively, while for 4-*N*-phenylethylcytosine there is no change. Thus, the amine group which acts as $\pi$-electron acceptor in the majority of molecular systems, becomes the electron donor in phenylcytosine and naphthalenecytosine. The aromatic rings which usually compensate changes in electron density in cytosine act as electron acceptors. When the aromatic substituents are separated by the $CH_2CH_2$ chain, the density redistribution is reduced, which is in agreement with the tendency observed for chlorobenzenes. Investigations of cytosine and its derivatives have shown that the cytosine derivatives with an aliphatic substituent do not show an anticancer activity, but those with an aromatic substituent do. However, the mechanism of this activity has not been explained. The results suggest that in the search for anticancer drugs from the group of 4-*N*-cytosine derivatives, the choice should be those in which the aromatic substituents are not separated by a $CH_2CH_2$ chain, whose presence induces a significant redistribution of $\pi$-electron density. The $^{14}N$ NQR frequencies recorded for cytosine derivatives were practically the same at ambient and liquid nitrogen temperatures, which means that no essential redistribution of electron density occurs in this temperature range.

The $^{14}N$ nucleus, because of its widespread occurrence in all types of systems (especially biologically active systems), is of particular interest in studying electron density distribution, molecular reorientations and intermolecular time-dependent interactions. It seems that such studies will acquire more and more importance in the future and will occur more frequently, especially with the availability of double resonance spectrometers and new data processing techniques such as the maximum entropy method. The examples discussed do not, of course, exhaust the potential of NQR as a tool for structure and chemical bonding. These are only simple illustrations of the applied aspects of NQR spectroscopy.

### Application of NQR to the detection of explosives, contraband etc.

The increasing use of plastic explosive and drugs has found experts and research facilities scrambling for new detection methods. One of these ‘new’ methods is NQR. In the past five years a number of researchers around the world have independently begun to reconsider NQR as a possible solution for the detection of plastic explosives and started developing it specifically for bomb and narcotics detection. The noninvasive nature of NQR (closely connected with the absence of magnets) gives it some advantages over other methods.

Pulsed-RF NQR produces single, or nearly single, peak signals at specific frequencies that depend on the specific bond environment surrounding an element in a given compound, usually a crystalline solid. Because the resonance frequency is almost unique to each compound, NQR exhibits great specificity for

analytes such as explosives and narcotics. The most useful elements to monitor by NQR are $^{14}N$, $^{35}Cl$ and $^{37}Cl$. Most high explosives contain 30–40% N and a large number of drugs such as cocaine and heroin are prepared as chloride salts. Most pure explosive such as RDX, HMX, TNT, C-4 are crystalline or semicrystalline compounds embedded in a polymer matrix, rather than pure polymeric compounds, so that they are immobilized and relatively ordered and as such give good NQR. Of course liquids and polymers are too disordered to give an NQR signal, although some monomers have shown detectable resonances.

NQR can be also used to differentiate between explosives, narcotics and benign nitrogen-containing compounds such as polyurethane foam or nylon. False alarms from these compounds are not the problem for NQR whereas, they might be for neutron activation or other generic nitrogen-detecting methods. It is well known that commercial and military explosives are physical mixtures of pure explosive compounds with some additive plasticizer or binder. Because NQR is so compound-specific, physical additives do not interfere with the signal for a target compound so NQR can be used to identify explosives that are not in a pure form. Moreover, it does not matter what form the explosive or drug is in – whether it be tin sheets or small pellets. NQR may eventually be used to detect bombs or narcotics with spatial resolution, in the same way as X-ray metal detectors.

NQR is inherently less flexible than NMR but when it works it is extremely attractive because of its specificity. NQR can work with slurries, aggregates and possibly even emulsions, as long as the molecular dynamics are slower than the NQR method time scale (the MHz range). The ongoing use of NQR as analytical tool for measuring solid-state phase transitions and order–disorder in materials may also be of interest.

## Summary

NQR is not as extensively useful at present as NMR spectroscopy. The best results on light nuclei, such for example as those on $^{27}Al$ nuclear coupling constants in mineral samples, have been made using NMR. Here, the changes in the NMR spectra were considered as a function of the orientation of a single crystal in external magnetic field. NQR could, however, directly measure the same nuclear coupling constant data using neither single crystals nor an external magnetic field. NQR measurements possess high spectral resolution, precision, specificity and speed of measurements. The reason for the relatively limited practical application of NQR seems to lie in the lack of sufficiently sophisticated equipment.

NQR applications can be divided into four groups:

1. Studies of the electron density distribution in a molecule – changes in orbital populations under substitution, and complexation.
2. Studies of molecular motions – reorientations, rotations, hindered rotations.
3. Studies of phase transitions.
4. Studies of impurities and mixed crystals.

## List of symbols

$e^2Qq_{zz}h^{-1}$ = quadrupole coupling constant; $e^2Qq_ph^{-1}$ = quadrupole coupling constant; $I$ = nuclear spin quantum number; $I_B$ = ionic character of a bond; $q_p$ = the electric field gradient produced by one unbalanced p electron; $q_{zz}$ = electric field gradient (EFG) at a quadrupole nucleus; $U_P^b$, $U_P^t$ = unpaired p-electron density of a bridging and terminal, halogen, respectively; $\eta$ = asymmetry parameter; $\gamma$ = Sternheimer antishielding factor.

*See also:* **Mossbauer Spectrometers; Mossbauer Spectroscopy, Applications; Mossbauer Spectroscopy, Theory; NQR, Theory; Nuclear Quadrupole Resonance, Instrumentation.**

## Further reading

Buslaev JA, Kolditz L and Kravcenko EA (1987) *Nuclear Quadrupole Resonance in Inorganic Chemistry*. Berlin: VEB Deutsche Verlag der Wissenschaften, 237 pp.

Das TP and Hahn EI (1958) *Nuclear Quadrupole Resonance Spectroscopy. Solid State Physics*, suppl. I. New York: Academic Press, 223 pp.

Gretschischkin WS (1973) *Yadernye Kwadrupolnye Vzaimodejstviya v Tverdych Telach*. Nauka: Moskva, 264 pp.

Lucken EAC (1969) *Nuclear Qudrupole Coupling Constants*. London: Academic Press, 360 pp.

Safin IA and Osokin D (1977) *Yadernyj Kvadrupolnyj Rezonans v Soedinieniach Azota*. Moskva: Nauka, 255 pp.

Semin GK, Babushkina TA and Yakobson GG (1975) NQR Group of INEOS AN SSSR, *Nuclear Quadrupole Resonance in Chemistry*, (English edition). London: Wiley, 334 pp.

Smith JA (1974–1983) *Advances in Nuclear Quadrupole Resonance*, Vol. 1–5. London: Heiden & Sons.

Townes CH and Dailey BP (1949) Determination of electronic structure of molecules from nuclear quadrupole effects. *Journal of Chemical Physics* **17**: 782–796.

Townes CH and Dailey BP (1955) The ionic character of diatomic molecules. *Journal of Chemical Physics* **23**: 118–123.

# Nuclear Quadrupole Resonance, Instrumentation

**Taras N Rudakov**, S.E.E. Corporation Ltd., Bentley, WA, Australia

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
**Methods & Instrumentation**

Nuclear quadrupole resonance (NQR) is a modern research method for the analytical detection of chemical substances in the solid state. NQR is a type of radiofrequency (RF) spectroscopy, and is defined as the phenomenon of resonance RF absorption or emission of electromagnetic energy. It is due to the dependence of a portion of the energy of the electron–nuclear interactions on the mutual orientation of asymmetrically distributed charges of the atomic nucleus and the atomic shell electrons, as well as those charges that are outside the atomic radius. Thus, all changes in the quadrupole coupling constants and NQR frequencies are due to their electronic origin. The nuclear electric quadrupole moment $eQ$ interacts with the electric field gradient $eq$, defined by the asymmetry parameter $\eta$. Therefore the nuclear quadrupole coupling constant $e^2Qq$ and the asymmetry parameter $\eta$, which contain structural information about a molecule, may be calculated from the experimental data. The main spectral parameters in NQR experiments are the transition frequencies of the nucleus and the line width $\Delta f$. In addition, measurement of the spin–lattice relaxation time $T_1$, spin–spin relaxation time $T_2$ and line-shape parameter $T_2^*$ (inversely proportional to $\Delta f$) is also of great value. These parameters must also be taken into consideration when choosing the experimental technique and equipment.

## NQR methods

In contrast to nuclear magnetic resonance (NMR) methods, NQR can operate without a strong external DC magnetic field. This technique is known as 'pure NQR', or direct NQR detection, and has many advantages for some applications, such as identification of specific compounds and remote NQR. In turn, direct NQR detection techniques can be subdivided into two main areas: oscillator-detector methods [continuous wave (CW) and superregenerative techniques] and pulsed methods. At present the pulse method is the most widely used direct NQR method. The use of various multipulse sequences permits a considerable increase in the sensitivity, and a decrease in the duration of the experiment.

Besides direct methods of detection, indirect NQR detection methods have also been developed. As a rule, they are used at low frequencies or in cases when the concentration of quadrupolar nuclei is not high, i.e. when the sensitivity of the direct methods is not sufficient. For most experiments, when using these methods, a constant magnetic field is needed. To apply these methods, it is necessary for the sample to have two spin systems connected by dipole–dipole interactions. Indirect methods can be conventionally divided into double-resonance techniques and the cross-relaxation spectroscopy method.

A block diagram summarizing the composition of the main technologies for NQR methods is given in **Figure 1**.

## Direct NQR detection (pure NQR) techniques and equipment

Direct methods are the basis of experimental methods in NQR. They are used in a wide frequency range from hundreds, sometimes even tens of kHz to hundreds of MHz. Historically these methods started to be applied earlier than indirect methods. The first observations of NQR were undertaken using CW and superregenerative techniques, which are simple and inexpensive. These methods and techniques were developed during the 1960s to 1980s. At present it is the pulsed techniques that are more widely used, as they permit a better sensitivity and are more convenient for measuring NQR parameters of the sample over a wide frequency range. The modern pulsed techniques use the latest signal processing methods, including fast Fourier transformation (FFT), digital filtering, and others.

### Oscillator-detector methods

**Continuous wave techniques** CW NQR spectrometers are based on the use of oscillator-detectors, which are built around the circuits of a marginal oscillator or a limited oscillator (Robinson oscillator). Such an oscillator-detector includes a tank circuit with a coil, into which the studied sample is inserted. When the frequency of the oscillator-detector coincides with the NQR frequency in the sample, the

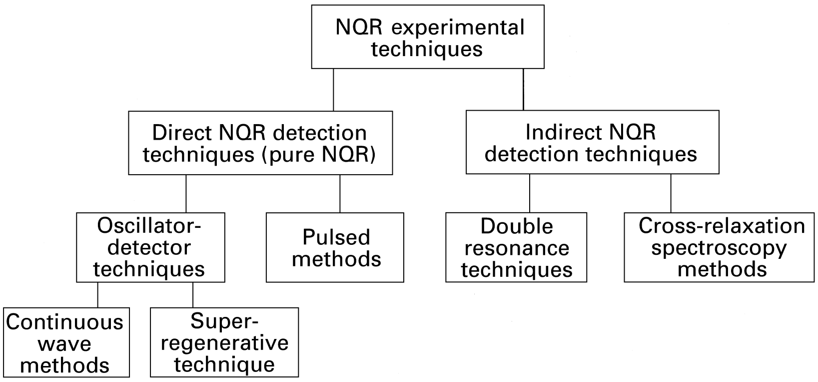



**Figure 1** Main categories of NQR experimental techniques.

sample starts to absorb energy and the amplitude of oscillations in the circuit decreases. This change in the amplitude is detected on the output of the oscillator-detector. The sensitivity of such spectrometers can be high and depends on the Q-factor of the coil, on the filling factor and on the noise temperature of the oscillator. Normally these oscillator-detectors cannot work at high levels of RF voltages and require accurate tuning of the electric circuit parameters. To avoid saturation of the sample during the detection process, they use either frequency modulation or Zeeman modulation, usually bisymmetric. Frequency modulation is easier to perform from the technical viewpoint, but it has the drawbacks of giving base-line drift and spurious signals.

In the literature one can find many descriptions of CW NQR spectrometers and oscillator-detectors for various purposes and frequency ranges. Particular interest is evoked by limited oscillators, which have a low noise level similar to that of marginal oscillators but do not need continual critical adjustment of circuit parameters. **Figure 2** shows a simplified block diagram of a CW spectrometer, containing an oscillator-detector with a Zeeman modulation system and a computer for signal processing and frequency control.

**Superregenerative technique** Superregenerative oscillators (SRO) have been widely used as oscillator-detectors in the study of NQR because of their high sensitivity at high RF levels, reliability of operation and simplicity of construction. The phenomenon of NQR was discovered by Demelt and Kruger using a SRO spectrometer. SRO is actually an oscillator-detector, with oscillation periodically quenched either with the help of an external quench oscillator (quencher mode), or with the help of internal circuits (self-quenched mode). A high sensitivity of the SRO is provided during its coherent work, when the oscillations are not quenched completely, i.e. not to the level of noise. Besides, SRO performs very well in logarithmic mode, when RF pulses are built up to a limiting amplitude. Detection of the NQR signal in SRO occurs as follows. Using RF pulses, the SRO excites an NQR signal in a sample. This signal, which occurs in the intervals between these pulses, is added up with the resilient voltage. In this case oscillations in the next pulse will start on a higher level of voltage and will reach the limiting amplitude sooner, i.e. the length of the pulse will increase. This change in the length of the RF pulses of the SRO is detected as the signal.

Usually the SRO has a number of significant disadvantages: (i) poor frequency stability; (ii) the RF spectrum contains components (sidebands) spaced at integral multiples of the quench frequency on either side of the fundamental frequency, each of them being capable of exciting the NQR signal; (iii) poor fidelity of line shape reproduction; and (iv) dependence of the centre frequency on the quench frequency. The sideband suppression is achieved by using the quench frequency modulation. The other disadvantages can be eliminated by using the quench frequency modulation. The other disadvantages can be eliminated by using the phenomenon of frequency (or injection) and phase-locking. In this case the frequency stability of the SRO will be determined by the stability of the external locking oscillator. To stabilize the shape of lines of the detected NQR signal, phase-locking is used. A simplified block diagram of such an NQR spectrometer is given in **Figure 3**. The frequency locking of the SRO is done by introducing a small voltage from the external adjustable oscillator, which switches SRO from the regime of noncoherent work to coherent. An injection signal can be introduced with the help of a small capacitor or by

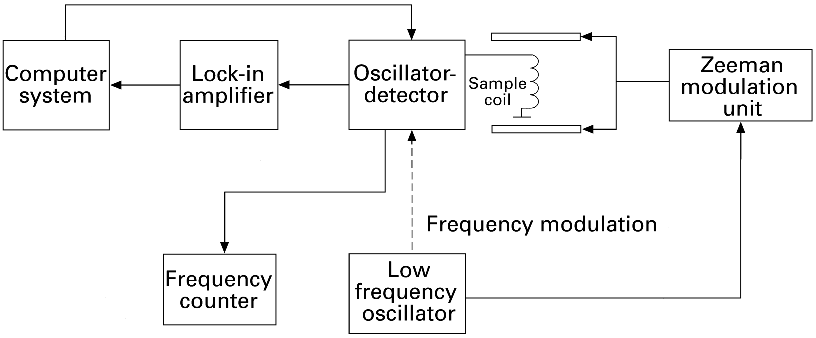



**Figure 2** Simplified block diagram of a CW NQR spectrometer.

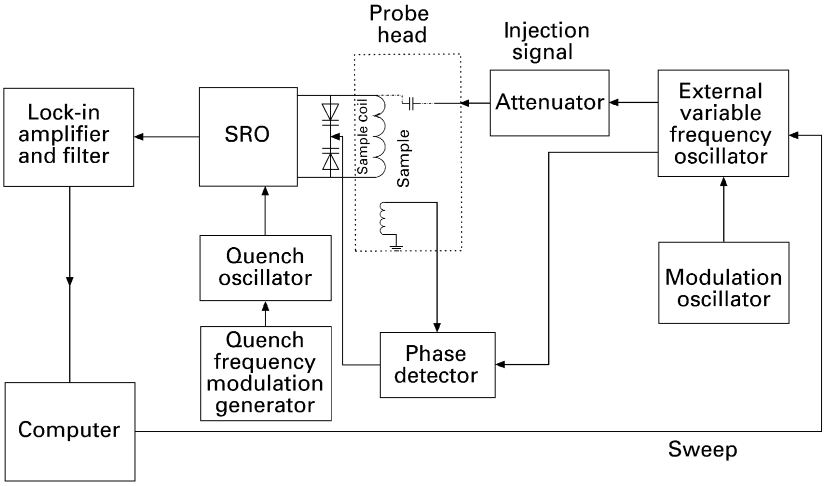



**Figure 3** Simplified block diagram of the frequency and phase-locked NQR spectrometer based on a SRO-type oscillator-detector.

simply using the inductance of the connection. As the SRO frequency is determined by the frequency of the external oscillator, the frequency modulation is applied to that oscillator. In the mode of frequency locking the change of the capacity of the SRO circuit leads to a change in the phase of the generator frequency, which causes a change in the line shape of the detected signal. Therefore to stabilize the shape of the NQR signal line on the output of the SRO, so-called phase locking is carried out, using an additional phase detector. Such a phase detector permits the control of a required phase shift between the SRO frequency and the external oscillator frequency by automatic tuning of the capacity of the SRO circuit.

### Pulsed methods

The essence of the pulse method approach consists of irradiating the spin-system by RF pulses with frequencies equal or close to the NQR transition frequency and followed by detection of signals induced by this spin-system. Direct pulsed NQR methods are of great interest for industrial, medical, chemical and biochemical investigations. They also seem to be an effective technique for detecting the presence of prohibited goods (narcotics), landmines and plastic bombs. This technique is very convenient and commonly used for the measurement of line widths and relaxation times. CW and SRO spectroscopy are quite inefficient with regard to measuring time and in particular for samples with long $T_1$ and short $T_2$. Nowadays, CW and SRO spectrometers have almost completely been replaced by pulsed FFT NQR instruments.

Multipulse techniques, widely used in magnetic resonance, are also very common in NQR spectroscopy. They are effectively used for increasing sensitivity, reducing the duration of the experiment, and for measuring relaxation times in the sample. In NQR such well-known sequences as spin-echo, Carr–Purcell (CP), Meiboom–Gill-modified CP, spin-locking sequence and phase-alternated pulse sequence are widely used. There is a large practical

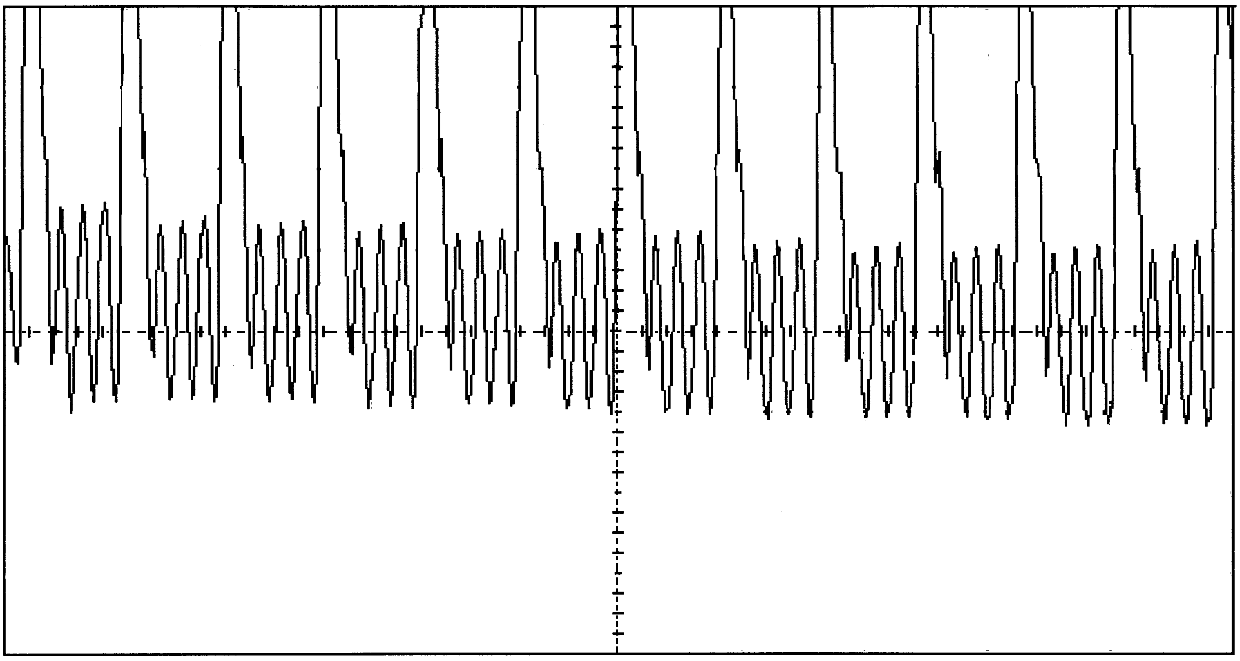


Figure 4 The response from nitrogen-14 in a sample of HMX to the SSFP pulse sequence with offset −5 kHz. The resonance frequency is 5.301 MHz at 295 K.

interest at present in pulse sequences of the steady-state free precession type (SSFP). The simplest version of this sequence is well known in NQR as 'strong off-resonance comb'. When certain requirements are complied with, these sequences give a stationary signal that does not die down while the sequence is in action. This method is very convenient for achieving fast coherent accumulation (averaging) of the signal.

At present, various modifications of the pulsed SSFP-type sequences have been developed or are being developed to increase sensitivity and eliminate some undesirable effects, such as intensity and phase anomalies. Typical results for $^{14}N$ responses to SSFP-type sequences in HMX ($C_4H_8O_8N_8$) are shown in **Figure 4**. To carry out multipulse experiments, the pulsed NQR spectrometer must have a programmable pulse generator with a variety of possibilities for changing pulse length and spacing, and a phase shifter, which permits a large variation of the phase of the RF pulse.

Now, pulse Fourier spectroscopy is the preferred experimental technique in NQR. It was made possible by the introduction of inexpensive computers and by the development of FFT algorithms. Note that the principles of Fourier spectroscopy are well developed and widely used in other fields of spectroscopy, including magnetic resonance, electron paramagnetic resonance etc. Several monographs are available on this subject. The reader is referred to the Further reading section for details of Fourier spectroscopy, including its application in NQR. To illustrate the FFT method in pulsed spectroscopy, **Figure 5** shows a typical $^{14}N$ NQR spectrum of powder RDX ($C_3H_6O_6N_6$), obtained from a free induction decay (FID).

One of the most important characteristics of a wide-band and pulsed NQR spectrometer is its

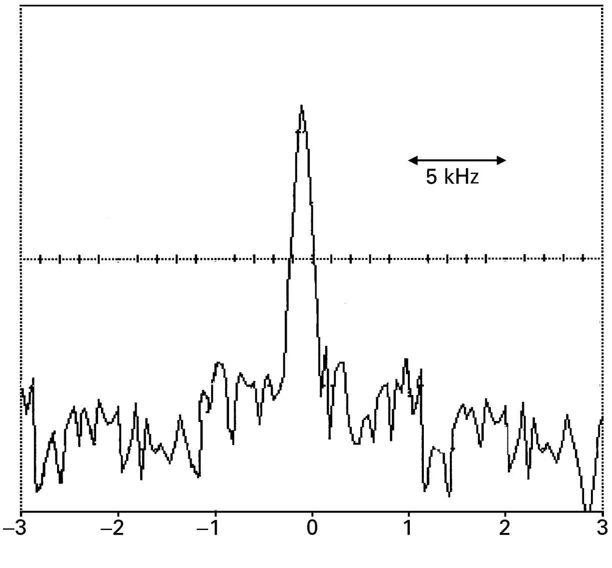



Figure 5 A typical nitrogen-14 NQR spectrum of powder RDX at 297 K, obtained from a FID. The resonance frequency is ~5.193 MHz.

sensitivity. This is because, in most cases, the intensity of NQR signals is very low. It is especially true for the low-frequency range, less than 10 MHz, but is also significant for higher frequencies. Therefore it is necessary for the receiving system of the spectrometer to have a low noise factor and to ensure matching of its input impedance with the circuit. Without this matching the noise factor can increase considerably. The signal-to-noise (S/N) ratio can also be increased by increasing the $Q$-factor of the circuit. The S/N ratio is proportional to $\sqrt{Q}$, because with the increase in quality, the amplitude of the signal voltage increases $Q$ times, while the effective noise voltage increases by $\sqrt{Q}$ only. However, with the increase of the quality of the circuit, transient processes known as 'ringing', after the irradiating signal, become longer. At low frequency the length of transient processes can reach hundreds of microseconds, which creates considerable difficulties for detecting the inductance signal and even spin-echo signals in samples with short $T_2$ and $T_2^*$ times. The problem of ringing can be solved by several methods. Most often $Q$-switching is used, i.e. during transmission the $Q$ factor is low, about 5–10, but during receiving it is high, typically > 100. In a number of practical applications of NQR, e.g. in remote NQR, or detection using large volumes, it is desirable to have a high $Q$-factor in the transmit regime as well. In this case methods of active damping of transient processes in the circuit are used, i.e. a short-term decrease of $Q$ directly after the irradiation pulse. There are many descriptions of circuits for this purpose in the literature. Good results are achieved when short antiphase RF pulses are used immediately after pulses of RF irradiation. This effect is shown in **Figure 6.**

Besides ringing, transient processes in the receiver module can also impede the observation of NQR signals in both high- and low-frequency parts of the receiver. To eliminate this, it is possible to use wide-band amplifiers in the high-frequency part, and to close the input of the receiver and the input of the synchronous detectors and switch off the reference frequency transmission during the excitation pulse. High sensitivity of an NQR spectrometer is achieved when using optimum filtering, which corresponds best to the parameters of the detected signal. For this purpose Bessel and Butterworth filters are usually used. As different substances have resonance lines of different width, it is preferable to have a receiver with a variable bandwidth. Most often receivers of NQR spectrometers are built on the basis of superheterodyne techniques. These techniques are very efficient for wide frequency band spectrometers operated at 10 MHz. At low frequencies though (0.5–10 MHz) straight receiver systems are successfully used too. They consist of an RF amplifier, synchronous detectors, low-pass filter and an output video amplifier. As a rule, the principle of quadrature detection is used in all these receivers. Digital receivers, where analogue-to-digital conversion is carried out either at the intermediate frequency, or directly on the resonance frequency, have recently become more widely used. Evidently in the future this will be the most promising direction in the development of receiver systems of NQR spectrometers.

The system of exciting the NQR signal in a sample is also an important part of a spectrometer. Besides a power amplifier, it contains a means of generating RF pulses (gate), a programmable pulse generator (sequencer) and a stable variable oscillator (synthesizer). To form various multipulse sequences it is also necessary to use phase shifters. A standard block diagram for a pulsed NQR instrument is given in **Figure 7.**

In the literature more complicated NQR spectrometers with frequency conversion in both the receiver and the transmitter (irradiation) modules have been described. It should be noted that in wide-band NQR spectrometers several receivers and power amplifiers are used, each of them intended for work in a specific frequency range. Many laboratories involved in NQR research use spectrometers that are either home-made or made to order by specialized companies. However, companies specializing in the technology for magnetic resonance also often produce small batches of high-quality NQR equipment. For example, the equipment produced by Tecmag, Inc. is extremely good for carrying out the most complicated NQR experiments. It contains devices for forming all kinds of common and rarely used pulse sequences. It also contains such vital modules as a digital signal processor, a pulse programmer, a signal averager with complex memory, a special frequency synthesizer, a digital receiver and the latest model of a PC with specialized software.

The development of electronics and new experimental techniques has allowed the considerable broadening of the areas of practical application for pulsed NQR. Important NQR applications include the detection of drugs, landmines and plastic explosives based on signals from the nuclei of nitrogen and chlorine, in which the NQR signal at low frequencies of 0.5–6 MHz is detected. The weakness of NQR signals hinders further development of this method, especially in this frequency range. Two different NQR techniques are being applied to detect these substances, detection using large ordinary coils and remote detection using special coils (antennae). The first technique is used to scan packages. It is a development of ordinary pulsed NQR techniques for

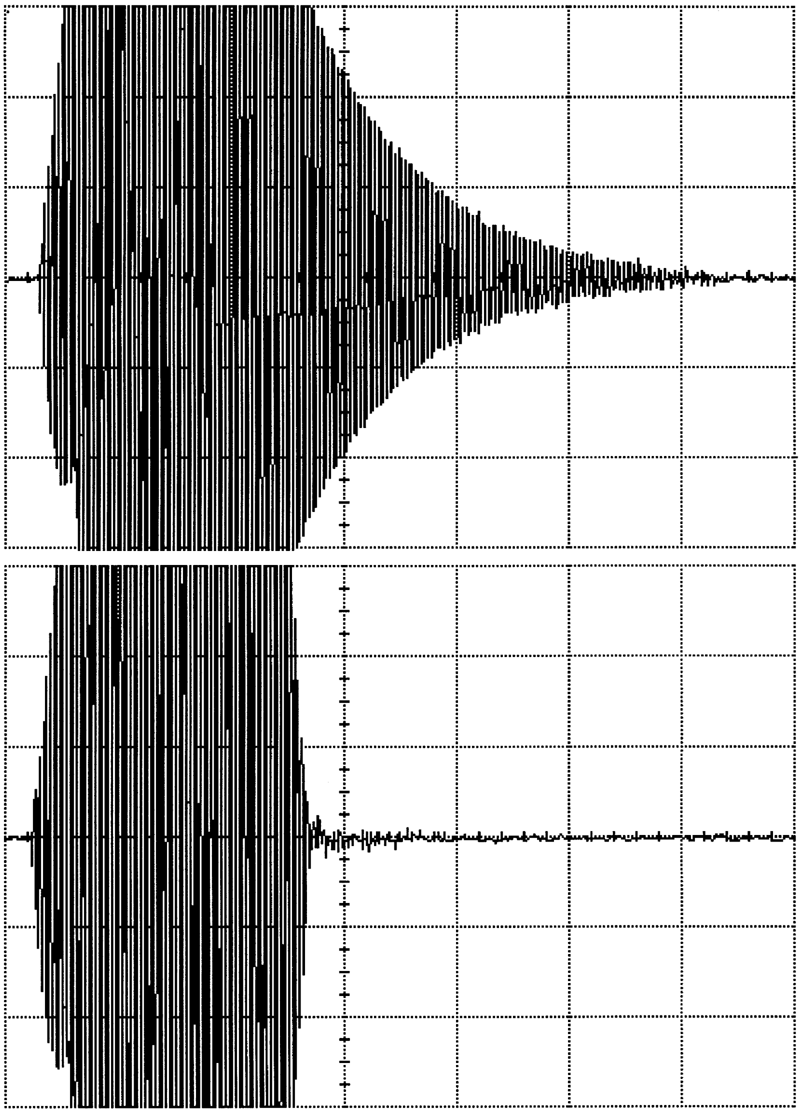



**Figure 6** An RF pulse in the NQR spectrometer coil: (A) without ring damping, (B) after ring damping.

use with large volumes and cannot be utilized to detect samples beyond a certain distance from the spectrometer coil (e.g. landmines). At the same time, the remote NQR technique is more universal. However, besides extremely complicated highly sensitive receiving equipment, it requires the use of specific signal processing methods for eliminating external interference. It is also necessary to establish the required RF field at a definite distance from the coil. In remote NQR, special flat-surface coils of various designs are used, with irradiating and receiving coils (antennae) sometimes separated.

## Indirect NQR detection technique

The indirect NQR detection technique was developed as a result of wide interest in important chemical compounds containing light nuclei. As a rule, these are organic compounds that are interesting from the point of view of biology and medicine. As NQR frequencies of light nuclei are located in the low-frequency range (less than 1 MHz), the intensity of the signals from these nuclei when using direct techniques is very low. The indirect NQR technique permits high sensitivity for detecting many light elements, which is very convenient for locating unknown resonances. This technique cannot replace completely the direct method for NQR detection, but is important as a complementary technique.

### Double resonance

The principle of double resonance (DR) consists of the detection of a weak signal from one type of

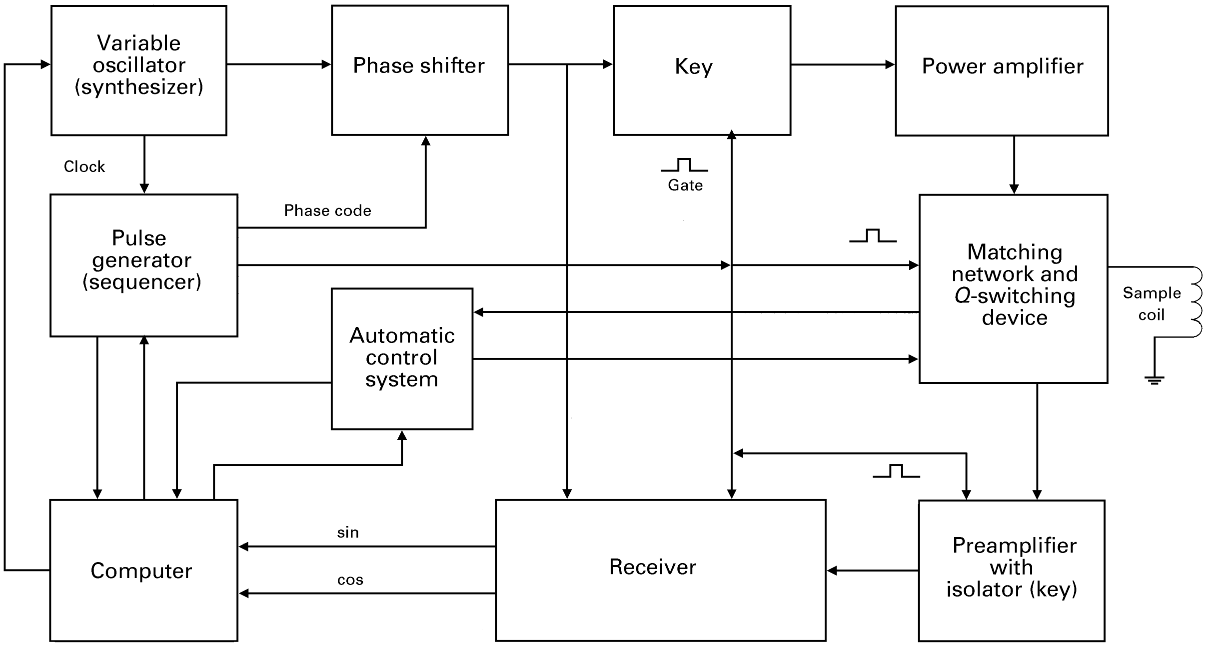



**Figure 7** Block diagram of a pulsed NQR instrument.

nucleus when directly detecting changes in the strong signal from another type of nucleus. Thus, the studied sample must contain two types of nuclei, $N_Q$ and $N_P$, with one of them (e.g. $N_P$) having a strong nuclear resonance signal. Spin systems $N_Q$ and $N_P$ must be connected by dipole–dipole coupling, so that the exposure of one system is reflected in the state of the other. Nuclei with a strong NQR signal could be used as the $N_P$ nuclei, but most often nuclei with a strong NMR signal, e.g. protons are used for this purpose.

The DR techniques allow the successful detection of the NQR signal from such nuclei as $^{2}H$, $^{10}B$, $^{11}B$, $^{14}N$, $^{17}O$, $^{23}Na$, $^{25}Mg$ and $^{27}Al$, which can be present in biologically important molecules. The DR methods have developed and improved quickly, and now a whole range of methods has appeared. Usually DR methods for NQR are divided into the following main groups: DR in the rotating frame, DR in the laboratory frame, DR by level crossing, DR by continuous coupling and DR by solid-state effect. The areas and details of the use of these methods can be found in the reviews given in the Further reading section. It should be noted that the choice of a specific method depends on the type of the studied nuclei and the parameters of the sample, with relaxation times of both spin systems being very important.

In spite of a wide variety of different DR methods, the main stages of the experiment can very briefly be described as follows, using the widely used concept of spin temperature in magnetic resonance. Initially the $N_P$ spins are prepared in a polarized state. Then the polarization of the $N_Q$ spins occurs through thermal contact between the two spin systems. Another method widely used is resonant heating of the $N_Q$ spins by irradiating them with a RF field. As a result of the new thermal contact, the spin temperature of the $N_P$ spins changes, and this change is then registered as the NQR signal.

The experimental equipment for NQR DR consists first of all of a spectrometer for detecting nuclear resonance signals from the nuclei of $N_P$. As a rule, this is an ordinary NMR spectrometer and a DC magnet, which permits the detection of FID signals with 90° pulses. In addition, a DR spectrometer contains a sample-transfer system, an irradiating system (for the $N_Q$ nuclei), a cryostat, a controlling and automatic signal processing unit, including a PC computer, and also some additional modules depending on the chosen methods of detection, e.g. a source of additional magnetic fields.

**Figure 8** shows a typical block diagram of an NMR/NQR DR spectrometer. Such spectrometers require a special sample transit system. Systems of various design can be used for this purpose, including those with linear induction motors, step motors or compressed air systems. As a source of DC magnetic field, permanent magnets or electromagnets are used. Initially the sample is located in the field of this magnet, and the NMR signal is detected (in the P-coil), then the sample is transferred into the Q-coil,

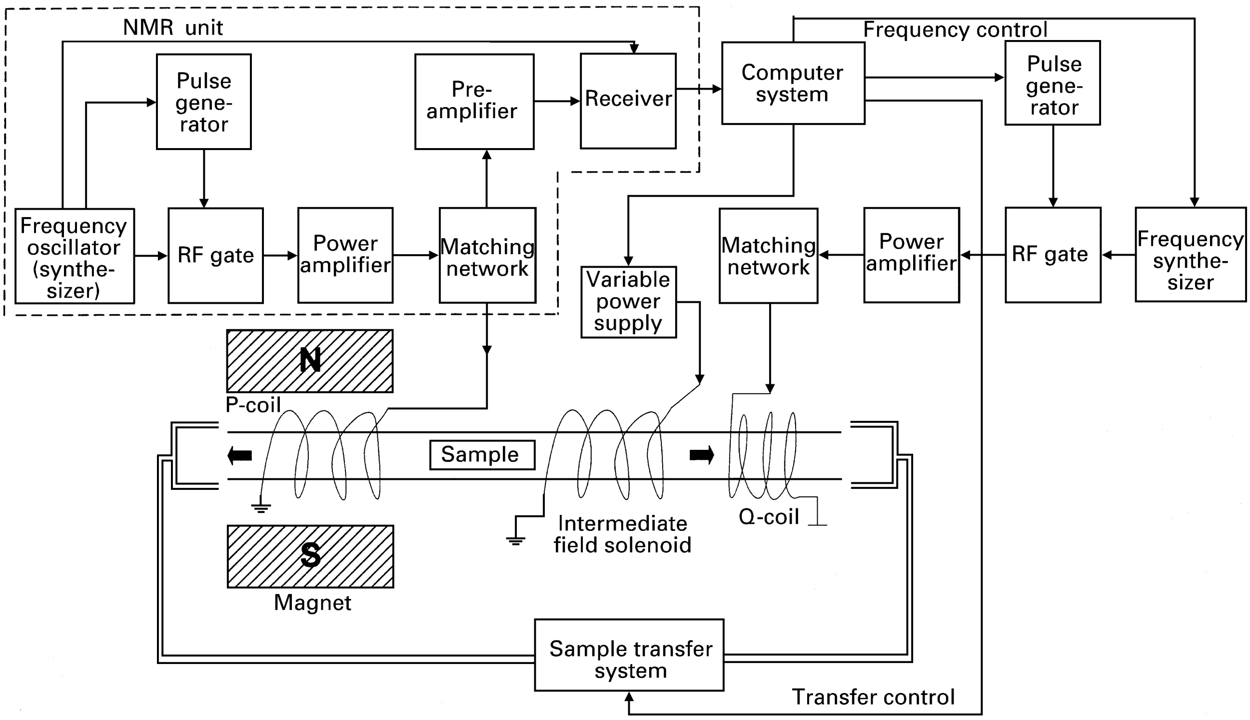



**Figure 8** Simplified block diagram of the double resonance (NQR/NMR) spectrometer.

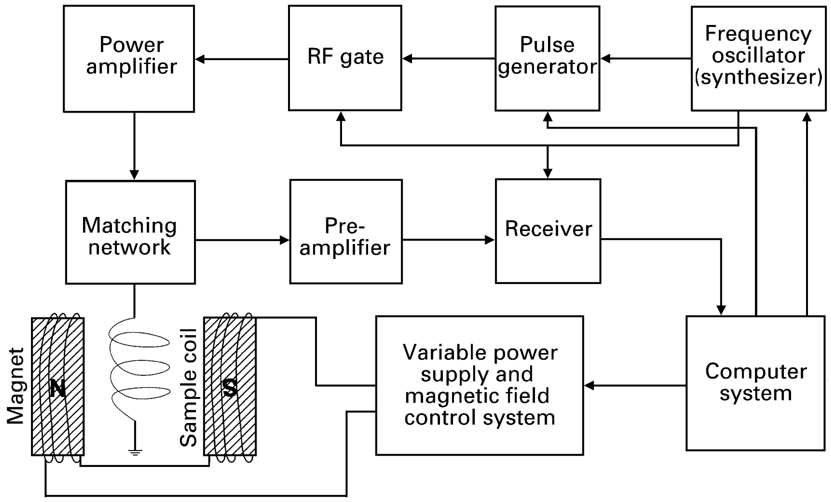



**Figure 9** Simplified block diagram of a cross-relaxation NQR spectrometer.

where it is irradiated with RF pulses at the frequency of quadrupole nuclei. After this, the sample is returned into the magnetic field, where the NMR signal is detected again. The change in the amplitude of the NMR signal detects the NQR. For some experiments an intermediate magnetic field can be used. It is also possible instead of the mechanical transfer of the sample from the magnetic field, to switch the field. For experiments in quadrupole (NQR/NQR) DR the external magnetic field need not necessarily be used. This refers to spin-echo DR, or DR in the rotating frame, for example.

### Cross-relaxation spectroscopy method

The cross-relaxation (CR) method is well known in magnetic resonance spectroscopy and can be successfully used for the detection of weak NQR signals in the low-frequency area in particular. A characteristic of this method is the absence of irradiation of the spin system at NQR frequencies. In principle this method can also be regarded as a version of DR. A block diagram of a spectrometer for implementing the method of CR spectroscopy, is given in **Figure 9**. It includes an NMR spectrometer, a controlling unit for regulating the magnetic field, and a computer system. The experiment is usually carried out in the following way. At first, as when using an ordinary DR method, the sample is acted on by a strong DC magnetic field, and NMR signals from the $N_P$ nuclei are detected. Then the magnetic field is changed to satisfy the conditions of CR between the levels of the $N_P$ and $N_Q$ nuclei. If the conditions of CR are satisfied, the amplitude of the NMR signal detected at the end of the measurement cycle will change, and these changes are then registered as the NQR signal.

## The use of amplifiers based on a DC SQUID

One of the new methods that permits effective detection of weak NQR signals at low frequencies, is the use of amplifiers based on a DC superconducting quantum interference device (SQUID). The operation of this is based on the use of two superconducting phenomena: flux quantization and Josephson tunnelling. In fact a SQUID represents a flux-to-voltage transducer with a very low noise temperature as compared with regular amplifiers, which is particularly important for low frequencies. The SQUID can be used in various methods for NQR detection, including CW methods as well as pulsed methods. Also, a SQUID amplifier does not have an input resonance circuit, and therefore permits the detection of signals over a broad bandwidth, from close to zero. Readers interested in the principles of operation, design and specificity of application for SQUIDs, can find more information in the literature, included in the Further reading section. NQR spectrometers using SQUIDs, are of great interest for many practical applications, particularly in areas where the use of indirect methods is complicated from the technical point of view. This technique is developing rapidly at present.

## List of symbols

$eq$ = electric field gradient; $eQ$ = nuclear electric quadrupole moment; $e^2Qq$ = nuclear quadrupole coupling constant; $Q$ = quality factor; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $T_2^*$ = line-shape parameter; $\Delta f$ = line width; $\eta$ = asymmetry parameter.

*See also:* **Fourier Transformation and Sampling Theory; NMR in Anisotropic Systems, Theory; NMR of Solids; NMR Pulse Sequences; NMR Spectrometers; Nuclear Quadrupole Resonance, Applications; Solid State NMR, Methods.**

## Further reading

Blinc R (1975) Double resonance detection of nuclear quadrupole resonance spectra. In: Smith JA (ed) *Advances in Nuclear Quadrupole Resonance*, Vol 2, pp 71–90. London: Heyden.

Clarke J (1993) SQUIDs: Theory and practice. In: Weinstock H and Ralston RW (eds) *The New Superconducting Electronics*. Dordrecht: Kluwer Academic.

Edmonds DT (1977) Nuclear quadrupole double resonance. *Physics Reports (Section C of Physics Letters)* **29**: 233–290.

Ernst RR, Bodenhausen G and Wokaun A (1987) In: Rowlinson JS, Green MLH, Halpern J, Mukaiyama T, Schowen RL, Thomas JM and Zewail AH (eds) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford: Clarendon Press.

Greenberg Ya S (1998) Application of superconducting quantum interference device to nuclear magnetic resonance. *Reviews of Modern Physics* **70**: 175–222.

Klainer SM, Hirschfeld TB and Marino R (1982) Fourier transform nuclear quadrupole resonance spectroscopy. In: Marshall AG (ed) *Fourier*, *Hadamard and Hilbert Transforms in Chemistry*, pp 147–182. New York: Plenum Press.

Read M (1974) A frequency and phase-locked super-regenerative oscillator spectrometer for nuclear quadrupole resonance at 200 MHz. In: Smith JA (ed) *Advances in Nuclear Quadrupole Resonance*, Vol 1, pp 203–226. London: Heyden.

Semin GK, Babushkina TA and Iakobson GG (1975) Pick AJ (ed) *Nuclear Quadrupole Resonance in Chemistry*. New York: Wiley.

Smith JAS (1986) Nuclear quadrupole interactions in solids. *Chemical Society Reviews* **15**: 225–260.

# Nuclear Quadrupole Resonance, Theory

**Janez Seliger**, University of Ljubljana and 'Jožef Stefan' Institute, Ljubljana, Slovenia

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Theory**

## Introduction

Nuclear quadrupole resonance (NQR) offers a unique means for the study of less structure, dynamics and chemical bonding in solids. A number of atomic nuclei possess nonzero electric quadrupole moments in the ground state. The interaction of a nuclear quadrupole moment with the local inhomogeneous electric field removes the degeneracy of the nuclear ground state. The transition frequencies between the nuclear quadrupole energy levels, which are typically in the MHz frequency region, depend on the nuclear quadrupole moment and on the electric field gradient (EFG) tensor at the nucleus. The electric quadrupole moment is a well-defined property of a nucleus in its ground state, whereas the EFG tensor depends on the electric charge distribution around the nucleus. The NQR frequencies thus indirectly provide valuable information on the local structure around the observed atom and its chemical bonding. Thermal motions in solids partially average out the EFG tensor, whereas thermal motions in isotropic liquids average the EFG tensor to zero. The temperature dependence of the NQR frequencies thus provides valuable information on the thermal behaviour of solids.

Nuclear constants and the natural abundance of naturally occurring quadrupolar nuclei are listed in **Table 1.**

## The Hamiltonian

We start by considering a nuclear electric charge distribution $\rho(\boldsymbol{r})$ placed in an external inhomogeneous electric field with the potential $V(\boldsymbol{r})$. The origin, $\boldsymbol{r} = 0$, is at the centre of gravity of the nucleus. The interaction energy $E$ is given by

$$E = \int \rho\,(\boldsymbol{r})V(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} \qquad [1]$$

where the integral is taken over the volume of $\rho(\boldsymbol{r})$. When the electrostatic potential varies weakly over the nuclear electric charge distribution, it can be expanded in a Taylor series about the origin to give the familiar multipole expansion of the energy:

$$E = V(0)\int \rho\,(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} + \sum_k \left(\frac{\partial V}{\partial x_k}\right)_0 \int x_k \rho(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} + \frac{1}{2}\sum_{k,l}\left(\frac{\partial^2 V}{\partial x_k\,\partial x_l}\right)_0 \int x_k\,x_l \rho(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} + \cdots \qquad [2]$$

The first term is the energy of the nuclear 'monopole' and does not depend on its orientation. It is thus of no interest to us. The second term is the dipole contribution, which vanishes because a nucleus in its ground state has definite parity and therefore zero electric dipole moment. In fact all odd nuclear electric multipole moments vanish by the same argument. The strongest term representing the orientation-dependent energy of a nucleus is thus the third, quadrupole, term in Equation [2]. The symmetric traceless second-rank tensor $(\partial^2 V/\partial x_k \partial x_l)_0 = V_{kl}$ is called the electric-field-gradient (EFG) tensor. Introducing a symmetric traceless second rank tensor $Q_{kl}$,

$$Q_{kl} = \int (3x_k\,x_l - r^2\delta_{kl})\rho\,(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} \qquad [3]$$

we may rewrite the quadrupole term $E_Q$ in Equation [2] as

$$E_Q = \frac{1}{6}\sum_{kl} Q_{kl}V_{kl} \qquad [4]$$

Here $Q_{kl}$ are the elements of the nuclear electric quadrupole moment tensor. Equation [4] is of course a classical expression. A quantum-mechanical expression for the quadrupole Hamiltonian $H_Q$ is obtained from Equation [4] by replacing the tensor elements $Q_{kl}$ by the operators $Q_{kl}^{(\mathrm{op})}$:

$$H_Q = \frac{1}{6}\sum_{kl} Q_{kl}^{(\mathrm{op})}V_{kl} \qquad [5]$$

The degeneracy of the nuclear ground state with a spin $I$ is $2I + 1$. $H_Q$ may be treated as a small

**Table 1** Nuclear constants and natural abundance of naturally occurring quadrupolar nuclei

| *Isotope* | *Natural abundance (%)* | *Spin* | *Magnetic dipole moment ($\mu n$)* | *NMR Frequency in B = 1 T (MHz)* | *Electrical quadrupole moment Q ($10^{-28}m^2$)* |
|---|---|---|---|---|---|
| $^{2}$H | 0.0156 | 1 | +0.8574 | 6.54 | 0.00286 |
| $^{6}$Li | 7.43 | 1 | +0.8220 | 6.27 | −0.00083 |
| $^{7}$Li | 92.57 | 3/2 | +3.2564 | 16.55 | −0.0406 |
| $^{9}$Be | 100 | 3/2 | −1.1778 | 5.98 | +0.053 |
| $^{10}$B | 18.83 | 3 | +1.8006 | 4.58 | +0.08472 |
| $^{11}$B | 81.17 | 3/2 | +2.6886 | 13.66 | +0.04065 |
| $^{14}$N | 99.64 | 1 | +0.4038 | 3.08 | +0.0193 |
| $^{17}$O | 0.037 | 5/2 | −1.8938 | 5.77 | −0.02578 |
| $^{21}$Ne | 0.257 | 3/2 | −0.6618 | 3.36 | +0.103 |
| $^{23}$Na | 100 | 3/2 | +2.2175 | 11.26 | +0.1006 |
| $^{25}$Mg | 10.05 | 5/2 | −0.8555 | 2.61 | +0.201 |
| $^{27}$Al | 100 | 5/2 | +3.6415 | 11.09 | +0.150 |
| $^{33}$S | 0.74 | 3/2 | +0.6438 | 3.27 | −0.076 |
| $^{35}$Cl | 75.4 | 3/2 | +0.8219 | 4.17 | −0.08249 |
| $^{37}$Cl | 24.6 | 3/2 | +0.6841 | 3.47 | −0.06493 |
| $^{39}$K | 93.08 | 3/2 | +0.3915 | 1.99 | +0.049 |
| $^{41}$K | 6.91 | 3/2 | +0.2149 | 1.09 | +0.060 |
| $^{43}$Ca | 0.13 | 7/2 | −1.3176 | 2.87 | −0.049 |
| $^{45}$Sc | 100 | 7/2 | +4.7565 | 10.33 | −0.22 |
| $^{47}$Ti | 7.75 | 5/2 | −0.7885 | 2.40 | +0.29 |
| $^{49}$Ti | 5.51 | 7/2 | −1.1042 | 2.40 | +0.24 |
| $^{50}$V | 0.24 | 6 | +3.3457 | 4.25 | +0.209 |
| $^{51}$V | 99.76 | 7/2 | +5.1487 | 11.19 | −0.052 |
| $^{53}$Cr | 9.54 | 3/2 | −0.4754 | 2.41 | −0.15 |
| $^{55}$Mn | 100 | 5/2 | +3.4532 | 10.55 | +0.33 |
| $^{59}$Co | 100 | 7/2 | +4.627 | 10.10 | +0.404 |
| $^{61}$Ni | 1.25 | 3/2 | −0.7500 | 3.81 | +0.162 |
| $^{63}$Cu | 69.09 | 3/2 | +2.2233 | 11.29 | −0.211 |
| $^{65}$Cu | 30.91 | 3/2 | +2.3817 | 12.09 | −0.195 |
| $^{67}$Zn | 4.12 | 5/2 | +0.8755 | 2.66 | +0.150 |
| $^{69}$Ga | 60.2 | 3/2 | +2.0166 | 10.22 | +0.168 |
| $^{71}$Ga | 39.8 | 3/2 | +2.5623 | 12.98 | +0.106 |
| $^{73}$Ge | 7.61 | 9/2 | −0.8795 | 1.49 | −0.173 |
| $^{75}$As | 100 | 3/2 | +1.4395 | 7.29 | +0.314 |
| $^{79}$Br | 50.57 | 3/2 | +2.1064 | 10.67 | +0.331 |
| $^{81}$Br | 49.43 | 3/2 | +2.2706 | 11.50 | +0.276 |
| $^{83}$Kr | 11.55 | 9/2 | −0.9707 | 1.64 | +0.253 |
| $^{85}$Rb | 72.8 | 5/2 | +1.3534 | 4.11 | +0.23 |
| $^{87}$Rb | 27.2 | 3/2 | +2.7518 | 13.39 | +0.127 |
| $^{87}$Sr | 7.02 | 9/2 | −1.0936 | 1.85 | +0.335 |
| $^{91}$Zr | 11.23 | 5/2 | −1.3036 | 3.96 | −0.206 |
| $^{93}$Nb | 100 | 9/2 | +6.1705 | 10.41 | −0.32 |
| $^{95}$Mo | 15.78 | 5/2 | −0.9142 | 2.77 | −0.022 |

**Table 1** *Continued*

| *Isotope* | *Natural abundance (%)* | *Spin* | *Magnetic dipole moment ($\mu n$)* | *NMR Frequency in B = 1 T (MHz)* | *Electrical quadrupole moment Q ($10^{-28}m^2$)* |
|---|---|---|---|---|---|
| $^{97}$Mo | 9.60 | 5/2 | −0.9335 | 2.83 | +0.255 |
| $^{99}$Ru | 12.81 | 5/2 | −0.641 | 1.95 | +0.079 |
| $^{101}$Ru | 16.98 | 5/2 | −0.7188 | 2.19 | +0.457 |
| $^{105}$Pd | 22.23 | 5/2 | −0.642 | 1.96 | +0.660 |
| $^{113}$In | 4.16 | 9/2 | +5.5289 | 9.31 | +0.799 |
| $^{115}$In | 95.84 | 9/2 | +5.5408 | 9.33 | +0.86 |
| $^{121}$Sb | 57.25 | 5/2 | +3.3634 | 10.25 | −0.36 |
| $^{123}$Sb | 42.75 | 7/2 | +2.5498 | 5.52 | −0.49 |
| $^{127}$I | 100 | 5/2 | +2.8133 | 8.52 | −0.79 |
| $^{131}$Xe | 21.24 | 3/2 | +0.6919 | 3.49 | −0.120 |
| $^{133}$Cs | 100 | 7/2 | +2.5820 | 5.59 | −0.00371 |
| $^{135}$Ba | 6.59 | 3/2 | +0.8379 | 4.23 | +0.160 |
| $^{137}$Ba | 11.32 | 3/2 | +0.9374 | 4.73 | +0.245 |
| $^{138}$La | 0.089 | 5 | +3.7136 | 5.62 | +0.45 |
| $^{139}$La | 99.911 | 7/2 | +2.7830 | 6.01 | +0.20 |
| $^{141}$Pr | 100 | 5/2 | +4.2754 | 13.04 | −0.0589 |
| $^{143}$Nd | 12.20 | 7/2 | −1.065 | 2.32 | −0.63 |
| $^{145}$Nd | 8.30 | 7/2 | −0.656 | 1.43 | −0.33 |
| $^{147}$Sm | 15.07 | 7/2 | −0.8148 | 1.77 | −0.26 |
| $^{149}$Sm | 13.84 | 7/2 | −0.6717 | 1.46 | +0.075 |
| $^{151}$Eu | 47.77 | 5/2 | +3.4717 | 10.59 | +0.903 |
| $^{153}$Eu | 52.23 | 5/2 | +1.5330 | 4.64 | +2.412 |
| $^{155}$Gd | 14.68 | 3/2 | −0.2591 | 1.32 | +1.30 |
| $^{157}$Gd | 15.64 | 3/2 | −0.3398 | 1.73 | +1.36 |
| $^{159}$Tb | 100 | 3/2 | −2.014 | 10.23 | +1.432 |
| $^{161}$Dy | 18.73 | 5/2 | −0.4803 | 1.46 | +2.507 |
| $^{163}$Dy | 24.97 | 5/2 | +0.6726 | 2.05 | +2.648 |
| $^{165}$Ho | 100 | 7/2 | +4.132 | 9.00 | +3.58 |
| $^{167}$Er | 22.82 | 7/2 | −0.5639 | 1.23 | +3.565 |
| $^{173}$Yb | 16.08 | 5/2 | −0.6799 | 2.07 | +2.80 |
| $^{175}$Lu | 97.40 | 7/2 | +2.2327 | 4.86 | +3.49 |
| $^{176}$Lu | 2.60 | 7 | +3.1692 | 3.45 | +4.92 |
| $^{177}$Hf | 18.39 | 7/2 | +0.7935 | 1.73 | +3.365 |
| $^{179}$Hf | 13.78 | 9/2 | −0.6409 | 1.09 | +3.79 |
| $^{181}$Ta | 99.99 | 7/2 | +2.3705 | 5.16 | +3.28 |
| $^{185}$Re | 37.07 | 5/2 | +3.1871 | 9.59 | +2.18 |
| $^{187}$Re | 62.93 | 5/2 | +3.2197 | 9.68 | +2.07 |
| $^{189}$Os | 16.1 | 3/2 | +0.6599 | 3.31 | +0.856 |
| $^{191}$Ir | 38.5 | 3/2 | +0.1507 | 0.81 | −0.816 |
| $^{193}$Ir | 61.5 | 3/2 | +0.1637 | 0.83 | +0.751 |
| $^{197}$Au | 100 | 3/2 | +0.1457 | 0.73 | +0.547 |
| $^{201}$Hg | 13.24 | 3/2 | −0.5602 | 2.85 | +0.385 |
| $^{209}$Bi | 100 | 9/2 | +4.1106 | 6.84 | −0.37 |
| $^{235}$U | 0.71 | 7/2 | −0.38 | 0.83 | +4.55 |

perturbation that at least partially removes the degeneracy of the nuclear ground state. To calculate the energies of the nuclear quadrupole energy levels and the corresponding NQR frequencies, it is necessary to evaluate the matrix elements $\langle I,m\,|H_Q|\,I,m'\rangle$. Here $I$ is the nuclear spin and $m$ and $m'$ are the magnetic quantum numbers. To evaluate the above matrix elements of $H_Q$, it is necessary to evaluate the matrix elements

$$\left\langle I,m|Q_{kl}^{(\mathrm{op})}|I,m'\right\rangle \quad [6]$$

of the quadrupole moment tensor. Using the Wigner–Eckart theorem, we may replace the matrix elements given by Equation [6] by the matrix elements

$$\begin{aligned}&\langle I,m|\tilde{Q}_{kl}|I,m'\rangle\\&=\left\langle I,m\middle|C\left(\frac{3}{2}(I_kI_l+I_lI_k)-\delta_{kl}\boldsymbol{I}^2\right)\middle|I,m'\right\rangle\end{aligned} \quad [7]$$

of a symmetric, traceless, second rank tensor $\tilde{Q}$ composed of the components of the nuclear angular momentum vector $\boldsymbol{I}$. We define a scalar constant $eQ$ called the nuclear quadrupole moment as $eQ = \langle I,I\,|\tilde{Q}_{zz}|\,I,I\rangle$ and express the proportionality $C$ as $C = eQ/I(2I{-}1)$. The quadrupole Hamiltonian may thus be expressed as

$$H_Q=\frac{eQ}{6I(2I-1)}\sum_{kl}V_{kl}\left[\frac{3}{2}(I_kI_l+I_lI_k)-\delta_{kl}\boldsymbol{I}^2\right] \quad [8]$$

The elements $V_{kl}$ depend on the choice of the coordinate system. The simplest situation arises in a coordinate system where the coordinate axes point along the principal directions of the EFG tensor. In this coordinate system the EFG tensor is diagonal. Let us denote the three principal values of the EFG tensor as $V_{XX}$,$V_{YY}$ and $V_{ZZ}$ ($|\,V_{XX}|\le|\,V_{YY}|\le|\,V_{ZZ}|$)| and the corresponding three principal directions as $X$,$Y$ and $Z$. The three principal values of the traceless EFG tensor are not independent. They can be described by two parameters:

$$eq=V_{ZZ} \qquad \text{and} \qquad \eta=(V_{XX}-V_{YY})/V_{ZZ} \quad [9]$$

The quantity $eq$ is thus the largest principal value of the EFG tensor measuring its magnitude, whereas the asymmetry parameter $\eta$ measures the departure of the EFG tensor from axial symmetry. The asymmetry parameter $\eta$ ranges between 0 and 1. Let us further choose the principal axis $Z$ as the quantization axis and rewrite the quadrupole Hamiltonian as

$$H_Q=\frac{e^2qQ}{4I(2I-1)}\left[3I_Z^2-\boldsymbol{I}^2+\frac{\eta}{2}(I_+^2+I_-^2)\right] \quad [10]$$

The quantity $e^2qQ$ divided by the Planck constant $h$ is called the quadrupole coupling constant. It measures the magnitude of the nuclear quadrupole interaction. The NQR frequencies of a given nucleus depend on its spin $I$ and on two parameters, $e^2qQ/h$ and $\eta$, related to the EFG tensor.

In solid-state NMR of quadrupolar nuclei, $H_Q$ often represents a perturbation of the Zeeman interaction. It is then natural to choose the quantization axis $z$ along the direction of the external magnetic field and rewrite Equation [8] in a convenient form:

$$\begin{aligned}H_Q=\frac{eQ}{4I(2I-1)}&\left[V_0(3I_z^2-\boldsymbol{I}^2)+V_{+1}(I_-I_z+I_zI_-)\right.\\&\left.+V_{-1}(I_+I_z+I_zI_+)+V_{+2}I_-^2+V_{-2}I_+^2\right]\end{aligned} \quad [11]$$

Here

$$V_0=V_{zz},\ V_{\pm1}=V_{xz}\pm \mathrm{i}V_{yz}$$

and

$$V_{\pm2}=(V_{xx}-V_{yy})/2\pm \mathrm{i}V_{xy}$$

## Nuclear quadrupole energy levels and resonance frequencies

It is obvious from Equation [10] that $H_Q$ mixes the states with $\Delta m = 2$ when $\eta \neq 0$ and the quantization axis points along the $Z$ principal axis of the EFG tensor. We shall from now on assume that the quantization axis points along the $Z$ principal axis of the EFG tensor and expand the eigenstates $|\psi\rangle$ of $H_Q$ in the representation of the eigenstates $|I, m\rangle$ of $I_Z$. The set of states $|I, m\rangle$ can first be divided into two subsets: (a) $|I, I\rangle$, $|I, I{-}2\rangle$, etc. and (b) $|I, I{-}1\rangle$, $|I, I{-}3\rangle$, etc.

Each eigenstate of $H_Q$ is represented either by the states of the subset (a) or by the states of the subset (b). $H_Q$ is also invariant to time inversion. Let $|\psi,t\rangle$ be an eigenstate of $H_Q$. Owing to the time-inversion symmetry of $H_Q$ the state $|\psi,-t\rangle$ obtained after the inversion of time is also an eigenstate of $H_Q$ with the same energy as $|\psi, t\rangle$. A state $|I, m\rangle$ transforms

under the time inversion into the state $(-1)^{I-m}||I, -m\rangle$. In the case of an integer spin nucleus the states $|I\ m\rangle$ and $|I, -m\rangle$ belong to the same subset of states and $|\psi,-t\rangle$ is generally equal to $|\psi, t\rangle$ multiplied by a phase factor. The energy levels are thus nondegenerate, except for $\eta = 0$. In the case of half-integer spin nuclei the states $|I, m\rangle$ and $|I, -m\rangle$ belong to two different subsets of states and $|\psi, -t\rangle$ is different from $|\psi, t\rangle$. The energy levels are generally doubly degenerate. This is called Kramers degeneracy. The inhomogeneous electric field thus completely removes the degeneracy of the nuclear ground state only for integer spin nuclei when $\eta \neq 0$. In case of a half-integer spin nucleus with a spin $I$, we observe only $I + 1/2$ doubly degenerate nuclear quadrupole energy levels.

We shall first consider nuclei with an integer spin 1, 2 and 3. In all these cases the energies of the nuclear quadrupole energy levels can be calculated analytically.

## Spin 1

The energies $E$ of the three nuclear quadrupole energy levels and the expansion coefficients $c^m$ of the corresponding eigenstates $|\psi\rangle = \Sigma_m C^m|1, m\rangle$, in the representation of the eigenstates of $\boldsymbol{I}_z$ are given in **Table 2**.

The upper energy level in the case of $\eta = 0$ is doubly degenerate and only one NQR frequency, $\nu_Q = 3e^2qQ/4h$, is observed. For a nonzero $\eta$, the three nuclear quadrupole less resonance frequencies $\nu_+, \nu_-$ and $\nu_0$ $(\nu_+ \geq \nu_- > \nu_0)$ are

$$\begin{aligned}\nu_+ &= \frac{e^2qQ}{4h}(3+\eta)\\ \nu_- &= \frac{e^2qQ}{4h}(3-\eta)\\ \nu_0 &= \nu_+ - \nu_- = \frac{e^2qQ}{2h}\eta\end{aligned} \quad [12]$$

The energies of the nuclear quadrupole energy levels and the NQR frequencies as functions of $\eta$ are shown in **Figure 1**.

**Table 2** Energies $E$ in units of $e^2qQ/4$ and the expansion coefficients $c^m$ of the eigenstates of $\boldsymbol{H}_Q$ for a nucleus with $I = 1$ in the representation of the eigenstates of $\boldsymbol{I}_Z$

| $E$ | $c^1$ | $c^0$ | $c^{-1}$ |
|---|---|---|---|
| $1+\eta$ | $1/\sqrt{2}$ | 0 | $1/\sqrt{2}$ |
| $1-\eta$ | $1/\sqrt{2}$ | 0 | $1/\sqrt{2}$ |
| −2 | 0 | 1 | 0 |

None of the NQR transitions is generally forbidden except for special orientations of the radiofrequency magnetic field. The NQR line at $\nu = \nu_+$ vanishes when the resonant radiofrequency magnetic field points perpendicularly to the principal direction $X$. Similarly, the NQR line at $\nu = \nu_-$ vanishes when the resonant radiofrequency magnetic field points perpendicularly to the principal direction $Y$, and the NQR line at $\nu = \nu_0$ vanishes when the resonant radiofrequency magnetic field points perpendicularly to the principal direction $Z$. In a powder sample all three NQR lines can be observed, whereas in a single crystal the orientation dependence of the intensity of the NQR lines may be used to determine the orientation of the principal axes of the EFG tensor in a crystal-fixed coordinate system. The quadrupole coupling constant is usually calculated from the NQR frequencies as $e^2qQ/h = 2(\nu_+ + \nu_-)/3$, and the asymmetry parameter $\eta$ is equal to $\eta = 3(\nu_+ - \nu_-)/(\nu_+ + \nu_-)$.

## Spin 2 and 3

The energies of the nuclear quadrupole energy levels and expansion coefficients $c^m$ of the corresponding eigenstates of $H_Q$ in the representation of the eigenstates of $I_Z$, $|I, m\rangle$ for spins 2 and 3 are given in **Table 3** and **Table 4**, respectively.

Nuclei with an integer spin larger than 1 are seldom observed in practice. **Tables 3** and **4** are therefore included only for completeness.

## Spin $\frac{3}{2}$

As seen from **Table 1**, a half-integer nuclear spin is in practice much more common than an integer nuclear spin. Nuclei with a half-integer spin are often observed in practice. As already mentioned, the nuclear quadrupole energy levels of the half-integer spin nuclei are generally doubly degenerate. The two eigenstates of $H_Q$, $|\psi_+\rangle$ and $|\psi_-\rangle$, corresponding to the same doubly degenerate energy level are generally expressed as

$$\begin{aligned}|\Psi_+\rangle &= \sum_{k=0}^{I-1/2} c^k|I, I-2k\rangle\\ |\Psi_-\rangle &= \sum_{k=0}^{I-1/2} c^k|I, -I+2k\rangle\end{aligned} \quad [13]$$

The energies of the nuclear quadrupole energy levels and the corresponding eigenstates of $H_Q$ can in the general case $(\eta \neq 0)$ be expressed analytically

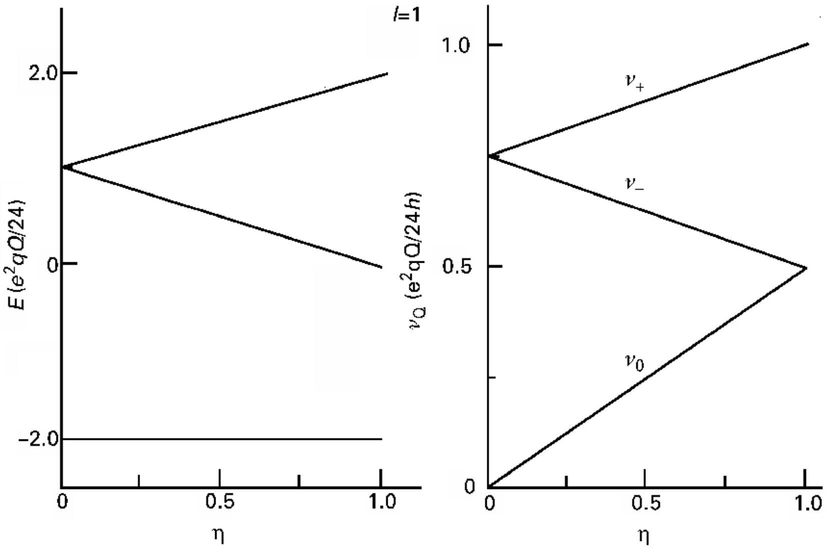



**Figure 1** Energy levels and NQR frequencies for $I = 1$.

only for $I = \frac{3}{2}$, where

$$E_{\pm 3/2} = \pm \frac{e^2qQ}{4}\sqrt{1+\frac{\eta^2}{3}} \qquad [14]$$

and the eigenstates of $H_Q$ are

$$\left|\Psi_{\pm 3/2}\right\rangle = \sqrt{\frac{z+3}{2z}}\left|\frac{3}{2},\pm\frac{3}{2}\right\rangle + \sqrt{\frac{z-3}{2z}}\left|\frac{3}{2},\mp\frac{1}{2}\right\rangle$$

$$\left|\Psi_{\pm 1/2}\right\rangle = -\sqrt{\frac{z-3}{2z}}\left|\frac{3}{2},\pm\frac{3}{2}\right\rangle + \sqrt{\frac{z+3}{2z}}\left|\frac{3}{2},\mp\frac{1}{2}\right\rangle$$

$$z = \sqrt{9+3\eta^2} \qquad [15]$$

Only one NQR frequency $\nu_Q$,

$$\nu_Q = \frac{e^2qQ}{2h}\sqrt{1+\frac{\eta^2}{3}} \qquad [16]$$

is observed in this case. The quadrupole coupling constant $e^2qQ/h$ and the asymmetry parameter $\eta$ cannot be determined separately from the NQR frequency. The problem is usually solved by the application of a weak magnetic field or by the application of two-dimensional NQR techniques.

## Spin $\frac{5}{2}$

The energies $E$ of the three nuclear quadrupole energy levels are obtained from the secular equation

$$x^3 - 7(3+\eta^2)x - 20(1-\eta^2) = 0 \qquad [17]$$

Here energy $E$ is given as $E = (e^2qQ)/20x$, where $x$ is a solution of the secular equation. The energies are usually labelled as $E_m$, where $m$ is the magnetic quantum number which can be assigned to a given energy level when $\eta = 0$. The three NQR frequencies are labelled as $\nu_{5/2-1/2}$, $\nu_{5/2-3/2}$ and $\nu_{3/2-1/2}$ ($\nu_{5/2-1/2} > \nu_{5/2-3/2} \geq \nu_{3/2-1/2}$). The energies $E_m$, and the NQR frequencies are shown in **Figure 2**.

The NQR line at the frequency $\nu_{5/2-1/2}$, $\nu_{5/2-1/2} = \nu_{5/2-3/2} + \nu_{3/2-1/2}$ is generally weaker than the other two NQR lines and cannot be observed when $\eta = 0$. The asymmetry parameter $\eta$ is in practice calculated from the ratio $R = \nu_{3/2-1/2}/\nu_{5/2-3/2}$ which ranges from $R = 0.5$ for $\eta = 0$ to $R = 1$ for $\eta = 1$. When $\eta$ is known, the quadrupole coupling constant can be calculated from any NQR frequency, most precisely from the highest NQR frequency $\nu_{5/2-1/2}$.

**Table 3** Energies of the nuclear quadrupole energy levels in units of $e^2qQ/8$ and the expansion coefficients $c^m$ for $I = 2$

| $E$ | $c^2$ | $c^1$ | $c^0$ | $c^{-1}$ | $c^{-2}$ |
|---|---|---|---|---|---|
| 2z | $\sqrt{\frac{z+1}{4z}}$ | 0 | $\sqrt{\frac{z-1}{2z}}$ | 0 | $\sqrt{\frac{z+1}{4z}}$ |
| 2 | $\frac{1}{\sqrt{2}}$ | 0 | 0 | 0 | $-\frac{1}{\sqrt{2}}$ |
| $-(1-\eta)$ | 0 | $\frac{1}{\sqrt{2}}$ | 0 | $\frac{1}{\sqrt{2}}$ | 0 |
| $-(1+\eta)$ | 0 | $\frac{1}{\sqrt{2}}$ | 0 | $-\frac{1}{\sqrt{2}}$ | 0 |
| $-2z$ | $-\sqrt{\frac{z-1}{4z}}$ | 0 | $\sqrt{\frac{z+1}{2z}}$ | 0 | $-\sqrt{\frac{z-1}{4z}}$ |

$z = \sqrt{1+\eta^2/3}$

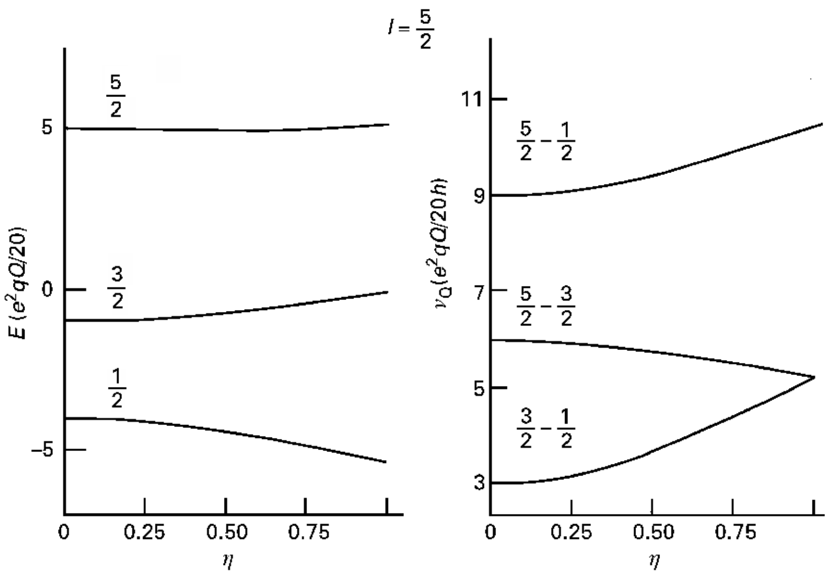



**Figure 2** Energy levels and NQR frequencies for $I = \frac{5}{2}$.

## Spin $\frac{7}{2}$

The four energies $E$ of the nuclear quadrupole energy levels are calculated from the secular equation

$$x^4 - 42\left(1 + \frac{\eta^2}{3}\right)x^2 - 64(1-\eta^2)x + 105\left(1 + \frac{\eta^2}{3}\right)^2 = 0 \qquad [18]$$

where $E = e^2qQx/28$. They are again labelled as $E_m$, $m = \frac{7}{2}, \frac{5}{2}, \frac{3}{2}, \frac{1}{2}$. The dependence of the energies $E_m$ and of the NQR frequencies $\nu_{m-(m-1)} = (E_m - E_{m-1})/h$ on the symmetry parameter $\eta$ is shown in **Figure 3**.

The three NQR frequencies corresponding to $\Delta m = 1$ give the strongest NQR signals. The NQR signals at the frequencies corresponding to $\Delta m = 2$ and $\Delta m = 3$ are also observed for large values of $\eta$, but their intensities are lower than the intensities of the NQR lines corresponding to $\Delta m = 1$. As seen from **Figure 3**, the NQR frequency $\nu_{3/2-1/2}$ depends strongly on $\eta$, whereas the $\eta$-dependence of the

**Table 4** Energies of the eigenstates of $H_Q$ in units of $e^2qQ/20$ and the expansion coefficients $c^m$ for $I = 3$

| $E$ | $c^3$ | $c^2$ | $c^1$ | $c^0$ | $c^{-1}$ | $c^{-2}$ | $c^{-3}$ |
|---|---|---|---|---|---|---|---|
| $1+\eta+4x$ | $\sqrt{\frac{4x+4-\eta}{16x}}$ | 0 | $\sqrt{\frac{4x-4+\eta}{16x}}$ | 0 | $\sqrt{\frac{4x-4+\eta}{16x}}$ | 0 | $\sqrt{\frac{4x+4-\eta}{16x}}$ |
| $1-\eta+4y$ | $\sqrt{\frac{4y+4+\eta}{16y}}$ | 0 | $\sqrt{\frac{4y-4-\eta}{16y}}$ | 0 | $-\sqrt{\frac{4y-4-\eta}{16y}}$ | 0 | $-\sqrt{\frac{4y+4+\eta}{16y}}$ |
| $2z-2$ | 0 | $\sqrt{\frac{z+1}{4z}}$ | 0 | $\sqrt{\frac{z-1}{2z}}$ | 0 | $\sqrt{\frac{z+1}{4z}}$ | 0 |
| 0 | 0 | $\frac{1}{\sqrt{2}}$ | 0 | 0 | 0 | $-\frac{1}{\sqrt{2}}$ | 0 |
| $1+\eta-4x$ | $-\sqrt{\frac{4x-4+\eta}{16x}}$ | 0 | $\sqrt{\frac{4x+4-\eta}{16x}}$ | 0 | $\sqrt{\frac{4x+4-\eta}{16x}}$ | 0 | $-\sqrt{\frac{4x-4+\eta}{16x}}$ |
| $1-\eta-4y$ | $-\sqrt{\frac{4y-4-\eta}{16y}}$ | 0 | $\sqrt{\frac{4y+4+\eta}{16y}}$ | 0 | $-\sqrt{\frac{4y+4+\eta}{16y}}$ | 0 | $\sqrt{\frac{4y-4-\eta}{16y}}$ |
| $-2-2z$ | 0 | $-\sqrt{\frac{z-1}{4z}}$ | 0 | $\sqrt{\frac{z+1}{2z}}$ | 0 | $-\sqrt{\frac{z-1}{4z}}$ | 0 |

$x = \sqrt{1 - \frac{\eta}{2} + \frac{\eta^2}{6}}$, $y = \sqrt{1 + \frac{\eta}{2} + \frac{\eta^2}{6}}$, $z = \sqrt{1 + \frac{5\eta^2}{3}}$

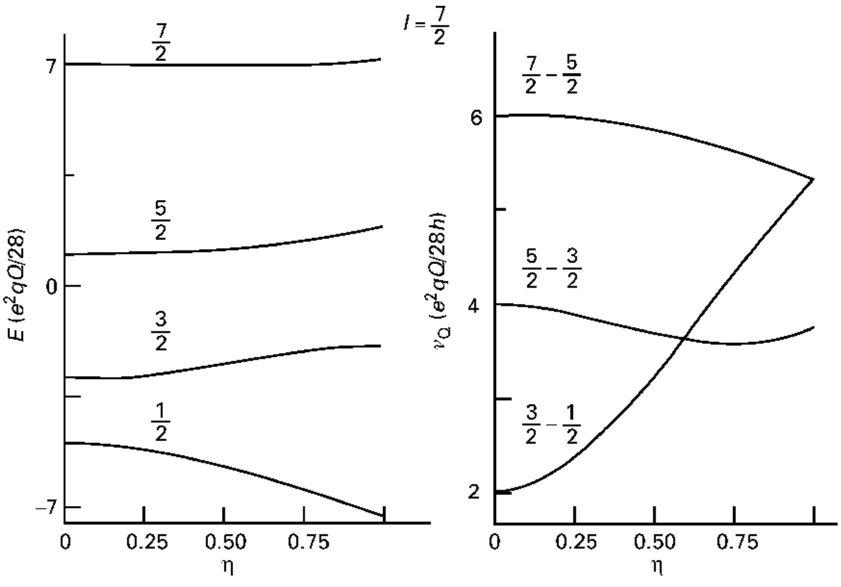



**Figure 3** Energy levels and NQR frequencies $\nu_{m-(m-1)}$ for $I = \frac{7}{2}$.

other two NQR frequencies is weaker. The asymmetry parameter $\eta$ is in practice determined either from the ratio $\nu_{3/2-1/2} / \nu_{5/2-3/2}$ or from the ratio $\nu_{3/2-1/2}/\nu_{7/2-5/2}$. When $\eta$ is known, the quadrupole coupling constant is calculated from any NQR frequency, most precisely from the highest NQR frequency observed.

### Spin $\frac{9}{2}$

The highest half-integer nuclear spin of a stable nucleus is $I = 9/2$. The energy $E$ of a nuclear quadrupole energy level is given as $E = e^2qQx/24$, where $x$ is a solution of the secular equation

$$x^5 - 11(3+\eta^2)x^3 - 44(1-\eta^2)x^2 + \frac{44}{3}(3+\eta^2)^2x + 48(3+\eta^2)(1-\eta^2) = 0 \qquad [19]$$

The energies of the nuclear quadrupole energy levels are again labelled as $E_m$, with $m$ being the magnetic quantum number assigned to a quadrupole energy level when $\eta = 0$. The dependence of the energies $E_m$ and of the NQR frequencies $\nu_{m-(m-1)}$ on the asymmetry parameter $\eta$ is shown in **Figure 4**.

The lowest NQR frequency $\nu_{3/2-1/2}$ also in this case exhibits the strongest dependence on $\eta$. The asymmetry parameter $\eta$ is in practice determined from a ratio of the NQR frequencies, say $\nu_{3/2-1/2} / \nu_{5/2-3/2}$. When $\eta$ is known, the quadrupole coupling constant $e^2qQ/h$ is calculated from any NQR frequency.

## Application of a weak magnetic field: Zeeman perturbed NQR

A weak static magnetic field is often used in NQR. In a powder sample it may cause broadening of a NQR line and consequently the disappearance of a NQR signal. In a single crystal a weak external magnetic field removes the degeneracy of the doubly degenerate quadrupolar energy levels. In the case of a half-integer quadrupolar nucleus, each NQR line splits into a quartet. The splitting depends on the orientation of the external magnetic field in the principal coordinate system of the EFG tensor. The orientation dependence of the splitting of the NQR lines gives the orientation of the principal axes of the EFG tensor in a crystal-fixed coordinate system and, for the case $I = 3/2$, also the value of the asymmetry parameter $\eta$. When $I$ is integer, the external magnetic field slightly shifts the resonance frequencies. The orientation dependence of the frequency shift makes it possible to determine the orientation of the principal axes of the EFG tensor in a crystal-fixed coordinate system. In both cases the multiplicity of the resonance lines in nonzero magnetic field gives the number of magnetically nonequivalent nuclei in the crystal unit cell. Here we treat in detail only the situation for two nuclear spin systems $I = 1$ and $I = \frac{3}{2}$.

### Spin 1

The Hamiltonian is

$$H = H_Q - h\nu_L(\boldsymbol{In}) \qquad [20]$$

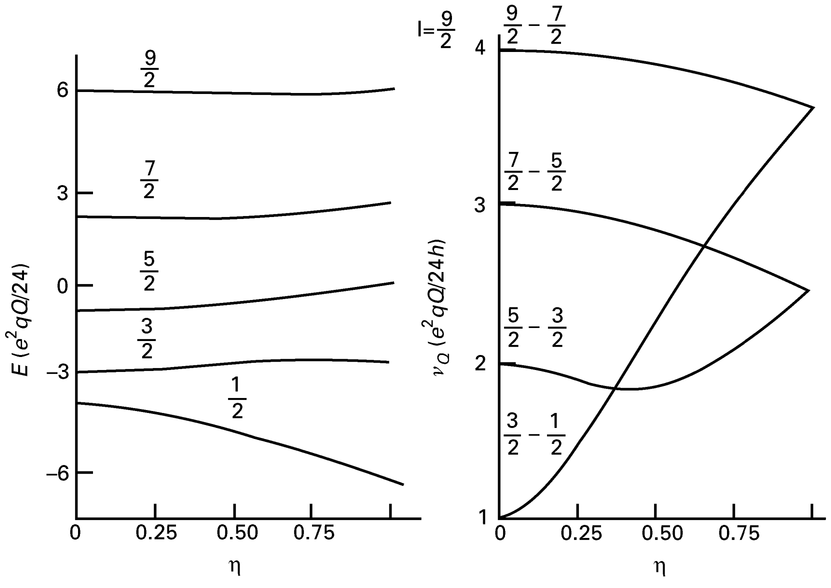



**Figure 4** Energy levels and NQR frequencies $\nu_{m-(m-1)}$ for $I=\frac{9}{2}$.

Here $H_Q$ is given by Equation [10], $\nu_L = \gamma B/2\pi$ is the Larmor frequency of a nucleus in the external magnetic field $\boldsymbol{B}$ and $\boldsymbol{n}$ is a unit vector in the direction of $\boldsymbol{B}$. We assume that the second term in equation [20] may be treated as a perturbation and that $\nu_L \ll \nu_0$. The eigenstates of $H_Q$ and their energies are given in **Table 2** for the case of $I=1$. The expectation value $\langle\psi|I|\psi\rangle$ for any nondegenerated eigenstate of $H_Q$ is identically equal to zero when $I$ is integer. The first-order perturbation corrections of the energies are thus equal to zero. The lowest-order nonzero terms are thus the second-order terms. Let us denote the resonance frequencies of $H$ as $\nu_+^*$, $\nu_-^*$ and $\nu_0^*$. Using the second-order perturbation theory, we obtain

$$\begin{aligned}
\nu_+^* &= \nu_+ + \nu_L^2\left\{\frac{1}{\nu_0} - \sin^2(\theta)\left[\left(\frac{1}{\nu_0} - \frac{2}{\nu_+}\right)\right.\right.\\
&\qquad \left.\left. + \left(\frac{2}{\nu_+} - \frac{1}{\nu_-}\right)\sin^2(\theta)\right]\right\}\\
\nu_-^* &= \nu_- + \nu_L^2\left\{-\frac{1}{\nu_0} + \sin^2(\theta)\left[\left(\frac{1}{\nu_0} + \frac{1}{\nu_+}\right)\right.\right.\\
&\qquad \left.\left. + \left(\frac{2}{\nu_-} - \frac{1}{\nu_+}\right)\sin^2(\theta)\right]\right\}\\
\nu_0^* &= \nu_0 + \nu_L^2\left\{\frac{2}{\nu_0} - \sin^2(\theta)\left[\left(\frac{2}{\nu_0} - \frac{1}{\nu_+}\right)\right.\right.\\
&\qquad \left.\left. + \left(\frac{1}{\nu_+} + \frac{1}{\nu_-}\right)\sin^2(\theta)\right]\right\}
\end{aligned} \qquad [21]$$

Here $\theta$ is the angle between the unit vector $\boldsymbol{n}$ and the principal axis $Z$ of the EFG tensor, whereas $\phi$ is the angle between the projection of $\boldsymbol{n}$ on the $X$–$Y$ plane and the principal axis $X$. It is clear from Equation [21] that the magnetic shifts of the NQR frequencies depend on the orientation of the external magnetic field with respect to the principal axes of the EFG tensor. In general in a magnetic field a NQR line splits into more lines. The multiplicity of the resonance line gives the number of crystallographically equivalent but magnetically nonequivalent positions of the studied atom in the unit cell. The crystallographically equivalent nuclei occupy positions where the principal values of the EFG tensor are equal. The magnetically nonequivalent sites differ in the orientation of the principal axes of the EFG tensor. An orientation dependence of the frequency shifts makes it possible to determine for each nuclear position the orientation of the principal axes $X$, $Y$ and $Z$ in a crystal-fixed coordinate system.

In the case of a low value of $\eta$ with $\nu_0 \ll \nu_+, \nu_-$ when the lowest NQR frequency $\nu_0$ may be comparable to the Larmor frequency $\nu_L$, the first-order perturbation calculation gives

$$\begin{aligned}
\nu_\pm^* &= 3K \pm \sqrt{\eta^2K^2 + \nu_L^2\cos^2\theta}\\
\nu_0^* &= 2\sqrt{\eta^2K^2 + \nu_L^2\cos^2\theta}
\end{aligned} \qquad [22]$$

Here $K = e^2qQ/4h$. A splitting of the upper energy level, and consequently three resonance frequencies, are observed even when $\eta = 0$ if $B \neq 0$.

## Spin $\frac{3}{2}$

The Hamiltonian is given by Equation [20] where the second term is treated as a perturbation. The energies and eigenstates of the unperturbed Hamiltonian $H_Q$

are given by Equations [14] and [15]. By first-order perturbation theory, the upper energy level splits into two energy levels with the energies $E_{3/2} \pm h\delta_{3/2}$ where

$$\delta_{3/2} = \frac{\nu_L}{2z}\sqrt{\begin{array}{l}(z+6)^2 + 3(z^2 - 6z - 18)\sin^2\theta) \\ +6\eta(z-3)\sin^2(\theta)\cos(2\phi)\end{array}} \qquad [23]$$

Here $z = \sqrt{(9 + 3\eta^2)}$. The lower energy level splits into two energy levels with the energies $E_{1/2} \pm h\delta_{1/2}$ where

$$\delta_{1/2} = \frac{\nu_L}{2z}\sqrt{\begin{array}{l}(z-6)^2 + 3(z^2 + 6z - 18)\sin^2(\theta) \\ -6\eta(z+3)\sin^2(\theta)\cos(2\phi)\end{array}} \qquad [24]$$

The NQR line splits a magnetic field into four lines. In a crystal a set of four lines is observed for each magnetically nonequivalent atomic site. From the orientation dependence of the splitting it is again possible to determine the orientation of the principal axes of the EFG tensor in a crystal-fixed coordinate system.

In practice, NQR spectroscopists often locate zero splitting where the frequency of two inner resonance lines is equal to $\nu_\theta$. This happens when $\delta_{3/2} = \delta_{1/2}$ or, consequently,

$$3\cos^2(\theta_0) - 1 + \eta\sin^2(\theta_0)\cos(2\phi) = 0 \qquad [25]$$

For zero $\eta$, zero splitting is observed when $\boldsymbol{B}$ lies on the surface of a cone around the principal axis $Z$ with the angle $\theta_0 = -\boldsymbol{B}$, $Z$=54.44°.

For nonzero $\eta$, the directions of the external magnetic field which leave an unsplit component at $\nu_\theta$ form an elliptical cone around the $Z$ axis. The angle $\theta_0$ is maximum when $\phi = 0°(\boldsymbol{B}\perp Y)$ and minimum when $\phi = 90°(\boldsymbol{B}\perp X)$. From an experimentally determined elliptical cone it is then easy to find the directions of the principal axes of the EFG tensor. The knowledge of $\theta_0(\phi = 0°)$ and $\theta_0(\phi = 90°)$ makes it possible to calculate the asymmetry parameter $\eta$:

$$\eta = \frac{3[\sin^2\ \theta_0(0°) - \sin^2\ \theta_0(90°)]}{\sin^2\ \theta_0(0°) + \sin^2\ \theta_0(90°)} \qquad [26]$$

Many of the features, of the $I = \frac{3}{2}$ case also appear in cases of $I = \frac{5}{2}, \frac{7}{2}$ and $\frac{9}{2}$. However, in these cases the asymmetry parameter $\eta$ is determined simply from the ratio of the NQR frequencies, whereas the external magnetic field is used to determine the number of magnetically nonequivalent nuclei in the unit cell and the orientation of the principal axes of the EFG tensor in a crystal-fixed coordinate system.

## List of symbols

$c^m$ = expansion coefficients of eigenstates of the quadrupolar Hamiltonian; $E_m$ = nuclear quadrupolar energy levels; $I$ = nuclear spin quantum number; $m$ = nuclear magnetic quantum number; $Q_{kl}$ = nuclear quadrupole moment tensor; $\eta$ = asymmetry parameter; $V(\boldsymbol{r})$ = electric potential; $V_{kl}$ = electric field gradient tensor; $\nu$ = NQR frequency; $\nu_L$ = Larmor frequency; $\rho(\boldsymbol{r})$ = nuclear electric charge distribution.

*See also:* **NMR Principles; NQR Applications; Nuclear Quadrupole Resonance, Instrumentation.**

## Further reading

Das TP and Hahn EL (1958) Nuclear quadrupole resonance spectroscopy. In F. Seitz and D. Turnbull (eds) *Solid State Physics* vol. 5. New York: Academic Press.

Grechishkin VS (1973) *Jadernie Kvadrupolnie Vzaimodejstvija v Tverdih Telah*. Moscow: Nauka.

Lucken EAC (1969) *Nuclear Quadrupole Coupling Constants*. New York: Academic Press.

Raghavan P (1989) Table of nuclear moments. *Atomic Data and Nuclear Data Tables* **42**: 189–291.

Schempp E and Bray PJ (1970) Nuclear quadrupole resonance spectroscopy. In Henderson D (ed) *Physical Chemistry*, pp 521–632. New York: Academic Press.

Semin GK Babushkina TA and Jakobson GG (1972) *Primenenie Jadernogo Kvadrupolnogo Rezonansa v Himii*. Leningrad: Himija. [English translation (1975): *Nuclear Quadrupole Resonance in Chemistry*. New York: Wiley.

Slichter CP (1963) *Principles of Magnetic Resonance*. New York: Harper & Row.

# Nucleic Acids and Nucleotides Studied Using Mass Spectrometry

**Tracey A Simmons**, **Kari B Green-Church** and **Patrick A Limbach**, Louisiana State University, Baton Rouge LA, USA

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Applications**

Mass spectrometry is a powerful tool for the characterization of biomolecules including nucleotides, oligonucleotides and nucleic acids. The advantages of mass spectrometry are high sensitivity, high mass accuracy and, more importantly, structural information. Historically, oligonucleotides and nucleic acids have proved difficult to characterize using mass spectrometry. Problems often arise with impure samples, low ion abundance for analysis owing to inefficient ionization processes and low mass accuracy for higher-molecular-mass compounds. Advances in the development of electrospray ionization and matrix-assisted laser desorption/ionization now permit the analysis of oligonucleotides and nucleic acids with high sensitivity and good mass accuracy. These improvements allow the use of mass spectrometry for the identification of unknown oligonucleotide structures and are suitable for applications focused on acquiring sequence information on the samples of interest.

## Nucleic acids and nucleotides

Nucleic acids are high-molecular-mass biopolymers composed of repeating units of nucleotide (nt) residues (**Figure 1**). The three major substituents of a nucleotide residue are the heterocyclic base, a sugar and a phosphate group. The five most common heterocyclic bases are adenine (Ade), cytosine (Cyt), guanine (Gua), thymine (Thy) and uracil (Ura). Cytosine, thymine and uracil are classified as pyrimidine bases, and adenine and guanine are classified as purine bases. Adenine, cytosine, guanine and thymine are the major bases found in deoxyribonucleic acids (DNA), and adenine, cytosine, guanine and uracil are the major bases found in ribonucleic acids (RNA). Nucleic acids are identified as either RNA or DNA, depending on the identity of the sugar. The sugar is a 2′-deoxy-D-ribose in DNA and a D-ribose in RNA. The phosphate group is usually attached through the 5′ or 3′ hydroxyl groups of the sugar.

Mononucleotides are typically represented by a shorthand notation that identifies the sugar, the base and the number of phosphate groups. For example, dNMP represents any 2′-deoxynucleoside monophosphate, dCDP represents 2′-deoxycytidine diphosphate and GTP represents guanosine triphosphate. Unfortunately, this convention is rarely followed in the mass spectrometry literature, where nucleosides and mononucleotides are denoted by their single letter base abbreviation and a preceding or following ‘p’ to represent the location of the phosphate group on the mononucleotides (for example, dC, dCp, pT and T for 2′-deoxycytidine, 2′-deoxycytidine 3′-monophosphate, thymidine 5′-monophosphate and thymidine respectively). This designation may be a source of confusion for those not familiar with this nomenclature and care should be used whenever such designations are used, especially when one wishes to distinguish nucleobases, nucleosides and nucleotides.

Oligonucleotides are most commonly joined together through the 3′ and 5′ sites of each nucleoside by a phosphodiester linkage. The primary sequence of an oligonucleotide is by definition determined from the 5′ to the 3′ end. Usually the sequence is written in shorthand notation, using the single letter abbreviations for the nucleobases; e.g., 5′-d(pTCAG)-3' as in **Figure 2**. In the case where the base composition is known, but the primary sequence is not, the unknown sequence regions are enclosed in parenthesis; e.g. 5′-dACT(GGCT)AAT-3′. Oligonucleotides of a specific length are typically referred to as ‘*n*-mers’, where *n* is the number of oligonucleotide residues.

Modified nucleosides are designated in the overall sequence by their chemical symbol (e.g., Am for 2′-O-methyladenosine). Unknown modified nucleosides often are designated by an ‘X’ in the overall sequence. Modified linkages are generally identified by describing the type of modified oligonucleotide followed by the usual sequence notation: a methylphosphonate 10-mer d(ACACGTTGAC) or a phosphorothioate 12-mer d(GCGCATATGCGC). Occasionally, phosphorothiates are identified by an ‘s’ internucleotide linkage, such as d(AsCsTsAsG).

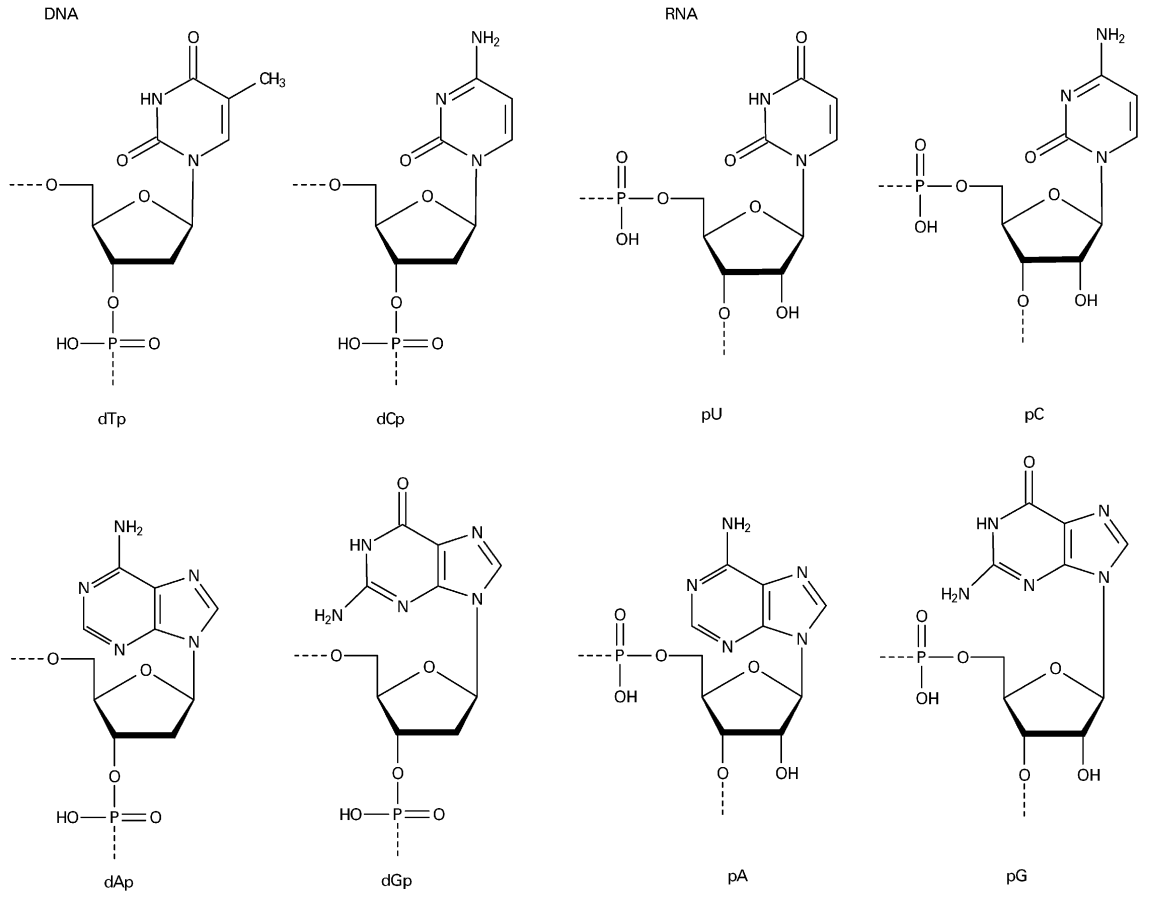



**Figure 1** Subunit structures of the major nucleotides from deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). T, C and U nucleobases are pyrimidine derivatives, and A and G nucleobases are purine derivatives.

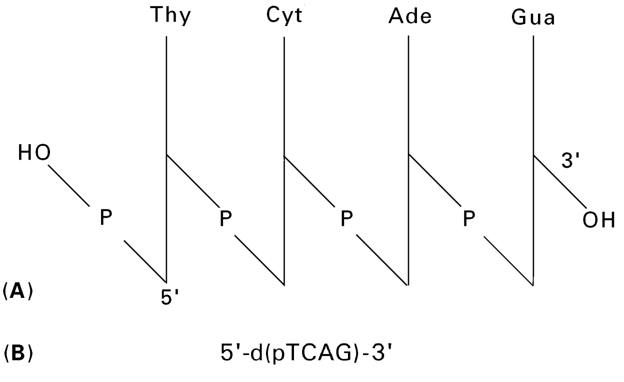



**Figure 2** Shorthand notations of oligonucleotides: (A) line notation of oligonucleotides, and (B) one-letter sequence notation.

## Ionization and mass analysis

Mass spectrometric measurements of oligonucleotides and nucleic acids involve the determination of molecular mass, or of the masses of dissociation products of gas-phase ions, which can then be related to various structural properties, such as sequence. In either case, the crucial step lies in the conversion of liquid- or solid-phase solutions of the analytes into gaseous ions. Common ionization sources used in mass spectrometry for nucleic acid and nucleotide experiments are fast-atom bombardment (FAB), electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). FAB was historically the ionization method of choice, but has largely been replaced by ESI or MALDI. ESI- and MALDI-based methods have particularly aided the analysis of nucleic acids because these techniques accomplish the otherwise experimentally difficult task of producing gas-phase ions from solution species that are both very highly solvated and highly ionic. Mass analysis of oligonucleotides and nucleic acids up to the 100-mer level has been demonstrated using ESI-MS. Molecular ions from significantly larger oligonucleotides (up to the 400-mer level) can be generated using MALDI-MS, but realistic mass analysis is limited to oligonucleotides and nucleic acids at the 125-mer level.

### Fast-atom bombardment

Oligonucleotide samples used in FAB must be soluble in a liquid matrix and are spotted onto a probe tip. Matrices selected are liquids with low vapour pressures and include glycerol, or an equal volume mixture of dithioerythritol and dithiothreitol (the 'magic bullet' matrix). Oligonucleotide samples are typically prepared as saturated solutions in water, and are spotted on the probe tip with neat matrices, prior to instrument insertion. FAB-MS of oligonucleotides is typically performed on a sector or quadrupole mass spectrometer and is limited to oligonucleotides that are no greater than 5000 Da in mass (**Figure 3**).

### Electrospray ionization

ESI requires sample introduction in liquid form and thus is convenient for the analysis of oligonucleotides. Samples are introduced into the instrument through a narrow capillary into a strong electrostatic field. Mixtures of an organic solvent and water are used in ESI-MS of oligonucleotides. Typical organic solvents are isopropanol, methanol or acetonitrile. Oligonucleotide sample concentrations for ESI experiments are prepared at the micromolar and nanomolar levels depending on the type of electrospray apparatus utilized. A mass spectrum of an oligonucleotide generated using ESI is characterized by a series of multiply charged negative ions that differ from one another by the removal of a single proton. The multiple charging effect from ESI-generated ions is advantageous for mass spectral analysis because mass analysers with limited upper mass ranges can be used (**Figure 4**).

### Matrix-assisted laser desorption/ionization

For MALDI-MS, oligonucleotide samples are mixed with a UV-absorbing matrix, spotted on a sample plate, and allowed to air dry. The common oligonucleotide MALDI matrices are 3-hydroxypicolinic acid, 2,4,6-trihydroxyacetophenone and 6-aza-2-thiothymine. Sample solutions for MALDI are at the micromolar or nanomolar level, with only a small fraction of the sample utilized in the MALDI experiment. A MALDI-generated mass spectrum of an oligonucleotide typically contains a singly negatively charged peak at the mass of the deprotonated ion. The analysis of MALDI-generated oligonucleotides is limited to either time-of-flight (TOF) mass spectrometers, which have an extended upper mass range but limited mass accuracy, or Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers, which have a limited upper mass range but high mass accuracy (**Figure 5**).

## Measurement of molecular mass

In ESI, both DNA and RNA are typically converted to gas-phase negative ions. Although the ionization efficiency is different for different homopolymers ($dT_{10} > dA_{10} \approx dC_{10} \approx dG_{10}$), both these and mixed bases are readily detectable. In MALDI, both DNA and RNA can be converted into gas-phase positive or negative ions. The quality of mass spectra from

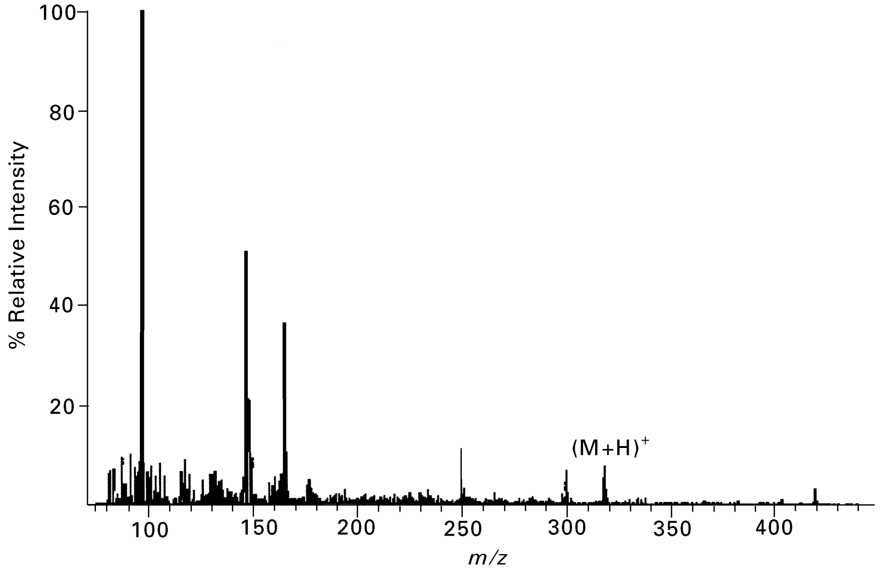



**Figure 3** Positive ion mode FAB mass spectrum using a double focusing sector mass analyser. The sample is dpC with piperidine in nitrobenzyl alcohol matrix. The parent ion mass, $(M + H)^+$, is 307.3 Da.

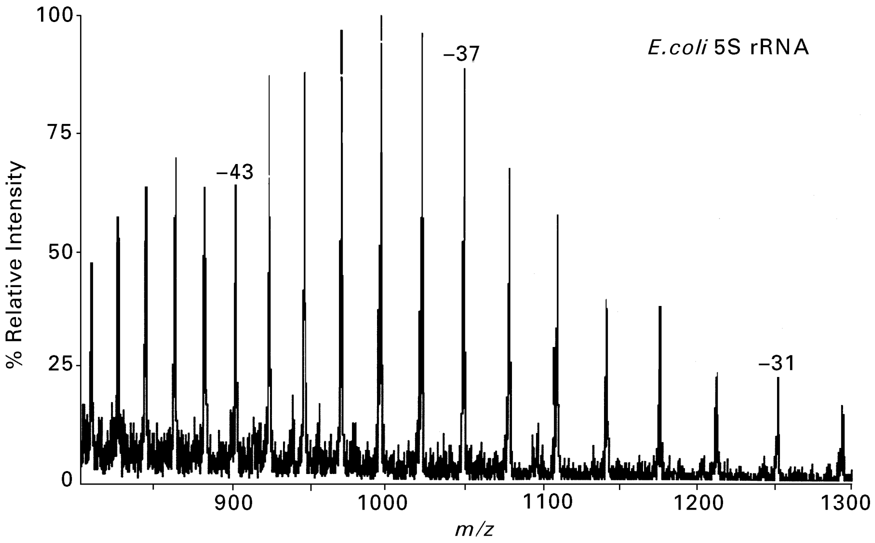



**Figure 4** Negative ion mode ESI spectrum of *Escherichia coli* 5S rRNA treated with *trans*-1,2-diaminocylohexane-*N*,*N*,*N′*,*N′*-tetraacetic acid (CDTA) and triethylamine (TEA) additives.

MALDI analyses is profoundly influenced by base composition. Thus, polythymidylic acids are readily analysed, but polydeoxyguanylic acids are a distinct challenge and mixed-base oligomers represent an intermediate situation. As a consequence of the increased stability of the glycosidic bond in RNA, it is more readily analysed than DNA.

Relative molecular mass is an intrinsic molecular property, which, when measured with high accuracy, becomes a unique and unusually effective parameter for characterization of synthetic or natural oligonucleotides. Mass spectrometry-based methods can be broadly applied not only to normal (phosphodiester) nucleotides but also to phosphorothioates, methylphosphonates and other derivatives. The level of mass accuracy will depend on the capabilities of the mass analyser used. Quadrupole and TOF instruments yield lower mass accuracies than sector or FT-ICR instruments. High mass accuracy is necessary not only for the qualitative analysis of nucleotides present in a sample but to provide unambiguous peak identification in a mass spectrum. For example, uridine monophosphate and cytidine monophosphate are separated by 1 Da (306.26 u, and 305.25 u, respectively) and oligonucleotides differing by the number of uridine and cytidine

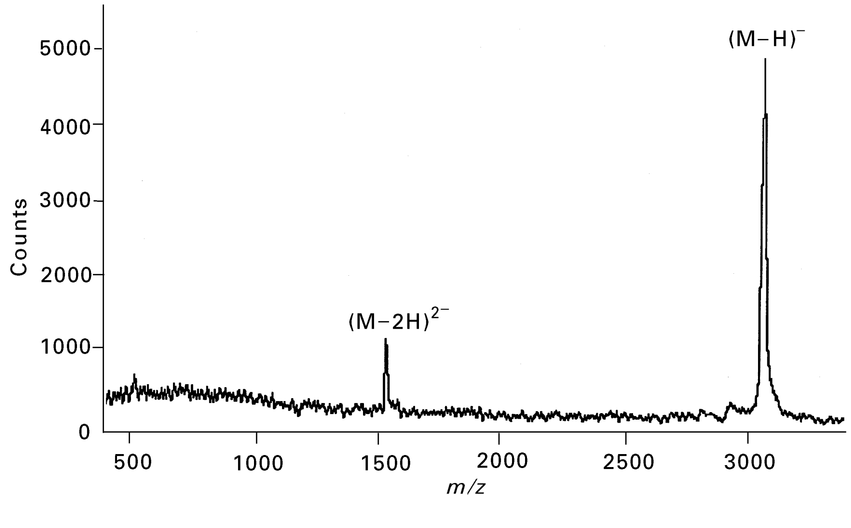



**Figure 5** MALDI-TOF mass spectrum, in negative ion mode, of $dA_{10}$ in 2,4,6-trihydroxyacetophenone matrix. The parent mass is seen as both singly and doubly charged ions.

nucleotides may be difficult to distinguish in a mixture. In such cases, the confidence of assigning a putative oligonucleotide base composition from the molecular mass measurement is dependent on the type of mass analyser used.

In ESI, mass measurement errors of 0.03% and lower are typical and, with appropriate sample clean-up, values lower than 0.01% can be obtained routinely. In contrast, for MALDI, typical measurement errors are between 0.03% and 0.05% and, with use of internal standards, mass measurement errors on the order of 0.01% can be obtained for small oligonucleotides.

The primary challenge to accurate molecular mass measurements of oligonucleotides and nucleic acids is reduction or complete removal of salt adducts. In solution, the phosphodiester backbone is completely ionized at pH > 1 and the solvent acts as a dielectric shield to reduce the repulsive Coulombic charging. During the ionization process, this Coulombic protection is lost, which eventually results in the adduction of any cations that may be present in solution. These cations (usually $Na^+$) reduce the polarity of this highly charged backbone, improving the production of gas-phase ions. It is these adducts that shift the centroid to higher values and limit the ability to determine molecular masses accurately.

Ammonium (or tetraethylammonium) are preferred salt forms for both ESI and MALDI. High-performance liquid chromotography (HPLC) purification is a suitable approach for sample purification, especially when ammonium buffers are used. Solvent additives, such as ethylenediaminetetraacetic acid and triethylamine, can also be added to the sample solution in ESI or MALDI to help alleviate the interference of alkali salt ions from the spectra (**Figures 4** and **5**). In MALDI, cation-exchange resin beads are effective at generating the ammonium salts of the oligonucleotides prior to mass analysis.

## Determination of nucleotide sequence by gas-phase sequencing

### General fragmentation processes

Dissociation of oligonucleotides can occur as a result of the excess energy that is imparted to the analyte during the desorption/ionization process. This dissociation is relatively fast (i.e. timescales shorter than the total mass spectral analysis time), resulting in ions that are generally difficult to identify accurately. Electrospray ionization produces stable, intact molecular ions except when the source conditions are adjusted to impart excess translational energy into the analyte, leading to its dissociation in the electrospray interface region (so-called nozzle–skimmer dissociation, discussed below). Thus, most dissociations that are desorption/ionization-induced are seen only in MALDI or FAB spectra. There are essentially four differing timescales for desorption/ionization-induced dissociations: prompt, fast, fast metastable and metastable.

- Prompt dissociations occur on a timescale equal to or less than the desorption event.
- Fast dissociations occur after the desorption event, but before or at the beginning of the acceleration event.
- Fast metastable decays occur on the timescale of the acceleration event.
- Metastable decays occur after the acceleration event, during the field-free flight time of the ion.

In theory, dissociations occurring during any one of these timescales will generate fragment ions that could be used to determine the sequence of the oligonucleotide. In practice, unless certain instrumental parameters are manipulated, most of these fragments result in a broadening of the molecular ion peak with a concomitant loss of resolution and sensitivity.

Owing to these drawbacks and because such a method for sequencing provides little control over the extent of fragmentation, there are few reports of using desorption/ionization-induced fragmentation to determine oligonucleotide sequences. Prompt fragment ions will be detected at their appropriate *m/z* values, using linear and reflectron time-of-flight mass spectrometers, with the reflectron geometry yielding higher resolution than the linear geometry, presumably owing to the interference of fast fragment ions in the latter case. Also, it is evident that formation of prompt fragment ions is matrix- and sequence-dependent, but is independent of the laser wavelength.

**Figure 6** illustrates the possible fragments one might see during the dissociation of an oligonucleotide. $B_1$, $B_2$ and $B_3$ represent nucleobases numbered consecutively from the 5′-terminus of the molecule. Fragment ions containing the 5′-terminus of the original molecule are labelled a, b, c and d, and differ by the site of bond cleavage along the oligonucleotide backbone. Fragment ions containing the 3′-terminus of the original molecule are labelled w, x, y and z, and, as the 5′-fragments, they differ by the site of bond cleavage along the oligonucleotide backbone. The sequence location of bond cleavage for the a–d and w–z fragment ions is denoted by a subscript reflecting the number of nucleobases remaining in the fragment ion (**Figure 7**).

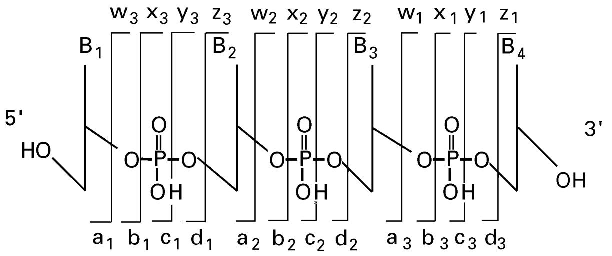



**Figure 6** Representative fragmentation pattern of oligonucleotides.

### Intentional dissociation methods

Intentional fragmentation of oligonucleotides may be caused by several methods including nozzle–skimmer dissociation (NS), photodissociation (PD), and collision-induced dissociation (CID). Intentional fragmentation can be made to occur within or near to the ionization source, or within a specified region of the mass analyser. NS dissociation occurs in ESI-MS when the translationally excited molecular ion collides with the background neutrals that are present in the high-pressure ESI source interface. NS dissociation generates fragment ions similar to those found in traditional CID spectra for small ($n \leq 15$) oligonucleotides. NS dissociation can also be used to obtain sequence information on larger ($n \geq 50$) oligonucleotides. However, complete sequence information on oligonucleotides of this size is not possible using this technique. Infrared multiphoton dissociation (IRMPD) has been used to characterize the sequence of small and large oligonucleotides. As with NS dissociation, complete sequence information can be obtained for small oligonucleotides, but this technique does not yield complete sequence information from larger oligonucleotides.

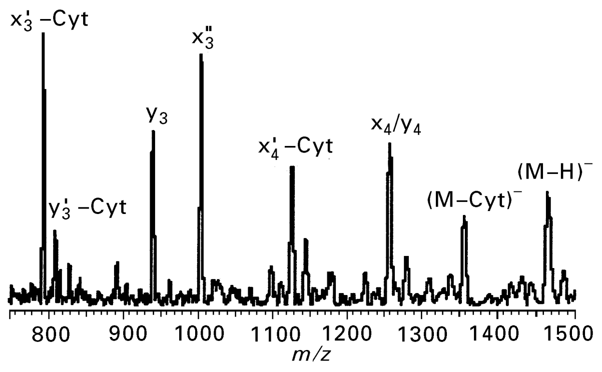



**Figure 7** MALDI mass spectrum of d(CATCG) with 2,4,6-trihydroxyacetophenone as the matrix. The laser-induced fragment ions are labelled corresponding to the fragmentation nomenclature shown in **Figure 6.**

### Tandem mass spectrometry (MS/MS)

An important goal in the mass spectrometric analysis of oligonucleotides has been to develop methods to determine structural information (such as sequence) through the dissociation of a molecule in the gas phase and interpretation of the resulting fragmentation pattern. The development of FAB-MS increased greatly the ability to perform such sequence analysis. However, it was not until the later development and application of primarily ESI-MS (and to a small extent, MALDI-MS) that meaningful tandem mass spectrometry studies were performed on oligonucleotides. In general, the fragmentation patterns of oligonucleotides are similar in spite of the different ionization techniques or mass analysers used.

Mass accuracy is essential for sequencing of unknowns using MS/MS experiments. If the oligonucleotide length is known, possible combinations of the four nucleotides may be calculated corresponding to the parent ion mass, and reasonable fragment ions can be assigned. If the length of the oligonucleotide is not known, working with high mass accuracy can help to reduce the possible sequence combinations calculated to match the parent ion mass. This technique is powerful because the specific fragment ions for a specific molecule can be identified and interference from other ions is eliminated. MS/MS also helps to reduce the problem of increased combination sequences as a result of increased sample size. When using mass analysers such as FT-ICR-MS or quadrupole ion traps, subsequent collisions and mass analysis ($MS^n$) may be conducted repeatedly, increasing the structural information obtained with each step.

## Determination of nucleotide sequence by solution-phase sequencing

### Mass ladders

The utility of any mass spectrometric sequencing method that relies on consecutive backbone cleavages depends on the formation of a mass ladder. The sequence information is obtained by determining the mass difference between successive peaks in the mass spectrum. In the case of oligodeoxynucleotides, the expected mass differences between successive peaks will correspond to the loss of dC = 289.25 u, dT = 304.26 u, dA = 313.27 u and dG = 329.27 u (all values are atomic mass-based). With oligoribonucleotides, the mass differences will be rC = 305.25 u, rU = 306.26 u, rA = 329.27 u and rG = 345.27 u (all values are atomic mass-based).

As is the case for all sequence determination methods that rely on the mass measurements of successive *n*-mers, oligodeoxynucleotides are easier to characterize owing to the relatively large differences in mass among the four oligodeoxynucleotide residues. The ribonucleotide residues, owing to the small mass difference between U- and C-containing nucleotides (a mass difference of only 1 Da), require a higher mass accuracy measurement to correctly distinguish U from C. Further, all mass ladder methods have a distinct advantage for sequence determination because it is the difference in two mass measurements that results in the desired information – the identity of the nucleotide residue. Because the accuracy of each individual oligonucleotide mass measurement is not critical, provided that the same bias exists for all ions detected, this procedure can be performed routinely on instruments with inherently lower mass accuracies; in particular, on TOF mass spectrometers.

### Failure sequence analysis

One of the simplest methods for characterizing the chain length and sequence of synthetic oligonucleotides is the detection of failure sequences from the original synthesis step. This procedure takes advantage of the fact that automated solid-phase synthesis of oligonucleotides, especially those that contain modified internucleotide linkages such as methylphosphonates or phosphorothioates, is not 100% efficient. For example, the solid-phase synthesis of phosphodiester-linked oligonucleotides generally produces a yield of 99–100% for each synthesis cycle. Synthesis of a 10-mer at a 99.5% yield per cycle would result in an overall yield of 95.6%, with the remaining 4.4% being failure products that terminated before completion of the synthesis. Synthesis of larger oligonucleotides, naturally, will result in lower overall yields. For example, the synthesis of a 50-mer oligonucleotide at the 99.5% yield per cycle level would result in an overall yield of only 78.2%; nearly one-quarter of the sample would be failure sequences.

One potential drawback to this method is the reduction in yield of the failure sequences of larger oligonucleotides. For example, assuming that the synthesis yield per cycle is 95%, the 5-mer failure product will be present in the final solution at a concentration four times greater than a 30-mer failure product. As the coupling yields increase, this difference becomes smaller. If the coupling yield increases to 99%, the 5-mer will be only 1.28 times as concentrated as the 30-mer. However, the total amount of failure sequence products will also decrease in this case. These factors place some upper limit on the size of the oligonucleotide that is amenable to sequencing and to the applicability of this approach for very efficient synthesis protocols.

Sequence determination of oligonucleotides by an analysis of the failure sequences is an extremely simple and straightforward method. The mass spectrum will contain a series of peaks that corresponds to the final product and to each one of the failure sequences, each of which differs in mass by the appropriate nucleotide residue value. The sequence of the oligonucleotide is determined in the 5′ to 3′ direction from the mass ladder of the synthesis failure products.

### Exonuclease digestion

The use of exonucleases to generate mass ladders of oligonucleotides that are suitable for analysis by mass spectrometry is now a standard method for sequencing small to moderate-length oligonucleotides. Unlike the failure sequence analysis method described above, this approach is suitable for the sequence analysis of naturally occurring oligonucleotides. Enzymes that hydrolyse the oligonucleotide consecutively from either the 5′- or 3′-terminus may be combined with unknown nucleotide samples and, over the reaction time, oligonucleotide ion signals are monitored by mass spectrometry. As the reaction progresses with the loss of one nucleotide at a time, the parent ion decreases in abundance and a 'mass ladder' is generated with each peak of lower mass corresponding to the sequential loss of a terminal nucleotide. Oligonucleotides up to the 40-mer level can be characterized readily by this approach.

### Chemical digestion

An alternative to the enzymatic approach is the use of chemical agents to cleave the original oligonucleotide into smaller components for mass analysis. Although the number of applications of chemical digestion methods for sequence determination of oligonucleotides is far smaller than the number of applications of the failure sequence method or enzymatic approach, there are several important cases where neither of the latter two approaches is suitable for oligonucleotide analysis. These cases generally occur when the oligonucleotide backbone has been modified, such as in the case of antisense oligonucleotides, where modification is chosen specifically to reduce the susceptibility of the oligonucleotide to enzymatic digestion. The principles of analysis for the majority of the chemical digestion protocols are identical to those in failure sequence analysis and enzymatic digestion in that the goal is to generate a mass ladder of peaks that differ from one another by a nucleotide residue. As such, all of the general

concerns for sequence determination by formation of mass ladders will apply to these methods.

We can classify chemical cleavages by their specificity for oligonucleotide base composition and/or linkage identity. Nonspecific chemical cleavages are reactions such as acid and/or base hydrolysis and alkylation reactions. In each of these reactions, the cleavage of the phosphodiester backbone does not depend on the base composition of the oligonucleotide. Acid hydrolysis is more specific for DNA, base hydrolysis is more specific for RNA and methylphosphonate-linked oligonucleotides, and alkylation is more specific for phosphorothioate-linked oligonucleotides. However, there is little to no base specificity for these chemical approaches.

The generation of a mass ladder of oligonucleotides for sequence determination using a chemical digestion approach can be complicated by the nonspecific nature of the chemical cleavage reaction. Any linkage site can potentially be cleaved by the chemical agent. If a single cleavage site is generated randomly for each oligonucleotide, then two sequence-specific fragments are produced, one from the 5′- and another from the 3′-terminus. If each of these fragments is detected in the resulting mass spectrum, then there is twice the necessary amount of information needed for the sequence to be determined. Although this additional information may facilitate sequence determination, especially if complete sequence coverage is unobtainable in any one direction, these two ion series can be a source of confusion.

*See also:* **Biochemical Applications of Mass Spectrometry; Fast Atom Bombardment Ionization in Mass Spectrometry; Fragmentation in Mass Spectrometry; Metastable Ions; Quadrupoles, Use of in Mass Spectrometry; Sector Mass Spectrometers; Time of Flight Mass Spectrometers.**

## Further reading

Chapman JR (1993) *Practical Organic Mass Spectrometry*, 2nd edn. Chichester: Wiley.
Dienes T, Pastor SJ, Schurch S, *et al* (1996) *Mass Spectrometry Reviews* **15**: 163–211.
Limbach PA (1996) *Mass Spectrometry Reviews* **15**: 297–336.
Limbach PA, Crain PF and McCloskey JA (1995) *Current Opinion in Biotechnology* **6**: 96–102.
McCloskey JA (ed) (1990) *Methods in Enzymology*, Vol 193. San Diego: Academic Press.
McNeal CJ (ed) (1986) *Mass Spectrometry in the Analysis of Large Molecules*. Chichester: Wiley.
Nordhoff E, Kirpekar F and Roepstorff P (1996) *Mass Spectrometry Reviews* **15**: 67–138.
Saengar W (1984) *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.

# Nucleic Acids Studied Using NMR

**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

## Introduction

The methods used to analyse and assign the NMR spectra of nucleotides and nucleic acids (DNA and RNA) are very similar to those for peptides and proteins. These are the subject of other articles in this Encyclopedia. Individual residues in DNA comprise deoxyribose sugar units connected at C-1′ to either cytosine (C) or thymine (T) pyrimidine bases, or to adenine (A) or guanine (G) purine bases. For RNA, the corresponding sugar is ribose and uracil (U) is found instead of thymine. The residues are connected *via* phosphate linkages between the 3′- and 5′-positions of the sugar rings.

Unlike proteins where use is generally made of $^{1}H$, $^{13}C$ and $^{15}N$ NMR data, the NMR spectroscopy of nucleotides and nucleic acids uses $^{1}H$, $^{13}C$ and $^{31}P$ NMR spectra. In protein NMR studies, it is usual nowadays to prepare the protein extensively labelled with both $^{13}C$ and $^{15}N$ and also sometimes with $^{2}H$. These techniques have not been applied so extensively in nucleic acid NMR spectroscopy.

Similar sample preparation methods are used for nucleic acids as for proteins. Nucleic acids are typically dissolved in 90% $H_2O$ – 10% $D_2O$ with the corresponding use of appropriate solvent resonance suppression methods for the measurement of $^{1}H$ NMR spectra.

## Assignment of NMR spectra

The usual approach used for proteins is also applied to NMR spectra of nucleic acids. 1D $^{1}$H NMR spectra can be supplemented with COSY or TOCSY spectra to aid spin system connectivities. Separate signals for the exchangeable amino and imino protons can often be observed and the properties of these signals can be very indicative of nucleotide structure. For example, imino protons involved in Watson–Crick base pairing such as A with T generally resonate in a distinctive window between δ13.0 and δ14.5. The $^{1}$H chemical shifts in duplex structures can be different to those for bases in single strands and thus the unwinding of a duplex can be monitored using these shift changes.

Because base protons and sugar protons are separated by a minimum of four bonds, spin couplings are not usually observed between these units and recourse is made to the use of NOE measurements often as 2D NOESY studies. Thus, for example, NOEs observed between the sugar anomeric proton and H-6 and H-8 of a base serve to identify the base and sugar units of a single residue. NOE measurement can also be used to gain information on the sequence of residues in a nucleic acid.

In addition, heteronuclear coupling between $^{1}$H and $^{31}$P can be used to make sequential connectivities between residues. This is possible because there is a continuous relay of a series of homonuclear and heteronuclear couplings along the nucleotide backbone.

Variable-temperature studies can be very informative as they give information on the melting and denaturation of duplex structures.

## Structural aspects of nucleic acids from NMR spectroscopy

It is possible to extract information on a wide variety of structural features using NMR spectroscopy of nucleic acids. These include characterization of base pairing, the conformation of the ribose and deoxyribose sugar rings, the torsion angles between sugar rings and the phosphate groups, the orientation of the sugar rings relative to the nucleotide bases, discrimination of right- and left-handed helices and the overall helix structure. In addition, information can be obtained on the location, dynamics and stoichiometry of binding of drugs to nucleic acids. Also, for duplex structures, it is possible to investigate mismatched pairings.

The overall aim is usually to determine the complete 3D structure of a nucleic acid and this is achieved in the same fashion as for protein structures. Thus, the NMR spectra are interpreted in terms of qualitative or semi-quantitative distance information which is then used as a set of constraints for a theoretical calculation of the structure generally based on distance geometry or restrained molecular dynamics calculations. One of the main problems with this approach, unlike for proteins, is the lack of long-range distance constraints for double-helix structures.

Despite these limitations, a combination of NMR spectroscopy and modelling has been used to derive structures for oligonucleotides, both DNA and RNA fragments, RNA–DNA hybrids and drug–DNA complexes. The review by Rizo and Bruch in the Further reading section provides references to a wide range of structural and dynamic studies of nucleotides and nucleic acids using NMR spectroscopy.

*See also:* **Carbohydrates Studied By NMR; Macromolecule–Ligand Interactions Studied By NMR; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Peptides and Proteins Studied Using Mass Spectrometry; Proteins Studied Using NMR Spectroscopy; Solvent Suppression Methods in NMR Spectroscopy; Two-Dimensional NMR, Methods.**

## Further reading

Davies DB and Veselkov AN (1996) Structural and thermodynamic analysis of molecular complexation by $^{1}$H NMR spectrocopy – intercalation of ethidium bromide with the isomeric deoxytetra-nucleoside triphosphates 5′-d (GpCpGpC) and 5′-d (CpGpCpG) in aqueous solution. *Journal of the Chemical Society, Faraday Transactions 2*, **92**: 3545–3557.

Pardi A, Walker R, Rapoport H, Wider G and Wuthrich K (1983) Sequential assignments for the $^{1}$H and $^{31}$P atoms in the backbone of oligonucleotides by two-dimensional nuclear magnetic resonance. *Journal of the American Chemical Society* **105**: 1652–1653.

Patel D, Pardi A and Itakura K (1982) DNA conformation, dynamics and interactions in solution. *Science* **216**: 581–590.

Rizo J and Bruch MD (1996) Structure of biological macromolecules. In: Bruch MD (ed.) *NMR Spectroscopy Techniques*, Chapter 6, p. 285. New York: Marcel Dekker.

Searle MS, Hall JG, Denny WA and Wakelin LPG (1988) NMR studies of the interaction of the antibiotic nogalamycin with the hexadeoxyribonucleotide duplex D(5′-GCATGC)$_2$. *Biochemistry* **27**: 4340–4349.

# O

# Optical Frequency Conversion

**Christos Flytzanis**, Laboratoire d'Optique Quantique du C.N.R.S., Palaiseau, France

Copyright © 1999 Academic Press

ELECTRONIC SPECTROSCOPY
Methods & Instrumentation

## General aspects

The introduction of intense coherent light sources such as lasers has drastically and irreversibly modified the field of optics in all its aspects, spectroscopy in particular. The emergence of nonlinear optics strikingly illustrates this revolution. In turn, the impact of nonlinear optics on the development of coherent light sources, since the very inception of lasers, has also been tremendous and to an extent that one can hardly now draw a demarcation line between them. Nonlinear optical processes, besides being at the very heart of laser action, are also increasingly integrated with laser systems to improve the spatiotemporal characteristics of the optical beams and pulses, to control their intensities and most importantly to provide access to frequency ranges that are still inaccessible by primary laser sources. This latter aspect, which concerns nonlinear optical processes where radiation is generated at frequencies that are different from those of the incident radiation and usually goes under the denomination of *optical frequency conversion*, has had wide and profound implications in the evolution of spectroscopy and remains central to research and development in nonlinear optics and laser physics. In this article we will sketch and summarize the main schemes of optical frequency conversion with some indications regarding their performances, achievements and tendencies.

Nonlinear optics concerns effects in the response of a material medium that are nonlinear in the optical field intensity. By response we refer to the radiation sources that are set up inside the medium because of the forced motion of bound and unbound charges (actually electrons). There are different ways to classify such effects, depending on the emphasis one wishes to put on a particular aspect of the nonlinear optical response and the validity of certain criteria. Thus as long as the field intensity is below the cohesive field intensity $E_c$ of a sample, which essentially keeps electrons bound to their atomic core (typically of the order $10^{10}\,\mathrm{V\,cm^{-1}}$) and concerns the reversible behaviour of bound electrons, the most common and straightforward classification is according to the powers of the electric field intensity $\boldsymbol{E}$ of the applied electromagnetic field. Within this context and referring to the Fourier decomposition of the incident field intensity one can also introduce the distinction between frequency-preserving and frequency non-preserving processes, the latter being the basis of the optical frequency conversion. Notwithstanding the complexity of the nonlinear optical effects the development here follows some rather well-classified patterns and goals. For higher intensities than those of the cohesive field $E_c$ the electrons can be forced out of the field of the atomic core and the classification of their response in powers of the electric field intensity loses its utility, as also does the distinction between frequency-preserving and frequency non-preserving effects. Here different approaches, in particular non-perturbative ones, are being pursued, some of them still tentative, as are also the experimental schemes that are being used; we expect in this regime of extreme nonlinear optics new and exciting developments, in particular regarding X-ray generation, as our understanding and description of the effects progresses.

Most of this article is devoted to summarizing the main achievements and state-of-art in optical frequency conversion within the regime where the perturbative approach is valid, namely $E < E_c$, and only at the very end we will take up the regime of extreme nonlinear optics, namely $E > E_c$. The progress

in the field of frequency conversion will continue as it is closely connected with progress in the performances of lasers, optical materials (linear and nonlinear ones) and improvements in the efficiency and understanding of the underlying nonlinear optical processes. Although some specific aspects may be modified in the near future, shifting the emphasis to new directions, the main patterns outlined here are expected to remain relevant.

## Nonlinear polarization sources – conceptual background

To keep track of the origin and interconnection of the nonlinear optical processes we recall that these are related to the different polarization terms in the power series expansion of the polarization $P$ induced in the medium by an incident electric field $E$ which in the dipolar approximation takes the form

$$\begin{aligned} P &= P^{(1)} + P^{(2)} + P^{(3)} + \dots \\ &= \chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \dots \end{aligned} \quad [1]$$

where the coefficient $\chi^{(n)}$ is the dipolar susceptibility of order $n$, a tensor of rank $(n+1)$, that measures the deformability of bound electron clouds to external fields as long as the latter do not exceed the cohesive field $\boldsymbol{E}_c$ intensity which is typically in the range of $10^{10}$–$10^{11}$ V cm$^{-1}$ for atoms and molecules. Introducing the Fourier analysis of these polarization terms and fields one can also write

$$P_i^{(1)}(\omega) = D_1\, \chi_{ij}^{(1)}(\omega)\, E_j(\omega) \quad [2]$$

$$P_i^{(2)}(\omega_1 + \omega_2) = D_2\, \chi_{ijk}^{(2)}(\omega_1, \omega_2)\, E_j(\omega_1)\, E_k(\omega_2) \quad [3]$$

$$\begin{aligned} P_i^{(3)}(\omega_1 + \omega_2 + \omega_3) &= D_3\, \chi_{ijkl}^{(3)}(\omega_1, \omega_2, \omega_3) \\ &\quad \times E_j(\omega_1)\, E_k(\omega_2)\, E_1(\omega_3) \end{aligned} \quad [4]$$

for the linear, second and third order terms, respectively, and similarly for the higher order ones. $D_n$ are degeneracy factors equal to the number of distinct permutations of the applied field Fourier components (modes) involved in the interaction; with this convention the value of $\chi^{(n)}$ for $\omega_i \to 0$ is independent of the chosen frequency path. Far from resonances in the transparency range of the materials the magnitude of $\chi^{(n)}$ is of the order of $1/E_c^{n-1}$ or $10^{-7}$ esu and $10^{-14}$ esu for $\chi^{(2)}$ and $\chi^{(3)}$ respectively.

The precise behaviour of the $\chi^{(n)}$, which is a tensor of rank $(n+1)$, and in particular their frequency dependence can be obtained by quantum-mechanical perturbation approaches. These coefficients possess some general symmetry properties related to the invariance laws such as causality, intrinsic symmetrization, time and space symmetries. The latter can lead to a drastic reduction of the number of independent components of the susceptibility tensor of a given order/rank. In particular, we recall that all even order dipolar susceptibilities $\chi^{(2n)}$ vanish for a medium with inversion symmetry, for instance for a gas, a liquid or an amorphous solid. Other considerations such as the Kleinman relations lead to some simplifications and reduction of the number of independent components in certain cases. Their quantum-mechanical expressions show that these coefficients can be enhanced by single and multiple resonances or can be greatly reduced because of destructive interference of such resonances. In **Figure 1** we indicate some schemes for frequency conversion. These coefficients are generally complex and both their magnitude and phase are relevant in assessing the energy transfer efficiency either among the electromagnetic field components involved in the nonlinear interaction process or between the electromagnetic field and matter. In the previous development, Equations [1]–[4], which is valid in the dipolar approximation we neglected nonlocal contributions that involve spatial or temporal derivatives of the electric fields or equivalently quadrupolar and magnetic dipolar contributions. These are in general, but no always, weaker than the electric dipolar ones and in certain cases introduce complications with boundary conditions; although interesting at a fundamental level they are not of much use for efficient frequency conversion and will not be considered here.

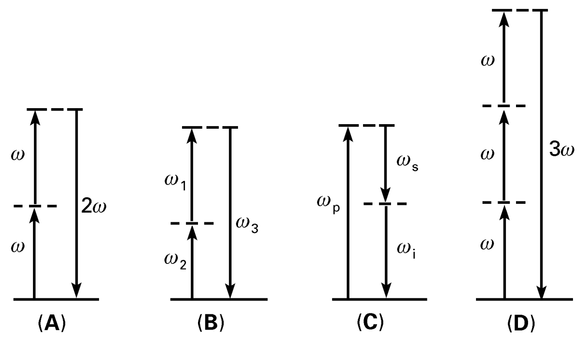



**Figure 1** Main nonlinear frequency conversion schemes. (A) Second harmonic, (B) sum frequency, (C) parametric emission and (D) third harmonic generation.

## Nonlinear propagation

The nonlinear polarization sources generate electromagnetic fields inside the medium that extract and transfer energy from the incident fields. These new fields can have markedly different spatiotemporal features from the incident ones. In particular, fields at new frequencies are generated such as harmonics of the incident frequencies or combinations thereof with conversion efficiencies close to one in certain cases. The assessment of the behaviour and efficiency of the frequency conversion process proceeds from the inhomogeneous propagation equation that contains the nonlinear polarization terms in Equation [1] as sources. Taking into account that the intense coherent light sources used in frequency conversion deliver pulses or wave packets whose envelope has a finite spatiotemporal extension one may set any of the nonlinear polarization terms in the form

$$P_{\mathrm{NL}}(\boldsymbol{r},t) = \mathrm{Re}\left\{\mathcal{P}_{\mathrm{NL}}(\boldsymbol{r},t)\ \mathrm{e}^{\mathrm{i}(\boldsymbol{k}_{\Sigma}\cdot\boldsymbol{r}-\omega t)}\right\} \qquad [5]$$

where $\mathcal{P}_{\mathrm{NL}}(\boldsymbol{r},t)$ is its envelope that varies slowly over the carrier period $T = 2\pi/\omega$ and $\boldsymbol{k}_{\Sigma}$ is the vector sum of the wave vectors of all field components that interact to set up the nonlinear polarization source (Eqn [5]). The electric field generated in the same frequency $\omega$ can be written in the form

$$\boldsymbol{E}(\boldsymbol{r},t) = \mathrm{Re}\left\{\hat{e}\,A(\boldsymbol{r},t)\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right\} \qquad [6]$$

where $\boldsymbol{k}$ is related to $\omega$ through the dispersion relation

$$k^2 = \omega^2(\hat{e}\,\varepsilon(\omega)\,\hat{e})/c^2 = \omega^2\,n_{\omega}^2/c^2 \qquad [7]$$

and in general $\boldsymbol{k} \neq \boldsymbol{k}_{\Sigma}$; $A(\boldsymbol{r},t)$ is the envelope of the electric field amplitude and varies slowly over the wavelength $\lambda = 2\pi/k$ and over the period $T = 2\pi/\omega$ and $\varepsilon(\omega) = 1 + 4\pi\chi^{(1)}(\omega)$.

With the slowly varying envelope approximation (SVEA) which is valid for almost all presently realistic cases one derives the equation

$$\frac{\partial A}{\partial z} + \frac{1}{2\mathrm{i}k}\Delta_{\perp}A - \frac{k''}{2\mathrm{i}}\frac{\partial^2 A}{\partial \tau^2} = \frac{\mathrm{i}k}{2\varepsilon(\omega)}\mathcal{P}_{\mathrm{NL}} \qquad [8]$$

with $\tau = t - z/v_{\mathrm{g}}$ where the group velocity is defined by $v_{\mathrm{g}} = -(\partial k/\partial\omega)^{-1}$, $k'' = (\partial^2 k/\partial\omega^2) = -(\partial v_{\mathrm{g}}/\partial\omega)/v_{\mathrm{g}}^2$ is its dispersion, $\Delta_{\perp} = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is connected with diffraction and $\Delta\boldsymbol{k} = \boldsymbol{k}_{\Sigma} - \boldsymbol{k}$ is the wave vector mismatch in the direction of propagation. For simplicity we have neglected absorption loss which can be taken into account by including an *ad hoc* term $\alpha A$ in the left-hand side of Equation [8] where $\alpha$ is the linear absorption coefficient. In general one is actually faced with a set of coupled equations which, however, can be simplified if one assumes that the incident (pump) beams are undepleted (parametric approximation). The solution can be obtained numerically but in several specific but useful cases one can obtain analytical solutions as well that elucidate several aspects of more general and complex configurations. The role played by the different terms in Equation [8] is evident (diffraction, group velocity dispersion, pulse broadening, etc.) and one can also make provisions to include the effect of anisotropy to account for the beam walk off and other related effects; the paraxial approximation is a good starting approach. The case of focused Gaussian beams has been extensively studied in harmonic generation as it has important implications on the phase matching configuration.

Several aspects regarding the efficiency of the frequency conversion processes can be grasped by restricting the analysis within the stationary and plane wave regime in an isotropic medium in which case Equation [8] reduces to

$$\frac{\partial A}{\partial z} = \mathrm{i}\frac{2\pi k}{\varepsilon(\omega)}\mathcal{P}_{\mathrm{NL}}\mathrm{e}^{\mathrm{i}\Delta kz} = \mathrm{i}\frac{4\pi}{n\lambda}\mathcal{P}_{\mathrm{NL}}\mathrm{e}^{\mathrm{i}\Delta kz} \qquad [9]$$

which can actually be used even for pulses in the range of few nanoseconds. Again linear absorption at the frequency $\omega$ has been neglected in Equation [9] but can be easily accounted for as previously indicated.

It is evident from Equation [9] that other factors being equal the efficiency of the conversion process increases with frequency and the strength of the nonlinear polarization source; we also notice in Equation [9] that the conversion starts with a phase jump of $\pi/2$ with respect to the incident field at the boundary. The most drastic impact on the conversion efficiency, however, stems from the presence of the factor $\exp(\mathrm{i}\Delta kz)$ in Equation [9]. To appreciate its impact let us assume that the incident 'pump' beam remains undepleted so that $\mathcal{P}_{\mathrm{NL}}$ can be assumed to remain unaffected during the conversion process. The

solution of Equation [9] is then simply

$$A = \frac{4\pi}{n_\omega \lambda} \mathscr{P}_{\mathrm{NL}} e^{i(\pi+\Delta kz)/2} \sin c(\Delta kz/2) \qquad [10]$$

or in terms of the transmitted intensity at $z = L$.

$$I = (cn_\omega/2\pi)|A|^2 \sim (\mathscr{P}_{\mathrm{NL}} L/n_\omega \lambda)^2 \sin c^2(\Delta kz/2) \qquad [11]$$

where $\sin cx = \sin x/x$ and we assumed that $A(0) \approx 0$ which is equivalent to neglecting the reflected beam; a more rigorous approach shows that the latter is reduced by a factor $(\Delta k/k)^2$ with respect to the transmitted beam. It is quite evident that $A(z)$ is an oscillating function of $z$ with a period $l_c = \pi/\Delta k$, the coherence length, with its amplitude scaling as $(1/\Delta k)$ as long as $\Delta k \neq 0$: the larger the wave vector mismatch $\Delta k$ the smaller the amplitude $A$ and the coherence length $l_c$ and the conversion efficiency is dramatically reduced. For high conversion rates one must either achieve $\Delta k = 0$ namely $\boldsymbol{k}_\Sigma = k$ along the propagation direction or artificially compensate for $\Delta k$. Both approaches are presently used for the most commonly exploited case of second harmonic generation but for more complex cases only the former is exploited.

Once the phase matching is achieved one can easily see from Equations [10] and [11] that $|A|$ will grow as $z$ (or equivalently $I \sim z^2$) at the initial stage of the conversion process but this cannot continue indefinitely because of the compound impact of pump depletion and back reaction of the created beam on the incident pump beams. This problem has been extensively studied in the case of second harmonic and sum frequency generation where analytical solutions in terms of elliptic integrals can be obtained. In principle, conversion close to 100% for the fundamental to the harmonic can be achieved over a propagation length of the order of a few interaction lengths $L_i = (n_1^2 n_2^2/8\pi^3\omega^2\chi_e^{(2)} I_p)^{1/2}$; more precisely for the simplest case of second harmonic generation in the phase matched configuration one finds

$$I_2 = I_1 \tan^2 h(L/L_i) \qquad [12]$$

where $I_1$ is the intensity of the fundamental. Similar considerations apply for the sum frequency generation, difference frequency generation and other second order parametric processes. These considerations for second order processes are also qualitatively corroborated by the Manley–Rowe conservation relations for the photon fluxes valid in the transparency range of any medium. The previous considerations valid for plane wave have also been numerically extended to Gaussian and Bessel beams.

## Wave vector mismatch

From the previous general considerations it is quite evident that among the different factors that affect the frequency conversion efficiency the wave vector mismatch $\Delta \boldsymbol{k} = \boldsymbol{k}_\Sigma - \boldsymbol{k}$ has by far the most dramatic effect and the achievement of phase matching, namely the cancellation of $\Delta \boldsymbol{k}$, is an indispensable condition for a nonlinear beam configuration to be considered for frequency conversion. Parenthetically we point out here that the phase mismatch is a manifestation of nonlocality in electromagnetic propagation; in the phase matched configuration this nonlocality is suppressed and the interaction effectively becomes local. Because of the key role played by the phase matching in all frequency conversion processes we shall outline and classify below the main phase matching schemes. These fall in either of the following two classes:

(i) compensation of the refractive index dispersion, for instance by exploiting the natural or artificial birefringence of the nonlinear optical material to compensate for the refractive index mismatch so that one can achieve $\boldsymbol{k}_\Sigma = \boldsymbol{k}$ or equivalently $\Delta \boldsymbol{k} = 0$.
(ii) artificial compensation of the wave vector mismatch by introducing an appropriate spatial (or eventually spatiotemporal) modulation of the nonlinear polarization source with a grating period $\Lambda = 2ml_c = 2m\pi/\Delta k$ where $m = 1,3,5$, etc.

### Dispersion compensation

This is the first experimentally implemented scheme for phase matching. It is the most widely used because it is flexible and can be extended to a wide class of nonlinear interactions. It relies on the fact that the two polarization eigenstates of an electromagnetic wave propagating in a given direction inside a birefringent medium experience different refractive indices (different phase velocities). One may then chose the propagation configuration (polarization states, crystal orientation, etc.) of the incident and created beams so that the birefringence compensates the dispersion. This also sets the main restriction on the use of this method the others being the crystalline symmetry, the transparency range, etc.

Because of these restrictions, at present only linearly birefringent (or optically anisotropic) media

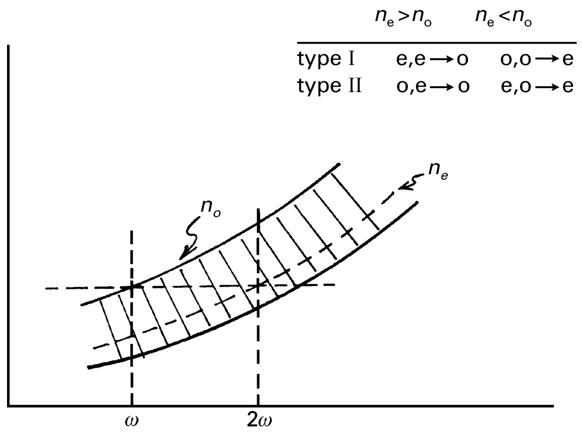



**Figure 2** Phase-matching in a birefringent crystal by dispersion compensation, showing the two types of phase-matching for second order processes.

are used for phase matching over restricted spectral regions. For second-order effects one distinguishes the type I and II phase matching configurations. We recall that in such media for each propagation direction there are two initially orthogonal linear polarization eigenstates with different refractive indices, the one direction that is angle independent (o, ordinary) and the other direction that is angle dependent (e, extraordinary); the principle of phase matching is depicted in **Figure 2**, where the two types are defined. Natural linear birefringence is the most widely used but in some cases externally induced linear birefringence can also be used for finer tuning (for instance electrooptic or Pockels effect).

Several well-characterized crystals for frequency doubling or frequency sum up to 0.25 μm which corresponds to the fourth harmonic of the YAG laser frequency (1.053 μm), are proposed by crystal manufacturers with outstanding conversion efficiencies. This upper frequency limitation around 0.25 μm results from the onset of absorption above this frequency and concomitant heating in practically all efficient doubling crystals; there are similar limitations towards the mid- and far-IR spectrum because of strong absorption bands there due to IR active lattice modes (phonons, vibrations) and one resorts here to other nonlinear processes and coupling schemes.

In principle, circular birefringence, natural (rotatory power or optical activity) or artificial (Faraday effect), can also be used but circular birefringence is in general very weak in compensating dispersion and no useful results have been obtained yet.

For frequency conversion above 0.25 μm the previous schemes cannot be used because as previously stated in all presently available nonlinear crystals absorption sets in above this frequency with concomitant heating of the crystal that can provoke beam instabilities and crystal damage. Here one proceeds instead with multiple (odd order) harmonic generation in gases, actually a two-component gas mixture, and exploits the anomalous dispersion to find coincidences of the refractive index for the fundamental and its multiple odd order harmonic; by controlling the relative concentration of the two gas components the dispersion can be 'tuned' to reach such a coincidence. The conversion efficiency is certainly lower than in crystals (<10%) and one needs long gas cell dimensions but in contrast to the previous cases one can now use very high beam intensities and in addition approach resonances to enhance the conversion efficiency; an additional advantage is the inherent isotropy of the medium. The beam focusing configuration plays a key role here and has been extensively studied for Gaussian beams and to some extent Bessel beams.

### Wave vector mismatch compensation

This is actually the first proposed phase matching scheme but its implementation in practice had to wait the development of appropriate artificial crystal growth techniques; it has emerged as the most efficient and robust technique for the development of compact devices that can be integrated with a laser, a situation that could not be envisaged with the previously described class of phase matching schemes. This technique, which nowadays goes under the denomination of quasi-phase-matched (QPM), has now been applied for second harmonic generation but its extension in other situations involving second-order effects (frequency sum/difference) can also be considered.

In its simplest version the QPM technique for second harmonic generation (SHG) consists in alternatively reversing the polar axis direction in successive thin layers of the same thickness $d$ of a crystalline material. This leads to a reversal of sign of $\chi^{(2)}$ in successive layers of the layered material and one can easily see that if $d = ml_c = m\pi/\Delta k$ then the conversion efficiency increases with the number of layers. This actually amounts to periodically modulating $\chi^{(2)}$ in the nonlinear optical source $\mathcal{P}^{(2)} = \chi^{(2)}\boldsymbol{EE}$ in Equation [9] with a wave vector that precisely matches and cancels $\Delta k$ in the propagation direction; note that the linear refractive index is unaffected and remains the same as in the bulk crystal while $I_2$ grows as $z^2$ (**Figure 3**). In a certain sense the quasiphase-matched systems can also be viewed as a nonlinear photonic crystal.

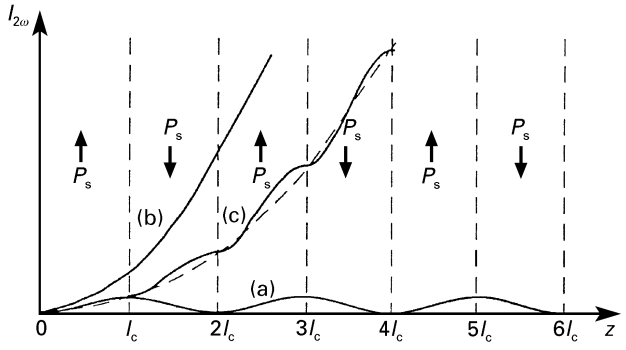



**Figure 3** Wave vector compensation by quasi-phase matching (a) non-phase matched, (b) quasi-phase matched and (c) phase matched.

This technique has already been applied with outstanding results using certain ferroelectrics (domain reversal either by poling or doping) and holds much promise for the development of compact integrated lasing and doubling devices. It can also be used with poled polymers but their transparency range is not as extended as in the ferroelectrics. Attempts are being made to use this technique with cubic heteropolar semiconductors, ex:GaAs, but the use of molecular beam epitaxy techniques here to make the layered structure is costly and time consuming because of the required layer thickness. The technique can also be extended to cascading of SHG processes for higher frequency conversion within the transparency range of the material.

## Nonlinear optical materials

Along with the phase matching problem the choice of the nonlinear material for frequency conversion (essentially SHG) has a major impact on the efficiency and robustness of the process. To the extent that cascading second-order effects can effectively and practically cover all cases of third-order processes as well, attention nowadays is being concentrated on non-centrosymmetric materials, inorganic or organic, with particular emphasis on oxygen polyhedra based inorganics such as $LiNbO_3$, $BaBO_4$, (BBO), $LiIO_3$, KDP, ADP, KTP, $KNbO_3$ and derivatives by substitution or specific semiconductors such as $AgGaGe_2$, $ZnGeP_2$ etc., certain organic crystals (urea, MAP, POM, MNA, NPP) or poled polymers. In view of the widespread uses of such materials not only in frequency conversion but also in light modulation and other optoelectronic applications in devices, and the need of their artificial growth, many factors must be taken into consideration when making the choice of a particular material, their nonlinear properties per se not being the sole criterium. Aspects such as optical anisotropy and transparency range, multiphoton losses, threshold of optical breakdown, photochemical stability and mechanical and thermal properties are important in the choice of the appropriate material, as are the material growth technique, doping and ageing considerations, processability, interfacing and packaging and other aspects since for a wide class of materials the nonlinear figures of merit are of comparable magnitude. **Table 1** gives indicative values of the nonlinear optical coefficients of some crystals.

## Cavity enhanced frequency conversion

The efficiency of frequency conversion processes can be substantially enhanced and the characteristics of the generated beams drastically improved by an appropriate choice of the geometry of the interaction configuration. In this context the insertion of the frequency conversion process in a cavity or guide can have a major impact. Different schemes have been used for second-order processes, such as intracavity frequency doubling, external cavity frequency doubling, quasi-phase-matched doubling and optical parametric oscillation. In all these cases one essentially exploits the fact that the power inside a cavity can be much higher than outside and the effective interaction length very long and this enhances the efficiency of the conversion process; together with other factors these substantially improve the generated beam quality. These schemes are briefly described below.

### Intracavity SHG

The doubler crystal is inserted into the laser cavity with the output coupler replaced by a mirror that is highly reflective/transmittive in the fundamental/harmonic respectively. With appropriate choice of the transmittivity and coupling parameter the fundamental can be completely converted into the harmonic. High quality nonlinear optical crystals are required here and specific optical arrangements must be inserted to eliminate spatial hole-burning, amplitude fluctuations, etc. Clearly the ultimate goal here could be to compound the lasing gain and doubling processes in one and the same crystal, which requires appropriately doped nonlinear optical crystals such as rare earth doped ferroelectrics or dye doped nonlinear organic crystals or poled polymers. Some interesting results have already been achieved in this direction.

**Table 1** Second-order nonlinear coefficients of some crystals

| | | | | $d_{14}$ | $d_{33}$ | | | $I_{ob}$ | *Transm.* |
|---|---|---|---|---|---|---|---|---|---|
| *Crystal* | *Symmetry* | $n_0$ | $d_{21}$ | ($10^{-12}$ m $V^{-1}$)[a] | | $d_{eff}$ | 6 | (GW $cm^{-2}$)[b] | (μ)[c] |
| $LiNbO_3$ | 3m | 2.232 | –2.1 | 0 | –27 | 5.1 | 70 | 10 | 0.35–5 |
| $BaBO_4$ | 3m | 1.655 | –2.3 | 0 | 0 | 1.9 | 16 | 14 | 0.2–2.6 |
| $LiIO_3$ | 6m | 1.857 | 0 | 0 | 4.5 | 1.8 | 13 | 2 | 0.34–4 |
| KDP | $\bar{4}2m$ | 1.493 | 0 | 0.37 | 0 | 0.35 | 1 | 5 | 0.18–1.8 |
| ADP | $\bar{4}2m$ | 1.509 | 0 | 0.47 | 0 | 0.39 | 1.2 | 6 | 0.18–1.5 |
| $AgGaGe_2$ | $\bar{4}2m$ | 2.594 | 0 | 33 | 0 | 28 | 81 | 0.3 | 0.78–18 |
| $ZnGeP_2$ | $\bar{4}2m$ | 3.073 | 0 | 69 | 0 | 70 | 292 | 0.05 | 0.74–12 |
| KTP | mm2 | 1.737 | 0 | 0 | 8.3 | 3.2 | 47 | 15 | 0.35–4.5 |
| $KNbO_3$ | m2 | 2.119 | 0 | 0 | –19.5 | –11 | 312 | 7 | 0.4–5.5 |
| | | | | | $d_{11}$ | | | | |
| Urea | | | | | 12 | | | 5 | 0.2–1.4 |
| MAP | | 1.508 | 16.8 | | | | | 3 | 0.5–2.5 |
| POM | | 1.663 | | 9.2 | | | | 2 | 0.5–1.7 |
| MNA | | 2.0 | | | 168 | | 1000 | | 0.48–2.0 |
| NPP | | | | | 85 | | | 0.05 | 0.48–2.0 |

[a] Divide each value by $4.2 \times 10^{-4}$ to convert into esu.
[b] Optical breakdown.
[c] Transmission range.

### External cavity SHG

For low gain lasers there is an advantage in using an external optical cavity configuration instead of the intracavity one as this also permits one to independently optimize the laser and doubling cavities. The main problems in this scheme are the requirements of frequency matching between the laser frequency and the doubling cavity resonance on one hand, and that of 'impedance matching' on the other. Excellent performances have been achieved here with monolithic cavities fabricated from a single piece of doubler crystal with mirrors directly coated on the crystal surfaces (other features have also been integrated to improve the efficiency and quality of the process). Quasi-phase-matched configurations along the lines previously discussed have also been considered here.

### Optical parametric oscillators (OPOs)

This is an alternative solution to a tunable laser source in the optical frequency range. Here a powerful single frequency radiation (pump) is parametrically converted into equally powerful coherent radiation tunable over a wide optical frequency range (signal and idler). In its simplest version, which is also the most useful one, it is based on the optical parametric amplification of 'noise' photons of frequencies $\omega_1$ and $\omega_2$ provided by the dissociation of a pump photon of frequency $\omega_p$ such that $\omega_p = \omega_1 + \omega_2$, the selection of a particular frequency pair being made through the phase-matching condition $k_p = k_1 + k_2$. Thus any one of the previous phase-matching schemes, combined with additional features regarding the cavity configuration, mirrors etc., can be used to select a particular frequency for the output radiation (or its complementary) and also to improve the output beam quality. The present surge in the development of OPOs stems from the improvements in pump beam quality and nonlinear crystal performances.

Optical parametric oscillators with continuous wave (CW) or pulsed regime operation are presently available with outstanding performances for the most demanding spectroscopic studies; particularly impressive has been the development of OPOs operating on the few-femtoseconds pulse regime.

## Stimulated scattering processes – frequency shifters

The previous nonlinear optical processes for frequency conversion are coherent and essentially instantaneous as they do not involve any dynamics of material excitations and all energy transfer and storage only occur within the electromagnetic field modes; material resonances eventually serve to enhance the efficiency of the process.

There is a whole class of nonlinear optical processes involving real multiphoton transitions inside the medium where the incident beams are converted into beams of different frequencies at the same time as the material system undergoes a transition between two

different energy levels and here the dynamics (relaxation) of this transition play a key role. Several such schemes can be conceived and exploited for frequency conversion, and they can be classified in two categories: stimulated scattering processes and up-conversion lasing processes. **Figure 4** shows some characteristic configurations. A characteristic feature of all these processes is that they exhibit an exponential growth with the spatial extension of the process as in the case of lasing processes and in contrast on the previously discussed frequency conversion schemes that exhibit a power law dependence on the interaction distance. In fact in all these schemes the generated beams are built up on spontaneous processes and can be viewed as generalized lasing processes that exploit multiphoton pumping schemes.

### Stimulated Raman frequency shifter

This is the simplest and most used scheme of stimulated scattering process in which an incident wave at frequency $\omega$ is converted into a scattered wave at frequency $\omega_s$, the difference in photon energy $\hbar(\omega_i - \omega_s) = \pm\hbar\omega_R$, being taken up (Stokes component) or supplied (anti-Stokes component) by the nonlinear medium, which undergoes a transition between two energy levels $E_a$ and $E_b$ such that $\omega_R = |E_a - E_b|/\hbar$. Various types of such transitions can be involved, such as electronic ones, vibrations, rotations etc., namely the same ones that give rise to spontaneous light scattering, the extreme case being that of Brillouin and Rayleigh scattering with negligible or no frequency shift. Actually the spontaneous process provides the 'seeding photons' from which the stimulated process builds up. The amplitude of the generated beam and the gain in the stationary regime, can be derived from Equation [9] after insertion of the appropriate nonlinear polarization source, $\mathcal{P}^{(3)} = \chi_R^{(3)} E_L E_L^* E_s$ where $\chi_R^{(3)}$ is the $|\omega_i - \omega_s - \omega_R|$ – resonant part of $\chi^{(3)}(\omega_L, -\omega_L, \omega_s)$ or $\chi_R^{(3)} = i\chi''$ and $\chi''$ is positive/negative for the Stokes/anti-Stokes component. The process is in fact automatically adjusted to satisfy the phase matching condition. Clearly the Stokes component has a gain and grows exponentially with the interaction distance as in the case of lasing action.

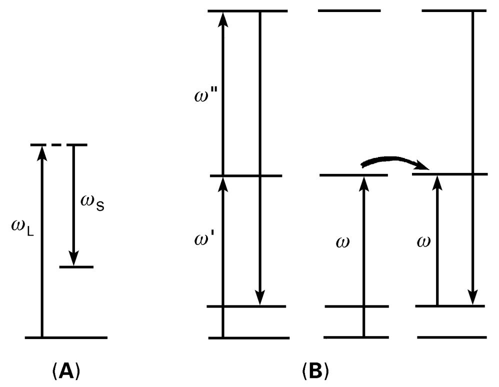



**Figure 4** Main schemes of nonlinear stimulated processes for frequency conversion. (A) Stimulated Raman and (B) up-conversion lasing schemes.

In the stimulated Raman Stokes scattering process the frequency shift $\omega_R$ is characteristics of the material medium, and by changing the medium new coherent sources with different down-shifted frequencies can be obtained. If the conversion to the Stokes wave is appreciable, generation of multiple Stokes waves $\omega - \omega_R$ can take place by a cascading process, the first one acting as the pump for the next one and so on. Although the Raman jumps are discrete, by using an initially tunable source one can have a tunable output as well one shifted en bloc by the Raman frequency. Note that because the process involves a material transition (excitation) its dynamics are conditioned by that transition. The previous considerations essentially concern the stationary regime where the pulse length is larger than the characteristic times of the material resonance, otherwise the transient regime must be used with much lower conversion efficiencies and other complications. This restricts the use of stimulated Raman frequency conversion technique to a certain pulse length regime.

Although the direct anti-Stokes process shows no gain from the outset, the generated Stokes field can, by interacting with the pump field, generate the anti-Stokes field through a phase matched four-wave mixing process. Here again one can have multiple anti-Stokes coherent generation through cascading processes. Because the phase-matching condition must be satisfied at each step the anti-Stokes output is generated in the form of cones around the direction of the Stokes output, which coincides with the direction of maximum gain or equivalently that of largest interaction distance between pump and Stokes. Actually the build up of the anti-Stokes field at the expense of the Stokes via a four-wave interaction process that requires the phase-matching to be satisfied is the simplest case of an 'optical balance'. In an 'optical balance' either configuration of two nonlinear optical processes sharing the same multiphoton resonance can be enhanced at the expense of the other by exploiting a phase matching condition: this is precisely the case of the Stokes/anti-Stokes generation.

Although the stimulated Raman conversion techniques have been superseded by others they are still useful in many instances, in particular because of their easy and economical implementation. One typically reaches 20% conversion efficiencies, or

even higher, but then the pump depletion must be taken into account.

### Upconversion lasers

This is the generic denomination of a large class of stimulated lasing emission schemes between two widely separated energy levels $E_a$ and $E_b$, with $(E_b - E_a)/\hbar$ in the UV or higher ($E_b > E_a$), where population inversion is reached by appropriate real multiple step transitions involving several photons of energy $\hbar\omega_i$ lower than $E_b - E_a$, usually provided by a powerful infrared pump. The key point here is that intermediate metastable states are strongly populated and used as a reservoir for the pump energy. By appropriate choice of the metastable states with regard to the transition oscillator strength and relaxation times a population inversion can be achieved between two states with widely spaced energies $E_a$ and $E_b$ or $E_b - E_a > \hbar\omega_i$ and subsequent lasing emission between them occurs when inserted in a cavity.

To the extent that the pump photon energy is stored at the intermediate metastable levels, which are then strongly populated, phase matching considerations are in general irrelevant in these schemes as in all lasing schemes, although in some cases 'optical balance' configurations may take place and population build-up may be suppressed if a phase matching condition is reached that results in a suppression of the population inversion.

There is a wide range of processes that can be exploited for up-conversion lasing, some of them involve transitions and levels within one atom or molecule but most frequently these schemes evolve through cooperative processes that involve two atoms or molecules. Many of these schemes in fact rely on the cooperative transitions originally predicted for impurity atom (or ion) pairs in crystals and subsequently extended in gases in the so-called laser-assisted van der Waals collisions and other similar schemes.

These processes in doped crystals, glasses, fibres or gases have been used to provide up-conversion schemes where IR laser radiation is converted into UV radiation with good efficiency. The main drawback of several of these schemes is the need of low temperature instrumentation because of the crucial role played by the line widths, relaxation and collision processes.

## Terahertz radiation

Much effort is presently concentrated to extend the nonlinear frequency conversion techniques to cover the far-IR or terahertz region, roughly the spectral region 0.1–10 THz, where along with the XUV and X-ray regions to be discussed below there is a lack of coherent sources. In the past attention was concentrated on the frequency difference generation in noncentrosymmetric crystals the extreme case being the optical rectification effect. In several other cases a frequency down-conversion through a succession of stimulated Raman processes such as spin-flip or vibrational–rotational in cascade in various gases has been exploited but the efficiency drastically drops with the generated frequency unless one compensates this reduction with intermediate resonant enhancement. The best results were obtained with a combination of stimulated vibrational and rotational scattering in molecular gases which, however, cannot lead to compact far-IR sources.

Recent efforts to generate THz radiation have been essentially based on the optical rectification effect in ferroelectrics or the generation of transient photogenerated currents in semiconductors, both with femtosecond pulses; some variants of these approaches, also termed terahertz time-domain spectroscopy (THz-TDS), have produced very promising results and open the way to the development of compact coherent sources that emit pulses in the THz-region with durations close to the single terahertz optical cycle. Many technical problems, however, still remain before these schemes can be implemented in reliable devices for tunable THz-radiation.

## Extreme nonlinear optics – X-ray generation

In the previously discussed frequency conversion schemes the laser field amplitude $E_L$ is well below the cohesive field intensity $E_c$ so that the perturbative approach, as exemplified by the power series development Equation [1], can be used to single out the appropriate nonlinear optical process for frequency conversion. In particular the order of the nonlinear process that contributes is well identified. If $E_L$ exceeds $E_c$, which is typically of the order of $10^{10}$ V cm$^{-1}$, one may expect photoionization to take place, liberating electrons from the neutral atoms with concomitant depletion of the frequency conversion and irreversible damage to the material through optical breakdown. With presently available commercial amplified fs-lasers the intensities can easily be in the range of $10^{17}$ W cm$^{-2}$ and with appropriate arrangements the pulse duration can be as short as the light cycle, typically 2–3 fs in the visible. The photo-ionization is initiated by the suppression of the Coulomb potential through the strong electric laser field, allowing the electrons to tunnel out of the atomic core within the oscillation period of the laser

field. The full ionization of the atom can, however, be avoided if the intense light pulse is short enough so that the electrons are not fully liberated from the atomic core. Rather, the quasi-free electron is accelerated in the laser field and subsequently may re-encounter and recombine with the parent ion, emitting a photon, or collide with the surrounding atoms, causing electron ejection and photon emission through the inverse bremsstrahlung process. The two processes are laser polarization state dependent (linear and circular respectively). Both processes can lead to compact coherent X-ray sources although still many technical problems are to be surmounted before such sources become reliable for spectroscopic studies or other applications.

In this regime of extremely intense optical fields where the effective interaction time is close to the single light period and irreversible damage is avoided, the power series expansion [1] breaks down as the different terms acquire comparable magnitude. Non-perturbative approaches have been developed to treat this regime, predicting the generation of X-ray pulses from gas targets illuminated with intense ultrashort laser pulses. Several pulse and beam characteristics have a key impact on the efficiency of the X-ray generation process. Phase matching plays an important role in the process as also does the pulse duration and pulse phase. Because of the symmetry of the interaction, odd order harmonics are present in the emitted spectrum with a strength that drastically decreases above a given order.

The results obtained with rare gases and also some metallic surfaces are very promising and open the way to a generation of compact coherent XUV and X-ray radiation with easily accessible femtosecond pump sources.

## Future prospects

The progress in frequency conversion techniques has been tremendous since the first observation of second harmonic generation in the early 1960s. This article gives only a rough picture of this progress which still continues. At present one can state without exaggeration that starting from a high quality primary laser that is tunable in a narrow region of the optical range and using the previously discussed optical frequency conversion schemes one can cover the whole spectral range from 10–0.1 μm with outstanding characteristics regarding the generated radiation, such as frequency tunability and stability, beam intensity and quality, pulse duration and shape. These schemes are now being implemented in compact and robust devices that greatly facilitate and widen their uses in different areas and in spectroscopy in particular. The introduction of all-solid-state pump sources will greatly facilitate this trend.

This trend will continue unabated as the demand for tunable and versatile coherent light sources is continuously growing in many areas. In the near future important breakthroughs are expected in the two extreme frequency ranges, i.e. the far-IR and THz spectral ranges and the XUV and X-ray ranges. The expected applications in these frequency ranges are very important as are also the fundamental studies that such sources will generate.

## List of symbols

$A$ = amplitude, $d$ = thickness of crystalline material, $D_n$ = degeneracy factors, $E_a$, $E_b$ = energy levels, $E_c$ = cohesive electric field intensity, $E_L$ = laser field amplitude, $I_1$ = intensity of the fundamental, $\mathbf{k}_\Sigma$ = vector sum of the wave vectors, $l_c$ = coherence length, $L_i$ = interaction length, $P$ = polarization induced by an incident electric field, $T$ = carrier period, $v_g$ = group velocity, $\alpha$ = linear absorption coefficient, $\chi^{(n)}$ = dipolar susceptibility of order $n$, $\Lambda$ = grating period, $\omega$ = frequency, $\omega_s$ = scattered wave frequency, $\hbar$ = Planck's constant/2π.

*See also:* **Laser Applications in Electronic Spectroscopy; Laser Magnetic Resonance; Laser Microprobe Mass Spectrometers; Laser Spectroscopy Theory; Nonlinear Optical Properties; Nonlinear Raman Spectroscopy, Theory; X-Ray Spectroscopy, Theory.**

## Further reading


Armstrong JA, Bloembergen N, Ducuing J and Pershan P (1962) *Physics Review* **127**: 1918.

Auston DA, Glass AM and Ballman AA (1972) *Physics Review Letters* **28**: 897.

Bordui F and Fejer MM (1993) *Annual Reviews Material Science* **23**: 321.

Digiorgio V and Flytzanis C (eds) (1994) *Nonlinear Optical Materials. Principles and Applications.* Amsterdam: North-Holland.

Franken PA, Hill AE, Peters CW and Weinreich G (1961) *Physics Review Letters* **7**: 118.

Giordemaine J (1962) *Physics Review Letters* **8**: 19.

Grischkowsky D, Keiding S, van Exter M and Fattinger C (1990) *Journal of the Optical Society of America* **B7**: 2006.

Krausz F, Brabec T, Schnürer M and Spielman C (1998) *Optics and Photonics News*, July.

Maker PD, Terhune RW, Nisenoff M and Savage CM (1962) *Physics Review Letters* **8**: 21.

Miles RB and Harris SE (1971) *Applied Physics Letters* **19**: 385.

Mittleman D (1998) *Laser Focus World*, May issue, p 191.

Rabin H and Tang CL (eds) *Quantum Electronics: A Treatise*, Volumes 1 and 2. New York: Academic Press.

Reintjes J (1984) *Nonlinear Optical Parametric Processes in Liquids and Gases*. New York: Academic Press.
Shen YR (ed) (1977) *Nonlinear Infrared Generation*. Berlin: Springer Verlag.
Shen YR (1984) *Principles of Nonlinear Optics*, p 46. New York: Wiley.
Simon U and Tittel FK (1997) In *Experimental Methods in the Physical Sciences*, Vol 29C.
Yablonovitch E, Flytzanis C and Bloembergen N (1972) *Physics Review Letters* **29**: 865.
Yariv A (1990) *Quantum Electronics*, Ch 9. New York: Wiley, New York: Academic Press.

# Optical Spectroscopy, Linear Polarization Theory

**Josef Michl**, University of Colorado, Boulder, CO, USA

Copyright © 1999 Academic Press

ELECTRONIC SPECTROSCOPY
Theory

The degree of linear polarization of light arriving at a detector from a sample contains information on (i) the anisotropy of molecular optical properties and (ii) the molecular orientation distribution. In the following, we outline a general theoretical framework for extracting this information, equally useful for electronic and vibrational spectroscopy. We apply it to processes that involve the interaction of one photon or two photons (successively, as in photoluminescence and photoinduced dichroism; or simultaneously as in two-photon absorption and Raman scattering) with an isotropic or partially aligned solution of molecules that interact with light one at a time. We emphasize uniaxial assemblies (e.g. solutes in an electric field, in nematic liquid crystals, in stretched polymers, or in membranes, and solutes 'photoselected' with linearly polarized light).

## Probabilities of optical events

Amplitudes of molecular optical events are proportional to off-diagonal matrix elements of interaction operators between the wavefunctions of the initial, final, and possibly also intermediate states of the molecule, $|0\rangle$, $|f\rangle$, and $|j\rangle$ respectively. These operators are projections of molecular transition vector and tensor operators onto the polarization directions of photons created or annihilated in the event (**Table 1**). The amplitudes depend on the wavenumber of the light used and can be real or complex. The probability of an optical event *W* is proportional to the square of the absolute value of its amplitude (**Table 2**). The proportionality constant is of no consequence in polarization spectroscopy, which deals with ratios of probabilities observed for the same collection of molecules under different conditions of light polarization. In the ratios, the proportionality constant cancels (except in strongly birefringent solvents, for which corrections are necessary).

### Polarized light

We label the laboratory system of axes *X*, *Y*, *Z*. Linearly polarized light (electromagnetic radiation) is characterized by a real unit vector $\boldsymbol{\varepsilon}^U$ oriented along the direction *U* of its electric field, and perpendicular to the propagation direction. Circularly polarized light is characterized by a complex unit vector. For right-handed light (*Z* component of photon angular momentum equal to $-\hbar$) propagating in the positive direction of *Z*, the polarization vector is $(\boldsymbol{\varepsilon}^X + i\boldsymbol{\varepsilon}^Y)/\sqrt{2} = \boldsymbol{\varepsilon}^+$, and for left-handed light, it is $(\boldsymbol{\varepsilon}^X - i\boldsymbol{\varepsilon}^Y)/\sqrt{2} = \boldsymbol{\varepsilon}^-$. Circularly polarized light is essential in chiroptical spectroscopy, and important in two-photon and Raman spectroscopy.

### Molecular transition moments and spectra

In electronic and vibrational spectroscopy we can neglect both molecular dimensions relative to the wavelength and the effects of the magnetic field of light relative to those of its electric field (electric dipole approximation). Then, the interaction of a single *U*-polarized photon with a molecule is described by the projection of the electric dipole moment vector operator ***M*** (**Table 3**) into $\boldsymbol{\varepsilon}^U$ (photon creation) or $\boldsymbol{\varepsilon}^{U*}$ (photon annihilation). Creation

**Table 1** Amplitudes of molecular optical events[a]

| | |
|---|---|
| 1. One-photon events (photon polarization: $U$) | |
| Annihilation | Creation |
| $\varepsilon^{U*}\cdot\boldsymbol{M}(0f)$ | $\varepsilon^{U}\cdot\boldsymbol{M}(f0)$ |
| 2. Successive two-photon events (photon polarizations: $U, V$) | |
| Annihilation + annihilation | Annihilation + creation |
| $\varepsilon^{U*}\cdot\boldsymbol{M}(0j)\boldsymbol{M}(j'f)\cdot\varepsilon^{V*}$ | $\varepsilon^{U*}\cdot\boldsymbol{M}(0j)\boldsymbol{M}(j'f)\cdot\varepsilon^{V}$ |
| 3. Simultaneous two-photon events (photon polarizations: $U, V$) | |
| Annihilation + annihilation | Annihilation + creation |
| $\varepsilon^{U*}\cdot\mathbf{T}(j,f)\cdot\varepsilon^{V*}$ | $\varepsilon^{U*}\cdot\alpha'(j,f)\cdot\varepsilon^{V}$ |

[a] $\boldsymbol{M}$ is the electric dipole transition moment vector; $\mathbf{T}$ is the two-photon absorption tensor; $\alpha'$ is the Raman scattering tensor; and $\varepsilon^U$ is a unit vector in the direction of light polarization $U$.

**Table 2** Probabilities of molecular optical events

| | |
|---|---|
| 1. One-photon events (tensor $\mathbf{O}$, photon polarization: $U$) | |
| Absorption | Emission |
| $[\boldsymbol{M}(0f)]^{U*}[\boldsymbol{M}(0f)]^{U}$ | $[\boldsymbol{M}(0f)]^{U}[\boldsymbol{M}(0f)]^{U*}$ |
| 2. Successive two-photon events (tensor ${}^{(4)}\mathbf{O}$, photon polarizations: $U, V$) | |
| Photoinduced dichroism | Photoluminescence |
| $[\boldsymbol{M}(0j)]^{U*}[\boldsymbol{M}(j'f)]^{V*}$ $[\boldsymbol{M}(0j)]^{U}[\boldsymbol{M}(j'f)]^{V*}$ | $[\boldsymbol{M}(0j)]^{U*}[\boldsymbol{M}(j'f)]^{V}$ $[\boldsymbol{M}(0j)]^{U}[\boldsymbol{M}(j'f)]^{V*}$ |
| 3. Simultaneous two-photon events (photon polarizations: $U, V$) | |
| Two-photon absorption | Raman |
| $[\mathbf{T}(0f,\tilde{\nu}_1)]^{U^*V^*}[\mathbf{T}(0f,\tilde{\nu}_1)^*]^{UV}$ | $[\alpha'(0f,\tilde{\nu}_1)]^{U^*V}[\alpha'(0f,\tilde{\nu}_1)^*]^{UV^*}$ |

and/or annihilation events involving two photons are described by projections of tensor operators of rank two, etc. (**Tables 1–3**). The off-diagonal matrix elements of $M$ between molecular states, $M(0f)$, are the (electric dipole) transition moments. The analogous matrix elements of tensor operators $\mathbf{T}$ and $\alpha'$ (**Table 3**) are the two-photon absorption tensor $\mathbf{T}(0f, \tilde{\nu}_1)$ and the Raman scattering tensor $\alpha'(0f, \tilde{\nu}_1)$.

The one-photon events in question (**Table 1**) are (i) the annihilation of a photon of wavenumber $\tilde{\nu}_1$ by the initial state 0 to yield an excited state $f$ of energy $\tilde{\nu}_1$ (ordinary absorption), and (ii) the creation of a photon of wavenumber $\tilde{\nu}_1$ by an excited state $f$ to yield a lower-energy state 0 (luminescence). The simultaneous two-photon events that we consider are (i) annihilation of two photons of wavenumbers $\tilde{\nu}_1$ and $\tilde{\nu}_2$ by the initial state 0 to yield an excited state of energy $\tilde{\nu}_1 + \tilde{\nu}_2$ (two-photon absorption), and (ii) annihilation of photon of wavenumber $\tilde{\nu}_1$ and creation of a photon of wavenumber $\tilde{\nu}_2$ to yield a (usually vibrationally) excited state of energy $\tilde{\nu}_1 - \tilde{\nu}_2$ above the initial state 0 (Raman scattering). The successive two-photon events that we consider are (i) annihilation of a photon of wavenumber $\tilde{\nu}_1$ that converts the initial state 0 to an intermediate state $j$ followed later by the annihilation of a photon of wavenumber $\tilde{\nu}_2$ that converts the intermediate state to the final state $f$ (photoinduced dichroism), and (ii) annihilation of a photon of wavenumber $\tilde{\nu}_1$ that converts the initial state 0 to an intermediate state $j$ followed later by the creation of a photon of wavenumber $\tilde{\nu}_2$ that converts the intermediate state to the final state $f$, often another (or even the same) vibrational level of the initial state 0 (photoluminescence). For both of these, the intermediate state $j$ may develop into state $j'$ during the interval between the first and the second event: the molecule may undergo internal conversion or intersystem crossing to another electronic state, it may vibrationally relax, it may change its conformation or chemical structure, etc. Especially important types of change are molecular rotation and the transfer of excitation energy to another molecule of the same kind, which may be differently oriented. For both compound events the tensor operator $\mathbf{M}$ is the direct product of the electric dipole moment operator $M$ with itself, $\mathbf{M} = MM$ (**Table 3**).

Optical transitions are associated with line shapes. The overall spectrum of a molecule is given by a sum of contributions provided by transitions from the initial state 0 to all possible final states $f$ plotted against $\tilde{\nu}_1$, $\tilde{\nu}_2$, or in two dimensions against $\tilde{\nu}_1$ and

**Table 3** Definition of transition moment vectors and tensors[a]

Electric dipole transition moment vector[b]

$$\boldsymbol{M}(0f) = \langle f|\, -|e|\sum_{l=1}^{n} \boldsymbol{r}_l + |e|\sum_{k=1}^{N} Z_k\boldsymbol{R}_k|0\rangle$$

Two-photon absorption tensor[c]

$$\mathbf{T}(0f,\tilde{\nu}_1) = \sum_j [\boldsymbol{M}(0j)\boldsymbol{M}(jf)/(\tilde{\nu}_j - \tilde{\nu}_1) + \boldsymbol{M}(jf)\boldsymbol{M}(0j)/(\tilde{\nu}_j - \tilde{\nu}_2)]$$

Raman scattering tensor[d]

$$\alpha'(0f,\tilde{\nu}_1) = \langle f|\sum_j [\boldsymbol{M}(0j)\boldsymbol{M}(0j)/(\tilde{\nu}_j + \tilde{\nu}_1) + \boldsymbol{M}(0j)\boldsymbol{M}(0j)/(\tilde{\nu}_i - \tilde{\nu}_1)]|0\rangle$$

[a] For transitions from an initial vibronic state 0 to the final state $f$. $Z_k$ is the atomic number of nucleus $k$, $e$ is the charge of the electron, $n$ is the number of electrons, $N$ is the number of nuclei, $\boldsymbol{r}_l$ is the position vector of the $l$th electron, and $\boldsymbol{R}_k$ is the position vector of the $k$th nucleus.

[b] In vibrational transitions, 0 and $f$ belong to the same electronic state and differ in the vibrational part, whereas in electronic transitions, they describe different electronic states.

[c] The sum is over all electronic states of the molecule $j$; $\tilde{\nu}_1 + \tilde{\nu}_2 = \tilde{\nu}_f$, where $\tilde{\nu}_f$ is the wavenumber of excitation to the final state $f$.

[d] Here, 0 and $f$ on the outside refer to the initial and final vibrational wavefunctions, $\tilde{\nu}_1$ is the excitation wavenumber, $\nu_j$ is the wavenumber of excitation to state $j$, the integration in the evaluation of $\boldsymbol{M}(0j)$ is only over the electronic degrees of freedom, and the summation is over all electronic states of the molecule $j$.

$\tilde{\nu}_2$. In vibrational spectra, the lines are usually sufficiently narrow that the contributions of the individual transitions do not overlap. It is then easy to measure the ratio of intensities with which any one line appears for various choices of photon polarization $U$ (or $U$ and $V$). When contributions from transitions provided by two or more excited states $f$ overlap, as often happens in electronic spectra, it is sometimes still possible to derive the ratios of spectral intensities for different choices of $U$ (or $U$ and $V$) using the stepwise reduction procedure.

Polarized intensity ratios equal the ratios of the probabilities listed in **Table 2**, and reveal the macroscopic optical anisotropy of a sample. To derive molecular optical anisotropy, the molecular orientation distribution in the laboratory system of axes needs to be considered.

## Description of molecular alignment

Polarization spectroscopy is most commonly performed on three types of samples: (i) isotropic fluid or rigid solutions (only for two-photon and higher-order processes); (ii) uniaxial partially aligned solutions possessing a unique direction ($Z$), with all directions perpendicular to it equivalent (any orthogonal pair can be chosen for $X$ and $Y$); and (iii) crystals, which contain molecules all oriented in the same way or in a small number of distinct ways. Isotropic samples are the easiest to study, since they do not change the state of polarization of light that propagates through them. This is also true of uniaxial samples if the electric vector of the light is either parallel or perpendicular to $Z$ (other orientations are best avoided). High-symmetry crystals, e.g. cubic, also satisfy this condition, but most low-symmetry crystals do not. Crystals are otherwise generally easier to handle since no or little orientation averaging is involved, but their spectra are often complicated by strong intermolecular interactions, absent in dilute solutions. In the following, we treat dilute isotropic or uniaxial samples with light polarization direction $U$ (and/or $V$) along $X$, $Y$ or $Z$.

To describe molecular orientation in a large assembly, we choose an arbitrary set of molecular axes $x'$, $y'$, $z'$ associated rigidly with the molecular framework (the primes indicate the arbitrariness of the choice). The orientation of a molecule in the sample is then described by values for the three rotational degree of freedom (e.g., the three Euler angles $\alpha'$, $\beta'$, $\gamma'$) that convert the laboratory directions $X$, $Y$, $Z$ into the molecular directions $x'$, $y'$, $z'$, and denote them collectively by $\Omega'$. The probability that a molecule has an orientation between $\Omega'$ and $\Omega' + \mathrm{d}\Omega'$ is $f(\Omega')\mathrm{d}\Omega'$. The orientation distribution function $f(\Omega')$ is normalized and permits the calculation of the orientation average of any angle-dependent quantity $a(\Omega')$:

$$\int f(\Omega')\,\mathrm{d}\Omega' = 1$$

$$\int a(\Omega')f(\Omega')\,\mathrm{d}\Omega' = \langle a \rangle \qquad [1]$$

In uniaxial samples, $f(\Omega')$ has the same value for all angles $\alpha'$ and depends only on $\beta'$ and $\gamma'$. In the special case of molecules for which all angles of rotation about their own $z'$ axis are equally likely, $f(\Omega')$ also has the same value for all angles $\gamma'$.

A redundant but convenient set of quantities that specify the orientation of the molecular $x'$, $y'$, $z'$ axes relative to the laboratory $X$, $Y$, $Z$ axes are the nine angles between the two sets. In uniaxial samples the angles $\eta_{x'}$, $\eta_{y'}$ and $\eta_{z'}$ between $x'$, $y'$, $z'$, respectively, and the $Z$ axis are sufficient, since rotation of the sample about $Z$ makes no observable difference. Only two of them are independent:

$$\begin{aligned} \cos\eta_{x'} &= -\sin\beta' \cos\gamma' \\ \cos\eta_{y'} &= \sin\beta' \sin\gamma' \\ \cos\eta_{z'} &= \cos\beta' \\ \cos^2\eta_{x'} + \cos^2\eta_{y'} &+ \cos^2\eta_{z'} = 1 \end{aligned} \qquad [2]$$

There are several equivalent choices of statistical parameters for the description of molecular alignment. We use one of these (orientation factors), and give formulae for conversion to two others.

### Orientation factors

Orientation factors are orientation averages of products of the direction cosines $\cos\eta_u$ of the molecular axes with respect to the unique sample axis $Z$. For one- and two-photon events only the second-order factors $[\mathbf{K}]_{uv}$ and fourth-order factors $[\mathbf{L}]_{stuv}$ are needed (the indices $s$, $t$, $u$, $v$ can acquire one of the values $x'$, $y'$ or $z'$, and their order is immaterial):

$$[\mathbf{K}]_{uv} = \langle \cos\eta_u \cos\eta_v \rangle \qquad [3]$$

$$[\mathbf{L}]_{stuv} = \langle \cos\eta_s \cos\eta_t \cos\eta_u \cos\eta_v \rangle \qquad [4]$$

Like the angles $\eta_{x'}$, $\eta_{y'}$ and $\eta_{z'}$ themselves, the orientation factors are redundant:

$$[\mathbf{K}]_{x'x'} + [\mathbf{K}]_{y'y'} + [\mathbf{K}]_{z'z'} = 1 \qquad [5]$$

$$[\mathbf{L}]_{x'x'uv} + [\mathbf{L}]_{y'y'uv} + [\mathbf{L}]_{z'z'uv} = [\mathbf{K}]_{uv} \quad [6]$$

and there are only $2j+1$ independent orientation factors of order $j$.

The orientation factors are elements of orientation tensors $\mathbf{K}$ (rank 2), $\mathbf{L}$ (rank 4), etc. The freedom available in the arbitrary choice of the axes system $x'$, $y'$, $z'$ in the molecular framework can be used to diagonalize the tensor $\mathbf{K}$. The principal axes of this tensor are called the molecular orientation axes $x$, $y$, and $z$. The eigenvalues are the molecular orientation factors $K_x = K_{xx}$, $K_y = K_{yy}$ and $K_z = K_{zz}$ ($K_{yz} = K_{zx} = K_{xy} = 0$), where $K_{uv} = [\mathbf{K}]_{uv}$. The axes are labelled such that $K_z \geq K_y \geq K_x$. The $z$ axis is the effective molecular orientation axis and is the direction in the molecular frame that is aligned best with $Z$, while the $x$ axis is the direction that makes the largest angle with $Z$. In a plot of $K_z$ against $K_y$, all orientation distributions are represented by points that lie in the 'orientation triangle', defined by its vertices ($K_z$, $K_y$): (1,0), molecular axis $z$ aligned perfectly with $Z$, $x$ and $y$ equivalent; (1/2, 1/2), the $yz$ plane aligned perfectly with $Z$, $y$ and $z$ equivalent; and (1/3, 1/3), $x$, $y$, and $z$ all equivalent. The sides of the triangle represent special orientation distributions: $K_y = K_x$, all those with $x$ and $y$ equivalent (rod-like alignment, all angles $\gamma$ equally likely); $K_z = K_y$, all those with $y$ and $z$ equivalent (disc-like alignment); $K_y = 1 - K_z$, $K_x = 0$, those with $yz$ plane along $Z$ (all Euler angles $\gamma$ equal to 90°).

If the molecular framework possesses planes or axes of symmetry, the orientation axes $x$, $y$, $z$ lie in them or perpendicular to them. For example, in molecules of $C_{2v}$ or $D_{2h}$ symmetry the positions of the $x$, $y$, $z$ axes are fully determined and only $1+j/2$ nonzero independent orientation factors of order $j$ remain. Using $L_{stuv} = [\mathbf{L}]_{stuv}$ and choosing the three independent fourth-order factors as $L_z = L_{zzzz}$, $L_y = L_{yyyy}$, and $L_x = L_{xxxx}$, we have for $L_{uv} = L_{uvuv}$ (with $u$ and $v$ in any order),

$$L_{uv} = [(K_u - L_u) + (K_v - L_v) - (K_w - L_w)]/2 \quad [7]$$

For examples of the values of molecular orientation factors $K$ and $L$ for selected orientation distributions, see **Table 4**.

### Saupe orientation matrices

The molecular orientation factor matrices $\mathbf{K}$ and $\mathbf{L}$ provide the simplest expressions for polarized spectral intensities but acquire awkward values for isotropic samples, $K_u = 1/3$ and $L_u = 1/5$. This shortcoming is removed by transformation to Saupe matrices:

$$\begin{aligned} S_{uv} &= \tfrac{1}{2}(3K_{uv} - \delta_{uv}) \\ S_{stuv} &= \tfrac{1}{8}[35L_{stuv} - 5(K_{st}\delta_{uv} + K_{su}\delta_{tv} + K_{sv}\delta_{tu} \\ &\quad + K_{tu}\delta_{sv} + K_{tv}\delta_{su} + K_{uv}\delta_{st}) \\ &\quad + (\delta_{st}\delta_{uv} + \delta_{su}\delta_{tv} + \delta_{sv}\delta_{tu})] \end{aligned} \quad [8]$$

where $\delta_{ij}$ is the Kronecker delta: $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. In isotropic solution, $S_{uv} = S_{stuv} = 0$.

In the molecular orientation system of axes $x$, $y$, $z$, the second-order Saupe matrix is diagonal

**Table 4** Orientation factors for some uniaxial orientation distributions

| | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* | *I* | *J* |
|---|---|---|---|---|---|---|---|---|---|---|
| $K_x$ | 0 | $(1-K_z)/2$ | 1/5 | 3/10 | 1/3 | 1/3 | 1/3 | 1/5 | $K_x$ | 0 |
| $K_y$ | 0 | $(1-K_z)/2$ | 1/5 | 3/10 | 1/3 | 1/3 | 1/3 | 2/5 | $(1-K_x)/2$ | 1/2 |
| $K_z$ | 1 | $K_z$ | 3/5 | 2/5 | 1/3 | 1/3 | 1/3 | 2/5 | $(1-K_x)/2$ | 1/2 |
| $L_x$ | 0 | $3(1-2K_z + L_z)/8$ | 3/35 | 6/35 | 1/5 | 1/9 | 1/3 | 3/35 | $L_x$ | 0 |
| $L_y$ | 0 | $3(1-2K_z + L_z)/8$ | 3/35 | 6/35 | 1/5 | 1/9 | 1/3 | 9/35 | $3(1-2K_x + L_x)/8$ | 3/8 |
| $L_z$ | 1 | $L_z$ | 3/7 | 9/35 | 1/5 | 1/9 | 1/3 | 9/35 | $3(1-2K_x + L_x)/8$ | 3/8 |
| $L_{xy}$ | 0 | $(1-2K_z + L_z)/8)$ | 1/35 | 2/35 | 1/15 | 1/9 | 0 | 2/35 | $(K_x - L_x)/2$ | 0 |
| $L_{xz}$ | 0 | $(K_z - L_z)/2$ | 3/35 | 1/14 | 1/15 | 1/9 | 0 | 2/35 | $(K_x - L_x)/2$ | 0 |
| $L_{yz}$ | 0 | $(K_z - L_z)/2$ | 3/35 | 1/14 | 1/15 | 1/9 | 0 | 3/35 | $(1-2K_x + L_x)/8$ | 1/8 |

For these orientation distributions, other orientation factors $K$ and $L$ vanish. (*A*) perfect alignment of $z$ axis; (*B*) rod like; (*C*) photoselected with $Z$-polarized light (absorption $z$-polarized); (*D*) photoselected with natural light propagating along $Z$ (absorption $xy$-polarized); (*E*) random; (*F*) $x,y,z$ axes all at magic angle (54.7°) to $Z$; (*G*) any one of $x,y,z$ axes aligned with $Z$, with equal probabilities; (*H*) photoselected with natural light propagating along $Z$ (absorption $x$-polarized), or photoselected with $Z$-polarized light (absorption $yz$-polarized); (*I*) disc-like; (*J*) perfect alignment of $yz$ plane, $y$ and $z$ equivalent.

($S_{yz} = S_{zx} = S_{xy} = 0$). The vertices of the orientation triangle are ($S_{zz}$, $S_{yy}$): (−1/2, 1), (1/4, 1/4), and (0, 0).

If a uniaxial alignment of a set of local axes $Z''$ relative to $Z$ is described by $S_{Z''Z''}$ and a uniaxial alignment of a set of axes $z$ relative to each $Z''$ is described by $S''_{zz}$, the uniaxial alignment of the axes $z$ with respect to $Z$ is given by $S_{zz} = S''_{zz}S_{Z''Z''}$.

### Order parameters

The use of Wigner matrices as an orthonormal basis set for the expansion of $f(\Omega)$ has the advantage of providing a nonredundant set of expansion coefficients. The coefficients define the order parameters $A$ and $B$, related to the orientation factors by

$$\begin{aligned}
A_0{}^{(0)} &= 1\\
A_0{}^{(2)} &= (1/2)(3K_z - 1)\\
A_1{}^{(2)} &= -(3/2)^{1/2}K_{xz}\\
B_1{}^{(2)} &= (3/2)^{1/2}K_{yz}\\
A_2{}^{(2)} &= (1/2)(3/2)^{1/2}(K_x - K_y)\\
B_2{}^{(2)} &= -(3/2)^{1/2}K_{xy}\\
A_0{}^{(4)} &= (1/8)(35L_z - 30K_z + 3)\\
A_1{}^{(4)} &= -(1/4)5^{1/2}(7L_{xzzz} - 3K_{xz})\\
B_1{}^{(4)} &= (1/4)5^{1/2}(7L_{yzzz} - 3K_{yz})\\
A_2{}^{(4)} &= (1/4)(5/2)^{1/2}[7(L_y - L_x) - 6(K_y - K_x)]\\
B_2{}^{(4)} &= -(1/2)(5/2)^{1/2}(7L_{xyzz} - K_{xy})\\
A_3{}^{(4)} &= (1/4)35^{1/2}(3L_{xyyz} - L_{xxxz})\\
B_3{}^{(4)} &= (1/4)35^{1/2}(3L_{xxyz} - L_{yyyz})\\
A_4{}^{(4)} &= (1/8)(35/2)^{1/2}(L_x + L_y - 6L_{xy})\\
&= (1/8)(35/2)^{1/2}[4(L_x + L_y) + 3(1 - L_z)\\
&\quad - 6(K_x + K_y)]\\
B_4{}^{(4)} &= (1/2)(35/2)^{1/2}(L_{xyyy} - L_{xxxy})
\end{aligned} \qquad [9]$$

For the $x$, $y$, $z$ axes, $A_1{}^{(2)} = B_1{}^{(2)} = B_2{}^{(2)} = 0$. Orientation triangle vertices are [$A_0{}^{(2)}$, $A_2{}^{(2)}$]: [1,0], [0, − $(1/4)(3/2)^{1/2}$], and [0,0]. In isotropic solutions, $A_0{}^{(0)} = 1$ and the other order parameters vanish.

### Polarized intensities for aligned samples

Expressions for polarized spectral intensities are of the form $\boldsymbol{\varepsilon}^U \cdot \mathbf{O}(0f) \cdot \boldsymbol{\varepsilon}^U$ for one-photon processes and of the form $\boldsymbol{\varepsilon}^U\boldsymbol{\varepsilon}^V \cdot {}^{(4)}\mathbf{O}(0f) \cdot \boldsymbol{\varepsilon}^U\boldsymbol{\varepsilon}^V$ for two-photon processes (**Table 2**). The nature and possible complex conjugation of the unit vectors $\boldsymbol{\varepsilon}^U$ and $\boldsymbol{\varepsilon}^V$ and the nature of the second-rank tensor $\mathbf{O}(0f)$ and the fourth-rank tensor ${}^{(4)}\mathbf{O}(0f)$ depend on the polarization of the photons and the specific measurement in question. Knowledge of the nine elements of $\mathbf{O}(0f)$ or the 81 elements of ${}^{(4)}\mathbf{O}(0f)$ thus permits the calculation of event probability for any choice of polarizations $U$ and $V$.

If all molecules were lined up with their $z'$ axis along $Z$, $y'$ axis long $Y$, and $x'$ axis along $X$, both $\mathbf{O}(0f)$ and ${}^{(4)}\mathbf{O}(0f)$ would have the same elements in the laboratory and molecular system of axes, $[\mathbf{O}(0f)]^{ZZ} = [\mathbf{O}(0f)]_{z'z'}$, etc. When the molecular axes are rotated to a general orientation $\Omega'$, the tensor elements in laboratory axes become linear combinations of those in molecular axes:

$$\begin{aligned}
[\mathbf{O}(0f)]^{UV} &= \Sigma_{uv}[\mathbf{D}^{UV}]_{uv}[\mathbf{O}(0f)]_{uv}\\
[{}^{(4)}\mathbf{O}(0f)]^{STUV} &= \Sigma_{stuv}[{}^{(4)}\mathbf{P}^{STUV}]_{stuv}[{}^{(4)}\mathbf{O}(0f)]_{stuv}
\end{aligned} \qquad [10]$$

The measured elements of the optical tensor in the laboratory $XYZ$ system are related to its elements in molecular $x'y'z'$ system by ensemble averaging:

$$[\,\langle \mathbf{O}(0f) \rangle\,]^{UV} = \Sigma_{uv}[\,\langle \mathbf{D}^{UV} \rangle\,]_{uv}[\mathbf{O}(0f)]_{uv} \qquad [11]$$

$$[\langle {}^{(4)}\mathbf{O}(0f) \rangle\,]^{STUV} = \Sigma_{stuv}[\,\langle {}^{(4)}\mathbf{P}^{STUV} \rangle\,]_{stuv} \times [{}^{(4)}\mathbf{O}(0f)]_{stuv} \qquad [12]$$

For light electric field along or perpendicular to $Z$, we need two $\langle \mathbf{D}^{UV} \rangle$ and seven $\langle {}^{(4)}\mathbf{P}^{STUV} \rangle$:

$$\begin{aligned}
&\langle \mathbf{D}^{ZZ} \rangle,\ \langle \mathbf{D}^{YY} \rangle = \langle \mathbf{D}^{XX} \rangle,\\
&\langle {}^{(4)}\mathbf{P}^{ZZZZ} \rangle,\ \langle {}^{(4)}\mathbf{P}^{YYYY} \rangle = \langle {}^{(4)}\mathbf{P}^{XXXX} \rangle,\\
&\langle {}^{(4)}\mathbf{P}^{YZYZ} \rangle = \langle {}^{(4)}\mathbf{P}^{XZXZ} \rangle,\\
&\langle {}^{(4)}\mathbf{P}^{ZYZY} \rangle = \langle {}^{(4)}\mathbf{P}^{ZXZX} \rangle,\\
&\langle {}^{(4)}\mathbf{P}^{XYXY} \rangle = \langle {}^{(4)}\mathbf{P}^{YXYX} \rangle,\\
&\langle {}^{(4)}\mathbf{P}^{++++} \rangle,\ \langle {}^{(4)}\mathbf{P}^{+-+-} \rangle
\end{aligned} \qquad [13]$$

Equations [11] and [12] separate the information on anisotropic molecular optical properties ($\mathbf{O}(0f)$ and ${}^{(4)}\mathbf{O}(0f)$ from that on orientation distribution and polarizer orientation ($\langle \mathbf{D}^{UV} \rangle$ and $\langle {}^{(4)}\mathbf{P}^{STUV} \rangle$), which can be expressed through the parameters chosen to describe the orientation distribution: the orientation factors [3], [4], the Saupe matrix elements [8], or the order parameters [9].

For uniaxial samples, the two required $\langle \mathbf{D} \rangle$ tensors are

$$\langle \mathbf{D}^{ZZ} \rangle = \mathbf{K}$$
$$\langle \mathbf{D}^{YY} \rangle = \langle \mathbf{D}^{XX} \rangle = (\mathbf{1} - \mathbf{K})/2 \qquad [14]$$

where $\mathbf{1}$ is the unit tensor. The five (in the $xyz$ system of axes, two) orientation factors $K$ describe these tensors fully. The tensors $\langle {}^{(4)}\mathbf{P} \rangle$ also contain the nine independent factors $L$:

$$[\langle {}^{(4)}\mathbf{P}^{ZZZZ} \rangle]_{stuv} = L_{stuv}$$
$$[\langle {}^{(4)}\mathbf{P}^{YYYY} \rangle]_{stuv} = (1/8)\{\delta_{st}\delta_{uv}(2\delta_{su} + 1) \times(1 - K_{st} - K_{uv}) + (\delta_{st} - \delta_{uv})[2\delta_{st}(\delta_{s,|u\times v|} - 3)K_{uv} - \delta_{uv}(2\delta_{|s\times t|,u} - 3)K_{st}] + [2\delta_{su}\delta_{tv} + \delta_{sv}\delta_{tu}) - 1] \times K_{|s\times t|,|u\times v|} + 3L_{stuv}\}$$
$$[\langle {}^{(4)}\mathbf{P}^{YZYZ} \rangle]_{stuv} = (1/2)(\delta_{su}K_{tv} - L_{stuv})$$
$$[\langle {}^{(4)}\mathbf{P}^{ZYZY} \rangle]_{stuv} = (1/2)(\delta_{tv}K_{su} - L_{stuv})$$
$$[\langle {}^{(4)}\mathbf{P}^{XYXY} \rangle]_{stuv} = (1/8)\{\delta_{st}\delta_{uv}(2\delta_{su} - 1) \times(1 - K_{st} - K_{uv}) + (\delta_{st} - \delta_{uv})[\delta_{st}(2\delta_{s,|u\times v|} - 1)K_{uv} - \delta_{uv}(2\delta_{|s\times t|,u} - 1)K_{st}] + [3(3\delta_{su}\delta_{tv} - \delta_{su} - \delta_{tv}) - (3\delta_{sv}\delta_{tu} - \delta_{sv} - \delta_{tu})]K_{|s\times t|,|u\times v|} + L_{stuv}\}$$
$$[\langle {}^{(4)}\mathbf{P}^{++++} \rangle]_{stuv} = (1/4)\{\delta_{st}\delta_{uv}(2\delta_{su} - 1) \times(1 - K_{st} - K_{uv}) + (\delta_{st} - \delta_{uv})[\delta_{st}(2\delta_{s,|u\times v|} - 1)K_{uv} - \delta_{uv}(2\delta_{|s\times t|,u} - 1)K_{st}] + [2\delta_{su}\delta_{tv} + \delta_{sv}\delta_{tu}) - 1] \times K_{|s\times t|,|u\times v|} + L_{stuv}\}$$
$$[\langle {}^{(4)}\mathbf{P}^{+-+-} \rangle]_{stuv} = (1/4)\{\delta_{st}\delta_{uv}(1 - K_{st} - K_{uv}) + (\delta_{st} - \delta_{uv})[\delta_{uv}K_{st} - \delta_{st}K_{uv}] + [2(\delta_{su}\delta_{tv} + \{\delta_{sv} - \delta_{tu}\}^2) - 1]K_{|s\times t|,|u\times v|} + L_{stuv}\}$$
$$[\langle {}^{(4)}\mathbf{P}^{UVUV} \rangle]_{stuv} = [\langle {}^{(4)}\mathbf{P}^{UVUV} \rangle]_{uvst} \qquad [15]$$

where the cross-product of two indices is defined in a manner similar to the vector product of basis vectors ($|x\times x| = 0$, $|x\times y| = z$, $|x\times z| = y$, etc.) and a $K$ with at least one subscript equal to zero vanishes.

For general orientation distributions, Equations [11] and [12] still apply, but $\langle \mathbf{D} \rangle$ and $\langle {}^{(4)}\mathbf{P} \rangle$ then also contain general orientation factors (averages of direction cosines with axes other than $Z$).

## One-photon processes

### Polarized absorption and emission

The same formulae apply in both cases, provided that the emission is excited in an isotropic fashion (e.g., chemiluminescence). Measurements on isotropic samples provide the same result for light polarized along any direction $U$ and for unpolarized light, and useful information is obtained only if the samples are aligned. The probability of absorption and emission of linearly $U$-polarized light is proportional to $[\langle \mathbf{O}(0f) \rangle]^{UU}$, with $\mathbf{O}(0f) = M(0f)M(0f)$ (**Table 2**). From Equations [11] and [14], for a uniaxial sample the dichroic ratio $d_f$ is

$$d_f = E^Z(0f)/E^Y(0f) = [\boldsymbol{M}(0f)]^Z[\boldsymbol{M}(0f)]^Z/[\boldsymbol{M}(0f)]^Y[\boldsymbol{M}(0f)]^Y = \Sigma_{uv}K_{uv}[\boldsymbol{M}(0f)]_u[\boldsymbol{M}(0f)]_v/\Sigma_{uv}(1/2) \times(\delta_{uv} - K_{uv})[\boldsymbol{M}(0f)]_u[\boldsymbol{M}(0f)]_v \qquad [16]$$

where the indices $u$ and $v$ run over $x'$, $y'$ and $z'$ and $E^U(0f)$ is proportional to the contribution of the $f$th transition to the observed $U$-polarized absorbance (or emission intensity). When the components of the transition moment $\boldsymbol{M}(0f)$ are expressed in the molecular orientation axis system $x$, $y$, $z$,

$$d_f = 2\{K_z[\boldsymbol{M}(0f)]_z{}^2 + K_y[\boldsymbol{M}(0f)]_y{}^2 + K_x[\boldsymbol{M}(0f)]_x{}^2\}/(1 - K_z)[\boldsymbol{M}(0f)]_z{}^2 + (1 - K_y)[\boldsymbol{M}(0f)]_y{}^2 + (1 - K_x)[\boldsymbol{M}(0f)]_x{}^2\} \qquad [17]$$

Simplifications occur in the presence of molecular symmetry. If it is sufficient (e.g. $C_{2v}$, $D_{2h}$) to define the location of the $x$, $y$, $z$ axes, and forces $\boldsymbol{M}(0f)$ to lie in one of these axes, say $u$, the result for a purely polarized (nonoverlapping) transition is that anticipated from the proportionality of absorption probability to the cosine squared of the angle between $U$ and $\boldsymbol{M}(0f)$,

$$d_f = 2K_u/(1 - K_u)$$
$$K_u = d_f/(d_f + 2) \qquad [18]$$

Thus, a measurement of $d_f$ for a $u$-polarized transition in a molecule of this symmetry determines $K_u$ and only two such measurements for distinct axes $u$ and $v$ are needed to determine $\mathbf{K}$. If the orientation distribution is such that two molecular axes are equivalent, say $K_x = K_y$, a single measurement suffices (e.g., that of $d_z$, which yields $K_z$, and $K_x = K_y = (1-K_z)/2$).

Total $u$-polarized absorption spectra $A_u(\tilde{\nu})$, $u = x, y, z$, can be obtained in special cases, but not in general, since three such spectra are possible but only two linearly independent spectra $E^Z$ and $E^Y = E^X$ can be measured. For instance, when $A_u(\tilde{\nu})$ is negligible in a spectral region, as low-energy out-of-plane polarized absorption is in aromatic hydrocarbons,

$$A_z(\tilde{\nu}) = [(1 - K_y)E^Z(\tilde{\nu}) - 2K_yE^Y(\tilde{\nu})]/(K_z - K_y)$$
$$A_y(\tilde{\nu}) = [2K_zE^Y(\tilde{\nu}) - (1 - K_z)E^Z(\tilde{\nu})]/(K_z - K_y) \quad [19]$$

In molecules of lower symmetry, the direction cosines $\cos\phi_u^f$ of $M(0f)$ in an $x'$, $y'$, $z'$ frame can be determined in favourable circumstances from relations such as

$$\Sigma_{uv}\cos\ \phi_u^f K_{uv}\ \cos\ \phi_v^f = d_f/(d_f + 2) \quad [20]$$

which simplifies in the $x$, $y$, $z$ system of axes to

$$\Sigma_u[K_u - d_f/(d_f + 2)]\ \cos^2\phi_u^f = 0 \quad [21]$$

#### Photoselection

Light polarized along $Z$, or, less effectively, unpolarized or circularly polarized light propagating along $Z$, selects molecules for excitation by the orientation of their absorbing transition moment. This process is known as photoselection, and because of it Equations [16]–[21] do not apply to emission by samples excited by a beam or another anisotropic light source. Explicit expressions can be written for the orientation distribution of the molecules excited from an isotropic solution and those left behind. They are simple only if the depletion of the ground state is small, but the remaining molecules can be aligned much more highly if the depletion is significant.

The results for molecules photoselected from an isotropic sample with $Z$-polarized light are particularly useful in two cases: (i) the molecule is of high enough symmetry that each $M(0f)$ must lie in one of the orientation axes $x$, $y$, $z$, but the transitions may overlap, and (ii) only one transition ($0 \rightarrow f$) absorbs at $\tilde{\nu}_0$ (**Table 4**). In the former case,

$$K_u(\tilde{\nu}_0) = [1 + 2r_u(\tilde{\nu}_0)]/5$$
$$L_{u\nu}(\tilde{\nu}_0) = (1 + 2\delta_{u\nu})[1 + 2r_u(\tilde{\nu}_0) + 2r_\nu(\tilde{\nu}_0)]/35 \quad [22]$$

where $r_u$ is the fraction of absorption polarized along $u$.

Using the orientation factors Equation [22], Equations [16]–[21] provide expressions for polarized absorption and emission by molecules photoselected from isotropic samples and represent a special case of the treatment given below, which handles both one-photon events simultaneously and describes photoinduced absorption and luminescence from all uniaxial samples.

## Two-photon processes

Polarized intensities of all successive and simultaneous two-photon events on isotropic or partially aligned optically inactive samples are given by Equation [12], using the same tensor $\langle {}^{(4)}\mathbf{P}\rangle$, with different choices for the tensor ${}^{(4)}\mathbf{O}$ (**Table 2**). For successive events, measurements with circularly polarized light do not offer any additional information and will not be considered, whereas for simultaneous two-photon processes they are indispensable in certain cases.

For uniaxial samples and molecules whose symmetry dictates the orientation axes $x$, $y$, $z$, there are only three independent orientation factors $L$ and Equation [15] simplifies:

$$\begin{aligned}
[\langle {}^{(4)}\mathbf{P}^{ZZZZ}\rangle]_{stuv} &= \delta_{|s\times t|,|u\times v|}L_{stuv}\\
[\langle {}^{(4)}\mathbf{P}^{YYYY}\rangle]_{stuv} &= (1/8)[\delta_{st}\delta_{uv}(2\delta_{su}+1)\\
&\quad\times(1-K_{st}-K_{uv}) + (\delta_{su}\delta_{tv}+\delta_{sv}\delta_{tu})\\
&\quad\times K_{|s\times t|,|u\times v|} + 3\delta_{|s\times t|,|u\times v|}L_{stuv}]\\
[\langle {}^{(4)}\mathbf{P}^{YZYZ}\rangle]_{stuv} &= (1/2)(\delta_{su}\delta_{tv}K_{tv} - \delta_{|s\times t|,|u\times v|}L_{stuv})\\
[\langle {}^{(4)}\mathbf{P}^{ZYZY}\rangle]_{stuv} &= (1/2)(\delta_{su}\delta_{tv}K_{su} - \delta_{|s\times t|,|u\times v|}L_{stuv})\\
[\langle {}^{(4)}\mathbf{P}^{XYXY}\rangle]_{stuv} &= (1/8)[\delta_{st}\delta_{uv}(2\delta_{su}-1)\\
&\quad\times(1-K_{st}-K_{uv}) + (3\delta_{su}\delta_{tv}-\delta_{sv}\delta_{tu})\\
&\quad\times K_{|s\times t|,|u\times v|} + \delta_{|s\times t|,|u\times v|}L_{stuv}\\
[\langle {}^{(4)}\mathbf{P}^{++++}\rangle]_{stuv} &= (1/4)[\delta_{st}\delta_{uv}(2\delta_{su}-1)\\
&\quad\times(1-K_{st}-K_{uv}) + (\delta_{su}\delta_{tv}+\delta_{sv}\delta_{tu})\\
&\quad\times K_{|s\times t|,|u\times v|} + \delta_{|s\times t|,|u\times v|}L_{stuv}]\\
[\langle {}^{(4)}\mathbf{P}^{+-+-}\rangle]_{stuv} &= (1/4)[\delta_{st}\delta_{uv}\\
&\quad\times(1-K_{st}-K_{uv}) + (\delta_{su}\delta_{tv}-\delta_{sv}\delta_{tu})\\
&\quad\times K_{|s\times t|,|u\times v|} + \delta_{|s\times t|,|u\times v|}L_{stuv}]\\
[\langle {}^{(4)}\mathbf{P}^{UVUV}\rangle]_{stuv} &= [\langle {}^{(4)}\mathbf{P}^{UVUV}\rangle]_{uvst}
\end{aligned} \quad [23]$$

Unlike one-photon measurements, two-photon measurements are useful even on isotropic samples. For these, the orientation factors $K$ and $L$ have the values given in **Table 4**. The results then are especially simple and this practically important case will be described separately.

### Photoinduced absorption and emission (static)

Identical formulae apply in both cases. When excited molecules do not rotate or transfer excitation energy to differently oriented molecules between the initial absorption event and the later absorption (photoinduced dichroism) or emission (photoluminescence) event, their orientation factors are time independent. Negligible ground-state depletion is assumed in the first step.

When $M(0j)M(j'f)M(0j)M(j'f)$ is substituted for $^{(4)}\mathbf{O}$ is Equation [12], only 36 of the 81 terms in the quadruple sum remain. The contribution to the overall event probability that is provided by a combination of transition $0 \to j$ involving a $U$-polarized photon in the first step with transition $j' \to f$ involving a $V$-polarized photon in the second step is proportional to $F^{UV}(0j, j'f)$:

$$F^{UV}(0j, j'f) = \Sigma_{(su),(tv)}[2(2 - \delta_{su} - \delta_{tv}) + \delta_{su}\delta_{tv}] \times[\langle {}^{(4)}\mathbf{P}^{UVUV} \rangle]_{stuv}[\boldsymbol{M}(0j)]_s[\boldsymbol{M}(j'f)]_t[\boldsymbol{M}(0j)]_u[\boldsymbol{M}(j'f)]_v \quad [24]$$

where each of the unordered pairs $(su)$ and $(tv)$ can acquire the values $(x'x')$, $(y'y')$, $(z'z')$, $(y'z')$, $(z'x')$, and $(x'y')$. In molecules of high enough symmetry for transition moments to lie in the $x$, $y$, $z$ axes, only the nine terms with $s = u$ and $t = v$ contribute:

$$F^{UV}(0j, j'f) = \Sigma_{st}[\langle {}^{(4)}\mathbf{P}^{UVUV} \rangle]_{stst} \times[\boldsymbol{M}(0j)]_s^2[\boldsymbol{M}(j'f)]_t^2 \quad [25]$$

If several transitions are present, their contributions can be combined into purely $s$-polarized excitation spectra $A_s(\tilde{\nu}_1)$ and purely $t$-polarized photoinduced absorption or emission spectra $B_t(\tilde{\nu}_2)$:

$$F^{UV}(\tilde{\nu}_1, \tilde{\nu}_2)/F^{U'V'}(\tilde{\nu}_1, \tilde{\nu}_2) = \Sigma_{st}[\mathbf{S}^{UV}]_{st}A_s(\tilde{\nu}_1)B_t(\tilde{\nu}_2)/\Sigma_{st}[\mathbf{S}^{U'V'}]_{st}A_s(\tilde{\nu}_1)B_t(\tilde{\nu}_2) \quad [26]$$

where

$$[\mathbf{S}^{ZZ}]_{st} = L_{st}$$
$$[\mathbf{S}^{YY}]_{st} = (1/8)[(1 + 2\delta_{st})(1 - K_s - K_t) + 3L_{st}$$
$$[\mathbf{S}^{YZ}]_{st} = (1/2)(K_t - L_{st})$$
$$[\mathbf{S}^{ZY}]_{st} = (1/2)(K_s - L_{st})$$
$$[\mathbf{S}^{XY}]_{st} = (1/8)[(3 - 2\delta_{st})(1 - K_s - K_t) + L_{st}] \quad [27]$$

If only a single $u$-polarized transition $0 \to j$ contributes at $\tilde{\nu}_1$ and a single $\nu$-polarized transition $j' \to f$ at $\tilde{\nu}_2$,

$$F^{UV}(\tilde{\nu}_1, \tilde{\nu}_2)/F^{U'V'}(\tilde{\nu}_1, \tilde{\nu}_2) = [\mathbf{S}^{UV}]_{uv}/[\mathbf{S}^{U'V'}]_{uv} \quad [28]$$

The 'stepwise reduction' procedure may yield the ratio $[\mathbf{S}^{UV}]_{uv}/[\mathbf{S}^{U'V'}]_{uv}$ even in the case of overlapping transitions. Overlapping phosphorescent emissions from the three triplet sublevels can be separated at very low temperatures when their populations do not equilibrate and lifetimes differ.

### Photoinduced absorption and emission in rigid isotropic solution (static)

If the starting orientation distribution is random (ordinary photoselection), the orientation factors are known (**Table 4**). Equation [24] then yields for $F^{\parallel}$ (both polarization directions equal) and $F^{\perp}$ (polarizations mutually perpendicular):

$$F^{\parallel}(\tilde{\nu}_1, \tilde{\nu}_2)/F^{\perp}(\tilde{\nu}_1, \tilde{\nu}_2) = [\Sigma_{jf}r_j(\tilde{\nu}_1)q_f(\tilde{\nu}_2) \times(1 + 2\Sigma_u\cos^2\phi_u^f\cos^2\phi_u^j)]/[\Sigma_{jf}r_j(\tilde{\nu}_1)q_f(\tilde{\nu}_2) \times(2 - \Sigma_u\cos^2\phi_u^f\cos^2\phi_u^j)] \quad [29]$$

where $r_j(\tilde{\nu}_1)$ is the fraction of the absorption at $\tilde{\nu}_1$ due to the $j$th transition and $q_f(\tilde{\nu}_2)$ is the fraction of absorption or emission at $\tilde{\nu}_2$ due to the $f$th transition. The degree of anisotropy $R$ is

$$R(\tilde{\nu}_1, \tilde{\nu}_2) = (F^{\parallel} - F^{\perp})/(F^{\parallel} + 2F^{\perp}) = \Sigma_{jf}r_j(\tilde{\nu}_1)q_f(\tilde{\nu}_2)(3\Sigma_u\cos^2\phi_u^f\cos^2\phi_u^j - 1)/5 \quad [30]$$

If symmetry constrains transition moments to the $x$, $y$, $z$ axes,

$$F^{\parallel}(\tilde{\nu}_1, \tilde{\nu}_2)/F^{\perp}(\tilde{\nu}_1, \tilde{\nu}_2) = [\Sigma_{uv}(1 + 2\delta_{uv}) \times A_u(\tilde{\nu}_1)B_v(\tilde{\nu}_2)][\Sigma_{uv}(2 - \delta_{uv})A_u(\tilde{\nu}_1)B_v(\tilde{\nu}_2)] \quad [31]$$

$$R(\tilde{\nu}_1, \tilde{\nu}_2) = (3/5)[\Sigma_u r_u(\tilde{\nu}_1)q_u(\tilde{\nu}_2) - (1/3)] \quad [32]$$

In a molecule of any symmetry, when the $j$th and $f$th transitions do not overlap with others,

$$R(j, f) = (3\cos^2\phi - 1)/5 = (2/5)P_2(\cos\phi) \quad [33]$$

where $\phi$ is the angle between the transition moments $\boldsymbol{M}(0j)$ and $\boldsymbol{M}(j'f)$ and $P_2$ is the second Legendre polynomial. The limiting values are reached when $\phi$ equals 0 (parallel transitions, $R = 2/5$) or 90° (perpendicular transitions, $R = -1/5$). Molecules with a threefold or higher symmetry axis may have degenerate states. If one transition is polarized parallel to the high-order axis and the other in the plane perpendicular to it, $R = -1/5$. If both are polarized in this plane, $R = 1/10$.

### Photoinduced absorption and emission in fluid isotropic solution (dynamic)

If the excited molecule rotates during its lifetime, or transfers its excitation to another differently oriented molecule, the polarization of the photoinduced absorption or luminescence becomes a function of time $\tau$ that has elapsed after the initial excitation event. We assume that the initial orientation distribution is isotropic and that the optical properties of the molecule are known from static studies. The steps required are an evaluation of the rotational correlation functions and their interpretation in terms of a suitable model for diffusion or energy transfer.

If symmetry constrains the directions of transition moments and the principal axes of the orientation tensor to the symmetry axes $x$, $y$, $z$, the results are the simplest. If the first photon is absorbed in a transition purely polarized along $u$ and the second is absorbed or emitted in a transition purely polarized along $v$, the degree of anisotropy shows a double exponential decay in time:

$$\begin{aligned} R(u,u,\tau) &= (3/10)\{[2/3 + (D - D_u)/\Delta] \\ &\quad \times \exp[-2(3D+\Delta)\tau] + [2/3 - (D - D_u)/\Delta] \\ &\quad \times \exp[-2(3D-\Delta)\tau]\} \\ R(u,v,\tau) &= (3/10)\{[1/3 + (2D - D_u - D_v)/\Delta] \\ &\quad \times \exp[-2(3D+\Delta)\tau] + [1/3 - (2D - D_u - D_v)/\Delta] \\ &\quad \times \exp[-2(3D-\Delta)\tau]\} \end{aligned} \quad [34]$$

where the second equation applies if $u \neq v$, and $D = (1/3)\Sigma_u D_u$, $\Delta = [(1/2)\Sigma_{uv}(3\delta_{uv} - 1)D_u D_v]^{1/2}$, and $D_u$ are the diagonal elements of the diffusion tensor. In lower-symmetry cases, up to five simultaneous exponential decays are expected, not counting the decay of the excited state itself.

If the diffusion tensor is isotropic, only a single exponential decay of $R$ is observed even in low-symmetry molecules, and the rotational relaxation time is $\tau_R = 1/6D$:

$$R(\tilde{\nu}_1, \tilde{\nu}_2, \tau) = R(\tilde{\nu}_1, \tilde{\nu}_2)\exp(-6D\tau) \quad [35]$$

where $R(\tilde{\nu}_1, \tilde{\nu}_2) = R(\tilde{\nu}_1, \tilde{\nu}_2, 0)$ is the degree of anisotropy at time $\tau = 0$ after excitation, given for purely polarized absorbing and emitting transitions by Equation [33]. In this limit, the degree of anisotropy $\bar{R}(\tilde{\nu}_1, \tilde{\nu}_2)$ in a steady excitation experiment on the same sample is

$$\bar{R}(\tilde{\nu}_1, \tilde{\nu}_2) = R(\tilde{\nu}_1, \tilde{\nu}_2)/(1 + 6D\tau_0) \quad [36]$$

where $\tau_0$ is the excited state lifetime.

### Two-photon absorption and Raman scattering

In uniaxial samples, seven distinct combinations of photon polarization are important. The observed intensities are given by $I^{ZZ}$, $I^{YY} = I^{XX}$, $I^{YZ} = I^{XZ}$, $I^{ZY} = I^{ZX}$, $I^{XY} = I^{YX}$, and two intensities involving circularly polarized light propagating along $Z$, $I^{CON}$ and $I^{DIS}$ (or $I^{SYN}$ and $I^{ANTI}$). In a measurement of $I^{CON}$ ($I^{DIS}$) the tip of the electric vector viewed along $Z$ rotates in the same (opposite) sense for both photons, regardless of whether they propagate parallel or antiparallel to each other. In a measurement of $I^{SYN}$ ($I^{ANTI}$) the two photons are of the same (opposite) handedness regardless of the sense of their propagation. If the photons propagate in the same direction, SYN = CON, ANTI = DIS. If they propagate in opposite directions, SYN = DIS, ANTI = CON.

In Raman spectroscopy, all seven measurements are available. In two-photon spectroscopy, all seven are accessible only if the two photons are taken from different beams, and otherwise only $I^{ZZ}(\tilde{\nu}_1, \tilde{\nu}_1)$, $I^{YY}(\tilde{\nu}_1, \tilde{\nu}_1) = I^{XX}(\tilde{\nu}_1, \tilde{\nu}_1)$, and $I^{CON}(\tilde{\nu}_1, \tilde{\nu}_1)$ can be measured.

Up to a proportionality constant, the contribution of the $f$th transition to the observed intensity $I(0f, \tilde{\nu}_1)$ is again given by Equation [12], with proper choice of the tensor $^{(4)}\mathbf{O}$ (**Table 2**). $\mathbf{T}(0f, \tilde{\nu}_1)$ is symmetric if the two photons have the same wavenumber and $\boldsymbol{\alpha}'(0f, \tilde{\nu}_1)$ is symmetric for nonresonant Raman. The elements of $\langle {}^{(4)}\mathbf{P}^{UVUV} \rangle$ for circularly polarized light differ:

Two-photon absorption:

$$\langle {}^{(4)}\mathbf{P}^{CON} \rangle = \langle {}^{(4)}\mathbf{P}^{++++} \rangle \qquad \langle {}^{(4)}\mathbf{P}^{DIS} \rangle = \langle {}^{(4)}\mathbf{P}^{+-+-} \rangle$$

Raman:

$$\langle {}^{(4)}\mathbf{P}^{\mathrm{CON}}\rangle = \langle {}^{(4)}\mathbf{P}^{+-+-}\rangle \qquad \langle {}^{(4)}\mathbf{P}^{\mathrm{DIS}}\rangle = \langle {}^{(4)}\mathbf{P}^{++++}\rangle \tag{37}$$

Simplification occurs if the molecule has high enough symmetry to dictate the positions of the $x$, $y$, $z$ axes. Then, only two orientation factors $K$ and three $L$ are independent, the tensors $\langle {}^{(4)}\mathbf{P}^{UVUV}\rangle$ have 21 nonzero elements of the types $[\langle {}^{(4)}\mathbf{P}^{UVUV}\rangle]_{sstt}$, $[\langle {}^{(4)}\mathbf{P}^{UVUV}\rangle]_{stst}$, and $[\langle {}^{(4)}\mathbf{P}^{UVUV}\rangle]_{stts}$, and the quadruple sum in Equation [12] is reduced to three double sums. The two-photon absorption cross-section is proportional to

$$\begin{aligned} I^{UV}(0f,\tilde{\nu}_1) = \Sigma_{st}\{&[\mathbf{R}_F{}^{UV}]_{st}[\mathbf{T}(0f,\tilde{\nu}_1)]_{ss}[\mathbf{T}^*(0f,\tilde{\nu}_1)]_{tt} \\ &+ [\mathbf{R}_G{}^{UV}]_{st}[\mathbf{T}(0f,\tilde{\nu}_1)]_{st}[\mathbf{T}^*(0f,\tilde{\nu}_1)]_{st} \\ &+ [\mathbf{R}_H{}^{UV}]_{st}[\mathbf{T}(0f,\tilde{\nu}_1)]_{st}[\mathbf{T}^*(0f,\tilde{\nu}_1)]_{st}\} \end{aligned} \tag{38}$$

and for Raman the result is identical except that $\alpha'(0f, \tilde{\nu}_1)$ replaces $\mathbf{T}(0f, \tilde{\nu}_1)$. The matrices $\mathbf{R}_F{}^{UV}$, $\mathbf{R}_G{}^{UV}$, and $\mathbf{R}_H{}^{UV}$ are analogous to the matrix $\mathbf{S}^{UV}$ introduced in Equation [27]:

$$\begin{aligned} &\mathbf{R}_F{}^{UU} = \mathbf{R}_G{}^{UU} = \mathbf{R}_H{}^{UU} = \mathbf{R}^{UU} \\ &\mathbf{R}_F{}^{UV} = \mathbf{R}_F{}^{VU} = \mathbf{R}_H{}^{UV} = \mathbf{R}_H{}^{VU} \quad \text{if } U \neq V \\ &\mathbf{R}_F{}^{SYN} = \mathbf{R}_H{}^{ANTI} = 2\mathbf{R}_F{}^{XY} = 2\mathbf{R}_H{}^{XY} \\ &\mathbf{R}_F{}^{ANTI} = \mathbf{R}_G{}^{ANTI} = \mathbf{R}_G{}^{SYN} = \mathbf{R}_H{}^{SYN} \\ &[\mathbf{R}^{ZZ}]_{st} = [\mathbf{R}^{ZZ}]_{ts} = (1/3)(3-2\delta_{st})L_{st} \\ &[\mathbf{R}^{YY}]_{st} = [\mathbf{R}^{YY}]_{ts} = (1/8)[1-K_s-K_t \\ &\qquad + (3-2\delta_{st})L_{st}] \\ &[\mathbf{R}_F{}^{YZ}]_{st} = [\mathbf{R}_F{}^{YZ}]_{ts} = (1/6)[\delta_{st}K_s - (3-2\delta_{st})L_{st}] \\ &[\mathbf{R}_G{}^{YZ}]_{st} = (1/6)(3-2\delta_{st})(K_t - L_{st}) \\ &[\mathbf{R}_G{}^{ZY}]_{st} = (1/6)(3-2\delta_{st})(K_s - L_{st}) \\ &[\mathbf{R}_F{}^{XY}]_{st} = [\mathbf{R}_F{}^{XY}]_{ts} = (1/24)[(4\delta_{st}-3) \\ &\times(1-K_s-K_t) + (3-2\delta_{st})L_{st}] \\ &[\mathbf{R}_G{}^{XY}]_{st} = [\mathbf{R}_G{}^{XY}]_{ts} = (1/24)[(9-8\delta_{st}) \\ &\times(1-K_s-K_t) + (3-2\delta_{st})L_{st}] \\ &[\mathbf{R}_F{}^{ANTI}]_{st} = [\mathbf{R}_F{}^{ANTI}]_{ts} = (1/12)(3-2\delta_{st}) \\ &\times(1-K_s-K_t+L_{st}) \end{aligned} \tag{39}$$

The presence of molecular symmetry also simplifies the molecular tensors $\mathbf{T}(0f, \tilde{\nu}_1)$ and $\alpha'(0f, \tilde{\nu}_1)$. They are commonly written as a sum of a spherically symmetric isotropic part $\mathbf{T}^{(0)}$ or $\alpha'^{(0)}$, a symmetric traceless part $\mathbf{T}^{(s)}$ or $\alpha'^{(s)}$, and an antisymmetric part $\mathbf{T}^{(as)}$ or $\alpha'^{(as)}$:

$$\begin{matrix} (0) & (s) & (as) \\ \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} & \begin{pmatrix} b & c & d \\ c & e & f \\ d & f & -b-e \end{pmatrix} & \begin{pmatrix} 0 & g & h \\ -g & 0 & i \\ -h & -i & 0 \end{pmatrix} \end{matrix} \tag{40}$$

If the axes $x$, $y$, $z$ are symmetry-determined, Equation [38] simplifies, as the tensor has either (i) only two symmetrically disposed possibly unequal off-diagonal elements (forms $\mathbf{T}^{(s)} + \mathbf{T}^{(as)}$ or $\alpha'^{(s)} + \alpha'^{(as)}$), or (ii) only up to three possibly different diagonal elements (forms $\mathbf{T}^{(0)} + \mathbf{T}^{(s)}$ or $\alpha'^{(0)} + \alpha'^{(s)}$).

### Two-photon absorption in isotropic solution

While measurement in an isotropic solution does not differentiate among the molecular $x$, $y$, $z$, axes, it does permit a separation of the $F$, $G$, and $H$ parts of the intensity. Three of the four distinct measurements are linearly independent:

$$\begin{aligned} I^{\parallel}(0f,\tilde{\nu}_1) &= I^{ZZ}(0f,\tilde{\nu}_1) = I^{YY}(0f,\tilde{\nu}_1) \\ &= (1/30)(2\delta_F + 2\delta_G + 2\delta_H) \\ I^{\perp}(0f,\tilde{\nu}_1) &= I^{YZ}(0f,\tilde{\nu}_1) = I^{ZY}(0f,\tilde{\nu}_1) = I^{XY}(0f,\tilde{\nu}_1) \\ &= (1/30)(-\delta_F + 4\delta_G - \delta_H) \\ I^{SYN}(0f,\tilde{\nu}_1) &= (1/30)(-2\delta_F + 3\delta_G + 3\delta_H) \\ I^{ANTI}(0f,\tilde{\nu}_1) &= (1/30)(3\delta_F + 3\delta_G - 2\delta_H) \end{aligned} \tag{41}$$

where the rotational invariants of the tensor are defined by

$$\begin{aligned} \delta_F &= \Sigma_{st}[\mathbf{T}(0f,\tilde{\nu}_1)]_{ss}[\mathbf{T}(0f,\tilde{\nu}_1)^*]_{tt} \\ \delta_G &= \Sigma_{st}[\mathbf{T}(0f,\tilde{\nu}_1)]_{st}[\mathbf{T}(0f,\tilde{\nu}_1)^*]_{st} \\ \delta_H &= \Sigma_{st}[\mathbf{T}(0f,\tilde{\nu}_1)]_{st}[\mathbf{T}(0f,\tilde{\nu}_1)^*]_{ts} \end{aligned} \tag{42}$$

If $\mathbf{T}(0f, \tilde{\nu}_1)$ has the form $\mathbf{T}^{(0)}$, $\delta_F = 3\delta_G$ and $\delta_H = \delta_G$. For the form $\mathbf{T}^{(s)}$, $\delta_F = 0$ and $\delta_H = \delta_G$, and for $\mathbf{T}^{(as)}$, $\delta_F = 0$ and $\delta_H = -\delta_G$. Circularly polarized light is needed to determine all three invariants. When both photons are taken from the same beam, only $I^{\parallel}(0f, \tilde{\nu}_1)$ and $I^{SYN}(0f, \tilde{\nu}_1)$ can be measured. However, $\mathbf{T}(0f, \tilde{\nu}_1)$ is then symmetric, $\delta_H = \delta_G$, and the two remaining invariants can be determined.

**Raman scattering in isotropic solution**

The $x$, $y$, $z$ axes cannot be distinguished, and three rotational invariants of $\alpha'(0f, \tilde{\nu}_1)$ result, usually defined as the isotropic part $\bar{\alpha}^2$, symmetric anisotropy $\gamma_s^2$, and antisymmetric anisotropy $\gamma_{as}^2$:

$$\begin{aligned}\bar{\alpha}^2 &= (1/9)\Sigma_{st}[\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)]_{ss}[\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)^*]_{tt}\\ \gamma_s^2 &= (1/2)\Sigma_{st}|[\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)]_{ss} - [\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)]_{tt}|^2\\ &\quad + (3/4)\Sigma_{s\neq t}|[\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)]_{st} + [\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)]_{ts}|^2\\ \gamma_{as}^2 &= (3/4)\Sigma_{st}|[\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)]_{st} - [\boldsymbol{\alpha}'(0f,\tilde{\nu}_1)]_{ts}|^2 \end{aligned} \quad [43]$$

These are related to $\delta_F$, $\delta_G$ and $\delta_H$ through

$$\begin{aligned}\bar{\alpha}^2 &= (1/9)\delta_F\\ \gamma_s^2 &= (3/4)(\delta_G + \delta_H) - (1/2)\delta_F\\ \gamma_{as}^2 &= (3/4)(\delta_G - \delta_H)\end{aligned} \quad [44]$$

and have particularly simple values for the tensors $\alpha'^{(0)}$, $\alpha'^{(s)}$, and $\alpha'^{(as)}$:

$$\begin{aligned}\boldsymbol{\alpha}'^{(0)}&: \quad \bar{\alpha}^2 = (1/3)\delta_G, \gamma_s^2 = \gamma_{as}^2 = 0\\ \boldsymbol{\alpha}'^{(s)}&: \quad \gamma_s^2 = (3/2)\delta_G, \bar{\alpha}^2 = \gamma_{as}^2 = 0\\ \boldsymbol{\alpha}'^{(as)}&: \quad \gamma_{as}^2 = (3/2)\delta_G, \bar{\alpha}^2 = \gamma_{as}^2 = 0\end{aligned} \quad [45]$$

The observable intensities are proportional to

$$\begin{aligned}I^{\|}(0f,\tilde{\nu}_1) &= I^{ZZ}(0f,\tilde{\nu}_1) = I^{YY}(0f,\tilde{\nu}_1)\\ &= \bar{\alpha}^2 + (4/45)\gamma_z^2\\ I^{\perp}(0f,\tilde{\nu}_1) &= I^{YZ}(0f,\tilde{\nu}_1) = I^{ZY}(0f,\tilde{\nu}_1)\\ &= I^{XY}(0f,\tilde{\nu}_1) = (1/15)\gamma_s^2 + (1/9)\gamma_{as}^2\\ I^{SYN}(0f,\tilde{\nu}_1) &= (2/15)\gamma_s^2\\ I^{ANTI}(0f,\tilde{\nu}_1) &= \bar{\alpha}^2 + (1/45)\gamma_s^2 + (1/9)\gamma_{as}^2\end{aligned} \quad [46]$$

Unless $\gamma_{as}^2 = 0$ (nonresonant Raman), measurement with circularly polarized light is needed to determine the three invariants. The results are usually expressed as the depolarization ratio $I^{\perp}(0f, \tilde{\nu}_1)/I^{\|}(0f, \tilde{\nu}_1)$. For nonresonant Raman scattering this equals 3/4 for $\alpha' = \alpha'^{(s)}$, is lower otherwise, and reaches zero if $\alpha' = \alpha'^{(0)}$ (totally symmetric vibration). In resonant Raman, $\gamma_{as}^2$ can be different from zero, and the depolarization ratio can exceed 3/4 ('anomalous polarization').

## List of symbols

$A,B$ = order parameters; $D_u$ = diagonal elements of the diffusion tensor; **D** = rank 2 orientation transformation tensor; $\hbar$ = Planck constant/2π; $I$ = light intensity; **K** = orientation tensor, rank 2; $K_{uv}$ = orientation factor, rank 2; $L_{stuv}$ = orientation factor, rank 4; **L** = orientation tensor, rank 4; $M$ = electric dipole moment operator; **M** = tensor operator = $MM$; **O** = one-photon event tensor; $^{(4)}$**O** = two-photon event tensor; $P_2$ = second Legendre polynomial; $^{(4)}$**P** = rank 4 orientation transformation tensor; $q_f(\tilde{\nu}_2)$ = fraction of absorption or emission at $\tilde{\nu}_2$ due to the $f$th transition; $r_u$ = fraction of absorption polarized along $u$; $R$ = degree of anisotropy; **R** = simplified form of orientation transformation tensor for Raman and two-photon absorption spectroscopy; $r_j(\tilde{\nu}_1)$ = fraction of absorption $\tilde{\nu}_1$ due to the $j$th transition; $s$, $t$, $u$, $v$ = indices taking values $x$, $y$, $z$ or $x'$, $y'$, $z'$; **S** = simplified form of orientation transformation tensor for fluorescence and transient dichroism spectroscopy; $S_{uv}$, $S_{stuv}$ = Saupe matrix elements; **T** = two-photon absorption tensor; $U$ $V$ = photon electric field vector direction (polarization); $x$, $y$, $z$ = molecular orientation axes; $x'$, $y'$, $z'$ = arbitrary molecular axes; $X$, $Y$, $Z$ = laboratory axes; $\alpha'$, $\beta'$, $\gamma'$ = Euler angles; $\bar{\alpha}^2$ = isotropic part of $\alpha'$; $\alpha'$ = Raman scattering tensor; $\gamma_{as}^2$ = symmetric anisotropy; $\gamma_{as}^2$ = antisymmetric anisotropy; $\delta_F$, $\delta_G$, $\delta_H$ = rotational invariants of the two-photon absorption tensor; $\delta_{ij}$ = Kronecker delta; $\boldsymbol{\varepsilon}^U$ = real unit vector along $U$; $\boldsymbol{\varepsilon}^+$, $\boldsymbol{\varepsilon}^-$ = polarization vector for left-, right-handed light; $\eta_{x'}$, $\eta_{y'}$, $\eta_{z'}$ = angles of $x'$, $y'$, $z'$ with $Z$; $\tilde{\nu}$ = wavenumber; $\tau$ = time after initial excitation event; $\tau_0$ = excited-state lifetime; $\tau_R$ = rotational relaxation time; $\phi$ = angle between $M(0j)$ and $M(j'f)$; $\phi_u f$ = direction cosines of $M(0f)$ in $x'$, $y'$, $z'$; $\Omega'$ = $\alpha'$, $\beta'$, $\gamma'$.

*See also:* **Chiroptical Spectroscopy, General Theory; Electromagnetic Radiation; Fluorescence Polarization and Anisotropy; IR Spectroscopy, Theory; Linear Dichroism, Applications; Rayleigh Scattering and Raman Spectroscopy, Theory; Symmetry in Spectroscopy, Effects of.**

## Further reading

Clarke D and Grainger JF (1971) *Polarized Light and Optical Measurement*. Oxford: Pergamon Press, Oxford.

Dörr F (1971) In: Lamola AA (ed) *Creation and Detection of the Excited State*, Vol 1, Chapter 2, pp 53–122. New York: Marcel Dekker.

Krause S (ed) (1981) *Molecular Electro-Optics*. New York: Plenum Press.

Michl J and Thulstrup EW (1995) *Spectroscopy with Polarized Light*, 2nd edn. New York: VCH Publishers.
Thulstrup EW and Michl J (1989) *Elementary Polarization Spectroscopy*. New York: VCH Publishers.
Zannoni C (1979) In: Luckhurst GR and Gray GW (eds) *The Molecular Physics of Liquid Crystals*, Chapter 3. New York: Academic Press.

# ORD and Polarimetry Instruments

**Harry G Brittain**, Center for Pharmaceutical Physics, Milford, NJ, USA

Copyright © 1999 Academic Press

ELECTRONIC SPECTROSCOPY
**Methods & Instrumentation**

## Introduction

Chiral molecules interact with electromagnetic radiation in exactly the same fashion as do achiral molecules in that they will exhibit optical absorption, have a characteristic refractive index, and can scatter oncoming photons. Optically active compounds are also capable of an additional interaction with light whose electric vectors are circularly polarized. One particular manifestation of this property is an apparent rotation of the plane of linearly polarized light upon passage through a medium containing the optically active agent. When this property is measured at a single wavelength, the phenomenon is commonly referred to as polarimetry. The wavelength dependence of polarimetric response is termed optical rotatory dispersion (ORD). Charney, who has provided the most readable summary of the history and practice associated with chiroptical spectroscopy, has provided a general introduction to optical activity. A number of other monographs have been written which concern various applications of chiroptical spectroscopy, ranging from the very theoretical to the very practical, and these texts contain numerous references suitable for those unfamiliar with the field.

## Polarization properties of light

An understanding of the polarization properties of light is essential to polarimetry and optical rotatory dispersion, and considerable insight can be gained by considering the characteristics of electric vectors. Unpolarized light propagating along the $z$-axis will contain electric vectors whose directions span all possible angles within the $x$–$y$ plane. Linearly polarized light represents the situation where all the transverse electric vectors are constrained to vibrate in a single plane. The simplest way to produce linearly polarized light is by dichroism, where the polarization is obtained by passage of the incident light beam through a material that totally absorbs all electric vectors not lying along a particular plane. The other elements suitable for the production of linearly polarized light are crystalline materials that exhibit optical double refraction, and these are the Glan, Glan–Thompson, and Nicol prisms.

As with any vector quantity, the electric vector describing the polarization condition can be resolved into projections along the $x$ and $y$ axes. For linearly polarized light, these will always remain in phase during the propagation process unless passage through another anisotropic element takes place. Attempted passage of linearly polarized light through another polarizer (referred to as the analyzer) results in the transmission of only the vector component that lies along the axis of the second polarizer. If the incident angle of polarization is orthogonal to the axis of transmission of the analyzer, then no light will be transmitted.

Certain crystalline optical elements have the property of being able to alter the phase relationships existing between the electric vector projections. When the vector projections are rendered 90° out of phase, the electric vector executes a helical motion as it passes through space, and the light is now denoted as being circularly polarized. Since the helix can be either left- or right-handed, the light is referred to as being either left- or right-circularly polarized. It is preferable to consider linearly polarized light as being the resultant formed by combining equal amounts of left- and right-handed circularly polarized components, the electric vectors of which are always exactly in phase. When the phase angle

between the two vector components in any light beam lies between 0 and 90°, the light is denoted as being elliptically polarized. The production of a 90° phase shift is termed quarter wave retardation, and an optical element that effects such a change is a quarter wave-plate. Passage of linearly polarized light through a quarter wave place produces a beam of circularly polarized light, the sense of which depends on whether the phase angle has been either advanced or retarded by 90°. The passage of circularly polarized light through a quarter wave-plate will produce linearly polarized light, whose angle is rotated by 90° with respect to the original plane of linear polarization.

## Circular birefringence (optical rotation)

The study of molecular optical activity can be considered as beginning with the work of Biot, who demonstrated that the plane of linearly polarized light would be rotated upon passage through an optically active medium and designed a working polarimeter capable of quantitatively measuring the effect. Mitscherlich introduced the use of calcite prisms in 1844, and Soleil devised the double-field method of detection in 1845. Since these early developments, many advances have been made in polarimetry and a large number of detection schemes are now possible. An extensive summary of methods is available in Heller's comprehensive review.

The phenomenon of optical activity is determined by the relative indices of refraction for left- and right-circularly polarized light within the medium under consideration. In an optically inactive medium, the refractive indices of left- and right-circularly polarized light are equal. Since the two polarization senses would remain in phase at all times during passage through the medium, the resultant vector leaving the medium would be unchanged with respect to the vector that entered the medium. In an optically active medium, the refractive indices for left- and right-circularly polarized light are no longer equal, so the components are no longer coherent when they leave the medium. These phenomena are illustrated in **Figure 1**. When viewed directly along the direction of the oncoming beam, the resultant vector appears as a rotation of the initial plane of polarization (see **Figure 2**). Optical activity is therefore a manifestation of circular birefringence.

In principle, the measurement of optical rotation is extremely simple, and a suitable apparatus is shown in schematic form in **Figure 3**. The incident light is collimated and plane-polarized, and allowed to pass

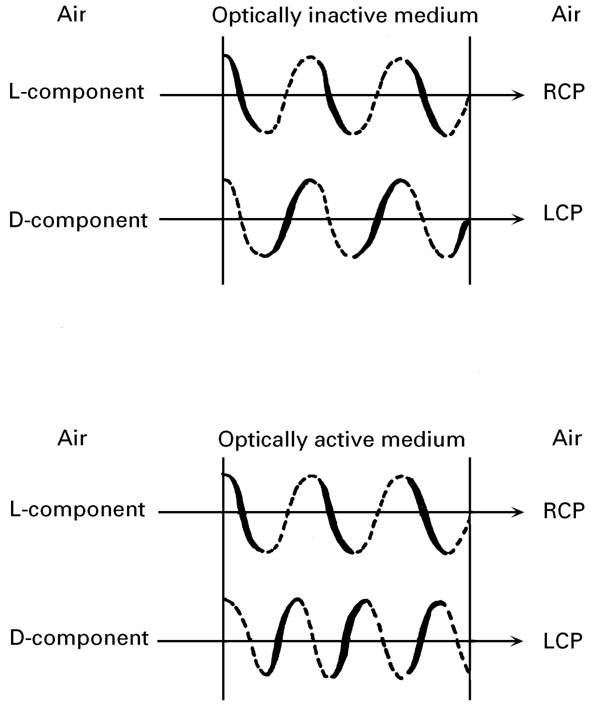



**Figure 1** Behaviour of the right-circularly polarized (RCP) and left-circularly polarized (LCP) components of linearly polarized light as they pass through an optically inactive medium (upper diagram) or thorough an optically active medium (lower diagram).

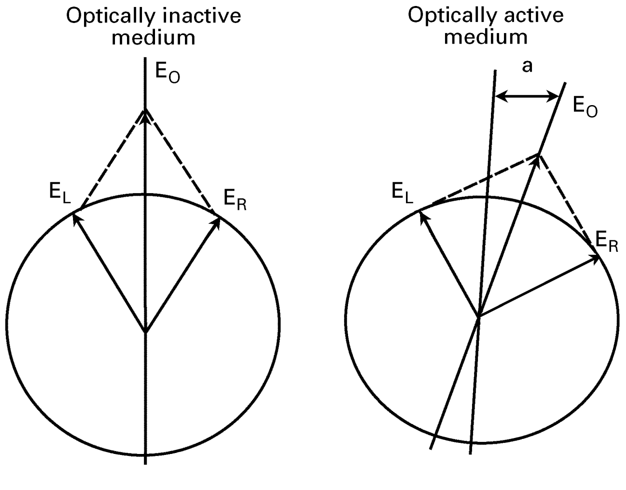



**Figure 2** Phase relations associated with the electric vectors of linearly polarized light as it is passed through various media. For optically inactive media, recombination of left-circularly polarized ($E_L$) and right-circularly polarized ($E_R$) components yields linearly polarized light whose resultant electric vector ($E_O$) is unchanged with respect to the incident axis. For optically active media, recombination of the $E_L$ and $E_R$ components yields a resultant $E_O$ vector rotated from this incident axis by the angle α.

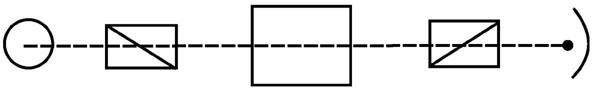



**Figure 3** Block diagram of a simple polarimeter. Monochromatic light from the source is linearly polarized by the initial polarizer, and then allowed to pass through the sample medium. The angle of polarization associated with the light leaving the medium is determined by rotating the analyzer polarizer to the new null position. In automatic operation, the observed angle of rotation is determined using a photoelectric determination of the null points.

through the medium of interest. In most common measurements, the medium consists of the analyte dissolved in an appropriate solvent. The plane of the incident light is specified, and then the angle of rotation is defined with respect to this original plane. This is carried out by first determining the orientation of polarizer and analyzer for which no light can be transmitted (the null position). The medium containing the optically active material is then introduced between the prisms, and the analyzer is rotated until a null position is again detected. The observed angle of rotation is taken as the difference between the two null angles.

The measurement of the null angle as a minimum in the transmitted light is experimentally difficult to observe when conducted visually. For this reason, a more efficient detection mechanism was developed, which is commonly known as the ‘half-shade’ technique. The half-shade is a device that transforms the total extinction point into an equal illumination of two adjacent fields. This mode of detection was superior for manual polarimeters, since the human eye is much more adept at balancing fields of transmitted light than at detecting a minimum in light intensity. In most cases, setting an appropriate device in front of a simple analyzing prism, such as a Lippich prism, brings about the half-shade effect.

With the development of photoelectric devices, the manual detection of null positions in polarimetry became superseded by instrumental measurements of the endpoint. As would be anticipated, measurements can be made far more easily and accurately using photoelectric detection of the null position. Early versions of automated polarimeters used the half-shade method, but the two light intensities were measured using photomultiplier tubes. The position of the analyzer was rotated until the difference in signals detected by the two detectors reached a minimum. Other methods have made use of modulated light beams, and variations on the method of symmetric angles. A wide variety of detection methods have been developed, enabling accurate measurements of optical rotation to be made on a routine basis. The influence of polarimeter design on the observed signal-to-noise characteristics has been discussed by a number of investigators.

The velocity of light ($v$) passing through a medium is determined by the index of refraction ($n$) of that medium:

$$v = c/n \qquad [1]$$

where $c$ equals to the velocity of light in vacuum. For a non-chiral medium, the refractive index will not exhibit a dependence on the sense of the polarization state of the light. For a chiral medium, the refractive index associated with left-circularly polarized light will not normally equal the refractive index associated with right-circularly polarized light. It follows that the velocities of left- and right-circularly polarized light will differ on passage through a chiral medium. Since linearly polarized light can be resolved into two in-phase, oppositely-signed, circularly polarized components, the components will not longer be in phase once they pass through the chiral medium. Upon leaving the chiral medium, the components are recombined, and linearly polarized light is obtained whose plane is rotated (relative to the original plane) by an angle equal to half the phase angle difference of the circular components. This phase angle difference is given by:

$$\beta = \frac{2\pi b'}{\lambda_o}(n_L - n_R) \qquad [2]$$

In Equation [2], $\beta$ is the phase difference, $b'$ is the medium path length (in cm) $\lambda_o$ is the vacuum wavelength of the light used (in cm), and $n_L$ and $n_R$ are the refractive indices for left- and right-circularly polarized light, respectively. The quantity ($n_L$–$n_R$) defines the circular birefringence of the chiral medium, and is the origin of what is commonly referred to as optical rotation. The observed rotation of the plane-polarized light is given in radians by:

$$\phi = \frac{\beta}{2} \qquad [3]$$

The usual practice is to express rotation in terms of degrees, and in that case Equation [2] becomes:

$$\alpha = \frac{1800\, b}{\lambda_o}(n_L - n_R) \qquad [4]$$

In Equation [4], $b$ represents the medium path length in *decimeters*, which is the conventional unit. The optical rotation (α) of a chiral medium can be either positive or negative depending on the sign of the circular birefringence. It is not practical to measure the circular birefringence directly, since the magnitude of ($n_L - n_R$) exhibits typical magnitudes of between $10^{-8}$ and $10^{-9}$.

The optical rotation exhibited by a chiral medium depends on the optical path length, the wavelength of the light used, the temperature of the system, and the concentration of dissymmetric analyte molecules. If the solute concentration ($c$) is given in terms of grams per 100 mL of solution, then the observed rotation (an extrinsic quantity) can be converted into the specific rotation (an intrinsic quantity) using:

$$[\alpha] = \frac{100\ \alpha}{b\ c} \quad [5]$$

The molar rotation [$M$], is defined from:

$$[M] = \frac{(\mathrm{FW})\ \alpha}{b\ c} \quad [6]$$

$$= \frac{(\mathrm{FW})\ [\alpha]}{100} \quad [7]$$

where FW is the formula weight of the dissymmetric solute. When the solute concentration is given in units of molarity $C$, Equation [6] becomes:

$$[M] = \frac{10\ \alpha}{b\ C} \quad [8]$$

The temperature associated with a measurement of the specific or molar rotation of a given substance must be specified. Thermal volume changes or alterations in molecular structure (as induced by a temperature change) are capable of producing detectable changes in the observed rotations. In the situations where solute–solute interactions become important at high concentrations, it may be observed that the specific rotation of a solute is not independent of concentration. It is acceptable practice, therefore, to obtain polarimetry data at a variety of concentration values to verify that a true molecular parameter has been measured.

The verification of polarimeter accuracy is often not addressed on a routine basis. The easiest method to verify the performance of a polarimeter is to use quartz plates that have been cut to known degrees of retardation. These plates are normally certified to yield a specified optical rotation at a specific wavelength. A reasonable criterion for acceptability is that the observed optical rotation should be within ±0.5% of the certified value. The quartz plates are available from a variety of sources, including manufacturers of commercial polarimeters or houses specializing in optical components.

Another possibility of verify the accuracy of polarimetry measurements is to measure the optical rotation of a known compound. Owing to the stability of their rotatory strengths once dissolved in a fluid medium, steroids are probably most suitable for this purpose. Chafetz has provided a compilation of the specific rotation values obtained for a very extensive list of steroids. When possible, data have been reported for alternate solvents, as well as for wavelengths other than 589 nm.

Optical rotation measurements are most commonly used to confirm the enantiomeric identities of resolved enantiomers. When reference standards of totally resolved materials are available, polarimetry can be used to determine the enantiomeric purity of samples of defined composition.

## Optical rotatory dispersion

Aside from the intrinsic molecular contribution, the most important parameter in determining the magnitude of optical rotation is the wavelength of the light used for the determination. Generally, the magnitude of the circular birefringence increases as the wavelength becomes shorter, so specific rotations increase in a regular manner with a decrease in wavelength. This behaviour persists until the light is capable of being absorbed by the chiral substance, whereupon the refractive index exhibits anomalous behaviour. The variation of specific or molar rotation with wavelength is termed optical rotatory dispersion (ORD).

Biot observed that the optical rotation of tartaric acid solutions was a function of the wavelength used for the determination. When measured outside of absorption bands, an ORD spectrum (as illustrated in one half of **Figure 4**) will consist either of a plain positive or plain negative dispersion curve. When measured inside of an absorption band, the ORD will exhibit anomalous dispersion, which is referred to as a Cotton effect. A positive Cotton effect consists of positive ORD at long wavelengths, and negative ORD at shorter wavelengths (see **Figure 4**). In the simplest measurement of ORD, the fixed wavelength source of **Figure 3** is replaced by a tunable source (such as a xenon arc combined with a

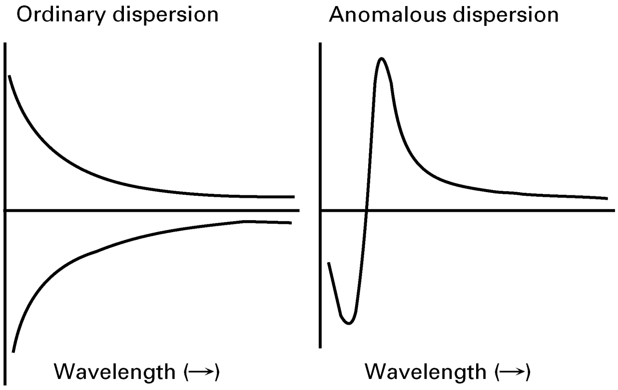



**Figure 4** Optical rotatory curves, as would be obtained for simple dispersion (both positive and negative curves are shown) and anomalous dispersion (illustrating a positive Cotton effect).

monochromator), and the angle of rotation is automatically determined as the wavelength is swept.

The anomalous dispersion observed in ORD spectra arises since the refractive index of material is actually the sum of a real and imaginary part:

$$n = n_o + \mathrm{i}\,k \qquad [9]$$

where $n$ is the observed refractive index at some wavelength, $n_o$ is the refractive index at infinite wavelength, and $k$ is the absorption coefficient of the substance. It is evident that if $(n_L - n_R)$ does not equal zero, then $(k_L - k_R)$ will not equal zero either. The relation between the various quantities was first conceived by Cotton, and is illustrated schematically in **Figure 5**. The effect of the differential absorption $(k_L - k_R)$ (i.e. circular dichroism) is to render the incident linearly polarized light into an emergent beam that is elliptically polarized, for which the major axis of the ellipse will be rotated with respect to the incident vector.

The earliest ORD spectrometer was the non-recording instrument designed by Rudolph, consisting merely of a manual polarimeter through which light of various wavelengths could be passed. The optical rotation was measured at each incident wavelength, and the ORD spectrum obtained by plotting the observed values. In an improvement, the incident polarizer was rocked through a small angle by an electric motor at low frequency, the transmitted light measured by a photoelectric cell, and the analyzer rotated until a null position was again achieved.

Recording ORD spectrometers have been developed which either use the Faraday effect to obtain both modulation and balance, or use servo-driven analyzers to determine the null position. A block diagram illustrating the latter mode of operation is given in **Figure 6**. Comparison of **Figure 3** and **Figure 6** reveals that the only real difference between an ordinary polarimeter and such an ORD spectrometer is the additional servo-mechanisms and associated feedback circuits which automatically and continuously determine the null positions as the incident wavelength is changed. Currently available ORD instrumentation is, however, invariably based on adapted circular dichroism spectrometers with their enhanced performance through phase modulation. The calibration of ORD spectrometers is verified using the same standards as used for the calibration of ordinary polarimeters.

The earliest investigative work involving chiral organic molecules was entirely based on ORD methods, since little else was available at the time. One of

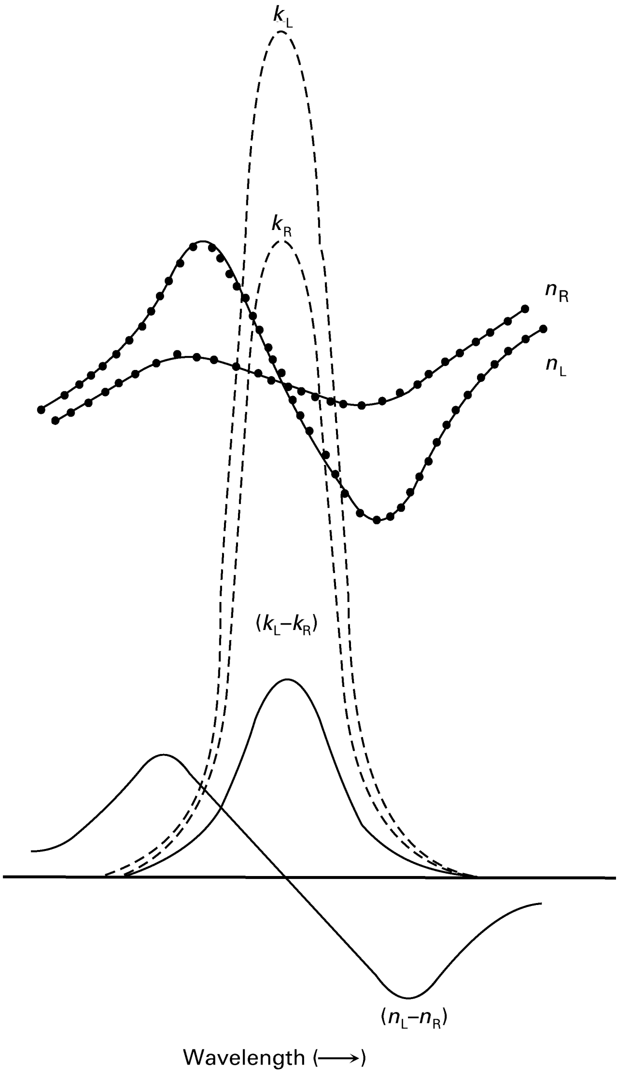



**Figure 5** Schematic diagram of the Cotton effect, illustrating the effects of circular birefringence and circular dichroism within an isolated absorption band.

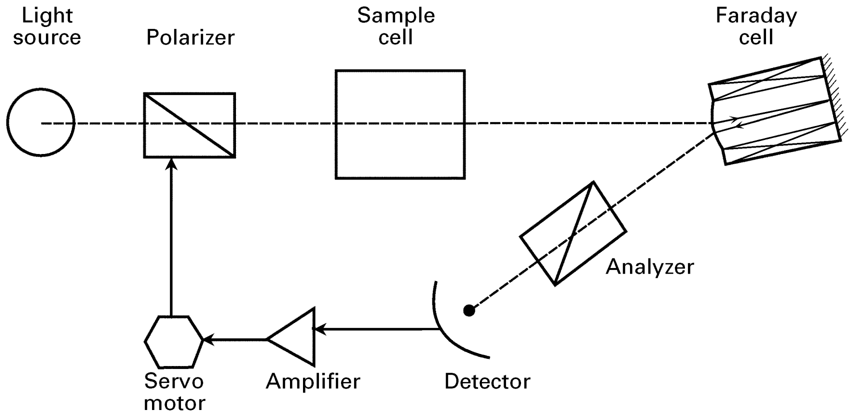



**Figure 6** Block diagram of a servo-driven ORD spectropolarimeter. Monochromatic light from the source is linearly polarized by the initial polarizer, and then allowed to pass through the sample medium. Further modulation is effected by the Faraday cell, and then the angle of polarization associated with the light leaving the medium is determined by rotating the analyzer polarizer to the new null position. The observed angle of rotation is determined using a photoelectric determination of the null points.

the largest data sets collected to date concerns the chirality of ketone and aldehyde groups, which eventually resulted in the deduction of the octant rule. The octant rule was an attempt to relate the absolute stereochemistry within the immediate environment of the chromophore with the sign and intensity of the ORD Cotton effects. To apply the rule, the circular dichroism and/or ORD within the n–π* transition and 300 nm is obtained, and its sign and intensity noted. The rule developed by Djerassi and co-workers states that the three nodal planes of the n- and π*-orbitals of the carbonyl group divide the molecular environment into four front octants and four back octants. A group or atom situated in the upper-left or lower-right rear octant (relative to an observer looking at the molecule parallel to the C=O axis) induces a positive Cotton effect in the n–π* band. A negative Cotton effect would be produced by substitution within the upper-right or lower-left back octant. This has been illustrated in **Figure 7** for the particular example of a generic 3-hydroxy-3-alkyl-cyclohexanone.

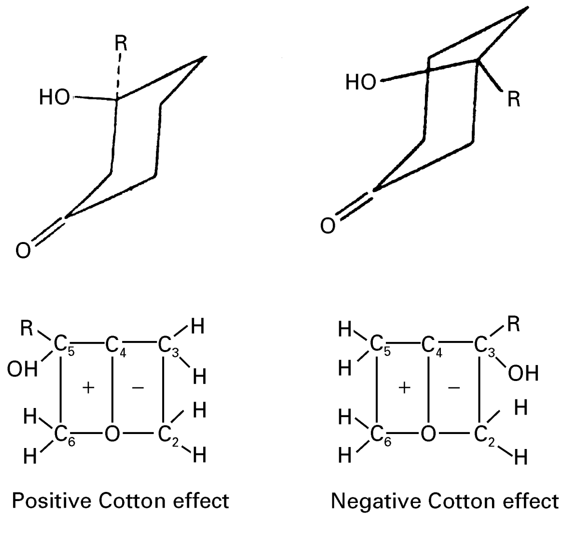



**Figure 7** Application of the octant rule, applied to the enantiomers of 3-hydroxy-3-alkyl-cyclohexanone.

Although exceptions to the octant rule have been shown, the wide applicability of the octant rule has remained established. The ability to deduce molecular confirmations in solution on the basis of ORD spectra data has proven to be extremely valuable to synthetic and physical organic chemists, and enabled investigators of the time to develop their work without requiring the use of more heroic methods.

*See also:* **Biomacromolecular Applications of Circular Dichroism and ORD; Chiroptical Spectroscopy, Emission Theory; Chiroptical Spectroscopy, General Theory; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism; Light Sources and Optics; Luminescence, Theory; Nonlinear Optical Properties; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## List of symbols

$b$ = path length in decimetres; $b'$ = pathlength in cm; $c$ = velocity of light *in vacuo*; $c$ = concentration in grams per 100 mL; C = molarity concentration; $E_L$ = left circularly polarized electric vector; $E_R$ = right circularly polarized electric vector; $E_0$ = resultant electric vector; FW = formula weight; $k_L/k_R$ = absorption coefficient for left/right circularly polarized light;

$k$ = absorption coefficient; $[M]$ = molar radiation; $n$ = index of refraction; $n_L/n_R$ = index of refraction for left/right polarized light; $n_0$ = refractive index at infinite wavelength; $\alpha$ = optical rotation in degrees; $[\alpha]$ = specific rotation; $\beta$ = phase angle of difference; $\phi$ = optical rotation in radians; $\lambda_0$ = vacuum wavelength of light; $v$ = velocity of light in medium.

## Further reading

Barron L (1982) *Molecular Light Scattering and Optical Activity*. Cambridge: Cambridge University Press.

Caldwell DJ and Eyring H (1971) *The Theory of Optical Activity*. New York: Wiley-Interscience.

Chafetz L (1993) Implementing changes in polarimetry. *Pharmacopeial Forum* **19**: 6159–6162.

Charney E (1979) *The Molecular Basis of Optical Activity*. New York: Wiley.

Crabbe P (1972) *ORD and CD in Chemistry and Biochemistry*. New York. Academic Press.

Djerassi C (1960) *Optical Rotatory Dispersion: Applications to Organic Chemistry*. New York: McGraw Hill.

Heller W (1972) Optical Rotation - Experimental Techniques and Physical Optics. In: Weissberger A and Rossiter BW (eds) *Physical Methods of Chemistry*, Vol I, Chapter 2. New York: John Wiley.

Kankare J and Stephens R (1986) The influence of optical design on the signal-to-noise characteristics of polarimeters. *Talanta* **33**: 571–576.

Kirk DN (1986) The chiroptical properties of carbonyl compounds. *Tetrahedron* **42**: 777–818.

Mason S (1982) *Molecular Optical Activity and the Chiral Discrimination*. Cambridge: Cambridge University Press.

Moffitt W, Woodward RB, Moscowitz A, Klyne W and Djerassi C (1961) Structure and the optical rotatory dispersion of saturated ketones. *Journal of the American Chemical Society* **83**: 4013–4018.

Purdie N and Brittain HG (1994) *Analytical Applications of Circular Dichroism*. Amsterdam: Elsevier Science.

Snatzke G (ed) (1967) *Optical Rotatory Dispersion and Circular Dichroism in Organic Chemistry*. London: Heyden.

Velluz L, Legrand M and Grosjean M (1965) *Optical Circular Dichroism: Principles, Measurements, and Applications*. Weinheim: Verlag Chemie.

Viogtman E (1992) Effect of source 1/f noise on optical polarimeter performance. *Analytical Chemistry* **64**: 2590–2595.

Yeung ES (1985) Signal-to-noise optimization in polarimetry. *Talanta* **33**: 1097–1100.

# ORD Spectroscopy of Biomacromolecules

*See* **Biomacromolecular Applications of Circular Dichroism and ORD.**

# Organic Chemistry Applications of Fluorescence Spectroscopy

**Stephen G Schulman**, **Qiao Qing Di** and **John Juchum**, University of Florida, Gainesville, FL, USA

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| Applications |

## Introduction

The fluorescence of organic molecules consists of the emission of light by molecules, which have previously absorbed visible or ultraviolet radiation. The measurement of fluorescence often permits very low analyte detection limits ($10^{-15}$–$10^{-4}$ mol dm$^{-3}$) and is widely employed in quantitative analysis, especially as a detection and quantitation method in liquid chromatography. Most applications are derived from the relationship between analyte concentration and fluorescence intensity and are, therefore, similar in

concept to other spectrochemical methods of analysis. However, as well as spectral intensity, other features of fluorescence spectral bands of organic molecules, such as position in the electromagnetic spectrum, emission lifetime and excitation spectrum are exquisitely sensitive to the molecular environment and to molecular structure and, therefore, are also analytically useful, especially for probing the environment of the fluorophore.

This article will deal with the nature of organic molecular fluorescence, its dependence upon molecular structure, reactivity and interactions with the environment and its utility in the trace analysis of organic compounds.

## The origin of molecular fluorescence

The electronic excitation of molecules occurs as the result of the absorption of near ultraviolet or visible light.

Subsequent to excitation, the loss of excess vibrational energy, known as vibrational relaxation, takes place in about $10^{-12}$ s, the excess energy being lost to inelastic collisions with solvent molecules.

There is also an efficient radiationless pathway for the demotion of the excited molecule from higher to lower electronically excited states called internal conversion. In aliphatic molecules which have a high degree of vibrational freedom, vibrational relaxation and internal conversion may return the excited molecule to the ground electronic state radiationlessly within $10^{-12}$ s after excitation, in which case fluorescence is not observed.

However, in aromatic molecules, the degree of vibrational freedom is restricted. In this case, the excited molecule may, radiationlessly, arrive in the lowest vibrational level of the lowest electronically excited singlet state. Subsequently (after $10^{-11}$–$10^{-7}$ s), it may return to the ground electronic state by emitting near ultraviolet or visible fluorescence.

With very few exceptions, fluorescence always originates from the lowest excited singlet state. This means that only one fluorescence band may be observed from a given molecule, even though it will usually have several absorption bands. Therefore, the observation of several fluorescence bands in a solution of pure sample, suggests the occurrence of a chemical reaction in either the ground or excited state, resulting in two or more fluorescent species. Alternatively, the purity of the sample must be questioned. Because the fluorescence spectrum of any one organic compound can demonstrate only one fluorescence band, a band which is usually broad and lacking features, the fluorescence spectrum does not reveal the detailed information about molecular structure that NMR, IR or mass spectrometry does. Nevertheless, fluorescence spectra can give information about the molecular environment that is unobtainable by other methods.

## Chemical structural effects upon fluorescence

Most fluorescence spectra arise from functionally substituted aromatic molecules. Consequently, the compounds of interest in this article are those derived from organic compounds that possess aromatic rings, such as benzene, naphthalene or anthracene, or their heteroaromatic analogues pyridine, quinoline, acridine, etc. The fluorescence spectra of these substances may often be understood in terms of the electronic interactions between the simple aromatic structures and their substituents.

### Chemical structure and fluorescence intensity

The intensity of fluorescence observable from a given molecular species depends upon the probability of light absorption and the probability of fluorescence. The molar absorptivity is a macroscopic manifestation of the probability of light absorption by molecules in the optical path. For most aromatic molecules, the $\pi$, $\pi^*$ absorption bands lying in the near-ultraviolet and visible regions of the spectrum have molar absorptivities of $1 \times 10^3$–$1 \times 10^5$ $dm^3\ mol^{-1}\ cm^{-1}$ so that the appropriate choice of the transition to excite can influence the intensity of fluorescence by about two orders of magnitude. If the absorbance of the potentially fluorescing species, at the wavelength of excitation, is below 0.02 the intensity of fluorescence is proportional to the molar absorptivity at that nominal wavelength. At high absorbances the sample is not equally illuminated along the optical path, giving rise to ‘self-shadowing’ effects.

More important is the quantum yield or efficiency of fluorescence, which may affect the intensity of fluorescence over about four orders of magnitude and may determine whether fluorescence is at all observable. The quantum yield of fluorescence is dependent upon the rates of processes competing with fluorescence for the deactivation of the lowest excited singlet state.

Aromatic molecules that contain lengthy aliphatic side-chains generally tend to fluoresce less intensely than those without the side-chains because of greater opportunity for vibrational deactivation (the ‘loose-bolt’ effect). In unsubstituted aromatic molecules, the rigidity of the aromatic ring results in lower probabilities of vibrational deactivation and hence, higher quantum yields.

The fluorescence of organic molecules is quenched (diminished in intensity) by heavy atom substituents such as $–As(OH)_2$, Br and I and by certain other groups such as –CHO, $–NO_2$ and nitrogen in six-membered heterocyclic rings (e.g. quinoline). These substituents cause mixing of the spin and orbital motions of the valence electrons. Spin–orbital coupling obscures the distinct identities of the singlet and triplet states and, thereby, enhances the probability or rate of singlet $\rightarrow$ triplet intersystem crossing. This process favours population of the lowest triplet state at the expense of the lowest excited singlet state and thus decreases the fluorescence quantum yield. Consequently, aromatic arsenites, nitro compounds, bromo and iodo derivatives, aldehydes, ketones and N-heterocyclics tend to fluoresce very weakly or not at all.

### Chemical structure and position of the fluorescence

The energies of the ground and excited states of fluorescing molecules are affected by molecular structure. This is reflected in the positions of the fluorescence maxima in the spectrum.

$$E = h\nu = hc/\lambda \qquad [1]$$

According to Equation [1], where $E$ is the energy, $\nu$ the frequency and $\lambda$, the wavelength of fluorescence and $h$ and $c$ are, respectively, Planck's constant and the velocity of light, the greater the separation between the ground and excited states the greater will be the frequency and the shorter will be the wavelength of fluorescence. This separation depends upon the energy difference between the highest occupied and lowest unoccupied molecular orbitals and the repulsion energy between the electronic configurations corresponding to the ground and excited states. In aromatic hydrocarbons, the greater the degree of linear annulation the closer together will be the highest occupied and lowest unoccupied orbitals. Consequently, benzene fluoresces at shorter wavelengths than naphthalene which fluoresces at shorter wavelengths than anthracene. Phenanthrene, which is angularly annulated, emits at wavelengths between those of naphthalene and anthracene. In functionally substituted aromatic molecules the substituents with lone electron pairs, e.g. $–NH_2$, –OH, will have highest occupied orbitals, higher in energy than those of the unsubstituted hydrocarbons while the substituents with vacant $\pi^*$ orbitals, e.g. –CHO, $–CO_2H$, will have $\pi^*$ orbitals lower in energy than those of the unsubstituted hydrocarbons. This means that in substituted aromatic molecules fluorescence will be at wavelengths longer than in the unsubstituted molecules. This is so regardless of whether the substituents are electron donating or electron withdrawing.

## Influence of the chemical and physical environment on fluorescence spectra

### The solvent

The solvents in which fluorescence spectra are observed play a role secondary only to molecular structure in determining the spectral positions and intensities with which fluorescence bands occur and occasionally determine whether fluorescence is observed.

The electronic transition accompanying excitation entails a change in electronic charge distribution. If the excited state is more polar than the ground state a more polar solvent will stabilize the excited state more than the ground state and cause the fluorescence to shift to longer wavelengths relative to that observed in a less polar solvent. If, however, the ground state is more polar than the excited state, which is rarely the case, the fluorescence will tend to shift to shorter wavelengths upon going to a more polar solvent.

Hydrogen bonding in the lowest excited singlet states occasionally results in a decrease in fluorescence quantum yield upon going from hydrocarbon to hydrogen bonding solvents. Many arylamines and phenolic compounds demonstrate this behaviour which is due to internal conversion enhanced by coupling of the vibrations of the molecule of interest to those of the solvent.

Molecules having the lowest excited singlet states of the $n,\pi^*$ type rarely fluoresce in hydrocarbon solvents because the $n,\pi^*$ singlet state is efficiently deactivated by intersystem crossing. However, in polar, hydrogen bonding solvents, such as ethanol or water, these molecules often become fluorescent. This results from the stabilization of the lowest singlet $\pi,\pi^*$ state relative to the lowest $n,\pi^*$ state by hydrogen bonded interaction. If the interaction is sufficiently strong, the lowest $\pi,\pi^*$ state drops below the $n,\pi^*$ state in the strongly solvated molecule, becoming the lowest excited singlet state and permitting intense fluorescence. Quinoline and 1-naphthaldehyde, for example, do not fluoresce in cyclohexane but do so in water.

### The influence of pH

The spectral shifts accompanying protonation or dissociation of basic or acidic functional groups depend upon whether the functional group undergoing

protonation or dissociation is directly coupled to the aromatic system and whether it gains or loses electronic charge upon going from the ground to the excited state. The higher the positive charge on electron-attracting groups and the higher the negative charge on electron-donating groups, the lower, in general, will be the energy of fluorescence. Thus, the protonation of electron-withdrawing groups, such as carboxyl, carbonyl and pyridinic nitrogen, causes shifts of fluorescence spectra to longer wavelengths while the protonation of electron-donating groups, such as the amino group, produces spectral shifts to shorter wavelengths. The protolytic dissociation of electron-donating groups, such as hydroxyl, sulfhydryl or pyrrolic nitrogen, produces spectral shifts to longer wavelengths while the dissociation of electron-withdrawing groups, such as carboxyls, shifts the fluorescence spectra to shorter wavelengths.

In some molecules, the presence of nonbonded electrons obviates the occurrence of fluorescence. However, protonation of the functional group possessing the non-bonded electron pair raises the $n,\pi^*$ lowest excited singlet state above the lowest excited singlet $\pi,\pi^*$ state and thereby allows fluorescence to occur. Benzophenone, for example, does not fluoresce as the neutral molecule but does so, moderately intensely, as the cation in concentrated sulfuric acid.

An interesting aspect of acid–base reactivity of fluorescent molecules is derived from the occurrence of protonation, and dissociation, during the lifetime of the lowest excited singlet state and is occasionally observed in the pH dependence of the fluorescence spectrum.

The lifetimes of molecules in the lowest excited singlet state are typically $10^{-10}$–$10^{-7}$ s. Typical rates of proton transfer reactions are $\leq 10^{10}\ s^{-1}$. Consequently, excited state proton transfer may be much slower, much faster, or competitive with radiative deactivation of the excited molecules. If excited state proton transfer is much slower than fluorescence, the relative fluorescence intensity will vary with pH exactly the same way as does the absorbance, reflecting only the ground-state acid–base equilibrium. If excited state proton transfer is much faster than fluorescence, the fluorescence intensity will vary with pH in a way that reflects the acid–base equilibrium in the lowest excited singlet state. Equilibrium in the excited state is a rare phenomenon and will not be dealt with further here.

If the rate of proton transfer, in the excited state, is comparable to the rates of photophysical deactivation of excited acid and conjugate base, the variations of the fluorescence intensities of acid and conjugate base, with pH, will be governed by the kinetics of the excited state proton transfer reactions and the fluorescence of acid and conjugate base will be observed over a wide pH range.

### The influence of high solute concentrations

Several types of excited state solute–solute interaction are common at high solute concentrations. The aggregation of excited solute molecules with unexcited molecules of the same type may result in a new excited molecule called an excimer, which may either not luminesce or may luminesce at lower frequency than the monomeric excited molecule. Because excimer formation takes place in the excited state it is sometimes demonstrable as a shifting of the fluorescence spectrum. However, after fluorescence, the deactivated polymer, which is unstable in the ground state, rapidly decomposes. Hence, the absorption spectrum does not reflect the presence of the excited state complexes. Occasionally, excited state complex formation may occur between two different solute molecules. The term 'exciplex' has been coined to describe a heteropolymeric excited state complex. Excimer and exciplex formation are usually observed only in fluid solution because diffusion of the excited species is necessary to form the excited complexes. One concentration effect that is observed in molecules in fluid or rigid media is resonance energy transfer. Energy transfer entails the excitation of a molecule which, during the lifetime of the excited state, transmits its excitation energy to a nearby molecule. The probability of resonance energy transfer decreases as the inverse sixth power of the distance between donor and acceptor and can occur between molecules which are separated by up to 100 nm. Because the mean distance between molecules decreases with increasing concentration, energy transfer is favoured by increasing the concentration of the acceptor. For energy transfer to occur between two dissimilar molecules, the fluorescence spectrum of the energy donor must overlap the absorption spectrum of the energy acceptor.

Fluorescence may be diminished in intensity or eliminated due to the deactivation of the lowest excited singlet state of the analyte by interaction with other species in solution. This is called quenching of fluorescence. Mechanisms of quenching appear to entail internal conversion, intersystem crossing, electron-transfer and photodissociation as modes of deactivation of the excited fluorescer–quencher complexes.

Quenching processes may be divided into two broad categories. In dynamic or diffusional quenching, interaction between the quencher and the

potentially fluorescent molecule takes place during the lifetime of the excited state. As a result, the efficiency of dynamic quenching is governed by the rate constant of the quenching reaction, which is usually typical of that for a diffusion-controlled reaction, the lifetime of the excited state of the potential fluorescer and the concentration of the quenching species. Interaction between quencher and fluorophore results in the formation of a transient excited complex which is non-fluorescent and may be deactivated by any of the usual radiationless modes of deactivation of excited singlet states. Because interaction occurs only after excitation of the potentially fluorescing molecule, the presence of the quenching species has no effect on the absorption spectrum of the fluorophore. Many aromatic molecules, for example, the quinolines and acridines, are dynamically quenched by halide ions such as $Cl^-$, $Br^-$ and $I^-$. Static quenching is characterized by complexation in the ground state between the quenching species and the molecule which, when excited alone, should fluoresce. The complex is generally not fluorescent and, as a result, the ground state reaction diminishes the intensity of fluorescence of the potentially fluorescent species. The quenching of the fluorescence of *o*-phenanthroline by complexation with iron(II) is an example of static quenching.

## Applications

### Native fluorescence of organic compounds

Numerous organic compounds are intrinsically fluorescent and so may be assayed directly, eliminating any need of derivatization or labelling. This section lists some examples of types of compounds which possess native fluorescence and some factors which exert influence on that fluorescence. The assay of organic compounds by fluorescence spectroscopy is covered in greater detail by Guilbault and that of organic natural products by Wolfbeis. Simple fluorescence spectra of 2000 compounds have been published by Sadtler Research Laboratories.

Three of the twenty common amino acids exhibit fluorescence. These are phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp). Phenylalanine, as the name implies, consists of alanine with a benzene ring attached. It is weakly fluorescent at $\lambda_f = 282$ nm ($\lambda_{ex} = 260$ nm) and cannot be detected in the presence of Tyr or Trp. Tyrosine is phenolic and fluoresces ($\lambda_{ex} = 275$ nm; $\lambda_f = 303$ nm) with much greater intensity than Phe just as phenol fluoresces far more intensely than benzene. The phenolic group, which ionizes at pH above 10, introduces the need to control pH in the assay. The phenolate form of Tyr fluoresces with far less intensity than the nonionized form and redshifts to 345 nm. Tryptophan has a quantum yield very close to that of Tyr but absorbs far better and so fluoresces ($\lambda_{ex} = 287$ nm; $\lambda_f = 348$ nm) more intensely. The nitrogen-containing indole moiety in Trp becomes protonated at low pH thus decreasing its fluorescence. The fluorescence of Trp is also quenched at high pH and so Trp fluoresces best over a range in pH from 4 to 9. The fluorescence properties of these amino acids as mentioned apply only to the free amino acids in solution. Such properties can change significantly when they become part of a protein.

The indole moiety seen in tryptophan is a common structure in nature and is seen in abundance in alkaloids. Indole in water fluoresces at $\lambda_f = 352$ nm ($\lambda_{ex} = 275$ nm). Many simple natural derivatives of indole fluoresce in the 330–350 nm range and can be maximally exited at 270–290 nm. The ‘indole alkaloid’ is a major class of alkaloids, which include a number of well-known drugs, legal as well as illicit. The ergot alkaloids are in this class. The famous hallucinogen lysergic acid diethylamide (LSD) at pH 7 fluoresces at 365 nm with an excitation maximum at 325 nm. Bromolysergic acid diethylamide is not hallucinogenic. It is brominated at the 2 position on the indole moiety. It fluoresces ($\lambda_{ex} = 315$ nm; $\lambda_f = 460$ nm) at pH 1 with far less intensity than LSD.

Other classes of alkaloids, which exhibit fluorescence, include the quinoline and isoquinoline alkaloids. Quinoline is weakly fluorescent. The antimalarial drug quinine includes a methoxy substituent on the 6 position of the quinoline moiety and fluoresces very intensely. Quinine, in sulfuric acid solution, is often used as a standard in fluorescence spectroscopy for determining a quantum yield. Its fluorescent properties are sensitive to pH. At pH 2 it has an excitation maximum of 347 nm with fluorescence at 448 nm. At pH 7 the peaks shift to absorb at 331 nm and emit at 382 nm. In hydrochloric acid solution, absorption is unaffected but fluorescence intensity is quenched greatly by the halide anions.

Isoquinoline alkaloids can be subdivided into two groups. Those alkaloids which preserve the isoquinoline moiety and those which have a reduced ring structure, the tetrahydroisoquinolines. Papaverine, a smooth muscle relaxant, possesses the isoquinoline moiety. In chloroform, papaverine fluoresces at $\lambda_f = 347$ nm ($\lambda_{ex} = 315$ nm). The addition of some trichloroacetic acid to this protonates the isoquinoline moiety and causes a shift in absorption and fluorescence peaks ($\lambda_{ex} = 415$ nm; $\lambda_f = 452$ nm). Berberine is an isoquinoline alkaloid which has a quaternary nitrogen with hydroxide as the standard counterion. In ethanol it has excitation maxima of

432 and 352 nm and fluoresces at 548 nm. A change of solvent to DMF causes a shift of the excitation maximum to 380 nm and emission to 510 nm. Berberine deposited on a TLC plate shows excitation maxima of 433 and 353 nm and fluoresces at 510 nm. 13-Methoxyberberine in ethanol has one excitation peak, at 433 nm, and its emission peak shifts to 562 nm.

Tetrahydroisoquinoline alkaloids have a benzenoid fluorophore rather than isoquinoline and so fluoresce at shorter wavelengths. The opiate alkaloids are of this class of alkaloids. Morphine and codeine differ only at the 3 position. Morphine has a hydroxy substituent, making it phenolic, and codeine has a methoxy substituent. In water at pH 1 both compounds fluoresce at 350 nm. Both are excited at 285 nm. Codeine has an additional excitation peak at 245 nm. They can be assayed in admixture because morphine loses its fluorescence under basic conditions by way of phenolate formation but codeine retains its fluorescence at high pH.

Caffeine is a member of the xanthine alkaloids. Caffeine in water fluoresces at $\lambda_f = 303$ nm ($\lambda_{ex} = 270$ nm). Caffeine is 1,3,7-trimethylxanthine. Xanthine in water at pH 1 fluoresces at $\lambda_f = 435$ nm ($\lambda_{ex} = 275$ nm). The fluorophore of caffeine and xanthine is that of purine ($\lambda_{ex} = 300$ nm; $\lambda_f = 360$ nm at pH 13).

Purine is, likewise, the fluorophore associated with the purine bases adenine and guanine. Adenine at pH 1 fluoresces at $\lambda_f = 380$ nm ($\lambda_{ex} = 265$ nm). Adenosine and its various phosphates all fluoresce at $\lambda_f = 390$ nm ($\lambda_{ex} = 272$ nm) in 5 M sulfuric acid. Guanine at pH 1 fluoresces at $\lambda_f = 350$ nm ($\lambda_{ex} = 275$ nm) and at pH 11 at $\lambda_f = 360$ nm ($\lambda_{ex} = 275$ nm). Guanosine and GMP (guanosine 5′-phosphate) at pH 1 fluoresce at $\lambda_f = 390$ nm ($\lambda_{ex} = 285$ nm).

The other bases associated with DNA and RNA are the pyrimidines: cytosine, thymine and uracil. They also exhibit fluorescence. Cytosine in water fluoresces at $\lambda_f = 313$ nm ($\lambda_{ex} = 267$ nm) but CMP (cytidine 5′-phosphate) in water fluoresces at $\lambda_f = 330$ nm ($\lambda_{ex} = 248$ nm). Thymine in water at pH 7 fluoresces at $\lambda_f = 320$ nm ($\lambda_{ex} = 265$ nm) but TMP (thymine 5′-phosphate) fluoresces at $\lambda_f = 330$ nm ($\lambda_{ex} = 248$ nm). Uracil in water at pH 7 fluoresces at $\lambda_f = 309$ nm ($\lambda_{ex} = 258$ nm) but UMP fluoresces at $\lambda_f = 320$ nm ($\lambda_{ex} = 248$ nm). The quantum yields of all the bases, nucleosides and nucleotides are extremely low so extraordinary conditions must be applied to the fluorescence analysis of DNA, RNA and their component parts.

Examples of compounds possessing native fluorescence listed so far have had a benzene ring or a nitrogen-containing heterocycle as the fluorophore. There is also a great, number of oxygen-containing heterocyclic compounds that fluoresce. Coumarins and flavonoids are the two largest classes of oxygen heterocycles. Coumarins fluoresce more intensely under basic conditions where flavones fluoresce weakly. In 30% sulfuric acid solution, flavones fluoresce intensely and coumarins do not.

Coumarin is not normally fluorescent but hydroxy substitution in any position except 8 leads to intense fluorescence at room temperature. In methanol, 3-hydroxycoumarin fluoresces at $\lambda_f = 372$ nm ($\lambda_{ex} = 316$ nm), 4-hydroxycoumarin fluoresces at $\lambda_f = 357$ nm ($\lambda_{ex} = 300$ nm), 6-hydroxycoumarin fluoresces at $\lambda_f = 431$ nm ($\lambda_{ex} = 341$ nm) and 7-hydroxycoumarin fluoresces at $\lambda_f = 392$ nm ($\lambda_{ex} = 333$ nm). Other oxygen-containing heterocycles are also of interest. Cannabinols in ethanol fluoresce at $\lambda_f = 318$ nm ($\lambda_{ex} = 280$ nm). Tocopherols (vitamin E) in ethanol fluoresce at $\lambda_f = 340$ nm ($\lambda_{ex} = 295$ nm).

Other vitamins and coenzymes exhibit fluorescence. Riboflavin (vitamin $B_2$) is a three-ring heterocycle which in water at pH 7 fluoresces at $\lambda_f = 565$ nm ($\lambda_{ex} = 370$ or 440 nm). The various forms of vitamin $B_6$ all have native fluorescence: pyridoxine ($\lambda_{ex} = 340$ nm; $\lambda_f = 400$ nm), pyridoxal ($\lambda_{ex} = 330$ nm; $\lambda_f = 385$ nm) and pyridoxamine ($\lambda_{ex} = 335$ nm; $\lambda_f = 400$ nm). Vitamin $B_{12}$, as cyanocobalamin, has a porphyrin moiety (cf. haem and chlorophyll) and fluoresces at $\lambda_f = 305$ nm ($\lambda_{ex} = 275$ nm). All D vitamins fluoresce when treated with strong acid but that is due to a degradation product. Vitamin $D_2$ (calciferol) has been reported exhibit native fluorescence in ethanol at $\lambda_f = 420$ nm ($\lambda_{ex} = 348$ nm). Calciferol's fluorescence is due to a rigid conjugated triene rather than an aromatic ring. Vitamin A (retinol) also has no aromatic ring but rather has an extended conjugation of $\pi$-bonds which leads to fluorescence. Retinol in ethanol fluoresces at $\lambda_f = 470$ nm ($\lambda_{ex} = 325$ nm) and in pentane–hexane it fluoresces at $\lambda_f = 513$ nm ($\lambda_{ex} = 321$ nm).

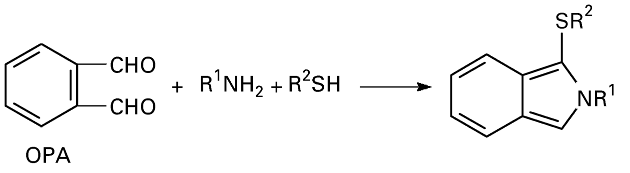
## Fluorescent derivatization

Not all substances are fluorescent. Non-fluorescent substances may be analysed by indirect methods of which there are several.

(1) Some organic compounds, themselves non-fluorescent, can be converted into fluorescent compounds by a chemical reaction with another organic compound which is itself also non-fluorescent.

o-Phthalaldehyde (OPA), itself non-fluorescent, is one of the most widely used reagents for the assay of amines, amino acids, peptides, amino carbohydrates, etc. It reacts with the primary amino group, in the presence of a thiol (usually 2-sulfanylethanol),

to give strongly fluorescent condensation products. The fluorescence is measured at ~455 nm with the excitation wavelength at ~340 nm. The assay can be conducted in the nanomole range.

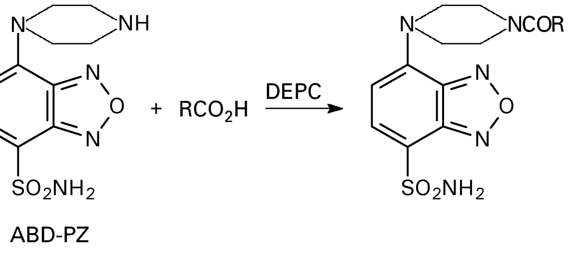



For example, this method has been employed for analysis of carbamate pesticides in surface water after the pesticides are hydrolysed in strong base to yield methyl amine and phenols.

New reagents such as naphthalene-2,3-dicarboxaldehyde (NDA), 1-phenylnaphthalene-2,3-dicarboxaldehyde (ϕNDA), and anthracene-2,3-dicarboxaldehyde (ADA) have been synthesized as improved OPA/thiol type reagents. While similar in many respects, there are important differences in these isoindole products. For example, the products formed with NDA, ϕNDA or ADA are considerably more stable than the corresponding OPA derivatives and possess substantially higher fluorescence quantum efficiencies and minimal interference compared with the latter.

(2) Some non-fluorescent organic compounds can react with fluorescent dyes to give fluorescent products which usually show altered fluorescence properties with regard to the free dye.

4-(Aminosulfonyl)-7-(1-piperazinyl)-2,1,3-benzoxadiazole (ABD-PZ), (maximum wavelength 565 nm with excitation at 413 nm), has been synthesized as a fluorescent reagent for determination of carboxylic acids. It reacts with carboxylic acids in the presence of diethyl phosphorocyanidate (DEPC) to produce fluorescent adducts with fluorescence at longer wavelengths. For example, the maximum wavelength of fluorescence of arachidic acid labelled with ABD-PZ is 580 nm with excitation at 440 nm. This method has been applied to a reversed-phase HPLC column for fatty acid mixture analysis. The detection limits for eight fatty acids are in the 10–50 fmol range.




When the piperazinyl group in ABD-PZ is substituted by a chiral group such as the 3-aminopyrrolidinyl group, it becomes a chiral derivatization reagent (D-ABD-APy or L-ABD-APy). This chiral derivatization reagent reacts with carboxylic acid enantiomers to form diastereomers for fluorescence detection. The diastereomers derived from antiinflammatory drugs and *N*-acetylamino acids are efficiently resolved by a reversed-phase column after they react with D-ABD-APy. The detection limit, for example, of ABD-APy-Naprofen on HPLC chromatograms is 30 fmol.

(3) Fluorometric enzyme assay involves a reaction catalysed by an enzyme, in which the product must show different fluorescence properties compared with those of the substrate. One example is the fluorometric peroxygenase assay for lipid hydroperoxides in meats and fish. In the reaction, catalysed by pea peroxygenase, the lipid hydroperoxide is reduced to an equimolar amount of alcohol during hydroxylation of the substrate, 1-methylindole, which shows no fluorescent product in the absence of the peroxygenase. The maximum wavelengths of excitation and emission for the hydroxylated product, 3-hydroxy-1-methylindole, appear at 410 and 485 nm, respectively, in *n*-butanol. The detectabilities of hydroperoxides are in the range of 25–150 nmol and α-tocopherol, an antioxidant at levels equivalent to those in meats and fish, did not affect the peroxygenase reaction. The method enables determination of total lipid hydroperoxides in sample homogenates without extracting total lipids from meats and fish.

(4) Fluoroimmunoassay involves attaching a fluorescent-labelled antibody to its specific antigen or vice versa, making use of a complexation reaction between the antigen and the antibody, for fluorescence detection at nanogram and lower levels. The specific affinity reactions may include the following: enzyme–substrate, hormone–receptor, neurotransmitter–receptor, pharmacological agent–receptor, etc. Such fluoroimmunoassays are divided into two categories – heterogeneous assays, which involve physical separation of the assay mixture before detection, and homogeneous assays, in which no separation steps are required. The most common fluorescent labels employed for fluoroimmunoassay are fluorescein isothiocyanate (FITC), which emits apple-green fluorescence (520 nm) when excited by ultraviolet or, preferably, by blue light (494 nm), and tetramethylrhodamine isothiocyanate (TRITC), which emits orange fluorescence (575 nm) when excited by ultraviolet or, preferably, by green light (550 nm).

One example is the fluoroimmunoassay for the routine detection of buprenorphine in urine samples of persons suspected of Temgesic® abuse. Buprenorphine antibody is labelled with pseudobuprenorphine, the dimer of buprenorphine. In this case, pseudobuprenorphine has a higher affinity for the antibody than that of FITC–norbuprenorphine. Pseudobuprenorphine shows an intense blue fluorescence with maximum at 435 nm when excited at 326 nm. The minimum detectable dose of buprenorphine by the fluoroimmunoassay is calculated to be 20 ng $mL^{-1}$.

## List of symbols

$c$ = velocity of light; $E$ = energy; $h$ = Planck's constant; $\lambda_{ex}$ = excitation wavelength; $\lambda_f$ = fluorescent wavelength; $\nu$ = frequency.

*See also:* **Biochemical Applications of Fluorescence Spectroscopy; Fluorescence Microscopy, Applications; Fluorescent Molecular Probes; Fluorescence Polarization and Anisotropy; Inorganic Condensed Matter, Applications of Luminescence Spectroscopy; UV-Visible Absorption and Fluorescence Spectrometers; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Baeyens WRG, de Keukeleire D, Korkidis K (ed) (1991) *Luminescence Techniques in Chemical and Biochemical Analysis*. New York: Marcel Dekker.

Brand L and Johnson ML (1997) (ed) *Fluorescence Spectroscopy*. San Diego: Academic Press.

Eastwood D and Love LJC (1988) (ed) *Progress in Analytical Luminescence*. Philadelphia: ASTM.

Goldberg MC (ed) (1989) *Luminescence Applications in Biological, Chemical, Environmental, and Hydrological Sciences*. Washington, DC: American Chemical Society.

Guilbault GG (ed) (1990) Assay of organic compounds. In: *Practical Fluorescence*, 2nd edn, pp 231–366. New York: Marcel Dekker.

Lakowicz JR (ed) (1991) *Topics in Fluorescence Spectroscopy*, Vol 1–5. New York: Plenum Press.

Lumb MD (ed) (1978) *Luminescence Spectroscopy*. London and New York: Academic Press.

Mason WT (ed) (1993) *Fluorescent and Luminescent Probes for Biological Activity: a Practical Guide to Technology for Quantitative Real-time Analysis*. London, San Diego: Academic Press.

Rendell D (1987) *Fluorescence and Phosphorescence Spectroscopy*. Published on behalf of ACOL, London by Wiley, Chichester and New York, 1987.

Sadtler Research Laboratories (1974–1976) *Fluorescence Spectra*, Chapter 3, Vol 1–8. Philadelphia.

Schulman SG (1977) *Fluorescence and Phosphorescence Spectroscopy: Physicochemical Principles and Practice*. New York: Pergamon Press.

Schulman SG (ed) (1985–1988) *Molecular Luminescence Spectroscopy: Methods and Applications*, Part I–II. New York: Wiley.

Soper SA, Warner IM and McGown LB (1998) Molecular fluorescence, phosphorescence and chemiluminescence. *Analytical Chemistry* 70: 477R–494R.

Winefordner JD, Schulman SG and O'Haver TC (1972) *Luminescence Spectrometry in Analytical Chemistry*. New York: Wiley-Interscience.

Wolfbeis OS (1985) The fluorescence of organic natural products. In: Schulman SG (ed) *Molecular Luminescence Spectroscopy: Methods and Applications*, Part I, Chapter 3, pp 167–370. New York: Wiley.

# Organic Chemistry Applications of NMR Spectroscopy

*See* **Structural Chemistry Using NMR Spectroscopy, Organic Molecules.**

# Organometallics Studied Using Mass Spectrometry

**Dmitri V Zagorevskii**, University of Missouri-Columbia, Columbia, MO, USA

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Applications**

Mass spectrometry is playing a significant role in organometallic chemistry. One of the most important applications of mass spectrometry is the determination of the molecular mass and elemental composition of metal compounds, and the identification of their structure. Mass spectrometry was first applied to the analysis of relatively volatile metal complexes using electron impact and chemical ionization techniques. Further development of ionization techniques has made it possible to analyse and to study the structure of a wide range of organometallics, including nonvolatile, ionic, multiply charged, polymetallic, and high-molecular-mass derivatives. Some of these ionization techniques allow identification of metal-containing intermediates directly from the condensed phase, providing beneficial information about reaction mechanisms.

Mass spectrometry is a unique method that allows study of the reactivity of isolated metal-containing ions in the gas phase in the absence of solvent. A number of fundamental thermochemical characteristics of organometallic molecules and ions, such as ionization energies, proton affinities, electron affinities and metal–ligand bond dissociation energies, can be determined from mass spectrometry experiments.

One of the most promising applications of mass spectrometry to the chemistry of metal compounds is in the investigation of the reactivity of metal-containing ions and molecules in the gas phase. Information about transformations of molecules on metal centres can be provided by these experiments. This is especially profitable in the study of mechanisms of reactions involving elusive intermediates (catalytic processes, interstellar chemistry, etc.) that cannot be isolated or characterized by 'traditional' spectral methods.

## Basics of the mass spectral analysis of metal complexes

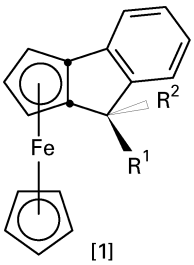
The following types of molecular characterization and recognition of metal complexes can be performed using a combination of mass spectrometric methods. The determination of the molecular mass of the analyte is a common first step in the mass spectral analysis of an unknown (organometallic) compound. A variety of ionization techniques can be employed to obtain molecular ions or charged adducts with a specific ionizing reactant. The isotope distribution in the molecular ion cluster (most metal atoms have a well-recognized isotope pattern) allows a thorough preliminary evaluation of the kind and number of metal atoms present. High-resolution experiments (accurate mass measurement) on molecular ions or adducts provide information about the elemental composition of the compound of interest, in many instances replacing relatively expensive and time-consuming traditional elemental analyses.

Detailed information about the structure of metal complexes may be obtained from the reactivities of their ions, including dissociation processes and ion-molecule reactions. A general problem in the elucidation of the structure of organometallic compounds is the determination of ligands surrounding the central metal atom(s). A common approach to this problem is to break metal–ligand bond(s) by allowing the ion to dissociate. This dissociation may occur upon ionization in (molecular) ions having an excess of internal energy. A sequential loss of metal–ligand bonds also can be achieved by activation of the ion. The observed mass loss leads to the deduction of the elemental composition of the ligand. Similar information can be obtained from ligand substitution ion–molecule reactions. Note however, that the latter method is limited to fragment ions having a coordinatively unsaturated metal atom. Molecular ions are usually inert to neutral reactants.

The recognition of atom connectivity in ligands is not always a direct task and a combination of experimental data should be considered. Reactions involving losses of radical and neutral molecules as well as migration of groups to the metal atom are the most informative for the recognition of the structure and location of substituents in the ligand. The most abundant processes are those involving atoms and groups located in $\alpha$ and $\beta$ positions of the substituent in ligands. It is useful to compare the mass spectral behaviour of the unknown compound with the reactivity of similar derivatives of well-established structure.

Some isomeric organometallic complexes may be recognized using mass spectral methods. The

difference in the mass spectra of isomers usually comes from the interaction of the metal atom with electronegative groups or unsaturated bonds in the substituent. These interactions increase the bond strength between the metal atom and the ligand, resulting in a preferential loss of other ligands. If an electronegative group is involved in such an interaction, it may migrate to the metal atom. The extent of the latter reaction depends on how close the migrating group is to the metal atom. Both effects can be illustrated by the electron impact ionization-induced fragmentation of complexes [1], having a hydroxyl group in the *exo* or *endo* position relative to the metal atom. The intensities of peaks due to the loss of the unsubstituted cyclopentadienyl ring and the migration of OH to the Fe atom were higher for the *endo* isomer (**Table 1**) in which the hydroxy group is in close proximity to the metal atom.




[1]

Intramolecular interaction of the substituent with the metal atom makes it possible to distinguish *ortho*- from *meta*- and *para*-ferrocenylbenzenes, $C_5H_5FeC_5H_4C_6H_4R$ ($R = NH_2$, COOH, $COOCH_3$, $COCH_3$, etc.). The dissociation of molecular ions ($P^+$) of *ortho* isomers gave rise to a strong peak due to the loss of the unsubstituted cyclopentadienyl ring, whereas their *meta* and *para* analogues produced no or very low intensity $[P-C_5H_5]^+$ ions.

The interaction of the metal atom may also induce loss of a radical from the substituent. This process is usually more intense if the lost group is in the *exo* position (**Table 1**).

Another effect that may result in differentiation of isomers is a metal-catalysed loss of neutral molecules from ligands. The isomer having an electronegative group in closer proximity to the metal atom usually loses this group more readily as a part of a neutral molecule. For example, the electron-impact ionization mass spectrum of $C_5H_5FeC_5H_4CH(OH)CH_3$ displays a stronger peak corresponding to the loss of $H_2O$ than the peak of the same mass in the EI mass spectrum of $C_5H_5FeC_5H_4CH_2CH_2OH$.

**Table 1** Relative intensities (relative to the molecular ion, P) of some ions in the mass spectra of *exo*-[1] and *endo*-[1]

| *Isomer* | | $[P\text{-}C_5H_5]^+$ | $C_5H_5FeOH^+$ | $[P\text{-}OH]^+$ |
|---|---|---|---|---|
| *Exo:* | $R^1$=H $R^2$=OH | 0.1 | 0.1 | 1.6 |
| *Endo:* | $R^1$=OH $R^2$=H | 0.6 | 0.33 | 1.7 |

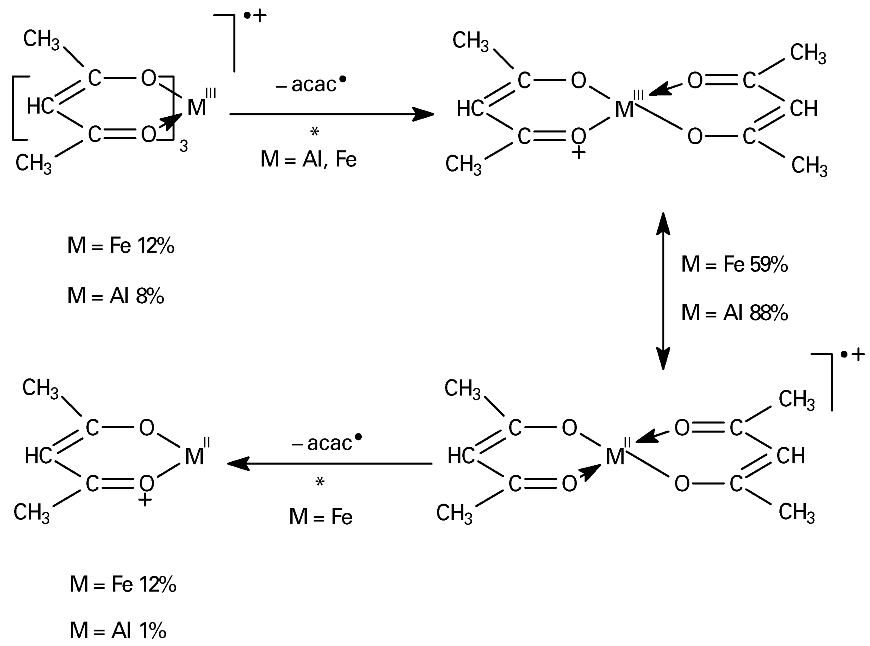
Metal-containing ions are useful reactants for identification of organic compounds. The formation of metal adducts is especially advantageous when traditional methods of ionization (electron-impact (EI) and chemical ionization (CI)) do not result in stable molecular ions or protonated species. Chemical ionization with metal and metal-containing ions provides high selectivity and sensitivity to specific types of analytes (unsaturated and functionalized hydrocarbons, peptides, crown ethers, polymers, etc.) and can be successfully used in GC-MS experiments. Regiospecific and stereospecific dissociation of metal adducts allows isomers to be distinguished. For example, the stereochemistry of the hydrogenated $D_8$-naphthalene, $C_{10}D_8H_4$, was deduced from the reactivity of the adduct with gas-phase $Co^+$ ions. Its collisional activation resulted in losses of $H_2$, $2H_2$, $D_2$ and $2D_2$, but no elimination of HD had been observed. These results were consistent with *cis* orientation of all hydrogen atoms in the organic molecule.

## Interpretation of mass spectra of metal compounds

Two formal approaches are used for the interpretation of mass spectra of organometallic compounds. The concept of valence-change has been applied to rationalize and to predict the type of fragmentation for a variety of $\sigma$ complexes and metal chelates, whereas the concept of charge localization on the metal atom was more advantageous for the interpretation of mass spectra of $\pi$ complexes of transition metals.

The valence-change concept makes the following assumptions:

(1) The metal atom contains an even number of electrons in the parent neutral molecule.
(2) Stabilization of molecular ions occurs when the metal atom can increase its oxidation state (OS). Fragmentation of these complexes involves a loss of even-electron species (molecules).
(3) If the metal atom has only one stable OS then molecular ions have a low intensity and their fragmentation will probably involve a loss of a radical followed by losses of neutral molecules.
(4) If the metal atom can easily reduce OS then the dissociation of (relatively unstable) molecular ions will involve a predominant loss of two radicals.

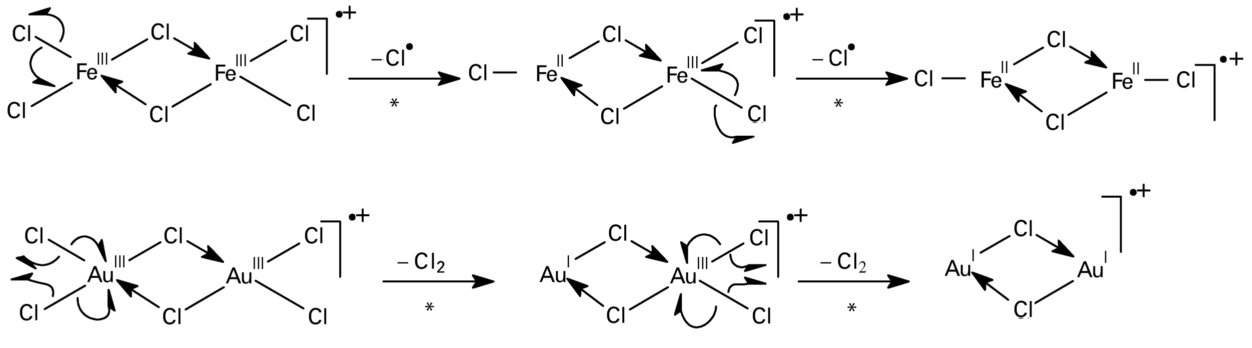



**Scheme 1** Comparative fragmentation pathways of the Al(III) and Fe(III) acetylacetonates. Reproduced with permission of John Wiley & Sons Limited from Lacey MJ and Shannon JS (1972) Valence-change in the mass spectra of metal complexes. *Organic Mass Spectrometry* **6**: 931–937.

The valence-change effect is illustrated by the behaviour of aluminium (one stable OS: III) and iron (two stable OS: III and II) trisacetylacetonates, M(acac)$_3$ (**Scheme 1**). Both complexes produced modest molecular ions in the EI mass spectra with the dominant loss of one β-diketonate ligand producing $M(acac)_2^+$ ions in which metal atoms retained their stable OS III. The Fe-derivative also underwent a significant loss of the second ligand, giving rise to ions having the metal atom in the OS II. The abundance of this process in the EI mass spectrum of Al(acac)$_3$ was very low owing to the low stability of Al(II).

The dissociation of complexes having two metals with two stable oxidation states follow the above rules. The most abundant reactions result in ions containing the metal atom in stable oxidation states and ions having two different (stable) OS can be easily formed (**Scheme 2**). In spite of much formalism, this concept has made a positive contribution to the interpretation and prediction of mass spectra of inorganic and organometallic compounds.

The most successful interpretation of mass spectra of transition metal π complexes is based on the concept of charge localization on the metal atom. The ionization energies of the majority of organometallic



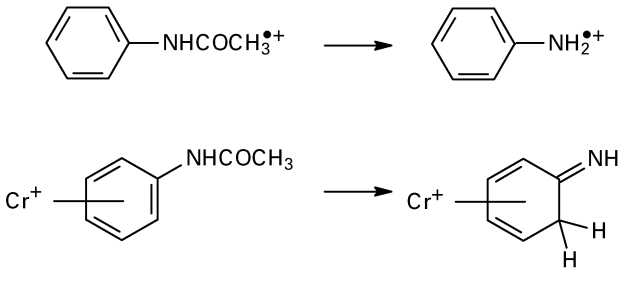

**Scheme 2** Comparative ion fragmentations of $Fe_2Cl_6$ and $Au_2Cl_6$. Reproduced with permission of John Wiley & Sons Limited from Lacey MJ and Shannon JS (1972) Valence-change in the mass spectra of metal complexes. *Organic Mass Spectrometry*, **6**: 931–937.

molecules are lower than the ionization energies of the ligands. They differ only slightly from the ionization energies of the free metal atoms. Accordingly, the ionization of transition metal complexes probably involves the removal of an electron from the metal atom. As a result, on the decomposition of molecular ions of transition metal–hydrocarbon compounds, the positive charge usually remains on the metal-containing fragment. The other consequence of charge localization on the metal atom is that the central metal atom controls the dissociation of organometallic ions. Unlike ionized metal-free organic molecules, their metal complexes lose even-electron neutrals rather than radicals. Also, the mechanisms of similar reactions in the coordinated and metal-free organic molecules are often different. Participation ('catalysis') of the metal atom in transformations of the ligand regulates these mechanisms as illustrated in **Scheme 3**. The loss of ketene from the ionized acetamide involves a 4-centre cyclic transition state, whereas its adduct with $Cr^+$ ions dissociates via a (formally) 6-centre intermediate.

Strong support for charge localization on the metal atom is provided by the reactivity of organic molecules upon their interaction with metal ions. Similarly to the unimolecular dissociation of coordinated ligands, most of these reactions result in the loss of even-electron neutrals formed via direct interaction of leaving group(s) with the metal atom.

The formation of neutral metal-containing species upon fragmentation of organometallic ions does not contradict the initial charge localization at the metal atom. The positively charged metal atom is easily attacked by electronegative groups (as anions) present in ligands, forming stable metal-containing molecules:

$$C_5H_5FeC_5H_4COCl^+ \rightarrow C_5H_5Fe^{II}Cl + C_5H_4CO^+ \quad [1]$$




**Scheme 3** Mechanisms of the loss of ketene from a metal-free and coordinated phenylacetamide.

This effect is common in the interaction of metal ions with functionalized organic reactants:

$$Mg^{\bullet +} + ClCH_2CH_2Cl \rightarrow Mg^{II}Cl_2 + C_2H_4^{\bullet +} \quad [2]$$

## Thermochemistry of organometallic molecules and ions

Mass spectrometry can be used to determine fundamental thermochemical properties of organometallic molecules such as ionization energy (IE), electron affinity (EA) and proton affinity (PA). IEs can be obtained by determining ionization thresholds of electron induced ionization (adiabatic IE) or photoionization (vertical IE). Electron exchange between a positively charged ion and a neutral molecule (electron transfer bracketing) allows the estimation of adiabatic IEs. In this approach, various reference compounds are introduced and the observation of (direct or reverse) electron transfer reactions indicates which molecule has the lower IE. If it is possible to establish accurately a pressure for the neutral reactant, then the electron transfer equilibrium can be measured to give IE values with high accuracy. This method requires the use of reference compounds having well-established ionization energies.

Similar experiments, involving electron transfer between an anion and neutral molecule, yield relative or absolute EAs. The method has been used to determine relative free energies for electron attachment for a variety of metallocenes and β-diketonate molecules. The results for tris(hexafluoracetylacetonate) and tris(acetylacetonate) showed a distinctive metal effect on EAs in the order Cr < Fe < Co < Mn. Measurements of electron attachment energies provide an important component of the thermochemical cycles involving oxidation/reduction of metal complexes. The latter may be used to obtain values of average heterolytic metal–ligand gas-phase bond energies and homolytic metal–ligand bond energies and to understand the thermochemistry of metal ion–solvent interaction.

The analysis of substituent effects in metal complexes is one of the goals of electron transfer equilibrium studies. For example, alkyl substitution in metallocenes predictably decreases their IEs. At the same time it has been shown that the gas-phase EAs can be increased by a larger alkyl substituent in the hydrocarbon ligand (**Figure 1**).

Anion transfer reactions to/from metal complexes are sources for anion affinities of organometallic molecules. To illustrate, the hydroxide affinity of $(CO)_5Fe$ has been determined by measurement of the

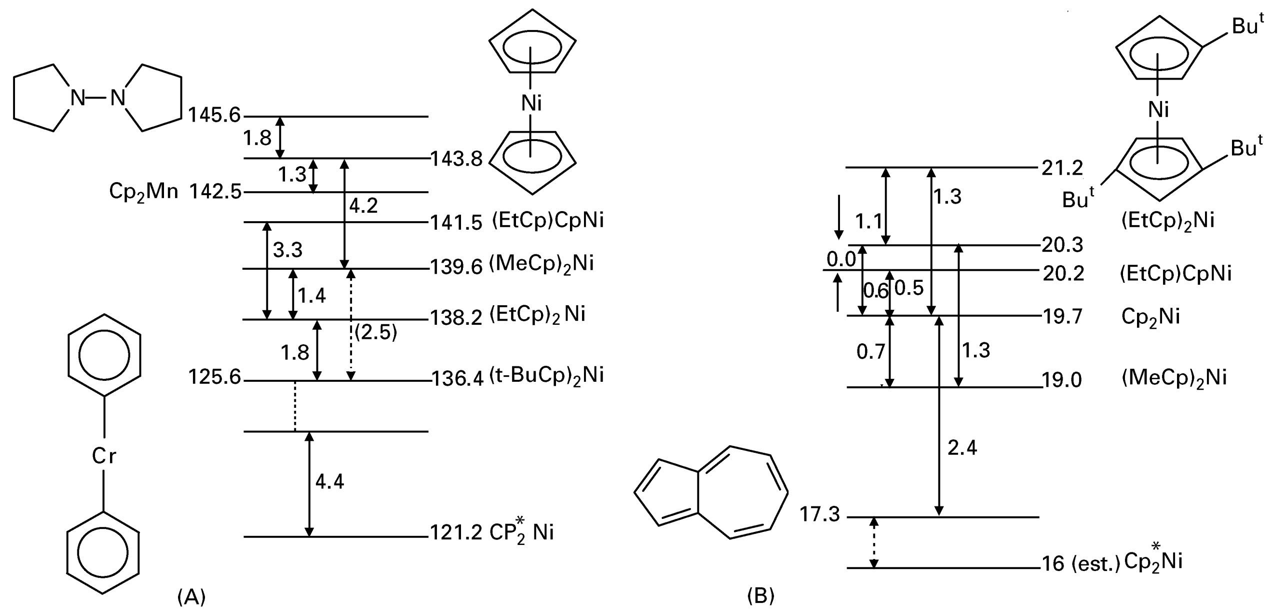


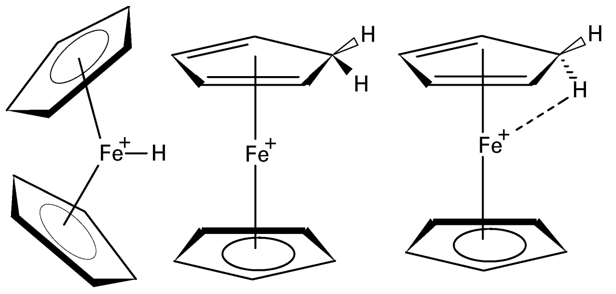

**Figure 1** Electron transfer equilibrium ladders showing free energies of ionization (A) and electron attachment (B). Cp* denotes the pentamethylcyclopentadienyl ligand. Reproduced with permission of the American Chemical Society from Richardson DE, Ryan MF, Khan MDNI and Maxwell KA (1992) Alkyl substituent effects in cyclopentadienyl metal complexes; trends in gas-phase ionization and electron attachment energetics of alkylnickelocenes. *Journal of the American Chemical Society*, **114**: 10482–10485.

equilibrium constant for hydroxide transfer exchange between $(CO)_5Fe$ and $SO_2$. This value was used to estimate the heat of formation of $(CO)_4FeCOOH^-$. Ion–molecule reactions of these ions and their collision-induced dissociation gave rise to a variety of negatively charged species having a coordinatively unsaturated metal atom. Study of their reactivity is a good source for obtaining thermochemical characteristics of elusive metal complexes.

Proton transfer equilibrium measurements and proton transfer bracketing methods are the sources for proton affinity values of organometallic complexes. The determination of the site of protonation, i.e. metal atom versus a ligand, is a fundamental dilemma of any study on the protonation of metal complexes. It was demonstrated, for example, that $Fe(CO)_5$ was protonated exclusively at the metal atom, whereas the results for the proton transfer to ferrocene can be explained by the formation of a metal-protonated compound [2a] and a ring-protonated [2b] form. The observation of hydrogen (deuterium) atom exchange between the cyclopentadienyl rings in $[(C_5H_5)_2Fe]D^+$ and $[(C_5D_5)_2Fe]H^+$ ions was rationalized by the agostic interaction with the metal atom (structure [2c]).

## Chemistry of metal–ligand bonds

Knowledge of metal–ligand bond energies is fundamental information for organometallic chemistry. It is essential for the understanding of catalytic reaction mechanisms, which often involve the cleavage or the formation of these bonds. Mass spectrometry offers a series of experimental methods for determining absolute and relative bond strengths between a




[2a] [2b] [2c]

positively (or negatively) charged metal centre and organic (or inorganic) ligands. Direct determination of metal–ligand bond dissociation energies can be performed by measuring appearance energies (AEs) of the molecular and fragment ions. The best results are obtained from AEs of ions produced by metastable dissociation of mass-selected precursors. The kinetic energy release distribution (KERD) during metastable dissociation of ions is another source for the quantitative characterization of metal–ligand bonds. The experimental results obtained by this method require theoretical calculation to extract the information on the enthalpy change for the observed processes. Photoelectron–photoion coincidence is also a useful source for determining the thermochemistry of metal-containing ions.

A large number of both relative and absolute bond energies in metal-containing ions have been measured by the method that depends upon the abundances of products formed by competitive ligand loss (the kinetic method). The metal–ligand bond enthalpies can be determined from the metastable and collision-induced dissociation of $LM^+L'$ ions, where L and L′ are different molecules. The general trend in metal–ligand bond dissociation energies is that the larger alkyl derivatives are more strongly bound to the metal atom than are their smaller homologues. A study of systems containing three and four ligands at the metal atom, $L_2M^+L'$ and $L_2M^+L'_2$ reveals information about relative molecular pair and molecular triplet metal ion affinities. A particular advantage of the kinetic and some other methods is that they can probe ions containing thermally unstable ligands whose chemistry is difficult to study in the condensed phase.

A variety of other methods for obtaining metal–ligand bond dissociation characteristics employ activation of the ion of interest to initiate its dissociation. In addition to activation by collisions with a target gas, the dissociation of metal–containing ions of interest can be induced by photoexcitation, colliding them into a surface, or bombarding them with electrons.

Ion–molecule ligand exchange reactions are convenient for obtaining relative and absolute metal–ligand bond strengths. Using this approach, the affinities of molecules for a variety of 'bare' and ligated metal ions have been determined. From these data the order of relative softness of the Lewis acids was $H^+ < Al^+ << Mn^+ \leq FeBr^+ < Co^+ \leq C_5H_5Ni^+ < Ni^+ < Cu^+$. The metal ion affinities obtained represented relative metal–ligand binding energies, which can also be measured for two-ligand systems, $D^0(M^+\text{-}2L)$, as well as for negatively and doubly charged positive ions. An important application of ligand exchange reactions is the determination of relative metal ion affinities for molecules of biological origin, e.g. peptides and amino acids.

Monitoring of the thresholds of endothermic reactions by the guided ion beam technique is one of the methods used to obtain the most accurate bond energies. Similar measurements can be performed employing (Fourier transform) ion cyclotron resonance (FT-ICR) spectroscopy. General types of endothermic reactions studied to obtain bond dissociation energies are shown in Equations [3]–[5] ($M^+$ is a 'bare' or ligated metal ion):

$$M^+ + RX \longrightarrow MX^+ + R \quad [3]$$

$$M^+ + RX \longrightarrow MX + R^+ \quad [4]$$

$$MY^+ + RX \longrightarrow MX^+ + RY \quad [5]$$

If the thermochemistry of the reactants ($M^+$, $MY^+$, RX) and one of the reaction products (R, $R^+$, RY) is well established, then using the experimentally measured threshold energy of reaction one can calculate $M^+$–X and M–X bond strengths. Important organometallic intermediates, such as metal hydrides, carbynes, carbenes and metal-alkyls, have been characterized by the above methods and their ionic and neutral heats of formation were deduced from metal–ligand bond energies. An extensive study of exothermic and thermoneutral ion–molecule reactions provide complementary and in many cases unique information on the thermochemistry of metal-containing species. All of these results have made a great contribution to the understanding of metal–ligand bonding and of periodic trends in transition metal–ligand bonds, and to the evaluation of the multiplicity of metal–ligand bonds and other fundamental topics of organometallic chemistry.

The results on metal–ligand bond dissociation energies (strengths) obtained by different methods are usually in good agreement. For example, the $Fe^+$–CO bond dissociation energies determined from KERD, ion beam and FT-ICR experiments ($32 \pm 5$, $31.3 \pm 1.8$ and $37 \pm 6$ kcal mol$^{-1}$, respectively) were very similar and close to the theoretically calculated value of 30.3 kcal mol$^{-1}$, whereas the appearance energy measurements for $FeCO^+$ and $Fe^+$ ions in the EI mass spectrum of $Fe(CO)_5$ gave significantly higher values (between 41.5 and 51.4 kcal mol$^{-1}$). Good agreement

between two or more methods supported by the results of high-level theoretical calculations is usually a sufficient criterion for which number should be used in thermochemical calculations (33 kcal mol$^{-1}$ is probably a good estimate for $Fe^+$–CO). However, a critical analysis of the experimental techniques, understanding their limitations and sources of errors, as well as knowledge of trends in changing bond dissociation energies is highly recommended before a specific thermochemical value is accepted.

## Transformations of ligands on charged metal centres

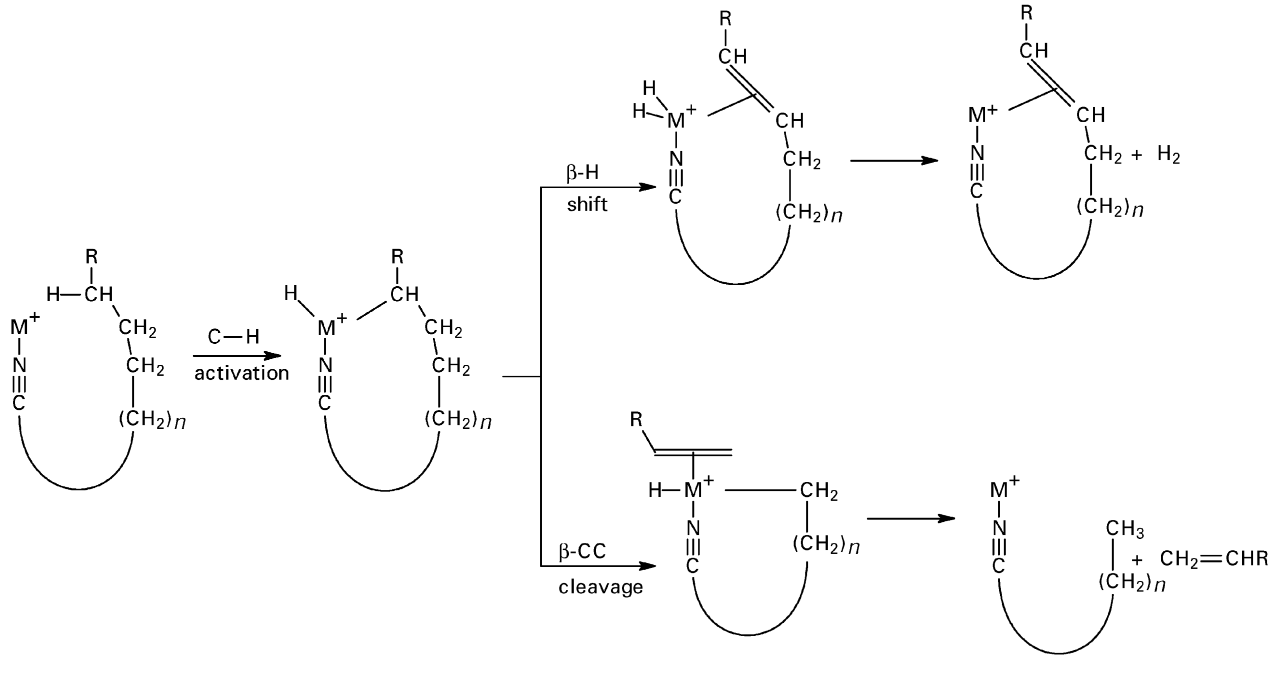
Mass spectrometry is widely used for studying reaction mechanisms involving metal-containing reaction intermediates. A majority of these studies involve the investigation of transformations of organic molecules on ligated or 'bare' metal ions. Reactions [1]–[5] provide a few examples from a large number of processes involving bond cleavage within ligands. Particular interest in this field of research is focused on intrinsic mechanisms of metal ion-induced C–H, C–C and C–heteroatom bond activation reactions. The results of these studies help the understanding and modelling of the elementary stages of important homogeneous and heterogeneous catalytic processes, metal ion biochemistry, synthesis of electronic and ceramic materials, metal–solvent interactions, reactions taking place in interstellar systems, etc.

A series of experiments on the reactivity of metal ions with nitriles (RCN) led to the discovery of the remote functionalization mechanism. The initial interaction of the metal ion involves coordination at the nitrile group. The insertion of the metal atom into a C–H or C–C bond occurs only after the alkyl chain becomes long enough (at least 3 or 4 methylene groups) to interact with a remote bond. The dissociation of the metal–hydride (**Scheme 4**) or metal–alkyl intermediate results in a loss of $H_2$, alkene or alkane molecules depending on the structure of the hydrocarbon group R.

Other important reactive metal-containing intermediates (e.g. metal–benzene complexes) and processes (e.g. decarboxylation of ketones by metal ions) of practical interest have been characterized using various mass special methods and provide insight into the mechanisms of organometallic reactions.

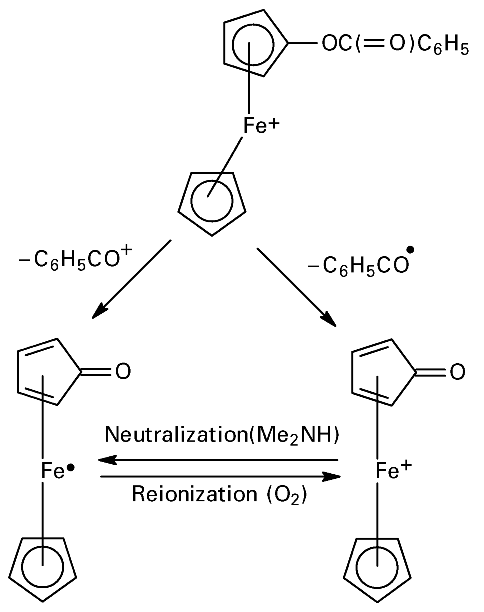
## Elusive neutral organometallics generated from ions

Neutralization–reionization mass spectrometry (NRMS) is a unique mass spectral technique that allows the generation of *neutral* species from their charged counterparts. The major application of NRMS is to produce unstable reaction intermediates that cannot be isolated or characterized by other means to yield new, previously unknown, molecules and radicals. The method has been used successfully




**Scheme 4** Generalized mechanism for the 'remote funtionalization' of a C–H bond. Reproduced with permission of the American Chemical Society from Eller K and Schwartz H (1991) Organometallic chemistry in the gas phase. *Chemical Reviews* **91**: 1121–1127.

to generate a variety of organometallic species. A wide range of elusive metal-containing molecules (AuF, PrF, $CH_2SiH_2$, $C_5H_5$Rhacac) and radicals (NiCCH, $C_5H_5FeC_5H_4CO$, $(C_5H_5)_2Zr$ and others) have been generated and characterized for the first time using the NRMS method. The observation of these species in the gas phase suggests their possible formation in other than gas-phase experimental conditions, at least as short-lived reaction intermediates. This information is used by chemists to confirm and evaluate reaction mechanisms, including elementary steps in catalytic transformations on metal centres. **Scheme 5** represents a common sequence in the NRMS experiment for the generation of elusive organometallic species. Complementary information about the intrinsic stability of molecules and radicals can also be obtained by identifying the neutral products of ion fragmentation using the collision-induced dissociative ionization method followed by the detection of its positively charged counterpart. This technique was used, for example, for the observation of the neutral ferrocenyloxy radical originating from the dissociation of ferrocenylbenzoate to benzoyl ion (see **Scheme 5**).




**Scheme 5** Generation of ferrocenyloxy radical from the ionized ferrocenylbenzoate.

*See also:* **Fragmentation in Mass Spectrometry; Ion Energetics in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Metastable Ions; Neutralization-Reionization in Mass Spectrometry; Photoelectron-Photoion Coincidence Methods in Mass Spectrometry (PEPICO); Photoionization and Photodissociation Methods in Mass Spectrometry; Proton Affinities.**

## Further reading

Bruce MI (1968) Mass spectra of organometallic compounds. In: Stone FGA and West R (eds) *Advances in Organometallic Chemistry*, Vol 6, pp 273–333. New York: Academic Press.

Cais M and Lupin MS (1970) Mass spectra of metallocenes and related compounds. In: Stone FGA and West R (eds) *Advances in Organometallic Chemistry*, Vol 8, pp 211–333. New York: Academic Press.

Charalambous J (ed) (1975) *Mass Spectrometry of Metal Compounds*. London: Butterworths.

Freiser BS (1994) Selected topics in organometallic ion chemistry. *Accounts of Chemical Research* **27**: 353–360.

Freiser BS (ed) (1996) *Organometallic Ion Chemistry*. Dordrecht: Kluwer.

Eller K and Schwarz H (1991) Organometallic chemistry in the gas phase. *Chemical Reviews* **91**: 1121–1177.

Lacey MJ and Shannon JS (1972) Valence-change in the mass spectra of metal complexes. *Organic Mass Spectrometry* **6**: 931–937.

Litzov MR and Spalding PH (1973) *Mass-Spectrometry of Inorganic and Organometallic Compounds*. Amsterdam: Elsevier.

Marks TJ (ed) (1990) *Bonding Energetics in Organometallic Chemistry*, Washington, DC: American Chemical Society.

*Mass Spectrometry* (Specialist Periodic Report Vols 1–10) (1971–1989). London: The Chemical Society.

Muller J (1972) Decomposition of organometallic complexes in the mass spectrometer. *Angewandte Chemie, International Edition in English***11**: 653–665.

Squires RR (1987) Gas-phase transition-metal negative ion chemistry. *Chemistry Reviews* **87**: 623–646.

Zagorevskii DV and Holmes JL (1994) Neutralization-reionization mass spectrometry applied to organometallic and coordination chemistry. *Mass Spectrometry Reviews* **13**: 133–154.

# P

# $^{31}$P NMR

**David G Gorenstein** and **Bruce A Luxon**, University of Texas Medical Branch, Galveston, TX, USA

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Applications |

## Introduction

Although $^{31}$P NMR spectra were reported as early as 1951, it was the availability of commercial multinuclear NMR spectrometers in about 1955 that led to the application of $^{31}$P NMR as an important analytical tool for structure elucidation. Early spectrometers generally required neat samples in large nonrotating tubes (8–12 mm OD). By the mid-1960s NMR became more sensitive and the availability of higher field electromagnets and signal averaging spurred rapid growth in the number of reported $^{31}$P spectra and the publication of the first monograph devoted to this field.

With the introduction of Fourier-transform (FT) and high-field superconducting-magnet NMR spectrometers in about 1970, $^{31}$P NMR spectroscopy expanded beyond the study of small organic, organometallic and inorganic compounds to biological phosphorus-containing compounds as well. The latest multinuclear FT NMR spectrometers have reduced if not eliminated the serious limitation to the widespread utilization of phosphorus NMR, which is the low sensitivity of the phosphorus nucleus (6.6% at constant field compared to $^{1}$H NMR; magnetogyric ratio $\gamma$, $10.839 \times 10^7$ rad $T^{-1}$ $s^{-1}$; NMR frequency at 4.7 T, 80.96 MHz). Today, routinely, millimolar (or lower) concentrations of phosphorus nuclei in as little as 0.3 mL of solution are conveniently monitored. The $^{31}$P nucleus has other convenient NMR properties making it suitable for FT NMR: spin $\frac{1}{2}$ (which avoids problems associated with quadrupolar nuclei), 100% natural abundance, moderate relaxation times (providing relatively rapid signal averaging and sharp lines), and a wide range of chemical shifts (>2000 ppm).

In this article the interpretation of various $^{31}$P NMR spectroscopic parameters, particularly chemical shifts and coupling constants will be described. A major emphasis will be placed on newer developments in $^{31}$P NMR methods which have considerably expanded the utility of this important spectroscopic probe in organic and biological structure determination.

## Practical considerations

Today's commercial NMR spectrometers cover the $^{31}$P frequency range from 24 to 323 MHz (on a 800 MHz $^{1}$H NMR spectrometer). Generally, for small, phosphorus-containing compounds, high signal-to-noise and resolution requirements dictate use of as high a magnetic field strength as possible since both sensitivity and chemical-shift dispersion increase at higher operating frequency (and field). However, consideration must be given to field-dependent relaxation mechanisms such as chemical shift anisotropy which can lead to substantial line broadening of the $^{31}$P NMR signal at high fields. Indeed, especially for larger biomolecules, sensitivity is often poorer at very high fields because of considerably increased line widths. The latest probe designs provide a remarkable improvement in signal-to-noise. Indirect detection 2D NMR experiments have further improved sensitivity.

Typical acquisition times for the 1D $^{31}$P free induction decay (FID) following a 90° radiofrequency

pulse are 1–8 s depending on the required resolution (dictated by the line width of the signal, $1/\pi T_2^*$, where $T_2^*$ is the time constant for the FID). Waiting longer than $2T_2^*$ will generally not improve the signal-to-noise ratio (S/N). Additional consideration for optimization of the S/N must be given to the time it takes for the $^{31}$P spins to return to thermal equilibrium after a 90° radiofrequency pulse, which is roughly 3 times the spin–lattice relaxation time ($T_1$). If $T_1 \approx T_2^*$, as would be true for small phosphorus-containing molecules where magnetic field inhomogeneity and paramagnetic impurities do not lead to any additional line broadening, then a waiting period between pulses of $3T_2^*$ provides a good compromise between adequate resolution and signal sensitivity. If $T_1 > T_2^*$, as is often the case in larger biomolecular systems, then waiting only $3T_2^*$ does not allow the magnetization to return to equilibrium and an additional delay must generally be introduced so that the total time between pulses is ~$3T_1$. This wait can be substantially shortened if the Ernst relationship is used to set the pulse flip angles to < 90°. At low field, 60–70° pulses, 4 to 8 k data points and 2.0–5.2 s recycle times are generally used. The spectra are generally broadband $^1$H decoupled.

The $^{31}$P spectra are generally referenced to an external sample of 85% $H_3PO_4$ or trimethylphosphate which is ~3.46 ppm downfield of 85% $H_3PO_4$. Note that throughout this review the IUPAC convention is followed so that *positive values are to high frequency* (low field). One should cautiously interpret reported $^{31}$P chemical shifts because the early literature (pre-1970s) and even many later papers use the opposite sign convention.

### Quantification of peak heights

The intensity of a resonance can be measured in several ways: (1) peak heights and areas obtained from the standard software supplied by the spectrometer manufacturer, (2) peak heights measured by hand, (3) peaks cut and weighed from the plotted spectrum, and (4) peaks fitted to a Lorentzian line shape. For flat baselines, intensity measurements are generally straightforward. However, in the event of curved baselines the measurements are somewhat uncertain and manual measurements are generally more reliable than intensity values obtained from computer software.

It is often necessary that experiments be carried out without allowing time for full recovery of longitudinal magnetization between transients because of the limited availability of spectrometer time or of the limited lifetime of the sample. Because of variations in $T_1$ between different phosphates and variation in the heteronuclear NOE to nearby protons, care should be made in interpretation of peak area and intensities. Addition of a recycle delay of at least $5 \times T_1$ between pulses and gated decoupling only during the acquisition time to eliminate the $^1$H–$^{31}$P NOE largely eliminates quantification problems.

## $^{31}$P chemical shifts

### Introduction and basic principles

The interaction of the electron cloud surrounding the phosphorus nucleus with an external applied magnetic field $B_0$ gives rise to a local magnetic field. This induced field shields the nucleus, with the shielding proportional to the field $B_0$ so that the effective field, $B_{\text{eff}}$, felt by the nucleus is given by

$$B_{\text{eff}} = B_0(1 - \sigma) \qquad [1]$$

where $\sigma$ is the shielding constant. Because the charge distribution in a phosphorus molecule will generally be far from spherically symmetrical, the $^{31}$P chemical shift (or shielding constant) varies as a function of the orientation of the molecule relative to the external magnetic field. This gives rise to a chemical-shift anisotropy that can be defined by three principal components, $\sigma_{11}$, $\sigma_{22}$ and $\sigma_{33}$ of the shielding tensor. For molecules that are axially symmetrical, with $\sigma_{11}$ along the principal axis of symmetry, $\sigma_{11} = \sigma_{\parallel}$ (parallel component), and $\sigma_{22} = \sigma_{33} = \sigma_{\perp}$ (perpendicular component). These anisotropic chemical shifts are observed in solid samples and liquid crystals, whereas for small molecules in solution, rapid tumbling averages the shift. The average, isotropic chemical shielding $\sigma_{\text{iso}}$ (which would be comparable to the solution chemical shift) is given by the trace of the shielding tensor or

$$\sigma_{\text{iso}} = \frac{1}{3}(\sigma_{11} + \sigma_{22} + \sigma_{33}) \qquad [2]$$

and the anisotropy $\Delta\sigma$ is given by

$$\Delta\sigma = \sigma_{11} - \frac{1}{2}(\sigma_{22} + \sigma_{33}) \qquad [3]$$

or, for axial symmetry,

$$\Delta\sigma = \sigma_{\parallel} - \sigma_{\perp} \qquad [4]$$

### Theoretical $^{31}$P chemical shift calculations and empirical observations

Three factors appear to dominate $^{31}$P chemical shift differences $\Delta\delta$, as shown by

$$\Delta\delta = -C\Delta\chi_X + k\Delta n_\pi + A\Delta\theta \qquad [5]$$

where $\Delta\chi_X$ is the difference in electronegativity in the P–X bond, $\Delta n_\pi$ is the change in the π-electron overlap, $\Delta\theta$ is the change in the σ-bond angle, and *C*, *k*, and *A* are constants.

As suggested by Equation [5], electronegativity effects, bond angle changes, and π-electron overlap differences can all potentially contribute to $^{31}$P shifts in a number of classes of phosphorus compounds. While these semiempirical isotropic chemical-shift calculations are quite useful in providing a chemical and physical understanding for the factors affecting $^{31}$P chemical shifts, they represent severe theoretical approximations. More exact *ab initio* chemical-shift calculations of the shielding tensor are very difficult although a number of calculations have been reported on phosphorus compounds. Whereas the semiempirical theoretical calculations have largely supported the importance of electronegativity, bond angle, and π-electron overlap on $^{31}$P chemical shifts, the equations relating $^{31}$P shift changes to structural and substituent changes unfortunately are not generally applicable. Also, because $^{31}$P shifts are influenced by at least these three factors, empirical and semiempirical correlations can only be applied to classes of compounds that are similar in structure. It should also be emphasized again, that structural perturbations will affect $^{31}$P chemical shift *tensors*. Often variations in one of the tensor components will be compensated for by an equally large variation in another tensor component with only a small net effect on the isotropic chemical shift. Interpretation of variations of isotropic $^{31}$P chemical shifts should therefore be approached with great caution.

Within these limitations, a number of semiempirical and empirical observations and correlations, however, have been established and have proved useful in predicting $^{31}$P chemical-shift trends. Indeed, unfortunately, no single factor can readily rationalize the observed range of $^{31}$P chemical shifts (**Figure 1**).

**Bond angle effects** Changes in the $\sigma$-bond angles appear to make a contribution ($|A|$, Equation [5]) to the $^{31}$P chemical shifts of phosphoryl compounds, although electronegativity effects apparently predominate. Empirical correlations between $^{31}$P chemical shifts and X–P–X bond angles can be found, although success here depends on the fact that these correlations deal with only a limited structural variation: in the case of phosphate esters, it is the number and chemical type of R groups attached to a tetrahedron of oxygen atoms surrounding the phosphorus nucleus. For a wide variety of different alkyl phosphates (mono-, di-, and triesters, cyclic and acyclic neutral, monoanionic, and dianionic esters), at bond angles < 108° a decrease in the smallest O–P–O bond

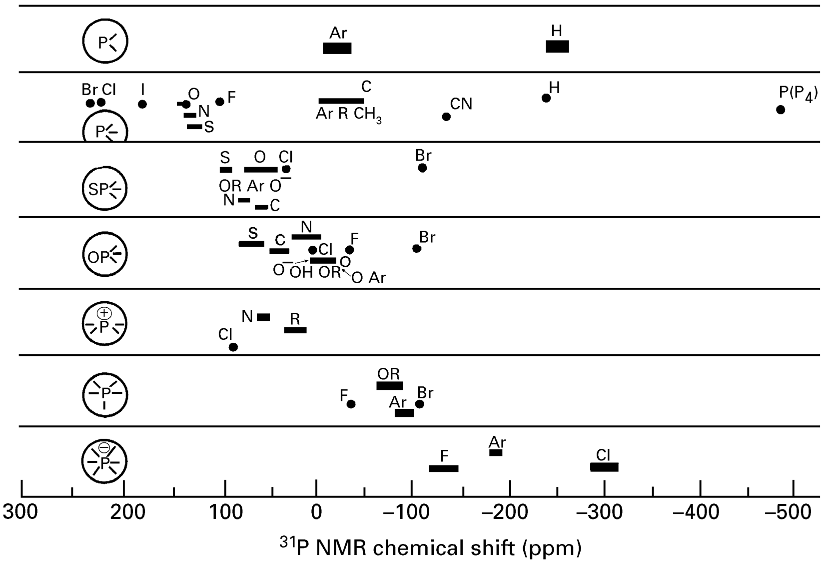



**Figure 1** Typical $^{31}$P chemical shift ranges for phosphorus bonded to various substituents in different oxidation states. (P– indicates the $P_4$ molecule.)

angle in the molecule generally results in a deshielding (downfield shift) of the phosphorus nucleus.

**Torsional angle effects on $^{31}$P chemical shifts** Semi-empirical molecular orbital calculations and *ab initio* gauge-invariant-type molecular orbital, chemical-shift calculations suggested that $^{31}$P chemical shifts are also dependent on P–O ester torsional angles which has been shown to be of great value in analysis of DNA structure (see below). The two nucleic acid P–O ester torsional angles, ζ (5′-O–P) and α (3′-O–P) are defined by the (5′-O–P–O-3′) backbone dihedral angles. These chemical-shift calculations and later empirical observations indicated that a phosphate diester in a $B_I$ conformation (both ester bonds gauche(−) or −60°) should have a $^{31}$P chemical shift 1.6 ppm upfield from a phosphate diester in the $B_{II}$ conformations (α = gauche(−); ζ = trans or 180°).

## $^{31}$P signal assignments

If the proton spectra of the molecule has been previously assigned, then 2D $^{31}$P–$^{1}$H heteronuclear correlation NMR spectroscopy can generally provide the most convenient method for assigning $^{31}$P chemical shifts in complex spectra. Whilst application of these experiments to DNA is clear, the 2D methods will of course equally apply to organophosphorus compounds as well.

Conventional 2D $^{31}$P–$^{1}$H heteronuclear shift correlation (HETCOR) NMR spectroscopy, the 2D long-range COLOC (correlation spectroscopy via long range coupling) experiment and indirect detection ($^{1}$H detection) HETCOR experiments can be used to assign multiple $^{31}$P signals in complex spectra such as those of oligonucleotide duplexes. Additional 2D heteronuclear *J* cross-polarization hetero TOCSY (TOCSY = total correlation spectroscopy), 2D heteronuclear TOCSY-NOESY (NOESY = nuclear Overhauser effect spectroscopy), and even a 3D hetero TOCSY-NOESY experiment can be used if additional spectral dispersion, by adding a third frequency dimension, is desirable. This may prove to be extremely valuable for ribo-oligonucleotides where very little $^{1}$H spectral dispersion in the sugar proton chemical shifts is unfortunately observed.

Generally these 2D experiments correlate $^{31}$P signals with coupled $^{1}$H NMR signals. Assuming the $^{1}$H NMR spectra have been assigned, these methods allow for direct assignment of the $^{31}$P signals. The HETCOR measurements, however, suffer from poor sensitivity as well as poor resolution in both the $^{1}$H and $^{31}$P dimensions, especially for larger biomolecular structures. The poor sensitivity is largely due to the fact that the $^{1}$H–$^{31}$P scalar coupling constants are generally about the same size or smaller (except for organophosphorus molecules with directly bonded hydrogens) than the $^{1}$H–$^{1}$H coupling constants. Sensitivity is substantially improved by using a heteronuclear version of the 'constant time' coherence transfer technique, referred to as COLOC and originally proposed for $^{13}$C–$^{1}$H correlations.

An example of a 2D HETCOR spectrum of the self-complementary 14-base-pair oligonucleotide duplex d(TGTGAGCGCTCACA)$_2$, is shown in **Figure 2**. The cross-peaks represent scalar couplings between $^{31}$P nuclei of the backbone and the H3′ and H4′ deoxyribose protons. Assuming that the chemical shifts of these protons have been assigned (by $^{1}$H–$^{1}$H NOESY and COSY spectra) the $^{31}$P signals may be readily assigned (COSY = homonuclear chemical shift correlation spectroscopy).

## Coupling constants

### Directly bonded phosphorus coupling constants $^{1}J_{PX}$

One bond P–X coupling constants ($J_{PX}$) have generally been rationalized in terms of a dominant Fermi-contact term

$$J_{PX} = \frac{Aa_P^2 a_X^2}{1 + S_{PX}^2} + B \qquad [6]$$

where *A* and *B* are constants, $a_P^2$ and $a_X^2$ are percentage s character on phosphorus and atom X, respectively, and $S_{PX}$ is the overlap integral for the P–X bond. Because the Fermi-contact spin–spin coupling mechanism involves the electron density at the nucleus (hence the s-orbital electron density), an increase in the s character of the P–X bond is generally associated with an increase in the coupling constant. The percentage s character is determined by the hybridization of atoms P and X, and as expected sp$^3$-hybridized atoms often have $^{1}J_{PX}$ larger than p$^3$ hybridized atoms. Thus $^{1}J_{PH}$ for phosphonium cations of structure $PH_nR_{4-n}^+$ with sp$^3$ hybridization are ~500 Hz, whereas $^{1}J_{PH}$ for phosphines $PH_nR_{3-n}$ with phosphorus hybridization of approximately p$^3$ are smaller, ~200 Hz. Furthermore, as the electronegativity of atom X increases, the percentage s character of the P–X bond increases, and the coupling constant becomes more positive. In many cases, however, these simple concepts fail to rationalize experimental one-bond P–X coupling constants (**Table 1**) because other spin–spin coupling mechanisms can also contribute significantly to the coupling constant. For tetravalent phosphorus, a very good correlation

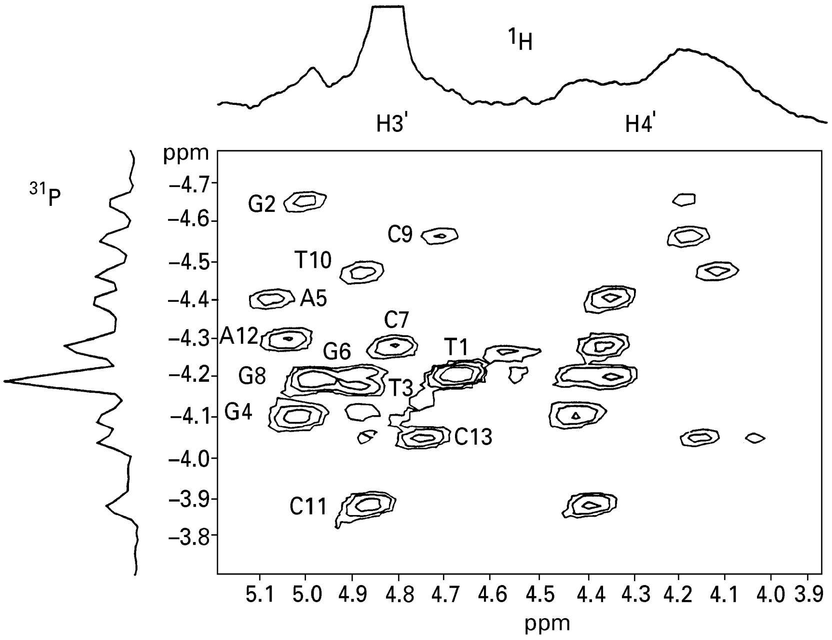



**Figure 2** Pure absorption phase $^{31}$P–$^{1}$H heteronuclear correlation spectrum of tetradecamer duplex d(TGTGAGCGCTCACA)$_2$ at 200 MHz ($^{1}$H). $^{31}$P chemical shifts are reported relative to trimethyl phosphate which is 3.456 ppm downfield from the 85% phosphoric acid. Reproduced with permission.

is found between $^{1}J_{PC}$ and the phosphorus 3s–carbon 2s bond orders, the percentage s in the P–C bonding orbital in going from alkyl to alkenyl to alkynyl ($sp^3 \rightarrow sp^2 \rightarrow sp$), and $^{1}J_{PC}$. Calculations and empirical observations on trivalent phosphorus compounds are *not* successful however, and suggest that the Fermi-contact contribution only dominates tetravalent phosphorus compounds.

One-bond P–H coupling constants appear always to be positive and vary from about +120 to +1180 Hz. Other heteroatom one-bond P–X coupling constants vary over a similar wide range and can be either positive or negative. The expected range of values is given in **Table 1**.

**Table 1** One-bond phosphorus spin–spin coupling constants $^{1}J_{PX}$

| *Structural class (or structure)* | *$^{1}J$ (HZ)$^{a}$* | *Structural class (or structure)* | *$^{1}J$ (Hz)$^{a}$* |
|---|---|---|---|
| P(II) | | P(IV) (continued) | |
| —$PH_2^-$ | 139 | >P(O)H | 460–1030 |
| >PH | 180–225 | —P(M)F (with \ and /) | 1000–1400 |
| >P—C | 0–45 | (M=O, S) | |
| >PF | 820–1450 | >P(S)H | 490–650 |
| >P—P< | 100–400 | P(V) | |
| P(IV) | | —PH (with \| and /\|) | 700–1000 |
| —$PH^+$ (with \ and /) | 490–600 | P—F (with \| and /\|) | 530–1100 |
| —PC (with \ and /) | 50–305 | $PF_5$ | 938 |
| $^{+}P(CH_3)_4$ | +56 | P(VI) | |
| | | $PF_6^-$ | 706 |

$^{a}$For structural classes, only absolute value for $J$ is given.

## Two bond coupling constants: $^{2}J_{PX}$

Two-bond $^{2}J_{PX}$ coupling constants may be either positive or negative and are generally smaller than one-bond coupling constants (**Table 2**). The $^{2}J_{PCH}$ and $^{2}J_{PCF}$ constants are stereospecific and a Karplus-like dihedral dependence to the two-bond coupling constant (H or F)–C–P–X (X = lone pair or heteroatom) has been found. Thus in the *cis*- and *trans*-phosphorinanes, the $^{2}J_{PC}$ constants are 0.0 and 5.1 Hz in the cis- and trans-isomers, respectively.

## Three-bond coupling constants, $^{3}J_{PX}$

Three-bond coupling constant, $^{3}J_{PX}$, through intervening C, N, O, or other heteroatoms are generally

**Table 2** Two-bond phosphorus spin–spin coupling constants $^{2}J_{PX}$

| Structural class ( or structure) | $^{3}J$ (HZ)[a] | Structural class (or structure) | $^{3}J$ (HZ)[a] |
|---|---|---|---|
| P(III) | | P(IV) (Continued) | |
| >PCH | 0–18 | —$P^{+}$CH | 12–18 |
| $P(CH_3)_3$ | +2.7 | —$P^{+}$CC | 0–40 |
| >PCF | 40–149 | $P^{+}(C_2H_5)_4$ | −4.3 |
| $P(CF_3)_3$ | 85.5 | —POC | −6 |
| >PNH | 13–28 | P(V) | |
| >PCC | 12–20 | >PCH | 10–18 |
| $P(C_2H_5)_3$ | +14.1 | >PCF | 124–193 |
| >POC | 10–12 | P(VI) | |
| >PXP (X=S, C) | 70–90 | —PCF | 130–160 |
| P(IV) | | | |
| >P(O)CH | 7–30 | | |
| $P(O)(CH_3)_3$ | −12.8, −13.4 | | |

[a]For structural classes, only the absolute value for $J$ is given.

<20 Hz (**Table 3**). The dihedral-angle dependence of vicinal $^{3}J_{POCH}$ coupling $^{3}J_{PCCH}$ and $^{3}J_{PCCC}$ has been demonstrated. The curves may be fitted to the general Karplus equation

$$J(\phi) = A\cos^2\phi + B\cos\phi + C \qquad [7]$$

where $\phi$ is the dihedral angle and $A$, $B$ and $C$ are constants for the particular molecular framework. Caution is recommended when attempting to apply these Karplus equations and curves to classes of phosphorus compounds that have not been used in establishing these relationships because separate correlations and values for the constants $A$, $B$, and $C$ in Equation [7] probably exist for each structural class. In all cases, a minimum in these Karplus curves is found at ~90°.

## Applications to nucleic acid structure

The Karplus-like relationship between HCOP and CCOP dihedral angles and $^{3}J_{HP}$ and $^{3}J_{CP}$ three-bond coupling constants, respectively, has been used to determine the conformation about the ribose–phosphate backbone of nucleic acids in solution. Torsional angles about both the C3′–O3′ and C5′–O5′ bonds in 3′,5′-phosphodiester linkages have been determined from the coupled $^{1}$H and $^{31}$P NMR spectra.

Within the limitations just described for the general application of the Karplus relationship, the best Karplus relationship for the nucleotide H3′–P coupling constants appears to be

$$J = 15.3\cos^2(\theta) - 6.1\cos(\theta) + 1.6$$

**Table 3** Three-bond phosphorus spin–spin coupling constants $^{3}J_{PX}$

| Structural class ( or structure) | $^{3}J$ (HZ)[a] | Structural class (or structure) | $^{3}J$ (HZ)[a] |
|---|---|---|---|
| P(III) | | P(IV) (Continued) | |
| >POCH | 0–15 | >P(O)OCH | 0–13 |
| $P(OCH_3)_3$ | 10.8–11.8 | $P(O)(OCH_3)_3$ | 10.2–11.4 |
| >PCCH | 10–16 | >P(M)CCH | 14–25 |
| >PNCH | 3–14 | (M=O, S) | |
| $P[N(CH_3)_2]_3$ | 8.8–9.0 | —POCC | |
| P(IV) | | P(V) | |
| —$P^{+}$OCH | 7–11 | >PCCH | 20–27 |
| —$P^{+}$CCH | 15–22 | >POCH | 12–17 |
| —PSCH | 16–20 | | |

[a]For structural classes, only the absolute value for $J$ is given.

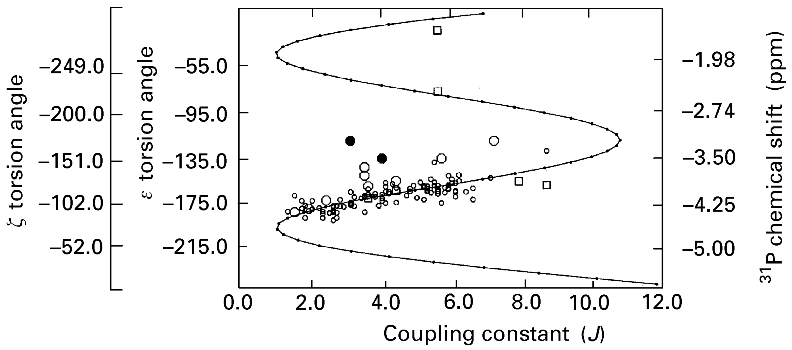



**Figure 3** Plot of $^{31}$P chemical shifts for duplex oligonucleotide sequences (O) and an actinomycin D bound d(CGCG)$_2$ tetramer complex (□) with measured $J_{H3'-P}$ coupling constants (●, phosphates in a tandem GA mismatch decamer duplex which shows unusual, slowly exchanging signals). Also shown are the theoretical $\varepsilon$ and $\zeta$ torsion angles (solid curve) as a function of the coupling constant derived from the Karplus relationship ($\varepsilon$) and the relationship $\zeta = -317 - 1.23\varepsilon$. $^{31}$P chemical shifts are reported relative to trimethyl phosphate. Reproduced with permission.

From the H3′–C3′–O–P torsional angle $\theta$, the C4′–C3′–O–P torsional angle $\varepsilon$ ($= -\theta - 120°$) may be calculated.

The $J_{H3'-P}$ coupling constants in larger oligonucleotides cannot generally be determined from the coupled 1D $^{31}$P or $^{1}$H spectra because of spectral overlap. 2D *J*-resolved long-range correlation pulse sequences can be used to overcome this limitation. The Bax–Freeman selective 2D *J* experiment with a DANTE (delays alternating with nutations for tailored excitation) sequence for a selective 180° pulse on the coupled protons can be readily implemented on most spectrometers. This is particularly useful for measuring phosphorus–H3′ coupling constants in duplex fragments, which can vary from ~1.5 to 8 Hz in duplexes as large as tetradecamers. There is a strong correlation ($R = -0.92$) between torsional angles C4′–C3′–O3′–P ($\varepsilon$) and C3′–O3′–P–O5′ ($\zeta$) in the crystal structures of various duplexes. Thus both torsional angles $\varepsilon$ and $\zeta$ can often be calculated from the measured P–H3′ coupling constant.

Coupling constants of both 5′ protons are analysed in order to determine conformations about the C5′–O bond. Unfortunately, these $\beta$ torsional angles have in practice been generally unobtainable even in moderate-length duplexes. Selective 2D *J*-resolved spectra generally fail for H4′, H5′, or H5″ coupling to $^{31}$P because the spectral dispersion between these protons is so limited. However, with either $^{13}$C labelling or even natural abundance $^{13}$C methods, it is possible to measure not only the $^{1}$H–$^{31}$P but also the $^{13}$C–$^{31}$P coupling constants. Analysis of the 2D multiplet pattern, especially the 'E. COSY' pattern of the $^{1}$H–$^{13}$C HSQC spectrum, has allowed extraction of many carbon (C3′,C4′,C5′) and proton (H3′,H4′,H5′,H5″) coupling constants to phosphorus. The larger line widths of longer duplexes limit measurement of the small coupling constants.

As shown in **Figure 3**, the Karplus relationship provides for four different torsional angle solutions for each value of the same coupling constant. Although all four values are shown in **Figure 3**, the limb which includes $\varepsilon$ values between 360° and –270° is sterically inaccessible in nucleic acids. As shown in **Figure 3**, nearly all of the phosphates for normal Watson–Crick duplexes fall along only a single limb of the Karplus curve. Thus, for 'normal' B-DNA geometry, there is an excellent correlation between the phosphate resonances and the observed torsional angle, while phosphates that are greatly distorted in their geometry must be more carefully analysed.

It is clear from **Figure 3** that $^{31}$P chemical shifts and coupling constants provide probes of the conformation of the phosphate ester backbone in nucleic acids and nucleic acid complexes. It is important to remember that $^{31}$P chemical shifts are dependent on factors other than torsional angles alone. As noted above, $^{31}$P chemical shifts are very sensitive to bond angle distortions as well. It is quite reasonable to assume that backbone structural distortions as observed in unusual nucleic acid structures also introduce some bond angle distortion as well. Widening of the ester O–P–O bond angle indeed is expected to produce an upfield shift, while narrowing of this bond angle causes a downfield shift, and it is possible that this bond angle effect could account for the anomalous shifts. Indeed, very large $^{31}$P chemical shift variations (~3–7 ppm) are observed in transfer RNA and hammerhead RNA phosphates, and are probably due to bond angle distortions in these tightly folded structures.

Generally the main-chain torsional angles of the individual phosphodiester groups along the oligonucleotide double helix are responsible for sequence-specific variations in the $^{31}$P chemical shifts. In duplex B-DNA, the gauche(−), gauche(−) (g⁻, g⁻; ζ, α) (or $B_I$) conformation about the P–O ester bonds in the sugar phosphate backbone is energetically favoured, and this conformation is associated with a more shielded $^{31}$P resonance. In both duplex and single stranded DNA the trans, gauche(−) (t, g⁻; ζ, α) (or $B_{II}$) conformation is also significantly populated.

The $^{31}$P chemical shift difference between the $B_I$ and $B_{II}$ phosphate ester conformational stages is estimated to be 1.5–1.6 ppm. As the result of this sensitivity to the backbone conformational state, $^{31}$P chemical shifts of duplex oligonucleotides have been shown to be dependent both upon the sequence and the position of the phosphate residue.

The possible basis for the correlation between local helical structural variations and $^{31}$P chemical shifts can be analysed in terms of deoxyribose phosphate backbone changes involved in local helical sequence-specific structural variations. As the helix winds or unwinds in response to local helical distortions, the length of the deoxyribose phosphate backbone must change to reflect the stretching and contracting of the deoxyribose phosphate backbone between the two stacked base pairs. To a significant extent, these changes in the overall length of the deoxyribose phosphate backbone 'tether' are reflected in changes in the P–O ester (as well as other) torsional angles. These sequence-specific variations in the P–O (and C–O) torsional angles may explain the sequence-specific variations in the $^{31}$P chemical shifts.

## $^{31}$P NMR of protein complexes

$^{31}$P NMR spectroscopy has proven to be very useful in the study of various protein complexes. **Table 4** provides an indication of the range of $^{31}$P chemical shifts and the titration behaviour of various phosphoprotein model compounds. Two examples of such studies are described below.

**Table 4** Chemical shifts and pH titration data for representative model compounds

| *Compound* | *Chemical shift (ppm)*[a] | *Titratable*[b] | $pK_a$ |
|---|---|---|---|
| Phosphomonoesters | | | |
| Phosphoserine | 4.6 | + | 5.8 |
| Phosphothreonine | 4.0 | + | 5.9 |
| Pyridoxal phosphate | 3.7 | + | 6.2 |
| Pyridoxamine phosphate | 3.7 | + | 5.7 |
| Flavin mononucleotide | 4.7 | + | ~6.0 |
| Phosphodiesters | | | |
| RNA, DNA, phospholipids | 0 to −1.5 | − | |
| Diphosphodiesters | | | |
| Flavin adenine dinucleotide | −10.8 to −11.3 | − | |
| Phosphotriesters | | | |
| Dialkyl phosphoserine | 0 to −3.0 | − | |
| Phosphoramidates | | | |
| *N*³-Phosphohistidine | −4.5 | − | |
| *N*¹-Phosphohistidine | −5.5 | − | |
| Phosphoarginine | −3.0 | + | 4.3 |
| Phosphocreatine | −2.5 | + | 4.2 |
| Acyl phosphates | | | |
| Acetyl phosphate | −1.5 | + | 4.8 |
| Carbamyl phosphate | −1.1 | + | 4.9 |

[a] All chemical shifts are reported with respect to an external 85% $H_3PO_4$ standard; upfield shifts are given a negative sign.

[b] Titrability: + indicates that changes are observed in the chemical shift on changes in pH: for phosphomonoesters this change is 4 ppm; for phosphoramidates 2.5 ppm; for acyl phosphates 5.1 ppm; − indicates no change observed.

### Ribonuclease A

Secondary ionization of a phosphate monoester produces approximately a 4 ppm down field shift of the $^{31}$P signal. Thus the pH dependence of the $^{31}$P signal of various phosphate monoesters bound to proteins can provide information on the ionization state of the bound phosphate ester. For example, pyrimidine nucleotides, both free in solution and when bound to bovine pancreatic ribonuclease A (RNase A), demonstrate this point. The $^{31}$P chemical shift of free solution cytidine 3′-monophosphate (3′-CMP) follows a simple titration curve, and the ionization constant derived form the $^{31}$P shift variation agrees with potentiometric titration values. The $^{31}$P chemical shift titration curve for the 3′-CMP·RNase A complex, however, cannot be analysed in terms of a single ionization process. Two inflections observed in this titration indicated two ionizations with $pK_1 = 4.7$ and $pK_2 = 6.7$.

These results suggest that the nucleotide binds at around neutral pH in the dianionic ionization state. Thus the 3′-CMP·RNase A complex $^{31}$P resonance is shifted upfield less than 0.3 ppm from the free 3′-CMP between pH 6.5 and 7.5, whereas monoprotonation of the free dianion results in a 4 ppm upfield shift. Furthermore, the addition of the first proton to the nucleotide complex ($pK_2 = 6.0$–$6.7$) must occur mainly on some site other than the dianionic phosphate because the $^{31}$P signal is shifted upfield by only 1–2 ppm. The addition of a second proton

(p$K_1$ = 4.0–5.7) to the complex shifts the $^{31}$P signal further upfield so that at the lowest pH values, the phosphate finally appears to be in the monoanionic ionization state.

On the basis of X-ray and $^{1}$H NMR studies, it is known that the nucleotides are located in a highly basic active site with protonated groups histidine-119, histidine-12 and, probably, lysine-41, quite close to the phosphate. This suggests that p$K_1$ is associated with ionization of a protonated histidine residue which hydrogen bonds to the phosphate. This highly positive active site, which is capable of perturbing the p$K$ of the phosphate from 6 to 4.7, must have one or more hydrogen bonds to the phosphate over the entire pH region. Yet at the pH extrema, little if any perturbation of the $^{31}$P chemical shift is found. Apparently, the $^{31}$P chemical shift of the phosphate esters is largely affected by the protonation state and not by the highly positive local environment of the enzyme.

### Two-dimensional exchange $^{31}$P NMR of phosphoglucomutase

Phosphoglucomutase (PGM) catalyses the interconversion of glucose 1-phosphate and glucose 6-phosphate. The enzyme has 561 residues on a single polypeptide chain with molecular weight 61 600 Da. Catalysis proceeds via a glucose 1,6-bisphosphate intermediate where the formation and breakdown of this intermediate results from two phosphate transfer steps involving a single enzymic phosphorylation site, Ser-116. A metal ion is required for activity and the most efficient metal ion is the physiological activator, $Mg^{2+}$. The phosphate transfer steps are shown below.

$$\mathrm{E_P + Glc1P^*} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{E_D \cdot Glc1P^*6P} \underset{k_6}{\overset{k_{-6}}{\rightleftharpoons}} \mathrm{E_{P^*} + Glc6P} \qquad [8]$$

$$\mathrm{E_{P^*} + Glc1P} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{E_D \cdot Glc1P6P^*} \underset{k_6}{\overset{k_{-6}}{\rightleftharpoons}} \mathrm{E_P + Glc6P^*} \qquad [9]$$

$E_P$ and $E_D$ are the phospho and dephospho forms of the enzyme, respectively, Glc1P is glucose-1-phosphate and Glc6P is glucose-6-phosphate.

Metal-free PGM and complexes with a variety of metal ions, substrates, and substrate analogues have been studied by $^{31}$P NMR. Under conditions where the enzyme is inactive, each of the three enzyme-bound intermediates in the above scheme can be studied.

Exchange processes can be detected by 2D $^{31}$P NMR in addition to the conventional 1D methods. The 2D exchange experiment (NOESY) described by Ernst and co-workers involves three 90° pulses. Nuclei are frequency labelled by a variable delay time ($t_1$) separating the first and second pulses. The mixing time is between the second and third pulses, and the detection of transverse magnetization as a function of time ($t_2$) follows the third pulse. During the mixing time, nuclei labelled in $t_1$ with a frequency corresponding to one site are converted by the exchange processes to a second site and evolve in $t_2$ with the frequency of the second site, giving rise to cross-peaks in the 2D spectrum. A 2D $^{31}$P exchange spectrum of PGM shows cross-peaks indicating exchange between bound Glc6P and free Glc6P and between the two bound phosphorus sites, indicating transfer through free $E_P$ involving a full catalytic cycle (see Eqn [8]).

## Medical applications of $^{31}$P NMR

*In vivo* $^{31}$P NMR and $^{31}$P magnetic resonance imaging are also important applications of this nucleus. $^{31}$P signals from inorganic phosphate, adenosine triphosphate, adenosine diphosphate, creatine phosphate, and sugar phosphates can be observed in whole-cell preparations, intact tissues, and whole bodies and can provide information on the viability of the cells and tumour localization. Low sensitivity continues to be a problem in widespread application of these techniques. Additional details can be found in several of the entries in the Further reading section.

## Conclusions

$^{31}$P NMR has become an indispensable tool in studying the chemistry and reactivity of phosphorus compounds, as well as in studying numerous biochemical and biomedical problems. Newer NMR instrumentation has enormously enhanced the sensitivity of the experiment and allowed 2D NMR studies to provide new means of signal assignment and analysis. Through 2D and 3D heteronuclear NMR experiments it is now possible to unambiguously assign the $^{31}$P signals of duplex oligonucleotides and other phosphate esters. Both empirical and theoretical correlations between measured coupling constants, $^{31}$P chemical shifts, and structural parameters have provided an important probe of the conformation and dynamics of nucleic acids, protein complexes, and small organophosphorus compounds.

## List of symbols

$a^2$ = percentage s character; $J$ = coupling constant; $S$ = overlap integral; $t_1$ = delay time; $t_2$ = observe time in 2D NMR; $T_1$ = spin–lattice relaxation time; $T_2^*$ = time constant for the FID; $\Delta\delta$ = chemical shift difference; $\Delta\theta$ = change in the $\sigma$-bond angle; $\Delta\chi_X$ = electronegativity difference in the P–X bond; $\sigma$ = shielding constant; $\sigma_{iso}$ = isotropic shielding constant; $\sigma_{\parallel}$ = parallel component of shielding constant; $\sigma_{\perp}$ = perpendicular component of shielding constant; $\sigma_{11}, \sigma_{22}, \sigma_{33}$ = components of shielding tensor; $\phi$ = dihedral angle.

*See also:* **Cells Studied By NMR; *In vivo* NMR, Applications, $^{31}$P; NMR Pulse Sequences; Nuclear Overhauser Effect; Nucleic Acids Studied Using NMR; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Parameters in NMR Spectroscopy, Theory of; Perfused Organs Studied Using NMR Spectroscopy; Proteins Studied Using NMR Spectroscopy; Two-Dimensional NMR Methods.**

## Further reading

Burt CT (1987) *Phosphorus NMR in Biology*, pp. 1–236. Boca Raton, FL: CRC Press.

Crutchfield MM, Dungan CH, Letcher LH, Mark V and Van Wazer JR (1967) Topics in phosphorus chemistry. In: Grayson M and Griffin EF (eds) *Topics in Phosphorous Chemistry*, pp. 1–487. New York: Wiley (Interscience).

Gorenstein DG (1984) *Phosphorus-31 NMR: Principles and Applications*, pp. 1–604. Orlando, FL: Academic Press.

Gorenstein, DG (1992) Advances in P-31 NMR. In: Engel, R (ed) *Handbook of Organophosphorus Chemistry*, pp. 435–482. New York: Marcel Dekker.

Gorenstein DG (1994) Conformation and Dynamics of DNA and Protein–DNA Complexes by $^{31}$P NMR, *Chemical Reviews* **94**: 1315–1338.

Gorenstein DG (1996) Nucleic Acids: Phosphorus-31 NMR. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, pp. 3340–3346. Chichester: Wiley.

Karaghiosoff K (1996) Phosphorus-31 NMR. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance,* pp. 3612–3618. Chichester: Wiley.

Mavel G (1973) *Annual Reports on NMR Spectroscopy* **5B**: 1–350.

Quin LD and Verkade JG (1994) *Phosphorus-31 NMR Spectral Properties in Compound Characterization and Structural Analysis*, p. 1. New York: VCH.

Tebby JC (1991) *Handbook of Phosphorus-31 Nuclear Magnetic Resonance Data*, p. 1. Boca Raton FL: CRC Press.

Verkade JG and Quin LD (1987) *Phosphorus-31 NMR Spectroscopy in Stereochemical Analysis; Organic Compounds and Metal Complexes*, pp. 1–455. Deerfield Beach, FL: VCH.

# Palladium NMR, Applications

*See* **Heteronuclear NMR Applications (Y–Cd).**

# Parameters in NMR Spectroscopy, Theory of

**GA Webb**, University of Surrey, Guildford, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Theory |

## Introduction

High-resolution NMR provides spectra that consist of a number of lines and bands whose frequency, relative intensity and shape may be analysed to yield molecular parameters. The NMR parameters in questions are the nuclear shielding, $\sigma$, which describes the shielding of the nucleus from the applied magnetic field by the surrounding electrons and gives rise to chemical shifts; $J$, which relates to nuclear spin–spin coupling and depends upon relative nuclear orientations; and the times $T_1$ and $T_2$ which refer to the relaxation processes encountered by the nuclei excited in the NMR experiment.

Both the nuclear shielding and spin–spin coupling interactions are interpreted within the framework of quantum chemistry, whereas a quasi-classical form of mechanics is usually adopted to describe the nuclear relaxation interactions.

## Nuclear shielding (chemical shifts)

For an NMR experiment the basic resonance condition is given as

$$\omega = \gamma B_0 \qquad [1]$$

where $B_0$ is the applied magnetic field in which the experiment is performed, $\gamma$ is the magnetogyric ratio of the nucleus in question and $\omega$ is the angular frequency of the radiation producing the NMR transition. From this expression it follows that all nuclei with a given value of $\gamma$, e.g. protons, will produce a single absorption in the NMR spectrum. In such a situation NMR spectroscopy would not be of much chemical interest. In reality the expression for the resonance condition needs to be modified to include the fact that the value of the magnetic field experienced by the resonating nuclei is usually less than $B_0$ owing to shielding of the nucleus in a molecule by the surrounding electrons. Thus the expression for the resonance condition becomes

$$\omega = \gamma B_0(1 - \sigma) \qquad [2]$$

where $\sigma$ is the nuclear shielding. In NMR experiments the resonance frequencies are normally reported relative to that of a given nucleus in a standard molecule added to the experimental sample as a reference. The shielding difference, or chemical shift $\delta$, is then defined as the difference in shielding between the given nucleus in the reference compound, $\sigma_{ref}$, and that of the nucleus of interest, $\sigma_{sample}$. Namely

$$\delta = (\sigma_{ref} - \sigma_{sample})$$

From which it follows that a shift of resonance to high frequency, denoted by an increase in $\delta$, corresponds to a decrease in $\sigma_{sample}$.

In seeking a molecular interpretation for $\sigma$ it is important to realize that the nuclear shielding is represented by a second-rank tensor. Many NMR experiments are performed on nonviscous solutions, or sometimes on gaseous samples, in which case rapid, and random, molecular motion ensures that the nuclear shielding experienced is the scalar corresponding to one-third of the trace of the tensor.

NMR measurements taken on the solid and liquid crystal phases can yield values for the individual components of the shielding tensor and its anisotropy, $\Delta\sigma$. For linear and symmetric-top molecules,

$$\Delta\sigma = \sigma_{\parallel} - \sigma_{\perp} \qquad [3]$$

where $\sigma_{\parallel}$ refers to the shielding component along the major molecular axis and $\sigma_{\perp}$ is that in the direction perpendicular to it. For less symmetrical molecules,

$$\Delta\sigma = \sigma_{\alpha\alpha} - \frac{1}{2}(\sigma_{\beta\beta} + \sigma_{\gamma\gamma}) \qquad [4]$$

where the $\sigma_{ii}$ are the principal tensor components taken in accordance with the convention $\sigma_{\alpha\alpha} > \sigma_{\beta\beta} > \sigma_{\gamma\gamma}$.

The first report on the theory of nuclear shielding appeared in 1950; since then many reports have appeared of attempts to calculate shieldings, most of them within the framework of molecular orbital

(MO) theory. Some of the earlier results, particularly those based upon semiempirical MO methods, are at best indicative of shielding trends in series of closely related molecules. In general these are unsuitable for predictive purposes. In recent years this situation has changed dramatically and *ab initio* MO calculations of nuclear shielding are routinely providing satisfactory results. In principle, quantum chemistry can provide a full account of all molecular properties. In practice, various approximations are introduced into the calculations to make them tractable. Such approximations tend to produce limitations on the results obtained from the calculations. For example, calculations at the Hartree–Fock (HF) level involve a single determinant for a rigid isolated molecule; consequently, the effects of electron correlation, variations in geometry and media influences on nuclear shielding are ignored. Normally such effects are considered separately as are possible relativistic effects on the shielding of heavier nuclei. Calculations of molecular magnetic properties, such as nuclear shielding, can suffer from all of these limitations and an additional one known as the gauge problem. This arises from the use of perturbation theory to describe the rather small contribution to the total electronic energy of the molecule provided by the applied magnetic field in the NMR experiment. The magnetic perturbation is described by the orbital angular momentum operator. Since this operator is not invariant with respect to translations, its influence depends upon the position at which it is evaluated. Consequently, the result obtained for the calculated nuclear shielding depends upon the choice of origin for the calculation. This theoretical artefact has to be dealt with before comparison takes place between experimental and theoretical shielding data.

One way to combat the gauge problem in nuclear shielding calculations is to employ large basis sets in calculations using the coupled Hartree–Fock (CHF) approach. If smaller basis sets are employed, the shielding results obtained are gauge-dependent unless the gauge origin is taken to be at the nucleus in question; these are referred to as common-origin calculations. An example of $^{13}C$ nuclear shielding calculations of this type is provided by buckminsterfullerene, $C_{60}$: all of the carbon atoms are equivalent in this molecule and thus symmetry arguments can be used to reduce the number of integrals to be evaluated. If a relatively modest basis set, such as 6-13G*, is used in the calculation of the nuclear shielding, then about ten days of CPU time is required on a DEC 8400 computer. Consequently, it seems unlikely that common-origin nuclear shielding calculations will become widely affordable.

An alternative to using large basis sets to overcome the gauge problem is to introduce gauge factors either into the atomic orbitals of the basis set or into the MOs of a CHF calculations of nuclear shielding. The inclusion of gauge factors in the atomic orbitals used gives rise to the gauge-included atomic orbital (GIAO) method. In contrast the IGLO (individual gauges for localized orbitals) method employs individual gauge origins for different localized molecular orbitals. Both the GIAO and IGLO methods are referred to as local origin variants of the CHF method.

An alternative to the CHF calculations of second-order magnetic properties is to use the random-phase approximation within the equations of motion procedure. This has developed into a method using localized MOs with local origins (LORG). The LORG method results in a localization of the MOs used and provides a pathway to the decomposition of the calculated nuclear shielding into individual local bond and bond–bond contributions. Thus the LORG and IGLO methods of calculating nuclear shieldings are analogous to each other. The results of some $^{13}C$ shieldings and their anisotropies, produced by IGLO, LORG and GIAO calculations, are given in **Table 1**.

The results given are obtained by the use of medium-sized basis sets; e.g. sets of triple zeta quality with a set of d polarization functions for the heavy atoms. In general, the calculated and experimental results are in satisfactory agreement. In comparing the relative merits of the GIAO, IGLO and LORG methods, it appears that the GIAO procedure is the more efficient in terms of the convergence of the shielding value with respect to the size of basis set used. However, the IGLO and LORG calculations

**Table 1** Comparison of some $^{13}C$ shieldings and their anisotropies $\Delta\sigma$ (in ppm) produced by IGLO and GIAO calculations and experimental values

| *Molecule and* $\Delta\sigma$ | *IGLO* | *LORG* | *GIAO* | *Experimental* |
|---|---|---|---|---|
| $CH_4$ | 196.7 | 196.0 | 193.0 | 195.1 |
| HCN | 72.9 | 77.3 | 74.8 | 82.1 |
| $\Delta\sigma$ | 306 | 301 | 304.8 | 316.3 ± 1.2 |
| $C_2H_2$ | 116.4 | 122.3 | 118.3 | 117.2 |
| $\Delta\sigma$ | 243.3 | 235.0 | 241.0 | 240 ± 5 |
| CO | −6.0 | – | −21.3 | 1.0 |
| $\Delta\sigma$ | 420.0 | – | 439.0 | 405.5 ±1.4 |
| $C_2H_6$ | 183.5 | 184.7 | 181.2 | 180.0 |
| $\Delta\sigma$ | 13.5 | 8.0 | 11.3 | – |
| $CH_3OH$ | 157.2 | 145.7 | 134.6 | 136.6 |
| $\Delta\sigma$ | – | 60.5 | 77.3 | 63.0 |
| $H_2CO$ | −3.8 | 4.0 | 2.6 | −8.0 |
| $\Delta\sigma$ | 183.8 | 183.0 | 196.5 | – |

produce shielding contributions that may be attributed to specific molecular regions. Since the GIAO, IGLO and LORG calculations can all be performed with different levels of basis set quality, the results obtained are found to be dependent upon the choice of basis set. An example of the dependence of $^{13}C$ shieldings calculated by the GIAO method upon the choice of geometry and basis set is provided in **Table 2**.

The use of experimental geometries is indicated by NOOPT (none optimized geometry); also included are optimized geometries from *ab initio* MO calculations using 4-31G and 431G**basis sets, whereas the $^{13}C$ shieldings are calculated with 3-21G, 4-31G and 4-31G** basis sets. The best agreement with the experimental shieldings is given by the calculations using experimental molecular geometries and the 4-31G** basis set; in this case the average least-squares error for the set of molecules studied is 2.7 ppm.

In general, GIAO, LORG and IGLO calculations are capable of producing shieldings for nuclei from the first and second long rows of the periodic table to within about 3 or 4% of an element's shielding range. Thus these calculations can be used for predictive purposes as well as providing some information on the molecular electronic factors that determine the extent of nuclear shielding and its variations.

Another method of tackling the gauge problem in nuclear shielding calculations is to employ individual gauges for atoms in molecules (IGAIM). This procedure differs from the GIAO, IGLO and LORG procedures in that the gauge origins in IGAIM are

**Table 2** Calculated and observed isotropic $^{13}C$ chemical shifts (in ppm from $CH_4$) of the resonant nuclei (*C) using NOOPT/4-31G, 4-31G/4-31G, NOOPT/3-21G, NOOPT/4-31G** and 4-31G**/4-31G** basis sets

| *Molecule* | *NOOPT/4-31G* | *4-31G/4-31G* | *NOOPT/3-21G* | *NOOPT/4-31G*** | *4-31G**/4-31G*** | *Experimental* |
|---|---|---|---|---|---|---|
| $CH_4$ | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| $C_2H_6$ | 4.2 | 4.8 | 2.9 | 5.0 | 5.5 | 8.0 |
| $C_2H_4$ | 131.2 | 127.4 | 119.2 | 125.4 | 121.5 | 125.4 |
| $^*CH_3CH_2CH_3$ | 16.1 | 15.4 | 13.5 | 16.6 | 16.3 | 17.7 |
| $CH_3{}^*CH_2CH_3$ | 16.7 | 13.8 | 13.7 | 18.6 | 16.4 | 18.2 |
| *cis*-$^*CH_3CH{=}CHCH_3$ | 12.4 | 12.9 | 10.6 | 12.3 | 12.6 | 12.7 |
| *cis*-$CH_3{}^*CH{=}CHCH_3$ | 130.8 | 130.1 | 119.0 | 126.3 | 124.9 | 125.9 |
| *trans*-$^*CH_3CH{=}CHCH_3$ | 19.9 | 18.6 | 17.4 | 19.5 | 18.4 | 19.4 |
| *trans*-$CH_3{}^*CH{=}CHCH_3$ | 134.2 | 129.6 | 121.2 | 129.3 | 124.5 | 127.2 |
| cyclo-$C_3H_6$ | −2.0 | −2.6 | −2.6 | −1.4 | −2.2 | 0.1 |
| cyclo-$C_6H_{12}$ | 29.5 | 25.5 | 24.9 | 30.3 | 26.2 | 29.9 |
| $C_6H_6$ | 133.2 | 132.1 | 119.7 | 129.2 | 127.3 | 130.0 |
| $^*CH_2{=}CHCH_3$ | 123.7 | 128.3 | 112.8 | 118.2 | 115.1 | 117.5 |
| $CH_2{=}{}^*CHCH_3$ | 138.7 | 138.0 | 124.9 | 133.7 | 131.2 | 137.8 |
| $CH_2{=}CH{}^*CH_3$ | 17.5 | 22.9 | 15.2 | 17.1 | 19.3 | 20.8 |
| $^*CH{\equiv}CCH_3$ | 74.5 | 74.6 | 69.2 | 69.6 | 69.0 | 69.0 |
| $CH{\equiv}{}^*CCH_3$ | 79.2 | 80.4 | 71.7 | 74.1 | 74.5 | 82.0 |
| $CH{\equiv}C{}^*CH_3$ | 5.0 | 4.0 | 4.3 | 4.9 | 3.9 | 4.0 |
| Toluene (C-1) | 141.3 | 140.2 | 126.3 | 137.7 | 136.4 | 140.8 |
| Toluene (C-2) | 133.3 | 132.5 | 119.9 | 128.9 | 128.1 | 131.4 |
| Toluene (C-3) | 134.5 | 133.6 | 120.6 | 130.6 | 129.6 | 132.2 |
| Toluene (C-4) | 130.2 | 129.3 | 117.1 | 126.1 | 125.1 | 128.5 |
| Toluene ($CH_3$) | 21.9 | 22.1 | 19.6 | 21.1 | 21.5 | 24.3 |
| $^*CH{\equiv}C{-}C{\equiv}CH$ | 68.4 | 69.8 | 63.6 | 65.1 | 66.0 | 69.3 |
| $CH{\equiv}{}^*C{-}C{\equiv}CH$ | 71.9 | 73.0 | 65.4 | 66.5 | 66.8 | 71.0 |
| cyclo-$C_3H_4(CH_2)$ | – | −3.7 | – | – | – | 3.0 |
| cyclo-$C_3H_4({=}C)$ | – | 117.1 | – | – | – | 108.0 |
| Averaged least-squares error (ppm) | 3.16 | 3.24 (without cyclo-$C_3H_4$) | 7.34 | 2.72 | 3.20 | – |
| Correlation coefficient | 0.998 | 0.995 | 0.997 | 0.998 | 0.999 | |
| Slope | 0.96 | 0.97 | 1.06 | 1.00 | 1.02 | |

determined by properties of the charge density in real space rather than by the behaviour of the chosen basis functions in the Hilbert space of the molecular wavefunction. The results of some IGAIM $^{13}$C shielding calculations are given in **Table 3**, where they are compared with the results obtained from conventional CHF calculations using the same basis set.

In the CHF calculations, the common gauge origin is placed at the nucleus whose shielding is being deduced. **Table 3** shows that the IGAIM results are in much better agreement with experiment than those produced by the CHF calculations. The absence of electron correlation effects from HF calculations is most noticeable in cases of 'electron-rich' molecules containing, for example, multiple bonding and lone pair electrons. It is possible to enhance the local origin methods for calculating nuclear shieldings by including some electron correlation effects. The GIAO method has been extended by means of many-body perturbation theory (MBPT). The results of some GIAO and GIAO-MBPT calculations of $^{17}$O shieldings are compared in **Table 4**, where the effect of including electron correlation is seen to lead to an increase in the calculated values of the shieldings. This increase usually results in a closer agreement between the calculated and experimental shieldings.

Electron correlation effects have been included in the IGLO method by means of a nonperturbative multiconfiguration extension to give the MC-IGLO method. **Table 5** shows the results of some MC-IGLO calculations of $^{1}$H, $^{13}$C, $^{17}$O, $^{19}$F and $^{31}$P nuclear shieldings in comparison with comparable results from experiment and from the self-consistent fixed (SCF)-IGLO method.

The effects of electron correlation on the calculated nuclear shieldings are shown to be small for methane and phosphine but much more significant for fluorine, carbon monoxide and ozone, which are 'electron rich' molecules. For the central oxygen atom of ozone, the effect of including electron correlation in the shielding calculations is to produce an increase by over 2000 ppm.

The LORG method of calculating nuclear shieldings has been combined with the second-order polarization propagator (SOPPA) technique to produce the second-order LORG or SOLO procedure. The results of some LORG and SOLO $^{15}$N shielding calculations are compared with experiment and with some IGLO results in **Table 6**.

The conjugated heterocycles chosen for the study represent cases where electron correlation effects are predicted to be significant. In general, the inclusion of electron correlation leads to an increase in the calculated nitrogen shieldings and usually to an improved agreement with the experimental results. A comparison of the LORG and SOLO data given in **Table 6** shows root mean square errors of 49.2 and 19.9 ppm, respectively.

**Table 3** Comparison of absolute $^{13}$C shieldings in (ppm) calculated by the IGAIM method and the CHF procedure, using the same 6-31G**(2d, 2p) basis set, with experimental values taken as thermal averages at 300 K in the limit of zero gas density

| *Molecule* | *IGAIM* | *CHF* | *Experimental* |
|---|---|---|---|
| $CH_4$ | 197.4 | 198.5 | 195.1 |
| HCN | 79.9 | 89.5 | 82.1 |
| $C_2H_2$ | 119.3 | 127.0 | 117.2 |
| $C_2H_4$ | 66.4 | 73.4 | 64.5 |
| $C_2H_6$ | 186.3 | 192.3 | 180.9 |
| $C_3H_4$ (C-1) | 119.5 | 130.2 | 115.2 |
| $C_3H_4$ (C-2) | –34.8 | –22.4 | –29.3 |
| $C_6H_6$ | 61.5 | 82.1 | 57.9 |
| CO | –7.4 | –11.9 | 1.0 |
| $CO_2$ | 57.9 | 78.9 | 58.8 |
| $CS_2$ | –41.1 | 51.9 | –8.0 |
| CSO | 21.9 | 78.2 | 30.0 |
| $CH_3NH_2$ | 167.0 | 173.9 | 158.3 |
| $CH_3OH$ | 148.0 | 155.8 | 136.6 |
| $CH_3F$ | 130.2 | 140.0 | 116.8 |
| $CF_4$ | 86.0 | 122.3 | 64.5 |
| HCOOH | 32.2 | 50.2 | 23.7 |

**Table 4** Comparison of some $^{17}$O shieldings (in ppm) produced by GIAO and GIAO-MBPT calculations and experimental values

| *Molecule* | *GIAO* | *GIAO-MBPT* | *Experimental* |
|---|---|---|---|
| $H_2O$ | 323.18 | 339.79 | 357 |
| $H_2O_2$ | 139.01 | 150.88 | 134 |
| CO | –113.47 | –54.06 | –40.1 ± 17.2 |
| $H_2CO$ | –471.40 | –345.02 | –312.1 |
| $CH_3OH$ | 341.55 | 354.41 | 344.9 |
| $CO_2$ | 200.37 | 236.37 | 243.4 |
| $OF_2$ | –471.13 | –465.53 | –473.1 |
| NNO | 107.54 | 192.12 | 200.5 |

**Table 5** Comparison of some $^{1}$H, $^{13}$C, $^{17}$O, $^{19}$F and $^{31}$P shieldings (in ppm) produced by SCF-IGLO and MC-IGLO calculations and experimental values

| *Molecule* | *Nucleus* | *SCF-IGLO* | *MC-IGLO* | *Experimental* |
|---|---|---|---|---|
| $CH_4$ | C | 193.8 | 198.4 | 198.7 |
| | H | 31.22 | 31.13 | 30.61 |
| $PH_3$ | P | 583.4 | 598.2 | 594.4 |
| | H | 29.43 | 29.65 | 29.28 |
| $F_2$ | F | –165.3 | –204.3 | –192.8 |
| CO | C | –23.4 | 13.4 | 3.0 |
| | O | –83.9 | –36.7 | –42.3 |
| $O_3$ | O (central) | –2730.1 | –657.7 | –724.0 |
| | O (terminal) | –2816.7 | –1151.8 | –1290.0 |

**Table 6** Comparison of some $^{15}N$ shieldings (in ppm) produced by IGLO, LORG, and SOLO calculations on some conjugated heterocycles and experimental values

| *Molecule* | *IGLO* | *LORG* | *SOLO* | *Experimental values* |
|---|---|---|---|---|
| Sym-Triazine | −41 | −33 | −28 | −39 |
| Pyrimidine | −71 | −58 | −45 | −51 |
| Pyridine | −104 | −94 | −72 | −73 |
| Pyrazine | −121 | −136 | −102 | −90 |
| Sym-Tetrazine | −221 | −213 | −159 | −141 |
| Pyridazine | −240 | −235 | −197 | −156 |
| 1,2,4-Triazine | | | | |
| N-3 | | −76 | −42 | −54 |
| N-2 | | −171 | −151 | −134 |
| N-1 | | −255 | −207 | −178 |

Density functional theory (DFT) is an alternative to HF methods for describing molecular electronic structure. Electron correlation effects are explicitly included in DFT calculations. Coupled DFT (CDFT), together with the IGLO method, has been used in some nuclear shielding calculations and some results for $^{13}C$, $^{15}N$, $^{17}O$, $^{19}F$ and $^{31}P$ are shown in **Table 7**. For comparison purposes, the results of some GIAO calculations, not including electron correlation, and experimental results are given. The CDFT results are seen to be in much better agreement with experiment than are those from the GIAO calculations.

## Spin–spin couplings

Many NMR signals appear as multiplets, the structure of which arises from spin–spin coupling interactions with other nuclei in the molecule. The separation between adjacent members of a multiplet can give the value of *J*, the spin–spin coupling interaction between the spin coupled nuclei. For nuclei whose spin is $\frac{1}{2}$, the relative signal intensities of the members of a given first-order multiplet are given by the factors of a binomial expansion. If A and B are the two spin-$\frac{1}{2}$ coupled nuclei then the NMR signal for A will consist of a multiplet with $n + 1$ lines due to spin–spin coupling to $n$ equivalent B nuclei, provided the chemical shift between A and B is large relative to $J_{AB}$.

**Table 7** Comparison of some $^{15}C$, $^{15}N$, $^{17}O$, $^{19}F$ and $^{31}P$ shieldings in (ppm) produced by GIAO and CDFT calculations and experimental values

| *Molecule* | *Nucleus* | *GIAO* | *CDFT* | *Experimental* |
|---|---|---|---|---|
| PN | P | −15.8 | 42.1 | 53 |
| | N | −409.4 | −347.3 | −349 |
| $P_2H_2$ | P | −294.2 | −190.9 | −166 |
| CO | C | −8.0 | −0.3 | 1 |
| | O | −61.3 | −63.4 | −42.3 |
| NNO | N (terminal) | 89.0 | 97.0 | 99.5 |
| | N(central) | −2.0 | 5.9 | 11.3 |
| | O | 219.4 | 185.4 | 200.5 |
| $H_2O_2$ | O | 191.5 | 157.2 | 133.9 |
| $N_2$ | N | −80.0 | −69.3 | −61.0 |
| $N_2CO$ | C | 14.2 | −12.3 | −1 |
| | O | −406.2 | −362.6 | −312.1 |
| $F_2$ | F | −181.4 | −197.8 | −193.8 |

As for $\sigma$, the value of *J* depends upon the chemical environment of the nuclei concerned. Hence values of *J* are of use in molecular structure determinations. Unlike the case for nuclear shieldings, values of *J* are independent of the magnitude of the applied magnetic field used in the NMR experiment; thus the gauge problem does not arise when considering quantum-chemical calculations of *J*. Nuclear spin–spin couplings arise from indirect interactions between the spin, *I*, of neighbouring nuclei. The spin orientation information is transmitted from one nucleus to the other by means of both bonding and nonbonding electrons encountered on the spin coupling pathway.

Values of *J* are usually given in Hz as is apparent from the following definition of the energy, $E_{AB}$, of the coupling interaction between nuclei A and B:

$$E_{\mathrm{AB}} = hJ_{\mathrm{AB}}I_{\mathrm{A}}I_{\mathrm{B}} \qquad [5]$$

As in the case of nuclear shielding, $J_{AB}$ is a scalar quantity; an estimate of the anisotropy of the corresponding second-rank tensor may be forthcoming from measurements on oriented samples. The theoretical aspects of spin–spin coupling are based upon three types of electron-coupled interactions between the electrons and nuclei of the molecule concerned. Normally the largest of these is the contact (C) interaction between the electron and nuclear spins; the second one is a magnetic dipolar (D) interaction between the electron and nuclear spins; finally there is the orbital (O) interaction between the magnetic field produced by the orbital motion of the electrons and the nuclear magnetic dipole.

Accurate calculations of spin–spin couplings provide a challenge to the theoretician. Reliable results are difficult to obtain for molecules of chemical interest, because spin–spin couplings rely upon subtle aspects of molecular electronic structure. Consequently, a deeper understanding of the relationships between spin–spin couplings and molecular structure could considerably enhance the application of high-resolution NMR spectroscopy to the elucidation of molecular electronic structure. At present the theoretical analysis of spin–spin couplings is advancing

in two different directions. For small molecules with light atoms, i.e. those up to the second row, highly accurate *ab initio* MO calculations are being applied. Alternatively, rather simple semiempirical calculations are used to provide some understanding of possible relationships between physical phenomena and experimental data.

At the HF level, the C contribution to spin–spin couplings is the most difficult to evaluate accurately owing to the poor description provided of the electron spin densities at the coupled nuclei. Consequently, it becomes necessary to include electron correlation effects to provide accurate calculations of spin–spin couplings. Many-body perturbation theory can be used to introduce some electron correlation into calculations of the C contribution to spin–spin couplings. Using this approach for some first-row hydrides, where the C contribution is expected to dominate, satisfactory agreement is found between calculated and observed values of one-bond couplings. However, the calculated values of $^2J$(H–H) are much too large, which suggests that electron correlation effects beyond second order are important in determining the magnitudes of spin–spin couplings.

The use of multiconfiguration linear response (MCLR) theory is another approach to the calculation of spin–spin coupling interactions. As is usual for *ab initio* MO calculations, the results obtained are found to be basis set dependent. In general, satisfactory agreement with the available experimental data is achieved.

Other *ab initio* MO calculations of spin–spin couplings include those based upon polarization propagator methods, e.g. RPA, SOPPA and the coupled cluster single and double polarization propagator approximation (CCSDPPA). These three methods have been used to calculate the C contributions to the values of $^1J$(C–H) and $^2J$(H–H) for methane as functions of bond length variation in the region of the equilibrium geometry, as shown in **Figures 1** and **2**, where $S_1$ represents the symmetric stretching coordinate. In the case of the CCSDPPA result, about 91% of the correlation contribution to the value of $^1J$(C–H) is recovered, whereas the corresponding figure for the SOPPA calculation is about 79%. For the calculations on $^2J$(H–H), the corresponding recoveries are 88% and 79% for the CCSDPPA and SOPPA methods, respectively.

Semiempirical MO calculations of spin–spin couplings are often used in conjunction with conformational analysis studies. In general, the investigations are based upon a dihedral angle dependence of the $^3J(^{13}C–^1H)$ values. However, calculations of longer-range couplings can also play a role in understanding molecular structure.

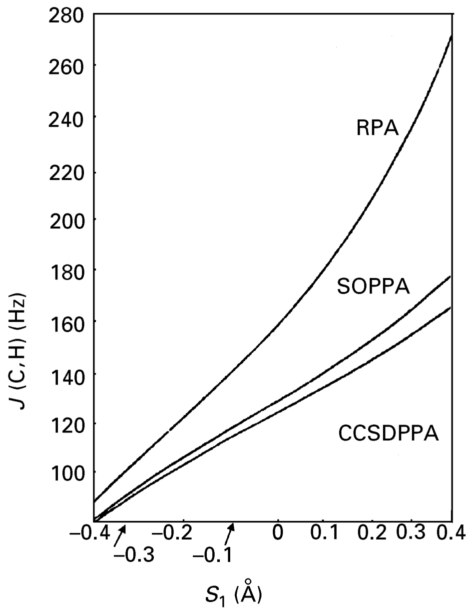



**Figure 1** Dependence of the contact contribution to $^1J$(C–H) on the symmetric stretching coordinate $S_1$ of methane. Results are given at the RPA, SOPPA and CCSDPPA levels of theory.

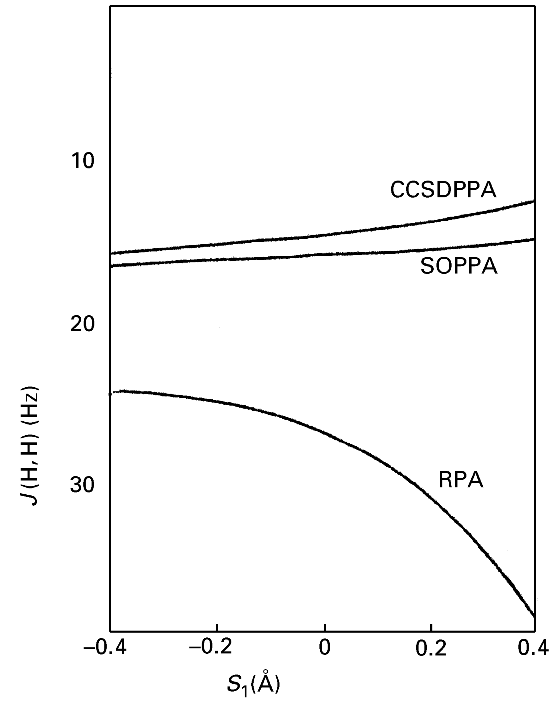



**Figure 2** Dependence of the contact contribution of $^2J$(H–H) on the symmetric stretching coordinate $S_1$ of methane. Results are given at the RPA, SOPPA and CCSDPPA levels of theory.

Self-consistent perturbation theory (SCPT) semiempirical calculations have been used in a study of the effects of the oxygen lone pair electrons on $^1J$(C–C) values in furan derivatives. The results show that the effects of the lone pairs on the spin–spin couplings,

and the changes due to protonation, are similar to those resulting from the lone pair electrons on the nitrogen atom in imines.

## Nuclear spin relaxation

The time taken for nuclear spin relaxation to occur constitutes the third type of chemically interesting NMR parameter. Since NMR is normally observed in the radiofrequency region of the electromagnetic spectrum it involves rather low-energy transitions; consequently spontaneous emission tends to be of negligible importance for NMR relaxation. Nuclear spin relaxation may be characterized by two relaxation times, $T_1$ and $T_2$. The spin–lattice relaxation time, $T_1$, relates to the exchange of nuclear magnetization in a direction parallel to that of the applied magnetic field. $T_2$, the spin–spin relaxation time applies to the exchange of magnetization in directions perpendicular to that of the applied magnetic field.

The ideal NMR line shape is Lorentzian and its full width at half-height, $W_{1/2}$, is controlled by $T_2$:

$$W_{1/2} = \frac{1}{\pi T_2} \qquad [6]$$

For nonviscous liquids, $T_1$ and $T_2$ are usually equal; thus comments made about $T_1$ apply equally to $T_2$.

A number of mechanisms may contribute to nuclear spin relaxation times. These mechanisms operate in chemically distinct ways, such that the identification of which particular mechanism(s) is operative can be of chemical interest. For any mechanism to be operative in producing spin relaxation it must produce an oscillating magnetic field at the nuclear site. The frequency of this local magnetic field must be equal to the resonance frequency of the nucleus to be relaxed. If this situation occurs, then a relaxation transition may be induced.

The microdynamic behaviour of molecules in fluids is attributed to Brownian motion, and the frequency distribution of the components of the local fluctuating magnetic field is expressed by a power spectral density. The component of this spectral density at the resonance frequency is responsible for nuclear relaxation. The magnitude of this component, taken together with the energy of interaction between the nuclear spin system and the molecular motions, determines the value of $T_1$.

In discussing nuclear relaxation phenomena it is normally assumed that the motional narrowing limit applies:

$$\omega_0{}^2\tau_0{}^2 \ll 1$$

where $\omega_0$ refers to the resonance frequency and $\tau_0$ is the correlation time characterizing the appropriate molecular motion. For the motional narrowing limit to apply, the molecules in question must be tumbling rapidly; this implies small molecules in a low-viscosity medium and a relatively high temperature. As shown in **Figure 3**, under these conditions $T_1$ becomes frequency independent and equal to $T_2$. Larger molecules may not satisfy the motional narrowing limit, for example macromolecules, in which case $T_1$ and $T_2$ are almost certain to be unequal and to have different frequency dependences.

Provided the extreme narrowing conditions are satisfied, then the left-hand side of **Figure 3** is the appropriate one for further discussion of the various mechanisms that contribute to $T_1$.

Nuclear magnetic dipole relaxation interactions may occur with other nuclei, or with unpaired electrons. These processes usually dominate the relaxation of spin-$\frac{1}{2}$ nuclei. Both intra- and intermolecular interactions may contribute to dipole–dipole nuclear relaxation times. The value of $T_1$ due to the intramolecular dipole–dipole process is proportional to the sixth power of the internuclear separation. Consequently, this process becomes rather inefficient in the absence of directly bonded magnetic nuclei. However, it follows that a measurement of $T_1$ can be provide an estimate of internuclear separation that can be of chemical interest. The nuclear Overhauser effect (NOE) depends upon the occurrence of dipole–dipole relaxation processes and can similarly provide an estimate of internuclear separation.

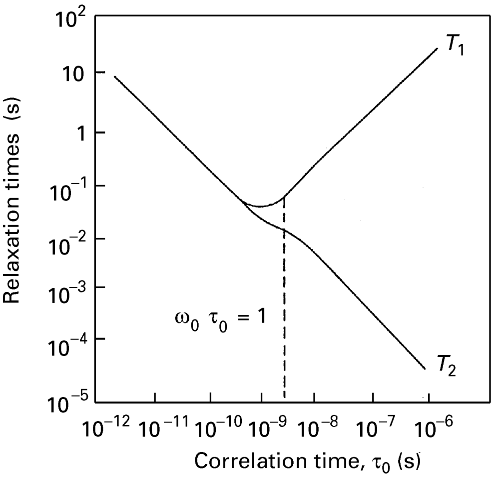



**Figure 3** Schematic representation of the nuclear relaxation times $T_1$ and $T_2$ as functions of the correlation time $\tau_0$.

The large magnetogyric ratio of the proton coupled with its common molecular occurrence ensures that dipole–dipole interactions with protons frequently dominate the relaxation of other spin-$\frac{1}{2}$ nuclei such as $^{13}C$ and $^{15}N$.

The electron has a magnetogyric ratio that is more than 600 times larger than that of the proton; thus, if unpaired electrons are present their dipole–dipole interaction with a given nucleus normally controls the relaxation of that nucleus. Consequently, paramagnetic centres may be introduced to override nuclear–nuclear relaxation processes in certain cases, for example to reduce embarrassingly long relaxation times and to remove NOEs in cases where they are not required.

Nuclei with a spin $I > \frac{1}{2}$ have electric quadrupole moments in addition to the magnetic dipole moments required for the NMR experiment. The quadrupole moment may interact with a local electric field gradient to provide a very efficient nuclear relaxation process, and thus broad NMR signals. The value of $T_1$ for the quadrupolar relaxation process depends critically upon the electronic environment of the nucleus in question. This is demonstrated by the NMR line widths of about 10 Hz for $^{35}Cl$ in NaCl and about 10 kHz for $^{35}Cl$ in $CCl_4$. In the former example, the electronic environment of the chloride ion is approximately spherical, thus there is only a small field gradient, at best, at the site of the chlorine and the line width is controlled by the less efficient dipole–dipole process. For covalently bonded $CCl_4$, the large field gradients at the chlorine nuclei give rise to rapid quadrupolar relaxation.

Spin–rotation interactions may also produce nuclear relaxation. These arise from interactions between nuclear magnetic moments and rotational magnetic moments of the molecules containing the nuclei in question. A direct transfer occurs of nuclear spin energy to the molecular motion. This contrasts with the dipole–dipole and quadrupole mechanisms, which operate via an indirect energy transfer. The value of $T_1$ due to spin–rotation interactions decreases as the temperature increases, which is in contrast to the other nuclear relaxation mechanisms. Hence the observed temperature dependence of $T_1$ may be used to demonstrate the contribution or absence of spin–rotation interaction processes to the nuclear relaxation. Spin–rotation relaxation is most likely to be dominant for small molecules tumbling rapidly at high temperatures. Thus it is likely to be of particular importance for vapour-phase studies.

Anisotropy of the nuclear shielding tensor may also contribute to nuclear relaxation. Brownian motion can modulate the nuclear shielding tensor and thus provide a fluctuating magnetic field. The corresponding relaxation times depend inversely upon the square of the applied magnetic field and the square of the shielding anisotropy. Thus this relaxation process is likely to be of most importance at very high magnetic field strengths and for heavier nuclei, which tend to have very large shielding anisotropies, e.g. $^{195}Pt$ and $^{199}Hg$. The fact that $T_1$ values for this process depend upon the strength of the applied magnetic field provides a means of determining the contribution or absence of nuclear shielding anisotropy to the relaxation of a given nucleus.

If chemical exchange or internal rotation causes the spin–spin coupling interaction between two nuclei to become time dependent, then scalar relaxation of the first kind can occur. Scalar relaxation of the second kind relates to the case where the relaxation rate of a coupled nucleus is fast compared with $2\pi J$. Coupling to a quadrupolar nucleus can give rise to this relaxation mechanism. For scalar coupling relaxation to be operative, it is generally important that the resonance frequencies of the coupled nuclei be similar. This is, perhaps, the least common of the nuclear spin relaxation processes considered.

## List of symbols

$B_0$ = applied magnetic field strength (flux density); $E_{AB}$ = energy of couplings interaction between nuclei A and B; $J$ = spin coupling constant; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $W_{1/2}$ = full width at half-height of NMR line; $\delta$ = chemical shift; $\omega$ = angular frequency of applied radiation; $\Delta\sigma$ = shielding anisotropy; $\sigma$ = nuclear shielding parameter; $\tau_0$ = molecular correlation time.

*See also:* **$^{13}C$ NMR, Parameter Survey; Chemical Shift and Relaxation Reagents in NMR; Gas Phase Applications of NMR Spectroscopy; NMR in Anisotropic Systems, Theory; NMR Principles; NMR Relaxation Rates; Nuclear Overhauser Effect.**

## Further reading

Abragam A (1961) *The Principles of Nuclear Magnetism*. Oxford: Clarendon Press.

Ando I and Webb GA (1983) *Theory of NMR Parameters*. London: Academic Press.

Contreras RH and Facelli JC (1993) Advances in theoretical and physical aspects of spin–spin couplings. In: Webb GA (ed) *Annual Reports on NMR*, Vol 27, p 255. London: Academic Press.

de Dios AC (1996) Ab initio calculations of the NMR chemical shift. *Progress in NMR Spectroscopy* **29**: 229.

*Specialist Periodical Reports on NMR*, published annually by the Royal Society of Chemistry, Webb GA (ed),

contain chapters dealing with all aspects of NMR Parameters. Latest edition is Vol 29 (1999).
Webb GA (1978) Background theory of NMR parameters. In: Harris RK and Mann BE (eds) *NMR and the Periodic Table*, p 49. London: Academic Press.
Webb GA (1993) An overview of nuclear shielding calculations. In: Tossell JA (ed) *Nuclear Magnetic Shieldings and Molecular Structure*, p 1. Dordrecht: Kluwer.

# Peptides and Proteins Studied Using Mass Spectrometry

**Michael A Baldwin**, University of California, San Francisco, CA, USA

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
| --- |
| Applications |

Thirty years ago it was impossible to ionize and analyse even a small peptide by mass spectrometry unless it was first made volatile by derivatization, such as acetylation and/or permethylation. In recent years 'soft ionization' methods have made mass spectrometric analysis of peptides and proteins a routine activity. Such methods employed for ionization and analysis of peptides and proteins have included field desorption (FD) from a heated emitter by high electric fields, direct chemical ionization (DCI) by the interaction of a hot plasma with a solid sample, fast atom bombardment (FAB) involving bombardment of an analyte solution with high energy xenon atoms or caesium ions, plasma desorption (PD) using nuclear fission fragment bombardment of a sample on a solid support such as nitrocellulose, electrospray ionization (ESI) by evaporation of charged droplets of analyte solution, and matrix-assisted laser desorption/ionization (MALDI) by laser irradiation of crystals of a matrix doped with analyte. Several of these are still in limited use but the almost universal utility of ESI and MALDI for the analysis of macromolecules of virtually unlimited mass range with extreme sensitivity has caused these two methods to supplant all other techniques, so only these methods will be discussed further.

At its simplest level, MS measures molecular masses. With calibration it can also determine quantities on a relative or absolute scale for 'pure' compounds and, with varying degrees of success, for components in a mixture. The analysis of complex mixtures such as a protein digest may require coupling with a separative method such as chromatography (GC-MS or LC-MS) or electrophoresis, either off-line (SDS-PAGE) or on-line (CE-MS). Further experiments can provide detailed structural information, e.g. peptides can be sequenced by collision-induced dissociation (CID) of their molecular ions and tandem MS (MS/MS). MS may also be used in conjunction with chemical modification or enzymatic digestion of a protein to aid its identification and/or sequence analysis. In practice, the diverse techniques available for ionization and mass analysis allow experiments to be optimized to answer very specific questions.

## Mass spectrometry

### Sample preparation and ionization methods

Optimization of sample preparation depends upon the nature of the sample, the information required and the type of mass spectrometer available. It is desirable to minimize salts and detergents, and if buffers are unavoidable these should be volatile whenever possible, e.g. ammonium formate or ammonium bicarbonate. In general MALDI is more tolerant of impurities than ESI. It may be essential to remove salts and detergents by dialysis, precipitation, absorption/elution from beads or a membrane, or absorption onto a small column and elution into the mass spectrometer. Achieving such separations without substantial losses is frequently complicated by limited amounts of material, sample aggregation, hydrophobicity and binding to surfaces.

Most peptides and proteins contain readily protonated basic sites, suitable for positive ion MS. Analytes are ionized directly from liquid solution for ESI but from the solid state for MALDI, consequently sample handling is fundamentally different for these alternative methods. In ESI-MS, liquid is usually introduced in a continuous stream, ideal for direct coupling with reversed phase high performance

liquid chromatography (RP-HPLC), which is applicable to the separation of most peptide mixtures and many proteins. However, trifluoroacetic acid (TFA), widely used to optimize separations by RP-HPLC, can inhibit ionization in ESI-MS. Solvent systems developed for LC-MS replace TFA by formic acid, alternatively low flow rates from capillary HPLC columns can be supplemented with solvents more compatible with ESI-MS. An alternative to an externally pumped system is provided by nanospray, in which a small quantity of sample solution is placed in a capillary tube drawn to a fine tip. Liquid flows out at ~20–50 nL $min^{-1}$ under the combined influence of capillary action and an applied electric field, allowing each sample to be studied for an hour or more, which assists studies on mixtures such as protein digests. Fortunately, peak intensity in ESI or nanospray is largely independent of flow rate; consequently low flow rates efficiently conserve samples that are difficult to isolate and purify.

For MALDI-MS the analyte as a pure compound or a mixture is co-crystallized with a matrix that absorbs laser radiation and promotes ionization. Matrix materials ideal for peptides and proteins are aromatic acids such as sinapinic acid, 2,4-dihydroxybenzoic acid, and $\alpha$-cyano-4-hydroxycinnamic acid, each having slightly different ionization characteristics. Published protocols for optimization of sample preparation try to achieve multiple, evenly distributed, small crystals. It is often necessary to remove salts by washing the crystals with water after they have been deposited.

### Measurement of molecular mass

MS separates ions according to mass/charge (*m*/*z*). Peptides and proteins ionized by ESI under acidic conditions acquire multiple charges, *z* being roughly proportional to *m*, with *m*/*z* in the range 500–1500. In practice a distribution of charges gives multiple peaks in the mass spectrum, the spacing of which allows *z* to be calculated. The raw data for a pure compound can be deconvoluted to a zero-charge profile of the molecular mass, although this is more difficult for mixtures. An advantage of multiple peaks is the statistical improvement in mass accuracy. Multiple charging allows the *m*/*z* range of the mass spectrometer to be modest, even for large proteins. Mass analysers for ESI are mostly quadrupoles and ion traps with *m*/*z* ranges of 2000–3000, but orthogonal acceleration TOFs, hybrid quadrupole-TOFs, sector instruments and FTICRs of higher mass range are all available with ESI sources.

MALDI attaches only a single charge or a small number of charges to a peptide or protein, consequently the *m*/*z* range for a suitable mass analyser must be much greater. Potentially a linear TOF instrument, with or without a reflectron, has unlimited mass range. Mass separation is based on ion velocity; slower ions take longer to arrive at the detector, therefore mass range is limited only by the observation time. In practice, factors such as detector design may inhibit the effective observation of the most massive species, but a mass range of several hundred thousand daltons is attainable.

MS methods for analysing peptides and proteins have two different operating regimes, which can be called low mass and high mass. Low mass describes the range where individual isotopic contributions to the overall molecular ion signal can be resolved as separate peaks. This is 1–2 kDa for a quadrupole of modest performance, perhaps 5 kDa for a high performance MALDI-TOF or ESI orthogonal-acceleration TOF, and significantly higher for FT-ICR. This regime which applies to peptides rather than proteins gives narrow peaks. With internal calibration 'monoisotopic' masses can be measured to 5–20 ppm for the ions containing the lowest mass isotopes, including $^{1}H$, $^{12}C$, $^{14}N$ and $^{16}O$. Only for the smallest species is this sufficient for an unambiguous isotopic assignment but it frequently differentiates between alternative isobaric species (ions of the same nominal mass). For multiply charged ions, the spacing of adjacent peaks within an isotopic cluster is equal to the reciprocal of the charge (1/*z*), thus *z* can be determined from a single peak. This is useful for complex spectra with multiple peaks that would otherwise be difficult to assign. In the high mass regime the isotopic clusters are not resolved and the 'average' molecular mass is obtained. Here mass spectrometer resolving power has less effect on overall mass accuracy, although any factor that broadens or distorts peak envelopes will introduce errors. This can include small covalent modifications such as methionine oxidation (+16 Da) or addition of a cation such as sodium (+23 Da) rather than a proton. These should be clearly resolved for small proteins of perhaps 20 kDa, but not for large proteins of say 100 kDa with inherently broad peaks. The best mass accuracy likely to be achieved with standard instrumentation is approximately 0.1–0.3 Da at 10 000 Da or 1–3 at 100 000 Da.

FT-ICR with a high field magnet represents a divergence from the above statement as this can have extraordinarily high resolving power. **Figure 1** shows the resolved isotopic cluster for the +49 charge state of bovine serum albumin (molecular mass 66.4 kDa), measured with a resolving power of 370 000 using a 11.5 T magnet. Thus, with such an instrument, almost any sample can give isotopic resolution.

### Sensitivity of detection

Soft ionization usually gives molecular ions but not fragment ions, thus ion current is concentrated into a single peak or isotopic cluster. Although the ion yield of the ionization methods is relatively low (~1 ion per 1000 neutral molecules), MS is highly sensitive. Less than 100 ions are sufficient to define a mass spectrometric peak, i.e. ~$10^5$ molecules or 0.1 attomole. To exploit this inherent sensitivity it is necessary to integrate the entire ion signal, rather than scan a spectrum in which only a small fraction of the ions is monitored while most go unobserved. This is achieved by MALDI-TOF as each laser shot forms a packet of ions which are accelerated into the mass analyser to ultimately arrive at the detector. MALDI also has the advantage that a discrete quantity of sample on the target is available for analysis for as long as the experimenter chooses to select a new region to investigate or until the sample is exhausted, a dried spot from 1 μL of sample being sufficient for several thousand laser shots. By contrast ESI is used mostly with scanning instruments and spectra are recorded during the limited time the analyte enters the ionization region. This is relatively inefficient and ESI has been regarded as less sensitive than MALDI. A new generation of TOF instruments compatible with ESI integrate the signal, and nanospray is more like MALDI as sample is retained throughout the experiment, providing a substantial sensitivity enhancement. MALDI and nanospray both provide detection limits in the low femtomole region or better.

## Additional techniques

### Direct analysis of mixtures versus LC-MS

Because MALDI gives predominant singly charged molecular ions with few fragments, analysis of multicomponent mixtures such as protein digests is readily achieved. Each peak corresponds to a separate peptide and can be selected for ‘post-source decay’ or PSD. However, some components in a mixture may not compete effectively for the available charges and may be weak or absent, e.g. tryptic peptides terminating in lysine rather than arginine. ESI is less suitable for direct analysis of mixed peptides as each component gives several multiply charged peaks that cause complex spectra. However, ESI is ideal for LC-MS and is less discriminatory as components elute separately, giving more comprehensive coverage of the original protein. Although unimportant for identification of a protein in a database, this is essential to find protein modifications or mutations. Automation is available from some instruments manufacturers for MS-MS analysis on each molecular species eluting

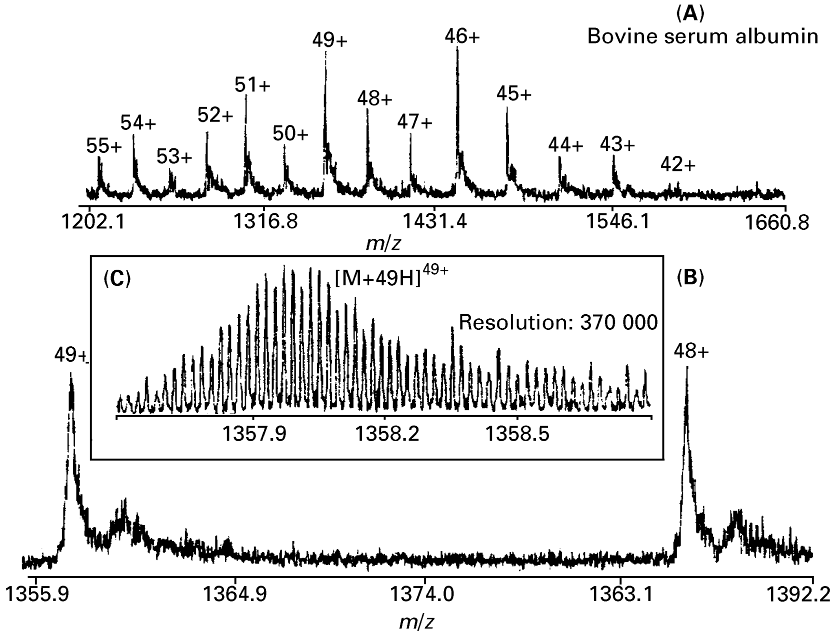



**Figure 1** (A) ESI-FTICR spectrum of bovine serum albumin recorded using an 11.5 T magnet; (B) An expansion of the 49+ and 48+ charge states; (C) A further expansion of the 49+ charge state recorded at 370 000 resolving power showing isotopic separation. Reproduced with permission of Elsevier Science from Gorshkov MV, Tolić LP, Udseth HR, et al. (1998) Electrospray ionization – Fourier transform ion cyclotron resonance mass spectrometry at 11.5 Tesla: instrumental design and initial results. *Journal of the American Society for Mass Spectrometry* **9**: 692–700.

from the chromatograph. Disadvantages of LC-MS include added cost and complexity of the instrumentation and additional time spent equilibrating columns and waiting for components to elute.

### Peptide sequencing by CID

Peptides and proteins can be sequenced by Edman chemistry at levels down to 1–10 pmol, depending on the number of residues to be determined. This requires a free amino terminus and is ineffective for modified amino acids unless appropriate standards are available. Through the use of CID and tandem mass spectrometers, MS has established itself as a more sensitive, faster alternative to Edman sequencing, although it cannot handle an intact undigested protein. Unlike Edman sequencing, fragment data can be obtained on each component of a mixture without separation. The efficiency of peptide sequencing by MS depends greatly upon the type of instrument available. Tandem mass spectrometers for CID include multisector instruments, triple quadrupoles (QQQs) and hybrid quadrupole-sectors. The less expensive and easier to operate QQQ is widely used. A molecular ion selected in Q1 fragments in Q2 through low energy collisions with a gas, then fragment ions are analysed in Q3. The same experiment, and even MS$^n$, can be carried out in a relatively modest ion trap, or in an FT-ICR with higher resolution but at considerably greater expense. Hybrid QQ-TOFs offer substantially superior performance to QQQs, giving high sensitivity, high resolution data. A MALDI-TOF instrument equipped with a reflectron is equally capable of MS-MS by PSD.

Interpretation of spectra from CID-MS-MS of peptides of up to ~20–25 residues has been well documented. Peptides have a repeating linear backbone with sidechains defining the constituent amino acids. Backbone fragmentation of a singly charged peptide ion gives two species, an ion and a molecule. Retention of the proton by the *N*-terminal fragment gives an **a**, **b** or **c** ion, whereas *C*-terminal ions are classified as **x**, **y** or **z** (**Figure 2**). The most common cleavage at the amide bonds gives **b** or **y** ions; tryptic peptides with a *C*-terminal basic residue generally exhibit predominant **y** ions. Because ionization occurs by proton addition, this gives an ion with no odd electrons, which can be more stable than a corresponding radical cation. Subsequent cleavage of a backbone bond is associated with transfer of a hydrogen radical to prevent the thermodynamically unfavourable formation of a radical cation and a neutral radical. In forming a **b** ion this hydrogen moves to the neutral, whereas a **y** ion gains one hydrogen in addition to that added during ionization; thus **y** ions are sometimes designated as **y″** or **y+2**. Ion masses are calculated as the sum of the amino acid residues involved plus 1 H for **b** ions or plus

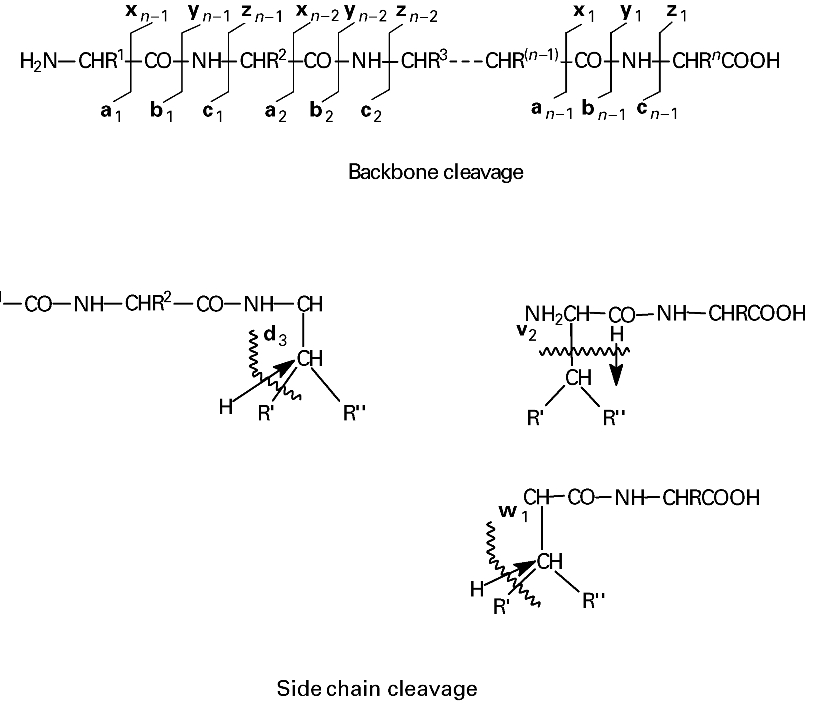



**Figure 2** Collision induced fragmentation scheme for peptides and proteins. The initial ionizing proton is not shown and proton transfers are not shown for backbone cleavages (see text).

$H_3O$ for **y** ions. Theoretically, cleavage can occur at each amide bond, giving a series of ions defining the amino acid sequence. In **Figure 2**, subscripts attached to the ion types identify which bonds have broken to form the fragment ions, e.g. for a peptide of *n* amino acids, the *C*-terminal ionic fragment formed by loss of the most *N*-terminal amino acid is designated $\mathbf{y}_{n-1}$. Some ions are formed by further cleavages at the sidechains, giving peaks referred to as **d**, **v** and **w** ions, some of which can identify specific amino acids and differentiate isomeric amino acids such as leucine and isoleucine. The **d** ions represent loss of a group from the $\beta$-carbon of an **a** ion, **w** ions are formed by the equivalent loss from a **z** ion, and **v** ions represent loss of the intact sidechain at the $\alpha$-carbon of a **y** ion. Multiple bond cleavages also give internal fragments, including individual amino acids that appear in the low mass region of the spectrum as immonium ions $^{+}NH_2{=}CHR$ that are valuable for diagnostic purposes. Note that $\mathbf{a}_1$ is the immonium ion for residue 1.

Interpretation of MS-MS spectra of multiply charged ions from ESI-MS is complicated by the different charge states possible amongst the fragment ions, unless high mass resolution is available. The most comprehensive fragment ion spectra are generally obtained from high energy CID, such as from sector instruments. Programs exist for both the straightforward prediction of spectra and the more difficult interpretation of experimentally obtained spectra.

### Chemical derivatization

Many straightforward chemical reactions can enhance the quality and utility of MS data from peptides and proteins. Before the advent of ESI or MALDI, polar groups in biomolecules were often derivatized to increase volatility, e.g. by permethylation or silylation. This could provide additional information on the number of replaceable hydrogens of a given type. Such procedures are still useful. Acetylation with acetic anhydride adds 42 Da for each free amino group, confirming whether the amino terminus is free or blocked, and can distinguish isobaric glutamine from lysine, the latter becoming acetylated. An equimolar mixture of perdeutero and protonated reagent gives double peaks separated by 3 Da for *N*-terminal but not *C*-terminal fragment ions. Esterification with acetyl chloride/methanol adds 14 Da per carboxylic acid and provides similar information about the *C*-terminus and the location of glutamate or aspartate residues. Trypsin digestion in $H_2{}^{16}O/H_2{}^{18}O$ differentially labels the *C*-termini of all resulting peptides, except the original protein *C*-terminus. All of these techniques employing stable heavy isotopes enhance the information content in MS/MS as the *N*- and *C*-terminal fragments are readily distinguishable. Other derivatizations to improve MS/MS spectra include the addition of a permanent positive charge at one or other terminus, usually the *N*-terminus, which directs the fragmentation and aids spectral interpretation.

Protein disulfide bonds can be reduced with dithiothreitol and alkylated with a reagent such as iodoacetic acid before digestion, adding 58 Da per cysteine. Although MS is replacing Edman sequencing to a significant degree, Edman chemistry is used for ladder sequencing, in which phenyl thiocyanate is included at each cycle with the normal Edman reagent, phenylisothiocyanate. This blocks the *N*-terminus of a small fraction of the analyte molecules and prevents further cleavage, giving mixed products differing from each other by single amino acids. The sequence is read directly from the MALDI spectrum of the unseparated mixture.

## Applications

### Quality control of synthetic peptides and recombinant proteins

MS analysis has become a routine aid in the purification of synthetic peptides and recombinant proteins and plays an essential role in quality control of materials required to be of high purity. Following cleavage from the solid-phase resin, high quality peptides are purified by RP-HPLC. Fractions collected from an analytical run can be surveyed by MS to identify the elution profile of the desired product, with a minimum of impurities. Fractions may be dried down and used directly or they may act as a guide to fraction collection for a larger scale separation. The presence of unwanted side products such as those formed by amino acid deletions, incomplete removal of protecting groups and chemical modifications should be immediately apparent from the measured masses. If careful attention was paid to the sequence of amino acids loaded onto the synthesizer, the observation of the desired molecular mass should be sufficient to confirm the anticipated product. If necessary, MS/MS can be used to confirm the sequence.

MS and MS/MS are particularly useful for identifying and locating heavy isotopes. For example, **Figure 3** shows a portion of the MS/MS spectrum of three versions of a 14-residue peptide prepared for an NMR study, two of which contain $^{13}C$ labels at a carbonyl and an $\alpha$-carbon in two alanine residues. Mass differences between successive **a** and **b** ions allow the positions of these residues to be determined

precisely. As $\mathbf{b}_6$ is at *m/z* 656 for all three species including the unlabelled control, no $^{13}C$ labels are present in the first six residues. However, $\mathbf{a}_7$ for compound (C) is 1 Da higher than the equivalent ion for (A) or (B), therefore the $\alpha$-carbon of the seventh residue in (C) is $^{13}C$. Similar logic allows each of the other labels to be identified.

A number of potential chemical modifications can cause recombinant proteins to differ from the desired product. Cysteine-containing proteins may show a time-dependent shift of both chromatographic retention time and mass as aerobic oxidation causes disulfide formation (–2 Da per disulfide). Methionine oxidation to the sulfoxide is quite common and is readily identified (+16 Da). *N*-terminal glutamine may eliminate ammonia to form pyroglutamic acid (–17 Da), especially if stored in acidic solution. Harder to detect may be deamidation of asparagine to form a succinimide intermediate that is then hydrolysed to aspartate, or its isomer isoaspartate (+1 Da). Posttranslational modifications such as glycosylation occurring in mammalian systems are rarely observed in proteins expressed in bacteria but processes such as phosphorylation are not unknown. Enzyme impurities from the expression system may be responsible for numerous reactions, including the total degradation of the desired product. Carboxypeptidases and aminopeptidases can result in unexpected trimming of the intact sequence. *N*-terminal methionine is quite often observed to be partly or completely absent (–131 Da). This list of potential variants is far from complete but it gives an indication of the role that MS can play in their identification.

### Protein identification by in-gel digestion and database searching

The closing decade of the twentieth century witnessed the initiation of a major concerted programme to sequence the human genome, and genomes for several other organisms are already completed. This effort is yielding a vast array of information about genes, but this will be the tip of the iceberg compared with the unanswered questions relating to proteins, including cellular and tissue-specific variations in levels of expression, posttranslational modifications, and their associations to form functional multimolecular units. MS will play an essential role in the elucidation of this information, often referred to as proteomics. Techniques are now available for the analysis of the major proteins in specific cell types. At present the most productive methods link 2D electrophoresis with high sensitivity MS and database searching, sometimes referred to as 'mass fingerprinting'.

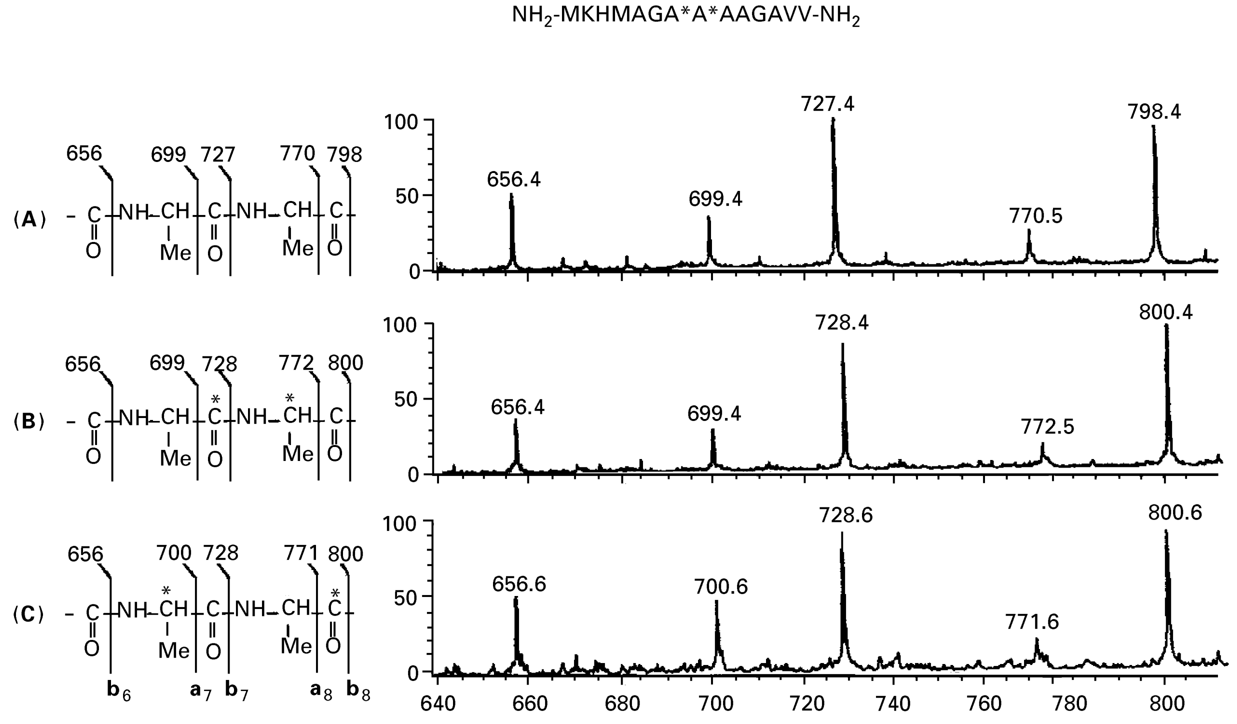



**Figure 3** Partial CID-MS/MS spectrum obtained on a tandem 4-sector mass spectrometer for a 14-residue synthetic peptide. (A) without $^{13}C$ labels; (B) and (C) with $^{13}C$ labels as indicated by asterisks.

As many as 2000 proteins from a cell digest may be separated on a 2D gel as discrete spots stained with Coomassie blue (100 ng sensitivity), silver (1–10 ng) or a fluorescent dye (<1 ng). These proteins may be difficult to elute from the gel in high yield for direct analysis by MS. Furthermore, if it could be determined, even the precise atomic composition would be insufficient information to identify a protein. However, cutting the spot from the gel, digesting the protein with an enzyme such as trypsin, eluting the peptides and analysing them by MS can provide sufficient information to unambiguously identify any known protein in a standard protein database, particularly if ion masses are measured accurately. Tryptic peptides are advantageous as they all terminate in a readily protonated basic residue, lysine or arginine. They can either be analysed in the intact mixture, usually by MALDI or nanospray, or separated and analysed by LC-MS. Several programs are freely available via the World Wide Web for database searches. As yet the databases are incomplete and many proteins remain unidentified, but ultimately all possible proteins will be predictable for an entire genome. As databases increase in size it will become necessary to increase the number of peptides identified or further improve the accuracy of mass measurement. Alternatively, sequence information from CID and MS/MS will greatly increase the reliability of identification.

### Identification of genetic mutations

Most genetic mutations are identified by molecular biological techniques including gene sequencing. However, mass spectrometry of protein digests was a valuable additional method even before the development of ESI and MALDI. The molecular mass of an intact protein can now be determined by ESI or MALDI, which compared with a normal protein or the predicted sequence from a c-DNA will indicate the presence of a mutation. The precise location and identity can be confirmed by digestion and further analysis of the peptides by LC/MS or LC/MS/MS. The identification of haemoglobin variants represents one of the best known examples of mutational analysis by MS. Most patients are heterozygous, i.e. blood samples contain both the normal and variant forms of the protein. With a molecular mass in the region 15–16 kDa, monomeric haemoglobin is a relatively small protein and the differences between normal and variant forms are easily measured with high precision. Mass spectrometry complements the normal electrophoretic detection of the numerous haemoglobin variants responsible for a number of serious diseases including sickle-cell anaemia.

### Posttranslational modifications

Gene sequences specify the initial amino acid sequences of proteins as expressed by the ribosomes. However, complex enzymic reactions in the cell can greatly modify the chemical composition of mature proteins. There are numerous possible modifications such as trimming of the expressed sequence to remove signal peptides, cysteine oxidation to cystine and addition of functionalities such as phosphates, sulfates, oligosaccharides and lipids. Understanding these modifications may be fundamental to understanding the biological action of a protein. To some extent they may be predictable from the existence of a consensus sequence such Asn-X-Thr or Asn-X-Ser for glycosylation of asparagine, but the extent of the modification may range from 0–100% of protein molecules. Also, because a process such as glycosylation is carried out by a complex cascade of enzymes, the glycoforms at a single site are frequently diverse in both structure and molecular mass. MS of a glycopeptide or glycoprotein will reveal the presence of the oligosaccharide by an increase in mass, but the potential heterogeneity may make it difficult to resolve the various glycoforms. Furthermore, mass alone will not distinguish the many isomeric forms of the individual monosaccharide units, their sequences or any branching patterns. The sugars can also be modified, e.g. by presence of phospholipids as in the glycosyl phosphatidylinositol-anchored proteins.

In order to determine the mass of the amino acid sequence of the protein alone the oligosaccharides can be removed chemically or with a glycosidase enzyme such as Endo-H or peptide *N*-glycosidase F (PNGase F). The released sugars can also be studied although their complete characterization may be more difficult than that of the peptide or protein. Tandem MS may identify the glycopeptides in a complex peptide mixture as CID gives low mass ions characteristic of the monosaccharide units, including $m/z$ 163 (hexose), 204 (*N*-acetylhexose) and 292 (sialic acid). These masses are ideal for LC-MS-MS precursor ion scans to identify the glycopeptides and define their masses. Methods have been developed for the detection and analysis of *N*-linked and *O*-linked oligosaccharides and the more recently discovered but widespread *N*-acetyl glucosamine modification of serine (and threonine), which apparently plays a regulatory role similar to that of phosphorylation.

**Figure 4** illustrates the complexity that can arise due to glycosylation. A 75-residue glycopeptide derived from the prion protein by digestion with endopeptidase Lys-C contained a single glycosylation site. The oligosaccharides attached to this site were highly heterogeneous, giving the complex ESI

spectrum in **Figure 4A**, with the deconvoluted spectrum in **Figure 4B**. Treatment with PNGase F removed the sugars completely, leaving just the amino acid chain. This gave a simplified spectrum from which a molecular mass of 8607.8 was determined. Subtraction of the mass of the deglycosylated peptide from those of the prominent glycopeptides gave the masses of the sugars, which could be related to partial structures for the glycoforms determined previously.

Modifications such as phosphorylation are chemically simpler and are more easily identified by MS, although the addition of 80 Da can be due to either a phosphate ($HPO_3$) or a sulfate ($SO_3$). Serine and threonine phosphates are acid sensitive and may not survive LC/MS whereas tyrosine phosphates are generally more robust. Sulfate is also labile and may not be observed in positive ion MALDI but it usually can be identified by ESI or negative ion MALDI. The importance of phosphorylation in signalling and the control of cellular events has greatly increased interest in identification of protein phosphorylation sites. A valuable technique to locate phosphopeptides in a protein digest uses LC-ESI-MS-MS precursor ion scans of *m*/*z* 79 in the negative ion mode to identify all ions forming the characteristic fragment $PO_3^-$. The extent of phosphorylation may be much less than stoichiometric, so the same peptides may occur with and without phosphate. The identified phosphopeptides can be sequenced by MS-MS to identify the specific residues carrying the phosphate groups. A current challenge is to increase the sensitivity for detection and identification of such modifications in complex mixtures of proteins extracted from cells and tissues.

Disulfide bond analysis by MS is usually straightforward for proteins but is challenging for small

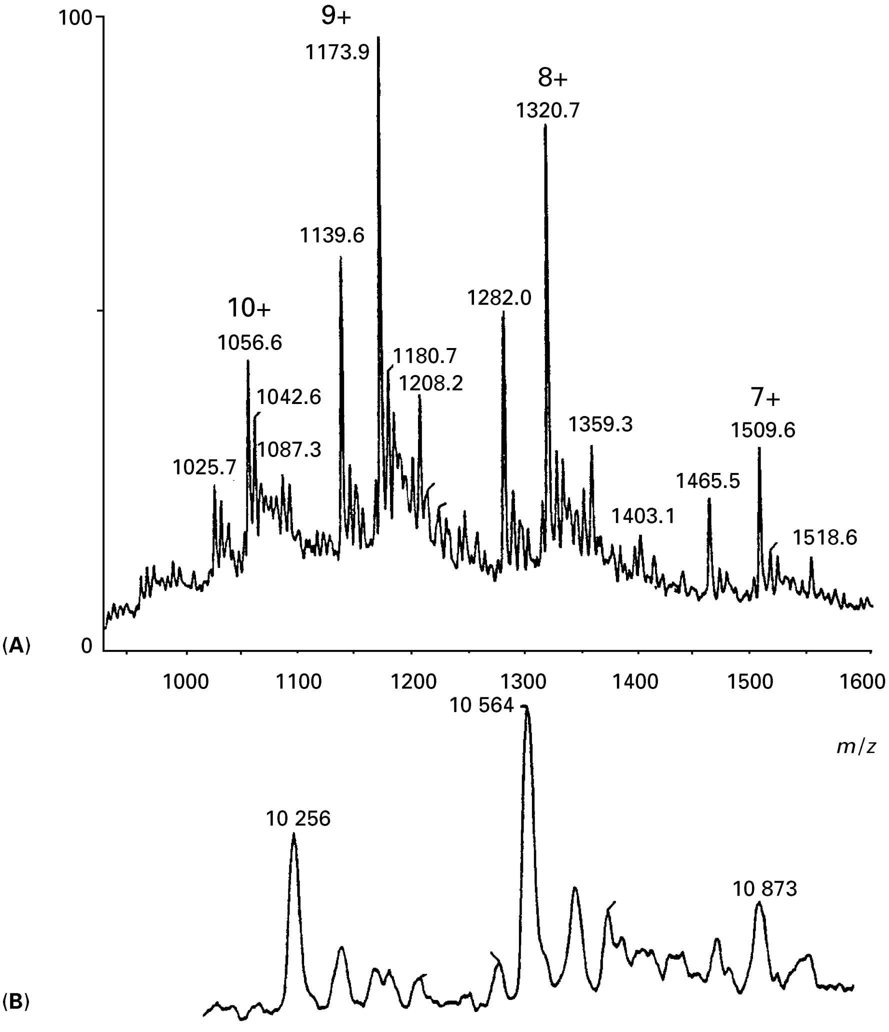



Figure 4 (A) ESI-MS spectrum of a 75-residue glycopeptide from the prion protein showing multiple peaks due to oligosaccharide heterogeneity. (B) Deconvolution of the spectrum in (A) showing molecular masses rather than *m*/*z*.

cysteine-rich proteins and peptides such as the conotoxins which may contain as many as 6 cysteine residues within a peptide of only 20–30 amino acids. For a protein an effective strategy is to carry out chemical or enzymic digestion and peptide mapping on the native material after reduction with dithiothreitol and alkylation of the cysteines using a reagent such as iodoacetic acid. The addition of a carboxymethyl group to a cysteine increases the mass by 58 Da, so cysteine-containing peptides are readily detected by comparison with a digest of the unmodified protein. A protein digest with the disulfide bonds intact will result in some peptides linked by the disulfides. If these are not identifiable on the basis of mass alone, they may be separated by HPLC, reduced and then analysed as two separate peptides.

## Noncovalent interactions and analysis of higher order structure

Peptides and proteins do not carry out their biological functions in isolation. Their active forms may be associated with metal ions, small molecules, macromolecules including proteins and DNA, and they may form complex multiprotein assemblies. Biological processes occur in aqueous solution in the presence of salts and other ions whereas MS is used to probe substances isolated *in vacuo*, conditions that could be incompatible. Despite this, the use of MS to study these interactions is growing rapidly. MALDI, which ionizes directly from the solid phase, does not lend itself to studies that mirror solution conditions, but ESI, which abstracts ions directly from solution by rapid evaporation of nebulized droplets, can monitor noncovalent associations very successfully.

The most successful studies have been on systems involving electrostatic rather than hydrophobic interactions. The metal-binding properties of a peptide (or protein) can be monitored directly from aqueous solution of perhaps 10 μM peptide containing the metal cation salt at various concentrations. The pH is controlled by the use of a weak buffer such as 1–10 mM ammonium acetate. An equilibrium in solution between free peptide and the complex can be preserved during the rapid drying of the microdroplets, occurring within 0.5 ms. Low micromolar dissociation constants derived from the peak intensities are similar to those obtained by conventional methods. Although removal of the dielectric shielding effect of water might encourage random binding of cations to negative amino acids, strong specific binding can be clearly distinguished from weak nonspecific effects. Peptides homologous to some ATP-ases containing His-X-His or His-X-X-His bound either $Cu^{2+}$ or $Ni^{2+}$ but not $Zn^{2+}$, whereas analogous peptides with a single histidine bound only $Cu^{2+}$. By contrast a peptide corresponding to a region of the Alzheimer's precursor protein having the motif His-X-His-X-His was seen to bind all three divalent cations.

**Figure 5** shows partial mass spectra at two time points for a metal ion transfer reaction between metallothionen coordinated with 7 zinc ions and carbonic anhydrase. The conversion of the apo form to the holo form was monitored over a wide pH range for up to 120 min. It was also seen that zinc ion uptake was associated with addition of a hydroxyl radical or water molecule. Such an experiment gives direct information about the stoichiometry of reaction intermediates and complexes and can yield kinetic and thermodynamic data.

Complexes between the trp repressor (TrpR) and its specific operator DNA were monitored in a competition experiment. When TrpR was mixed with an equimolar mixture of DNA containing two consensus sequences separated by 2, 4 or 6 base pairs, 1:1 protein:DNA complexes formed only with DNA having the 4-bp spacer. The dimerization of the AIDS

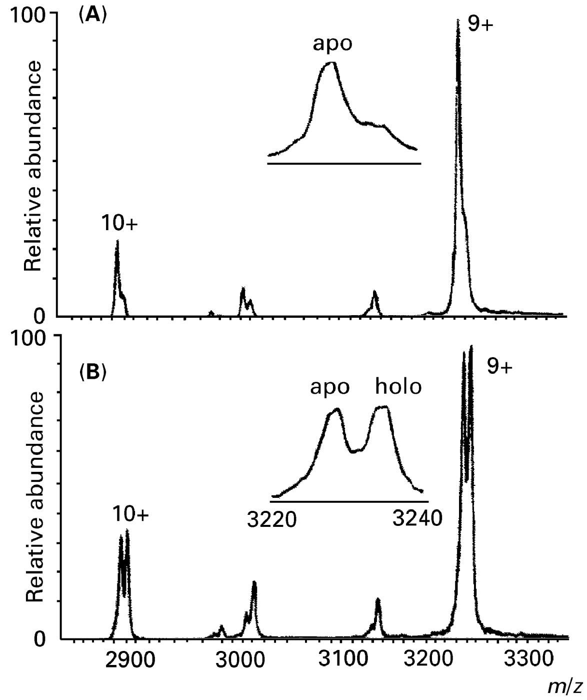



**Figure 5** A metal transfer reaction between metallothionen-$Zn_7$ and carbonic anhydrase at pH 7.5 monitored by ESI-MS showing the major peaks for carbonic anhydrase, (A) after 3.5 min, and (B) after 43.5 min. Reproduced with permission of Cambridge University Press from Zaia Z, Fabris D, Wei D, Karpel RL and Fenselau C (1998) Monitoring metal ion flux in reactions of metallothionen and drug-modified metallothionen by electrospray mass spectrometry. *Protein Science* **7**: 2398–2404.

protease has also been demonstrated by ESI but hydrophobic interactions such as this are stabilized by the dielectric of the solvent and by the presence of salts. This stabilization is lost on transfer to the gas phase, necessitating very mild conditions for the nebulization, evaporation and transfer of such delicate complexes into the vacuum system.

ESI-MS studies at neutral pH of proteins in their native state usually results in the attachment of far fewer protons compared with ionization from acidified solution. Consequently *m*/*z* values are frequently 5000 or more and may range up to 10 000. Such species are beyond the normal working range of many mass spectrometers including most quadrupoles and ion traps, but they can be studied with TOF or high field magnetic instruments.

### Hydrogen–deuterium exchange

MS is obviously well suited to establishing the primary amino acid sequence and posttranslational modifications of peptides and proteins and it also has a role in determining the quaternary structure, i.e. the noncovalent association of subunits in multiprotein assemblies. What is less apparent is that hydrogen–deuterium exchange allows secondary and tertiary structure to be probed. Other techniques commonly used for such biophysical studies include optical spectroscopy, fluorescence, circular dichroism, light scattering, NMR and X-ray crystallography. Only the last two of these can give a detailed 3D molecular structure, the others give integrated views that obscure individual features. MS requires much less material than any other method and can follow the dynamics of rapid protein folding and unfolding.

The amide hydrogens along the protein backbone and in some sidechain positions are labile and undergo rapid exchange with solvent protons unless they are involved in hydrogen bonding. Secondary structural elements such as helices and sheets are maintained by hydrogen bonds, as are many aspects of tertiary structure. The rate of exchange of such protected hydrogens is usually reduced dramatically. Usually a peptide or protein can be deuterated fully by lengthy exchange in $D_2O$, each deuteron increasing the mass by 1 Da, so that the total increase gives a measure of the number of exchangeable hydrogen atoms. The accessibility of these atoms is assessed by measuring the rate of back exchange with $H_2O$. Above pH 3 the rate of hydrogen exchange for freely accessible hydrogens is proportional to $[OH^-]$, i.e. is approximately 10 000 times faster at physiological pH than at pH 3. After carrying out the exchange reaction at pH 7.4, the reaction can be effectively frozen by taking aliquots at different times and rapidly dropping both the pH and the temperature. The aliquots can then be analysed by ESI-MS to determine the degree of exchange. Alternatively the exchange process occurs in a flowing stream in which the reagents mix for a controlled period of time before being quenched by mixing with an acidified solution then injected directly into the ESI source.

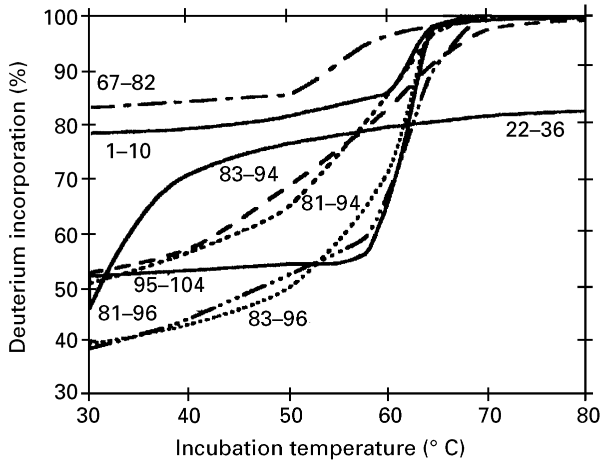



**Figure 6** Deuterium incorporation into specific segments of cytochrome C as a function of temperature, showing thermal melting at 58–64°C of $\alpha$-helical regions. Reproduced with permission of Cambridge University Press from Zhang Z and Smith DL (1993) Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Science* **2**: 522–531.

Peptides can be sequenced by MS-MS to determine the precise location of the deuterons. However it is usually essential to employ intact proteins rather than shorter peptides to address questions concerning secondary and tertiary structure, but the location of the deuterons in a protein requires digestion to peptides after exchange. Fortunately the non-specific protease pepsin is most active at the low pH values at which further exchange is minimized. Samples of the protein that have been subjected to back exchange can be digested quickly using a high enzyme:substrate ratio, then separated and analysed rapidly by LC-MS using fast flow rates through chilled, highly porous HPLC columns. Deuterium incorporation into specific segments of cytochrome C as a function of temperature is illustrated in **Figure 6**. This shows strong evidence for thermal denaturation with a sharp transition at 58–64°C that affects only the most structured regions known to be involved in $\alpha$-helices.

## Conclusions

MS is a critically important method for the identification and characterization of peptides and proteins. This role will continue to grow as proteomic studies

spur the demand for high throughput protein analysis and characterization. Many mass spectrometric techniques lend themselves to rapid analysis using automation and robotics, particularly MALDI for which it is already easy to deposit hundreds or even thousands of samples on a single target plate and to programme the mass spectrometer to collect a spectrum for each one. One serious factor limiting further exploitation is the availability of suitable computerized data systems to analyse and summarize the enormous quantities of data such automation can yield. However, it is likely that the concerted efforts of the academic community, mass spectrometer manufacturers and major users such as multinational pharmaceutical companies, will successfully solve such problems quite rapidly.

*See also:* **Atmospheric Pressure Ionization in Mass Spectrometry; Biochemical Applications of Mass Spectrometry; Chromatography-MS, Methods; Laboratory Information Management Systems (LIMS); Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Proteins Studied Using NMR Spectroscopy; Time of Flight Mass Spectrometers.**

## Further reading

Burlingame AL, Boyd RK and Gaskell SJ (1998) Mass spectrometry. *Analytical Chemistry* **70**: 647R–716R.
Burlingame AL, Carr SA and Baldwin MA (1999) *Mass Spectrometry in Biology and Medicine*. Totowa, NJ: Humana Press.
Chapman R (1996) *Protein and Peptide Analysis by Mass Spectrometry. Methods in Molecular Biology*, Vol 61. Totowa, NJ: Humana Press.
Cole RB (1997) *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications*. New York: Wiley.
Cotter RJ (1997) *Time-of-Flight Mass Spectrometry: Instrumentation and Applications in Biological Research*, ACS Symposium Series 549. Washington DC: American Chemical Society.
Ens W, Standing KG and Chernushevich IV (1998) *New Methods for the Study of Biomolecular Complexes*. NATO ASI Series. Dordrecht: Kluwer Academic.
Loo JA (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews* **16**: 1–23.
McCloskey JA (1990) Mass spectrometry. *Methods in Enzymology* **193**.

# Perfused Organs Studied Using NMR Spectroscopy

**John C Docherty**, National Research Council of Canada, Winnipeg, Manitoba, Canada

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

Studies of isolated, perfused organs have been invaluable in furthering our knowledge of metabolism and function. This experimental system allows the direct observation of the effects of an intervention on an organ in the absence of neurohumoral control mechanisms. Experimental conditions are simpler to control, and to modify, in isolated organs than in the whole animal. Isolated organs also maintain important intercellular interactions that do not take place in isolated cells, subcellular organelles or cellular homogenates. Nuclear magnetic resonance (NMR) spectroscopy of these organs has greatly increased the versatility and power of such studies. NMR spectroscopy allows data to be collected nondestructively throughout the experimental protocol in discrete time segments. The types of data that can be acquired include information on energy status, ion homeostasis, substrate utilization and enzyme activities (**Table 1**). Most studies have been performed using isolated hearts. Liver, kidney and skeletal muscle have been used to a lesser extent. In addition to furthering knowledge of biochemistry and physiology, such studies have provided a basis for the development of *in vivo* NMR spectroscopy as a tool for assessing normal and pathophysiological function in humans. This article will describe the uses of $^{31}P$, $^{23}Na$, $^{13}C$, $^{1}H$, $^{19}F$, $^{87}Rb$ and $^{7}Li$ NMR spectroscopy in cardiac physiology and pathophysiology. The application of these techniques to studies of isolated liver and kidney will also briefly be described.

## Studies of isolated hearts

NMR spectroscopy has been performed on hearts isolated from a wide variety of species including dog, pig, sheep, rabbit, turkey, ferret, guinea-pig, rat and

mouse. Hearts from the larger animals have been studied in horizontal-bore magnets with a clear bore of 20–30 cm and operating at field strengths of 4.7–7.0 T. Rodent hearts have generally been studied in vertical-bore magnets with a clear bore of approximately 8 cm and operating at field strengths of 4.7, 8.7, 9.4 and 11.75 T. In most studies perfusion is performed with a modified Krebs–Henseleit buffer with the approximate composition (mM) NaCl 118, KCl 4.7, $MgSO_4$ 1.2, $NaHCO_3$ 25 and with $CaCl_2$ varied between 1.2 and 2.5 mM depending upon the species. Ethylenediaminetetraacetic acid (EDTA) is often added at a concentration of 0.5 mM to chelate heavy-metal contaminants. For studies involving $^{31}P$ spectroscopy, phosphate is normally omitted from the perfusate to allow identification and quantification of the intracellular inorganic phosphate ($P_i$) peak. For other studies, perfusate may be supplemented with phosphate at a final concentration of 1.2 mM. The perfusate is aerated with 95% $O_2$/5% $CO_2$ to achieve a pH of 7.4, a $p_{O_2}$ > 600 mmHg and a $p_{CO_2}$ of 35–45 mmHg. Glucose (11 mM) is often supplied as the sole substrate source, although in many situations a mixture of glucose and pyruvate (or other substrates) is used. For studies of hearts from the larger species, the perfusate is supplemented with serum albumin to minimize the oedema that results from perfusion with crystalloid solutions. The use of crystalloid perfusate results in coronary flows ~3 times those observed under conditions where whole blood or buffers supplemented with washed red cells are used.

Hearts can be perfused in two modes; the working heart preparation and the isovolumic (or Langendorff) preparation. Both methods allow assessment of cardiac mechanical function throughout the experimental protocol. For both preparations, the ascending aorta is cannulated and retrograde perfusion of the aorta is initiated. In Langendorff preparations, perfusion is continued in this manner and the coronary arteries are continuously perfused throughout the protocol. Perfusion is performed under conditions of constant pressure (60–80 mmHg) or of constant flow by means of a pump. A compliant water-filled balloon is inserted into the left ventricle and connected to a pressure transducer in order to measure left ventricular pressure. The balloon is inflated to achieve a relevant end-diastolic pressure (generally in the region of 10 mmHg). In the working heart preparation, following stabilization in the Langendorff mode, the left atrium is cannulated and perfusion is continued through this chamber. The perfusate enters the left ventricle and is ejected into the aorta. Perfusion of the coronaries occurs during diastole when the aortic valve closes. Cardiac function is assessed on the basis of cardiac output (measured with flow probes) and aortic pressures. Owing to the physical constraints imposed by working within a magnet, most MR spectroscopy studies are performed using the Langendorff preparation. Temperature regulation is achieved using water-jacketed perfusion lines, by immersing the heart in the perfusion buffer and by means of a flow of warm air within the bore of the magnet.

Hearts isolated from rodents can be contained within commercially available NMR tubes (20–30 mm) and make use of commercially available broad-band or nuclei-specific NMR probes. Studies on hearts from larger mammals generally require an organ bath that incorporates a custom-built NMR coil within its structure or a surface coil attached to the left ventricular wall.

**Table 1** NMR-visible nuclei relevant to the study of perfused organs

| *Nucleus* | *Information obtained* |
|---|---|
| $^{1}H$ | Levels of lactate and creatine. Changes in lipid |
| $^{7}Li$ | Congener of $Na^+$. Measure $Na^+$ fluxes |
| $^{13}C$ | Substrate selection. Citric acid cycle activity |
| $^{19}F$ | Measure intracellular $Ca^{2+}$ using fluorinated $Ca^{2+}$ probes |
| $^{23}Na$ | Measure intracellular $Na^+$ levels |
| $^{31}P$ | Assess energy status |
| | Measure intracellular pH (from chemical shift of $P_i$)[a] |
| | Measure enzyme kinetics – saturation transfer for creatine kinase reaction |
| $^{87}Rb$ | $K^+$ congener. Measure $K^+$ fluxes |

[a]$P_i$ = inorganic phosphate.

## $^{31}P$ NMR spectroscopy

$^{31}P$ NMR spectroscopy is widely used for studies of isolated hearts. Using the endogenous $^{31}P$ signal arising from the tissue, it is possible to obtain information about the energy status of the heart and also to determine the intracellular pH (from the chemical shift of $P_i$). Assessment of extracellular pH is also possible using phosphonates that are confined to the extracellular space (e.g. phenylphosphonic acid or methylphosphonic acid). The heart metabolizes substrates (fatty acids, ketones, lactate, glucose, etc.), with the resultant energy being stored in the high-energy phosphate compound adenosine triphosphate (ATP). Most of this ATP is formed by mitochondrial oxidative phosphorylation. The phosphocreatine shuttle is responsible for transferring the energy from this mitochondrial ATP to sites of energy expenditure at the myofibrils and sarcolemma. **Figure 1** shows a typical $^{31}P$ spectrum obtained from an isolated guinea-pig heart. The phenylphosphonic

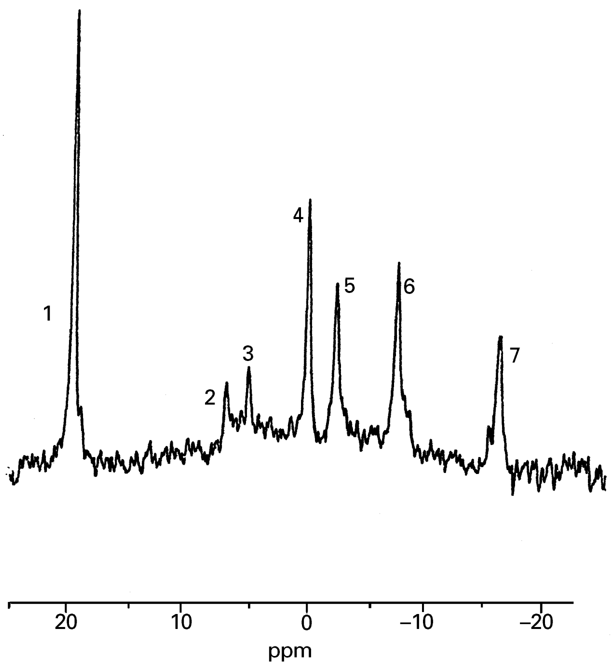



**Figure 1** $^{31}$P NMR spectrum of a guinea-pig heart perfused with Krebs–Henseleit solution. The spectrum was acquired at 8.7 T using a broad-band probe tuned to 145.8 MHz. Peak assignments are: 1, phenylphosphonic acid (external reference); 2, phosphomonoesters; 3, inorganic phosphate ($P_i$); 4, phosphocreatine (PCr); 5, $\gamma$-phosphorus of adenosine triphosphate (ATP); 6, $\alpha$-phosphorus of ATP; 7, $\beta$-phosphorus of ATP. The spectrum was acquired in 2.5 min by summing 72 free induction decays (FIDs) with a 35 µs pulse and a repetition time of 2 s. Prior to Fourier transformation the FID was subjected to exponential multiplication with a 20 Hz line broadening factor.

acid, which acts as an external standard, is contained within a capillary tube placed alongside the heart and contained entirely within the coil of the NMR probe. $^{31}$P NMR can detect phosphorus-containing compounds that are present in the fluid phase at concentrations of 0.6 mM or greater. The compounds visible by this technique are inorganic phosphate ($P_i$), phosphomonoesters (in this case sugar phosphates), phosphocreatine (PCr) and adenosine triphosphate (ATP). Adenosine diphosphate (ADP) is not visible because most of this nucleotide is protein bound within the cardiomyocytes.

Spectra are routinely collected with a repetition time of approximately 2 s, permitting the acquisition of data with adequate signal-to-noise in 2–5 min. This leads to 10–20% saturation of the PCr signal ($T_1 \approx 3$ s in rat heart at 8.7 T). The free induction decays are normally subjected to Fourier transformation following exponential multiplication using an appropriate line broadening (5–20 Hz). In many situations, alterations in the high-energy phosphate content of the heart over the course of an experiment are expressed as changes relative to the starting level. Absolute quantification of metabolite levels can be achieved by use of an appropriate external reference (e.g. phenylphosphonic acid in **Figure 1**) and correction for the partial saturation of the PCr signal. ATP contents are determined from the integral of the $\beta$ ATP peak; the $\alpha$ and $\gamma$ ATP peaks overlap with resonances from other molecular species. The PCr and $\beta$ ATP phosphates are 100% NMR-visible under aerobic conditions. The free ADP concentration can be calculated from the creatine kinase equilibrium equation,

$$K_{\mathrm{eq}} = \frac{[\mathrm{Cr}][\mathrm{ATP}]}{[\mathrm{PCr}][\mathrm{ADP}][\mathrm{H}^+]} \qquad [1]$$

where total creatine is normally determined biochemically and free [Cr] is determined from the difference between total [Cr] and [PCr]. Literature values for $K_{\mathrm{eq}}$ are ~$10^9$. The free energy of hydrolysis of ATP may be calculated from

$$\Delta G_{\mathrm{ATP}} = \Delta G_0 + RT \ln \frac{[\mathrm{ADP}][\mathrm{P_i}]}{[\mathrm{ATP}]} \qquad [2]$$

where $\Delta G^0$ is taken to be −30.5 kJ mol$^{-1}$. It is often suggested that $\Delta G_{\mathrm{ATP}}$ more accurately reflects the energetic capabilities of tissue than does a determination of the levels of high-energy phosphates.

$^{31}$P NMR can also be used to determine intracellular pH from the pH-dependent chemical shift of $P_i$ using the following formula based on the Henderson–Hasselbach equation:

$$\mathrm{pH} = \mathrm{p}K + \log\,[(\delta - 3.27)/(5.69 - \delta)] \qquad [3]$$

where p$K$ = p$K_2$ of inorganic phosphate (6.75). This technique yields an intracellular pH of 7.10–7.20 when the heart is perfused with buffer at pH 7.4.

$^{31}$P NMR spectroscopy has been applied to questions relating both to normal and to pathophysiological conditions. $^{31}$P NMR spectroscopy has been used in the normal heart to investigate the regulation of cardiac energy supply in response to increased demand. This study showed that there is no simple equilibrium between the phosphorylation potential and the mitochondrial redox state and that other factors are involved in coordinating energy supply and demand.

$^{31}$P spectroscopy has also been used to study the mechanisms responsible for myocardial ischaemia–

reperfusion injury. During ischaemia, blood (or perfusate) flow is restricted or totally occluded, resulting in an insufficient supply of oxygen to support oxidative metabolism. Anaerobic metabolism, in the form of glycolysis, is stimulated but is not adequate to maintain the energy balance. This leads to a depletion of high-energy phosphates. In addition, intracellular acidosis develops as a result of ATP hydrolysis and the accumulation of acidic end products of glycolytic metabolism. $^{31}P$ spectroscopy can follow the time course of changes in intracellular pH and high-energy phosphates during ischaemia and reperfusion (**Figure 2**). The effects of drug interventions on these profiles can provide insights into the mechanisms responsible for any observed cardioprotection conferred by the drug. Such studies also provide essential information on the roles of high-energy phosphate depletion and intracellular acidosis in ischaemia–reperfusion injury.

### $^{23}Na$ NMR spectroscopy

Ionic concentration gradients exist across cell membranes and are responsible for maintaining the resting membrane potential. Intracellular and extracellular $Na^+$ are approximately 10 mM and 140 mM, respectively. The converse is true for $K^+$, with an intracellular concentration of 130–140 mM and an extracellular concentration of 4–5 mM. Sodium enters the cells of excitable tissue during the up stroke of the action potential and potassium leaves the cell during the repolarization phase. The gradients are maintained by the operation of a $Na^+K^+$ ATPase (the sodium pump) that exchanges intracellular $Na^+$ for extracellular $K^+$. In the heart these ionic gradients can be disrupted by factors that prevent full activity of the sodium pump such as ischaemia or drugs (e.g. the cardiac glycosides related to digitalis). The ability to measure intracellular $Na^+$ levels in the intact heart makes $^{23}Na$ NMR spectroscopy a very powerful technique for assessing the role of altered $Na^+$ homeostasis in disease states.

Intracellular water represents about half the total water of the intact heart, the exact proportion being dependent on the species and perfusion conditions. This fact and the low intracellular concentration mean that, of the total $Na^+$ signal from the heart, less than 3% originates from the intracellular $Na^+$. Several studies have used double- or triple-quantum filtering techniques to discriminate this small intracellular $Na^+$ signal from the dominant extracellular $Na^+$ signal. Most studies, however, make use of non-cell-permeant paramagnetic reagents to 'shift' the extracellular peak and allow quantification of the intracellular $Na^+$ signal. These shift reagents are anionic chelates of lanthanide ions that do not cross membranes and thus are excluded from the intracellular space. The original agent used for this purpose, dysprosium bis(triphosphate), $Dy(PPP)_2^{7-}$ possesses the largest paramagnetic shift of any such complex for $Na^+$ or $K^+$. However, this reagent is quite sensitive to $Ca^{2+}$ and $Mg^{2+}$ and the shifts produced are greatly

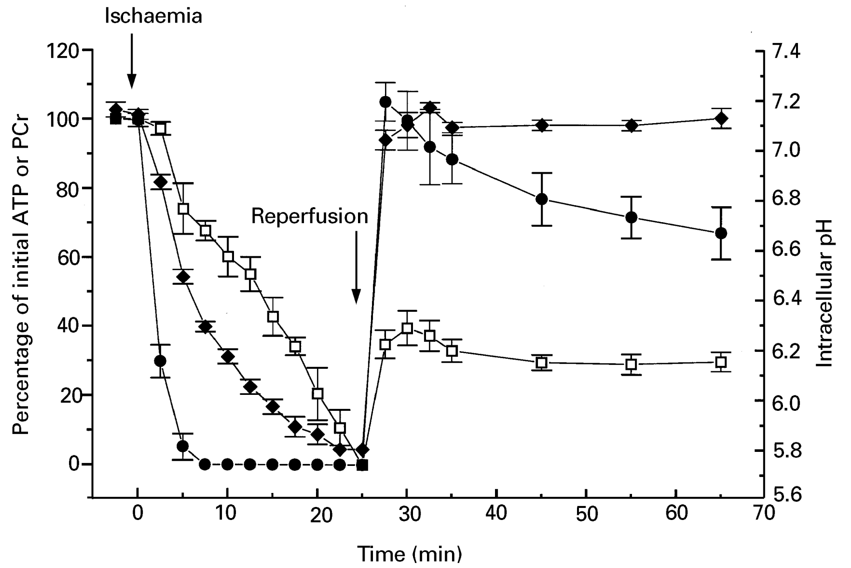



**Figure 2** Time course of changes in intracellular pH (♦), ATP(□) and PCr (•) in a rat heart subjected to 25 min of total global ischaemia followed by 30 min of reperfusion. pH was determined from the chemical shift of $P_i$. Changes in ATP and PCr are expressed as percentage change from the basal levels measured prior to ischaemia. Total global ischaemia was achieved by stopping all flow of perfusate to the heart.

reduced by the presence of these ions. This fact precludes the use of $Dy(PPP)_2^{7-}$ in the intact heart, which requires both $Ca^{2+}$ and $Mg^{2+}$ for full functional integrity. The triethylenetetraminehexaacetic acid chelate of dysprosium, $Dy(TTHA)^{3-}$, produces smaller shifts but is much less sensitive to the effects of $Ca^{2+}$ and $Mg^{2+}$. For studies in intact hearts that use $Dy(TTHA)^{3-}$, perfusates containing the shift reagent must be supplemented with $Ca^{2+}$ to offset the $Ca^{2+}$ -chelating properties of the shift reagent. This is well tolerated by the heart and results in adequate mechanical function. $Dy(TTHA)^{3-}$ at 5 mM causes a significant shift in the extracellular peak but also results in considerable line broadening, which still makes it somewhat difficult to fully resolve the small intracellular peak without specific processing strategies to maximize the resolution (**Figure 3**). Increasing the concentration of the shift reagent will cause a larger shift; however, the benefit is offset by an increase in line broadening.

The most recent shift reagent to be introduced is the tetraazacyclododecane-1,4,7,10-tetrakis(methylenephosphonate) chelate of thulium, $Tm(DOTP)^{5-}$ (**Figure 4**). This shift reagent also chelates $Ca^{2+}$ and the perfusate must be supplemented with $Ca^{2+}$ to maintain adequate mechanical function. At 4–5 mM $Tm(DOTP)^{5-}$ causes a significant shift in the extracellular $Na^+$ signal with very little line broadening. This latter property of the shift reagent makes it possible to perform interleaved $^{31}P$ and $^{23}Na$ NMR spectroscopy (in conjunction with a switchable NMR probe) in the presence of $Tm(DOTP)^{5-}$. Such studies are not possible in the presence of $Dy(TTHA)^{3-}$ owing to the excessive line broadening effects.

This strategy has been successfully applied to studies on isolated rat hearts to determine the involvement of the $Na^+$–$H^+$ exchanger in ischaemia–reperfusion. This sarcolemmal protein exchanges one $Na^+$ for one $H^+$. It is thought that this exchanger may contribute to myocardial ischaemia–reperfusion injury. During ischaemia, intracellular acidosis develops and this activates the $Na^+$–$H^+$ exchanger. This leads to an increase in intracellular $Na^+$ as intracellular $H^+$ is exchanged for extracellular $Na^+$. The increased intracellular $Na^+$ may activate the $Na^+$–$Ca^{2+}$ exchanger, with intracellular $Na^+$ exchanging with extracellular $Ca^{2+}$. The end result is an increase in intracellular $Ca^{2+}$, which may be a major factor in the deleterious effects of ischaemia–reperfusion injury. $^{31}P$ and $^{23}Na$ NMR experiments performed in the presence of $Tm(DOTP)^{5-}$ provided data on intracellular pH and the $Na^+$ content. Inclusion of a relatively specific inhibitor of the $Na^+$–$H^+$ exchanger (ethyl isopropyl amiloride) in the perfusate partially attenuated the changes in pH and $Na^+$ and significantly decreased mechanical dysfunction following ischaemia–reperfusion. This provided good evidence for the involvement of the $Na^+$–$H^+$ exchanger in ischaemia–reperfusion injury and confirmed the presumed mechanism of action of the drug. In most

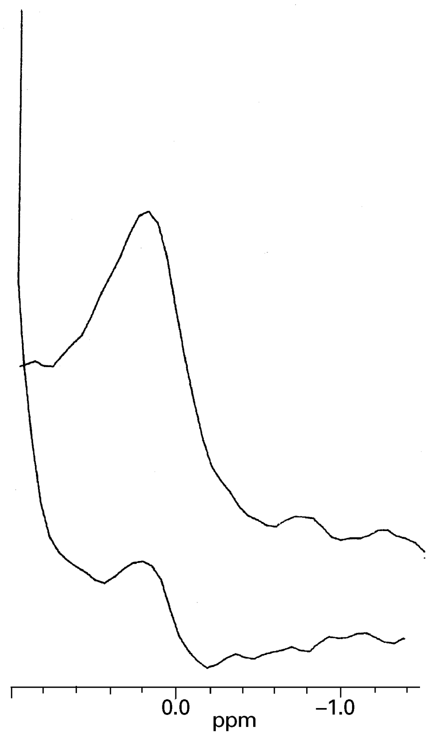



**Figure 3** $^{23}Na$ NMR spectra of a rat heart in the presence of 5 mM $Dy(TTHA)^{3-}$. Spectra were acquired at 8.7 T using a broadband probe tuned to 95.25 MHz. The lower trace is a spectrum acquired during normal perfusion. The addition of shift reagent to the perfusate has shifted the large extracellular $Na^+$ peak 2 ppm downfield and has also caused a 0.2 ppm shift in the smaller intracellular $Na^+$ peak. The upper trace is a spectrum acquired following 25 min of total global ischaemia. The intracellular $Na^+$ peak has grown substantially, reflecting the intracellular $Na^+$ accumulation that occurs during ischaemia. Resolution of the peaks was enhanced using Gaussian multiplication with line broadening of −25 Hz and GB parameter of 0.15.

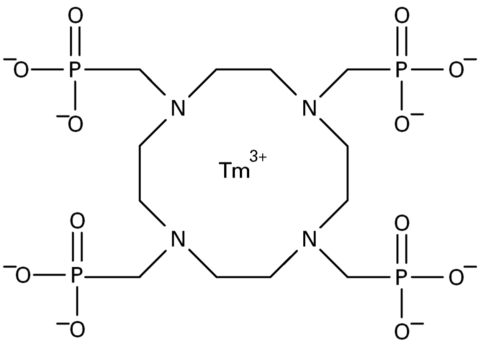



**Figure 4** $Tm(DOTP)^{5-}$.

studies only relative changes in $Na^+$ levels are reported rather than intracellular concentrations. This is in large part due to the need to make assumptions regarding the visibility of the $Na^+$ NMR signal under various experimental conditions.

### $^{13}C$ NMR spectroscopy

Most studies on isolated hearts use glucose as the sole energy source. Normally the heart would be exposed to a variety of substrates including glucose, pyruvate, lactate, acetoacetate and a mixture of fatty acids. $^{13}C$ NMR spectroscopy has been used to demonstrate that fatty acids and acetoacetate are the preferred substrates for the heart under normal physiological conditions. It is also important to determine how substrate selection and the efficiency with which the heart metabolizes these substrates are altered under pathological conditions. The citric acid cycle is the central pathway for energy production in the heart and it is critical to determine how the flux of metabolites through this pathway is altered in diabetes and cardiomyopathy and during reperfusion of the ischaemic myocardium. Citric acid cycle flux has been determined indirectly by measuring the enrichment of $^{13}C$ into glutamate from α-ketoglutarate by the action of aspartate aminotransferase. For these studies, hearts are provided with substrate, or substrate mixtures, highly enriched with $^{13}C$ at specific carbon atoms (e.g. [1-$^{13}C$]glucose, [1,2-$^{13}C$]acetate, [3-$^{13}C$]lactate, etc.). The contribution of selected substrates to overall citric acid cycle activity may then be determined by isotopomer and multiplet analyses of $^{13}C$ enrichment in glutamate. Such studies have been performed under steady-state and non-steady-state conditions. Analyses are most usually performed by high-resolution spectroscopy on trichloroacetic acid extracts of hearts perfused with $^{13}C$-enriched substrates, although useful data can be obtained by performing spectroscopy on intact beating hearts.

### $^{1}H$ NMR spectroscopy

The abundance of $^{1}H$ in water forms the basis for most magnetic resonance imaging. However, $^{1}H$ NMR spectroscopy has not found such universal applicability to studies of isolated hearts. This technique has been used to quantify the total content of creatine, which may provide a useful index of tissue viability. $^{1}H$ NMR spectroscopy has also been used to determine lactate levels during ischaemia and following interventions designed to modulate metabolism. The lactate methyl group and lipid methylene groups resonate at 1.3 ppm. These resonances can be differentiated and quantified using spin-echo spectral editing techniques.

### $^{19}F$ NMR spectroscopy

The most important use of $^{19}F$ NMR spectroscopy in studies of isolated hearts is to measure intracellular $Ca^{2+}$ levels. This is based on the use of fluorinated derivatives of calcium chelators. The extracellular $Ca^{2+}$ concentration is ~1.2 mM. The intracellular $Ca^{2+}$ concentration at diastole is less than 100 nM. This increases to several hundred nM during cardiac excitation. This elevated $Ca^{2+}$ (or calcium transient) is responsible for contractile activity at each heartbeat. Efficient relaxation at each beat depends upon the intracellular $Ca^{2+}$ being restored to diastolic levels. Most of the cytosolic $Ca^{2+}$ enters the cell through voltage-regulated $Ca^{2+}$ channels during the plateau phase of the action potential or is released from the intracellular organelle, the sarcoplasmic reticulum. Diastolic $Ca^{2+}$ levels are restored by active pumping of the $Ca^{2+}$ back into the sarcoplasmic reticulum and by activation of the sarcolemmal $Na^+$–$Ca^{2+}$ exchanger, exchanging intracellular $Ca^{2+}$ for extracellular $Na^{2+}$. Thus, the level of $Ca^{2+}$ within the cardiac cell is tightly controlled under normal physiological conditions. If the intracellular level of $Ca^{2+}$ rises significantly above normal physiological limits, consequences may be deleterious as a result of activation of $Ca^{2+}$-dependent proteases and phospholipases and also due to mitochondrial damage.

Intracellular $Ca^{2+}$ has been measured by $^{19}F$ NMR spectroscopy of intact hearts loaded with the 5,5′-difluoro derivative of 1,2-bis(*o*-aminophenoxy)ethane-N,N,N′,N′-tetraacetic acid (5F-BAPTA) (**Figure 5**). 5F-BAPTA is loaded into the heart as the cell-permeant acetoxymethyl (AM) ester. Esterases within the cardiomyocyte hydrolyse the AM ester to the free acid, which, being charged and therefore unable to cross the cell membrane, is trapped within the cell. The calcium-bound and calcium-free 5F-BAPTA species undergo slow exchange resulting in two NMR-visible peaks. The intracellular $Ca^{2+}$ concentration is

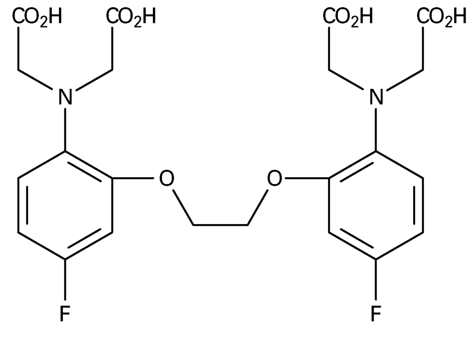



**Figure 5** 5F-BAPTA.

calculated using the relation

$$[\mathrm{Ca}^{2+}] = K_\mathrm{d}\frac{[\mathrm{B}]}{[\mathrm{F}]} \quad [4]$$

where $K_d$ is the dissociation constant of $Ca^{2+}$–5F-BAPTA (literature values of 537, 500, 635 and 285 nM), [B] is the area under the $Ca^{2+}$– 5F-BAPTA peak and [F] is the area under the 5F-BAPTA peak.

A $K_d$ of approximately 500 nM makes 5F-BAPTA an ideal probe for measuring intracellular $Ca^{2+}$ in the physiologically relevant range (100 nM to 1 μM). Unfortunately, a $K_d$ of 500 nM also causes 5F-BAPTA to provide excellent buffering to intracellular $Ca^{2+}$ levels. This causes a decrease in mechanical function of the heart. When 5F-BAPTA is present within cardiomyocytes at sufficiently high concentration (300 μM) to provide adequate signal to noise in the $^{19}F$ NMR spectrum (> 10:1), developed pressure is reduced by 75–80%. This results from an increase in diastolic pressure and a decrease in systolic pressure. Thus, although studies using 5F-BAPTA can provide valuable information regarding changes in intracellular $Ca^{2+}$, it must always be borne in mind that these results were acquired under conditions of severely compromised cardiac function.

A BAPTA derivative with a higher $K_d$ overcomes the problem of buffering of intracellular $Ca^{2+}$. The recently developed analogue, 1,2-bis(2-amino-5,6-difluorophenoxy)ethane-N,N,N′,N′-tetraacetic acid (TF-BAPTA) (**Figure 6**) has a $K_d$ of 65 μM. Loading of intact hearts with this derivative causes <10% decrease in mechanical function. This high $K_d$ causes TF-BAPTA to be a less accurate probe for measuring basal $Ca^{2+}$ levels but does make it more suitable for measuring intracellular $Ca^{2+}$ under pathophysiological conditions (e.g. ischaemia) where $[Ca^{2+}]$ may rise above 2 μM. The $Ca^{2+}$-bound and $Ca^{2+}$-free TF-BAPTA species are in intermediate–fast exchange. This leads to a single resonance with its chemical shift position being dependent upon the extent of $Ca^{2+}$ binding (analogous to the pH-dependent shift in the $P_i$ peak). Binding of $Ca^{2+}$ to TF-BAPTA does not alter the chemical shift of the fluorine in the 6 position of TF-BAPTA. The fluorine in the 5 position shifts downfield upon binding of $Ca^{2+}$. The chemical shift difference between the 5 and 6 fluorines is used to determine intracellular $[Ca^{2+}]$ (with corrections for the effects of pH and $[Mg^{2+}]$). TF-BAPTA accumulates in the sarcoplasmic reticulum (SR) and this property has been used to determine $[Ca^{2+}]$ in this subcellular organelle in intact rabbit hearts. A 5-F resonance at 5 ppm (with the 6-F resonance set at 0 ppm) corresponds to a time-averaged basal cytosolic $[Ca^{2+}]$ of 600 nM. A second 5-F resonance at 14 ppm corresponds to a SR $[Ca^{2+}]$ of 1.5 mM. This technique may prove to be invaluable in assessing the role of alterations in SR $Ca^{2+}$ handling in various pathological conditions.

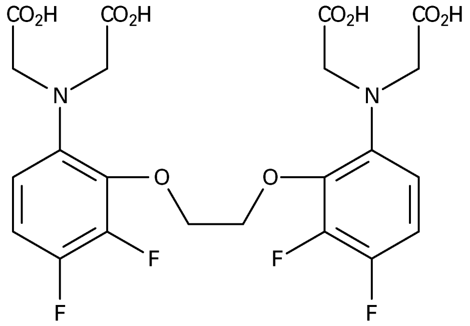



**Figure 6** TF-BAPTA.

### $^{87}$Rb and $^{7}$Li NMR spectroscopy

$^{23}Na$ spectroscopy is useful for measuring steady-state levels of intracellular sodium but is not suitable for measuring fluxes of this cation. The contribution of various ion channels, exchangers and pumps to the total movement of ions across the sarcolemmal membrane can be assessed by measuring ion fluxes in the presence and absence of selective inhibitors of these membrane proteins. It is possible to acquire this type of information using electrophysiological techniques. These data would be complemented by the use of techniques that can measure ion fluxes in the intact heart. $^{87}Rb$ and $^{7}Li$, congeners of $K^+$ and $Na^+$, respectively, have been used to assess fluxes of $K^+$ and $Na^+$ in intact hearts and vascular tissue. For $^{87}Rb$ spectroscopy, 20% of the perfusate $K^+$ content can be replaced with $Rb^+$ with no effects on function. $^{87}Rb$ NMR spectroscopy can be used to determine rate constants for the uptake of $Rb^+$ on switching from $Rb^+$-free to $Rb^+$-supplemented perfusate. The converse, switching from $Rb^+$-supplemented to $Rb^+$-free perfusate, can be used to determine the rate constant for $Rb^+$ washout. This approach has been used to demonstrate that under normal conditions the bulk of $Rb^+$ uptake occurs through the $Na^+K^+$ ATPase rather than the $Na^+$–$K^+$–$2Cl^-$ exchanger or $K^+$ channels. $^{7}Li$ NMR spectroscopy has been used in similar types of experiments to study $Na^+$ channel activity in intact hearts. For these studies an excellent signal could be achieved by substituting a modest amount of perfusate $Na^+$ with $Li^+$ (15 mM). Similar to the $^{87}Rb$ studies, the kinetics of release of $Li^+$ could be determined in washout experiments. These studies demonstrated that $Li^+$ efflux from cardiomyocytes is predominantly through $Na^+$ channels.

## Studies of isolated livers

Livers isolated from rats or mice may conveniently be studied by NMR spectroscopy. The organs are perfused through the portal vein with solutions similar to those described for isolated heart studies (supplemented with albumin). The methods described above for $^{31}P$ and $^{23}Na$ NMR spectroscopy have been successfully applied to the isolated liver. The $^{31}P$ NMR spectrum of the isolated liver differs from the heart spectrum in that it lacks the PCr peak observed in spectra obtained from hearts. This technique has been utilized to examine various aspects of ethanol metabolism in the liver, including the effects of chronic ethanol exposure on subsequent acute ethanol exposure and hypoxia. These studies revealed that chronic ethanol exposure caused an adaptation in the liver such that it becomes more resistant to acute ethanol exposure and also to hypoxia.

$^{15}N$ NMR spectroscopy has been used to study the urea cycle in isolated liver. $^{15}N$ was provided to the liver in the form of $^{15}NH_4Cl$ or [$^{15}N$]alanine in the presence and absence of unlabelled lactate or ornithine. Proton-decoupled spectra were obtained from the intact liver, from the perfusing medium or from tissue extracts and yielded peaks corresponding to glutamine, arginine, urea, citrulline, glutamate, alanine and ammonia. Such studies may prove to be useful in the *in vivo* liver as a means of assessing the effects of disease states on urea cycle activity.

## Studies of isolated kidneys

Kidneys isolated from rodents are suitable for study by NMR spectroscopy following cannulation of the renal artery and perfusion with an albumin-supplemented Krebs–Henseleit solution. Combined $^{31}P$ and $^{23}Na$ NMR spectroscopy has been used to determine the energetic cost of $Na^+$ transport in the kidney. Similar multinuclear techniques have been used in studies investigating the factors influencing renal function during the progression from pre-hypertension to hypertension in a spontaneously hypertensive rat model. $^{31}P$ NMR spectroscopy is being applied to studies into the viability of kidneys used for transplant. At present, donor availability is the limiting factor for transplant programmes. As the use of non-heart-beating donors increases, there is a greater need for a rapid, reliable noninvasive technique for assessing organ viability. $^{31}P$ NMR spectroscopy is showing promise in this regard.

## List of symbols

[A] = concentration of species A; $K_d$ = dissociation constant; $K_{eq}$ = equilibrium constant; $p_{A=}$ (partial) pressure of species A; $P_i$ = inorganic phosphate; $\delta$ = chemical shift; $\Delta G^0$ = free energy change under standard conditions; $\Delta G_{ATP}$ = free energy of hydrolysis of ATP.

*See also:* **Cells Studied By NMR; Chemical Shift and Relaxation Reagents in NMR; $^{13}C$ NMR, Methods; *In Vivo* NMR, Applications, Other Nuclei; *In Vivo* NMR, Applications, $^{31}P$; $^{31}P$ NMR.**

## Further reading

Balaban RS (1989) MRS of the kidney. *Investigative Radiology* **24**: 988–992.

Barnard ML, Changani KK and Taylor-Robinson SD (1997) The role of magnetic resonance spectroscopy in the assessment of kidney viability. *Scandinavian Journal of Urology and Nephrology* **31**: 487–492.

Deslauriers R, Kupryianov VV, Tian G *et al* (1996) Heart preservation: magnetic resonance studies of cardiac energetics and ion homeostasis. In: Dhalla NS, Beamish RE, Takeda N and Nagano M (eds) *The Failing Heart*, pp 463–487. Philadelphia: Lippincott-Raven.

Elgavish GA (1993) Shift reagent-aided $^{23}Na$ nuclear magnetic resonance spectroscopy. In: Pohost GM (ed) *Cardiovascular Applications of Magnetic Resonance*, pp 371–391. Mount Kisco, NY: Futura.

Evanochko WT and Pohost GM (1993) $^{1}H$ NMR studies of the cardiovascular system. In: Schaefer S and Balaban RS (eds) *Cardiovascular Magnetic Resonance Spectroscopy*, pp 185–193. Norwell, MA: Kluwer Academic.

Ingwall JS (1993) Measuring sodium movement across the myocardial cell wall using $^{23}Na$ NMR spectroscopy and shift reagents. In: Schaefer S and Balaban RS (eds) *Cardiovascular Magnetic Resonance Spectroscopy*, pp 195–213. Norwell, MA: Kluwer Academic.

Kusuoka H, Chacko VP and Marban E (1993) Measurement of intracellular $Ca^{2+}$ in intact perfused hearts by $^{19}F$ nuclear magnetic resonance. In: Pohost GM (ed) *Cardiovascular Applications of Magnetic Resonance*, pp 393–401. Mount Kisco, NY: Futura.

Malloy CR, Sherry AD and Jeffrey FMH (1993) $^{13}C$ nuclear magnetic resonance methods for the analysis of citric acid cycle metabolism in heart. In: Pohost GM (ed) *Cardiovascular Applications of Magnetic Resonance*, pp 261–270. Mount Kisco, NY: Futura.

Ugurbil K and From AHL (1993) Nuclear magnetic resonance studies of kinetics and regulation of oxidative ATP synthesis in the myocardium. In: Schaefer S and Balaban RS (eds) *Cardiovascular Magnetic Resonance Spectroscopy*, pp 63–92. Norwell, MA: Kluwer Academic.

# PET, Methods and Instrumentation

**TJ Spinks,** Hammersmith Hospital, London, UK

Copyright © 1999 Academic Press



## Introduction

This article discusses the ways in which the design of instrumentation for positron emission tomography (PET) has evolved to provide data with ever greater accuracy and precision. Even though a PET tracer may be intrinsically capable of providing highly specific biochemical and physiological information, this is of little use if the radiation detection system is inadequate. The central aims of PET instrumentation are to increase the number of unscattered photons detected, making maximum use of the amount of tracer it is permissible to administer, and to resolve with greater accuracy their point of origin. A complementary aim is to reduce statistical noise in the data. In addition, developments in the measurement of tracer in the blood are important. A central component of quantification in PET is the combined use of tracer input to the tissue (arterial blood) and tissue uptake (determined from the image). Later in the article a brief discussion will be made of the methods used to obtain physiological, biochemical and pharmacological parameters using different mathematical models of the biological system under investigation.

## Detector materials and signal readout

The overall aim in the search for new detectors is to maximize photon stopping power (efficiency) and the signal generated and to minimize response time. The latter property is required to cut down losses due to dead time or the time during which the detection system is 'busy' dealing with one interaction and thereby misses others that might occur. Theoretical predictions can be made of these properties for different materials, but good detectors have generally been found by a process of trial and error. The dominant detector material in PET today is bismuth germanate ($Bi_4Ge_3O_{12}$, or BGO for short), although sodium iodide (NaI(Tl), activated with thallium) is still quite widely used. A comparison of the principal properties of these substances and those of a promising new material that is being developed (lutetium orthosilicate, LSO) is shown in **Table 1**. LSO is very attractive because it has a similar density to BGO but also has about five times its scintillation efficiency and a very much shorter scintillation decay time. In consequence, its energy resolution and timing resolution are higher.

**Table 1** Characteristics of some scintillation detectors used in PET

| *Property* | *Sodium iodide (NaI(Tl))* | *Bismuth germanate (BGO)* | *Lutetium orthosilicate (LSO)* |
|---|---|---|---|
| Density (g cm$^{-3}$) | 3.7 | 7.1 | 7.4 |
| Effective atomic number | 51 | 75 | 66 |
| Scintillation efficiency (% of NaI(Tl)) | 100 | 15 | 75 |
| Scintillation decay time (ns) | 230 | 300 | 40 |
| Hygroscopic? | Yes | No | No |

The method by which the response of a detector is recorded and analysed (the 'readout') has undergone several developments, and refinements are still being made. Most PET tomographs at present utilize the BGO block detector, a device introduced in the mid-1980s. This was revolutionary in the sense that it enabled a multiple-ring tomograph to be produced at a reasonable cost. Earlier designs had utilized one-to-one coupling of crystal and photomultiplier (PMT), but the multiplicity of coincidence circuits had restricted commercial scanners to at most two detector rings. Furthermore, because of the desire to reduce detector size, the dimensions of readily available PMTs made one-to-one coupling impractical. The schematics of the block detector are shown in **Figure 1**.

A BGO crystal is divided into a number of elements (in modern scanners $8 \times 8$ with dimensions $4 \times 4$ mm$^2$ face $\times$ 30 mm depth) which are viewed by four relatively large, standard PMTs. The cuts between the elements contain light-reflecting material and are made to different depths across the crystal. This partial separation controls the amount of scintillation light reaching the PMTs and leads to a ratio of signals characteristic of each detector element. The $x$ and $y$ coordinates of the interaction are calculated by the formulae given in **Figure 1**. A ring of such block detectors would constitute eight rings of closely packed detector elements and multiple rings

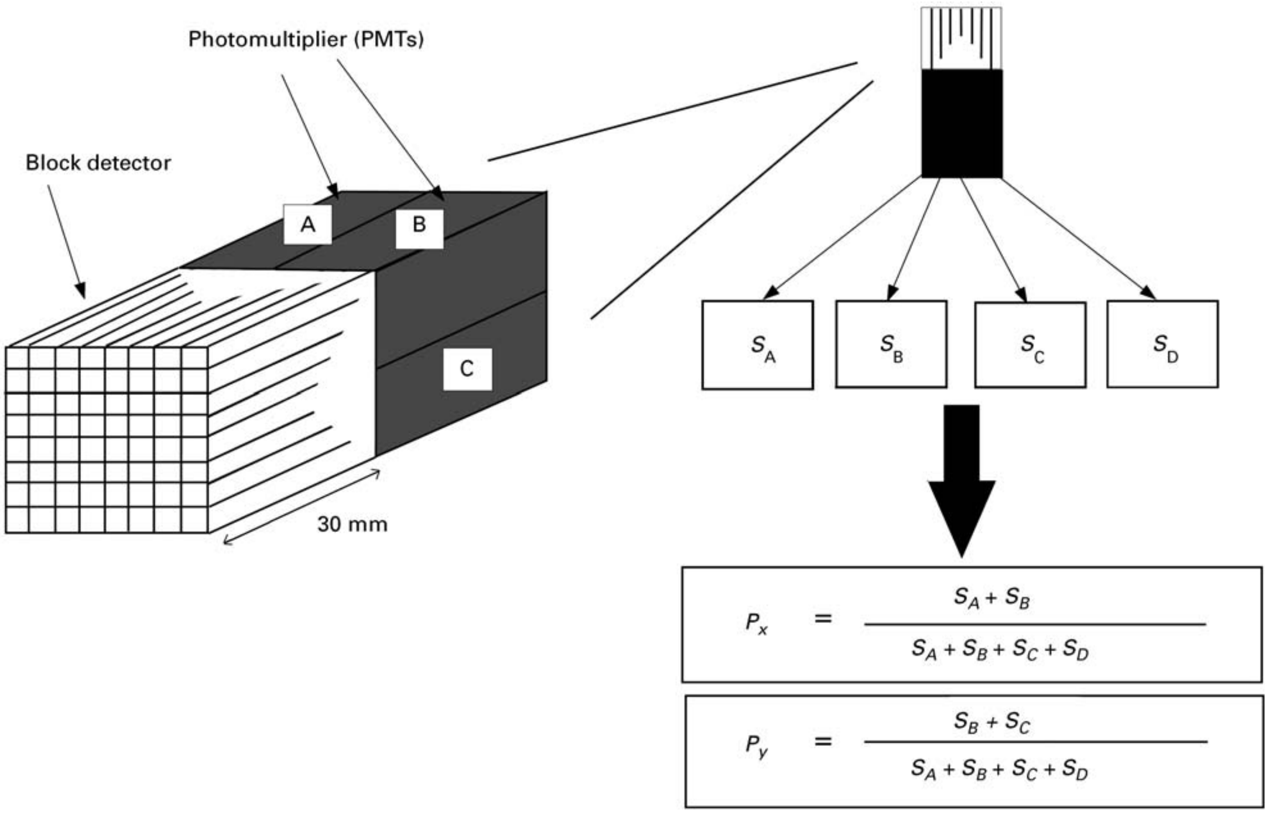



**Figure 1** Principle of the block detector.

of blocks can be placed adjacent to each other to give the desired field of view (FOV) (within financial constraints). With this arrangement, the degree of multiplexing of signals is greatly reduced. On the other hand, the block detector has poorer dead-time properties than an individual crystal–PMT pairing. When one element is hit by a photon, all other elements are effectively 'dead' until the original interaction has been analysed. In some experimental systems avalanche photodiodes (diode detectors with internal amplification) are used instead of PMTs. Their availability and performance still does not compete with those of PMTs, but developments continue and their great advantage is their very compact size and ability to view small detector elements.

Full collection of scintillation light from a photon interaction results in a charge pulse from the PMT output whose height is proportional to the energy deposited. The timing of a given event is determined by the point at which the pulse crosses a certain voltage level. The statistical nature of pulse generation and the range of pulse heights encountered give rise to a spread or uncertainty in this timing. An ingenious electronic device known as the constant fraction discriminator greatly reduces this spread by shaping the pulse so that triggering is made at a constant fraction of the pulse height. The rise and fall of scintillation light determine timing resolution $\tau$ (how accurately the time of a single interaction can be determined) and integration time (how much time is required to measure the scintillations produced and hence energy deposited). For BGO, the timing resolution is about $6 \times 10^{-9}$ s (6 ns) and the integration time about $10^{-6}$ s (1 µs). The time spectrum of two detectors (the measured time differences between the arrival of two events) shows a peak superimposed on a constant background (for a given activity). The peak corresponds to true coincidences and the background to random coincidences. The detection of coincidence events involves a convolution of the timing resolutions of each detector and the coincidence timing window is thus set at twice the timing resolution (i.e. $2\tau$). If this window is made narrower, the random events will fall linearly but the true events will decrease more rapidly. It is observed experimentally that a window of 12 ns is optimum for BGO detectors.

Even if a beam of monoenergetic 511 keV photons is incident on a crystal, the different scattering and absorption interactions occurring lead to different amounts of energy being deposited and result in output signals with a range of intensities. **Figure 2** is a sketch of the energy spectrum from a BGO detector. The photopeak at the right represents events that

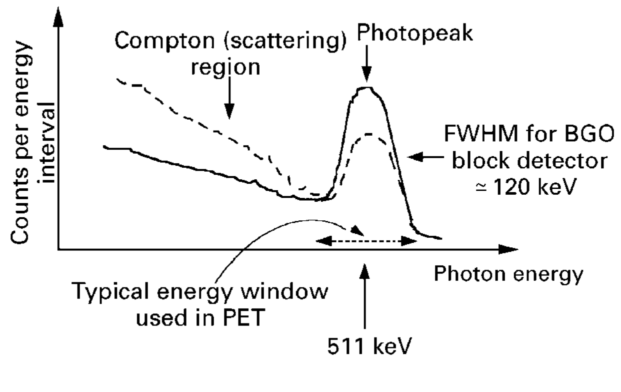



**Figure 2** Shape of energy spectrum from a BGO detector for monoenergetic 511 keV photons (——) and those from a scattering medium (– – –).

have been totally absorbed (photoelectric effect), while the broad continuum corresponds to partial energy deposition (Compton scattering). If the source is within a scattering medium such as the body, the scattering region of the spectrum will be enchanced and the photopeak depressed. The width of the photopeak or the energy resolution (full width at half maximum, FWHM) is about 120 keV for a BGO block detector. (It is only about half this for an uncut block owing to better light transferrence to the PMT.) In a similar way as for the timing spectrum, the acceptance of events is restricted by setting an energy window of about 2 × FWHM (about 350–650 keV). This scheme attempts to minimize detection of incident scattered (lower energy) photons but, although BGO has high stopping power, its scintillation efficiency and hence energy resolution is relatively poor and a fraction of scattered photons are inevitably accepted. Methods for subtracting these events are discussed later in this article.

## 2D and 3D acquisition modes

The great advantage or annihilation coincidence detection is the automatic or 'electronic' collimation that it provides, but scattered photons are an ever-present complication, the effects of which need to be minimized. The first PET tomographs consisted of 'area' detectors such as the large sodium iodide crystals of gamma cameras that were already a standard device for radioisotope imaging. However, it became clear that such scanners had a high sensitivity to scattered photons and random coincidences. Subsequent designs were based on rings of individual detectors with tight collimation using lead shielding. The first commercial scanner comprised a single ring of NaI(Tl) crystals with variable lead collimators in front of the detectors and heavy shielding on either side of the ring. Multiple-ring designs arose from the inevitable demand for a larger axial FOV and in these a form of side shielding known as septa, consisting of lead or tungsten annuli, were inserted between the rings as shown in **Figure 3A**. In such an arrangement, data are acquired as a number of contiguous transaxial 'planes' or 'slices' and it is, for this reason, termed the '2D mode'. Each plane consists of data from coincidences either within an individual ring or in closely adjacent rings. The example given in **Figure 3B** shows a maximum ring difference ($rd_{max}$) of 1, but, as axial detector width became narrower to improve resolution, greater values of $rd_{max}$ were used in order to maintain efficiency. However, it can be seen that for large ring differences coincidence counts would be severely attenuated by the septa (as shown by the dotted line). The approach to statistical image noise has emphasized the importance of maximizing

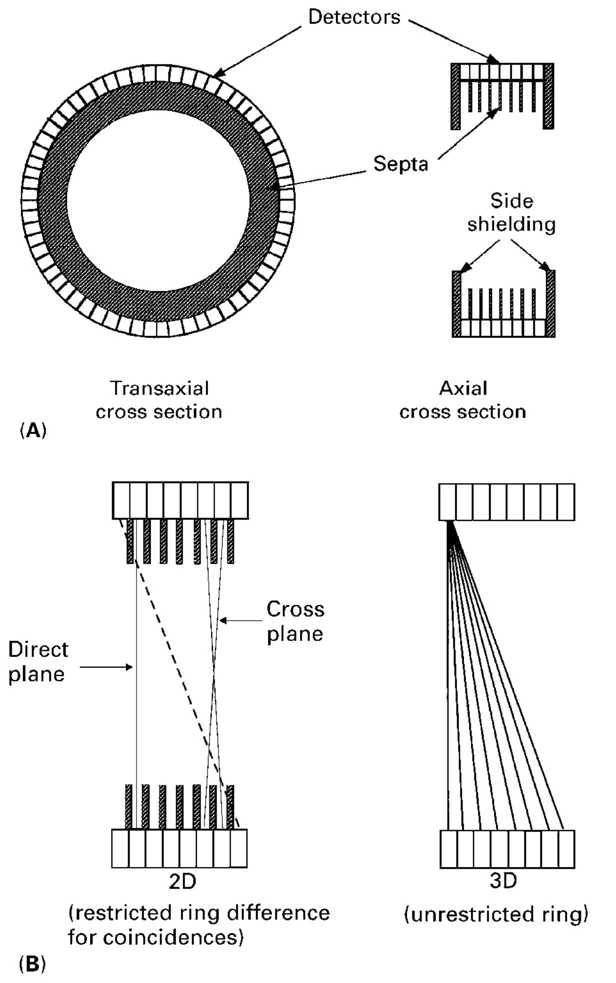



**Figure 3** (A) Arrangement of inter-ring septa in multi-ring scanners. (B) Axial cross section of a multiple-ring tomograph illustrating inter-ring coincidence combinations for 2D and 3D modes.

detection efficiency. The restriction of data to '2D' slices tends to go against this and so there have been increasing moves during the 1990s to acquire data without septa inserted and with all possible coincidence lines-of-response (LORs) operational. Such a scheme is naturally known as the '3D mode'. Objections to 3D PET scanning stem, of course, from the original desire to reduce the effects of scattered and random coincidences, but the proponents of the method point to the significant increase in the efficiency of detection of unscattered true coincidences. For example, with a tomograph having eight crystal rings and a 2D $rd_{max}$ of 1 (**Figure 3B**), the total number of inter-ring combinations is 22 (8 direct $+ 2 \times 7$ cross), whereas for 3D the number is 64 $(8 \times 8)$. This threefold increase is further enhanced (by about a factor of 2) owing to 'shadowing' of the crystals by the septa.

This simple analysis is appropriate at low count rates when random events are negligible, but at higher rates the advantage must be defined in terms of improvement in statistical noise as embodied in the parameter noise-equivalent counts (NEC). It is found that there is an NEC gain with the 3D mode over the whole range of count rates. For brain studies this varies from about a factor of 5 at low rates (for example, in receptor binding studies with $^{11}C$-labelled compounds) to about 3 at high rates (such as are encountered in blood flow studies with $^{15}O$-labelled water).

## Data handling and image reconstruction

In addition to the reticence concerning the 3D mode on physical grounds, there are practical challenges. Current hardware designed to process coincidence events runs at rates sufficiently high to accommodate the tracer doses that can be administered, but a principal problem is one of data transfer, storage and reconstruction. 2D data are 'compressed' into a number of slices, while 3D data are conventionally acquired without any (or relatively little) compression. The increase in the number of (individual) LORs approaches an order of magnitude and these have to be backprojected during reconstruction and stored. However, despite these perceived disadvantages, the storage and archiving media (multi-gigabyte hard disks, DAT tape, optical disks etc.) and ever faster processors available today have lessened the practical burdens of 3D scanning.

A feature of data acquisition that is becoming of greater interest in PET is that of list mode. In this, the events are not stored over preselected time frames in the sinogram data matrices but instead each event is stored separately to disk. If a large number of counts are acquired over a short time, this is not necessarily an efficient method, but for studies following the uptake and clearance of tracer for 2 hours or more, it becomes much more efficient. However, even for shorter scanning times, list mode provides very high temporal ($\simeq$ 1 ms) resolution and the ability to utilize physiological gating, or separation of the data into specific phases, for example of the cardiac and respiratory cycles. Furthermore, list mode acquisition would be of great advantage in the correction for patient movement, which is still an area of development.

Readily available dedicated processing hardware can reconstruct a set of images from the largest 3D PET data volume in about 10 minutes but the ability of tomographs to acquire data in 3D mode somewhat preceded the development of appropriate reconstruction algorithms. The specific problem with filtered backprojection lay in the variation of the point response function (PRF). Reconstruction of 2D slices relies on the fact that the response to a point source (efficiency of detection) remains constant over the slice. This requirement needs to be met because the filtering process represents a convolution with the measured projection data. In the 3D mode the PRF is not constant over the FOV, being a maximum in the centre and falling steadily (axially) towards the edge (**Figure 4A**). A way of overcoming this difficulty, and one that is now commonly used, is the reprojection method. In this (**Figure 4B**), 2D images are first reconstructed from direct plane (single ring) data and then 'missing' projections are created by forward projection (images to views). The missing projections are those that would have been acquired if additional detector rings had been present and these complement the data to provide an invariant PRF over the (original) FOV.

## Alternative tomograph designs

Tomographs consisting of multiple rings of individual crystals are the most widely used. **Figure 5** shows the CTI/Siemens model 966 (covers removed) which is the most sensitive PET scanner yet constructed. However, there are a number of other designs in operation.

### Tomographs based on planar sodium iodide detectors

Large-area planar sodium iodide (NaI(Tl)) detectors, which are used routinely in gamma cameras for imaging of single-photon tracers, have seen a revival in

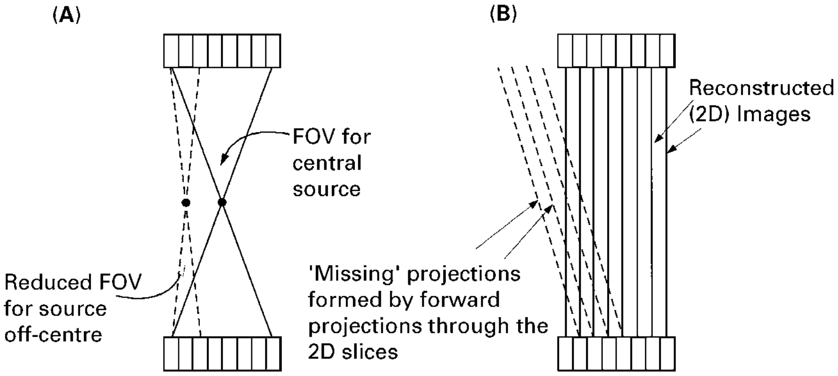



**Figure 4** (A) Variation of point response function (PRF) in a multiple-ring tomograph. (B) Creation of 'missing' projections for the reprojection reconstruction algorithm.

PET over the last few years. A commercial design consists of six planar detectors arranged in a hexagon and operated without septa (3D mode). The crystal is viewed by an array of PMTs and the point of photon interaction is determined, in a similar way to the block detector, by comparing signals between adjacent tubes. This device takes advantage of the high light output (**Table 1**) of NaI(Tl) and possesses a similar spatial resolution to BGO systems. On the other hand, its stopping power (efficiency) is significantly less than that of BGO. Indeed, the standard detector thickness used for single-photon imaging (at about 100 keV) of about 10 mm is increased to 25 mm for the 511 keV photons in PET. One advantage of this system is that large NaI(Tl) crystals can be produced at relatively low cost, giving a large FOV. This aspect and the increasing availability of the glucose analogue tracer [$^{18}$F]fluorodeoxyglucose (FDG, see later for clinical diagnostic imaging) have been responsible for its commercial success.

There is also increasing interest in 'dual-use' systems for both single-photon emission tomography (SPET) and PET, consisting of double-headed gamma cameras (operated with and without multihole lead collimators, respectively). The significantly lower efficiency for PET compared with purpose-built tomographs needs to be borne in mind, but such systems could be useful for specific diagnostic tests (e.g. detection of tumour metastases) and their flexibility for general nuclear medicine use is attractive.

### Partial-ring tomographs

As alluded to above, the expense of multiple-ring BGO tomographs has led to the development of lower-cost systems, principally those employing NaI(Tl) planar detectors. An alternative to these is the commercially available partial-ring scanner. In this device, two banks of BGO detector blocks (about 1/3 of a complete ring) rotate around the body (at 30 revolutions per minute) and the output of data is achieved via optical coupling. Voltage supplies are provided through slip rings.

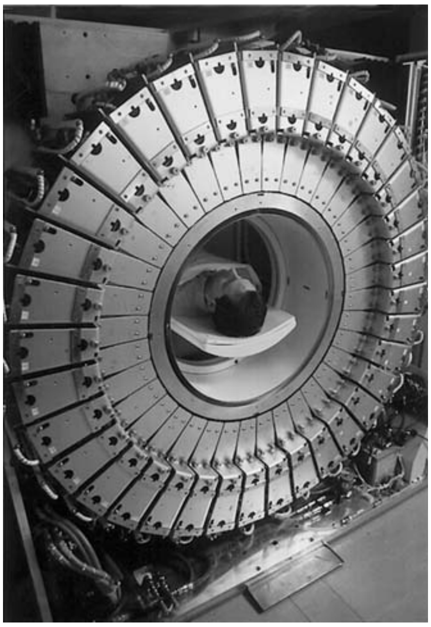


**Figure 5** A commercial PET scanner.

### Tomographs based on multiwire proportional chambers

The multiwire proportional chamber (MWPC) is a device used extensively in high-energy particle physics experiments that has been modified in different ways for use in PET. The principle of operation is the detection of an electron (i.e. an electron avalanche) on planes of closely spaced fine wires (~1 mm apart) held at a high electric potential. Cathode and anode wires are arranged orthogonally to

each other, thus providing the spatial localization, the wire spacing being the basic determinant of resolution. The electrons are produced in various ways, such as interaction of photons in thin sheets of lead or photoionization of a gas in the wire chamber by ultraviolet light from a barium fluoride ($BaF_2$) scintillator.

### Tomographs for experimental studies

A number of research centres have designed and implemented smaller-diameter tomographs (BGO, NaI(Tl) and multiwire chamber detectors have all been used) specifically for the scanning of small animals. Spatial resolution close to 1 mm has been achieved. Although this does not compete with the resolution of autoradiography or dissection, the time–activity curve can be followed in a single animal, which is a great advantage in many studies, such as the investigation of new tracers and the testing of models of disease processes. It is anticipated that important new information will also be forthcoming in the development of new pharmaceuticals.

## Data correction procedures

### Normalization

Variations in the fabrication of detectors and their geometrical arrangement in a tomograph inevitably lead to variations in efficiency for different LORs. For example, the detectors at the edge of a BGO block have lower efficiency than those at the centre and the LORs crossing the centre of the FOV will have a different solid angle of detection to those at the edge. If these effects are not corrected for, systematic errors (of both high and low spatial frequency) will occur in the reconstructed image. The process of correction is known as normalization. The basic data for normalization are acquired by exposing each LOR to the same activity, for example in the form of a thin planar source or a line source scanning across the FOV.

### Attenuation

The principles of attenuation and its correction in PET are outlined in an associated article. Here, some specific practical examples are given. Most PET tomographs utilize sources of $^{68}$Ge (half-life about 9 months), which are stored in shields within the gantry of the tomograph and moved into the FOV by remote control. In turn, a blank scan (empty FOV) and a transmission scan (patient in position, usually before tracer administration) are acquired. The ratio between these two provides the attenuation correction factors for each LOR. The logarithm of these ratios can also be backprojected to yield a transmission image (tissue density map). As for emisssion data, scattered radiation (see below) contaminates the data and one method of reducing this is electronic 'windowing'. This is illustrated in **Figure 6**. The transmission source in this case is a rotating rod whose position is encoded. If a photon from the rod is scattered in the object and measured, the event will be rejected because the resulting LOR (LOR2) does not pass through the rod. Unscattered events (e.g. LOR1) are accepted.

A continuing theme of PET is improvement in efficiency, and this is so for transmission scanning. Constraints of time and detector performance mean that transmission data can be suboptimal and this has led to the recent implementation of the single-photon transmission technique. This is similar to the process of X-ray computerized tomography and takes advantage of the fact that about two orders of magnitude more single-photon events are collected than coincidence events. A working mechanism employs a $^{137}$Cs point source (single-photon: energy 662 keV, half-life 30 years) rotating in a helical tube around the subject. However, as for the rotating $^{68}$Ge source in **Figure 6**, detection of scattered radiation will also contaminate the data in this case and lead to inaccurate correction. Recourse to the windowing method cannot be made for $^{137}$Cs, but another way to compensate for scatter is to create a transmission image that is then segmented into regions of similar density, and correct attenuation factors are assigned accordingly.

### Scattered radiation

The inclusion of scattered events in the projection data does not significantly affect the spatial resolution or 'sharpness' of the reconstructed image but it will lessen the contrast between different regions and cause inaccuracies in the measurement of activity concentration. The distributions of scatter due to a line source placed axially in the centre of a water-filled cylinder (20 cm diameter) are shown in **Figure 7** for 2D and 3D acquisition modes. The central peak corresponds to the position of the source and the broad continuum ('wings') on either side is due to scatter, which decreases with increasing angle. The scatter fraction (SF, integral of the 'wings' divided by the total events) for brain scanning is 10–15% with septa and 30–40% without septa; even larger values obviously occur in body scanning.

A number of methods have been used to correct for scatter, particularly for the 3D scanning mode. Correction schemes have broadly been based on the measured spatial or energy distribution (spectrum)

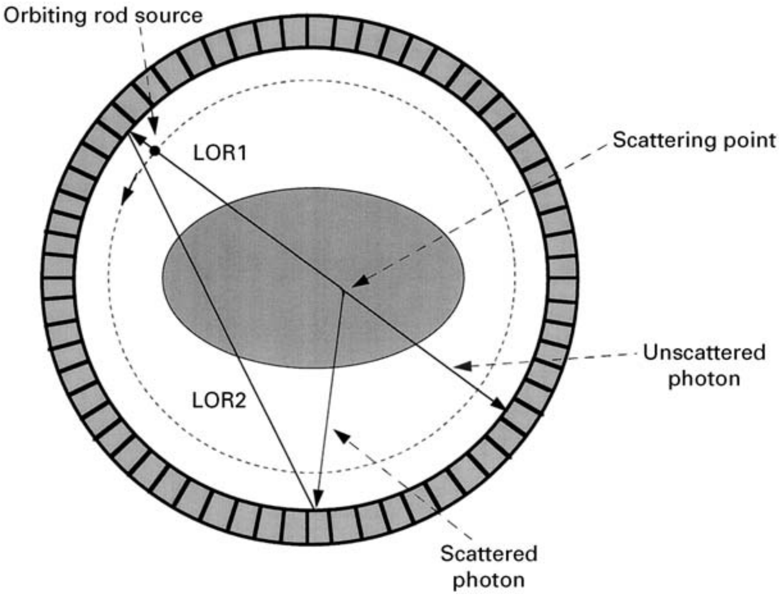



**Figure 6** Mechanism for attenuation correction using a rotating $^{68}Ge$ rod with 'windowing'.

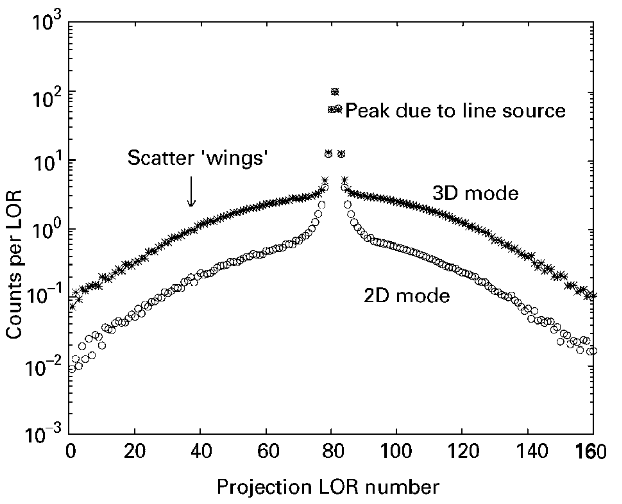



**Figure 7** Scatter distributions for a line source in the centre of a cylinder (2D and 3D); peaks are normalized to the same height.

of scattered events or, more recently, on calculation from first principles of the probability of scattering through different angles.

The first scatter correction methods, which are still employed, treated the scatter distribution as a convolution of the projection data with a function or 'kernel', the shape of which was derived from curves such as that in **Figure** 7. The simplest form of the kernel is $\alpha\exp(-\beta|x|)$, where $\alpha$ and $\beta$ are positive constants and $|x|$ is the absolute distance along a projection. In general, convolution is carried out iteratively because the measured projections contain scatter and so the initial calculation is an overestimate. A shortcoming of the simple kernel above is that the shape of the scatter distribution does not stay constant with the position of the source. As it moves to one side of the object, the distribution becomes more and more asymmetric. Strictly, in this case, an integral transform rather than a convolution should be employed since the kernel shape is position dependent. However, quite accurate scatter corrections for the head have been demonstrated, even in the 3D mode, with an invariant kernel.

A correction method based on the photon energy spectrum utilizes the varying proportions of scattered and unscattered events for different ranges or energy windows. This technique had its origins in single-photon tomography. Some PET tomographs have the ability to acquire data in two energy windows (over the photopeak and over part of the Compton continuum). A simple form of the dual-window correction is based on the assumption that the ratios are constant throughout the object. The dual-window method has been shown to give similarly accurate corrections (within about 6%) to the convolution method in 3D brain scanning but has slightly poorer noise characteristics.

The simplest method of scatter correction consists of fitting a function (typically a Gaussian) to the 'wings' (pure scatter) outside the object. This has the advantage of being based on a direct measurement of scatter for each object. On the other hand, assumptions must be made about the shape of the distribution *inside* the object. Again, accurate results are

obtained for the head, but for the chest a simple Gaussian is not necessarily a good choice of function.

The speed of current computers has made it feasible to calculate scatter distributions analytically from first principles for each set of data. This is carried out by taking the uncorrected tracer (emission) image and calculating the scatter that would arise from selected points given the transmission image. The probabilities of scattering through given angles are known precisely from physical principles. In practical implementations, a relatively coarse grid of points is selected. This saves time and utilizes the fact that the scatter distribution is smooth and is amenable to interpolation. The beauty of this technique is that it is makes few assumptions. Convolution and dual-window methods work well in a relatively uniform object such as the head but have more difficulty, for example, in the chest where there are abrupt density changes. A big challenge for PET is to obtain accurate quantification in 3D body scanning and it is likely that analytical methods such as this will prove the most valuable.

## Dead time and random coincidences

As the activity in the body increases, the problems of electronic dead time and registration of random coincidences, leading to a reduction in efficiency and an increase in statistical noise, become ever more pressing.

Dead time correction schemes have ranged from those that are founded on an intimate knowledge of the electronic circuitry of the scanner to those based on empirical curve-fitting using test objects. All commercial scanners provide automatic dead time correction. The variation of true and random coincidences with activity is shown in **Figure 8A**. The peak and steady falloff in the trues curve at high rates is described as paralysable dead time behaviour and is typical of PET tomographs. In this case, if an event occurs during the analysis of a previous event ('busy' period) the effect is a successive lengthening of the dead time. Theoretically it can be shown that measured ($N_m$) and corrected ($N_0$) count rates are related by

$$N_m = N_0 \exp(-N_0 \tau_d) \qquad [1]$$

in which case the measured rate can eventually go to zero (or be 'paralysed') for very high activities ($\tau_d$ is the system dead time). The fundamental determinant of dead time is the rate of single events that are striking the detectors. Tomographs continuously record this rate and use it as a basis of the calculation of dead time correction factors. It should be noted that, for coincidence counting, the overall dead time is the *product* of the dead times of opposing detectors. Generally speaking, correction is accurate to within 5% for the range of counting rates encountered in *in vivo* studies (**Figure 8B**), but it needs to be stressed that the magnitude of the correction factor should not be too large and that scanning should not be carried out near to or beyond the peak of the 'trues' curve.

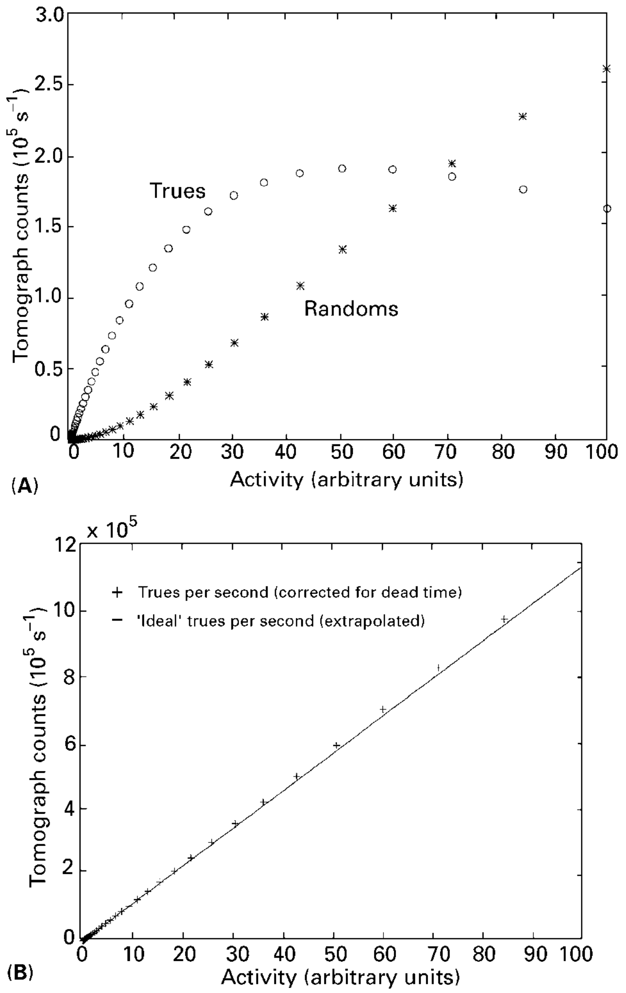



**Figure 8** (A) The variation of measured trues and randoms with activity in the FOV. (B) The correspondence between 'ideal' trues rates (extrapolated from low activities) and dead time-corrected trues.

The steep rise in random events with activity is, similarly to dead time, due to the product of the rates of single events on opposing detectors. Without dead time, trues rise linearly with activity, whereas randoms rise quadratically (the dead time for trues and randoms is the same because they are counted by the same coincidence circuits). This behaviour of randoms means that judicious administration of activity should be adhered to and/or good shielding of

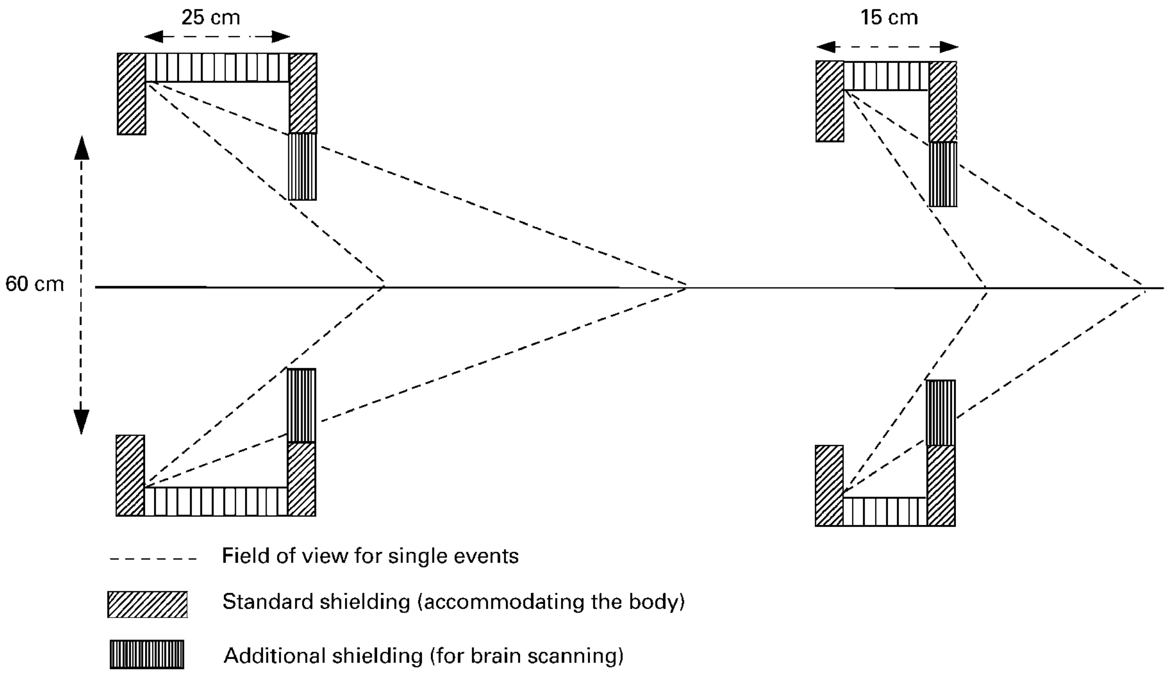



**Figure 9** The change in FOV for single events for different detector ring diameters and axial lengths.

activity outside the FOV should be provided. The move towards 3D PET and the desire to accommodate any part of the body and to increase the axial FOV pose problems in this regard. The dilemma can be illustrated by simple geometry (**Figure 9**). The largest axial FOV for a multi-ring tomograph is about 25 cm (left of **Figure 9**). With standard side shielding giving an aperture of 60 cm diameter, the FOV for single photons extends about 75 cm beyond the coincidence FOV. With the insertion of additional shielding appropriate for the brain (an aperture of 35 cm), the single-photon FOV is reduced significantly. The more common axial detector length of about 15 cm clearly gives a much reduced singles FOV, which for brain scanning barely extends beyond the side shielding. Providing effective shielding for body studies is not easy, but a way of processing the random events in order to lessen their effect is to apply some form of smoothing. As for scatter, the distribution of randoms is of a broad, low-frequency nature and thus amenable to smoothing. The future alternative to this is to use detectors with faster response, narrower coincidence window and proportionately lower randoms.

### Spatial resolution effects

The spatial spread of a point source in an image gives rise to a phenomenon known as spillover, in which activity in one region affects that measured in an adjacent region. A source of decreasing size in a 'background' of lower activity concentration will appear to have decreasing activity. This is purely an effect of finite resolution and not of efficiency and is often referred to as the partial volume effect (that is, the object only partially fills the detector resolution field-of-view and is 'mixed' with its surroundings); it is also expressed by saying that the recovery coefficient of the object is less than unity. This is a difficult problem, but one correction technique that has been applied in the brain relies on the much higher resolution of a magnetic resonance image to provide accurate anatomical data.

Coincidence detection gives good uniformity of resolution along an LOR, but for circular ring systems the resolution in the radial direction gradually worsens owing to the interaction of photons with detectors at increasingly oblique angles (**Figure 10**). This effect is magnified as the ring diameter decreases. The desire for smaller diameters (and hence less expensive systems) requires a remedy for this non-uniformity. A number of methods of correction are being tested, such as the dual use of photodiodes and PMTs on either ends of a detector, but no one method has yet gained general acceptance.

## Models of tracer kinetics

Even if all physical corrections have been applied and the image is an accurate representation of tracer distribution, the further big challenge in PET is to derive biochemical, physiological and pharmacological

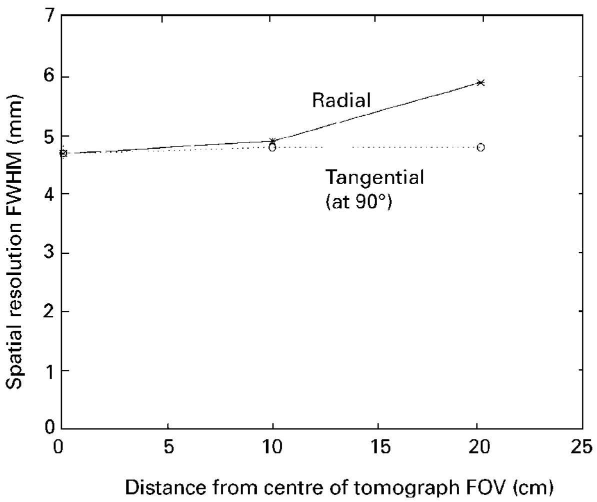



**Figure 10** Variation of spatial resolution in radial and tangential (at right angles) directions with distance from the centre of the tomograph FOV.

parameters from the data. This is carried out by diverse mathematical models of the biological system. What may be termed a conventional approach is to treat the system as a number of separate compartments within each of which the tracer is uniformly distributed. The passage of tracer between compartments is described by rate constants that have dimension per unit time ($time^{-1}$). An example of such a system is shown in **Figure 11**, which divides the volume under investigation into blood (plasma) and free and fixed tracer in tissue. In PET research blood activity is usually measured either by taking discrete samples or by on-line monitoring in a detector placed by the side of the patient. A refinement of this is analysis of the blood into different radioactive components or metabolites because the tracer itself is broken down in its passage through tissue.

A compartmental model describes the system by a number of differential equations from which the rate constants are estimated. A common application of the model in **Figure 11** is in the dynamics of glucose utilization. The most widely used tracer in PET is an analogue of glucose known as FDG ($^{18}$F-labelled fluorodeoxyglucose). It is transported into the cell and undergoes the biochemical process of phosphorylation similarly to glucose, but it then remains 'trapped' in the tissue. Because of this, good quality images of glucose utilization can be obtained. Over the period of a study (an hour or so), there is negligible release from the fixed compartment ($k_4$ in **Figure 11** is negligible) and this simplifies the model. In this case, the metabolic rate for glucose (MRGl) is given by

$$\mathrm{MRGl} = \frac{C_p k_1 k_3}{\mathrm{LC}(k_2 + k_3)} \qquad [2]$$

where $C_p$ is the (natural) glucose concentration in the plasma and LC is a constant that describes the difference in transport and phosphorylation rates between FDG and glucose. Examples of functional images of MRGl (two slices through the brain) are displayed in **Figure 12** (left) along with magnetic resonance (MRI) images showing anatomical detail (right) and coregistered PET and MRI images (centre). Coregistration covers a number of techniques used to overlay functional and anatomical images which, for example, make use of specific anatomical markers or the minimization of the variance between the two images.

Other types of tracer model do not seek to impose a compartmental structure but instead determine a combination of kinetic components which best fit the data. For example the technique of spectral analysis views the tissue response ($C_{tiss}$) as a convolution between the plasma/blood input ($C_p$) and a large but finite range of so-called basis functions of the form $\gamma\exp(-\delta t)$ where $t$ is time:

$$C_{tiss}(t) = C_p(t) \otimes \sum_{j=1}^{j=n} \gamma_j \exp(-\delta_j t) \qquad [3]$$

When the $\gamma_j$ are fixed and the $\delta_j$ are constrained to be zero or positive, it transpires that typically only two to four positive $\delta_j$ are given for the time–activity curve on each image pixel. The resulting parameters are used to generate the impulse response function (the tissue time–activity curve resulting from a unit pulse input at $t = 0$). The intercept of this function (for each pixel) at $t = 0$ gives an image of the clearance of tracer from blood to tissue (denoted $K_1$) and its integral gives the volume of distribution (concentration in tissue relative to blood, denoted $V_d$).

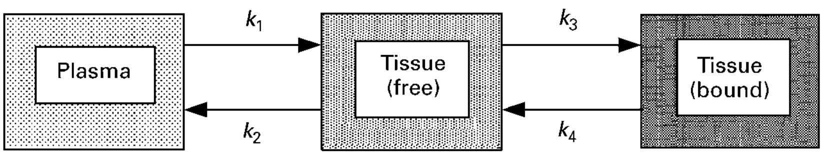



**Figure 11** Three-compartment tracer model.

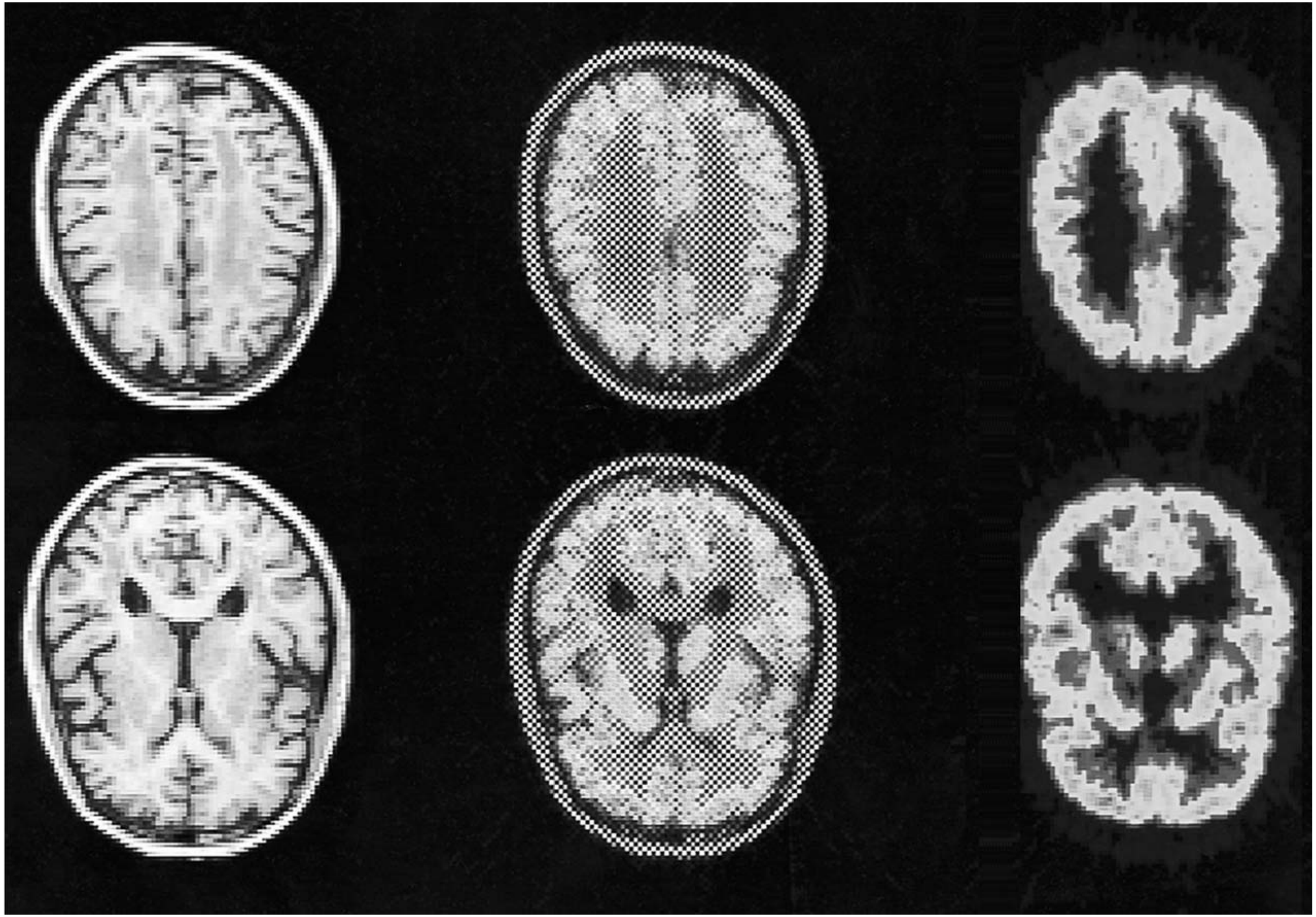


**Figure 12** PET functional images of glucose metabolic rate (MRGl) (right), MRI images (magnetic resonance images, left) showing anatomical detail, and coregistered (overlaid) PET/MRI (centre). (See Colour Plate 42).

Frequency analysis is an example of a more objective way of extracting kinetic information from PET time–activity data. Other forms of this are factor, principal components and cluster analyses. These attempt to define pixels in the projection or image data that have similar kinetic characteristics. Their fundamental aim is the derivation of more specific and objective images of function rather than purely radioactivity concentration, a task that is enhanced by the superior physical properties of positron annihilation coincidence detection.

## List of symbols

$C_p$ = plasma concentration, plasma response; $C_{tiss}$ = tissue response; $k_{1-4}$ = rate constants; $K_1$ = tracer clearance (blood to tissue); MRGl = metabolic rate for glucose; $N_0$ = corrected count rate; NEC = noise-equivalent counts; $N_m$ = measured count rate; $rd_{max}$ = maximum scanner ring difference; $t$ = time; $V_d$ = volume of distribution; $\delta_j$ and $\gamma_j$ = parameters of basis functions; $\tau$ = timing resolution; $\tau_d$ = system dead time.

*See also:* **MRI Applications, Biological; MRI Applications, Clinical; MRI Instrumentation; PET, Theory; SPECT, Methods and Instrumentation; Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules; Scattering Theory; SPECT, Methods and Instrumentation; Two-Dimensional NMR, Methods.**

## Further reading

Barrett HH and Swindell W (1981) *Radiological Imaging: The Theory of Image Formation, Detection, and Processing*. San Diego: Academic Press.

Bendriem B and Townsend DW (eds) (1998) *The Theory and Practice of 3D PET*. Dordrecht, Boston, London: Kluwer Academic.

Carson RE, Daube-Witherspoon ME, Herscovitch P (eds) (1998) *Quantitative Functional Brain Imaging with Positron Emission Tomography*. San Diego: Academic Press.

Casey ME and Nutt R (1986) A multicrystal two dimensional BGO detector system for positron emission tomography. *IEEE Transactions on Nuclear Science* NS-**33**: 570–574.

Cho ZH and Farukhi S (1977) New bismuth germanate crystal — a potential detector for the positron camera application. *Journal of Nuclear Medicine* **18**: 840–844.

Knoll GF (1979) *Radiation Detection and Measurement*. Chichester: Wiley.

Melcher CL and Schweitzer JS (1992) Cerium-doped lutetium orthosilicate: a fast, efficient new scintillator. *IEEE Transactions on Nuclear Science* **39**: 502–505.

Murray IPC, Ell PJ and Strauss HW (eds) (1994) *Nuclear Medicine in Clinical Diagnosis and Treatment*. Edinburgh: Churchill Livingstone.

Myers R, Cunningham V, Bailey D and Jones T (eds) (1996) *Quantification of Brain Function Using PET*. San Diego: Academic Press.

Phelps M, Mazziotta J and Schelbert H (eds) *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart*. New York: Raven Press.

Schwaiger M (ed) (1996) *Cardiac Positron Emission Tomography*. Boston: Kluwer Academic.

Spinks TJ, Jones T, Bailey DL *et al* (1992) Physical performance of a positron tomograph for brain imaging with retractable septa. *Physics in Medicine and Biology* **37**: 1637–1655.

# PET, Theory

**TJ Spinks**, Hammersmith Hospital, London, UK

Copyright © 1999 Academic Press

| SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS |
|---|
| Theory |

## Introduction

The acronym PET can be used to stand for both positron-emitting tracers and positron emission tomography; it is a representation of the radioactive isotopes used and the methods by which their distribution is visualized (tomography derives from the Greek tomos or ‘cut’ – i.e. an imaged slice through the body). PET is the most sensitive and specific method of studying molecular interactions and pathways in the living organism and is assuming ever greater importance in medical diagnosis and research, in the understanding of biochemistry and physiology in health and disease, and in the development of drugs. Most of the world’s PET centres are in North America, Europe and Japan but many other

countries are making plans for the installation of PET facilities. Most of the centres are dedicated to diagnosis, particularly in heart disease and cancer, but there are also a number of centres associated with large medical research institutes that concentrate purely on research. This article will deal with the theoretical aspects of PET: (a) isotope production, (b) radiation interactions and detection, (c) data acquisition and image formation, and (d) properties of the image of radioisotope distribution.

## Physics of the positron

A positron (positively charged electron) is a particle of so-called 'antimatter' that cannot coexist for long with the 'ordinary' matter of which we and all that surrounds us is made. Such particles were postulated in the 1930s by the physicist Paul Dirac, who pictured the vacuum as a 'sea' of electrons in negative energy levels that could be excited into positive energy levels by the absorption of quanta of energy. Although this concept was not readily accepted by most physicists, the existence of the positron was demonstrated experimentally by Anderson three years after the theoretical prediction. It was observed that a photon, of energy greater than or equal to twice the rest mass energy of the electron, in the field of the nucleus could give rise to the simultaneous appearance of a positron and an electron. This is known as the pair-production process. The positrons used in PET, however, arise from the disintegration of atomic nuclei that are unstable because they have an 'excess' of positive charge.

## Production of positron emitters

Positron-emitting atoms do not normally exist in nature. The radionuclides used for PET are usually produced by a cyclotron, which, by harnessing powerful electric and magnetic fields, accelerates charged particles (such as protons, deuterons or alpha particles) to high energies (about 2–5% of the speed of light); these then bombard stable atoms in a target to give rise to radioactive isotopes. **Table 1** gives examples of the reactions used to produce the principal radionuclides used in PET ($^{11}C$, $^{15}O$, $^{18}F$ and $^{13}N$). It can be seen from column 3 that there are generally more protons than neutrons in each of the product nuclei. This 'excess' charge is released during nuclear disintegration (beta decay) by the emission of a (positively charged) positron (or in a smaller fraction of cases by the capture of an orbiting electron). Two principal characteristics of the positron emitters in **Table 1** that are responsible for their success as *in vivo* radiotracers are that (a) they are radioisotopes of major body elements (or in the case of $^{18}F$ serve as *in vivo* analogues) and (b) they have short half-lives. These properties enable them to label biological molecules without altering the biochemical action and to be injected into a patient or normal volunteer in usable quantities with an acceptably low radiation dose. The physical characteristics of detection of these tracers, which will be described below, provide additional reasons for their pre-eminence in nuclear medicine. However, the short half-lives demand that the scanner (tomograph) and cyclotron are in close proximity. Such a necessity has given rise to the impression that PET is an expensive technique, but an increasing number of clinical centres are obtaining tracers from shared central cyclotron facilities and lower-cost scanner designs are commercially available.

## Positron annihilation

Positrons are emitted from nuclei of a given isotope with a range of energies up to a characteristic maximum 'end point' energy $E_{max}$ (**Table 2**), the mean energy being roughly one-third $E_{max}$. Positrons lose their energy by Coulomb interactions with atomic electrons, following a tortuous path until they are brought to rest within a precisely defined range (dependent on their energy and the effective atomic number of the medium). The ranges for mean and maximum energies in soft tissue are given in **Table 2**. When the energy of the positron is close to zero, the probability of interaction with an electron is highest. From direct interaction or after the formation of a transient system with an electron known as positronium, two photons, each of energy 511 keV (the rest mass energy of the electron or positron), are emitted in opposite directions with the disappearance (annihilation) of both particles. The 'back-to-back' photon emission arises from the conservation of momentum. However, there is only precisely 180° between the photon directions if the net momentum is zero at annihilation. The small residual momentum of the positronium system leads to an angular spread of about ± 0.3°. Positron range and the angular spread determine the physical limits of spatial resolution in PET.

## Annihilation coincidence detection

Simultaneous detection of the annihilation photons provides significantly greater efficiency and improved uniformity of spatial resolution than detection of individual photons. These points are illustrated schematically in **Figure 1**. Consider a point source positron emitter in air moving across the channel between the two detectors $D_1$ and $D_2$ (**Figure 1A**). The channel is conventionally termed a

**Table 1** Production and characteristics of principal isotopes used in PET

| *Positron emitting product* | *Stable element* | *Nuclear reaction* | *Number of protons (p) and neutrons (n) in the product nucleus* | *Half-life of product* (min) | *Stable nucleus after positron emission* |
|---|---|---|---|---|---|
| $^{11}C$ | Nitrogen($^{14}N$) | $^{14}N(p,\alpha)^{11}C$ | 6p, 5n | 20.4 | $^{11}B$ |
| $^{18}F$ | Oxygen($^{18}O$) | $^{18}O(p,n)^{18}F$ | 9p, 9n | 109.8 | $^{18}O$ |
| $^{15}O$ | Nitrogen($^{14}N$) | $^{14}N(d,n)^{15}O$ | 8p, 7n | 2.03 | $^{15}N$ |
| $^{13}N$ | Carbon ($^{12}C$) | $^{12}C(d,n)^{13}N$ | 7p, 6n | 10.0 | $^{13}C$ |

**Table 2** Positron ranges in soft tissue for the principal positron emitters

| | *Positron energy* (MeV) | | *Positron range in soft tissue* (mm) | | *Contribution to resolution* (mm FWHM) |
|---|---|---|---|---|---|
| *Positron emitter* | *Maximum* ($E_{max}$) | *Mean* | *Maximum* | *Mean* | |
| $^{18}F$ | 0.635 | 0.250 | 2.6 | 0.61 | 0.2 |
| $^{11}C$ | 0.970 | 0.386 | 4.2 | 1.23 | 0.3 |
| $^{13}N$ | 1.200 | 0.491 | 5.4 | 1.73 | 0.4 |
| $^{15}O$ | 1.740 | 0.735 | 8.4 | 2.97 | 1.2 |

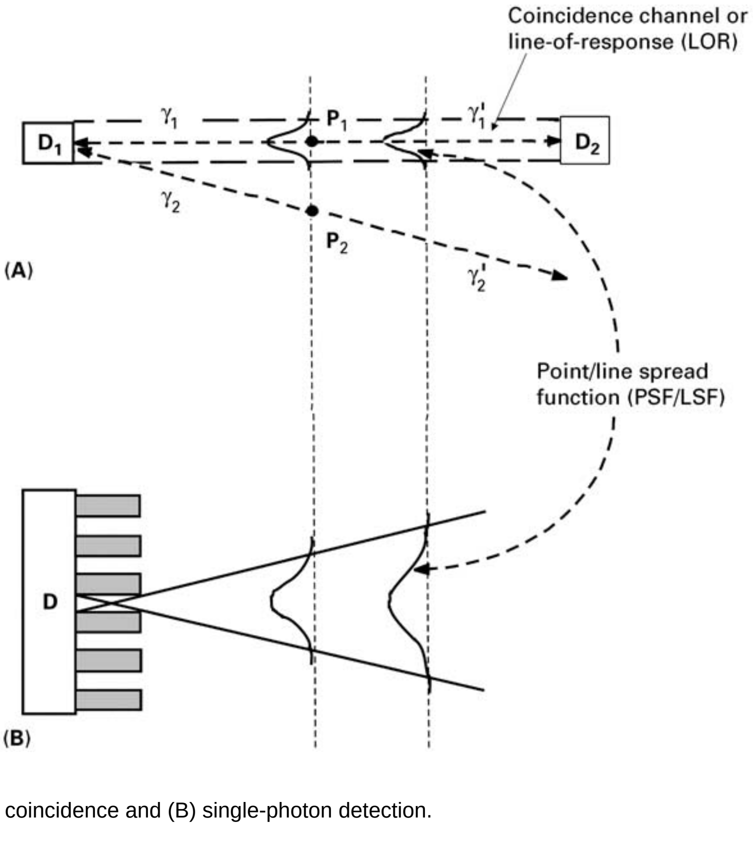



**Figure 1** (A) Annihilation coincidence and (B) single-photon detection.

'line-of-response' (LOR) and the detectors are connected such that a 'count' is only registered when a photon interaction is recorded in each detector at the same time (or within the time resolution of the detector, see below). Only when the source is within the LOR (e.g. position $P_1$) do annihilation photons such as $\gamma_1$ and $\gamma_1'$ have a chance of being detected as a coincident pair. With the source at position $P_2$ (e.g. photons $\gamma_2$ and $\gamma_2'$) there is no possibility of a coincidence event. Therefore, coincidence counting confers a natural collimation of the radiation, often referred to as electronic collimation and, in addition, the

resolution varies relatively little for different positions along the LOR. The variation in counts obtained by passing a point or (orthogonal) line source across the LOR is termed the point or line spread function (PSF, LSF) and is conventionally characterised by its full-width at half-maximum (FWHM). The PSF obtained in this way represents the intrinsic resolution of the detector pair, in other words, the best achievable.

To obtain spatial localization with a source emitting only single photons (γ rays from nuclear disintegration), a physical (lead) collimator has to be placed between source and detector D (**Figure 1B**). The solid angle subtended by the collimator aperture shows that the resolution in this case will vary with distance from the detector. The aperture can be made smaller to give a spatial resolution as high as desired, but this will be at the expense of detection efficiency and so some compromise is needed. Single-photon emission tomography (SPECT) systems normally consist of large-area detectors, whereas most modern PET systems consist of thousands of coincidence detector pairs and the detection efficiency in PET is about 100 times that of SPECT.

## Photon interactions and attenuation in scattering media

The previous discussion of coincidence detection becomes more complicated when the source is within a scattering medium (such as body tissue). At the energy of the annihilation photons (511 keV), the possible interactions (with atomic electrons) are photoelectric (total) absorption and coherent and incoherent (Compton) scattering. Scattering refers to change of direction without (coherent) or with (incoherent) energy loss. Compton scattering is overwhelmingly predominant at 511 keV in the body (the probability of photoelectric absorption and coherent scattering can be considered negligible) and is thus, practically speaking, totally responsible for the removal of a photon from a particular LOR. This is termed attenuation and is illustrated in **Figure 2A**. If the scattering medium is uniform, it will have a constant attenuation coefficient ($\mu$), dependent on the effective atomic number, defined as

$$\mathrm{d}N = \mu N_x \,\mathrm{d}x \qquad [1]$$

where d$N$ are the number of photons scattered within a distance $x$ to $x + \mathrm{d}x$ and $N_x$ is the number of unscattered photons at $x$. Integrating Equation [1] gives the following formula for the number of photons ($A_T$) left

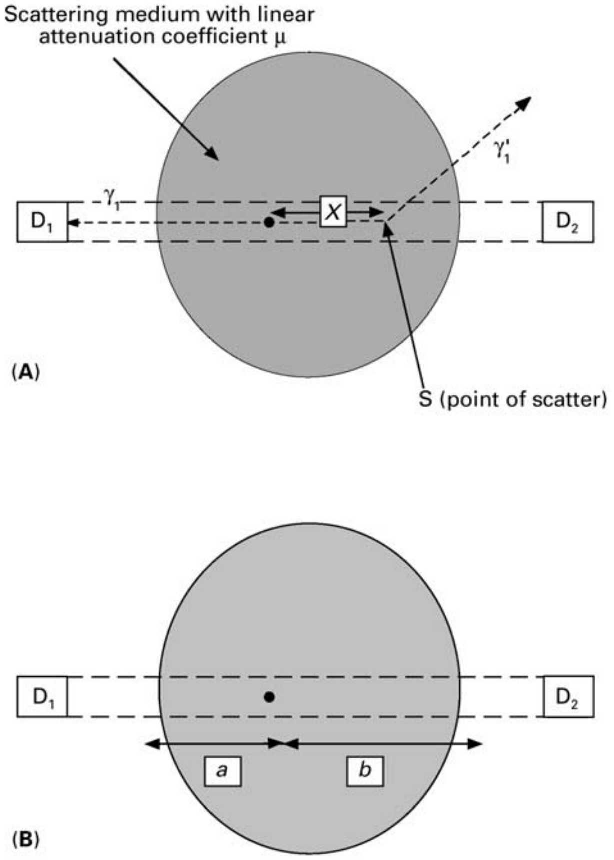



**Figure 2** Photon attenuation in a given LOR due to Compton scattering.

unscattered after a distance $T$ along the LOR:

$$A_T = A_0 \exp[-\mu T] \qquad [2a]$$

If the attenuation medium is not uniform, then Equation [2a] is modified to give

$$A_T = A_0 \exp[-\sum_i \mu_i \,\Delta T_i] \qquad [2b]$$

where $\mu_i$ = the attenuation coefficient within a small element $\Delta T_i$.

Equation [2] describes what is known as 'narrow beam' attenuation because if one is just interested in the LOR between two detectors, a scattered photon is completely lost. For an array of detectors (as in a PET scanner), however, it can be imagined that the scattered photon might be detected in another LOR; this is considered later.

The larger the angle ($\Theta$) through which a scattered photon is deflected (**Figure 3**), the greater is its loss of energy. From simple kinematics (conservation of

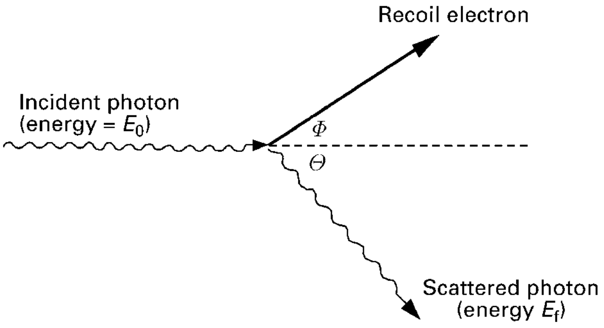



**Figure 3** Mechanism of photon (Compton) scattering.

energy and momentum), the relationship between the initial photon energy $E_0$ and final energy $E_f$ is

$$E_f = \frac{E_0}{1 + E_0(1 - \cos\Theta)/m_0c^2} \qquad [3a]$$

where $m_0c^2$ is the rest mass energy of the electron ($m_0$ = electron rest mass; $c$ = the speed of light). Since the rest mass energy is equal to 511 keV, this equation simplifies for PET to

$$E_f = \frac{E_0}{2 - \cos\Theta} \qquad [3b]$$

The maximum energy that can be transferred to an electron in this interaction is when the photon is deflected back along its original path ('backscattering'), that is when $\Theta = 180°$. In this case $E_f$ (minimum) = $E_0/3$ = 170 keV. Some other examples of scattering angles and energies are given in **Table 3**. However, Equation [3] only gives the resultant energy for a given angle; the probability of scattering through a particular angle is described by a much more complicated expression known as the Klein–Nishina formula. The relative probabilities of scattering are shown in the third column of **Table 3** (taking that for 0° as 100%). It can be seen that, even for a relatively large scattering angle, the photon still retains a significant fraction of its energy but that the probability of scattering falls quite quickly with increasing angle.

To obtain an accurate measurement of the regional distribution of isotope concentration in the body, the primary correction is that for attenuation. The basic principles of attenuation correction, relatively straightforward in PET, are as follows. Consider the attenuation lengths $a$ and $b$ on either side of the point source in **Figure 2B**. From Equation [2], attenuation along these two paths will be proportional to exp[$-\mu a$]

**Table 3** Examples of scattering angles and energies in Compton scatter of annihilation photons

| *Scattering angle* $\Theta$ (°) (*see* **Figure 3**) | *Energy of scattered photon* (keV) | *Probability of scatter* (%) (*0° = 100%*) |
|---|---|---|
| 30 | 451 | 68.7 |
| 60 | 341 | 31.5 |
| 90 | 256 | 18.8 |
| 180 (Back-scattering) | 170 | 18.5 |

and exp[$-\mu b$]. For coincidence counting, the total attenuation will be proportional to the product of these two factors: exp[$-\mu(a+b)$]. This expression is independent of the position of the source along the LOR and indeed would be so if the source were outside the object. If such an external source is measured with and without the (inactive) object in the LOR, the ratio of the measurements gives the attenuation correction factor along that line. For single-photon detection, it should be noted that attenuation is dependent on depth in the object and this makes correction more complicated. Specific methods of attenuation correction in PET are dealt with in the article on Instrumentation and Methods for PET.

## Detection of annihilation photons

Most PET scanners consist of individual detector elements arranged in a number of adjacent coaxial rings surrounding the patient. Each element is connected in coincidence with a number of other elements in both the same ring and any number of other rings, and modern scanners consist of thousands of detectors and millions of LORs. Specific scanner configurations are given in the article covering Instrumentation and Methods.

Detector materials must have a high atomic number to maximize their attenuation of annihilation photons and must also produce a measurable response. The great majority of detectors used in PET are scintillators, which respond to the absorption of photon energy by the emission of visible light. This occurs when electrons fall from excited energy levels to the ground state, a process often facilitated by the inclusion of a small amount of impurity (or activator) into the scintillation crystal. The light output rises rapidly to a peak and then falls with a characteristic 'decay time' and is converted into an electrical pulse by a photomultiplier tube (PMT) and subsequently amplifi.ed. The speed of response of the detector determines the width of this pulse and the precision with which the time of the interaction can be measured – the timing resolution, denoted by $\tau$. In

addition, the detector electronics take a finite time to process each event, during which other events go unrecorded. This time is known as dead time and corrections have to be made for accurate quantification. Based on the timing resolution, a coincidence time window is set within which events in two opposing detectors are regarded as constituting a coincidence event. This can be either a true event or a random (chance) event, depending on whether the photons came from the same or different annihilations. These collectively are known as prompt events (events are commonly termed 'prompts', 'trues' and 'randoms'). A distinction can be made between trues and randoms by counting the number of events occurring when the time window for one detector is delayed (by about 100 ns) relative to the other. A coincidence recorded in the 'delayed circuit' cannot be a true coincidence and it is assumed that the 'delayed events' are equal to the randoms recorded in the undelayed ('primary') circuit. Trues are therefore determined by subtracting the randoms from the prompts.

An alternative way of calculating random events is by recording the total rate of single photons striking each detector. If these singles rates for detectors 1 and 2 are $S_1$ and $S_2$, then the rate of random events ($R_{12}$) between the two detectors is given by

$$R_{12} = S_1 \cdot S_2 \cdot T \qquad [4]$$

where $T$, the coincidence time window, is twice the timing resolution (i.e $2\tau$). The distribution of randoms is quite uniform over the field-of-view (FOV) of the scanner but if not subtracted will impair the quantification of regional radiotracer concentration and reduce contrast in the image.

## Spatial sampling and resolution

One of the continuing aims in PET is to improve spatial resolution or the clarity of definition of isotope distribution in the body. The physical limit of resolution is dictated by positron range and non-collinearity of annihilation photons as outlined above. The ranges of positrons in the body for the most important isotopes are given in **Table 2**. These appear to be rather large, but it should be borne in mind that the contribution of positron range to resolution in the image is reflected in the average range and that positrons travel in all directions and not just orthogonally to an LOR. The net contribution to resolution (FWHM of the point spread function) is given in **Table 2**. The physical limitation imposed by the small angular spread around the 180° ('back-to-back') photon emission leads to an additional 'blurring' of resolution independent of positron energy but increasing with the diameter of the detector array. For a scanner of diameter 80 cm (common for imaging of humans) the fundamental limit imposed by these effects is about 2 mm, whereas for a diameter of 20 cm the limit is less than 1 mm.

To image the distribution of a radiotracer, the active volume must be sampled as finely as possible. In other words, measurements must be made along a number of LORs, as closely spaced as practicable, both across the object and at different angles. The basic property of a tomographic system with good resolution is the ability to distinguish changes in tracer concentration. Another way of expressing this is that the system has a good spatial frequency response. This may be envisaged analytically by presenting a 'bar pattern' of alternating white and black (active/inactive) stripes to the imaging system. As the width of the bars is reduced, there will come a point when the imager will no longer be able to reproduce the pattern; the input (object) frequency will no longer produce a faithful output (image). Mathematically this is expressed in terms of the modulation transfer function (MTF), which gives the fraction of signal amplitude that a system will transfer to the image at each spatial frequency. A broader MTF function will give a sharper resolution.

The requirement of an imaging system is expressed formally by the sampling theorem, which states that the highest frequency that can reliably be measured (known as the Nyquist frequency) is equal to $1/(2\Delta d)$, where $\Delta d$ is the sampling distance (the spacing between adjacent LORs). One of the principal aims in the design of a PET tomograph is to provide as high a degree of sampling as possible. However, a fundamental limit (apart from positron range and photon noncollinearity) is imposed by the detector width (the intrinsic detector resolution as defined above). Better sampling schemes will deliver an image resolution ever closer to the intrinsic resolution, but this cannot be exceeded without additional 'post-processing' of the image. Early tomographs of the late 1970s/early 1980s with relatively large (~25 mm) detectors, which were not densely packed, optimized their sampling by incremental linear and angular motion. Most tomographs in use today consist of circular rings of detectors of width 6 mm or less. For these, adequate linear and angular sampling is achieved without any motion, although in some designs a rotatory motion (known as 'wobble') is incorporated. This has largely been abandoned because the data volumes were increased typically by a factor of 4 without a great improvement in resolution.

The organization of data acquisition in a circular ring of detectors is shown in **Figure 4**. Each detector

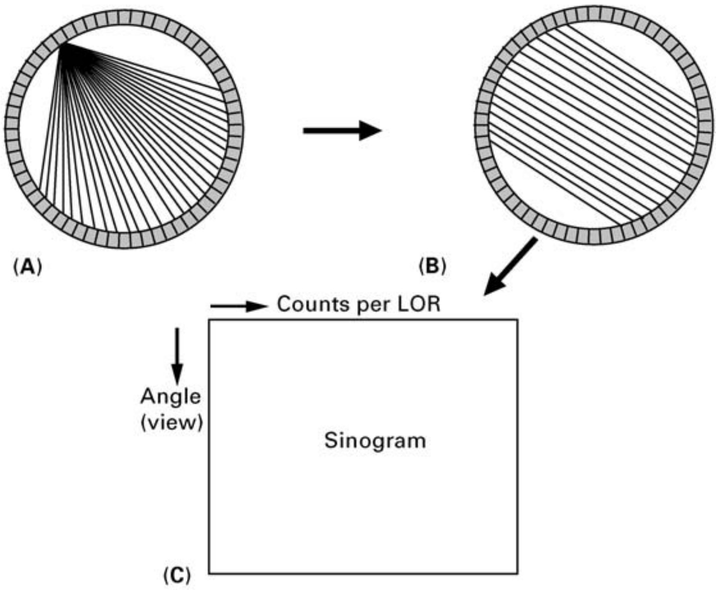



**Figure 4** Geometrical relation between LORs and the sinogram data matrix.

is connected in a coincidence circuit with a number of detectors on the other side of the ring (**Figure 4A**) (The more general case of multiple rings is discussed in the article on Instrumentation and Methods. The LORs that are parallel to each other are grouped together to form projections or views of the object at each angle (**Figure 4B**). The counts recorded in each of these LORs form one row of a data matrix called a sinogram (**Figure 4C**), each row corresponding to an angle of view. In a recent design there are 576 detectors in the ring and each detector is in coincidence with 288 other detectors. The number of views in the sinogram in this case is 576/2 = 288 (each separated by an angle 180°/288 = 0.625°) and the number of LORs in each view is 288.

Formation of an image can be achieved by a number of different methods, but all of them involve back-projection of the views across the FOV. For practical purposes the position of an annihilation event along an LOR is indeterminate. Some detectors, notably barium and caesium fluorides, are capable of modest time-of-flight resolution (FWHM ~5 cm) by measuring the time difference of the photon interactions, but this has not proved to give significant advantages because of their lower efficiency.) The raw process of image formation conventionally divides the 'image space' into a matrix of square boxes or pixels (picture elements) and places 'counts' in each pixel proportional to the counts recorded in the particular LOR and the area of overlap of LOR and pixel. This process is illustrated for a point source in **Figure 5A**. Only a small number of projections, each consisting of a peak corresponding to the source position, are shown for clarity. The back-projected profiles for each projection will intersect at the position of the point source, but the image will be a poor representation of the original because of the 'background' imposed. This may be described mathematically by saying that the high spatial frequency components of the source have been attenuated and low frequencies enhanced. This process is reversed by frequency filtering in which the Fourier transform is first applied to the projection data to give the magnitude of each component spatial frequency and a filter is applied with increasing weight given to higher frequencies. This filter is accordingly termed a ramp in frequency space. In real space, the form of the filter is as sketched in **Figure 5B**, which shows alternating positive and negative oscillations decreasing in intensity from its centre. When this is convolved with the projections and back-projection is performed, the effect is that the positive and negative components cancel each other out, so removing the low-frequency 'blur' and restoring the high-frequency nature of the object in the filtered image.

Filtered back-projection is the most commonly used method of image formation in PET because Fourier transforms can be calculated rapidly with modern computers. However, the method has its drawbacks and these are intimately allied to statistical variations or noise in the projection data. As the number of counts recorded by the system decreases, so the statistical uncertainty increases, as described by Poisson statistics, and 'star artefacts' (remnants of the back-projection process) become increasingly apparent. Filtering causes each pixel in the image to be

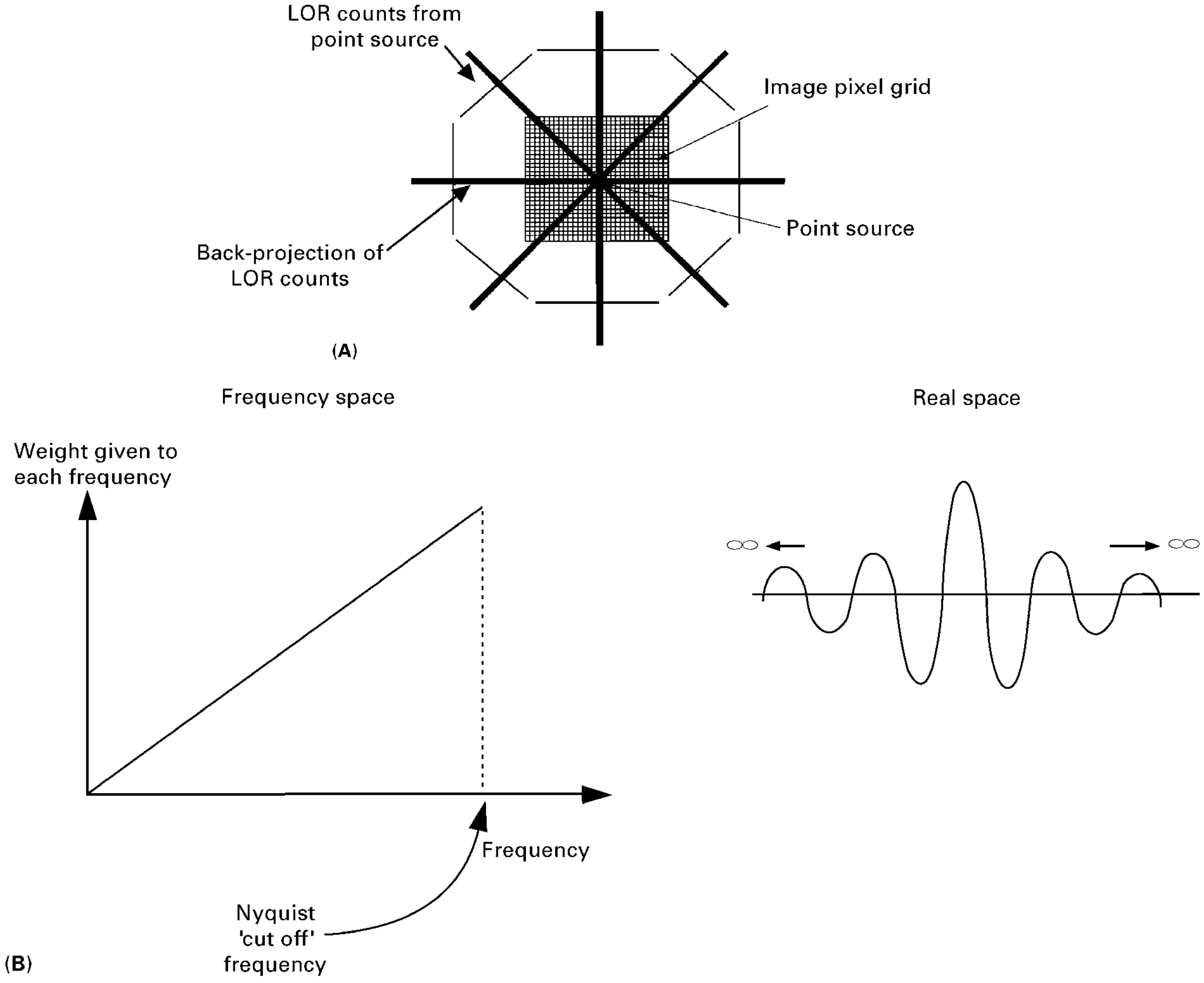



**Figure 5** (A) Formation of an image by filtered back-projection. (B) Shape of ramp filter in frequency space and real space.

correlated to some degree with every other pixel and poor statistics enhances this.

Inherently superior methods of image reconstruction are the so-called iterative methods in which the image is successively corrected to be consistent with the projection data to within desired error limits. These techniques involve iterative forward-projection and back-projection (between image and projections) and a number of different algorithms have been developed for the correction step at each iteration. Such methods have much more flexibility than filtered back-projection because models for the statistical nature of the data and the physical processes involved in data acquisition can be incorporated. For example if the projection data are assumed to obey Poisson statistics, the likelihood of the reconstructed image can be maximized according to this, giving the so-called EM-ML (expectation maximization–maximum likelihood) algorithm. In contrast to filtered back-projection, which produces a 'one-off' solution, iterative algorithms gradually converge to the desired solution but the number of iterations needs to be carefully assessed to optimize the signal-to-noise ratio in the image and regional quantification of tracer.

## Detection efficiency and noise

The efficiency of detection of annihilation photons obviously depends on the geometry of the tomograph and the ability of the detector material to absorb the photon energy (detector 'stopping power'). Considering a single ring of detectors of radius $r$ and width (axial thickness) $t$, decreasing the radius increases the solid angle of detection (with $r^2$) but the detector volume decreases with $r$. Therefore overall efficiency increases with $r$. However, as $r$ decreases, the solid angle for detection of random and scattered events increases, and clearly $r$ must be large enough to provide the desired FOV (for head or body). In addition, spatial resolution worsens in the radial direction as $r$ decreases. This implies that some compromise must be reached to balance these factors.

Although the 'raw' efficiency, or number of true coincidences acquired, is a basic determinant of the quality of a PET scanner, the complicating factors of scattered and random events and dead time have to be brought into the analysis. Details of the distribution of scattered events and correction methods can be found elsewhere, but for the present purposes it can be stated that scattered radiation (or 'scatter' for short) produces a relatively flat 'background' on the projection and image data, impairing contrast and reducing quantitative accuracy. The quantity of scatter detected, the scatter fraction (SF), is expressed simply in terms of the total true (unscattered + scattered) events ($T_{tot}$) and the scattered events ($S$) by

$$\mathrm{SF} = S/T_{\mathrm{tot}} \qquad [5]$$

A similar reduction in contrast and quantification would result if random events were not subtracted from the data. As discussed above, subtraction of randoms is usually carried out 'on-line' during data acquisition and, although accurate quantitatively, imposes a statistical penalty on the net counts. If the numbers of prompt and random events are designated by $P$ and $R$, then

$$T_{\mathrm{tot}} = P - R \qquad [6]$$

Assuming that Poisson statistics apply to these counts (e.g. the standard error of $P = \sqrt{P}$), the standard error of $T_{tot}$ is

$$\sqrt{\left[(\sqrt{P})^2 + (\sqrt{R})^2\right]} = \sqrt{(P + R)} = \sqrt{(T_{\mathrm{tot}} + 2R)} \qquad [7]$$

Thus the standard error of $T_{tot}$ is larger than $\sqrt{(T_{tot})}$ and increases with the rate of random events. Any procedure that reduces randoms, such as better shielding of extraneous radiation or reduction in the coincidence time window (Eqn [4]), will lower this uncertainty. The resultant standard error is conventionally described by the noise equivalent count (NEC) which is defined as follows. The fractional error of $T_{tot}$ is $\sqrt{(T_{tot} + 2R)}/T_{tot}$. If a number (NEC) of hypothetical counts are collected (free of background) the fractional error is $(\sqrt{\mathrm{NEC}})/\mathrm{NEC} = 1/\sqrt{\mathrm{NEC}}$. Equating this to the fractional error of $T_{tot}$ gives

$$\mathrm{NEC} = \frac{T_{\mathrm{tot}}^2}{T_{\mathrm{tot}} + 2R}$$

However, this expression is modified to take into account (a) the subtraction of scatter and (b) the fact that random events are spread fairly uniformly over the whole FOV of the PET tomograph, whereas unscattered true events are confined to the limits of the object. If the fraction of the FOV subtended by the object is $f$ then the final expression for NEC is

$$\mathrm{NEC} = \frac{[T_{\mathrm{tot}}(1 - \mathrm{SF})]^2}{T_{\mathrm{tot}} + 2fR} \qquad [8]$$

Another assumption implicit in this is that the subtraction of scattered events does not lead to an increase in noise. This would be so if, for example, a mathematical function were used to describe the distribution of scatter, but is not quite true in all methods.

NEC provides an overall factor for the determination of statistical quality, but it is of importance to investigate noise (statistical 'ripple') in the reconstructed image. As mentioned above in the discussion of reconstruction, image noise is greater than expected purely from the Poisson statistics of the projection data, owing to the filtering process. If there are an adequate number of projection angles and sampling points (LORs) per projection, then it can be shown that the noise/signal ratio (% root mean square) for a uniformly labelled object is given by

$$\mathrm{N/S}(\%) = \frac{120\{N_{\mathrm{re}}\}^{3/4}}{N^{1/2}} \qquad [9]$$

where $N$ is the total number of counts (trues) acquired and $N_{re}$ is the number of 'resolution elements' contained within the object. The resolution element is defined as a region of dimensions (sampling distance)$^2$ or $(\Delta d)^2$. If the object is a disc (or a slice through a cylinder) of diameter 200 mm, the resolution FWHM = 6mm and $10^6$ counts are acquired, N/S is 19%. This is about 6 times what would be expected purely from Poisson statistics. Furthermore, if the resolution decreased to 3 mm FWHM, the number of counts $N$ would have to increase by a factor of 8 to keep the N/S per resolution element constant. These simple calculations illustrate the importance of efficiency in PET. Resolution might be technically improved by using narrower detectors, but the advantage will be lost if efficiency of detection is not increased. This point is of fundamental importance and is discussed more fully in the article on Instrumentation and Methods.

## List of symbols

$A_T$ = number of photons left unscattered after distance $T$ along the LOR; $A_0$ = original number of photons; $c$ = speed of light; $E_0$, $E_f$ = initial and final energy of scattered photon; $f$ = fraction of FOV subtended by the object; $m_0$ = rest mass of the electron; $N$ = total number of counts (trues); $N_{re}$ = number of resolution elements; $N_x$ = number of unscattered photons at $x$; $P$ = number of prompt events; $r$ = detector radius; $R$ = number of random events, detector ring radius; $S$ = total number of scattered events; $t$ = detector width (axial thickness); $T$ = distance along LOR, coincidence time window; $T_{tot}$ = total number of true (unscattered + scattered) events; $\Delta d$ = sampling distance = spacing between adjacent LORs; $\Theta$ = scattering angle of photon; $\mu$ = attenuation coefficient; $\tau$ = timing resolution.

*See also:* **Fourier Transformation and Sampling Theory; PET, Methods and Instrumentation; Scattering Theory; Statistical Theory of Mass Spectra.**

## Further reading

Barret HH and Swindell W (1981) *Radiological Imaging: The Theory of Image Formation, Detection, and Processing*. San Diego: Academic Press.

Bendriem B and Townsend DW (eds) (1998) *The Theory and Practice of 3D PET*. Dordrecht: Kluwer Academic.

Herman GT (1980) *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*. New York: Academic Press.

Jones T (1996) The imaging science of positron emission tomography. *European Journal of Nuclear Medicine* **23**: 807–813.

Kinahan PE and Rogers JG (1989) Analytic 3D image reconstruction using all detected events. *IEEE Transactions on Nuclear Science* **NS-36**: 964–968.

Knoll GF (1979) *Radiation Detection and Measurement*. New York: Wiley.

Murray IPC, Ell PJ and Strauss HW (eds) (1994) *Nuclear Medicine in Clinical Diagnosis and Treatment*. Edinburgh: Churchill Livingstone.

Phelps M, Mazziotta J and Schelbert H (1986) *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart*. New York: Raven Press.

Shepp LA and Vardi V (1982) Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, **MI-1**: 113–122.

Webb S (ed) (1988) *The Physics of Medical Imaging*. Bristol: Institute of Physics Publishing.

# Pharmaceutical Applications of Atomic Spectroscopy

**Nancy S Lewen** and **Martha M Schenkenberger**,
Bristol-Myers Squibb, New Brunswick, NJ, USA

Copyright © 1999 Academic Press

**ATOMIC SPECTROMETRY**
**Applications**

## Introduction

The United States Federal Food and Drug Administration (FDA) regulations require the complete characterization of drug compounds. Since most pharmaceutical agents are organic compounds, much of this characterization involves various chromatography-based analytical techniques, as well as NMR, IR and various physical testing methods (such as DSC, TGA and XRD). The field of atomic spectroscopy has not traditionally played a major role in the characterization of pharmaceutical products, but on closer inspection it is clear that absorption and emission spectroscopic techniques can play a valuable role in the process of drug development and the quality of the product that finally reaches consumers.

From drug synthesis to quality control (QC) monitoring of over-the-counter medications, metals are found in all phases of the drug development process. Many metal-based products are used as imaging agents, and metals are used in the synthesis of drug substances, as excipients in tablets, capsules and liquids. In addition, trace metals can arise from the equipment used to manufacture a drug substance or compound. Because of the prevalence of metals associated with the drug development and manufacturing process, various atomic and emission-based techniques are often used to help fully characterize

pharmaceutical products. The wide variety of pharmaceutical dosage forms and matrices, such as tablets, capsules, injectables, liquids, effervescing compounds, ointments and creams, makes the development of analytical methods and the analysis of samples a challenging and interesting process. In this article we will describe the types of situations in the pharmaceutical industry where an analyst is likely to use atomic spectroscopy to solve the analytical problem and meet regulatory requirements. The pharmaceutical development process described will be based on the regulations and requirements in the USA.

## Techniques of interest in the analysis of pharmaceutical products

The need for the determination of metallic constituents or impurities in pharmaceutical products has, historically, been addressed by ion chromatographic methods or various wet-bench methods (e.g. the USP heavy metals test). As the popularity of atomic spectroscopy has increased, and the equipment has become more affordable, spectroscopy-based techniques have been routinely employed to solve analytical problems in the pharmaceutical industry. **Table 1** provides examples of metal determinations in pharmaceutical matrices, using spectroscopic techniques, and the reasons why these analyses are important. Flame atomic absorption spectrometry (FAAS), graphite furnace atomic absorption spectrometry (GFAAS), inductively coupled plasma-atomic emission spectroscopy (ICP-AES – also referred to as inductively coupled plasma-optical emission spectroscopy, or ICP-OES) and inductively coupled plasma-mass spectrometry (ICP-MS) are all routinely utilized in pharmaceutical applications. While there are other techniques of note available, such as microwave induced plasma (MIP) or direct coupled plasma (DCP), they have not been routinely used in the pharmaceutical industry, and will, therefore, not be discussed here. The theories involved in the use of FAAS, GFAAS, ICP and ICP-MS may be found in other articles of this Encyclopedia.

The first atomic spectroscopic techniques to see increased usage in the pharmaceutical field were FAAS and GFAAS. Among the current instrumental techniques available, they are among the most inexpensive, and have seen considerably more usage in all fields of endeavour, thus availing the pharmaceutical analyst of a vast array of knowledge upon which to draw and develop analytical methods. Because of the relatively low cost of the instrumentation, as well as its ease of use, QC laboratories in the pharmaceutical industry are more likely to have this type of atomic spectroscopy equipment than any other type. The speed and sensitivity of FAAS for elements such as Na, K and Li make it superior to wet-bench techniques. Examples of pharmaceutical products which require Na, K or Li determinations are nafcillin sodium (an antibiotic), oral solutions of potassium chloride (an electrolyte replenisher) and Lithane® (a psychotropic drug).

**Table 1** Examples of metals that are determined in pharmaceutical analyses

| *Element* | *Reason for assay/therapeutic area of use* | *Suggested analytical technique* |
|---|---|---|
| Ag | 1. Determination of geographical origin of illicit drugs | 1. Graphite furnace AA |
| | 2. Complex formation with drug for indirect determination (e.g. tetracycline) | 2. ICP-AES |
| | 3. Monitor Ag content of material (e.g. antiseptic creams, ophthalmic solutions) | 3. ICP-AES, graphite furnace AA |
| Al | 1. Monitor Al in antihaemophilia preparations, which are sometimes precipitated with aluminium hydroxide. | 1. Graphite furnace AA |
| | 2. Determine geographical origin of illicit drugs | 2. Graphite furnace AA |
| | 3. Monitor Al concentratiosns in dialysis solutions. | 3. ICP-AES |
| Au | 1. Monitor Au concentration in arthritis drugs | 1. ICP-AES |
| B | 1. Monitor for presence of B in regrents used in synthesis | 1. ICP-AES, ICP-MS |
| | 2. Monitor for leaching of B from glass vials, containers | 2. ICP-AES, ICP-MS |
| Ba | 1. Determination of geographical origin of material (e.g. illicit drugs) | 1. ICP-MS |
| | 2. Monitor Ba content in materials (e.g. used for diagnostic imaging) | 2. Flame AA, ICP-AES |
| Bi | 1. Indirect determination of cocaine | 1. ICP-AES |
| | 2. Monitor Bi content in materials (e.g. antacid products) | 2. ICP-AES |
| Br | 1. Determination of geographical origin of material (e.g. illicit drugs) | 1. ICP-MS |
| | 2. Monitor for presence of reagents used in synthesis, or as part of the compound | 2. ICP-MS |
| Ca | 1. Determination of Ca in calcium supplements and vitamins | 1. ICP-AES, graphite furnace AA |
| | 2. Monitor Ca impurities in magnesium oxide (often used as an excipient in pharmaceutical preparations) | 2. Flame AA |
| | 3. Determination of geographical origin of material (e.g. illicit drugs) | 3. ICP-MS |

**Table 1** *Contd.*

| *Element* | *Reason for assay/therapeutic area of use* | *Suggested analytical technique* |
|---|---|---|
| Cd | 1. Monitor Cd in dialysis solutions | 1. Graphite furnace AA |
| | 2. Monitor heavy metals content of medicinal plants or herbal drugs | 2. Graphite furnace AA, flame AA |
| | 3. Determination of geographical origin of material (e.g. illicit drugs) | 3. ICP-MS. |
| | 4. Monitor trace metals content in materials (e.g. penicillin G) | 4. ICP-AES |
| Co | 1. Complexing agent for indirect determination of drug (e.g. salicylic acid, lidocaine) | 1. Flame AA |
| | 2. Monitor Co in dialysis solutions | 2. Graphite furnace AA |
| | 3. Monitoring trace metals content in materials (e.g. penicillin G) | 3. ICP-AES |
| | 4. Determination of B-vitamins | 4. HPLC-FAAS |
| Cr | 1. Complexing agent for indirect determination of drug (e.g. thioridazine, amitriptyline, imipramine, orphenadrine) | 1. Flame AA |
| | 2. Determine geographic area of origin of illicit drugs | 2. Graphite furnace AA |
| | 3. Monitor trace metals content in materials (e.g. penicillin G) | 3. ICP-AES |
| | 4. Monitor Cr content in vitamins | 4. Graphite furnace AA |
| Cs | 1. Monitor for presence of reagents used in synthesis | 1. Flame AA |
| Cu | 1. Complexing agent for indirect determination of drug (e.g. lincomycin, isonicotinic acid hydrazid, ethambutol hydrochloride, neomycin, streptomycin) | 1. Flame AA |
| | 2. Monitor Cu in dialysis solutions | 2. Flame AA, graphite furnace AA |
| | 3. Moinitor heavy metals content in medicinal plants | 3. Graphite furnace AA |
| | 4. Determine Cu concentrations in vitamins | 4. ICP-AES |
| | 5. Monitor Cu in herbal drugs | 5. Flame AA |
| | 6. Monitor trace metals content in materials (e.g. penicillin G) | 6. ICP-AES |
| | 7. Determination of synthetic route and geographical origin of material (e.g. illicit drugs or to prevent patent infringement) | 7. ICP-MS |
| Fe | 1. Monitor Fe in dialysis solutions | 1. Flame AA |
| | 2. Monitor Fe contamination in magnesium oxide (often used as excipient in pharmaceutical preparations) | 2. Flame AA |
| | 3. Monitor Fe concentrations in vitamins | 3. ICP-AES |
| | 4. Determination of geographical origin of illicit drugs | 4. ICP-MS |
| | 5. Determination of trace metals content of materials (e.g. penicillin G) | 5. ICP-AES |
| | 6. Monitor Fe concentrations in imaging agents. | 6. ICP-AES, flame AA |
| Gd | 1. Monitor Gd content in imaging agents (e.g. Prohance®) | 1. ICP-AES |
| Hg | 1. Monitor Hg content of materials (e.g. antiseptic solutions and creams, ophthalmic solutions) | 1. Flame AA |
| | 2. Monitor heavy metals content of materials | 2. ICP-MS |
| I | 1. Determination of geographical origin of illicit drugs | 1. ICP-MS |
| In | 1. Monitor trace metals concentration in final drug substance | 1. ICP-MS |
| K | 1. Monitor for presence of reagents used in synthesis | 1. Flame AA |
| | 2. Monitor salt counter-ion concentration | 2. Flame AA, ICP |
| Li | 1. Monitor for presence of reagents used in synthesis | 1. Flame AA, ICP-MS |
| | 2. Monitor Li concentration in drug (e.g. lithium-based psychotropic drugs for treatment of manic/depressive disorder) | 2. Flame AA |
| Mg | 1. Determine Mg concentrations in vitamins | 1. ICP-AES |
| | 2. Determination of geographical origin of illicit drugs | 2. ICP-MS |
| | 3. Monitoring magnesium stearate content or magnesium oxide (used as lubricant, sorbent, respectively, in pharmaceuticals) | 3. ICP-AES, flame AA |
| Mn | 1. Detemination of geographical origin of illicit drugs. | 1. Graphite furnace AA |
| | 2. Monitor trace metals content of materials (e.g. penicillin G) | 2. ICP-AES, ICP-MS |
| Mo | 1. Monitor trace metals content of materials (e.g. penicillin G) | 1. ICP-AES |
| Na | 1. Determination of synthetic route and geographical origin of material (e.g. to prevent patent infringement; illicit drugs) | 1. Flame AA, ICP-MS |
| | 2. Monitor salt counter-ion concentration of salt content (e.g. in diagnostic agents, in electrolyte replenishing solutions, in cathartics) | 2. Flame AA, ICP-AES |
| Ni | 1. Monitor Ni in dialysis solutions | 1. Graphite furnace AA |
| | 2. Determination of geographical origin of illicit drugs | 2. Graphite furnace AA |

**Table 1** *Contd.*

| *Element* | *Reason for assay/therapeutic area of use* | *Suggested analytical technique* |
|---|---|---|
| | 3. Monitor trace metals content of materials (e.g. penicillin G) | 3. ICP-AES, ICP-MS, graphite furnace AA |
| P | 1. Determine P concentration of vitamins | 1. ICP-AES |
| | 2. Determination of geographical origin of illicit drugs | 2. ICP-MS |
| | 3. Determination of constituents of materials (e.g. alendronate sodium) | 3. ICP-AES |
| Pb | 1. Determination of Pb in calcium supplements | 1. ICP-AES, graphite furnace AA |
| | 2. Monitor Pb in dialysis solutions | 2. Flame AA |
| | 3. Monitor heavy metals content in medicinal plants. | 3. Graphite furnace AA |
| | 4. Determination of geographical origin of illicit drugs | 4. ICP-MS |
| | 5. Monitor trace metals content of materials (e.g. penicillin G) | 5. ICP-AES |
| Pd | 1. Determination of residual catalyst in pharmaceuticals (e.g. fosinopril, semisynthetic penicillin) | 1. ICP-MS, graphite furnace AA |
| | 2. Determination of geographical origin of illicit drugs | 2. ICP-MS |
| Pt | 1. Speciation of Pt-containing compounds (cisplatin, transplatin, carboplatin, JM-216) | 1. HPLC-ICP-MS, graphite furnace AA |
| | 2. Monitor for residual catalysts | 2. ICP-MS, graphite furnace AA |
| Rh | 1. Monitor for residual catalysts used in synthesis of pharmaceuticals | 1. ICP-MS |
| Sb | 1. Determination of synthetic route or geographical origin of material (e.g. for illicit drugs, or to prevent patent infringement) | 1. ICP-MS |
| | 2. Monitor Sb content in materials (e.g. final drug substances) | 2. ICP-MS |
| Se | 1. Monitor for presence of reagents used in synthesis | 1. Graphite furnace AA, ICP-MS |
| | 2. Monitor Se concentration in vitamins | 2. Graphite furnace AA, ICP-MS |
| | 3. Monitor Se concentration in anti-fungal and anti-seborrhoeic products | 3. Graphite furnace AA, ICP-MS |
| Si | 1. Determination of geographical origin of illicit drugs | 1. ICP-MS |
| | 2. Monitor for Si contamination from silicone-based compounds used in packaging processes | 2. ICP-MS, ICP-AES |
| | 3. Monitor for silica gel (used to prevent caking or as a suspending agent) | 3. ICP-AES |
| Sn | 1. Monitor for presence of reagents used in synthesis | 1. Graphite furnace AA, ICP-MS |
| | 2. Monitor heavy metals content of materials | 2. Graphite furnace AA, ICP-MS |
| Sr | 1. Determination of geographical origin of illicit drugs | 1. Graphite furnace AA, ICP-MS |
| Ti | 1. Determine Ti concentration in sunscreens (titanium dioxide is often used in sunscreens) | 1. Flame AA |
| | 2. Monitor trace metals content of materials (e.g. penicillin G) | 2. ICP-AES |
| Zn | 1. Monitor heavy metals content in medicinal plants | 1. Graphite furnace AA |
| | 2. Determine Zn concentration in vitamins | 2. ICP-AES |
| | 3. Determination of geographical origin and synthetic route of material (e.g. illicit drugs or to prevent patent infringement) | 3. ICP-MS |
| | 4. Monitor trace metals of content of materials (e.g. penicillin G) | 4. ICP-AES |
| | 5. Monitor Zn content of materials (e.g. insulin, antibiotics, sunscreens) | 5. ICP-AES, flame AA |

The speed of FAAS is, undeniably, a tremendous asset of the technique. Sample analysis times of less than 1 min per sample enables the analyst to process numerous samples in a given day by FAAS. FAAS is particularly useful when analysing a trace level analyte in the presence of another metal whose concentration is very high. This situation is encountered when analysing products which incorporate a metal into the drug substance, such as Platinol® (an oncology agent), Prohance® (an imaging agent) and Myochrysine® (a product used for the treatment of rheumatoid arthritis). These products contain high concentrations of platinum, gadolinium and gold, respectively. These elements have very rich spectra, with numerous spectral lines, which may overlap with the spectral lines of the analyte elements. In such cases, FAAS is a better choice for trace metals determinations than ICP-AES, since coincident line overlap is not a problem with the former technique, but presents a considerable problem for the latter.

GFAAS is also commonly found in pharmaceutical company laboratories, owing to the affordability of this spectroscopic instrumentation. GFAAS is ideally suited for the analysis of samples which are available only in small quantity, because it requires considerably less sample for a given analysis than FAAS or any of the plasma-based techniques (e.g. 20 μL per determination, versus 3 mL per determination). Additionally, GFAAS has the ability to remove the sample matrix before atomization of the sample for analyte

determination, thus affording the analyst great versatility in the analysis of samples which are composed of rich organic matrices. As with FAAS, GFAAS is also well suited to those pharmaceutical applications where a low concentration analyte is determined in the presence of a high concentration metal.

Second in popularity to atomic absorption based techniques for applications in the pharmaceutical industry is ICP-AES. This instrumentation affords the analyst greater flexibility, with a wider dynamic range and a broader range of elements which can be analysed in a single run. It is often employed to simultaneously determine metals such as Al, Cr, Fe, Mn, Ni, Zn, P, B, Pd and Pt in pharmaceutical matrices. Though FAAS and GFAAS may also be used to monitor these elements, ICP-AES can scan all of these elements in a single analysis (either by scanning, or by the use of a simultaneous unit). In addition, ICP-AES can monitor multiple wavelengths for each element for confirmation of its presence, making it an attractive alternative to either FAAS or GFAAS. The wide linear range of ICP-AES is quite useful in the analysis of pharmaceutical samples, owing to the time saved in developing methods. In addition, the need to make multiple dilutions of a sample is eliminated, as is the need to run multiple standard concentrations within an analysis.

Depending on the stage of development of a pharmaceutical product, the decision for selecting ICP-AES over an AAS technique may simply be the amount of sample available for a set of analyses. The advent of axial ICP-AES systems, with their increased sensitivity, makes ICP-AES an excellent choice where large amounts of sample are not available. The axial ICP-AES system allows the analyst to use considerably less sample than in the past, while achieving the same detection limits and minimum quantifiable limits.

Owing to its ability to monitor multiple wavelengths for a given analyte, and its wide linear range, ICP-AES is well suited for identity testing. An identity test is one in which the analyst is only confirming or denying the presence of a given analyte. In some cases, a compound may have a sufficiently high concentration of a given metal, making it possible to monitor the metal to determine if the compound is authentic. Monitoring of multiple wavelengths is often used to positively confirm the identity of the analyte metal, thus fulfilling the needs of an identity test.

ICP is seeing more use as a sample introduction system for various hyphenated techniques. New to the pharmaceutical industry is the use of inductively coupled plasma-mass spectrometry (ICP-MS). ICP-MS offers excellent versatility and sensitivity to the analyst, and greatly complements any pharmaceutical atomic spectroscopy laboratory. The sensitivity of the technique and its scanning capabilities make it an ideal choice for the analysis of pharmaceuticals in the early stages of development, when sample material may be in extremely short supply, as the chemists try to optimize and change the synthesis. ICP-MS has been used in our laboratories as an alternative to the United States Pharmacopeia (USP) heavy metals test, providing more accurate, element-specific results for several very toxic metals. The USP test requires a minimum of 1 g of material to perform the nonspecific sulfate ashing procedure. In comparison, the ICP-MS procedure requires only 25 mg and provides element-specific information on 14 different metals. Additionally, ICP-MS is able to examine different isotopes of a given metal present in a sample. This can be quite useful when studying imaging agents, which may be formulated with radioisotopes as part of the desired active ingredient. ICP-MS is very useful in the analysis of trace metals in a matrix containing another metal at high concentrations. As noted before, coincident lines may cause problems for ICP-AES determinations in these cases, and the sensitivity or need for individual lamps may slow or preclude the use of FAAS or GFAAS for these determinations as well. With ICP-MS, the high concentration metal may be ‘skipped’, while several analyte metals present at trace concentrations may be examined in a single analysis.

All of the techniques discussed have been used to analyse pharmaceutical products which have no chromophores and cannot be analysed by traditional UV-based chromatographic systems. In these cases, metallic complexes are formed with the compounds of interest and then indirectly determined by FAAS, GFAAS, ICP-AES or ICP-MS analysis. This approach can provide valuable information in a short time, one of the chief advantages of spectroscopic techniques when compared with a chromatographic technique, which may take several minutes to an hour per sample analysis. In addition to the situations where a pharmaceutical product is complexed with a metal before analysis by these techniques, FAAS, ICP-AES and ICP-MS are also used in concert with various chromatographic techniques, such as LC-ICP-AES, LC-ICP-MS, IC-ICP-AES, IC-ICP-MS, LC-FAAS and IC-FAAS. The coupling of chromatographic systems with FAAS, ICP-AES and ICP-MS instruments has provided the pharmaceutical analyst with tools which can be used to speciate metallic constituents in drug products, to achieve even lower detection limits, and to examine the different isotopes of metallic constituents present in a sample. Indeed, the sensitivity, flexibility and speed of each of these techniques prove to be valuable in the pharmaceutical industry.

The plasma-based techniques can also serve as detectors for laser ablation (LA) and electrothermal vaporization (ETV). These techniques are well-suited for the analysis of solid samples. ETV can also be used to analyse liquid and slurry samples. Both techniques use small quantities of material and, when interfaced with ICP-MS, are quite sensitive. A cool plasma accessory can also be interfaced with the ICP-MS. This allows for the removal or minimization of interferences caused by the formation of molecular species in the plasma, permitting the determination of Li, Na, Ca, K, Fe and Cr which can not be analysed successfully by conventional ICP-MS. Such analyses exhibit the same sensitivity as afforded by FAAS.

## How is an analytical technique selected in a pharmaceutical laboratory?

### The stages of the drug development process – some background

The role that atomic spectroscopy plays in the pharmaceutical industry may be directly linked to the various stages of the drug development process. To understand how these techniques might be encountered it is important to examine, in closer detail, what happens during each step of the drug development process.

From the time a potential new drug candidate is identified to the time that it reaches the market it undergoes considerable testing and evaluation. It is imperative that the testing and evaluation of a new drug candidate be completed as quickly as possible, since the pharmaceutical company's patent on a drug has a finite life. The patent gives the pharmaceutical company exclusive rights to the production and sale of the drug once it is approved by the FDA. Once the drug goes off patent, other pharmaceutical companies are allowed to produce a generic form of the drug. The sale of these generic forms can have a great impact on the sales of the originator company product. One of the goals of drug development is to maximize the length of time in which the company has exclusive marketing rights. It is not uncommon for the sales of a major pharmaceutical product to exceed $1 billion a year at the time the exclusivity period expires. This translates to ~$100 million per month or more. Thus, each month the company can reduce from the development cycle can literally be worth hundreds of millions of dollars.

SmithKline-Beecham's product, Tagamet®, illustrates this well. The earnings from the sale of Tagamet® were £484 million in 1994 (last year of exclusivity). The earnings in 1995, the first full year Tagamet® was off patent, were £286 million, a drop of almost 41%. Therefore, it is beneficial to the company to reduce the time and expense required to get a drug product through the discovery and development phases to market. The steps and the goals of each phase in the development process are quite specific and well defined by the FDA, and are summarized in **Table 2**.

### Preclinical (discovery) testing

The earliest stage of drug development is the preclinical or discovery phase. During this phase, a potential drug candidate is identified, and work begins on developing an optimal synthesis. Preliminary assessments are made regarding the safety and biological activity of the potential drug candidate in laboratory and animal studies. At this stage of development, the synthetic chemist has very little experience with the molecule and may utilize exotic catalysts to produce the first few grams of the materials. Analyses of preliminary batches of the drug candidate and key intermediates are performed to ensure the preliminary safety data is reflective of the drug candidate and not impurities generated by the synthetic process. Since

**Table 2** The phases of the drug development process

| *Step in the drug development process* | *Goal/objective of the step in the process* |
|---|---|
| Preclinical (discovery) phase | Assess the drug's safety and biological activity in the laboratory and animal studies |
| Phase I clinical trials (IND phase ) | Establish the safety and bioavailability of the compound in humans. This is typically done in studies using healthy volunteers |
| Phase II clinical trials (IND phase) | Determine proof-of-principle for the drug's mode of action. Monitor for any possible side effects and evaluate the drug's effectiveness. This study uses patient volunteers who have the disease / condition for which the drug is targeted |
| Phase III clinical trials (IND phase) | Establish dose form and dosage strength for registrational filing. Continue to monitor possible side effects and adverse reactions. Verify the effectiveness of the drug in the targeted patient population. This study involves many more patient volunteers than the Phase II study |
| FDA review/approval process | The FDA reviews the new drug application (NDA). If it is fully approved, the drug may proceed to market. If the NDA is not fully approved, the FDA may require additional testing, answers to questions or it may reject the application |

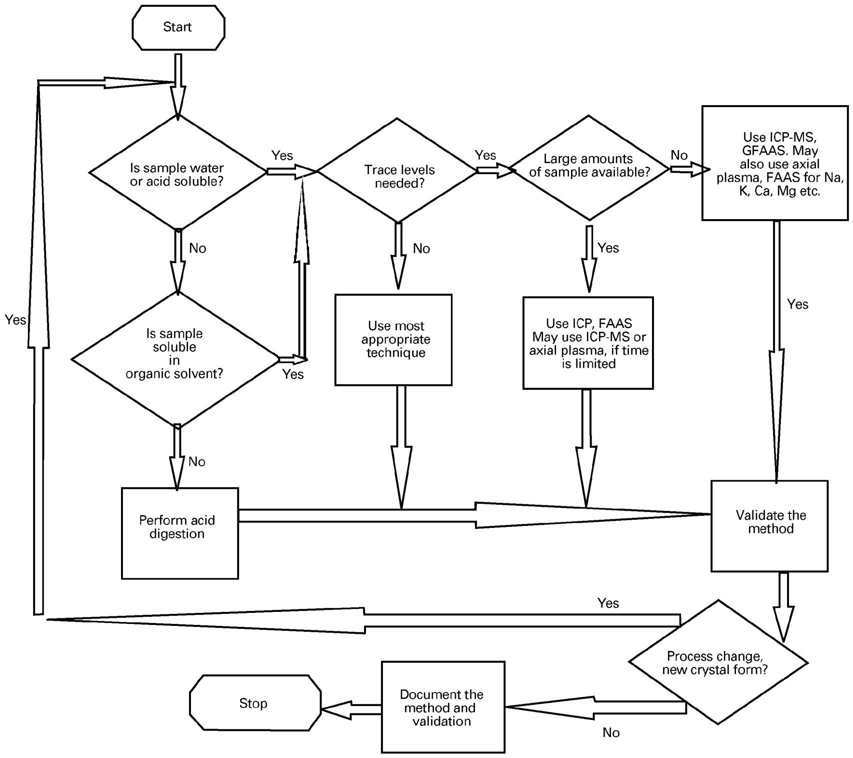



**Figure 1** Flow chart of method development decision-making process. Reproduced with permission of the editor from *Atomic Spectroscopy: Pharmaceutical Applications of Atomic Spectroscopy* **12**(9): 14–23 (1997) published by Advanstar Communications.

patents have finite lifetimes, time is of the essence, especially at this early stage, when the drug has not yet been evaluated in man. During this phase of development the salt and/or crystal form of the drug substance may not have been selected. As a result, samples analysed can vary in solubility properties, pH, purity, etc. Analyses that are typically performed include counterion, trace metals (from equipment sources, e.g. stainless steel) and trace catalyst determinations.

The flow chart given in **Figure 1** illustrates the thought processes involved in the selection of a given analytical technique for the determination of metals in pharmaceutical related samples.

The atomic spectroscopist is typically involved in supporting a new potential drug candidate before the final salt and/or crystal form have been selected. Several forms of the drug substance (salt forms, and/or polymorphs) are considered during the discovery phase and are generated in small laboratory batches via several synthetic pathways or crystallization procedures. The atomic spectroscopy laboratory plays an important role in the selection of the final form by assaying these samples for trace metals, salt counter-ions and trace catalysts used in the syntheses. Once a final form has been selected, testing continues to support the optimization of the synthetic process.

The selection of an appropriate analytical technique is highly dependent upon the time constraints that pharmaceutical companies set for the complete development of a drug product. As the costs of developing drugs has risen, the push within the pharmaceutical industry has been to reduce the time from discovery to the clinical studies as much as possible. This is driven by the fact that somewhere between one in seven and one in ten potential drug candidates in development actually make it to market. Therefore, results on preclinical samples are usually required in a short time (a few days, or hours) so that refinements to the synthetic process can be made, if necessary. The technique that is chosen for the assay must be rapid, meet the sensitivity requirements and generally consume small quantities of material. To expedite the analysis, it is prudent to perform as many determinations as possible in one assay. ICP-AES and ICP-MS are well suited for this type of determination; however, they are not ideal for all trace metals. The elements Na and K must be analysed by FAAS unless they are present at concentrations high enough for ICP-AES.

Once the final form has been selected and the synthesis refined, methods are developed and validated for metals that are present in reagents used in the synthesis, metals that may arise from the equipment used in the synthesis and metals that are incorporated into

the active ingredient in the final drug product. Validated methods are often required for synthetic intermediates as well. In addition, the FDA requires that a method be validated for the determination of heavy metals (i.e. lead, mercury, etc.) in the final drug substance. The USP heavy metals test requires one gram of sample for each determination. This method is non-specific and is based on a sulfate ashing of the sample, followed by a colorimetric comparison with a lead solution standard. Since this much material is usually not available in this phase of drug development, ICP-MS has been demonstrated to be an excellent technique for the determination of heavy metals in early development drug candidates. ICP-MS offers element-specific information and utilizes substantially less sample.

In addition to support of the drug substance, analyses are performed on starting materials and any raw materials used in the last step of the synthesis. Small batches of the drug substance which will be used in animal toxicology and pharmacokinetic studies must also be analysed. Once a synthetic process and final form have been selected, an investigational new drug (IND) application is filed with the FDA.

### Clinical (IND) phase

The IND contains information regarding the drug's composition and synthesis and lists all specifications that have been set for the drug substance. Specifications are set for many tests, which may include trace metals. All subsequent batches of the drug that will be used in clinical studies must meet these specifications before their release. The IND contains information regarding animal toxicology study data and protocols for clinical trails. The IND clinical study protocol for a new drug candidate consists of three clinical phases (**Table 2**).

Optimization and refinement of the synthetic process continues during the IND phase. The synthetic chemists scale up the synthesis to produce kilogram size batches. Support of this stage of drug development is similar to that performed to support synthesis optimization on small laboratory batches during the preclinical phase. Validated methods must be refined as the synthesis is refined, because even the slightest change in the synthesis can have a profound effect on whether a previously validated method will continue to be adequate for the trace metal determination. The use of a different solvent or reagent can be sufficient to invalidate a method.

During the preclinical stage of drug development, speed and sample consumption are typically the most important factors when selecting a technique; however, as the compound moves through the clinical phase of development there are other factors to consider. First, the atomic spectroscopist must consider the analytes of interest and the sensitivity that is required. Speed is still an important issue; however, sample consumption is less of a critical factor, since batch sizes of several hundred grams to several kilograms are routinely being produced. One must consider whether the method will be transferred to a QC laboratory since the instrumentation within their laboratory will often dictate which spectroscopic technique is used. If a QC laboratory will be performing the analysis, then usually either FAAS or GFAAS will be preferred since this instrumentation is typically found in QC laboratories, owing to the lower cost, compared with plasma-based instrumentation. This poses a challenge when one requires the sensitivity of GFAAS but must dissolve the drug substance in an organic solvent that is too viscous for GFAAS systems to handle. ICP-AES or ICP-MS would be the ideal alternative, but, most QC laboratories cannot afford such instrumentation.

Once the development of the drug candidate passes into the clinical Phase II and III studies, the demand for bulk substance increases. The synthesis is scaled up in the pilot plant to batch sizes ranging from ten to several hundred kilograms, and eventually to final production size batches of the final drug substance. The atomic spectroscopist will sometimes be called upon to help troubleshoot the process during the scale-up. Troubleshooting samples come in a variety of forms: discoloured drug substance or intermediate; scrapings from the equipment used in the synthesis; reagents used in the synthesis; filters used in the synthesis; liquid streams from the processing or slurries that were produced owing to a malfunctioning of the equipment. Sometimes the chemist will have an idea as to why the process failed and can help narrow down the investigation for the analyst. The cause of a process excursion can range from the use of a reagent contaminated with metals to equipment failure, such as a lubrication oil or coolant leak or the corrosion of the stainless steel equipment by the reaction by-products. ICP-MS is an excellent tool for assessing the problem quickly by performing qualitative or semi-quantitative scans of the periodic table. If these scans indicate that any metals are present at concentrations high enough for concern (several parts per million), alternative techniques, such as ICP-AES or FAAS, are used to confirm and quantitate their presence in the sample. Usually, but not always, sample consumption is not of great concern, but the speed of the technique is critical, since the chemist cannot proceed with the processing of the batch(es) until the source of the problem is identified. In analysing oils, contaminated filters, discoloured drug substance or intermediates,

the analyst will often focus on the possible presence of wear metals from lubricating oils (Al, Cr, Cu, Fe, Pb, Sn and Mo), metals from coolant contamination (Na, K and B) and metals found in stainless steel (Ni, Cd, Pb, Al, Fe, Cr, Cu, Mn and Zn). ICP-AES is sensitive enough for most of these metals and, because it is capable of multielement analyses, it is also rapid enough to satisfy the short turn-around-times required for processing these samples. Na and K must be assayed by FAAS unless they are present at high enough concentrations for ICP-AES.

LA-ICP-MS can be used to quickly analyse solid samples, such as filters or scrapings from the equipment. A minute amount of sample is consumed and analysis is fast, as no sample preparation is required. LA-ICP-MS is especially useful when analysing solid samples with distinct discolourations, since the laser can be focused on the area of interest to increase sensitivity. Each discoloured area can be ablated and assayed separately to determine its metallic composition. Often, these areas are caused by contamination from an oil or coolant that has leaked from the equipment. This affords the spectroscopist great selectivity over a conventional dilute-and-shoot method in which the small discoloured areas cannot be analysed separately.

Before the introduction of cool plasma ICP-MS, qualitative ICP-MS scans did not provide accurate information on Li, Na, K, Ca, Cr and Fe (owing to spectral interferences or the element being easily ionized). Therefore, ICP-AES or FAAS assays were required for accurate information on these elements. The advent of cool plasma ICP-MS makes it possible to quickly analyse all metals using only one spectroscopic technique, which is important when only a small amount of sample is available.

### NDA phase (Phase IV)

In the final stage of drug development, a new drug application (NDA) is filed with the FDA. Animal and clinical studies continue during the NDA phase. Stability tests of the drug substance and product continue, including studies of the commercial formulation in the market packaging. The spectroscopy laboratory supports this stage of development by performing analyses that are included in the specifications that have been set for the drug substance and product, using the methods filed with the NDA. The spectroscopist may see samples during this phase that are generated when the process of the drug substance or product is transferred to a new production facility. The steps that are taken in selecting and using a spectroscopic technique are the same as those described in the previous section.

**Table 3** Examples of pharmaceutical compounds which contain metals

| *Metal* | *Examples of uses in pharmaceutical compounds* |
|---|---|
| Na, Mg, Ca, K | Used in various excipients, in vitamins, in dialysis solutions and Eye Stream®, a liquid used for irrigating eyes, contains Na, Mg and K |
| Pt | Used in several oncology drugs: Paraplatin®, Platinol® and Cisplatin® |
| Zn | Used in Insulin and in Cortisporin® ointment (a steroid–antibiotic ointment) |
| Li | Used in Lithobid® and Cibalith-S®, both of which are used for the treatment of manic-depressive psychosis |
| Al | Often used in antacid preparations, such as AlternaGel™ or Mylanta® |
| Ag | Used as a topical antimicrobial for the treatment of burns in Silvadene® cream, 1% |
| Au and Na | Used in Myochrysine® injection, for the treatment of rheumatoid arthritis |
| Fe | Active ingredient in Chromagen®, a drug used in the treatment of anaemia. Also, sometimes used in pigments for printing tablets or capsules |
| Se | Selsun Blue®, a dandruff shampoo |
| Mn | Active ingredient in LumenHance®, an imaging agent |

The last stage in this phase consists of FDA inspections and a review of the NDA. During these inspections, the auditors may examine instrument calibration records and previous batch results to ensure that they were collected under Good Laboratory Practices (GLP) and/or Good Manufacturing Practices (GMP). Based on the inspection, the FDA will either approve the new drug candidate, require additional testing, request answers to questions and concerns they have or reject the drug product. Some examples of pharmaceutical compounds that contain metals are given in **Table 3**.

The role that the atomic spectroscopy laboratory plays in the drug development process is an important one. It helps ensures the safety and quality of the drug products that are approved by the FDA.

*See also:* **Atomic Absorption, Methods and Instrumentation**; **Atomic Absorption, Theory**; **Atomic Emission, Methods and Instrumentation**; **Biomedical Applications of Atomic Spectroscopy**; **Forensic Science, Applications of Atomic Spectroscopy**; **Hyphenated Techniques, Applications of in Mass Spectrometry**; **Inductively Coupled Plasma Mass Spectrometry, Methods**; **Inorganic Chemistry, Applications of Mass Spectrometry**.

## Further reading

Ali SL (1983) Atomic absorption spectrometry in pharmaceutical analysis. *Journal of Pharmaceutical and Biomedical Analysis* **1**: 517–523.

Lewen N, Schenkenberger M, Larkin T, Conder S and Brittain H (1995) The determination of palladium in

Fosinopril sodium (Monopril) by ICP-MS. *Journal of Pharmaceutical and Biomedical Analysis* **13**: 879–883.

Lewen N, Schenkenberger M, Raglione T and Mathew S (1997) The application of several atomic spectroscopy techniques in a pharmaceutical analytical research and development laboratory. *Spectroscopy* **12**: 14–23.

Ma TS (1990) Organic elemental analysis. *Analytical Chemistry* **62**: 78R–84R.

(1992) *Physician's Desk Reference*, 46th edn. Medical Economics Data, a division of Medical Economics Company. 1992.

Rousselet F and Thuillier F (1979) Atomic absorption spectrometric determination of metallic elements in pharmaceutical products. *Progress in Analytical Atomic Spectroscopy* **1**: 353–372.

Schulman SG and Vincent WR (1984) Atomic spectroscopy (in pharmaceutical analysis). *Drugs Pharmaceutical Science* **11**: 359–399.

Taylor A, Branch S, Crews HM and Halls DJ (1993) Atomic spectroscopy update – clinical and biological materials, foods and beverages. *Journal of Analytical Atomic Spectrometry* **8**: 79R–149R.

United States Pharmacopeial Convention (1975). *The United States Pharmacopeia*, nineteenth revision.

http://www.searlehealthnet.com/pipeline.html

http://www.allp.com/drug _dev.html

# Photoacoustic Spectroscopy, Applications

**Markus W Sigrist**, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY/ VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Applications**

## Introduction

In conventional absorption spectroscopy the measurement of absorption is transferred to a measurement of the radiation power transmitted through the sample. On the contrary, in photoacoustic spectroscopy, the absorbed power is determined directly via its heat and hence the sound produced in the sample. Photoacoustics, also known as optoacoustics, was pioneered by AG Bell, in 1880. The photoacoustic (PA) effect concerns the transformation of modulated or pulsed radiation energy, represented by photons, into sound. In general, two aspects have to be considered: first, the heat production in the sample by the absorption of radiation; and secondly, the resulting generation of acoustic waves. Closely related to the PA effect are photothermal (PT) phenomena which are caused by the original heating via absorption of radiation. While the PA effect is detected via acoustic sensors such as microphones, hydrophones or piezoelectric devices, the PT phenomena are sensed via the induced changes of the refractive index of the media by probe beam deflection, thermal lensing or, also, PT radiometry. Both PA and PT spectroscopy are widely used today in many applications. Experimental aspects are outlined in a separate article while this article discusses the main characteristics of this spectroscopic tool. The great potential is illustrated with examples from applications on solids, liquids and gases as well as in life sciences.

## Spectroscopic applications

PA and PT phenomena are widely used for numerous non-spectroscopic applications such as the determination of thermal diffusivity, non-destructive testing of materials (in particular the probing of sub-surface defects) by thermal wave imaging, time-resolved studies of de-excitation processes or on biological photoreceptors, studies of phase transitions, etc. Here, only spectroscopic applications are considered that demonstrate the main characteristics and the potential of photoacoustic spectroscopy (PAS). In the following, illustrative examples are presented for solids, liquids, gases, biological and medical samples.

### Studies on solids

A main advantage of PAS applied to solids is the fact that no elaborate sample preparation is required and unpolished sample surfaces pose no problems. Since the PA signal is proportional to the *absorbed* energy, even spectra of strongly scattering samples, e.g. powders, can easily be measured. However, it should be

mentioned that owing to the complex nature of the signal generation involving interstitial gas expansion, etc., PA studies on powders are usually only qualitative. Another advantage is the high sensitivity that is achieved because the PA detection is a null method for measuring absorption. Hence, absorbances as low as $10^{-7}$ can be detected. For modulated radiation a simple theoretical model has been developed which is based on the fact that the acoustic signal is due to a periodic heat flow from the solid to the surrounding gas, as the solid is cyclically heated by the absorption of the chopped light. Six different cases are distinguished, depending on the optical and thermal properties of the solid samples. This allows the unique feature of measuring totally opaque materials which is impossible by conventional transmission measurements. Hence PAS is technique to study weak bulk and surface absorption in crystals and semiconductors, to evaluate the level of absorbed energy in thin films, to measure the spectra of oxide films in metals, various powders, organic materials, etc. and also to investigate multi-layered samples.

An early example is shown in **Figure 1** for the insulator $Cr_2O_3$. Spectrum (A) depicts the normalized PA spectrum of $Cr_2O_3$ powder in the 200 to 1000 nm region. In comparison spectrum (B) shows an optical absorption spectrum obtained on a 4.4 μm thick bulk crystal, taken parallel and vertical to the crystal *c*-axis whereas spectrum (C) represents a diffuse reflection spectrum of $Cr_2O_3$ powder. The advantage of PAS is obvious in that the two crystal-field bands of the $Cr^{3+}$ ion at 460 and 600 nm are almost as clearly resolved in the PA spectrum of the powder as they are in the crystal spectrum, and substantially better resolved than in the diffuse reflectance spectrum. It should be noted, however, that the theoretical description of the PA effect in strongly scattering media is not straightforward and quantitative data are therefore difficult to determine from such spectra.

Another example concerns adsorbates on the surfaces of solids. PAS is expected to be rather sensitive to surface adsorption, especially if the substrate is transparent or highly reflective in the wavelength region in which the adsorbate absorbs. Both sinusoidal modulation of the incident laser beam and pulsed lasers have been used for this purpose. An interesting version is the modulation of the laser beam polarization to suppress the background signal that originates from substrate absorption. A fraction of only 0.005 of a monolayer of ammonia ($NH_3$) adsorbed on a cold silver substrate in ultrahigh vacuum was detected. An example is shown in **Figure 2** where the PT signal is recorded as a function of time as ammonia is slowly admitted to the system and condenses

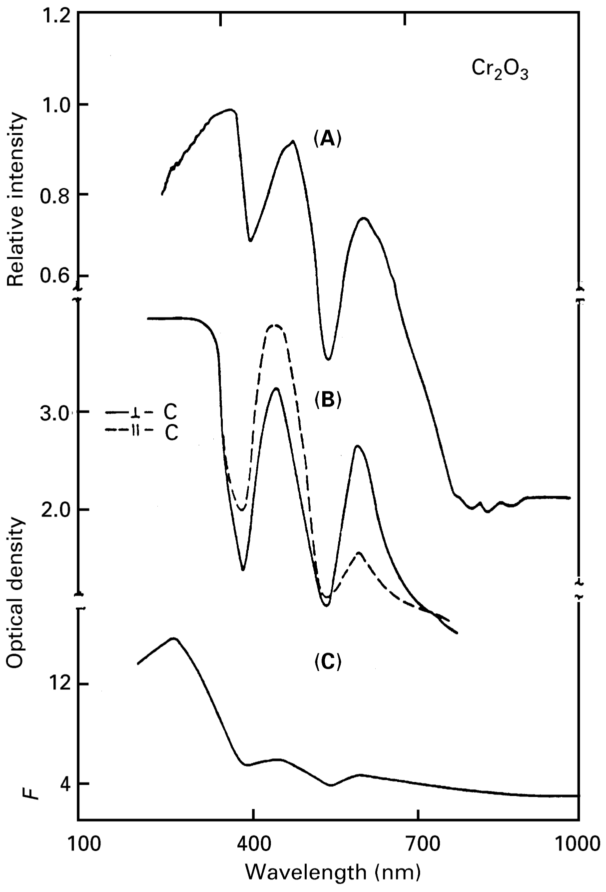



**Figure 1** (A) Normalized PA spectrum of $Cr_2O_3$ powder, (B) optical transmission spectrum of a 4.4 μm thick $Cr_2O_3$ crystal, (C) diffuse reflectance spectrum of $Cr_2O_3$ powder. All spectra were taken at 300 K. Reproduced with permission of Academic Press from Rosencwaig A (1977). In: Pao Y-H (ed) *Optoacoustic Spectroscopy and Detection.* New York: Academic Press.

on the silver substrate. The signal of a microbalance as indicator of molecular coverage is monitored simultaneously. Later studies were aimed at investigating the kind of adsorption in more detail, e.g. to differentiate between chemisorption and physisorption, by combining the high spectral resolution and high sensitivity offered by pulsed laser PAS.

In other studies, the wide free spectral range offered by Fourier transform infrared (FTIR) spectroscopy combined with the step-scan methods has been increasingly applied in conjunction with PA detection for infrared spectral depth profiling of laminar and otherwise optically heterogeneous materials. IR spectra that are often unavailable by use of other techniques become accessible from samples that are strongly absorbing or even opaque, from strongly light-scattering samples and from samples in situ. The scheme is also applied as an analytical tool for chemical characterization and quantification, e.g. of

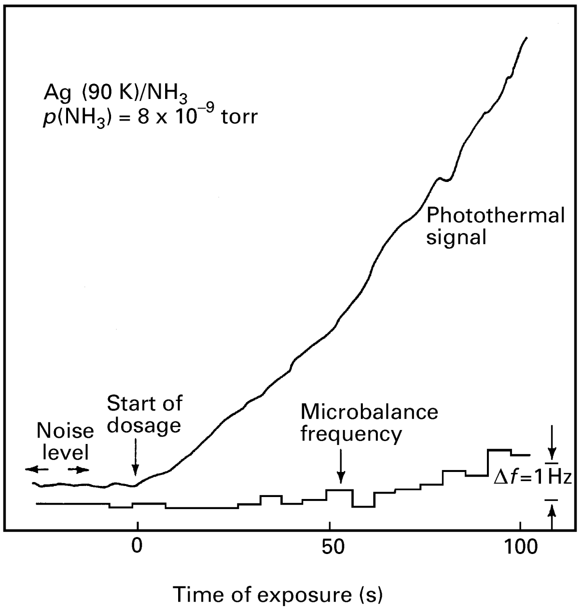



**Figure 2** Photothermal signal and microbalance record versus exposure time as ammonia molecules are slowly adsorbed on a silver substrate. The maximum coverage is 0.8 monolayers, the ammonia partial pressure in the system is $8 \times 10^{-9}$ torr. The noise level (left) indicates the signal from the clean substrate. Reproduced with permission of Elsevier from Coufal H, Trager F, Chuang T and Tam A (1984) *Surface Science* **145**: L504.

polychlorinated biphenyls (PCBs) in industrial waste management such as PCB contamination of soils.

Finally, the available spectral range for PA studies on solids has been extended to the X-ray region by using hard X-rays from synchrotron radiation. As an example, the X-ray absorption near Cu K-edge regions has been measured on copper (Cu), Cu alloys (brass) and Cu compounds (CuO, $Cu_2O$ and $CuInSe_2$) with a PA detector and compared with the usual X-ray absorption (10 μm thick Cu and brass foils and < 50 μm thick powdered samples of CuO, $Cu_2O$ and $CuInSe_2$ put on Scotch tape were used as specimens). It was found that the energy peak values derived from the PA spectra agree with those deduced from optical density spectra, suggesting that the heat production processes are also reflected in the absorption spectra. A more detailed insight is obtained by dividing the PAS data by the optical density data, i.e. by forming the ratios PAS:log($I_0/I_t$), which are proportional to the heat production efficiency. In **Figure 3** these ratios are plotted for Cu, $Cu_2O$ and $CuInSe_2$ versus the photon energy near the K-edge of Cu. The results clearly indicate differences between X-ray absorption and PA spectra and hence imply a spectral variation of the heat production efficiency. Obviously, the heat production process is also different in $Cu_2O$ compared with the other Cu compounds.

### Studies on liquids

Experimental and theoretical PA and PT studies on liquids comprise a wide absorption range from 'transparent' to opaque liquids.

For investigations on weakly absorbing media a flash-lamp-pumped dye laser with pulse energies of 1 mJ was used as excitation source and a submersed piezoelectric transducer for detecting the generated acoustic signals. The high sensitivity permits, e.g. the recording of the water spectrum in the visible range where accuracies of other techniques such as long-path absorption measurements are often limited. Another example concerns the study of weak overtones of the C–H stretch absorption band of hydrocarbons up to the 8th harmonic. In **Figure 4** the absorption band of the 6th harmonic at 607 nm of benzene dissolved in $CCl_4$ is plotted for different dilution ratios of benzene. With increasing dilution, the absorption peak is obviously blue-shifted and both the line width and the line asymmetry decrease. These and other results demonstrate that PAS permits the measurement of minimum absorption coefficients of $10^{-6}$ cm$^{-1}$, corresponding to absorbed laser pulse energies of only 1 nJ.

Another field of interest concerns analytical investigations on pollutants in liquids. Detection limits in the sub-ppb range were achieved by PAS, e.g. for carotene or cadmium in chloroform or for pyrene in heptane. More recently, pesticides in aqueous solutions have attracted interest. Different experimental arrangements with pulsed or CW pump lasers and various PA and PT lens detection schemes were used in these studies. Limits of detection are down to below $10^{-6}$ cm$^{-1}$, corresponding to ppb concentrations. An example is presented in **Figure 5** where the calibration curves for the detection of the dinitrophenol herbicide DNOC in aqueous solutions are compared with the untreated standard solution. The techniques used involved PT techniques, namely PT deflection spectroscopy (PDS, **Figure 5A**), thermal lensing (TL, **Figure 5B**), PT interferometric spectroscopy (PIS, **Figure 5C**), PAS (**Figure 5D**) and a conventional spectrophotometer (Cary 2400, **Figure 5E**). Obviously, the detection limit of the spectrophotometer in the low ppb (μg kg$^{-1}$) range is exceeded by the PA and PT methods. In particular, TL and PDS appear superior in the determination of environmental pollutants. It should be noted that both US and EU standards require detection limits of 0.1 μg L$^{-1}$ for pesticides in drinking water.

On the other end of the scale are opaque or strongly absorbing liquids. PA spectroscopy offers the great advantage that absorption coefficients that are two to three orders of magnitude higher than is

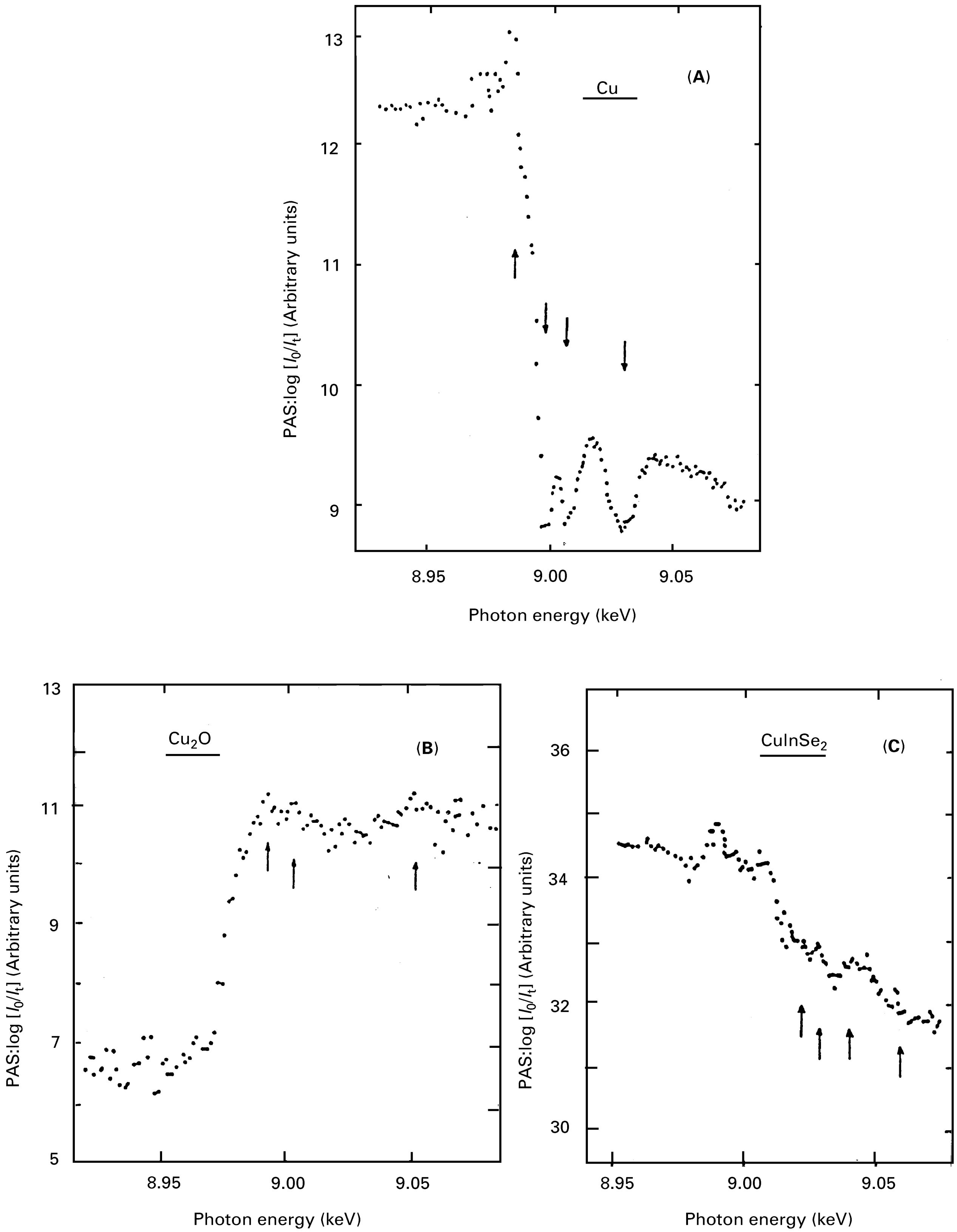



**Figure 3** X-ray PA spectra normalized with optical transmission spectra (PAS : log$I_0/I_t$), where $I_0$ and $I_t$ denote the incident and transmitted intensity, respectively, at the K-edge region for different copper compounds. (A) Pure Cu, (B) $Cu_2O$ and (C) $CuInSe_2$. Reproduced with permission of IGP AS, Trondheim, Norway from Toyoda T, Masujima T, Shiwaku H and Ando M (1995) *Proceedings of the 15th International Congress on Acoustics*, Vol I, 443.

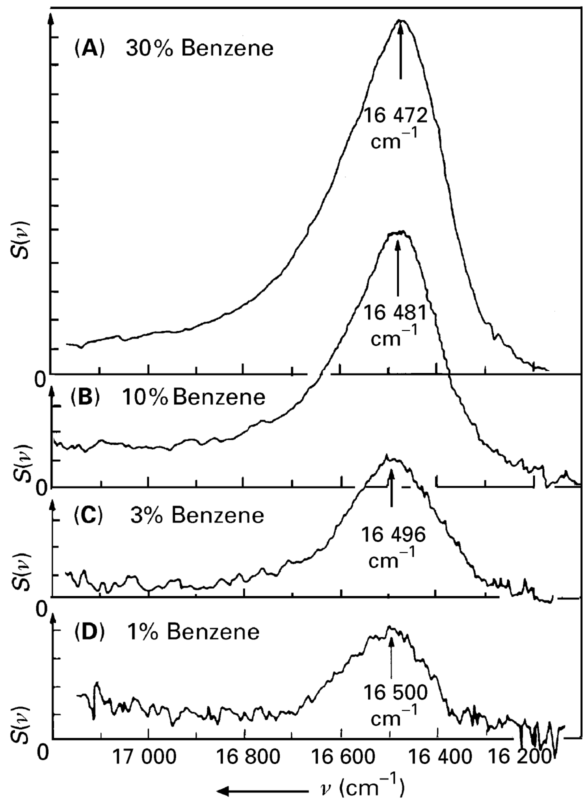



**Figure 4** PA spectra of the 6th harmonic absorption of the C–H bond of benzene dissolved in carbon tetrachloride ($CCl_4$) in arbitrary linear units, as the volume dilution ratios indicate. The positions of the absorption peaks are given. Reproduced with permission of the Optical Society of America (OSA) from Tam AC, Patel C and Kerl R (1979) *Optics Letters* **4**: 81.

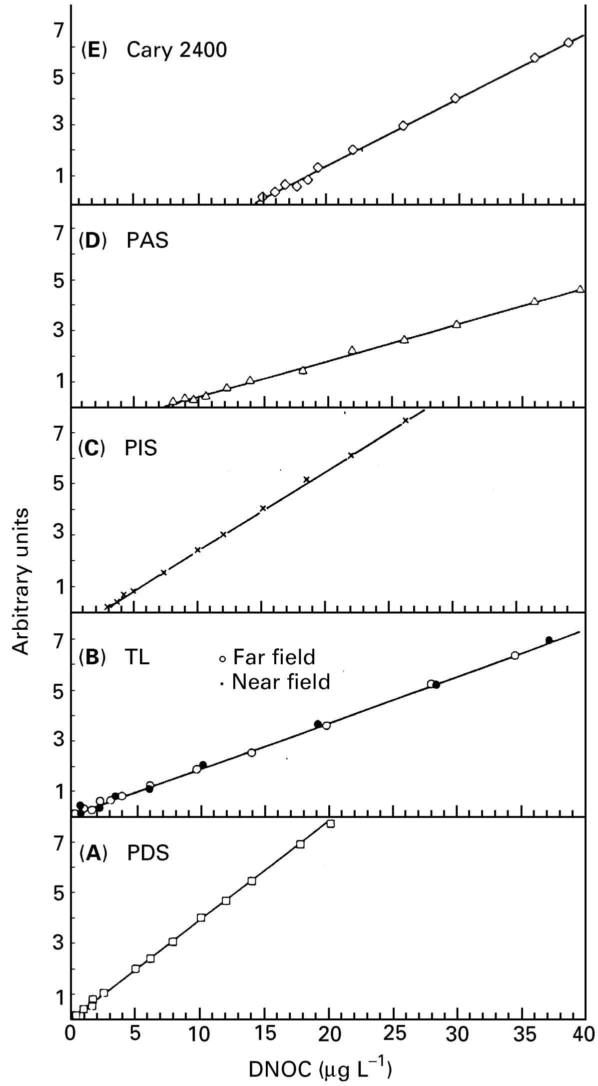



**Figure 5** Calibration curves for the dinitrophenol herbicide DNOC in aqueous solution when using different techniques: (A) PDS: Photothermal deflection spectroscopy, (B) TL: thermal lensing, (C) PIS: Photothermal interferometric spectroscopy, (D) PAS: a photoacoustic spectroscopy, (E) a conventional spectrophotometer Cary 2400. Reproduced with permission of SPIE from Faubel W (1997) Detection of pollutants in liquids and gases. In: Mandelis A and Hess P (eds) *Life and Earth Sciences.* Progress in Photothermal and Photoacoustic Science and Technology, Vol III, Chapter 8. Bellingham: SPIE.

accessible by conventional transmission spectroscopy can be determined without difficulties. Various schemes have been proposed for this case including the optothermal window. As example, the *trans*-fatty acid (TFA) content of margarine was determined using a $CO_2$ laser and the optothermal window. Good agreement with alternative techniques such as FTIR, gas–liquid chromatography and thin-layer chromatography was obtained.

### Studies on gases

Early PAS studies on gases had already demonstrated the high sensitivity that is achieved with a rather simple setup and have subsequently favoured further developments in trace gas monitoring. In comparison with conventional optical absorption measurements, PAS offers the following main advantages: (i) only short pathlengths are required which enables measurements at wavelengths outside of atmospheric transmission windows, (ii) the microphone as detector represents a simple room-temperature device with a wavelength-independent responsivity, (iii) scattering effects are less important, and (iv) the dynamic range comprises at least five orders of magnitude. Measurements are generally performed with the gas either contained in or flowed through a specially designed PA cell. Typically, a minimum detectable absorption coefficient $\alpha_{min}$ of the order of $10^{-8}$ cm$^{-1}$ atm$^{-1}$, corresponding to ppb ($10^{-9}$) concentrations, i.e. densities of µg m$^{-3}$, is achieved with laser-based setups. At the cost of dynamic range this limit can be lowered further to the <100 ppt range by

**Table 1** List and detection limits of selected gas species monitored by laser PAS under interference-free conditions

| *Species* | *Type of laser* | *Spectral region* (μM) | *Detection limit* [ppb] |
|---|---|---|---|
| Formic acid | Dye | 220[a] | 140 |
| Sulfur dioxide | Dye | 290–310[a] | 0.12 |
| Formaldehyde | Dye | 303.6[a] | 50 |
| Nitrogen dioxide | $Kr^+$ | 406.8[a] | 2 |
| Methane | DF | 3.8 | Few |
| Nitrous oxide | DF | 3.8 | Few |
| Carbon monoxide | $PbS_{1-x}Se_x$ | 4.6 | 40 |
| Nitric oxide | CO–SFR | 5.3 | < 0.1 |
| Phosgene | CO | 5.45 | Few |
| Acetaldehyde | CO | 5.66 | 3 |
| Carbon disulfide | CO | 6.48 | 0.01 |
| Ethane | CO | 6.7 | 1 |
| Pentane | CO | 6.8 | 0.1 |
| Trimethyl amine | CO | 6.93 | 10 |
| Dimethyl sulfide | CO | 6.95 | 3 |
| Acetylene | CO | 7.2 | 1 |
| Hydrazines | $CO_2$ | 9–11 | <10 |
| Freons | $CO_2$ | 9–11 | < 4 |
| Explosives | $CO_2$ | 9–11 | 0.2–25 |
| Ammonia | $CO_2$ | 9.22 | 0.4 |
| Ethanol | $CO_2$ | 9.46 | 17 |
| Ozone | $CO_2$ | 9.50 | 13 |
| Methanol | $CO_2$ | 9.68 | 5 |
| Ethylene | $CO_2$ | 10.53 | 0.3 |
| Sulfur hexafluoride | $CO_2$ | 10.59 | 0.01 |
| Vinyl chloride | $CO_2$ | 10.61 | 20 |

[a] In nm.

operating the PA cell intracavity. **Table 1** lists some gaseous compounds, laser sources used and (extracavity) detection limits achieved.

In practice, one usually deals with multicomponent samples. The analysis is done on the basis of the individual spectra and measurements performed at properly selected wavelengths to reduce absorption interferences. Apart from the PA signal amplitude the PA phase yields additional information for the analysis. A broad tuning range of the laser source and a narrow line width are advantageous for obtaining a high selectivity which is further enhanced for species with well structured spectra.

Most PA studies on trace gases have been devoted to laboratory investigations on collected air samples of different origin such as vehicle exhausts or industrial emissions. If the temporal evolution of the gas composition is of interest, the air is flowed continuously through the PA cell and the laser is switched repeatedly between appropriate wavelengths that are characteristic for the absorption of the gases to be recorded. In addition to laboratory analyses field studies yielding temporally and spatially resolved data on ambient trace gas concentrations and on their distributions are required to obtain a profound knowledge of atmospheric chemistry as well as of emission processes. Unlike lidar or long-path absorption measurements in the open atmosphere, PA schemes are not suited for remote studies but are applied to in situ measurements for the simultaneous monitoring of various compounds. Apart from non-laser based PA gas sensors only a few mobile laser PA systems have been operated so far. Examples include a balloon-borne system equipped with a spin-flip Raman CO laser PA spectrometer that was employed for recording stratospheric diurnal NO concentration profiles. Furthermore, a waveguide $CO_2$ laser PA system was applied successfully to in situ measurements at a power plant where the commonly used scheme for the reduction of the emission of nitric oxides ($NO_x$) by injection of ammonia ($NH_3$) into the combustion process was to be tested. This required a reliable, fast and selective monitoring of $NH_3$ down to the 1 ppm level under rough measurement conditions. Studies of this type at power-, incineration- or industrial-plants play a key role for the evaluation of the pollution, for the emission control and the testing of remedial strategies.

During the last few years a fully automated $CO_2$ laser PA spectrometer, which is installed in a trailer, has been developed and which is operated unattended for longer time periods. The computer control ensures a proper laser wavelength selection with a long-term frequency stability of $10^{-3}$ $cm^{-1}$ for ~70 laser transitions between 9.2 and 10.8 μm using $^{12}C^{16}O_2$ or for ~65 transitions between 9.6 and 11.4 μm with a $^{13}C^{16}O_2$ laser tube. The resonant PA cell is connected to a gas flow system and the air to be analysed is pumped continuously through the cell at atmospheric pressure and with a flow rate of typically 0.5–1 L $min^{-1}$. The air is not pretreated by any means except for measurements in dusty environments where a micropore filter is inserted into the gas stream at the air inlet. The humidity and $CO_2$ monitor in the air stream allow independent measurements of water vapour and $CO_2$ concentrations for comparison. Furthermore, the trailer is equipped with meteorological devices for wind and solar irradiance measurements. Hitherto, we have applied this mobile system to industrial stack emission sensing and ambient air monitoring in urban and rural environments. The stack emission measurements demonstrated the good time resolution and the high selectivity that are achieved in multicomponent analyses. In certain cases it could even be differentiated between isomers, e.g. between *o*- and *m*-dichlorobenzene, among various

other compounds, mostly VOCs, at ppm concentration levels. The selectivity is determined by the compound to be measured, and the tuning characteristics and bandwidth of the laser source.

The air in urban environments often contains numerous pollutants with rather high and varying concentrations. More recently, the mobile system was used in a harsh and noisy environment at the exit of a freeway tunnel to record gases emitted by road traffic during one week. The polluted air, filtered by a Teflon dust filter with a porosity of 1 µm, was flowed continuously through the PA cell and the laser was tuned sequentially to wavelengths characteristic for ammonia ($NH_3$), ethylene ($C_2H_4$) and $CO_2$ absorption as well as to reference wavelengths with no appreciable absorption by these compounds. Concentration profiles of these three species could thus be recorded almost simultaneously with a time resolution of 10 min. As **Figure 6** shows, the temporal concentration data are clearly correlated with the independently monitored CO concentration and the traffic density. Rather high concentrations are recorded, particularly for ammonia which are most probably caused by the majority of those cars that are equipped with catalytic converters. Based on the gas concentrations and the calculated air flow through the tunnel the corresponding emission factors (mass of an exhaust component per vehicle and kilometre) were determined. These factors are 15 mg km$^{-1}$ for ammonia, 26 mg km$^{-1}$ for ethylene,

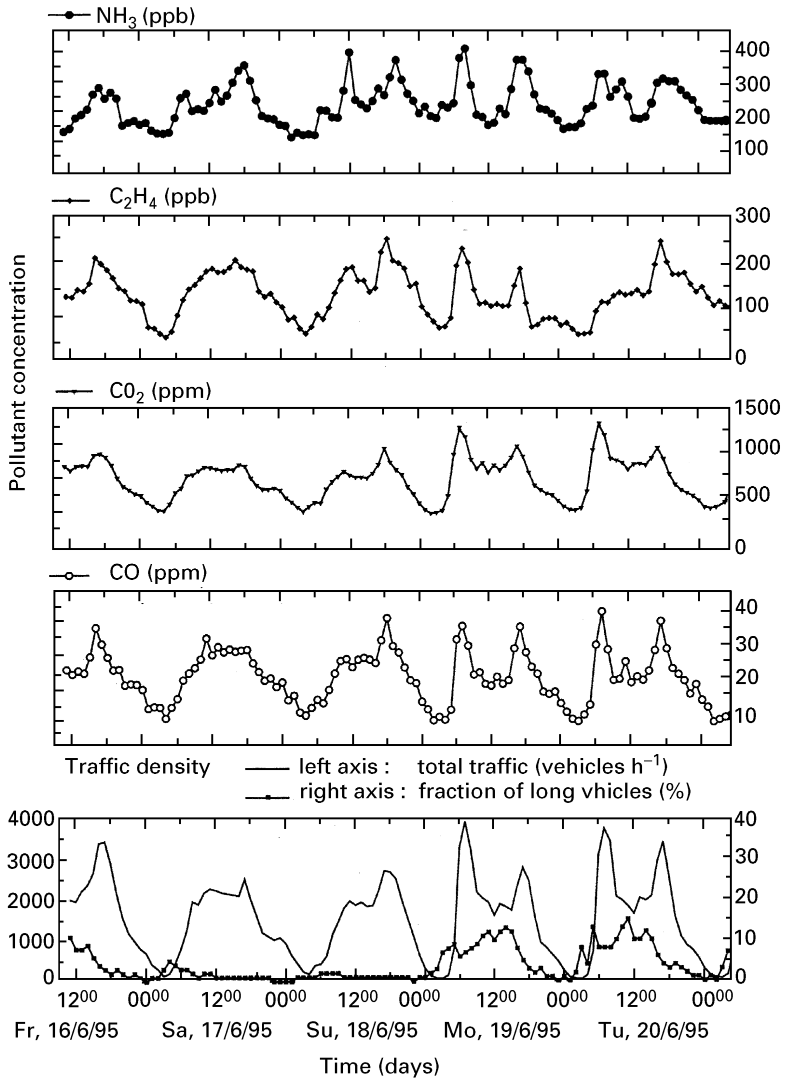



**Figure 6** Temporal concentration profiles of four air pollutants during 5 days taken at the exit of a freeway tunnel. The $NH_3$, $C_2H_4$ and $CO_2$ were measured photoacoustically, CO was recorded with a commercial IR gas analyser. Total traffic density and fraction of long vehicles are recorded at the bottom. Reproduced with permission of the American Chemical Society from Moeckli M, Fierz M and Sigrist M (1996) *Environmental Science and Technology* **30:** 2864.

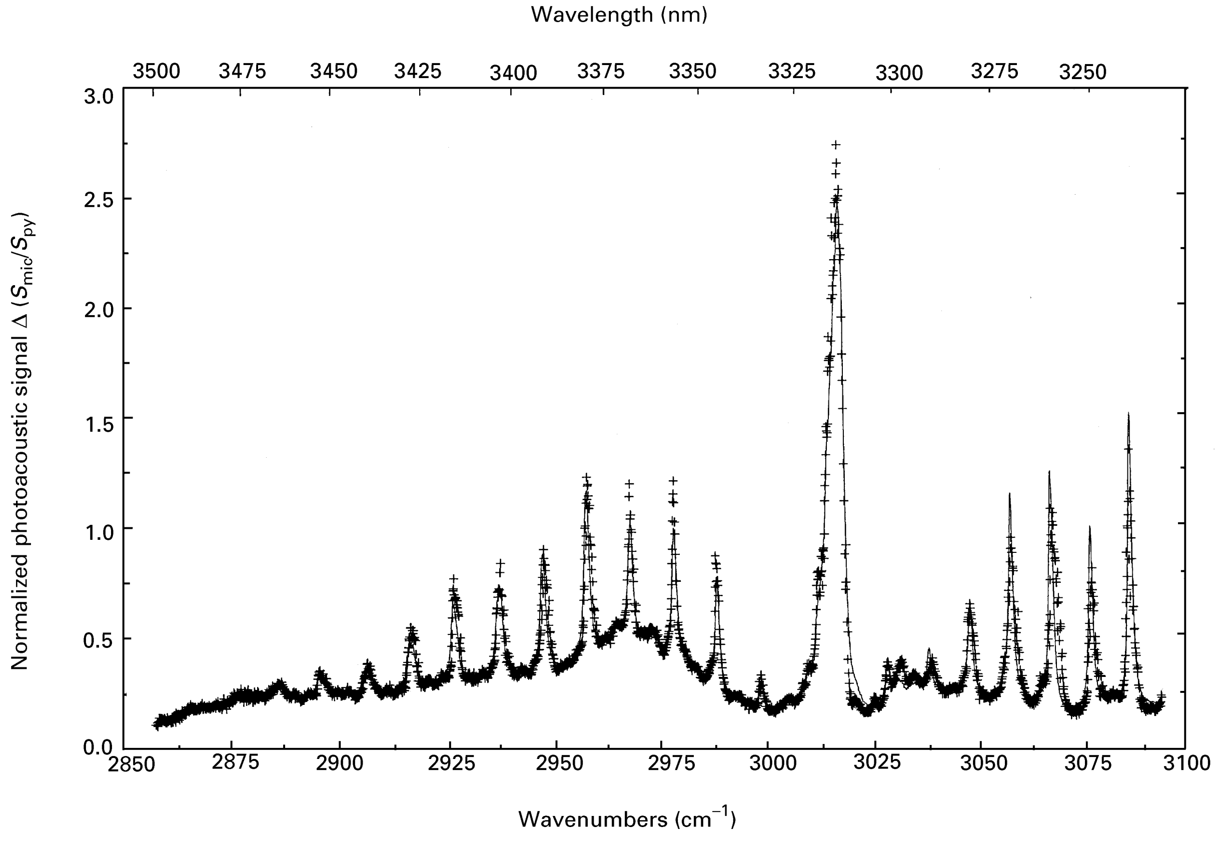



**Figure 7** Spectrum of a gas mixture containing methane, methanol, ethanol, isopentane, benzene and toluene, all at ppm concentrations, buffered to 960 mbar total pressure with synthetic air (80% $N_2$, 20% $O_2$). The measured spectrum (+) was taken photoacoustically with a difference frequency laser spectrometer based on an optical parametric oscillator (OPO) with a line width of 0.2 $cm^{-1}$. Excellent agreement is obtained with the superimposed fitted spectrum (——) using the HITRAN database for methane and the previously recorded reference spectra for the other substances. Reproduced with permission of Elsevier from Bohren A and Sigrist M (1997) *Infrared Physics and Technology* **38**: 423.

5.2 g $km^{-1}$ for CO and 201 g $km^{-1}$ for $CO_2$. Such data are valuable for the estimation of the total annual emission of certain compounds from road traffic and their fractions of the total load in certain area.

Ambient monitoring in rural air requires detection schemes with very high sensitivity. So far, studies of PA sensing have concentrated on a few gases. An example concerns the in situ recording of ambient $NH_3$, $H_2O$ vapour and $CO_2$ in a heath in the central Netherlands for several months in 1989. The mobile system was applied to measurements in a rural location in central Switzerland where $H_2O$ vapour, $CO_2$, $NH_3$, $O_3$ and $C_2H_4$ were recorded simultaneously with a time resolution of 10 min by using nine carefully selected $CO_2$ laser transitions.

A key issue in analysing multicomponent samples is the detection selectivity, which is strongly influenced by the tuning characteristics and line width of the source as well as by the measuring conditions themselves (e.g. reduced gas pressure). Recent laser developments can enhance the performance substantially. A continuously tunable narrowband high pressure $CO_2$ laser has, for example, enabled the analysis of a mixture of six $CO_2$ isotopes. In one study, presented in **Figure** 7, a mixture of six hydrocarbons at ppm concentrations has been analysed by an all-solid-state laser PA spectrometer.

### Studies in life sciences

The inherent high light scattering and the often strongly varying depth structure render biological and medical samples rather difficult for investigations with conventional spectroscopic tools. As various researchers, however, have demonstrated, PA and PT techniques can successfully by applied to media such as skin tissue, blood or plants. *In vivo* studies were performed on human skin with specially designed PA cells that allowed the study on living skin by avoiding the noise induced by pulsating blood. In particular, the absorption of UV light by protein (α-keratin) and the application of sunscreens to protect skin from UV damage were studied spectroscopically by the PA

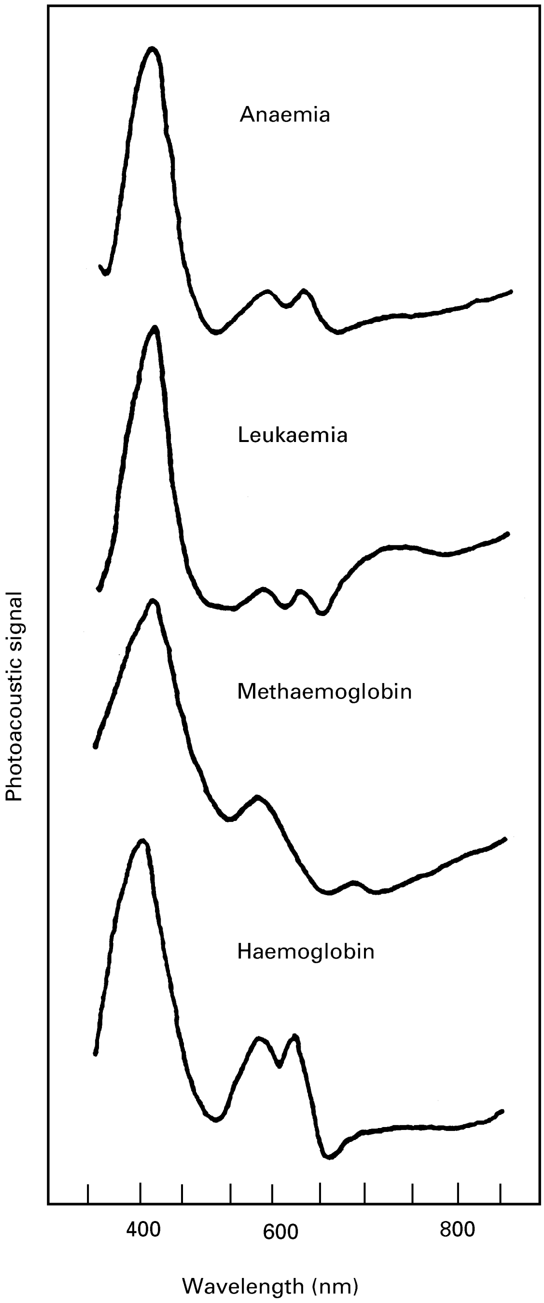



**Figure 8** Photoacoustic spectra of haemoglobin and of the blood from anaemia, leukaemia and methaemoglobin patients. Reproduced with permission of Springer from Pan Q, Qui S, Zhang S, Zhang J and Zhu S (1987) Springer Series in Optical Sciences, Vol 58, 542.

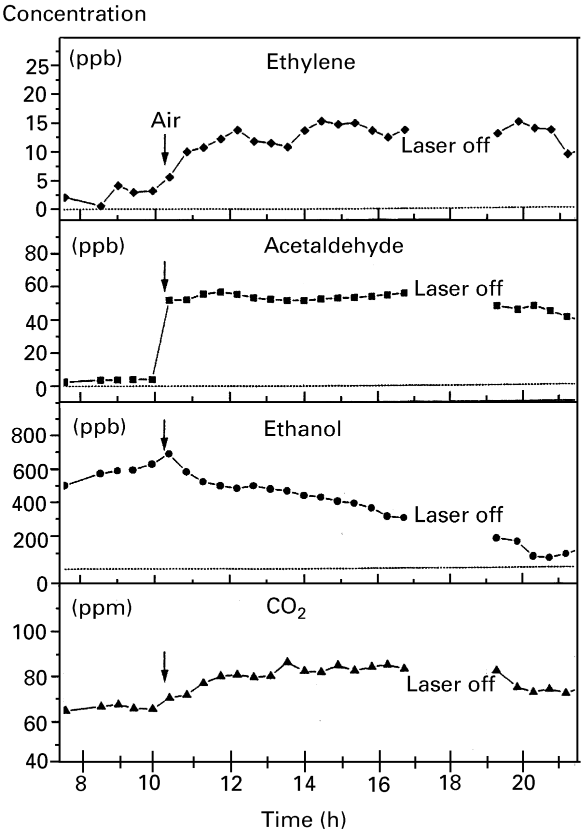



**Figure 9** Temporal evolution of the emission of different gases (ethylene, acetaldehyde, ethanol and $CO_2$) from a cherry tomato subjected to a change from anaerobic to aerobic conditions at $t = 10.2$ h (indicated by the arrow). The emission of a reference fruit kept under aerobic conditions was subtracted to obtain the data in this figure. Reproduced with permission of SPIE from Harren F and Reuss J (1997). Applications in plant physiology, entomology, and microbiology; gas exchange measurements based upon spectral selectivity. In: Mandelis A and Hess P *Life and Earth Sciences.* Progress in Photothermal and Photoacoustic Science and Technology, Vol III, Chapter 4. Bellingham: SPIE.

technique. The dependence on the kind of sunscreen and the amount applied (in μg cm$^{-2}$) were investigated as well as the rate of penetration into the skin and the time of residence in different skin layers.

The ability of PAS to detect specific compounds present in a highly diffusive medium is a great advantage for spectroscopic studies on blood. The protein haemoglobin is responsible for the $O_2$–$CO_2$ exchange between blood and tissue. It exhibits three absorption bands, at 415, 540 and 580 nm, that are caused by the tetraporphyrin cycle bound to its amino acid skeleton. In **Figure 8** the typical haemoglobin spectrum is compared with blood spectra of patients suffering from anaemia, leukaemia and methaemoglobin. The deviations from normal blood are clearly visible and can complement other diagnostic results. The noncontact character of PA investigations and the fact that only minor sample amounts are required is another advantage that permits further studies on the same sample by alternative techniques.

It should be noted that no complicated preparation, treatment or any purification processes are required before measurements. This is advantageous

also for studies on living plants. Photosynthetic activities of plants and the influence of environmental stress have been investigated by various research groups. As example, a PA apparatus has been developed to measure the oxygen evolution rate directly at a single leaf of a living plant. Environmental factors such as the effects of water stress, temperature extremes, varying light flux and gaseous pollutants were studied in details. A study on plant physiology is presented in **Figure 9**. The measurements were performed with a CO-laser intracavity PA arrangement. To enhance the detection specificity selective trapping was applied by leading the incoming air stream over different temperature levels of a cold trap. In the experiment, three cherry tomatoes were kept under pure nitrogen in a cuvette for 10 h before switching back to an air flow. This re-exposure obviously caused significant changes of the production rates of the different compounds. This example demonstrates that the PA technique is well suited for the study of fast responses of plant tissues to changing ambient conditions.

Another area of interest where PA spectroscopy is a valuable tool concerns food science. Examples include the determination of the iron content in milk powder concentrate, moisture in instant skim milk powder, stem in ground pepper or the detection of adulterated powdered coffee. An example of a deliberate adulteration in spices concerns the contamination of ground red paprika spice by red lead ($Pb_3O_4$) which enhances the colour of paprika but also adds to its total weight. PA studies on ground sweet red paprika have demonstrated that PA spectroscopy can be recommended as a method for rapid and gross screening for $Pb_3O_4$ adulterant. The current limit of detection of 2% w/w) is, however, above the internationally adopted maximum permissible level and inferior to that of established techniques like atomic absorption spectroscopy (AAS) or inductively-coupled plasma spectroscopy (ICPS).

*See also:* **Environmental Applications of Electronic Spectroscopy; IR Spectrometers; Laser Applications in Electronic Spectroscopy; Photoacoustic Spectroscopy, Theory; Photoacoustic Spectroscopy, Applications; Photoacoustic Spectroscopy, Methods and Instrumentation; Surface Studies By IR Spectroscopy; Zeeman and Stark Methods in Spectroscopy, Applications.**

## Further reading

Almond DP and Patel PM (1996) *Photothermal Science and Techniques*. London: Chapman & Hall.

Bialkowski SE (1996) *Photothermal Spectroscopy Methods for Chemical Analysis*, Chemical Analysis Series, Vol 134. New York: Wiley.

Bicanic D (ed.) (1998) Photoacoustic, Photothermal and related methods as problem solvers in agricultural and environmental sciences. Special Issue of *Instrumentation Science and Technology* **26** (2&3).

Hess P (ed) (1989) *Photoacoustic, Photothermal and Photochemical Processes in Gases*. Topics in Current Physics, Vol **46**. Berlin: Springer.

Hess P (ed) (1989) *Photoacoustic, Photothermal and Photochemical Processes at Surfaces and in Thin Films*, Topics in Current Physics, Vol. **47**. Berlin: Springer.

Mandelis A (ed) (1982) *Principles and Perspectives of Photothermal and Photoacoustic Phenomena*, Progress in Photothermal and Photoacoustic Science and Technology, Vol I. New York: Elsevier.

Mandelis A and Hess P (eds) (1997) *Life and Earth Sciences*, Progress in Photothermal and Photoacoustic Science and Technology, Vol III. Bellingham, Washington: SPIE.

Meyer PL and Sigrist MW (1990) Atmospheric pollution monitoring using $CO_2$-laser photoacoustic spectroscopy and other techniques. *Reviews of Scientific Instruments* **61**: 1779.

Pao Y-H (ed) (1977) *Optacoustics Spectroscopy and Detection*. New York: Academic Press.

Rosencwaig A (1978) Photoacoustic spectroscopy. In: *Advances in Electronics and Electron Physics*, Vol 46, Chapter 6.

Sell JA (ed.) (1989) *Photothermal Investigations of Solids and Liquids*. San Diego: Academic Press.

Sigrist MW (1986) Laser generation of acoustic waves in liquids and gases. *Journal of Applied Physics* 60: R83–R121.

Sigrist MW (1994) Air monitoring by laser photoacoustic spectroscopy. In: Sigrist MW (ed) *Air Monitoring by Spectroscopic Techniques*, Chemical Analysis Series, Vol 127, Chapter 4. New York: Wiley.

Tam AC (1983) Photoacoustics: spectroscopy and other applications. In: Kliger DS (ed) *Ultrasensitive Laser Spectroscopy*, Chapter 1. New York: Academic Press.

Tam AC (1986) Application of photoacoustic sensing techniques. *Reviews of Modern Physics* **58**: 381–431.

Zharov VP and Letokhov VS (1986) *Laser Optacoustics Spectroscopy*, Springer Series in Optical Sciences, Vol 37. Berlin: Springer.

# Photoacoustic Spectroscopy, Methods and Instrumentation

**Markus W Sigrist**, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

Copyright © 1999 Academic Press

ELECTRONIC SPECTROSCOPY
Methods & Instrumentation

## Introduction

Since the discovery of the photoacoustic (PA) effect by Bell in 1880, who used the Sun as radiation source, a foot-operated chopper for modulation and an earphone as acoustic detector, the PA effect has found numerous applications as a sensitive and rather simple technique for determining optical, thermal and mechanical properties of all kinds of samples. This article focuses on methods and instrumentation employed in *spectroscopic* applications. Since *photothermal* (PT) spectroscopy is discussed elsewhere in the encyclopedia, PT schemes are only briefly mentioned here, whereas emphasis is put on instrumentation used in *photoacoustic* spectroscopy.

Although the technique of photoacoustic spectroscopy has existed for more than a century it is particularly the advent of lasers as radiation sources with high spectral brightness that has initiated a renaissance of the PA effect. In the meantime a great variety of experimental schemes have been developed which render the PA method a very versatile spectroscopic tool.

## Photoacoustic and photothermal schemes

### Experimental arrangements

As the photoacoustic (and related photothermal (PT)) phenomena comprise a large diversity of facets there exist various detection techniques which rely on the acoustic or thermal disturbances caused by the absorbed radiation. The selection of the most appropriate scheme for a given application depends on the sample, the sensitivity to be achieved, ease of operation, ruggedness, and any requirement for noncontact detection, e.g. in aggressive media or at high temperatures and/or pressures. **Figures 1–3** present the most typical arrangements applied for solid, liquid and gaseous samples, respectively.

Experimental schemes for PA studies on solid samples include the measurement of the generated pressure wave either directly in the sample with a piezoelectric sensor for the pulsed regime, or indirectly in the gas which is in contact with the sample by a microphone. These most widely used setups are depicted in **Figures 1A** and **1B**. The indirect detection of the generated acoustic wave in the gas phase with the microphone is inevitable if a direct contact with the sample is not readily possible, e.g. for samples such as powders, gels or grease. The properties of piezoelectric transducers and microphones as pressure sensors are discussed below.

If the use of pressure sensors is not appropriate, because, for example, a piezoelectric transducer cannot be attached to the sample or measurements need to be done at high temperatures which hinders the application of microphones, noncontact techniques are to be applied. As shown in **Figure 1C** the induced spatial and temporal gradient of the refractive index can be sensed in this case by monitoring the deflection of a probe beam either within the (transparent) sample or directly above the (plane) sample surface (PT beam deflection or so-called mirage effect). Changes of the surface reflectivity or slight deformations of the surface (PT beam displacement) can also be detected in a noncontact manner by a probe beam. Finally, as depicted in **Figure 1D**, variations of the thermal radiation from the surface can be monitored with an infrared detector (PT radiometry). This method is of particular interest for measurements at elevated temperatures owing to the increased radiation intensity according to the Stefan–Boltzman law. Still other techniques include pyroelectric detection in thin films, thermal lensing and interferometric methods.

The typical experimental arrangement for absorption spectroscopy in weakly absorbing liquids is shown schematically in **Figure 2A**. The beam of a pulsed tunable laser is directed through the PA cell that contains the sample under study. The generated acoustic waves are detected by a piezoelectric transducer with fast response time. Usually, only the first peak of the ringing acoustic signal is taken and further processed. Pulse-to-pulse variations of the laser power are accounted for by normalizing the

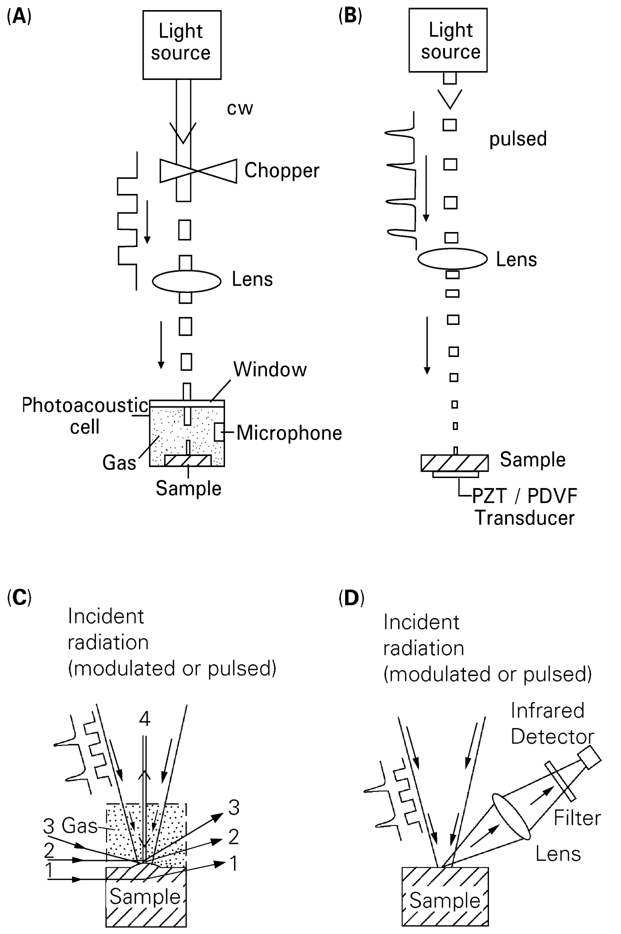



**Figure 1** Typical experimental arrangements used for photoacoustic (PA) and photothermal (PT) studies on solids. As indicated one differentiates between modulated ( ) or pulsed ( ) incident radiation. (A) Indirect PA detection by microphone in the gas phase. (B) Direct PA detection with PZT transducer or PVDF foil attached to solid. (C) PT sensing of the gradient of the refractive index with probe beam deflection in the (transparent) sample (probe beam 1), or above the sample surface (Mirage effect, probe beam 2). Monitoring of the generated surface displacement (probe beam 3) or of the change of surface reflectivity (probe beam 4). (D) PT radiometry senses the induced change of the IR radiation that is radiated off the sample surface.

piezoelectric signal with the laser power measured with the power meter after the cell.

Another area of interest is the measurement of opaque or strongly absorbing liquids. A simple open PA cell called an optothermal window was developed for this purpose as displayed in **Figure 2B**. It essentially consists of an uncoated ZnSe window to which an annular lead zirconate titanate (PZT) piezotransducer is glued from the bottom. The excitation beam from a laser passes unobstructed through this window and is absorbed by a droplet of the sample deposited on the other side of the ZnSe disc. The generated heat diffuses into the disk which expands. The induced stress is then recorded by the PZT transducer. Unlike in conventional transmission spectroscopy where the cell thickness is the restricting factor in dealing with strongly absorbing samples, the magnitude of the optothermal window signal depends solely on the product between the absorption coefficient and the thermal diffusion length whereby the latter can be adapted via the modulation frequency.

The typical setup for gas phase measurements is shown in **Figure 3A**. A tunable laser with narrow line width, or a conventional (broadband) radiation source followed by optical filters, is used. In general, amplitude-modulated (or sometimes pulsed) radiation is directed through the PA cell. The acoustic sensor is usually a commercial electret microphone or a condenser microphone. These devices are easy to use and sensitive enough for trace gas studies with very low absorptions. Often, the detection threshold is neither determined by the microphone responsivity $R_{mic}$ itself nor by the electrical noise but rather by other sources (absorption by desorbing molecules from the cell walls, window heating, ambient noise, etc.). However, if this latter background is known from reference measurements, the ultimate detection sensitivity is determined solely by fluctuations of the radiation intensity, and by microphone and amplifier noise. The frequency dependence of $R_{mic}$ is usually rather small and the temperature dependence may have to be taken into account in special cases only. If modulated radiation is employed the microphone signal is fed to a lockin amplifier locked to the modulation frequency. Since, according to theoretical

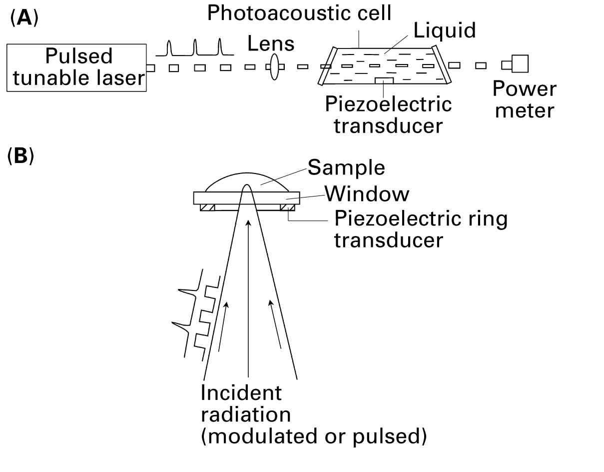



**Figure 2** Typical experimental arrangements used for PA detection in liquids. (A) PZT detection of acoustic wave generated by pulsed radiation in weakly absorbing liquid. (B) Optothermal window setup applied for studies on strongly absorbing or opaque liquids with modulated or pulsed radiation.

considerations, the microphone signal amplitude is proportional to the absorbed power for weakly absorbing media, the average radiation power is recorded simultaneously by a power meter for normalization.

If pulsed radiation is employed, the microphone bandwidth is often not sufficient to resolve the temporal shape of the generated acoustic pulses. However, common microphones can still be used even for nanosecond laser pulses because the length of a single acoustic pulse is essentially determined by the transit time of the acoustic wave across the beam radius. Normalized PA amplitudes are obtained by dividing the microphone signal peaks by the corresponding laser pulse energy that is recorded by a sensor such as a pyroelectric detector. Averaging over several pulses improves the signal-to-noise ratio. Another approach for pulsed radiation consists of using an acoustic resonator with high Q-factor as the gas cell, recording the microphone signals in the time domain but analysing the PA signal amplitudes after Fourier transformation in the frequency domain. The excited cell resonances then appear in the PA frequency spectra.

An important issue for many applications concerns the calibration of the entire PA or PT detection system. Since the PA signal depends on many factors that are not known with sufficient accuracy, a straightforward calibration is often achieved by employing a reference sample with known absorption. As an example, certified gas mixtures (trace gas diluted in a nonabsorbing buffer gas) or well characterized dye solutions in the case of liquids are used. The situation is more difficult with solid and biological samples, particularly layered media, powders, gels or tissue. In such cases, quantitative data are difficult, if not impossible, to obtain. But even qualitative instead of quantitative spectra are often valuable, especially when other spectroscopic techniques fail owing to opaqueness or strong scattering of the sample.

It should be emphasized that numerous different versions and modifications of these general schemes have been presented in the literature. In particular, combinations of conventional methods, such as Fourier-transform IR (FT-IR) or gas chromatography (GC), with PA detection have been reported. Some types of PA detection schemes are also implemented in commercial spectrometers.

In the following, the different components of PA spectrometers are briefly discussed.

## Radiation sources

In commercial PA spectrometers, incoherent sources such as lamps are employed in combination with filters or with an interferometer. Devices equipped with a small light bulb, with either a chopper or direct current modulation as modulated radiation source and appropriate filters to avoid absorption

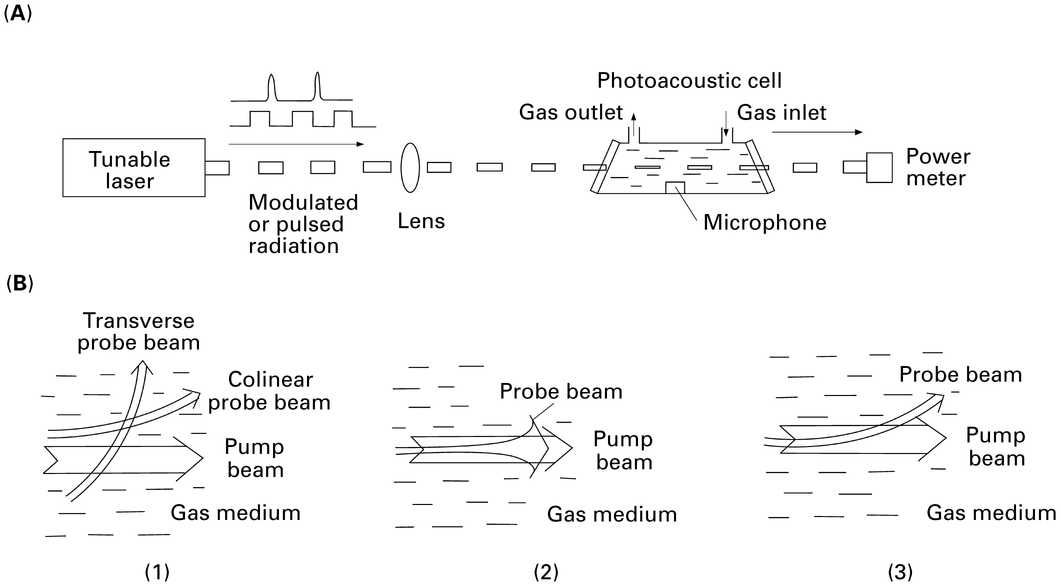



**Figure 3** Typical experimental PA and PT arrangements used for gas monitoring with tunable laser sources. (A) PA detection with conventional microphone in resonant gas cell for modulated cw radiation or in nonresonant cell for pulsed radiation. (B) Noncontact refractive index sensing schemes with displaced colinear or transverse probe beam (PA deflection, 1), thermal lensing (2), or colinear probe beam (PT deflection, 3).

interferences with other species, are used as compact gas sensors, e.g. for indoor $CO_2$ monitoring.

However, since the generated PA signal is proportional to the absorbed (and thus to the incident) radiation power, powerful radiation sources, particularly lasers offering high spectral brightness, are advantageous for achieving high detection sensitivity and selectivity in spectroscopic applications. In the UV and visible spectral range, excimer and dye lasers have been employed, whereas in the midinfrared (fundamental or mid-IR) wavelength range line-tunable $CO_2$ and CO lasers dominate the applications. Diode lasers have so far only rarely been employed in PA spectroscopy owing to their limited power. This situation may, however, change with the ongoing developments in this field. On the one hand, near infrared diode lasers with sufficient power for PAS are available for monitoring overtones and combination bands of molecular fundamental absorptions. On the other hand, current efforts focus on the implementation of widely tunable narrowband all-solid-state laser devices in the mid-IR region for accessing the (much stronger) fundamental absorptions. Optical parametric oscillators (OPOs) and difference frequency generation (DFG) in nonlinear crystals are certainly of great interest for compact spectrometers. Furthermore, recent developments in quantum cascade lasers look very promising in this respect.

### Modulation schemes

Modulation schemes can be separated into the modulation of the incident radiation and the modulation of the sample absorption itself. The first technique includes the most widely used amplitude modulation (AM) of continuous radiation by mechanical choppers, electrooptic or acoustooptic modulators as well as the modulation of the source emission itself by current modulation or pulsed excitation. In comparison to AM, frequency (FM) or wavelength (WM) modulation of the radiation may improve the detection sensitivity by eliminating the continuum background caused by a wavelength-independent absorption, e.g. of the cell windows, known as window heating. This type of modulation is obviously most effective for absorbers with narrow line width and most easily performed with radiation sources whose wavelength can rapidly be tuned within a few wavenumbers. Pulsed excitation is often applied for liquids but is also of interest for gaseous samples because it permits time gating and the excitation of acoustic resonances.

In certain cases the modulation of the absorption characteristics of the sample itself is advantageous. In gas studies the Stark or Zeeman effect has been employed, i.e. by applying a modulated electric or magnetic field to the sample. The result is a suppression of the continuum background and an enhancement of detection selectivity in multicomponent samples because, for example, Stark modulation only affects molecules with a permanent electric dipole moment like ammonia ($NH_3$) or nitric oxide (NO) while other, possibly interfering molecules, are not affected. Finally, combinations of both amplitude and sample absorption modulation have been successfully applied, e.g. for the sensitive detection of ammonia in the presence of absorbing water vapour and carbon monoxide.

### Photoacoustic cell designs

The PA cell serves as a container for the sample under study and for the microphone or some other device for the detection of the generated acoustic wave. An optimum design of the PA cell represents a crucial point when background noise ultimately limits the detection sensitivity. In particular, for trace gas applications many cell configurations have been presented including acoustically resonant and nonresonant cells, single- and multipass cells, as well as cells placed intracavity. Nonresonant cells of small volume are mostly employed for solid samples with modulated excitation or for liquids and gaseous samples with pulsed laser excitation. As a unique example, a small-volume cell equipped with a 'tubular' acoustic sensor consisting of up to 80 single miniature microphones has been developed. These microphones are arranged in eight linear rows with ten microphones in each row. The rows are mounted in a cylindrical geometry parallel to the exciting laser beam axis and located on a circumference around the axis. This configuration is thus ideally adapted to the geometry of the generated acoustic waves.

Resonant cells, in combination with modulated excitation, are normally applied for gas monitoring. These cells are operated on longitudinal, azimuthal, radial, or Helmholtz resonances. The signal enhancement by the Q-factor (usually >100) is often advantageous. Resonance frequencies lie in the kHz range resulting in resonance widths of a few Hz. Furthermore, the gas handling for the cell can be designed in such a way that the gas inlets and outlets are located at pressure nodes of the acoustic resonance which allows measurements in flowing gas with flow rates of the order of 1 L $min^{-1}$ without increasing the noise level.

Finally, cells developed for special purposes have been suggested, such as windowless cells equipped with acoustic baffles to reduce the influence of the ambient noise or heatable cells for studying liquid samples with low volatility.

### Detection sensors

As mentioned above the acoustic disturbances generated in the sample are detected by some kind of pressure sensor. In contact with liquid or solid samples these are piezoelectric devices such as lead zirconate titanate (PZT), $LiNbO_3$ or quartz crystals with a typical responsivity $R$ in the range of up to V $bar^{-1}$ or thin polyvinylidene-difluoride ($PVF_2$ or PVDF)-foils with lower responsivity. These sensors offer fast response times and are thus ideally adapted for pulsed photoacoustics.

For studies in the gas phase, commercial microphones are employed. These include miniature electret microphones such as Knowles or Sennheiser models with typical responsivities $R_{mic}$ of 10–20 mV $Pa^{-1}$ as well as condenser microphones, e.g. Brüel & Kjær models with typical $R_{mic}$ of 100 mV $Pa^{-1}$. Usually $R_{mic}$ depends only weakly on frequency. The electret microphones produced for hearing aids exhibit a rather flat frequency response between, say, 20 Hz and 20 kHz whereas the bandwith $\Delta\nu$ of condenser microphones extends to frequencies of 100 kHz. All these microphones are thus well suited for typical modulation frequencies in the 100 Hz to kHz range. For pulsed applications, the general relation between $R_{mic}$ and $\Delta\nu$, as well as the occurrence of external noise implies a reduction of the signal-to-noise ratio for very large bandwidths so that miniature electret microphones are often appropriate detectors also in this case. The detection sensitivity can be enhanced by adding the signals of several microphones. In such a configuration the signal increases with the number of microphones used, whereas the microphone random noise decreases with the square root of their number. Since electret microphones are small and cheap, a number of them can be arranged in a still compact geometry. A further improvement of sensitivity is expected from the insertion of an electrical filter that cuts the low-frequency components below, say, 1 kHz of the signal because these components contribute less to the increase of the signal-to-noise ratio than the higher-frequency components. Finally, an adaption of the frequency response of the microphone preamplifier and amplifier stages to that of the microphone is advantageous to fully exploit all the sensed frequency contributions except noise components at frequencies not contributing to the acquired signal.

If there is a need for noncontact detection, refractive index sensing, notably thermal lensing and both PA and PT deflection, are employed. These methods use a pump beam and a probe beam (HeNe or diode laser) in either colinear or transverse arrangement as shown in **Figure 3B**. In comparison to the conventional PA method with pressure sensors, these schemes offer similar sensitivity but require a somewhat more sophisticated setup and imply a more difficult calibration.

## List of symbols

$R$ = responsivity; $R_{mic}$ = microphone responsivity; $\Delta\nu$ = bandwidth.

*See also:* **Environmental and Agricultural Applications of Atomic Spectroscopy; Environmental Applications of Electronic Spectroscopy; IR Spectrometers; Photoacoustic Spectroscopy, Applications; Photoacoustic Spectroscopy, Theory; Surface Studies By IR Spectroscopy; X-Ray Absorption Spectrometers.**

## Further reading


Almond DP and Patel PM (1996) *Photothermal Science and Technique. London*: Chapman & Hall.

Mandelis A (ed) (1992) Principles and perspectives of photothermal and photoacoustic phenomena. In: *Progress in Photothermal and Photoacoustic Science and Technology*, Vol 1. New York: Elsevier.

Meyer PL and Sigrist MW (1990) Atmospheric pollution monitoring using $CO_2$-laser photoacoustic spectroscopy and other techniques. *Review of Scientific Instruments* **61**: 1779–1807.

Pao, Yoh-Han (ed) (1977) *Optoacoustic Spectroscopy and Detection*. New York: Academic Press.

Rosencwaig A (1980) Photoacoustics and photoacoustic spectroscopy. In *Chemical Analysis Series*, Vol 57. New York: Wiley.

Sigrist MW (1986) Laser generation of acoustic waves in liquids and gases. *Journal of Applied Physics* **60**: R83–R121.

Sigrist MW (1994) Air monitoring by laser photoacoustic spectroscopy. In: Sigrist MW (ed) *Air Monitoring by Spectroscopic Techniques*, Chemical Analysis Series, Vol 127, Chapter 4. New York: Wiley.

Tam AC (1983) Photoacoustics: spectroscopy and other applications In: Kliger DS (ed) *Ultrasensitive Laser Spectroscopy*, Chapter 1. New York: Academic Press.

Tam AC (1986) Applications of Photoacoustic Sensing Techniques. *Review of Modern Physics* **58**: 381–431.

Zharov VP and Letokhov VS (1986) *Laser Optoacoustic Spectroscopy*. Springer Series in Optical Sciences, Vol 37. Berlin: Springer.

# Photoacoustic Spectroscopy, Theory

**András Miklós**, **Stefan Schäfer** and **Peter Hess**,
University of Heidelberg, Germany

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY/ VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Theory**

## Introduction

The phenomenon of the generation of sound when a material is illuminated with nonstationary (modulated or pulsed) light is called the photoacoustic (PA) effect. Photoacoustic spectroscopy (PAS) is the application of the PA effect for spectroscopic purposes. It differs from conventional optical techniques mainly in that, even though the incident energy is in the form of photons, the interaction of these photons with the material under investigation is studied not through subsequent detection and analysis of photons after interaction (transmitted, reflected or scattered), but rather through a direct measurement of the effects of the energy absorbed by the material. Since photoacoustics measures the transient internal heating of the sample, it is clearly a form of calorimetry as well as a form of optical spectroscopy.

In the following discussion the complex PA effect will be divided into three steps:

- Heat release in the sample material due to optical absorption.
- Acoustic and thermal wave generation in the sample material.
- Determination of the PA signal in a PA detector.

A quantitative analysis of the PA signal is possible only when all three steps can be described quantitatively. Up to now this could be achieved only in a few special cases. Excitation of rovibrational levels leads to a complete transformation of the absorbed radiation into heat, whereas for vibronic excitation competing channels exist such as emission of radiation and photoinduced reactions. In this latter case, additional information on these competing channels is needed for the theoretical description.

A quantitative treatment of acoustic and thermal wave generation is usually possible only for simple excitation geometries, as can be realized by laser excitation, using the basic equations of fluid mechanics and thermodynamics.

The PA signal finally detected with a calibrated microphone for quantitative analysis is also influenced by the shape and nature of the photoacoustic cell, which impose boundary conditions on the evolution of the generated acoustic waves. However, only for highly symmetric resonant setups with high $Q$ factors does a theoretical analysis seem to be possible. This means that for most experimental arrangements used in PAS a quantitative signal analysis cannot be achieved and more or less drastic approximations have to be introduced into the signal analysis.

## Heat release in the sample material due to optical absorption

The interaction of photons with the material may produce a series of effects (**Figure 1**). If any of the incident photons are absorbed by the material, internal energy levels (rotational, vibrational, electronic) within the sample are excited. The excited state may lose its energy by radiation processes, such as spontaneous or stimulated emission, and by nonradiative

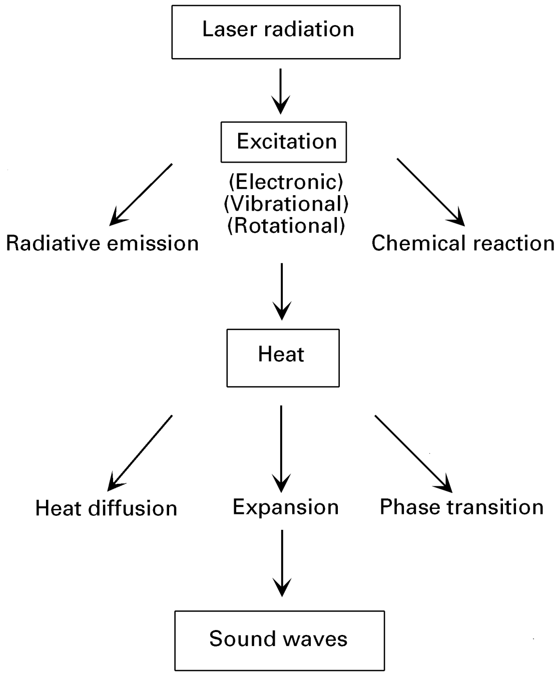



**Figure 1** Elementary processes occurring during PA signal generation. The absorbed photon energy is partly transformed into heat and acoustic energy.

deactivation, which channels at least part of the absorbed energy into heat. In a gas this energy appears as kinetic energy of the gas molecules, while in a solid it appears as vibrational energy of ions or atoms (phonons). In the case of vibrational excitation of gas molecules, radiative emission and chemical reactions do not play an important role, because the radiative lifetime of vibrational levels is long compared with the time needed for collisional deactivation at ordinary pressures and the photon energy is too small to induce reactions. Thus the total absorbed energy is released as heat. However, in the case of electronic excitation, the emission of radiation and chemical reaction processes may compete efficiently with collisional deactivation. Chemical reactions may also contribute to the release of heat, and thus they may increase the PA effect. If photodissociation occurs, for example, the local increase of the number of molecules and the thermalization of the recoil energy of the fragments generates a local pressure and temperature rise.

The heat release due to optical absorption in the sample material can be modelled by a rate equation. If it is assumed that the thermalization of the absorbed photon energy can be described by a simple linear relaxation process, the heat released per unit volume and time can be determined by solving the rate equation. If the near-resonant vibration–vibration (V-V) and the vibration–translation (V-T) relaxation are the fastest processes (as, for example, in many gases at atmospheric pressure), the heat power density will be proportional to the absorption coefficient and to the incident light intensity. In cases of very short laser pulses or high light intensity, optical saturation may occur. Then the heat production will be a nonlinear function of the light intensity and absorption coefficient. The timescales of the temporal evolution of the processes involved are summarized in **Figure 2**. The heat release may be delayed in gas mixtures if the excess energy of the excited molecule can be channelled by collisions to a long-lifetime transition of another species.

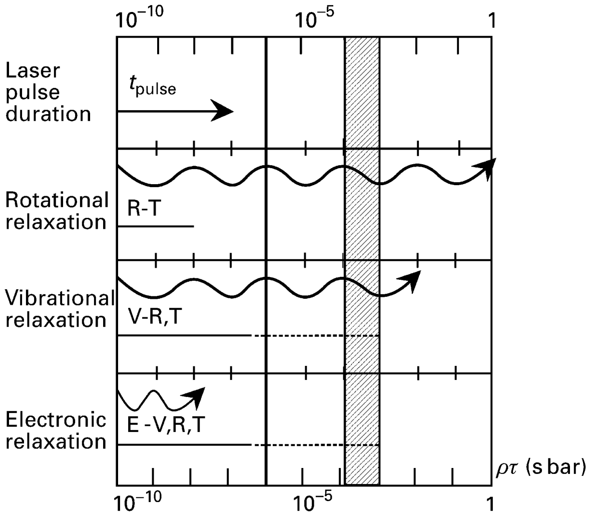



**Figure 2** Timescales for the radiative emission and relaxation processes. The shaded area indicates the typical response time of a PA resonator equipped with a microphone. The thick line represents the acoustic transit time. The wavy lines depict the radiative emission and the horizontal lines the range of relaxation processes characterized by the relaxation time $\tau$.

## Acoustic and thermal wave generation in the sample material

Sound and thermal wave generation can be theoretically described by classical disciplines of physics such as fluid mechanics and thermodynamics. The governing physical laws are the energy, momentum and mass conservation laws, given in the form of the heat-diffusion, Navier–Stokes and continuity equations, respectively. The physical quantities characterizing the PA and photothermal (PT) processes are the temperature $T$, pressure $P$ (mechanical stress in case of solids), density $\rho$ and three components of the particle velocity vector $\boldsymbol{v}$. As five equations are not enough to determine the above six quantities, a sixth equation is added, the thermodynamic equation of state in the form of $\rho = \rho\ (P,\ T)$. As the changes of $\rho$, $P$ and $T$ induced by light absorption are usually very small compared to their equilibrium values, the equations can be linearized by introducing the deviations from the equilibrium values as new variables, by neglecting the products of the small variables and by regarding the equilibrium values as constants. Moreover, the velocity vector $\boldsymbol{v}$ can always be separated into two components $\boldsymbol{u}$ and $\boldsymbol{w}$, where curl $\boldsymbol{u} = 0$ and div $\boldsymbol{w} = 0$. As the heat-diffusion and continuity equations are coupled only by the nonrotational component $\boldsymbol{u}$ to the Navier–Stokes equation, the $\boldsymbol{w}$ component of the particle velocity can be omitted. The governing equations may be written then as follows:

$$\frac{\partial \theta}{\partial t} + \frac{\gamma - 1}{\beta} \operatorname{div} \boldsymbol{u} - \kappa_{\mathrm{V}}\ \nabla^2 \theta = \frac{H(\boldsymbol{r}, t)}{\rho C_{\mathrm{V}}} \qquad [1]$$

$$\frac{\partial}{\partial t}\left(p - \frac{\beta}{K_{\mathrm{T}}}\theta\right) + \frac{\operatorname{div} \boldsymbol{u}}{K_{\mathrm{T}}} = 0 \qquad [2]$$

$$\rho \frac{\partial \boldsymbol{u}}{\partial t} + \operatorname{grad} p - \frac{4\eta}{3}\ \nabla^2 \boldsymbol{u} = 0 \qquad [3]$$

where $\theta = T - T_0$, $p = P - P_0$, $\gamma$, $\beta$, $\kappa_{\mathrm{V}}$, $C_{\mathrm{V}}$, $H$, $K_{\mathrm{T}}$ and $\eta$ are the new temperature and pressure variables, the

adiabatic coefficient, the heat expansion coefficient, the thermal diffusivity and heat capacity at constant volume, the density of the deposited heat power, the isothermal compressibility and the dynamic viscosity, respectively. For solid materials the Navier–Stokes equation is replaced by the wave equation of the longitudinal elastic waves.

The heat power density $H$, released in the material as the result of all nonradiative de-excitation processes, appears as the source term on the right hand side of the heat-diffusion equation. The spatial size and shape of the source volume depend on the light-beam geometry and on the absorption length in the material. Similarly, the time dependence of the heat source is determined by the time evolution of the light excitation and by the relaxation processes in the material. A photoacoustic effect can be generated by modulated radiation as well as by pulsed radiation. As the theoretical treatments of the two cases are different, they will be discussed separately.

### Modulated PAS

In this case the intensity (or the wavelength) of the incident light beam is modulated by an angular frequency $\omega$ to generate the acoustic signal. As the modulation frequency is usually in the audio frequency range, the time delay between heat release and light intensity may be negligible. In this case the source term has the same time dependence as the light intensity. Assuming an $\exp(\mathrm{i}(\omega t - \boldsymbol{kr}))$ dependence of the variables, $\theta$, $p$ and $\boldsymbol{u}$, two independent plane wave solutions of Equations [1] and [3] can be derived: a thermal wave and a sound wave. The wave-lengths of the two plane waves can be determined from the corresponding eigenvalues of the wave vector $\boldsymbol{k}$, taking into account that the length $|\boldsymbol{k}| = k$ of the wave vector (called wavenumber) is inversely proportional to the wavelength $\lambda$ ($k = 2\pi/\lambda$). The eigenvalues of the wavenumber for the thermal and sound wave may be given as $k^2_{\mathrm{th}} \cong -\mathrm{i}\omega/\kappa_{\mathrm{P}}$ and $k^2_{\mathrm{s}} \cong \omega^2/c^2(1 - \mathrm{i}\omega\nu/c^2)$, respectively, where $\kappa_{\mathrm{P}}$, $\nu = 4\eta/3\rho$ and $c$ are thermal diffusivity at constant pressure, the effective kinematic viscosity and the sound velocity, respectively. As the orders of magnitude of $\kappa_{\mathrm{P}}$ and $c$ are $10^{-5}$–$10^{-7}$ m$^2$ s$^{-1}$ and $10^2$–$10^4$ m s$^{-1}$ respectively, the wavelength of the thermal wave is much shorter than that of the sound wave. That is, two types of waves are simultaneously generated: a very strongly damped thermal wave with submillimetre wavelength, and a slightly damped sound wave with wavelengths in the centimetre to metre range. The thermal wave corresponds more or less to an isobaric thermal expansion, i.e. the changes of the temperature and density are much larger than that of the pressure. Because of the large damping coefficient, this wave cannot propagate far away from the heated region; it appears only in the neighbourhood of the exciting light beam. In the sound wave a quasi-adiabatic state change propagates with the velocity of sound; here the orders of magnitude of the relative changes in pressure, temperature and density are the same. The amplitude of the periodic pressure change is proportional to the time-varying (AC) component of the released heat power density and inversely proportional to the modulation frequency.

As the average of the intensity of a modulated light beam is nonzero, the heat energy in the illuminated volume will rise continuously. Therefore, the temperature will slowly increase and the density will decrease until the heat deposition rate is equal to the loss rate due to heat conduction. This process is also governed by the heat-diffusion equation. For a closed cell the average density is constant; therefore a pressure rise will occur. This DC component of the released heat power density changes the thermodynamic state of the material, in particular in very small PA detectors ($\approx$ cm$^3$).

The amplitude of the periodic temperature change ($\theta$) is also proportional to the AC component of the heat power density and inversely proportional to the modulation frequency. This means that the PA and PT effects are inherently coupled; the heat deposited by the interaction of radiation with the material generates a localized temperature rise, a thermal wave and a propagating sound wave. The first two effects will result in a periodically pulsating temperature distribution.

### Pulsed PAS

In the case of pulsed excitation, the absorption of photons ceases when the laser pulse is over, but the relaxation processes will continue until the full thermalization of the absorbed energy is achieved. Therefore, the duration of the thermal pulse is always longer than that of the light pulse. Nevertheless, the heat pulse may be regarded as instantaneous acoustic excitation if the characteristic time of the acoustic event is much larger than the duration of the heat pulse. Two quantities characterize the acoustic process, namely the transit time $\tau_{\mathrm{s}}$ of the sound through the heated volume and the response time of the PA detector. The value of $\tau_{\mathrm{s}}$ is usually in the microsecond range (**Figure 2**), because the laser beam diameter usually does not exceed a few millimetres in PAS. The PA response time depends on the period of the eigenmodes of the detector and is usually in the millisecond range (**Figure 2**). As the pulse duration of most pulsed lasers is in the nanosecond range, and in many cases, such as V-V and V-T

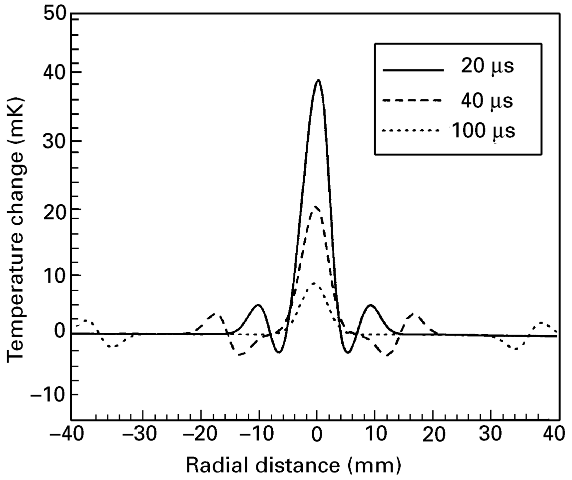



**Figure 3** Modelling of the gas temperature distribution as a function of the radial distance from the light excitation source. The diameter of the light beam was 6 mm and the pulse duration 20 ns. The three curves represent three different time delays after illumination.

relaxation in gases, the relaxation time is also in the nanosecond to microsecond range, the source term of Equation [1] may be regarded as a Dirac-delta pulse. The governing equations (Eqns [1] to [3]) can be solved by the Green's function technique, taking into account the Green's functions of both the thermal and acoustic problems. The solution is composed of a slowly broadening quasi-Gaussian temperature distribution and an outward propagating sound pulse of duration $\geq 2\tau_s$. After a short time the sound pulse will be separated from the thermal distribution, allowing the measurement of both features separately (**Figure 3**). The spatial symmetry of the solution depends on the shape and size of the heated region. In strong absorbers, the heated region is usually small compared to the sound wavelength; therefore mostly spherical sound waves are produced. In weakly absorbing gases, cylindrical acoustic waves are generated. The outward-propagating primary wavefront can be detected by appropriate high-frequency pressure sensors. As the achievable sensitivity and signal-to-noise (S/N) ratio are small, the direct detection of the primary waves has no practical significance. In practice the acoustic excitation takes place inside a PA detector, and the time evolution of the acoustic signal is strongly influenced by the properties of this detector.

## Determination of the PA signal in a PA detector

The actual solutions of the governing equations (Eqns [1] to [3]) depend on the boundary conditions determined by the type and geometry of the PA detector. In gas-phase photoacoustics, the PA detector consists of a cavity and a microphone to monitor the acoustic signal. From an acoustic point of view, the PA detector is a linear acoustic system that responds in a characteristic way to an excitation. The acoustic properties of the PA detector can be determined independently using acoustic modelling techniques or by measuring them in an acoustic laboratory. Once the acoustic properties are known, the response of the PA detector for any kind of PA excitation can be determined by calculation. Although PA detectors can usually be used with both modulated and pulsed excitation, the theoretical description of the two cases will be presented separately.

### PA detectors excited by modulated light

A simple PA detector consists of a cavity and a microphone to monitor the acoustic signal (**Figure 4A**). Even such a simple system has acoustic resonances. If the modulation frequency is much smaller than the lowest resonance, the PA detector or PA cell operates in a nonresonant mode. Such a PA cell is frequently called a 'nonresonant' cell in the literature. In this case the sound wavelength is much larger than the cell dimensions and thus the sound cannot propagate. The average pressure in the detector will oscillate with the modulation frequency. The amplitude of the oscillation may be determined by integrating Equations [1] to [3] over the volume of the detector. The pressure amplitude will be inversely proportional to the volume of the cell. The photoacoustically generated pressure can be approximated by the expression

$$p(\omega) = \frac{(\gamma - 1)\alpha l W_L}{i\omega V_{cell}} \qquad [4]$$

where $\alpha$, $l$, $W_L$, $\omega$, $V_{cell}$ and $\gamma$ denote the absorption coefficient of the material at the light pass length, the incident light power, the modulation frequency, the cell volume and the adiabatic coefficient of the material, respectively. In the case of a small cell and low modulation frequency, the signal may be quite large. The PA signal has a 90° phase lag with respect to the light intensity. Unfortunately, the noise also increases with decreasing frequency and volume; thus the S/N ratio will usually decrease.

As mentioned, the acoustic and thermal processes are inherently coupled. Until the cell dimensions are much larger than the size of the pulsating thermal distribution, the simple model described above can be applied. The thermal wave is usually not completely

damped at the walls of the cell in the case of a very small cell and very low modulation frequency. Then both the average temperature and the amplitude of the temperature oscillation will be influenced by the heat conduction through the cell walls to the environment. As in a closed cell the pressure is proportional to the temperature, the PA signal will also depend on the heat conduction through the walls. Since the theoretical modelling of this effect is practically impossible, such small cells and very low modulation frequencies allow only a qualitative signal analysis.

A special case should be mentioned here, one of the oldest arrangements among PA detectors. Termed the 'gas-microphone cell' it is used for investigating samples of condensed materials (**Figure 4B**). It usually consists of a small cylinder equipped with a sample holder, a microphone and a window. The gas in the cell is nonabsorbing; only the sample and the backing material (in the case of an optically thin sample) absorb the incident radiation. The PA signal is produced in an indirect way; the thermal waves generated in the solid sample are transmitted to the gas, and the periodic heat expansion of the gas above the sample surface acts as a piston and produces pressure oscillations in the closed cell volume. As the penetration depth of the thermal wave is usually much smaller than the diameter of the light beam, one-dimensional theoretical modelling is possible.

In resonant photacoustics, an acoustic resonator with an optimized geometry such as a cylinder or sphere is used as the gas cell. Such an acoustic system has several eigenresonances, whose frequencies depend on the geometry and size of the cavity. For a lossless cylinder, the resonance frequencies of the different eigenmodes can be calculated from the equation

$$f_{n,m,n_z} = \frac{c}{2}\sqrt{\left(\frac{\chi_{m,n}}{R}\right)^2 + \left(\frac{n_z}{L}\right)^2} \qquad [5]$$

where $c$, $R$ and $L$ are the sound velocity of the gas in the cavity, the radius and the length of the cylinder, respectively. The indices $m$, $n$ and $n_z$ may take the values 0, 1, 2, etc. The quantity $\chi_{m,n}$ is the $n$th zero of the derivative of the $m$th-order Bessel function divided by $\pi$.

When only one index is nonzero, the eigenmodes separate into longitudinal ($n_z \neq 0$), radial ($n \neq 0$) and azimuthal ($m \neq 0$) modes (**Figure 5**). In the other mixed eigenmodes the spatial distribution of the sound pressure is much more complicated.

The modulation frequency may be tuned to one of the eigenresonances of the PA detector. Such 'resonant PA cells' or detectors have found widespread application in PA trace gas measurements. Since the eigenmodes of a lossless cavity are orthogonal, the sound pressure field in a lossy cavity can be approximated in the form of a series expansion of the eigenmodes, where the amplitudes of the terms depend on frequency. In fact, each amplitude function is a resonance curve characterized by the corresponding resonance frequency and quality factor ($Q$ factor). The first term of the series expansion determines the sound pressure below the lowest resonance. To aid understanding of the behaviour of a resonant PA cell a computer simulation is shown in **Figure 6**, for which the frequency dependence of the PA signal at one end of a closed cylinder filled with a strongly absorbing gas was calculated. It can be seen that several resonances (the odd-numbered longitudinal ones) lie on a nonresonant curve. For well-separated sharp resonances, only two terms of the series expansion are necessary to determine the sound pressure around a certain resonance. These are the first nonresonant term, which is independent of the spatial coordinates but depends inversely on the frequency, and the resonant term, corresponding to the selected eigenmode. This term depends on the overlap of the acoustic mode pattern and the light beam distribution, on the position of the microphone and on the

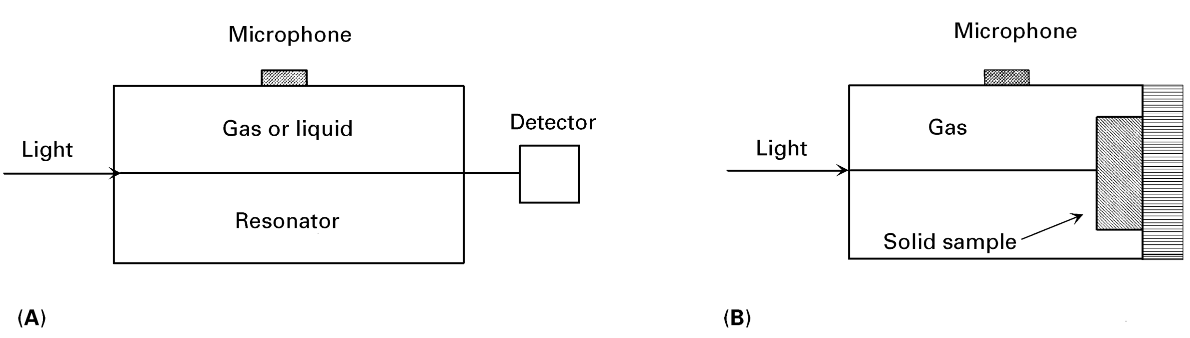



**Figure 4** PA setups for monitoring the PA signal with a microphone in a resonator: (A) gas or liquid, (B) solid sample in a gas microphone cell.

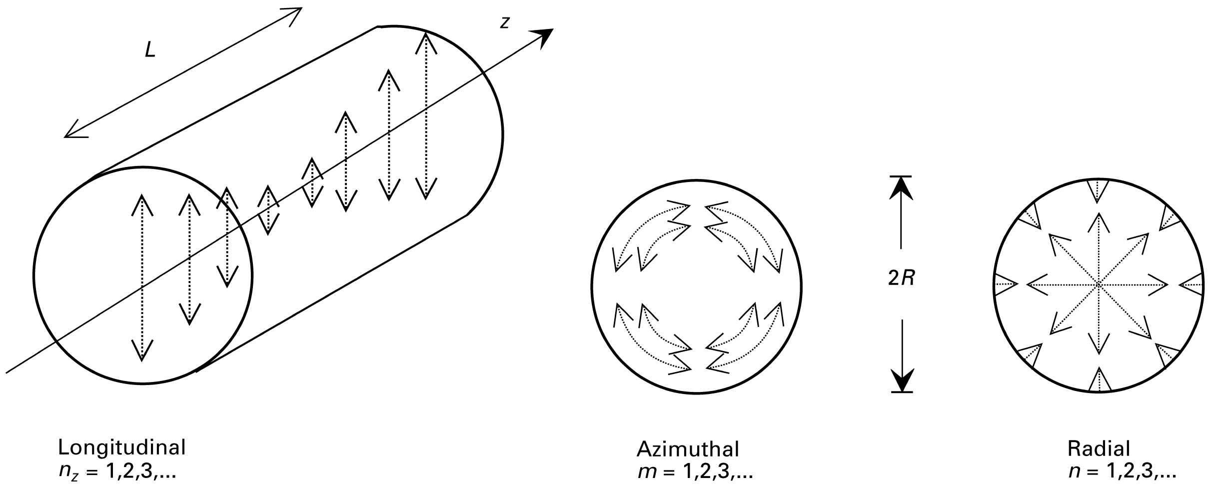



**Figure 5** Schematic representation of the longitudinal, azimuthal and radial acoustic modes in a cylindrical resonator.

quality factor of the eigenresonance. In the case of high $Q$ factors ($Q > 100$), the contribution of the nonresonant term can be neglected and the PA signal amplitude at the resonance frequency $\omega_j$ and at the position $\boldsymbol{r}_{\mathrm{M}}$ of the microphone can be calculated as

$$p(\boldsymbol{r}_{\mathrm{M}}, \omega_j) = \frac{(\gamma - 1)\, l\, U_j\, p_j(\boldsymbol{r}_{\mathrm{M}}) Q_j}{V_{\mathrm{cell}}\, \omega_j\, D_j} \alpha W_{\mathrm{L}} \quad [6]$$

where $l$, $p_j$, $U_j$, $Q_j$, $D_j$, $V_{\mathrm{cell}}$, $\alpha$ and $W_{\mathrm{L}}$ are the length of the light path within the cell, the pressure distribution of the $j$th eigenmode of the cell, the overlap integral of the light intensity distribution with $p_j$, the $Q$ factor and the normalization factor of the $j$th eigenmode, the cell volume, the absorption coefficient and the incident light power, respectively. As the quantities in the first term of Equation [6] are independent of the light power and the absorption coefficient, the first term may be regarded as a characteristic setup quantity, called the 'cell constant', of the PA arrangement.

Since the resonance frequency depends on the speed of sound, a temperature drift of the gas inside the cavity causes a shift of all resonance frequencies. In modulated PAS this problem can be solved by temperature stabilization, by continuous monitoring of the gas temperature or by synchronizing the modulation frequency to the resonance peak using appropriate electronics.

As the modulation frequency of a CW light source may be arbitrary, a given PA detector can operate in both the nonresonant and resonant modes. The main advantage of resonant operation is the amplification of the PA signal by the $Q$ factor of the resonator if the modulation frequency of the incoming light is properly tuned to the selected acoustic resonance ('acoustic amplifier').

If the modulation frequency is so low that the corresponding wavelength is much longer than the dimensions of the detector, Equation [3] may be used to calculate the PA signal. Nonresonant operation is also possible by tuning the modulation frequency away from a resonance. In such cases the PA signal can be estimated by substituting $Q = 1$ into Equation [6], but a better accuracy may be achieved by taking into account the nonresonant term and at least the neighbouring eigenmodes in the series expansion.

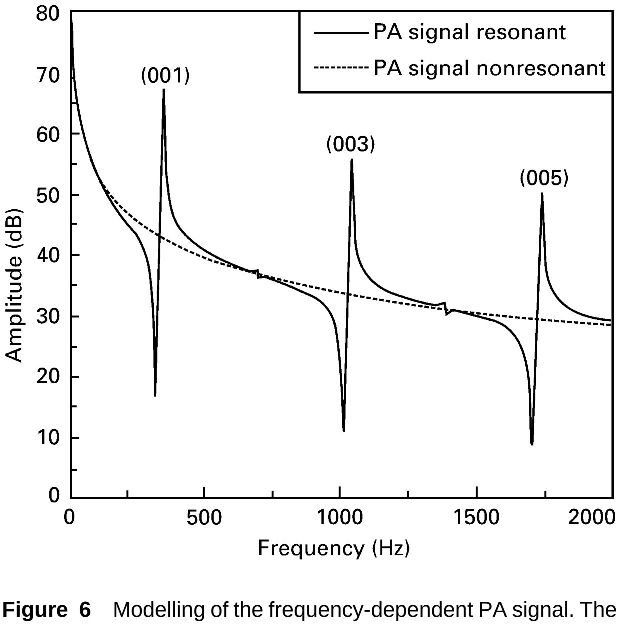



**Figure 6** Modelling of the frequency-dependent PA signal. The excited modes are the first (001), third (003) and fifth (005) longitudinal modes of the cavity. The dotted line shows the nonresonant contribution.

### PA detectors excited by light pulses

The absorption of a light pulse generates a primary acoustic pulse inside the PA detector. This pulse acts as a broadband acoustic source for the PA detector, exciting all eigenmodes simultaneously. If the acoustic transit time is much shorter than the period of the detected eigenmode of the resonator, the excitation of the PA signal can be regarded as instantaneous (**Figure 2**). In this case the slow PA detector responds to the excitation similarly to the way in which a ballistic pendulum or galvanometer responds to force or charge pulses; a sudden rise of the PA signal followed by a slow decay can be observed. The amplitude of the first period of the sound pressure oscillation will be proportional to the released heat energy. Thus, a pulsed PA detector can be used for absolute (calorimetric) measurement of that part of the absorbed light energy that was converted to heat. As calibrated microphones are available, the absolute measurement of the photoacoustically generated sound pressure is possible. The solutions of Equations [1] to [3] can again be given in the form of a series expansion of the orthogonal eigenmodes, but in this case the amplitudes will depend not only on the frequency but also on time. A PA cell for pulsed operation is designed for optimal excitation of a selected eigenmode. This mode should be well separated from the neighbouring ones and should have a high $Q$ factor. As the PA response in the time domain shows a very complicated behaviour, it is much better to evaluate the PA signal by converting the time signal to the frequency domain using Fourier transformation. A part of the frequency spectrum measured in a cylindrical cell, optimized for the first radial mode, is shown in **Figure 7**. The PA signal amplitude at the peak of the selected resonance can be determined from the theory as

$$p(\boldsymbol{r}_{\mathrm{M}}) = \frac{(\gamma - 1)\, lU_j\, p(\boldsymbol{r}_{\mathrm{M}})}{V_{\mathrm{cell}}\, D_j} \alpha E_{\mathrm{L}} \qquad [7]$$

where $E_{\mathrm{L}}$ is the pulse energy. As the fast Fourier transform (FFT) algorithm applied for calculating the spectrum delivers the average spectrum of the signal over the recorded time window, the amplitude of the resonance peak has to be corrected in order to obtain the value of $p(\boldsymbol{r}_{\mathrm{M}})$. The ratio of the corrected PA amplitude and laser pulse energy depends only on the product of several geometry factors and the absorption coefficient $\alpha$ of the absorbing component. In contrast to Equation [6], the PA signal amplitude $p(\boldsymbol{r}_{\mathrm{M}})$ in Equation [7] does not depend on the $Q$ factor of the cavity, which cannot be calculated with high accuracy theoretically. Thus, pulsed PAS is an absolute method for measuring the absorption coefficient. Since $\alpha$ is given as the product of the number density $N$ and the absorption cross section $\sigma$ of the absorbing molecules, pulsed PAS can be applied for both spectroscopic studies (known $N$) and trace gas analysis (known $\sigma$).

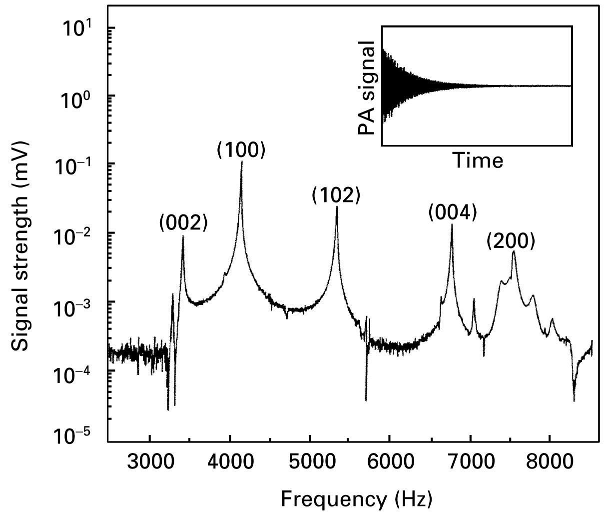



**Figure 7** Time-dependent PA signal recorded by a microphone (inset) and corresponding frequency spectrum of a cylindrical resonator. In the displayed frequency range, the second longitudinal (002), first radial (100), combination (102), fourth longitudinal (004) and second radial (200) modes of the resonator are detected.

## Summary

The theory of PAS is sufficiently complicated that only an overview could be presented. In the theoretical description of a given PAS experiment, a separated treatment of the three physical processes, as presented here, is often not possible. Moreover, important quantities of the theory, such as the $Q$ factor and the overlap integral $U_j$, are very difficult to keep under control. Small changes in the experimental adjustment (e.g. microphone position or beam focusing) may cause considerable changes, so that the agreement between theory and experiment may be degraded. As the PA signal does not depend on the $Q$ factor in pulsed PAS, and the $Q$ factor, which is needed only for calculating the correction factor, can be derived from the measured spectrum, pulsed PAS is more suitable for quantitative measurements than is modulated PAS. On the other hand, the probability of optical saturation is quite high in pulsed PAS, since the pulse energy of the available laser sources is usually in the millijoule range and the corresponding instantaneous power in the kilowatt to megawatt range. Therefore,

the linear dependence of the PA signal on the pulse energy must always be checked in pulsed PAS.

## List of symbols

$c$ = sound velocity; $C_V$ = heat capacity at constant volume; $D_j$ = normalization factor of $j$th eigenmode; $E_L$ = pulse energy; $H$ = heat power density; $k = |\boldsymbol{k}|$; $\boldsymbol{k}$ = wave vector; $K_T$ = isothermal compressibility; $L$ = cavity length; $l$ = light path length; $P$ = pressure; $p = P–P_0$; $Q_j = Q$ factor of $j$th eigenmode; $\boldsymbol{r}$ = position vector; $R$ = cavity radius; $T$ = temperature; $U$ = overlap integral; $\boldsymbol{u}$, $\boldsymbol{v}$ = particle velocity vector; $W_L$ = incident light power; $\alpha$ = absorption coefficient; $\beta$ = thermal expansion coefficient; $\gamma$ = adiabatic coefficient; $\eta$ = dynamic viscosity; $\theta = T - T_0$; $\kappa_V$, $\kappa_P$ = thermal diffusivity at constant volume, pressure; $\lambda$ = wavelength; $\nu$ = effective kinematic viscosity = $4\eta/3\rho$; $\rho$ = density; $\tau_s$ = transit time of sound; $\chi_{m,n}$ = the $n$th zero of the derivative of the $m$th-order Bessel function × (1/π); $\omega$ = angular frequency of modulation.

*See also:* **Laser Spectroscopy Theory; Light Sources and Optics; Photoacoustic Spectroscopy, Applications.**

## Further reading

Diebold GJ (1989) Application of the photoacoustic effect to studies of gas phase chemical kinetics. In: Hess P (ed) *Photoacoustic, Photothermal and Photochemical Processes in Gases*, pp 125–170, Vol 46 in Topics in Current Physics. Berlin: Springer.

Fiedler M and Hess P (1989) Laser excitation of acoustic modes in cylindrical and spherical resonators: theory and applications. In: Hess P (ed) *Photoacoustic, Photothermal and Photochemical Processes in Gases*, pp 85–121, Vol 46 in Topics in Current Physics. Berlin: Springer.

Hess P (1992) Principles of photoacoustic and photothermal detection in gases. In: Mandelis A (ed) *Progress in Photothermal and Photoacoustic Science and Technology*, Vol 1, pp 153–204. New York: Elsevier.

Morse PM and Ingard KU (1986) *Theoretical Acoustics*. Princeton, NJ: Princeton University Press.

Pao YH (1977) *Optoacoustic Spectroscopy and Detection*. New York: Academic Press.

Rosencwaig A (1980) *Photoacoustics and Photoacoustic Spectrosopy*. New York: Wiley.

Schäfer S, Miklós A and Hess P (1997) Pulsed laser resonant photoacoustics/applications to trace gas analysis. In: Mandelis A and Hess P (eds) *Progress in Photothermal and Photoacoustic Science and Technology*, Vol 3, pp 254–289. Bellingham, WA: SPIE.

Schäfer S, Miklós A and Hess P (1997) Quantitative signal analysis in pulsed resonant photoacoustics. *Applied Optics* **36**: 3202–3211.

West GA, Barrett JJ, Siebert DR and Reddy KV (1983) Photoacoustic spectroscopy. *Review of Scientific Instruments* **54**: 797–817.

Zharov VP and Letokhov VS (1986) *Laser Optoacoustic Spectroscopy*. Berlin: Springer.

# Photoelectron Spectrometers

**László Szepes** and **György Tarczay**,
Eötvös University, Budapest, Hungary

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Methods & Instrumentation**

This article is centred on one of the most important molecular spectroscopic applications of photoionization. First, conventional photoelectron spectrometers based on the pioneering work of Turner, Vilesov and Siegbahn are reviewed, the basic building elements are shown and the principles of operation are discussed. Besides the conventional photoelectron experiment, threshold analysis and photoelectron-photoion coincidence spectroscopy as well as conceptually new techniques related to laser photoionization are discussed briefly.

## Conventional photoelectron spectrometers

Photoelectron spectroscopy is a molecular spectroscopic method which is based on photoionization. If an atom or molecule (M) is irradiated with photons of larger energy than the ionization energy of the particle ionization may occur. The fundamental process and its energetics are given by the following equations:

$$\mathrm{M} + h\nu = \mathrm{M}_i^+ + \mathrm{e}^- \qquad [1]$$

$$h\nu = \mathrm{IE}_i + \mathrm{KE}_i \qquad [2]$$

where $\mathrm{IE}_i$ is the $i$-th ionization energy of the system studied and $\mathrm{KE}_i$ is the kinetic energy of the ejected electron. Complementary to mass spectrometry, where the fate of the ion is investigated, in the case of photoelectron spectroscopy the other ‘reaction product’, the electron, is the source of the information. As a consequence, photoelectron spectroscopy normally requires a monoenergetic radiation source, a sample inlet system, a target chamber where photon–atom/molecule interaction occurs, an electron kinetic energy analyser, a detector and a recording system. A further requirement comes from the fact that electrons lose energy through collisions with gas molecules, thus electrons can only be handled in high vacuum where the mean free path is of the same order as the characteristic geometric dimension of the spectrometer. This is normally achieved at a pressure equal to or less than $10^{-5}$ mbar ($10^{-3}$ Pa). A block diagram of a typical experimental arrangement is shown in **Figure 1**.

Two types of photoelectron spectrometers are distinguished on the basis of the energy of the radiation sources. Energy needed to eject electrons from the valence shell (above 6 or so eV) corresponds to photon wavelengths in the vacuum ultraviolet (VUV) region of the electromagnetic spectrum. This branch of instruments is known as VUV photoelectron spectrometers, normally abbreviated as UPS. More energetic X-ray photons, commonly in the range of 1000–2000 eV, are used for core electron ionizations. This type of photoelectron experiment is called X-ray photoelectron spectroscopy, XPS, also named ESCA (electron spectroscopy for chemical analysis).

As far as operation is concerned there is no principle difference between the above two methods. However, the diverse fields of application – gas-phase, molecular studies on the valence shell versus solid-phase surface studies based on core electron ionizations – make sense of this distinction.

The term ‘conventional photoelectron spectrometers’ designates those UPS instruments which are technically, based on the invention of Turner and Vilesov and operate with an energy resolution between 5 and 30 meV. As far as the main building units are concerned, an XPS spectrometer has a similar experimental arrangement, but lower energy resolution (0.5–1 eV) is obviously achieved.

## Light sources

A photon source used in photoelectron (PE) spectroscopy must fulfil at least two basic requirements. First, the incident radiation must be monochromatic; the second requirement is high photon intensity. In order to achieve sufficient electron flux at the detector a source with a typical intensity of $10^{10}$–$10^{12}$ photons $s^{-1}$ is needed. The most commonly used source in the XPS technique is the X-ray tube, while in valence shell PE spectroscopy resonance lamps are applied.

In the X-ray tube, radiation is generated by the bombardment of the anode with high-energy (10–15 keV) electrons. The schematic view of the widely used dual anode source is seen in **Figure 2**. In this set-up the two filaments, which are at ground potential, can be heated separately. Electrons from the hot cathode hit the corresponding positive anode face. The emerging characteristic X-ray radiation leaves the source through a thin aluminium window. As the two anode faces are covered by two different metal layers, the photon energy can be changed without a break. There are many metals which can be used as anodes (see

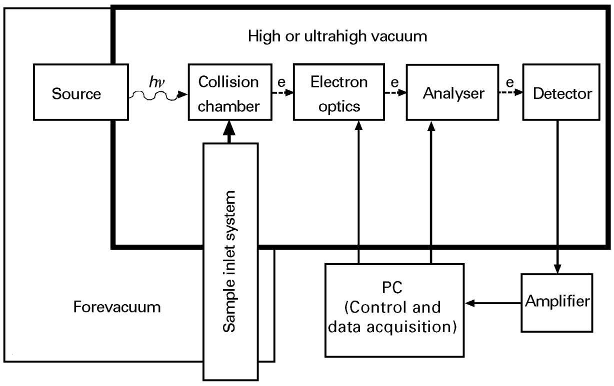



**Figure 1** Block diagram of a photoelectron spectrometer.

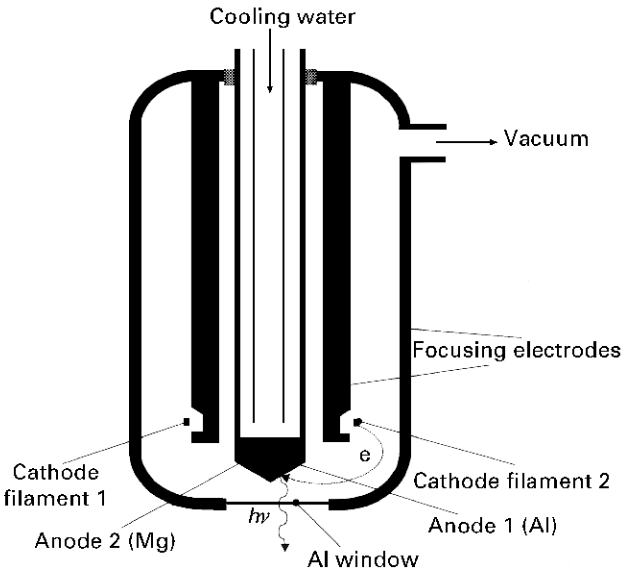



**Figure 2** A dual anode X-ray tube.

Table 1) but in most cases Mg or Al are used due to their narrow line width. With the use of a monochromator it is possible to filter out the satellite lines as well as the continuous radiation (the Bremsstrahlung) and to decrease the line half width to about 0.2 eV.

Ultraviolet resonance radiation may be produced by spark, direct current (d.c.) or microwave/radio frequency discharge through a rare gas. The most common type in use is the low-pressure helium d.c. discharge lamp (**Figure 3**). The discharge is initiated by applying a high voltage (~5 kV) across the discharge capillary. After the ignition period the voltage falls to 600–700 V and the gas pressure can be optimized for the generation of He(I) (21.21 eV), or He(II) radiation (40.81 eV). (The Roman numerals denote the emitting species which can be a neutral atom (I) or a singly (II) or doubly (III) ionized atom). Since there is no transmitting material in this energy region, the beam has to pass from the cathode area into the collision chamber through a windowless collimating capillary. To prevent self-absorption and to keep the gas out of the high vacuum side, differential pumping of the lamp is necessary. The line half width of this radiation is about 1 meV.

Besides helium some other noble (or atomic) gases can also be used. Their resonance energies are summarized in **Table 2**.

**Table 1** Some X-ray lines used in X-ray photoelectron spectroscopy

| *Anode* | *Transition* | *Energy* (eV) | *Line width* (eV) |
|---|---|---|---|
| Be | K | 108.9 | 5.0 |
| Ni | $L_\alpha$ | 851.5 | 2.5 |
| Cu | $L_\alpha$ | 929.7 | 3.8 |
| Mg | $K_\alpha$ | 1253.6 | 0.7 |
| Al | $K_\alpha$ | 1486.6 | 0.85 |
| Zr | $L_\alpha$ | 2042 | 1.7 |
| Ti | $K_\alpha$ | 4510 | 2.0 |
| Cr | $K_\alpha$ | 5417 | 2.1 |
| Cu | $K_\alpha$ | 8048 | 2.6 |

**Table 2** Resonance energies of some discharge lamps

| *Gas* | *Line* | *Energy* (eV) |
|---|---|---|
| He | He(I)$\alpha$ | 21.2175 |
| | He(I)$\beta$ | 23.0865 |
| | He(II)$\alpha$ | 40.8136 |
| Ne | Ne(I)$\alpha$ | 16.6704, 16.8476 |
| Ar | Ar(I)$\alpha$ | 11.6233, 11.8278 |
| Xe | Xe(I) | 8.4363, 9.5695 |
| H | Lyman$\alpha$ | 10.1986 |
| | Lyman$\beta$ | 12.0872 |
| | Lyman$\gamma$ | 12.7482 |

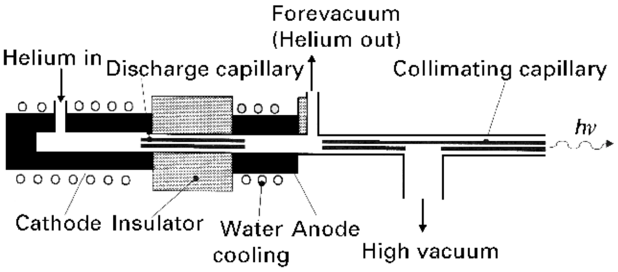



**Figure 3** Low-pressure helium d.c. discharge lamp.

For some special studies (autoionization, predissociation, etc.) it is favourable to use a tunable source, and it is indispensable for threshold PE spectrometer experiments. The simplest way to generate tunable radiation is the use of a many-line or a continuous radiation source with a monochromator. A many-line source is usually a low-pressure hydrogen lamp while the other can be continuous helium or synchrotron radiation. Synchrotron radiation is produced by electron accelerators or storage rings as a consequence of radial acceleration of the electrons in a magnetic field. The main advantage of synchrotron radiation is its high intensity and broad-range tunability.

## Sample inlet systems

Although both XPS and UPS techniques can be adapted for the investigation of either condensed or gas-phase samples, the former technique is almost exclusively applied for surface analysis while UPS is considered mainly as a gas-phase electron structure elucidation method.

The ideal solid sample is approximately 1 cm square with a mirror-smooth surface. Lumpy samples can also be measured directly while powders can be stuck onto a plastic tape or pressed into a pellet.

Regardless of the physical appearance of the solid sample, it is usually pretreated. The simplest treatment is washing the surface with organic solvents. The main drawback of this method is that the composition of the surface may change during the treatment: on the one hand, some ions may be washed out from the surface layer, on the other hand, the solvent, or even elemental carbon, may be adsorbed on the surface. These solvents, together with adsorbed water and gases can be removed in the instrument by baking the sample in ultrahigh vacuum conditions. The most sophisticated surface cleaning method is electron or ion bombardment, and laser vapourization can also be used. These techniques are also appropriate for the etching of the surface layer-by-layer, providing the possibility for depth analysis.

In the case of a gas-phase sample a large pressure gradient must be sustained between the interaction region and the rest of the system in order to obtain a sufficient electron current and to avoid the scattering of the electrons on the sample. The easiest way to achieve this gradient is the application of a collision chamber which may or may not be differentially pumped. In this setup the sample and the UV light pass into the chamber through a capillary or a bore and the electrons exit through a third slit directed toward the electron optics and the analyzer. The geometry of the chamber is not significant but its surface must be well cleaned, smoothed and uncharged. Local surface charges or the adsorption of dipolar molecules on the surface will shift the spectrum and decrease the resolution. These effects are most easily minimized by surface coatings of gold or colloidal graphite.

It is also possible to investigate less volatile solid samples in the gas phase. In this case the sample is evaporated close to (or in) the collision chamber. The sample can be heated by circulating hot liquid, by the heat of the discharge lamp or by a resistance oven. The temperature usually attainable by these methods is about 300°C. The temperature can be increased further up to 1000°C by special oven designs. In these constructions the disturbing magnetic and electric fields of the resistance heater are reduced by the use of noninductively wound and shielded wires. Even higher temperatures (2000–2500°C) can be achieved by laser vapourization.

Short-lived reaction products prepared *in situ* can also be studied. The sample molecules can undergo reaction shortly before ionization. Monomolecular reactions are commonly induced by UV photolysis or pyrolysis, while the feeding of highly reactive radicals into the sample vapour can cause bimolecular reactions. A well-tried variant of the latter method is the fluorine atom reaction where atoms generated from fluorine molecules in a microwave discharge outside the instrument are injected into the sample just before the ionization chamber. These methods qualify PE spectroscopy for the investigation of highly reactive molecules, transients and radicals not stable under normal conditions.

## Analysis of energetic electrons

Basically, there are three ways of measuring the kinetic energy of electrons. The first is to measure the time needed to traverse a known distance. In the second method a retarding potential is applied to the electrons to be analysed. The third method is based on the deflection of the electrons in an electrostatic or magnetic field.

The time-of-flight (TOF) of an energetic electron over a distance of a few centimetres is very short, consequently, electronics with a response time of the order of nanoseconds are needed. This sort of kinetic energy analysis has not gained widespread application in the conventional techniques.

In retarding-field analysers, only those electrons are permitted to reach the detector which have higher energy than the retarding potential. This type of energy analysis can be performed by placing a grid in front of the detector and varying its potential. Many variations of retarding-field analysers are known; cylindrical and spherical arrangements are the most common. Analysers based on this principle yield a stepwise integral energy distribution curve in which a rise in the detected electron current indicates a group of energetically different electrons. The energy distribution can be obtained by differentiating this plot of detector signal versus retarding potential. The advantages of this type of analyser are an almost equal transmission of electrons of all energies and a large acceptance angle and consequently high sensitivity. One drawback of this method is that the examination of low-energy electrons is sometimes difficult. Furthermore, any contamination of the grids may give rise to surface potential variations which result in the loss of resolution. Retarding-field analysers can be very easily constructed but problems related to their performance limit their use for high-resolution studies.

Finally, the energies of electrons can be determined by passing them through an electrostatic or magnetic field where the deflection of the electron paths is a function of their energy. Since it is generally easier to produce a uniform electric field than a uniform magnetic field, electrostatic analysers have come to predominate in photoelectron spectroscopy. Two types of dispersive analysers are generally distinguished. In deflectors, electrons follow equipotential lines during analysis, while in mirrors electrons cross equipotential lines on moving between two electrodes.

The following part is restricted to the description of some frequently used electrostatic analysers, namely the *radial cylindrical*, the *hemispherical* and the *cylindrical-mirror* analysers.

One of the most important parameters of an analyser is the energy resolution. It may be defined as a measure of the ability of an analyser to resolve two adjacent peaks in the spectrum separated by $\Delta E$. The resolution for deflection analysers at energy $E$ is given by an equation of the form

$$\Delta E/E = a\omega + b(\Delta\alpha)^2 + c(\Delta\beta)^2 \qquad [3]$$

where $a$, $b$ and $c$ are constants characteristic of the particular analyser; $\omega$ is the entrance and exit slit width; and $\Delta\alpha$ and $\Delta\beta$ are the angular deviation of the electron beam in the plane of deflection and the perpendicular plane, respectively.

**Figure 4** shows a 127° electrostatic cylindrical analyser, one of the most frequently used in gas-phase molecular photoelectron spectroscopy. Electrons produced in the ionization chamber pass through the entrance slit ($S_1$) and enter the analyser where they are deflected by a radial electric field produced by the electric potentials $V_1$ and $V_2$ placed on the concentric cylindrical electrodes. According to charged particle optics, focusing is achieved after an angular deflection of 127°. If $E$ is the energy of electrons following the central path transmitted through the analyser and is related to the electrical potential $V$ by $E = qV$, then the potentials to be applied to the outer and inner electrodes are

$$V_2 = V + 2V \ln(R_2/R) \quad [4]$$

and

$$V_1 = V + 2V \ln(R_1/R) \quad [5]$$

where $R_2$ and $R_1$ are the radii of the outer and the inner electrodes while $R$ is the mean radius.

The majority of commercial electron spectrometers – either UPS or XPS instruments – are equipped with some form of spherical deflector analyser. The 180° spherical sector – shown in **Figure 5** – is often used because of its compactness and good resolving power.

In order to transmit electrons of energy $E = qV$ along a circular path of radius $R$, electrical potentials applied to the outer ($V_2$) and inner ($V_1$) electrodes are given by

$$V_2 = V[2(R/R_2) - 1] \quad [6]$$

and

$$V_1 = V[2(R/R_1) - 1] \quad [7]$$

where $R_2$ is the radius of the outer hemisphere while $R_1$ is that of the inner one.

For a hemispherical deflector, the coefficient $c$ in the expression of resolution is equal to zero. This means that exact focusing is achieved in the plane perpendicular to the plane of deflection.

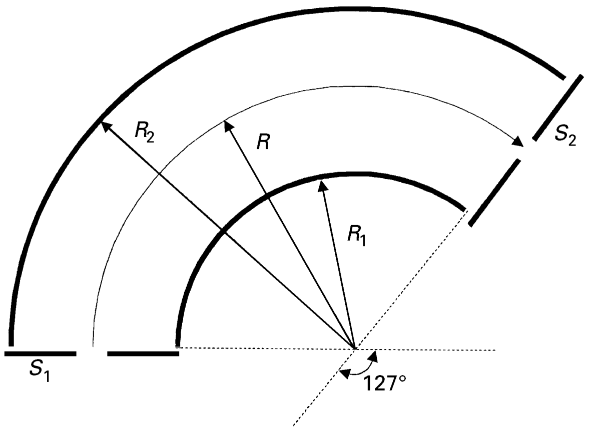



**Figure 4** The 127° radial cylindrical analyser $R_1$, $R$, $R_2$ are the radius of the inner electrode, the midradius, and the radius of the outer electrode, respectively, $S_1$ and $S_2$ are the entrance and exit slits, respectively.

The cylindrical mirror analyser (CMA) is shown in **Figure 6**. In this analyser, electrons emerging from the ionization region (located on the axis) pass through an annular slit into the space between two concentric cylinders. Two arrangements may be distinguished for this analyser system: axial focusing and slit-to-slit focusing. In the former case, electrons of energy $E$ are deflected so that after leaving the exit slit they are focused along the axis. Here, focusing occurs in both the deflection plane and the perpendicular plane. A source of modest size is expected for this design. Further parameters of operation are: the injection angle is 42.3°; the inner cylinder is earthed together with the ionization region, and the potential on the outer cylinder is

$$V = 1.3\, E \ln(R_2/R_1) \quad [8]$$

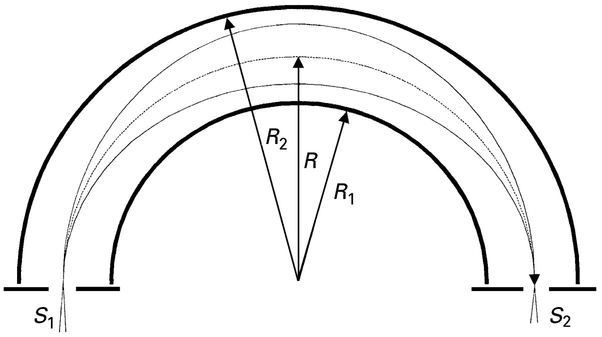



**Figure 5** The hemispherical analyser. $R_1$, $R$, $R_2$ are the radius of the inner electrode, the midradius, and the radius of the outer electrode, respectively. $S_1$ and $S_2$ are the entrance and exit slits, respectively.

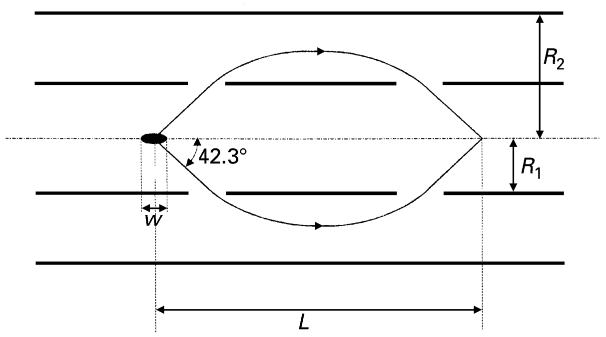



**Figure 6** The axial focusing cylindrical mirror analyser. $R_1$ and $R_2$ are the radius of the inner and the outer cylinder, respectively; $w$ is the axial extent of the source, $L$ is the source–detector distance.

The resolution is approximately

$$\Delta E/E = 1.09(w/L) \quad [9]$$

where $w$ is the axial extent of the source and $L$ is the distance from the source to the detector.

If the source is not small and well defined, slit-to-slit focusing can be chosen. In this arrangement energy-resolving slits in the inner electrode are used and the image occurs on the surface of the inner cylinder. This design yields large signals but focusing is obtained in the deflection plane only. Slit-to-slit focusing has been successfully used in high temperature UPS studies where small signals are expected.

CMA systems consisting of two axial-focusing analysers in series have been applied for XPS instruments. This combination gives an improved resolution at the cost of sensitivity.

In all electrostatic-deflection analysers the width of the bandpass $\Delta E$ is proportional to the transmitted energy $E$. That is why the absolute energy resolution can be improved by preretardation of the electrons prior to energy analysis. Most spectrometers employ this technique using appropriate electron optical elements.

### Detection and registration

An electron current of less than $10^{-14}$ A, which is typical in photoelectron spectroscopy, can be detected with an electron multiplier. The multipliers used earlier consisting of Cu–Be dynodes have now been replaced by channeltrons. This variant of the electron multipliers is a curved glass or ceramic tube representing a continuous dynode system based on its semiconducting inner surface. The resistance between the two ends of channeltron is about $10^9$ Ω and a potential of about 3 kV is applied across it in operation. The attainable gain is about $10^8$. Channeltrons have the advantage of small size, low cost and ruggedness. Multichannel plates operating by the same principle are also in use. In this case, electrons within a range of energies are collected simultaneously at the focal plane of a suitably constructed analyser.

Following detection, data are collected and processed by a PC.

### Operational considerations

In photoelectron spectroscopy one must compromise between energy resolution and the intensity of the detected signal ($S$) relative to the noise ($N$). (Both quantities expressed in counts per second.) The resolution may be enhanced at a lower signal count rate and vice versa. The maximum attainable signal-to-noise ratio is

$$(S/N)_{\max} = S^{1/2}t^{1/2} \quad [10]$$

where $t$ is the data collection time. The above relationship indicates that theoretically any desired resolution and signal-to-noise ratio can be achieved if sufficient data collection time is provided.

Practically, there are two modes of operation of PE spectrometers. The first is scanning the voltage $V$ applied to the analyser electrodes. In this case, the energy resolution, $\Delta E/E$, is constant and consequently peak widths vary across the spectrum. Alternatively, if the analyser voltages are fixed while the retarding potential is scanned by the electron optics, the peak width does not change throughout the spectrum. Finally, it has to be mentioned that an accurate electron energy cannot be determined from the voltage applied to the analyser because an energy shift will be present due to contact potentials within the analyser and other part of the spectrometer. Therefore, energy calibration by introducing an inert gas together with the sample is necessary for precise ionization energy determinations.

## Threshold photoelectron spectroscopy

As described above, conventional photoelectron spectrometers operate at fixed photon energy while the electron energy is scanned. If a continuous light source is available the photon energy may be scanned while electrons of essentially zero kinetic energy, called threshold electrons, are detected. This is the operational principle of threshold photoelectron spectroscopy (TPES). Threshold electron detection

uses a simple technique which discriminates strongly against energetic electrons while electrons formed with zero, or near zero kinetic energy are collected very effectively with small electric fields. In practice, threshold electrons are selected by a steradiancy analyser which consists of a stainless steel collimated hole system with a specific length-to-diameter ratio. During the analysis, threshold electrons are directed to pass through the tubes while those with appreciable energy are likely to hit the tube walls and be lost. Considering the fact that energetic electrons with initial velocity vector directed toward the detector are not discriminated by the analyser, the energy resolution using a discharge lamp is not better than 2–4 meV. This type of photoelectron spectroscopy is suited for experiments with synchrotron radiation sources as well as for photoelectron–photoion coincidence (PEPICO).

## Photoelectron–photoion coincidence spectroscopy

In VUV photoelectron spectroscopy, the kinetic energy spectrum of electrons is measured which provides information about the ionic states of atoms and molecules. However, very limited information is given about the fate of the excited species formed in the photoionization process. By the use of photoelectron–photoion coincidence (PEPICO) spectroscopy a detailed analysis of the dissociation of molecular ions is possible: the fragmentation of a particular ionic state – identified through the photoelectron energy – can be directly studied. So, if ions are detected in coincidence with electrons of a given kinetic energy, the unimolecular decay of molecular ions in selected energy states can be investigated. There are two types of PEPICO instruments. In one case a light source of fixed energy, usually He(I), is used together with a dispersive electron energy analyser. Two modes of operation of this experimental setup are feasible: (i) the mass spectrum is scanned at a fixed electron energy, or (ii) a particular ion is chosen and the PE spectrum in coincidence is measured. According to the other approach, a continuum light source in conjunction with a vacuum monochromator is used and threshold electrons are collected in coincidence with parent and fragment ions, so that the ion internal energy is given by $h\nu$–IE. In a typical PEPICO arrangement, which proved to be extremely powerful in the field of dissociation dynamics, threshold electrons not only serve to identify the formation of an ion of known internal energy but also provide a start signal for measuring the time-of-flight (TOF) of the ion. The ion TOF distribution contains all dynamical information such as the kinetic energy released in polyatomic dissociations, metastable ion lifetimes and collision dynamics of ion–molecule interactions.

## High-resolution photoionization techniques

Since the 1980s, some conceptually new photoionization techniques have appeared in the field of gas-phase investigations whose resolution is better than that of the conventional UPS technique by some orders of magnitude. Two technical novelties preceded the appearance of these techniques: the introduction of supersonic jets in photoelectron spectroscopy and the availability of high-energy lasers.

Supersonic jets have been used from the early 1950s in the fields of beam studies. According to this technique, the sample vapour is adiabatically expanded from several atmospheres into a high vacuum through a nozzle which has a diameter between 50 and 300 μm. As a result, a molecular beam with low translational (~1 K), rotational (~10 K) and vibrational (~100 K) temperatures is produced. This means that, according to the Boltzmann distribution, the non-ground rotational and vibrational states will be barely populated and the sample molecules are obtained in a well-defined energetic state. Feeding a noble gas to the sample beam as a carrier enhances this cooling effect. At an early stage this method was carried out by installation of huge pumps. Nowadays, the use of a pulsed nozzle is more common. In this case the nozzle opens repeatedly for some 100 μs, and the pump has to be large enough to evacuate the chamber between two pulses only. The diverging molecular beam can be collimated or skimmed. For this purpose, separate pumping of the additional chamber (between the nozzle and the skimmer) is also needed.

The next step toward the new techniques was the availability of high-power UV lasers. In a typical setup the UV laser beam is generated as follows: an excimer or a YAG laser produces monochromatic photons, the photon energy is then tuned by a dye laser and at the last stage the selected frequency may be doubled by a crystal (e.g. by β-barium borate, BBO). The largest photon energy attainable by this method is ~6.5 eV. Higher energies can be reached by frequency mixing in gaseous media. Four-wave mixing and the third harmonic generations are the two typical approaches to this problem. These relatively new techniques extend the laser photon energy into the VUV region up to 19 eV.

## Resonance enhanced multiphoton ionization

The resonance enhanced multiphoton ionization (REMPI) technique was originally developed as a sensitive method for the investigation of excited states. In the REMPI experiment a laser excites the molecule, then either the same laser (one-colour REMPI or 1+1 REMPI) or another with a different photon energy (two-colour or 1+1′ REMPI) ionizes the molecule from the excited state. If the photon density in the well-focused laser beam is high enough, nonresonant two-photon excitation may occur and then, similar to the previous case, the same laser, or a second one, can ionize the molecule (2+1 and 2+1′ REMPI). By scanning the wavelength of the first laser and measuring the amount of ions formed, information can be derived about the excited states.

Since excitation and ionization are usually carried out by pulsed lasers the obvious detection method is TOF-mass spectrometry. Formerly, linear TOF tubes were in use but they have now been replaced by reflectron TOF tubes. These tubes are folded at the middle and the ions are reflected backward by positively charged rings in the fold (see **Figure** 7). This type of TOF detection has two advantages. Firstly, ions of the same mass and charge but different kinetic energy will hit the detector at the same time. Secondly, the parent ions and the ions formed by fragmentation in the drift (or in the accelerating) region can be separated by tuning the reflectron voltage.

Beside ions, the detection of electrons is possible. If electrons are analysed according to their kinetic energy the result is a new type of photoelectron spectroscopy, called REMPI-PES. This technique usually uses TOF analysers. Prolongation of the flight time of the electrons, and consequently enhanced effectiveness of separation, can be achieved by the use of an external magnetic field.

Information about the ionic states can be gained more easily by scanning with the second (ionizing) laser and monitoring the total ion intensity. This method is called REMPI-photoionization efficiency (REMPI-PIE) spectroscopy. The REMPI-PIE spectrum is a staircase function, and the conventional photoelectron spectrum can be obtained from its differentiation.

## Zero kinetic energy spectroscopy and mass-analysed threshold ionization

In 1984 Müller-Dethlefs and Schlag pioneered a new laser photoionization method called zero kinetic energy (ZEKE) spectroscopy. The original idea was to detect only the strictly zero kinetic energy electrons in the following way. After the laser pulse the zero kinetic energy electrons remain in the ionization chamber for some microseconds. During this time the nonzero kinetic energy electrons have left this region. After the delay the zero kinetic energy electrons are extracted by a pulse of electric field (~0.1–1 V $cm^{-1}$) toward the flight tube and the detector (**Figure** 7). The latter is gated for the time of arrival of these electrons. However, practice has shown that an apparatus operating along these principles is unable to prevent the detection of electrons having a relatively small but still measurable kinetic energy (the so-called near-threshold electrons) which has a detrimental effect on the resolution. As it turned out, the mechanism described above – although it had been supposed since the late 1980s – only accurately describes

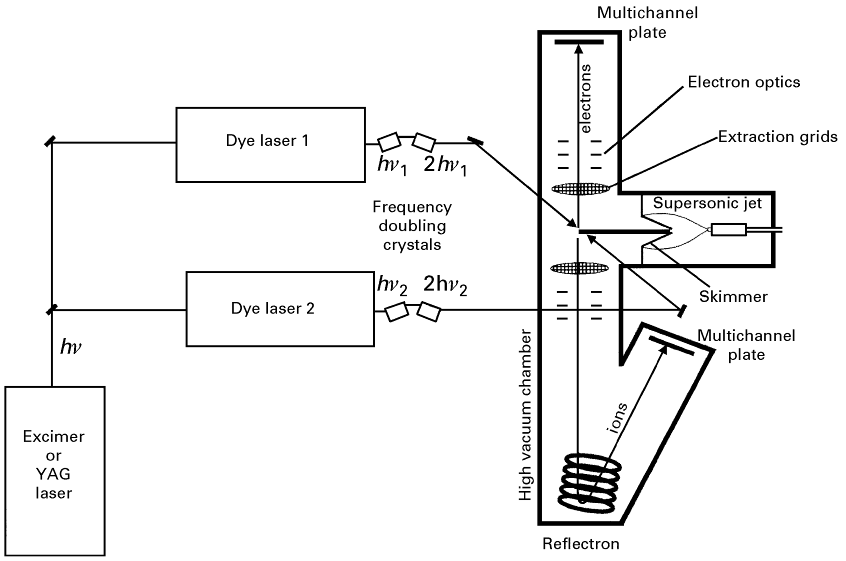



**Figure 7** Schematic illustration of an experimental setup capable of performing REMPI, ZEKE and MATI measurements.

the process related to the electron removal from anions (electron detachment spectroscopy). In 1988 Reiser showed that when neutral molecules are ionized, the applied pulsed field extracts the electrons from long-lived Rydberg states formed by laser excitation. (This is the so-called pulsed field ionization). Current ZEKE instruments, utilizing the recognition of this mechanism, employ a weak electric field pulse (usually opposite in sign to the extracting field) just after the laser pulse, followed by a delay of some microseconds and then the extracting pulse. The first pulse removes the near-threshold electrons and also those in the highest-lying Rydberg states. The main extracting pulse then ionizes the remaining molecules in next-to-highest states; therefore the measured ionization energy will be somewhat smaller than the real value. The exact ionization energies can be obtained by repeated measurements at different fields followed by extrapolation to zero field. More commonly, the energies are corrected using the following equation:

$$\delta E \approx 4\sqrt{F} \qquad [11]$$

where $F$ is the applied field. This method provides photoelectron spectra with a resolution of 0.1–0.5 $cm^{-1}$ (~0.05 meV).

The extension of this technique to ions is somewhat complicated. The difficulty is due to the higher mass of the ions. In order to distinguish between ions formed by the electric field pulse and those formed by the laser pulse, longer pulses are needed. This spectroscopy is called mass-analysed threshold ionization (MATI). The main benefit of this method is the possibility of recording the spectra of various fragment ions together with that of the parent ion.

Technically, laser excitation to Rydberg states can be performed either directly or, similarly to the REMPI technique, via an excited state, using two lasers. The advantage of the first method is that we need no prior information about the excited states. The advantage of the second method is that investigation of chemically impure samples can be carried out since one can select of the molecule of interest by varying the excitation laser energy.

## List of symbols

$E$ = energy; $I$ = extracting field; IE = ionization energy; KE = kinetic energy; $L$ = distance from source to detector; $N$ = noise; $R$ = mean radius; $R_1$ = radius of inner electrode; $R_2$ = radius of outer electrode; $S$ = signal; $V_1$ = potential applied to inner electrode; $V_2$ = potential applied to outer electrode; $w$ = axial extent of source; $\Delta E$ = resolution; $\Delta\alpha$ = angular deviation of electron beam in plane of deflection; $\Delta\beta$ = angular deviation of electron beam in perpendicular plane; $\omega$ = entrance and exit slit width.

*See also:* **Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Light Sources and Optics; Multiphoton Excitation in Mass Spectrometry; Multiphoton Spectroscopy, Applications; Optical Frequency Conversion; Pharmaceutical Applications of Atomic Spectroscopy; Photoelectron Spectroscopy; Photoionization and Photodissociation Methods in Mass Spectrometry; Pyrolysis Mass Spectrometry, Methods; Time of Flight Mass Spectrometers; X-Ray Spectroscopy, Theory; Zero Kinetic Energy Photoelectron Spectroscopy, Applications; Zero Kinetic Energy Photoelectron Spectroscopy, Theory.**

## Further reading

Baer T (1979) State selection of photoion-photoelectron coincidence. In: *Gas Phase Ion Chemistry*. New York: Academic Press.

Briggs D (ed). (1977) *Handbook of X-ray and Ultraviolet Photoelectron Spectroscopy*. London: Heyden.

Brundle CR and Baker AD (eds) (1977) *Electron Spectroscopy: Theory, Techniques and Applications*. Vol 1–4 London: Academic Press.

Eland JHD (1974) *Photoelectron Spectroscopy*. London: Butterworth.

Hollas JM (1997) Photoelectron and related spectroscopies; lasers and laser spectroscopy. In: *Modern Spectroscopy, Third Edition*. Chichester: Wiley.

Kamke W (1993) Photoelectron-photoion coincidence studies of clusters. In: *Cluster Ions*. Chichester: Wiley.

Martensson N, Baltzer P, Brühwiler PA *et al* (1994) A very high resolution electron spectrometer. *Journal of Electron Spectroscopy and Related Phenomena* **70**: 117–128.

Müller-Dethlefs K and Schlag W (1998) Chemical applications of zero kinetic energy (ZEKE) photoelectron spectroscopy. *Angewandte Chemie, International Edition in English*. **37**: 1346–1374.

Powis I, Baer T and Ng CY (eds) (1995) *High Resolution Laser Photoionization and Photoelectron Studies*. Chichester: Wiley.

Rabalais JW (1977) *Principles of Ultraviolet Photoelectron Spectroscopy*. New York: Wiley.

Schlag EW (1998) *ZEKE Spectroscopy*. Cambridge: Cambridge University Press.

Schlag EW, Peatman WB and Müller-Dethlefs K. (1993) Threshold photoionization and ZEKE spectroscopy: a historical perspective. *Journal of Electron Spectroscopy and Related Phenomena* **66**: 139–149.

# Photoelectron Spectroscopy

**John Holmes**, University of Ottawa, Ontario, Canada

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| Theory |

## Photoelectron spectroscopy

The Encyclopedia includes a specialized article entitled 'Zero Kinetic Energy Photoelectron Spectroscopy' that describes a modern development of a long-established technique.

The principle upon which photoelectron spectroscopy (PES) is based is simple. If a molecule is excited by a high-energy photon in the ultraviolet region of the spectrum that has sufficient energy to ionize the molecule, the excited species will eject electrons. PES is the analysis of the kinetic energies of the ejected electrons. For a given excitation energy, the energy distribution of the ejected electrons reflects the distribution of accessible energy levels of the excited (ionized) molecule.

$$E_{\mathrm{ion}} = h\nu - E_{\mathrm{electron}} = h\nu - \tfrac{1}{2} m_{\mathrm{e}} v_{\mathrm{e}}^2$$

The first experiments in this field were performed in Russia in 1961 and in England in 1962.

With the measurement of highly resolved electron energy distributions, the orbitals from which the electrons are lost can be identified as well as vibrational progressions for the various excited ionic states.

The most commonly used UV photons are those from the He(I) line ($1s^2 2p^1 \rightarrow 1s^2$) from a helium discharge lamp, at 58.43 nm, corresponding to an energy of 21.22 eV, well above the ionization energy of organic compounds. For a thorough review of the technique and its results, the book by Turner, a founding father of the subject, should certainly be consulted.

## List of symbols

$E_X$ = energy of species X; $h$ = Planck constant; $m_e$ = electron mass; $v_e$ = electron velocity; $\nu$ = photon frequency.

*See also:* **Photoelectron Spectrometers; Photoelectron Spectroscopy, Applications; Photoelectron Spectroscopy, Theory.**

## Further reading

Brundle CR (1993) UPS at the beginning. *Journal of Electron Spectroscopy and Related Phenomena* **66**: 3–17.

Eland JHD (1983) *Photoelectron Spectroscopy; an Introduction to UV Photoelectron Spectroscopy in the Gas Phase*, 2nd edn. London: Butterworths.

Turner DW, Baker AD, Baker C and Brundle CR (1970) *Molecular Photoelectron Spectroscopy*. London: Wiley Interscience.

# Photoelectron–Photoion Coincidence Methods in Mass Spectrometry (PEPICO)

**Tomas Baer**, University of North Carolina, Chapel Hill, NC, USA

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Methods & Instrumentation |

## Introduction

Photoelectron–photoion coincidence (PEPICO) is a method for energy selecting ions and studying their reaction dynamics. This subfield of mass spectrometry provides information about the mechanism and dissociation dynamics of ions. It is also a subfield of reaction kinetics because it provides a simple method

for investigating reaction rates with energy-selected ions. Mass spectrometry is a useful analytical tool for investigating the structure of molecules. It is based on the principle that a molecule such as A-B-C-D-E-F once ionized in the source of a mass spectrometer, breaks apart into the ionic fragments A-B-C$^+$, C-D-E-F$^+$, A-B$^+$, etc. By piecing together the various units whose masses have been measured, it is possible to reconstruct the original molecule. Such an approach works very well if the molecule is well behaved and breaks apart in a manner suggested above. However, it is common for ions to rearrange to new structures prior to dissociation. In those cases, the fragment ions observed in the mass spectrometer may not have any simple relationship to the original molecule. An example is acetol, which reacts at low energies in the following manner:

$$CH_3C(O)-CH_2OH \rightarrow CH_3CO^+ + CH_2OH^\bullet \quad [1]$$

$$CH_3C(O)-CH_2OH \rightarrow C_2H_5O^+ + HCO^\bullet \quad [2]$$

The first dissociation path yields a product ion that is clearly related to the original acetol molecule. However, the second dissociation path, the loss of $HCO^\bullet$, is not possible from the original structure. Indeed, a rather major isomerization of the acetol ion must precede the loss of $HCO^\bullet$. Isomerization reactions often dominate the dissociation dynamics at low energies, when the ion is formed with an energy just above the dissociation limit. At higher energies, reaction channels that are dominant near threshold are often superseded by other dissociation paths. Thus, learning about reaction mechanisms and their variation with the ion internal energy becomes an important part of understanding the final mass spectrum. Information of this sort is provided by the PEPICO technique.

## Methods for ion energy selection

### Review of the methods

A number of methods for ion energy selection have been developed. They can be divided into two categories. In one, ions are first produced in the ground (or near ground) state and then excited by photons, usually in the form of laser light. Such experiments have been carried out in ion cyclotron resonance (ICR) mass spectrometers, in sector instruments, and in laser-based time-of-flight (TOF) instruments. The advantage of this method is that laser methods can be used which provide good intensity and good energy resolution. Furthermore, it is possible to investigate very fast rate processes. The disadvantage is that parent ions are sometimes difficult to produce with little excess energy so that the final ion energy may be uncertain.

The other method involves direct ionization of the molecule to the ion energy of interest. This can be done optically by photoelectron–photoion coincidence (described below), or by charge exchange, which is a form of chemical ionization. Common chemical ionization methods include charge transfer and proton transfer.

$$Xe^{\bullet+} + CH_3OH \rightarrow Xe + CH_3OH^{\bullet+} \quad [3]$$

$$CH_4^{\bullet+} + CH_3OH \rightarrow CH_3^\bullet + CH_3OH_2^+ \quad [4]$$

Both methods produce ions with little internal energy. Charge transfer between a rare gas atom and a molecule deposits all of the reactant ion energy into the molecular ion as internal energy, with small amounts going into the translations. Charge transfer with ions of various ionization energies has thus been used to energy select ions. The major shortcoming of this method is that only a limited number of atomic ions can be used for this purpose.

### Energy selection by photoelectron photoion coincidence

*The photoionization process* Ionization by photons takes place by both direct ionization and by autoionization.

$$AB + h\nu \rightarrow AB^{\bullet+} + e^- \quad [5]$$

$$AB + h\nu \rightarrow AB^* \rightarrow AB^{\bullet+} + e^- \quad [6]$$

The processes differ in that the transition probability in the former is determined by the ionization continuum, whereas it is governed by the transition to the excited neutral state, AB*, in the latter. Of interest to the PEPICO experiment is the decay of these autoionizing resonances and their effect on the ability to select electrons of a given energy. It has been found that, because of autoionization, ions can be prepared in Franck–Condon gap regions that are normally not accessible by direct photon excitation from either the ground neutral or the ground ionic state.

**The principle of energy selection by PEPICO** When a molecule absorbs a vacuum ultraviolet (VUV) photon with an energy ($h\nu$) above the molecule's ionization energy (IE), the ejected electron can have energy that ranges from 0 to IE – $h\nu$. Thus, the ion is produced with an internal energy given by

$$E_{\mathrm{ion}} = h\nu - \mathrm{IE} - E_{\mathrm{el}} + E_{\mathrm{thermal}} \qquad [7]$$

where $E_{el}$ is the energy carried away by the electron and $E_{thermal}$is the thermal energy of the molecule prior to ionization. It is evident from Equation [7] that an ion of internal energy $E_{ion}$ is associated with an electron of a given kinetic energy. Thus, by detecting only ions that are collected in coincidence with an electron of a given energy, it is possible to selectively investigate energy-selected ions. This is done by placing a small electric field across the ionization region so that electrons and ions are accelerated in opposite directions. The electrons are passed through an appropriate energy analyser and collected by an electron multiplier. In the meantime, the much heavier ion requires much longer to travel down the drift tube of a TOF mass spectrometer. The time difference between the arrival time of the electron and the ion is the ion TOF. Thus, the basic information is contained in the ion TOF distribution in which the energy analysed electron provides the start signal and its corresponding ion provides the stop signal. **Figure 1** shows a schematic of a typical threshold PEPICO apparatus. The start and stop signals in **Figure 1** are usually sent to a time-to-pulse height converter (TPHC) followed by a multichannel pulse height analyser that stores and displays the ion PEPICO spectrum.

Two types of PEPICO experiments have been carried out. In one the light source has a fixed energy (e.g. the He(I) source at 21.2 eV) and the electron energies are selected by a dispersive analyser such as a hemispherical analyser. In the more versatile approach (**Figure 1**), the light source energy is tunable and the electrons of interest have zero energy. Thus the ion energy is selected by varying the photon energy while the electron energy analyser remains fixed to pass initially zero energy, or threshold electrons. The major advantages of the latter approach are (a) the better energy resolution possible with threshold electron detection, (b) the much higher collection efficiency for threshold than for energetic electrons, (c) the ability to select ions in Franck–Condon gap regions (due to autoionizing resonances), and (d) the lower level of false coincidence signals. The major disadvantages are (a) the cost and complexity of the tunable VUV source and (b) the problem with discrimination against energetic electrons whose initial velocity vector is directed towards the electron multiplier. The latter problem is suitably solved only with the use of pulsed synchrotron radiation light sources.

## Electron energy analysis

### Energetic electrons

When a light source of fixed energy is used, the electron energy selection is made by a dispersive energy analyser such as a hemispherical analyser. The resolution of such an analyser is given by $(d/r)E_e$ in which $d$ is the diameter of the input aperture, $r$ is the average radius of the inner and outer sphere in between which the electrons pass, and $E_e$ is the energy of the electron as it passes through the analyser. Although such analysers can be run with a resolution of 10 meV, this is not really practical for coincidence experiments because of a number of factors. If an electric field of $\varepsilon$ V cm$^{-1}$ is applied across the ionization region in order to extract ions, then the electrons will experience a voltage drop of $\varepsilon d_{ph}$ across the photon beam of width $d_{ph}$ thus limiting the resolution to $\varepsilon d_{ph}$. In order to avoid this, the ionization region must be kept field free and the ions extracted by a pulsed electric field applied whenever an electron is detected.

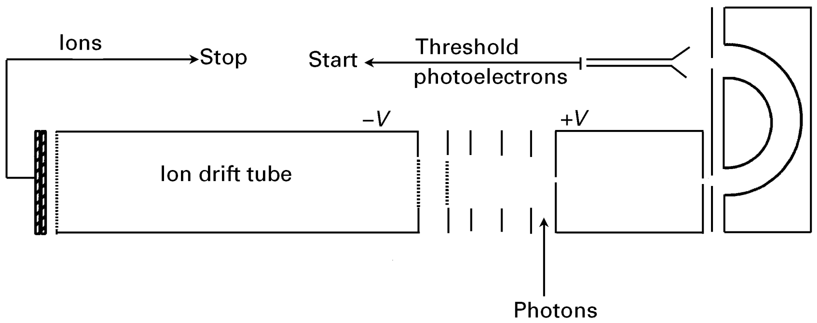



**Figure 1** Schematic diagram of a threshold photoelectron photoion coincidence experiment. The electron resolution is determined primarily by the length of the electron drift tube and the size of the apertures.

### Threshold electrons

The principle of threshold electron detection is totally different from the energy analysis of energetic electrons. A threshold has no initial kinetic energy, and thus must be extracted from the ionization region with an electric field. Because such an electron can be accelerated directly towards the detector, it can be collected with near 100% efficiency. On the other hand, energetic electrons will have a distribution of initial ejection angles which, as a first approximation, can be assumed to be isotropic. The application of a small electric field will not be sufficient to bend the electron trajectories towards the detector; thus most of the energetic electrons will be lost. Only those electrons whose initial velocity is directed toward the detector can be collected. (These are the electrons that are collected in a fixed photon energy PEPICO experiment.) It is possible to collect threshold electrons with good discrimination against energetic ones by simply passing them through small apertures in a long pipe. No real energy analysis is required. An important benefit for the PEPICO experiment is that a constant electric field can be applied to the ionization region, which gently extracts the threshold electrons and at the same time extracts the ions.

The electron energy resolution, or the electron collection efficiency, $R_{el}(E_0)$, as function of the electron's initial energy can be derived for an analyser that consists of a single acceleration region and a long drift region of length $l$ and with apertures of diameter, $d$. If we assume a point source for the electrons within this electric field, the electron collection efficiency is given by

$$R_{el} = 1 - \left[1 - \left(\frac{E_g}{E_0} + 1\right) \sin^2 \phi_c\right]^{1/2} \qquad E_0 > E_g \tan^2 \phi_c \qquad [8]$$

$$R_{el} = 1 \qquad E_0 \leq E_g \tan^2 \phi_c \qquad [9]$$

where $E_g$ is the energy gained by the electron as it is accelerated out of the ionization region, $E_0$ is the electron's initial kinetic energy, and $\phi_c$ is the planar critical collection angle, which for large $l$ is approximately $d/2l$.

A resolution of 25 meV is readily attained with electric fields of 10 V cm$^{-1}$, an electron drift tube of 10 cm and apertures of 3 mm. The real power of this method is realized when the light source is pulsed with high repetition rates, as is possible with radiation from a storage ring of a synchrotron. In that case, an additional energy analysis, TOF, can be used to improve the resolution the resolution to a few meV.

## Counting statistics — real and false coincidences

If an electron of the appropriate energy is collected but its corresponding ion is (for whatever reason) not collected, then it is possible that some other, totally unrelated ion can provide the stop signal. Such a coincidence event is called a 'false' coincidence because the electron and ion come from different precursor molecules. In a DC extraction field, these false coincidences appear at random times and thus provide a uniform background to the coincidence TOF mass spectrum. These can be easily distinguished from the real coincidences, which appear at a TOF that is determined by the ion's mass. However, if the ions are extracted with a voltage pulse, then both real and false ions will have the same TOF and are correspondingly more difficult to distinguish.

The number of false coincidence events depends upon the total ionization rate. If the total ionization rate is $10^6$ ions s$^{-1}$, then an electron and ion pair are born every microsecond. Because most electrons are of the wrong energy, they are rejected so that the observed electron count rate is much smaller than $10^6$. However, all ions can in principle be detected and, since it takes about 5 µs to extract an ion from the ionization region, it is clear that there are many ions that are able to provide stop pulses. A total ionization rate of $10^6$ is not unreasonably high for a fixed light source PEPICO experiment because the helium resonance lamp is very intense and thus generates a very large ion signal.

PEPICO experiments work best with a continuous light source that is relatively weak, and with very high collection efficiencies for both electrons and ions. A pulsed source such as a 10 Hz pulsed VUV laser would not work because each laser pulse would produce many electrons and ions within a 10 ns interval. Thus, it is impossible to distinguish which electron is associated with which ion.

One of the unique features of PEPICO is the ability to determine the collection efficiencies and thus the absolute total ionization rate, $N_T$. The ion and electron collection efficiencies ($E_i$ and $E_e$) can be calculated from the coincidence count rates ($N_c$) and the observed ion and electron count rates ($N_i$ and $N_e$). They are given by

$$\text{Ion collection efficiency} = E_i = N_c/N_e \qquad [10]$$

$$\text{Electron collection efficiency} = E_e = N_c/N_i \quad [11]$$

Once the efficiencies are known, the total ion production rate is given by $N_T = N_i/E_i$, or $N_e/E_e$. These calculations make one important assumption, which is that the electrons and ions extracted from the ionization region originate from the same ionization volume.

## The ion TOF distribution

The information content of the PEPICO technique lies in the ion TOF distribution. Not only does the TOF distribution disperse ions according to their masses, but it also provides information about the kinetic energy released in a dissociation, and about the dissociation rate of the ion if it is in the range of $10^4$–$10^7$ s$^{-1}$. Examples of such data are discussed below.

### The ion breakdown diagram

A breakdown diagram is a plot of the fractional abundances of ion masses as a function of the ion internal energy. An example is shown in **Figure 2** for the case of chromium hexacarbonyl. When the photon energy is above the molecule's ionization energy but below the first dissociation limit, the TOF distribution shows only one peak, the parent ion. Once the ion internal energy exceeds the dissociation limit, the parent ion signal disappears and in its place the first fragment ion appears, $Cr(CO)_5^+$, which is associated with the loss of a CO molecule. The cross-over energy at which parent and fragment ion intensities are equal provides a very precise means for determining this first dissociation limit.

As the photon energy is increased, the excess energy is partitioned between the translational, rotational and vibrational energies of the fragments. The rotational and vibrational energy remaining in the $Cr(CO)_5^+$ increases with photon energy until it exceeds the next dissociation limit for the loss of a second CO fragment. Because in the energy partitioning most of the energy remains in the vibrational modes of the larger fragment, the $Cr(CO)_5^+$ signal decreases to nearly zero while the $Cr(CO)_4^+$ ion takes its place. This continues up to the final loss of the last CO unit. Because of energy partitioning, each new onset is more gradual than the previous one. It must be pointed out that this breakdown diagram is very peculiar and not at all typical. It is characterized by the following successive dissociation paths:

$$Cr(CO)_6^{\bullet+} \rightarrow Cr(CO)_5^{\bullet+} \rightarrow Cr(CO)_4^{\bullet+} \rightarrow Cr(CO)_3^{\bullet+} \rightarrow Cr(CO)_2^{\bullet+} \rightarrow Cr(CO)^{\bullet+}$$

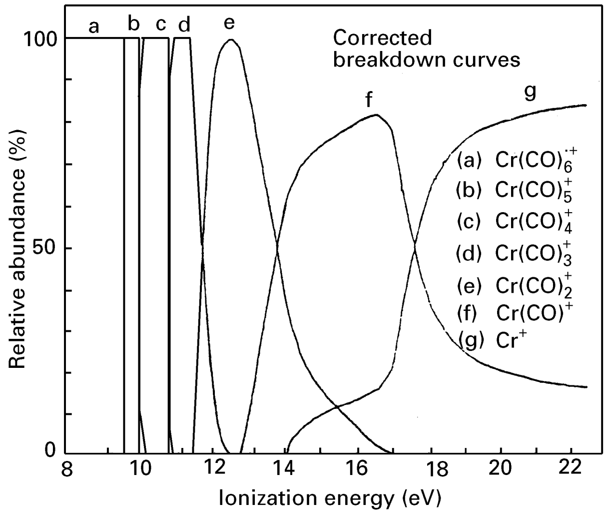



**Figure 2** A breakdown diagram of chromium hexacarbonyl ions obtained by PEPICO. This ion dissociates by a sequential mechanism. The heats of formation of the various ions can be determined from the energy onsets of their formation. Reproduced with permission from Das PR, Nishimura T and Meisels GG (1985) Fragmentation of energy selected hexacarbonylchromium ion. *Journal of Physical Chemistry* **89**: 2808–2812.

Most organic ions dissociate not in sequential manners but in a combination of parallel and sequential reactions such as those for nitrobenzene:

$$C_6H_5NO_2^+ \rightarrow C_6H_5O^+ + NO^\bullet \qquad AE = 10.98\ \text{eV} \quad [12]$$

$$\hookrightarrow C_5H_5^+ + CO \qquad AE = 11.30\ \text{eV} \quad [13]$$

$$\hookrightarrow C_4H_3^+ + C_2H_2O \qquad AE = 11.40\ \text{eV} \quad [14]$$

$$\rightarrow C_6H_5O^\bullet + NO^+ \qquad AE = 11.04\ \text{eV} \quad [15]$$

$$\rightarrow C_6H_5^+ + NO_2^\bullet \qquad AE = 11.14\ \text{eV} \quad [16]$$

Breakdown diagrams such as those in **Figure 2** provide an overview of the dissociation dynamics for ions as a function of the internal energy. Generally, fragments produced by low-energy rearrangement reactions appear first, while the direct bond cleavage products appear at higher energies. An example is the loss of $NO_2^\bullet$ from the nitrobenzene ion at 11.14 eV. Lower-energy dissociation paths involving the production of $NO^\bullet$ and $NO^+$ involve rearrangements.

An important aspect of the breakdown diagram is the cross-over energy for the first dissociation limit where the parent and daughter ion signals are equal. For samples at a temperature, $T$, this energy is located below the 0 K dissociation onset. However, there is no ambiguity in identifying this energy in the breakdown diagram if the sample temperature and its vibrational frequencies are known. To convert the 298 K cross-over energy to the cross-over energy expected for reactants initially at 0 K, one simply

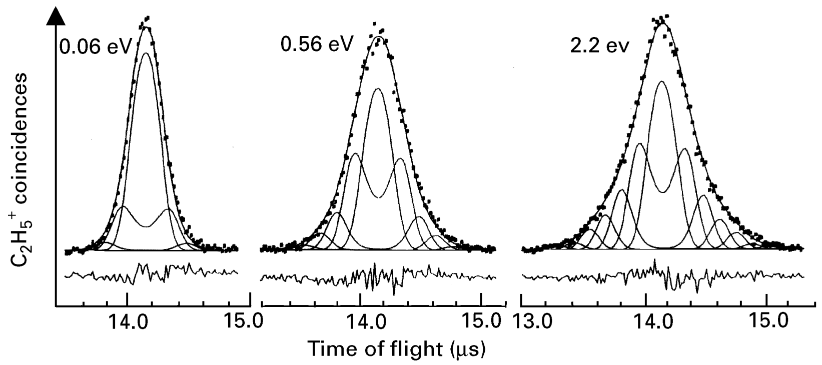



**Figure 3** The PEPICO TOF distribution of $C_2H_5^+$ ions from energy selected $C_2H_5I^{\bullet+}$ ions at various energies above the dissociation limit for $I^{\bullet}$ loss. The solid lines that fit the experimental points are single energy release distributions. From these data, the whole distribution of product translational energies could be determined. Reproduced with permission from Baer T, Büchler U and Klots TCE (1980) Kinetic energy release distributions for the dissociation of internal energy selected $C_2H_5I^{\bullet+}$ ions. *Journal de Chimie Physique* **77**: 739–743.

adds the median thermal energy to the 298 K crossover energy.

### Kinetic energy release in the dissociation

Because the timing in PEPICO experiments is very precise, all of the broadening in TOF peaks can be associated with the kinetic energy of the parent or the fragment ions. Consider two ions that are ejected along the TOF axis but in opposite directions with initial velocity, $\pm v_0$ where the positive sign signifies the direction toward the ion TOF tube. The negative-going ion will be decelerated and turn around and come back to its original point of formation but with its velocity vector now converted from – to +. From this point, the ion will have the identical TOF to the ion whose initial velocity vector was positive. Thus, the TOF difference between + and – ions is just the 'turn-around' time of the negative-going ion. This time is readily calculated from Newton's equation, $F = ma = q\varepsilon$, where $q$ is the ion charge, $m$ is the ion mass and ε is electric field, and is given by $2v_0m/q\varepsilon$, where again $v_0$ is the initial velocity of the fragment ion. In a polyatomic molecule, much of the energy in excess of the dissociation limit remains in the vibrational modes. In that case, the ion fragments with a *distribution* of translational energies that is often well modelled by a Maxwell–Boltzmann distribution, as predicted by the statistical theory of unimolecular decay. If the three-dimensional distribution in the centre of mass is a Maxwell–Boltzmann type, the TOF peak will appear Gaussian, and the average kinetic energy release (KER) will be given by

$$\langle \text{KER} \rangle = \frac{M}{(M-m)\,m}\frac{3}{16\ln 2}(q\varepsilon)^2(\text{FWHM})^2 - \frac{m}{M-m}\langle E_{\text{th}} \qquad [17]$$

where $M$ and $m$ are the masses of the parent and fragment ions, and FWHM is the TOF peak width. The first term is a result of kinetic energy release in the dissociation, while the second term subtracts the width due to the thermal motion of the parent ion. An example of this sort of statistical distribution is shown in **Figure 3** for the case of $C_2H_5I^+$ dissociation at three energies above the dissociation limit. As the ion energy increases, the peak widths broaden. The solid lines in this figure are TOF distributions expected for single kinetic energy releases of $9n^2$ ($n = 2, 3, 4,..$) meV. Because of apertures, ions with significant off-axis velocities will be stopped. Thus, each TOF peak for ions with KER greater than $n = 2$ will appear as a doublet, which accounts for the forward and backward ejected ions.

One of the most useful applications of peak widths in the TOF distributions is in distinguishing an ionization from a dissociative ionization process. This is particularly useful in the case of photoionization of clusters. An example of such a study is the case of $ArCO^+$, which can arise from a variety of process in a molecular beam expansion of Ar and CO gases, including those shown in Equations [18]–[20].

$$\text{ArCO} + h\nu \rightarrow \text{ArCO}^{\bullet+} + \text{e}^- \qquad [18]$$

$$\text{Ar}_2\text{CO} + h\nu \rightarrow \text{ArCO}^{\bullet+} + \text{Ar} + \text{e}^- \qquad [19]$$

$$\text{Ar(CO)}_2 + h\nu \rightarrow \text{ArCO}^{\bullet+} + \text{CO} + \text{e}^- \qquad [20]$$

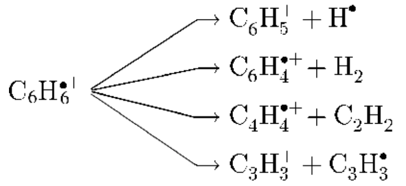
If one is interested in the appearance energy or the thermochemistry of the $ArCO^+$ ion, it is necessary to know by which process this ion is created. In each of the three reactions, the product ion mass is the same.

However, the first reaction can be distinguished from the dissociative ionization processes because its TOF peak is expected to be narrow, whereas the latter two reactions will impart kinetic energy to their fragments and thus result in broadened TOF peaks. This is illustrated in **Figure 4**. The narrow peak with a width of 34 ns is due to reaction [18], while the broad peak with a width of 160 ns is due to the dissociative processes. By this approach, an onset of 13.03 eV for the $ArCO^+$ ion from ArCO could be measured. In addition, by measuring the width of the broad peak as a function of the cluster ion energy, it was possible to extrapolate the kinetic release in the second reaction down to the threshold (as shown in **Figure 5**).

A similar study found that $Ar_n^{\bullet+}$($n$>2) ions produced in a supersonic expansion of argon originated exclusively by dissociative photoionization from higher-order clusters. Only in the case of $Ar_2$ could ions of that mass be produced without the accompanying dissociation. This was obvious from the broad peak widths, which showed no evidence of any narrow components for $Ar_n^{\bullet+}$ other than for $n$=2.

### Ion dissociation rate measurements

The dissociation rates of metastable ions can be measured by PEPICO because the ion TOF distribution is very sensitive to the position of dissociation in the acceleration region. Ions that dissociate rapidly gain the full kinetic energy in the acceleration region, whereas ions that dissociate some distance into the acceleration region end up with less kinetic energy because it is partitioned between the ion and neutral fragments. Typical ion TOF distributions of products formed from metastable $C_6H_6^{\bullet+}$ ions are shown in **Figure 6**. The benzene ion dissociates via four major paths at low ion energies:



[21]

Only the latter two fragments are shown in **Figure 6** because the mass difference between the parent ion and the H and $H_2$ loss channels are not sufficient to be resolved in this low-resolution TOF spectrum. However, what is evident is the asymmetry of the $C_4H_4^{\bullet+}$ and $C_3H_3^+$ ion TOF distributions. The peak shapes can be modelled (solid lines in **Figure 6**) knowing the mass of the ions, the electric field, the length of the acceleration and drift distances, and the ion dissociation rate, which is an adjustable parameter. As the ion energy is increased from 14.85 eV to 15.32 eV, the rate constant increases from 0.16 to ~1.2 $\mu s^{-1}$.

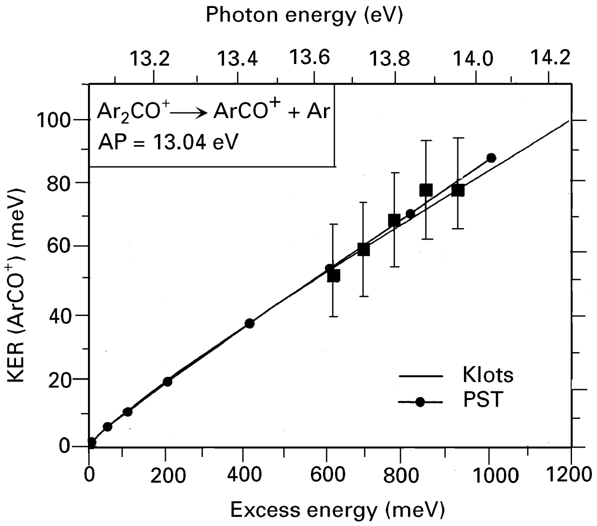



**Figure 5** The derived kinetic energy release from the energy-selected $Ar_2CO^{\bullet+}$ ions as a function of the trimer ion internal energy. The solid line is a calculated kinetic energy release based on the statistical theory of dissociation: phase space theory (PST) or the version of PST due to C.E. Klots. AP is the threshold energy for $ArCO^+$ formation. The onset leads to a heat of formation of the trimer ion. Reproduced with permission from Mahnert J, Baumgartel H and Weitzel KM (1997) The formation of $ArCO^+$ ions by dissociative ionization of argon/carbon monoxide clusters. *Journal of Physical Chemistry* **107**: 6667–6676.

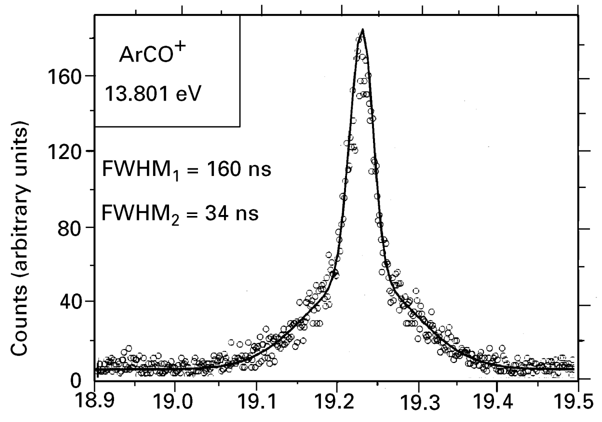



**Figure 4** The PEPICO TOF distribution of $ArCO^{\bullet+}$ ions from various precursor cluster ions. The sharp peak is due to the direct ionization of ArCO dimers, whereas the broad peak with a width of 160 ns is due to dissociative ionization of $Ar_2CO$ trimers. Reproduced with permission from Mahnert J, Baumgartel H and Weitzel KM (1997) The formation of $ArCO^+$ ions by dissociative ionization of argon/carbon monoxide clusters. *Journal of Physical Chemistry* **107**: 6667–6676.

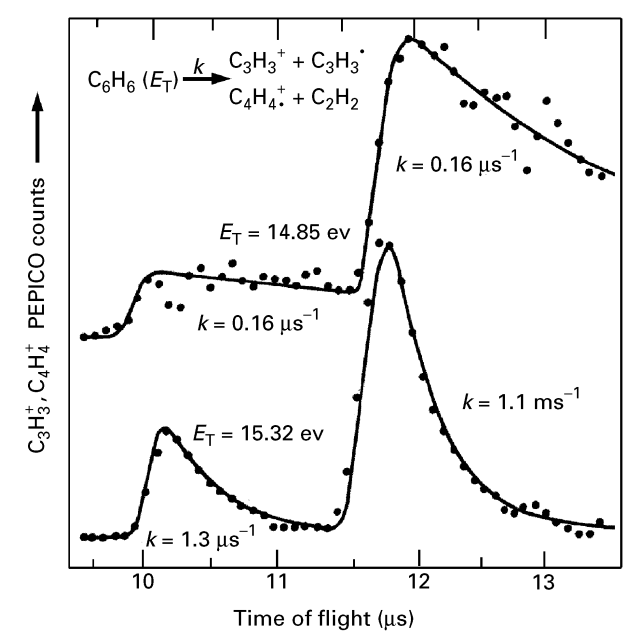



**Figure 6** ($k$ = rate constant; $E_T$ = total ion energy) The PEPICO TOF distribution of $C_3H_3^+$ and $C_4H_4^{\cdot+}$ ions from energy-selected $C_6H_6^{\cdot+}$ ions. The asymmetric TOF distributions are a result of the slow reaction of the metastable ions. The solid lines are calculated distributions using the mean ion dissociation rate as an adjustable parameter. Reproduced with permission from Baer T, Willet GD, Smith D and Phillips JS (1979) The dissociation dynamics of internal energy selected $C_6H_6^+$. *Journal of Chemical Physics* **70**: 4076–4085.

Such data collected at a number of ion energies leads to a $k(E)$ vs $E$ curve shown in **Figure 7**. The solid lines in this figure are the results of a statistical theory calculation using the Rice–Rampserger–Kassel–Marcus (RRKM) formulation. Studies such as this one on many molecules have shown that the RRKM theory is extremely useful and accurate in predicting the dissociation rate constants of ionic reactions. It is also very useful for determining the onset of dissociation. This is of interest both for fundamental reasons as well as for a very practical one. Dissociation onsets are often used to determine thermochemical properties of ions and free radicals. Consider the reaction

$$AB + h\nu \rightarrow A^+ + B^\bullet \qquad [22]$$

in which the stable molecule is dissociatively ionized to the ion $A^+$ and a free radical, $B^\bullet$. By energy conservation, the heats of formation of the three species in Equation [22] are related to the minimum energy of dissociation (DE) by

$$DE = \Delta H_f(A^+) + \Delta H_f(B^\bullet) - \Delta H_f(AB) \qquad [23]$$

Thus, if any two of the three heats of formation are known, the third can be calculated from the measured dissociation energy, DE. Many heats of formation derived using this approach have been reported. However, the validity of this method depends upon a rapid dissociation of the molecular ion. What is actually measured is the appearance energy, AE, of the product ions which is always greater than DE. Ions prepared just above their dissociation threshold often fragment very slowly (e.g. $k<10^4\ s^{-1}$) so that no $A^+$ ions will be observed since ions are collected in just a few microseconds after their formation. It is not until the ion energy reaches a level at which the dissociation rate constant, $k$, exceeds $10^4 s^{-1}$ that $A^+$ ions are observed. An effective means for circumventing this problem is to measure the dissociation

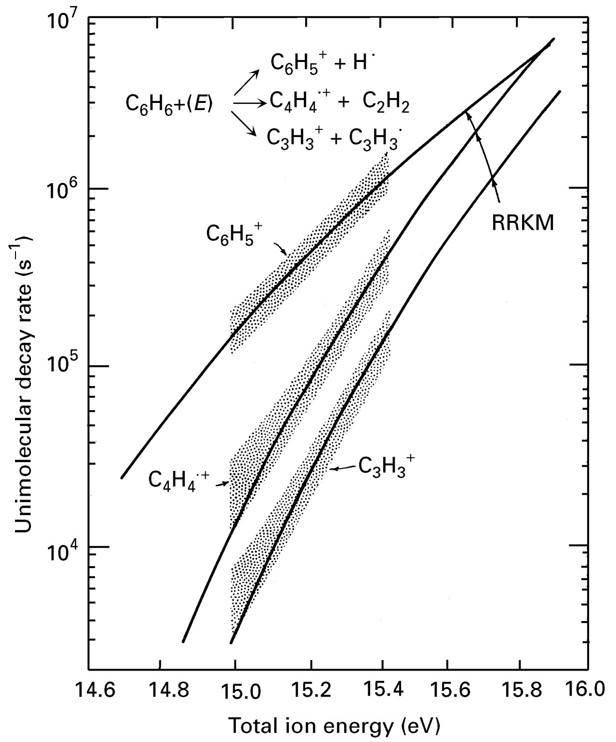



**Figure 7** The dissociation rate constants of $C_6H_6^{\cdot+}$ ions as a function of the ion internal energy. The solid lines are calculated rate constants using the statistical theory of unimolecular decay (RRKM). Reproduced with permission from Baer T, Willet GD, Smith D and Phillips JS (1979) The dissociation dynamics of internal energy selected $C_6H_6^+$. *Journal of Chemical Physics* **70**: 4076–4085.

rate constant as a function of the ion internal energy and to extrapolate this $k(E)$ curve to the dissociation onset by use of the statistical theory of unimolecular decay. The extrapolations in **Figure 7** predict that the true onsets for these dissociation channels for loss of $H^{\bullet}$, $C_2H_2$, and $C_3H_3^{\cdot}$ lie at 13.7, 14.32, and 14.47 eV, respectively. It is apparent that these onsets are well below the data in **Figure 7**, which were all collected above 15 eV because below this energy no fragment ions were observed. This is an example of the 'kinetic shift' that has shifted the onset towards higher energy because the dissociation rate constant at threshold is too slow.

Dissociation rate measurements are also very useful for determining the mechanism of ionic dissociation reactions. If the ion isomerizes prior to dissociation, the rate constant will be slower than predicted by the RRKM theory. Often it has proved possible to determine experimentally the energy of the isomerized structure by modelling the $k(E)$ curve with the RRKM theory. The reliability of such studies is greatly enhanced by the use of *ab initio* molecular orbital theory, which provides valuable input for the RRKM theory calculations.

## List of symbols

$d$ = diameter of analyser input aperture; diameter of analyser drift region aperture; $d_{ph}$ = photon beam width; DE = dissociation energy; $E_0$ = electron initial kinetic energy; $E_e$ = electron collection efficiency; $E_e$ = energy of electron as it passes through the analyser; $E_g$ = electron energy gain on acceleration from ionization region; $E_i$ = ion collection efficiency; $h$ = Planck constant; IE = ionization energy; $k$ = dissociation rate constant; KER = kinetic energy release; $l$ = length of analyser drift region; M, $m$ = masses of parent and fragment ions, respectively; $N_c$ = coincidence count rate; $N_e$ = electron count rate; $N_i$ = ion count rate; $N_T$ = total ionization rate; $q$ = ion charge; $r$ = average radius of inner and outer spheres between which electrons pass in the analyser; $R_{el}$ = electron collection efficiency; $\Delta H_f$ = heat of formation; $\varepsilon$ = electric field strength; $v_0$ = initial ion velocity; $\phi_c$ = analyser planar critical collection angle.

*See also:* **Ion Dissociation Kinetics, Mass Spectrometry; Ion Energetics in Mass Spectrometry; Ionization Theory; Metastable Ions; Photoionization and Photodissociation Methods in Mass Spectrometry; Statistical Theory of Mass Spectra.**

## Further reading

Baer T (1986) The dissociation dynamics of energy selected ions. *Advances in Chemical Physics* **64**: 111–202.

Baer T and Hase WL (1986) *Unimolecular Reaction Dynamics: Theory and Experiments*. New York: Oxford University Press.

Baer T and Mayer PM (1997) Statistical RRKM/QET calculations in mass spectrometry. *Journal of the American society of Mass Spectrometry* **8**: 103–115.

Baer T, Peatman WB and Schlag EW (1969) Photoionization resonance studies with a steradiancy analyzer. II. The photoionization of $CH_3I$. *Chemical Physics Letters* **4**: 243–247.

Baer T, Willett GD, Smith D and Phillips JS (1979) The dissociation dynamics of internal energy selected $C_6H_6^+$. *Journal of Chemical Physics* **70**: 4076–4085.

Baer T, Buchler U and Klots CE (1980) Kinetic energy release distributions for the dissociation of internal energy selected $C_2H_5I^+$ ions. *Journal de Chemie Physique* **77**: 739–743.

Berkowitz J (1979) *Photoabsorption, Photoionization, and Photoelectron Spectroscopy*. New York: Academic Press.

Dannacher J, Rosenstock HM, Buff R *et al* (1983) Benchmark measurement of iodobenzene ion fragmentation rates. *Chemical Physics* **75**: 23–35.

Das PR, Nishimura T and Meisels GG (1985) Fragmentation of energy selected hexacarbonylchromium ion. *Journal of Physical Chemistry* **89**: 2808–2812.

Eland JHD (1972) Predissociation of triatomic ions studied by photoelectron photoion coincidence spectroscopy and photoion kinetic energy analysis. *International Journal of Mass Spectrometry and Ion Processes* **9**: 397–406.

Mahnert J, Baumgartel H and Weitzel KM (1997) The formation of $ArCO^+$ ions by dissociative ionization of argon/carbon monoxide clusters. *Journal of Physical Chemistry* **107**: 6667–6676.

Ng CY (1998) State-selected and state to state ion-molecule reaction dynamics by photoionization methods. In Farrar JM and Saunders WHJ (eds) *Techniques for the Study of Ion–Molecule Reactions,* 417–488. New York: Wiley.

Stockbauer R (1977) A threshold photoelectron photoion coincidence mass spectrometer for measuring ion kinetic energy release on fragmentation. *International Journal of Mass Spectrometry and Ion Physics* **25**: 89–101.

# Photoionization and Photodissociation Methods in Mass Spectrometry

**John C Traeger**, La Trobe University, Bundoora, Victoria, Australia

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
Methods & Instrumentation

The formation of ions in a mass spectrometer ion source can be accomplished by a variety of ionization methods. Photoionization of a molecule or atom involves the absorption of a vacuum ultraviolet (VUV) or soft X-ray photon. Depending on the wavelength (λ) of the photon, whose energy $E$ is given by the Planck–Einstein relationship,

$$E(\mathrm{eV}) = \frac{hc}{\lambda} = \frac{1239.85}{\lambda(\mathrm{nm})} \qquad [1]$$

both ionization and subsequent fragmentation of the ionized molecule may occur. It is also possible to induce fragmentation of an initially formed ion by absorption of a single photon, or by the sequential absorption of a number of photons of equal or different wavelengths, in a photodissociation process. One particularly useful advantage of using a photon beam for ionization of a neutral, or dissociation of an ion, is that the amount of energy deposited depends only on the wavelength, which can be precisely measured and controlled. For this reason, the techniques of photoionization and photodissociation mass spectrometry are primarily used for fundamental energetic, kinetic and structural studies, although there are various analytical applications that can also benefit from these particular experiments. Because mass spectrometers used for photoionization and photodissociation experiments are not widely available, they are generally designed and constructed in the laboratory engaging in the particular studies.

A major source of energetic information is the photoionization efficiency (PIE) curve in which the relative numbers of photoions produced per number of photons transmitted is measured as a function of photon wavelength (energy). A typical PIE curve is shown in **Figure 1**.

Apart from yielding accurate thermochemical data, the PIE is also a convenient means of studying the energy dependence of a range of gas-phase unimolecular processes, which include those involving weakly bound clusters and ion–neutral complex intermediates. The detection of near zero-energy

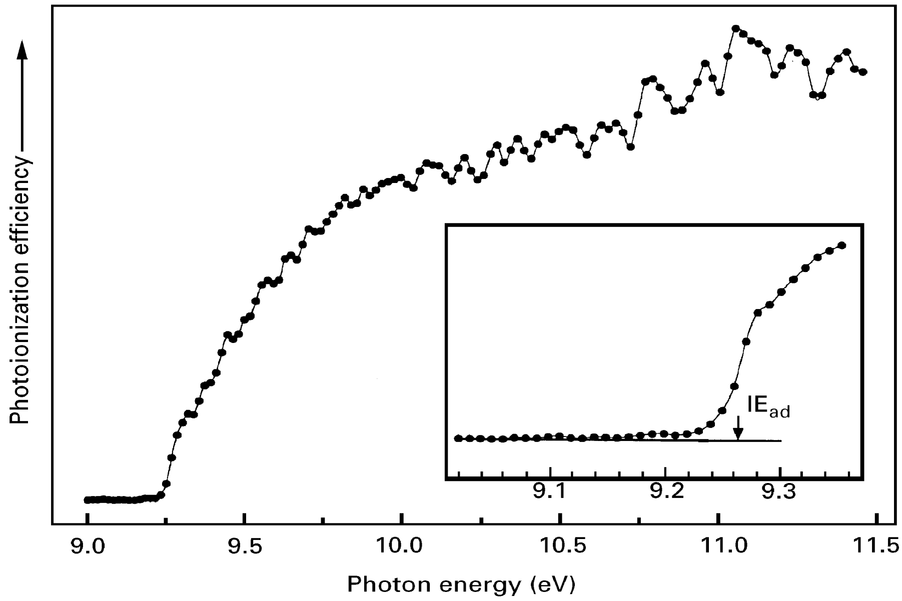



**Figure 1** Photoionization efficiency curve with (inset) an expanded threshold region for the molecular ion from cyclopentanone. The photoions detected below the adiabatic ionization energy $IE_{ad}$ (indicated by the arrow) are due to hot bands.

photoelectrons coincident with photoion formation (threshold PEPICO (photoelectron–photoion coincidence)) provides a direct measure of the internal energy of the initially formed ion; for many systems similar information can be obtained from the corresponding first differential PIE curve.

Photodissociation experiments involve the photon-assisted decomposition of a preformed ion beam with subsequent detection of the resulting fragment ions. To observe photodissociation with a mass spectrometer it is necessary that the photoabsorption process produces an excited state of the precursor ion above its lowest dissociation threshold. The dissociation should also be relatively fast compared to any other competing process, such as relaxation. The photodissociation spectrum of an ion, which is a measurement of the photodissociation efficiency as a function of photon energy, is related to the photoelectron spectrum for the corresponding neutral molecule and is comparable to the ion absorption spectrum. However, unlike photoabsorption experiments, photodissociation is a more sensitive technique for obtaining structural information because any small changes are not superimposed on a relatively large signal. There is also less interference from other contaminating species, such as ion–molecule reaction products.

Photodissociation spectra of excited ionic species that predissociate contain more structural information than those resulting from a direct dissociation. Unfortunately, there are few ions that undergo suitable predissociation processes. Time-resolved photodissociation, which involves the monitoring of fragment ion build-up following photoexcitation, provides a direct measure of the bond cleavage kinetics, from which bond strength information can be obtained. These experiments, in conjunction with Rice–Ramsperger–Kassel–Marcus (RRKM) calculations, can also help to determine dissociation thresholds.

## Photoionization

### Sample introduction

The simplest method for introducing a sample into an ion source is via an effusive beam through a pinhole or needle valve. For some experiments it is advantageous to cool the sample prior to introduction to minimize problems associated with the thermal population of vibrationally and rotationally excited species. Because of condensation problems, this method is not applicable to molecules with relatively high freezing points. However, the use of a supersonic molecular beam for sample introduction can result in a substantial internal cooling effect, with rotational and vibrational temperatures below 20 K being achieved for polyatomic molecules. This significantly reduces the interference from low-energy tailing, called hot band structure, that is usually observed in the pre-threshold region of a room-temperature PIE curve (see **Figure 1**). Various free radicals and transient molecules have been studied by photoionization, although they often require special methods of preparation and introduction. Because these species are usually present at very low concentrations, it is preferable to have access to a high-intensity photon source. A molecular beam can also facilitate the production of clusters for investigation by photoionization.

### Photon sources

Several different laboratory light sources can be used for VUV studies. These usually involve a gas discharge, which may produce either a continuum or a discrete-line emission spectrum. Although the output of light is generally much greater for a line spectrum, these cannot be used for very high-resolution studies because PIE measurements are then restricted to the finite number of emitted photon energies available. A commonly used photon source is the hydrogen many-lined spectrum (pseudocontinuum), which is shown in **Figure 2**. This is produced by a simple DC discharge in hydrogen at a relatively low pressure (<1 kPa), and produces usable photons in the energy range of 7–14 eV.

Care must be exercised with the data processing when using such a discrete-line source as it is possible for false structure to be generated in the PIE curve, particularly in the low photon intensity regions (e.g. 9.3–9.6 eV) of the lamp.

Two continuum sources that have been used for photoionization studies are the argon continuum (8–12 eV) and the Hopfield helium continuum (12–21 eV). Both of these involve emission from a transient diatomic molecule generated via a pulsed high-voltage, high-pressure (>10 kPa) rare gas discharge. As for the hydrogen pseudocontinuum, a typical photon intensity is approximately $10^{9}$–$10^{10}\,s^{-1}$, which is substantially lower than the photon intensity obtained with either an electron monochromator or a conventional electron ionizing source.

Experiments involving continuous emission at energies above 21 eV require the use of synchrotron radiation, in which light is emitted from rapidly moving electrons accelerated in a circular path. The spectral distribution and intensity available from the electron storage ring vary depending on a number of different operating parameters. However, usable photon energies typically extend from the visible into the X-ray region, with intensities significantly greater

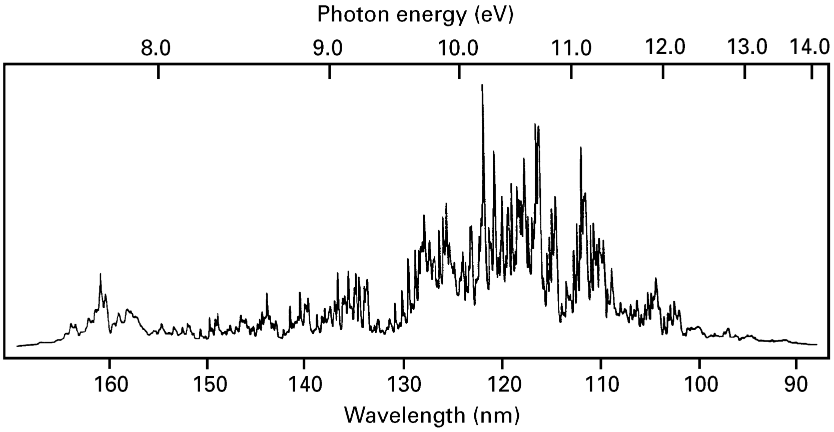



**Figure 2** The many-lined molecular hydrogen pseudocontinuum.

than those of comparable gas discharge lamps. The high intensity of synchrotron radiation has proved particularly successful for photoionization studies involving low concentrations of free radicals.

The use of lasers as a photoionization light source has been limited by the availability of a suitable tunable VUV laser. The fundamental problem in developing such a high-energy laser is that the laser action becomes increasingly difficult to sustain as the photon energy is increased. Consequently, most laser-based photoionization experiments have involved multiphoton techniques in which the energies of several photons are combined. A commonly used method for the ionization of molecules is resonance-enhanced multiphoton ionization (REMPI), in which the laser frequency is tuned to an excited neutral electronic state. This greatly enhances the cross section for ionization. One photon is involved in the excitation process while a second photon, which may have a different wavelength (colour), is used to ionize the excited molecule.

### Monochromators

For energetic studies it is necessary to select a given photon wavelength from the light source with a monochromator. Although there are numerous types available, it is preferable to use a design in which the entrance and exit slits remain in a fixed position relative to the mass spectrometer. The image at the exit slit should also stay in focus. Two monochromators that essentially satisfy these requirements are the Seya–Namioka and the near-normal-incidence type. The Seya–Namioka monochromator, in which wavelength selection is accomplished by the simple rotation of a diffraction grating about its centre, unfortunately suffers from an astigmatic and curved image at the exit slit, which result in a decrease in both intensity and resolution. Despite its more complex scanning mechanism, the near-normal-incidence monochromator is better suited for high-resolution studies. Photoionization experiments with monochromator resolutions better than 0.002 nm have been reported, although a typical operating resolution is 0.1 nm, which corresponds to 0.008 eV for a photon with an energy of 10 eV. Wavelength calibration is normally done using atomic lines generated in the emission spectrum of the lamp.

The concave diffraction gratings used in these monochromators are usually optimized to maximize the light intensity over a selected operating wavelength region. This includes the selection of a blaze angle and a particular overcoating of the ruled surface. As shown in **Figure 3**, an aluminium surface with a thin film of $MgF_2$ or LiF provides the highest reflectance for photon energies below 12.0 eV. Metal coatings, such as platinum, are best used for experiments that extend to higher energies. The dispersion (resolution) of a grating depends on both its radius of curvature and the ruled line density. Typical values for these parameters are 1 m and 1200 grooves $mm^{-1}$ respectively. Ghosting effects and the amount of scattered light can be reduced with a grating that has been holographically ruled.

If a high-pressure gas discharge is being used as the VUV light source, it is preferable to isolate the monochromator from the lamp. Apart from the problems associated with interfacing the lamp and monochromator to the vacuum environment of a mass spectrometer, self-absorption of light by the discharge gas greatly reduces the available photon intensity. In some cases the installation of an optical window at the entrance slit can be used to avoid this, although the choice of window material is somewhat restricted. A LiF filter can be used for photon energies up to 12.0 eV, although for photoionization

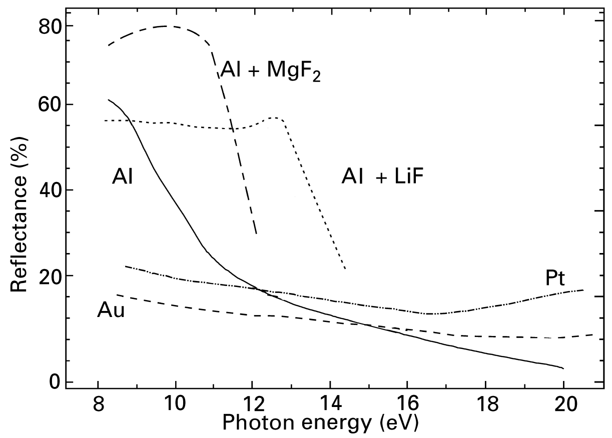



**Figure 3** Vacuum ultraviolet reflectance curves for different grating surfaces.

studies above this sharp cut-off energy the monochromator needs to be windowless. This requires the use of substantial vacuum pumping systems to maintain a satisfactory pressure differential between the lamp and the monochromator, especially when a wide entrance slit is being used. Because the storage ring of a synchrotron light source is operated under a high vacuum, it does not suffer from this particular interfacing problem. However, the wide energy range of the generated light can result in contamination of a PIE curve by second-order or higher-order diffraction, requiring analytical removal during the data reduction. Alternatively, for low-energy studies an appropriate filter, such as a LiF window, can be used to effectively remove this interference.

### Photon detectors

When conducting PIE experiments it is essential to monitor the total amount, or a representative sample, of light that is being used for the ionization process. The simplest device involves counting the photoelectrons emitted from either a simple irradiated metal plate, or, for low levels of light, via a discrete or continuous dynode (channeltron) electron multiplier. However, both methods require calibration of the photocathode spectral response. Electron multipliers have also been developed in conjunction with special photosensitive cathodes for operation well into the vacuum ultraviolet region. Because they are relatively insensitive to light of wavelengths greater than 300 nm, which is approximately the limit of solar radiation reaching the earth's surface, they are often referred to as 'solar blind' photomultipliers. Again, the photoelectric yield is wavelength dependent and must be calibrated.

An alternative approach to photon detection is to use a phototransducer in conjunction with a conventional photomultiplier. This involves measuring the fluorescence produced following irradiation with VUV light. A particularly convenient and widely used phosphor is sodium salicylate, which has a high fluorescence efficiency and nearly constant quantum yield over a wide photon energy range (5–35 eV). Periodic replacement of this phosphor is required as its low-wavelength efficiency deteriorates with time in a typical vacuum environment.

### Mass analysers

Several different types of mass analyser have been used for photoionization studies, including quadrupole, magnetic sector and time-of-flight (TOF) mass spectrometers. The technique of time-resolved photoionization mass spectrometry uses a quadrupole ion trap to store ions for a selected time prior to detection, which greatly increases the threshold sensitivity for slow fragmentation processes. Many experiments require extended periods of data acquisition, so that a high mass stability is vital. In addition, because photoionization produces only small numbers of ions in the ion source, it is important to ensure that the mass analyser has a high ion transmission efficiency. The probability of photoionization can be enhanced by increasing the sample pressure in the ion source. However, to minimize absorption corrections and collision-induced effects, such as ion–molecule reactions and collisional activation, maximum pressures are generally restricted to below 0.1 Pa.

The quadrupole mass filter is an inexpensive and compact device in which the ion source operates at ground potential, making it convenient for interfacing to a monochromator. The ion optics for entry to the quadrupole RF/DC field are relatively straightforward, while it is a simple procedure to alter the mass resolution and maximize the ion count rate for a particular experiment.

A magnetic mass analyser is considerably more expensive and physically larger than its quadrupole counterpart. In normal operation the ion source operates at a high positive voltage, so that the monochromator interface region must be carefully designed to ensure that photoelectrons ejected from either the photodetector or any metal surfaces are not accelerated into the ion source to produce spurious ionization. For maximum ion transmission it is necessary to use an ion-optical lens, such as an electrostatic quadrupole, to focus the ions from a wide area into the narrow entrance slit of the mass spectrometer. The magnetic mass spectrometer, unlike a quadrupole

mass filter, is able to detect metastable ions, which makes it particularly useful for the study of unimolecular fragmentation processes.

The TOF mass spectrometer is well suited to the mass analysis of ions produced from a pulsed source. In contrast to scanning mass analysers, the TOF detects all ions sampled from the ion source. For this reason it is widely used for coincidence PEPICO measurements and various laser ionization experiments. The pulsed nature of synchrotron radiation also lends itself to TOF photoionization studies. Peak shapes for different ions in a TOF mass spectrum can provide valuable information about their kinetic energy distributions.

## Ion detectors

Various ion multipliers have been used to detect mass-selected photoions. These include the discrete dynode electron multiplier, the continuous dynode channeltron and the Daly detector, which uses a combined scintillator/photomultiplier to monitor the fluorescence produced from secondary electrons generated by ion bombardment of a high-voltage conversion electrode (**Figure 4**). Because of the small numbers of ions generated in photoionization experiments, pulse counting, rather than analogue, techniques are usually employed. The digital method of detection is more immune to noise, although it is preferable to use a multiplier that has a low background pulse count rate, typically less than one pulse every 10 s. Stray external magnetic fields, such as those encountered with a magnetic mass analyser, can impair the multiplier performance and require appropriate shielding.

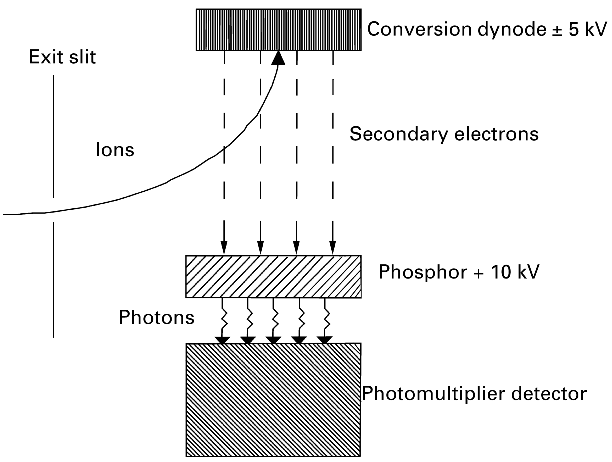



**Figure 4** Schematic representation of a Daly scintillator/photomultiplier ion detector.

## Data processing

Most photoionization experiments require extended data acquisition to increase both the ion count statistics and the signal-to-noise ratio to acceptable levels, particularly in the threshold region where the photoion count rate can be extremely low. It is not unusual for some experiments to last for 24 hours or more. However, as the reduction in random noise only increases as the square root of the accumulated counts, there is a practical limit to the improved quality of data obtained using this technique.

The PIE curve provides the basis for obtaining precise gas-phase thermochemical information, such as bond energies, proton affinities, and cationic and neutral heats of formation. Because the ionization process is governed by the Franck–Condon principle, the most probable process will involve a vertical transition. Depending on the geometry change following ionization, this may not always correspond to the adiabatic (minimum) ionization energy (IE), as shown in **Figure 5**.

Because the precursor molecules have a thermal energy distribution, ionization will not all occur from the neutral ground state, resulting in ions being produced below the true ionization threshold energy (**Figure 5**). This is observed as pre-threshold hot band structure in the PIE curve and will typically extend over a range of 0.1 to 0.2 eV for a polyatomic molecule at room temperature (see **Figure 1**).

When variable-energy photons are used for ionization it is possible for autoionization to occur. This is

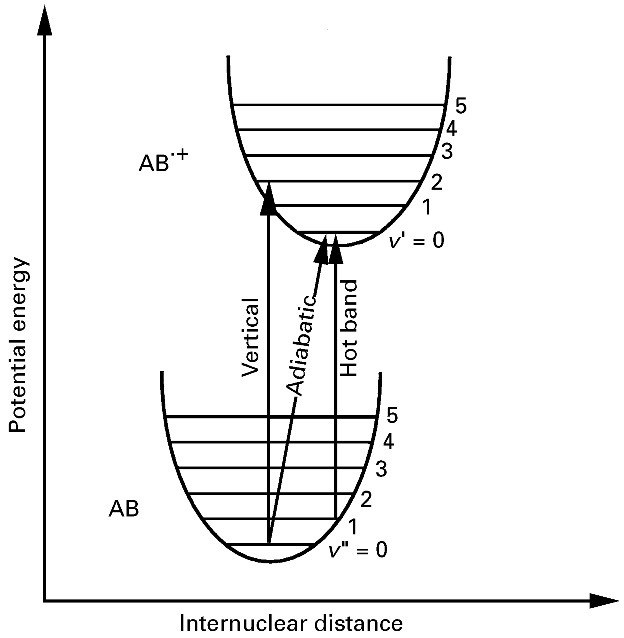



**Figure 5** Single ionization of a molecule AB, showing the adiabatic, the vertical and a hot band transition.

the result of ionization from an excited neutral Rydberg state above the ionization threshold. Autoionization processes are observed as peaks in the PIE curve and can often complicate the assignment of the adiabatic ionization energy. **Figure 6** shows the effect of autoionization in the PIE curve for xenon. The structural profiles are similar to those observed in comparable photoabsorption experiments.

Apart from the higher resolution relative to similar electron ionization experiments, photoionization has an additional advantage in that there is a finite cross section at threshold, making it easier to detect the actual ionization onset. This also applies to unimolecular fragmentation reactions and is a result of the different threshold laws for the two ionization processes. An ionization efficiency curve produced by photoionization is comparable to a first derivative of its electron-ionization counterpart.

The appearance energy (AE) is the minimum energy required to photoionize and fragment the molecule AB in the reaction shown in Equation [2] and, in the absence of any excess internal energy, can provide useful thermochemical information about the reactant and products.

$$AB + h\nu \rightarrow AB^{\bullet +} + e^- \rightarrow A^+ + B^\bullet + e^- \qquad [2]$$

However, the AE will not provide a true thermochemical value if there is any reverse activation energy ($E_{rev}$) associated with the decomposition. Furthermore, if the rate of fragment ion production is low for photon energies just above threshold, there will be a kinetic shift ($E_{kin}$). This is the energy in excess of the thermochemical threshold, including any reverse activation energy, for the ion to be detectable by the mass spectrometer. These effects are illustrated in **Figure 7**.

A kinetic shift can be minimized by increasing the detection sensitivity or by trapping the dissociating precursor ion and extending the time prior to mass analysis. An additional complication will arise when there is competition with another more favourable fragmentation process. This can also cause an increase in the observed AE and is often referred to as a competitive shift.

Unless reaction [2] is carried out at 0 K, the products will not be formed at any well-defined equilibrium thermodynamic temperature because the fragmentation is a unimolecular process with no means for the products to thermalize with their surroundings. Provided that there is no excess internal energy, the experimental AE for reaction [2] at a given temperature is given by

$$\begin{aligned} AE = {} & \Delta H_f^0(A^+) + \Delta H_f^0(B^\bullet) + \Delta H_f^0(e^-) \\ & - \Delta H_f^0(AB) - \Delta H_{cor} \end{aligned} \qquad [3]$$

where $\Delta H_{cor}$ is a thermal energy correction term to convert the products to the same temperature as the neutral precursor. $\Delta H_{cor}$ varies with different fragments, but is typically in the 10–30 kJ mol$^{-1}$ range at 298 K.

For threshold studies it is convenient to consider the ejected electron at rest at all temperatures, i.e. $\Delta H_f^0(e^-) = 0$. This stationary electron convention is

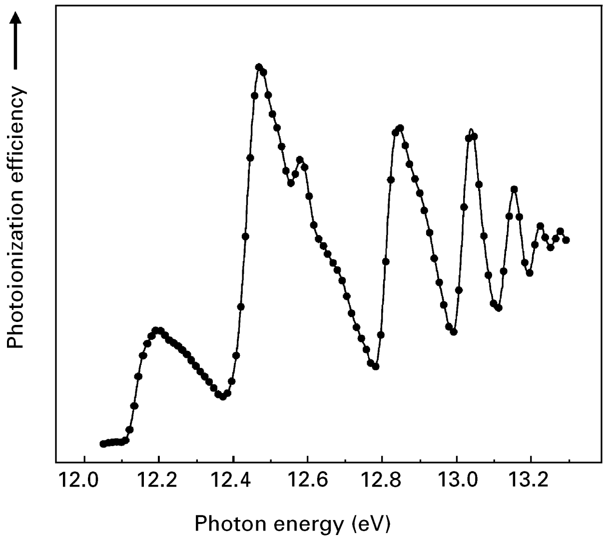



**Figure 6** Photoionization efficiency curve for $Xe^+$ showing the extensive autoionization structure above the ionization energy of 12.13 eV.

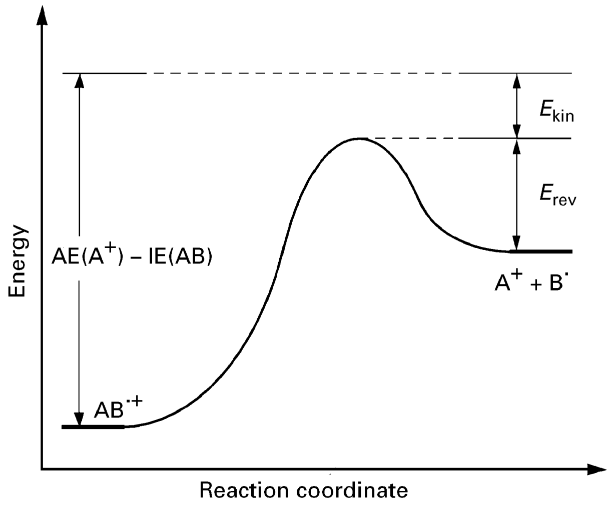



**Figure 7** A unimolecular reaction involving both a reverse activation energy and a kinetic shift.

most commonly used for cationic heats of formation derived from mass spectrometry experiments. However, there are many other data compilations that invoke a thermal energy convention, in which the electron has a translational energy commensurate with the reaction temperature. At 298 K this corresponds to a difference of 6.2 kJ $mol^{-1}$. Consequently, when using such information it is important to know which particular convention has been adopted.

## Photodissociation

### Sample introduction

The precursor ions used in a photodissociation experiment are usually generated via a mass spectrometer. This minimizes complications that can arise when ion–molecule reactions occur in the ion source. The most common method of initial ion production is electron ionization as this invariably produces larger numbers of primary ions than either photoionization or collisional activation techniques. The ionizing electron energy is kept as low as practicable to minimize the initial internal energy of the precursor ions, which are then mass filtered before introduction into the photodissociation region of the instrument. Variation of the ionizing energy can be used to prepare ions with different internal energies and also facilitates the determination of the electronic state(s) involved in the photoinduced transition(s). Because of space-charge effects the precursor ion concentration is considerably lower than that of a comparable beam of neutral molecules.

### Light sources

A high-intensity source of photons is required to produce detectable numbers of photofragment ions from a mass-selected ion beam. For this reason, most experiments have used a laser for the photodissociation process. Continuous lasers generally produce the highest average power outputs, with an excellent stability and a very narrow photon energy bandwidth (~$10^{-5}$ eV). They may be used to pump a dye laser to produce light from the near-infrared to the visible region of the spectrum. The use of a pulsed excimer laser to pump the dye laser can extend the wavelength range to cover the 1.5–3.5eV energy region, although this requires the use of over twenty different dye solutions. Frequency doubling and etalon bandwidth narrowing techniques can be used to further enhance the laser performance. Accurate wavelength calibration presents a problem for high-resolution studies. One procedure used in the visible region of 1.7–2.5 eV is to simultaneously record the iodine absorption as a reference, producing calibrations to better than $10^{-6}$ eV. For near UV studies, optogalvanic lines provide one of the few convenient methods available.

The interaction between the photon and ion beams can be achieved by either a crossed or a coaxial configuration. The crossed arrangement produces a better definition of the interaction region and is used when absolute cross section and angular distribution measurements are required. However, a better overlap of the two beams, and hence improved sensitivity, results from an arrangement in which both beams travel in parallel along the interaction region. This also permits the use of a Doppler tuning technique to improve the photon energy resolution and to reduce the problems associated with the ion beam thermal energy.

### Mass analysers

Photodissociation experiments are usually performed with either a beam or an ion-trap mass spectrometer. The beam instrument is a tandem mass analyser in which the first stage is used to mass select a given precursor ion. This ion beam is passed through a laser interaction region and the resulting photofragments are then energy or mass analysed by a second stage. A schematic diagram of a triple-quadrupole mass spectrometer with a coaxial configuration is shown in **Figure 8**.

Apart from restricting the laser–ion interaction time, fast ion beams, particularly those produced with a magnetic sector instrument, essentially restrict the lifetimes of ions that can be detected to less than 10 μs. This is not a limitation for an ion cyclotron resonance (ICR) or quadrupole ion storage (QUISTOR) trap, where ions can be stored almost indefinitely. However, the presence of relatively high concentrations of neutral species may result in interfering ion–molecule reactions taking place in the cell. The use of an ion trap with an external ion source can help to remove these competing processes and also allows photodissociation studies of precursor ions formed by ionization techniques other than electron ionization.

### Data processing

As in threshold photoionization experiments, photofragment ion concentrations are extremely low and it is necessary to utilize signal-enhancement techniques. Although phase-sensitive detection has been used, it is more usual to employ signal-averaged pulse counting. For photodissociation with a pulsed laser, an appropriate gating system synchronized

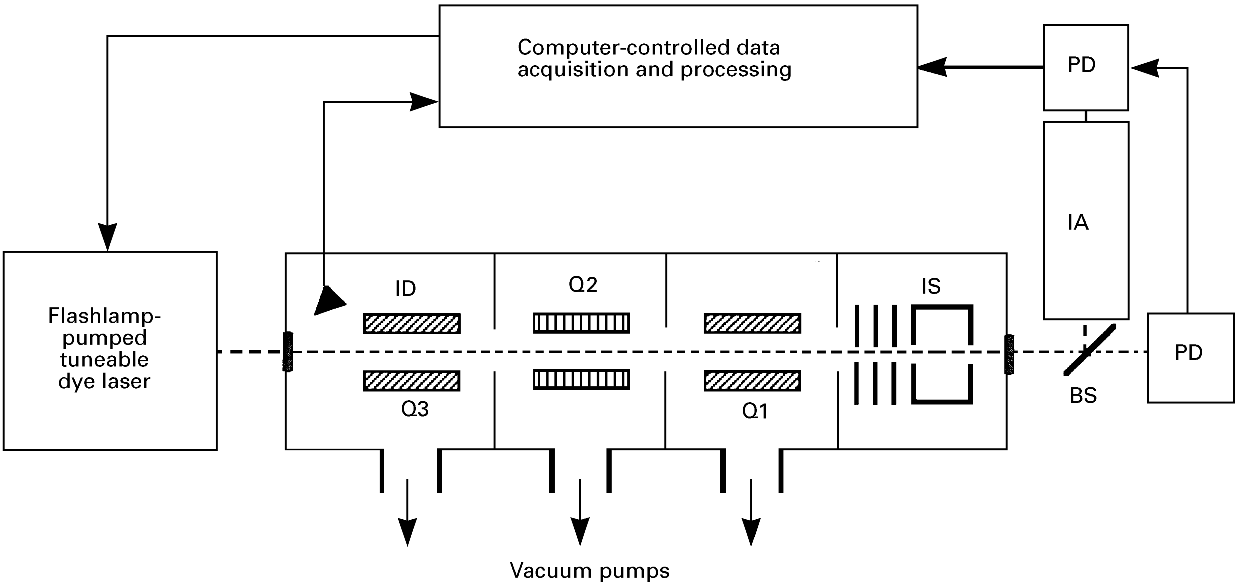



**Figure 8** Schematic diagram of a triple-quadrupole laser-induced photodissociation mass spectrometer operated in a coaxial configuration. IS is the ion source; Q1 is the precursor ion mass filter; Q2 is the ion photodissociation region; Q3 is the photofragment ion mass filter; ID is the off-axis ion detector; PD is a photon detector; BS is a beam splitter; and IA is an iodine absorption cell.

with each laser pulse can further improve the signal-to-noise ratio.

## List of symbols

$c$ = speed of light; $E$ = energy; $h$ = Planck's constant; $\Delta H_{cor}$ = thermal energy correction term; $\Delta H_f^0$ = standard enthalpy of formation; $\lambda$ = wavelength.

*See also:* **Ion Dissociation Kinetics, Mass Spectrometry; Ion Energetics in Mass Spectrometry; Ionization Theory; Multiphoton Excitation in Mass Spectrometry; Photoelectron–Photoion Coincidence Methods in Mass Spectrometry (PEPICO); Spectroscopy of Ions; Time of Flight Mass Spectrometers.**

## Further reading


Becker U and Shirley DA (1996) *VUV and Soft X-ray Photoionization*. New York: Plenum Press.

Berkowitz J (1979) *Photoabsorption, Photoionization, and Photoelectron Spectroscopy*. New York: Academic Press.

Bowers MT (1984) Ions and light. *Gas Phase Ion Chemistry*, Vol 3. New York: Academic Press.

Dunbar RC (1979) Ion photodissociation. In: Bowers MT (ed) *Gas Phase Ion Chemistry*, Vol **2**, pp 181–220. New York: Academic Press.

Dunbar RC (1996) New approaches to ion thermochemistry via dissociation and association. In: Adams NG and Babcock LM (eds) *Advances in Gas Phase Ion Chemistry*, Vol **2**, pp 87–124. London: JAI Press.

Futrell JH (1986) *Gaseous Ion Chemistry and Mass Spectrometry*. New York: Wiley.

Moseley JT (1985) Ion photofragment spectroscopy. *Advances in Chemical Physics* **60**: 245–298.

Ng CY (1983) Molecular beam photoionization studies of molecules and clusters. *Advances in Chemical Physics* **52**: 263–362.

Ng CY (1991) *Vacuum Ultraviolet Photoionization and Photodissociation of Molecules and Clusters*. Singapore: World Scientific.

Reid NW (1971) Photoionization in mass spectrometry. *International Journal of Mass Spectrometry and Ion Physics* **6**: 1–31.

Samson JAR (1967) *Techniques of Vacuum Ultraviolet Spectroscopy*. New York: Wiley.

Tecklenburg RE and Russell DH (1990) An evaluation of the analytical utility of the photodissociation of fast ion beams. *Mass Spectrometry Reviews* **9**: 405–451.

van der Hart WJ (1989) Photodissociation of trapped ions. *Mass Spectrometry Reviews* **8**: 237–268.

# Plasma Desorption Ionization in Mass Spectrometry

**Ronald D Macfarlane**, Texas A&M University,
College Station, TX, USA

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Californium-252 is the key ingredient in a mass spectrometric method known as $^{252}$Cf-plasma desorption ($^{252}$Cf-PD). The $^{252}$Cf prefix to 'plasma' distinguishes this method from other mass spectrometric methods that also use a plasma to generate ions. The method uses the energy released in nuclear fission to produce mass spectra of a wide variety of materials ranging from refractory inorganic matrices to samples of biological materials.

A high-energy fission fragment from the spontaneous nuclear fission of $^{252}$Cf, penetrating a sample of insulin, penicillin or an unknown protein, decomposes most of what is in its path, but occasionally, an intact molecular ion of the sample is ejected from the surface of the fission track into the vacuum of a mass spectrometer where its molecular mass is measured. This article outlines the essential features of the method including the nuclear properties of $^{252}$Cf, the interaction of fission fragments with solids, ion emission from fission tracks in solids, the design of the mass spectrometer, sample preparation, and how data are acquired. Some applications have been selected that have utilized some of the unique features of the method.

## A brief history

The development of mass spectrometry for the study of molecules of biological origin did not progress until the problem of forming intact gas phase molecular ions was solved. Volatilization of the molecules by heating led to decomposition. Field desorption ionization introduced in the late 1960s was the first solution to the volatilization–decomposition problem. The discovery of the $^{252}$Cf-PD method in the 1970s expanded the spectrum of molecules that could be studied by mass spectrometry. One of the first applications was determining the molecular mass of pharmaceuticals and natural products. The structures of these species were determined by a combination of other spectroscopic techniques, and the molecular mass of the intact molecule was used to verify the atomic composition of the structure that was deduced. In several cases, the molecular mass did not correlate with the proposed structure and a re-examination of the data from the other spectroscopic measurements led to the final structure. Some of the pharmaceuticals currently in use had their molecular masses first determined by $^{252}$Cf-PD. These include vancomycin, amphotericin, thiostreptin, and bleomycin, and the marine toxins, palytoxin, tetrodotoxin and the red-tide toxin.

With the introduction of fast-atom bombardment (FAB) in 1982, and matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI), most of the biomedical applications have been directed towards these methods. The $^{252}$Cf-PD method has been found to have wide applicability, including the study of refractory materials, catalysts, semiconductors and frozen gases. Electronics capable of measuring the timing of events with subnanosecond resolution (the time it takes for a single photon to travel 1 cm) is used by this method as well as event-by-event data acquisition using the computer to make decisions at the molecular level, the basis of correlation mass spectrometry, a unique feature of $^{252}$Cf-PD.

## Properties of californium-252

### Radioactive decay properties

Californium-252 is produced in a high-flux nuclear reactor by multiple neutron capture from $^{239}$Pu. It has a half-life of 2.6 years and decays by alpha-particle emission (97%) and spontaneous fission. In the fission process, two high-energy ions (fission fragments) are emitted in opposite directions (180°). (This feature of fission is utilized in the mass spectrometer.) Most of the energy released in the fission process appears in the kinetic energy of the fission fragments, 90–130 million electron volts (MeV). The high kinetic energy coupled with a high charge (15–25 times the electron charge) is responsible for what happens when a fission fragment interacts with a solid. An example of one of the more abundant fission fragments is the technetium ion with a kinetic energy of 100 MeV and charge of +20.

### The californium-252 source

Sources of $^{252}$Cf specifically designed for $^{252}$Cf-PD applications are commercially available. A carrier-free sample of $^{252}$Cf is electroplated onto a thin nickel foil in the form of a small circular spot with a diameter of 3 mm and is then sealed with a thin nickel foil to contain the $^{252}$Cf. (Fission fragment tracks formed in the source eject $^{252}$Cf atoms from the source; a process called self-transfer.) A source used in $^{252}$Cf-PD contains about 0.5 μg of $^{252}$Cf and, when located to within a few millimetres of a sample, will deliver 2000–3000 fission fragments into the sample per second. The useful life of the source in a $^{252}$Cf-PD mass spectrometer is of the order of 4 years.

## Properties of the fission track

In a $^{252}$Cf-PD measurement, fission fragments form a fission track in the sample to be analysed. The range of the fission fragment is of the order of 10 μm. If the sample is thinner, the fission fragment passes through the sample depositing 100 eV nm$^{-1}$ of energy in the form of electron excitation and ionization. **Figure 1** depicts the time evolution of the fission track for a 100 MeV Tc ion passing through a thin film and emerging with an energy of 70 MeV. Positive ions and electrons are formed creating a cylindrical high-temperature plasma that lasts for $10^{-15}$ s. If the material is a conductor, the electrons and positive ions (holes) recombine as shown on the right side of **Figure 1** and the energy is dissipated as heat. In the initial stages of the track development, atomic and molecular ions and photons are ejected from both ends of the fission track. After a period of a few nanoseconds, the track region returns to close to its original state with some evidence of dislocation of some of the atoms in the matrix.

If the matrix is an insulating material, the evolution of the fission track follows a different course.

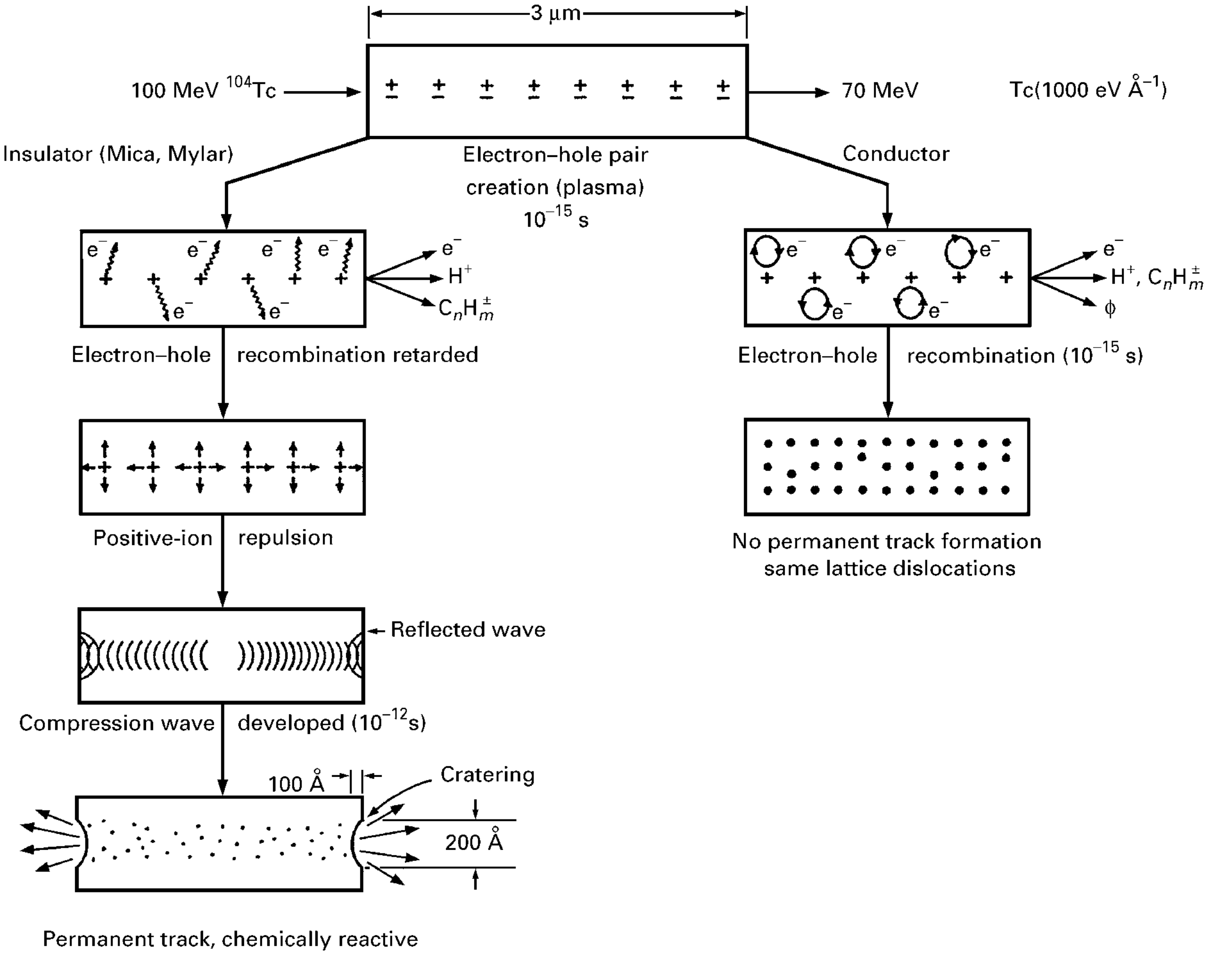



**Figure 1** Time evolution of a fission track in an insulator and conductor. A typical PD MS experiment will have an insulator as the energy deposition medium. Protonated molecules from a biological matrix are probably formed during the last stage when material ablation from the surface takes place. The craters depicted here have been observed using scanning tunnelling microscopy.

The electrons ejected from the fission track in the primary ionization are much less mobile and, as a result, the cylinder of positive charge expands due to Coulomb repulsion resulting in the development of a pressure pulse that propagates to the ends of the fission track. The high-pressure gradient generated at the surface of the sample results in ablation of material from the surface, leaving a crater approximately 20 nm in diameter and 10 nm deep. These craters have been studied in detail by scanning tunnelling microscopy. In the $^{252}$Cf-PD measurement, the molecules to be analysed are located on the surface of the matrix, and those that are located close to the fission track will be desorbed from the surface within a few trillionths ($10^{-12}$) of a second after the fission fragment passed through the sample. A permanent damage track is formed in the insulating matrix that is chemically reactive and can be etched to increase the size of the track for microscopic analysis. It is not known when in the sequence of events species are ejected that are detected that contribute to the mass spectrum. For penicillin, 1500 neutral molecules, on average, are ejected from a single fission track and, typically, 1 track in 10 produces an intact penicillin molecular ion.

## Subnanosecond chemistry and the fission track

Ions that are formed as a result of excitation of the sample by fission fragments are detected in the $^{252}$Cf-PD measurement. They are formed in various stages of the evolution of the fission track, from the surface, and in the gas phase immediately above the surface (called the selvedge region). The high density of ejected species close to the surface supports gas phase chemical reactions and the ejected clusters of molecules in an excited state are also a source of ions that comprise the mass spectrum. **Figure 2** summarizes the processes that can occur leading to the formation of gas phase ions from a perspective that depicts the surface region around the track. The primary fission track with a diameter of 0.1 nm is at the centre of the track. Emanating from the track is the region excited by secondary electrons extending to 6 nm diameter. The range of vaporization extends to 20 nm.

It is likely that the species contributing to the formation of molecular ions of penicillin or insulin are located in the peripheral regions of the fission track. The molecular ions that are detected include radical cations and anions of redox species such as chlorophyll, protonated and deprotonated molecules

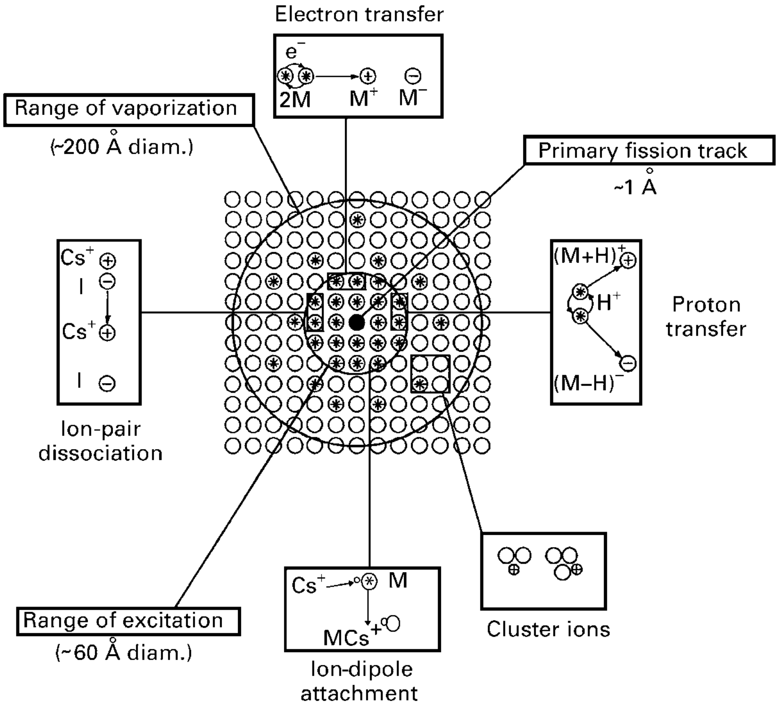



**Figure 2** Depiction of the surface of a nascent fission track showing the various modes of ionization processes taking place based on the nature of the gas phase ions observed using different matrices.

such as in the case of insulin, and cations formed from the addition of Na ions to the molecule. The mechanism of the formation of these molecular ions is an open question. There is strong evidence that the important molecular-ion-producing chemistry is taking place in the gas phase, in the selvedge region.

## Spectrum measurement

### Sample preparation

Most of the molecules of the species to be analysed (analyte) that contribute to the mass spectrum are located in the surface region of the sample. By using adsorption at a solution–solid interface, a uniform layer of the analyte molecules can be made on the surface of a substrate. The substrate is chosen depending on the affinity of the analyte. For example, nitrocellulose is an effective matrix for adsorbing proteins from solution. Hydrophobic substrates, and cationic and anionic substrates have also been used. The sequence for sample preparation is outlined in **Figure 3**.

A thin metallized polymer film is stretched over a sample holder providing mechanical support for the sample to be analysed. (In the $^{252}$Cf-PD measurement, the conducting layer is used to establish an electric field for accelerating the ejected ions into the mass spectrometer.) The adsorbing substrate is then deposited as a thin layer on the metallized surface. If nitrocellulose is used, it is electrosprayed from an acetone solution onto the surface. A solution of the analyte is then deposited onto the adsorbing layer. The solution is then removed, and the adsorbed layer washed to remove species that do not adsorb to the surface. The sample is then mounted into the mass spectrometer and analysed.

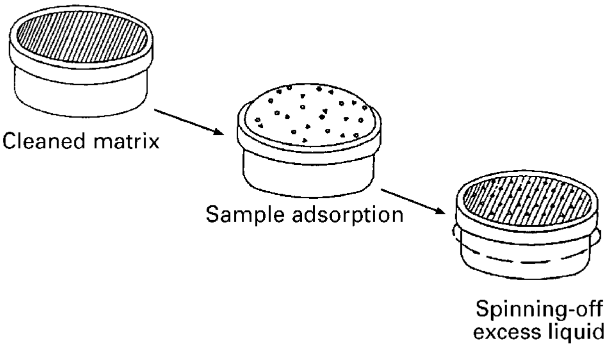



**Figure 3** The steps involved in the preparation of a sample for PD MS analysis. A matrix with a high adsorption affinity for the analyte is first deposited on the surface of a thin support film. A solution containing the analyte is deposited on the matrix layer and analyte molecules are adsorbed on the surface. The solution is then removed and the matrix–analyte layer washed to remove impurities. Reproduced from Macfarlane, R.D. (1988) Trends in Analytical Chemistry, **7(5)**, 179–183 with permission from Elsevier Science.

### The mass spectrometer

The arrangement of $^{252}$Cf source and sample foil in the mass spectrometer is shown in **Figure 4**. The sample is positioned 4 mm in front of the $^{252}$Cf source and behind a high-transmission metal grid at ground potential. In high vacuum, a voltage is then applied to the sample foil. When a fission fragment from the $^{252}$Cf source passes through the sample foil ions are ejected from the surface of the sample foil and are accelerated through the grid with the same energy to charge ratio into a 1 m long tube called the flight tube. An ion detector, located at the end of the flight tube generates an electronic pulse every time an ion hits the detector.

The mass of an ion is determined by measuring how long it takes for the ion to travel the distance through the flight tube to the detector, since all ions have been accelerated through a common electric field and have the same energy. (Some ions, particularly protein molecular ions, are multiply charged. In these cases, mass to charge ratio is measured and the energy of the accelerated ion is the voltage + charge product.) The travel time is referred to as the time-of-flight (TOF). To measure the TOF of an ion, a ‘time zero’ marker must be generated. The complementary fission fragment, the companion of the fission fragment passing through the sample foil, provides this marker. Travelling in the opposite direction, it is detected by an ion detector located behind the $^{252}$Cf source that first identifies it as a fission fragment and not an alpha-particle by the amplitude of the electronic pulse that is generated, and then sends a timing pulse to an electronic module called a time interval digitizer. For the fission track depicted in **Figure 4**, five ions were ejected from an adsorbed layer of insulin molecules. This particular track, which released a protonated insulin molecule, two fragment ions from the β-chain, a sodium ion and a hydrogen ion were identified as contributors to the mass spectrum of the insulin sample under study. For illustration, what happens to the data generated by these five ions as the analysis passes into the data acquisition phase, is followed in the subsequent sections.

### Time-of-flight measurement and spectrum

The time interval digitizer (TID) is a fast electronic clock that has a resolution of 80 ps. The electronic pulse from the fission detector travels to the fast clock through a wire at the speed of light and arrives at the

TID while the ions are still travelling down the flight tube. The hydrogen ion arrives first at the ion detector at the end of the flight tube because it has the lowest mass and highest velocity. The electronic signal generated by this ion is sent to the TID which records the time lapsed since the arrival of the time-zero marker from the fission fragment. In this particular case, the lapsed time was determined to be 977.260 ns as shown in **Figure 5**. The clock keeps running and as the other three ions arrive, their flight times are also recorded in the same data set. After the maximum time has lapsed, the clock is reset for the next event and the data set is sent to an interrogation module, an interface between the clock and a computer, for the initial procession step. This scenario is repeated 2000 times a second as 2000 fission fragments pass through random sites in the insulin monolayer, generating the ions whose TOF values will comprise the mass spectrum for this sample.

### Event-by-event analysis

The interrogation module is the link between experimentalist and the computer. It can be programmed to provide a variety of functions. Some of these options are indicated in the flow chart shown in **Figure 6**. The first option is the sorting of those ions used in obtaining a mass calibration, converting TOF values into *m*/*z* values. In the experiment used in the illustration, a decision was made to use the sodium and hydrogen ions as calibration ions, and it was determined by previous measurements that the TOF values for these ions were in time windows of $980 \pm 10$ ns and $1880 \pm 10$ ns respectively. TOF values were stored 'as is' as a separate file. Counting the number of ions detected for a particular event, truncation of the significant figures for a TOF value, and detecting whether particular sets of species of ions are formed in the same event are some of the options that have been used in $^{252}$Cf-PD studies. Examining the ions formed from a single fission track and interacting with the data set at the molecular level is a unique feature of the $^{252}$Cf-PD method.

### Generating a mass spectrum

What comes out of the interrogation interface is stored in a computer. For the insulin measurement being used for illustration, the TOF values for the calibration ions are identified by the interrogation module and stored in a separate computer file which sorts by TOF value and records the frequency distributions of TOF values. All other TOF values are truncated to 1 ns resolution and stored in a file that will be used to generate the mass spectrum over the recorded mass range. The sequence of events is depicted in **Figure 7**.

The TOF spectrum is a composite of contributions from $10^6$ to $10 \times 10^6$ fission fragments passing through the sample foil in a typical measurement requiring 10–100 min of running time. After the measurement is completed, the accumulated TOF spectra for the calibration ions is displayed as shown in **Figure 7** and the mean TOF values are determined. A calibration equation is calculated using the known masses of the calibration ions. The accumulated TOF spectrum for the range of TOF values set for the time interval digitizer (e.g. 32 μs) is then processed. A portion of the TOF spectrum is shown in the lowest

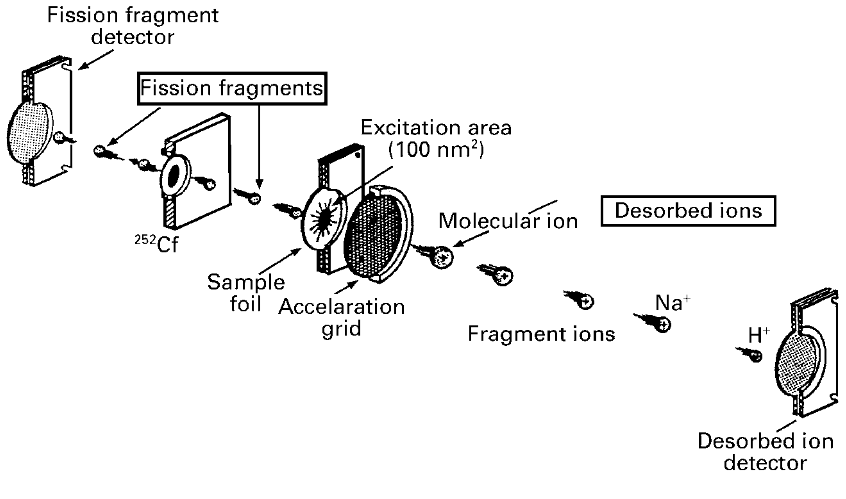



**Figure 4** The geometrical arrangement of the essential components of a PD MS system. In the event portrayed here, five ions were ejected from the fission track. Subsequent measurement and processing of the data for these five ions are illustrated in **Figures 5**, **6** and **7**.

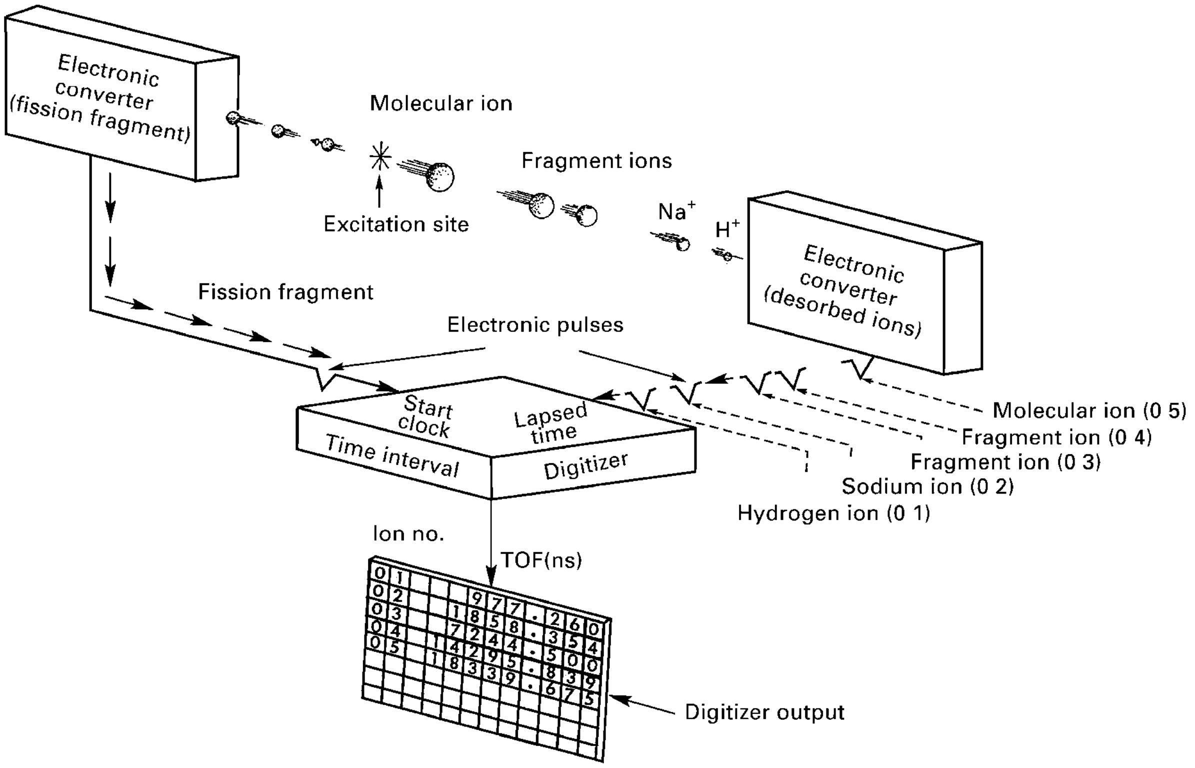



**Figure 5** Conversion of ion detection into electronic pulses that are transmitted to a fast clock. The measured time-of-flight for the five ions displayed at the digitizer output is then transferred to a computer for processing.

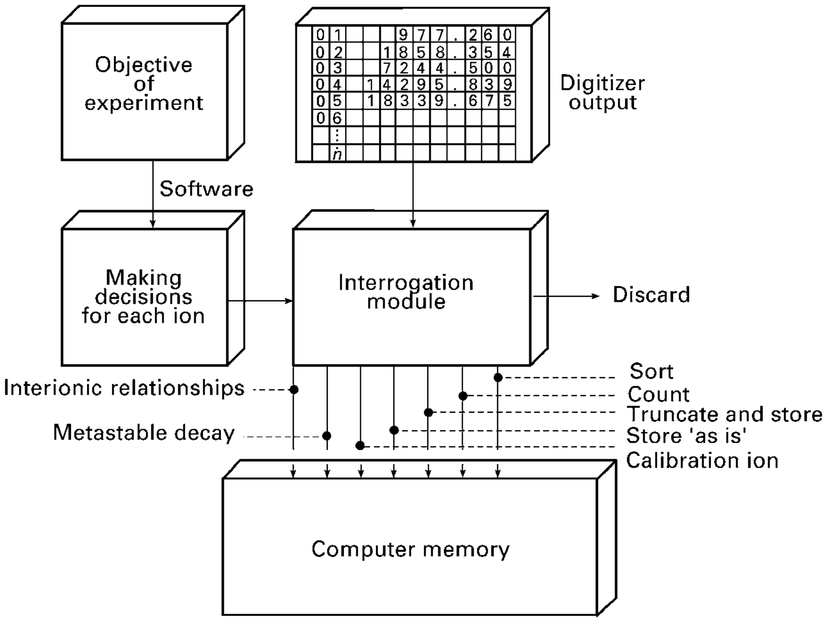



**Figure 6** Features of event-by-event analysis. The data from the digitizer are sorted according to the objective of the experiment. A standard PD MS analysis involves identifying the calibration ions and storing the data with 78 ps resolution. The data for other ions are truncated to 1 ns resolution and stored. Examples of other software-controlled options are also listed.

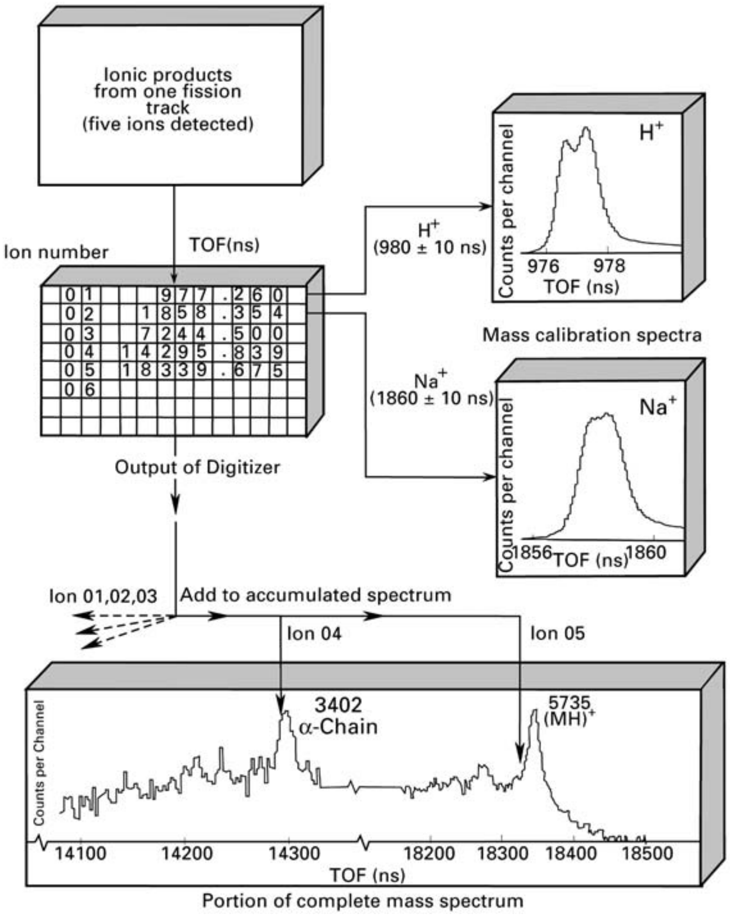



**Figure 7** Example of the ion sorting process. The software has identified ions 1 and 2 in the event depicted in **Figure 4** as $H^+$ and $Na^+$ calibration ions and has stored these data in separate files. It then truncates the TOF values for all five ions and stores these data in a separate file that will be used to generate the full mass spectrum. This process is repeated 2000 times per second over a period of a few hours. The mass spectrum is a composite of the accumulation of the ions emitted from $10^6$ to $10 \times 10^6$ fission tracks.

block in **Figure 7**. Peaks in the spectrum are identified, and mean TOF values for each peak determined. These TOF values are converted into *m*/*z* ratios and are tabulated. From this analysis, it was determined that ion number 5 depicted in **Figure 4** was from a protonated insulin molecule, and ion numbers 3 and 4 were fragment ions from the β-chain of a neighbouring insulin molecule close enough to be ejected by the same fission track. It is possible that the two insulin molecules were ejected from the surface as part of a cluster of molecules and, in the region just above the surface, processes took place within the cluster resulting in the formation of the protonated insulin molecule with the protonated second insulin ion acquiring enough internal excitation to dissociate into a protonated fragment of the β-chain and, presumably, a neutral fragment of the α-chain. The formation of these two ions from the same fission track is an example of a correlation phenomenon and is the consequence of two insulin molecules being within nanometres of each other on the surface of the matrix.

## Characteristics of a mass spectrum

### Common features for all matrices

Although ions from the fission track are emitted from both ends, generally only the ions emitted from one end are used to produce a mass spectrum. Some of the ions emitted are not structurally related to the sample molecules being studied but are to high-energy processes that are characteristic of a high-temperature plasma. These ions include $H^+$, the dominant ion in the spectrum, $H_2^{\bullet +}$, $H_3^+$, electrons, $H^-$, and a set of positive and negative hydrocarbon ions in the *m*/*z* range 12–100. These ions originate from impurity molecules on the surface of the sample and are formed in high-energy gas phase processes near the surface.

### Involatile organic samples

Most of the studies carried out for organic samples have involved involatile, highly polar biological molecules because this class of compound was not generally amenable to mass analysis prior to the development of $^{252}$Cf-PD. Peptides, proteins, synthetic oligonucleotides are some of the classes of compound that were characterized by this method. These compounds produce protonated molecules and ions formed by alkali metal ion attachment to the molecule. In addition, many of these species form deprotonated negative molecular ions as well. By measuring the *m/z* values of these species, the $M_r$ of the parent molecule can be identified. In addition to the protonated molecule, an extensive pattern of fragment ions is also formed due to the dissociation of ions that have acquired a high level of internal excitation in the desorption/ionization process. This feature of $^{252}$Cf-PD, the identification of $M_r$ and acquisition of an extensive high-excitation fragmentation pattern in a single measurement is unique amongst the desorption/ionization mass spectrometric methods.

### Involatile inorganic species

The $^{252}$Cf-PD mass spectrum of metal halides and oxides consists of a family cluster ions of these compounds extending to over *m/z* 10 000, produced by the ejection of small domains of the crystal lattice in the region around the fission track. In addition, cluster ions are also observed that do not correlate with the composition of the crystal lattice, indicating that some of the cluster ions are involved in gas phase reactions in the desorption plume. One of the unique applications of $^{252}$Cf-PD is the elucidation of the composition of large transition metal cluster compounds with $M_r$ values approaching $10^5$.

## Correlations and ion multiplicity

While most of the $^{252}$Cf-PD measurements have involved the recording of mass spectra, a higher level analysis is made possible by the event-by-event feature of the data acquisition, as described above. In this section, two examples are given that make use of this feature. The evolution of the fission track and the processes leading to ion formation are complex and variable. Each fission track is unique. Some tracks generate a large number of ions while others produce none. The generation of a protonated insulin molecule is a rare event, observed in one in a hundred fission tracks. To learn more about those tracks which produce this ion, the interrogation module can be programmed to determine how many other ions are formed in the event that produced the protonated insulin molecule. This experiment is outlined in **Figure 8**.

A time window is set up in the interrogation module encompassing the TOF values for the protonated insulin molecule. The data file formed for each event is interrogated to determine the number of ions detected. Two scenarios are portrayed, one in which four ions are detected but none with the TOF of insulin. The multiplicity of this event (4) is stored in a data file. In the second event, 10 ions were detected, and one was within the TOF window of insulin. The multiplicity of this event was then stored in a separate data file. After $10^6$ fission fragments had passed through the sample, two multiplicity histograms were generated; the blocks shown in **Figure 8**. By recording data at the event-by-event level, it was possible to determine that some fission tracks evolve in a manner that enhances the ion emission probability and these tracks have a greater probability for generating the large protonated insulin molecule.

### $^{252}$Cf-PD as a microprobe for heterogeneity at the nanometre level using correlation analysis

Most of the samples that have been studied by $^{252}$Cf-PD have been chemically homogeneous, pure samples of a biological molecule or inorganic matrix. Under these circumstances, if two or more ions are detected in a single fission fragment event, desorption/ionization is correlated because they were formed from the same desorption plume. The plume contains approximately 1500 identical molecules that were in close proximity in the sample matrix and within the excitation volume of the fission track. However, if the sample is not chemically homogeneous, microscopic heterogeneity at the nanometre level can be measured by studying the composition of the desorption plume at the event-by-event level. If the sample consists of two components that are homogeneous at the nanometre level, ions from both components are ejected from the same fission track. Event-by-event analysis will show that when ions from one of the components are desorbed it is highly likely that ions from the other component will be desorbed from the same track because both components are in close proximity at the nanometre level. If the components exist in separate domains separated by more than 10 nm within the matrix, fission tracks formed in one domain only desorb ions characteristic of that domain. Although the total spectrum is a composite of ions from both domains, the event-by-event analysis makes it possible to determine what fraction of the two matrices reside with the 10 nm dimensions of the fission track.

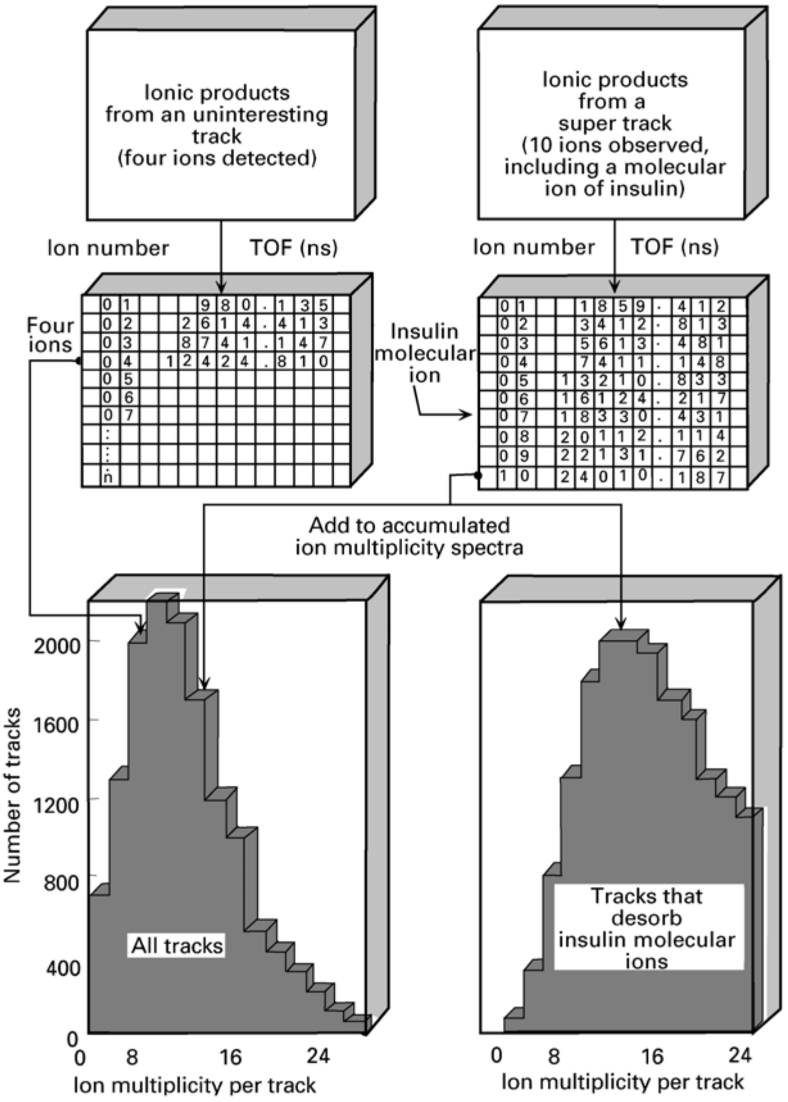



**Figure 8** Example of a correlation measurement. The experiment was to determine how many ions are emitted from a fission track that produces a protonated insulin molecule. The interrogation module was programmed to count the number of ions detected for each event where an insulin molecular ion was detected, and store that sum in a separate file. The ion multiplicity for all other events was then stored in a separate file. The conclusion of this study was that tracks that desorb large molecular ions also desorb a much larger number of other ions than average. Other types of correlation measurements include sorting events where one or two ions of a particular type are emitted from the same track.

## A historical perspective (1999)

The $^{252}$Cf-PD method has been largely supplanted by MALDI and ESI MS for the analysis of complex biological molecules because these methods are more efficient and widely applicable. The method has stimulated considerable interest in the field of particle-induced desorption and understanding of the mechanisms of the desorption/ionization process. Scanning tunnelling microscopy has been used to characterize the craters produced by fission fragments and high-energy heavy ions from particle accelerators have been used to study the influence of charge, mass and energy on the desorption/ionization process. High-energy cluster ions, including $C_{60}$ fullerene ions, are being used to study the chemistry and physics of particle-induced energy deposition and transfer and ion emission from solids.

*See also:* **Fast Atom Bombardment Ionization in Mass Spectrometry; Proton Microprobe (Method and Background); Time of Flight Mass Spectrometers.**

## Further reading


Da Silveira EF, Duarte SB and Schweikert EA (1998) Multiplicity analysis: a study of secondary particle correlation. *Surface Science* **408**: 28–42.

Della Negra S and Le Beyec Y (1983) Secondary ion emission from surfaces of solids bombarded by high energy heavy ions. Applications in analytical mass spectrometry and studies with beams from accelerators. *Nuclear Science Applications* **1**: 569–590.

Demirev P (1995) Particle-induced desorption in mass spectrometry. Part I Mechanism and processes. *Mass Spectrometry Reviews* **14**: 279–308.

Demirev P (1995) Particle-induced desorption in mass spectrometry. Part II Effects and applications. *Mass Spectrometry Reviews* **14**: 309–326.

Eriksson J, Rottler J and Reimann CT (1998) Fast-ion-induced surface tracks in bioorganic films. *International Journal of Mass Spectrometry and Ion Processes* **175**: 293–308.

Fritsch H-W, Schmidt L, Koehl P, Jungclas H and Duschner H (1993) Application of 252-Cf-PDMS in dental research. *International Journal of Mass Spectrometry and Ion Processes* **126**: 191–196.

Jonsson GP, Hedin AB, Hakansson PL *et al* (1986) Plasma desorption mass spectrometry of peptides and proteins adsorbed on nitrocellulose. *Analytical Chemistry* **58**: 1084–1087.

Jungclas H, Koehl P, Schmidt L and Fritsch H-W (1993) Quantitative matrix assisted plasma desorption mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **126**: 157–161.

Macfarlane RD (1983) Californium-252 plasma desorption mass spectrometry – Large molecules, software, and the essence of time. *Analytical Chemistry* **55**: 1247A–1283A.

Macfarlane RD, Hill JC, Jacobs DL and Geno PW (1989) 252-Cf-Plasma desorption mass spectrometry – Past and present. *Advances in Mass Spectrometry* **11**: 3–21.

Matthaus R, Moshammer R, von Hayn G, Wien K, Della Negra S and Le Beyec Y (1993) Secondary ion emission from various metals and semiconductors Si and GaAs induced by MeV ion impact. *International Journal of Mass Spectrometry and Ion Processes* **126**: 45–58.

Roepstorff P and Sundqvist B (1986) Plasma desorption mass spectrometry of high molecular weight biomolecules. In: Gaskell, SJ (ed) *Mass Spectrometry in Biomedical Research*, pp 269–285. Chichester, UK: Wiley.

Sundqvist B and Macfarlane RD (1985) 252-Cf-Plasma desorption mass spectrometry. *Mass Spectrometry Reviews* **4**: 421–460.

Van Stipdonk MJ and Schweikert EA (1995) High energy chemistry caused by fast ion–solid interactions. *Nuclear Instruments and Methods in Physics Research B* **96**: 530–535.

Van Stipdonk MJ, Schweikert EA and Park MA (1997) Coincidence measurements in mass spectrometry. *Journal of Mass Spectrometry* **32**: 1151–1161.

Wien K (1989) Fast heavy ion induced desorption. *Radiation Effects and Defects* **109**: 137–167.


# Platinum NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# Polarimeters

*See* **ORD and Polarimetry Instruments.**

# Polymer Applications of IR and Raman Spectroscopy

**CM Snively** and **JL Koenig**, Case Western Reserve University, Cleveland, OH, USA

Copyright © 1999 Academic Press



The two techniques of Raman and infrared (IR) spectroscopy have some similarities, yet are quite different in a number of ways. They provide complementary vibrational information. The reader is referred to the articles on the fundamentals of these techniques for details. Raman and IR will here be collectively referred to as VS (Vibrational Spectroscopy), except where details of a particular technique warrant discussion. It should also be noted that most of the techniques mentioned can also be used with near-IR radiation, but specific examples will not be cited.

Among the large number of applications of VS, applications to polymeric systems are especially interesting because of the wide variety of chemical structures and physical ordering that is present in polymer systems. The article is arranged as follows. First, the application of VS to the determination of chemical properties of polymeric systems will be illustrated. In particular, the use of VS as an identification tool for complex polymeric systems and the application of VS to the various chemical reaction processes will be detailed. Then the application of VS to the determination of polymer structure on a wide range of length scales will be explained, with particular emphasis on the determination of stereoregularity, chain conformation and crystallinity. The remainder of the article will focus on the study of dynamic properties of polymers such as diffusion and rheological properties and current topics such as millisecond time-resolved and microimaging applications.

## Spectroscopic considerations unique to polymers

In contrast to small-molecule compounds, in polymer molecules the atoms are all linked together to form long chains. The presence of such long chains causes additional vibrational modes to be present that do not exist in small-molecule analogues. These arise owing to the vibrations of the chain as a whole. This topic is best treated using classical physics and normal coordinate analysis, which is beyond the scope of this article. A more indepth discussion of these topics can be found in the books by Koenig listed in Further reading. In addition, the long chains can possess ordering along the chain as well as between neighbouring chains. This quality is also unique to polymers and is responsible for many of the physical properties that make polymers the material of choice for a wide variety of applications.

Several properties of polymers complicate their analysis via VS. First, a problem unique to transmission IR spectroscopy is that polymers are very strong absorbers of IR radiation. Therefore, in order to be within the linear region of Beer's law, an extremely thin polymer film must be used in transmission. A good rule of thumb is to keep the thickness below 5 μm. While the production of such thin films is possible in the laboratory, it must be remembered that most commonly encountered polymer systems are much thicker than this. As a result, the most commonly used industrial IR techniques are reflectance techniques such as attenuated total reflectance (ATR) or reflection–absorption spectroscopy (RAS), which have much smaller effective optical path lengths, typically on the order 1 μm or below.

A problem unique to Raman spectroscopy is the fact that most polymers fluoresce strongly when exposed to laser radiation. This problem can be reduced by using Fourier transform and resonance Raman techniques. Because of this and other difficulties associated with Raman spectroscopy, the quality of Raman spectra of polymers is typically less than that of IR spectra. Therefore, it is not surprising that a quick search of the literature reveals many more quantitative studies of polymers using IR than using Raman.

Using a combination of IR and Raman spectroscopies can yield valuable information. IR spectroscopy is sensitive to any chemical groups that possess a significant dipole moment, such as C–H, and C=O, which are commonly found in polymer side groups. Raman spectroscopy is more sensitive to highly polarizable groups such as C–C and C=C, which are commonly found in the polymer chain backbone. Thus, when used together, IR and Raman spectroscopy can be used to gain more information than is available from either of the individual techniques.

## Complete structure determination

Vibrational spectroscopy can be used in the complete chemical and physical structure determination of polymers. Four size scales of structure and orientation found in polymers are covered. The most basic is the chemical identity of the chains, including chemical groups present, monomer sequences and stereoregularity. This is followed by studies of the local conformation of individual chains and interactions between chains, and finally orientation induced by the application of macroscopic forces.

### Chemical identity of polymer chains

The techniques used for identification of the basic chemical structure of polymer chains are similar to those used for small molecules. The difference is that there is not a single chemical structure present; rather there is a distribution of chemical structures. The semirandom statistics involved in the polymerization reaction produce a distribution of molecular masses. Therefore, an average molecular mass is reported. Although chromatography and light scattering are most commonly used for molecular mass determination, IR can provide complementary chemical structure information. A number-average relative molecular mass $M_n$, can be obtained by determining the number of end groups present for a given amount of polymer by

$$M_n = \frac{2}{E_{\text{end}}}$$

where $E_{\text{end}}$ is the number of mass equivalent weights of end groups per gram of polymer, as determined from a simple Beer's law treatment. This method works if the structure of the end groups and the nature of the polymerization process are known.

Another unique property of polymerization and copolymerization reactions is the presence of a distribution of chemical connectivities in the final polymer, which come about from the orientation of monomer addition. This is commonly referred to as regiochemistry or regioisomerism. In a typical monosubstituted vinyl polymerization, the monomer has two options of addition to the growing chain, commonly referred to as head or tail addition, depending on whether the propagating chain attacks the monomeric carbon with the pendant group or the unsubstituted carbon, respectively. The ratio of head-to-head, head-to-tail, etc. groups present in the final polymer can be determined by comparing the ratios of spectral bands associated with these linkages. In a similar manner, copolymer systems can be analysed to determine the statistics of the polymerization process. NMR is the most commonly used technique for this type of analysis, but not all polymers can be analysed using this method. VS can be used in these cases and also to verify the NMR results.

Perhaps the most common application of VS in the determination of chemical makeup in polymeric systems is the identification of components in complex polymer mixtures. Polymeric products are rarely composed of a single component. There are always additives present that aid in processing, appearance, adhesion, chemical stability or other properties important to the function of the final product. In an industrial setting, it is important to be able to determine both the identity and quantity of polymers and additives in a specific formulation for quality control purposes. This can be a fairly routine operation if tools such as spectral libraries are utilized. In this method, a computer search algorithm compares a spectrum with a catalogue of standard spectra to determine the identity of the compound or compounds present. Advanced statistical techniques, such as partial least squares (PLS) and principal-component analysis (PCA), are also often used to identify known and unknown components in polymeric systems. The details of these methods are described elsewhere in the Encyclopedia.

In certain polymerization processes, the final polymer may contain side branches. The most notable example of this is polyethylene, which commonly contains short branches that result from chain transfer reactions during the polymerization process. For each type of branch, up to six carbons in length, the methyl rocking band around the 900 $cm^{-1}$ has a unique frequency. By comparing the intensities of these branch peaks to peaks associated with the main chain, the relative amount of each type of branch can be determined.

Owing to the length of typical polymer chains, the arrangements of side groups along the chain becomes important in the determination of the chemical properties of the polymer. The polymerization process permanently locks in this arrangement of side groups. The persistence of a specific pattern along the entire chain is known as stereoregularity. Stereoregular effects appear in vibrational spectra as spectral peak splitting or other changes in band shape. An example is shown in **Figure 1**, which shows the IR spectra from the three pure stereoisomers of polystyrene. As can be seen, the bands around 1070, 550 and 900 $cm^{-1}$ are quite different for each stereoisomer, revealing that the local molecular environments are different in each case.

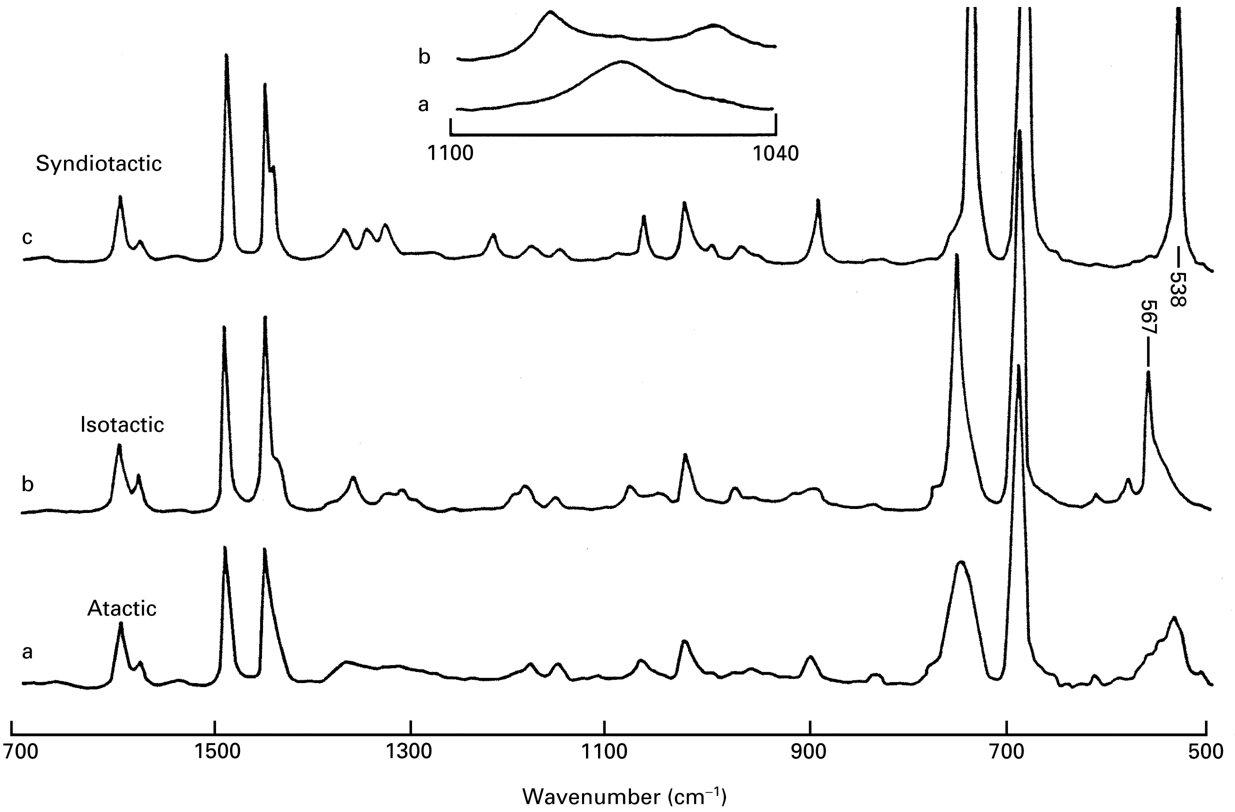



**Figure 1** Infrared spectra of the pure stereoisomers of polystyrene.

### Orientation of a single polymer chain

The next highest level of order in polymer chains is the presence of rotational isomers, which result from stable configurations of the polymer chain. The presence of *cis* and *trans* configurations and the relative ratios of each can be determined using VS because each of these structures possess a unique spectral band. Owing to the complementary nature of IR and Raman spectroscopy, a wealth of information about the chain conformation can be determined. Raman is particularly useful in this application owing to its enhanced sensitivity to the local environments around carbon–carbon bonds in the chain backbone. Using a simple thermodynamic treatment, this technique can be extended to determine quantitatively the activation energy barrier between the two conformers. The kinetic rate constant, $k$, can be determined from spectral information using the expression:

$$k = \frac{A_1/a_1}{A_2/a_2}$$

where $A_1$ and $A_2$ are the absorbances of the bands associated with each conformer, and $a_1$ and $a_2$ are the corresponding extinction coefficients of the bands. By acquiring spectra at a range of different temperatures, the activation energy, $\Delta G$, can be determined from the well-known relation

$$\Delta G = RT \ln k$$

where $R$ is the gas constant and $T$ is the temperature.

Stable crystalline forms arise from the propagation of a particular sequence of conformations along the entire length of the chain, resulting in a helical arrangement. The most commonly used method for analysing crystalline materials is X-ray diffraction, which can detect long-range crystalline order, like that found in semicrystalline polymers such as poly(vinylidene fluoride) and polyethylene. VS probes the local molecular environment, and therefore has the advantage over X-ray methods that it can detect the short-range order present in the amorphous phase. The presence of helical order shows up in a predictable manner in both IR and Raman spectra.

VS has been applied to polystyrene, which can possess several localized crystalline forms. In this particular case, order results from a planar zigzag

arrangement of the chains that pack together to form structures that possess either trigonal or orthorhombic symmetries. As subtle as this ordering may seem, it can be detected effectively using VS. This can be seen in **Figure 2**, which shows that the pure crystalline forms of syndiotactic polystyrene have distinct IR spectra. This particular example illustrates the spectral subtraction technique used to determine the spectra of pure crystalline forms of polymers. The spectrum from a completely amorphous sample is subtracted from the spectrum of a partially crystalline sample to yield the 'pure' crystalline spectrum. This is repeated for all crystalline forms. Using these spectra, the crystal forms present in any sample can then be determined. This method of spectral subtraction is commonly used to isolate the spectral contributions of individual components to determine which structures are present.

### Interactions between chains

To achieve desirable properties, polymers are often blended to form stable mixtures. The formation of successful polymer blends depends on the presence of favourable interactions, such as hydrogen bonding, between chains. These strong interactions show up in the vibrational spectra as the appearance or disappearance of peaks or as peak shifts. The most commonly studied blend systems involve polymers with carbonyl groups, which show different carbonyl bands for hydrogen-bonded and nonbonded groups. In these studies, the degree of compatibility can be determined by monitoring the spectral shifts and relative peak intensities of the bonded and nonbonded forms for a range of concentrations and temperatures. With the proper mathematical treatment, thermodynamic and kinetic parameters for these processes can be determined. The details are described in depth in the book by Coleman *et al.* (see Further reading). This technique has also been used to probe the interactions between polymers that do not contain such obvious interacting groups, such as the classic totally miscible pair polystyrene and poly(phenylene oxide).

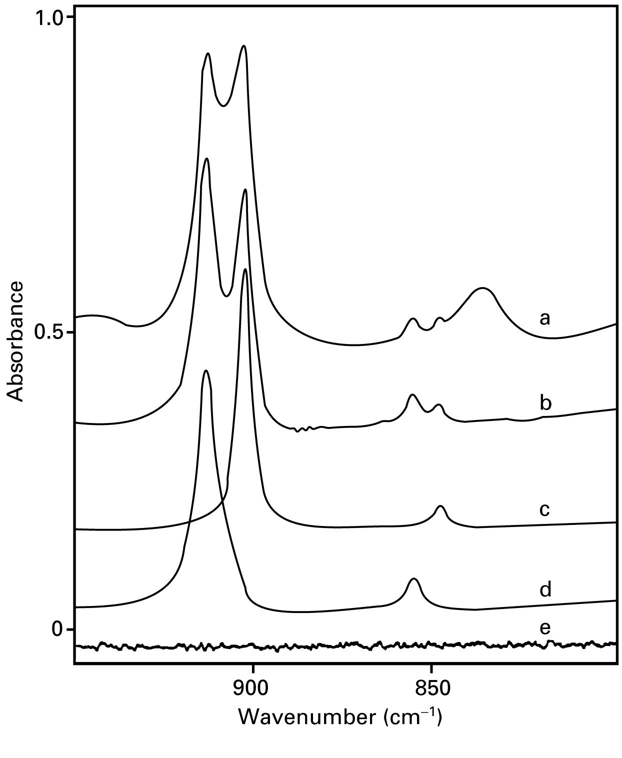



**Figure 2** Infrared spectra of polystyrene with (curve a) a mixture of crystalline and amorphous forms, (curve b) a mixture of crystalline forms only, (curve c) one 'pure' crystalline form obtained by spectral subtraction, (curve d) another 'pure' crystalline form obtained by spectral subtraction and (curve e) subtraction result of curves b – c – d. Reproduced with permission of John Wiley & Sons Limited from Musto P, Tavone S, Guerra G and De-Rosa C (1997) Evaluation by Fourier transform infrared spectroscopy of the different forms of syndiotactic polystyrene samples. *Journal of Polymer Science B* **35**: 1055–1066.

### Orientation induced by processing

The production of polymer products typically involves processes in which the polymer is raised above its glass transition temperature ($T_g$) and forced into some desired shape. The most common of these methods include injection moulding and blow moulding and depend on the desired geometry of the final product. During this process, the highly flexible polymer chains in different parts of the mould orient according to the local shear and elongational stresses. These stresses cause orientation of both the crystalline and amorphous regions of the polymer and can produce ordered regions with undesirable anisotropic mechanical properties in the final products. This orientation can be measured using a variety of methods involving transmission, ATR or RAS depending on the sample geometry and transparency. These reflectance methods are especially useful for determining the orientation at or near the surface of the finished product. The three-dimensional orientation of the chains near the surface can be quantified in this manner.

The most commonly used method for quantifying the extent of orientation in polymeric systems is the determination of the dichroic ratio. Two spectra are collected: one with radiation polarized parallel to a reference direction and one perpendicular to this direction. The reference direction is most commonly chosen along the direction of orientation. The ratio of these two spectra, often called the dichroic ratio spectrum, can then be used to characterize the orientation in the system. Dichroic ratios greater

than unity indicate orientation along the reference direction, while those less than unity indicate orientation in the orthogonal direction. A value of 1 is indicative of a lack of orientation.

It should also be noted that the presence of orientation in samples can give rise to errors in quantitative studies. The presence of orientation changes the measured absorbance values, which affects the results of quantitative analysis. One way around this is to calculate the equivalent absorbance value with no orientation, the so-called structure factor $A$:

$$A = \frac{A_x + A_y + A_z}{3}$$

where $A_x$, $A_y$, and $A_z$ represent the absorbances obtained with $x$, $y$, and $z$ polarization, respectively. The structure factor is then used in place of absorbance in quantitative studies.

The behaviour of polymers can also be followed using a time-resolved technique known as rheooptical spectroscopy or dynamic IR linear dichroism (DIRLD). In this technique, a strain is applied to a polymer sample, and the spectral changes are monitored over time. This strain is usually either a sinusoid or a step function, these being the easiest to reproduce experimentally. With the inclusion of stress and strain gauges, the spectral changes can be directly related to the mechanical properties. This technique has been used to directly study strain-induced conformational changes. A more advanced application is the study of the mechanical behaviour of polymer blends. When a miscible blend is examined, it can be seen that the spectral bands corresponding to different chains respond in phase, implying that they are molecularly mixed and act as one unit. The spectral responses have different phases for an immiscible blend, which shows that each polymer functions alone.

## Chemical property determination

### Polymerization reactions

Perhaps the most obvious chemical application is the monitoring of polymerization reactions involved in the production of polymers. Like most chemical reactions, the polymerization process shows up very well in vibrational spectra as the simultaneous disappearance and appearance of spectral bands. For example, in the polymerization of a typical vinyl monomer, carbon–carbon double bonds are broken and carbon–carbon single bonds are formed. This can be seen clearly in the spectra as the disappearance of the vinyl stretch band around 1500 cm$^{-1}$ and the simultaneous appearance of carbon–carbon single-bond stretching bands and aliphatic carbon–hydrogen stretching bands. This technique can be used to determine both the kinetic order of the polymerization reaction and the kinetic rate constants.

On-line monitoring has also received much recent interest in this area. This application allows spectra to be collected during the polymerization reaction without the necessity of stopping the reaction or performing the experiment under restrictive laboratory conditions. With the advent of low-loss fibreoptics, it is possible to monitor a polymerization reaction in a variety of harsh conditions far away from the spectrometer. This is particularly useful for large production settings commonly found in the polymer industry. An example of the quality of spectra that can be obtained using this method is shown in **Figure 3**, which shows several spectra from different times during a copolymerization reaction between styrene and 2-ethylhexyl acrylate. As can be seen, the quality of the spectra is quite good despite the drastic experimental conditions.

Several types of reactions involving polymers themselves have been studied using VS, including cross-linking, vulcanization, and degradation. The same methods are used for these processes as are used to study polymerization reactions.

### Time-dependent phenomena and spatially resolved studies

Vibrational spectroscopy is perhaps the most frequently used technique for the study of diffusion in polymeric systems because it provides a rapid way to quantitatively describe this phenomenon. The two areas of this type of research include the diffusion of small molecules into polymers and polymer–polymer interdiffusion. The easiest and most commonly used technique for this purpose is ATR. In this experiment, a polymer film is placed in contact with an ATR crystal, and the diffusing species is placed on top of the polymer. As diffusion progresses, the diffusing species moves closer to the ATR crystal and shows up in the spectrum obtained as an increase in the diffusant specific spectral band. This spectral change can be used to determine the diffusion coefficient of the system with the appropriate diffusion equation. This technique is limited to IR because of the optics and sample geometry required.

A more modern approach to this problem is to use microscopic techniques, such as an IR microscope or Raman microprobe, which directly monitor the movement of diffusing species. In this technique, the

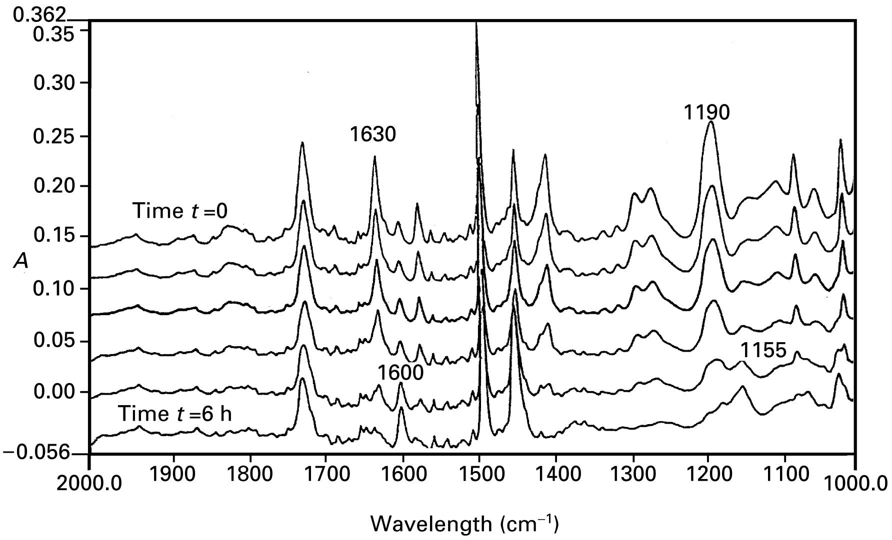



**Figure 3** *In situ* infrared spectra measured during the progress of a copolymerization reaction using a fibreoptic probe. Reproduced with permission of John Wiley & Sons Limited from Chatzi EG, Kammona O and Kipanssides (1997) Use of a midrange infrared optical-fiber probe for the on-line monitoring of 2-ethylhexyl acrylate/styrene emulsion copolymerization. *Journal of Applied Polymer Science* **63**: 799.

contact method is used, in which a thin film of polymer is placed into edge-on contact with the diffusing species. After diffusion has been allowed to progress for a certain period of time, the diffusion process is stopped by quenching, and spectra are collected along the diffusion direction to yield spatially resolved concentration information. As in the ATR method, this data are fitted to the appropriate diffusion equation to yield a diffusion coefficient. This method is more difficult to apply than ATR because of the difficulty in sample preparation, but it has the advantage of utilizing a simplified diffusion equation.

A recent trend is the incorporation of two-dimensional detectors into IR and Raman microscopes. For IR studies, focal plane array detectors are used, while charge-coupled device (CCD) detectors are used for Raman studies. When combined with the appropriate hardware, these systems are capable of collecting images of high spatial and spectral resolution in a matter of a few minutes. This allows systems that vary in time to be monitored *in situ*, and in real time depending on the rapidity of the process under study. This technique has the advantage of producing spatially resolved, chemically specific spectral images, which aids in the visualization of the process, in addition to providing high-fidelity quantitative information. An example of the quality of data that can be obtained is shown in **Figure 4**, which shows a diffusion profile obtained using an IR microscope equipped with a $64 \times 64$ element focal plane array detector. The normalized absorbance values of the diffusant peak are plotted against diffusion distance and fitted to a Fickian diffusion profile, which is also shown. This data was extracted from an image that was acquired in less than 3 minutes.

The determination of the spatial distribution of chemical species in polymeric systems is perhaps the most basic and most commonly encountered use of microspectroscopy. This technique is frequently used for the identification of defects in finished polymer products and for the identification of phase-separated regions of polymer blends. Polymer laminate films, which typically consist of layers between 2 and 10 μm thick, are also frequently studied using this technique. Raman techniques are typically more useful than IR

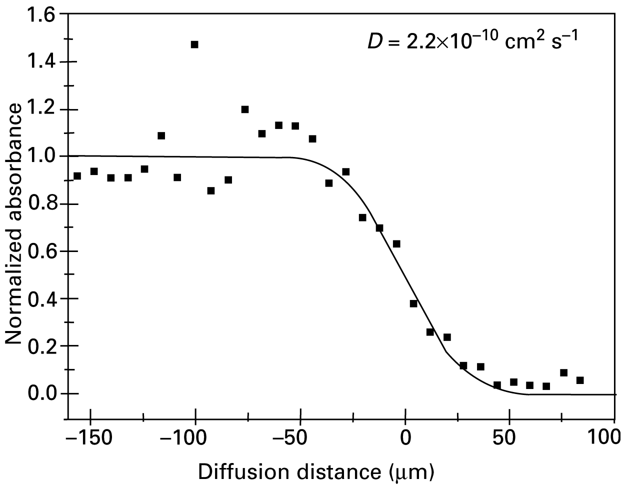



**Figure 4** Diffusion profile obtained from an infrared image along with the fit to the diffusion equation. $D$ = diffusion coefficient.

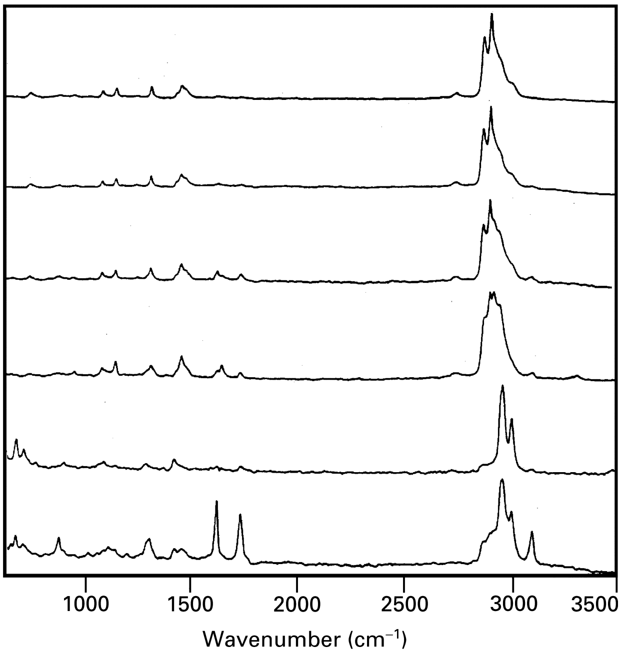



**Figure 5** Raman spectra taken at 10 μm increments from a multilayer polymer laminate film. Reproduced with permission of Elsevier Science Limited from Xue G (1997) Fourier transform Raman spectroscopy and its application for the analysis of polymeric materials. *Progress in Polymer Science* **22**: 313–406.

techniques owing to the higher spatial resolution attainable. This is shown in **Figure 5**, which shows Raman spectra obtained from 10 μm regions of a laminate film. As can be seen, the identity of each layer can be distinguished from the others by the spectral features. Another technique, known as confocal Raman microscopy, allows the acquisition of spectra from thin regions, typically around 1 μm, through the depth of the sample. This method has advantages over other methods because it can be used as a nondestructive quality control test to determine the thickness and chemical identity of the individual layers of a thin film.

An additional technique that can obtain spatially resolved spectral information is step-scan photoacoustic IR spectroscopy. Spectra can be obtained from different depths of a layered sample quickly and with higher spatial resolution (in some cases less than 1 μm) than the diffraction-limited optics of IR microscopes are capable of obtaining.

## Mechanical property determination using Raman spectroscopy

An interesting application of Raman spectroscopy is in the determination of the modulus of pure crystalline forms of a polymer. Raman is very sensitive to the longitudinal acoustic vibrational modes of simple polymer chains. Using a combination of normal coordinate analysis and experiment, the modulus of pure crystalline forms of simple straight-chain polymers, such as polyethylene and poly(oxymethylene) have been determined. Similar techniques have been used to study the pressure-dependent band shifts that occur when a polymer sample is placed under stress. This technique is complementary to mechanical analysis because it gives insight into what is occurring at the molecular level during mechanical deformation.

## List of symbols

$a_1$, $a_2$ = extinction coefficients of bands associated with conformers 1, 2; $A$ = structure factor; $A_1$, $A_2$ = absorbances of bands associated with conformers 1, 2; $A_{x,y,z}$ = absorbances obtained with $x$,$y$,$z$ polarized radiation; $k$ = kinetic rate constant; $M_n$ = number-average relative molecular mass; $R$ = gas constant; $T$ = temperature; $T_g$ = glass transition temperature; $\Delta G$ = activation energy.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; IR Spectral Group Frequencies of Organic Compounds; Nuclear Quadrupole Resonance, Instrumentation; Photoacoustic Spectroscopy, Applications; Rayleigh Scattering and Raman Spectroscopy, Theory.**

## Further reading

Bark LS and Allen NS (eds) (1982) *Analysis of Polymer Systems*. London: Applied Science.

Bower DI and Maddams WF (1989) *The Vibrational Spectroscopy of Polymers*. New York: Cambridge University Press.

Coleman, MM, Graf JF and Painter PC (1991) *Specific Interactions and the Miscibility of Polymer Blends*. Lancaster: Technomic.

Fawcett AH (ed) (1996) *Polymer Spectroscopy*. New York: Wiley.

Griffiths PR and DeHaseth JA (1986) *Fourier Transform Infrared Spectrometry*. New York: Wiley.

Koenig JL (1980) *Chemical Microstructure of Polymer Chains*. New York: Wiley.

Koenig JL (1992) *Spectroscopy of Polymers*. Washington DC: American Chemical Society.

Painter PC, Coleman MM and Koenig JL (1982) *The Theory of Vibrational Spectroscopy and Its Application to Polymeric Materials*. New York: Wiley.

Siesler HW and Holland-Moritz K (1980) *Infrared and Raman Spectroscopy of Polymers*. New York: Dekker.

Spells SJ (ed) (1994) Characterization of Solid Polymers. New York: Chapman & Hall.

Zbinden R (1964) *Infrared Spectroscopy of High Polymers*. New York: Academic Press.

# Porosity Studied By MRI

*See* **MRI of Oil/Water in Rocks.**

# Positron Emission Tomography, Methods and Instrumentation

*See* **PET, Methods and Instrumentation.**

# Positron Emission Tomography, Theory

*See* **PET, Theory.**

# Potassium NMR Spectroscopy

*See* **NMR Spectroscopy of Alkali Metal Nuclei in Solution.**

# Powder X-Ray Diffraction, Applications

**Daniel Louër**, Université de Rennes, CNRS, France

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| Applications |

## Introduction

X-ray powder diffraction is a nondestructive technique widely used for the characterization of micro-crystalline materials. The method has been traditionally applied for phase identification, quantitative analysis and the determination of structure imperfections. In recent years, applications have been extended to new areas, such as the determination of moderately complex crystal structures and the extraction of three-dimensional microstructural properties. This is the consequence of the higher resolution of modern diffractometers, the advent of high-intensity X-ray sources and the development of line-profile modelling approaches to overcome the line overlap problem arising from the one-dimensional data contained in a powder diffraction pattern. The method is normally applied to data collected at

room temperature. Nevertheless, it is also used with data collected *in situ* as a function of an external constraint (temperature, pressure, electric field, atmosphere, etc.), offering a useful tool for the interpretation of chemical reaction mechanisms and materials behaviour. Various kinds of microcrystalline materials may be characterized from X-ray powder diffraction, such as inorganic, organic and pharmaceutical compounds, minerals, catalysts, metals and ceramics. For most applications, the amount of information which can be extracted depends on the nature and magnitude of the microstructural properties of the sample (crystallinity, structure imperfections, crystallite size), the complexity of the crystal structure and the quality of the experimental data (instrument performance, counting statistics).

## Line profile parameters

The observed diffraction line profiles are distributions of intensities $I(2\theta)$ defined by several parameters. The most commonly used measure for the reflection angle is the *position* $2\theta_0$ of the maximum intensity ($I_0$). It is related to the lattice spacing $d$ of the diffracting *hkl* plane and the wavelength $\lambda$ by Bragg's law.

$$\lambda = 2d \sin\theta$$

The *dispersion* of the distribution, or diffraction line broadening, is measured by the full width at half the maximum intensity (FWHM) or by the integral breadth ($\beta$) defined as the integrated intensity ($I$) of the diffraction profile divided by the peak height ($\beta = I/I_0$). Line broadening arises from the convolution of the spectral distribution with the functions of instrumental aberrations and sample-dependent effects (crystallite size and structure imperfections). Since Fourier-series methods play an important role in X-ray diffraction by imperfect solids, the coefficients ($A_n$, $B_n$) of the Fourier series used to represent a line profile are also characteristics of the line broadening. The line *shape factor* is described by the ratio $\phi$ of the FWHM to the integral breadth ($\phi$ = FWHM/$\beta$). There are alternative shape factors according to the analytical functions $\Phi$, given in **Table 1**, commonly used for modelling individual diffraction lines, e.g. the mixing factor $\eta$ for the pseudo-Voigt and the exponent $m$ for the Pearson VII. The *area* is defined by the integrated intensity $I$ of the diffraction line. It is related to the atomic content and arrangement in the unit cell, to the amount of diffracting sample and to angle-dependent factors ($Lp$). It is proportional to the square of the structure factor amplitude $|F_{hkl}|$:

$$F_{hkl} = \Sigma_j N_j f_j \exp 2\pi i(hx_j + ky_j + lz_j) \times \exp\left(-B_j \sin^2\theta/\lambda^2\right)$$

where $N_j$ is the site occupation factor, $f_j$ is the atomic scattering factor, $B_j$ is the isotropic atomic displacement (thermal) parameter, $h$, $k$ and $l$ are the Miller indices, and $x_j$, $y_j$ and $z_j$ are the position coordinates of atom $j$ in the unit cell. **Table 2** lists the specific applications of the powder diffraction method according to the line profile parameters.

## Diffraction geometry and data collection

For most applications it is essential that the powder diffraction data be collected appropriately. Therefore, it is of prime importance to spend time optimizing the adjustment of the diffractometer, the quality of the radiation employed and the randomization of

**Table 1** Some flexible line-profile functions $\Phi(x)$ used to model powder diffraction line profiles ($L$ and $G$ denote the Lorentzian and Gaussian functions, respectively)

| *Name* | *Function* | *Shape factor* |
|---|---|---|
| Pseudo-Voigt (p-V) | $C_1[\eta L + (1-\eta)G]$ | *Mixing factor* $\eta$ |
| | ($\eta, C_1$: adjustable parameters) | $\eta = 1$ Lorentzian shape |
| | | $\eta = 0$: Gaussian shape |
| | | $\eta > 1$: super-Lorentzian shape |
| Pearson VII (PVII) | $C_2(1+Cx^2)^{-m}$ | *Exponent m* |
| | ($m$, $C$, $C_2$: adjustable parameters) | $m = 1$: Lorentzian shape |
| | | $m = \infty$: Gaussian shape |
| | | $m < 1$: super-Lorentzian shape |
| Voigt (V) | $C_3 \int L(x')G(x - x')\mathrm{d}x'$ | $\phi$ = FWHM/$\beta$ |
| | ($C_3$: adjustable parameter) | $\phi$ = 0.636 6: Lorentzian shape |
| | | $\phi$ = 0.939 4: Gaussian shape |

**Table 2** Main applications of X-ray powder diffraction

| *Diffraction line parameter* | *Applications* |
|---|---|
| Peak position | Unit-cell parameter refinement |
| | Pattern indexing |
| | Anisotropic thermal expansion |
| | Homogeneous stress |
| | Phase identification |
| Line intensity | Phase abundance |
| | Chemical reaction kinetics |
| | Crystal-structure determination and refinement (whole pattern) |
| | Search/match (*d*–*I*) |
| | Space-group determination ($2\theta_0$–absent $I_{hkl}$) |
| | Preferred orientation |
| Line width and shape | Instrumental resolution function |
| | Microstructure |
| Line-profile broadening | Microstructure (crystallite size, size distribution, lattice distortion, structure mistakes, dislocations, composition gradient), crystallite growth kinetics |

the crystallites in the sample. There are several designs for X-ray powder diffractometers (reflection or transmission modes), each of them having advantages and disadvantages. **Figure 1** shows an optics commonly used with conventional divergent-beam X-ray sources, based on the Bragg–Brentano parafocusing geometry. The beam converges on the receiving slit after diffraction by the sample. The geometry is characterized by two circles, the goniometer circle with a constant radius *R* and the focusing circle with a radius dependent on $\theta$. It uses a flat sample which lies tangent to the focusing circle. The main advantage of this reflection geometry is that no absorption correction has to be made if an 'infinitely' thick sample is used. Disadvantages are that by using a flat sample, preferred orientation effects are increased and at low angles the illuminated area can become larger than the sample. Preferred orientation effects can be reduced by using side-loaded sample holders. The most frequent angular errors arise from a shift of the zero-$2\theta$ position and a displacement $\delta$ of the specimen from the goniometer axis of rotation ($\Delta 2\theta = 2\delta \cos\theta / R$). Transmission geometry, often combined with a position-sensitive detector (PSD), can be employed with thin flat samples or capillaries. Main advantages are the small amount of sample used and, with capillaries, the reduction of preferred orientation. Transmission optics are ideal for 'transparent' materials, containing only light atoms, but for highly absorbing materials, patterns are difficult to measure. X-ray synchrotron radiation presents some important advantages over conventional X-ray sources. The excellent angular collimation combined with the wavelength tunability and high brightness of the synchrotron source make it ideal for many types of X-ray diffraction experiments. Moreover, with parallel-beam optics, sample-displacement aberrations in diffraction lines are completely eliminated.

Although a basic requirement of the Bragg law is the use of monochromatic radiation, the doublet $K\alpha_1$–$K\alpha_2$ from copper is the most popular wavelength

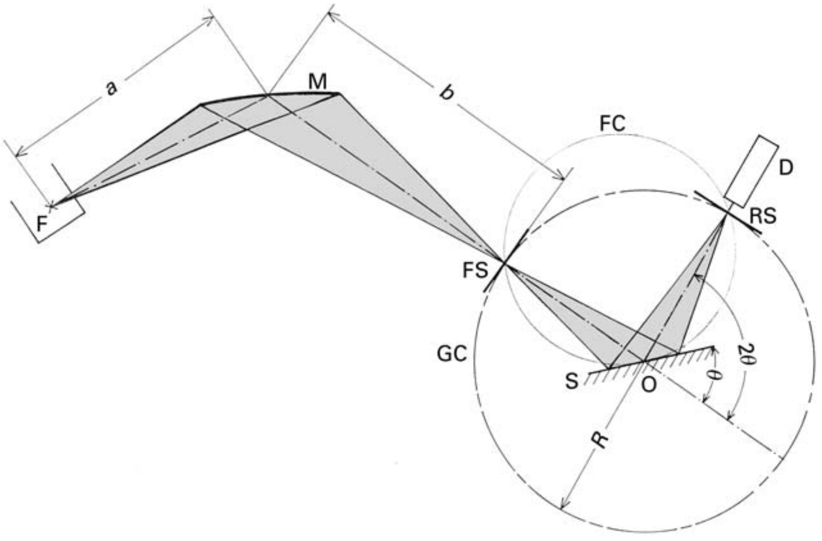



**Figure 1** Optics of a conventional focusing powder diffractometer with monochromatic X-rays (Bragg–Brentano geometry with reflection specimen): F line focus of X-ray tube, M incident-beam monochromator, *a* short focal distance, *b* long focal distance, FS focal slit, S flat specimen, GC goniometer circle, O goniometer axis, R goniometer radius, FC focusing circle, $\theta$ Bragg angle, $2\theta$ reflection angle, RS receiving slit, D detector.

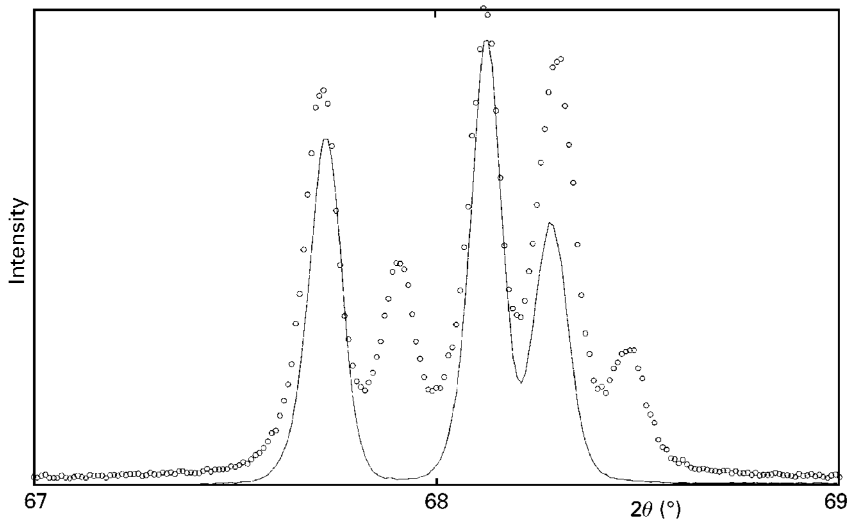



**Figure 2** The quartz cluster of reflections 212, 203 and 301 collected with Cu$K\alpha_1$ radiation (continuous line) and with the Cu$K\alpha_1$–$K\alpha_2$ doublet (circles).

used with laboratory diffractometers. However, for the more demanding applications it is desirable to remove the $K\alpha_2$ component with a monochromator located in the incident beam (**Figure 1**). There is some reduction in the intensity of $K\alpha_1$, but the number of reflections in the pattern is halved and thereby the degree of line overlap is reduced. **Figure 2** shows the improvement in resolution achieved with the monochromatic Cu$K\alpha_1$ radiation ($\lambda$ = 1.5406 Å) with respect to the $K\alpha_1$–$K\alpha_2$ doublet.

The performances of powder diffractometers are determined by the precision on peak position measurements and the instrument resolution function (IRF), normally expressed by the angular dependence of the FWHM obtained with a reference material. The most widely used standard reference materials (SRMs) for diffractometer characterization are those from the National Institute of Standards and Technology (NIST). For instance, powders of Si (SRM 640b, $a$ = 5.430 940 ± 0.000 035 Å) and fluorophlogopite mica (SRM 675, interlayer spacing $d_{001}$ = 9.981 04 ± 0.000 07 Å) are proposed as $d$-spacing standards, while $LaB_6$ (SRM 660) is recommended as a line-profile standard. With laboratory X-ray diffractometers errors on peak positions can be less than 0.01°(2$\theta$) and the IRF has typically a minimum around 0.06°(2$\theta$) at intermediate angles, increasing to twice this value at ~130°(2$\theta$) as a consequence of spectral dispersion. The best instrumental resolution (~0.01–0.02°2$\theta$) is obtained with the synchrotron parallel beam optics with a crystal analyser mounted in the diffracted beam. This is of particular interest in some applications, such as the study of complex crystal structures.

## Pattern modelling

Pattern modelling techniques are used in most current applications. The intensity at point $x_i$ in the calculated pattern is given by

$$y_{\mathrm{cal}}(x_i) = \Sigma_k\ I_k \Phi(x_i - x_k) + b(x_i)$$

where $I_k$ is the integrated intensity of reflection $k$, $\Phi$ is one of the normalized profile functions given in **Table 1** and $b(x_i)$ is the background contribution. The summation is over all reflections contributing to the intensity at point $x_i$. Parameters defining the model are refined until the quantity

$$S = \Sigma\ w(x_i)[y_{\mathrm{obs}}(x_i) - y_{\mathrm{cal}}(x_i)]^2$$

is a minimum, the summation being over all data points in the diffraction pattern with $w(x_i)$ being the appropriate weighting factor. Although a visual inspection of the difference curve between observed and calculated patterns is the best way to judge the quality of the fit over the whole angular range, numerical factors are used to assess the quality of the final refinement. These are listed in **Table 3**.

### Pattern decomposition

This is a systematic procedure for decomposing a powder pattern into its component Bragg reflections without reference to a structure model and, thereby,

**Table 3** Some numerical criteria of fit used in pattern-fitting methods.

| | |
|---|---|
| *R*-profile | $R_p = \frac{\sum \lvert y_i(\text{obs}) - y_i(\text{cal}) \rvert}{\sum y_i(\text{obs})}$ |
| *R*-weighted profile | $R_{wp} = \left[ \frac{\sum \lvert w_i [y_i(\text{obs}) - y_i(\text{cal})] \rvert^2}{\sum w_i [y_i(\text{obs})]^2} \right]^{1/2}$ |
| *R*-structure factor | $R_F = \frac{\sum \lvert I_{hkl}(\text{'obs'})^{1/2} - I_{hkl}(\text{cal})^{1/2} \rvert}{\sum I_{hkl}(\text{'obs'})^{1/2}}$ |
| *R*-Bragg factor | $R_B = \frac{\sum \lvert I_{hkl}(\text{'obs'}) - I_{hkl}(\text{cal}) \rvert}{\sum I_{hkl}(\text{'obs'})}$ |

line-profile parameters ($2\theta_0$, $I_0$, $I$, FWHM, $\beta$, shape factors) are extracted. **Figure 3** shows the fitting of pseudo-Voigt functions to the individual lines of the pattern of a ZnO sample displaying diffraction line broadening. However, for extracting integrated intensities over the complete pattern, for a subsequent structural analysis, an approach incorporating constrained peak positions, according to the refined unit-cell dimensions and to the space group, allows the generation of only space-group allowed intensities.

## The Rietveld method

In the Rietveld method an observed and a calculated powder diffraction pattern are compared and the difference is used to refine the atomic coordinates of the structure model. The calculated intensity at point $x_i$ is given by the equation

$$y_{\text{cal}}(x_i) = s\Sigma_k\, m_k (Lp)_k |F_k|^2 P_k \Phi(x_i - x_k) + b(x_i)$$

where $s$ is a scale factor, $m_k$ is the reflection multiplicity and $P_k$ is a function to deal with the preferred orientation of the crystallites. They are two groups of refined parameters arising from the structure model (atomic coordinates $x_j$, $y_j$ and $z_j$, atomic displacement parameters, unit-cell dimensions) and the instrumental model (angular dependence of the profile parameters, FWHM and shape factors, 2θ-zero position, preferred orientation, etc.). Recommendations for Rietveld-refinement strategies have been formulated by the Commission on Powder Diffraction of the International Union of Crystallography. **Figure 4** shows a typical final Rietveld plot obtained from powder data collected with the capillary method. The precision of a structure refinement depends on many factors, e.g. the number of parameters to be refined, data quality (preferred orientation, counting statistics, anisotropic line broadening), the contrast between atoms and the size of the unit cell. Moreover, structure refinement from X-ray diffraction data is strongly influenced by the fall-off of the scattering atomic factors $f_j$ with $d^{-1}$ (or $2\theta$). This is in contrast with neutron-diffraction data, for which the scattering length does not fall off significantly over the range of observations. Another important difference between neutron and X-ray

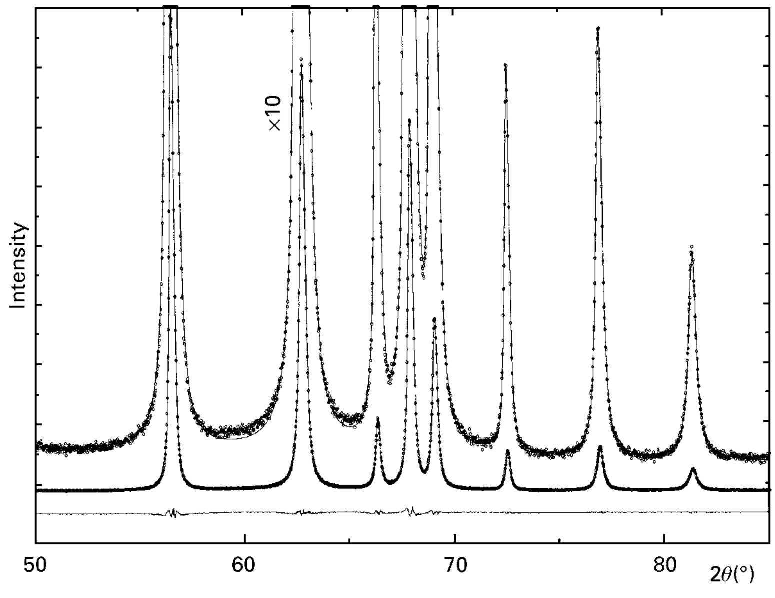



**Figure 3** Fitting of a part of the pattern of a sample of nanocrystalline ZnO, using pseudo-Voigt functions for modelling individual diffraction lines, Cu*K*$\alpha_1$ radiation, $R_{wp}$ = 1.2%. The observed intensity data are plotted as circles and the calculated pattern is shown as a continuous line. The lower trace is the difference curve. The ×10 scale expansion shows the fit in the line-profile tails.

diffraction is that the relative scattering powers of atoms for neutrons and X-rays are significantly different. With neutrons there are pronounced differences in the scattering lengths of neighbouring elements. In the case of X-rays, there is a monotonic variation in X-ray scattering factors and hence light atoms are weak X-ray scatterers. For instance, in the case of $U(UO_2)(PO_4)_2$ the X-ray powder data are dominated by the U atom, the ratio of the scattering factors of U to O being greater than 11.5, while with neutrons oxygen is a relatively strong scatterer and then the ratio of the appropriate scattering lengths is 1.45. This property explains the major role played by neutron powder diffraction in determining oxygen content and position in high $T_c$ cuprates in the presence of heavy atoms such as Ba, Hg, Tl and Bi.

## Qualitative and quantitative analyses

### Phase identification

Qualitative phase identification is traditionally based on a comparison of observed data with interplanar spacings $d$ and relative intensities $I_0$ compiled for crystalline materials. The Powder Diffraction File (PDF), edited by the International Centre for

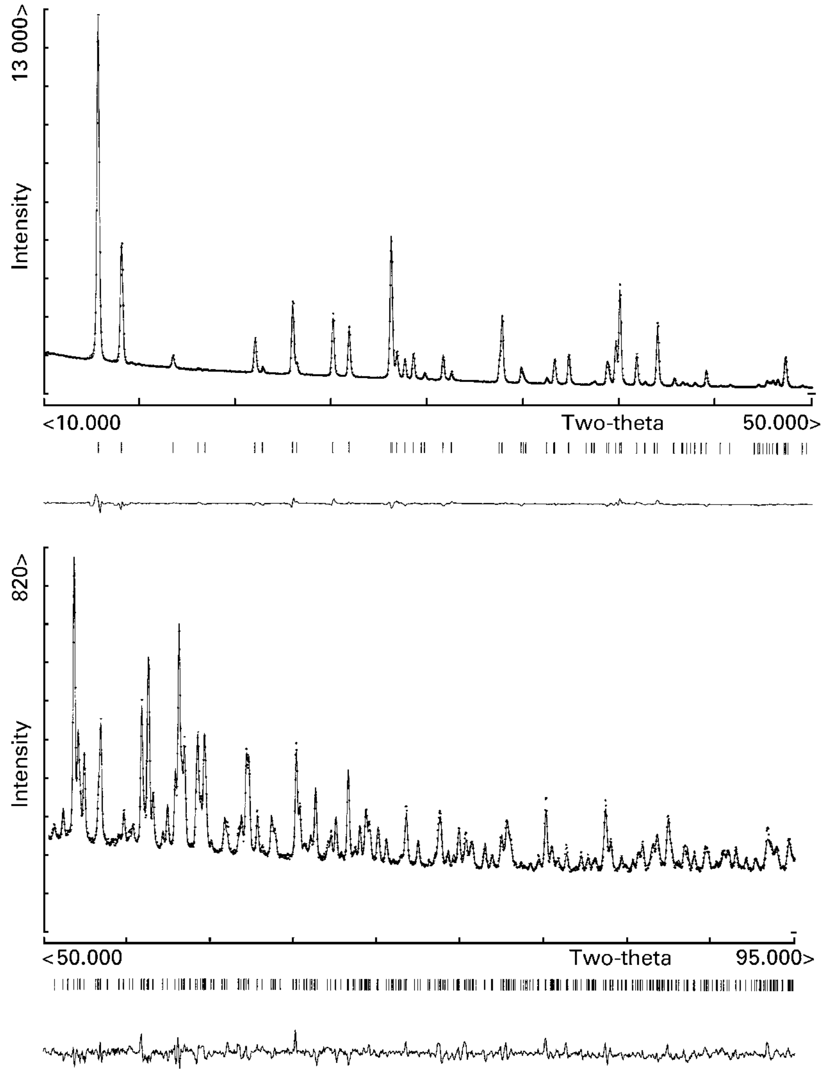



**Figure 4** Example of a typical Rietveld plot. The powder data of $LiB_2O_3(OH).H_2O$ were collected with the Debye–Scherrer (capillary) geometry using monochromatic Cu$K\alpha_1$ radiation and a curved position-sensitive detector. The observed data are plotted as points and the calculated pattern as a continuous line. The lower trace is a plot of the difference of observed minus calculated. The vertical markers indicate the positions of calculated Bragg reflections. The intensity scale is magnified for the high-angle part, where the intensity of observed diffraction lines is low. Reprinted with permission of the International Union of Crystallography from Louër D, Louër M and Touboul M (1992) Crystal structure determination of lithium diborate hydrate, $LiB_2O_3(OH){\cdot}H_2O$, from X-ray powder diffraction data collected with a curved position-sensitive detector. *Journal of Applied Crystallography* **25**: 617–623.

Diffraction Data (ICDD), contains powder data for more than 106 000 (sets 1–48) substances, including approximately 38 000 calculated patterns from the Inorganic Crystal Structure Database (ICSD). The Boolean search program supplied by the ICDD with the database offers great flexibility in phase identification and characterization. In a more recent search/match procedure, complete digitized observed diffraction patterns are used, instead of simply a list of extracted *d*s and *I*s. To decide whether or not the pattern contains the data of a particular PDF entry, its data are compared with parts of zero intensity in the pattern for the unknown substance. The method can be used for multiphase patterns. As each phase is identified, the diffraction lines can be removed and the procedure is repeated by using the remaining regions of the pattern. The method is very efficient, even for the identification of minor phases if some known chemical constraints are introduced into the search.

### Quantitative phase abundance

Quantitative phase analysis is the determination of the amounts of different phases present in a sample. The powder diffraction method is widely used to determine the abundance of distinct crystalline phases, e.g. in rocks and in mixtures of polymorphs, such as zirconia ceramics. The principle of the method is straightforward; the integrated intensity ($I$) of the diffraction lines from any phase in a mixture is proportional to the mass of the phase present in the sample. One analytical approach is based on a reference-intensity ratio (RIR) defined as the integrated intensity of the strongest reflection for the phase of interest to the strongest line of a standard (usually the 113 reflection of corundum) for a 1:1 mixture by weight. The measurement and use of RIRs is straightforward for materials for which the intensity does not vary with composition and preferred orientation. Because of the potential health hazard of respirable crystalline silica, X-ray powder diffraction is also used for detecting, identifying and quantifying the crystalline and amorphous silica of all types of samples from airborne dusts to bulk commercial products. SRMs 1878 (α-quartz) and 1979 (cristobalite) from NIST are certified with respect to amorphous content for analysis of silica-containing materials in accordance with health and safety regulations.

An extension of the Rietveld method is its application to multiphase samples for the determination of phase abundance. There is a simple relationship between the individual scale factors determined in a Rietveld analysis and the weight fractions ($W_i$) of the phase concentration in a multicomponent mixture:

$$W_i = s_i(ZMV)_i/\Sigma_j s_j(ZMV)_j$$

where $s_i$, $Z_i$, $M_i$ and $V_i$ are the scale factor, the number of molecules per unit cell, the mass of the formula unit and the unit-cell volume of phase $i$ and the summation is over all phases present. A requirement of the method is that the crystal structure is known for each phase in the mixture. The use of all reflections in the selected angular range is a great advantage, since the uncertainty on phase abundance is reduced by minimizing preferred orientation effects. In general, phase-analysis results obtained with X-ray data are inferior to those obtained from neutron data. This is related to residual errors arising from an imperfect preferred orientation modelling and from microabsorption effects.

## *Ab initio* structure determination

Among the most recent advances of the powder method is the determination of crystal structures from powder diffraction data. It is an application for which the resolution of the pattern is of prime importance. A series of successive stages are involved in the analysis, including the determination of cell dimensions and identification of the space group from systematic reflection absences, the extraction of structure factor moduli $|F_{hkl}|$, the solution to the phase problem to elaborate a structure model and, finally, the refinement of the atomic coordinates with the Rietveld method.

### Pattern indexing

The purpose of pattern indexing is to reconstruct the three-dimensional reciprocal lattice of a crystalline solid from the radial distribution of lengths $d^*(=1/d)$ of the diffraction vectors. The basic equation used for indexing a powder diffraction pattern is obtained by squaring the reciprocal-lattice vectors $\mathbf{d}_{\mathrm{hkl}}^*$ $(=h\mathbf{a}^*+k\mathbf{b}^*+l\mathbf{c}^*)$, expressed in terms of the basis vectors of the reciprocal lattice ($\mathbf{a}^*$, $\mathbf{b}^*$, $\mathbf{c}^*$) and $hkl$ Miller indices,

$$Q(hkl) = h^2Q_A + k^2Q_B + l^2Q_C + 2klQ_D + 2hlQ_E + 2hkQ_F$$

where $Q(hkl) = 1/d^2$, $Q_A = a^{*2}$, $Q_B = b^{*2}$, $Q_C = c^{*2}$, $Q_D = b^*c^*\cos\alpha^*$, $Q_E = c^*a^*\cos\beta^*$, $Q_F = a^*b^*\cos\gamma^*$; $a^*$, $b^*$, $c^*$ are the linear parameters and $\alpha^*$, $\beta^*$, $\gamma^*$ the angles of the reciprocal unit cell. This quadratic

form corresponds to the triclinic crystal symmetry. Only four parameters are required for a monoclinic cell, three for an orthorhombic cell, two for tetragonal and hexagonal cells, and one for a cubic cell. Indexing a powder pattern consists in finding the linear and angular dimensions of the unit cell, from which a set of Miller indices *hkl* can be assigned to each observed line $Q_{obs}$, within the experimental error on the observed peak positions. Automatic procedures are available to index a powder diffraction pattern. They are based on three main approaches, regardless of symmetry: the zone-indexing method, the index-permutation method and the successive-dichotomy method. The use of pattern-indexing methods requires a high precision on peak positions ($|\Delta 2\theta| < 0.03°2\theta$). The assessment of the reliability of an indexed pattern is carried out with the de Wolff figure of merit $M_{20}$:

$$M_{20} = Q_{20}/2 \langle \Delta Q \rangle N_{cal}$$

where $Q_{20}$ corresponds to the 20th observed line, $\langle \Delta Q \rangle$ is the average absolute discrepancy between $Q_{k\text{obs}}$ and the nearest $Q_{k\text{cal}}$ value and $N_{cal}$ is the number of distinct calculated $Q$ values smaller than $Q_{20}$, not including any systematic absences if they are known. $M_{20}$ is greater when $\langle \Delta Q \rangle$ is small and $N_{cal}$ is as close as possible to 20. A related figure of merit, $F_N$, introduced for the evaluation of powder data quality, is also used. It is reported as 'value ($\langle \Delta 2\theta \rangle$, $N_{cal}$)', where $\langle \Delta 2\theta \rangle$ is the average angular discrepancy. A solution with $M_{20}$ greater than 20 is generally correct, although some geometrical ambiguities or the presence of a dominant zone can mask the true solution. An additional check of the reliability of the solution consists in indexing all measurable peak positions in the pattern, from which systematic absent reflections can be detected and, then, possible space groups are proposed. With synchrotron parallel-beam optics, higher figures of merit are obtained, since angular precision and resolution (more lines are observed) are considerably improved. There is no particular problem for indexing, from conventional X-ray data, patterns of materials with moderate cell volumes and, for instance, patterns of monoclinic compounds with volumes up to 3000 Å$^3$ can be normally handled.

## Structure solution

Following the determination of the unit cell and space group assignment, integrated intensities ($I_{hkl}$) (or structure factor amplitudes) are extracted with a pattern-decomposition technique. For clusters of overlapping lines an equipartition of the overall intensity is generally applied. Therefore, intensity data sets contain a limited number of unambiguously measured reflections. An estimation of the proportion of statistically independent reflections can be calculated from an algorithm based on line width and reflection proximity. It is a useful indicator which depends on the selected angular range. The methods for solving the phase problem used with single-crystal data (Patterson methods and direct methods) are generally applicable with powder data. Computer programs have been adapted to the powder diffraction case. In general, only fragments of the structure are found from these methods; the remaining atoms are then obtained from subsequent Fourier calculations. However, with good data and a favourable proportion of unambiguous reflections, the success of the direct methods can be very high and even complete models (non-H atoms) may be revealed from one calculation. Considerable effort has been devoted to the development of new approaches for the treatment of powder-diffraction data. They include the Monte Carlo and the simulated annealing approaches, the maximum-entropy method, the atom–atom potential method and genetic algorithms. Computer-modelling approaches operate in direct space. Trial crystal structures are generated independently of the observed powder diffraction data. The suitability of each structure model is assessed by comparison between the calculated and observed diffraction patterns and is quantified using an appropriate profile *R*-factor (see **Table 3**). The direct-space methods present the advantage of avoiding the critical stage of extracting the individual intensities from pattern decomposition. They are suitable for organic molecules, but a detailed knowledge of the expected molecule (bond lengths and torsional angles) is a basic requirement. An example is the structure determination of a metastable form of piracetam, $C_6H_{10}N_2O_2$, whose lifetime at room temperature is only 2 h, solved from powder data collected with a PSD (**Figure 5**). Structure determination from powder data is used in varied fields of materials science, including inorganic, organometallic and organic chemistry, mineralogy and pharmaceutical science. Although the analysis is generally applicable to moderately complex structures, it has been successful for solving structures containing, for instance, 29 atoms in the asymmetric unit, e.g. $Ga_2(HPO_3)_3.4H_2O$, $Ba_3AlF_3$, $Bi(H_2O)_4(OSO_2CF_3)$, or having a high unit-cell volume, e.g. 7471 Å$^3$ for $[(CH_3)_4N]_4Ge_4S_{10}$. With the ultra-high resolution available with X-ray synchrotron radiation, determination of more complex structure can now be undertaken, particularly with materials containing light atoms.

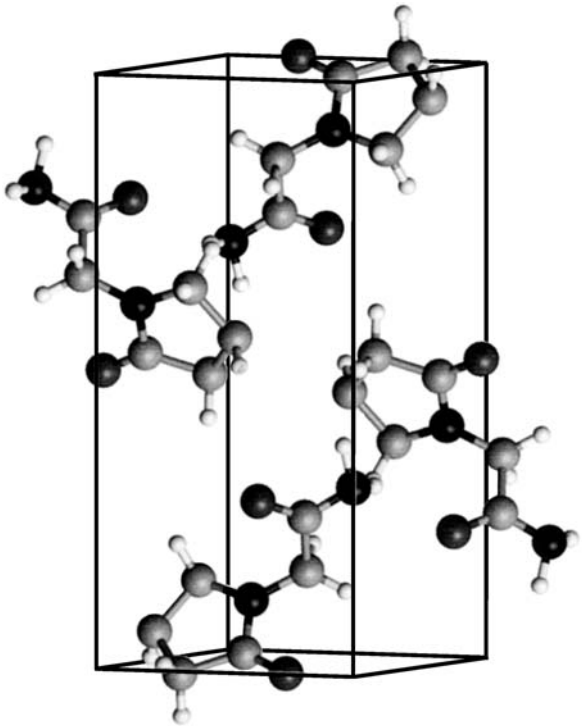


**Figure 5** View of the crystal structure of a metastable phase of piracetam, $C_6H_{10}N_2O_2$, solved from the atom–atom potential method and refined with the Rietveld method. The crystal symmetry is monoclinic, space group $P2_1/n$ (the unit cell is shown with the *a* axis horizontal and *b* axis vertical) (crystal data are reported in *Acta Crystallographica* **B51**: 182, 1995). Powder diffraction data were collected from a conventional X-ray source.

For specific applications, the tunability of synchrotron radiation sources allows the X-ray wavelength to be changed readily, and this can be exploited to enhance the contrast between close elements in the periodic table. These studies are termed anomalous (resonant) scattering experiments. The atomic scattering factor for X-rays is defined as

$$f = f_0(\theta) + f'(E) + \mathrm{i}f''(E)$$

where $f_0$ varies only with $\sin\theta/\lambda$ and $f'$ and $f''$ are the energy-dependent real and imaginary parts of the anomalous contribution. By selecting a wavelength close to the absorption edge of an element the scattering factor may change by a few electrons. By comparing powder data collected near-edge and off-edge, the technique can be used to determine the distribution of cations with similar atomic numbers over crystallographically distinct sites. Applications can be extended to mixed-oxidation states of an absorbing element if these are ordered within a crystal structure, e.g. $Eu^{2+}$ and $Eu^{3+}$ in $Eu_3O_4$ or $Ga^+$ and $Ga^{3+}$ in $GaCl_2$.

## Diffraction line broadening analysis

Microstructural imperfections (lattice distortions, stacking faults) and the small size of crystallites (i.e. domains over which diffraction is coherent) are usually extracted from the integral breadth or a Fourier analysis of individual diffraction line profiles. Lattice distortion (microstrain) represents departure of atom position from an ideal structure. Crystallite sizes covered in line-broadening analysis are in the approximate range 20–1000 Å. Stacking faults may occur in close-packed or layer structures, e.g. hexagonal Co and ZnO. The effect on line breadths is similar to that due to crystallite size, but there is usually a marked *hkl*-dependence. Fourier coefficients for a reflection of order $l$, $C(n,l)$, corrected from the instrumental contribution, are expressed as the product of real, order-independent, size coefficients $A^S(n)$ and complex, order-dependent, distortion coefficients $C^D(n,l)$ [$=A^D(n,l)+\mathrm{i}B^D(n,l)$]. Considering only the cosine coefficients $A(n,l)$ [$=A^S(n).A^D(n,l)$] and a series expansion of $A^D(n,l)$, $A^S(n)$ and the microstrain $\langle e^2(n)\rangle$ can be readily separated, if at least two orders of a reflection are available, e.g. from the equation

$$A(n,l) = A^S(n) - A^S(n)2\pi^2 l^2 n^2 \langle e^2(n) \rangle$$

The inverse of the initial slope of the size coefficients $A^S(n)$ versus the Fourier harmonic number $n$ (or $L = n/\Delta s$, where $\Delta s$ is the range of the intrinsic profile in reciprocal units) is a measure of the Fourier apparent size, $\varepsilon_F$, defined as an area-weighted crystallite size. The second derivative of $A^S(n)$ is a measure of the crystallite size distribution in a direction perpendicular to the diffracting plane *hkl*, obtainable when strains are negligible. Fourier coefficient $C(n)$ are also related to the integral breadth, expressed in reciprocal units, $\beta^*$ [$=\Delta s/\Sigma_n|C(n)|$]. For strain-free materials the reciprocal of this quantity is the integral-breadth apparent size, $\varepsilon_\beta$, defined as a volume-weighted thickness of crystallite in the direction of the diffraction vector. For a spherical crystallite the direction of the diffraction vector is unimportant and there is a simple relation between the diameter $D$ and $\varepsilon_\beta$ ($=3D/4$) or $\varepsilon_F$ ($=2D/3$). The ideal ratio $\varepsilon_\beta/\varepsilon_F$ is 1.125. Any departure from this ratio is due to the presence of a distribution of crystallite sizes. Similar expressions have been derived for other crystallite shapes, e.g. the cylinder, for which the two limiting cases are acicular and disk-like forms. Anisotropic models require precise determination of apparent

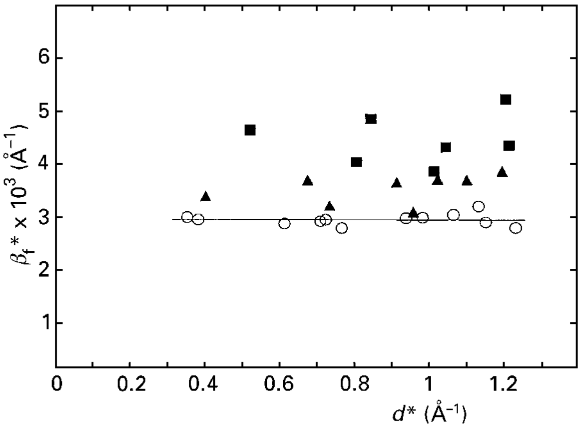



**Figure 6** Example of a Williamson–Hall plot for a sample of nanocrystalline ZnO powder exhibiting size and stacking-fault diffraction line broadening according to the *hkl* values. (O) reflections unaffected by mistakes (*hk*0 and *hkl* with *l* even, $h-k=3n$), (▲) first reflection set affected by stacking faults (*hkl* with *l* odd, $h-k=3n\pm1$), (■) second reflection set affected by stacking faults (*hkl* with *l* even, $h-k=3n\pm1$).

sizes in various crystallographic directions. Representative examples of average spherical and cylindrical crystallite shapes are found in $CeO_2$ and ZnO loose powders.

These methods can be applied to more diffraction lines when profile-fitting techniques are employed, provided the observed data are precisely modelled. A more-detailed (three-dimensional) characterization of the microstructure can then be obtained. In the Fourier approach, line profiles must be reconstructed analytically, from the line-profile parameters extracted from pattern-decomposition techniques, prior to the physical interpretation. With the integral-breadth method, the use of Voigt functions in the modelling of observed $h$ and instrumental $g$ profiles represents a definitive advantage, since the integral breadth $\beta_f$ of the intrinsic $f$ profile can be precisely determined. Whatever the diffraction-line-broadening analysis employed, an essential preliminary step in the analysis is to examine a (Williamson–Hall) plot giving the variation of $\beta_f$ (expressed in reciprocal units) as a function of $d^*$. Although this plot is not used for quantitative characterization, it gives an overview of the nature of the broadening due to sample imperfections and orients the subsequent analysis. **Figure 6** shows the *hkl*-dependence of the integral breadths for a sample of nanocrystalline ZnO. These plots can be varied according to the effects at the origin of line broadening and to their anisotropic nature. Since perfect modelling of line profiles are required in the Rietveld method, the technique can also, in principle, be used to extract microstructural properties. However, unless line broadening is isotropic, to obtain meaningful information the procedure must be able to handle the various possible sources of broadening present in the pattern, such as crystallite size, strain, stacking faults, dislocation density, and the residual errors related to the structure and preferred orientation models must be minimized.

## Dynamic and non-ambient diffraction

Time and temperature dependent X-ray diffraction includes the measurement of a series of diffraction patterns as a function of time, temperature or other physical constraint. The time required for collecting data decreases considerably with the availability of fast detectors, such as PSDs, and the brightness of the X-ray source. In principle, line-profile parameters can be extracted for each pattern and interpreted in structural (peak position and integrated intensity) and microstructural (breadths and line shapes) terms. Consequently, the structural and microstructural changes taking place as a function of the external constraint (temperature, atmosphere, pressure, electric field, etc.) are displayed from the successive patterns. The method affords the possibility of establishing the pathways during chemical solid-state reactions, such as phase transition and thermal decomposition, and to determine the kinetics of processes, e.g. the crystallization of nanocrystalline solids or phase transformations. Patterns can be collected on a time scale of a few minutes with conventional X-rays but the high brightness of synchrotron radiation makes it possible to measure diffraction data in very short time periods. A representative application is the investigation of fast and self-propagating solid combustion reactions on a subsecond time-scale. *In situ* powder diffraction can also be combined with other complementary techniques applied simultaneously, such as EXAFS (extended X-ray absorption fine structure). The two structural probes may provide long-range order information (powder diffraction) and short-range order details (EXAFS). For instance, *in situ* combined X-ray diffraction and EXAFS have been used for the study of the formation of heterogeneous catalysts.

## List of symbols

$A_n$, $B_n$ = Fourier coefficients; $A^D$ = distortion coefficient; $A^S$ = size coefficient; $B$ = isotropic atomic displacement; $d$ = interlayer spacing; $D$ = crystallite diameter; $e$ = microstrain; $f$ = atomic scattering factor; $F$ = structure factor; $h,k,l$ = Miller indices; $I$ = intensity; $Lp$ = angle-dependent factor; $m$ = Pearson exponent; $M_{20}$ = de Wolff figure of merit;

$m_k$ = reflection multiplicity; $N$ = site occupation factor; $R$ = goniometer circle radius; $R_p$ = profile factor; $s$ = scale factor; $W$ = weight fraction; $Z$ = number of molecules per unit cell; $\beta$ = integral breadth; $\delta$ = specimen displacement; $\varepsilon$ = apparent crystallite size; $\eta$ = pseudo-Voigt mixing factor; $\theta$ = diffraction angle; $\lambda$ = wavelength; $\phi$ = shape factor; $\Phi$ = line-profile function.

*See also:* **Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; Materials Science Applications of X-Ray Diffraction; Neutron Diffraction, Theory; Small Molecule Applications of X-Ray Diffraction; X-Ray Absorption Spectrometers.**

## Further reading

Bish DL and Pose JE (eds) (1989) Modern powder diffraction. *Reviews in Mineralogy* **20**: 1–369.

Harris KDM and Tremayne M (1996) Crystal structure determination from powder diffraction data. *Chemistry of Materials* **8**: 2554–2570.

Jenkins R and Snyder RL (1996) *Introduction to X-ray Powder Diffractometry*. New York: Wiley.

Langford JI and Louër D (1996) Powder diffraction. *Reports on Progress in Physics* **59**: 131–234.

McCusker LB, Von Dreele RB, Cox DE, Louër D and Scardi P (1999) Rietveld refinement guidelines. *Journal of Applied Crystallography* **32**: 36–50.

Parrish W (1995) Powder and related techniques: X-ray techniques. In Wilson AJC (ed) *International Tables for Crystallography*, Vol. C, 42–79. Dordrecht: Kluwer Academic Publishers.

Smith DK (1997) Evaluation of the detectability and quantification of respirable crystalline silica by X-ray powder diffraction methods. *Powder Diffraction* **12**: 200–227.

Young RA (ed) (1995) *The Rietveld Method*. Oxford: IUCr-Oxford University Press.

# Product Operator Formalism in NMR

**Timothy J Norwood**, Leicester University, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
|---|
| Theory |

The introduction of the product operator formalism by Ernst and co-workers can be regarded as one of the milestones in the development of NMR. Unusually, this is not because it constitutes an advance in either theory or technique, for it is neither, but because it makes it possible for many of the users of NMR to understand the experiments they carry out. It is necessary to use quantum mechanics to explain the behaviour of nuclear magnetism and a density operator to describe the state of the system. While this approach is rigorous, it can often appear abstract and lacking in physical intuition to those who do not have a background in quantum mechanics. Furthermore, calculations often involve cumbersome matrix operations. While there is an alternative in the physically intuitive net-magnetization vector model, this classical approach is only really useful for explaining very simple experiments such as the spin echo, and breaks down when it comes to explaining phenomena such as coherence transfer and multiple-quantum coherence that are ubiquitous in modern NMR spectroscopy. The product operator formalism provides a half-way house between the two approaches; it retains both the rigour of the density matrix and the physical intuition and ease of use of the vector model. While with the density matrix it is usual to follow the evolution of the system as a whole during the course of an experiment, with the product operator formalism the fate of individual components can easily be traced. The product operator formalism can equally well be used either quantitatively to calculate the amplitude of the magnetization following particular coherence transfer pathways in the course of an experiment or qualitatively to sketch out those pathways. The product operator formalism has seen extensive use both as an education tool for explaining how pulse sequences work and as a research tool used as an aid for designing new ones. Its main limitation is that it is really only suitable for describing weakly coupled spin systems. When multiple-quantum coherence is under consideration, it can be usefully supplemented with raising and lowering operators.

# Theory

## Origin and definitions

The density operator $\sigma(t)$ can be expanded into a set of orthogonal base operators $\{B_s\}$:

$$\sigma(t) = \sum_{s=1}^{n^2} b_s(t) B_s \qquad [1]$$

Product operators are one of many possible sets of base operators that can be chosen; the choice usually depends on the application. Product operators are based upon the angular momentum operators $I_x$, $I_y$ and $I_z$ of individual spins, and for spin-$\frac{1}{2}$ nuclei are defined by

$$B_s = 2^{(q-1)} \prod_{k=1}^{N} (I_{k\alpha})^{a_{ks}} \qquad [2]$$

where $N$ is the number of spin-$\frac{1}{2}$ nuclei in the spin system, $k$ is the spin index, $\alpha = x$, $y$ or $z$, and $q$ is the number of operators in the product. The coefficient $a_{ks}$ has a value of 1 for spins in the product and 0 for all other spins.

For a system consisting of a single uncoupled spin $k$ there are four possible product operators $(\frac{1}{2})E$, $I_{kz}$, $I_{kx}$ and $I_{ky}$, where $E$ is the unity operator, which need not be considered further. The three remaining operators and their corresponding representations in the net-magnetization vector model are given in **Figure 1**. The operator $I_{kz}$ corresponds to the $k$-spin $z$ magnetization found at thermal equilibrium as a result of the Boltzmann distribution between the two spin states α and β. The operators $I_{kx}$ and $I_{ky}$ correspond to $x$ and $y$ components of $k$-spin magnetization in the $xy$ plane. These two operators can also be described as the $x$ and $y$ components of the in-phase magnetization or as the in-phase single-quantum coherence of spin $k$. The importance of making these qualifications will become apparent below. Neither $I_{kx}$ nor $I_{ky}$ is present at thermal equilibrium but they may be present after the application of a radiofrequency pulse to the equilibrium magnetization. The operators $I_{kx}$ and $I_{ky}$ are also important because they are the only operators that are directly observed in an NMR experiment; it is from these operators that the signal observed in the free induction decay arises.

Most systems of interest consists of more than one spin. The product operators for a two-spin system are listed in **Table 1**. Where applicable, representations of characteristic examples of these operators are given in the net-magnetization vector

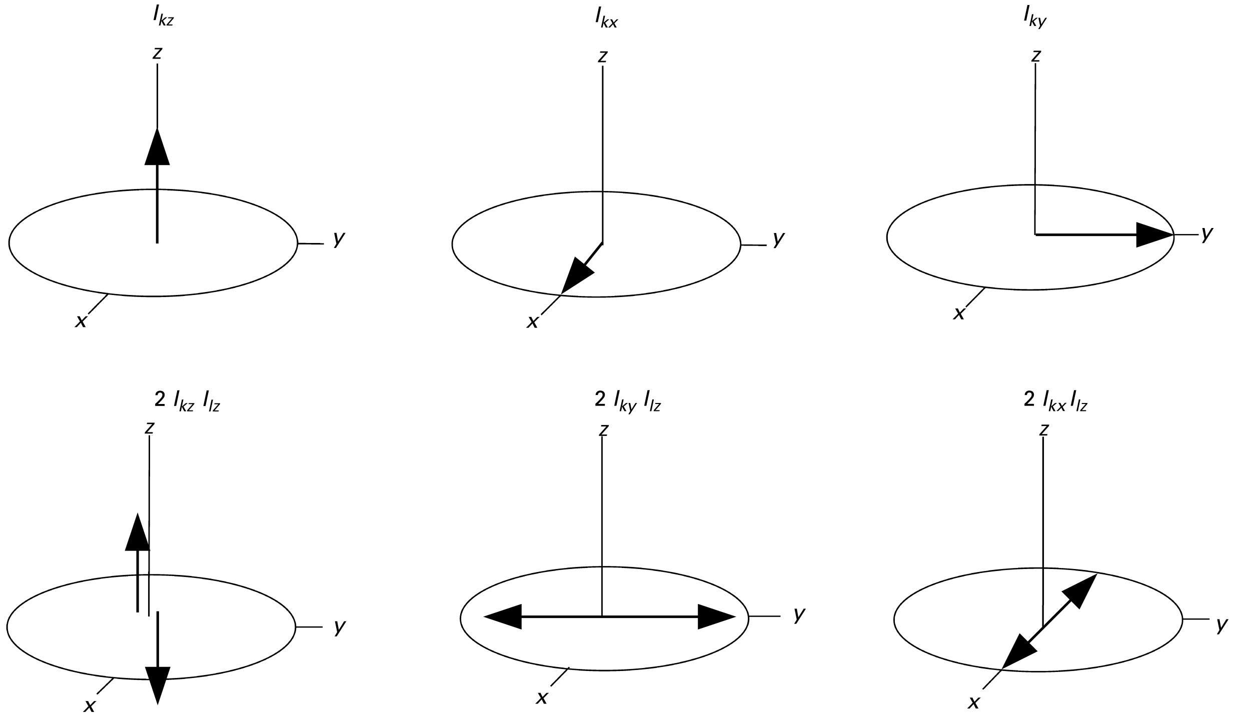



**Figure 1** Vector model representations of one-spin and two-spin product operators arising from a spin $k$ scalar coupled to a spin $l$. In the case of $2I_{kz}I_{lz}$ the $l$-spin vectors will be similarly arranged.

**Table 1** Product operators for two spin-½ nuclei

| *Product operator* | *Description* |
|---|---|
| $\frac{1}{2}E$ | $E$ = unity operator |
| $I_{kx}$, $I_{ky}$ | In-phase $x$ magnetization and $y$ magnetization, respectively, of spin $k$ |
| $I_{lx}$, $I_{ly}$ | In-phase $x$ magnetization and $y$ magnetization, respectively, of spin $l$ |
| $I_{kz}$, $I_{lz}$ | Longitudinal magnetization of spins $k$ and $l$, respectively |
| $2I_{kx}I_{lz}$, $2I_{ky}I_{lz}$ | $x$ and $y$ components, respectively, of $k$-spin magnetization antiphase with respect to $l$ |
| $2I_{kx}I_{lx}$, $2I_{kx}I_{ly}$, $2I_{ky}I_{lx}$, $2I_{ky}I_{ly}$ | Components of two-spin coherence between spins $k$ and $l$ |
| $2I_{kz}I_{lz}$ | Longitudinal two-spin order between $k$ and $l$ |

model in **Figure 1**. It should be noted that there are two methods for differentiating between spins. Different subscripts, $k$ and $l$ in the operator $2I_{kz}I_{lz}$, are commonly used to denote different spins in homonuclear systems, while in heteronuclear systems the operators themselves may be denoted by different letters, $I$ and $S$ in the operator $2I_zS_z$. In the latter case $I$ is usually taken to denote $^1H$. In the case of a doublet, longitudinal two-spin order arises when the equilibrium magnetization of one of its two transitions has been inverted. Antiphase magnetization arises when the two components of the doublet point in opposite directions in the $xy$ plane. The longitudinal component of the operator indicates the scalar coupling that separates the two components of the multiplet with opposite phases. The four operators denoting two-spin coherence consist of linear combinations of zero-quantum and double-quantum coherence; these will be considered below.

### Evolution

The evolution of product operators can be calculated in a similar way to that of the density matrix, by a series of transformations of the type

$$\exp\{-\mathrm{i}\phi B_r\}B_s\exp\{\mathrm{i}\phi B_r\} = \sum_t b_{ts}(r,\phi)B_t \qquad [3]$$

where $\phi B_r$, corresponds to the relevant Hamiltonian. This takes the form $(\omega_k\tau)I_{kz}$ for chemical shift evolution of a spin $k$ for a period $\tau$, $(\pi J_{kl}\tau)2I_{kz}I_{lz}$ for scalar coupling evolution between two weakly coupled spins $k$ and $l$ for a period $\tau$, and $(\beta)I_{kx}$ for a radiofrequency pulse applied to a spin $k$ producing a rotation of $\beta$ about the $x$ axis. In each case the bracket has been added to indicate the angle through which the affected operators will evolve. Since all terms in the free precession Hamiltonian commute, the evolution of individual chemical shifts and scalar couplings in multispin systems can be treated separately and in any order. In general, each process causes a rotation in a subspace, (**Figure 2**) corresponding to the evolution of one operator into another. In the case of a $\beta_x°$ pulse applied to the equilibrium magnetization of a spin $k$, a component of the magnetization may be rotated onto the $y$ axis:

$$I_{kz} \xrightarrow{\beta I_{kx}} I_{kz}\cos\beta - I_{ky}\sin\beta \qquad [4]$$

A vector description of this process is given in **Figure 3**. However, a $\beta_x°$ pulse has no effect on $I_{kx}$:

$$I_{kx} \xrightarrow{\beta I_{kx}} I_{kx} \qquad [5]$$

It is important to use the correct sign conventions, which are given in **Figure 2**. If a pulse acts on a product of several operators, its overall effect can be determined by calculating its effects on each individual operator and then multiplying out the results. For example

$$\begin{aligned} 2I_{kz}I_{lz} \xrightarrow{\beta(I_{kx}+I_{lx})}\; & 2(I_{kz}\cos\beta - I_{ky}\sin\beta) \\ & \times (I_{lz}\cos\beta - I_{ly}\sin\beta) \\ \equiv\; & 2I_{kz}I_{lz}\cos^2\beta \\ & - 2I_{kz}I_{ly}\cos\beta\sin\beta - 2I_{ky}I_{lz}\sin\beta \\ & \times \cos\beta + 2I_{ky}I_{ly}\sin^2\beta \end{aligned} \qquad [6]$$

Where a pulse affects all of the nuclear spins under consideration the shorthand '$\beta_x°$' is often written over the arrow.

Chemical shift evolution may be expressed in a similar fashion to that of radiofrequency pulses. For example, the chemical shift evolution of the operator $I_{ky}$ at a frequency of $\omega_k$ over a period $\tau$ can be written as

$$I_{ky} \xrightarrow{\omega_k\tau I_{kz}} I_{ky}\cos\omega_k\tau - I_{kx}\sin\omega_k\tau \qquad [7]$$

A vector model representation of this process is given in **Figure 3**.

Scalar coupling evolution results in the interconversion of components of in-phase and antiphase magnetization. In the case of a component of in-phase $k$-spin magnetization $I_{ky}$, evolution due to a scalar coupling to a spin $l$ during a period $\tau$ can be

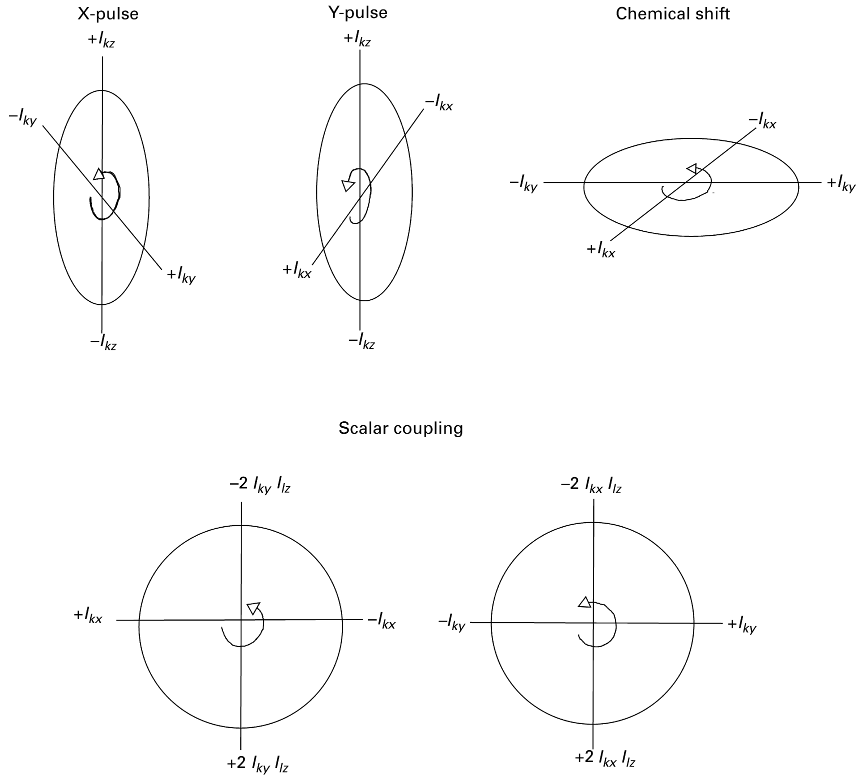



**Figure 2** Sign conventions for product operator evolution under radiofrequency pulse, chemical shift and scalar coupling evolution.

described by

$$I_{ky} \xrightarrow{\pi J_{kl}\tau 2I_{kz}I_{lz}} I_{ky}\cos\pi J_{kl}\tau - 2I_{kx}I_{lz}\sin\pi J_{kl}\tau \qquad [8]$$

This is represented pictorially in **Figure 3**. If the spin $k$ is scalar coupled to a further spin $m$, the effects of evolution due to the two scalar couplings can be calculated sequentially:

$$\begin{aligned} I_{ky} \xrightarrow{\pi J_{kl}\tau 2I_{kz}I_{lz}} & \xrightarrow{\pi J_{km}\tau 2I_{kz}I_{mz}} \\ & I_{ky}\cos\pi J_{kl}\tau\cos\pi J_{km}\tau \\ & - 2I_{kx}I_{mz}\cos\pi J_{kl}\tau\sin\pi J_{km}\tau \\ & - 2I_{kx}I_{lz}\sin\pi J_{kl}\tau\cos\pi J_{km}\tau \\ & - 4I_{ky}I_{lz}I_{mz}\sin\pi J_{kl}\tau\sin\pi J_{km}\tau \end{aligned} \qquad [9]$$

The term $4I_{ky}I_{lz}I_{mz}$ corresponds to a $y$ component of doubly antiphase $k$-spin magnetization; it is antiphase with respect to both $l$ and $m$. It is important to note that if a component of magnetization has become antiphase with respect to one spin, scalar coupling evolution with respect to another spin will not make it in-phase again; only a further period during which evolution occurs due to the first spin's scalar coupling will have that effect.

### Multiple-quantum coherence

The components of transverse magnetization considered above are all what are known as single-quantum coherences. Single-quantum coherences are phase coherences between states differing in overall magnetic quantum number by ±1. Phase coherences between states that do not fulfil this criterion are known as multiple-quantum coherence. Multiple-quantum coherences cannot be created by the application of a single nonselective radiofrequency pulse to the equilibrium magnetization and neither can they be detected directly since they have no net magnetization associated with them. Nevertheless, multiple-quantum coherences play an important role in many NMR experiments. The product operators $2I_{ky}I_{lx}$, $2I_{kx}I_{ly}$, $2I_{ky}I_{ly}$ and $2I_{kx}I_{lx}$ all describe linear combinations of multiple-quantum coherences; however, their precise meaning only becomes apparent when they are rewritten in terms of raising and lowering operators:

$$I_{kx} = \left(\frac{1}{2}\right)(I_k^+ + I_k^-) \qquad [10]$$

$$I_{ky} = \left(\frac{1}{2\mathrm{i}}\right)(I_k^+ - I_k^-) \qquad [11]$$

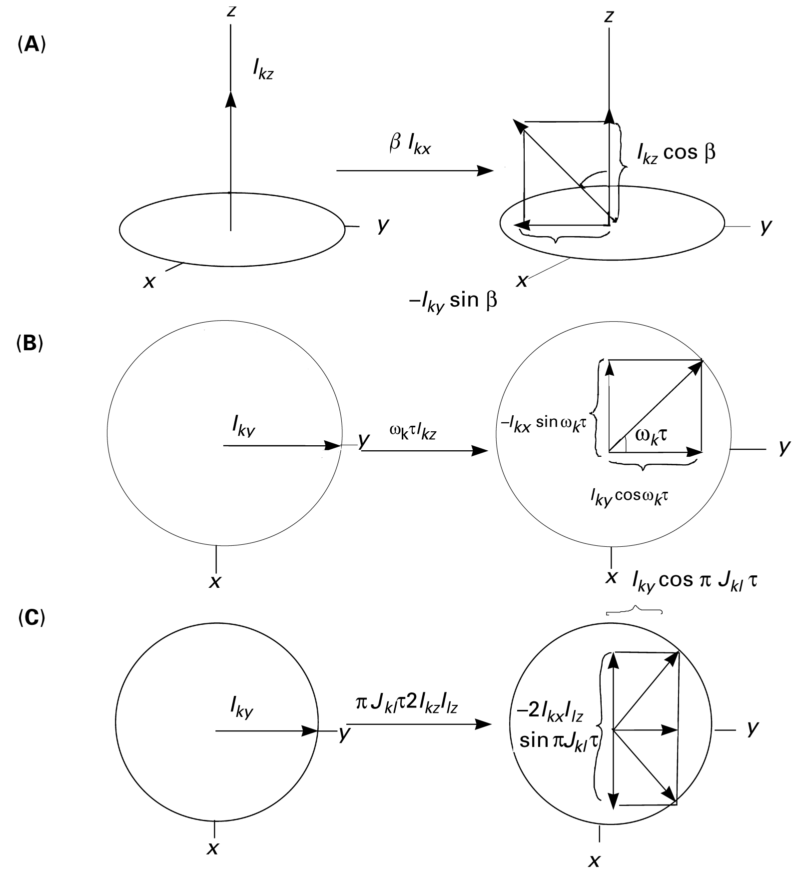



**Figure 3** Vector model representations of the evolution of *k*-spin magnetization. (A) The effect of a $\beta_x°$ pulse on *z* magnetization. (B) The effect of chemical shift evolution for a time $\tau$ on *y* magnetization. (C) The effect of scalar coupling evolution for a time $\tau$ on *y* magnetization.

The operators $I_k^+$ and $I_k^-$ are counterrotating components corresponding to +1 and −1 quantum coherence, respectively; the former precesses at a frequency of $+\omega_k$ and the latter at a frequency of $-\omega_k$. Re-writing $2I_{ky}I_{lx}$ in terms of raising and lowering operators we find that

$$2I_{ky}I_{lx} = \left(\frac{1}{2\mathrm{i}}\right)(I_k^+I_l^+ + I_k^+I_l^- - I_k^-I_l^+ - I_k^-I_l^-) \quad [12]$$

Clearly, $2I_{ky}I_{lx}$ consists of a linear combination of +2, 0, 0 and −2 quantum coherence. The other three operators can be re-written in a similar fashion. Pure zero-quantum coherence (ZQC) and double-quantum coherence (DQC) can be isolated by taking linear combinations of these four product operators:

$$\mathrm{ZQC}_x = \left(\frac{1}{2}\right)(2I_{kx}I_{lx} + 2I_{ky}I_{ly}) = \left(\frac{1}{2}\right)(I_k^+I_l^- + I_k^-I_l^+) \quad [13]$$

$$\mathrm{ZQC}_y = \left(\frac{1}{2}\right)(2I_{ky}I_{lx} - 2I_{kx}I_{ly}) = \left(\frac{1}{2\mathrm{i}}\right)(I_k^+I_l^- - I_k^-I_l^+) \quad [14]$$

$$\mathrm{DQC}_x = \left(\frac{1}{2}\right)(2I_{kx}I_{lx} - 2I_{ky}I_{ly}) = \left(\frac{1}{2}\right)(I_k^+I_l^+ - I_k^-I_l^-) \quad [15]$$

$$\mathrm{DQC}_y = \left(\frac{1}{2}\right)(2I_{kx}I_{ly} + 2I_{ky}I_{lx}) = \left(\frac{1}{2\mathrm{i}}\right)(I_k^+I_l^+ - I_k^-I_l^-) \quad [16]$$

The spins involved in a coherence, in this case *k* and *l*, are often referred to as the active spins. The evolution of these *x* and *y* components of multiple-quantum coherences can be calculated in the same fashion as single-quantum coherence ($I_{kx}$ and $I_{ky}$) with several important provisos:

(1) The precessional frequency of a multiple-quantum coherence is a linear combination of those of its active spins. A double-quantum coherence with active spins $k$ and $l$ will evolve at $(\omega_k + \omega_l)$, while the corresponding zero-quantum coherence will evolve at $(\omega_k - \omega_l)$

(2) Multiple-quantum coherences do not exhibit scalar couplings between their active spins. Their scalar couplings to passive spins are linear combinations of those of their individual active spins to the passive spin concerned. Thus a double-quantum coherence between spins $k$ and $l$ would exhibit a scalar coupling constant of $(J_{km}+J_{lm})$ to a passive spin $m$; the corresponding zero-quantum coherence scalar coupling constant would be $(J_{km}-J_{lm})$.

Consequently, the chemical shift evolution of the $y$ component of a zero-quantum coherence with active spins $k$ and $l$ during a period $\tau$ can be written as

$$\mathrm{ZQC}_y \xrightarrow{(\omega_k I_{kz}+\omega_l I_{lz})\tau} \mathrm{ZQC}_y \cos(\omega_k - \omega_l)\tau - \mathrm{ZQC}_x \sin(\omega_k - \omega_l)\tau \quad [17]$$

and the scalar coupling evolution of the corresponding double-quantum coherence to passive spin $m$ can be described by

$$\mathrm{DQC}_y \xrightarrow{(\pi J_{km} 2I_{kz}I_{mz} + \pi J_{lm} 2I_{lz}I_{mz})\tau} \mathrm{DQC}_y \cos(J_{km} + J_{lm})\pi\tau - \mathrm{DQC}_x 2I_{mz} \times \sin(J_{km} + J_{lm})\pi\tau \quad [18]$$

## Applications

The product operator formalism has been used to analyse most liquid-state NMR experiments since it was first introduced in 1983. The examples given here have been chosen either because of the way in which they exemplify the manipulation of product operators or because the pulse sequences concerned play an important role in contemporary NMR spectroscopy. The spin echo is considered because of its wide use and the opportunity it presents to undertake a relatively simple product operator calculation in full without resort to any of the short cuts that will be used to simplify later calculations. COSY (correlation spectroscopy) is also a widely used experiment and allows the introduction of the concept of coherence transfer, and its double-quantum filtered variant exemplifies some important points about how multiple-quantum coherence is handled in product operator calculations. The HSQC (heteronuclear single-quantum coherence) pulse sequence is widely used as a basis for many heteronuclear experiments, particularly for structural studies of proteins in solution, and DEPT (distortionless enhanced polarization transfer) is one of the mainstays of perhaps the largest group of NMR users, organic chemists.

### Spin echo

The spin echo is important both as an experiment in its own right for measuring transverse relaxation and as a building block of numerous other pulse sequences. The spin echo pulse sequence is

$$90^\circ_x - \tau - 180^\circ_y - \tau - \text{Acquisition}$$

The properties of this pulse sequence can be determined using the product operator formalism. Its effects on the equilibrium magnetization of a spin $k$ with a scalar coupling to a spin $l$ are calculated below. As an exercise the intermediate steps in the calculation are given in full (see Eqn [19] opposite).

It can be seen from the above that the product operators for even very simple systems can proliferate rapidly. Consolidation reduces the 16 terms present at the end of the calculation to two. It can be seen from the result that overall there is no evolution due to chemical shift (and by implication the effects of magnetic field inhomogeneities), which makes the pulse sequence useful for measuring transverse relaxation. However, scalar coupling evolution continues unaffected. In reality, it is rarely necessary to undertake a calculation of such complexity; the results of standard pulse sequences such as the spin echo are well known, so it is only necessary to write

$$I_{kz} \xrightarrow{90^\circ_x} -I_{ky} \xrightarrow{\tau,180^\circ_y,\tau} -I_{ky}\cos(2\pi J_{kl}\tau) + 2I_{kx}I_{lz}\sin(2\pi J_{kl}\tau) \quad [20]$$

In the remaining examples given below, the product operators will only be given at key stages during an experiment and any not following the selected coherence transfer pathway (and which are not observed) will be omitted.

### COSY

The COSY (correlation spectroscopy) experiment has become one of the mainstays of modern NMR spectroscopy: it is also one of the simplest

$$
\begin{aligned}
I_{kz} \xrightarrow{90^\circ_x} -I_{ky} \xrightarrow{\omega_k\tau I_{kz}} & -I_{ky}\cos(\omega_k\tau) + I_{kx}\sin(\omega_k\tau)\\
& \downarrow \pi J_{kl}\tau 2I_{kz}I_{lz}\\
& -I_{ky}\cos(\omega_k\tau)\cos(\pi J_{kl}\tau) + I_{kx}\sin(\omega_k\tau)\cos(\pi J_{kl}\tau)\\
& +2I_{kx}I_{lz}\cos(\omega_k\tau)\sin(\pi J_{kl}\tau) + 2I_{ky}I_{lz}\sin(\omega_k\tau)\sin(\pi J_{kl}\tau)\\
& \downarrow 180^\circ_y\\
& -I_{ky}\cos(\omega_k\tau)\cos(\pi J_{kl}\tau) - I_{kx}\sin(\omega_k\tau)\cos(\pi J_{kl}\tau)\\
& +2I_{kx}I_{lz}\cos(\omega_k\tau)\sin(\pi J_{kl}\tau) - 2I_{ky}I_{lz}\sin(\omega_k\tau)\sin(\pi J_{kl}\tau)\\
& \downarrow \omega_k\tau I_{kz}\\
& -I_{ky}\cos^2(\omega_k\tau)\cos(\pi J_{kl}\tau) - I_{kx}\sin(\omega_k\tau)\cos(\pi J_{kl}\tau)\cos(\omega_k\tau)\\
& +2I_{kx}I_{lz}\cos^2(\omega_k\tau)\sin(\pi J_{kl}\tau) - 2I_{ky}I_{lz}\sin(\omega_k\tau)\sin(\pi J_{kl}\tau)\cos(\omega_k\tau)\\
& +I_{kx}\cos(\omega_k\tau)\cos(\pi J_{kl}\tau)\sin(\omega_k\tau) - I_{ky}\sin(\omega_k\tau)^2\cos(\pi J_{kl}\tau)\\
& +2I_{ky}I_{lz}\cos(\omega_k\tau)\sin(\pi J_{kl}\tau)\sin(\omega_k\tau) + 2I_{kx}I_{lz}\sin^2(\omega_k\tau)\sin(\pi J_{kl}\tau)\\
& \downarrow \pi J_{kl}\tau 2I_{kz}I_{lz}
\end{aligned}
$$

$$
\begin{aligned}
& -I_{ky}\cos^2(\omega_k\tau)\cos^2(\pi J_{kl}\tau) - I_{kx}\sin(\omega_k\tau)\cos^2(\pi J_{kl}\tau)\cos(\omega_k\tau)\\
& +2I_{kx}I_{lz}\cos^2(\omega_k\tau)\sin(\pi J_{kl}\tau)\cos(\pi J_{kl}\tau) - 2I_{ky}I_{lz}\sin(\omega_k\tau)\sin(\pi J_{kl}\tau)\cos(\omega_k\tau)\cos(\pi J_{kl}\tau)\\
& +I_{kx}\cos(\omega_k\tau)\cos(\pi J_{kl}\tau)\sin(\omega_k\tau)\cos(\pi J_{kl}\tau) - I_{ky}\sin(\omega_k\tau)^2\cos^2(\pi J_{kl}\tau)\\
& +2I_{ky}I_{lz}\cos(\omega_k\tau)\sin(\pi J_{kl}\tau)\sin(\omega_k\tau)\cos(\pi J_{kl}\tau) + 2I_{kx}I_{lz}\sin^2(\omega_k\tau)\sin(\pi J_{kl}\tau)\cos(\pi J_{kl}\tau)\\
& +I_{kx}I_{lz}\cos^2(\omega_k\tau)\cos(\pi J_{kl}\tau)\sin(\pi J_{kl}\tau) - 2I_{ky}I_{lz}\sin(\omega_k\tau)\cos(\pi J_{kl}\tau)\cos(\omega_k\tau)\sin(\pi J_{kl}\tau)\\
& +I_{ky}\cos^2(\omega_k\tau)\sin^2(\pi J_{kl}\tau) + I_{kx}\sin(\omega_k\tau)\sin^2(\pi J_{kl}\tau)\cos(\omega_k\tau)\\
& +2I_{ky}I_{lz}\cos(\omega_k\tau)\cos(\pi J_{kl}\tau)\sin^2(\omega_k\tau) + 2I_{kx}I_{lz}\sin(\omega_k\tau)^2\cos(\pi J_{kl}\tau)\sin(\pi J_{kl}\tau)\\
& -I_{kx}\cos(\omega_k\tau)\sin(\pi J_{kl}\tau)\sin(\omega_k\tau)\sin(\pi J_{kl}\tau) + I_{ky}\sin^2(\omega_k\tau)\sin^2(\pi J_{kl}\tau)
\end{aligned}
$$

$$\equiv -I_{ky}\cos(2\pi J_{kl}\tau) + 2I_{kx}I_{lz}\sin(2\pi J_{kl}\tau) \qquad [19]$$

two-dimensional experiments:

$$90^\circ_x - t_1 - 90^\circ_\phi - \text{ Acquisition}$$

where $\phi = x, y$. The value of $\phi$ determines whether the $\sin(\omega_k t_1)$ or the $\cos(\omega_k t_1)$ modulated component of the data is observed; it is necessary to measure both if absorptive phase-sensitive spectra are to be obtained. The coherence transfer pathways detected in each case are given below. The magnetization of a spin $k$ coupled to a spin $l$ is considered. When $\phi = y$:

$$
\begin{aligned}
I_{kz} \xrightarrow{90^\circ_x} - I_{ky} \xrightarrow{t_1} & - I_{ky}\cos\omega_k t_1 \cos\pi J_{kl}t_1\\
& + 2I_{kx}I_{lz}\cos\omega_k t_1 \sin\pi J_{kl}t_1\\
& \downarrow 90^\circ_y\\
& - I_{ky}\cos\omega_k t_1 \cos\pi J_{kl}t_1\\
& - 2I_{kz}I_{lx}\cos\omega_k t_1 \sin\pi J_{kl}t_1
\end{aligned}
\qquad [21]
$$

and when $\phi = x$:

$$
\begin{aligned}
I_{kz} \xrightarrow{90^\circ_x} - I_{ky} \xrightarrow{t_1} & - I_{kx}\sin\omega_k t_1 \cos\pi J_{kl}t_1\\
& + 2I_{ky}I_{lz}\sin\omega_k t_1 \sin\pi J_{kl}t_1\\
& \downarrow 90^\circ_x\\
& - I_{kx}\sin\omega_k t_1 \cos\pi J_{kl}t_1\\
& - 2I_{kz}I_{ly}\sin\omega_k t_1 \sin\pi J_{kl}t_1
\end{aligned}
\qquad [22]
$$

These equations provide an introduction to the concept of coherence transfer, the process by which one coherence is transformed into another. In each case a component of $k$-spin magnetization is transformed into $l$-spin magnetization. In the first equation the coherence transfer process, brought about by the second 90° pulse, is

$$2I_{kx}I_{lz} \xrightarrow{90^\circ_y} - 2I_{kz}I_{lx} \qquad [23]$$

Coherence transfer plays a central role in many NMR experiments and usually involves antiphase magnetization. The components of in-phase magnetization present at the end of the schemes above have evolved at $\omega_k$ during $t_1$ and will evolve at $\omega_k$ during the acquisition time $t_2$: consequently they will give rise to a diagonal COSY peak with the coordinates $(\omega_k, \omega_k)$. However, the antiphase magnetization will have evolved at $\omega_k$ during $t_1$, but since it underwent coherence transfer from $k$ to $l$ at the second 90° pulse it will evolve at $\omega_l$ during the acquisition period. Consequently it will give rise to an off-diagonal peak in the COSY spectrum at the coordinates $(\omega_k, \omega_l)$. The components of magnetization giving rise to the diagonal and off-diagonal peaks are 90° out of phase; $I_{ky}$ and $2I_{kz}I_{lx}$ in the case of the $\cos(\omega_k, t_1)$ modulated components. This leads to diagonal peaks that are always dispersive when the off-diagonal peaks are phased to be absorptive, and vice versa.

### Double-quantum filtered COSY

Double-quantum filtered COSY is now the most widely used version of the COSY experiment. It has two main advantages over its precursor: diagonal and off-diagonal peaks can be phased to be absorptive simultaneously, and singlets are removed from the spectrum. The pulse sequence is

$$90^\circ_x - t_1 - 90^\circ_\phi - 90^\circ_x - \text{Acquisition}$$

where $\phi = x, y$ to select both $\sin(\omega_k, t_1)$ and $\cos(\omega_k, t_1)$ modulated components of the data. The selected pathway followed by the magnetization of a spin $k$ with a single scalar coupling partner $l$ is given below. When $\phi = x$:

$$\begin{gathered} I_{kz} \xrightarrow{90^\circ_x} -I_{ky} \xrightarrow{t_1} +2I_{kx}I_{lz}\cos\omega_k t_1 \sin\pi J_{kl}t_1 \\ \downarrow 90^\circ_x \\ \left(\frac{1}{2}\right)(2I_{ky}I_{lx} + 2I_{kx}I_{ly})\cos\omega_k t_1 \sin\pi J_{kl}t_1 \\ \downarrow 90^\circ_x \\ \left(\frac{1}{2}\right)(2I_{kz}I_{lx} + 2I_{kx}I_{lz})\cos\omega_k t_1 \sin\pi J_{kl}t_1 \end{gathered} \quad [24]$$

and when $\phi = y$:

$$\begin{gathered} I_{kz} \xrightarrow{90^\circ_x} -I_{ky} \xrightarrow{t_1} +2I_{ky}I_{lz}\sin\omega_k t_1 \sin\pi J_{kl}t_1 \\ \downarrow 90^\circ_y \\ \left(\frac{1}{2}\right)(2I_{ky}I_{lx} + 2I_{kx}I_{ly})\sin\omega_k t_1 \sin\pi J_{kl}t_1 \\ \downarrow 90^\circ_x \\ \left(\frac{1}{2}\right)(2I_{kz}I_{lx} + 2I_{kx}I_{lz})\sin\omega_k t_1 \sin\pi J_{kl}t_1 \end{gathered} \quad [25]$$

The second 90° pulse actually creates $2I_{ky}I_{lx}$ or $2I_{kx}I_{ly}$, depending on $\phi$. However, phase cycling (or magnetic field gradient pulses) select only that component of the operator corresponding to double-quantum coherence. Using Equations [13]–[16], these operators can be resolved into components of zero-quantum and double-quantum coherence. For example, in the case of $2I_{ky}I_{lx}$:

$$2I_{ky}I_{lx} = \left(\frac{1}{2}\right)(2I_{ky}I_{lx} + 2I_{kx}I_{ly}) + \left(\frac{1}{2}\right)(2I_{ky}I_{lx} - 2I_{kx}I_{ly}) \quad [26]$$

where the first linear combination on the right of the equation consists of pure double-quantum coherence and the second pure zero-quantum coherence. The diagonal and off-diagonal peaks now both arise from components of antiphase magnetization with the same phase.

### HSQC

The HSQC (heteronuclear single-quantum coherence) pulse sequence is an important tool in heteronuclear $^{1}H$–$^{15}N$ and $^{1}H$–$^{13}C$ NMR spectroscopy, particularly for larger molecules such as proteins. The pulse sequence is given below:

$$\begin{array}{ll} I & 90^\circ_x - \left(\frac{1}{4J_{IS}}\right) - 180^\circ_y - \left(\frac{1}{4J_{IS}}\right) - 90^\circ_y - \frac{t_1}{2} - \\ S & \qquad\qquad 180^\circ_y \qquad\qquad\qquad 90^\circ_y \end{array}$$

$$\begin{array}{l} 180^\circ_y - \frac{t_1}{2} - 90^\circ_y - \left(\frac{1}{4J_{IS}}\right) - 180^\circ_y - \left(\frac{1}{4J_{IS}}\right) - \text{Acquisition} \\ \qquad\quad 90^\circ_\phi \qquad\qquad\qquad 180^\circ_y \qquad\qquad\qquad \text{B.B. decoupling} \end{array}$$

where $\phi = x, y$ in alternate experiments. HSQC is used both in its own right and as a building block for

more sophisticated experiments. It produces two-dimensional spectra correlating the chemical shift of $^{1}$H ($I$) in the $F_2$ dimension with that of the bonded heteronucleus ($S$) in the $F_1$ dimension. To maximize sensitivity, $^{1}$H magnetization is both initially excited and detected. In the scheme given below, the magnetogyric ratios of both nuclei are assumed to have the same sign, and hence to rotate with the same sense in response to a radiofrequency pulse:

$$\begin{aligned}
&I_z \xrightarrow{90^\circ_x(I)} -I_y \xrightarrow{\frac{1}{4J_{IS}},180^\circ_y(I,S),\frac{1}{4J_{IS}}} +2I_xS_z \xrightarrow{90^\circ_y(I,S)} -2I_zS_x \\
&\downarrow \tfrac{t_1}{2},180^\circ_x(I),\tfrac{t_1}{2} \\
&+2I_zS_x\cos(\omega_S t_1) \;+\; 2I_zS_y\sin(\omega_S t_1) \\
&\quad 90^\circ_y(I,S) \;\; \text{(OR)} \;\; 90^\circ_x(S),90^\circ_y(I) \\
&-2I_xS_z\cos(\omega_S t_1) \;+\; 2I_xS_z\sin(\omega_S t_1) \\
&\downarrow \tfrac{1}{4J_{IS}},180^\circ_y(I,S),\tfrac{1}{4J_{IS}} \\
&-I_y\cos(\omega_S t_1) \;\;\text{(OR)}\;\; I_y\sin(\omega_S t_1) \qquad [27]
\end{aligned}$$

Since the length of each spin echo is ($1/2J_{IS}$), in-phase magnetization present at the begining of one will become completely antiphase by the end and vice versa. The scheme splits into two after the evolution period to show how both sine and cosine $t_1$-modulated data are acquired in consecutive experiments by changing the phase $\phi$ of the second 90°($S$) pulse; this is necessary to obtain $F_1$ spectra that are both phase sensitive and absorptive.

## DEPT

The DEPT (*d*istortionless *e*nhanced *p*olarization *t*ransfer) pulse sequence is probably one of the most widely used of all NMR experiments. It produces $^{13}$C subspectra edited according to the number of protons bonded to each $^{13}$C. The DEPT pulse sequence is

$$\begin{aligned}
&^{1}\text{H} \qquad 90^\circ_x - \left(\frac{1}{2J_{\text{CH}}}\right) - 180^\circ_x - \left(\frac{1}{2J_{\text{CH}}}\right) - \theta^\circ_y \\
&^{13}\text{C} \qquad\qquad\qquad\qquad 90^\circ_x \qquad\qquad\qquad 180^\circ_x \\
&\qquad - \left(\frac{1}{2J_{\text{CH}}}\right) - \text{B.B. decoupling} \\
&\qquad\qquad \text{Acquisition}
\end{aligned}$$

The experiment is repeated with $\theta$ set to 45°, 90° and 135° and linear combinations of the resulting data are taken to obtain the edited subspectra. The pulse sequence contains staggered $^{1}$H and $^{13}$C spin echos. This structure results in the cancellation of all chemical shift evolution by the start of acquisition, and consequently nothing is lost by omitting chemical shift evolution from the product operator description. The analysis for a $^{13}$CH group is given below:

$$\begin{aligned}
&I_z \xrightarrow{90^\circ_x(I)} -I_y \xrightarrow{\left(\frac{1}{2J_{\text{CH}}}\right)} +2I_xS_z \xrightarrow{90^\circ_x(S)} \\
&-2I_xS_y \xrightarrow{\left(\frac{1}{2J_{\text{CH}}}\right)} -2I_xS_y \\
&\downarrow \theta^\circ_y(I), 180^\circ_x(S) \\
&-2I_zS_y\sin\theta \\
&\downarrow \left(\tfrac{1}{2J_{\text{CH}}}\right) \\
&+S_x\sin\theta
\end{aligned} \qquad [28]$$

The 90°($S$) pulse converts the antiphase magnetization present after the first period of free precession into heteronuclear two-spin coherences. This corresponds to a linear combination of heteronuclear zero-quantum and double-quantum coherence. Since both are selected and both evolve in the same way under the conditions of the experiment, there is no need to separate them and consider their evolution separately.

As noted above, all chemical shift evolution cancels out and neither coherence will exhibit significant scalar coupling evolution during the subsequent ($1/2J_{\text{CH}}$) evolution period. The $\sin\theta$ dependence of the coherence that is ultimately observed on the angle of the subsequent $^{1}$H pulse is important, as will become evident below. This pulse converts part of the coherence into antiphase $^{13}$C single-quantum coherence which is observed after it has become in-phase.

The product operator description for a $^{13}$CH$_2$ group is the same except for the last three steps:

$$\begin{aligned}
&I_z \xrightarrow{90^\circ_x(I)\left(\frac{1}{2J_{\text{CH}}}\right),\,90^\circ_x(S)} -2I_xS_y \xrightarrow{\left(\frac{1}{2J_{\text{CH}}}\right)} \\
&+4I_xS_xI_z \xrightarrow{\theta^\circ_y(I),\,180^\circ_x(S)} 4I_zS_xI_z\sin\theta\cos\theta \\
&\downarrow \left(\tfrac{1}{2J_{\text{CH}}}\right) \\
&-S_x\sin\theta\cos\theta
\end{aligned} \qquad [29]$$

In this case the two-spin coherences generated by the 90°($S$) pulse both exhibit large scalar couplings to the remaining proton in the $^{13}$CH$_2$ group. These are

($J_{CH} - J_{HH}$) and ($J_{CH} + J_{HH}$) for the zero-quantum and double-quantum coherences, respectively. However, since scalar couplings are ineffective between equivalent spins, this reduces to $J_{CH}$ in each case. Consequently, the operator evolves to become completely antiphase with respect to the remaining proton in the $^{13}CH_2$ group during the subsequent ($1/2J_{CH}$) period. This time the observed magnetization has a $\sin\theta\cos\theta$ dependence on the subsequent $^1$H pulse.

For a $^{13}CH_3$ group:

$$I_z \xrightarrow{90^\circ_x(I)\ \left(\frac{1}{2J_{CH}}\right),\ 90^\circ_x(S)} -2I_xS_y \xrightarrow{\left(\frac{1}{2J_{CH}}\right)}$$
$$+8I_xS_yI_zI_z \xrightarrow{\theta^\circ_y(I),\ 180^\circ_x(S)} 8I_zS_yI_zI_z \sin\theta\cos^2\theta \quad [30]$$
$$\downarrow \left(\frac{1}{2J_{CH}}\right)$$
$$-S_x \sin\theta\cos^2\theta$$

The two-spin coherences generated by the 90°($S$) pulse exhibit scalar couplings of $J_{CH}$ to the two remaining protons in the $^{13}CH_2$ group and consequently evolve to become antiphase with respect to both of them during the subsequent ($1/2J_{CH}$) period. As a result, the observed magnetization will exhibit a $\sin\theta\cos^2\theta$ dependence. Since each group has a different dependence on the θ° pulse, it is possible to isolate the subspectra arising from each group by acquiring data with three values of θ and taking the appropriate linear combinations of the results.

## List of symbols

$B_s$ = orthogonal base operator; $E$ = unity operator; $I_x$, $I_y$, $I_z$ = angular momentum operators; $J$ = coupling constant; $k$, $l$, $m$ = spin indices; $N$ = number of spin-½ nuclei; $q$ = number of operators in product; $S_x$, $S_y$, $S_z$ = angular momentum operators; $t_1$ = evolution time; $t_2$ = acquisition time; θ = pulse angle; σ($t$) = density operator; τ = duration of evolution; ϕ = pulse phase; ω = angular frequency of spin (nucleus).

*See also:* **Two-Dimensional NMR Methods; $^{13}$C NMR, Methods; NMR Principles; NMR Pulse Sequences; Proteins Studied Using NMR Spectroscopy.**

## Further reading

Ernst RR, Bodenhausen G and Wokaun A (1987) *Principles Of Nuclear Magneic Resonance in One and Two Dimesions*. Oxford: Clarendon Press.

Kessler H, Gehrke M and Griesinger C (1988) Two-dimensional NMR spectroscopy: background and overview of the experiments. *Angewandte Chemie International Edition in English* **27**: 490–536.

Sorensen OW, Eich GW, Levitt MH, Bodenhausen G and Ernst RR (1983) Product operator formalism for the description of NMR pulse experiments. *Progress in NMR Spectroscopy* **16**: 163–192.

# Proteins Studied Using NMR Spectroscopy

**Paul N Sanderson**, GlaxoWellcome Research and Development, Stevenage, UK

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| Applications |

## Introduction

The study of proteins by NMR spectroscopy has gained great impetus in recent years, providing a focus for the proliferation of many new complex NMR experiments and, quite possibly, justification for the purchase of more high-field spectrometers than any other field within NMR. The first $^1$H NMR spectrum of a protein was published in 1957; this accurately reflected the amino acid composition but had neither the sensitivity nor resolution to yield further information. In the last two decades, however, many protein structures have been characterized and greater insight into the activity of proteins has been obtained from protein NMR studies. These achievements became possible as a result of the concurrent development of higher field (≥ 500 MHz) superconducting magnets, powerful computational hardware and software, complex multidimensional heteronuclear NMR experiments and isotopic labelling techniques. The quest for structural knowledge has been driven by the recognition that functions of biologically active proteins (such as enzymes, hormones and receptors) are fundamentally dependent on their three-dimensional structure. The development of NMR techniques, in parallel with X-ray crystallography, to obtain greater structural information from increasingly complex protein systems has resulted in increased understanding of biological processes. NMR also has a role in the characterization of protein interactions via, for example, titrations that map the binding surface of a protein through specific chemical shift changes. These binding studies, which are discussed in a separate article, benefit from prior characterization of solution structure by NMR and it is the structural aspects of protein NMR that are the focus of the present article. The theory and practical aspects of many areas of NMR spectroscopy which are of importance to protein NMR studies are described in detail elsewhere in this Encyclopedia; these are consequently mentioned only briefly here. Protein NMR has been thoroughly documented in many books and reviews and for more detailed discussion the reader is directed towards the Further reading section.

## What proteins are suitable for study by NMR?

Full structural characterization of proteins in solution is possible for proteins of up to ~100 amino acids using homonuclear proton NMR. For larger proteins, of up to ~30 kDa, isotopic enrichment with $^{15}$N and $^{13}$C is required. There is no clear 'cut-off' in terms of protein size; each protein has to be considered on its own merits. Often, the likelihood of a successful structure determination only becomes clear after considerable protein purification and preliminary one-dimensional (1D) $^1$H NMR data have been obtained.

Proteins must be non-aggregated and monomeric, or at least not present as large heteromultimers, under conditions of NMR measurements. Aggregated proteins give increased line widths and thus reduced spectral resolution and sensitivity. Internal mobility and the presence of multiple interconverting conformations also influence resonance line width.

## Protein structure

Proteins are composed of a linear sequence of L-amino acid residues linked via amide bonds. The 20 naturally occurring amino acids are distinguished by the chemical nature of their side-chains. The amide bonds, or peptide linkages, are essentially planar and provide structural rigidity to the protein backbone, the only freedom to rotate being around the bonds to the α carbon. The angles of rotation are called, in IUPAC nomenclature, phi ($\phi$) for the N–Cα bond [$C(O)_{i-1}$–$N_i$–$C\alpha_i$–$C(O)_i$] and psi ($\psi$) for the Cα–C(O) bond [$N_i$–$C\alpha_i$–$C(O)_i$–$N_{i+1}$] (**Figure 1**). If these two torsion angles are known for each amino acid the conformation of the whole polypeptide backbone is defined.

The linear sequence of amino acids represents the *primary structure* of the protein. Local regions within this sequence can adopt stable, defined, *secondary structure* such as α helices or β sheets. The packing of these secondary structural elements into compact domains gives rise to the protein's *tertiary structure*, such that distant regions of the peptide chain can be spatially close together. Multimeric

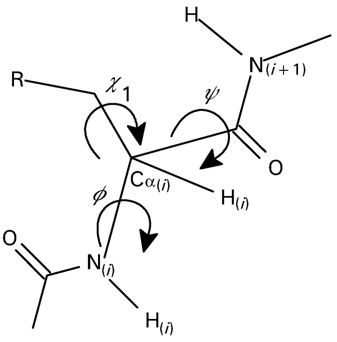



**Figure 1** Stylized representation of a portion of a polypeptide chain, indicating the nomenclature of the backbone torsion angles ($\phi$ and $\psi$) and the side-chain torsion angle ($\chi_1$) for the bonds emanating from the α carbon of an amino acid residue(*i*).

proteins are composed of several polypeptide chains arranged together in a *quaternary structure*. Adoption of the correct tertiary and quaternary structure is usually essential for biological function of a protein and a central dogma of biochemistry is that knowledge of a protein's structure leads to greater understanding of its activity.

## Isotopic labelling of proteins

The range of proteins that can be studied by NMR and the nature of the structural information available have been extended through the biosynthetic incorporation of $^{15}N$ and $^{13}C$ isotopes. Data on isotopes encountered most frequently in protein NMR are given in **Table 1.**

Most proteins for structural analysis are prepared in cultures of bacteria or yeast that have been genetically modified to overexpress the protein of interest. Typically, a primary expression of 20–100 mg of protein is required for a structural NMR study. *Escherichia coli* (*E. coli*) bacteria are preferentially used as they can be grown rapidly in large quantities on chemically defined media which can be supplemented with $^{15}N$-labelled ammonium salts and $^{13}C$-labelled compounds, such as glucose, acetate or glycerol, as required. It is essential to exclude all sources of natural abundance nitrogen and carbon from growth media to maximize incorporation of isotopic label. The production of $^{15}N$-labelled protein is reasonably straightforward and relatively inexpensive, provided it can be expressed in *E. coli*. Additional incorporation of $^{13}C$-label is more costly, but necessary for full analysis of larger proteins. The activity of expressed proteins should be validated to ensure correct folding and show that isotope incorporation has not impaired function.

**Table 1** Properties of isotopes most commonly encountered in protein NMR

| *Isotope* | *Natural abundance* (%) | *Resonance frequency at 14.0926 T* (MHz) | *Relative sensitivity* |
|---|---|---|---|
| $^{1}H$ | 99.98 | 600.00 | 1.0000 |
| $^{2}H$ | 0.015 | 92.10 | 0.00965 |
| $^{13}C$ | 1.108 | 150.86 | 0.0159 |
| $^{15}N$ | 0.37 | 60.80 | 0.00104 |
| $^{31}P$ | 100 | 242.88 | 0.0663 |

Mammalian proteins produced in bacteria or yeast may not undergo correct post-translational processing, i.e. disulfide cross-linking, protein folding, phosphorylation or glycosylation. This problem may be overcome by expression in baculovirus systems, in insect cells or in mammalian cell lines such as Chinese hamster ovary (CHO) cells. Mammalian cells will not grow on minimal media, so their growth media must contain appropriately labelled amino acids, which can be obtained, for example, from hydrolysates of labelled algae.

Isotopic labelling with deuterium ($^{2}H$) can be used to provide spectral editing. For example, specific incorporation of a deuterium atom into methylene or methyl groups of amino acids can be used to obtain detailed structural and dynamic information. Dramatic increases in proton NMR resolution for larger proteins can be achieved through random labelling of the protein with deuterium, at levels between 50 and 85%, by growth on substrates in which the ratio of $^{1}H$ to $^{2}H$ is controlled.

## Preparation of protein samples for NMR

A typical sample for protein NMR will contain 1 mM protein, of at least 95% purity, in 0.5 mL of aqueous solution. The final stages of purification may include desalting, buffer exchange, $^{2}H_2O$ exchange and concentration by lyophilization (if the protein is sufficiently stable), vacuum centrifugation or ultrafiltration. Sample volumes can be as small as 100 μL (for 2.5 or 3 mm diameter tubes) if only limited amounts of protein are available.

Data collection can take several days, during which the protein must remain stable; therefore oxidation, hydrolysis and microbial contamination must be minimized. Samples are not usually degassed, as protein solutions tend to 'froth' and paramagnetic broadening by oxygen can usually be ignored. Cysteine and methionine residues are susceptible to oxidation, so spontaneous formation of

non-native intra- or intermolecular disulfide bonds should be prevented by addition of low levels of dithiothreitol or β-mercaptoethanol. The growth of microorganisms can be prevented by sodium azide.

The quality of NMR spectra of proteins can be strongly influenced by sample pH, ionic strength, buffer, concentration and temperature; optimum conditions are generally determined empirically via 1D spectra. For $^{1}$H NMR, solutions are generally buffered, at 10–50 mM, with phosphate (which can cause precipitation or aggregation of some proteins) or with deuterated buffer salts (e.g. Tris). Deuterated reducing agents, cation chelators and proteolytic enzyme inhibitors may also be incorporated as necessary.

## Obtaining NMR data from proteins

Protein NMR spectra are usually recorded in aqueous solutions to obtain data from the exchangeable amide protons. In these spectra the solvent water signal is at least $10^{5}$ times more intense than the protein protons of interest and must be suppressed to allow detection of protein signals at an acceptable signal-to-noise ratio. The most common method for suppressing the water signal is presaturation, in which continuous low level irradiation is applied at the frequency of the water resonance. This is the most effective method if the effects of exchange between solvent and solute are not important. Many alternative methods that do not reduce the intensity of solute signals in exchange with the water are available; these include the use of pulsed-field gradients. Full discussion of solvent suppression techniques can be found in a separate article.

### One-dimensional NMR data

Initial assessment of the suitability of a protein sample and preliminary solution optimization can be achieved via 1D $^{1}$H NMR experiments; these should also be used to assess the consistency of a sample before and after lengthy acquisitions. Broad signals may indicate protein aggregation and suggest that further experiments are not worth pursuing on that sample. The presence (or absence) of protein structure can be assessed from the distribution of signals; typically, upfield shifted methyl signals (δ0), and downfield shifted alpha proton (δ5–6) and NH proton (>δ9) signals indicate that a protein is 'structured'. These regions are highlighted in a $^{1}$H NMR spectrum of lysozyme in **Figure 2**. For a denatured (or random coil) protein, such signals are absent from these regions.

### Multidimensional heteronuclear NMR experiments

Having established optimum sample conditions, all subsequent assignment and structural data are obtained from multidimensional NMR experiments. For smaller proteins, this can be achieved using two-dimensional (2D) homonuclear proton NMR experiments: correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY) to give scalar through-bond coupling data, and nuclear Overhauser effect spectrometry (NOESY) to provide 'through space' information. For larger proteins, resonance overlap is considerable, increased line width results in cancellation of COSY cross peaks and much signal loss occurs during long TOCSY mixing times. Full assignment and NOE analysis can, in these cases, be achieved by spreading data into three or four dimensions to increase signal resolution and by further spectral simplification through $^{15}$N and $^{13}$C isotope editing.

In heteronuclear correlation experiments, magnetization transfer between protons and heteronuclei can be via either heteronuclear single quantum coherence (HSQC) or heteronuclear multiple quantum coherence (HMQC) pathways. The HSQC sequence gives rise to narrower lines, but uses more pulses and requires a longer phase cycle than the HMQC. Thus, HSQC is used for 2D experiments where the highest resolution is required and HMQC is preferred for 3D sequences in which the experimental time is limited.

Labelling with $^{15}$N alone can be sufficient to overcome spectral overlap for proteins of up to 20 kDa and, for these proteins, virtually complete resolution can often be achieved for the backbone amide groups in 2D-$^{1}$H–$^{15}$N HSQC experiments. These experiments are very robust and can be used to determine amide proton exchange rates or chemical shift temperature coefficients. For high protein concentrations $^{1}$H–$^{15}$N HSQC data sets can be acquired rapidly; typically within 10 min for a 2 mM sample or 2–3 hr for a 0.2 mM protein sample. Consequently, this experiment has become the mainstay of NMR approaches to monitor the binding of ligands to $^{15}$N-labelled proteins through titration experiments.

Information on side-chain resonances or sequential connectivities can be obtained by converting the 2D-$^{1}$H–$^{15}$N heteronuclear correlation experiment into a 3D experiment by adding either a TOCSY or a NOESY step. These 3D experiments can be considered as a series of 2D homonuclear experiments in which each is edited by a different $^{15}$N frequency. Thus, having established amide $^{1}$H/$^{15}$N pairs via the HSQC, the TOCSY-HMQC correlates alpha protons

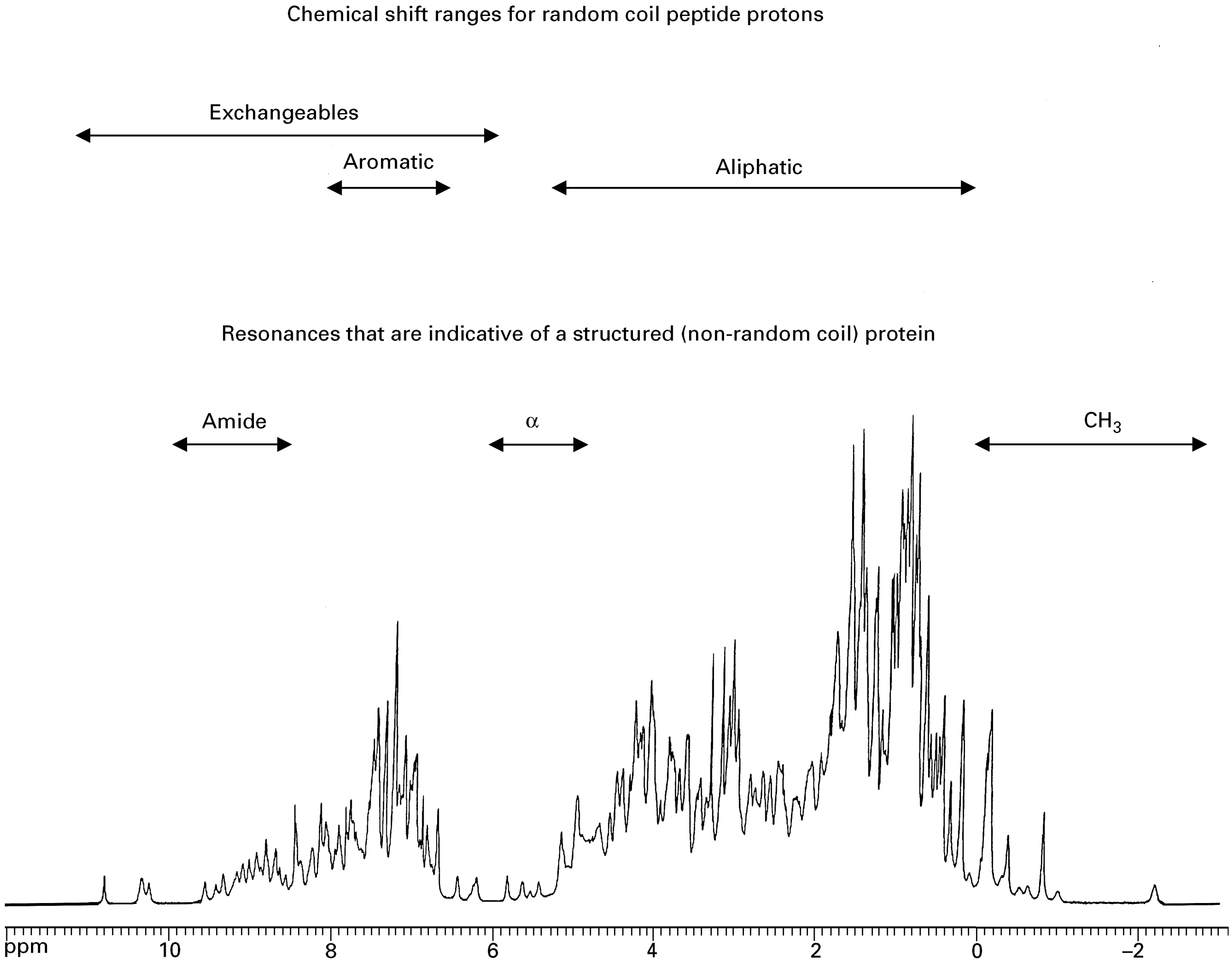



**Figure 2** One-dimensional $^1$H NMR spectrum of lysozyme in aqueous solution obtained with presaturation of the water signal, illustrating the range of proton chemical shifts expected for a random coil, or denatured, protein. The positions of upfield shifted methyl signals and downfield shifted alpha and amide proton signals that are indicative of a non-random, ordered protein structure are also shown.

(and in favourable cases side-chain protons) to these pairs and residue types can be assigned. Sequential assignment can then follow in a manner analogous to the 2D procedure (see below) via the NOESY-HMQC spectrum. Several other 3D experiments have been devised to facilitate assignments of $^{15}$N-labelled proteins. One such, which is of particular use for identifying helical regions of a protein, is the HMQC-NOESY-HMQC experiment, which allows identification of $^1$H–$^1$H NOEs between residues with degenerate amide proton resonances.

For larger, double-labelled proteins, assignments are made via a range of three- and four-dimensional triple resonance experiments in which sequential assignments are facilitated via magnetization transferred through $^{13}$C-couplings. These couplings are usually at least as large as the $^{13}$C line width for proteins of up to 30 kDa and are reasonably independent of the protein backbone conformation. A summary of the magnitudes of coupling constants utilized in correlation experiments is shown in **Figure 3**. These experiments each focus on a particular coupling network (their nomenclature usually reflects this in a logical manner) and several different experiments may perform the same correlation, but via a different route. Residue-specific experiments can extend assignments from the peptide backbone along amino acid side-chains, for example through to the guanidino group of arginine residues.

## Sequence-specific assignments

NMR spectra of small proteins can be fully assigned in a systematic manner by a sequential assignment procedure using well-resolved 2D homonuclear spectra and a knowledge of the amino acid sequence of

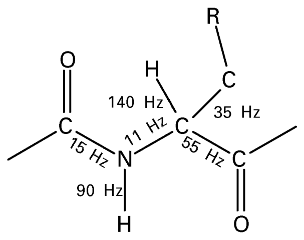



**Figure 3** Schematic representation of the peptide backbone, showing the magnitudes of the one-bond coupling constants that are utilized in multidimensional heteronuclear correlation experiments to provide sequential connectivities.

the protein. This procedure depends on the fact that for all of the sterically allowed values of the torsion angles $\phi$, $\psi$ and $\chi_1$ (**Figure 1**) NOEs will be observed for at least one of the interproton distances between NH, αH or βH of adjacent amino acids. Hence, spin systems of individual amino acids are first identified via spin–spin couplings in COSY and TOCSY spectra and are then connected to neighbouring amino acids across the peptide linkage via 'through-space' correlations in NOESY spectra. The latter are often obtained from a small region of the 2D NOESY spectrum, the 'fingerprint region'. This comprises amide proton frequencies in the directly detected dimension (F2) and α (and side-chain) proton frequencies in the indirectly detected dimension (F1) and contains both intra- and interresidue NH–αH connectivities, allowing sequential assignments to be made simply by 'walking' from peak to peak. The early stages of spin system identification are facilitated by the fact that many amino acids, either individually or in groups, give rise to characteristic peak patterns in COSY and TOCSY spectra and this feature provides the basis for automated assignment routines.

## Protein structure information from NMR parameters

Conformation-dependent data, primarily in the form of NOEs and scalar coupling constants, are obtained from essentially the same experiments (NOESY and COSY respectively) as the assignments. Additional structural information can also be extracted from chemical shifts, relaxation times and amide exchange data. It should be emphasized here that all of these NMR parameters are population-weighted and time averaged; thus, as molecular systems are inherently dynamic in nature, they do not in general represent single, precise values of interatomic distances and angles. Nevertheless these NMR parameters, which are considered below in terms of the structural information they represent, provide useful restraints to allow calculation of protein structures and, in the case of proteins that cannot be crystallized, access to structural information that cannot be obtained by other means.

### NOEs

A cross peak in a NOESY experiment indicates dipolar cross relaxation between two nuclei that are spatially close to one another. The cross peak intensity is dependent upon the inverse sixth power of the distance between the two nuclei. Thus, for two protons, $i$ and $j$ separated by a distance $r$, which give rise to a NOESY cross peak, the intensity of that cross peak $I$ is proportional to $r^{-6}$.

This simplified relationship assumes that the protein can be considered as a rotating rigid body in which correlation times for all proton pairs are the same, and are equal to the correlation time for the overall tumbling of the protein. For a more thorough analysis, account must be taken of, amongst other factors, internal mobility and spin diffusion. These are dealt with in greater detail in a separate article.

Interproton distances can, therefore, be determined from unambiguously assigned, well-resolved, high signal-to-noise NOESY data, by analysis of cross peak intensities. These may be obtained by volume integration and can be converted into estimates of interproton distances, using the equation above, for NOESY data at short mixing times. In this method, each proton pair is considered in isolation and NOESY cross peak intensities are compared with a reference cross peak from a proton pair of fixed distance, such as a geminal methylene proton pair or aromatic ring protons. A problem inherent to this approach is that the fixed distance is usually smaller than the unknown distance; this usually leads to systematic underestimation of the latter.

An alternative method of analysis of NOESY data, which is usually sufficient for resolved peaks with a digital resolution much greater than the intrinsic line width and coupling constants, is to measure the maximum peak amplitude or to count the number of contours. NOESY cross peaks can then be classified as strong, medium or weak and can be translated into upper distance restraints of around 2.5, 3.5 and 5.0 Å respectively. The lower distance constraint is usually the sum of the van der Waals radii (1.8 Å for protons). This simple approach is reasonably insensitive to the effects of spin diffusion or non-uniform correlation times and can usually lead to definition of the global fold of the protein, provided a sufficiently large number of NOEs have been identified.

Greater accuracy can be achieved by methods that involve calculation of a full relaxation matrix from the NOESY data to generate interproton distances. A model protein structure can then be iteratively refined by back calculation until differences in the empirical and calculated data are minimized. The resulting distances can be used as restraints for further refining the protein structure by distance geometry or molecular dynamics methods.

## Coupling constants

Geometric information, particularly for the bonds around the peptide backbone (**Figure 1**), can be obtained from vicinal spin–spin coupling constants. The magnitude of the coupling constant $J$ is dependent upon the dihedral angle $\theta$ as well as the nature and orientation of substituents in a manner that is defined by the Karplus relationship, which has the general form:

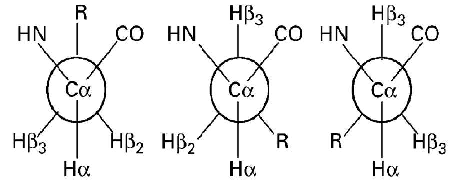
$$^3J = A\cos^2\theta - B\cos\theta + C$$

For a given coupling constant there are up to four valid solutions of the Karplus equation although knowledge of protein torsion angles (from protein structure databases) can be used to discount unlikely values. For example, the backbone torsion angle $\phi$ is usually negative, except in the case of asparagine, aspartate and glycine residues.

The $^3J_{NH,\alpha H}$ coupling constant, which is dependent upon the dihedral angle: [$HN_i$–$N_i$–$C\alpha_i$–$H\alpha_i$] ($\theta = \phi - 60°$ for L-amino acids), is commonly used for assessing secondary structure. Thus, a sequence of small (< 5 Hz) values indicates an α helix, whereas extended β structures have large (> 9 Hz) values that reflect the *trans* relationship of the NH and αH protons. Intermediate values are indicative of nonstandard structure or conformational averaging. The $^3J_{NH,\alpha H}$ values can be measured, in exceptional cases for small proteins, from high digital resolution 1D spectra but are more commonly obtained from 2D DQF-COSY spectra. If $^{15}N$-labelled protein is available they can be extracted from $^{15}N$-filtered correlation experiments and, in cases of signal overlap or insufficient digital resolution, from cross peak intensities in quantitative $J$-correlation experiments.

The $^3J_{\alpha H,\beta H}$ vicinal coupling constants can be determined from Exclusive COSY (E-COSY) type experiments and, together with intraresidue NOEs, can be used to obtain stereospecific assignments of β-methylene protons and side-chain $\chi$ angle restraints (**Table 2**).

**Table 2** NMR parameters that define conformations about the Cα–Cβ bond in amino acids




| *Parameter* | | | |
|---|---|---|---|
| Conformation | $g^-$ | $t$ | $g^+$ |
| $\chi_1$ | 60° | 180° | –60° |
| $d_{NH,H\beta 2}$ | 3.5–4.0 Å | 2.5–3.4 Å | 2.2–3.1 Å |
| NH–H$\beta_2$ NOE | Weak | Strong/medium | Strong |
| $d_{NH,H\beta 3}$ | 2.5–3.4 Å | 2.2–3.1 Å | 3.5–4.0 Å |
| NH–H$\beta_3$ NOE | Strong/medium | Strong | Weak |
| Hα–H$\beta_2$ NOE | Strong | Strong | Weak |
| Hα–H$\beta_3$ NOE | Strong | Weak | Strong |
| $^3J_{H\alpha,H\beta 2}$ | < 5 Hz | < 5 Hz | > 10 Hz |
| $^3J_{H\alpha,H\beta 3}$ | < 5 Hz | > 10 Hz | < 5 Hz |
| $^3J_{N,H\beta 2}$ | ~ 5 Hz | ~ 1 Hz | ~ 1 Hz |
| $^3J_{N,H\beta 3}$ | ~ 1 Hz | ~ 1 Hz | ~ 5 Hz |

## Chemical shifts

The chemical shift of a resonance reflects the chemical environment of the atom that gives rise to it. This is determined mostly by covalent bonding and to a smaller extent by the non-bonded environment. In unstructured peptides each amino acid exists in an ensemble of conformations and the random coil chemical shift represents the population-weighted mean value of these environments. The chemical shifts for amino acids in denatured proteins are close to random coil values; however, for structured proteins, many resonances are far removed from their random coil position (**Figure 2**). Even greater changes in proton chemical shift can result from the proximity of a proton to an aromatic ring (the ring current effect) or to a paramagnetic centre, as found in haem proteins such as haemoglobin and cytochromes.

The availability of chemical shift assignments for many proteins of defined structure has allowed changes in chemical shift from random coil values to be related to secondary structure and, consequently, to be used predictively. Thus, upfield shifts of ~0.3 ppm are characteristic for α protons in α helices, whereas α protons in β sheets experience downfield shifts of ~0.3 ppm. Similar effects are observed for $^{13}C$ chemical shifts of α carbons, which are shifted

downfield by ~3 ppm in α helices and upfield by ~1.5 ppm in β sheets. A sequence of similar changes can therefore help in the initial characterization of regions of secondary structure. Chemical shifts can also be used for refining protein tertiary structure; $^{13}C$ shifts, in particular, are sensitive to backbone geometry and can therefore help define backbone torsion angles.

### Relaxation times

The introduction of $^{15}N$ and/or $^{13}C$ labels into a protein facilitates the study of dynamic properties and, in particular, localized intramolecular motions. This arises because relaxation of these nuclei is usually dominated by dipole–dipole interactions with the directly bonded proton and this relaxation is dependent upon internuclear distance (which is fixed) and the rotational correlation time, which is only uniform throughout a 'rigid' protein. Proteins, however, usually contain regions that have greater flexibility, such as surface loops, which have different local correlation times that are reflected in heteronuclear relaxation times.

### Amide proton exchange and other exchange effects

Reduced rates of exchange of amide protons with the bulk solvent water indicate reduced solvent accessibility and potential involvement in hydrogen bonds. Almost all amide protons in regions of regular protein secondary structure (except for those near the edges) are hydrogen-bonded. A corollary of this is that fast amide exchange rates generally imply the absence of 'structure'. Measurement of amide proton exchange rates by following the time-course of the disappearance of signals in COSY, TOCSY or $^{1}H$–$^{15}N$ HSQC spectra therefore provides supportive evidence of secondary structure.

Other exchange phenomena that manifest themselves in NMR spectra include *cis–trans* isomerization of proline residues, aromatic ring 'flipping' and the rotation of primary amides of asparagine and glutamine.

## Deriving protein structures from NMR data

Extensive computational calculations are necessary to translate the information contained within NMR data into a protein structure. The quality of the structure obtained is dependent on the accuracy and, to a greater extent, quantity of the NMR data. As a rule of thumb, at least ten long-range NOE restraints

**Table 3** Characterization of protein secondary structure from NMR data

| *Parameter*[a] | α *Helix*[b] | β *Sheet*[c] |
|---|---|---|
| $\phi$ [C(O)$_{i-1}$–N$_i$–Cα$_i$–C(O)$_i$] | –57° | –139° |
| $^{3}J_{NH,\alpha H}$ | < 4 Hz | > 9 Hz |
| $d_{\alpha N}(i,i)$ (NOE intensity) | 2.6 Å (strong) | 2.8 Å (strong) |
| $d_{\alpha N}(i,i+1)$ (NOE intensity) | 3.5 Å (weak) | 2.2 Å (very strong) |
| $d_{\alpha N}(i,i+2)$ (NOE intensity) | 4.4 Å (weak) | – |
| $d_{\alpha N}(i,i+3)$ (NOE intensity) | 3.4 Å (medium) | – |
| $d_{\alpha\beta}(i,i+3)$ (NOE intensity) | 2.5–4.4 Å (medium) | – |
| $d_{\alpha N}(i,i+4)$ (NOE intensity) | 4.2 Å (weak) | – |
| $d_{NN}(i,i+1)$ (NOE intensity) | 2.8 Å (strong) | 4.3 Å (weak) |
| $d_{NN}(i,i+2)$ (NOE intensity) | 4.2 Å (weak) | – |
| $d_{\alpha\alpha}(i,j)$ (NOE intensity) | – | 2.3 Å (very strong) |
| $d_{\alpha N}(i,j)$ (NOE intensity) | – | 3.2 Å (medium) |
| $d_{NN}(i,j)$ (NOE intensity) | – | 3.3 Å (medium) |
| NH exchange rate | Slow | Slow |
| $^{1}$Hα Chemical shift change[d] | ~ –0.3 ppm | ~ +0.3 ppm |
| $^{13}$Cα Chemical shift change[d] | ~ +3 ppm | ~ –1.5 ppm |

[a] $d_{xy}(i, i+n)$ refers to the distance between proton $x$ in residue $i$ and proton $y$ in the residue $n$ positions from the C-terminus of residue $i$. In the case of the β sheet $j$ refers to the cross-strand partner.

[b] A $3_{10}$ helix differs from an α helix in that the NH of residue $i$ is hydrogen bonded to the carbonyl of residue $i$–4 in the α helix and of residue $i$– 3 in the $3_{10}$ helix. The consequences of this for differences in NMR data are small but are, most notably, a small decrease in $d_{\alpha N}(i, i+2)$, an increase in $d_{\alpha N}(i, i+4)$ to the point where a NOE is not observed and a small decrease in $^{3}J_{NH,\alpha H}$ for the $3_{10}$ helix.

[c] The values given above for a β sheet are for an antiparallel β sheet. Equivalent values for a parallel β sheet are essentially the same except for the interstrand distances, in particular $d_{\alpha\alpha}(i, j)$ is much larger (4.8 Å).

[d] $^{1}$Hα and $^{13}$Cα chemical shift changes are relative to random coil values.

are required for each amino acid residue to generate a 'reasonable' protein structure.

Regions of protein secondary structure are identified via characteristic short-range (i.e. ≤ 5 residues apart) NOEs, coupling constants, amide proton exchange and chemical shift data (**Table 3**). Longer-range NOEs then define the tertiary structure, which can be refined further against all available data.

Calculation of protein structures from NMR data requires conversion of the NMR data into distances and angles, usually in the form of allowed ranges, as

described above. These are incorporated as restraints into protein structure calculations, such that deviation from these values incurs an energetic penalty. In these calculations, distances and angles within a protein structure are optimized using a combination of distance geometry, molecular dynamics and simulated annealing procedures.

Distance geometry calculations aim to optimize all interatomic distances within a protein on the basis of the experimental restraints and are often used to define the global fold of a protein. This may then be refined using restrained molecular dynamics simulations in which structures evolve over time under the influence of a force field that contains potential energy terms for both covalent and non-bonded interactions and includes NMR restraints. A variation on molecular dynamics is simulated annealing, which differs in that normally 'prohibited' potential energy barriers can be crossed, allowing regions of a molecule to 'pass through' one another, thereby sampling larger regions of conformational space; this technique does not require a defined starting structure. After optimization of the NMR restraints by any or all of these techniques the potential energy of the resulting structure is minimized by molecular mechanics calculations to determine the lowest energy structure. The accuracy of the protein structure can be further increased by refinement against coupling constants, chemical shifts (both $^{1}$H and $^{13}$C), relaxation time rates ($T_1:T_2$) ratios and residual dipolar couplings.

Structural calculations usually result in an ensemble of protein structures that must be assessed to determine how well they satisfy the initial restraints. Violations in excess of 1 Å may indicate that regions of the structure are ill defined. NOESY spectra can be back calculated on the basis of the structures and compared with the experimental data to identify potential errors. Comparison of calculated and experimental NOEs can also lead to an '*R* factor' which gives an indication of the quality of the structures in a manner analogous to that used in X-ray crystallography. An alternative is to measure the RMS deviation of the ensemble of structures from the average structure. This measure should be used with care as it may indicate a high level of precision in the structures but not give a true indication of their accuracy.

## Other applications of protein NMR

### Protein dynamics

Static 3D protein structures cannot always explain biological processes or point the way, for example, to rational drug design; the dynamic properties of a protein may be of equal functional importance. Internal motions within proteins were recognized by early NMR experiments, and NMR spectroscopy has developed into a powerful technique for the study of dynamics, from the picosecond motions of bond vectors to millisecond motions. For labelled proteins, information about backbone dynamics and motions of nitrogen-containing side-chains can be obtained *via* $^{15}$N relaxation, which is dependent upon reorientation of the $^{15}$N–$^{1}$H bond vectors. A more complete description can result from the additional use of $^{13}$C relaxation data. Side-chain dynamics can be further characterized *via* deuterium relaxation by incorporation of a single deuterium atom into methyl or methylene groups and determining the attenuation of intensities in $^{1}$H–$^{13}$C-correlation experiments.

### Protein folding

A major advantage of NMR over X-ray crystallography is the ability to characterize the structure and dynamics of unfolded and partially folded states of proteins. These are of importance in protein folding and many cellular processes; indeed, many proteins or domains are intrinsically unstructured and only become structured upon binding other molecules. In these states, proteins exhibit rapid fluctuations between a range of conformations and insight into these processes can be gained from NMR, either through studying equilibrium states or through direct monitoring of kinetic folding events. Stabilized intermediates or fully denatured states at equilibrium can be characterized by essentially the same techniques as structured proteins, although chemical shift dispersion, with the exception of $^{15}$N and $^{13}$C-carbonyl resonances, is usually poor. The kinetics of protein folding can be monitored *via* the time-course of hydrogen–deuterium exchange in 2D experiments, by pulse labelling or by stopped-flow techniques. For $^{15}$N-labelled proteins, folding can also be monitored during the course of a single 2D-HSQC experiment by analysis of line shape changes.

## Concluding remarks

There are other areas of protein NMR that have not been considered here but are covered elsewhere in the Encyclopedia. These include the binding of ligands and the study of membrane-associated proteins in detergent micelles or phospholipid bilayers using solid state magic-angle spinning techniques.

The present article has focused on structural aspects of protein NMR. Determination of a de novo

protein structure by NMR takes longer, in general, than by X-ray crystallography – provided the protein can be crystallized. However, NMR is the only technique for obtaining high resolution structural data from proteins that cannot be crystallized and for which appropriate concentrations of non-aggregated protein can be achieved. Limitations of molecular size can be overcome, as with other biochemical and structural techniques, by considering protein fragments, or domains. Structural data for large proteins can thus be obtained for domains individually and the linkage and assembly of domains can then be determined. The extension of protein NMR to larger molecules may be further facilitated by measurement of residual anisotropic interactions, which give structural restraints that are orientational rather than distance based.

To summarize, NMR spectroscopy can provide a wealth of structural information about proteins and, leading from this, a greater insight into protein interactions and dynamic processes.

## List of symbols

$I$ = cross peak intensity; $J$ = coupling constant; $r$ = distance between two protons; $T_1$, $T_2$ = relaxation times; $\phi$ = dihedral angle defined by $H_i$–$N_i$–$C\alpha_i$–$H\alpha_i$; $\psi$ = dihedral angle defined by $H\alpha_i$–$C\alpha_i$–C = 0; $\chi_1$ = dihedral angle defined by $H\alpha_i$–$C\alpha_i$–Cβ–R.

*See also:* **Labelling Studies in Biochemistry Using NMR; Laboratory Information Management Systems (LIMS); Macromolecule–ligand Interactions Studied By NMR; Magnetic Field Gradients in High Resolution NMR; NMR Data Processing; NMR Pulse Sequences; NMR Relaxation Rates; Nuclear Overhauser Effect; Parameters in NMR Spectroscopy, Theory of; Solvent Suppression Methods in NMR Spectroscopy; Structural Chemistry Using NMR Spectroscopy, Peptides; Two-Dimensional NMR, Methods.**

## Further reading

Cavanagh J, Fairbrother WJ, Palmer AG and Skelton NJ (1996) *Protein NMR Spectroscopy: Principles and Practice*, 587 pp. San Diego: Academic Press.

James TL and Oppenheimer NJ (eds) (1989) *Methods in Enzymology,* Vol 176 (Nuclear Magnetic Resonance, Part A: Spectral Techniques and Dynamics) 530 pp and Vol 177 (Nuclear Magnetic Resonance, Part B: Structure and Mechanism) 507 pp. San Diego: Academic Press.

James TL and Oppenheimer NJ (eds) (1994) *Methods in Enzymology*, Vol 239 (Nuclear Magnetic Resonance, Part C) 813 pp. San Diego: Academic Press.

Leach AR (1996) *Molecular Modelling – Principles and Applications*, 595 pp. Harlow, UK: Longman.

*Nature Structural Biology* (NMR I Supplement) (1997) Vol 4, 841–866.

*Nature Structural Biology* (NMR II Supplement) (1998) Vol 5, 492–522.

Reid DG (ed) (1997) *Methods in Molecular Biology: Protein NMR Techniques,* 419 pp. New Jersey: Humana Press.

Roberts GCK (ed) (1993) *NMR of Macromolecules – A Practical Approach,* 399 pp. Oxford: Oxford University Press.

Wüthrich K (1986) *NMR of Proteins and Nucleic Acids*, 292 pp. New York: John Wiley & Sons.

# Proton Affinities

**Edward PL Hunter** and **Sharon G Lias**, National Institute of Standards and Technology, Gaithersburg, MD, USA

**MASS SPECTROMETRY**
**Theory**

## What is a proton affinity?

The 'proton affinity' and the related quantity 'gas-phase basicity', are defined thermodynamic quantities that enable us to assign numeric values to the tendency of a molecule to accept a proton in the gas phase, or conversely, the tendency of a positive ion to donate a proton. That is 'proton affinities' and 'gas-phase basicities' provide quantitative measures of the acid–base properties of positive ions and the corresponding conjugate-base neutral species in the gas phase. These quantities have proved to be of

scientific interest because of what they tell us about acid–base chemistry in the absence of a solvent, and of practical interest because of the applications of gas-phase proton transfer reactions in the analytical technique of chemical ionization mass spectrometry. This article will be concerned with defining these quantities, describing how they are determined experimentally, and discussing the current status of the collective database of proton affinity/gas-phase basicity data. For a discussion of the implications of those data in organic chemistry, the reader is referred to the available reviews of that subject, particularly to the work of R. W. Taft.

The proton affinity of species M, $PA_T(M)$, is defined as the negative of the enthalpy change, $\Delta H_T^0$, of the hypothetical gas-phase reaction at temperature $T$:

$$M(g) + H^+(g) \rightarrow MH^+(g) \quad [1]$$

$$\begin{aligned} PA_T(M) &\equiv -\Delta H_T^0 \text{ (Eqn [1])} \\ &= \Delta_f H_T^0(M) + \Delta_f H_T^0(H^+) - \Delta_f H_T^0(MH^+) \end{aligned} \quad [2]$$

while the gas-phase basicity, $GB_T(M)$, is the negative of the Gibbs free energy change of the same reaction, $\Delta G_T^0$, and therefore related to the proton affinity through the standard thermochemical expression:

$$\begin{aligned} GB_T(M) &\equiv -\Delta G_T^0 \text{ (Eqn [1])} \\ &= PA_T(M) + T\Delta S_T^0 \text{ (Eqn [1])} \end{aligned} \quad [3]$$

The entropy change of Equation [1], $\Delta S_T^0$, can be expressed in terms of absolute entropies of the species involved:

$$\begin{aligned} \Delta S_T^0 &\text{ (Eqn [1])} \\ &= S_T^0(MH^+) - S_T^0(M) - S_T^0(H^+) \quad [4] \\ &= \Delta S_{p,T}(M) - S_T^0(H^+) \quad [5] \end{aligned}$$

where $\Delta S_{p,T}(M)$ is the entropy of protonation of M:

$$\Delta S_{p,T}(M) = S_T^0(MH^+) - S_T^0(M) \quad [6]$$

Thus, the relationship between the gas-phase basicity, proton affinity and entropy of protonation is given by

$$GB_T(M) = PA_T(M) + T[\Delta S_{p,T}^0 (M) - S_T^0(H^+)] \quad [7]$$

In practice, proton affinity and gas-phase basicity values are used to predict the occurrence of bimolecular proton transfer reactions in the gas phase:

$$MH^+(g) + B(g) \rightarrow BH^+(g) + M(g) \quad [8]$$

(In the text that follows, all references to compounds in reactions or in thermochemical equations refer to species in the gas phase, although a specific designator, (g), is omitted in the interest of improving readability.) At any particular temperature, the species, $MH^+$ or $BH^+$, with the lower gas-phase basicity will transfer a proton to the conjugate base of the other (B or M), and the exact difference in the gas-phase basicities can be measured by determining the equilibrium constant, $K_{eq}$, of Equation [8].

$$K_{eq} = \frac{[BH^+][M]}{[MH^+][B]} \quad [9]$$

$$\begin{aligned} -RT \ln K_{eq} &= \Delta G_T^0 \text{ (Eqn [8])} \\ &= \Delta H_T^0 \text{ (Eqn [8])} \\ &\quad - T\,\Delta S_T^0 \text{ (Eqn [8])} \\ &= GB_T(M) - GB_T(B) \end{aligned} \quad [10]$$

Similarly, the enthaply change of Equation [8] is equal to the difference in the proton affinities of the two species,

$$\Delta H_T^0 \text{ (Eqn [8])} = PA_T(M) - PA_T(B) \quad [11]$$

and the entropy change is

$$\Delta S_T^0 \text{ (Eqn [8])} = \Delta S_{p,T}^0(B) - \Delta S_{p,T}^0(M) \quad [12]$$

Although the occurrence or nonoccurrence of the proton transfer depends, at any given temperature, on the relative gas-phase basicity values, rather than the proton affinities, entropy changes associated with Equation [8] are typically (although not always) small, and proton affinities are, in practice, often consulted to predict the occurrence or nonoccurrence of particular proton transfer reactions.

According to common usage, unless it is stated otherwise proton affinity values are assumed to refer to 298 K, unlike electron affinities and ionization energies, which are specifically referred to 0 K. Therefore, it is useful to question how absolute values of

proton affinities and protonation entropies vary with temperature. Differentiating and integrating Equation [2] with respect to temperature gives:

$$\mathrm{PA}_{T_2}(\mathrm{M}) - \mathrm{PA}_{T_1}(\mathrm{M}) = \int [C_\mathrm{p}(\mathrm{H}^+) + C_\mathrm{p}(\mathrm{M}) - C_\mathrm{p}(\mathrm{MH}^+)]\mathrm{d}T \qquad [13]$$

where the $C_p$ are the molar heat capacities at constant pressure of the parenthetically indicated species, and the integration is carried out from $T_1$ to $T_2$. At room temperature and above, $C_p(H^+)$ is taken as the classical value of $(5/2)R$, while $C_p(MH^+)$ will be close to, but greater than, $C_p(M)$. Thus, for example, the difference in the value of the absolute proton affinity of M at 298 K and 600 K will be less than $(5/2)\ R\ (600\ \mathrm{K}–298\ \mathrm{K}) = 6.2\ \mathrm{kJ\ mol^{-1}}$.

The relative proton affinities, $PA_T(M) - PA_T(B)$, of a pair of molecules, M and B in Equation [8], are essentially temperature independent, i.e.

$$-\Delta H^0_{T_1}\ (\mathrm{Eqn}\ [8]) = \mathrm{PA}_{T_1}(\mathrm{M}) - \mathrm{PA}_{T_1}(\mathrm{B}) \approx -\Delta H^0_{T_2}\ (\mathrm{Eqn}\ [8]) = \mathrm{PA}_{T_2}(\mathrm{M}) - \mathrm{PA}_{T_2}(\mathrm{B}) \qquad [14]$$

This can be shown more formally by differentiating and integrating Equation [11] with respect to temperature,

$$[\mathrm{PA}_{T_2}(\mathrm{M}) - \mathrm{PA}_{T_2}(\mathrm{B})] - [\mathrm{PA}_{T_1}(\mathrm{M}) - \mathrm{PA}_{T_1}(\mathrm{B})] = \int [C_\mathrm{p}(\mathrm{BH}^+) + C_\mathrm{p}(\mathrm{M}) - C_\mathrm{p}(\mathrm{MH}^+) - C_\mathrm{p}(\mathrm{B})]\mathrm{d}T \qquad [15]$$

and noting that, because of the structural similarities of reactants and products, the heat capacity terms of Equation [15] will essentially cancel. When a relative proton affinity is derived from a van't Hoff analysis of a proton transfer equilibrium over a suitable temperature range, it is safe to assume that $\Delta H^0_T$ (Eqn [8]) is independent of temperature over that range. The above discussion suggests that the temperature independence of $\Delta H^0_T$ (Eqn [8]) can be safely assumed throughout the range $298\ \mathrm{K} \leq T \leq 600\ \mathrm{K}$. The feature is a generally observed phenomenon for reactions in which the number of reactants and products is the same, as is the case for proton transfer reactions. Similar considerations also apply to relative protonation entropies, i.e.

$$\Delta S^0_{T_1}\ (\mathrm{Eqn}\ [8]) = \Delta S_{\mathrm{p},T_1}(\mathrm{M}) - \Delta S_{\mathrm{p},T_1}(\mathrm{B}) \approx \Delta S^0_{T_2}\ (\mathrm{Eqn}\ [8]) = \Delta S_{T_2}(\mathrm{M}) - \Delta S_{T_2}(\mathrm{B}) \qquad [16]$$

These considerations are important in establishing the validity of comparing data obtained at different temperatures, as must be done in evaluating gas-phase basicity and proton affinity data obtained under different experimental conditions.

## How are proton affinities and gas phase basicities determined?

The proton affinity of any species can easily be derived from Equation [2], provided that enthalpies of formation of all the relevant species, M, $H^+$ and $MH^+$, are known. Actually, in practice, there are relatively few species for which this condition is fulfilled. Usually, reliable data on the enthalpy of formation of $MH^+$ are lacking, since direct determination of this quantity requires either that the neutral precursor, MH, exists and is sufficiently stable that its ionization energy can be experimentally determined, or that the appearance energy of $MH^+$ formed in the dissociation of a larger molecule, MNH, can be determined:

$$\mathrm{MNH} \rightarrow \mathrm{MH}^+ + \mathrm{N} + \mathrm{e}^- \qquad [17]$$

Obviously, neutral analogues (MH) exist for only a very few protonated molecular species ($MH^+$). For example, although $CH_5^+$ (protonated methane) is a commonly used proton-donor species, the proton affinity of methane cannot be derived using this approach (Eqn [2]), since the molecule $CH_5$ does not exist as a stable species, nor is $CH_5^+$ formed as a fragment ion from any known dissociation of an organic molecular ion.

Most of the existing quantitative data on proton affinities and gas-phase basicities have been derived from determinations of equilibrium constants for proton transfer reactions (Eqn [8]) in the gas phase (Eqns [9] and [10]). Note that the measurement of a series of equilibrium constants at a single temperature will generate relative, rather than absolute, values for gas-phase basicities. In order to determine the enthalpy change of Equation [8] – that is, the relative

proton affinities of M and B (see Eqn [11]) – values for the entropy changes of the reaction must be obtained, through measurements of the equilibrium constant as a function of temperature (van't Hoff plot), through statistical-mechanical estimations or by *ab initio* calculations. Absolute values must then be assigned to the relative proton affinity scales using data for molecules whose position in the scale has been established, and for which absolute values of enthalpies of formation of both M and $MH^+$ are known from other measurements (Eqn [2]).

Equilibrium constants for gas-phase proton transfer reactions can be determined as follows. A mixture, of known composition, of gases M and B is introduced into a mass spectrometric instrument designed to allow multiple collisions of ions before sampling (that is, designed so that ion–molecule reactions can be observed). Most such measurements have been carried out using one of three types of mass spectrometer that operate in very different pressure regimes. An ion cyclotron resonance (ICR) spectrometer typically observes such equilibria at very low pressures (~$10^{-4}$ Pa), but at long times, so that there have been a sufficient number of collisions that an equilibrium can be established. Other measurements are done at higher pressures (100–1000 Pa) using a high-pressure mass spectrometer or a flow tube apparatus such as a flowing afterglow instrument.

If one or both of the parent molecular ions, $M^{\bullet+}$ or $B^{\bullet+}$, is a species that contains a labile proton that can be transferred to M or B, then Equation [8] may be observed in the system. When neither $M^{\bullet+}$ nor $B^{\bullet+}$ serves as a source of protons, the proton transfer reaction (Eqn [8]) may be initiated by the addition of a bath gas of some compound (methane, for example) that generates ions that transfer protons to the two subject compounds. If the reverse (endothermic) proton transfer reaction is fast enough to be observed on the timescale of the particular experiment, it is possible that a thermodynamic equilibrium can be established, and the equilibrium constant determined simply by observing the equilibrium ratio of the ion concentrations, $[MH^+]$ and $[BH^+]$; the composition of the mixture [M]/[B] does not change, since the neutral compounds are present in great abundance relative to the ions. Each measurement provides a value for the Gibbs free energy change of Equation [8] at a single temperature – that is, a value for the difference in gas basicities of molecules M and B (see Eqn [10]). There exist large interlocking scales of relative gas phase basicities (at a single temperature) determined by carrying out such measurements on series of molecules.

In certain cases, it is not possible to establish a proton transfer equilibrium, for example, if the subject compound, M, is unstable, or if $MH^+$ undergoes a fast reaction with M, or a reaction other than proton transfer with B. In these instances, other strategies have been adopted to determine relative gas-phase basicities or proton affinities. The simplest such approach is called the 'bracketing technique'; the ion $MH^+$ is generated and the occurrence or nonoccurrence of proton transfer with a series of molecules, $B_1$, $B_2$, etc., is observed. Reference compounds are chosen whose position in the relative scale of gas basicities is known.

$$MH^+ + B_1 \rightarrow \text{no proton transfer} \qquad [18]$$

$$MH^+ + B_2 \rightarrow B_2H^+ + M \qquad [19]$$

Under the assumption that proton transfer will be observed only if the reaction is associated with a negative value of the Gibbs free energy, the basicity of M is taken to be between the basicities of $B_1$ and $B_2$. In a variation of this approach proposed by Bouchoux, Salpin and Leblanc, called the 'thermokinetic method', trends in the reaction efficiency ($k_{Rn}$/$k_{Collision}$) with reactants of varying gas basicity are examined, and a correlation is used to predict the relative gas basicity.

Another approach, proposed by McLuckey, Cameron and Cooks, that is often used when compounds M and B are of low volatility is based on the observation of the collision-induced dissociation of proton-bound dimer ions, $M{\cdot}H^+{\cdot}B$, formed in association reactions:

$$M \cdot H^+ \cdot B + X \rightarrow MH^+ + B + X \qquad [20a]$$

$$\rightarrow BH^+ + M + X \qquad [20b]$$

A semiquantitative relationship between the ratios of the two product ions and the relative proton affinities has been developed, and can be used to derive relative values of the proton affinities provided that the entropy changes associated with processes [20a] and [20b] are similar.

Quantitative information about relative proton affinities can also be obtained through the determination of the energy barrier associated with endothermic proton transfer reactions through an Arrhenius treatment of the temperature dependence of the rate coefficients, although this approach has

rarely been used. In addition, determinations of the equilibrium constants of association reactions

$$AH^+ + B \rightarrow ABH^+ \qquad [21]$$

can provide values for enthalpies of formation of the product ion, $ABH^+$, provided the enthalpies of formation of $AH^+$ and B are known; if the enthalpy of formation of AB is also known, its proton affinity can be derived.

## The 1998 scale of gas-phase basicities and proton affinities

Beginning in 1971, when the first determinations of gas-phase proton transfer equilibria were published, several extensive scales of relative gas-phase basicities were generated in different laboratories. With the exception of a small number of entropy-change measurements made in the laboratory of Paul Kebarle at the University of Alberta, most of these initial studies were carried out at a single temperature; in most cases, entropy changes were estimated from statistical-mechanical considerations, usually making the simplifying assumption that the complete expression derived from the partition functions could be adequately approximated using the ratio of the rotational symmetry numbers ($\sigma$) of M and $MH^+$:

$$\Delta S^0_p(M) = R\ln[\sigma(M)/\sigma(MH^+)] \qquad [22]$$

Using such estimated entropy changes, the experimentally determined interlocking scales of relative basicities were converted to scales of proton affinities. By the early 1980s, values of proton affinities for about 800 molecules had been reported, but comparing data from different laboratories was not always straightforward because different researchers chose different primary standards to assign absolute values to the proton affinities, and proton affinity values assigned to these reference standards were not always internally consistent. Therefore, in 1984 the Ion Energetics Data Center at the National Bureau of Standards (now the National Institute of Standards and Technology) carried out a comprehensive evaluation of available data to put all data on the same basis, and provide an internally consistent scale.

That evaluated scale (sometimes referred to as 'the NBS scale') proved to be sufficiently useful that it continues to be cited years after its publication. However, in the intervening years, a large amount of new data has appeared in the literature, so the so-called 'NBS scale' (which will be referred to here as the '1984 scale') is seriously out of date, missing information on about 900 compounds. In addition, since 1991 several important publications (both experimental and theoretical) have presented data indicating that portions of the thermochemical scale, as evaluated in 1984, are in need of re-evaluation. Furthermore, recent publications from two laboratories (Mautner and Sieck at the National Institute of Standards and Technology, and Szulejko and McMahon at the University of Waterloo) provided extensive data sets in which equilibrium constants had been determined as a function of temperature, i.e. experimental entropy change determinations had become available.

The results of these recent experimental proton transfer equilibrium studies indicated that portions of the 1984 gas-phase basicity/proton affinity scale were constricted. For example, in the 1984 scale the difference between the proton affinities of isobutene and ammonia was given as 33.5 kJ mol$^{-1}$, while according to the newer results, this interval is actually about 50 kJ mol$^{-1}$. Also, these experimental results indicated that the portion of the 1984 scale representing gas-phase basicities/proton affinities higher than that of ammonia (a portion of the scale that was not 'anchored' by data for a primary standard) was seriously constricted.

Furthermore, according to the 1993 theoretical calculation carried out by Smith and Radom of the proton affinity of isobutene (one of the primary standards for the 1984 scale),

$$(CH_3)_2C = CH_2 + H^+ \rightarrow (CH_3)_3C^+ \qquad [23]$$

the value for the proton affinity of isobutene used in the 1984 evaluation was too high by about 16 kJ mol$^{-1}$; this implied that the previously accepted enthalpy of formation of the $(CH_3)_3C^+$ ion must be in error by this amount. New experimental determinations of that enthalpy of formation carried out by Baer and colleagues and by Traeger soon gave evidence that the theoretical prediction was correct. These experimental and theoretical results corroborated the new equilibrium constant data that indicated that the interval between the proton affinities of isobutene and ammonia was greater by 17 kJ mol$^{-1}$ than previously accepted. All these results had profound implications for the evaluation of the central portion of the thermochemical scales, which had been 'anchored' by taking the proton affinity of isobutene as reference standard.

For these reasons, the present authors undertook a re-evaluation of the entire corpus of data on gas-phase

basicities and proton affinities, which now includes data for about 1740 compounds. This effort involved an evaluation of the entire interrelated thermochemical scale. Users of proton affinity data are sometimes confused when a re-evaluation results in a change in the value assigned to a particular molecule, and inquire whether a particular proton affinity value has been re-determined; it is important to understand that, with the exception of the few standards used to 'anchor' the scales, a compound-by-compound evaluation of the scale of gas-phase basicities (or proton affinities) is not possible, and when the entire scale is expanded, contracted or shifted (owing to the appearance of newer, more reliable data), values assigned to individual compounds will change, even when the original experimental determinations (of relative gas-phase basicities/proton affinities) remain the only source of information about the proton affinity of the compound in question. Although users of the data attach importance to absolute values assigned to particular proton affinity values, it is usually the relative values that are of importance in designing experiments involving proton transfer reactions; it is of primary importance to ascertain that the values being used, be they relative or absolute, are internally consistent.

The evaluation of such a body of interrelated thermodynamic data involves first an evaluation of the thermochemical scales for internal consistency in the three parameters, $\Delta G^0$ (at different temperatures), $\Delta H^0$ and $\Delta S^0$. Final values assigned for the proton affinities and entropy changes must be consistent with what is known about the thermochemistry of M and $MH^+$. The lengths of segments of the scale linking different primary standards (compounds for which Eqn [2] can be used to derive an absolute proton affinity value) must of course match the known interval between the 'known' proton affinity values.

The 1998 evaluation of these data took as a starting point an examination of several extensive scales of gas-phase basicities obtained in different laboratories and using different kinds of instrumentation. The existence of such extensive independent data sets, each of which contained numerous checks on internal consistency, was invaluable in establishing confidence in the final evaluated scale. The gas-phase basicity scales were chosen as the starting point for the evaluation, rather than the newly published experimentally determined proton affinity scales (based on van't Hoff plots determined in high-pressure mass spectrometers over the temperature range ~450–650 K), because there was poor agreement between the relative proton affinity scales determined in the different studies. (There are several plausible explanations for the lack of agreement between these data sets determined in different laboratories, including pyrolysis of ions in the high-temperature region (> 600 K), clustering of neutral molecules to the ions in the low-temperature region (< 500 K), or poor characterization of the relative and absolute pressures of the reacting compounds in the high-pressure mass spectrometer ion sources, as demonstrated by Grimsrud and colleagues.)

These internally consistent scales of gas-phase basicity data on which the evaluation was primarily based included three data sets obtained at 600 K using high-pressure mass spectrometry, and a very extensive scale determined several years ago at 'ambient' temperature using ion cyclotron resonance spectrometry. These independent determinations of the extensive thermochemical ladder were, with a few minor exceptions, in excellent agreement over the central and upper parts of the energy range, except that the older data yielded scales that appeared to be slightly 'contracted' relative to the recent data, most likely as a result of problems in temperature measurements in the early experiments. A careful evaluation, involving comparisons with analogous experiments where the temperature was carefully measured, permitted a 'standardization' of the various data sets by making appropriate corrections for temperature.

From the evaluated composite gas-phase basicity scale, there was established a 'backbone' scale based on data for some 25 compounds, which included selected primary standards as well as a few other compounds for which entropy changes could be reliably assigned. Primary standards are compounds for which an absolute proton affinity can be established using known enthalpies of formation (Eqn [2]). Through the use of this 'backbone' scale, the established scale of relative gas-phase basicities could be tied to absolute proton affinity values. With the 'backbone' scale established, all other published gas basicity data (for some 1700 additional compounds) were related to that scale at appropriate temperatures. Then values for the entropy of protonation (Eqn [6]) were assigned, compound-by-compound, to generate the final complete scale of proton affinity values. Finally, the resulting data were examined to verify internal consistency and 'reasonableness' of all the proton affinity and entropy-change values.

In carrying out the evaluation, the scale of proton affinity values produced by Smith and Radom through *ab initio* calculations was an invaluable aid. That scale, of proton affinity values for 31 molecules covering an energy range of about 500 kJ mol$^{-1}$, effectively spanned most of the experimental scale reported from equilibrium constant determinations, and provided an independent check on the

**Table 1** 'Backbone' scale of proton affinities and gas-phase basicities

| *Molecule* | *1998 scale* | | | *Ab initio* | *1984 scale* | |
|---|---|---|---|---|---|---|
| | $GB_{298}$ | $PA_{298}$ | $\Delta S_{p,298}$ | $PA_{298}$ | $GB_{298}$ | $PA_{298}$ |
| $(CH_3)_3N$ | 918.1±6.0 | 948.9±6.0 | 5.6 | 951.6[a] | 909 | 942 |
| Pyridine | 898.1±6.0 | 930±6.0 | 2.0 | 929.8[a] | 892 | 924 |
| $(CH_3)_2NH$ | 896.5±6.0 | 929.5±6.0 | −2.0 | 931.7[a] | 890 | 923 |
| $C_2H_5NH_2$ | 878±6.0 | 912±6.0 | −5.1 | 941.0[a] | 872 | 908 |
| $CH_3NH_2$ | 864.5±6.0 | 899±6.0 | −7.0 | 901.0[a] | 861 | 896 |
| $NH_3$ | *819* | *853.6* | −6.4 | *853.6*[a] | *818* | *853.5* |
| $CH_2{=}C{=}O$[c] | *793.6±3.0* | *825.3±3.0* | *2.4* | *825.0*[a] | *793* | *828* |
| $CH_3COCH_3$ | 782.1±6.0 | 812±6.0 | 8.7 | 811.9[a] | 790 | 823 |
| $(CH_3)_2C{=}CH_2$[d] | *775.6±1.2* | *802.1±1.4* | *20* | *802.1*[a] | *784* | *820* |
| $(CH_3)_2O$ | 764.2±6.0 | 792±6.0 | 16.5 | 792.0[a] | 771 | 804 |
| $C_2H_5CN$ | 763±6.0 | 794.1±6.0 | 4.7 | 794.3[a] | 770 | 806 |
| $C_6H_5CH_3$ | 756.3±6.0 | 784±6.0 | 16 | – | 761 | 794 |
| $CH_2{=}CHCN$ | 753.7±6.0 | 784.7±6.0 | 4.9 | 784.7[a] | 761 | 794 |
| $HCOOCH_3$ | 751.5±6.0 | 782.5±6.0 | 5 | 782.2[a] | 757 | 790 |
| $CH_3CN$ | 748±6.0 | 779.2±6 | 4.3 | 780.0[a] | 756 | 788 |
| $CH_3CHO$[e] | *736.5±1.6* | *768.5±1.6* | *1.5* | *770.2*[a] | *747* | *781* |
| $CH_3OH$ | 724.5±6.0 | 754.3±6.0 | 9 | 754.3[a] | 728 | 761 |
| $CH_3CH{=}CH_2$[f] | *722.7±3.0* | *751.6±3.0* | *12* | *744.3*[a] | *718* | *751* |
| $CH_2{=}O$[g] | *683.3±1.1* | *712.9±1.1* | *9.5* | *711.8*[a] | *687* | *718* |
| $H_2S$[h] | *673.8±5.3* | *705.0±5.3* | *4.3* | *707.7*[a] | *681* | *712* |
| $H_2O$[i] | *660.0±3.0* | *691.0±3.0* | *5* | *688.4*[a] | *665* | *697* |
| $CS_2$ | 657.7±6.0 | 681.9±6.0 | 28 | 681.9[a] | 672 | 699 |
| $CH_2{=}CH_2$[j] | *651.5* | *680.5±1.7* | *11.5* | *681.9*[a] | *651* | *680* |
| CO[k] | *562.8±3.0* | *594.0±3.0* | *4.2* | *539.0*[a], *593.1*[b] | *562* | *594* |
| $CO_2$[l] | *515.8±2.0* | *540±2.0* | *26* | *539.3*[a], *541.0*[b] | *520* | *548* |

[a] Smith and Radom (1993).
[b] Kormornicki and Dixon (1992).
[c] $\Delta_f H(CH_3CO^+)$ from appearance energies in methyl ketones.
[d] $\Delta_f H((CH_3)_3C^+)$ from appearance energy determinations (Keister *et al* 1993).
[e] $\Delta_f H(CH_3CHOH^+)$ from appearance energy in $C_2H_5OH$.
[f] $\Delta_f H(CH_3)_2CH^+)$ from appearance energies in 2-halopropanes.
[g] $\Delta_f H(CH_2OH^+)$ from appearance energy in $CH_3OH$.
[h] $\Delta_f H(H_3S^+)$ from appearance energy in Van der Waals dimer $(H_2S)_2$.
[i] $\Delta_f H(H_3O^+)$ from appearance energy in Van der Waals dimer $(H_2O)_2$.
[j] $\Delta_f H(C_2H_5^+)$ from adiabatic ionization energy of the ethyl radical, and from appearance energy in $C_2H_5I$.
[k] $\Delta_f H(HCO^+)$ from appearance energy in HCOOH.
[l] $\Delta_f H(CO_2H^+)$ from appearance energy in HCOOH and other carboxylic acids.

evaluation. Theoretical proton affinity values from this study were available for the 'backbone' scale compounds, and served to verify the evaluations.

**Table 1** shows the results of the 1998 evaluation of the gas-phase basicity and proton affinity scales as exemplified by the 'backbone' scale. The results shown include the evaluated proton affinity, gas-phase basicity and entropy of protonation values, the *ab initio* value reported by Smith and Radom and, for comparison, the proton affinity value cited in the 1984 evaluation. Primary standards that anchor the scale are given in italic and the footnotes give an indication of the type of experiment that yielded a value for the enthalpy of formation of the protonated molecule, $MH^+$ (see Eqn [2]) used in establishing the absolute proton affinity value. Note that the values adopted for ammonia are essentially the same as those recommended in the 1984 evaluation, and that the portion of the scale above ammonia is expanded by about 8–9% compared to the

values assigned in that evaluation. The dramatic change in the proton affinity and gas-phase basicity of isobutene results in a general lowering of values assigned to compounds below isobutene in the scale. Below ethylene in the scale, the changes are less dramatic. It should be mentioned that the scale of gas basicities is not yet well established in the low-basicity region (below the basicities of $H_2S$ and $H_2O$), where there are major inconsistencies in the data reported from different laboratories.

In **Table 1**, the standard uncertainties assigned to the primary anchor molecules (italic entries) are the usual root-sum-of-squares combination of individual uncertainties associated with relevant enthalpies of formation and the uncertainty of some key measurement, such as an ionization or an appearance energy. The uncertainties assigned to all the other molecules are based on our best judgment using all the relevant information and a general knowledge of and experience with interlocking thermochemical scales.

## List of symbols

$BG_T(M)$ = gas-phase basicity of species M at temperature $T$; $C_p$ = molar heat capacity at constant pressure; $k$ = reaction rate constant; $K_{eq}$ = equilibrium constant; $PA_T(M)$ = proton affinity of species M at temperature $T$; $S_T^0$ (M) = absolute entropy of species M at temperature $T$; $T$ = temperature; $\Delta G_T^0(\text{Eqn}[n])$ = Gibbs free energy change of reaction (given as Eqn [$n$]) at temperature $T$; $\Delta H_T^0(\text{Eqn}[n])$ = enthalpy change of reaction (given as Eqn [$n$]) at temperature $T$; $\Delta_f H^0(M)$ = enthalpy of formation of species M at temperature $T$; $\Delta S_T^0(\text{Eqn}[n])$ = entropy change of reaction (given as Eqn [$n$]) at temperature $T$; $\Delta S_{p,T}(M)$ = entropy of protonation of species M at temperature $T$; $\sigma$ = rotational symmetry number.

*See also:* **Chemical Ionization in Mass Spectrometry; Ion Dissociation Kinetics, Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry.**

## Further reading

Bouchoux G, Salpin J-Y and Leblanc D (1996) A relationship between the kinetics and thermochemistry of proton transfer reactions in the gas phase. *International Journal of Mass Spectrometry and Ion Processes* **153**: 37–48.

East ALL, Smith BJ and Radom L ( 1997) Entropies and free energies of protonation and proton-transfer reactions. *Journal of the American Chemical Society* **119**: 9014–9020.

Hunter EP and Lias SG (1997) Proton affinity evaluation. In: Mallard WG and Linstrom PJ (eds) *NIST Standard Reference Database Number 69*. Gaithersburg, MD: National Institute of Standards and Technology (http://webbook.nist.gov).

Hunter EP and Lias SG (1998) Evaluated gas phase basicities and proton affinities of molecules: an update. *Journal of Physical Chemistry Reference Data* **27**: 413–656.

Kebarle P (1997) Ion thermochemistry and solvation from gas phase ion equilibria. *Annual Review of Physical Chemistry* **28**: 445–476.

Kebarle P, Yamdagni R, Hiroaka JK and McMahon TB (1976) Ion molecule reactions at high pressure: recent proton affinities, gas phase acidities and hydrocarbon clustering results. *International Journal of Mass Spectrometry and Ion Physics* **19**: 71–87.

Keister JW, Riley JS and Baer T (1993) The *tert*-butyl ion heat of formation and the isobutene proton affinity. *Journal of the American Chemical Society* **115**: 12 613–12 614.

Kormornicki A and Dixon DA (1992) Accurate proton affinities: *ab initio* proton binding energies for $N_2$, CO, $CO_2$, and $CH_4$. *Journal of Chemical Physics* **97**: 1087–1094.

Lau YK (1979) PhD thesis, University of Alberta.

Lias SG, Liebman JF and Levin RD (1984) Evaluated gas phase basicities and proton affinities of molecules; heats of formation of protonated molecules. *Journal of Physical Chemistry Reference Data* **13**: 695–808.

McGrew DS, Knighton WB, Bognar JA and Grimsrud EP (1994) Concentration enrichment in the ion source of a pulsed electron beam high pressure mass spectrometer. *International Journal of Mass Spectrometry and Ion Physics* **139**: 47–58.

McLuckey SA, Cameron D and Cooks RG (1981) Proton affinities from dissociations of proton bound dimers. *Journal of the American Chemical Society* **103**: 1313–1317.

Meot-Ner (Mautner) M and Sieck LW (1991) Proton affinity ladders from variable temperature equilibrium measurements. 1. A re-evaluation of the upper proton affinity range. *Journal of the American Chemical Society* **113**: 4448–4460.

Smith BJ and Radom L (1993) Assigning absolute values to proton affinities: a differentiation between competing scales. *Journal of the American Chemical Society* **115**: 4885–4888.

Szulejko J and McMahon TB (1991) A pulsed electron beam, variable temperature high pressure mass spectrometric re-evaluation of the proton affinity difference between 2-methylpropene and ammonia. *International Journal of Mass Spectrometry and Ion Processes* **109**: 279–294.

Szulejko J and McMahon TB (1993) Progress towards an absolute, proton affinity scale. *Journal of the American Chemical Society* **115**: 7839–7848.

Taft RW (1983) Protonic acidities and basicities in the gas phase and in solution: substituent and solvent effects. *Progress in Physical Organic Chemistry* **14**: 248–350.

Taft RW (1975) Gas phase proton transfer equilibria. In: Caldin EF and Gold V (eds) *Proton Transfer Reactions*, pp 31–78. New York: Wiley.

Traeger JC (1996) The absolute proton affinity for isobutene. *Rapid Communications in Mass Spectrometry* **10:** 119–121.

Williamson DH, Knighton WB and Grimsrud EP (1996) *International Journal of Mass Spectrometry and Ion Physics* **154:** 15–24.

Yamdagni R and Kebarle P (1976) Gas phase basicities and proton affinities of compounds between water and ammonia and substituted benzenes from a continuous ladder of proton transfer equilibrium measurements. *Journal of the American Chemical Society* **98:** 1320–1324.

# Proton Microprobe (Method and Background)

**Geoff W Grime**, University of Oxford, UK

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Methods & Instrumentation**

Although the techniques of ion beam analysis (IBA) have been used for many years with broad beams (5–10 mm diameter), it is only in the 1990s that the technology for focusing high-energy ion beams has developed to the point where the spatial resolution is comparable to that of other forms of probe beam microanalysis (around 1 µm). The interactions most commonly used require protons with energy of 1–4 MeV for optimum sensitivity, and in this form the instrument is commonly known as the proton or nuclear microprobe.

Using a combination of analytical techniques, the nuclear microprobe can provide simultaneous multi-elemental analysis over the entire Periodic Table with a spatial resolution of 1 µm, a minimum detection limit of 1–100 ppm depending on the conditions and a quantitative accuracy of 5–20% depending on the type of analysis. Although the penetration depth of MeV protons can be in the region of 100 µm in some materials, the nuclear microprobe is a surface-biased technique since signals are detected preferentially from the near surface region ($\lesssim$ 10 µm depth).

## Focusing systems

The high momentum of MeV ions which gives them their advantage for proton-induced X-ray emission (PIXE) analysis also makes them difficult to focus. In particular, the standard cylindrical magnetic lenses used in electron focusing columns are too weak to be used with MeV ions, and alternative technologies must be sought. Many different arrangements of apertures and electromagnetic fields have been proposed, but the system which has shown consistently the best performance is the magnetic quadrupole multiplet.

Quadrupole lenses have four magnetic poles arranged symmetrically N–S–N–S around the beam axis (**Figure 1A**). They have a strong focusing action on charged particle beams but, because of the antisymmetry, a single quadrupole lens converges the beam only in one plane and diverges in a plane normal to this. For this reason, two or more quadrupoles of alternating polarity are required to form a point focus. Combinations of two, three and four lenses have been used, as shown in **Figure 1B**.

Like conventional glass lenses, quadrupole lenses suffer from significant angle-dependent aberrations which increase the beam diameter and, together with the other parameters of the experiments, these set a limit to the smallest usable beam diameter that can be achieved. This dependence can be loosely expressed as follows:

$$d \propto \sqrt[n]{\frac{\delta R f(C)}{Y B \Omega}} \qquad [1]$$

where $R$ is the smallest count rate that is required to perform the experiment with good enough statistics in a reasonable time, $\delta$ is the energy spread of the beam, $Y$ is the normalized yield of the analytical reaction being used, $B$ is the brightness of the beam from the accelerator, $\Omega$ is the solid angle of the

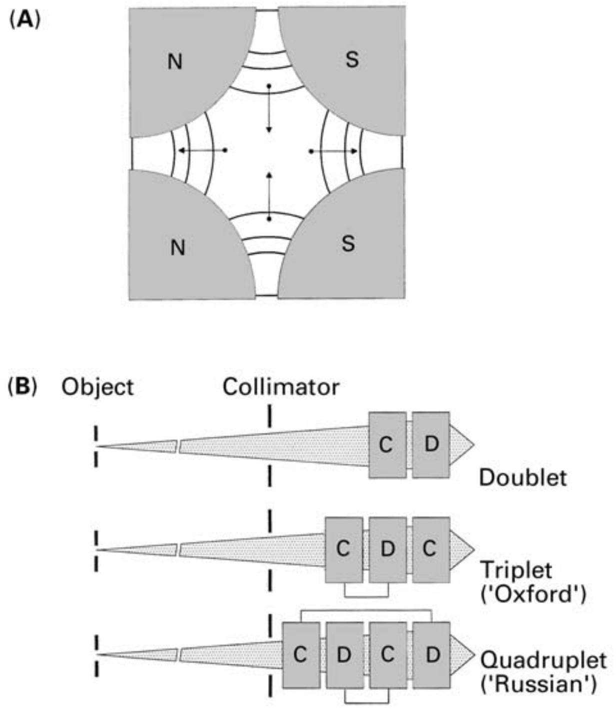



**Figure 1** (A) Cross-section of a quadrupole magnet used as a focusing lens for charged particles. Particles travelling in the vertical plane experience a force directed toward the axis; in the horizontal plane the force is directed away from the axis. (B) Combinations of quadrupoles which have been used for microbeam applications. In the triplet configuration of quadrupole lenses used in the Oxford nuclear microprobe the quadrupoles are excited alternately converging–diverging–converging and the first two magnets have the same excitation (for ease of alignment). Also shown is the object aperture (typically 50 × 10 μm) and the collimator aperture used to control the divergence of the beam (and hence the aberrations).

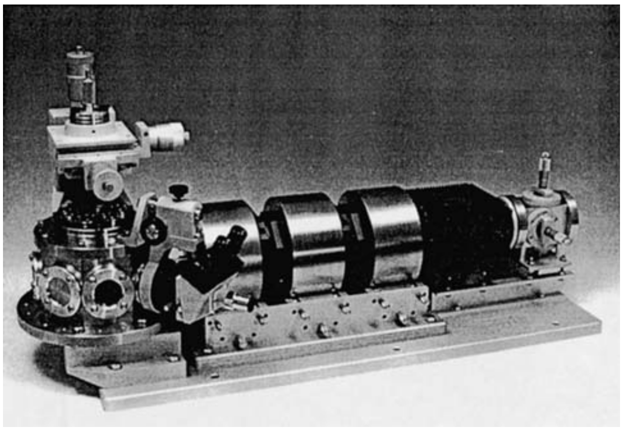


**Figure 2** Photograph of the end-stage of a commercially available microbeam focusing system using a triplet of magnetic quadrupoles. The beam enters the system from the right and passes through the collimator aperture, magnetic deflection coils to sweep the beam across the sample and the three quadrupole magnets before entering the target chamber where the sample is mounted on a micrometer controlled positioning stage. Courtesy *Oxford Microbeams Ltd.*

detector and $f(C)$ is some function of the aberration coefficients of the lens system, which in turn depend upon the precise shape and magnitude of the magnetic field in the system; $n$ varies between 2 and 4, depending on the type of the dominant aberration. From this it can be seen that for the best spatial resolution we must accept a low count rate with a high-yield reaction using a beam with small energy spread and high brightness and a detector with a large solid angle. In practice, there is little that can be done to improve these parameters. However, the aberration coefficients of the lens can be minimized to a certain extent by the configuration and precision of construction and alignment of the lens.

Apart from the intrinsic chromatic and spherical aberration present even in a perfect quadrupole field, numerical ray-tracing studies of quadrupole probe-forming systems show that the major contributions to the beam broadening are from small imperfections in the construction and alignment of the lenses, especially departures from exact four-fold symmetry in the poles of the quadrupole magnet and errors in the relative rotational alignment of the lenses. The lenses used in the most successful design are cut from a single piece of high-quality magnet iron to minimize errors due to the assembly of individual components and are mounted on tables permitting precise mechanical alignment to micrometer accuracy (**Figure 2**). Using this system, a spatial resolution of 1 μm can be achieved routinely with a beam current of around 150 pA of 3 MeV protons, sufficient for PIXE analysis. These lenses are available commercially and are now installed in a number of facilities world-wide. Other facilities differ only in scale or details of the layout, and the general principles remain the same.

## Analytical reactions used in nuclear microscopy

### Analytical reactions

Equation [1] indicates that for optimum spatial resolution, high-yield analytical reactions are required. Ion beam analysis (BA) techniques are discussed elsewhere in this volume and here it is shown that the highest cross-section reactions are PIXE and Rutherford backscattering (RBS), which together cover almost the whole of the periodic table. Nuclear reaction analysis has in general a much lower cross-section and is not routinely used for high-resolution microbeam applications.

**Table 1** Comparison between microPIXE and electron probe microanalysis

| | *MicroPIXE* | *Electron probe microanalysis* |
|---|---|---|
| Primary particle beam | 1–3 MeV protons from a nuclear accelerator focused to 1–10 μm with beam current 0.1–10 nA | 10–100 keV electrons focused to 10–100 nm with beam current 1 nA–1 μA |
| Minimum detection limit | 1–100 ppm depending on beam energy, sample matrix, elemental overlaps, etc. Limited by count rate | 100 ppm–1% depending on beam energy, sample matrix, elemental overlaps, etc. Limited by background |
| Beam penetration depth | 50–100 μm | 1–3 μm |
| Analysis depth | Depends on absorption of emerging X-rays; can be >50 μm | Similar to penetration depth |
| Spatial resolution | Equal to beam diameter | Determined by spreading in sample (typically 0.5–2 μm depending on sample thickness) |
| Quantitative accuracy | Moderate (depends on sample homogeneity). Standardless analysis possible | Good |
| Other options available | Light elements using RBS and NRA. In-air analysis using external beam facility | High-resolution imaging with Z contrast |
| Scale of equipment | Medium to large facility (possible shared access to accelerator) | Compact instrument |
| Availability | < 50 world-wide | Common |

RBS = Rutherford back scattering.
NRA = Nuclear reaction analysis.

**Table 1** presents a comparison of nuclear microbeam analysis using PIXE (microPIXE) with the analogous electron probe microanalysis.

### Other imaging techniques

Other non-analytical techniques are also available using MeV ions which can be used to give imaging contrast for a sample.

**Secondary electron imaging** MeV ions impinging on a surface liberate copious quantities of electrons, and these may be measured using a detector such as a channel electron multiplier to generate images of the sample. The energy of the emitted electrons is relatively low (typically < 1 keV) and so they are easily obstructed by surface irregularities. Detecting the electrons at a low (grazing) angle to the surface enhances this effect and allows clear topographic images to be obtained.

Secondary electrons resulting from proton bombardment do not carry any compositional information about the sample, and the spatial resolution is limited by the beam diameter of the ion beam to 1 μm or greater, which is significantly worse than the same technique used in the scanning electron microscope, but the technique can be valuable in helping to relate elemental distributions to the physical structure of the sample.

**Scanning transmission ion microscopy (STIM)** MeV ions in matter lose energy gradually by collisions with atomic electrons, so by measuring the energy of particles passing through thin (typically <30 μm) samples, a measure of the electron density along the ion path can be obtained. This can be used as a technique for mapping density fluctuation in samples thin enough to transmit the beam. STIM can be carried out in two modes: with the detector directly in the beam behind the sample (bright-field STIM), the contrast occurs solely due to proton energy loss in passing through the sample; with the detector mounted off-axis behind the sample (dark-field STIM), scattered particles are detected and contrast is due to a combination of small-angle scattering and energy loss in the sample. With bright-field STIM, each transmitted proton is a signal (i.e. the yield $Y$ in Equation [1] is 100%) and so the beam current can be reduced to around 1000–2000 particles $s^{-1}$ (~0.1 fA) to avoid detector damage and saturation of the acquisition electronics. This is normally carried out by reducing the beam defining apertures in the final lens, and so an added benefit of bright-field STIM is that the beam diameter is reduced and STIM imaging can be carried out at spatial resolutions of the order of 100 nm. For many applications, dark-field STIM is often used to map the sample to locate regions of interest prior to analysis, and this is normally carried out simultaneously with the PIXE and RBS mapping. A STIM image of two mouse red blood cells infected with malaria parasites on a thin plastic film is shown in **Figure 3**.

**Ion beam induced charge (IBIC)** Charged particles passing through the junctions of active semiconductor devices create electron–hole pairs which can be

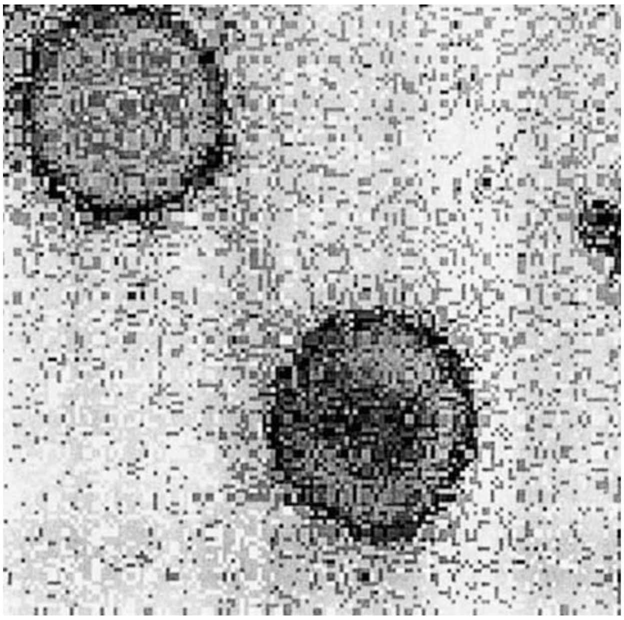


**Figure 3** STIM images of mouse red blood cells (5 µm in diameter), both infected with malaria parasites. The bottom one is well developed and shows the parasite coiled around the inside wall of the cell. Denser or thicker regions of the sample are indicted by lighter colours. Courtesy Professor F. Watt, National University of Singapore.

detected as charge pulses on the electrodes of the device. Thus scanning a semiconductor device using a microbeam gives the possibility of mapping the active regions of the device either for fault finding or for investigating sensitivity to radiation. This technique is more commonly used with electron beams but the use of MeV ions gives the advantage of the long range, which means that junctions may be studied which are covered by metallization or passivation layers. Like STIM, each particle creates a signal and so it is a high-yield technique with the possibility of high-resolution imaging. **Figure 4** shows IBIC images of a gallium arsenide transistor.

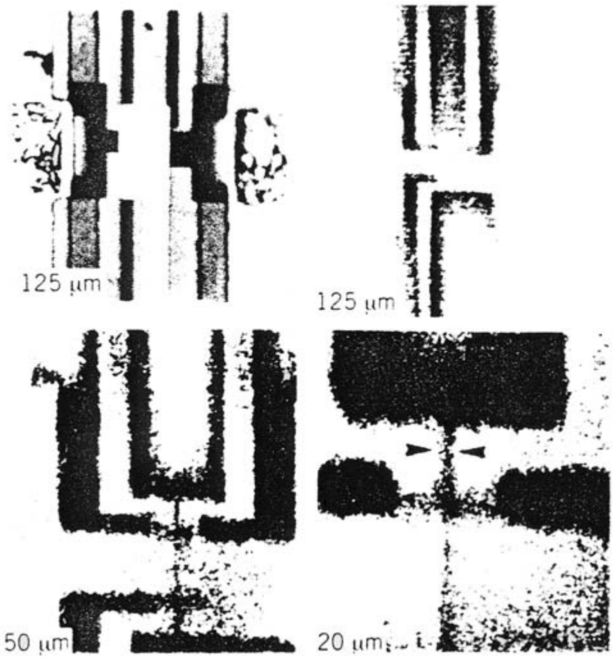



**Figure 4** Optical (top left) and IBIC images of a gallium arsenide transistor at three different magnifications (length of side of image area indicated on each area). The IBIC images show the intensity of charge collected between the p-type and n-type contacts, with the highest charge shown as darker. In the highest magnification image, the arrows indicate a 0.8 µm depletion region, showing the resolution that can be achieved using this technique. Reproduced from Breese MBH, Grime GW, Watt F and Blaikie RJ (1993) *Vacuum* **44**: 175, with permission.

### External beam milliprobe

One consequence of the long range of MeV ions is that they can be brought out of the vacuum system through a suitable exit foil into air, so that large or vacuum-incompatible objects can be analysed. Scattering in the foil and in the air means that the resolution is degraded and, according to the design of

**Table 2** Sample preparation techniques for nuclear microbeam analysis[a]

| *Type of sample* | *Preparation technique* |
|---|---|
| Biological cells | Cryo-fixation on to thin (<1 µm) plastic films |
| Biological tissue sections | Cryo-fixation followed by cryo-sectioning to a thickness of 5–10 µm and freeze-drying on thin (<1 µm) plastic films |
| Environmental particles, other particulates and powders | Small particles on filters: analysis *in situ* on filter membrane.<br>Small particles in bulk: dispersion in dilute resin solution which can be cast into a thin film<br>Large particles: adhesion to conducting adhesive pad |
| Mineralogical samples, metals, bone, etc. | Resin embedding and polishing as for electron probe analysis. Insulating samples may need carbon coating |

[a] Because IBA is predominantly an inner shell/nuclear technique with most of the signal coming from many atomic layers beneath the surface, there is no strong dependence on surface condition (although ideally the surface to be analysed should be flat). Because of this, many samples can be analysed with no further preparation (especially in the external beam). Other types of sample require some preparation, as summarized in this table.

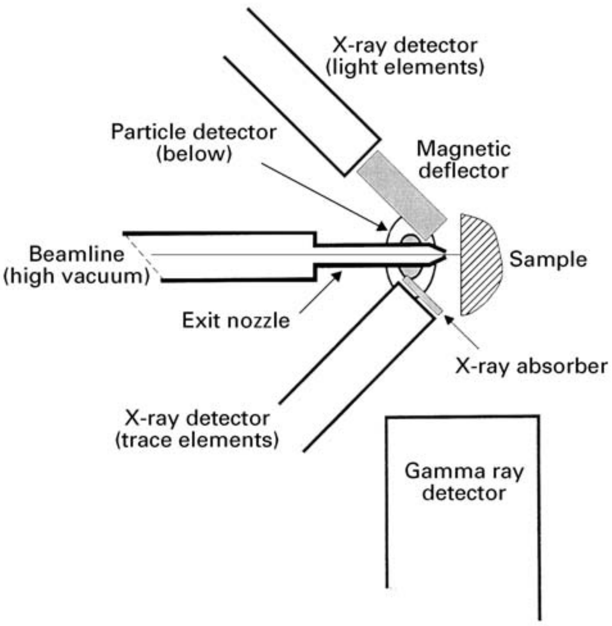



**Figure 5** Schematic plan view of the external beam analysis facility at Oxford. The beam emerges from the vacuum of the beamline through a 300 μm diameter hole sealed with a thin (8 μm) plastic foil. X-rays emitted from the sample are detected by two detectors mounted at 45° on either side of the beam. One detector is fitted with an X-ray absorber to kill intense low-energy X-ray lines from elements in the sample matrix to ensure good detection limits for the trace elements while the other detector has no absorber and is fitted with a magnetic deflector to ensure that recoiling high-energy protons do not reach the detector. A detector for gamma rays is mounted at 90° to the beam. Below the beamline, a detector for recoiling protons allows the total amount of charge falling on the sample to be monitored. Not shown in this diagram are a video microscope which uses a mirror to view the front surface of the sample during analysis and a low power alignment laser to assist in positioning the sample for analysis.

the exit aperture, spatial resolutions from 30 μm up to 1 mm can be achieved. The external beam facility used at Oxford University is shown schematically in **Figure 5**. PIXE is the technique most commonly used with external beams, although RBS can be used (with degraded depth resolution because of scattering from air molecules) and nuclear reaction analysis using gamma ray detection can be used to measure light elements because in general the beam currents are higher than in the microbeam. External beam milliprobes are of particular value for analysing archaeological or historical artifacts that it is not possible to sample or hydrated (or even living) biological samples.

### Practical aspects

**Sample preparation** Sample preparation for nuclear microscopy is similar to that required for electron microscopy with the benefit that because of the much longer range of protons in matter (typically 30–100 μm), there is no requirement for ultrathin samples, and in many cases samples can be analysed with little or no sample preparation. **Table 2** summarizes the main types of sample preparation that may be required for nuclear microscopy.

One important consideration for microPIXE analysis is the need to avoid sample treatments which may introduce trace element contamination at the ppm levels detectable using PIXE. This is a particular problem for biological sample preparation, where fixing and staining agents are normally used to preserve the structure of the sample and render visible the regions of interest. For microPIXE analysis of biological samples, the optimum preparation technique is cryo-fixation and sectioning. The structure of the sample must then be determined by a technique such as staining of adjacent sections or STIM imaging of the sample.

Unlike electron microscopy, there is no strong requirement to have a conducting sample, since the high energy of the proton beam will ensure that the beam will not be deflected by the relatively modest potentials encountered at the impact point. However, sample charging can be a problem for PIXE analysis, since secondary electrons present in the system can be accelerated to the positively charged impact points and create an enhanced bremsstrahlung background, which degrades the sensitivity. To reduce this effect, insulating samples (e.g. minerals, bone) are usually coated with a thin conducting film, such as carbon.

*See also:* **High Energy Ion Beam Analysis; NMR Microscopy; X-Ray Emission Spectroscopy, Applications; X-Ray Emission Spectroscopy, Methods.**

## Further reading

Breese MBH, Jamieson DN and King PJC (1996) *Materials Analysis Using a Nuclear Microprobe*. New York: Wiley.

Llabador Y and Moretto Ph (1996) *Nuclear Microprobes in the Life Sciences*. Singapore: World Scientific.

Watt F and Grime GW (1987) *Principles and Applications of High Energy Ion Microbeams*. Bristol: Adam Hilger.

# Pyrolysis Mass Spectrometry, Methods

**Jacek P Dwarzanski** and **Henk LC Meuzelaar**,
University of Utah, Salt Lake City, UT, USA

Copyright © 1999 Academic Press



Pyrolysis indicates decomposition caused by thermal energy and, as a process of covalent bond dissociation and rearrangement brought about by heat, represents a form of thermolysis. However, to indicate the relatively high temperature of the process, namely approaching the temperature of a fire, the use of the term 'pyrolysis' – that originates from the Greek: pyr, fire – is commonly used. Primary bond-scission occurring during such processes may generate reactive species which can be involved in further reactions called secondary pyrolysis processes or recombination reactions. The products of pyrolysis are composed of fragment molecules, called pyrolysate, and are generated using a family of devices for performing pyrolysis, namely, pyrolysers. An analytical technique based on direct coupling of a pyrolyser with a mass spectrometer that allows the detection and analysis on-line of that portion of the pyrolysate which has adequate vapour pressure to reach the detector has been named pyrolysis-mass spectrometry (Py-MS).

It is assumed that, under appropriate conditions, thermal degradation products reflect to a large extent the original structure of the pyrolysed material. Therefore, pyrolysis is utilized as a form of sample pretreatment and is considered as one of the basic approaches during mass spectrometric investigations of complex materials, because both techniques can be fully integrated.

The observation of primary reaction products of high activation energy processes requires both high temperatures and short residence times of the pyrolysate in the hot zone. Flash vacuum pyrolysis performed in proximity to the ion source of a mass spectrometer in many cases satisfies these criteria; however, the design of the pyrolyser plays an important role in the overall performance of the Py-MS system. Reduction of secondary reactions, which tend to decrease both the character and reproducibility of the pyrolytic products, may be achieved by providing short residence times of the pyrolysate in the hot pyrolysis region. The ultimate objective is to obtain precisely controlled and defined temperature/time profiles of the samples.

The basic steps used to obtain structural information about complex samples by Py-MS are summarized in **Figure 1**.

## Pyrolysis techniques

Pyrolysers can be divided into two main categories on the basis of their mode of operation, i.e. the continuous type, where the sample is supplied to a furnace preheated to the final temperature, and pulse mode reactors in which the sample is introduced into a cold furnace which is then heated to the final pyrolysis temperature. In the analytical pyrolysis of solid and some liquid materials mainly pulse mode pyrolysers are used and the following sections will focus on a few of the most popular pyrolysis techniques utilizing this mode of operation. However, for pyrolytic studies of liquid and gaseous samples continuous pyrolysers are applied.

### Direct probe pyrolysers

Samples to be analysed may be introduced into an MS ion source using heated probe units that allow the material to be deposited directly on straight, folded or coiled filaments or ribbons. Alternatively, a quartz tube or crucible can be placed inside a heater coil and subjected to a selected rate of temperature increase by resistive or inductive heating. The filament or coil is positioned close to the ion source to

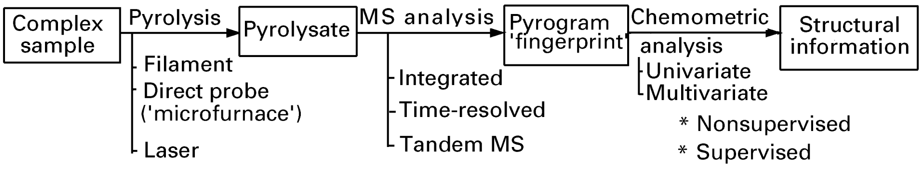



**Figure 1** Basic approach options to computerized pyrolysis mass spectrometry.

release pyrolysis products generated under high vacuum conditions.

In principle, any type of magnetic or quadrupole mass spectrometer can be utilized for the analytical pyrolysis of organic materials, if a direct introduction system capable of producing a desired temperature/time profile is available. For example, direct insertion probes (DIPs) and direct exposure probes (DEPs) are widely used for sample introduction and such probes are supplied with control units that allow heating and temperature programming of the sample up to 500–800°C. Therefore, such modules should be considered as the most readily available probes for Py-MS studies.

The temperature rise time (TRT), i.e. the time required for the pyrolyser temperature to be increased from its initial to the final temperature, can be chosen in the range from several milliseconds to several minutes. Simultaneously, the temperature time profile (TTP), representing temperature as a function of time for a particular pyrolysis experiment, may be easily programmable. Pyrolysis may be carried out at a fast rate of temperature increase, e.g. 10 000 K $s^{-1}$ in the case of flash pyrolysis, or the sample can be heated at a controlled rate over a temperature range in which pyrolysis occurs using a stepwise, linear or ballistic heating approach characterized by a total heating time (THT) of several minutes or even hours.

To illustrate the relationships among parameters that determine the heating profile for a pyrolytic experiment, a graphical representation of a TTP for an isothermal pyrolysis taking place at the equilibrium temperature is presented in **Figure 2**.

Time-resolved recording of pyrolysate spectra, sometimes referred to as a linear programmed thermal degradation mass spectrometry (LPTDMS), by using a slow heating rate adds additional information (**Figure 3A**) and enables one to perform pyrolysis close to the ion source (**Figures 4** and **5**). However, considerable contamination of the ion

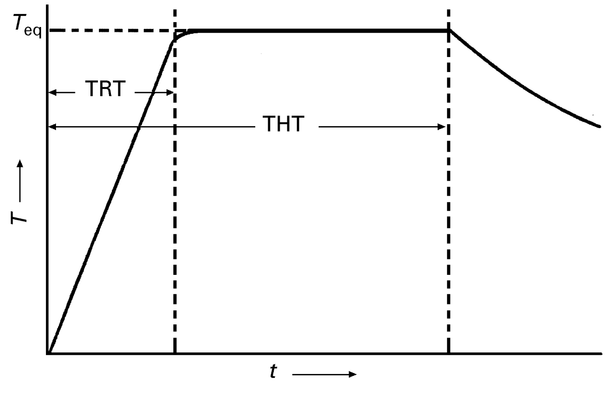



**Figure 2** Schematic representation of the parameters that determine the heating profile in filament pyrolysis, namely temperature rise time (TRT), equilibrium temperature ($T_{eq}$) and total heating time (THT).

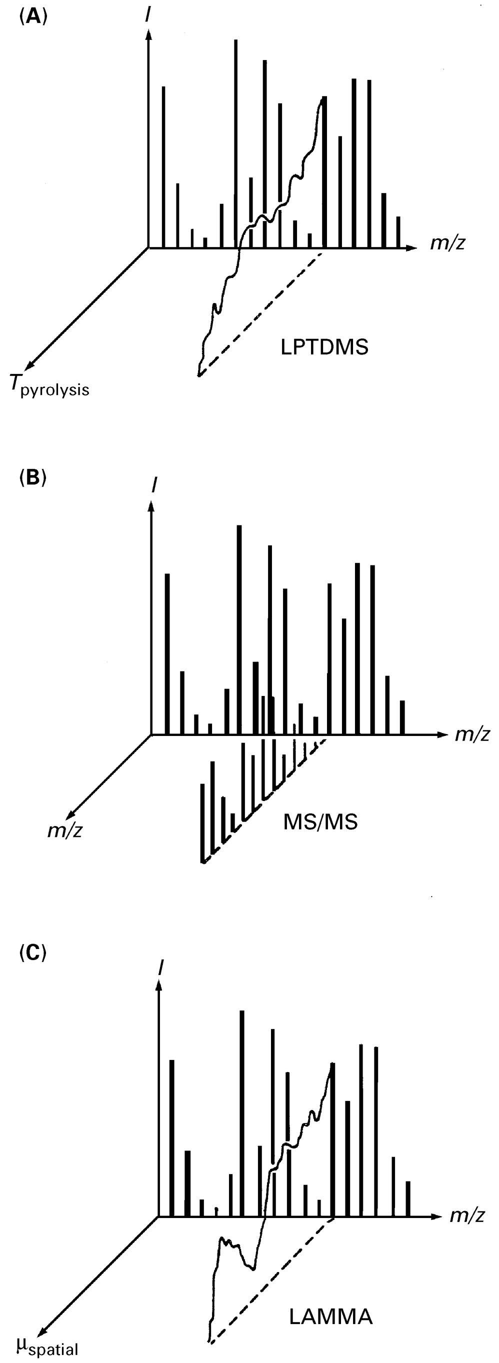



**Figure 3** Third dimension in pyrolysis mass spectrometry approaches: (A) linear programmed thermal degradation mass spectrometry [LPTDMS – third dimension = temperature]; (B) collisionally activated dissociation of 'parent' ions coupled with scanning of product ions using tandem mass spectrometry [MS/MS – third dimension = spectrum of product ions]; (C) laser microprobe mass analyser [LAMMA – third dimension = spatial resolution].

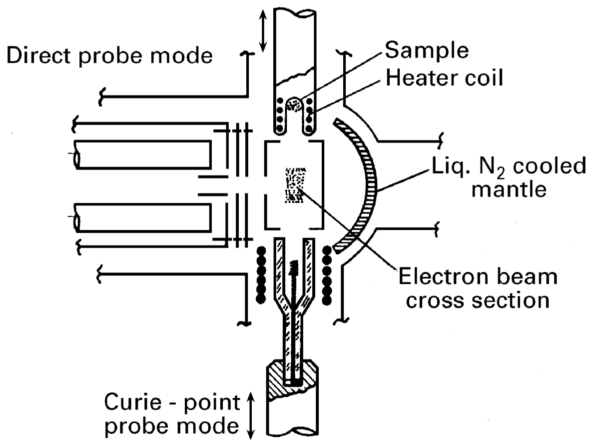



**Figure 4** Typical instrumental configurations for pyrolysis electron-impact ionization mass spectrometry: direct insertion probe pyrolysis mode (upper) and Curie-point pyrolysis mode (lower). Reproduced by permission of Elsevier Science from Meuzelaar HLC, Windig W, Huff SM and Richards JM (1986). *Analytica Chimica Acta* **190**: 119–132.

source, together with increased charring and the occurrence of secondary reactions, make this approach troublesome in many cases.

### Curie-point pyrolysers

Curie-point pyrolysis takes advantage of fast inductive heating of ferromagnetic materials placed in the high frequency (0.5–1.0 MHz) electromagnetic field generated inside the inductive coil surrounding the pyrolytic reaction tube. This produces fast heating of solid or liquid materials coated on – or gases flowing around – a ferromagnetic sample carrier in the shape of a filament, foil or tube.

When the high frequency field is switched on, the ferromagnetic carrier inductively heats to its Curie-point temperature ($T_c$), at which the metal loses its ferromagneticity, becomes paramagnetic and the heating effect decreases drastically owing to strongly reduced energy absorption from the high frequency field in the coil. This thermostatic effect ensures control of the final pyrolysis temperature ($T_{eq}$) determined by a sharply defined transition temperature that is specific for the composition of the alloy used, at the point where the residual energy absorption by eddy currents is balanced by the loss of heat through radiation and conduction.

The TRT for a pyrolyser reactor temperature (**Table 1**) is determined by the shape and diameter of the carrier, the composition of the alloy, the strength of the high frequency field and its frequency (**Figure 6**). Since under practical Py-MS conditions pyrolysis may take place before the wire reaches the final temperature, the heating rate is generally thought to have a critical influence on the pyrolysis patterns. However, other factors, such as the choice of the filament cleaning technique and of the solvent, as well as factors that govern either transfer of the pyrolysate to the ion source or influence ionization conditions, appear to determine the reproducibility of Py-MS more strongly than the temperature/time profile, especially with respect to long-term reproducibility.

Careful attention should be paid to the purity of the surfaces of the pyrolysis filament as well as of the

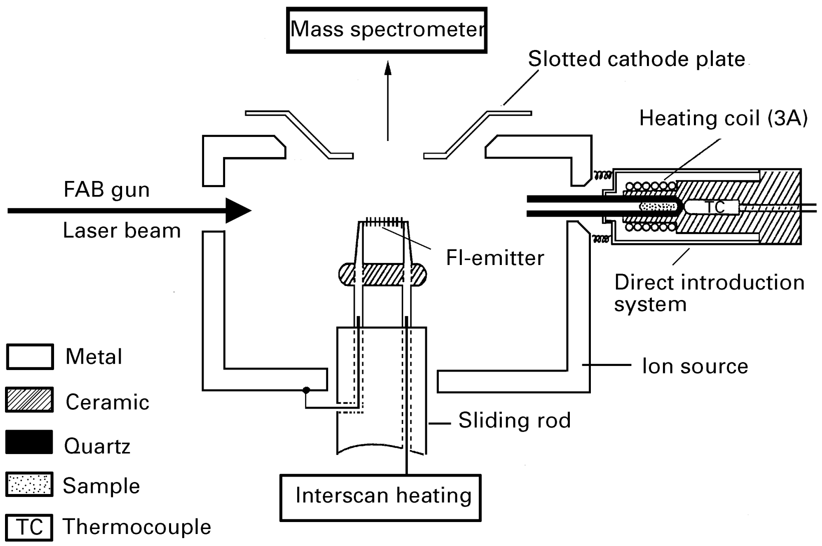



**Figure 5** Schematic drawing of the experimental setup for pyrolysis field ionization mass spectrometry. Reproduced by permission of the American Chemical Society from Schulten H-R, Simmleit N and Muller R (1987). *Analytical Chemistry* **59**: 2903–2908.

**Table 1** Temperature rise time (TRT) and Curie-point versus alloy composition

| *Alloy composition (%)* | | | | *Quoted TRTs (ms)* | |
|---|---|---|---|---|---|
| *Fe* | *Ni* | *Co* | *Curie-point (°C)* | *Fisher-Varian (1500 W)* | *Phillips (30 W)* |
| 0 | 100 | 0 | 358 | 300 | 1300 |
| 61.7 | 0 | 38.3 | 400 | 40 | 500 |
| 50.6 | 49.4 | 0 | 510 | 150 | 700 |
| 42.0 | 41.0 | 16.0 | 600 | 70 | 500 |
| 29.2 | 70.8 | 0 | 610 | 130 | 1150 |
| 33.0 | 33.0 | 33.0 | 700 | 90 | 1350 |
| 100 | 0 | 0 | 770 | 110 | 2100 |

Reproduced by permission of John Wiley & Sons from Grob RL (ed) (1995) *Modern Practice of Gas Chromatography*, 3rd edn. New York: John Wiley & Sons.

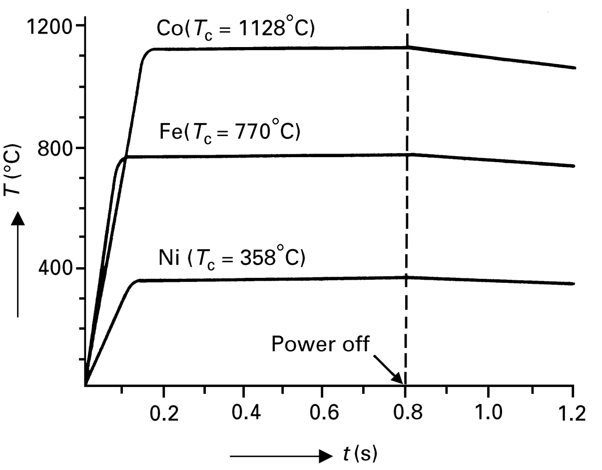



**Figure 6** Temperature/time profiles and Curie-point temperatures for pure Ni, Fe, and Co wires (diameter 0.5 mm) when using a 1.5 kW, 1.1 MHz high frequency power supply for pyrolysis/mass spectrometry studies. Reproduced by permission of Elsevier Science from Meuzelaar HLC, Haverkamp J and Hileman FD (1982), *Pyrolysis Mass Spectrometry of Recent and Fossil Biomaterials; Compendium and Atlas*. Amsterdam: Elsevier Science.

glass reaction tube. Although Curie-point wires are inexpensive and are usually discarded after use, glass or quartz reaction tubes should be carefully cleaned before their next use. The choice of the filament cleaning method may influence the pyrolysis pattern. For example, prolonged heating in water-saturated hydrogen, as a reductive cleaning technique, removes oxide layers on the filament surface which otherwise could have an oxidative or catalytic effect and change the emissivity of the filament surface. However, this cleaning technique can cause severe hydrogen absorption by the metal and this could conceivably influence the pyrolysis process by promoting catalytic hydrogenation reactions. Several other factors, such as substitution of a wire for a ferromagnetic tube, or wrapping the sample inside a metal foil, produce pronounced changes in the pyrolysis mass spectra of most compounds.

A pyrolysate composed of a multicomponent mixture, under ideal conditions, should allow the pyrolysis products to reach the ionization zone without any loss, degradation or recombination of products during transfer. However, in practice these conditions are never fulfilled owing to the dynamic character of the whole process, which is governed by a complex set of kinetic parameters related to every chemical reaction taking place during the pyrolysis stage as well as to heat and mass transfer conditions, and interactions with walls. In fact, pyrolytic products represent a mixture of compounds characterized by a broad range of relative molecular masses and vapour pressures. As a result, some pyrolysis products tend to remain on the filament in the form of macromolecular, nonvolatile chars and consequently are inaccessible for further analysis by mass spectrometry. It is well established that the amount of char formed is inversely proportional to the heating rate and can be minimized by avoiding excessively slow heating rates. Another group of compounds is volatile enough to escape from the hot pyrolysis carrier but condenses on the walls of the reaction chamber and transfer lines. Although pyrolysis directly in the ion source is possible, in practice it is difficult to avoid strong contamination of the ion source under these conditions. Therefore, the glass reaction tube surrounding the ferromagnetic carrier serves as a trap for relatively nonvolatile pyrolysis products which otherwise could contaminate the ion source. In addition, these tubes help to obtain maximum signal intensity by producing a forward oriented beam of volatile pyrolysis products which may directly enter the expansion chamber or the ion source, and the ions formed are mass analysed or studied using collisionally induced dissociation (**Figure 3B**).

To broaden the pressure–time profile in the ion source, and thus to enable the recording of a sufficient number of mass spectra generate a representative averaged mass spectrum of the pyrolysate, an expansion chamber is usually placed between a pyrolytic reactor and the ion source. The expansion chamber should be heated and have chemically inert walls to avoid interaction with pyrolysis products. Quartz or gold-plated expansion chambers, heated to about 150–200°C are frequently used to provide a buffer volume (**Figure 7**). However, the removal of the expansion chamber and positioning of the reaction tube directly in front of the ion source, together with slowing the heating rate of the wire, allow for

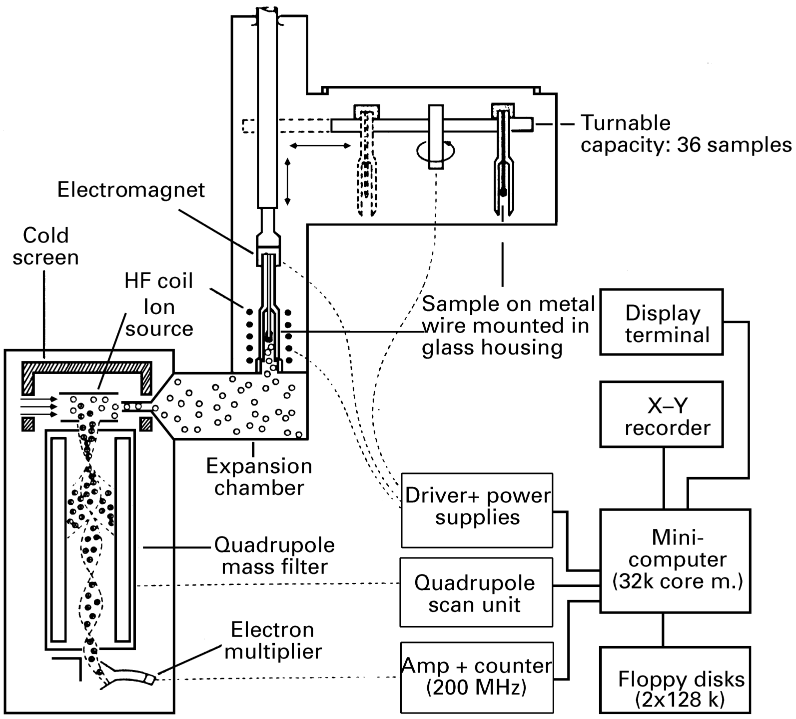



**Figure 7** Scheme of an automated Curie-point Py-MS system with a dedicated minicomputer. Samples are selected by a pickup arm from the turntable and positioned inside the high frequency coil. The pyrolysate diffuses via a buffer volume into the mass spectrometer where the molecules are ionized and mass-analysed. Reproduced by permission of Plenum Press from Wieten G, Meuzelaar HLC and Haverkamp J (1984). Analytical pyrolysis in clinical and pharmaceutical microbiology. In: Odham G, Larsson L and Mardh P-A (eds) *Gas Chromatography/Mass Spectrometry Applications in Microbiology*. New York: Plenum Press.

sufficient scanning of the pressure–time profile by the quadrupole mass spectrometers (**Figures** 5 and 8).

In the case of gas phase pyrolytic reactions the pyrolysis module works as a flow reactor with gaseous molecules passing along the preheated filament into a mass spectrometer. This technique allows quantitative detection of products with half-lives greater than 1 ms.

### Laser pyrolysers

Laser pyrolysers achieve very high TRTs together with the possibility of characterizing very small surfaces. With a spatial resolution as high as 0.5 μm (Figure 3C) and sample quantities in the range of picograms such instruments permit analysis of very small objects, for example a single bacterial cell. A high power (~$10^9$ W $cm^{-2}$) beam pulse from a commercially available laser microprobe mass analyser (LAMMA) is provided by Nd:YAG laser and focused on the sample by a microscope objective.

Short pulses (~15 ns) create a plasma of positive and negative ions which are mass analysed by a time-of-flight (TOF) mass spectrometer. Although the mechanism of laser-induced ionization and volatilization of solids is not well known, it is established that the duration and shape of the laser pulse affect the nature of the volatilization process. The region of direct impact of the laser on a sample is associated with high temperatures and consequently only small atomic and molecular fragments are emitted from this region. In the region immediately adjacent to the area of direct laser beam impact, a temperature gradient governed by a complex set of pyrolysis reactions will occur, resulting in emission of higher relative molecular mass products.

## Analytical systems based on pyrolysis-mass spectrometry

Analytical systems based on Py-MS are intended to perform pyrolytic reactions and to analyse the composition of the resultant pyrolysate by mass analysis of ions formed through the ionization of its components molecules. The structural complexity of the spectra recorded is usually considerable, owing to the formation of fragments with the same nominal mass from originally different components and the absence of reliable separation procedures. In addition, the observed complexity originates from secondary fragmentation processes, hence various techniques have been used that minimize these

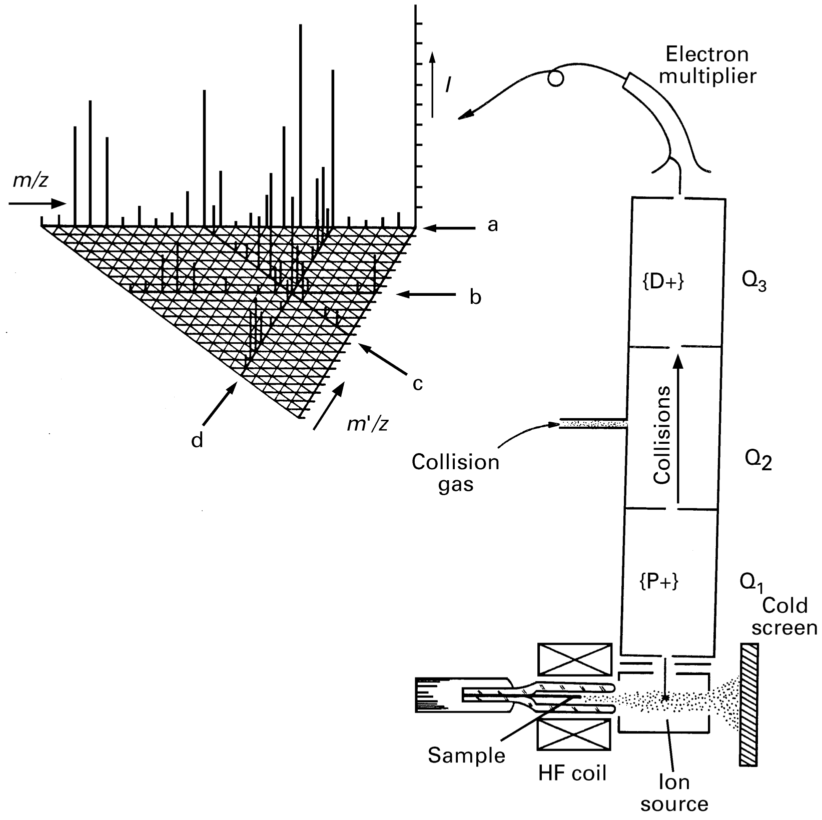



**Figure 8** Curie-point pyrolysis MS-MS system: pyrolysis products diffuses into the ion source where the ions are formed by electron-impact (EI) ionization and analysed by a triple quadrupole mass analyser (Q1–Q3). Different modes of ion analysis: (a) EI spectrum of a pyrolysate; (b) neutral loss spectrum; (c) single ion spectrum; and (d) collisionally induced dissociation spectrum of a 'parent' ion $\{P^+\}$.

effects. These include low-voltage electron ionization (LVEI) with 10–15 eV electrons instead of the standard 70 eV; chemical ionization (CI) with different reagent ions under vacuum or at atmospheric pressure and field ionization (FI). Despite the complexity of the mass spectra recorded, here referred to as mass pyrograms, even low-resolution spectra frequently provide structural information through the presence of characteristic ions or ion series. However, to provide clarity, many other techniques can be used, e.g. time/temperature resolved thermal degradation/pyrolysis MS, high-resolution mass spectrometry (HRMS) and tandem in space or tandem in time mass spectrometry ($MS^n$, $n \geq 2$) following collisionally activated dissociation of selected ions.

A schematic overview of some of the most frequently reported experimental arrangements of pyrolysis zones with respect to ionization regions is given in **Figure 9**. In configurations A–C the reaction zones are under vacuum whereas in D and E pyrolytic regions are kept at near-ambient pressures and for F at high pressure. Examples of integral zones include pyrolysis-field desorption in the ion source of a mass spectrometer (**Figure 10**), laser pyrolysis/photolysis/desorption/ionization mass analyser (LAMMA) and, to some extent, in-source pyrolysis chemical ionization MS. This type of configuration allows the detection of large, and frequently highly informative, molecular or fragment ions from nanogram quantities of a sample. The use of field-desorption ionization coupled with HRMS, as well as field ionization techniques which have been developed to produce a high intensity of molecular ions and to reduce fragmentation of polar compounds, are well suited for the molecular characterization of polymeric building blocks. This approach is characterized by a very short residence time of the pyrolytic sample in the reaction zone ($10^{-11}$ s), thus allowing for detection of short-lived primary products.

The other way to avoid losses of higher relative molecular mass components and less volatile thermal fragments, which are unable to enter the ion source owing to the condensation, is to place a pyrolysis probe very close to the ionization source, thus providing minimal separation between reaction and

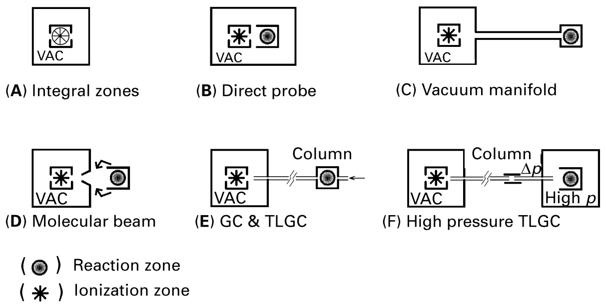



**Figure 9** Schematic representation of six on-line pyrolysis-mass spectrometry configurations. (*) Ionization zone; (•) pyrolysis zone.

ionization zones. Typical examples of this type of configuration are shown in **Figures 4**, **5**, **8** and **11**; however, vacuum TG-MS systems also fulfil the criteria of the configuration in **Figure 9B**. The so-called in-source pyrolysis produces compounds that can be ionized either under EI, CI or FI conditions. The application of such techniques for analysis of bacteria, lignocellulosic materials and polysaccharides has been documented, indicating additional advantages of this technique derived from the slower heating rate and the absence of mixing in an expansion chamber that represent the type of configuration shown in **Figure 9C** and includes a specially designed, fully automated Curie-point Py-MS system with an expansion chamber (**Figure 7**) but frequently represents the result of adapting an existing gas inlet manifold for on-line reaction studies.

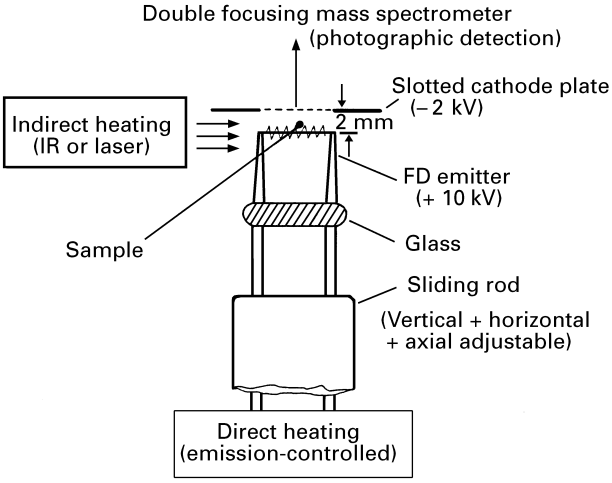



**Figure 10** Field desorption (FD) pyrolysis in the ion source of a mass spectrometer. The place of pyrolysis and ionization is identical. Reproduced by permission of Elsevier Science from Schulten H-R (1977) Pyrolysis field ionization and field desorption mass spectrometry of biomacromolecules, microorganisms, and tissue material. In: Jones CER and Cramers CA (eds) *Analytical Pyrolysis.* Amsterdam: Elsevier Science.

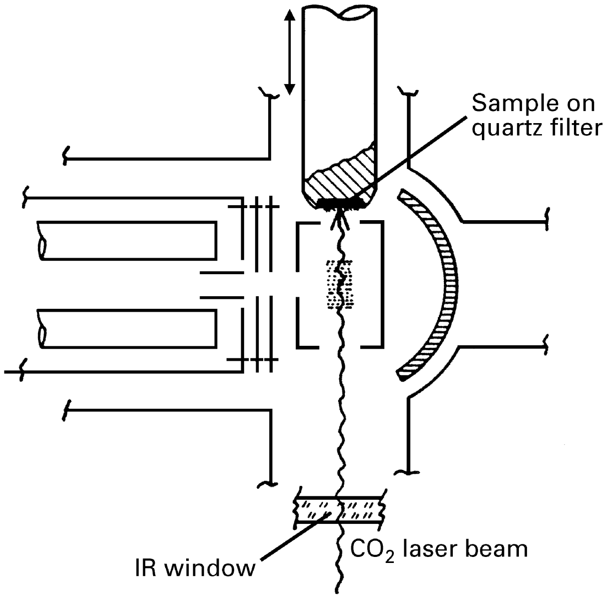



**Figure 11** Instrumental configuration for laser pyrolysis EI ionization mass spectrometry. Reproduced by permission of Elsevier Science from Meuzelaar HLC, Windig W, Huff SM and Richards JM (1986). *Analytica Chimica Acta* **190**: 119–132.

The temperature-resolved data acquisition mode of operation provides additional layers of information. The evaluation of time-resolved mass spectral data of temperature-programmed pyrolyses include the shape of the total ion current profile, variations of the average relative molecular mass of volatilized pyrolysis products as well as the selection of characteristic mass signals, typical for distinct pyrolysis intervals. In addition, the differentiating key signals can be easily found and used for further investigation, e.g. temperature-resolved profiles of single ion intensities or specific sets of ion profiles can be selected to study pyrolysis processes of distinct sample components. Such components can be selected using chemometric techniques, e.g. factor or principal component analysis of time-resolved series of spectra combined with the 'numerical extraction' of chemical components, as revealed by the calculated mass spectra.

The identity of mass peaks or assignment of chemical structures to mass signals is complicated by the presence of isobaric fragment ions and isomeric structures of molecular ions. However, the presence of isobaric ions with a different composition can be investigated by HRMS, and pyrolysis product identification may further be supported by using MS-MS and/or GC-MS methods. Nevertheless, the combination of various ionization techniques, HRMS and, especially, MS-MS (**Figure 8**) can provide the extra

specificity required for the analysis of complex materials.

If the pressure drop across the orifice or short capillary tube in a mass spectrometer sampling system exceeds a critical value of about 2.5, the flow velocity will reach the speed of sound and the flow beyond the exit expands into a supersonic 'free' jet, thereby producing a molecular beam (MB). This process is characterized by extreme collisional as well as internal energy (rotational and vibrational) cooling caused by isentropic and adiabatic expansion of a gas after crossing the sampling orifice (**Figure 9D**). A typical molecular-beam mass spectrometer (MBMS) sampling system coupled to a pyrolysis vapour generator is shown in **Figure 12**. It consists of a conical, extractive sampling probe, a skimmer to collimate the molecular beam and a line-of-sight MB inlet into the ion source of the mass spectrometer. Ions formed by EI ionization are directly mass analysed – or additionally studied using collisionally induced dissociation – with quadrupole mass filters or other mass spectrometric analysers.

The remaining two configurations (**Figures 9E** and **9F**) are characterized by inherently lower molecular conductances and are usually referred to as pyrolysis GC-MS or represent a form of mass spectrometric monitoring of high pressure reactors, including industrial-scale units.

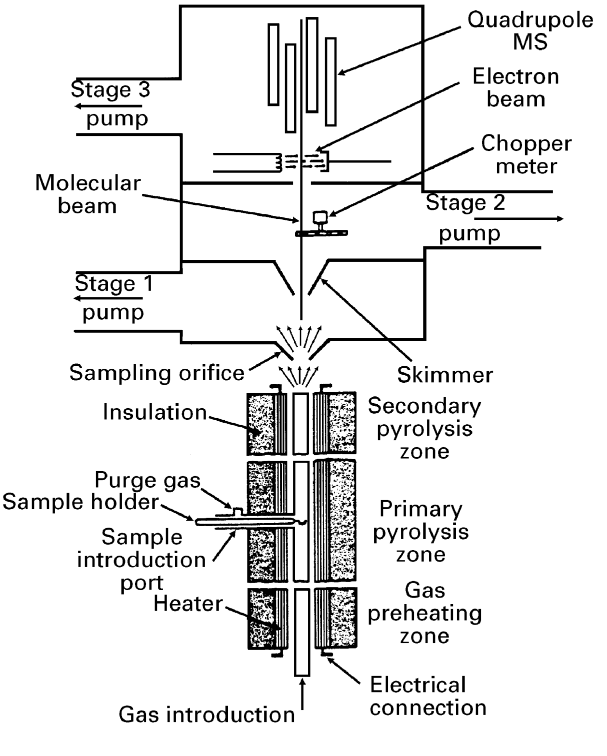



**Figure 12** Schematic of a pyrolysis vapour generator coupled to a molecular-beam mass spectrometer sampling system. Reproduced by permission of the American Chemical Society from Evans RJ and Milne TA (1987) *Energy and Fuels* **1**: 123–137.

## Reproducibility and data analysis procedures

Short-term reproducibility of pyrolysis-mass spectra recorded during Py-MS studies is usually good and variations of peak intensities to within 5% are readily achieved. However, interlaboratory and long-term reproducibilities are not well evaluated. Some of the factors that influence long-term reproducibility and which should be carefully evaluated in interlaboratory comparisons are presented in **Table 2**.

Progress in the design of mass spectrometers and the availability of online computer systems has allowed the integration of Py-MS data acquisition with multivariate mathematical data reduction methods into a single analysis technique. Such an approach combines rapid analysis capability with expert system or pattern recognition based data evaluation (**Figure 13**).

## Applications of Py-MS methods

### Natural products

Natural products in general, and organic natural products in particular, are characterized by a high degree of complexity, thus requiring highly specialized analytical methods which tend to be labour intensive, lengthy and expensive. However, in the early 1960s the search for extraterrestrial life, as part of the new space probe programmes, prompted the development and application of pyrolytic methods to complex biomaterials and microorganisms. The combination of pyrolysis and mass spectrometry was found to be an especially powerful tool

**Table 2** Factors that possibly influence long-term reproducibility

| *Sample preparation* | *Filament heating* | *Product transfer* | *Mass analysis* |
|---|---|---|---|
| Cleaning of sample carrier | Temperature rise time | Inlet temperature | Ionization |
| Solvent/suspending liquid | Equilibrium temperature | Residence time | Extraction |
| Sample size | Total heating time | Surface activity | Transmission |

Reproduced by permission of Elsevier Science from Meuzelaar HLC, Haverkamp J and Hileman FD (1982). *Pyrolysis Mass Spectrometry of Recent and Fossil Biomaterials; Compendium and Atlas*, Amsterdam: Elsevier Science

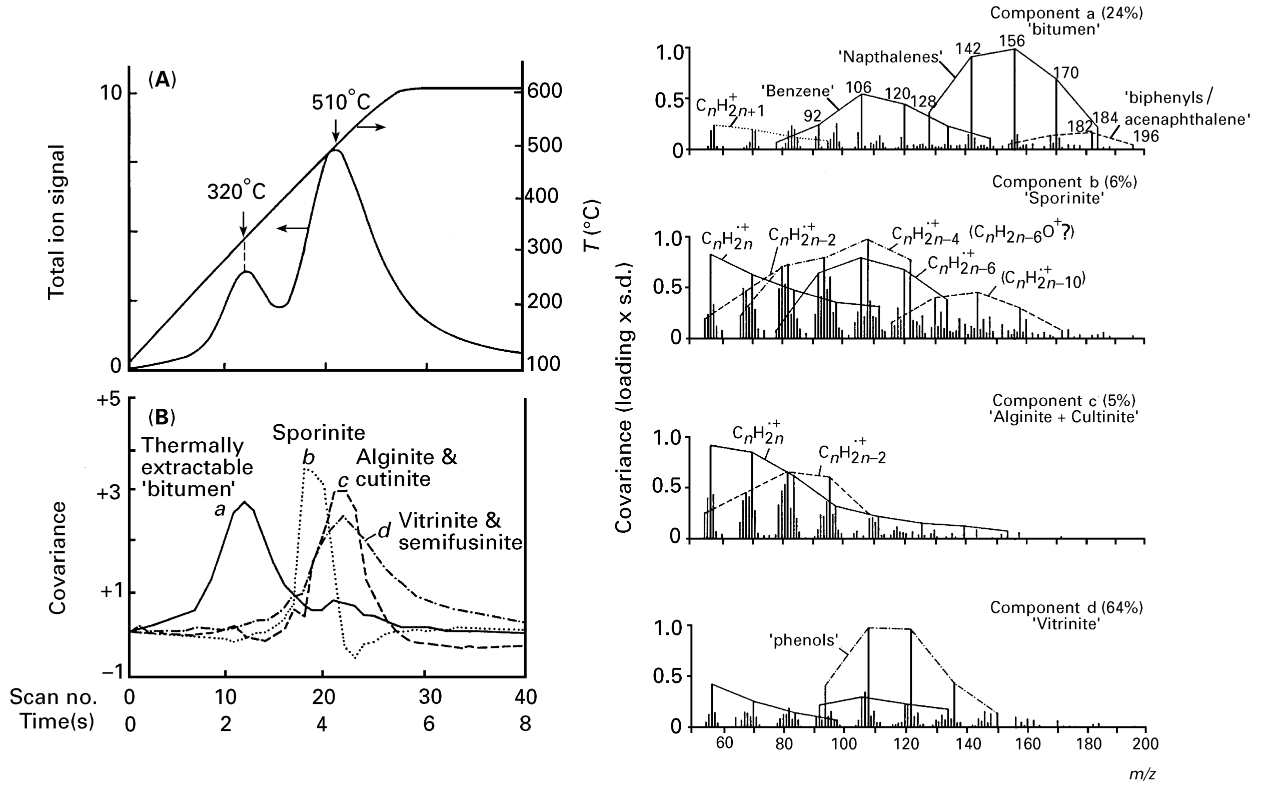



**Figure 13** Time-resolved Curie-point pyrolysis MS total ion signal profile of a coal sample obtained at a heating rate of 100 K $s^{-1}$ (A) and the 'deconvolution' of the second maximum (at 510°C) into at least three components (b–d) by means of factor analysis (B). Tentative interpretation of components a–d was based on comparison of numerically extracted spectra with the actual spectra of maceral concentrates. Numbers in parentheses show the percent of the total variance for each component. Reproduced by permission of Plenum Press from Meuzelaar HLC, Yun Y, Chakravarty T and Metcalf GS (1992) Computer-enhanced pyrolysis mass spectrometry: a new window on coal structure and reactivity. In: Meuzelaar HLC (ed) *Advances in Coal Spectroscopy*. New York: Plenum Press.

to produce structural information about the molecular building blocks of many natural products that are solid, high relative molecular mass, insoluble, homo- or heteropolymeric materials. Moreover, advantages of Py-MS methods include sensitivity, specificity and speed. However, it should be noted that inorganic substances, such as salts and metal cations, may have a catalytic influence on the formation of organic pyrolysis products and that highly heterogeneous, cross-linked materials will not yield 100% pyrolysate but will 'disproportionate' into volatile matter and char. For example, it was shown that alkali metals affect the thermal activation of polysaccharides, which can result in a total loss of oligomer information.

To make quantitative estimates of the portion of sample analysed, it is advisable to determine the amount and approximate composition of the pyrolysis residue. Generally, volatilization tends to decrease with the increasing polarity of a material – owing to the presence of additional intermolecular forces. As a consequence, pyrolysates of complex natural materials, e.g. soil particles, will under represent functional groups containing polarized bonds. To overcome these problems, various on-line and off-line derivatization techniques have been developed. For example, addition of a tetramethylammonium hydroxide (TMAH) solution to a sample deposited on a pyrolytic wire allows a wide range of O–, N–, S– and P–containing polar compounds to be analysed owing to the largely increased volatility of methylated products formed during the pyrolytic methylation step. However, in some cases, by using Py-MS systems with integrated reaction and ionization zones, high molecular mass oligomers can be recorded, as illustrated in **Figure 14**. In this case a series of ions representing 1,6-anhydro-oligosaccharides, ranging from monomer to dodecamer, obtained by in-source pyrolysis chemical ionization of cellulose can be observed.

In other cases, highly specific derivatives may be produced under pyrolytic conditions. For example, diketopiperazines (DKPs) are formed during the thermal decomposition of oligopeptides that range

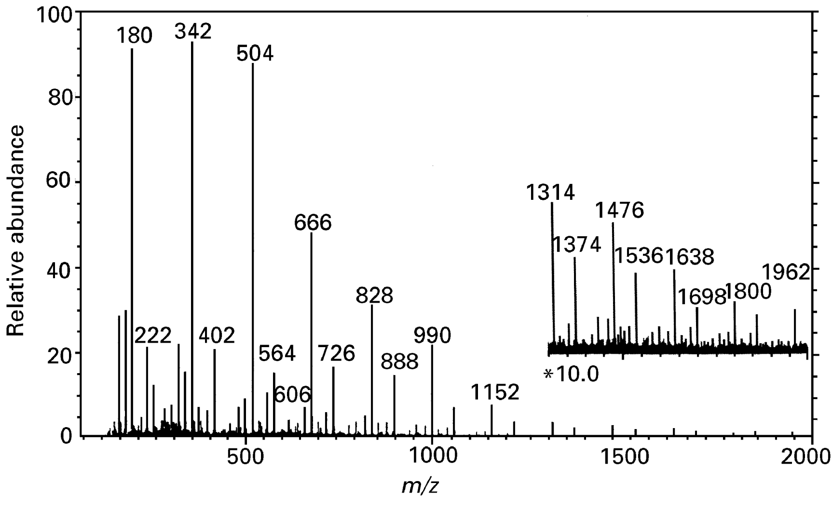



**Figure 14** In-source pyrolysis ammonia chemical ionization mass spectrum of cellulose, showing pseudomolecular ions $[MNH_4]^+$ of a series of 1,6-anhydro-oligosaccharides, ranging from the monomer (*m/z* 180) to the dodecamer (*m/z* 1962). The spectrum was obtained using a Pt–Rh filament probe and an E–B-type sector instrument. Reproduced by permission of Elsevier Science from Boon JJ (1992) Analytical pyrolysis mass spectrometry: new vistas opened by temperature-resolved in-source PYMS. *International Journal of Mass Spectrometry and Ion Processes* **118/119**: 755–787.

from two to six amino acids. In **Figure 15A** a pyrolysis mass spectrum of product ions derived from the parent ion at *m/z* 260, obtained during Py-MS of phenylalanyl-leucine is presented, whereas **Figure 15B** shows general fragmentation pathway of these dipeptide-derived ions. The results showed a variability in the relative intensities of the various DKPs which depends on the amino acids involved in the cyclization process and their position in the peptide. Identification of pure dipeptides from DKP formation observed in Py-MS is rapid and involves no derivatization or complicated sample preparation procedures.

## Bacteria

Classification and identification of microorganisms based on rapid and specific methods for determining the chemical composition have proved to be a valuable approach in microbiology. The identification of many chemical constituents that show a degree of specificity suitable for chemotaxonomic and/or diagnostic purposes made it clear that improvements in methods of sample preparation play an important role in fully utilizing the high speed and specificity offered by MS.

Py-MS is commonly applied for bacterial fingerprinting, i.e. classification of characteristic signal patterns followed the first reports published in the 1960s on the applicability of analytical pyrolysis techniques to clinical and pharmaceutical microbiology. Application of Py-MS as an independent tool for the characterization and identification of bacteria represents an alternative approach that may provide important information for the discrimination of closely related microbial strains that is difficult to obtain by other techniques. Hence, the objectives of most projects were focused on the differentiation and classification of bacterial strains at either the genus, species or subspecies level and subsequent identification of unknown strains. The structural information which can be obtained from the pyrograms and the quantitative nature of the data allow them to be successfully used not only for both the chemical characterization of whole cells and cellular components and also in quality control and screening.

Important advantages of Py-MS techniques are the small sample size required, the high sample throughput, and the ready compatibility of the data with computerized evaluation. To classify the Py-MS profiles of various bacteria an expert system was constructed using multivariate statistical analysis. However, various aspects should be taken into consideration during application of this technique in microbiology, i.e. instrument and biological. Standardization of instrumental conditions is essential to achieve an acceptable level of long-term and interlaboratory reproducibility while standardized culturing and sampling are required to minimize biological heterogeneity among the samples. Sampling directly from the solid medium, and thus avoiding further manipulations, proved to be superior to methods involving the washing of harvested bacteria to remove culturing media, often followed by lyophylization

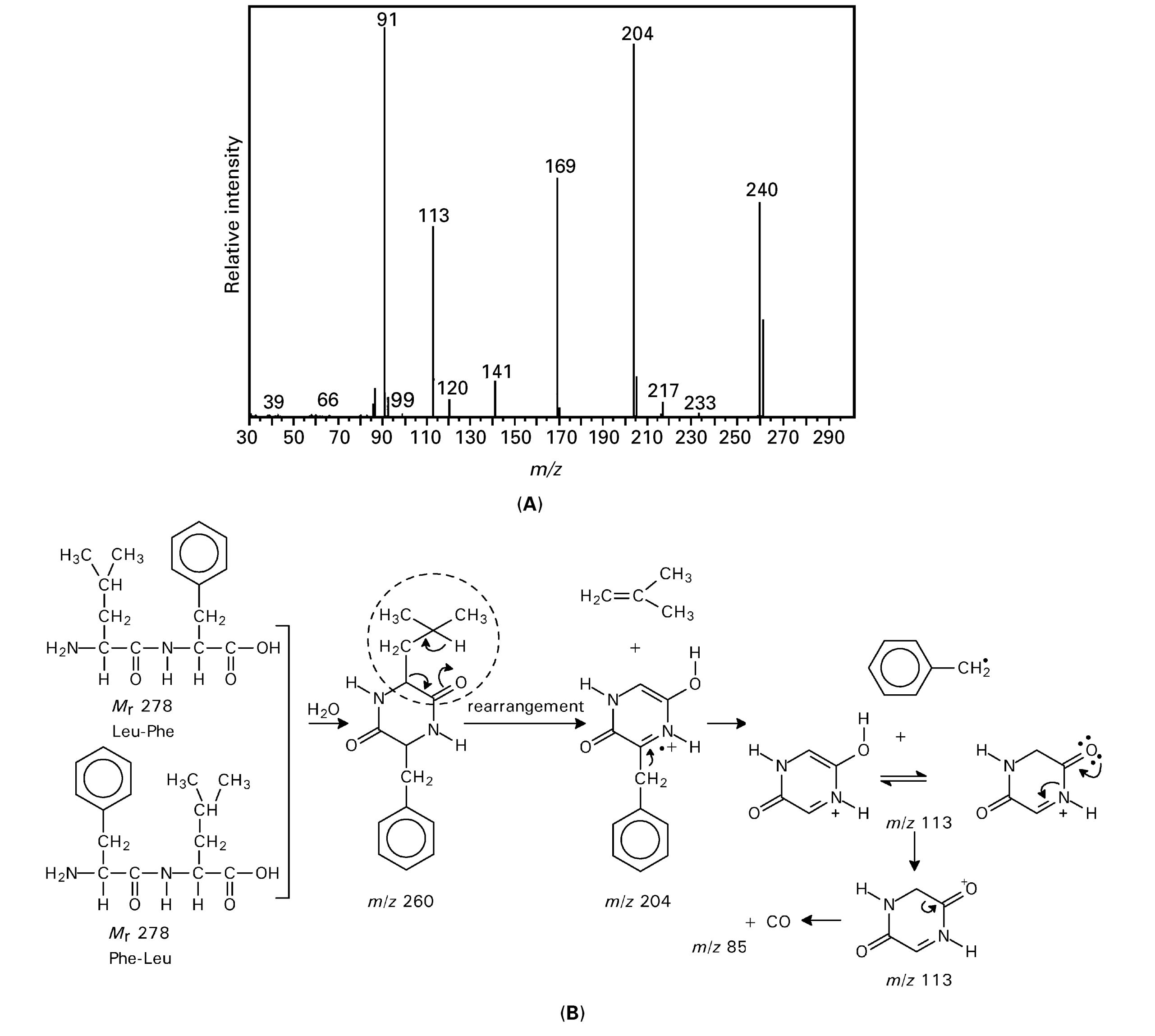



**Figure 15** Pyrolysis mass spectrum of product ions of *m/z* 260 of phenylalanyl-leucine [Phe-Leu] (A) and a general fragmentation pathway of Phe-Leu and leucyl-phenylalanine [Leu-Phe] (B). Reproduced by permission of the American Society for Mass Spectrometry from Noguerola AS, Murugaveri B, Voorhees KJ (1992), *Journal of the American Society for Mass Spectrometry* **3**: 750–756.

and resuspension of cells in sterile water. In addition, the pyrolytic derivatization approach allows one to utilize chemical derivatization for some groups of less volatile compounds to convert them into a suitable form for MS analysis.

In the 1990s the search for characteristic pyrolysis products of complex biomolecules with the use of GC-MS and MS-MS instruments has firmly established the usefulness of analytical pyrolysis techniques when used in a biochemical marker detection, rather than in a 'fingerprinting' mode. Either direct pyrolysis followed by analysis of thermal fragmentation products or, alternatively, chemical derivatization followed by analysis of the derivatized products can be used to transform nonvolatile polar and macromolecular components into characteristic chemical markers with sufficient volatility for MS analysis. The success of analytical pyrolysis was based on its abilities to detect compounds that are unique for particular groups of microorganisms.

For example, diacylpropenodiols, originating as pyrolytic products of bacterial phospholipids (PLs), under the standard conditions of EI ionization form molecular ions suitable for the evaluation of the acyl residue composition of PLs (number of carbon atoms and double bonds). In addition, the types of fatty acids involved in the composition of PLs may be determined using fragment ions of the general

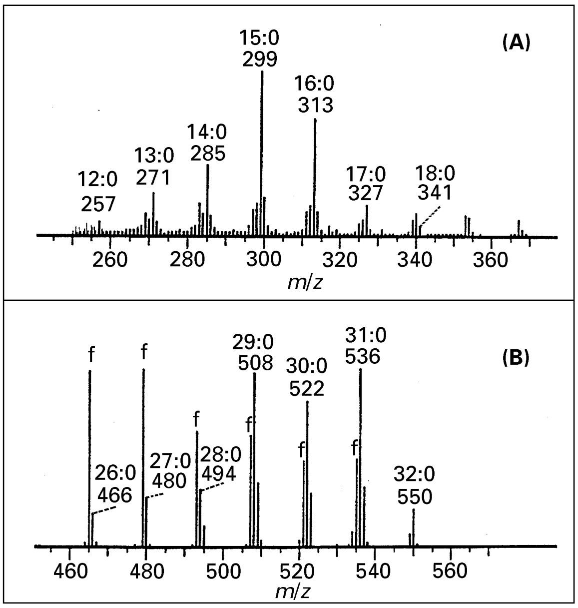



**Figure 16** Curie-point pyrolysis EI mass spectrum of *B. anthracis* showing (A) fragment ions of a series $[R–C=O + 74]^+$, representing the contribution of particular types of fatty acid residues to the overall profile of bacterial fatty acids, and (B) molecular ions of dehydrated diacylgycerols (diacylpropenodiols) originating from cellular phospholipids.

formula $[R–C=O + 74]^+$ (**Figure 16**). To perform detailed structural determinations of individual fatty acids, pyrolytic in situ hydrolysis-methylation of whole cells with TMAH was applied. This procedure, performed on a micro-scale on the surface of a pyrolytic wire, was shown to produce results which were qualitatively and quantitatively identical, or near-identical, to those from conventional methods. However, the Py-MS approach allows one to shorten the time needed to obtain this type of chemotaxonomic information by a few orders of magnitude. In addition, this method provides additional information, owing to the possibility for simultaneous analysis of other cell components, e.g. mycolic and mycocerosic acids from mycobacteria or of dipicolinic acid methyl ester from bacterial spores in addition to a plethora of pyrolytic products of proteins, sugars and nucleic acids.

### Synthetic polymers

Bond cleavages that occur as a result of thermal excitation of all vibrational modes of the polymer lead to the formation of macroradicals that undergo secondary reactions via intra- or intermolecular mechanisms. Investigations of polymer decomposition mechanisms using Py-MS have been widely used to characterize the polymer microstructure and the accumulated knowledge applied to unknown polymers and to investigations of polymer decomposition products.

Plastics and rubbers are mainly copolymers or blends of different polymers, and commercial products contain other additives, namely, plasticizers, antioxidants, pigments, cross-linking agents, flame retardants and many other components that are used to obtain desired physical properties of polymeric products. Hence, the Py-MS technique can be used as a method to quickly examine an unknown polymer sample to determine its composition, thus providing a rapid means of identifying copolymers, blends and additives in industrial rubbers and plastics. The investigation of polymer decomposition mechanisms allows for the qualitative and quantitative characterization of unknown polymers and Py-MS should be considered as a useful tool for studies of polymer micro-structure. In **Figure 17** the results of temperature-programmed pyrolysis-MS analysis of the ‘sheen’ on anti-static matting are presented – showing the capabilities of a Py-MS system to perform rapid structural investigations of polymers.

In spite of the complexity of the starting material, Py-MS allows one to analyse even microgram quantities of analytes thus providing the means for the investigation of heterogeneous products and for direct analysis, circumventing the need for dissolving or pre-separation of components. In addition, different parts of heterogeneous materials may be studied separately and pyrogram patterns may be used to compare related materials. Thus, the Py-MS technique can also be used for rapid investigation of

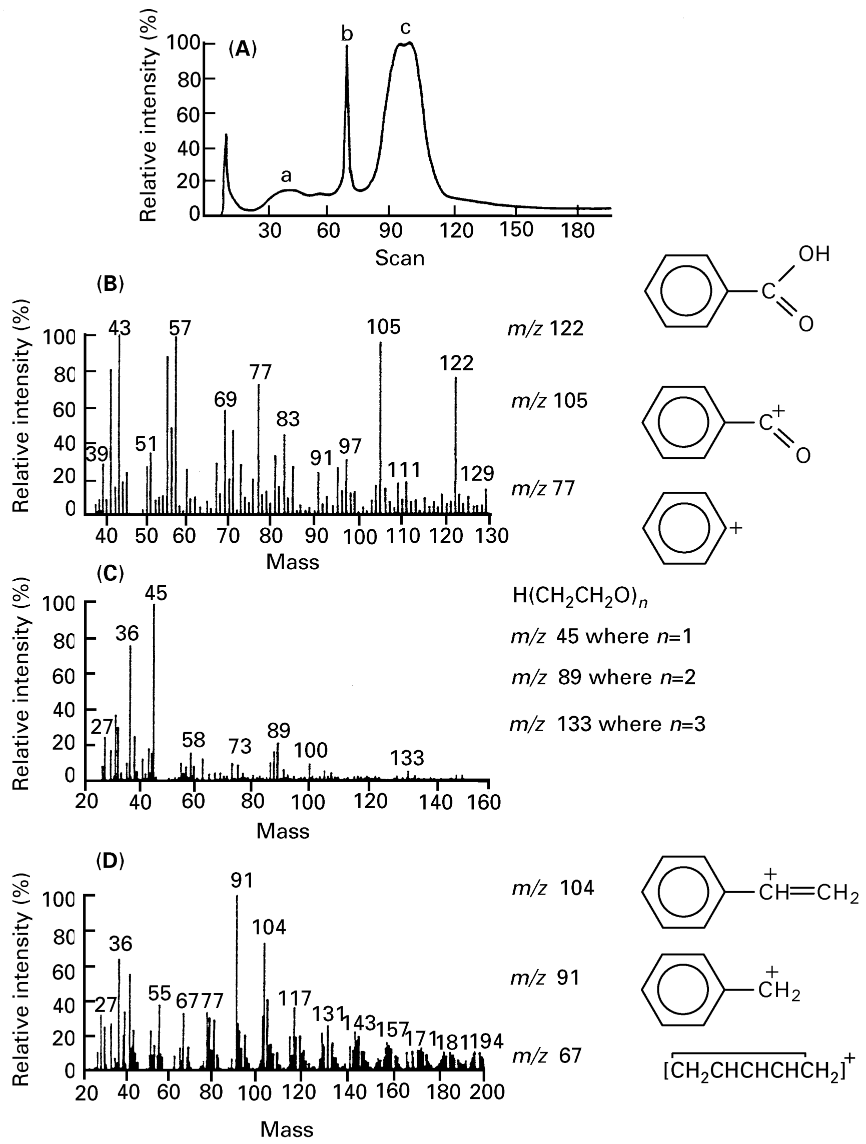



**Figure 17** Temperature-programmed pyrolysis-mass spectrometry analysis of the sheen on anti-static matting: (A) temperature-resolved total ion current signal profile of pyrolysis products obtained; (B) mass spectrum of component a; (C) component b; (D) component c. Reproduced by permission of Elsevier Science from Mundy SAJ (1993), *Journal of Analytical and Applied Pyrolysis* **25**: 317–324.

polymer sample to identify copolymers, blends, and selected additives for forensic applications.

Flash desorption of oligomers relies on kinetic competition between evaporation and thermal decomposition. Molecular ion signals from the MS have been used to determine the average molecular masses and distribution of oligomers, whereas expert systems can be used to establish mechanisms for the thermal degradation of polymers, e.g. to determine the relationships between polymer structure and the corresponding Py-MS spectra.

*See also:* **Atmospheric Pressure Ionization in Mass Spectrometry; Biochemical Applications of Mass Spectrometry; Chemical Ionization in Mass Spectrometry; Chemical Structure Information from Mass Spectrometry; Forensic Science, Applications of Mass Spectrometry; Hyphenated Techniques, Applications of in Mass Spectrometry; Laser Microprobe Mass Spectrometers; MS–MS and MS$^n$; Peptides and Proteins Studied Using Mass Spectrometry; Quadrupoles, Use of in Mass Spectrometry.**

## Further reading

Boon JJ (1992) Analytical pyrolysis mass spectrometry: new vistas opened by temperature-resolved in-source

PYMS. *International Journal of Mass Spectrometry and Ion Processes* **118/119**: 755–787.

Brown SD and Harper AM (1993) Multivariate analysis of time-resolved pyrolysis mass spectral data. In: Wilkins CL (ed) *Computer-Enhanced Analytical Spectroscopy*, Vol 4, pp 135–163. New York: Plenum Press.

Irwin WJ (1982) *Analytical Pyrolysis. A Comprehensive Guide*. New York: Marcel Dekker.

Meuzelaar HLC, Haverkamp J and Hileman FD (1982) *Pyrolysis Mass Spectrometry of Recent and Fossil Biomaterials*; *Compendium and Atlas*. Amsterdam: Elsevier Science.

Schulten H-R and Lattimer RP (1984) Applications of mass spectrometry to polymers. *Mass Spectrometry Reviews* **3**: 231–315.

Schulten H-R and Leinweber P (1996) Characterization of humic and soil particles by analytical pyrolysis and computer modeling. *Journal of Analytical and Applied Pyrolysis* **38**: 1–53.

Simmleit N and Schulten H-R (1989) Analytical pyrolysis and environmental research. *Journal of Analytical and Applied Pyrolysis* **15**: 3–28.

Simmleit N, Schulten H-R, Yun Y and Meuzelaar HLC (1992) Thermochemical analysis of U.S. Argonne premium coal samples by time-resolved pyrolysis-field ionization mass spectrometry. In: Meuzelaar HLC (ed) *Advances in Coal Spectroscopy*, pp 295–339. New York: Plenum Press.

Snyder AP (1990) Acrylic compound characterization by oxidative pyrolysis, atmospheric pressure chemical ionization-tandem mass spectrometry. *Journal of Analytical and Applied Pyrolysis* **17**: 127–141.

Voorhees KJ (ed) (1984) *Analytical Pyrolysis. Techniques and Applications*. London: Butterworths.

Voorhees KJ, Harrington PB, Street TE, Hoffman S, Durfee SL, Bonelli JE and Firnhaber CS (1990) Approaches to pyrolysis/mass spectrometry data analysis of biological materials. In: Meuzelaar HLC (ed) *Computer Enhanced Analytical Spectroscopy*, Vol 2, pp 259–275. New York: Plenum Press.

Wampler TP (ed) (1995) *Applied Pyrolysis Handbook*. New York: Marcel Dekker.

Wieten G, Meuzelaar HLC and Haverkamp J (1984) Analytical pyrolysis in clinical and pharmaceutical microbiology. In: Odham G, Larsson L and Mardh P-A (eds) *Gas Chromatography/Mass Spectrometry Applications in Microbiology*, pp 335–380. New York: Plenum Press.

# Q

# Quadrupoles, Use of in Mass Spectrometry

**PH Dawson**, Iridian Spectreal Technologies Ltd., Ottawa, Ontario, Canada
**DJ Douglas**, University of British Columbia, Vancouver, Canada

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Quadrupole mass spectrometers (often referred to as quadrupole mass filters because of the way they operate) are the most successful example of the class of mass spectrometers called 'dynamic'. Their performance depends upon a dynamic interaction of ions with time-varying electric fields. Most classes of spectrometers are 'static' and use either the interaction of fixed magnetic and electric fields or 'time-of-flight'.

The quadrupole mass spectrometer derives its name from the nature of the electric potential which is quadrupolar in the direction transverse to ion injection, i.e. dependent on the square of the distance from the centre of the field. The field is achieved by using four parallel rods as schematically illustrated in **Figure 1**.

This article begins with a brief history. It then explains the principles of operation of this mass spectrometer. The idealized view of its operation has to be tempered by consideration of some real-world situations that influence performance, such as the finite length of the field, the inevitability of fringing fields and field imperfections. Observations of performance and its limitations are illustrated. The applications section deals with residual gas analysis, gas chromatography and liquid chromatography mass spectrometry (GC-MS and LC-MS), collision induced dissociation using triple quadrupoles (MS/MS) and inductively coupled plasma mass spectrometry (ICP-MS) used for elemental analysis.

## A history of development

The possibility of using quadrupole radiofrequency fields for mass analysis was first suggested in 1953 by Paul and Steinwedel and in a US government report by Post. The first practical implementation was by Paul and co-workers in 1958. This work became the foundation for the field. Wolfgang Paul was awarded a share of the Nobel Prize for physics in 1989 for the development of the ion trap technique that also used quadrupole fields.

Early quadrupole mass filters were very limited in mass range and resolution but their physical simplicity and the absence of a magnet made them attractive for upper atmosphere and space applications. Major development occurred in the 1960s inspired by this application. This same period also coincided with a rising demand for residual gas analysers because ultrahigh vacuum technology began to have routine

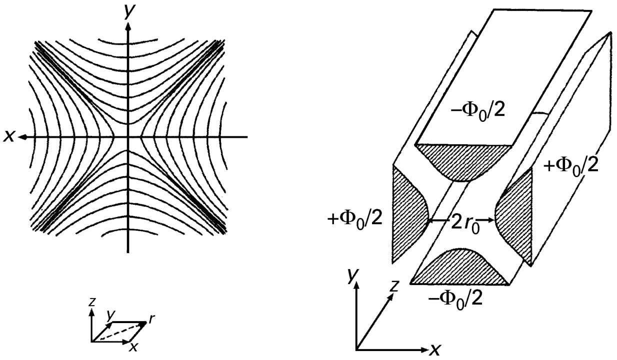



**Figure 1** A schematic illustration of a quadrupole mass spectrometer. Ions that are to be filtered to identify the presence of a particular mass are injected in the direction of the axis of the instrument. If the combination of radiofrequency and direct voltages applied to the rods is correctly chosen, only ions of one particular mass to charge ratio ($m/z$) will be successfully transmitted to the detector.

application, first in research laboratories and later in production equipment, especially in semiconductor processing. Gradually the quadrupole mass filter became the dominant instrument for this application and remains so today.

Quadrupole performance slowly improved but in the 1970s new applications to organic analysis and particularly the implementation of the combination of GC and MS placed increasing emphasis on better understanding of how the instruments worked in order to overcome their limitations. Computer simulations of performance became important. Combined with detailed experimental analysis, these led to a much improved knowledge of real-world quadrupoles with fringing fields and field imperfections. An important advance came with the application of phase space dynamics for calculating quadrupole performance. The new advances were incorporated in the classic textbook of the field written by Dawson and various collaborators and published in 1976. This book was re-issued by the American Institute of Physics in 1995 as a 'paper-back classic'.

In the 1980s and 1990s, the limits to quadrupole performance have been pushed back by precision manufacture and careful source design. A mass range of up to 2000 or more is commonly achieved with unit mass resolution. For LC-MS applications the mass range may reach 4000. There have been equally significant improvements in trace analysis capability based on a combination of sensitivity and more perfect peak shapes. High-performance quadrupole mass spectrometer manufacture has come to demand very high precision.

## Principles of operation

### The perfect field

In a perfect quadrupole mass filter field, motion in the $x$ and $y$ (transverse) directions is independent. There is no field in the axial direction and motion is unchanged along the axis. Both $x$ and $y$ motion are governed by the Mathieu equation (see the Further reading section for the derivation of these equations);

$$\mathrm{d}^2u/\mathrm{d}\xi^2 + (a_u - 2q_u \cos 2\xi)\, u = 0$$

i.e. where $u$ represents $x$ or $y$ and where

$$a_x = -a_y = 8zU/mr_o^2\omega^2, \quad q_x = -q_y = \frac{4zV}{mr_o^2\omega^2}$$

the applied voltages across the quadrupole are direct voltage $U$ and an alternating radiofrequency voltage of $V \cos \omega t$ and the minimum separation between opposite pairs of electrodes is $2r_0$. The charge to mass ratio of the ion is $z/m$. Time is expressed by $\xi = \omega t/2$, where $t$ is in seconds. For ion transmission the ions must have finite amplitude of oscillation in both $x$ and $y$ directions so that they do not strike the rods, i.e. the trajectories are both 'stable'.

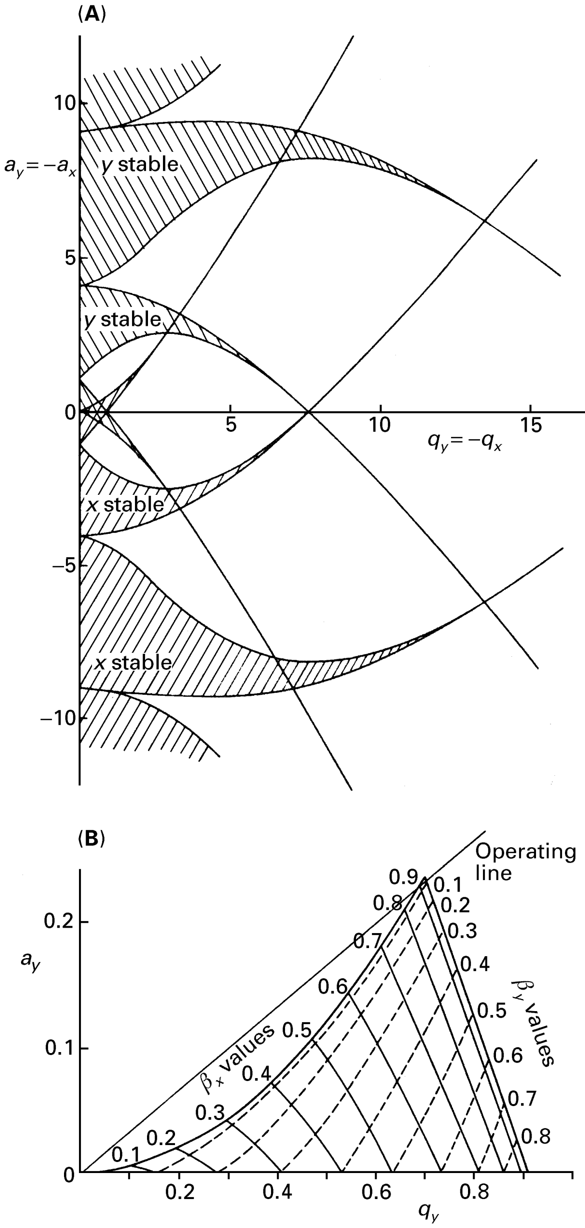



**Figure 2** Zones for stable trajectories in both $x$ and $y$ transverse directions expressed in terms of the parameters $a$ and $q$ which are related to the $m/z$ ratios of the ions. (A) General zones of stability, (B) a detail of the zone commonly used. The iso-beta lines are related to frequencies of ion oscillation.

Combinations of $a$ and $q$ values that give stable motion are shown in **Figure 2**. There are several areas of simultaneous stability. The one near the origin is commonly used but the higher zones have been examined both theoretically and experimentally especially in the search for peak shapes with more abrupt fall-off at the edges for applications in trace analysis. The quadrupole uses a sinusoidal alternating field. In principle any alternating field is possible but the higher harmonics may lead to complexity in the behaviour of the ions. The RF frequency is generally in the range of 1–2 MHz.

Mass selection is obtained by choosing a ratio of $q/a$ such that only a narrow region of the stability zone near the apex ($a = 0.23699$ and $q = 0.706$) is intersected by the operating line using $q/a$ = constant. For a given RF and DC voltage, ions of different mass to charge ratio ($m/z$) appear at different points along the line. Mass scanning is carried out by altering the values of $U$ and $V$ while maintaining their ratio constant to bring ions of different $m/z$ into the tip of the stability region. This would give a spectrum of constant resolution ($M/\Delta M$). In practice a resolution that increases with mass is preferred and the ratio $q/a$ is adjusted electronically throughout the mass scan in order to achieve more or less constant peak width. An alternative to voltage scanning would be to scan the frequency but this is rarely done because of technical difficulties in covering a large mass range.

This simple description implicitly assumes that the length of the quadrupole is infinite so that all ions with stable trajectories are differentiated from those with unstable trajectories. In practice, resolution may be limited by the length of the field. It is found that $R_{max} = n^2/h$, where $R_{max}$ is the maximum attainable resolution and $n$ is the number of RF cycles the ions spend in the field. The parameter $h$ depends upon the source and on the fringing fields. A value of 25 is not unusual; a value of 10 would be an excellent performance. High resolution is favoured for ions of low axial velocity (lower energy ions but, fortunately, also higher mass ions). One is limited, however, in reducing axial ion energies by the detrimental influence of fringing fields.

## The real world: transmission and fringing fields

In the real world, there are inevitably fringing fields at the entrance to and the exit from the mass filter. Ion motion in these fields can be very complex and the motion in each of the three coordinate directions becomes coupled. Very low energy ions may be reflected or even trapped in the fringing fields. However, a good theoretical approximation in most cases is to assume a linear increase of the $x$ and $y$ direction fields as the entrance is approached. The fringing fields extend over a distance comparable to the filter radius. The $y$-direction ($a$,$q$) values in the fringing field correspond to intrinsically unstable ion trajectories. If too much time is spent in the fringing fields, ion amplitudes will increase and the effective aperture of the spectrometer will be reduced.

The finite diameter of the field ($r_0$) means that ions are only 'accepted' for transmission when they enter the field with a small initial transverse displacement from the axis and small transverse velocities. The combination of displacement and transverse velocities that are possible defines the 'acceptance' of the instrument. The acceptance and so the overall sensitivity becomes smaller as the resolution is increased. The acceptance is calculated using phase space dynamics.

The sensitivity of the mass spectrometer is best expressed in terms of a phase space diagram as shown in **Figure 3**. This shows the initial combinations of

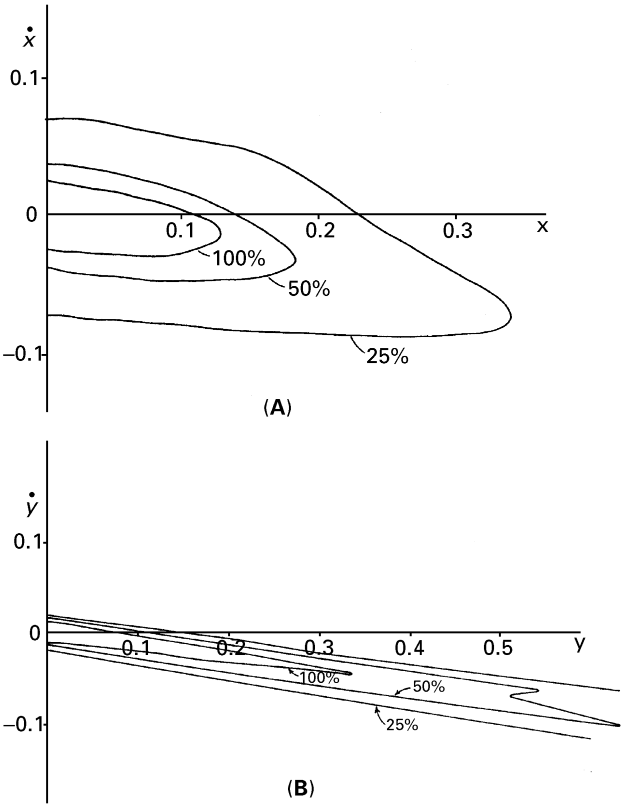



**Figure 3** Phase space acceptance diagrams showing the initial conditions of transverse displacement and velocities that lead to ion transmission at 100% and 50% of the initial phases of the RF field. (A) $x$-direction, (b) $y$-direction. This is for the centre of a peak and for ions spending two RF cycles passing through the fringing fields. The displacement from the axis is measured in units of $r_0$ and the velocity in terms of $r_0/\xi$.

transverse position and transverse velocity that will result in ion transmission. In practice, this depends upon the initial phase of the field when the ion first experiences the field. The Figure shows the $x$ and $y$ 'acceptances' for transmission in 100% of the initial phases, 50%, and so on. At higher resolutions the acceptance area decreases.

At a given resolution, combining the $x$ and $y$ acceptances together gives an indication of sensitivity for different numbers of RF cycles spent in the fringing fields giving a diagram such as that in **Figure 4**. This shows that fringing fields can be advantageous if properly chosen. There have been many attempts to tailor fringing fields for optimum performance, such as using retardation of ions after they enter the field or by adding RF only sections at the ion entry (the 'delayed DC ramp'), or by using specially shaped electrodes.

If the relative ion transmission is measured versus resolution for a particular mass number, there will be a curve such as that in **Figure 5**. If the source is evenly illuminated, at low resolution the source emittance will be less than the quadrupole acceptance and the transmission will vary little with resolution. Also the peaks will tend to be flat topped. At some point the acceptance will become limiting and transmission will tend to fall with the square of the resolution (taking fringing fields into account). Finally, at the length limitation of the quadrupole the transmission drops abruptly with resolution. Curves (a) and (b) illustrate different ion energies. In some cases, the source emittance may be rather diffuse or may not be evenly illuminated, then quite different transmission versus resolution behaviour will be observed.

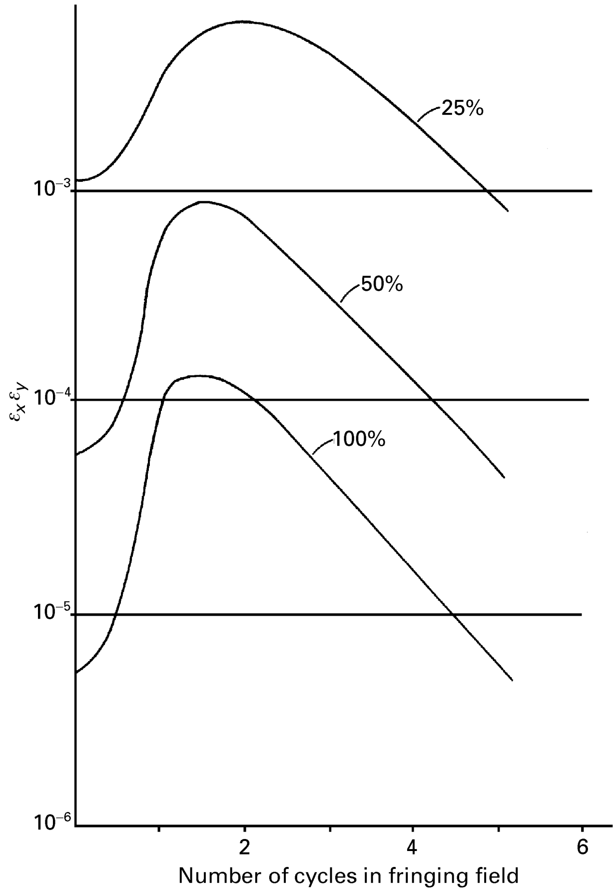



**Figure 4** Combined phase space acceptance areas for $x$ and $y$ directions as function of the length of the fringing field (expressed as the number of RF cycles that the ions spend within the fringing field). This illustrates how sensitivity may vary depending on ion velocity.

### The real world: field imperfections

There are inevitably other departures from the perfect fields and these become more and more important at high resolution. Displacement of one or more rods from the ideal position is the simplest case. This leads to higher order terms in the expression for the electric potential. One rod displaced would give predominantly third-order or hexapole correction terms. Opposite rods displaced gives fourth-order or octopole terms. These imperfections limit the resolution that is attainable, i.e. even if $n$ is increased, the resolution will not increase further. There have been various theoretical and more limited experimental studies of these effects. The field faults may also cause badly shaped or even split peaks because of nonlinear resonances in the ion oscillations at certain critical $(a, q)$ values.

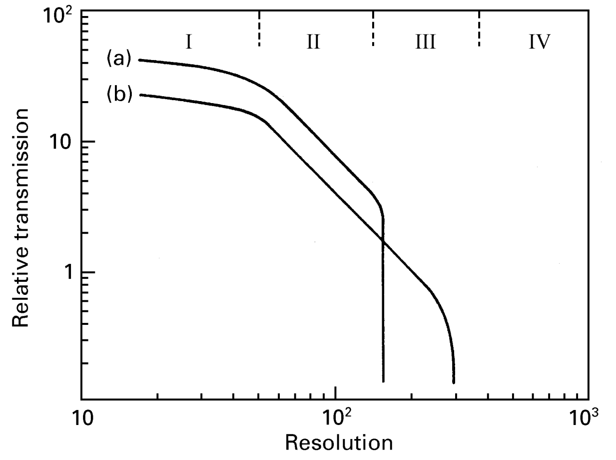



**Figure 5** A typical example of transmission efficiency versus resolution for an instrument with a well-defined source emittance. There are three regions of transmittance as resolution is increased. (I) Sensitivity is source limited, (II) sensitivity is filter acceptance limited, (III) resolution is length limited (number of cycles in the field) and (IV) resolution is limited by field imperfection.

Other field faults may arise from nonparallel rods (bending or bowing) or from errors in the electrical waveform. It is common to substitute round rods in the quadrupole for the ideal hyperbolic ones. Round rods are easier to precision-manufacture. If the diameter and positioning of the rods is correctly matched, field faults are minimized and only sixth-order distortions are produced. These are not expected to significantly influence performance. However, there is controversy over this choice of round and hyperbolic rods with many assertions being made but little solid experimental evidence. One confusing factor is that other field faults (e.g. from rod positioning) may be more serious when using round rods.

### Observations of limits to performance

**Figure 6** illustrates some of the data from an extensive examination of performance of a particular quadrupole. These data demonstrate most of the features discussed above. In addition, there is an ultimate limit to achievable resolution set by the perfection of the quadrupole field.

Similar performance data have been generated for operation using other regions of the stability diagram such as near $a = 0$, $q = 7.5$ or $a = 3$, $q = 3$. These regions may have interest for specialized application.

## Applications

### Residual gas analysis

One of the first uses of quadrupole mass filters was for residual gas analysis in high-vacuum chambers. An electron beam ionizes the background gas and a mass spectrum is recorded with the quadrupole. This allows determination of the composition of the background gas and, after calibration, the partial pressures of each of the components, such as $N_2$, $O_2$, $H_2O$, etc. Because only light gases are generally involved, the mass range of the quadrupole need only be 100–200 *m*/*z*. The compact and rugged construction of a quadrupole with purely electronic scanning make quadrupoles particularly attractive for this application. Often the quadrupole is mounted on a flange which is simply bolted onto the vacuum chamber. Sometimes a pressure reduction stage is used. Modern systems have computer controlled scanning and allow the possibility of searching libraries to match an unknown spectrum.

### GC-MS

The largest fraction of quadrupole mass spectrometers sold today are used as detectors for GC to identify

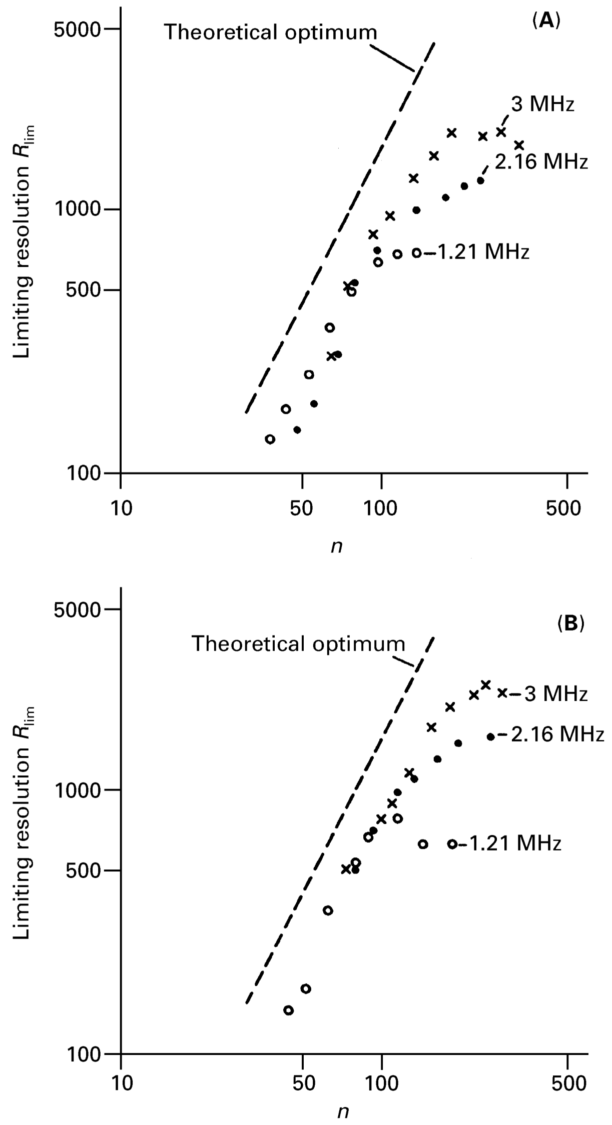



**Figure 6** Some examples of performance measurement for a particular quadrupole mass filter showing how the limiting resolution varied with the number of RF cycles in the field. The ultimate resolution reached was dependent on frequency and mass number.

and quantify trace levels of organic compounds. Environmental analysis and drug testing of athletes, for example, rely extensively on GC-MS. Organic compounds separated by a gas chromatograph elute into the ion source of a quadrupole mass filter. Ions are formed either by electron impact (EI) or chemical ionization (CI). In EI, positive molecular and fragment ions are usually formed. The resulting mass spectrum gives a fingerprint of the compound. Unknown compounds can be identified by searching a library of spectra. In CI, analytes react with a reagent ion present in excess to produce either positive or

negative molecular adduct ions, usually with minimal fragmentation. The combination of the retention time on the chromatograph and the appropriate molecular mass is often sufficient to identify trace analytes.

Quadrupole mass filters have become the standard for GC-MS because they are easily interfaced to computers, scan rapidly on a time scale compatible with peaks eluting from a GC, require only medium vacuum ($10^{-5}$ mbar), are compact, and are of comparatively low cost. Gas chromatography is restricted to relatively volatile compounds with moderate molecular masses and so the *m/z* range of a quadrupole used as a detector for GC is usually ~500–1000. A complete EI spectrum can be obtained on ~10 pg of an analyte. If only ions of one *m/z* are monitored ('single ion monitoring'), with CI, the detection limits can be lowered to low femtogram levels. Alternatively a few selected *m/z* values (say four) corresponding to the major peaks in the spectrum of a targeted compound can be monitored by switching the quadrupole between these *m/z* values without scanning intervening regions ('multiple ion monitoring'). The ability to switch a quadrupole from full scans to single ion monitoring to improve detection limits is an advantage over other methods where a complete spectrum must be acquired (TOF, ion trap ICR). A quadrupole system dedicated to GC/MS can be quite compact, often smaller than the GC itself.

## LC-MS

To separate and detect less volatile, more polar, more labile or higher molecular mass compounds, the GC is replaced by a LC. A number of ion sources have been used for LC-MS but two dominate today: atmospheric pressure chemical ionization (APCI) and electrospray ionization (ESI). In APCI, solvent and analytes eluting from the LC are sprayed into a heated tube at atmospheric pressure where they rapidly vaporize. Ions of the solvent are formed in a corona discharge. Typically in positive-ion mode these are protonated. These ions then transfer charge to analytes to produce molecular ions. In ESI the solution eluting from the LC is passed through a metal capillary that has a high voltage applied to it (3000–5000 V). Charged droplets emerge from the capillary tip at atmospheric pressure and lose solvent through evaporation, leading to the formation of gas phase ions characteristic of the ions in solution. Compounds that are present as simple ions in solution ($M^+$, $M^-$ or $MH^+$, $MH^-$) give the same ion in the gas phase. Protein ions with molecular masses 5000–100 000 acquire multiple charges to produce ions with *m/z* ratios of <4000 (**Figure** 7) that can still be analysed by a quadrupole. Molecules with intermediate molecular masses produce ions with a few charges, depending on the number of basic residues and their p*K*a values. Conventional ESI operates best at flow rates of 1–10 µL min$^{-1}$. LC flow rates are usually considerably higher (~1 mL min$^{-1}$) so the output of the LC is often split, with a fraction of the flow going to the ESI source.

APCI and ESI produce ions at atmospheric pressure. These are transferred into the vacuum system of a quadrupole mass filter with two or more stages of differential pumping. The ability to operate a quadrupole at moderate pressures of ~$10^{-5}$ torr means only modest, lower cost vacuum pumps are required. ESI and APCI generally produce protonated molecular ions. As in GC-MS, for targeted compound analysis the quadrupole is operated in 'single ion' mode to monitor the *m/z* of interest. Fragment ions can be formed by applying high electric fields to the ions in the ion sampling region. These fields accelerate the ions through the locally formed high density of gas causing collision-induced dissociation. If the system is operated in this mode a full scan over the spectrum of an analyte can be obtained. Alternatively multiple ion monitoring can be used for a few *m/z* values of interest and detection limits of a few picograms of organic compounds are possible.

A major application of LC-MS is to identify proteins. The molecular mass of the protein is

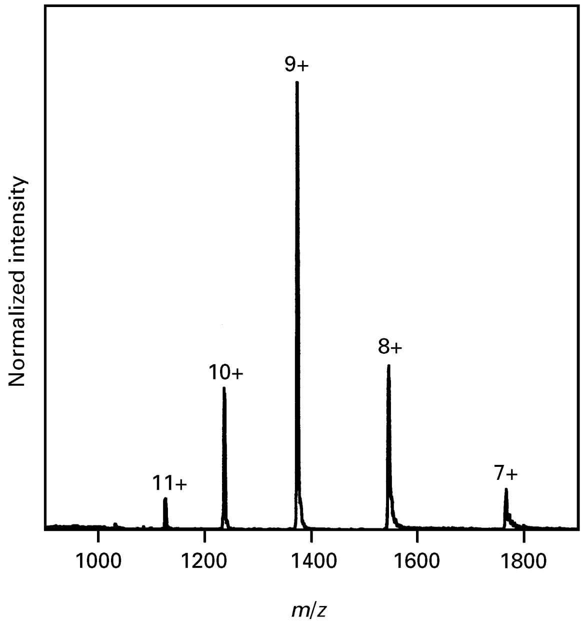



**Figure 7** The mass spectrum of the protein cytochrome c ($M_r$ 12 200) obtained with an ESI ion source and quadrupole mass filter. The isotopic structure of the peaks is not resolved. Each peak is identified by the number of protons attached to the protein.

determined by ESI-MS. A proteolytic enzyme such as trypsin is then used to cleave the protein into peptides. These are separated by LC and their molecular masses determined. These molecular masses, the molecular mass of the protein and specifying the residues cleaved by the enzyme can be used to search libraries to identify an unknown protein. If this is insufficient, some sequence information on the peptides can be obtained by tandem MS (see below).

Peptides often produce ions with two to four charges. Although quadrupoles are normally operated at unit resolution, sufficient to resolve peaks differing by one $m/z$, the resolution can be increased to separate the isotopic peaks of multiply charged ions. The spacing of these peaks allows the determination of the charge state directly (e.g. triply charged ions have isotopic peaks spaced by 0.33 $m/z$). Quadrupoles have demonstrated sufficient performance to resolve isotopic peaks of up to +4 ions in the range $m/z < 2000$. This is usually more than sufficient for peptide analysis. A higher mass range is required for LC-MS and this has pushed quadrupole performance to new limits. Current systems have an $m/z$ range of 2000–4000.

### Triple quadrupole mass spectrometers (MS/MS)

In tandem mass spectrometry (MS/MS) a first mass analyser selects an ion from a mixture, the ion is fragmented by collision, and a second mass analyser produces a spectrum of the fragment ions. MS/MS is used to determine ion structure and to detect and quantify targeted compounds in complex mixtures.

MS/MS can be carried out with a triple quadrupole system such as that shown in **Figure 8**. A first mass analysing quadrupole, Q1 mass selects a 'precursor' ion from the ESI source. The ion enters the collision cell with energies typically 10–500 eV. Here collisions with a neutral gas such as $N_2$ or Ar at a pressure $10^{-4}$ to $10^{-2}$ mbar transfer translational energy to the internal energy of the ion. It then undergoes unimolecular reaction to produce fragment or 'product' ions (collision-induced dissociation). Ions are confined to the collision cell by a quadrupole, Q2, operated with only a radiofrequency voltage between the poles. A broad range of ions with $q < 0.9$ have stable trajectories and are transmitted to the exit of the collision cell. They are then mass analysed in quadrupole Q3. The system in **Figure 8** shows an additional quadrupole Q0. This also operates in RF only mode and acts as an ion guide to transport ions from the skimmer to the first mass analyser.

Early triple quadrupoles were unable to efficiently extract higher mass ions from Q2 and at the same time to attain unit resolution in Q3 because of high fragment ion energies. A solution to this problem was found when it was recognized that by increasing the pressure in the collision cell, product ions have additional collisions with neutrals. This causes them to lose radial and axial kinetic energy, i.e. to 'cool', and to move to the centre of the RF quadrupole. They are then well within the acceptance of Q3 and the transmission increases. In addition, product ions emerge from Q2 with energies and energy spreads of only about 1 eV. With a modern triple quadrupole system, at least 50% of the precursor ions that are transmitted by Q1 can be converted to fragment ions that are transmitted through Q3 with unit resolution or better. Resolution of the isotopic peaks of up to +4 ions has been demonstrated (**Figure 9**). The high collision pressure also minimizes any ion focusing effects which could lead to variable transmission. Hexapole or octopole fields have also been used to confine ions in the collision region.

A triple quadrupole has many scan modes. In 'product ion' scans, Q1 mass-selects an ion from a

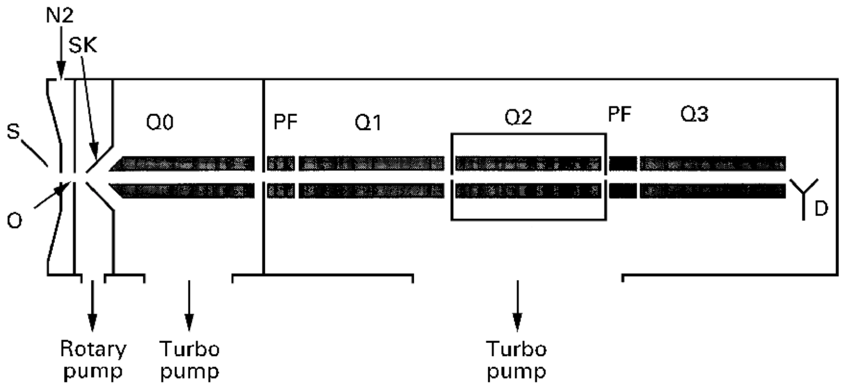



**Figure 8** A triple quadrupole mass spectrometer system with an electrospray ion source. S, electrospray source; N2, nitrogen curtain gas; O, ion sampling orifice; SK, skimmer; Q0, RF-only quadrupole ion guide; PF, delayed DC ramp 'prefilters'; Q1 mass analysing quadrupole; Q2, RF quadrupole enclosed in a collision cell; Q3 mass analysing quadrupole; D, ion detector.

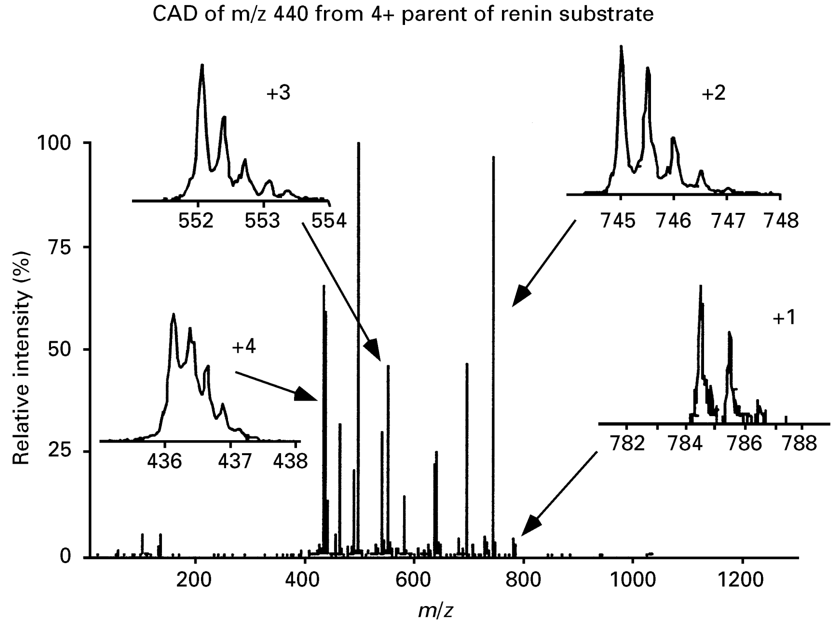



**Figure 9** Mass spectra of multiply charged fragment ions of the peptide renin substrate tetradecapeptide. The precursor was the $(M+4H^+)^{4+}$ ion at *m/z* 440. The insets show the resolution of the isotopic peaks for the +1 to +4 ions. Reproduced with permission of The American Chemical Society from Thomson BA, Douglas DJ, Corr JJ, Hager JW, Jolliffe CL (1995) *Analytical Chemistry* **67**: 1696–1704.

mixture, it is fragmented in Q2 and a mass spectrum of product ions is obtained by scanning Q3. This scan mode is useful to obtain structural information of an ion such as sequence information of a peptide. In 'precursor ion' scans, Q3 is fixed on a particular *m/z* value and Q1 scans through all the ions produced by the source. This scan is useful to identify those ions in the source that contain a particular functional group. In 'neutral loss' scans, Q1 and Q3 scan together with a constant difference in *m/z* that corresponds to loss of a given neutral group. For example, if the loss corresponds to $Cl_2$ (70 amu) this scan could identify ions that contain two or more chlorine atoms such as polychlorinated dioxins. For targeted compound analysis, the system can be run in 'single reaction monitoring' mode. Here Q1 is fixed at the *m/z* of the precursor ion of a targeted compound and Q3 is set to a *m/z* value of a major fragment ion of that compound. There is a massive discrimination against other compounds. If still greater selectivity is required, multiple reaction monitoring can be done. Here Q1 is set to the *m/z* value of the precursor and Q3 'peak hops' to several (say four) *m/z* values of fragments that come from the targeted compound. If the intensity ratios of the fragments are correct the compound is identified. If there is an interference on one of the fragments, the remaining fragments can be used to identify and quantify the compound. The ability to independently scan Q1 or Q3 under computer control makes a triple quadrupole MS/MS system a flexible tool for trace analysis. It has become the workhorse for LC-MS/MS and GC-MS/MS analysis.

An example of the use of a triple quadrupole MS/MS system to identify a protein is given here. The enzyme telomerase rebuilds the ends of chromosomes (telomeres) when cells divide. It consists of one RNA and two protein subunits of molecular mass 43 kDa and 123 kDa. The 123 kDa protein was separated on a 2D gel, extracted and digested to produce a mixture of peptides. Q1 of a triple quadrupole was scanned to produce a spectrum of all the peptides, shown in **Figure 10A**. To identify these peptides, Q1 was set to transmit a 2 *m/z* mass window and tandem mass spectra were collected with a 0.2 *m/z* step size. The fragment ion spectrum of a doubly charged peptide at *m/z* 830.4 is shown in **Figure 10B**. The doubly charged peptide ion fragments to singly charged ions to give fragments with *m/z* greater than the precursor. This mass spectrum along with esterification of the peptide allowed unambiguous assignment of the amino acid sequence. The sequences of eight of the peptides in the mass spectrum were determined. These amino acid sequences were then used to make DNA probes that led to identifying the gene containing the complete sequence of the protein. The total amount of

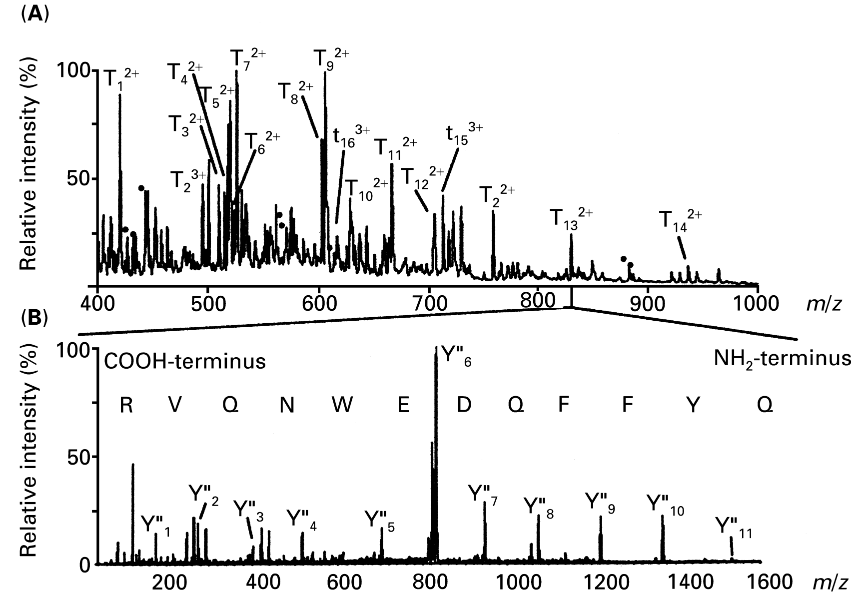



**Figure 10** (A) Mass spectrum of the peptides obtained from digesting the 123 kDa subunit of telomerase. Peptides were not separated by chromatography. Peptides that were sequenced fully or partially are marked by T or t, respectively. (B) Tandem mass spectrum of the peptide at *m/z* 830.4. Reproduced with permission of The American Association for the Advancement of Science from Lingner J, Hughes TR, Shevchenko A, Mann M, Lundbald V, and Cech TR (1997) *Science* **276**: 561–567.

protein available for this experiment was in the low picomole range.

## ICP-MS

Another major application area for quadrupoles is in ICP-MS systems for trace element analysis. A schematic of a system is shown in **Figure 11**. This ion source is an induction plasma in argon at atmospheric pressure with a temperature of 5000–7000 K contained in a torch. Samples are introduced to the plasma as aerosols, usually solutions that are sprayed. At the high plasma temperature dissolved solutes are vaporized, atomized and ionized. Most elements of the Periodic Table are present in the plasma as singly charged atomic ions (the degree of ionization is typically 90% or more). The plasma expands through an orifice about 1 mm in diameter into a region at a pressure of a few mbar, and the centreline flow then passes through a skimmer into a region at a pressure of about $10^{-4}$ torr. In this region ions are extracted from the rarefied plasma, pass through ion lenses and then are mass-analysed in a quadrupole.

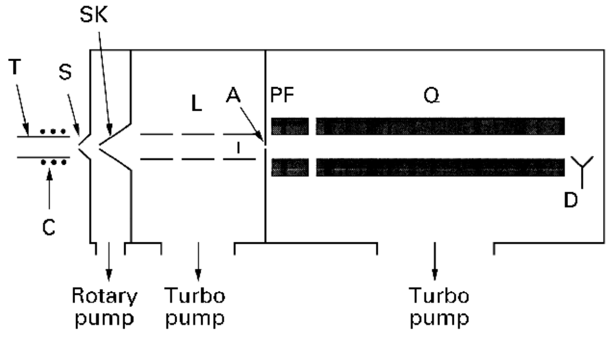



**Figure 11** A quadrupole ICP-MS system. T, torch; S, sampler; SK, skimmer; L, ion lenses; A, differential pumping aperture and lens; PF, delayed DC ramp 'prefilter'; Q, quadrupole mass filter; D, ion detector.

ICP-MS with a quadrupole gives simple mass spectra that are easy to interpret. **Figure 12** for example shows the mass spectrum at unit resolution of some transition metals at a concentration of 100 ng $mL^{-1}$. By scanning a quadrupole over the range $^{7}Li^{+}$ to $^{238}U^{+}$, over 70 elements can be determined in a 1 min scan. Alternatively the quadrupole can peak hop to selected isotopes or elements, to improve the duty cycle. Detection limits are typically in the 10 pg $mL^{-1}$ region for elements in solution. Isotopic information is inherent in ICP-MS. With a quadrupole the precision on isotope ratios is typically 0.2% with a measurement time of a few minutes. This is insufficient for many geological dating applications but is more than adequate for many isotopic tracing experiments in nutrition and other studies. In addition, it greatly facilitates quantification of trace elements by isotope dilution.

Ideally the ICP would produce only singly charged atomic ions of each element. However, it also

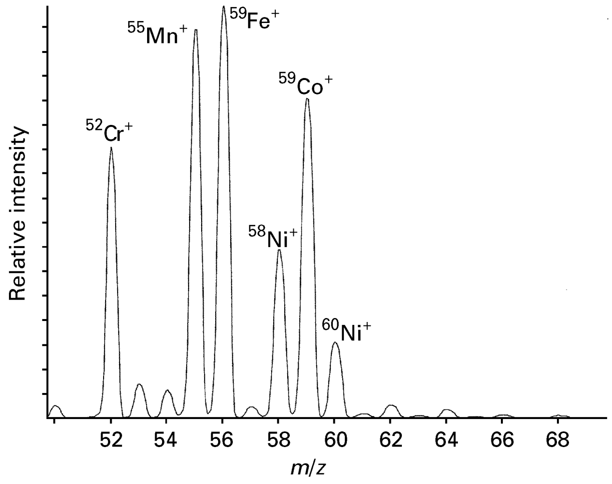



**Figure 12** Mass spectrum of transition metals each present at 100 ng ml$^{-1}$ in solution. The peak at *m*/*z* 56 also includes a contribution from $^{40}Ar^{16}O^+$. Small peaks at *m*/*z* 63–68 indicate contamination by Cu and Zn at concentrations of few ng mL$^{-1}$.

produces some molecular ions such as $ArO^+$ which interfere with $^{56}Fe^+$, or $Ar_2^+$ which interferes with $^{80}Se^+$. Such interferences are most common at *m*/*z* < 80. To separate these interferences requires a resolution that is beyond the capabilities of quadrupoles operated conventionally, although the use of alternative stability regions is being investigated. The interferences are not prohibitive because in many cases an alternative isotope can be found that is free of interference. As with many applications it is the comparatively low cost and electronic control of quadrupoles that make them attractive for ICP-MS.

## List of symbols

$a = 8zU/m\omega^2 r_o^2$; $f$ = RF Frequency; $m$ = mass of ion; $n$ = number of RF cycles; $q = 4zU/m\omega^2 r_o^2$; $r_o$ = separation between electrodes; $R_{lim}$ = limiting resolution; $R_{max}$ = maximum resolution; $t$ = time; $u = x$ or $y$ direction; $U$ = direct voltage; $V$ = voltage; $z$ = charge of ion; $\xi = \omega t/z$; $\omega = z\pi f$.

*See also:* **Atmospheric Pressure Ionization in Mass Spectrometry; Biomedical Applications of Atomic Spectroscopy; Chemical Ionisation in Mass Spectrometry; Chromatography-MS, Methods; Hyphenated Techniques, Applications of in Mass Spectrometry; Inductively Coupled Plasma Mass Spectrometry, Methods; Mass Spectrometry, Historical Perspective; MS-MS and MS*ⁿ*; Photoacoustic Spectroscopy, Applications.**

## Further reading


Bruins AP (1994) Atmospheric pressure ionization mass spectrometry. *Trends in Analytical Chemistry* **13**: 37–43; 81–90.

Busch KL, Glish GL and MacLuckey SA (1998) *Mass Spectrometry/Mass Spectrometry: Techniques and Applications of Tandem Mass Spectrometry*. Weinheim: VCH.

Cole RB (ed) (1997) *Electrospray Ionization Mass Spectrometry*. New York: Wiley.

Dawson PH (ed) (1976) *Quadrupole Mass Spectrometry and its Applications*. Amsterdam: Elsevier: (Reissued as a paperback Dawson PH (ed) *Quadrupole Mass Spectrometry and its Applications* (1995) New York: American Institute of Physics Press.

Dawson PH (1980) Mass filter design and performance. *Advances in Electronics and Electron Physics* **53**: 153.

Dawson PH and Bingqi Yu (1984) Performance comparison in conventional and higher stability regions. *International Journal of Mass Spectrometry and Ion Processes* **56**: 41.

Douglas DJ and Ying J-F (1996) High resolution ICP mass spectra with a quadrupole mass filter. *Rapid Communications in Mass Spectrometry* **10**: 649–652.

Du Z, Olney TH and Douglas DJ (1997) Inductively coupled plasma mass spectrometry with a quadrupole operated in the third stability region. *Journal of the American Society for Mass Spectrometry* **8**: 1230–1236.

Heumann K (1982) Isotope dilution mass spectrometry for micro- and trace-element determination. *Trends in Analytical Chemistry* **1**: 357–361.

Houk RS (1994) Elemental and isotopic analysis by inductively coupled plasma mass spectrometry. *Accounts of Chemical Research* **27**: 333–339.

Paul W, Reinard HP and von Zahn U (1958) *Zeitschrift für Physik* **152**: 143–182.

Yost RA and Enke CG (1979) Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Analytical Chemistry* **51**: 1251A–1264A.

# Quantitative Analysis

**T Frost**, Glaxo Wellcome, Dartford, UK

Copyright © 1999 Academic Press

**FUNDAMENTALS OF SPECTROSCOPY**
**Methods & Instrumentation**

Spectroscopic techniques are particularly useful for quantitation. They offer speed and great flexibility in instrumentation. UV spectrophotometry is widely used for quantitation using data at the maximum absorbance of a chromophor. Fluorescence spectrometry is also widely used for quantitation as it provides greater selectivity and sensitivity than UV spectrophotometry. The use of FT-IR spectrometry for quantitation is less common. The use of NIR for quantitation is an area of great activity and innovation. In addition, NMR spectroscopy can be used for quantitation. Here the criteria for success are very different from those for absorption or fluorescence spectroscopy. NMR quantitation is the subject of a separate article in this encyclopedia.

Quantitation can be carried out using either absorption or reflectance measurements and the laws governing these two types of measurements will be discussed first.

## Absorption of light

The absorption of light by a compound depends on its chromophor, the wavelength of the light and the thickness of the sample. Bouguer derived the relationship between absorption and the thickness of the sample. The integrated form of the equation is shown in Equation [1] where $I_0$ is the intensity of the incident radiation and $I$ the intensity of the transmitted radiation. The factor $a$ related to the absorptivity of the chromophor and $b$ is a measure of the sample thickness.

$$\log(I_0/I) = ab \quad [1]$$

Beer's law states that the absorption of monochromatic radiation by a sample is proportional to the concentration of the sample. The law is defined by Equation [2] where $a'$ is a constant and $c$ is the concentration.

$$\log(I_0/I) = a'c \quad [2]$$

Equations [1] and [2] are combined to give the Bouguer–Beer law shown in Equation [3] where $\varepsilon$ is the molar absorptivity, i.e. the absorbance of a one molar solution of the compound. Note that $\varepsilon$ is specific for each compound at particular wavelength.

$$\log(I_0/I) = \varepsilon bc \quad [3]$$

The term $\log(I_0/I)$ is the absorbance, $A$, of the sample and Equation [3] is presented more simply in Equation [4], often referred to as Beer's law.

$$A = \varepsilon bc \quad [4]$$

Spectrophotometers measure the ratio of the intensity of the incident and transmitted radiation, which is known as the transmittance, $T$.

$$T = I/I_o \quad [5]$$

Absorbance is then related to transmittance as shown in Equation [6].

$$A = -\log T \quad [6]$$

Logarithms to the base 10 are usually used.

## Reflection of radiation

The use of reflection techniques is becoming common for quantitation. Taking measurements from the surface of a sample avoids the need for sample preparation and thus provides a very rapid method of analysis. One drawback of the use of reflectance measurements is that it is necessary to assume that the surface is representative of the sample as a whole. Reflectance measurements are commonly used in the NIR and FT-IR regions.

In reflectance measurements, $\log(1/R)$, where $R$ is the reflectance of the sample, is proportional to concentration. The proportionality constant is not as universal as in absorption. The constant will depend

on factors such as the particle size of the sample and the moisture. The constant is thus unique for each sample and this makes quantitation using reflectance techniques very challenging.

### Fluorescence

Fluorescence quantitation can be described by Equation [7]

$$F = (I_o)(\Phi_f)(1 - e^{\varepsilon bc}) \qquad [7]$$

where $F$ is the fluorescence intensity of the sample and $I_o$ is the intensity of the incident light and $\Phi_f$ is the fluorescence quantum yield. The quantity $1–e^{\varepsilon bc}$ derives from Beer's law, Equation [4], and relates to the fraction of light absorbed by the fluorescing species. Equation [7] can be expanded to Equation [8]

$$F = I_o\Phi_f\left(1 - \left[1 - \varepsilon bc + \frac{(-\varepsilon bc)^2}{2} - \frac{(-\varepsilon bc)^3}{3} + \cdots\right]\right) \qquad [8]$$

At low concentrations with absorbance less that ~0.05, Equation [8] can be reduced to Equation [9] and fluorescence is linearly related to concentration.

$$F = I_o\Phi_f\,\varepsilon bc \qquad [9]$$

The approximations made to derive this equation must be remembered and at high concentrations fluorescence will not vary linearly with concentration and a curved calibration graph will be obtained.

## Quantitative techniques

### Data at a single wavelength

The simplest method of quantitation is to use data from a single wavelength. Ideally measurements should be taken at the wavelength of an absorption maximum because this avoids errors due to differences in wavelength calibration. This method of quantitation is usually used for absorption measurements rather than reflectance techniques which tend to require data from more than one wavelength. The absorbance of a solution of known concentration, the standard, is measured and compared to the absorbance of the sample whose concentration needs to be determined. Thus, if $A_{sam}$ is the absorbance of the sample and $A_{std}$ is the absorbance of the standard and $C_{std}$ is the concentration of the standard the concentration of the sample, $C_{sam}$ is given in Equation [10].

$$C_{sam} = A_{sam} \times C_{std}/A_{std} \qquad [10]$$

Two standards should be prepared and the responses compared to ensure they have been prepared correctly. Typically, the two responses should agree within 1 per cent. The mean response from the two standards can then be used to predict the concentration of the samples.

If the concentration of the samples varies over a wide range, a calibration curve, constructed from standards of varying concentration, can be used. The responses of each standard should be checked to ensure that they agree and the regression coefficient of the line should be close to 1.

Often the molar absorptivity of the compound under investigation is so well characterized that the use of a standard is not necessary and published values of the absorbance of a 1 per cent solution in a 1 cm cell, $A_1^1$ can be used. The use of $A_1^1$ values is particularly common in the pharmaceutical industry.

The drawback of using data at one wavelength is that the wavelength may not be specific for the compound of interest. The sample may contain other compounds that absorb at the same wavelength and interfere with the measurement. Thus quantitation using data at a single wavelength is limited to solutions of simple compounds. The technique is most commonly used in the ultraviolet and visible regions of the spectrum.

### Derivative spectroscopy

Derivative spectroscopy is useful in quantitative analysis both as a quantitative technique in its own right and for preprocessing data prior to analysis by some of the chemometric techniques described below.

A common problem in spectroscopic quantitation is that the band may overlap with a broad interfering band. The interfering band may arise from the sample matrix and give a sloping baseline making quantitation difficult. This is illustrated in **Figure 1** where the left-hand spectrum shows a band on a sloping baseline. The first derivative removes the sloping baseline provided it is linear. First derivative spectra are not always ideal for quantitation and often a second derivative spectrum, shown on the right of **Figure 1**, is used. The second derivative will remove any baseline curvature provided it can be fitted to a quadratic equation with respect to wavelength.

A first derivative spectrum is found by differentiating the absorbance spectrum and shows the rate of

change of the slope of the absorbance spectrum. The derivative spectrum is found by calculating $\mathrm{d}A/\mathrm{d}\lambda$, where $A$ is the absorbance and $\lambda$ is the wavelength.

Derivative spectra are normally calculated off-line on the digital spectral data. The use of the Savitzky–Golay routine for calculating smoothed derivatives is common but other methods are available. Instrumental parameters need to be optimized to obtain successful derivatives. In particular, the data interval used to record the spectrum, the scan speed and the interval used to calculate the derivative need to be optimized.

The spectrum must first be scanned using a suitable data interval. If the data interval is too large then vital high frequency features will be smoothed out when the spectrum is recorded and the power of derivative spectroscopy will be lost. For derivative work it is usually best to choose a relatively small data interval to scan the spectrum. Any noise recorded by the small interval can be smoothed in the calculation of the derivative by varying the derivative interval. The interval used for the calculation of the derivative will influence the spectrum. Some trial and error may be necessary to find the correct interval.

**Derivatives for quantitation** When Beer's law, Equation [4], is differentiated with respect to wavelength only $\varepsilon$, the molar absorptivity, will vary with wavelength. The concentration $C$ and the path length $b$ are constant. The first derivative is given in Equation [11].

$$\frac{\mathrm{d}A}{\mathrm{d}\lambda} = \frac{\mathrm{d}\varepsilon}{\mathrm{d}\lambda} Cb \qquad [11]$$

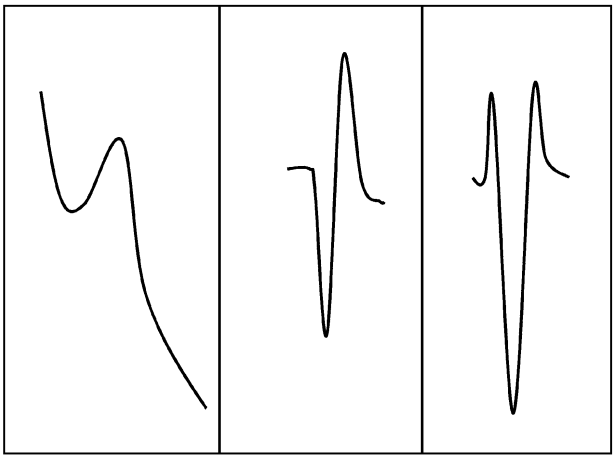


**Figure 1** Left: A typical absorption band; Centre: The first derivative of the absorption band; Right: The second derivative of the absorption band.

The derivative amplitude, $\mathrm{d}A/\mathrm{d}\lambda$, will be proportional to concentration and can be used for quantitation. Differentiating the first derivative spectrum forms derivative spectra of higher order. The second and fourth order derivatives are generally used for quantitation. **Figure 1** shows first and second derivative spectra.

Derivative spectroscopy can be used for analysis of mixtures when two components have different bandwidths. The technique is particularly useful when one compound has a spectrum with sharp features and another component has broad features. The sharp band can be emphasised at the expense of the broad band.

The use of derivative spectroscopy for direct quantitation is limited to compounds where the spectra show major differences. Quantitation can be carried out using data at single wavelengths or the whole spectrum can be used in some of the chemometric techniques described later.

## Chemometric techniques

Chemometric techniques have found widespread use in spectroscopic quantitation. The techniques are used when a single wavelength, that is specific for the compound of interest, cannot be found. Wavelengths have to be found where the species contributing to the spectrum have different absorption or reflectance. Mathematical approaches can then be used to unravel the contribution to the spectrum of the compounds of interest and hence deduce their contribution.

In a mixture of compounds the absorbance at a particular wavelength can be considered to arise from a linear combination of the absorbances from the individual components. If two components contribute to the spectrum we can expand Beer's law to reflect the individual contribution of each component as shown in Equation [12], where the subscripts 1 and 2 refer to the two components.

$$A = \varepsilon_1 b c_1 + \varepsilon_2 b c_2 \qquad [12]$$

If we know the molar absorptivities, $\varepsilon_1$ and $\varepsilon_2$, for each component and the path length, $b$, Equation [12] still cannot be solved for the concentrations, $c_1$ and $c_2$, because there are too many variables. To solve for the concentrations of the two components it is necessary to have data at two wavelengths where the molar absorptivities of the two components are different both from each other at both wavelengths. Two simultaneous equations can be constructed and solved for the two concentrations. If more than two components contribute to the spectrum it is necessary

to find extra wavelengths for each component where the molar absorptivities differ. Analogous equations can be written for reflectance measurements.

Rather than use simultaneous equations it is usual to resort to regression techniques. These fall into three main techniques multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS). All three techniques have found widespread use in spectroscopic quantitation using both absorption and reflectance techniques.

Multiple linear regression requires careful choice of wavelengths to find a quantitation model that is robust. PCR and PLS overcome problems of wavelength selection by using the full spectrum. PLS is often considered to give superior results to PCR. Full discussions of the algorithms and details of how to apply the techniques can be found in the monographs on chemometrics listed under Further reading.

## Multiple linear regression

Multiple linear regression can be used to solve for the constants in Equation [13], which can be described as a general equation for any *n* components as shown in Equation [13].

$$A = \varepsilon_1 bc_1 + \varepsilon_2 bc_2 + \varepsilon_1 bc_1 + \varepsilon_3 bc_3 + \cdots + \varepsilon_n bc_n \quad [13]$$

Matrix algebra is used to simplify the mathematics and Equation [13] is described in matrix terms in Equation [14]. This is a general equation for any spectroscopic data, be it absorption, reflection or derivative data.

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e} \quad [14]$$

The vector, **y**, is a column vector of spectroscopic data for each sample at one wavelength. The matrix **X** contains the concentrations of the samples. If an intercept is included in the model, then the first column of **X** must be a column of 1s. The vector **e** is a column of residuals associated with lack of fit in the model. The vector **e** contains the errors that are minimized in the regression analysis.

The matrix solution for **b** in Equation [14] is given in Equation [15].

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad [15]$$

where $\mathbf{X}'$ is the transpose of $\mathbf{X}$ and $(\mathbf{X}'\mathbf{X})^{-1}$ is the inverse of the covariance matrix $\mathbf{X}'\mathbf{X}$. The inversion of the covariance matrix is not always possible and this is the main drawback to MLR.

The practical steps involved in MLR are as follows. A set of calibration samples is selected and the concentration of the component of interest is determined by a reference technique. The spectroscopic data and the concentrations from the reference technique are fed into Equation [15] to give an equation that can be used to predict the concentration of the component in further samples.

MLR can result in simple equations that can be used for quantitation. For example, the constituents of food, such as moisture and fat, can be determined using spectroscopic data at a few NIR wavelengths. The disadvantages of MLR are that it requires the operator to select the wavelengths. The selection of inappropriate wavelengths can result in poor models that are mathematically unstable.

## PCR and PLS

PCR and PLS were developed to overcome the limitations of MLR. They use all the spectral data and so avoid the need for wavelength selection. PCR is essentially a mathematically more robust way of carrying out MLR. The regression is performed on the principal components of the data set. The principal components are determined from the data set with the specific aim that they will provide robust models. The principal components are linear combinations of the original measurements such that the first component explains the most variance in the data and subsequent components, all orthogonal, explain decreasing amounts of data variance.

The data set is broken down into loadings, which are analogous to abstract spectra, and scores, which describe the amount of each loading associated with each sample and are similar to abstract concentrations. Some of the loading spectra will describe noise or minor components and will not be relevant to the model and can be rejected. Determining which principal components can be rejected affects the robustness of the model.

A set of calibration samples is identified and the concentration of the components of interest is determined by a reference technique. The spectra are assembled into a data matrix often known as the calibration matrix. The data matrix is decomposed into the product of a score and a loading matrix using principal component analysis. Irrelevant principal components describing noise are rejected, typically those components of higher index.

Multiple linear regression is then used to relate the scores matrix to the matrix of concentrations determined by the reference technique.

The model is then complete and can be used to predict the concentration of the components of samples. The calibration set can be used many times provided the samples in the set have been chosen carefully. A good calibration set of samples will contain samples that arise from every source of variation likely to be encountered in further samples that the method will be used to quantify.

PCR takes no account of the concentration of the samples in the calibration set when finding the principal components and this has been considered a drawback. PLS was developed to use the concentration data when developing the mathematical model. PLS is thus claimed to be superior to PCR.

The practical application of PCR and PLS is very similar. Samples are selected for the calibration set and a reference technique used to determine the concentration of the component or components of interest. The spectra and concentrations then allow a calibration model to be set up for predictions of further concentrations.

Pretreatment of the spectroscopic data is usually carried out to improve the model. A first derivative can be applied to remove sloping backgrounds and offset errors. This is a very common method of pretreating NIR spectra. The data can be centred about its mean. Routines are available to smooth the data and remove spectroscopic scatter.

Once a set of calibration spectra have been set up the use of data pretreatment can be examined to find the optimum pretreatment. Similarly PCR and PLS can be compared to see which gives the best results.

## Validation

Whatever method is chosen to perform the quantitation, whether a simple single wavelength assay or a complex chemometric method, the method should be validated to ensure it produces meaningful data. Before the method can be validated, the exact procedure should be documented so that the final method is used in the validation experiments.

Validation should look at the following parameters:

- Linearity
- Specificity
- Accuracy
- Precision
- Robustness

### Linearity

The linearity of the method should be checked over and beyond the likely operating range of the method. Typically six concentrations of the standard solution over the range 25 to 150% of the normal working concentration should examined. A graph of absorbance against concentration can then be constructed. The intercept of the line should be examined to see how close to the origin the line crosses the *X* axis. A large intercept will indicate an offset error in the method or nonlinearity.

### Specificity

The specificity of the method can be checked by running a sample without the compound of interest, if one is available. The method can also be compared with other methods of known specificity. The acceptable limit of interference from the sample matrix will vary with the sample and the use of the results from the analysis. Typically, an interference of < 2% would be expected but higher values can often be tolerated.

### Accuracy

The accuracy of the method is a measure of how close to the true value the results from the method are. Accuracy is usually checked by preparing samples spiked with known amounts of the component of interest. The accuracy should be examined over a range that extends beyond the range of samples the method is likely to analyse. If the method is designed to measure samples at 100% of expected strength then it would be useful to check the accuracy at that value and at 80% and 120% of the expected strength. Results from the accuracy experiments should normally agree to within 2% of the true value.

When samples cannot be spiked with the component of interest then comparisons with a reference technique should be carried out.

### Precision

Precision is a measure of the spread of results from a method. Precision can be broken down into subsections. The first measurement of precision relates to how repeatable the spectroscopic measurements are. Taking ten readings and looking at the coefficient of variation can assess the repeatability of the spectroscopic measurement. Normally the coefficient of variation should be < 2%.

The repeatability of the sample and standard preparation can be assessed by preparing ten samples and standards from the sample source and examining the coefficient of variation of the response.

Depending on where the method is used the precision of different operators, equipment and laboratories should also be examined. Interlaboratory precision can be determined by having different laboratories analyse the same sample.

### Robustness

The robustness of the method should be examined by stressing the method parameters. Thus the effects of slight changes in wavelength, spectroscopic settings, sample preparation, etc. can be varied to see if any of the values are critical. The most effective way of doing this is to use experimental design so that parameters are varied in a planned way and statistics and response surfaces can be used to evaluate the outcome of the results.

## List of symbols

$a$ = constant; $A$ = absorbance; $A$ = absorptivity; $b$ = sample thickness, path length; $c$, $C$ = concentration; $F$ = fluorescence intensity; $I$ = intensity of transmitted radiation; $I_o$ = intensity of incident radiation; $R$ = reflectance; $T$ = transmittance; $\varepsilon$ = molar absorptivity; $\lambda$ = wavelength; $\Phi_f$ = fluorescence quantum yield.

*See also:* **Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy; Fourier Transformation and Sampling Theory; Multivariate Statistical Methods; UV-Visible Absorption and Fluorescence Spectrometers.**

## Further reading

Clark BJ, Frost T and Russell MA (eds) (1993) *UV Spectroscopy: Techniques, Instrumentation and Data Handling*. Chapman and Hall.

Geladi P and Kowalski BR (1986) Partial Least-Squares Regression: A Tutorial. *Analytica Chimica Acta* **185**: 1–17.

Martens H and Naes T (1989) *Multivariate Calibrations*. Wiley.

Madden HH (1978) Comments on the Savitzky–Golay Convolution Method for Least Squares Fit Smoothing and Differentiation of Digital Data. *Analytical Chemistry* **50**: 1383–1386.

Miller JC and Miller JN (1993) *Statistics for Analytical Chemistry*. Ellis Horwood.

Osbornce BG, Fearn T and Hindle PH (1993) *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*, 2nd edn. Longman.

Stavitsky A and Golay MJE (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **36**: 1627–1639.

# R

# Radiofrequency Field Gradients in NMR, Theory

**Daniel Canet**, Université H. Poincaré, Vandoeuvre, Nancy, France

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| **Theory** |

Any NMR experiment requires two magnetic fields, a static one denoted by $B_0$ which serves to polarize the nuclear spins (or to create two distinct energy levels in the case of spins $\frac{1}{2}$) and an alternating magnetic field $B_1$, also called the radiofrequency field (RF field) because of its usual frequency domain, whose purpose is to induce transitions. It can be mentioned that, in pulse NMR experiments, the RF field serves merely to rotate nuclear magnetization, or, in a more general way, to rotate spin operators involved in the description of the various populations or coherences that may exist or appear in the course of the experiment. In fact, these rotations occur in the so-called rotating frame, that is a frame which rotates at a constant angular velocity ($\omega_r = 2\pi\nu_r$, $\nu_r$ being the transmitter frequency) around the $Z$ direction defined by $B_0$. We shall denote $(x,y,z)$ as the rotating frame and $(X,Y,Z)$, with $Z \equiv z$, as the laboratory frame. In the rotating frame, the RF field is stationary and oriented along $x$,$y$ or any axis in the $(x,y)$ plane according to its phase. Finally, during receive operations, the rotating frame has still to be considered. This is due to the particular scheme usually employed in NMR that detects the signal, not at its own frequency $\nu_0$, but at a relative frequency $(\nu_0 - \nu_r)$ by means of appropriate devices (mixers). Moreover, by means of the so-called quadrature detection scheme, two signals can be acquired simultaneously: $\cos 2\pi\,(\nu_0 - \nu_r)t$ and $\sin 2\pi\,(\nu_0 - \nu_r)t$.

Most of the time, the experimentalist aims at uniform magnetic fields, either $B_0$ or $B_1$. For the former, this need arises from the basic Larmor equation which provides the resonance frequency $\nu_0$:

$$\nu_0 = \frac{\gamma B_0}{2\pi}(1 - \sigma) \qquad [1]$$

($\gamma$ is the gyromagnetic ratio and $\sigma$ the shielding coefficient, which determines the chemical shift of the considered nucleus). Thus a uniform $B_0$ field prevents line broadening or resonance spreading. For the latter, the need arises from the relation which provides the flip angle $\alpha$ (the angle by which a magnetization component or a spin operator rotates around the $B_1$ direction in the rotating frame) is

$$\alpha = \gamma B_1 \tau \qquad [2]$$

($\tau$ is the duration of the RF field application or more simply the duration of the RF pulse). A homogeneous $B_1$ field prevents a distribution of flip angles and, if the same coil is used for signal detection, a nonuniform receptivity by virtue of the reciprocity principle.

However, very early on it was recognized that some advantages may be drawn from using nonuniform magnetic fields. The simplest situation is a linear variation of the $B_0$ amplitude ($B_0$ still being oriented along $Z$) with respect to a spatial direction, say $X$. Let $g_0$ be the (uniform) slope of this variation, leading to an equation indicating how $B_0$ varies with

respect to $X$

$$B_0(X) = B_{00} + g_0 X \qquad [3]$$

Combining Equations [1] and [3], it is apparent that the resonance frequency also depends linearly on $X$ and therefore can provide information about any property which occurs that is spatially dependent. This simple feature led (1) to the capability of measuring self-diffusion coefficients, (2) to the various imaging methods, and (3) in a more subtle way, to the selection of coherence pathways.

Radiofrequency field gradients (RF gradients, or $B_1$ gradients) work in a different way (although strongly related) to $B_0$ gradients. In this case, it is the nutation frequency $\nu_1 = \gamma B_1/2\pi$, i.e. the frequency at which magnetization rotates in a plane perpendicular to $B_1$ (see Eqn [2]), which is spatially dependent; in other words, spatial encoding occurs via nutation and not via precession. It will be shown in this article that virtually all experiments that can be carried out with $B_0$ gradients can also be considered with $B_1$ gradients, sometimes with clear advantages, but unfortunately with some inconveniences which make the two approaches rather complementary. A common feature to these two types of gradients is the ability to *defocus* nuclear magnetization. This means that if a gradient of sufficient strength is applied for a sufficient time, it is capable of spreading out in the relevant plane all elementary magnetization vectors of a homogeneous sample so that the net result is zero.

## Creation of $B_1$ gradients and their specificity with respect to $B_0$ gradients

The first problem of concern is the uniformity of gradients. As will be discussed below, gradient uniformity appears to be mandatory for imaging experiments, recommended for self-diffusion measurements, but not really necessary in pure spectroscopic applications (selection of coherences, solvent suppression, etc.). The best way for creating a uniform $B_1$ gradient is a flat coil which may comprise a single or several turns. For a circular single-turn coil, it can be shown that good uniformity is obtained in a region extending from 0.2 $r$ to 0.9 $r$ from the coil centre ($r$ being the coil radius). The gradient strength is directly proportional to the $B_1$ field at the coil centre and thus depends on the coil quality factor and on the RF power it can hold. This is of course strongly related to the measurement frequency, low frequencies allowing stronger gradients. The state-of-the-art, as far as uniform gradients are concerned, corresponds to gradients up to 800 mT m$^{-1}$ at a measurement frequency of 100 MHz (obtained with a two-turn coil). It must be borne in mind that the gradient efficiency is in any event weighted (or enhanced) by the gyromagnetic ratio of the observed nucleus. As an example, **Figure 1** shows the experimental $B_1$ distribution in a 2 mm diameter sample obtained with a two-turn flat coil of 15 mm and 11 mm diameters respectively. Anyhow, the production of a $B_1$ gradient makes sense only if a detection coil with uniform receptivity exists within the probe. Because, most of the time, the detection coil is tuned at the same frequency as the gradient coil, this poses the leakage problem between the two coils. Ideally, they should be orthogonal. Achieving this orthogonality, although requiring fine mechanical adjustments, proves to be feasible. However, creation of RF gradients in the two complementary directions is actually a challenge. Let $X$ be the direction of the RF gradient created by the specific coil and let $Y$ be the axis of the detection coil (**Figure 2**). It is conceivable that the inherent inhomogeneity of the detection coil affords a $B_1$ gradient in the $Z$ direction (if, for instance, it is of the saddle-shape design), although uniformity is far from being warranted. What about a $B_1$ gradient in the $Y$ direction? If one imagines the installation in the probe of a second gradient coil, identical to the first one and orthogonal to it, its axis would be collinear to the detection coil with evidently a full leakage. This solution cannot be obtained in practice at present and whenever RF gradients in the $X$ and $Y$ directions appear to be necessary, the best way for tackling this problem is to rely on a single coil and to rotate the sample. This is clearly a major drawback of $B_1$ gradients vs. $B_0$ gradients which can be created independently in the three spatial directions without perturbing the RF system. On the other hand, RF gradients offer some distinct advantages over $B_0$ gradients which are summarized in **Table 1**.

In a general way, two points in favour of $B_1$ gradients must be emphasized: (i) a simple instrumentation (at least much simpler than the one currently employed for $B_0$ gradients); (ii) their double function, since they perform both spin excitation and spatial labelling. This latter feature may lead to a considerable simplification of experiments, usually requiring magnetic field gradients.

## Self-diffusion measurements

In order to illustrate the above statement, let us consider the $B_1$ gradient experiment used for measuring

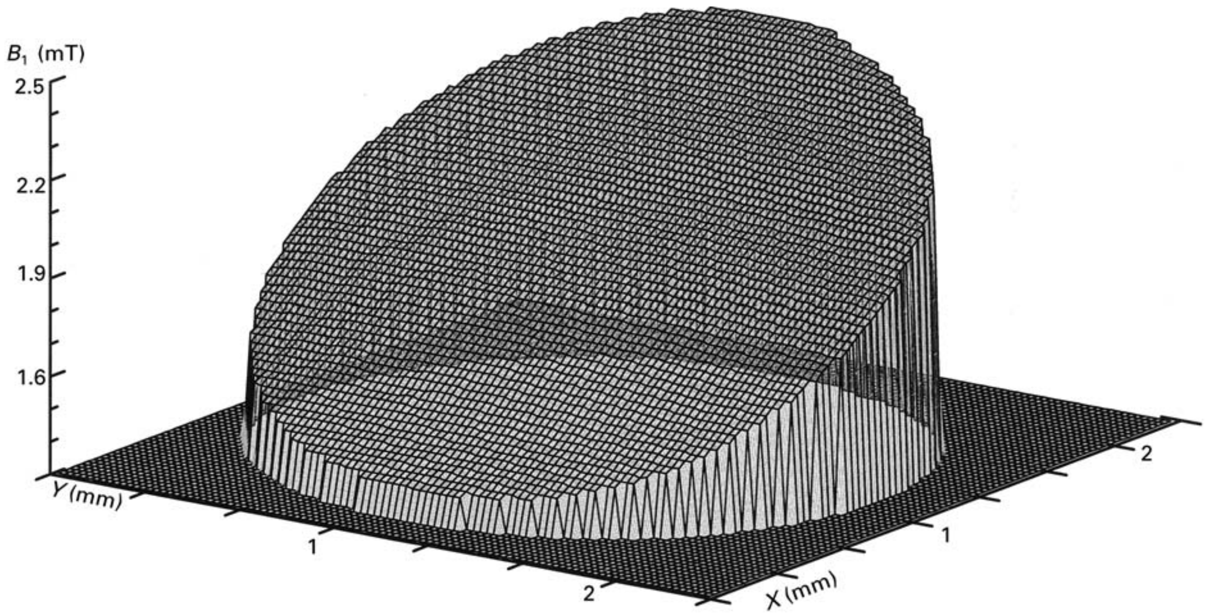



**Figure 1** Three-dimensional diagram demonstrating the excellent uniformity of a RF gradient created by a two-turn flat coil (respective diameters: 15 and 11 mm).

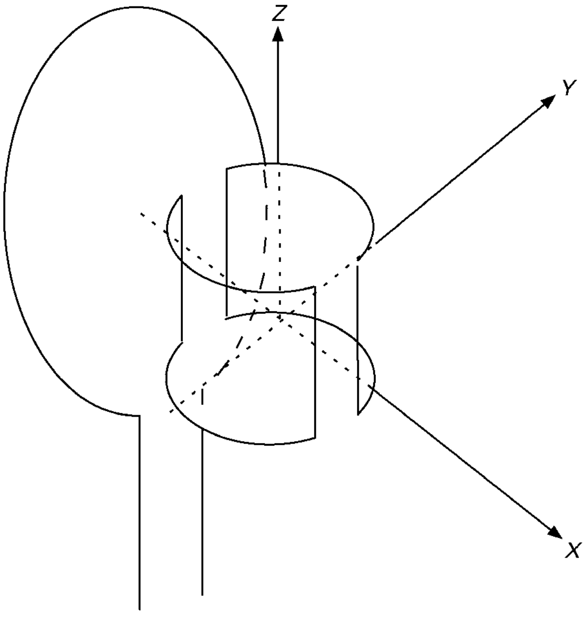



**Figure 2** Sketch of a probe possessing $B_1$ gradient capabilities. The single-turn coil serves for generating a RF gradient in the *X* spatial direction. The saddle coil (in principle electrically orthogonal to the latter) is used essentially for detection purposes. Possibly it can produce a RF gradient in the Z direction.

**Table 1** Properties of $B_1$ vs. $B_0$ gradients

| | *$B_1$ gradients* | *$B_0$ gradients* |
|---|---|---|
| Rise and fall times | < 1 μs | ≥ 100 μs |
| Eddy currents | None | Important, can be greatly attenuated by self shield |
| Perturbation of the static magnetic field and lock system | None | Strong |
| Effects of magnetic susceptibility variations across the sample | None | Important, must be compensated for |
| Effects of short $T_2$ | Reduced by a factor of two | Full |
| Maximum strength (1997) | 800 mT m$^{-1}$ | 10 T m$^{-1}$ |
| Spatial directions available | 1–2 | 3 |
| RF deposition | High (improper for medical applications) | Low (limited to short RF pulses) |

self-diffusion coefficients. The classical method is the PFGSE (pulsed field gradient spin echo) experiment which involves two pulses of $B_0$ gradient on both sides of the 180° RF pulse in a spin echo experiment. Here, with $B_1$ gradients, the experiment is especially simple and robust. It is sketched in **Figure 3** and starts with a $B_1$ gradient pulse $(g_1)_x$ along the $x$ direction of the rotating frame of duration $\delta$, separated by a 'diffusion interval' $\Delta$ from a second gradient pulse of duration identical to the first one ($\delta$), immediately followed by a read pulse which probes the longitudinal magnetization (the read pulse may be incorporated into the second gradient pulse). This experiment can be understood as follows. The first gradient pulse is assumed to totally defocus nuclear magnetization, while the

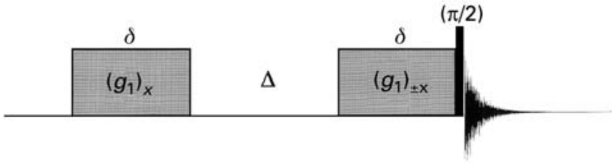



**Figure 3** The simplest experiment for measuring self-diffusion coefficients by RF gradient pulses (shaded rectangles). $+x$ and $-x$ denote the RF phases. The simple phase cycling ($\pm x$), corresponding to two successive experiments whose results are coherently added, eliminates any transverse component of magnetization. The ($\pi/2$) pulse is a standard read pulse.

second one would act in a refocusing fashion provided that no motion occurred during $\Delta$. Because of translational diffusion, the refocusing effect is not complete and some attenuation is observed at a rate depending on the self-diffusion coefficient $D$. The simple phase alternation $(g_1)_{\pm x}$ of the second gradient pulse eliminates any transverse magnetization contribution so that what is left is longitudinal magnetization whose amplitude decays according to (provided that $\Delta >> \delta$):

$$\exp(-\Delta/T_1)\exp(-\gamma^2 g_1^2 \Delta\delta^2 D) \qquad [4]$$

where $T_1$ is the longitudinal relaxation time and $g_1$ the gradient strength. In practice, this decay is monitored by the last $\pi/2$ pulse (which converts the longitudinal magnetization into observable transverse magnetization) as a function of $\delta^2$ and a semilogarithmic plot (for instance) provides $D$. An important point is that the additional attenuation due to relaxation involves $T_1$ and not $T_2$ as in the PFGSE experiment (in many systems of interest $T_1$ has a relatively high value while $T_2$ may become very small). In these latter cases, when employing $B_0$ gradients, stimulated echo experiments (more complicated and more instrumentally demanding) must be available to avoid an unacceptable attenuation by relaxation phenomena. Another point in favour of $B_1$ gradients is the variation of magnetic susceptibility inside heterogeneous samples which is responsible for internal $B_0$ gradients at interfaces. This feature obviously affects $B_0$ gradient experiments and requires the application of compensatory external $B_0$ gradient pulses. Such complications are by nature absent in $B_1$ gradient experiments.

## NMR imaging by radiofrequency field gradients

An early application of $B_1$ gradients in conjunction with spatial discrimination (imaging) is the so-called chemical shift imaging method; it is a two-dimensional experiment which produces chemical shift information in one dimension and spatial information (in the form of spin density) along the second dimension. This amounts to a one-dimensional experiment as far as spatial information is concerned. The experimental arrangement as well as the pulse sequence are very simple. It consists of using a surface coil (widely employed in biomedical applications of NMR), located upon the object under investigation and which therefore produces an inhomogeneous RF field along its axis. Applying a RF pulse of incrementable duration $t_1$ (this implies a series of experiments for each value of $t_1$) and acquiring the NMR signal immediately afterwards (according to the time variable $t_2$) generates a signal $S(t_1, t_2)$ modulated in $t_1$ according to the spin density distribution and in $t_2$ according to the chemical shifts of the different species existing within the object. Because these modulations are in the form of sine or cosine functions, a double Fourier transform with respect to $t_1$ and $t_2$ yields the result indicated above. However, because the $B_1$ gradient (originating from the inhomogeneity of the surface coil) is far from being uniform and because the receptivity of this coil (which serves as well for detection purposes) is by nature strongly non uniform, the results along the spatial dimension are obviously not quantitative. Nevertheless, the method has been widely used and is found very useful for obtaining the distribution of, for example, phosphorus metabolites. There are two instrumental prerequisites for obtaining quantitative spatial information: (i) a detecting coil with homogeneous receptivity and (ii) a coil delivering a uniform $B_1$ gradient. Clearly this corresponds to the arrangement of **Figure 2**. In practice, for making the experiment viable, the acquisition of spatial information must also be accomplished in a very short time (in a fraction of a second), as this is done by the read gradient in classical NMR imaging with $B_0$ gradients. These ideas led to the development of an experimental procedure adapted to $B_1$ gradients, which mimic the spatial encoding by a $B_0$ read gradient. The problem here is that acquisition of the NMR signal cannot take place while an RF field is on. The solution is indeed very simple and consists of applying the $B_1$ gradient in the form of short pulses separated by short intervals devoted to acquisition of the NMR signal. In practice a $B_1$ gradient pulse is followed by the acquisition of a single data point and the process is repeated until the NMR signal vanishes due to relaxation phenomena. The experiment is sketched in **Figure 4** and can be analysed as follows. Omitting

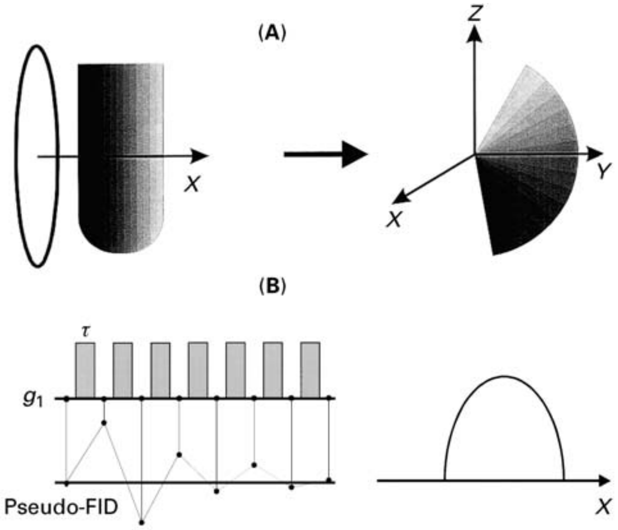



**Figure 4** (**A**) Left: Different elementary slices subjected to the decreasing RF field produced by the gradient coil. Right: corresponding nutations assuming that nuclear magnetization was initially at thermal equilibrium. (**B**) Left: The RF gradient pulse train with interleaved detection windows leading to the acquisition of a pseudo-FID. Right: the Fourier transform of this pseudo-FID yields a profile representative of the object shape (see A).

relaxation effects and defining, for the $l^{th}$ data point, the time $t$ as $l\tau$, where $\tau$ is the duration of each individual gradient pulse, we can write the amplitude of the NMR signal (which appears in the form of a 'pseudo-free induction decay', pseudo-FID) as

$$S(k) = \int_{\text{object}} \rho(X) \sin(2\pi X k) \mathrm{d}X \quad [5]$$

where $\rho(X)$ is the spin density at abscissa $X$ (the gradient is assumed to act along the $X$ spatial direction) and where the variable $k$, according to usual practice, is substituted to the variable $t$

$$k = (2\pi)^{-1} \gamma g_1 t \quad [6]$$

($g_1$ is the gradient amplitude). It appears that there exists a Fourier relationship between $S$ and $\rho$ which can therefore be deduced from the following integral (inverse Fourier transform)

$$\rho(X) = \int S(k) \sin(2\pi k X) \mathrm{d}k \quad [7]$$

yielding the object profile along $X$.

A two-dimensional ($X$,$Y$) image, possibly after a slice-selection procedure along the $Z$ direction, can be reconstructed from a series of profiles obtained at different orientations resulting from the sample rotation around $Z$. Moreover, every type of contrast used in MRI can also be considered with RF gradients. These contrasts can originate from relaxation times, diffusion coefficients and, most importantly, chemical shifts. This latter possibility is illustrated by **Figure 5** which shows separate images of two components in a solvent mixture, sorted according to their chemical shift.

## RF gradients in pure high-resolution NMR spectroscopy

It is now well established that the judicious use of $B_0$ gradient pulses in a complicated multipulse (and possibly multidimensional) NMR experiment can lead to invaluable improvements in the quality of spectra, eventually lifting barriers which were thought to be impossible to overcome. This is especially true for the suppression of huge peaks such as those of solvent or of protons bound to a majority isotope (e.g. $^{12}C$ vs. $^{13}C$ in natural abundance). Another attractive feature concerns the suppression of phase-cycling procedures which can be replaced by appropriate gradient pulses

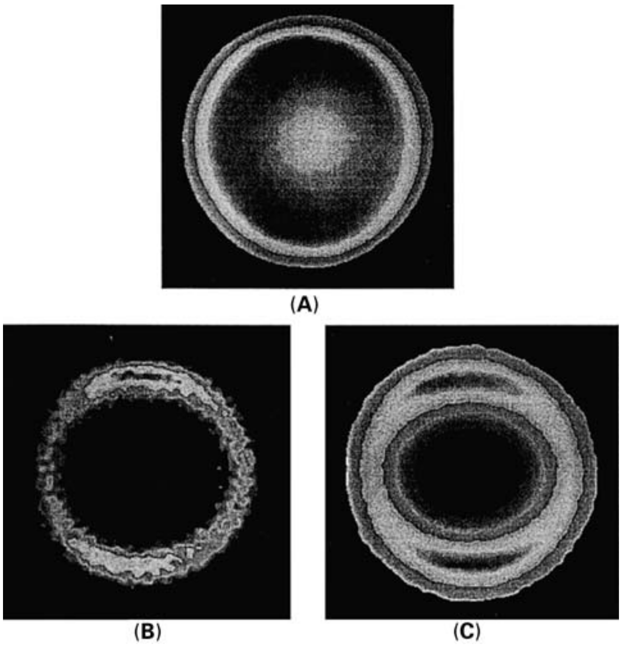



**Figure 5** (**A**) Full image of a polymer rod that has been immersed in a solvent mixture (isooctane, ethanol, toluene). (**B**,**C**) Chemical shift selected images of isooctane and toluene, respectively, showing the different degrees of penetration.

in a single experiment thus considerably reducing measuring times. Basically, the benefit of using gradients in high-resolution NMR relies on the possibility of defocusing all magnetizations and, by the end of the experiment, selectively refocusing the desired magnetization(s) or coherence(s). (Throughout, gradient strengths and gradient pulse durations are supposed to entail a complete defocusing, i.e. a complete disappearance of the NMR signal.) Actually refocusing can be achieved in two ways: (i) by applying a gradient pulse of the same strength and same duration with an inverse polarity, thus producing the same precession in the reverse direction and (ii) by applying exactly the same gradient pulse after a 180° RF pulse which changes the sign of magnetization and amounts to reversing the gradient sign. Now, if the 180° pulse is selective, the only magnetization which survives (i.e. which has been refocused) is the one corresponding to the bandwidth of the 180° selective pulse. These simple considerations show how it is possible to manipulate gradients in order to preserve one resonance while destroying all the others. In fact, gradients can accommodate much more complicated schemes involving coherences of different orders and selection of a given coherence pathway. It can be recalled that coherence orders stem from the state of the spin system in the course of the experiment. For instance, in the case of two weakly coupled nuclei of spin-$\frac{1}{2}$, denoted $A$ and $X$, the coherence orders are as follows ($I_+$ and $I_-$ are the classical raising and lowering operators): $p = 0$, longitudinal magnetization or zero quantum coherence (represented by e.g. $I_+^A I_-^X$); $p = \pm 1$, single quantum coherences corresponding to observable transverse magnetization; $p = \pm 2$, double quantum coherences (represented by e.g. $I_+^A I_+^X$).

Recognizing that a coherence of order 0 is not affected by $B_0$ gradients, that defocusing of a double quantum coherence is twofold, the defocusing of a one-quantum coherence and that the sense of defocusing is given by the sign of the coherence order, it can easily by recognized that the coherence pathways leading to a final observable signal must satisfy the simple equation

$$\sum_i p_i g_i = 0 \qquad [8]$$

where $p_i$ and $g_i$ are respectively the coherence order and the gradient strength relevant to the $i$th interval of the considered experiment. This equation is in fact valid for homonuclear experiments; in a heteronuclear experiment, gyromagnetic ratios must be introduced as multiplying factors. At this point, it must be emphasized that the above discussion applies to $B_0$ gradients, i.e. to precession (or rotation) in the transverse ($x$,$y$) plane. If one assumes that these $B_0$ gradients are sufficiently strong so that the natural precession (due to chemical shift) is negligible, a full equivalence with $B_1$ gradients can be established by recognizing that the latter correspond to the same type of rotation in the vertical plane of the rotating frame. Consequently, this equivalence amounts to flipping the vertical plane into the $xy$ plane and, after the $B_1$ gradient application, flipping it back to its initial position. This is achieved by clusters of the type

$$(\pi/2)_x(g_1)_y(\pi/2)_{-x};\ (\pi/2)_{-x}(g_1)_{-y}(\pi/2)_x;$$
$$(\pi/2)_{-y}(g_1)_x(\pi/2)_y;\ (\pi/2)_y(g_1)_{-x}(\pi/2)_{-x} \qquad [9]$$

where ($\pi/2$) stands for a standard hard pulse (homogeneous) and ($g_1$) for the application of a RF gradient of sufficient strength and duration (to produce a complete defocusing); RF phases are indicated as subscripts. By means of this equivalence, the numerous experiments involving $B_0$ gradients can easily be transposed with $B_1$ gradients and will not be discussed further.

The emphasis will rather be put on two simple and illustrative experiments which take advantage of the specific features afforded by RF gradients. In principle they could be converted into $B_0$ gradient experiments by using the reciprocal of Equation [9]; however, due to the inherent shortcomings of $B_0$ gradients (especially rise and fall times) and to the complexity of the resulting sequences, this is not practicable. The first example deals with solvent peak suppression. Among several possibilities allowed by $B_1$ gradients, the one using a train of short gradient pulses appears somewhat specific and difficult to convert into a $B_0$ gradient version. It works in the manner of a DANTE (delays alternating with nutations for tailored excitation) sequence. It can be recalled that the train of short, homogeneous RF pulses (each corresponding to a small flip angle) of a DANTE sequence is equivalent to a selective pulse because the effect of pulses is cumulative only for on-resonance magnetization. Magnetizations corresponding to other resonance frequencies are tipped back to the $z$ axis (to their equilibrium position) in the course of the pulse train. If, instead of homogeneous pulses, RF gradient pulses are used, selectivity is of course retained, but their effect is to defocus that magnetization which is on-resonance and thus to suppress the relevant signal. The efficiency of this procedure is illustrated in **Figure 6** where a sufficient number of cycles ($n$) effectively leads to complete on-resonance peak suppression.

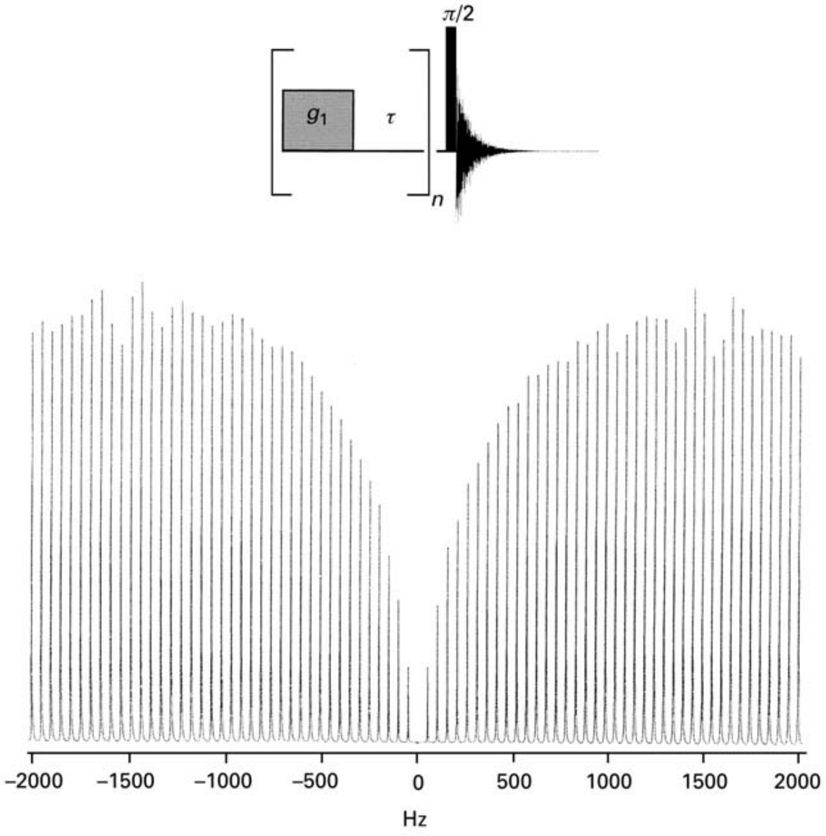



**Figure 6** Suppression profile resulting from the sequence shown at the top of the Figure and obtained by repeating the experiment for a series of transmitter frequencies. When the signal is on-resonance, it is selectively defocused by the train of RF gradient pulses ($g_1$). $\tau$ is chosen so that the first sideband ($2\pi$ precession) is outside the spectrum of interest.

Another appealing feature of RF gradient pulses is their ability to select coherences, and this will be discussed here for systems of spin-$\frac{1}{2}$ nuclei. A p-spin high-pass filter is indeed achieved by the following cluster of two $B_1$ gradient pulses of respective phases $x$ and $y$:

$$(g_1)_x(rg_1)_y; \quad r = p/(p-1) \qquad [10]$$

$rg_1$ denoting a second gradient pulse that is $r$ times longer than the first one. $r$ is actually set according to the desired order of filtering, meaning that one-spin systems are rejected for $r = 2$, one- and two-spin systems are rejected for $r = 3/2$ (**Figure 7**) and so on. Before going into detail for a particular spin system, it can be anticipated that the filtering properties should rest on antiphase coherences rather than on multiple quantum coherences as might be the case with $B_0$ gradients. To make this statement clearer, it can be recalled that an antiphase $A$ doublet (one line positive, the other negative) of a two-spin $AX$ system is represented by the operator product $2I_x^AI_z^X$ (provided the two $A$ lines are along the $x$ axis of the rotating frame). Now, $B_1$ gradients acting in vertical planes (e.g. $x,z$ or $y,z$) are prone to operate on such antiphase coherences rather than on multiple quantum coherences (e.g. of the type $2I_x^AI_y^X$). In order to clarify this point, consider the $A$ coherences which can exist after an evolution interval following a standard $\pi/2$ pulse: $I_x^A$, $I_y^A$, $2I_x^AI_z^X$, $2I_y^AI_z^X$ and let $\theta$ be the angle of nutation produced by $g_1$ for a given spatial location, acknowledging that, at the outcome, an average over all possible values of $\theta$ must be performed (for sufficient gradient strength and duration: $\langle\sin\theta\rangle = \langle\cos\theta\rangle = 0$, $\langle\sin^2\theta\rangle = \langle\cos^2\theta\rangle = 1/2$). From these considerations, it is easy to derive the way in which the various coherences transform under the gradient cluster $(g_1)_x(2g_1)_y$:

$$I_x^A \xrightarrow{(g_1)_x(2g_1)_y} 0$$

$$I_y^A \longrightarrow 0$$

$$2I_x^AI_z^X \longrightarrow 0$$

$$2I_y^AI_z^X \longrightarrow \left(2I_y^AI_z^X\right)/4 + \left(2I_z^AI_y^X\right)/4$$

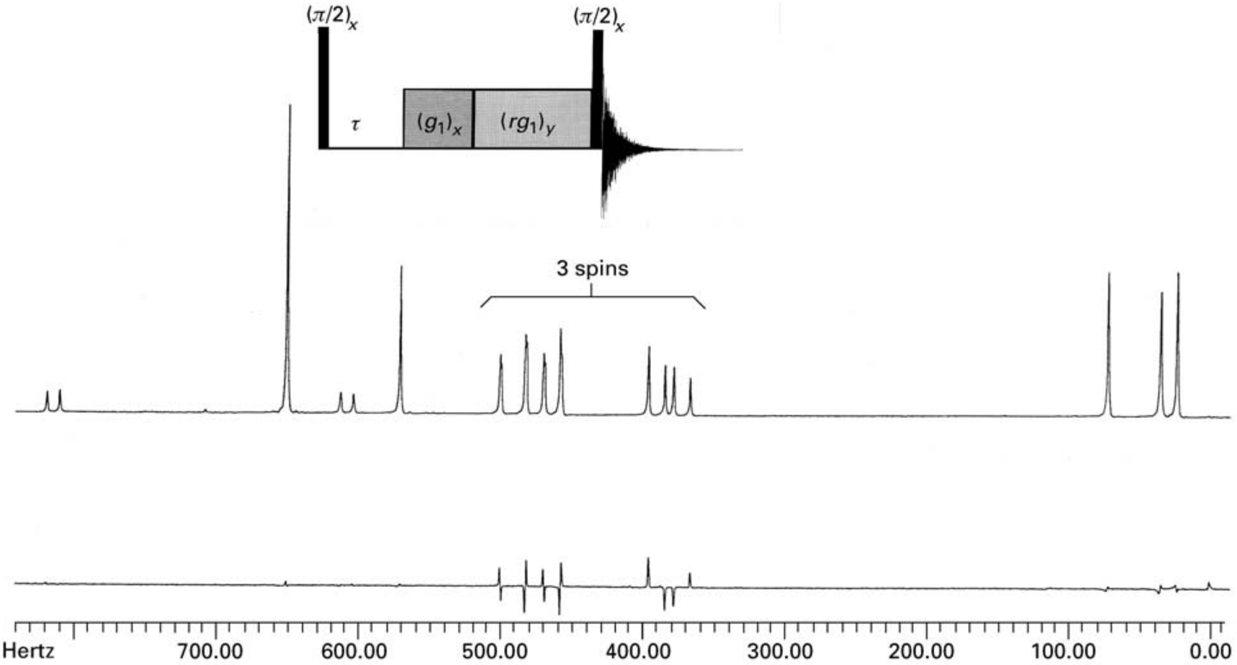



**Figure 7** Reference (top) and 3 spin-filtered (bottom) spectra. The latter has been obtained with the sequence shown at the top of the Figure ($r$ 3/2; see text). The time $\tau$ is chosen so that antiphase coherences can develop; the last $\pi/2$ pulse is for purging purposes.

Clearly, all quantities cancel (including those corresponding to a single spin system) except the antiphase coherence $2I_y^A I_z^X$, scaled down by a factor 4 and which leads to the corresponding quantity transferred to the $X$ nucleus. Thus, in addition to its filtering properties, this gradient cluster affords transfer capabilities suitable for two-dimensional correlation spectroscopy without the need of phase cycling procedures.

## List of symbols

$B_0$ = static magnetic field; $B_1$ = alternating magnetic field; $D$ = self-diffusion coefficient; $g_1$ = gradient amplitude; $I_+$ and $I_-$ = raising and lowering operators; $p_i$ = coherence order; $r$ = coil radius; $S$ = signal; $T_1$ = longitudinal relaxation time; $T_2$ = transverse relaxation time; $\alpha$ = flip angle; $\gamma$ = gyromagnetic ratio; $\delta$ = chemical shift; $\Delta$ = diffusion interval; $\theta$ = angle of nutation; $\nu_0$ = resonance frequency; $\nu_r$ = transmitter frequency; $\rho$ = spin density; $\sigma$ = shielding coefficient; $\omega_r$ = notating frame angular velocity.

*See also:* **Diffusion Studied Using NMR Spectroscopy; Magnetic Field Gradients in High Resolution NMR; MRI Theory; NMR Microscopy; NMR Principles; NMR Pulse Sequences; NMR Spectrometers; Product Operator Formalism in NMR; Two-Dimensional NMR Methods.**

## Further reading

Bosch CS and Ackerman JJH (1992) Surface coil spectroscopy. *NMR Basic Principles and Progress* **27**: 3–44.

Bottomley PA (1992) Depth resolved surface coil spectroscopy. *NMR Basic Principles and Progress* **27**: 67–102.

Callaghan PT (1993) *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford: Oxford University Press.

Canet D (1997) Radiofrequency field gradient experiments. *Progress in NMR Spectroscopy* **30**: 101–135.

Kimmich R (1997) *NMR Tomography, Diffusometry, Relaxometry*. Berlin: Springer.

Kormoroski RA (1993) Non medical applications of NMR imaging. *Analytical Chemistry* **65**: 1068A–1077A.

Maffei P, Mutzenhardt P, Retournard A *et al* (1994) NMR microscopy by radiofrequency field gradients. *Journal of Magnetic Resonance* **A107**: 40–49.

Norwood TJ (1994) Magnetic field gradients in NMR: friend or foe? *Chemical Society Reviews* **23**: 59–66.

Price WS (1996) Gradient NMR. *Annual Reports on NMR Spectroscopy* **32**: 51–142.

Price WS (1998) NMR imaging. *Annual Reports on NMR Spectroscopy* **35**: 139–216.

# Radiofrequency Spectroscopy, Applications

*See* **Microwave and Radiowave Spectroscopy, Applications.**

# Raman and Infrared Microspectroscopy

**Pina Colarusso**, **Linda H Kidder**, **Ira W Levin** and **E Neil Lewis**, National Institutes of Health, Bethesda, MD, USA

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**

**Methods & Instrumentation**

## Introduction

Vibrational Raman and infrared microspectroscopy are analytical tools that characterize both the chemistry and the physical structure of materials. These methods combine two separate approaches, vibrational spectroscopy and light microscopy, for elucidating chemical systems at the micron and the submicron levels. Whereas vibrational spectroscopy probes the details of molecular composition, light microscopy reveals sample morphology based on the variations in optical properties. Vibrational microsopy builds on both of these techniques, and thus provides chemically selective visualizations of microscopic samples. Domains within complex samples, ranging from diamond inclusions in a mineral to malignant cells in a tissue biopsy, may be examined with Raman and infrared techniques. The efficacy and flexibility of vibrational microspectroscopy is highlighted by its wide adoption in diverse areas such as geology, medicine, forensic science, and industrial process control.

Since the theoretical background and practical implementation of vibrational spectroscopy are described elsewhere in this encyclopedia, the emphasis here will be on the extension of Raman and infrared spectroscopy to the microscopic realm. It should be noted that many of the methods described here are applicable to other microspectroscopic methods, especially those based on optical phenomena such as fluorescence.

The balance between spectroscopy and microscopy distinguishes one vibrational microscopic technique from another. At one extreme, spectroscopic and microscopic analyses are carried out separately. In a typical arrangement, the sample is examined through a microscope under white-light illumination, and then either Raman or infrared spectra are recorded at one or more selected points. Approaches that more fully integrate the acquisition of morphological and spectroscopic data have also been developed. Mapping techniques record spectra at successive points or lines within the sample; the spectra are combined to provide a view of the sample at specific vibrational frequencies. Imaging techniques, by contrast, record spectra simultaneously for contiguous points within a given sample area.

Mapping and imaging methods allow for sample visualizations over sequential wavelength intervals. The data sets are represented by a three-dimensional 'image cube' having two spatial and one spectral dimension (see **Figure 1**). Depending on the orientation, two-dimensional slices along a particular axis yield either a set of image planes stacked as a function of wavelength or a group of spatially resolved spectra. Vibrational spectroscopic mapping and imaging combine the morphological and chemical analyses of microscopic systems and reveal trends that are often difficult to extract from bulk or isolated single-point measurements.

## Instrumentation

A standard Raman or infrared microspectrometer consists of an excitation source, a compound microscope, a spectrometer, and a detector. As in bulk techniques, the design of microspectroscopic experiments is guided by the sample composition as well as demands for frequency response, sensitivity, data acquisition rates, and spectral resolution. Factors specific to vibrational microspectroscopy include the spatial resolution and the optical throughput between the microscope and the spectrometer.

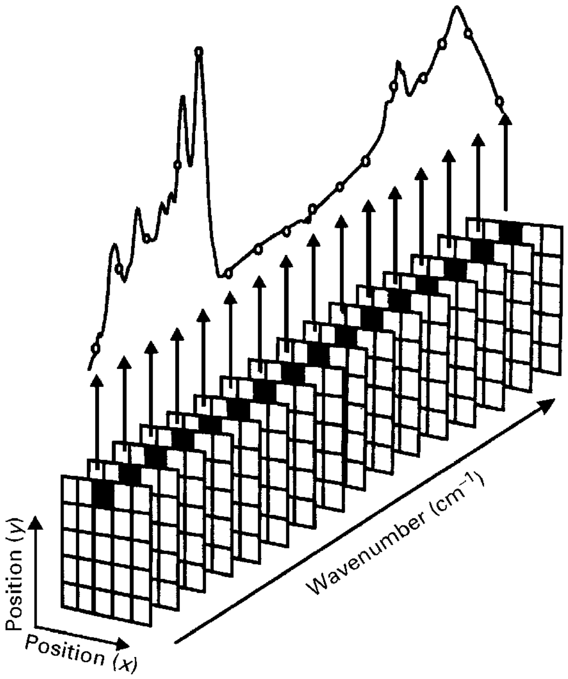



**Figure 1** Mapping and imaging data can be depicted as an 'image cube' having two spatial and one spectral dimensions.

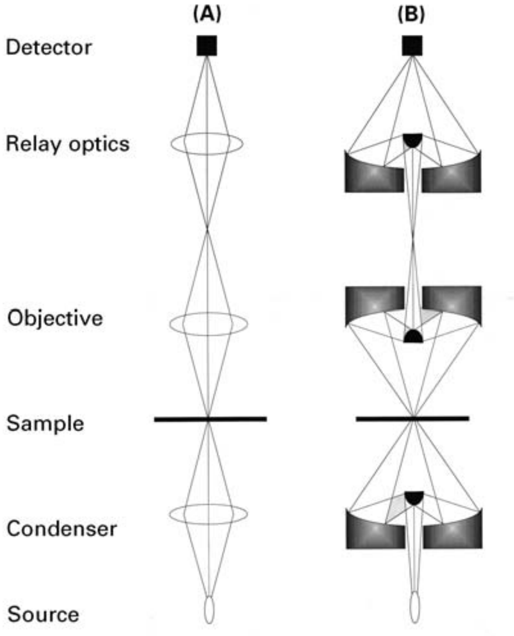



**Figure 2** Schematics of (A) refractive and (B) reflective microscopes.

## Sources

Raman microspectroscopy is usually implemented from the ultraviolet to near-infrared wavelength regions, from about 0.3 to 1.5 μm. Excitation with sources such as $Ar^+$, HeNe, and solid-state lasers is standard. Infrared microspectroscopy, by contrast, is usually carried out between 0.78 and 25 μm with broadband sources such as quartz lamps or ceramic globars. More recently, excitation with synchrotron radiation has also been demonstrated for infrared measurements.

## The compound microscope

The compound microscope is central to vibrational microspectroscopy. The optical components and light paths for both refractive and reflective microscopes are shown in **Figure 2**. As illustrated, a compound microscope contains a light source, a condenser, an objective, various apertures, and an ocular. The sample is first visualized under white light to aid in observation and alignment. The condenser illuminates the sample, and the transmitted light is collected by the objective. Light also can be reflected (or scattered) from the sample; in this configuration, which is known as epi-illumination, the objective also serves as the condenser. In either case, an inverted and usually enlarged image of the sample is formed at the back focal plane of the objective. This intermediate image is either relayed through collection optics to a video camera or through the ocular, which further magnifies the image for viewing. Microscopes tailored for operation at ultraviolet and infrared wavelengths usually contain a parfocal white-light path for sample observation.

Following the preliminary microscopic examination, the spectroscopic source is introduced on to the sample. In Raman microspectroscopy, the probed region is defined by the size of the impinging laser beam. For infrared applications, the radiation is localized within a given area in part by placing an aperture between the source and the sample. An image of the irradiated area, corresponding to the Raman or infrared signal, is formed behind the objective, and is diverted to the spectrometer for analysis. The optical path of the vibrational signal is configured such that it focuses in the same plane as the white-light image; the parfocal light paths ensure that identical sample regions are examined in both procedures, while maximizing the optical throughput.

**Microscope lenses** Microscope lenses are complex assemblies that are designed to balance the requirements for magnification, focal length, and light collection with various factors that degrade image fidelity (note that the optics in **Figure 2** are

simplified for clarity). Refractive and reflective optics exhibit aberrations that cause image blurring and deformation. These include effects such as spherical aberrations and astigmatism. Refractive optics also exhibit chromatic aberrations which causes light of different wavelengths to focus at separate points.

The highest quality images are generally obtained with light microscopes that contain refractive glass elements. Visible Raman microspectroscopy, in particular, benefits from the mature optical technology that has been developed for visible wavelengths. Refractive optics are available for vibrational microspectroscopic applications at other wavelengths; examples include quartz for the ultraviolet and germanium for the infrared. Reflective optics can also be tailored for different wavelength intervals and typically consist of glass substrates coated with a metallic layer. Since they do not exhibit chromatic aberration, reflective optics are particularly useful when large wavelength intervals are covered, as in mid-infrared microspectroscopy.

A standard optic in a reflective microscope is the Schwarzchild (also referred to as Cassegrainian) lens; it can be used as both a condenser and an objective. As depicted in **Figure 2B**, the Schwarzchild lens is centrally obscured, which reduces the amount of light collected compared with a refractive optic of similar dimensions. Since the optic is reflective, chromatic aberration is not a concern, but the spherical design can lead to image warping.

It is standard to label objectives with the numerical aperture, *NA*, which is a measure of the light-gathering capability of an optic:

$$NA = n \sin \phi_{\max} \qquad [1]$$

where $n$ is the refractive index of the medium between the sample and objective and $\phi_{\max}$ is the half-angle of the maximum cone of light collected by the lens (see **Figure 3**). In a given apparatus, the *NA* of the objective, relay optics, and spectrometer aperture are matched for optimum light collection.

Vibrational objectives typically operate in air, which has a refractive index near unity; the highest *NA* value obtained in air is about 0.95 using refractive visible objectives. By changing the surrounding medium, and thus $n$, it is possible to obtain values greater than 0.95. Immersion objectives operating under media such as water or oil achieve *NA* values greater than 1. For mid-infrared applications, the maximum *NA* for reflective objectives is about 0.65. For optimal image formation, it is important to follow the manufacturer's specifications for thickness and refractive index of slides and coverslips (if applicable).

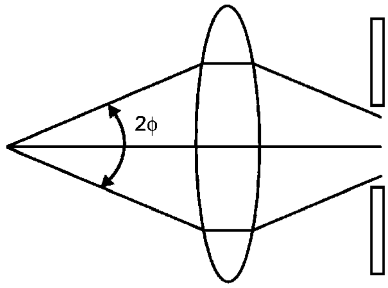



**Figure 3** The numerical aperture is defined as $n \sin \phi$, where $n$ is the refractive index of the surrounding medium and $\phi$ is the half-angle of the cone of light collected by the optic.

**Spatial resolution** Diffraction effects influence the spatial resolution of a vibrational microspectroscopic measurement. To illustrate consider, for example, monochromatic light from a point source as it propagates through the objective. The image that forms at the focal plane of the objective is not a point, but rather a diffraction pattern consisting of alternating light and dark concentric circles. The bright central disc in the pattern is known as the Airy disc, with a radius $r_A$ given by

$$r_A \approx 0.61\, \lambda / NA_{\text{obj}} \qquad [2]$$

where $\lambda$ is the wavelength of light and *NA* is the numerical aperture of the optic.

Two incoherent point sources of equal brightness lying in a plane perpendicular to the objective are just resolved when

$$R \approx 1.22\, \lambda / (NA_{\text{obj}} + NA_{\text{cond}}) \qquad [3]$$

where $R$ is the distance between the two points and $NA_{\text{obj}}$ and $NA_{\text{cond}}$ are the numerical apertures of the objective and condenser, respectively. Equation [3] is known as the Rayleigh criterion, and is a standard measure of the lateral spatial resolution. For an epi-illumination measurement, Equation [3] is recast as

$$R \approx 0.61\, \lambda / NA_{\text{obj}} \qquad [4]$$

From Equations [3] and [4], it is seen that the optimum spatial resolution is obtained when objectives with high *NA* are used for measurements at short-wavelengths.

The spatial resolution criteria given in Equations [3] and [4] apply for diffraction-limited microscope objectives in the far-field limit. In recent years, near-field techniques have been devised that exceed the diffraction limit by scanning a tapered light source, with a spot size less than the probe wavelength, in close proximity to the sample.

### Spectrometers

Dispersive elements and interferometers are widely used in vibrational microspectroscopy. As in bulk measurements, microscopic Raman studies are carried out with grating monochromators, spectrographs, or Fourier transform spectrometers, although Fourier transform instruments are usually limited to applications in the near-infrared spectral region. Infrared microspectroscopy, by contrast, is almost exclusively a Fourier transform technique.

At visible and near-infrared wavelengths, vibrational microspectroscopy is also implemented with solid-state filters, particularly in imaging applications. Electronically driven devices such as acousto-optic and liquid crystal filters provide the tunability and moderately narrow passbands required for spectroscopic imaging. These high-speed filters contain no moving parts and can be custom-built to provide spectral resolutions of less than 10 $cm^{-1}$ over large spectral ranges. Acousto-optic tunable filters (AOTFs) consist of a piezoelectric transducer bonded to a birefringent crystal such as $TeO_2$. When an RF frequency is applied to the crystal, an acoustic wave is generated that diffracts light over a narrow spectral interval. The AOTF passband is modified by varying the input RF frequency.

Another useful solid-state device is an interference filter fabricated from a set of liquid crystals. The birefringent properties of a liquid crystal tunable filter (LCTF) can be varied by applying an external voltage across a crystal axis. A filter is constructed from a series of polarizers and liquid crystals. A particular passband is selected by tuning the individual liquid crystal elements.

Thin-layer interference filters with passbands between 18 and 50 $cm^{-1}$ are also applied in microspectroscopic imaging. These devices can be tuned over large wavenumber ranges by varying the angle of incidence. Broader wavelength coverage may be obtained with a series of filters, which can be placed in a device such as a filter wheel.

### Detectors

The detectors used in vibrational microspectroscopy operate by transducing either the capture of a photon or a minute change in temperature into an electrical response. Both photon and thermal detectors are used individually or configured into an array. The single-point detectors that comprise the array detector are known as picture elements or pixels. **Table 1** lists the wavelength ranges and operating temperatures of several array detectors that are used in Raman and infrared microspectroscopy and imaging.

For the visible to near-infrared region, charge-coupled device (CCD) detectors have been widely adopted for single-point, mapping, and imaging Raman microspectroscopy. These photosensitive arrays, in most cases, are monolithic silicon devices that can be fabricated with millions of individual pixels. The wavelength response of CCDs generally declines in the red, cutting off at about 1.1 μm, the band-gap of silicon. For this reason, CCDs have limited application in near-infrared microspectroscopy.

For infrared microspectroscopy, single-element detectors are used for point and mapping measurements. More recently, array detectors have been applied for spectroscopic imaging in the infrared. In infrared focal plane arrays, the monolithic silicon design used in CCDs is replaced by a hybrid construction. In a hybrid detector, photon detection occurs in a semiconductor layer (indium antimonide, mercury cadmium telluride, and doped-silicon are typical detector materials), while the readout and amplification stages are carried out in a silicon layer. The two layers are electrically connected at each pixel through indium ‘bump-bonds’. Other innovations such as microbolometer arrays also show promise for spectroscopic imaging applications.

## Raman microspectroscopy

### Point microscopy

Raman microspectroscopy is routinely implemented using epi-illumination. As shown in **Figures 4A–4E**, the objective directs the laser excitation onto the

**Table 1** Properties of various array detectors

| | *CCD* | *InGaAs* | *Pt:Si* | *InSb* | *HgCdTe (MCT)* | *Si:As* | *Microbolometer* |
|---|---|---|---|---|---|---|---|
| Wavelength range (μm) | 300–1.1 | 0.9–1.7 | 1–5.7 | 0.5–5.4 | 0.8–12.5 | 1–25 | 8–14 |
| Operating temperature(K) | 77 | 300 | 77 | 77 | 77 | <10 | 300 |

sample and collects the light scattered from the surface. Often fibre optics are used for safe and convenient optical coupling of the microscope to the laser and spectrometer. In point microspectroscopy (**Figure 4A**), the Raman signal from a small spot on the sample is dispersed by a grating spectrometer and focused onto a CCD detector. The constituent wavelengths are mea-sured by one or more pixels along a long narrow strip on the CCD. The undesired Rayleigh scattered light, corresponding to the laser excitation wavelength, is removed by placing one or more holographic or di-electric notch rejection filters between the microscope and spectrometer. Various grating designs, such as dual-stage and subtractive monochromators, also may be used to remove the laser excitation, but are less optically efficient than instruments that incorporate notch rejection filters and single-stage monochromators.

### Confocal Raman microscopy

Optical sectioning of a sample may be achieved with confocal Raman microscopy. This technique rejects out of focus light by introducing pinhole apertures into the optical train of the microscope. A confocal scheme is shown in **Figure 5**, where one aperture is placed between the laser and the objective, and another is placed at the image plane of the objective. The apertures block the light scattered from regions outside the focal plane of the objective. Crisp images of thin optical sections may be obtained by mapping the sample point-by-point. Furthermore, a three-dimensional image of the sample may be constructed by carrying out measurements at different sample depths.

### Mapping

Mapping techniques scan the excitation source or the sample in a raster or line pattern (see **Figures 4B** and **4C**). In a raster scan, spectra may be recorded at successive points by moving either the sample or source on a motorized platform. Mapping may also be implemented by irradiating the specimen with a line source. In one method, a line is produced by rapidly sweeping the laser beam back and forth with a mirror powered by a piezoelectric transducer. It is also possible to focus the laser beam into a line with cylindrical optics. High-precision designs for mapping apparatus minimize mechanical instabilities and positioning errors.

For point mapping, the Raman signal is recorded as in single-point measurements. When the excitation is along a line, the Raman signal is collected through the objective, directed onto a grating and dispersed on to a CCD. The two-dimensional detector then records the spatial information along one axis and wavelength data along the other. An image cube is gradually built up as the line is moved across the full sample area.

### Imaging

As shown in **Figures 4D** and **4E**, point and line excitation are replaced by wide-field irradiation in both spatial encoding and direct imaging methods. In a common arrangement, a given sample area is illuminated by defocusing the laser through a beam expander prior to the objective. The sample illumination is not perfectly uniform in this configuration because the intensity distribution of the laser beam (typically Gaussian) is preserved. Other properties of monochromatic coherent excitation may also affect the image quality. Despite the experimental concerns, wide-field illumination is useful when dealing with fragile specimens, since the power density is reduced, thus minimizing damage from thermal or photolytic processes.

**Hadamard transform imaging** Spatial encoding methods such as Hadamard transform imaging also can be used for the spectroscopic visualization of samples. More recently, developments in digital microarray technology will likely provide a convenient new approach for spatial encoding from the mid-infrared to the ultraviolet. In one common arrangement, the entire sample area is irradiated with wide-field, epi-illumination (**Figure 4D**). Part of the Raman signal emanating from the sample is blocked with a mask containing a series of apertures. The spatially filtered signal is focused on to an entrance slit of a monochromator, which disperses the signal across a two-dimensional array detector. The slit preserves one image axis while the Hadamard mask is used to encode the other axis. Subsequent measurements are carried out with the mask in different positions. Each measurement corresponds to the Raman signal from the unmasked points on the sample along one spatial axis over the entire spectral range of interest. The experiment is designed such that the number of independent measurements equals the number of points on the sample. The spatially dependent images are then converted to spectroscopic images through a Hadamard transform. Unlike the other methods mentioned in this article, Hadamard spectroscopic imaging systems are limited to research activities and are not yet commercially available.

**Direct imaging** Raman microscopic imaging can also be carried out by imaging the sample through an interferometer or filter (shown in **Figure 4E**). The sample is irradiated with wide-field, epi-illumination and then modulated or filtered images are recorded

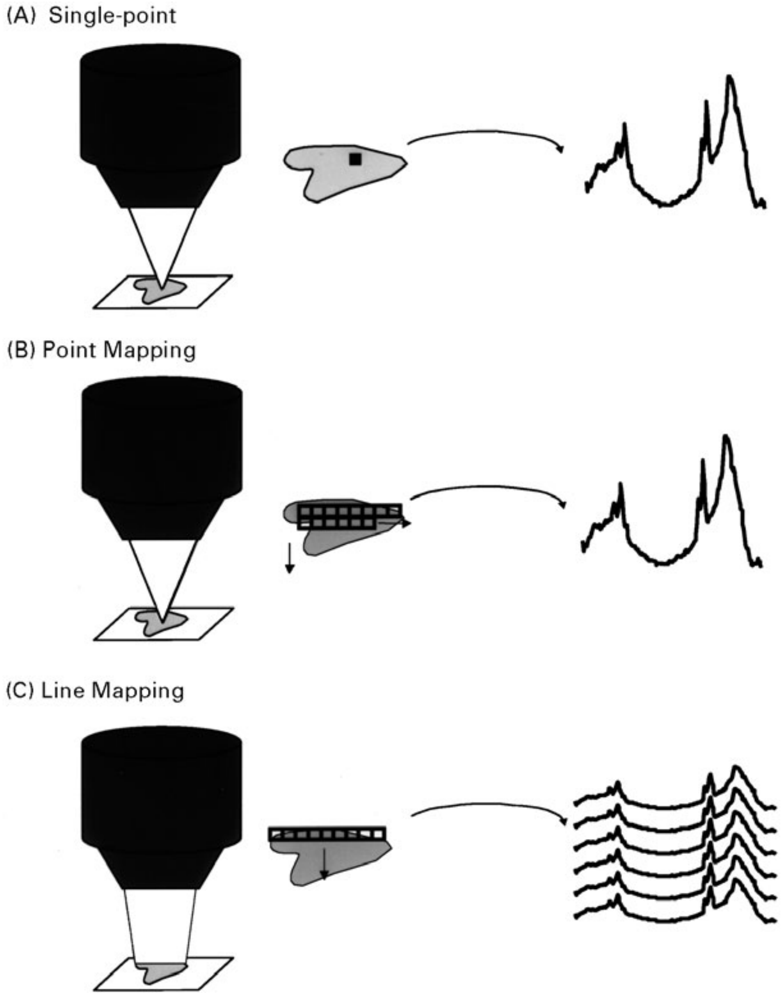



**Figure 4** Several different approaches employed in vibrational microspectroscopy. (A) point measurements: vibrational spectra are obtained at individual, selected points within the sample. (B) Point mapping: spectra are recorded at successive spots along within the sample. (C) Mapping is also carried out with line excitation. (D) Spatial encoding methods such as Hadamard transform imaging: a physical mask blocks part of the signal from reaching the detector. A series of images is obtained with the mask in different positions, and then the data are converted to wavelength-dependent images through a Hadamard transform. (E) Direct imaging: the spectroscopic image is obtained by recording the signal from all points on the sample simultaneously over a narrow spectral interval. A series of images at discrete wavelengths is recorded to provide spectrscopic information for each pixel.

with a CCD. Spectroscopic information is obtained by recording sequential images over a range of optical retardations in the case of the interferometer or a range of frequencies in the case of the tunable filter. In this way, spectra corresponding to various points within the sample are obtained at each detector pixel. Since the image is captured in its entirety, the results are not prone to mechanical positioning errors that may occur in point or line mapping methods. The image quality, in fact, is primarily limited by diffraction. When image shifts occur in direct imaging, they are usually predictable and can be compensated by either experimental or computer procedures.

The basic principles of Raman microscopy and imaging are encapsulated by **Figure 6**, which depicts a Raman image and a spectrum of a 1 μm diameter polystyrene bead. The data were obtained with a microscope coupled to a CCD array; individual images were recorded through an AOTF, which was successively tuned between 747 and 1363 cm$^{-1}$ from the exciting 647 nm $Kr^+$ laser line. The Raman image shown was recorded at 1000 cm$^{-1}$, corresponding to a symmetric aromatic ring vibration in polystyrene. The bright areas correspond to areas rich in polystyrene, thus revealing the distribution of beads in the sample. A spectrum for every pixel is available because images were recorded over a series of wavelengths. **Figure 6** illustrates a spectrum obtained from a single pixel in one of the bead centres.

## Infrared microspectroscopy

Infrared microspectroscopy is based on either transmission or reflection measurements. Transmission is

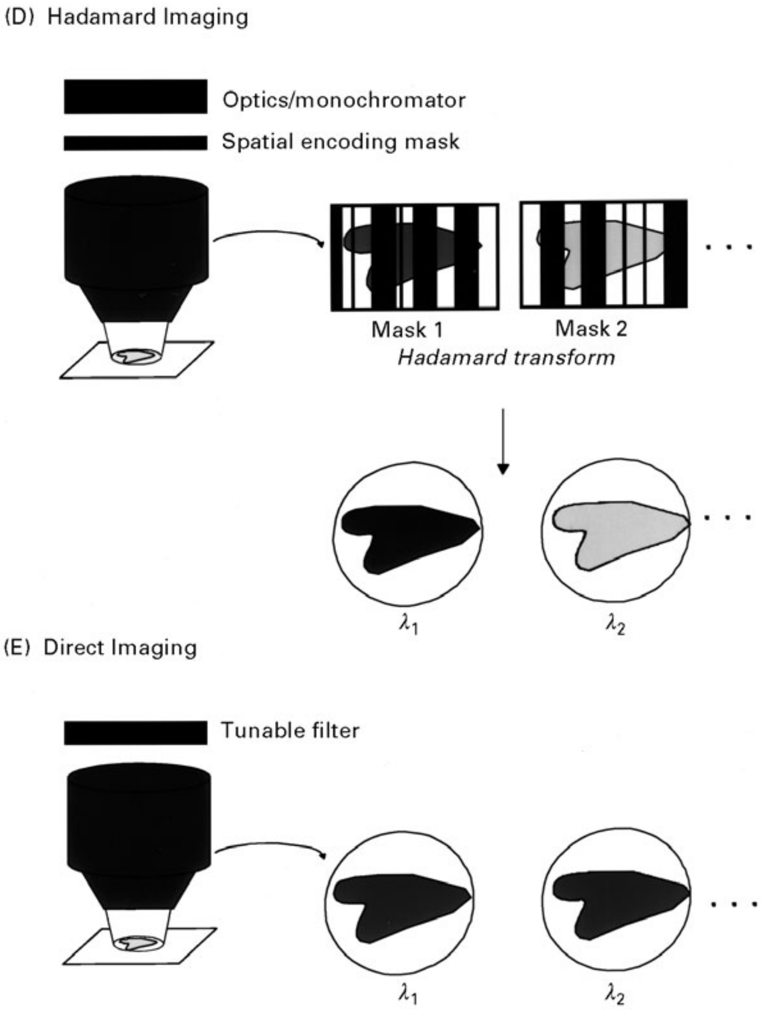



**Figure 4** (*Continued*)

the simplest implementation; the maximum sample thickness is sample and wavelength dependent, however, and is limited to ~15–20 μm for mid-infrared measurements. Reflection measurements are indispensable for samples that are thick or opaque. Various microscope objectives have been tailored for measurements based on the corresponding bulk infrared techniques, such as diffuse reflectance, attenuated total reflectance, and grazing angle reflectance.

Optical effects such as stray light or scattering are treated by placing one or more apertures in the optical train of the infrared microscope. For transmission measurements, one aperture is placed between the condenser and the sample and another is placed between the objective and the image plane. The two apertures are matched in size in order to optimize the signal. For reflectance measurements, only one aperture is used. This type of arrangement is widely employed in infrared spectroscopy, and is known as 'redundant aperturing'.

### Single-point measurements

Most single-point infrared microspectroscopy is implemented with Fourier transform spectrometers, for which the well known multiplex and throughput advantages generally apply. The multiplex advantage arises because all of the input radiation is detected over the entire scan time; however, it applies only when the dominant source of noise is the detector. When the detection is shot-noise limited, as with CCDs, the multiplex advantage does not hold. The throughput advantage for interferometers arises from the use of circular apertures. Dispersive instruments, by contrast, require narrow exit and entrance slits, particularly for higher spectral resolution. Furthermore, a circular aperture provides a more convenient geometry for coupling a microscope to the spectrometer.

Trace amounts of sample can be analysed with infrared microspectroscopy. The technique, for

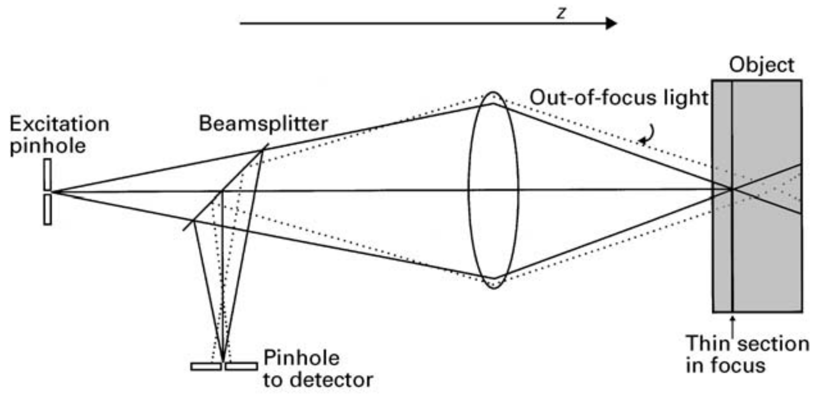



**Figure 5** Confocal microscopy: pinhole apertures placed in the optical train of the microspectrometer lead to rejection of out-of-focus light.

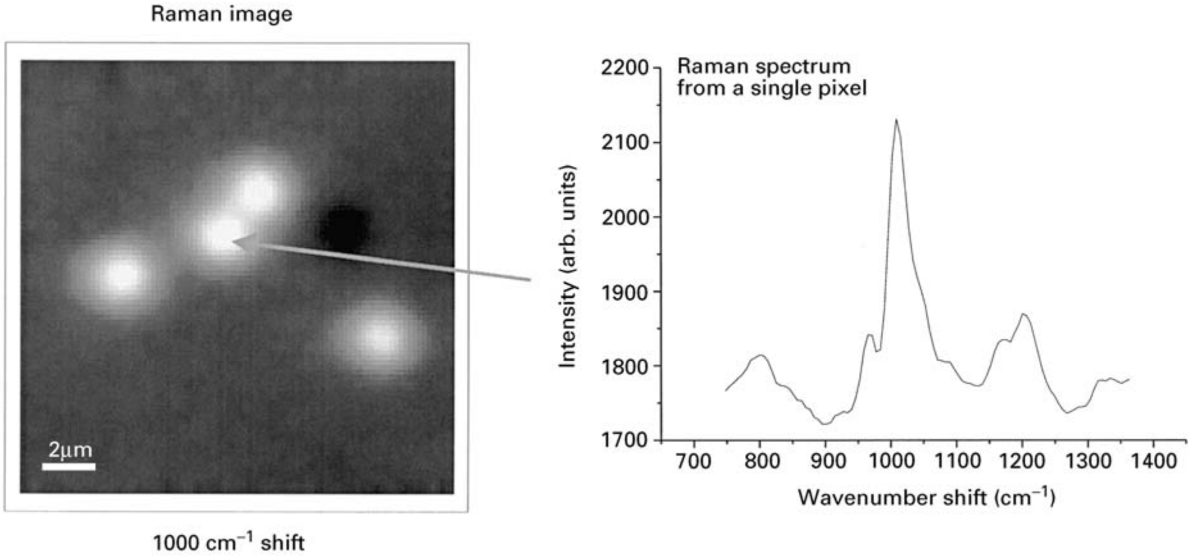



**Figure 6** A Raman spectroscopic image of 1 μm diameter polystyrene beads. The spectroscopic image on the left exhibits bright regions that correspond to the Raman signal at 1000 $cm^{-1}$ shift, a symmetric aromatic ring vibration. A series of images was recorded between 747 and 1363 $cm^{-1}$ shift; a spectrum is then available for every spatial resolution element in the image. The trace on the right corresponds to the spectrum of a single pixel from the centre of a single polystyrene bead.

example, is a standard tool in forensic science. **Figure 7** illustrates fibre spectra that were key evidence in a criminal investigation. A cotton fibre fragment, which was recovered from the nose of a bullet, was matched spectroscopically to fibres from the vest of a police officer who had been shot. Fibre samples represent a system that is difficult to examine with bulk techniques, but which is amenable to vibrational microscopic analysis.

### Mapping and imaging

Infrared mapping and imaging approaches are applied as in the corresponding Raman approaches, with the appropriate choices for source, spectrometer, and detector. For example, a combination of an acousto-optic filter and an InSb focal plane array can be used to record frequency-dependent images in the near-infrared. In the mid-infrared, direct images are recorded with a step-scan interferometer coupled to an infrared focal plane array detector. To illustrate, consider a transmission measurement in which a step-scan interferometer modulates the output from a blackbody source. The infrared radiation is directed on to the sample with the condenser. The transmitted infrared radiation is collected by the objective and then focused on to the array detector. Over the course of the scan, the movable mirror in the

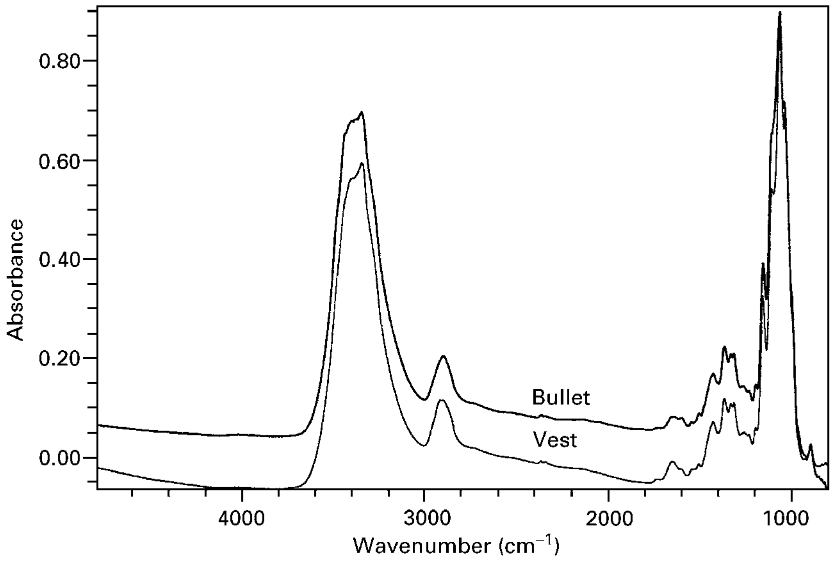



**Figure 7** Infrared microspectra of cotton fibres collected in a forensic application. The spectrum of a fibre retrieved from a bullet (top trace) is seen to match the spectrum of a fibre from the police officer's vest (bottom trace). Data courtesy of John A. Reffner, Spectra-Tech, Inc. and Ronald P. Kaufman, Maine State Police Crime Laboratory.

interferometer is paused for several milliseconds at every step in order to record the image. At the completion of a single scan, a complete interferogram is generated at each detector pixel. The interferograms are then converted to frequency by standard Fourier transform processing.

As an example of the Fourier transform infrared imaging technique, **Figure 8** depicts an infrared spectroscopic image of human breast cells which was recorded with a 64 × 64 mercury–cadmium–telluride (MCT) array. The data set then encompasses 4096 separate interferograms, one for each pixel. Spectra

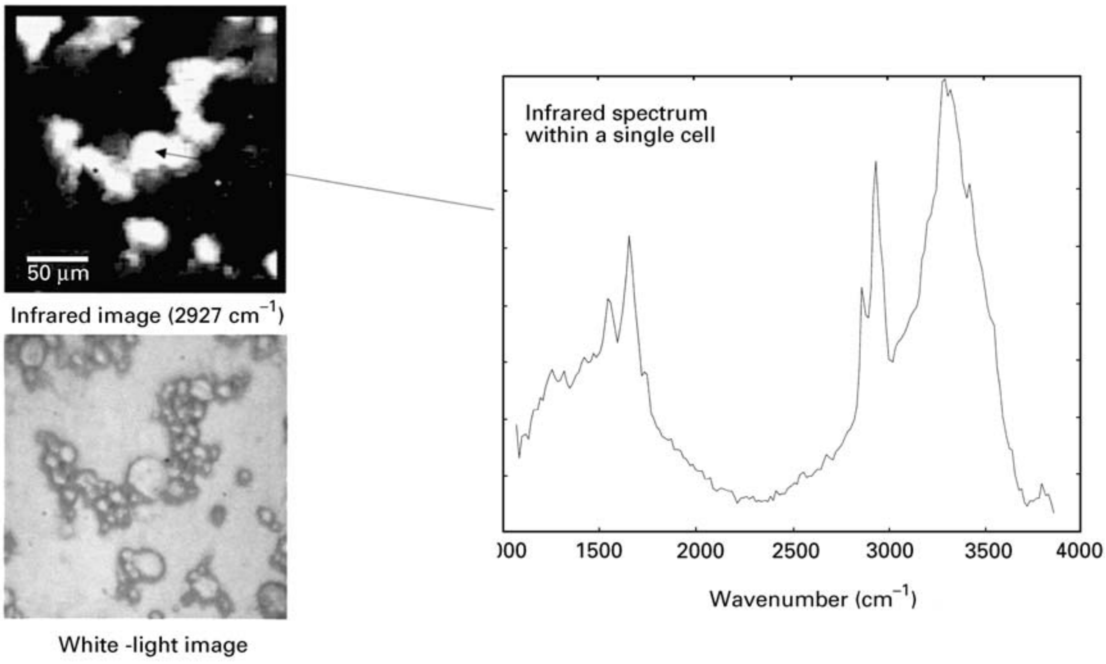



**Figure 8** An infrared spectroscopic image of human breast cells. The spectroscopic image on the left exhibits contrast based on the intensity of the absorption band centred at 2927 $cm^{-1}$ (antisymmetric $CH_2$ stretch), which contains contributions from both the lipid and protein fractions within the cell. Since images are obtained over a contiguous wavenumber interval, it is possible to construct a spectrum for every image pixel. The spectrum extracted from one of the cells is shown on the right.

are obtained through standard Fourier transform procedures. The resulting data cube contains the set of image planes stacked as a function of frequency. Each image exhibits contrast based on the differences in infrared spectral response, which in turn reflects the variation in the chemical composition within the sample. **Figure 8** depicts one such image plane at 2927 $cm^{-1}$, which corresponds to a protein and lipid infrared molecular marker ($CH_2$ antisymmetric stretch). In this way, the distribution of biochemical species can be visualized across the sample. More detailed chemical information can be obtained by examining the infrared spectra that are associated with each pixel in the image.

## Sampling considerations in image generation

The Nyquist theorem specifies that a sinuisoidal function in time or distance can be regenerated with no loss of information as long as it is sampled at a frequency greater than or equal to twice per cycle. The Nyquist theorem must be considered in direct imaging applications because the signal is sampled by the discrete pixel elements in an array. Consider a diffraction limited arrangement with a lateral spatial resolution $R_L$. If the total magnification $M_{tot}$ is a product of the magnifications of the microscope objective $M_{obj}$ and the projection lens $M_{proj}$, the Nyquist theorem requires

$$M_{tot} R_L \geq 2p \quad [5]$$

where $p$ is the pixel size. That is, the sampling interval must be at least twice the highest spatial interval. If the smallest resolvable feature is 5 μm, then each detector pixel must sample intervals that are ≤ 2.5 μm. As long as Equation [5] is obeyed, the spatial fidelity of the microscopic image is preserved and sampling artifacts are avoided. It follows that oversampling does not provide any additional information; this is also known as empty magnification.

*See also:* **IR Spectrometers; Raman Spectrometers; Scanning Probe Microscopes.**

## List of symbols

$n$ = refractive index; $NA$ = numerical aperture; $M$ = magnification; $p$ = pixel size; $r_A$ = radius of the Airy disc; $R$ = distance between two incoherent point sources; $R_L$ = lateral spatial resolution; $\lambda$ = wavelength of light; $\phi_{max}$ = half-angle of the maximum cone of light collected by the lens.

## Further reading


Humecki HJ (ed) (1995) *Practical Guide to Infrared Microspectroscopy*. New York: Marcel Dekker.

Inoué S and Oldenburg R (1995) Microscopes. In: van Stryland EW, Williams DR and Wolfe WL (eds) *Handbook of Optics*, Vol II, pp 17.1–17.52. New York: McGraw-Hill.

James J and Tanke HJ (1991) *Biomedical Light Microscopy*. Dordrecht: Kluwer Academic Publishers.

Katon JE and Summer AJ (1992) IR microspectroscopy. *Analytical Chemistry* **64**: 931A–940A.

Katon JE (1996) Infrared microspectroscopy: a review of fundamentals and applications. *Micron* **27**: 303–314.

Laserna JJ (ed) (1996) *Modern Techniques in Raman Spectroscopy*. Chichester: Wiley.

Lewis EN, Treado PJ, Reeder RC *et al* (1995) Fourier transform spectroscopic imaging using an infrared focal-plane array detector. *Analytical Chemistry* **67**: 3377–3381.

Messerschmidt RG and Harthcock MA (eds) (1998) *Infrared Microspectroscopy: Theory and Applications*. New York: Marcel Dekker.

Puppels GJ, de Mul FFM, Otto C *et al* (1990) Studying single living cells and chromosomes by confocal Raman spectroscopy. *Nature* **347**: 301–303.

Reffner JA (1998) Instrumental factors in infrared microspectroscopy. *Cellular and Molecular Biology* **44**: 1–7.

Rieke GH (1994) *Detection of Light: from the Ultraviolet to the Submillimeter*. Cambridge: Cambridge University Press.

Turrell G and Corset J (eds) (1996) *Raman Microscopy: Developments and Applications*. London: Academic Press.

# Raman Optical Activity, Applications

**Günter Georg Hoffmann**, Universität Essen, Germany

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
Applications

## Introduction

Raman optical activity (ROA), which is most often measured as the circular intensity difference (CID, the difference between a Raman spectrum excited with right circularly polarized light and that excited with left circularly polarized light), is a very small effect which can be easily obscured by artifacts. It can also be reported as the dimensionless circular intensity difference (CID or better DCID, $[I^R-I^L]/[I^R+I^L]$). The first ROA spectrometers were normal scanning Raman spectrometers modified by adding devices to modulate the polarization of the exciting laser beam. Measurements done on this instrumentation were plagued by the named artifacts and took up to 40 h to record a spectrum. Therefore relatively few spectra were measured before the advent of a new generation of spectrometers. Two main technical advances characterize the construction of these sensitive instruments: firstly the development of multichannel detectors culminating in backthinned CCDs which are practically ideal light detectors (with quantum efficiencies of about 80% at the wavelength of peak sensitivity), and secondly the development of the holographically manufactured line filter. The first improvement of shortened the time of measurement to about 20 min for pure samples and allowed the measurement of dilute samples in a few hours, the second allowed the replacement of multiple monochromators by a single polychromator with much higher optical throughput. Later the use of a backscattering arrangement was added to further improve the spectra. As these newer spectra are of much higher quality, mainly the relatively recent literature is covered in this article. For historic data, the interested reader is directed to earlier reviews given in the Further reading section.

In a typical modern ROA instrument, a linearly polarized laser beam is directed through a Pockels cell, which renders the beam circularly polarized. With a frequency of a few hertz the exciting radiation is switched between right circularly and left circularly polarized and focused onto the sample. In the latest spectrometers the Raman effect is observed in the so-called 180° or backscattering arrangement. This method of observation is achieved by drilling a hole into a plane mirror through which the exciting laser beam may pass. The mirror then reflects the backscattered part of the excited radiation through a highly efficient holographically manufactured notch filter to remove the Rayleigh radiation. The Raman radiation, having passed through a Lyot depolarizer to prevent polarization artifacts, can then be processed by a polychromator that consists of a slit and a single grating to finally reach the detector, a backthinned CCD chip.

As in normal Raman spectroscopy, two methods of detection are possible when not using the backscattering arrangement: through a linear polarizer horizontal to the scattering plane (those spectra are called depolarized) or through a vertical polarizer (polarized spectra).

To be useful for structural determinations, ROA spectra have to be compared with calculated spectra. Models confined to special functional groups or semiempirical calculations have been used with limited success for this purpose. During the last few years *ab initio* studies have become possible. These calculations, especially the time consuming ones with large basis sets or even with the inclusion of configurational interaction, allow for useful derivation of absolute configuration or even conformation of the molecules under study when compared with experimental data.

## Stereochemistry of small chiral molecules

One of the simplest chiral molecules one can imagine is bromochlorofluoromethane [1]. As a liquid which until now has only been obtained as an enriched but not pure substance, no crystal structure suitable to deduce the absolute configuration could be obtained. It is therefore a great success for ROA, together with *ab initio* calculations, that it is possible to deduce the absolute conformation of this molecule by comparing experimental and calculated spectra. **Figure 1** shows the depolarized Raman and ROA spectra of CHFClBr [1] (36% enantiomeric excess) together with the calculated spectrum of the (*R*)-enantiomer. Seven of the nine ROA band calculations arrive at the correct sign.

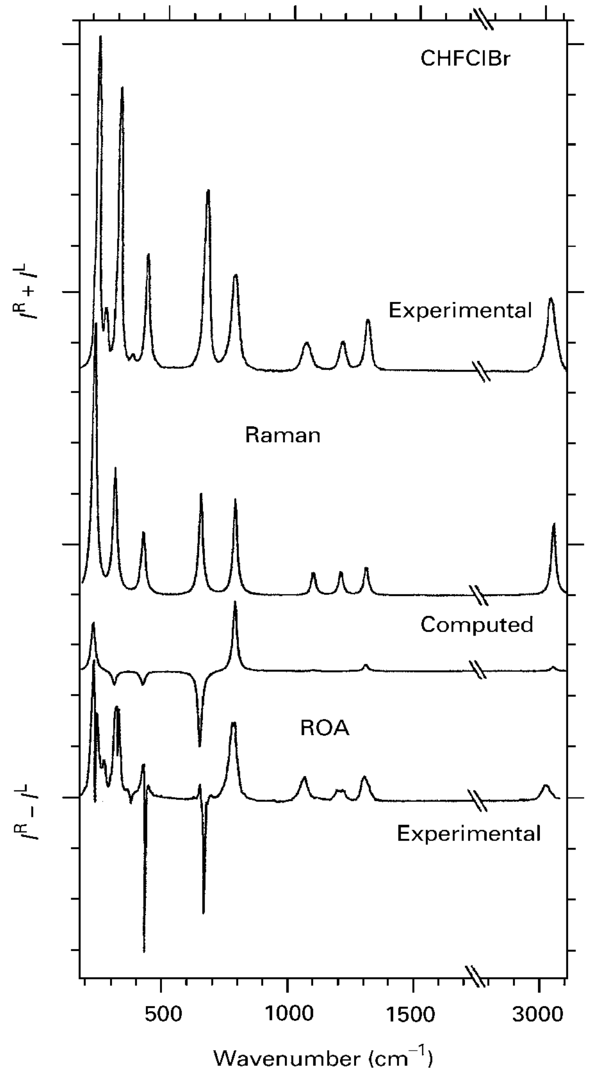
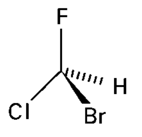
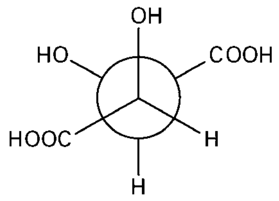



**Figure 1** Depolarized Raman (upper curve) and ROA (lower curve) spectra of (–)-CHFClBr together with the calculated spectra (middle curves) for the (*R*)-isomer. Reproduced with permission of Wiley-VCH from Costante J, Hecht L, Polavarapu PL, Collet A and Barron LD (1997) *Angewandte Chemie International Edition in English* **36**: 885–887. Copyright 1997 Wiley-VCH Verlag GmbH.



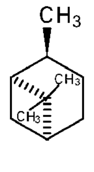
[1]

The well-known molecule (2*R*,3*R*)-(+)-tartaric acid [2] has been studied both as its $d_0$ and $d_4$ isotopomer. As small molecules are less demanding concerning computation time, the molecule could be studied at three theoretical levels: *ab initio* calculations using the 6-31G, 6-31G* and DZP basis sets were performed and compared with the experimental spectra. The molecule was found to exist as a single conformer in aqueous solution, i.e. the conformation with the carboxylic groups *trans* to each other.

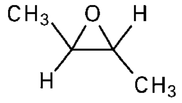


[2]

A comparison between VCD (vibrational circular dichroism) and ROA spectra of the four structurally similar molecules *trans*-pinane [3], *cis*-pinane, α- and β-pinene together with theoretical calculations show that the ratios of the respective parent technique (IR or Raman spectroscopy) to its chiral counterpart are by a factor of 3 in favour of ROA. This advantage is, unfortunately, cancelled by the somewhat more difficult method of measurement. Quantitative comparisons enlighten the supposition that VCD and ROA are highly complementary nonredundant stereochemical techniques.



[3]



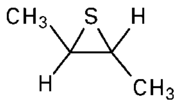
[4]

The experimental and calculated (6-31G* basis set) spectra of (*R*)-(+)-dimethyloxirane [4] in the 200–1500 cm$^{-1}$ region agree for the majority of bands. Depolarized, polarized and 'magic angle' Raman and ROA spectra have been measured. By setting the transmission axis of the analyser at the 'magic angle' of ± 35.26° to the vertical, only contributions from the electric dipole–magnetic dipole polarizability are present in the ROA spectra. As expected, calculations at higher theoretical levels give better results for the calculation of normal modes, but for the calculation of the polarizability derivatives they do not show any remarkable advantage. This is very important for practical reasons, as one can use the less time-consuming calculations for the computation of ROA.

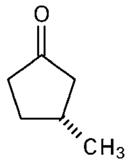
Methyl torsion Raman optical activity of *trans*-2,3-dimethyloxirane has been compared with that of *trans*-2,3-dimethylthiirane [5]. Torsion of the two methyl groups in these molecules can occur as in-phase and out-of-phase combinations. The former can be observed as a very weak and broad Raman band at ~200 $cm^{-1}$ in the dimethyloxirane with positive ROA of medium size, whereas it is observed as a shoulder at ~220 $cm^{-1}$ in the spectrum of the dimethylthiirane with large positive ROA. Contrary to the dimethyloxirane spectra, the dimethylthiirane spectra also contain the out-of-phase torsion: it shows up as a medium size Raman band at ~245 $cm^{-1}$ with zero or small negative ROA. This is in good agreement with *ab initio* calculations and even with the old inertial model for methyl torsions.

From the Raman optical activity of *trans*-2,3-dimethylthiirane the absolute configuration has been determined. Its experimental spectrum is very similar to that calculated (*ab initio*, basis set 6-31G*) for the 2*R*,3*R*-isomer [5] with best agreement in the skeletal vibration region.

The vibrational spectrum of (*S*)-3-methylcyclopentanone [6] has been calculated on the 6-31G and on the 6-31G** level, whereas Raman optical activity has only been predicted on the 6-31G**/6-31G and on the 6-31G level. The calculations with the larger basis set did not show any improvements.



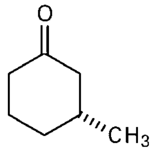

[5]



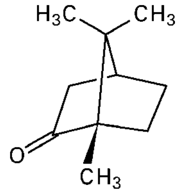

[6]

The ROA spectrum of (+)-3-methylcyclohexanone [7] shows some couplets, which gave rise to much discussion about their origin. Assignments of those vibrations were based on semiempirical calculations with limited reliability. The latest spectrum, however, together with an *ab initio* vibrational analysis, attributes, for example, the couplet at 494 and 518 $cm^{-1}$ to a ring bending mode and an antisymmetric in-plane C=O bending motion coupled to the bending motion of the methyl group relative to the ring, and correctly assigns the (*R*)-configuration to the molecule.

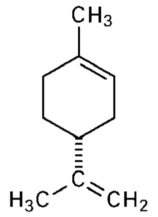
An extensive study has focused on terpenes with structures related to pinane, camphor [8], and limonene [9]. The in-phase dual circular polarization ($DCP_I$) ROA spectra and the normalized CIDs are presented for fourteen compounds. Correlations between ROA features and structural elements of the molecules are discussed. The study clearly demonstrates the advantage of the backscattering measurements, which for theoretical reasons should be about three times as intensive as the normal right-angle incident circular polarization (ICP) measurements.




[7]

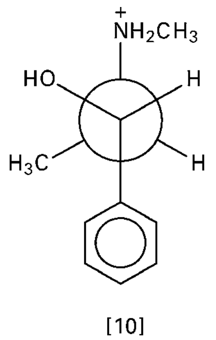



[8]



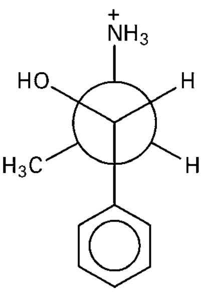

[9]

The Raman and ROA spectra of the hydrochloride salts of four medically applied ephedrine molecules in aqueous solution have been compared in the 700–1700 $cm^{-1}$ region. The salts of (1*S*,2*R*)-ephedrine [10], (1*S*,2*R*)-norephedrine [11], (1*S*,2*S*)-pseudoephedrine [12] and (1*S*,2*S*)-norpseudoephedrine [13] show some bands which with a high degree of




[10]

probability can be used as configurational ROA markers. For example, a band near 840 cm$^{-1}$ (strongly positive for [11–13]) reflects the chirality at C-1 and probably arises from the symmetric CCO stretch, while a band near 910 cm$^{-1}$, which is strongly negative in [11 and 12] arises from the *anti*symmetric CCO stretch at C-1. The chirality of the –C*HCH$_3$(NH$_2$R$^+$) moiety at C-2 may be reflected by the bands of the antisymmetric methyl deformation at 1450 cm$^{-1}$, as they are oppositely signed in the two (1*S*,2*R*)-ephedrines (both positive) and the two (1*S*,2*S*)-pseudoephedrines (both negative).




[11]

## Crystals

Optical activity in crystals is not only the effect of a single molecule but contains also the effects of a large array of similar entities. This is why the crystal ROA is about one order of magnitude larger and one can measure vibrations which cannot be observed in solutions or neat liquids. But one has to be careful not to observe artifacts produced by the linear birefringence of the crystals. An interesting class of crystals are the cubic crystals of the sodium halogenates. These are composed of achiral subunits that form a

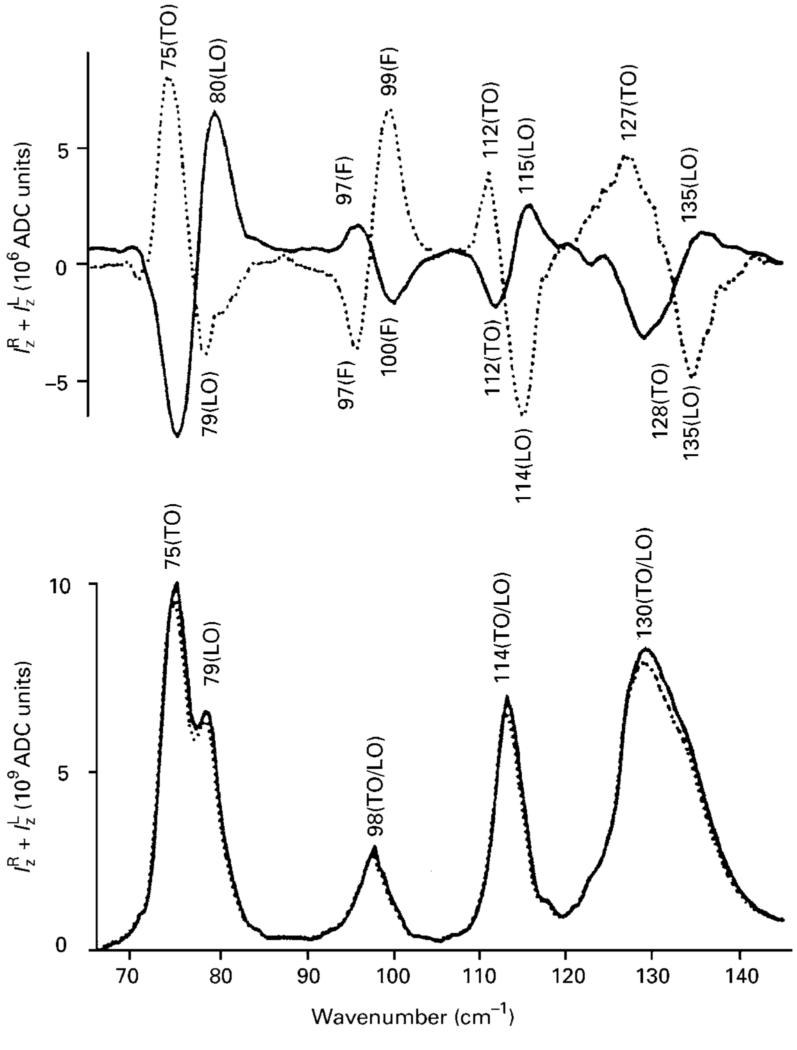
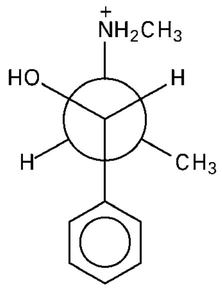
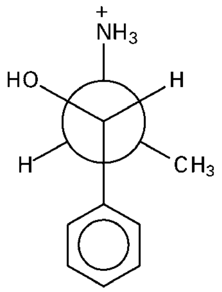



**Figure 2** Depolarized right-angle scattering Raman and ROA spectra for the lattice vibrations of (+)-NaBrO$_4$ (—) and (–)-NaBrO$_4$ (· · · ·), both 1.8 cm$^{-1}$ resolution. Reproduced with permission of John Wiley & Sons from Lindner M, Schrader B and Hecht L (1995) *Journal of Raman Spectroscopy* **26**: 877. Copyright 1995 John Wiley & Sons.



[12]



[13]

chiral array and therefore crystallize in enantiomorphic crystals, belonging to space group $T^4$ ($P2_13$). In the ROA spectra of sodium chlorate and bromate, longitudinal and transverse optical F phonons could be resolved. As an example, the depolarized right angle scattering Raman and ROA spectra of sodium bromate are shown in **Figure 2**. At a resolution of 1.8 $cm^{-1}$ the non-degenerate longitudinal (LO) and doubly degenerate transverse (TO) optical modes can be observed separately. The observed signs can be attributed to one of the two possible chiral arrays.

## Biochemical applications

One of the great advantages of Raman spectroscopy over IR spectroscopy is the ready applicability of water as a solvent. This is of course very important in biochemistry, where the compounds to be studied are, preferably, to be solved in water, as that is their natural medium. Nearly all biochemical compounds are chiral, so it is reasonable to study them not only by Raman spectroscopy but also by ROA, where additional features can be observed or resolved.

### Carbohydrates

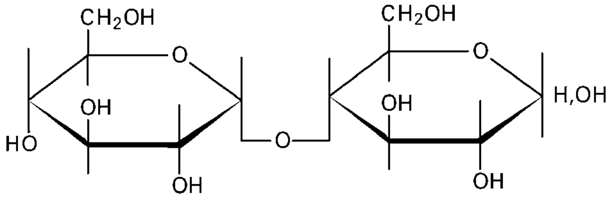
High quality spectra of carbohydrates can be measured in saturated aqueous solution, exhibiting clear bands that can be readily attributed to the orientation of OH substituents, anomeric preference and the conformation of exocyclic $CH_2OH$ groups. As 15 monosaccharides (e.g. arabinose, glucose, xylose, galactose and mannose) are examined, the conclusion that the ROA features reflect local stereochemical details seems to be based on enough material to be reliable.

In di- and polysaccharides ROA can probe the type and conformation of the glycosidic link. Early investigations on D-maltose [14] show a glycosidic couplet at 890–960 $cm^{-1}$ (centred at about 917 $cm^{-1}$) similar to that of D-glucose [15]. It is concluded that for the $\alpha(1\rightarrow4)$ glycosidic link the couplet is positive at lower and negative at higher wavenumbers.

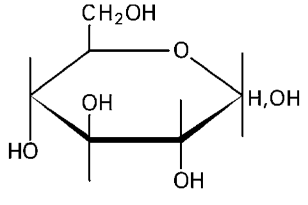
ROA observed at lower wavenumbers is also valuable for the determination of configuration. Compared with D-galactose, the disaccharide D-maltose and the polysaccharide laminarin [16] show a large ROA couplet centred at about 430 $cm^{-1}$. The configuration giving rise to a couplet that is negative at low, but positive at high wavenumbers is a β-glycosidic link (e.g. D-cellobiose, laminarin), whereas an α-glycosidic link (D-maltose) is positive at low, negative at high wavenumbers.

An extension of the investigations to the disaccharides D-maltose, D-maltose-O-$d_8$, D-cellobiose, D-isomaltose, D-gentobiose, D-trehalose and α-D-cyclodextrin confirms these results.

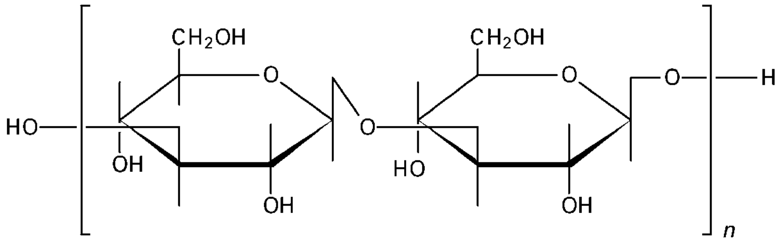
In cyclodextrins (CDs), conformational flexibility is studied by ROA. Using maltoheptose, β-cyclodextrin [17] and its derivatives with two- resp. three-times O-methylated sugar rings, an order of increasing flexibility of the CD ring could be established. Increases in couplet signal strength centred at ~915 $cm^{-1}$ are interpreted in terms of a reduction in conformational flexibility of the cyclodextrin ring. Maltoheptose (linear) is expected to be the most flexible of the compounds, as found in the



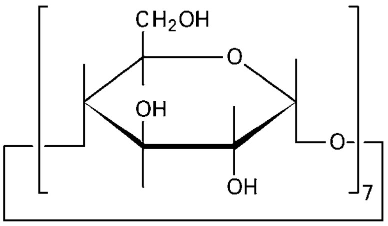
[14]



[15]




[16]



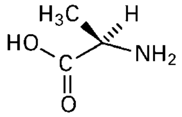

[17]

investigation. Good complexation of a guest in the CDs also lowers flexibility.

Other polysaccharides studied by ROA are laminarin [16] and pullulan [18]. Laminarin, which is a β(1→3) linked polymer of D-glucose, adopts a triple helix conformation. This is concluded by comparison with the ROA of the compound that corresponds to laminarin's dimer subunit, D-laminaribose. Pullulan is a linear polymer of D-glucose, consisting of D-maltotriose units connected through α(1→6) glycosidic links. It adopts a random coil structure in aqueous solution, as can be deduced by the similarity of its ROA spectrum to that of D-maltotriose.

### Amino acids

ROA spectra of L-(*S*)- as well as D-(*R*)-alanine [19] together with a detailed *ab initio* vibrational analysis of their zwitterionic structure at the 6-31G and the 6-31G* level have been reported. Though in the Raman spectra dependence on pH is clearly visible, the ROA spectra are not influenced by $H^+$ concentration. It is concluded from the fact that VCD spectra *do* respond to pH variation that ROA is more directly coupled to chirality and thus can yield more reliable information about absolute configuration.

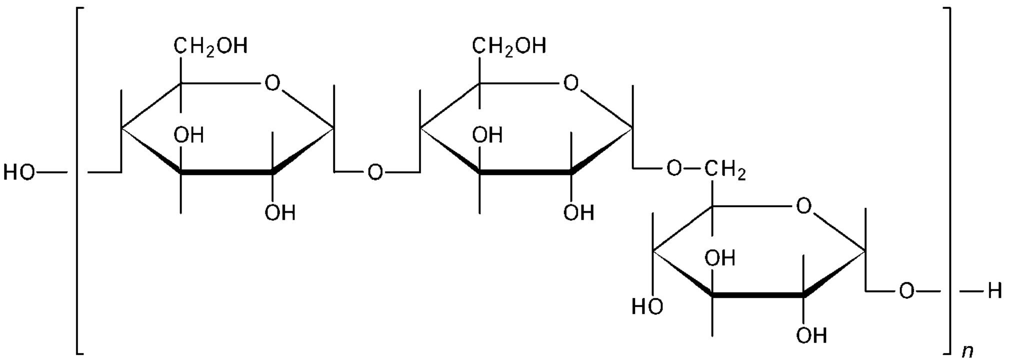



[19]

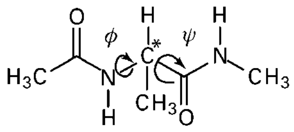
The simple amino acids L-serine, L-cysteine, L-valine, L-threonine and L-isoleucine have been studied in aqueous solution by backscattering, using the spectral range from 600–1600 cm$^{-1}$. Bands originating in $C_\alpha{}^*H$ deformation and the symmetric $CO_2^-$ stretch provide the best signatures for stereochemical studies. The ROA of the former vibration is positive for all acids studied and that for the latter negative. Isoleucine is an exception as it shows a nearly featureless ROA spectrum. The reason may be that it exists as an equal mixture of two rotamers in water.

### Peptides and proteins

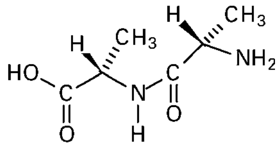
ROA spectra have proven to be of great value to study the secondary structure of proteins. By examining the




[18]




[20]




[21]

four different regions that contain the stretching vibrations of the backbone of proteins (the extended amide III region, the sidegroup and amide II region, and the amide I region), their α helix, β sheet, reverse turn and random coil content can quite readily be determined. This approximate division is shown in the lysozyme part of **Figure 4** (see below). Of course some sidegroups can also be observed by their Raman active vibrations, e.g. lysozyme and α-lactalbumin show tryptophan bands.

As a model for the peptide backbone *N*-acetyl-*N′*-methyl-L-alaninamide (NANMLA) [20] was subjected to a detailed ROA analysis. Its experimental spectrum was compared with the calculated spectra of nine different conformations. Both geometry optimizations and vibrational analyses using *ab initio* methods and the basis set 6-31G* were conducted on these conformers. As a result of these calculations the four conformers $C_5$, $C_{7,eq}$–$C_5$, $C_{7,eq}$ and $\alpha_R$ were found to play an important role in the room temperature equilibrium. These four conformers are characterized by torsion angles of $\phi$ = −157, −85, −99 and −60°, and of $\psi$ = 159, 136, 79 and −40°, respectively; here $\phi$ is the OC–NH–C*–CO backbone torsional angle and $\psi$ is the HN–C*–CO–HN angle. Though the comparison between experimental and calculated spectra does not give unambiguous results, it gives strong hints that the predominant conformer in $H_2O$, $D_2O$ and chloroform solution is the $C_{7,eq}$–$C_5$ conformer and that the $\alpha_R$ conformer is a minor component in $H_2O$ and chloroform solutions, whereas the $C_{7,eq}$ conformer could be present in small amounts in $D_2O$ solution.

The dipeptide L-alanyl-L-alanine [21] shows bands similar to the spectrum of L-alanine. In water, bands between 850 and 950 cm$^{-1}$ can be attributed to the $C^{\alpha}$–N stretch, symmetric $CO_2$ bend and the $C^{\alpha}$–C(O) stretch. In addition, the dipeptide shows a negative ROA band at ~1270 cm$^{-1}$ and a large positive ROA at ~1340 cm$^{-1}$. These vibrations occur in all known protein ROA spectra. According to newer investigations, the Raman band associated with the former effect is not only the amide III vibration but a superposition of three modes, of which only the $C_C$–$H^I$ deformation (which is the first of two possible orthogonal $C_C$–H deformations) contributes to the ROA. The positive ROA at ~1340 cm$^{-1}$ arises from the N–H in-plane deformation coupled to the $C_N$–$H^{II}$ deformation.

Looking at the ROA spectra of α helical and several unordered poly(L-lysine) preparations of increasing relative molecular mass ($M_r$), one not only finds bands at ~931 and 943 cm$^{-1}$, but one also finds a strong sharp positive ROA band at ~1340 cm$^{-1}$, which belongs to loops with a local order of a $3_{10}$ helix connecting regions of α helices. This feature, which is quite small at lower $M_r$ (26 000), becomes more prominent at increasing $M_r$ and becomes the dominant feature in a 268 000 Da protein, thus showing the $3_{10}$ helix content growing at the expense of the α helix.

Bovine serum albumin, which has a high α helix but no β sheet content, shows a strong positive band at 1340 cm$^{-1}$, probably connected to surface loop features forced mainly by its seventeen disulfide linkages. This prevents coupling of modes in this spectral region to other parts of the molecule and now the localized ‘amide III’ modes strongly resemble the vibrations in L-alanyl-L-alanine. The large globular protein ovalbumin, which has only a few disulfide bridges, lacks this large ROA feature.

α-Lactalbumin contains α helices as well as β sheet regions, and therefore shows bands at the corresponding wavenumbers in the 1020–1060 cm$^{-1}$ part of the $C^{\alpha}$–N stretch region. Large changes are induced in the extended amide III region as the hydrogen on nitrogen is replaced by deuterium, simplifying the spectra and thus allowing for easier identification of the remaining features.

It has been shown that ROA can detect residual amounts of structure in mainly unfolded proteins. Investigations on hen egg-white lysozyme and bovine ribonuclease A show that both proteins lose their natural structure to quite a large extent on reducing all the disulfide bonds (shown in **Figure 3** for ribonuclease A) in citrate buffer. The Raman and ROA spectra of the native proteins are shown in **Figure 4**. Lysozyme and ribonuclease A are known to contain both large α helix and β sheet regions, but on denaturation by reduction the two proteins behave differently. Whereas unfolded lysozyme loses nearly all of its structure at room temperature and only regains about 20% of its α helix content when cooled to 2°C, ribonuclease A still has about 50% of its secondary structure at room temperature.

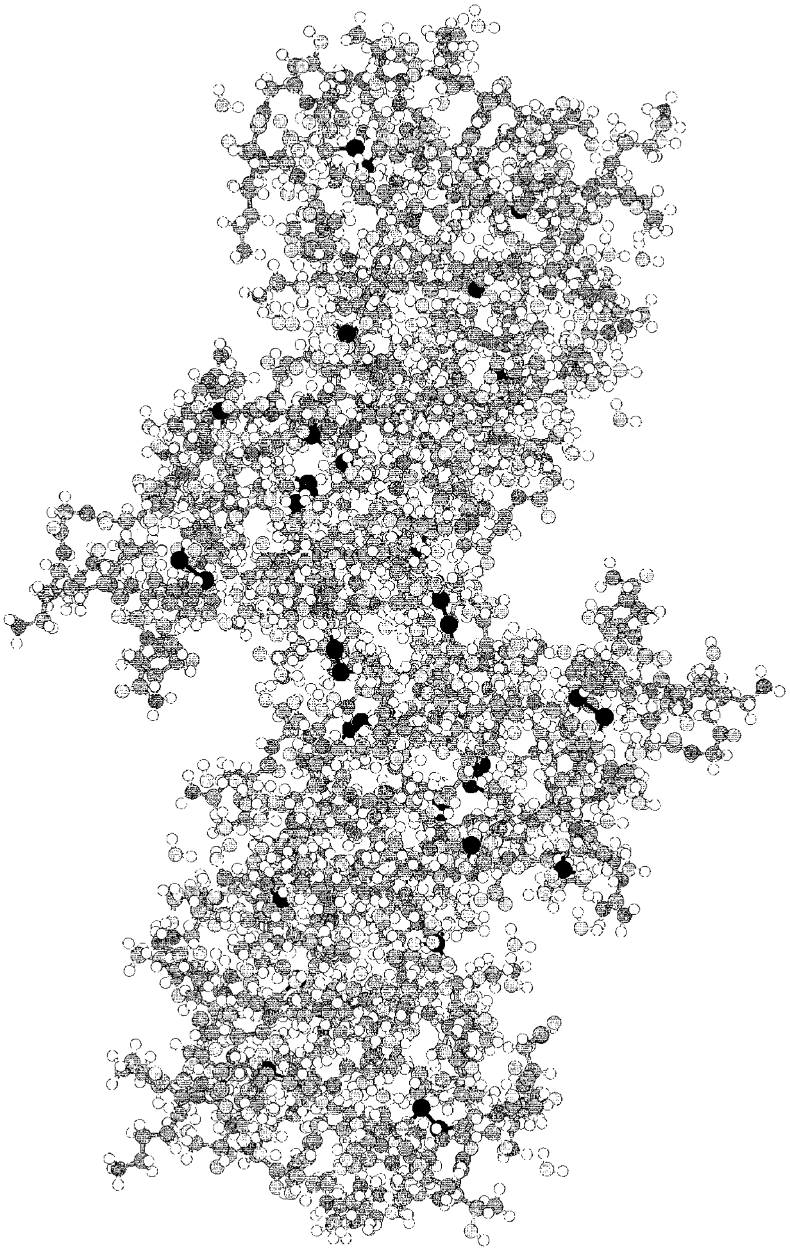


**Figure 3** Bovine seminal ribonuclease, disulfide bridges shown in black. Structure drawn with HyperChem 5.0 from PDB entry 1BSR of Mazzarella L, Capasso S, Demasi D, Di Lorenzo G, Mattia CA and Zagari A (1993) *Acta Crystallographica Section D* **49**: 389; The Protein Data Bank: Bernstein FC *et al* (1977) *Journal of Molecular Biology* **112**: 535–542.

A further study of native hen egg-white lysozyme, using the pH 5.4 buffer solution spectra of the molecule in the temperature range 2–50°C, revealed a new cooperative transition at ~12°C. A detailed analysis shows the protein's adoption of a more rigid structure, where tertiary loops and tryptophan side-groups lose their residual mobility.

The human $\alpha_1$-acid glycoprotein orosomucoid shows ROA bands that are characteristic of a high β-sheet content. It is to be expected from this single reported example that the ROA spectra of glycoproteins will yield valuable information about the conformation of the molecule, especially upon the mutual influence of the different protein and carbohydrate parts.

Another human protein that has been studied by ROA methods is the human serum albumin. Most prominent features in its ROA spectrum are bands

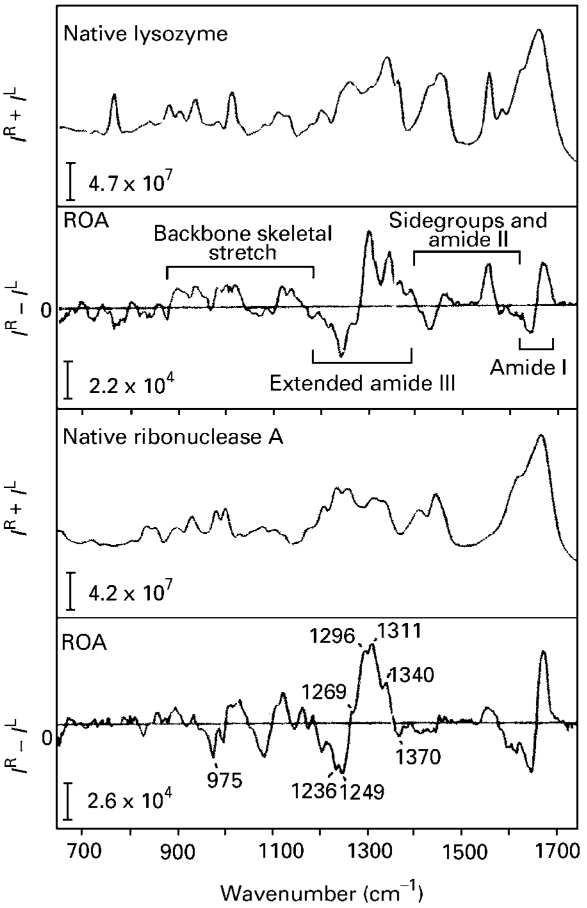
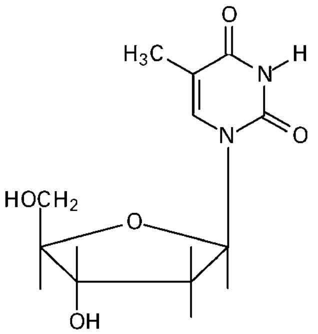



**Figure 4** Backscattered Raman and ROA spectra of native hen egg-white lysozyme (top pair) and of native bovine ribonuclease A (bottom pair), both in acetate buffer at pH 5.4. Reprinted with permission from Wilson G, Hecht L and Barron LD (1996) *Biochemistry* **35**: 12 518–12 525. Copyright 1996 American Chemical Society.

that arise from the large α helix content and the already discussed strong positive band at ~1340 $cm^{-1}$ which is characteristic for loop structures with local order of a $3_{10}$ helix. The latter band decreases to ~40% on reducing the pH to 3.4, where the molecule assumes its F state, which is a mildly denaturated form created by dissociation of the two halves of the heart-shaped molecule. These findings may be explained by a loss of rigidity of the molecule and are in accordance with X-ray analyses of the crystalline state and derived structure proposals for the F state.



[22]

### Nucleosides, nucleotides and nucleic acids

Pyrimidine nucleosides have been studied in backscattering. Unlike in normal Raman spectroscopy, where the sugar moiety only gives rise to weak signals, ROA detects signals of comparable intensity as well from the sugar as from the interaction of the sugars with the bases, but there are no signals from nearly achiral vibrations localized in the planar base rings. Of great diagnostic value are signals arising from the stretching at the C(1′)–N(1) glycosidic bond. These signals (~1200, 1230, 1380 and 1400 $cm^{-1}$) have opposite signs for α- and β-thymidine [22].

Of the three nucleotides adenylic acid, uridylic acid and cytidylic acid the three single-stranded polyribonucleotides as well as two double-stranded nucleotides were examined by ROA spectroscopy. Three regions were especially valuable for the analysis: 1550–1750 $cm^{-1}$ (called the base stacking region), 1200–1550 $cm^{-1}$ (sugar–base coupling region) and 950–1150 $cm^{-1}$ (sugar–phosphate backbone region). As the name suggests, the first region reflects base stacking. The second shows the orientation of the sugar and base rings. The third reflects the conformation of the sugar ring, even though the backbone may also be involved. Typical spectra in $H_2O$ as well as in $D_2O$ are shown in **Figure 5**. These spectra of polyadenylic acid show two large negative peaks in the sugar base region in the ROA. As expected, the base-dependent signatures lose much of their intensity on exchange with $D_2O$, whereas the sugar phosphate backbone region remains virtually unaltered.

## Other applications

Apart from probing molecular conformation and configuration, ROA has been employed in a couple of other interesting tasks.

The applicability of ROA for the determination of enantiomeric excess in mixtures of chiral enantiomers has been exploited. Using the test compound α-pinene an accuracy of 0.1% enantiomeric excess (ee) was achieved.

The possibility of observing Raman optical activity from chiral surfaces has been mentioned and a theory for the generation of second-harmonic optical activity (SHOA) from chiral surfaces and interfaces has been derived.

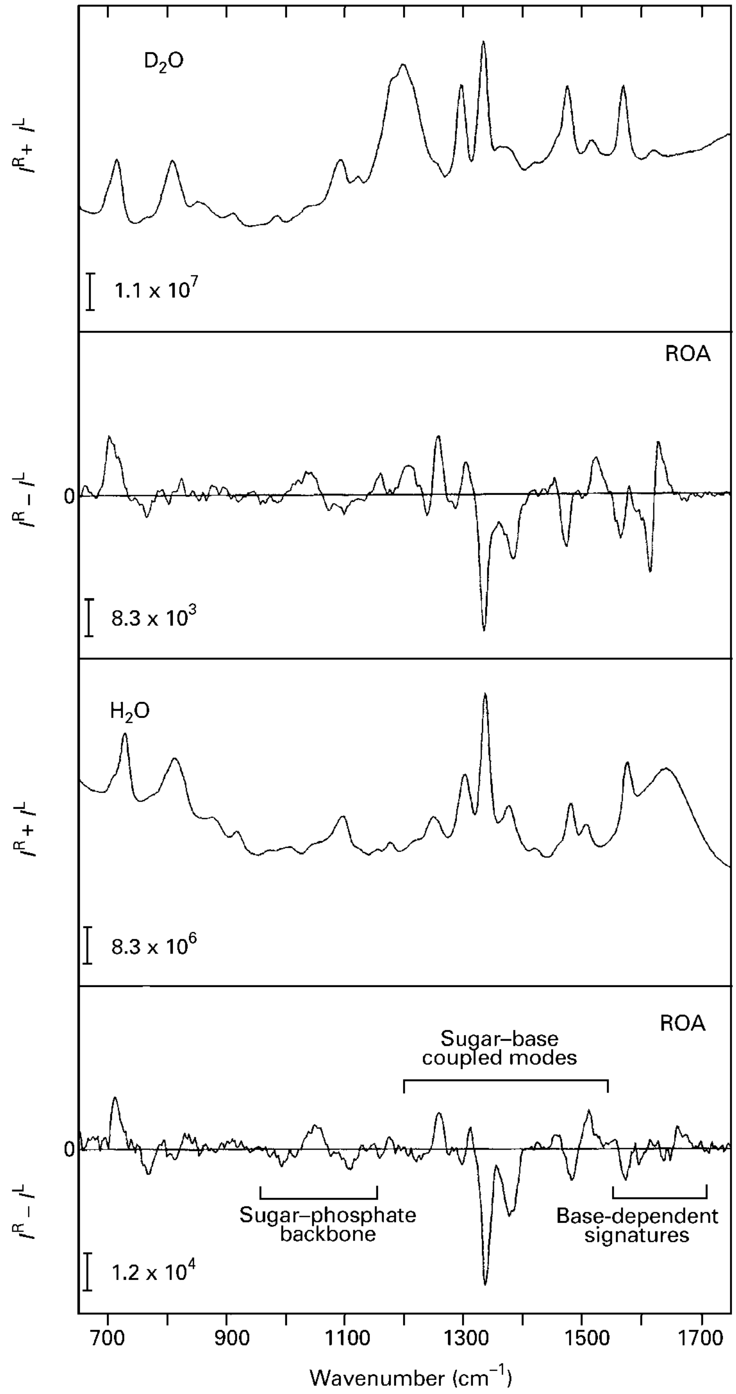



**Figure 5** Backscattered Raman ($I^R$+$I^L$) and ROA ($I^R$–$I^L$) spectra of poly(rA) in $H_2O$ (bottom) and $D_2O$ (top). Reprinted with permission from Bell AF, Hecht L and Barron LD (1997) *Journal of the American Chemical Society* **119**: 6006–6013. Copyright 1997 American Chemical Society.

Even theoretical aspects may be attacked by ROA spectroscopy. In benzene derivatives containing heteroatoms the question arises as to whether Rydberg transitions centred on the heteroatom are involved in the production of Raman spectra. Using a comparison between the polarized and depolarized ROA spectra, this contribution of Rydberg transitions could be detected in l-phenylethanol and l-phenylethylthiol, as large deviations from the factor of 2 in the ratio of polarized to depolarized spectra were found.

Protein folding is an area of much recent activity. It has been shown by ROA that water acts as lubricant in helix coil unfolding. Extremely fast conformational

fluctuations seem to occur with a frequency of about $10^{12}$ $s^{-1}$. These turnover rates, which are close to the theoretical limit of kinetics, could only be observed owing to the very short time scale (~$10^{-14}$ s) of Raman techniques.

The very high quality of ROA spectra makes it possible even to obtain difference ROA spectra of substances at two different temperatures. By this method the premelting of poly(rA)·poly(rU) was monitored. The compound, which has an A-type double helical structure, shows qualitatively the same ROA spectrum at both 20 and 45°C, while the intensities of the bands are lowered. From the fact that the difference spectrum is similar to the ROA spectra at both temperatures, it is concluded that the same average structure is maintained throughout all temperatures in the range examined. The investigation demonstrates the usefulness of the new technique to probe the dynamics of nucleic acids in aqueous solution.

## List of symbols

$I^L$ = intensity of left circularly polarized light; $I^R$ = intensity of right circularly polarized light, $\phi$ and $\psi$ = torsional angles.

*See also:* **Biochemical Applications of Raman Spectroscopy**; **Carbohydrates Studied By NMR**; **Chiroptical Spectroscopy, Emission Theory**; **Chiroptical Spectroscopy, General Theory**; **FT-Raman Spectroscopy, Applications**; **Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy**; **IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds**; **IR Spectral Group Frequencies of Organic Compounds**; **Matrix Isolation Studies By IR and Raman Spectroscopies**; **Non-linear Raman Spectroscopy, Applications**; **Non-linear Raman Spectroscopy, Instruments**; **Non-linear Raman Spectroscopy, Theory**; **Nucleic Acids and Nucleotides Studied Using Mass Spectrometry**; **Nucleic Acids Studied Using NMR**; **Polymer Applications of IR and Raman Spectroscopy**; **Proteins Studied Using NMR Spectroscopy**; **Raman and IR Microspectroscopy**; **Raman Optical Activity, Spectrometers**; **Raman Optical Activity, Theory**; **Raman Spectrometers**; **Stereochemistry Studied Using Mass Spectrometry**; **Vibrational, Rotational and Raman Spectroscopy, Historical Perspective**.

## Further reading


Barron LD and Hecht L (1994) Vibrational Raman optical activity: from fundamentals to biochemical applications. In: Nakanishi K, Berova N and Woody RW (eds) *Circular Dichroism – Principles and Applications*, pp 179–215. New York: VCH.

Barron LD, Hecht L and Polavarapu PL (1989) Polarized Raman optical activity in methyl antisymmetric deformation: influence of heteroatom Rydberg orbitals. *Chemical Physics Letters* **154**: 251–254.

Barron LD, Hecht L and Wilson G (1997) The lubricant of life: a proposal that solvent water promotes extremely fast conformational fluctuations in mobile heteropolypeptide structure. *Biochemistry* **36**: 13 143–13 147.

Bell AF, Hecht L and Barron LD (1997) Vibrational Raman optical activity of pyrimidine nucleosides. *Journal of the Chemical Society, Faraday Transactions* **93**: 553–562.

Bell AF, Hecht L and Barron LD (1997) Vibrational Raman optical activity as a probe of polyribonucleotide solution stereochemistry. *Journal of the American Chemical Society* **119**: 6006–6013.

Bell AF, Hecht L and Barron LD (1997) New evidence for conformational flexibility in cyclodextrins from vibrational Raman optical activity. *Chemistry – A European Journal* **3**: 1292–1298.

Costante J, Hecht L, Polavarapu PL, Collet A and Barron LD (1997) Absolute configuration of bromochlorofluoromethane from experimental and ab initio theoretical vibrational Raman optical activity. *Angewandte Chemie International Edition in English* **36**: 885–887.

Hecht L and Barron LD (1994) Rayleigh and Raman optical activity from chiral surfaces. *Chemical Physics Letters* **225**: 525–530.

Hecht L, Philips A and Barron LD (1995) Determination of enantiomeric excess using Raman optical activity. *Journal of Raman Spectroscopy* **26**: 727–732.

Hoffmann GG (1995) Vibrational optical activity (VOA). In: Schrader B (ed) *Infrared and Raman Spectroscopy – Methods and Applications*, pp 543–572. Weinheim: VCH.

Lindner M, Schrader B and Hecht L (1995) Raman optical activity of enantiomorphic single crystals. *Journal of Raman Spectroscopy* **26**: 877–882.

Polavarapu PL (1990) Ab initio vibrational Raman and Raman optical activity spectra. *Journal of Physical Chemistry* **94**: 8106–8112.

Qu X, Lee E, Yu G-S, Freedmann TB and Nafie LA (1996) Quantitative comparison of experimental infrared and Raman optical activity spectra. *Applied Spectroscopy* **50**: 649–657.

Teraoka F, Bell AF, Hecht L and Barron LD (1998) Loop structure in human serum albumin from Raman optical activity. *Journal of Raman Spectroscopy* **29**: 67–71.

Wilson G, Hecht L and Barron LD (1996) Residual structure in unfolded proteins revealed by Raman optical activity. *Biochemistry* **35**: 12 518–12 525.

Wilson G, Hecht L and Barron LD (1997) Evidence for a new cooperative transition in native lysozyme from temperature-dependent Raman optical activity. *Journal of Physical Chemistry, B* **101**: 694–698.

Yu G-S, Freedman TB and Nafie LA (1995) Dual circular polarization Raman optical activity of related terpene molecules: comparison of backscattering $DCP_I$ and right-angle ICP spectra. *Journal of Raman Spectroscopy* **26**: 733–743.

# Raman Optical Activity, Spectrometers

**Werner Hug**, University of Fribourg, Switzerland

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
Methods & Instrumentation

Raman optical activity (ROA) or, more precisely, spontaneous vibrational Raman optical activity scattering is, like vibrational circular dichroism (VCD), a spectroscopic method that directly probes the chirality, or handedness, of molecular vibrations. ROA and VCD therefore have an obvious stereochemical potential. That such phenomena could yield structural information not otherwise available was realized long before their measurement became feasible and the first observations of ROA and VCD date back only a quarter of a century. For ROA, measurement was preceded by a detailed theoretical analysis and, perhaps inevitably so in view of the experimental difficulties, some false claims of its observation.

Considerable progress on the collection of ROA data has been made since, but while at present such data can be reliably recorded for many samples it would be far from true to claim that the experimental situation is satisfactory. The currently most successful instruments are based on a ROA backscattering configuration first described at the start of the 1980s and characterized in ROA shorthand as 'ICP'. However, the field of ROA instrumentation is in full motion, and at this time it is not obvious which of several competing ROA variants will eventually emerge as the dominant method for routine ROA data collection.

We can conveniently divide the methodical advances, demonstrated or proposed, in the measurement of ROA into two categories. The first, and most extensively implemented, simply reflects the progress in general Raman instrumentation: multichannel detection systems based on backthinned CCD (charge coupled device) technology, and high luminosity spectrographs. The second consists of refinements specific to ROA. On one hand these are light gathering and data acquisition techniques which have no counterpart, or are of no interest, in ordinary spontaneous Raman spectroscopy. A typical example is a dual lens light collection system dating from the late 1970s. On the other hand are new ROA variants described in the late 1980s and early 1990s which have opened up new ways to collect ROA data, but so far it has not been shown convincingly that one of them may be better suited than the originally proposed techniques to conquer the two great enemies of ROA measurements: statistical noise and systematic spurious scattering differences.

## ROA variants

The reason why there is a whole set of different possibilities for measuring ROA arises from the fact that, as in any light scattering experiment, diffused light can be observed at different angles to the propagation direction of the incident, or exciting light, and that the intensity of the observed light depends on the polarization of the incident light and the eventual use, and the nature of, a polarization analyser for the scattered light. In all cases, however, a difference in the interaction of right and left circularly polarized light – chiral light – with chiral molecules of the sample is observed, i.e. there is what the chemist calls a diastereoisomeric interaction difference.

### Basic measurement arrangements

The scattering geometries of practical importance are right angle, or 90° scattering and collinear, or 0° and 180° scattering. The 0° and 180° scattering geometries yield totally different ROA information, and the magnitude of their ROA signals differs strongly. In general, ROA signals measured at 180° are larger than those measured at 0°, but this can depend on the particular vibrational Raman band under observation. For 90° scattering, ROA signal strengths are somewhere in between. Although the signal strength is lower at 90° than at 180°, backward scattering is not always more advantageous than right-angle scattering.

For any chosen scattering angle, ROA can be observed if either the incident light, or the scattered, polarization analysed and detected light, or both, are circularly polarized. In accordance with this one uses the terms ICP (incident circularly polarized), SCP (scattered circularly polarized) and DCP (dual circularly polarized). Two DCP variants obviously exist: in-phase ($DCP_I$) and out-of-phase ($DCP_{II}$), where in-phase means that if the incident light is right circularly polarized then the right circularly polarized component of the scattered light is detected, while in an out-of-phase experiment one would observe the

left circularly polarized component. Moreover, in ICP the scattered light can, or cannot, be polarization analysed with a linear polarization analyser, and in SCP the incident light can be natural or linearly polarized. Of the large number of ensuing experimental configurations **Table 1** shows those for which there are solid arguments that they will play a definite role in ROA instrumentation, either because of inherent advantages they have in measuring ROA or because they yield specific information not otherwise available.

### Comparison of data from different configurations

In the general non-resonant case there are three contributions to the observed ROA spectra: $\alpha G'$, the isotropic ROA invariant stemming from the electric dipole–magnetic dipole optical activity tensor, which is also responsible for the anisotropic invariant $\gamma^2$, and the anisotropic invariant $\delta^2$ due to the quadrupole transition tensor. The anisotropic invariants are also often written as $\gamma^2 = \beta(G')^2$ and $\delta^2 = \beta(A)^2$.

Relevant quantities of interest, e.g. differential scattered intensities $\Delta I$, are obtained from the formulae in **Table 1** by an appropriate choice of $K$. The difference between the intensity of right (R) and left (L) circularly polarized scattered light in a backscattering SCP experiment with unpolarized (u) exciting light from Equation [8] is

$$\Delta I = I_R^u - I_L^u = \frac{4K}{c}[24\gamma^2 + 8\delta^2] \qquad [10]$$

$$K = \frac{1}{90}\left(\frac{\omega^2\mu_0 E^{(0)}}{4\pi R}\right)^2 \qquad [11]$$

where $\omega$ is the angular frequency of the scattered light, $E^{(0)}$ is the magnitude of the electric field at the site (origin) of the scattering molecule, and $R$ is the distance to the observer.

The way in which the three invariants are associated is different for different scattering geometries. Thus, data measured for forward, backward and right-angle scattering cannot be directly compared with each other, even qualitatively. On the one hand, from a purely analytical point of view this is unfortunate. If very small samples only are available, then 90° is the only choice, as ROA requires fair sized scattering volumes and avoidance of skew angle reflections and refractions of the scattered light by the walls of the sample cell, while the measurement of dilute solutions is best performed by 180° scattering because of its higher values for the ratio $\Delta I/I$. On the other hand, from the point of view of the amount of information which can be extracted from ROA measurements this fact is fortunate, and there is, in my opinion, therefore a strong case for instruments which measure ROA under more than one scattering angle.

## Why ROA measurements are difficult

### Signal-to-noise considerations

Before lasers were available the recording of a Raman optical activity spectrum would have been impossible. To appreciate the signal-to-noise (S/N) problem it suffices to consider the typical depolarized 90° ICP variant. From Equations [1] and [6] in **Table 1**, neglecting the often small quadrupole contribution, one has

$$\frac{\Delta I}{I} \approx \frac{4\gamma^2}{c\beta^2} \qquad [12]$$

**Table 1** ROA variants most likely to play a role in future ROA instrumentation

| *Scattering angle* (°) | *Type* | *Polarization of exciting light* | *Polarization analyser detected light* | *Average quantity* | | *Difference (right minus left) quantity* | |
|---|---|---|---|---|---|---|---|
| 90 | ICP | ○ | ∥ | $K\,6\beta^2$ | [1] | $4\,(K/c)\;[6\gamma^2\text{-}2\delta^2]$ | [6] |
| | | | ⊥ | $K\,[45\,\alpha^2+7\beta^2]$ | [2] | $4\,(K/c)\;[45\alpha G'+7\gamma^2+\delta^2]$ | [7] |
| 180 | ICP | ○ | None | $K\,[90\,\alpha^2+14\,\beta^2]$ | [3] | $4\,(K/c)\;[24\gamma^2+8\delta^2]$ | [8] |
| | SCP | Unpolarized | ○ | | | | |
| | $DCP_I$ | ○ | ○ | $K\,12\,\beta^2$ | [4] | | |
| 0 | ICP | ○ | None | $K\,[90\,\alpha^2+14\,\beta^2]$ | [3] | $4K/c\,[180\alpha G'+4\gamma^2-4\delta^2]$ | [9] |
| | SCP | Unpolarized | ○ | | | | |
| | $DCP_I$ | ○ | ○ | $K\,[90\,\alpha^2+2\,\beta^2]$ | [5] | | |

All formulae are for the far from resonance limit only. ○: circularly polarized. ∥, ⊥: linearly polarized parallel and perpendicular to the scattering plane respectively. Other symbols explained in the text.

The value of Equation [12] is therefore at most a few times the ratio of a magnetic to an electric dipole transition moment. The largest observed values of $\Delta I/I$ are of the order of $5\times10^{-3}$, and values of $10^{-3}$ or less are more common.

In the shot noise limited case, where $I$ is represented by $N$ events, the statistics of a Poisson distribution hold. With $I_{\parallel}^{R} \approx I_{\parallel}^{L} \approx I$ the rms deviation of $\overline{\Delta N}$ is given by $\sqrt{2N}$. Thus, to recover $\overline{\Delta N} = a\overline{N}$ with a given S/N ratio one has

$$\overline{\Delta N} = a\overline{N} = (\mathrm{S/N})\sqrt{2N}$$
$$\overline{N} = 2\frac{(\mathrm{S/N})^2}{a^2} \qquad [13]$$

For a typical situation in ROA one may have $a = 5\times10^{-4}$, S/N = 10, which corresponds to some $10^9$ detected photons per spectral element of resolution for what might be, in ordinary Raman spectroscopy, an unremarkable depolarized Raman band. This number is put in perspective by comparing it with the value of $5.5\times10^4$ detected photons per second obtained in backscattering at 4 cm$^{-1}$ resolution with a modern commercial Raman instrument, equipped with a CCD detector and a 15 mW He-Ne laser, for the peak height of the intense polarized 992 cm$^{-1}$ band of benzene.

With modern slow-scan CCD detectors, read-out noise in a well-designed ROA spectrometer is negligible in comparison with photon shot noise. Yet, flicker, or $1/f$ noise, clearly plays its role. Unfortunately, to date no serious attempt to identify and reduce its influence has been made. What is well known since the first recording of entire ROA spectra in the mid 1970s is that dust can have a deleterious effect. In particular in microcapillaries, where sample volumes are small and convection essentially absent, single grains of dust sometimes appear to get trapped and gently oscillate in the modulated focused laser beam. Though not understood at that time, it probably was one of the first observations of the optical tweezer effect. Thus, careful sample preparation is important.

### Systematic spurious scattering differences

Also called artefacts, spurious scattering differences are arguably the biggest and best discussed nuisance in ROA spectroscopy. They can easily be understood and appreciated without any elaborate theoretical treatment, and the two optical solutions to the problem discussed in the section on light collection optics are based on simple practical considerations.

The easiest viewpoint is to think of a perfect, artefact-free instrument, and then to try to find means to systematically spoil its performance. How best to degrade instrument performance can depend on the ROA variant but the basic ideas are the same. We will, therefore, consider only the ICP backscattering case and for the sake of simplicity assume a fully polarized Raman band, but the insight gained is not limited to this particular situation. The sample is assumed to be achiral. Thus, our perfect instrument registers no scattering difference if modulation between right and left circularly polarized exciting light occurs.

Initially, we now assume the polarization of the exciting light remains perfect in the scattering zone, but we introduce a circular polarization analyser, which transmits only one circular polarization component, somewhere between the scattering zone and the detector. As the scattered light is purely left and right circularly polarized, huge scattering differences are obviously observed for the two modulation periods. Unfortunately, the elements for a partial circular polarization analyser can be present in a ROA instrument: birefringence in the walls of the scattering cell and in the optics, and polarization-dependent diffraction efficiency of the grating of the spectrograph, among other things.

In a second step we spoil the circular polarization of the exciting light in the scattering zone by placing a quarter-wave plate between it and the circular polarization modulator. The light in the scattering zone now becomes linearly polarized in orthogonal planes for the two modulation periods, as is also true for the scattered light. Again, large scattering differences will be observed due to the polarization-dependent efficiency of mirrors and gratings in the light collection and analysing system. A partial quarter-wave plate ahead of the scattering zone can easily be formed in a ROA instrument by birefringence in the windows of the scattering cell and in lenses, as well as by imperfections in the circular polarization modulator.

The simultaneous measurement of several Raman bands and the general increase in measurement speed have, in practical terms, been important for artefact reduction, perhaps more so than optical advances because of their easier implementation. Taken together these two advances allow one to check and adjust the baseline before, and if need be, during a ROA measurement.

## Instrument configurations

### General building blocks

All ROA instruments share the common features depicted schematically in **Figure 1** regardless of the

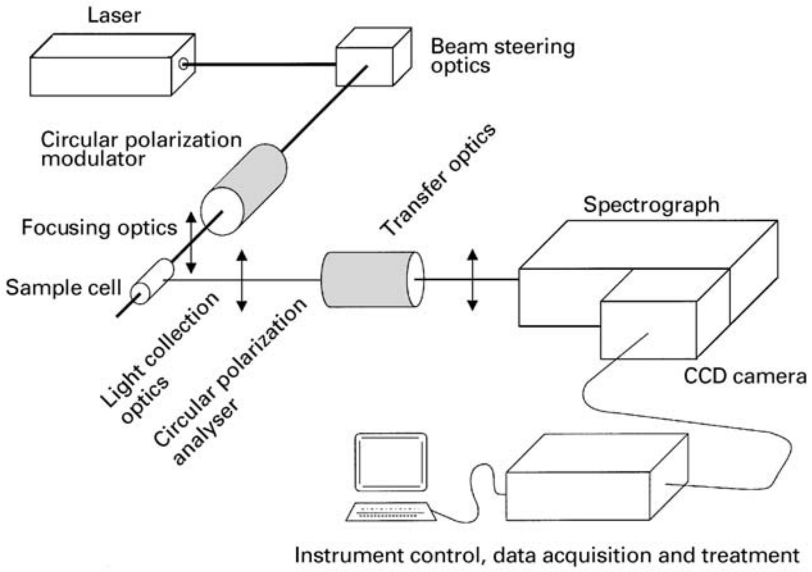



**Figure 1** Building blocks of Raman optical activity spectrometers. Either a circular polarizer (in ICP), analyser (in SCP) or both (in DCP) may be present.

particular ROA variant they measure. Many of these building blocks are also part of ordinary Raman instrumentation, but for ROA there are specific, often more exacting, demands.

**Lasers** ROA requires laser powers of 100 to 1000 mW at the sample, with the upper limit determined by what a sample can stand without serious degradation rather than by what would be desirable for the experiment. In view of the light losses by polarizers, modulators, beam splitters, lenses and mirrors, about 1.5 W of single line output power from the laser is desirable.

The wavelength of the laser has to meet two contradictory requirements. It should be as short as possible so that ROA intensities are maximized, but long enough so that absorption by the sample and the well-known fluorescence problem of Raman spectroscopy are avoided. The standard exciting wavelengths used to date are the 488 nm and, particularly, the 514.5 nm argon ion laser line. Attempts have been made to use the 647.1 nm krypton ion laser line but they proved unsuccessful. From an inspection of Equation [11] it transpires that, for a fixed intensity of the exciting light, the intensity of the scattered light varies as the fourth power of $1/\lambda$. For the ROA difference intensities, however, another factor of $1/\lambda$ is hidden in the expressions of the tensor invariants $\alpha G$, $\gamma^2$ and $\delta^2$ because ROA samples the variation of the vector potential of the light wave over the molecular dimension. On going from 514.5 to 647.1 nm the average Raman intensity therefore decreases by a factor of 2.5, but the ROA difference intensity drops by a factor of well over 3. Our new ROA spectrometer uses the 532 nm frequency doubled Nd:$YVO_4$ line. Average and difference intensities will be reduced by factors of 1.14 and 1.18 as compared with the 514.5 nm $Ar^+$ line, values which appear to be acceptable.

Laser stability is of crucial importance, power as well as beam pointing stability. Modern research argon ion lasers no longer present noticeable problems in this respect, but early models did. Frequency doubled Nd:$YVO_4$ lasers are expected to show even better performance.

**Scattering cells** Ordinary Raman scattering cells are generally unsuitable for ROA work, which requires cells of high optical quality without any strain induced birefringence in the windows through which the exciting light enters and the observed scattered light leaves. Moreover, care has to be taken to avoid the collection of Raman scattered light reflected from oddly oriented cell surfaces. High-quality fluorescence cells selected for their absence of birefringence are suitable for ROA work, but more sophisticated designs have also been developed. For right-angle scattering, microcapillary cells with sample volumes as low as 0.5 μL have been successfully used. Rotating cells that permit higher laser powers might be an option in backscattering if the optical problems can be solved.

**The dispersive system** The earliest ROA spectrometers were of the scanning type and used Czerny–Turner double monochromators. It took weeks to

record a single ROA spectrum. The scanning technology has now been completely superseded by multichannel instruments with spectrographs.

One particular ROA instrument, which dates from the late 1970s, fully exploited concave aberration corrected holographic grating technology to maximize throughput, and set a standard for the following decade. Its spectrograph is depicted in **Figure 2A**. At an average resolution of about 9 cm$^{-1}$, typical for ROA solution spectra, the field stop, represented by the entrance slit, has a surface area $S$ of $0.25 \times 17$ mm = 4.25 mm$^2$ at an average input focal-ratio of 6.33 with essentially no vignetting. The resulting étendue $G$ is so large that special light collection optics, described in later sections, are required to effectively fill it and amounts to

$$G = \pi S \sin^2 \alpha \approx \frac{\pi S}{(2\text{focal-ratio})^2} = 0.083 \text{ mm}^2 \text{ sr} \quad [14]$$

Sin $\alpha$ is the numerical aperture $NA$ of the system, and we have made the usual approximation that 2 $NA \approx 1/(\text{focal-ratio})$. The larger input than output focal length reduces the size of the slit image in the output focal plane to about half the length of the input slit. The larger kind of the currently available CCDs for spectroscopic use is therefore well suited for such a spectrograph design. In its original application, before such CCDs became available, an image intensifier was used in the output focal plane of the spectrograph, and a further, anamorphic image size reduction was performed between the image intensifier and the solid state detector.

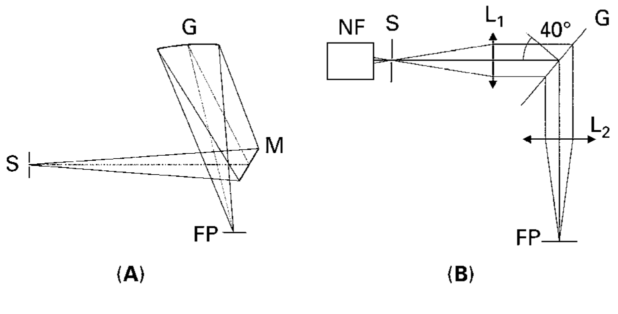



**Figure 2** (A) Concave aberration corrected holographic grating spectrograph. S: entrance slit. M: folding mirror. G: concave grating 1500 mm$^{-1}$, 110 mm diameter, $f_{in}$ = 658 mm, average $f_{out}$ = 356 mm. FP: output focal plane 9 × 50 mm. (B) Holographic transmission grating spectrograph. S: entrance slit. $L_1$: input lens, $f$ = 85 mm, $f$/1.8. G: grating 53 mm × 64 mm, 2400 mm$^{-1}$. $L_2$: output lens, $f$ = 85 mm, $f$/1.4. FP: output focal plane 6.6 × 27 mm (limited by vignetting and detector size; radius of curvature 39 mm of image of straight slit). NF: holographic notch filter assembly (in ROA integrated in light collection optics).

Spectrograph technology made another advance with the introduction of planar volume holographic transmission phase gratings in the 1990s. Such gratings are typically used in combination with holographic notch filters. The latest versions of these filters provide a Rayleigh light suppression of the order of $10^6$ and allow the observation of Raman bands down to 100 cm$^{-1}$ shifts. Commercial spectrographs use a simple back-to-back arrangement of two photographic single-lens reflex camera lenses with the grating placed in between (**Figure 2B**). The short focal length of these lenses, coupled with their high speed, keeps object and image size small while maintaining throughput. However, back-to-back configurations of photographic lenses are notorious for their vignetting problems and it does indeed reach 30% at the corners of a relatively small 6.6 × 27 mm CCD detector. Still, throughput is impressive. At 9.5 cm$^{-1}$ resolution a slit size of 0.1 × 6.6 mm is realized at a focal-ratio of about 1.8, which corresponds to an étendue of 0.16 mm$^2$ sr, 1.9 times the value of the concave grating spectrograph. In addition, holographic transmission gratings are claimed to exhibit close to 100% diffraction efficiency for s-polarized and up to 60% for p-polarized light, essentially double the average values of concave gratings, but precise data have not been forthcoming by the manufacturer.

A considerable disadvantage of such spectrographs is the curvature of the image of the entrance slit on the detector. In ROA this can limit the data acquisition rate.

**Detectors and their electronics** The first detectors used in multichannel Raman instruments were silicon intensified target tubes preceded by an image intensifier. Their large carry-over of information from one illumination period to the following severely limited their usefulness for ROA spectroscopy, where modulation takes place between right and left circularly polarized exciting or detected light. Thus, the first viable, shot noise limited multichannel ROA instrument used self-scanned diode array technology, coupled via high-speed ($f$/1.3) anamorphic optics to a high sensitivity, high gain ($10^6$) multistage image intensifier, with critical phosphor screens selected for their rapid decay times. While technically highly successful, the design was too elaborate to be widely adopted. The ready availability of backthinned CCDs with average quantum efficiencies of up to 70% for green to orange light has finally provided an easy solution to the detector problem. The high quantum efficiency more than makes up in ROA for the lack of shot noise limited performance, with CCD read-out noise of the order of 20 detected

photons or less. Still, to keep it well below photon shot noise, extremely low modulation speeds, of the order of a second per half-cycle, are presently used in ROA. The S/N degradation by flicker noise is bound to be an unavoidable consequence. Dark current with its associated shot noise can be kept low by cooling the detector to at least −70°C. Carry-over is inherently low and can be completely suppressed by rapid charge clearing cycles.

A detail of underestimated importance to ROA is data digitization. It is mandatory that precautions are taken to ensure that differences of less than 1 in $10^4$ for small Raman bands are reliably recovered by 14 to 16 bit AD converters by repeated digitization; if not, ROA difference spectra are simply swamped. The misunderstanding of this problem, particularly in recent high-throughput, low noise instruments has doubtless contributed to the abandonment of right-angle scattering. Systematic dither, i.e. small variations of the illumination time of the sample for consecutive modulation periods, compensated numerically in the data acquisition system, should solve this problem in future ROA instruments.

**Data acquisition and treatment** Two important additional requirements need to be met compared with ordinary Raman spectroscopy. These are the need to switch, synchronized to an optical modulator, the acquisition between a period for right and one for left circularly polarized light, and to recover, in addition to the average scattered intensity $I$, the difference intensity $\Delta I$, where $\Delta I$ is often less than 100 ppm of $I$. Modern PC-based data acquisition systems with a 32 bit word-length have no difficulty with this, and the vast amount of available memory capacity even allows storage of individual modulation cycles for entire acquisition runs. This facilitates checking data for corruption by cosmic ray events on the detector, as well as for sample degradation and similar problems. Cosmic ray events occur at a rate of about 2 $min^{-1}$ $cm^{-2}$ of CCD surface. Typically, at one second per half-cycle and a standard 6.6 mm × 27 mm CCD detector, the data for every 17th half-cycle have to be discarded. Unfortunately, cosmic ray interference is not yet systematically eliminated in current ROA instruments.

### Specific features

**Circular polarization modulators** Circularly polarized light is generated by passing linearly polarized light through a $\lambda/4$ retarder (**Figure 3**). Switching between right and left circular light occurs if the orientation of the fast and slow axis is interchanged. The exciting laser light is monochromatic and highly collimated. Fast switching longitudinal KDP ring-

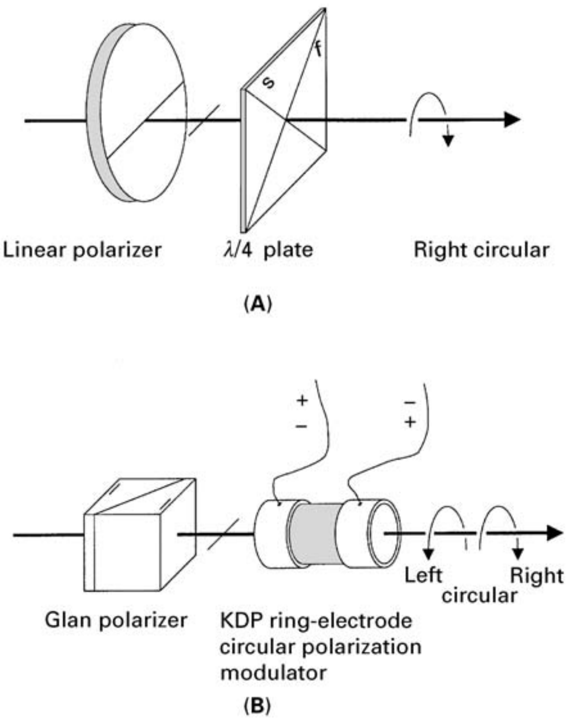



**Figure 3** (A) Basic circular polarizer. (B) Practical implementation most often used in ROA.

electrode modulators with their even retardation over the whole aperture have therefore been the preferred choice in ICP instruments. They have the inconvenience of requiring a quarter-wave voltage of about ± 1700 V for 532 nm light, suffer from residual stress-induced birefringence, and their retardation is strongly temperature dependent. Liquid crystal retarders, which require just a few volts of modulating voltage and are available with temperature compensation electronics, will doubtless be used in the future. Their slow switching speed can be accommodated by commuting them during the read-out time of the CCD detector.

**Circular polarization analysers** The circular polarization analysers used in SCP and DCP are modulators in reverse (**Figure 4**). Their practical implementation is complicated by the fact that the light passing through them is neither monochromatic nor, as it emanates from a scattering zone with finite dimensions, is it collimated. Mechanically rotated compensated zero-order and achromatic crystal quarter-wave plates have been the choice for the retarder. To date, in the one operating DCP instrument a true zero-order polymer quarter-wave plate with a much larger acceptance angle is now being used. Small as such advances might appear, they can be decisive for instrument performance.

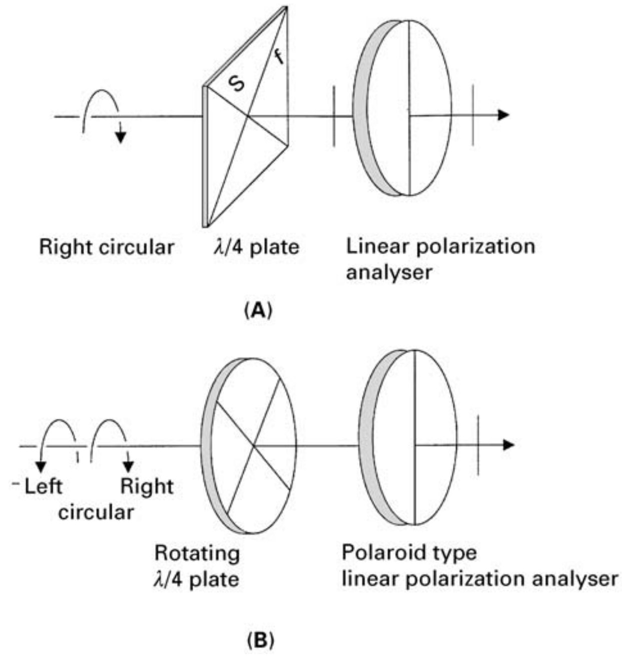



**Figure 4** (A) Basic circular polarization analyser. (B) Practical implementation currently used in ROA; future instruments may use LC retarders and thin film polarizers.

In view of the aperture size and the required acceptance angle the linear polarization analysers employed in SCP and DCP are presently of the dichroic variety. Given the advances in thin film technology, thin film polarizers with better transmission characteristics will appear in future ROA instruments. Their use might solve the intensity problem SCP suffers in comparison with ICP and DCP by allowing the simultaneous recording of right and left circularly polarized scattered light. Yet, in order to be on a par with ICP, such SCP instruments will require a spectrograph with twice the étendue of a comparable ICP instrument, but should enjoy the benefit of reduced flicker noise. With holographic transmission grating spectrographs the size of the étendue is no longer an issue. This is the justification for having included SCP in **Table 1**.

$DCP_I$ backscattering (**Figure 5**) is somewhat in a class of its own. The discarded polarization component corresponds to $DCP_{II}$ and carries no ROA information in the non-resonant case. For depolarized Raman bands ($\alpha^2 = 0$) $DCP_I$ should slightly outperform ICP backscattering with respect to shot noise, but it is for polarized bands where it is expected to shine. As seen from Equation [5] (**Table 1**), isotropic scattering does not contribute to the average scattering intensity. Unfortunately, this is also one of the prime reasons why $DCP_I$ is so artefact prone. Tiny differences in the leakage of isotropic scattering through the circular polarization analyser for the right and left circular modulation periods simply tend to swamp the ROA difference spectrum. $DCP_I$ will only become of general usefulness if active stabilization techniques for the circularity of the exciting light are developed.

**Light collection and transfer optics** High light-collection efficiency and avoidance of artefacts are prime concerns in ROA. Fortunately, well-designed light collection optics can combine them both.

In ICP right-angle scattering the perfect arrangement would be circular light collection around the sample. This is hard to realize, but a similarly advantageous practical approach consists in using two light-collection lenses placed under 90° to each other (**Figure 6**). It is the only light collection system capable of filling the étendue of high luminosity spectrographs with small sample volumes. Though

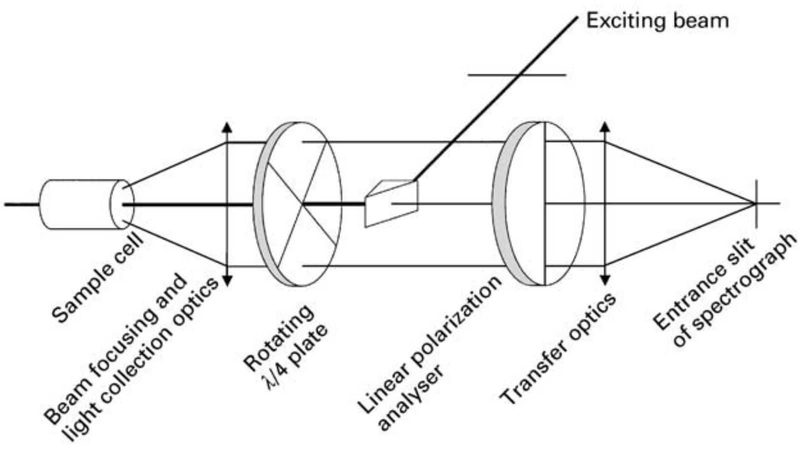



**Figure 5** $DCP_I$ backscattering arrangement. A single quarter-wave plate can function in a circular polarizer and analyser configuration. True zero-order retarders may be placed ahead of the light collection lens, albeit with loss of precision. The plate is rotated within 200 ms in the appropriate orientation and stopped during the data acquisition half-cycle.

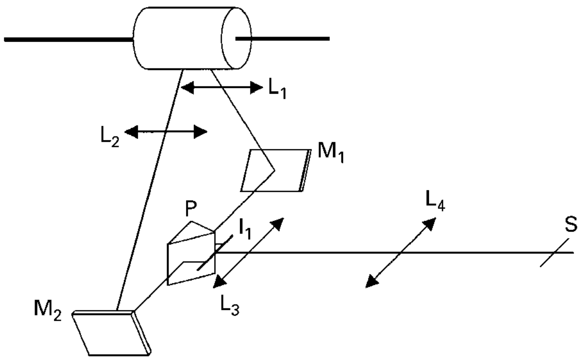



**Figure 6** Dual lens light collection system for right-angle scattering. The two light collection lenses $L_1$ and $L_2$ form a joint intermediate image $I_1$. Their optical axes intersect at 90° at the sample. $M_1$, $M_2$: 22.5° mirrors. $L_3$: field lens. $L_4$: lens which projects $I_1$ onto the entrance slit S of the spectrograph.

backscattering has the advantage of approximately double the value of $\Delta I/I$, right-angle scattering remains in this respect a far superior technique.

In forward and backward scattering, ICP, SCP and DCP, a regular light collection cone about the axis of the exciting light is the equivalent of circular light collection in right-angle scattering and provides similar benefits. The approach was demonstrated with the help of a fibre optics cross-section transformer in the first ROA backscattering instrument. Care has to be taken to keep focal-ratio degradation low. If this is not done, and if the fibre optics are not carefully matched to the remaining optics, then the light losses can be disastrous, and in later backscattering instruments the use of fibre optics was therefore dropped. Yet, there appears to be no alternative to fibres when filling the large étendue of modern spectrographs in collinear scattering is required, and future ROA instruments will doubtless revert to their use.

Artefact suppression is only substantial in the two lens system if the two light-collection branches are well balanced. Similarly, light collection in a regular cone in backscattering makes sense only if a direction perpendicular to the axis of the cone is not distinguished in some other way. Unfortunately, for ICP, the polarization dependence of the dispersive grating of the spectrograph does precisely that. For holographic transmission gratings, for example, a 50% higher diffraction efficiency is quoted for s polarization than for p polarization. The action of the grating therefore is akin to that of a low efficiency polarization analyser. Other optical elements, such as the 45° mirror in **Figure 7**, can add to or subtract from this effect. Depolarization of the scattered light therefore is an absolute necessity in ICP. A Lyot depolarizer placed directly after the sample cell into the divergent scattered light was used in the first instrument, complementing the effect of the fibre optics, and has remained the standard method ever since to depolarize quasi-monochromatic Raman light in all ICP ROA backscattering instruments.

## Instrument performance and the future

The rapid advance in the measurement of ROA is evident from the spectra recorded by the latest backscattering instruments equipped with backthinned CCDs and transmission phase gratings. To date, such spectra have been published for two ICP instruments operating at the University of Glasgow and for a $DCP_I$ instrument at Syracuse University. Depending on the earlier multichannel instrument taken for comparison, quoted increases in measurement speed vary from 5 to 100. Within the statistical noise limit, artefact control, achieved by nulling with achiral samples

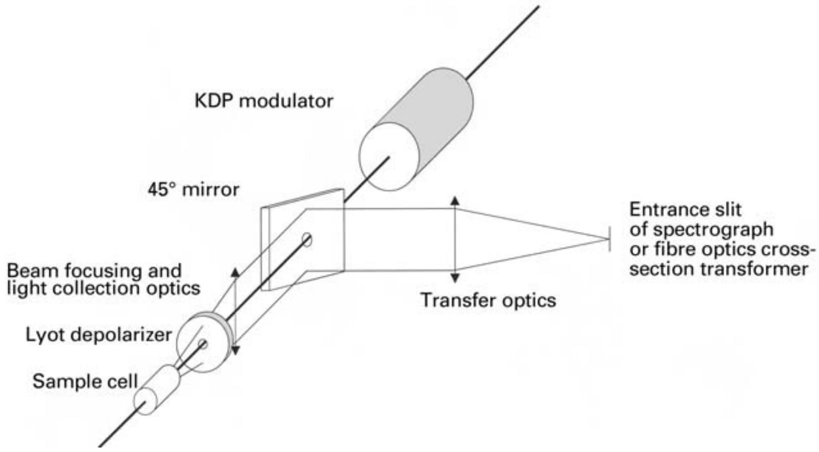



**Figure 7** Conventional ICP backscattering arrangement with a Lyot depolarizer in divergent light. Filling the étendue of a spectrograph requires a fibre optics cross-section transformer as used in the original design.

by, for example, slightly rotating the plane of polarization of the incident light ($DCP_I$) or the axes of the Lyot depolarizer (ICP), is essentially complete.

A SCP instrument constructed by the author at the University of Zürich uses even more advanced optics (**Figure 8**). It represents a further increase, by a factor of 4.4 (compared with the $DCP_I$ instrument) and 7.8 (ICP instrument), in light gathering and throughput capability (**Table 2**). Yet, its main attraction is dramatically reduced flicker noise and self-balancing of artefacts. This is shown in the very first ROA spectrum recorded on this instrument (**Figure 9**). Deliberately, no artefact nulling was performed, the scattering cell was not selected or individually calibrated, and the liquid crystal circular analyser was not stabilized, in contrast, for example, to the high degree of stabilization used for the retarder in ICP experiments. Likewise, no sample filtration was

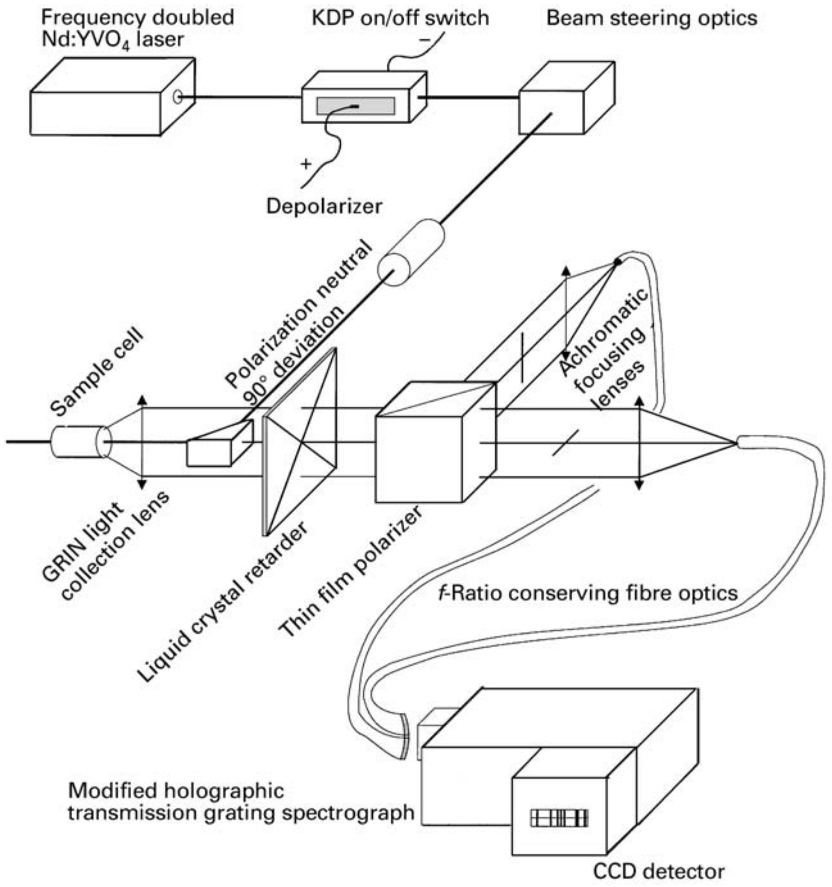



**Figure 8** A new generation SCP backscattering instrument. The right and left circular components are measured simultaneously, with the two channels being interchanged by switching the liquid crystal retarder. Sample illumination during read-out is suppressed by the KDP switch. The curved fibre optics output forms the entrance slit of the spectrograph and yields a straight line image on the detector. The notch filter is incorporated in the light collection/transfer optics.

**Table 2** Comparison of luminosity of current backscattering instruments as measured by the number of detected electrons per CCD column per joule of exciting energy at the sample; the values refer to the height of the depolarized 1436 cm $^{-1}$ band of α-pinene. Dispersion and CCD efficiency is similar for all instruments

| Instrument | *Simultaneous R/L detection* | Path length cell (mm) | *CCD column size* (mm$^2$) | *Exciting wavelength* (nm) | *Slit width* (μm) | *Detected charge* (electrons J$^{-1}$) |
|---|---|---|---|---|---|---|
| ICP (Glasgow) | No | 5 | 0.027 × 6.9 | 514.5 | 83 | $5.6 \times 10^{6}$ [a] |
| $DCP_I$ (Syracuse) | No | 10 | 0.024 × 7.9 | 514.5 | 83 | $10 \times 10^{6}$ [b] |
| SCP (Zürich) | Yes | 6 | 0.026 × 6.65 | 532 | 71[c] | $44 \times 10^{6}$ [d] |

[a] L Hecht, private communication.
[b] $DCP_I$ value multiplied by 7/6 for comparison.
[c] Slit width of the SCP instrument is the equivalent value calculated for the standard 85 mm focal length input lens of the spectrographs used in the other instruments.
[d] Limited by the sample cell.

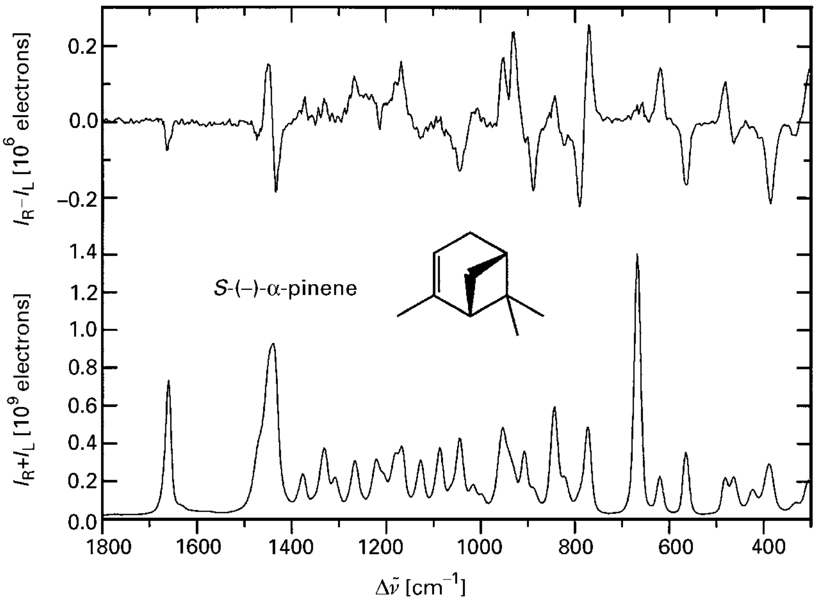



**Figure 9** ROA spectra of (–)-α-pinene recorded with the SCP instrument of **Figure 8**. The exciting energy was 21 J at the sample, yielding approximately $10^9$ electrons ( = detected photons) for the peak height of the 1436 cm$^{-1}$ band (see S/N discussion).

performed and no sample handling precautions were adopted. As a consequence, bright flashes of light due to grains of dust passing through the scattering zone of the exciting beam focused into the sample cell illuminated the walls of the laboratory during recording of the spectra. Ordinarily, such conditions would have totally invalidated any ROA measurement, and very large offsets were indeed observed in the two individual branches of the instrument. Yet, upon summing the data, offsets and artefacts essentially cancelled except at very low wavenumber shifts. The detector and electronics were probably temporarily overloaded at these wavelengths, and some Raman scattered light may also have been collected from the quartz windows of the sample cell, which is not yet optimized for the awesome light gathering power of the non-imaging optics of the instrument. These problems are amenable to straightforward solutions.

The spectrum in **Figure 9** is clearly not yet perfect but it is of acceptable quality, particularly if the short illumination time (300 s) and low laser power (70 mW at the sample) are also taken into account. Together with the above-cited ICP and $DCP_I$ data it lends strong support to the view that ROA instrumentation has finally progressed to the point where this powerful chiroptical method will become a generally useful analytical tool.

## List of symbols

$c$ = speed of light; $E^{(0)}$ = electric field at the scattering molecule; $f$ = frequency or focal length; $G$ = étendue; $I$ = scattered light intensity; $K$ = defined by Equation [11]; $N$ = number of events; NA = numerical aperture; $R$ = distance between scattering molecule and observer; $S$ = surface area; $S/N$ = signal to noise ratio; $\alpha$ = half angle of light cone; $\alpha^2$ = isotropic Raman invariant; $\alpha G'$ = isotropic ROA invariant due to the optical activity tensor; $\beta^2$ = anisotropic Raman invariant; $\gamma^2 = \beta(G')^2$ = anisotropic ROA invariant due to the optical activity tensor; $\delta^2 = \beta(A)^2$ = anisotropic ROA invariant due to the quadrupole tensor; $\mu_0$ = permeability of the vacuum; $\omega$ = angular frequency.

*See also:* **Fibre Optic Probes in Optical Spectroscopy, Clinical Applications; Light Sources and Optics; Raman Optical Activity, Applications; Raman Optical Activity, Theory; Raman Spectrometers; Vibrational CD Spectrometers.**

## Further reading

Barron LD and Hecht L (1994) Vibrational Raman optical activity: from fundamentals to biochemical applications. In: Nakanishi K, Berova ND and Woody RW (eds) *Circular Dichroism: Principles and Applications*, pp 179–215. New York: VCH.

Barron LD and Hecht L (1994) Recent developments in Raman optical activity instrumentation. *Faraday Discussions* 99: 35–47.

Barron LD and Hecht L (1996) Recent developments in Raman optical activity of biopolymers. *Applied Spectroscopy* 50: 619–629.

Greulich KO and Monajembashi S (1996) Laser microbeams and optical tweezers: how they work and why they work. In: *Optical and Imaging Techniques for Biomonitoring*, Vol 2628 of Proceedings of SPIE – International Society of Optical Engineering, 116–127. Bellingham: SPIE Press.
Hug W (1982) Instrumental and theoretical advances in Raman optical activity. In: Lascombe J and Huong PV (eds) *Raman Spectroscopy, Linear and Nonlinear*, pp 3–12. Chichester: Wiley-Heyden.
James JF and Sternberg RS (1969) *The Design of Optical Spectrometers*. London: Chapman & Hall.
Nafie LA (1996) Vibrational optical activity. *Applied Spectroscopy* **50**: 14A–26A.
Nafie LA and Zimba CG (1987) Raman optical activity and related techniques. In: Spiro TG (ed) *Biological Applications of Raman Spectroscopy, Vol 1, Raman Spectra and The Conformation of Biological Macromolecules*, pp 307–343. New York: John Wiley & Sons.
Vargek M, Freedman TB and Nafie LA (1997) Improved backscattering dual circular polarization Raman optical activity spectrometer with enhanced performance for biomolecular applications. *Journal of Raman Spectroscopy* **28**: 627–633.

# Raman Optical Activity, Theory

**Laurence A Nafie**, Syracuse University, New York, NY, USA

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Theory**

Raman optical activity (ROA) is defined as the difference in Raman scattering intensity for right minus left circularly polarized light. Along with optical rotational and circular dichroism, ROA is a form of natural optical activity. All forms of optical activity can be defined as the differential interaction of a molecule with right versus left circularly polarized radiation. Only chiral molecules exhibit natural optical activity and, for such molecules, the mirror image of the molecule cannot be superimposed on itself. The most common form of ROA is vibrational ROA. Vibrational ROA is also one of two form of vibrational optical activity. The other form is vibrational circular dichroism (VCD), which is the difference in the IR absorption of a molecule for left versus right circularly polarized radiation for a vibrational transition. VCD and ROA are complementary, non-redundant forms of vibrational optical activity in the same way that IR absorption and Raman scattering are complementary forms of ordinary vibrational spectroscopy.

There are many forms of ROA, depending on the choice of polarization modulation, scattering geometry and proximity of the exciting laser radiation to resonance with excited electronic states in the molecules. A general theory of ROA can be written from which all special cases can be derived. The first division of the theory is between circular polarization (CP) ROA and linear polarization (LP) ROA. To date only different forms of CP ROA have been measured experimentally. There are four forms of both CP and LP ROA. For CP ROA, the original and most common form is called incident circular polarization (ICP) ROA in which only the state of the incident laser radiation is modulated between right and left circular polarization (RCP and LCP). Analogously, the other three forms are called scattered circular polarization (SCP) ROA, where only the polarization of the Raman scattered radiation is sampled for RCP and LCP content, dual circular polarization one ($DCP_I$), where both the incident and scattered polarization states are modulated in-phase, and dual circular polarization two ($DCP_{II}$), where both states are modulated out-of-phase. Similar definitions have been provided for the four forms of LP ROA.

The two principal resonance limits of the theory of ROA are the far-from-resonance (FFR) limit, the original form of the theory of ROA, and the single-electronic-state (SES) limit, for the case of strong resonance between a single excited electronic state and the incident laser radiation. In the case of FFR ROA, *ab initio* calculations have been carried out for direct comparison with experiment. The SES theory is so simple that the complete SES-ROA spectrum can be predicted from the parent resonance Raman spectrum and the electronic circular dichroism spectrum of the resonant electronic state.

ROA is a new spectroscopic tool that has been applied with a high degree of success to the study of the structure of chiral molecules in solution. Areas of application include proteins, nucleic acids, carbohydrates, natural products, pharmaceuticals and other kinds of molecules of biological or therapeutic significance.

## Polarized light scattering

One of the essential properties associated with the scattering of light by molecules is the polarization of state of the light. Changes in the polarization state affect the nature and information content of the scattered light. This holds for Rayleigh scattering, where the frequency of the scattered radiation is unchanged from the exciting laser radiation, and for Raman scattering, where the scattered radiation differs from the incident laser radiation by a vibrational energy change in the molecule. The intensity of light scattering for any experiment can be expressed in terms of the general scattering tensor $\tilde{a}_{\alpha\beta}$ and the polarization vectors for the incident and scattered radiation $\tilde{\boldsymbol{e}}^{\mathrm{i}}_{\alpha}$ and $\tilde{\boldsymbol{e}}^{\mathrm{d}}_{\alpha}$, respectively, and is given by

$$I(\tilde{\boldsymbol{e}}^{\mathrm{d}}, \tilde{\boldsymbol{e}}^{\mathrm{i}}) = 90K\left\langle \left| \tilde{\boldsymbol{e}}^{\mathrm{d}*}_{\alpha} \tilde{a}_{\alpha\beta} \tilde{\boldsymbol{e}}^{\mathrm{i}}_{\beta} \right|^2 \right\rangle \qquad [1]$$

In this equation, $K$ is a constant, given below, that depends on among other things on the intensity of the incident laser radiation. The angular brackets designate an average over all angles of orientation of the molecule to the laboratory frame of reference. This is needed for liquid, solution or gaseous samples where there is no unique molecular axes relative to the laboratory axes. The polarization vectors have one Greek subscript and the scattering tensor has two. For repeated Greek subscripts, summation over the Cartesian directions $x$, $y$ and $z$ is implied. Hence, Equation [1] has nine terms within the vertical brackets, and these brackets designate the absolute value of the complex quantities within the brackets. The tilde above a quantity, such as a polarization vector or a scattering tensor, indicates that this quantity can be complex. The star superscript for the polarization vector of the scattered light designates complex conjugation. The constant $K$ is given by

$$K = \frac{1}{90}\left(\frac{\omega^2 \mu_0 \tilde{\boldsymbol{E}}^{(0)}}{4\pi R}\right)^2 \qquad [2]$$

where $\omega$ is the angular frequency of the scattered light, $\mu_0$ is the magnetic permeability $\tilde{\boldsymbol{E}}^{(0)}$ is the electric field strength of the incident laser radiation, and $R$ is the distance from the scattering to the detector. The general scattering tensor is given through its lowest-order tensors as

$$\tilde{a}_{\alpha\beta} = \tilde{\alpha}_{\alpha\beta} + \frac{1}{c}\left[\varepsilon_{\gamma\delta\beta}\boldsymbol{n}^{\mathrm{i}}_{\delta}\tilde{G}_{\alpha\gamma} + \varepsilon_{\gamma\delta\alpha}\boldsymbol{n}^{\mathrm{d}}_{\delta}\tilde{G}_{\gamma\beta} + \frac{1}{3}\omega(\boldsymbol{n}^{\mathrm{i}}_{\gamma}\tilde{A}_{\alpha,\gamma\beta} - \boldsymbol{n}^{\mathrm{d}}_{\gamma}\tilde{A}_{\beta,\gamma\alpha})\right] \qquad [3]$$

where the first tensor is simply the polarizability tensor that is responsible for ordinary Raman (and Rayleigh) scattering. The four tensors in square brackets are the ROA tensors. The first two are magnetic dipole–electric dipole ROA tensors and the second two are electric quadrupole–electric dipole ROA tensors. The vectors $\tilde{\boldsymbol{n}}^{\mathrm{i}}_{\alpha}$ and $\tilde{\boldsymbol{n}}^{\mathrm{d}}_{\alpha}$ are the propagation vectors for the incident and scattered light, respectively, and $\varepsilon_{\alpha\beta\gamma}$ is the unit antisymmetric tensor that is +1 for even permutations of the order $x$, $y$, $z$, –1 for odd permutation of this order, and zero if any two directions are the same.

The Raman polarizability tensor is given by

$$\tilde{\alpha}_{\alpha\beta} = \frac{1}{\hbar}\sum_{j\neq m,n}\left[\frac{\langle m|\hat{\boldsymbol{\mu}}_{\alpha}|j\rangle\langle j|\hat{\boldsymbol{\mu}}_{\beta}|n\rangle}{\omega_{jn} - \omega_0 + i\Gamma_j} + \frac{\langle m|\hat{\boldsymbol{\mu}}_{\beta}|j\rangle\langle j|\hat{\boldsymbol{\mu}}_{\alpha}|n\rangle}{\omega_{jm} - \omega_0 + \mathrm{i}\Gamma_j}\right] \qquad [4]$$

where $\hbar$ is Planck's constant divided by $2\pi$, and the summation is over all excited electronic states, $j$, except the initial and final states, $n$ and $m$, respectively. The states $n$ and $m$ differ by a vibrational quantum of energy. The denominators contain frequency terms, and $\omega_{jn}$ is the angular frequency difference between the state $j$ and $n$. The terms $i\Gamma_j$ are imaginary terms proportional to the width of the electronic state $j$, and hence inversely proportional to its lifetime. The first term in Equation [4] is called the resonance term since the frequency difference between the $jn$-transition frequency and the laser frequency vanishes at the resonance condition. The quantities in angular brackets are quantum mechanical matrix elements with electric dipole moment operators $\mu_{\alpha}$ given by

$$\hat{\boldsymbol{\mu}}_{\alpha} = \sum_{k} e_k r_{k\alpha} \qquad [5]$$

which is simply the summation over the charge and position in the $\alpha$ direction of all particles $k$, in the molecule, electrons and nuclei.

The matrix element in Equation [4] involving the operator $\hat{\mu}_{\beta}$ describes the interaction of the molecule with the incident radiation while the matrix elements

with the operator $\hat{\mu}_\alpha$ describes the interaction of the molecule with the scattered radiation. The matrix element products in each term can be read from right to left in a time-ordered sense, and hence the resonance term describes the molecule interacting first with a laser photon and subsequently creating a scattered photon, whereas the non-resonance terms reverses the natural order of the those two events.

The four ROA tensors differ from the Raman polarizability tensor by substitution of a higher-order operator for an electric dipole operator in Equation [4]. The two operators needed for ROA are the magnetic dipole moment operator and the electric quadrupole moment operator given respectively by

$$\hat{\boldsymbol{m}}_\alpha = \frac{1}{2}\sum_k \frac{e_k}{m_k}\varepsilon_{\alpha\beta\gamma} r_{k\alpha} p_{k\gamma} \qquad [6]$$

$$\hat{\Theta}_{\alpha\beta} = \frac{1}{2}\sum_k e_k(3r_{k\alpha}r_{k\alpha} - r_k^2\delta_{\alpha\beta}) \qquad [7]$$

The relationships between the operator substitutions and the resulting ROA tensors in Equation [4] is given by

$$\alpha_{\alpha\beta} \quad \tilde{G}_{\alpha\beta} \quad \text{when} \quad \hat{\mu}_\beta \quad \hat{m}_\beta \qquad [8]$$

$$\alpha_{\alpha\beta} \quad \tilde{G}_{\alpha\beta} \quad \text{when} \quad \hat{\mu}_\alpha \quad \hat{m}_\alpha \qquad [9]$$

$$\alpha_{\alpha\beta} \quad \tilde{A}_{\alpha,\beta\gamma} \quad \text{when} \quad \hat{\mu}_{\beta\gamma} \quad \hat{\Theta}_\beta \qquad [10]$$

$$\alpha_{\alpha\beta} \quad A_{\beta,\alpha\gamma} \quad \text{when} \quad \hat{\mu}_{\alpha\gamma} \quad \hat{\Theta}_\alpha \qquad [11]$$

The expressions given above provide the theoretical formalism for the description of all forms of polarized Raman scattering through first-order in the magnetic dipole and electric quadrupole interaction of light with matter. This is sufficient to describe the various forms of ROA within the assumptions given above.

## ROA observables

By specifying the desired polarization states of the incident laser radiation and the scattered radiation, it is possible to construct theoretical expressions to describe the various ROA observables that can be measured in the laboratory. If we restrict our attention to circular polarization ROA experiments, the expressions can be obtained from pairs of intensity expressions that differ only in the change in the circular polarization state of one or both the light beams from right circular to left circular, or vice versa.

The fundamental ROA observables are classified by polarization and scattering angle, $\xi$. There are four different forms of CP ROA given by

$$\text{ICP ROA:} \qquad \Delta I_\alpha(\xi) = I_\alpha^{\mathrm{R}}(\xi) - I_\alpha^{\mathrm{L}}(\xi) \qquad [12]$$

$$\text{SCP ROA:} \qquad \Delta I^\alpha(\xi) = I_{\mathrm{R}}^{\alpha}(\xi) - I_{\mathrm{L}}^{\alpha}(\xi) \qquad [13]$$

$$\text{DCP}_{\mathrm{I}}\ \text{ROA:} \qquad \Delta I_{\mathrm{I}}(\xi) = I_{\mathrm{R}}^{\mathrm{R}}(\xi) - I_{\mathrm{L}}^{\mathrm{L}}(\xi) \qquad [14]$$

$$\text{DCP}_{\mathrm{II}}\ \text{ROA:} \qquad \Delta I_{\mathrm{II}}(\xi) = I_{\mathrm{L}}^{\mathrm{R}}(\xi) - I_{\mathrm{R}}^{\mathrm{L}}(\xi) \qquad [15]$$

These different forms of ROA are illustrated in **Figure 1**. In the case of ICP- and SCP-ROA, the polarization states $\alpha$ is any fixed value. The standard

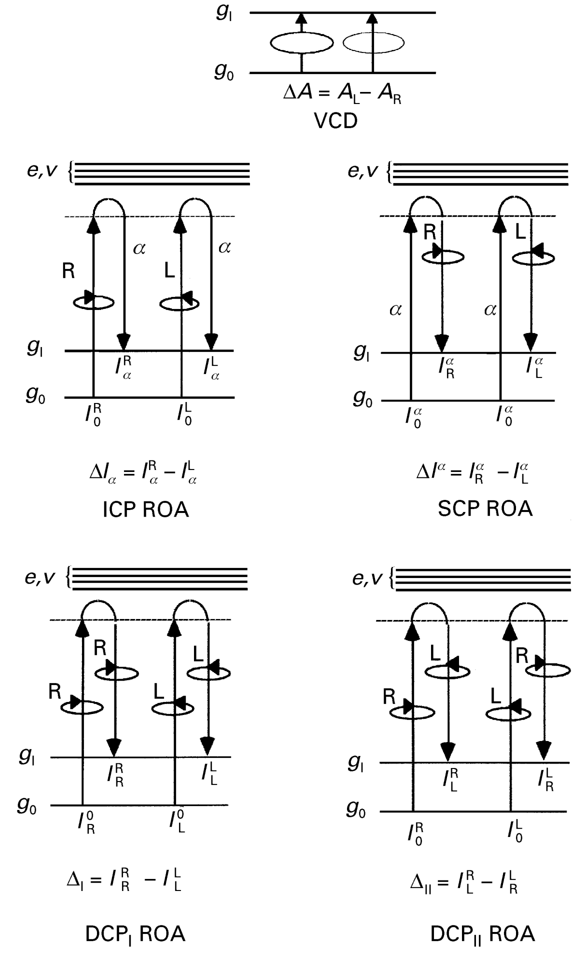



**Figure 1** Energy level diagram showing the transitions and polarization states associated with the definition of VCD and the various forms of ROA.

choices are unpolarized, linearly polarized parallel to the scattering plane (depolarized) or linearly polarized perpendicular to the scattering plane (polarized). The standard scattering angles are 90° (right-angle scattering) 180° (backscattering), and 0° (forward scattering).

As an example of an ROA observable we present in **Figure 2** the backscattering DCP-ROA and Raman spectra of neat (−)-β-pinene in the region between 200 and 1700 cm$^{-1}$. The stereospecific structure of this molecule is given in the figure. The opposite enantiomer, the mirror-image molecule, would yield an ROA spectrum in which the signs of all the bands would be reversed, i.e. a 'mirror-image' ROA spectrum. Note that the ROA spectrum is approximately three orders of magnitude smaller than the corresponding Raman spectrum. Each parent Raman band has associated with it an ROA band of a particular sign and magnitude. There is no correlation between strong Raman bands and strong ROA bands.

## General theory of ROA

The general, complete theory of ROA embraces all possible polarization experiments, scattering geometries and degress of resonance Raman intensity enhancement. Because of this generality, the level of theory is complex and too lengthy to describe in the present context. Instead, we provide a verbal description of the formalism and refer the interested reader to a comprehensive review by Nafie and Che (1994) of the theory and measurement of ROA.

ROA and Raman intensity are proportional to the square of a tensor quantity, as expressed in Equation [1]. For Raman scattering only the square of the polarizability is needed, whereas ROA intensity arises from the product of the polarizability and an ROA tensor. The ROA tensor are approximately three orders of magnitude smaller than the polarizability, and hence an ROA spectrum is approximately three orders of magnitude smaller than its parent Raman spectrum. As noted above, the Greek subscripts of the tensor refer to the molecular axis system. However, for both Raman and ROA, linear combinations of products of tensors can be found that do not vary with the choice of the molecular coordinate frame. Such combinations are called invariants. All Raman intensities from samples of randomly oriented molecules can be expressed in terms of only three invariants, called the isotropic invariant, the symmetric

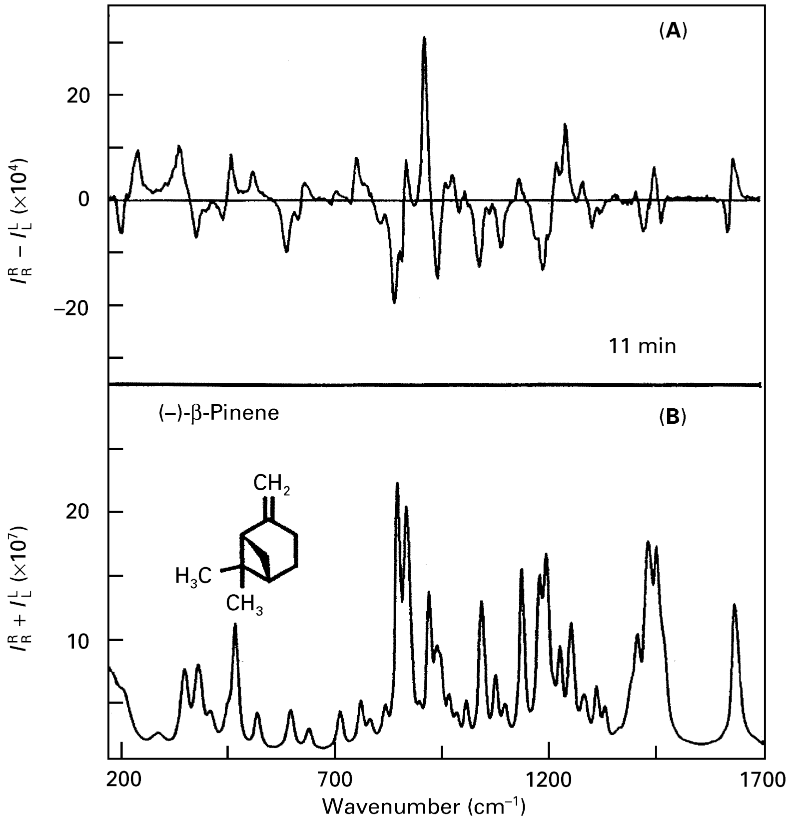



**Figure 2** DCP$_I$ Raman (A) and ROA (B) spectra for (−)-β-pinene.

anisotropy and the antisymmetric anisotropy, given by

$$\alpha^2 = \frac{1}{9}\mathrm{Re}[(\tilde{\alpha}_{\alpha\alpha})^{\mathrm{s}}(\tilde{\alpha}_{\beta\beta})^{\mathrm{s}*}] \quad [16]$$

$$\beta_{\mathrm{s}}(\alpha)^2 = \frac{1}{2}\mathrm{Re}[3(\tilde{\alpha}_{\alpha\beta})^{\mathrm{s}}(\tilde{\alpha}_{\alpha\beta})^{\mathrm{s}*} - (\tilde{\alpha}_{\alpha\alpha})^{\mathrm{s}}(\tilde{\alpha}_{\beta\beta})^{\mathrm{s}*}] \quad [17]$$

$$\beta_a(\alpha)^2 = \frac{1}{2}\mathrm{Re}[3(\tilde{\alpha}_{\alpha\beta})^{\mathrm{a}}(\tilde{\alpha}_{\alpha\beta})^{\mathrm{a}*}] \quad [18]$$

where the symmetric and anti-symmetric forms of the polarizability tensors are given by

$$(\tilde{\alpha}_{\alpha\beta})^{\mathrm{s}} = \frac{1}{2}[(\tilde{\alpha}_{\alpha\beta}) + (\tilde{\alpha}_{\beta\alpha})] \quad [19]$$

$$(\tilde{\alpha}_{\alpha\beta})^{\mathrm{a}} = \frac{1}{2}[(\tilde{\alpha}_{\alpha\beta}) - (\tilde{\alpha}_{\beta\alpha})] \quad [20]$$

For ROA there are ten invariants, five associated with the Roman (font) tensors, [$\alpha G$, $\beta_{\mathrm{s}}(\tilde{G})^2$, $\beta_{\mathrm{a}}(\tilde{G})^2$, $\beta_{\mathrm{s}}(\tilde{A})^2$ and $\beta_{\mathrm{a}}(\tilde{A})^2$ and five with the Arial tensors, [$\alpha\mathsf{G}$, $\beta_{\mathrm{s}}(\tilde{\mathsf{G}})^2$, $\beta_{\mathrm{a}}(\tilde{\mathsf{G}})^2$, $\beta_{\mathrm{s}}(\tilde{\mathsf{A}})^2$ and $\beta_{\mathrm{a}}(\tilde{\mathsf{A}})^2$]. All of the different ROA experiments can be expressed in terms of these invariants. The ROA intensity for each experiment is expressed as a linear combination of some or all of the ten invariants. Although sets of experiments can be devised to isolate all three ordinary Raman invariants, only six distinct combinations of ROA invariants can be isolated.

## Far from resonance theory of ROA

The theory of ROA simplifies drastically in the limit, where the exciting laser radiation is far from the lowest allowed excited state in the molecule, and the interaction of the light with the molecule is approximately the same for both the incident and the scattered radiation. This symmetry reduces the number of Raman invariants from three to two, the isotropic and (symmetric) anisotropic invariants, and the number of ROA invariants from ten to three. The relationships that reduces these thirteen Raman and ROA invariants to only five are

$$\beta_{\mathrm{a}}(\alpha)^2, \beta_{\mathrm{a}}(\tilde{G})^2, \beta_{\mathrm{a}}(\tilde{A})^2, \beta_{\mathrm{a}}(\tilde{\mathsf{G}})^2, \beta_{\mathrm{a}}(\tilde{\mathsf{A}})^2 = 0 \quad [21]$$

$$\beta_{\mathrm{s}}(\alpha)^2 = \beta(\alpha)^2 \quad [22]$$

$$\alpha G = -\alpha\mathsf{G} = \alpha G' \quad [23]$$

$$\beta_{\mathrm{s}}(\tilde{G})^2 = -\beta_{\mathrm{s}}(\tilde{\mathsf{G}})^2 = \beta(G')^2 \quad [24]$$

$$\beta_{\mathrm{s}}(\tilde{A})^2 = \beta_{\mathrm{s}}(\tilde{\mathsf{A}})^2 = \beta(A)^2 \quad [25]$$

where the following definition of the real and imaginary parts of complex tensor has been used

$$\tilde{T} = T - \mathrm{i}T' \quad [26]$$

The equations for the two Raman invariants and three ROA invariants are

$$\alpha^2 = \frac{1}{9}\alpha_{\alpha\alpha}\alpha_{\beta\beta} \quad [27]$$

$$\beta(\alpha)^2 = \frac{1}{2}(3\alpha_{\alpha\beta}\alpha_{\alpha\beta} - \alpha_{\alpha\alpha}\alpha_{\beta\beta}) \quad [28]$$

$$\alpha G' = \frac{1}{9}\alpha_{\alpha\alpha}G'_{\beta\beta} \quad [29]$$

$$\beta(G')^2 = \frac{1}{2}(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta}) \quad [30]$$

$$\beta(A)^2 = \frac{1}{2}\omega_0\alpha_{\alpha\beta}\varepsilon_{\alpha\gamma\delta}A_{\gamma,\delta\beta} \quad [31]$$

where the FFR polarizability and optical activity tensors are given by

$$\tilde{\alpha}_{\alpha\beta} = \frac{2}{\hbar}\sum_{j\neq n}\frac{\omega_{jn}}{\omega_{jn}^2 - \omega_0^2}\mathrm{Re}\left[\langle n|\hat{\boldsymbol{\mu}}_\alpha|j\rangle\langle j|\hat{\boldsymbol{\mu}}_\beta|n\rangle\right] \quad [32]$$

$$G'_{\alpha\beta} = \frac{-2}{\hbar}\sum_{j\neq n}\frac{\omega_0}{\omega_{jn}^2 - \omega_0^2}\mathrm{Im}\left[\langle n|\hat{\boldsymbol{\mu}}_\alpha|j\rangle\langle j|\hat{\boldsymbol{\mu}}_\beta|n\rangle\right] \quad [33]$$

$$A_{\alpha\beta\gamma} = \frac{2}{\hbar}\sum_{j\neq n}\frac{\omega_{jn}}{\omega_{jn}^2 - \omega_0^2}\mathrm{Re}\left[\langle n|\hat{\mu}_\alpha|j\rangle\langle j|\hat{\Theta}_{\beta\gamma}|n\rangle\right] \quad [34]$$

Using these invariants, we can write intensity expressions for ROA and Raman that cover all possible polarization modulations and scattering geometries in the FFR approximation. These expressions are:

$$I(R) - I(L) = \frac{8K}{c}[D_1\alpha G' + D_2\beta(G')^2 + D_3\beta(A)^2] \quad [35]$$

$$I(R) + I(L) = 4K[D_4\alpha^2 + D_5\beta(\alpha)^2] \quad [36]$$

The values of the constants $D_1$ through to $D_5$ are given in **Table 1**.

Most of the ROA spectra measured to date have been for one of three kinds of experiments, right-angle depolarized ICP ROA, backscattering unpolarized ICP ROA, and backscattering $DCP_I$ ROA. The early work on ROA was almost exclusively right-angle depolarized ICP ROA where the ROA and Raman intensity expressions from **Table 1** are

$$I_z^R(90°) - I_z^L(90°) = \frac{8K}{c}[3\beta(G')^2 - \beta(A)^2] \quad [37]$$

$$I_z^R(90°) - I_z^L(90°) = 4K[3\beta(\alpha)^2] \quad [38]$$

The corresponding polarized ICP ROA experiment included isotropic ROA invariants and was more difficult to measure without interference from polarization artifacts. In the late 1980s, the virtues of backscattering ROA were implemented on a routine basis. Two forms of backscattering ROA, each having their own theoretical or experimental advantages, have been used extensively. They are unpolarized backscattering ICP ROA, given by

$$I_u^R(180°) - I_u^L(180°) = \frac{8K}{c}[12\beta(G')^2 + 4\beta(A)^2] \quad [39]$$

$$I_u^R(180°) + I_u^L(180°) = 4K[45\alpha^2 + 7\beta(\alpha)^2] \quad [40]$$

and backscattering $DCP_I$ ROA, given by

$$I_R^R(180°) - I_L^L(180°) = \frac{8K}{c}[12\beta(G')^2 - 4\beta(A)^2] \quad [41]$$

$$I_R^R(180°) + I_R^L(180°) = 4K[6\beta(\alpha)^2] \quad [42]$$

From these expressions, the advantages of $ICP_u$ and $DCP_I$ ROA in backscattering are apparent. The ROA invariants have a multiplicative advantage of 4 in backscattering and are additive rather than subtractive between the magnetic-dipole and electric-quadrupole ROA invariants. Comparing right-angle $ICP_z$ and $DCP_I$ Raman, both represent depolarized scattering with the backscattering stronger by a factor of 2. It can also be seen that the ROA expressions for backscattering $ICP_u$ and $DCP_I$ ROA are identical even though the $ICP_u$ Raman intensity represents classical polarized Raman scattering and $DCP_I$ Raman intensity the corresponding depolarized scattering. The additional Raman intensity in $ICP_u$ Raman $4K[45\alpha^2 + \beta(\alpha)^2]$ carries no ROA intensity since this additional strongly polarized Raman intensity corresponds to $DCP_{II}$ ROA which has not intensity in the FFR approximation as shown in **Table 1**. Thus backscattering $DCP_I$ ROA discriminates against $DCP_{II}$ ROA by analysing the circular polarization of the scattering light, whereas both of these intensities are present in $ICP_u$ ROA which has no such discrimination.

Another interesting property of these expressions is the possibility of isolating the ROA spectra for the magnetic-dipole and electric-quadrupole ROA invariants. By proper experimental normalization of the depolarized Raman spectra given in Equations [38] and [42], the corresponding and suitably scaled $ICP_z(90°)$ and $DCP_I$ (180°) ROA spectra can be added and subtracted to yield these invariants. This has been accomplished for the molecule (+)-*trans*-pinane as shown in **Figure 3**.

Of the two backscattering ROA schemes predominantly in use these days, the $ICP_u$ approach enjoys an advantage of experimental simplicity, while the $DCP_I$ enjoys an advantage of higher ROA intensity per Raman intensity, particularly in regions of strongly polarized Raman bands where no such intensity enters the $DCP_I$ Raman spectrum.

**Table 1** Values of Raman and ROA invariant coefficients for the far-from-resonance ROA and Raman intensity expressions in Equations [41]

| | | *Raman* (4*K*) | | *ROA* (8*K*/*c*) | | |
|---|---|---|---|---|---|---|
| $\xi$(°) | *Form* | $\alpha^2$ | $\beta(\alpha)^2$ | $\alpha G'$ | $\beta(G')^2$ | $\beta(A)^2$ |
| 0 | $ICP_u$ | 45 | 7 | 90 | 2 | −2 |
| | $SCP_u$ | 45 | 7 | 90 | 2 | −2 |
| | $DCP_I$ | 45 | 1 | 90 | 2 | −2 |
| | $DCP_{II}$ | | 6 | | | |
| 90 | $ICP_p$ | $\frac{45}{2}$ | $\frac{7}{2}$ | $+\frac{45}{2}$ | $+\frac{7}{2}$ | $+\frac{1}{2}$ |
| | $ICP_d$ | | | | +3 | −1 |
| | ICP* | $\frac{45}{3}$ | $\frac{10}{3}$ | $+\frac{45}{3}$ | $+\frac{10}{3}$ | |
| | $SCP_p$ | $\frac{45}{2}$ | $\frac{7}{2}$ | $+\frac{45}{2}$ | $+\frac{7}{2}$ | $+\frac{1}{2}$ |
| | $SCP_d$ | | 3 | | +3 | −1 |
| | SCP* | | $\frac{10}{3}$ | $+\frac{45}{3}$ | $+\frac{10}{3}$ | |
| | $DCP_I$ | $\frac{45}{4}$ | $\frac{13}{4}$ | $+\frac{45}{2}$ | $+\frac{13}{2}$ | $-\frac{1}{2}$ |
| | $DCP_{II}$ | $\frac{45}{4}$ | $\frac{13}{4}$ | | | |
| 100 | $ICP_u$ | 45 | 7 | | +12 | +4 |
| | $SCP_u$ | 45 | 7 | | +12 | +4 |
| | $DCP_I$ | | 6 | | +12 | +4 |
| | $DCP_{II}$ | 45 | 1 | | | |

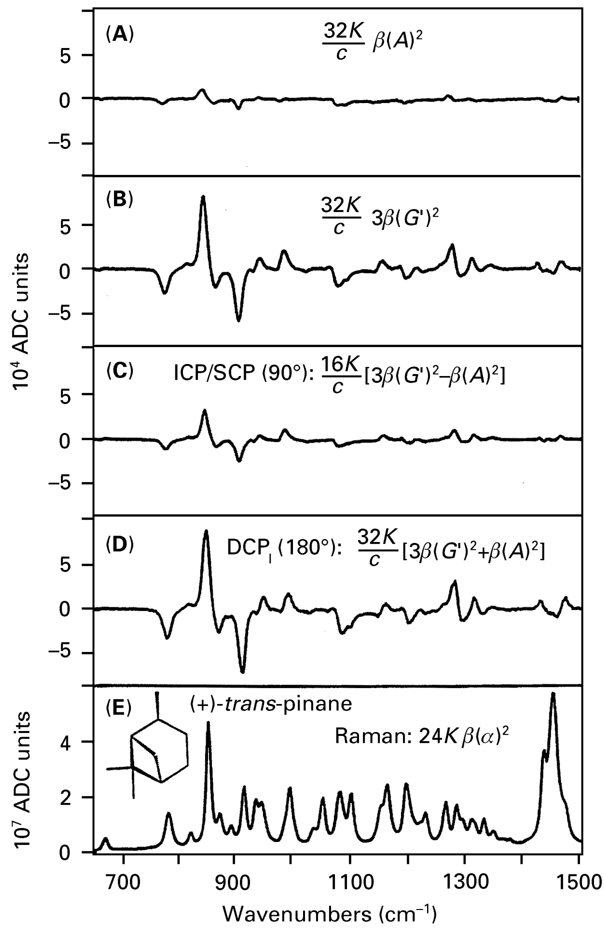



**Figure 3** Depolarized Raman and ROA spectra of (+)-*trans*-pinane, showing the decomposition of right-angle depolarized ICP and DCP$_I$ ROA into their magnetic-dipole and electric-quadrupole anisotropic ROA invariants.

## Single electronic state theory of resonance ROA

When the frequency of the incident laser radiation in a Raman scattering experiment is in resonance with a single electronic state (SES), it is well known that strong enhancement of the Raman scattering occurs. This is because the denominator of the resonant term in the polarizability expression, Equation [4] approaches zero and the value of the polarizability can increase by several orders of magnitude. This resonance condition brings simplifying conditions to the theory of Raman scattering as it becomes known as resonance Raman (RR) scattering. Under conditions of strong resonance, the non-resonant terms can be dropped and the contributions of all other electronic states can be dropped as too small to consider. The Raman polarizability in Equation [4] for a 0 to 1 vibrational transition in the ground electronic state, $g_0$ to $g_1$, becomes

$$(\tilde{\alpha}_{\alpha\beta})_{g_1,g_0} = \frac{1}{\hbar}\sum_{\nu}\frac{\langle g_1|\hat{\boldsymbol{\mu}}_\alpha|e\nu\rangle\langle e\nu|\hat{\boldsymbol{\mu}}_\beta|g_0\rangle}{\omega_{e\nu,g_0}-\omega_0+\mathrm{i}\Gamma_{e\nu}} \quad [43]$$

If the transition moment of the resonant electronic state is taken to lie in the $z$-direction, the general set of three Raman invariants and ten ROA invariants reduces effectively to only one Raman invariant and one ROA invariant, as

$$\alpha^2 = \frac{1}{9}|(\tilde{\alpha}_{zz})_{g_1,g_0}|^2 \quad [44]$$

$$\beta_s(\alpha)^2 = |(\tilde{\alpha}_{zz})_{g_1,g_0}|^2 \quad [45]$$

$$\alpha G = -\alpha \mathsf{G} = \frac{1}{9}\mathrm{Im}\Big[(\tilde{\alpha}_{zz})_{g_1,g_0}(\tilde{G}_{zz})^*_{g_1,g_0}\Big] \quad [46]$$

$$\beta_s(\tilde{G})^2 = \beta_s(\tilde{\mathsf{G}})^2 = \mathrm{Im}\Big[(\tilde{\alpha}_{zz})_{g_1,g_0}(\tilde{G}_{zz})^*_{g_1,g_0}\Big] \quad [47]$$

$$\beta_s(\tilde{A})^2 = \beta_s(\tilde{\mathsf{A}})^2 = 0 \quad [48]$$

$$\beta_a(\alpha)^2 = \beta_a(\tilde{G})^2 = \beta_a(\tilde{\mathsf{G}})^2\beta_a(\tilde{A})^2 = \beta_a(\tilde{\mathsf{A}})^2 = 0 \quad [49]$$

In addition, it can be shown the Raman invariant is proportional to the square of the electronic absorption strength for the resonant electronic state, and the ROA is proportional to the product of the electronic circular dichroism (CD) and the electronic of this state through the relationships,

$$|(\tilde{\alpha}_{zz})_{g_1,g_0}|^2 = (1/\hbar)|(\vec{\mu})^0_{eg}|^4 U(\omega_0) \quad [50]$$

$$\mathrm{Im}\Big[(\tilde{\alpha}_{zz})_{g_1,g_0}(\tilde{G}_{zz})^*_{g_1,g_0}\Big] = (1/\hbar)|(\vec{\mu})_{eg}|^2\mathrm{Im}\Big[(\vec{\mu}^0_{ge}\cdot(\vec{m})^{0*}_{eg}\Big]U(\omega_0) \quad [51]$$

where $\vec{\mu}^{\,\nu}_{eg}$ and $\vec{m}^{\,\nu}_{eg}$ are the electric-dipole and magnetic dipole transition moments, respectively, between the ground and resonant electronic states, respectively. In the case of resonance ROA in the SES limit, the most efficient form of ROA is backscattering $\mathrm{DCP_I}$ where the intensity expressions are:

$$I^{\mathrm{R}}_{\mathrm{R}}(180°) - I^{\mathrm{L}}_{\mathrm{L}}(180°) = \frac{96K}{c}\mathrm{Im}[(\tilde{\alpha}_{zz})_{g_1,g_0}(\tilde{G}_{zz})^*_{g_1,g_0}] \quad [52]$$

$$I_{\mathrm{R}}^{\mathrm{R}}(180^\circ) + I_{\mathrm{R}}^{\mathrm{L}}(180^\circ) = 24K|(\tilde{\alpha}_{zz})_{g_1,g_0}|^2 \qquad [53]$$

From these relationship emerges a deep connection between RROA in the SES limit and the electronic CD of the resonant electronic state. Since the anisotropy ratio, $g_{eg}$ is defined as the ratio of the CD intensity to the parent intensity, the following expression is found

$$\frac{I_{\mathrm{R}}^{\mathrm{R}}(180^\circ) - I_{\mathrm{L}}^{\mathrm{L}}(180^\circ)}{I_{\mathrm{R}}^{\mathrm{R}}(180^\circ) + I_{\mathrm{L}}^{\mathrm{L}}(180^\circ)} = -\left(\frac{4}{c}\right)\frac{\mathrm{Im}\left[(\vec{\mu}_{ge}^{0}\cdot(\vec{m})_{eg}^{0}\right]}{|(\vec{\mu})_{eg}|^2} = -g_{ge} \qquad [54]$$

The minus sign arises from the definition of ROA being right minus left circular polarization intensities, whereas the corresponding definition for CD is left minus right. Resonance ROA promises to open up new applications for ROA in the same way that resonance Raman spectroscopy extended the reach of Raman spectroscopy, particularly for biological applications.

## Simple models of ROA

Before the development of molecular orbital approaches to the calculation of ROA, a number of simple models of ROA were advanced to provide a conceptual basis for understanding ROA spectra. In various ways, these models arise from considering the polarizability of the molecule $\alpha_{\alpha\beta}$ as the sum of local polarizability units, such as those associated with bonds or atoms. While the polarizability and its individual local components are independent of the location of the origin of the molecule, the magnetic dipole and electric quadrupole ROA tensors do depend on this origin. In the FFR approximation, we can express these tensors in terms of local contributions as

$$\alpha_{\alpha\beta} = \sum_i \alpha_{i,\alpha\beta} \qquad [55]$$

$$G'_{\alpha\beta} = \sum_i G'_{i,\alpha\beta} + \sum_i \tfrac{1}{2}\omega\varepsilon_{\beta\gamma\delta}R_i\alpha_{i,\alpha\delta} \qquad [56]$$

$$A_{\alpha\beta\gamma} = \sum_i A_{i,\alpha\beta\gamma} - \sum_i \tfrac{1}{2}[3\boldsymbol{R}_{i,\beta}\alpha_{i,\alpha\delta} + 3\boldsymbol{R}_{i,\gamma}\alpha_{i,\alpha\beta} - 2\boldsymbol{R}_{i,\delta}\alpha_{i,\alpha\delta}\delta_{\beta\gamma}] \qquad [57]$$

where $R_{i,\alpha}$ is the location of the $i$th polarizability unit relative to the origin. In the application of ROA models it is assumed that only the local polarizability units are non-zero and that contributions from the local ROA tensors, $G'_{i,\alpha\beta}$ and $A_{i,\alpha\beta\gamma}$, are zero. The simplest of these models is the two-group model which describes the ROA from the two local symmetric polarizability groups in a molecule that are twisted in a chiral sense with respect to one another. ROA arises from the interference of the independent Raman scattering from each of these two groups. Only limited success has been achieved in the use of these models to understand the details of ROA spectra. For a quantitative understanding, one must use *ab initio* molecular orbital methods.

## Ab initio calculations of ROA

The first calculations of ROA using *ab initio* molecular orbital methods have been carried out recently. This has been achieved by starting with expressions for the polarizability and Rayleigh optical activity tensors in the zero-frequency limit of the FFR approximation as

$$\alpha_{\alpha\beta} = \frac{2}{\hbar}\sum_{j\neq n}\frac{\mathrm{Re}\left[\langle n|\hat{\boldsymbol{\mu}}_\alpha|j\rangle\langle j|\hat{\boldsymbol{\mu}}_\beta|n\rangle\right]}{\omega_{jn}} \qquad [58]$$

$$G'_{\alpha\beta} = -\frac{2\omega_0}{\hbar}\sum_{j\neq n}\frac{\mathrm{Im}\left[\langle n|\hat{\boldsymbol{\mu}}_\alpha|j\rangle\langle j|\hat{\boldsymbol{m}}_\beta|n\rangle\right]}{\omega_{jn}^2} \qquad [59]$$

$$A_{\alpha\beta\gamma} = \frac{2}{\hbar}\sum_{j\neq n}\frac{\mathrm{Re}\left[\langle n|\hat{\boldsymbol{\mu}}_\alpha|j\rangle\langle j|\hat{\Theta}_{\beta\gamma}|n\rangle\right]}{\omega_{jn}} \qquad [60]$$

where the distinction between the initial and final states is not needed. ROA and Raman intensities can be obtained from these tensors by calculating their variation with the normal coordinates of vibrational motion. The method by which these tensors have been calculated is the field perturbation approach. The summation over all the excited states $j$ can be avoided by substituting field perturbed wavefunctions in first-order perturbation theory for their nonperturbed counterparts as

$$\alpha_{\alpha\beta} = 2\sum_{j\neq n}\mathrm{Re}\left[\langle n|\hat{\boldsymbol{\mu}}_\alpha|n'(E_\beta)\rangle\right] \qquad [61]$$

$$G'_{\alpha\beta} = -2\hbar\omega_0\mathrm{Im}\left[\langle n'(E_\alpha)|n'(B_\beta)\rangle\right] \qquad [62]$$

$$A_{\alpha\beta\gamma} = 2\langle n|\Theta_{\beta\gamma}|n'(E_\alpha)\rangle \qquad [63]$$

where $E_\alpha$ and $B_\alpha$ are components of the electric and magnetic fields, and the prime on the wavefunction indicates the first derivative with respect to the field. In **Figure 4** we show the results of a comparison of *ab initio* ROA calculations for the molecule L-alanine in aqueous solution. A high degree of correspondence has been achieved between theory and experiment. This demonstrates that ROA can be calculated with success for small chiral molecules of biological significance.

## Applications to biological molecules

Experimental measurements of ROA were first achieved in the mid-1970s. Since that time ROA spectra of many classes of molecules of biological significance have been published. The include terpenes, amino acids, sugars, carbohydrates, peptides, proteins, and nucleic acids. Through these studies, ROA can be seen to be a sensitive probe of the stereo-conformational detail of these molecules in their native environments.

As described in this article, the theory of ROA is rich in content, offering experimentalists many options in the measurements of ROA spectra. These include the wavelength of the exciting radiation and its proximity to allowed transitions to excited electronic states in the molecules. Also of importance is the polarization modulations scheme and the scattering geometry. Recent studies have established that the optimum polarization and scattering conditions for biological applications are either unpolarized ICP or in-phase DCP in backscattering geometry.

The theoretical understanding of ROA is well in hand. What remains to be demonstrated is the ability to calculate ROA intensities accurately across a wide range of theoretical limits, including the most general cases, for most molecules of biological interest. To date, only the simplest level of theory has been used

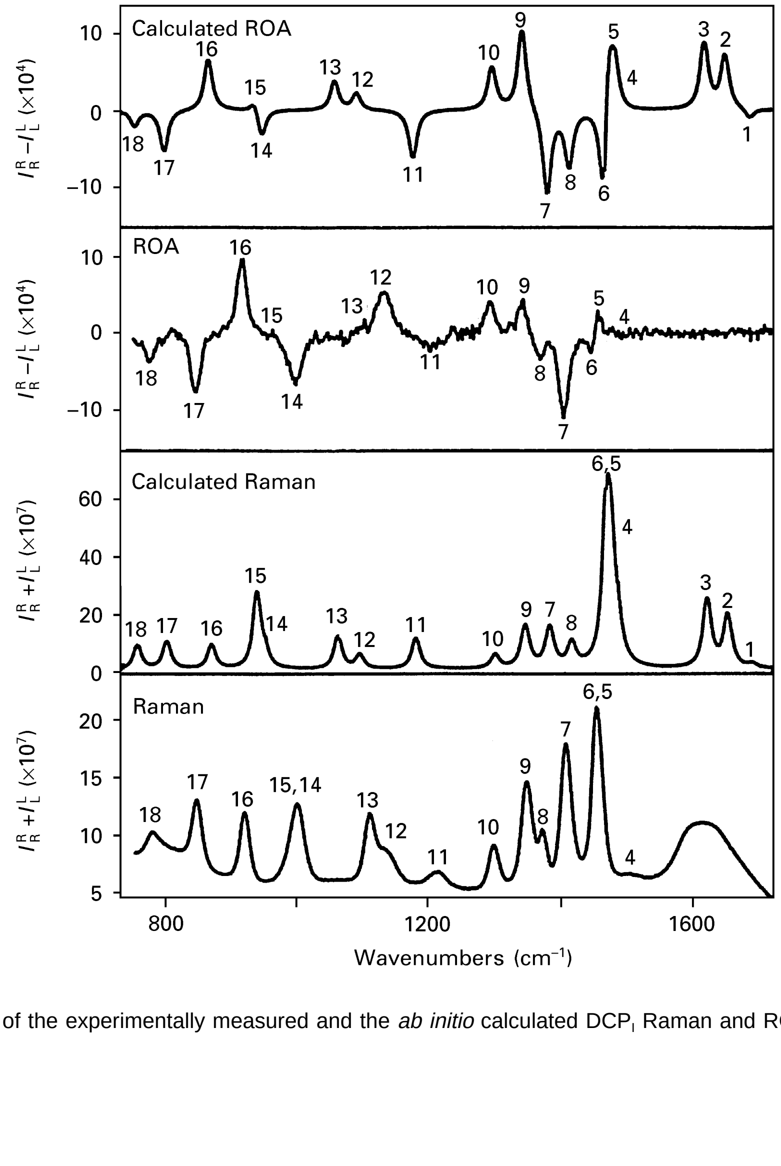



**Figure 4** Comparison of the experimentally measured and the *ab initio* calculated $DCP_I$ Raman and ROA spectra of L-alanine in aqueous solution.

for *ab initio* calculations of relatively simple biomolecules. Beyond the zero-frequency limit of the FFR approximation are the dynamic frequency limit, the near resonance conditions, and the strong resonance conditions involving more than one electronic state. As demonstrated above, the case of strong resonance with a single electronic state is trivial in that the ROA spectrum is completely predicted from the resonance Raman spectrum and the electronic CD of the resonance electronic state.

The problem of extending ROA calculations to molecules of increasing complexity will accompany the steady increase in the power of computational calculations made possible by advances in the speed and memory capacity of computers.

With the realization of improvements in the measurements and theoretical calculation of intensities, ROA will assume a place of special importance among our spectroscopic probes of the structure and dynamics of molecules of biological interest.

## List of symbols

$\tilde{a}_{\alpha\beta}$ = general scattering tensor; $D_1$–$D_5$ = constants; $\tilde{e}^{i}_{\alpha}$ = polarization vector for incident light; $\tilde{e}^{s}_{\alpha}$ = polarization vector for scattered light; $\tilde{E}^{(0)}$ = electric field strength of incident radiation; $g_{eg}$ = anisotropy ratio; $\hbar$ = Planck's constant/2π; $I$ = intensity of light scattering; $K$ = a constant; $n$ = initial electronic state; $\tilde{\boldsymbol{n}}^{i}_{\alpha}$, $\tilde{\boldsymbol{n}}^{d}_{\alpha}$ = propagation vectors for incident and scattered light, respectively; $m$ = final electronic state; $(\vec{m})^{v}_{eg}$ = magnetic dipole transition moment; $R$ = distance from the scattering to the detector; $R_{i,\alpha}$ = location of the $i$th polarizability relative to the origin; $\alpha^2$ = the isotropic invariant; $\tilde{\alpha}_{\alpha\beta}$ = Raman polarizability tensor; $\beta_a(\alpha)^2$ = antisymmetric anisotropy; $\beta_s(\alpha)^2$ = symmetric anisotropy; $\varepsilon_{\alpha\beta\gamma}$ = the unit antisymmetric tensor; $\xi$ = scattering angle; $\mu_0$ = magnetic permeability; $(\vec{\mu})^{v}_{eg}$ = electric dipole transition moment; $\omega$ = angular frequency.

*See also:* **Biochemical Applications of Raman Spectroscopy; Chiroptical Spectroscopy, Emission Theory; Chiroptical Spectroscopy, General Theory; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Electromagnetic Radiation; ORD and Polarimetry Instruments; Raman Optical Activity, Applications; Raman Optical Activity, Spectrometers; Raman Spectrometers; Scattering Theory; Vibrational CD Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading

Barron LD (1982) *Molecular Light Scattering and Optical Activity*. Cambridge: Cambridge University Press.

Barron LD and Hecht L (1994) Vibrational Raman optical activity: from fundamentals to biochemical applications. In: Nakanishi K, Berova ND and Woody RW (eds) *Circular Dichroism: Principles and Applications*. New York: VCH.

Barron LD, Hecht L, Bell AF and Wilson G (1996) Recent developments in Raman optical activity of biopolymers. *Applied Spectroscopy* **50**: 619–629.

Nafie LA and Che D (1994) Theory and measurement of Raman optical activity. In: Evans M and Kielich S (eds) *Modern Nonlinear Optics, Part 3*, Vol 85, pp 105–149. New York: Wiley.

Nafie LA, Yu G-S, Qu X and Freedman TB (1994) Comparison of IR and Raman forms of vibrational optical activity. *Faraday Discussions* **99**: 13–34.

Nafie LA, Yu G-S and Freedman TB (1995) Raman optical-activity of biological molecules. *Vibrational Spectroscopy* **8**: 231–239.

Nafie LA (1995) Circular polarization spectroscopy of chiral molecules. *Journal of Molecular Structure* **347**: 83–100.

Nafie LA (1996) Vibrational optical activity. *Applied Spectroscopy* **50**: 14A–26A.

Nafie LA (1996) Theory of resonance Raman optical activity: The single-electronic state limit. *Chemical Physics* **205**: 309–322.

Nafie LA (1997) Infrared and Raman vibrational optical activity: theoretical and experimental aspects. *Annual Reviews in Physical Chemistry* **48**: 357–386.

# Raman Spectrometers

**Bernhard Schrader**, Universität Essen, Germany

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Methods & Instrumentation**

## Synopsis

Raman spectrometers are quite different from 'ordinary' spectrometers. In Raman spectra the very weak Raman lines are accompanied by the extremely strong Rayleigh line. The stray light of it produces a background in the spectrometer which may be more intense by orders of magnitude than the Raman lines. Therefore, a Raman spectrometer has to combine the elimination of the Rayleigh line with the spectral dispersion and isolation of the Raman lines. Additionally, the necessary resolving power of Raman spectrometers has to be considerably higher compared with 'ordinary', e.g. infrared spectrometers.

This article describes the elements of Raman spectrometers for routine analyses which are available commercially. Instruments designed only for special research are not covered. Only spectrometers for 'classical' (linear) Raman scattering are mentioned, not those for observing resonance Raman scattering (RRS), surface-enhanced Raman scattering (SERS) and all nonlinear Raman techniques; they are described elsewhere in this Encyclopedia.

## Rayleigh and Raman scattering, the Raman spectrum

Raman spectra are complementary to infrared spectra. Both are composed of lines (or bands of lines) which are images of the vibrations of molecules. The intensity of Raman lines represents the change of the molecular polarizability by a vibration, while the intensity of infrared lines represents the change of the molecular dipole moment (actually the square of the change of the molecular polarizability or the molecular dipole moment). The Raman lines are accompanied by the 'Rayleigh line' at the wavelength of the exciting radiation; its intensity is proportional to the square of the molecular polarizability. In addition, the intensity of this 'unshifted' line is enhanced further by the exciting radiation which is directly scattered at the surfaces of the particles of powders or at the windows of the sample cuvettes.

When monochromatic exciting radiation of light quanta $h\nu_0$ hits a molecule, an *elastic scattering process*, i.e. *Rayleigh scattering* of quanta having the energy $h\nu_0$ has the highest probability. The *inelastic scattering process*, during which vibrational energy $h\nu_s$ is exchanged with the molecules, has a much lower probability, and is called *Raman scattering*. It emits quanta of energy $(h\nu_0 \pm h\nu_s)$. At ambient temperature most molecules are in their vibrational ground state. According to Boltzmann's law, a much smaller number are in their vibrationally excited state. Therefore, Raman scattering of quanta of energy $(h\nu_0 - h\nu_s)$ has much higher probability than the Raman scattering of quanta of energy $(h\nu_0 + h\nu_s)$. While studying fluorescence spectra Stokes, in 1852, postulated that the wavelength of light produced by photoluminescence, fluorescence or phosphorescence, is usually longer than that of the exciting light. In analogy, Raman lines are referred to as *Stokes lines* and *anti-Stokes lines*. Stokes lines are caused by the quanta of lower energy. Since their intensities are higher than those of anti-Stokes lines, only they are usually recorded as Raman spectrum. The line intensities are usually drawn over the Raman shift, $\nu_s$, in wavenumbers. The intensity ratio of Stokes and anti-Stokes lines of the same Raman shift allows, by employing the Boltzmann equation, the determination of the temperature of the sample under illumination.

The line width $\Delta\nu_s$ of the infrared and Raman lines is about the same. However, the necessary resolving power in the infrared spectrum for isolating an infrared line of width $\Delta\nu_s = 10$ cm$^{-1}$ at a wavelength of 10 µm equivalent to a wavenumber of $\nu_s = 1000$ cm$^{-1}$ is $R_{IR} = \nu_s/\Delta\nu_s = 100$. In the Raman spectrum excited with radiation of wavelength of 500 nm equivalent to $\Delta\nu_0 = 20\,000$ cm$^{-1}$ the same vibration, at $\nu_s = 1000$ cm$^{-1}$, is recorded at an absolute wavenumber of $(\nu_0 - \nu_s) = 20\,000 - 1000 = 19\,000$ cm$^{-1}$. Therefore, the necessary resolving power is $R_{RA} = 19\,000/10 = 1900$, this is larger by a factor of 19 than that necessary for recording the infrared line. This shows that Raman spectrometers have to supply a considerably large resolving power than infrared spectrometers.

The intensity of the Raman lines is given by the square of the change of the molecular polarizability by the vibrations. There is an analogy between the molecular polarizability and the molecular volume:

vibrations that modulate the molecular volume are *Raman active*.

**Figures 1A** and **1B** show a potential energy diagram of a molecule with the vibrational ground state and one vibrational excited state. The direct transition can be observed (**Figure 1A**) by an absorption of a light quantum from the infrared spectral range. There is another way of exciting vibrational states (**Figure 1B**). The molecule is illuminated with light quanta of higher energy, from the visible or near-infrared range. These quanta are, on the one hand, scattered elastically, producing the Rayleigh line. On the other hand, the inelastic scattering process emits a light quantum of the energy of the exciting radiation which is reduced by the vibrational energy. Therefore, infrared absorption and Raman scattering may produce—by different mechanisms—molecules in exactly the same vibrational excited state. The molecular symmetry may 'allow' or 'forbid' the activity in the Raman or infrared spectrum. An example is the *rule of mutual exclusion*: If a molecule has a centre of symmetry, then infrared active vibrations are forbidden in the Raman spectrum and vice versa.

For the observation of 'classical' Raman spectra exciting radiation from the visible part of the spectrum is used. When molecules have electronic energy levels, which can be excited by light quanta from the visible range of the spectrum, the molecules may undergo a transition into an electronic excited state. As a consequence the molecule emits fluorescence radiation (**Figure 1C**). Since this process has a quantum yield much larger than the Raman effect (of the order of 1 versus $10^{-6}$–$10^{-11}$) fluorescence is quite strong, much stronger than the Raman lines. Therefore, Raman lines are then overlaid by the strong (quasi-continuous) fluorescence radiation. This process may hinder the observation of Raman lines. The fluorescing molecules may be impurities, which can be removed by purification (by distillation, sublimation, or recrystallization). However, the fluorescing molecules may be normal constituents of the sample. This is true for all living cells or tissues. They have a chemical 'machinery', composed of enzymes, which are absorbing visible radiation. 'Purification' by destroying the fluorescing molecules will destroy the cells or tissues. The only way to prevent superimposing by fluorescence is excitation of Raman spectra with light quanta having a low energy which is not sufficient to excite electronic states (**Figure 1D**), this is done by using the Nd:YAG laser radiation at 1064 nm.

**Figure 2** shows the Stokes and anti-Stokes part of the Raman spectrum of coumarin, excited by the Ar-ion laser line at 514.5 nm. Three abscissa scales are drawn: the wavelength scale, the absolute wavenumber scale and the Raman shift, the energy difference between the exciting energy and the energy of the light quanta, scattered by the Raman effect. This is the only scale which is usually used to draw Raman spectra.

## The components of Raman spectrometers

A Raman spectrometer analyses the radiation scattered by molecules, when they are illuminated with monochromatic exciting radiation. The scattered radiation is composed of the strong Rayleigh line and the very weak Raman lines. The Rayleigh line has a radiant power that may exceed that of the Raman lines by about $10^6$ up to $10^{15}$. The electric signal $S$ produced by a radiation detector is proportional to

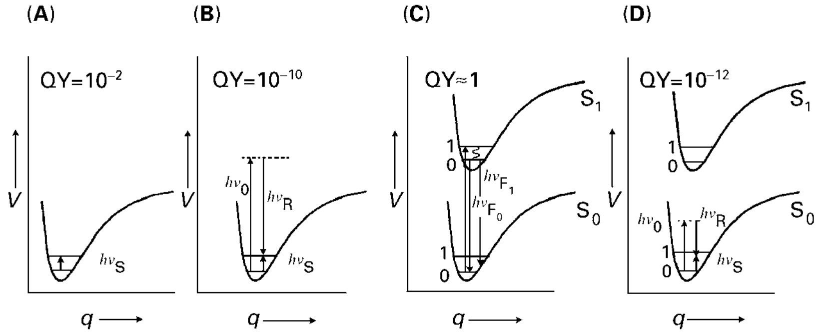



**Figure 1** Observation of the excitation of a vibrational state in the electronic ground state $S_0$ by (A) infrared absorption; (B) Raman scattering; excitation in the visible range ($\lambda$ = 488 nm); (C) absorption of the exciting radiation with subsequent fluorescence, (D) Raman scattering, excitation in the near-infrared range, ($\lambda$ = 1064 nm), the energy of the exciting light quanta is only 46% of that of (B), $V$ = potential energy, QY = order of magnitude of the quantum yield, $q$ = normal coordinate (describing the vibrational motion), $S_0$ = electronic ground level, $S_1$ = electronic excited level. Reproduced from Schrader B (ed) (1995) *Infrared and Raman Spectroscopy*. Weinheim: VCH Publishers, with permission of VCH.

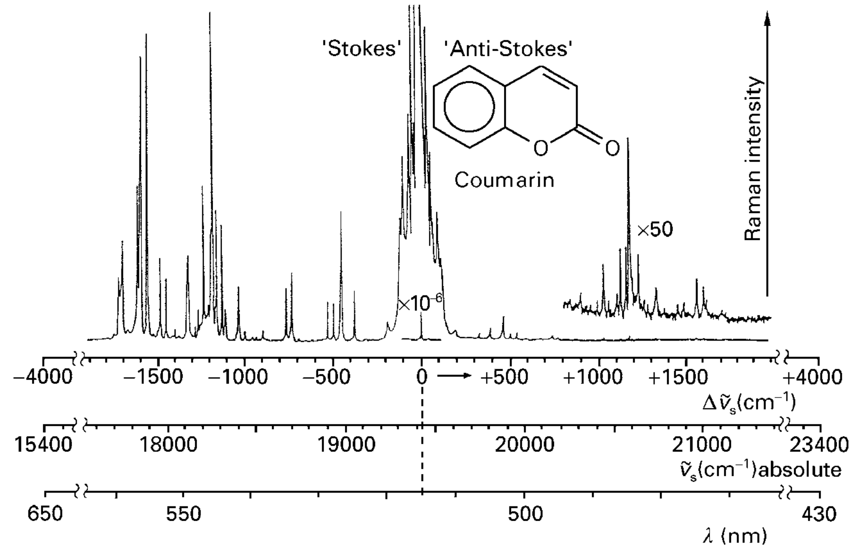



**Figure 2** Raman spectrum of coumarin, excited with the radiation of the $Ar^+$ laser at $\lambda = 514.53$ nm equivalent to $\nu = 19\,430$ cm$^{-1}$. Reproduced from Schrader B (ed) (1995) *Infrared and Raman Spectroscopy*. Weinheim: VCH Publishers, with permission of VCH.

the radiant power $\Phi$ (W), received by the detector. A Raman spectrometer has to facilitate recording of the Raman lines with a high signal-to-noise ratio ($S/N$) sufficiently resolved.

The radiant power transmitted $\Phi$ by any optical instrument is given by

$$\Phi = LG\tau$$

Here, $L$ describes the *radiance* of the radiation source, power per area and solid angle (W cm$^{-2}$ sr$^{-1}$), $G$ is the *optical conductance*, solid angle times area (sr cm$^2$), and $\tau$ represents the *transmission* of the whole system. In order to record Raman spectra with a large $S/N$, all factors, $L$, $G$ and $\tau$, have to be maximal. $L$ is optimized by appropriate sample arrangements. $G$ is maximal when the spectrometer uses a maximal area of the essential elements (the prism or grating or the beam splitter of an interferometer and the entrance aperture). A maximal value of $\tau$ is guaranteed by a proper design of the instrument, especially by antireflection coating of all glass/air interfaces and by a maximal reflectivity of all mirrors. As an approximation the optical conductance of a pair succeeding elements of a spectrometer constructed by apertures, lenses or mirrors having an area $F_i$, $i = 1$ and 2 at distance of $a_{12}$ is given by

$$G \approx F_1 F_2 / a_{12}^2$$

The flux through a spectrometer is appropriately described by

$$\Phi_\nu = L_\nu G_\nu (\Delta\nu^2)\tau$$

Here the subscript v stand for ‘per wavenumber’. $L_v$, the spectral radiance, is a property of the radiation source, $G_v$ the spectral optical conductance and $\Delta\nu$ the bandwidth of the instrument (in wavenumber units), $\tau$ is the overall transmission factor of the entire instrument.

In every well-constructed optical instrument the optical conductance of all pairs of succeeding elements should be the same. The overall optical conductance of any instrument is given by the smallest optical conductance of any succeeding pair of elements.

The practical resolving power is given by the ratio of the wavenumber (or wavelength) and the bandwidth. Its upper limit is the theoretical resolving power, $R_0$ determined by the properties of the dispersive elements, the number of grating rules in a grating spectrometer or the pathlength difference of the interfering rays in an interferometer (**Figures 3** and **4**).

The ratio of the spectral optical conductance of an interferometer, compared to a grating spectrometer (the Jacquinot advantage) is given by

$$G_\mathrm{v}^\mathrm{I} / G_\mathrm{v}^\mathrm{G} \approx 2\pi f / h$$

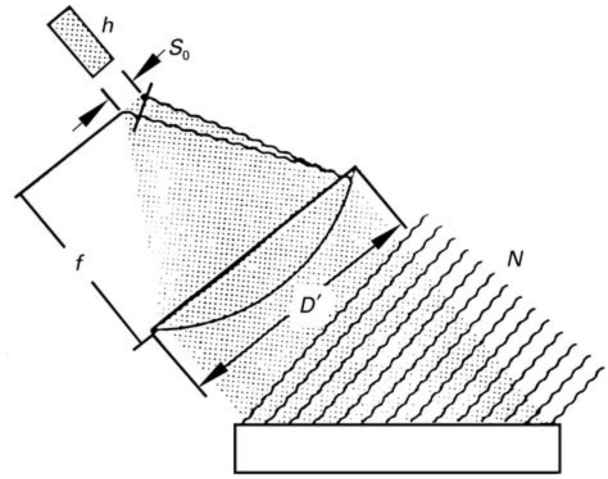



**Figure 3** The main components of a grating spectrometer: *N* is the number of interfering rays, given by the number of rules; $S_0$ is the halfwidth of the diffraction pattern of the collimator lens with diameter *D′* and focal length *f*, which determines the 'optimal' slit width, *h* is the slit length. Reproduced from Schrader B (ed) (1995) *Infrared and Raman Spectroscopy*. Weinheim: VCH Publishers, with permission of VCH.

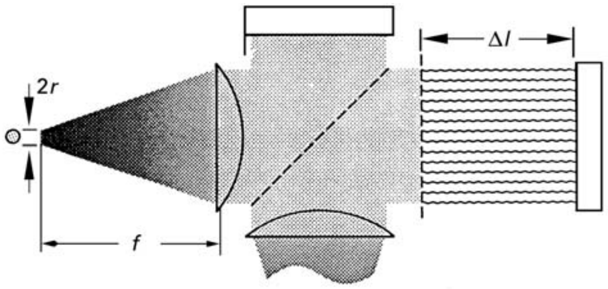



**Figure 4** The significant features of an interferometer: Δ*l* displacement of the moving mirror, 2*r* diameter of the Jacquinot stop. Reproduced from Schrader B (ed) (1995) *Infrared and Raman Spectroscopy*. Weinheim: VCH Publishers, with permission of VCH.

with *f* the focal length of the collimator and *h*, the slit height of the grating spectrometer. For spectrometers with the same beam area at the interferometer and the grating this factor amounts to about 500. However, common interferometers generally have a smaller beam area than grating instruments, therefore the Jacquinot advantage of actual instruments is of the order of 100. This can be compensated by using array detectors with grating polychromators, employing the multichannel advantage.

## Light sources

Raman spectroscopy began in 1928 by using the lines of a mercury discharge lamp at 435.8 or 404.7 nm as the exciting radiation. Since 1960 lasers have been available as ideal monochromatic sources of exciting radiation. The ruby laser (694.0 nm), HeNe-laser (632.8 nm), $Ar^+$ laser (488.0 and 514.5 nm), the GaAs diode laser (780 nm) and the Nd:YAG laser (1064 nm) are mainly employed. The light flux necessary for recording of Raman spectra is of the order 10 up to 1000 mW.

All laser lines in the visible range of the spectrum, especially the $Ar^+$ laser lines, but also the line at 780 nm, may excite fluorescence spectra, overlaying the Raman spectra. Excitation by radiation with longer wavelengths reduces the danger of fluorescence. With the exciting radiation of 1064 nm the minimal probability of fluorescence is reached.

## Sample arrangements

'Classical' Raman arrangements use the Raman radiation which is emitted at an angle of about 90° relatively to the direction of the exciting radiations (90° arrangement). This is the straightforward arrangement to illuminate the entrance slit of a grating spectrometer by the Raman radiation excited by a focused laser beam in a liquid. However, when interferometers having a circular entrance aperture are employed it has been proven to be superior to analyse the Raman intensity emitted at about 180° to the exciting radiation (180° or back-scattering, arrangement).

In order to record Raman spectra of a sufficiently large *S/N* in an acceptable time the spectral radiance of the sample has to be maximal. Since Raman spectra are very weak and an excessive large power of the exciting radiation could destroy the sample, sample arrangements have to be designed very carefully. In particular, the ratio of the usable intensity of the Raman radiation at the entrance aperture of a spectrometer versus the intensity of the available exciting radiation, the figure of merit of the sample cell, has to be optimized. This can be done, on the one hand, by observing the Raman radiation produced from a large thickness of the sample, for instance with *end-on* capillary cells (**Figure 5D**) or, on the other hand, by employing different kinds of multiple reflection cells. The largest part of the exciting radiation in cells as in **Figure 5A** is not used for producing Raman radiation, scattered in a direction to the spectrometer. Since the exciting radiation just passes the sample once, more than 99% of the exciting radiation is lost. With spherical mirrors surrounding the sample cell this radiation is reflected back to the sample. Also, the Raman radiation, which is not taken up by the spectrometer is reflected back to the sample. This is done by cells as shown in **Figure 5B**. A 'notch filter'

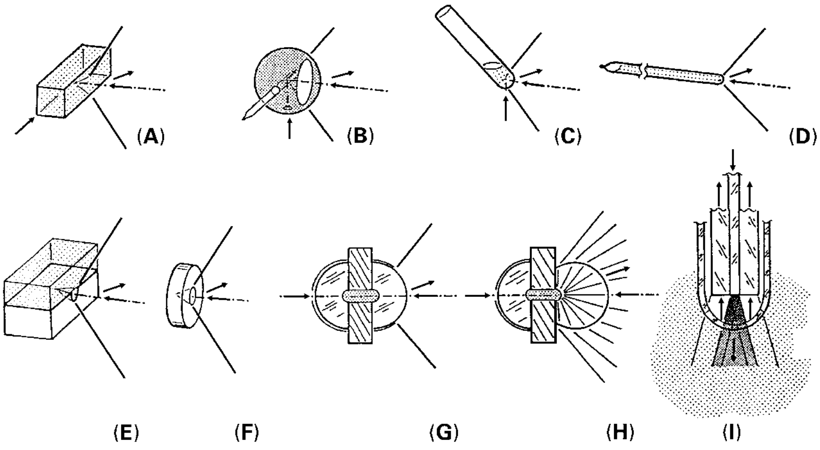



**Figure 5** Sample arrangements for Raman spectroscopy: (A) rectangular cell; (B) spherical cell for liquids and powders which are in a melting point capillary at the centre of the sphere with a reflecting surface; (C) liquid in an NMR tube, the axis of which has angle of 45° relative to the axis of the entrance optics. The cuvettes (A)–(C) may be used in a 90° or a 180° arrangement; (D) light pipe cuvette; (E) arrangement for solids or surface layers: the sample is placed upon a block of aluminium or stainless steel with a polished conical indentation, providing a multiple reflection arrangement; (F) tablet with a cone-shaped bore; (G) arrangement for Raman spectroscopy of powders in 0° and 180° arrangement, the half-spheres have a reflecting surface which reflects the exciting and Raman radiation back to the sample increasing the usable Raman intensity; (H) same as (G), but one half-sphere is exchanged against a Weierstrass lens (Weierstrass 1856), which collects the radiation emitted into a solid angle of ~2π; (I) sample head for a two-way bundle of optical fibres for spectroscopy of liquids or powders. The head can be introduced directly into the sample, it is protected by a cover in order to prevent sticking and pyrolysis of the sample at the central fibre which transports the exciting radiation from the laser to the sample. Reproduced from Schrader B (ed) (1995) *Infrared and Raman Spectroscopy.* Weinheim: VCH Publishers, with permission of VCH.

just behind the entrance optics reflects the exciting radiation emerging in the direction as the Raman radiation recorded by the spectrometer, while it transmits the Raman radiation. When the overall reflectivity of the multiple reflection cell for the exciting radiation is $\rho$, the increase of the observable Raman intensity is given by $I = 1/(1 - \rho)$. With $\rho$ = 90%, the Raman intensity will be larger by a factor of 10, compared to the original intensity. Since also the intensity of the effective Raman radiation is increased by the multiple reflection cell, the intensity of the Raman spectrum increases further (to a maximal of $I^2$). **Figure 5** gives examples of different sample cells for liquids and powders which have been used successfully. Rectangular sample cells, which are supplied with most Raman spectrometers, have a very small figure of merit, since they do not employ multiple reflection enhancement or the observation of an increased effective thickness of the sample.

In **Figure 5I** an arrangement for investigation of remote samples by fibre optics is shown. The high transmittance of optical fibres in the NIR range makes possible the recording of Raman spectra of samples, which are located up to several hundred metres away from the spectrometer. The quartz fibres, developed for telecommunications, have a large transmission in the spectral range of Raman spectra excited in the NIR—more than 80% per km! One fibre transports the exciting radiation from the laser to the sample, a second fibre or fibre bundle transports the Raman radiation from the sample to the spectrometer. However, it cannot be avoided that the Raman lines of quartz are excited in both fibres. Therefore, the sample head has to be combined with a transmission filter which eliminates, on the one hand, the Raman line of the quartz from the exciting radiation and, on the other hand, the exciting radiation coming back from the sample on the way to the spectrometer (see the next section: Rayleigh filters). Some companies supply such sample heads which allow remote product or production control.

The laser radiation can be focused to a diameter of the order of a few multiples of its wavelength. Therefore, even by using the normal entrance optics of the spectrometer one can investigate Raman spectra of micro samples. However, in order to be able to exactly adjust the area from which the Raman spectrum is taken, common microscopes, able to adjust the sample by observation with visible light are modified for the excitation and observation of Raman spectra. Such microscopes may even allow confocal observation with 3-dimensional spatial resolution. It must be remembered that the optical conductance of microscopes is quite small, therefore much longer observation times are needed for Raman spectroscopy of micro samples. This cannot be compensated by a

larger power of the laser radiation, since this may then overheat the sample.

### Rayleigh filters

A spectrometer set to pass radiation of a particular wavelength band always has a small amount of stray radiation of other wavelengths. The Rayleigh line is stronger by $10^6$–$10^{12}$ than the Raman lines. In ordinary spectrometers the Rayleigh scattering produces stray radiation, which conceals the Raman spectrum completely. Therefore, the Rayleigh radiation has to be eliminated in order to record Raman lines with a maximal signal/background ratio. Different principles are employed for this purpose:

Absorption filters composed of solutions with appropriate absorption bands or glass filters with absorption edges are able to absorb the spectral band of the Rayleigh radiation. Especially interesting is the elimination of the exciting line by atomic absorption of the same element, which produces the exciting radiation. Rasetti, in 1930, used the low-pressure mercury arc to excite rotational Raman spectra of gases with the mercury resonance line at 253.6 nm. With a drop of mercury in the spectrograph he could completely absorb the Rayleigh radiation. Similar procedures use the resonance absorption of rubidium and caesium.

Interference line filters reflect the whole spectrum with a high reflectivity except for the spectral line where the filter has the largest transmittance. The Rayleigh line intensity can thus be reduced by reflection on the filter. Combinations of such filters in a row make up a very efficient Rayleigh filter.

The same filters are employed satisfactorily to reduce the unwanted plasma lines of the laser radiation for the excitation of the Raman spectra, also, Raman lines of the optical fibres transporting the exciting radiation to the sample can be eliminated using interference transmission filters.

Volume holographic optical elements developed for military uses, are successfully applied for Raman spectroscopy: a so-called notch filter reflects a laser line virtually completely, its intensity is reduced by >6 orders of magnitude, with a transmission of about 70% at the other wavelengths.

Holographic laser bandpass filters reflect the laser radiation by >90% at an angle of exactly 90° while reducing the intensity of all unwanted plasma or Raman lines (of an optical fibre) considerably.

A similar construction, a so-called Holoplex transmission grating, disperses the Raman spectrum for recording on a CCD (charge-coupled devices), which are very sensitive detector arrays. They can be fabricated with 2-dimensional dispersion allowing the instantaneous recording of several spectral channels on a 2-dimensional CCD detector as for an echelle spectrograph.

### Dispersive mono- and polychromators, interferometers

When illuminated with monochromatic radiation a single monochromator usually shows continuous stray light of the order of $10^{-5}$ of the intensity of the monochromatic radiation. Therefore, 2 or 3 monochromators in series combined with additive dispersion reduce the stray radiation by about $10^{-10}$ or $10^{-15}$, respectively. However, the intensity of the Raman lines is also reduced when passing a monochromator: Since every monochromator has a transmittance only of about 30%, this means that a double monochromator has only a transmittance of 9%, a triple monochromator of 2.7%. Such monochromators are usually very voluminous and expensive. However, they are widely used for recording of Raman spectra with single detectors (photomultipliers).

When array detectors are used to record simultaneously the whole spectrum, or parts of it, three monochromators are used in a special arrangement. The first two monochromators are combined in a subtractive arrangement. A diaphragm at the middle slit blocks out the Rayleigh line. The third grating instruments acts as a polychromator, it produces a spectrum directly upon the elements (pixels) of a CCD.

Interferometers record an interferogram of a spectrum. By applying the Fourier transformation, the original spectrum is calculated. Jacquinot has pointed out in 1954, that interferometers have a considerably higher optical conductance, by a factor of 100–500, compared to prism or grating spectrometers. Interferometers make use of *the multiplex advantage* when compared to dispersive spectrometers. By recording all channels simultaneously with the same single detector, the detector noise is therefore distributed over all spectral channels. It is not recorded at every spectral channel separately as for prism or grating spectrometers. This increases the *S/N* considerably.

For recording the weak Raman spectrum excited by the Nd:YAG laser line at 1064 nm, interferometers are successfully used. They have to be combined with very powerful Rayleigh filters. The quantum noise of every strong line is distributed over the whole interferogram. By Fourier transformation it is distributed as white noise over the whole spectrum. To avoid this *multiplex disadvantage* all strong lines have to be removed from the spectrum to be analysed.

### Detectors

The first generation of Raman spectroscopists used photoplates for the recording of Raman spectra. Later, photomultipliers were used as very powerful single channel detectors at the exit slit of a monochromator recording the Raman spectrum sequentially.

They are now being replaced by metal oxide semiconductors where charge produced by light quanta is stored in a small area, a pixel. Arrays composed of many independent pixels store the charge pattern corresponding to the irradiation pattern. These arrays of 'charge-coupled devices', CCDs, can be linear or two-dimensional, thus storing spectra or images of the sample. Since the number of pixels is limited (usually to about 1024 in a row) spectra can be recorded either *completely, in low resolution*, or *in high resolution sequentially*. Combined with an echelle spectrograph several channels representing different spectral orders of grating can be recorded simultaneously in high resolution.

Since the spectral information 'seen' by the individual pixel is recorded simultaneously, an array with $n$ pixels is equivalent to $n$ spectrometers working separately or one spectrometer working $n$ times. Simultaneous recording of $n$ pixels provides the *multichannel advantage*.

Cooling reduces thermal noise of CCD detectors, so that integration times may be long, up to days. Thus, very faint Raman lines can be recorded.

Interferometers work with single detectors. For the NIR range InGaAs or Ge semiconductor detectors are used. They have to be extremely sensitive, since the intensity of the Raman lines decreases with the fourth power of its absolute frequency (the $\nu^4$ factor). In order to reduce their thermal noise they are cooled by Peltier elements or liquid nitrogen.

## Complete Raman spectrometers

Complete Raman spectrometers are produced by several companies. Due to limited space and to the fact that the market is changing continuously, only the names of the main producers can be given here: Andor, Bio-Rad, Bruker*, Dilor, Instruments S.A, Jobin-Yvon, Kaiser Optical Systems, Nicolet*, Ocean Optics, Perkin-Elmer*, Renishaw, Sentronik, Spex. The companies marked with a * supply Raman spectrometers with excitation at 1064 nm able to record 'fluorescence-free' Raman spectra. Most companies supply Raman microscopes.

## List of symbols

$a$ = distance; $f$ = collimator focal length; $F$ = area; $G$ = optical conductance; $G_\nu$ = spectral optical conductance; $h$ = slit height; $h\nu_0$ = quantum of energy; $L$ = radiance; $L_\nu$ = spectral radiance; $N$ = number of pixels; $q$ = normal coordinate; QY = quantum yield; $2r$ = diameter of Jacquinot stop; $R_0$ = theoretical resolving power; $R_{IR}$ = resolving power (infrared range); $R_{RA}$ = resolving power (Raman range); $S$ = electric signal; $S/N$ = signal-to-noise ratio; $\Delta\nu$ = line width (bandwidth); $\Delta l$ = displacement of mirrors; $\nu_0$ = frequency; $\nu_s$ = frequency for inelastic scattering; $\rho$ = overall reflectivity; $\tau$ = transmission; $\Phi$ = radiant power; $\Phi_\nu$ = flux.

*See also:* **Biochemical Applications of Raman Spectroscopy; FT-Raman Spectroscopy Applications; Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy; IR Spectral Group Frequencies of Organic Compounds; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; Matrix Isolation Studied By IR and Raman Spectroscopies; Nonlinear Raman Spectroscopy, Applications; Nonlinear Raman Spectroscopy, Instruments; Nonlinear Raman Spectroscopy, Theory; Raman and IR Microspectroscopy; Rayleigh Scattering and Raman Spectroscopy, Theory; Surface-Enhanced Raman Scattering (SERS), Applications; Vibrational, Rotational and Raman Spectrocopy, Historical Perspective.**

## Further reading

International Union of Pure and Applied Chemistry (1998) *Compendium of Analytical Nomenclature, Definite rules 1997*. Oxford: Blackwell Science.

Jacquinot P (1954) The luminosity of spectrometer with prisms, gratings or Fabry–Perot étalons. *Journal of the Optical Society of America* **44**: 761–765.

Schrader B (ed) (1985) *Infrared and Raman Spectroscopy*. Weinheim: VCH Verlagsgesellschaft.

Schrader B (ed) (1989) *Raman/Infrared Atlas of Organic Compounds*. Weinheim: VCH Verlagsgesellschaft.

Schrader B and Moore DS (1997) Laser-based molecular spectroscopy for chemical analysis – Raman scattering processes (IUPAC Recommendations 1997). *Pure and Applied Chemistry* **69**: 1451–1468.

# Raman Spectroscopy in Biochemistry

*See* **Biochemical Applications of Raman Spectroscopy.**

# Rayleigh Scattering and Raman Effect, Theory

**David L Andrews**, University of East Anglia, Norwich, UK

Copyright © 1999 Academic Press.

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
|---|
| Theory |

Rayleigh scattering, the commonplace phenomenon which accounts for the brightness of the sky (amongst many other familiar aspects of the world we inhabit) and the Raman effect, a weaker analogue seen only at high intensities, are closely similar processes in which light is scattered by atoms or molecules. The interactions each entails differ in that the Rayleigh process is technically *elastic* whilst its Raman counterpart is *inelastic* – all of the features in which the two processes significantly differ owe their origin to that fundamental difference in the energetics. Matter responsible for Rayleigh scattering neither loses nor gains energy thereby – and so the scattered light has the same frequency as the radiation from which it is produced. However, atoms or molecules engaged in Raman scattering either gain or lose energy in the process, so that the frequency of the emergent light differs from that impinging on them – by conservation of energy, the emergent light has either a lower or higher frequency, respectively, as a result. The two types of Raman process, known as Stokes and anti-Stokes, are illustrated schematically in the energy level or ladder diagrams of **Figures 1A** and **1B**; the Stokes process results in a molecular transition to a state of higher energy, its anti-Stokes counterpart is a transition to a state of lower energy. Rayleigh scattering processes are represented by **Figures 1C** and **1D**.

A simple picture widely used for didactic purposes portrays Rayleigh scattering in terms of the electric field of impinging radiation generating, through its interaction with the electron cloud of the scattering molecule, an outgoing field that oscillates at the same frequency. The Raman process is considered to be the generation of an emergent field modulated by molecular vibrations. However, theory cast at that level is of severely limited value – it fails, for example, to address the relative magnitudes of the Stokes and anti-Stokes Raman signals; and it is not well suited to processes involving electronic transitions. To fully understand those aspects we have to look further into the theory of the fundamental interactions involved in the scattering of light.

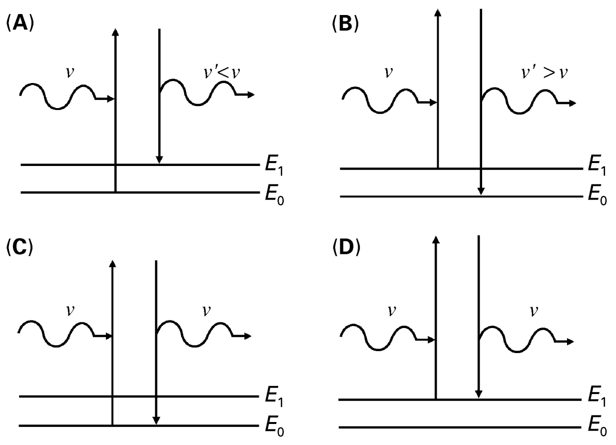



**Figure 1** Energy level diagrams illustrating Raman and Rayleigh scattering, with incoming radiation on the left, scattered radiation emergent on the right. Only energy levels directly involved are depicted: (A) Stokes Raman transition, (B) anti-Stokes Raman transition; (C) and (D) Rayleigh scattering.

## Rayleigh scattering

The blue of the sky attests to more efficient Rayleigh scattering at the higher frequency, shorter wavelength end of the visible spectrum – the familiar red sky before dusk and at dawn manifesting the loss of bluer light to skies in other parts of the world. In fact

the efficiency of scattering has a cubic dependence on the optical frequency, the scattered intensity having a fourth power dependence because the photon energy itself then enters into the equation. Although it is a moot point for Rayleigh scattering, these dependences actually relate to the frequency of the scattered rather than the incident light, a distinction which nonetheless becomes significant in the case of Raman scattering where the same power laws apply.

The mechanism for Rayleigh scattering entails the electronic polarizability $\alpha$ of the molecules of the sample. The polarizability itself is essentially a measure of how easily the molecular charge distribution is shifted through its interaction with electromagnetic radiation. At simplest, and in an isotropic system such as an atom, the polarizability is a quantity which represents the constant of proportionality between the electric field $\boldsymbol{E}$ of the radiation and the electric dipole moment it induces in the same direction, $\mu_x^{\mathrm{ind}} = \alpha \boldsymbol{E}$. A further guide to its nature can be gained from the polarizability volume, $\alpha' = \alpha/4\pi\varepsilon_0$ (where $\varepsilon_0$ is the vacuum permittivity), which casts this constant in units of volume and often yields a value similar in magnitude to the molecular volume. This correctly suggests that systems such as large atoms and aromatic molecules tend to have large polarizabilities. Nonetheless in all but the highest symmetry molecules the ease of charge displacement within the molecule varies with direction, so that in general the induced dipole moment is not parallel with the applied electric field, but slanted towards the direction of least resistance. Then the polarizability is a second rank tensor and we have;

$$\mu_i^{\mathrm{ind}} = \sum_j \alpha_{ij} E_j \qquad [1]$$

where $i$ and $j$ represent Cartesian coordinates – for example, the dipole moment induced in the $x$-direction is determined by $\mu_x^{\mathrm{ind}} = \alpha_{xx}E_x + \alpha_{xy}E_y + \alpha_{xz}E_z$.

It is the development of a fully quantum theoretical depiction of scattering at the molecular level which leads to the detailed structure of the electronic polarizability. Here, each Rayleigh (or Raman) scattering event is understood as involving the absorption of one photon of the incoming radiation, accompanied by the emission of one photon. It is important to recognize that the absorption and emission take place together in one concerted process; there is no measurable time delay between the two events. The energy–time uncertainty relation $\Delta E \Delta t \geq h/2\pi$ allows for each process to take place even when there is no energy level to match the energy of the absorbed photon, as indicated by the absence of any level at the upper end of the arrows in **Figure 1**. In other words the absorption does not populate a real intermediate state, since it is accompanied by emission. Thus it is, for example, that Rayleigh scattering of visible light takes place even in transparent media.

Despite their widespread adoption and utility, energy diagrams such as **Figure 1** are potentially misleading for any such processes involving the concerted absorption and/or emission of more than one photon – for in this case they incorrectly suggest that emission takes place subsequent to absorption. All such processes are best described with the aid of time-ordered diagrams which symbolically represent such interactions as a series of photon absorptions and emissions, and which lead to a more correct theoretical representation. Both for Rayleigh and Raman scattering there are two possible sequences, depending on whether the absorption or the emission comes first. These two cases are illustrated by the time-ordered diagrams of **Figures 2A** and **2B**, in which the vertical line represents the successive states of the molecule, and the wavy lines photons, the sequence of interactions being read upwards. Thus in both diagrams the molecular progress from an initial state $m$ to a final state $n$ proceeds via an intermediate state $r$; in (A) the transition from $m$ to $r$ is accompanied by absorption of a photon $h\nu$ from the incident beam, and the transition from $r$ to $n$ by emission of a photon $h\nu'$; in (B) this ordering of absorption and emission is reversed. In reality these processes are not separable; the diagrams simply assist development of the theory. Therefore, although the overall process is subject to energy conservation, i.e. $E_m + h\nu = E_n + h\nu'$, energy need not be conserved in the individual absorption and emission stages. For

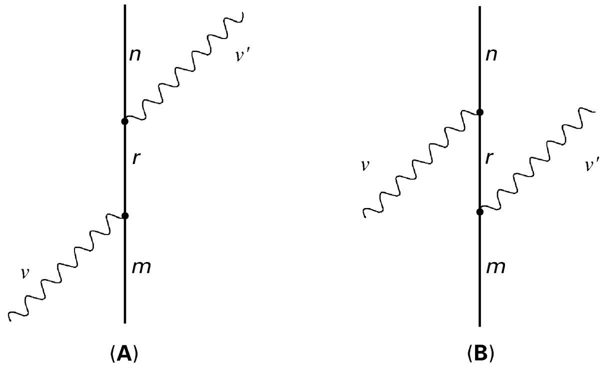



**Figure 2** Time-ordered diagrams for light scattering, with time progressing upwards. Taken together, each applies to both Rayleigh and Raman processes; in the Rayleigh case the initial state $m$ and the final state $n$ of the molecule become the same, and $\nu' = \nu$.

this reason the state $r$ is often referred to as a virtual state, and all possible energy levels must be taken into account.

Interpreting the time-ordered diagrams of **Figure 2** by the rules with which they are associated, and given that the states $m$ and $n$ can be identified with the electronic ground state with wavefunction $\varphi_0$ and energy $E_0$, the following result for the polarizability is obtained:

$$\alpha_{ij} = \sum_r \left[ \frac{\langle \varphi_0 \mid \hat{\mu}_i \mid \varphi_r \rangle \langle \varphi_r \mid \hat{\mu}_j \mid \varphi_0 \rangle}{E_r - E_0 - h\nu - ih\Gamma_r} + \frac{\langle \varphi_0 \mid \hat{\mu}_j \mid \varphi_r \rangle \langle \varphi_r \mid \hat{\mu}_i \mid \varphi_0 \rangle}{E_r - E_0 + h\nu - ih\Gamma_r} \right] \quad [2]$$

where the two complete terms on the left and right of the plus sign correspond directly to **Figures 2A** and **2B** respectively. In Equation [2], $\varphi_r$ is the wavefunction of state $r$ with energy $E_r$ and line width $\frac{1}{2}\Gamma_r$ and $\hat{\mu}_i$ is the $i$th component of the electric dipole moment operator. In frequency regions close to an optical absorption band, one of the states in the summation over $r$ will be such that $E_r - E_0 \approx h\nu$, so that the first term of Equation [2] will have a small denominator, approximating to the line width factor $-ih\Gamma_r$, and the term as a whole – see **Figure 2A** – will overwhelm all else. However, in more common off-resonant circumstances the line width factor can be neglected in each denominator. Moreover for most electronic states the Dirac brackets $\langle \varphi_0, |\hat{\mu}_i| \varphi_r \rangle$ and $\langle \varphi_r, |\hat{\mu}_i| \varphi_0 \rangle$ are identical and can be expressed more concisely as components of the transition dipole moment $\mu^{r0}$. Again for conciseness, defining $h\nu_{r0} = E_r - E_0$, Equation [2] then finally reduces to

$$\alpha_{ij} \approx \frac{2}{h} \sum_r \left( \frac{\nu_{r0}}{\nu_{r0}^2 - \nu^2} \right) \mu_i^{r0} \mu_j^{r0} \quad [3]$$

Rayleigh scattering generally produces radiation with a changed polarization state; polarized incident light is to some extent depolarized by the process whilst unpolarized light is to some extent polarized. Both effects are normally characterized by a depolarization ratio defined as the intensity ratio of plane polarized components of the scattered light. For right-angled scattering of light polarized in the $z$-direction and incoming along the $y$-direction, as shown in **Figure 3**, the depolarization ratio $\rho_l$ of light scattered in the $x$-direction is calculated as

$$\rho_l = I(z \to y)/I(z \to z) \quad [4]$$

where the subscript on the $\rho$ denotes linear polarization (synonymous with plane polarization). The value of $\rho_l$ depends on the molecules responsible for the scattering and is directly expressible in terms of polarizability parameters. Specifically, if we define the polarizability mean $\bar{\alpha}$ and anisotropy $\gamma$ through

$$\bar{\alpha} = \tfrac{1}{3} \sum_{i=x,y,z} \alpha_{ii}; \qquad \gamma^2 = -\tfrac{9}{2}\bar{\alpha}^2 + \tfrac{3}{2} \sum_{i=x,y,z} \sum_{j=x,y,z} \alpha_{ij}^2 \quad [5]$$

then for scattering by a gas or liquid we find

$$\rho_l = \frac{3\gamma^2}{4\gamma^2 + 45\bar{\alpha}^2} \quad [6]$$

with a value in the interval (0, 0.75). The lower limit corresponds to scattering with full retention of linear polarization and corresponds to $\gamma^2 = 0$, a case which occurs only for molecules of very high symmetry. Although introduced here for right-angled scattering, the above result is in fact independent of scattering angle. In contrast the extent of polarization introduced by the scattering of unpolarized light is an angle-dependent quantity. For right-angled scattering where the effect is largest, the corresponding 'depolarization ratio' is given as

$$\rho_n = I(n \to y)/I(n \to z) \quad [7]$$

(where the subscript of the $\rho$ stands for natural), giving

$$\rho_n = \frac{6\gamma^2}{7\gamma^2 + 45\bar{\alpha}^2} \quad [8]$$

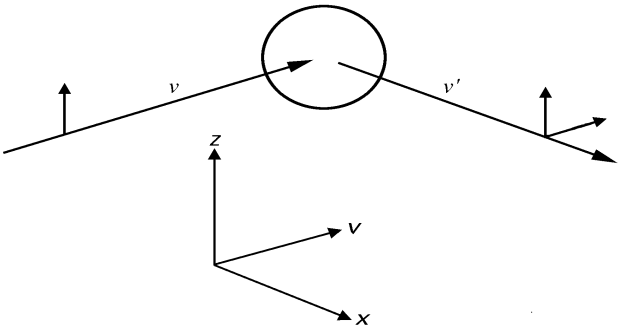



**Figure 3** Scattering geometry for the usual measurement of depolarization ratios; incident radiation is $z$-polarized and light scattered at right-angles is resolved for its $y$- and $z$-polarization components.

For scattering at other angles $\theta$ (where $\theta = 0$ relates to forward scattering) the result can be written as

$$\rho_\mathrm{n}(\theta) = \cos^2\theta + \rho_\mathrm{n} \sin^2\theta \qquad [9]$$

The angle dependence of the polarization which Rayleigh scattering confers on unpolarized light is immediately evident on viewing a clear daytime sky through polarizing spectacles.

## The Raman effect

The Raman effect was one of the first processes whose explanation, largely through the work of Placzek in 1934, exploited and vindicated the still nascent quantum theory. As for Rayleigh scattering, each scattering event involves concerted processes of photon absorption and emission, without the need for an energy level to match the absorbed photon. The difference is that as a result of this process, the scatterer undergoes an overall transition from one energy level to another, as depicted in **Figures 1A** and **1B**. The Raman effect is a very weak phenomenon; typically only one incident photon in ~$10^7$ produces a Raman transition, and observation of the effect thus requires a very intense source of light.

Raman scattering generally involves transitions amongst energy levels that are separated by much less than the photon energy of the incident light. The two levels, denoted by $E_0$ and $E_1$ in **Figure 1**, for example, are most often vibrational levels, whilst the energies of the absorbed and emitted photons are commonly in (or near to) the visible range – hence the effect provides the facility for obtaining vibrational spectra using visible light. In general the Stokes Raman transition from level $E_0$ to $E_1$ results in scattering of a frequency given by

$$\nu_\mathrm{s} = \nu - \Delta E/h \qquad [10]$$

where $\Delta E = (E_1 - E_0)$, and the corresponding anti-Stokes transition from $E_1$ to $E_0$ produces a frequency

$$\nu_\mathrm{AS} = \nu + \Delta E/h \qquad [11]$$

Thus each allowed Raman coupling generally produces two frequencies in the spectrum of scattered light, shifted to the negative and positive sides of the dominant Rayleigh line by the same amount, $\Delta\nu = \Delta E/h$. For this reason Raman spectroscopy is concerned with measurements of frequency shifts, rather than absolute frequencies. In most cases a number of Raman transitions can take place, involving various molecular energy levels, and the spectrum of scattered light contains a range of frequencies shifted away from the irradiation frequency. In the particular case of vibrational Raman transitions, the shifts can be identified with vibrational frequencies.

Although each Stokes line and its anti-Stokes counterpart are equally separated from the Rayleigh line, they are not of equal intensity. This is because the intensity of each transition is proportional to the population of the energy level from which the transition originates; under equilibrium conditions the ratio of populations is given by the Boltzmann distribution. With the fourth-power dependence on the scattering frequency, the ratio of intensities of the Stokes line and its anti-Stokes partner in a Raman spectrum is given by

$$I_\mathrm{AS}/I_\mathrm{S} = \{(\nu + \Delta\nu)/(\nu - \Delta\nu)\}^4 (g_1/g_0) \exp(-h\Delta\nu/kT) \qquad [12]$$

where $g_0$ is the degeneracy of the ground state and $g_1$ that of the upper level – and hence the anti-Stokes line is almost invariably weaker in intensity. The dependence of this ratio on the absolute temperature $T$ can, for example, be employed as a means of flame thermometry. However, since the Stokes and anti-Stokes lines give precisely the same information on molecular frequencies, it is usually only the stronger (Stokes) part of the spectrum that is recorded.

A development of detailed theory, again based on the time-ordered diagrams of **Figure 2**, establishes the dependence of Raman scattering on a transition tensor which takes on the same role as the polarizability in Rayleigh scattering. Once the usual Born–Oppenheimer separation of electronic and vibrational wavefunctions has been effected, then for the vibrational Raman transition v → v′ involving a normal mode of vibration $\lambda$, this tensor takes the form

$$\alpha_{ij}^{v'_\lambda v_\lambda} = \sum_{r,v''(\mu)} \left[ \frac{\langle \chi_0^{v'_\lambda} | \langle \varphi_0 | \hat{\mu}_i | \varphi_r \rangle | \chi_r^{v''_\mu} \rangle \langle \chi_r^{v''_\mu} | \langle \varphi_r | \hat{\mu}_j | \varphi_0 \rangle | \chi_0^{v_\lambda} \rangle}{E_r^{v''_\mu} - E_0^{v_\lambda} - h\nu - ih\Gamma_r^{v''_\mu}} \right. \qquad [13]$$

$$\left. + \frac{\langle \chi_0^{v_\lambda} | \langle \varphi_0 | \hat{\mu}_j | \varphi_r \rangle | \chi_r^{v''_\mu} \rangle \langle \chi_r^{v''_\mu} | \langle \varphi_r | \hat{\mu}_i | \varphi_0 \rangle | \chi_0^{v_\lambda} \rangle}{E_r^{v''_\mu} - E_0^{v_\lambda} + h\nu - ih\Gamma_r^{v''_\mu}} \right]$$

where, for example, the vibrational wavefunction $\chi_r^{v''_\mu}$ denotes a state with a quantum number $v''$ in the vibrational mode $\mu$, within the set of levels associated with electronic state $r$. Here also $E_r^{v''_\mu}$ and $\Gamma_r^{v''_\mu}$ relate to the total (electronic plus vibrational) energy and the damping, respectively, of that state. Away

from resonance, in other words when using frequencies $\nu$ well removed from any optical absorption bands, then the vibrational energy contributions in each denominator term of Equation [13] can safely be neglected. Then, using the completeness relation of quantum mechanics, the $|\chi_r^{v_\mu''}\rangle\langle\chi_r^{v_\mu''}|$ sum can be effected to give

$$\alpha^{v_\lambda' v_\lambda} = \sum_r \left[ \frac{\langle\chi_0^{v_\lambda'}|\langle\varphi_0|\hat{\mu}_i|\varphi_r\rangle\langle\varphi_r|\hat{\mu}_j|\varphi_0\rangle|\chi_0^{v_\lambda}\rangle}{E_r - E_0 - h\nu - ih\Gamma_r} + \frac{\langle\chi_0^{v_\lambda'}|\langle\varphi_0|\hat{\mu}_j|\varphi_r\rangle\langle\varphi_r|\hat{\mu}_i|\varphi_0\rangle|\chi_0^{v_\lambda}\rangle}{E_r - E_0 - h\nu - ih\Gamma_r} \right] = \langle\chi_0^{v_\lambda'}|\alpha_{ij}(Q_\lambda)|\chi_0^{v_\lambda}\rangle \quad [14]$$

using Equation [2]. The result thus involves the dependence of the electronic polarizability on the nuclear coordinate $Q_\lambda$ relating to the excited vibration. Although all molecules have a finite polarizability, that is not the case for $\alpha_{ij}^{v_\lambda' v_\lambda}$ – but no Raman signal will emerge when the latter is zero. Here a powerful symmetry rule emerges: any Raman-active vibration must transform under an irreducible representation spanned by components of the polarizability tensor (transforming as the quadratic variables $x^2$, $xy$, etc., or one of their combinations). Some of the broad implications of this are highlighted below.

To obtain the major selection rules for Raman scattering we can first expand Equation [14] in a Taylor series about the equilibrium configuration $Q_e$;

$$\alpha_{ij}(Q_\lambda) = \alpha_{ij}(Q_e) + \left.\frac{\partial\alpha_{ij}}{\partial Q_\lambda}\right|_{Q_e}(Q_\lambda - Q_e) + \cdots \quad [15]$$

and hence we have

$$\alpha_{ij}^{v_\lambda' v_\lambda} = \alpha_{ij}(Q_e)\delta_{v'v} + \left.\frac{\partial\alpha_{ij}}{\partial Q_\lambda}\right|_{Q_e}\langle\chi_0^{v_\lambda'}|(Q_\lambda - Q_e)|\chi_0^{v_\lambda}\rangle + \cdots \quad [16]$$

The first term on the right is non-zero only when the initial and final states are identical – which relates back to Rayleigh scattering. It is the second term which is significant for the Raman process and its detailed form establishes two rules governing Raman-allowed transitions, since both of its factors must then be non-zero. For the Dirac bracket to be non-zero dictates $v' = v \pm 1$, as in infrared absorption spectroscopy. For the polarizability derivative to be non-zero, the polarizability must change during the vibration, as the molecule passes through its equilibrium configuration. This is the key selection rule for the Raman effect, illustrated for $CO_2$ in **Figure 4**.

It is immediately apparent that Raman transitions are governed by different selection rules from absorption or fluorescence. The case of $CO_2$ illustrates a general principle applicable to all centrosymmetric molecules, which is that only gerade vibrations (those which are even with respect to inversion symmetry) appear in the Raman spectrum, whilst only ungerade vibrations (odd with respect to inversion) show up in infrared absorption. This illustrates the so-called mutual exclusion rule for centrosymmetric molecules, which states that vibrations active in the infrared spectrum are inactive in the Raman, and vice versa. Even for complex polyatomic molecules lacking much symmetry, the intensities of lines resulting from the same vibrational transition may be very different in the two types of spectrum, so that in general there is a useful complementarity between the two methods. Generally it is the vibrations of the most polarizable groups which are strongest in the Raman spectrum, those of the most polar groups being strongest in the infrared, as nicely illustrated in the spectra of the drug acetaminophen (UK paracetamol; *p*-hydroxyacetanilide) shown in **Figure 5**.

Further information on the symmetry properties of Raman-active molecular vibrations can be obtained by measurement of the depolarization ratios of the lines in the Raman spectrum – see **Figure 3** and Equation [4]. Interpretation of the results here invokes Equation [17]:

$$\rho_l = \frac{3\gamma'^2}{4\gamma'^2 + 45\bar{\alpha}'^2} \quad [17]$$

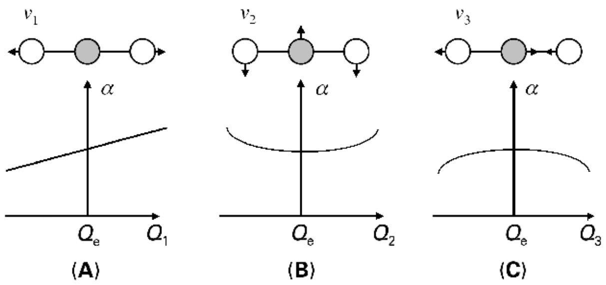
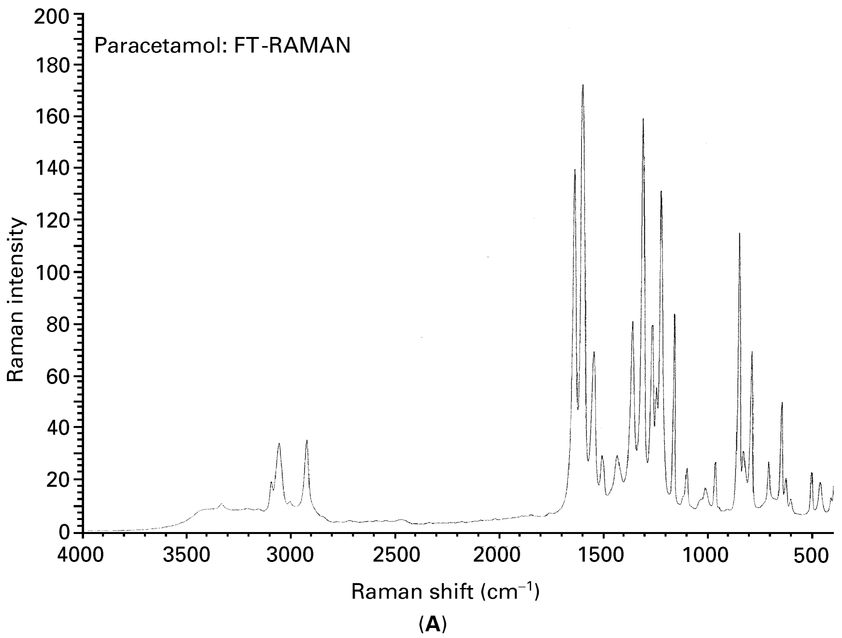



**Figure 4** Variation of polarizability in the course of three normal modes of vibration of carbon dioxide: (A) symmetric stretch, (B) bending mode and (C) antisymmetric stretch. The slope $\partial\alpha/\partial Q_\lambda$ on crossing the vertical axis is non-zero only for the symmetric stretch, and hence only this vibration gives a Raman signal.



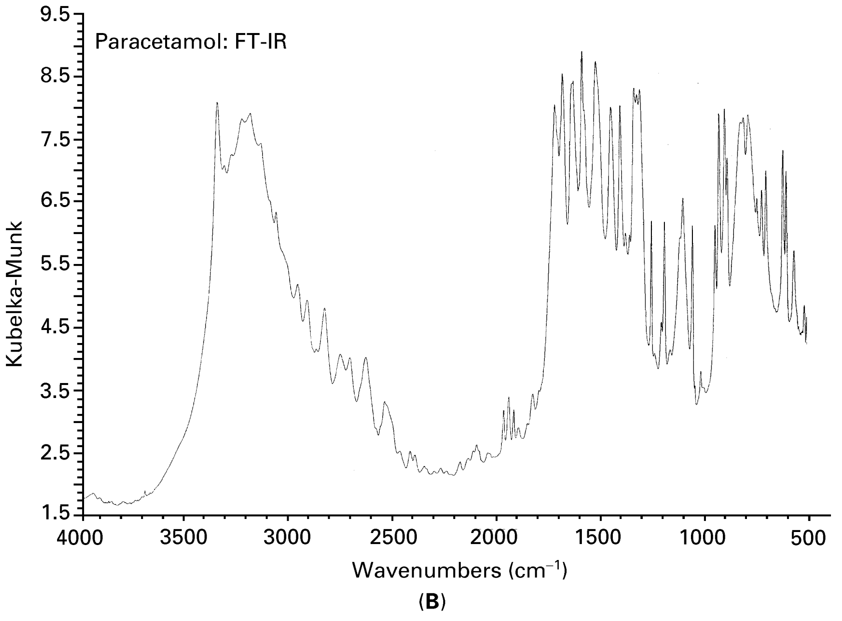




**Figure 5** Fourier-transform spectra of paracetamol: (A) Raman, and (B) infrared. Stretch vibrations of the non-polar C–H groups, close to 3000 $cm^{-1}$, show up well in the Raman spectrum. In the infrared spectrum this whole region is dominated by stretching vibrations of the highly polar O–H and N–H groups, much broadened through association with hydrogen bonding. Reproduced with permission of Nicolet Instruments.

The prime on the mean and anisotropy parameters, $\overline{\alpha}'$ and $\gamma'$, respectively, denote values obtained from the polarizability derivative – defined in the sense of Equation [5], but in components of the tensor $\partial\alpha_{ij}/\partial Q_\lambda$ rather than $\alpha_{ij}$ itself. In the case of gases and liquids, $\rho_l$ is lower than $\frac{3}{4}$ for vibrations that are totally symmetric (vibrations transforming under the totally symmetric representation of the molecular point group), but exactly for $\frac{3}{4}$ for other vibrations that lower the molecular symmetry, since $\overline{\alpha}'$ is then zero.

Although the frequency of radiation used for the study of Raman scattering is generally well removed from any absorption band of the sample, to forestall problems associated with absorption and subsequent fluorescence, special features become apparent on irradiation at a frequency close to a broad and intense optical absorption band. Quite simply, the closer the approach, the greater is the intensity of the Raman spectrum. Spectra obtained under such conditions are known as resonance Raman spectra. In the case

of large polyatomic molecules where any electronic absorption band may be due to localized absorption in a particular chromophore, the vibrational Raman lines which experience the greatest amplification are those of the appropriate symmetry involving vibrations of nuclei close to the groups responsible for the resonance.

Equation [13] correctly represents the Raman tensor even under resonance or pre-resonance conditions – and the resonance enhancement is clearly attributable to the fact that if there is an excited state for which $E_r^{v''_\mu} - E_0^{v_\lambda}$ is close to $h\nu$, the first term of that equation has a denominator of greatly diminshed magnitude. However, the subsequent development of theory leading to Equation [16] is no longer valid under such conditions; for example, the $v' = v \pm 1$ selection rule breaks down and overtones commonly appear. Other vibrations (those which transform like the rotations $R_x$, $R_y$ and $R_z$) can also become active through changed selection rules, associated with the fact that the Raman tensor is no longer real and index-symmetric, but complex and non-symmetric. As a result the equations for the Raman depolarization ratio also require modification to the following form

$$\rho_l^{\mathrm{res}} = \frac{3\gamma'^2 + 5\delta'^2}{4\gamma'^2 + 45\overline{\alpha}'^2} \qquad [18]$$

where

$$\delta'^2 = \tfrac{1}{4} \sum_{i=x,y,z} \sum_{j=x,y,z} \left|\alpha'_{ij} - \alpha'_{ji}\right|^2 \qquad [19]$$

is a measure of the degree of antisymmetry in the Raman tensor. One consequence of including this factor in Equation [18] is the possibility of depolarization ratios exceeding the normal upper bound of $\frac{3}{4}$ – in some cases indeed yielding an infinite result (complete depolarization).

Generally, the use of circularly polarized light in studies of Rayleigh or Raman scattering offers no further information beyond that provided by plane polarizations. However, optically active (chiral) compounds in the liquid or solution state respond differentially to circularly polarized light, according to its handedness, making it possible to obtain a spectrum showing a marginal difference in the Raman intensity $I^{\mathrm{R}}-I^{\mathrm{L}}$ as a function of scattering frequency. The extent of this differential for each molecular vibration is directly related to the detailed stereochemical structure responsible for the manifestation of chirality. In particular, the extent of differential scattering in a region of the spectrum associated with a particular group frequency can be interpreted in terms of the chiral environment of the corresponding functional group. In contrast to the theory developed here, the scattering entails not only electric dipole but also the much weaker magnetic dipole and electric quadrupole interactions.

At the high intensities now available from laser sources, numerous other variants of the Raman effect can be observed, many associated with optically nonlinear behaviour. For analytical purposes the most important of these come under the heading of coherent Raman spectroscopy, of which the process known as CARS (coherent anti-Stokes Raman spectroscopy) is the most common. Here, two beams are directed into the sample: one has a fixed frequency playing the role of $\nu$ and the other, a frequency $\nu'$ tunable across the Stokes range. As $\nu'$ tunes into each Stokes frequency $\nu_{\mathrm{S}}$ a four-photon process occurs, essentially combining all the elements of **Figures 1A** and **1B**, and generating coherent emission at the corresponding anti-Stokes frequency $\nu_{\mathrm{AS}}$. The laser-like nature of this output facilitates its collection for spectroscopic analysis, and permits the analysis of microscopic samples.

## List of symbols

$\boldsymbol{E}$ = electric field; $E_m$ = energy of level $m$; $g_0$ = degeneracy of ground state; $g_1$ = degeneracy of upper level; $h$ = Planck's constant; $I$ = intensity; $k$ = Boltzmann's constant; $Q$ = Nuclear coordinate; $t$ = time; $T$ = absolute temperature; $\boldsymbol{\alpha}$ = electronic polarizability; $\overline{\alpha}$ = mean polarizability; $\gamma$ = polarizability anisotropy; $\Gamma_r$ = damping of level $r$; $\varepsilon_0$ = vacuum permittivity; $\theta$ = scattering angle; $\lambda$ = normal mode of vibration; $\mu^{\mathrm{ind}}$ = induced electric dipole moment; $\mu^{r0}$ = transition dipole moment; $\nu$ = frequency of incident radiation; $\nu'$ = frequency of emitted radiation; $\nu_{\mathrm{S}}$ = Stokes frequency; $\nu_{\mathrm{AS}}$ = anti-Stokes frequency; $\rho$ = depolarization ratio; $\phi_0$ = wavefunction with energy $E_0$; $\chi$ = vibrational wavefunction.

*See also:* **Biochemical Applications of Raman Spectroscopy; Nonlinear Optical Properties; Raman Optical Activity, Applications; Raman Optical Activity, Spectrometers; Raman Optical Activity, Theory; Raman Spectrometers.**

## Further reading

Andrews DL (1997) *Lasers in Chemistry*, 3rd edn, pp 128–149. Berlin: Springer-Verlag.

Barron LD (1982) *Molecular Light Scattering and Optical Activity*. Cambridge: Cambridge University Press.
Craig DP and Thirunamachandran T (1984) *Molecular Quantum Electrodynamics*. London: Academic Press.
Long DA (1977) *Raman Spectroscopy*. New York: McGraw-Hill.
Placzek G (1934) Rayleigh-Streuung und Raman-Effekt. In: Marx E (ed) *Handbuch der Radiologie*, Vol. 6, Part 2, pp 205–374. Leipzig: Akademische Verlag.
Raman CV and Krishnan KS (1928) A new type of secondary radiation. *Nature* **121**: 501.
Sheppard N (1990) Chemical applications of molecular spectroscopy – A developing perspective. In: Andrews DL (ed) *Perspectives in Modern Chemical Spectroscopy*, pp 1–41. Berlin: Springer-Verlag.

# Regulatory Authority Requirements

*See* **Calibration and Reference Systems (Regulatory Authorities).**

# Relaxometers

**Ralf-Oliver Seitter** and **Rainer Kimmich**,
Universität Ulm, Germany

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Methods & Instrumentation**

## Purpose and classification of NMR relaxometers

Nuclear magnetic relaxation, that is, thermal equilibration of the spin systems with respect to longitudinal or transverse magnetization components, multiple-quantum spin coherences and longitudinal dipolar, quadrupolar or scalar order, comprises a vast variety of different experimental protocols and phenomena. In a typical relaxation experiment, one first produces a nonequilibrium population of the spin states, often combined with spin coherences. It is then a matter of the fluctuations of the spin couplings to induce spin transitions towards thermal equilibrium. 'Equilibrium' means (i) populations following Boltzmann's distribution, and (ii) completely vanishing spin coherences. Consequently there are three elements inherent in a typical relaxation experiment: *Preparation* of a nonequilibrium state of the spin systems; the variable *evolution* interval allowing for the induction of spin transitions; and the *detection* of the populations, longitudinal order, or coherences after the evolution interval. The time constants with which the 'observable' approaches equilibrium during the evolution interval are the relaxation times, such as the spin–lattice relaxation time $T_1$, the transverse relaxation time $T_2$, the rotating-frame relaxation time $T_{1\rho}$, the dipolar-order relaxation time $T_d$, and so on. In the following we will focus on $T_1$ in particular.

The spin–lattice relaxation rate of dipolar coupled homonuclear two-spin $I$ systems, for instance, is given by

$$\frac{1}{T_1} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{3}{2}\gamma^4\hbar^2 I(I+1)\left[J^{(1)}(\omega) + J^{(2)}(2\omega)\right] \quad [1]$$

where $\mu_0$ is the magnetic field constant, $\gamma$ is the gyromagnetic ratio, and $J^{(i)}(\omega)$ is the intensity function of the Larmor frequency, $\omega = \gamma B_0$, depending on the flux density of the external magnetic field, $B_0$. The intensity function is given as the Fourier transform of the dipolar autocorrelation function $G_i(\tau)$,

$$J^{(i)}(\omega) = \int_{-\infty}^{\infty} G_i(\tau)\exp(-\mathrm{i}\omega\tau)\mathrm{d}\tau \quad [2]$$

The dipolar autocorrelation function is defined by

$$G_i(\tau) = \left\langle F^{(i)}(0)F^{(-i)}(\tau)\right\rangle \quad [3]$$

where

$$F^{(1)} = \frac{1}{r^3}\sin\vartheta\cos\vartheta\exp(-i\varphi),$$
$$F^{(2)} = \frac{1}{r^3}\sin^2\vartheta\exp(-2i\varphi) \quad [4]$$

The polar coordinates $r$, $\vartheta$, $\varphi$ define the internuclear vector of the spin system. That is, any molecular motion affecting these coordinates leads to fluctuations of the functions $F^{(i)}$, so that the spin–lattice relaxation rate (Eqn [1]) directly reflects these motions via the intensity and autocorrelation functions. The prominent goal of NMR relaxometry, hence, is to monitor the features of molecular dynamics. This is best done by recording the frequency dependence of spin–lattice relaxation.

Variation of the Larmor frequency means variation of the external flux density $B_0$. The range within which one can do that using conventional NMR spectrometers is very limited. The reason is that the flux density has a fixed value given by the magnet in use. Typical values range from 1 to 20 T. The radio frequency (RF) part of conventional NMR spectrometers is tuned to resonance in the particular flux density provided by the magnet. That is, there is no reasonable way to study frequency dependences using spectrometers with fixed flux densities, because the signal-to-noise ratio, $S/N \propto B_0^{3/2}$, and the RF bandwidth deteriorates the lower the flux density and the frequency become.

The solution of the problem is field-cycling NMR relaxometry. The essence of this technique is that the magnetic flux density relevant for relaxation is different from that during signal detection. The 'relaxation field' may be varied over several orders of magnitude while the 'detection field' is kept fixed at the highest possible value. That is, the RF console is tuned to this particular detection field. The design of a field-cycling relaxometer thus implies the possibility to switch the flux density rapidly and precisely between different levels with very good stability. A typical field cycle is shown in **Figure 1**.

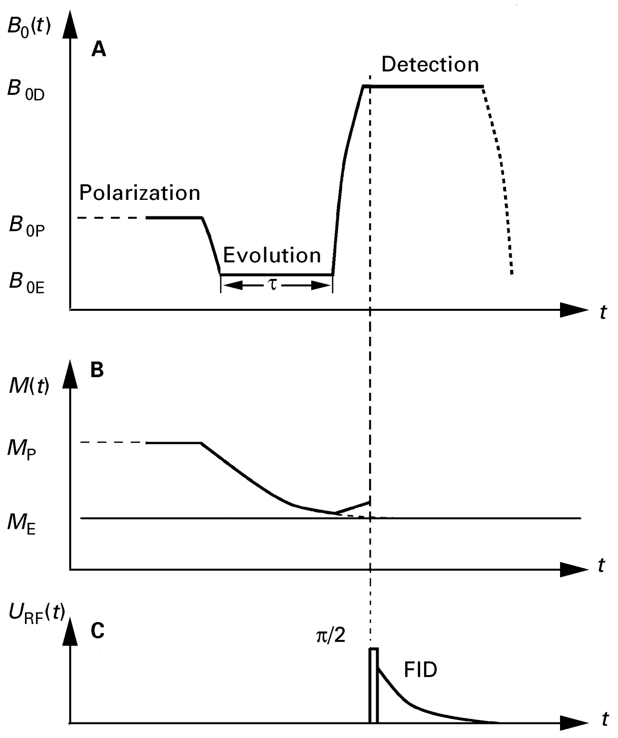



**Figure 1** Typical cycle of a field-cycling NMR relaxometer serving for spin–lattice relaxation experiments: (A) External magnetic flux density $B_0 = B_0(t)$, (B) longitudinal magnetization $M = M(t)$ in the sample, (C) RF pulse with free-induction decay (FID) during the detection period. Note that the field level in the (short) detection period is much higher than in the (long) polarization interval. This ensures equivalent heat production in both intervals.

## Field-cycling (FC) NMR relaxometry

### Principle

A typical field cycle consists of a preparation interval, an evolution interval and a detection interval. In the preparation interval the magnetic flux density, $B_P$, is chosen as high as technically possible. The purpose is to polarize the sample so that the magnetization becomes as large as feasible.

A nonequlibrium state is then produced by switching the magnetic field to another level, namely to that of the evolution interval, $B_E \neq B_P$. If the magnetic field is switched fast enough, so that the magetization cannot follow the instantaneous equilibrium value given by Curie's law, we have a nonequilibrium situation at the beginning of the evolution interval. That is, the magnetization starts to relax towards the equilibrium value corresponding to the field level of the evolution interval.

The relaxation curve is then probed point-by-point by switching the field up to the detection field level and by acquiring the NMR signal using a RF pulse or an echo RF pulse sequence. The detection field is chosen as high as possible, and is the same for all experiments. This is the field to which the RF console is tuned. The amplitude of the NMR signal represents the magnetization at the end of the evolution interval, in which relaxation takes place. Thus

varying the evolution interval, permits one to record the whole relaxation curve to which the spin–lattice relaxation time can be fitted. Assuming mono-exponential relaxation and neglecting losses during the switching intervals, the detected magnetization is given by

$$M(\tau) = M_{\mathrm{E}} + [M_{\mathrm{P}} - M_{\mathrm{E}}]\exp[-\tau/T_1(B_{\mathrm{E}})] \quad [5]$$

where $M_E$ and $M_P$ are the Curie magnetizations corresponding to the magnetic flux density in the evolution and polarization intervals, respectively. Incrementing the evolution field level, $B_E$, in subsequent experiments permits one to measure such relaxation curves at field levels not accessible to conventional NMR. That is, the whole frequency dispersion of spin–lattice relaxation can be probed with a relaxometer tuned to a fixed resonance frequency. This implies that the signal-to-noise ratio remains the same for all Larmor frequencies.

#### Requirements and limitations

**Signal-to-noise ratio** According to Curie's law, the magnetization reached in the polarization field, $B_P$, is $M_P \propto B_P$. The induction signal increases proportionally to the detection field, $B_D$, whereas noise is proportional to the square root of the carrier frequency, i.e. the square root of the detection field. The signal-to-noise ratio is thus determined by

$$S/N \propto B_{\mathrm{D}}^{1/2} M(\tau) \quad [6]$$

The magnetization at the end of the evolution interval obeys

$$M(\tau) \geq B_{\mathrm{P}}\exp[-\tau/T_1(B_{\mathrm{E}})] \quad [7]$$

so that the signal-to-noise ratio is limited by

$$S/N \propto B_{\mathrm{D}}^{1/2} B_{\mathrm{P}}\exp[-\tau/T_1(B_{\mathrm{E}})] \quad [8]$$

For optimum signal-to-noise ratio, this suggests field levels $B_D$ and $B_P$ as high as technically feasible.

**Feasibility of phase-sensitive detection** An insufficient signal-to-noise ratio, as is likely with deuteron relaxation or with very dilute systems and small samples, can be improved with signal accumulation. This, however, requires a detection field stability permitting phase-sensitive detection. That is, after having switched the field up to the detection level, resonance must be reproducibly reached in all cycles of the experiment with an accuracy of $10^{-5}$ before signal acquisition. This is a rather demanding condition requiring a technical solution like the battery driven power supply described below. By contrast, the accuracy of the polarization and evolution levels is much less critical. A few percent may be acceptable.

**Thermal stability** The *polarization field*, $B_P$, is applied as long as needed for reaching thermal equilibrium, typically a period $t_P \approx 5 \times T_1(B_P)$. The *detection field*, $B_D$, is applied only in the very short interval $t_D$ needed for acquiring the free-induction decay. That is,

$$t_{\mathrm{P}} \gg t_{\mathrm{D}} \quad [9]$$

Apart from a few superconducting versions reported in the literature, the magnet coils of field-cycling relaxometers are resistive, so that Joule's heat produced in the course of a field cycle affects the thermal stability of the system. If the magnet coil gets unduly warm during the polarization interval, thermal drifts of the field levels are unavoidable. Actually, the cooling properties of the magnet coil form a crucial factor limiting the applications. Considering the inequality given in Equation [9] it becomes obvious that it is more favourable to restrict the polarization field to a level still compatible with thermal stability, whereas the detection field should be as high as the magnet power supply permits. This is the reason why the field levels in the scheme shown in **Figure 1** are chosen in such a way that $B_D \gg B_P$.

**Accuracy of the evolution field level** On a logarithmic field or frequency scale covering several orders of magnitude, an accuracy of better than 10% may be adequate. However, at low magnetic fields, relaxation times less than a millisecond may occur. That is, a stability of the evolution field level of less than 10% must be reached in a ring-down time of less than 1 ms after switching. The evolution field level must begin with a sharp edge in the time dependence of the magnetic flux density. This is another demanding specification which requires special current control measures (see the design described below). It is also an essential condition to check the switch interval from the polarization level to that of the evolution field with the aid of a suitable Hall probe and a fast high-resolution (e.g. 500 kHz/16 bit) analogue-to-digital converter (**Figure 2**).

**Influence of the switching times** Provided that the time dependence of the magnetic flux density

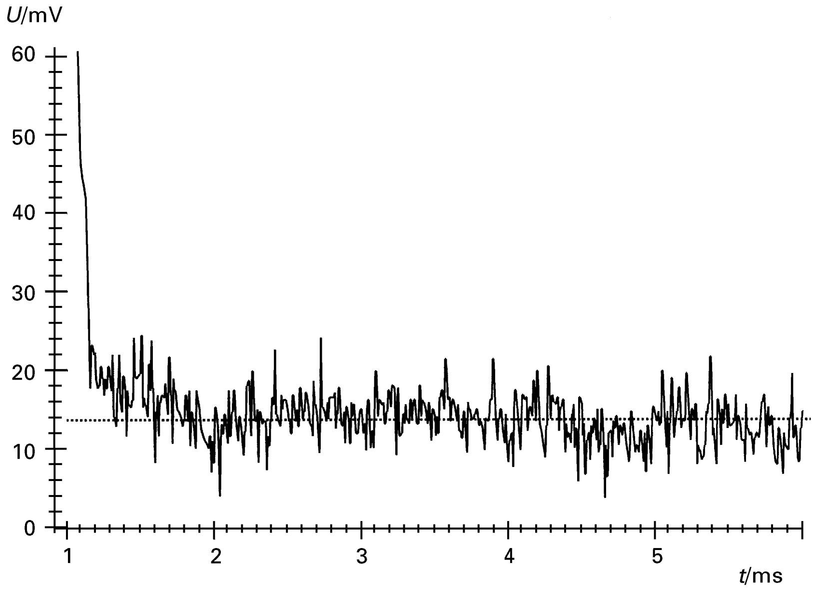



**Figure 2** Precision of the initial stabilization of the evolution field. The plot represents the voltage transient recorded with a Hall probe (6 μT/mV) and a 16 bit analogue-to-digital converter. The field level corresponds to a proton Larmor frequency of 3.5 kHz. The field-cycling relaxometer is described below. This plot demonstrates that the evolution field can be stabilized within ±1 digital unit in a ring-down time less than 1 ms.

(see **Figure 1**) is reproducible in an experiment for a given evolution flux density, there is no influence on the measured spin–lattice relaxation time even when the switching intervals are comparable with the low-field relaxation times. However, the dynamic range of the magnetization variation in a relaxation experiment is reached only in full if the switching intervals are much shorter than the low-field relaxation times. That is, the best measuring accuracy is achieved if the relaxation losses during the switching intervals are negligible.

**Lower limit of the evolution field** When the evolution field approaches the earth field, i.e. the field when the magnet is switched off, a compensation of this zero-field with the aid of correction coils becomes necessary. In this way, flux densities less than $5 \times 10^{-5}$ T can reliably be reached. Another limit is due to the local field within the sample. The local field arises because of secular spin interactions. It can therefore not be compensated by external correction coils. If the local field exceeds the external evolution field, two situations can arise. Firstly, the local field is approached by adiabatically switching off the polarization field. That is, the local magnetization vectors follow the instantaneous field direction which is finally given by the local field. In this case, dipolar or, in the case of quadrupole nuclei, quadrupolar order is produced. The relaxation time measured under such circumstances is the dipolar-order spin–lattice relaxation time. In the opposite case, the polarization field is switched off nonadiabatically so that the magnetization partly becomes transverse to the local field. This is the situation of the so-called zero-field NMR spectroscopy probing the evolution of spin coherences in the local field.

**Field homogeneity** Fast field switching stipulates small magnet coils so that the field energy to be varied in a cyclic way is low. On the other hand, the field homogeneity of small magnets tends to be poor. For field-cycling relaxometry the condition is that the homogeneity within the sample should correspond to the stability (or reproducibility) of the detection field. That is, if the detection field can be reproduced with an accuracy of $10^{-5}$, the field homogeneity in the sample need not to be better than $10^{-5}$ using suitable magnet designs this can be achieved easily.

## Crucial specifications of field-cycling NMR relaxometers

From the above outline of the factors limiting field-cycling NMR relaxometry, it becomes obvious that the following specifications of a relaxometer are most crucial. (a) The polarization and detection fields should each be as high as compatible with the thermal stability of the magnet coil. This normally means that the detection field is much higher than the polarization field. (b) The evolution field must achieve stability within a few percent after a stabilization time of less than the shortest relaxation time

to be measured. The total field switching intervals should preferably be shorter than the typical relaxation times of interest. (c) The detection field should be reproducible within $10^{-5}$ after a stabilization period of much less than the high-field relaxation time. This is the condition permitting phase-sensitive detection and, hence, automatic signal accumulation. (d) The field homogeneity in the sample should correspond to the reproducibility of the detection field, i.e. a relative field variation of less than about $10^{-5}$.

## Typical setup of a field-cycling relaxometer

Historically the first field-cycling instruments consisted of two magnets of varying flux densities. The sample was 'shuttled' between a resonant magnet and a magnet producing the variable evolution field. With modern relaxometers the field is varied much faster by controlling the current through the magnet coil electronically. This, of course, stipulates that the use of iron magnets is excluded. The magnets in use are resistive or even superconducting magnet coils with solenoid-like geometries.

The magnet current can be controlled with the aid of gate-turn-on thyristors (GTOs), metal-oxide semiconductor field-effect transistors (MOSFETs) or insulated gate bipolar transistors (IGBTs). The latter are most powerful with respect to maximum current and breakthrough voltage. A corresponding field-cycling setup is described in the following block diagram **Figure 3**.

### Field-cycling magnet

In order to facilitate fast field transitions, the total field energy, which is to be changed in a field cycle, should be as small as possible. That is, the magnet should be very compact. It is also desirable to restrict the maximum induction voltages occurring in a field cycle. Therefore, the inductance should be small.

The magnet coil may be cut out of a solid copper or aluminium block. In this way, a current density distribution for optimum homogeneity can be provided. A less demanding but nevertheless operational construction is to wind the magnet from ordinary copper wires. The magnet coil of the system described here, for instance, is composed of six double-winding layer split solenoids, so that refrigerated cooling oil can be pressed through the 1 mm spacings between the winding layers. The copper wire used has a cross section of $2.34 \times 1.35$ mm$^2$. The four inner layers consist of $2 \times 21$ windings; the two outer layers with $2 \times 18$ windings have a 10 mm gap in the middle. The inductance is 3.2 mH; the room temperature ohmic resistance is 0.46 Ω. The room temperature bore is 28 mm and the outer diameter is 70 mm. The axial length is 98 mm. A cross section is shown in **Figure 4**. The magnetic field in the sample

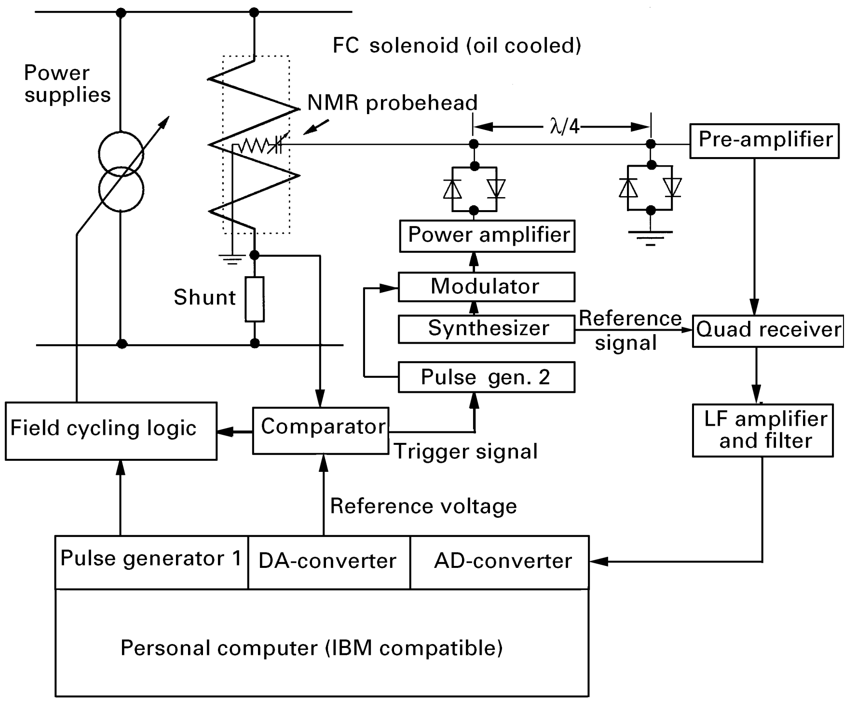



**Figure 3** Block diagram of a typical field-cycling NMR relaxometer.

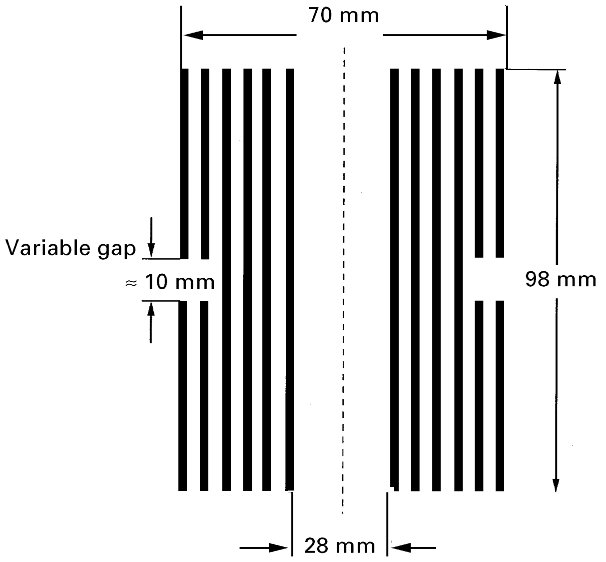



**Figure 4** Cross section of an easy-to-make and efficiently cooled field-cycling magnet. The solid vertical lines represent the six double-layers of windings of the magnet. The cooling medium is pressed through between these double layers.

volume can be homogenized by adjusting the position of the two outermost layers correspondingly. The relative inhomogeneity of the flux density is much less than $\Delta B/B = 1 \times 10^{-4}$. The current/field ratio amounts to 188.73 A/T.

## Digital control

The pulse sequence for the field cycle is generated with a personal computer equipped with a SMIS MR3020 pulse programmer board. The actual current through the magnet coil is sensed by a temperature compensated precision shunt. This current is compared with the set point given by the programmer board. The comparator unit needs an analogue reference voltage which is generated with the aid of a digital-to-analogue conversion board (Keithley Metrabyte DDA012) in the computer. For example, when the detection field level is reached, the comparator triggers the RF-pulse so that the signal acquisition is always performed at the same magnetic flux density within the accuracy of the voltage comparison (ca. $10^{-5}$). The free-induction decay recorded in phase quadrature in order to permit phase-sensitive detection is then digitized using an 8 bit analogue-to-digital converter board in the computer.

## Magnet power supply system

The magnet power supply is the most crucial part of a field-cycling relaxometer. It is a complicated system of commercial laboratory power supplies and certain measures for switching between different power supplies are optimized for different current ranges, and for connecting driver high-voltage capacitors to the magnet current circuit. The polarization and evolution fields above 4.2 mT are produced by a parallel combination of four modules of Kepco ATE 75-15M. The maximum current in these intervals is 60 A. The resolution of evolution fields below 4.2 mT is improved by using a special low-current power supply (< 800 mA) temporarily replacing the Kepco system.

The detection field is most demanding with respect to strength, stability and reproducibility. A series of 13 ordinary 12 V car batteries turned out to be best in these regards. In this way an extremely stable and reproducible current of about 300 A can be generated easily.

The transitions from high to low and from low to high magnetic fields are accelerated with the aid of a capacitor precharged to a voltage of 600 V. The polarity by which this driver capacitor is connected to the magnet current circuit depends on the direction of the transition.

For security, the system is surveyed by a unit checking the most important instrument parameters such as the cooling-oil temperature and pressure, and the storage capacitor voltage. In case the parameters deviate from the set point adjustment, the complete high-power part of the relaxometer is switched off within milliseconds.

In passive operation the time constant of the magnet coil is given by

$$\tau_{\mathrm{m}} = \frac{L}{R} \qquad [10]$$

where $L$ and $R$ are the inductance and the resistance of the coil, respectively. Employing semiconductor power switches as mentioned before permits one to actively control the field transition intervals. The following measures are in use.

- Reducing the flux density can be accelerated by dissipating the field energy in an ohmic damping resistor, $R_{\mathrm{d}}$, connected to the magnet in series during the switching-down process. The magnet current then obeys the differential equation

$$-L\frac{\mathrm{d}I}{\mathrm{d}t} = I(R + R_{\mathrm{d}}) - U_0 \qquad [11]$$

where $U_0$ is the voltage induced by the current variation in the magnet coil. With the initial condition

$I(0) = I_{max} = U_0/R$ the solution becomes

$$I_d(t) = \frac{U_0}{R + R_d} + \left[\frac{U_0}{R} - \frac{U_0}{R + R_d}\right] \exp(-t/\tau_d) \quad [12]$$

with the new time constant

$$\tau_d = \frac{L}{R + R_d} \quad [13]$$

- The damping resistors can be efficiently supplemented by voltage suppressor diodes which cut off all voltages above a certain nominal value. The field energy is then dissipated in the diodes as well.
- The third possibility is to transfer the field energy temporarily to a capacitor. With opposite polarity, such a capacitor can be used to drive the field up as mentioned before. The principle is illustrated in **Figure 5**. The switches are single IGBT modules (Semikron SKM 400 GA) with maximum currents of 400 A at 1200 V. The switching rate can be changed by altering the gate resistor $R_G$. Induction-voltage peaks due to fast switching can be reduced by increasing $R_G$. The IGBT switches are controlled by a Semikron SKHI 20 module, which galvanically separates the input TTL signal from the gate. Additionally it contains a small surveillance unit for the IGBT (overvoltage and overtemperature protection), which feeds an error signal back to the Control Logic so that the cycle is immediately stopped in the fault case. During the *polarization period* the switch $S_1$ (IGBT 1) is closed, and $G_1$ drives the magnet current. Simultaneously the capacitor $C$ is precharged to about 600 V. In the switching interval *polarization field → evolution field*, the digital-to-analogue converter defines a new set value for $G_1$. The switch $S_1$ opens, and the network operates as a series resonance circuit. That is, the magnet energy is transferred to the capacitor. After the desired evolution field level has been reached, the switch $S_1$ closes again. In the *evolution interval*, either $G_1$ or $G_2$ controls the magnet current depending on the adjusted field level. The two alternate current amplifiers are used in order to improve the field resolution, limited by that of the 12 bit digital-to-analogue converter. In the switching interval *evolution field → detection field*, the switch $S_2$ is closed, and the precharged capacitor provides the peak power for driving the current up to the detection field level. $S_3$ then closes after a 1 ms delay. When the level of the detection field is reached, $S_2$ opens and the batteries supply a current of 300 A required for the detection field. Finally, in the *switching-off interval*, the switch $S_1$ opens, and the magnet energy is transferred to the capacitor again. When the current has decayed to zero, $S_1$ is closed. Before starting a new polarization/evolution/detection field cycle, a settling time of 2 to 3 s is allowed.

## The probehead

As the bore of field-cycling magnets are kept as small as possible, extremely compact designs of the RF probehead are unavoidable. Typical diameters range between 20 and 30 mm. All electrically conducting

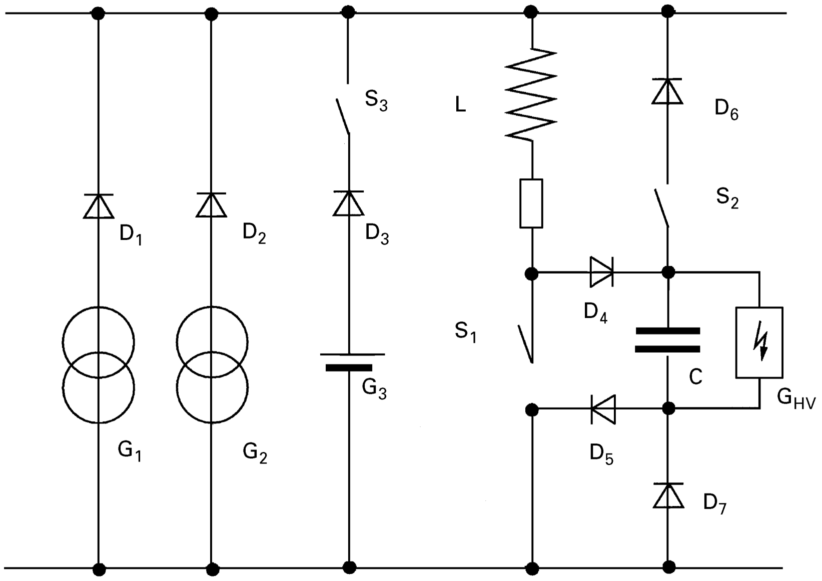



**Figure 5** Simplified scheme of the magnet current switching circuit for field-cycling purposes.

parts must be thoroughly slit in order to prevent the induction of eddy currents during the field cycle. The materials and the dimensions must be chosen in a way that the ohmic resistance is as high as possible. On the other hand, one must keep in mind that every slit works like an antenna, which might pick up undesired RF from other sources. Multiple layer shielding with interdigitated slits turn out to be best.

**Figure 6** schematically illustrates the 'heart' of the probe, the sample compartment with the RF coil and the temperature sensor. It is arranged closely below the tune and match capacitors, and is surrounded by the RF shielding. The outer diameter is 23 mm, the length of the sample coil 17 mm, with a maximal diameter of 12 mm for the sample containers. The coil is tuned either for proton or for deuteron resonance. That is, a 5 turn coil for 62 MHz or a 70 turn coil for 9.5 MHz, respectively. A glass tube supplies the air stream for heating or cooling of the sample. The temperature is controlled using a Pt100 resistor with an accuracy of ±1°C.

### General specifications

The total transition and ring-down time from the polarization field to a value of the evolution field, stable within ±10%, is less than 3 ms at all evolution field levels. Resonance in the detection field is reached within $\pm 10^{-5}$, in a transition time of 2.5 ms. The field switching rates involved are 750 T $s^{-1}$. Typical field values are listed in **Table 1**.

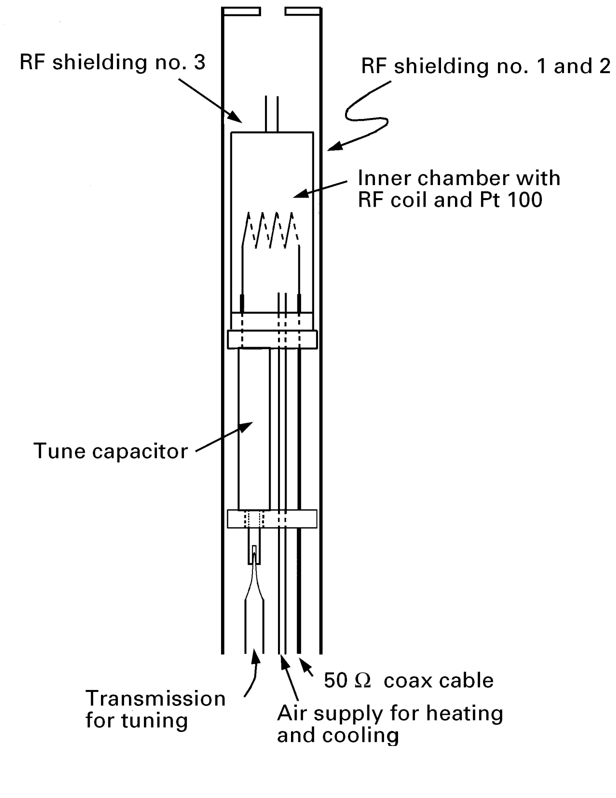



**Figure 6** Cross section through a probehead for field-cycling NMR relaxometry.

## Typical spin–lattice relaxation dispersion curves

Field-cycling NMR relaxometry is a versatile and powerful method for investigating molecular dynamics over a large range of timescales. It has been applied to manifold materials which show broad distributions of molecular motions, for example proteins, liquid crystals, synthetic polymers, and liquids confined in porous materials. **Figure 7A** represents an example for the investigation of polymer dynamics. The $T_1$-dispersion curve in the double-logarithmic scale shows the typical slopes observed in polyethylene oxide melts above the critical molecular

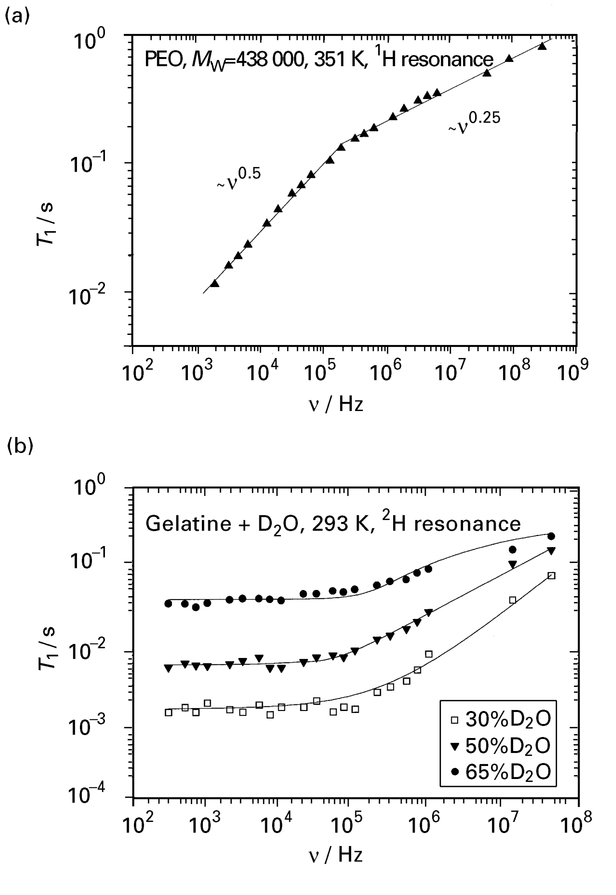



**Figure 7** Typical $^1$H and $^2$H spin–lattice relaxation curves. The power laws in (A) are discussed in Kimmich, Fatkullin, Weber and Stapf (1994) (see Further reading). The solid lines in (B) correspond to a fit of 'recorientations mediated by translational displacements (RMTD) model described in Kimmich (1997) (see Further reading).

**Table 1** Field values of the home-built relaxometer, expressed by the current and the Larmor frequencies of protons and deuterons.

| *Field* | *Current* (A) | $\nu_L(^1H)$ (Hz) | $\nu_L(^2H)$ (Hz) | *Rel. deviation* |
|---|---|---|---|---|
| Polarization | 55 | $12.4 \times 10^6$ | $1.904 \times 10^6$ | $2 \times 10^{-4}$ |
| Evolution (max.) | 28.8 | $6.51 \times 10^6$ | $1.0 \times 10^6$ | $3 \times 10^{-4}$ |
| Evolution (min.) | 0.009 | $2.0 \times 10^3$ | 300 | 0.2[b] |
| Detection[a] | 267–273 | 60–61.5 ($\times 10^6$) | 9.2–9.5 ($\times 10^6$) | $1.5 \times 10^{-5}$ |

[a] Depending on the state of charge of the batteries.
[b] Accuracy of the earth field compensation.

mass $\propto \nu^{0.5}$ and $\propto \nu^{0.25}$. The data can be explained by the renormalized Rouse theory. **Figure 7B** shows the $^2$H-dispersion of $D_2O$ in gelatine. Using deuterated water helps to clarify the mechanisms of water relaxation in biological tissue.

## List of symbols

$B_0$ = external magnetic flux density; $B_D$ = detection field; $B_E$ = magnetic flux density, evolution interval; $B_P$ = magnetic flux density, preparation interval; $G_i(\tau)$ = dipolar autocorrelation function; $J^{(i)}(\omega)$ = intensity function of the Larmor frequency; $M_E$ = Curie magnetization, evolution interval; $M_P$ = Curie magnetization, preparation interval; $S/N$ = signal-to-noise ratio; $T_1$ = spin–lattice relaxation time; $T_d$ = dipolar-order relaxation time; $T_{1\rho}$ = rotating-frame relaxation time; $T_2$ = transverse relaxation time; $\gamma$ = gyromagnetic ratio; $\mu_0$ = magnetic field constant.

*See also:* **Liquid Crystals and Liquid Crystal Solutions Studied By NMR; Magnetic Field Gradients, in High Resolution NMR; NMR Relaxation Rates; NMR Spectrometers; Proteins Studied Using NMR Spectroscopy.**

## Further reading

Kimmich R (1980) Field-cycling in NMR relaxation spectroscopy: Applications in biological, Chemical and polymer physics. *Bulletin on Magnetic Resonance* **1**: 195.

Kimmich R, Fatkullin N, Weber HW and Stapf S (1994) Nuclear spin–lattice relaxation and theories of polymer dynamics. *Journal of Non-Crystalline Solids* **172–174**: 689.

Kimmich R (1997) *NMR Tomography, Diffusometry, Relaxometry*. Springer-Verlag: Berlin.

Koenig SH and Brown RD (1990) Field-cycling relaxometry of Protein solutions and tissue: Implications for MRI. *Progress in NMR Spectroscopy* **22**: 487.

Noack F (1986) NMR field-cycling spectroscopy: Principles and applications. *Progress in NMR Spectroscopy* **18**: 171.

Noack F, Notter M and Weiss W (1988) Relaxation dispersion and zero-field spectroscopy of thermotropic and lyotropic liquid crystals by fast field-cycling NMR, *Liquid Crystals* **3**: 907.

Stapf S, Kimmich R and Seitter R.-O. (1995) Proton and deuteron field-cycling NMR relaxometry of liquids in porous glasses: Evidence for Lévy-Walk statistics. *Physical Review Letters* **75**: 2855.

# Rhenium NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# Rigid Solids Studied Using MRI

**David G Cory**, Massachusetts Institute of Technology, Cambridge, MA, USA

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

The function of many rigid solids is tied to their structure and hence there are questions in the study of synthetic, natural and biomaterials that can be addressed by imaging methods. In fact, the first nuclear magnetic resonance (NMR) images of rigid solids were acquired by Peter Mansfield and co-workers in 1973, the same year that Paul Lauterbur introduced the field of NMR imaging. NMR imaging of liquids rapidly advanced to where very sophisticated methods are used for clinically important diagnosis in medicine along with related techniques for the study of microscopy and complex flow dynamics. Progress in the study of rigid solids by NMR imaging has been somewhat slower, due partially to the priorities given to health care issues, but also due to technical challenges introduced by the broad NMR resonance lines of rigid solids. From a spectroscopy point of view the difference between a liquid and rigid solid is tied to the orientational dependence of NMR parameters such as chemical shifts and dipolar couplings. Mobile liquids have sufficient rotational motion that orientational dependencies are averaged to their isotropic value (hence the dipolar interaction vanishes since it has no isotropic component). For imaging purposes the sharper line widths seen in liquids leads to higher spatial resolution with simple imaging methods, however, the tensor properties of the full interactions permit more interesting forms of contrast to be achieved for solids. The goal then in NMR imaging of solids has been to develop methods that provide high quality images in the presence of broad lines and which can also take advantage of the greater array of contrast mechanisms. These methods fall into two general classes, wide-line methods that achieve spatial resolution by overwhelming the broad resonance line with a stronger magnetic field gradient, and coherent averaging approaches that carry out the imaging experiment in an interaction frame where the line width is reduced.

## Coherent and incoherent imaging

The basic ideas of NMR imaging are frequently presented as a simple result of the following three concepts:

1. the NMR resonance frequency is directly proportional to the applied magnetic field strength,
2. the NMR signal intensity is directly proportional to the spin concentration,
3. time-dependent linear magnetic field gradients can be applied along any arbitrary direction.

The simple idea is that a spatially varying magnetic field encodes the positions of the spins in their resonance frequencies, and thus the number of spins at any given location may be directly measured as the intensity of the NMR signal at the corresponding resonance frequency.

This real space picture of imaging corresponds best to an incoherent measurement where each spatial location is interrogated sequentially and the full image of the sample is built up from a sequence of such measurements. In such an approach the relative NMR signals from two different spatial locations have no phase (or coherent) relationship to one another. Incoherent imaging methods are occasionally employed in solid state imaging with the stray field imaging (STRAFI) method being the best known example (see below).

Most solid state imaging (and virtually all modern liquid state imaging) is carried out as coherent measurements where a specific phase relation is established between the signal from spins at different locations. It is this phase that is modulated throughout the course of the measurement. Such approaches have the advantage that the entire sample is measured simultaneously, leading to a multiplexing gain in the sensitivity. Coherent imaging is best introduced in a space reciprocal to real space, and discussed in terms of magnetization gratings.

## Magnetic field gradients, gratings, and *k*-space

The presence of a spatially varying magnetic field across the sample results in spins at different locations precessing at different rates, and so after some time they rotate to be out of phase with each other. This establishes a relative phase shift that depends

on the difference in the magnetic fields that the spins see and the evolution time. In most imaging, the spatially varying field is arranged to be a linear magnetic field gradient so that the difference in fields is directly proportional to the separation of the spins, with the proportionality constant being the strength of the magnetic field gradient (an experimentally controlled parameter).

Differential spin precession sets up a magnetization grating across the sample corresponding to a linear ramp of the phase of the nuclear spin (see **Figure 1**). Grating pictures of the spin dynamics emphasise the coherent nature of the imaging experiment, making it clear that there is a phase relation between any two given spins. The magnetization grating is most conveniently characterized by its corresponding wavenumber, $k$, which is,

$$k = 2\pi/\lambda \qquad [1]$$

where $\lambda$ is the spatial period of the grating – identified in **Figure 1**.

As the spins evolve in the magnetic field gradient, they are wound up into progressively tighter spirals and the pitch of the spiral continuously decreases, corresponding to an increasing wavenumber. In a constant magnetic field gradient the wavenumber increases linearly with time,

$$k(t) = k(0) + \gamma \partial B_z / \partial Z t \qquad [2]$$

For simplicity the gradient has been written as though it varies along the $z$-axis, in practice this could be any direction, and $\gamma$ is the gyromagnetic ratio.

The key relationship between the grating and the imaging measurement is that the NMR signal is the spin magnetization integrated over the sample. The grating thus is a spatial modulation of the spin density, $\rho$ $(x, y, z)$, at a given spatial period. The signal as a function of the wavenumber, $s(k)$, is a Fourier component of the local spin density of the sample,

$$s(\boldsymbol{k}) = \int \rho(\boldsymbol{r}) \exp(\boldsymbol{k} \cdot \boldsymbol{r}) \mathrm{d}\boldsymbol{r} \qquad [3]$$

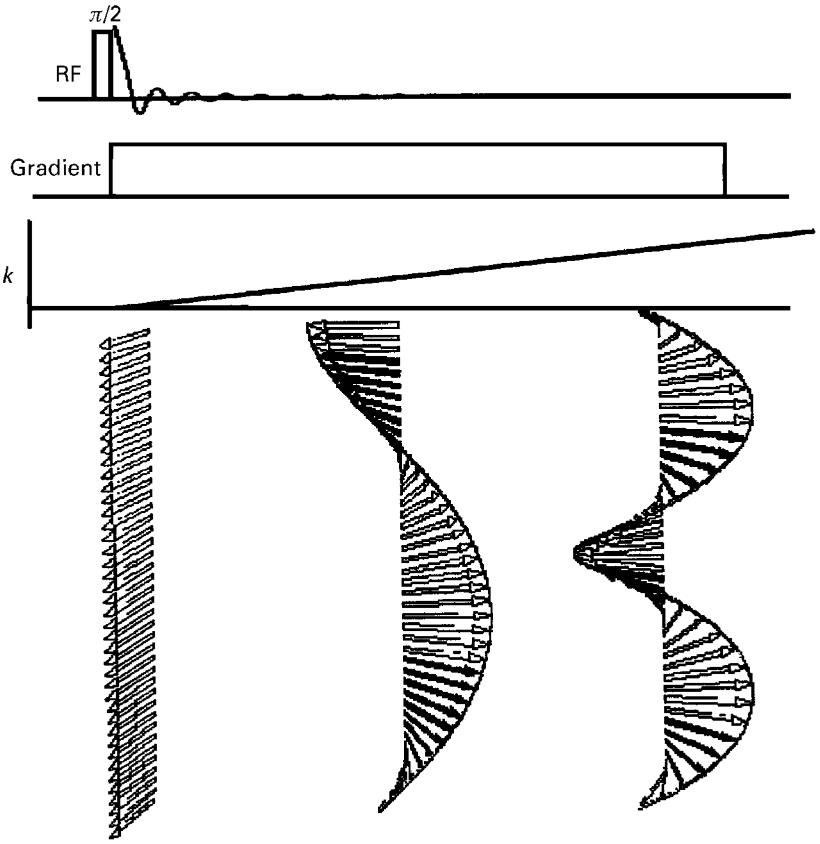



**Figure 1** A schematic representation of a one-dimensional imaging measurement and the corresponding magnetization grating. Notice that the pitch of the spatial helix becomes finer over time resulting in measurements of progressively higher wavenumbers, or Fourier components of the spin density. In a multidimensional imaging measurement wave vectors in 2 or 3 dimensions are indepentlymodulated.

where $\boldsymbol{k}$ is the reciprocal variable of the real space laboratory frame and $\boldsymbol{r}$ is a vector in 3D space.

In a coherent picture then the imaging experiment reduces to a series of measurements as a function of the wavenumber in three-dimensional space and taking the inverse Fourier transform of the NMR signal returns the real space image. At this level of approximation, the field of view and spatial resolution are given directly by the Nyquist sampling theorem.

There are a wealth of schemes for how to efficiently sample $k$-space in two or three dimensions. Since these are not unique to solid state studies the reader is referred to the descriptions in MRI theory.

## The complications from rigid solids

The picture set up above is complete when the gradients are the sole source of spin evolution, there is no relaxation and the spins do not move. In the liquid state the gradients can often be made sufficiently strong that relaxation and chemical shifts are minor complications and the limits to resolution and image distortions arise from limited NMR sensitivity and molecular transport. For rigid solids, molecular transport is not an issue, but the short lifetime of the transverse spin magnetization is.

The transverse relaxation time of proton-rich, rigid solids can be as short as tens of microseconds resulting in only this very limited time to impose a grating across the sample and to measure a given Fourier component. As seen from Equation [2] a given wavenumber (and hence spatial resolution) can still be reached, but as the lifetime of the transverse spin magnetization decreases the gradient strength must correspondingly increase. To obtain a wavenumber of $2\pi/20$ μm in 10 μs a gradient of 1000 G cm$^{-1}$ is needed. Typical gradients for small samples are of the order of 100 G cm$^{-1}$ although gradients as strong as 100 000 G cm$^{-1}$ have been employed in NMR.

The gradient strength is only a technical issue and large gradients are certainly available, their use in imaging, however, is limited by a corresponding loss in sensitivity. NMR is a very insensitive spectroscopy, with the most significant noise source being white noise from the receiver coil. As the gradient strength is increased the effective frequency bandwidth of the smallest volume element (voxel) also increases and the noise increases as the square root of the frequency spread (or gradient strength). Thus, as the resolution is made finer, and the voxel shrinks, the signal decreases with the number of spins in the voxel and the noise increases with the gradient strength. The overall result is that wide-line methods that rely upon large gradients to overwhelm the broad NMR line width typically have low resolution and low signal-to-noise ratios.

Formally, since the gradient evolution and the natural evolution of the spin system (that due to chemical shifts and dipolar couplings, etc.) commute with each other, the NMR spectrum blurs the image as a convolution. Of course the spectrum is in frequency units, and the image is in spatial units, so one must convert between these, and the gradient strength provides the necessary scaling,

$$\text{image}(x, yz) = \rho(x, y, z) \otimes \text{spectrum}\left(\frac{2\pi\nu}{\gamma(\partial B_z/\partial z)}\right) \qquad [4]$$

The comparison between liquid and solid state imaging is shown schematically in **Figure 2**.

## Methods

While obtaining an image is relatively straightforward and can be accomplished with methods that are closely related to those successfully employed in liquid state studies, the broad line width of rigid solids will result in low resolution and sensitivity unless special care is taken. This can be accomplished in three ways,

1. wide-line methods can be employed but with multiple sampling to increase the limited sensitivity,
2. constant-time methods can be used that accept the loss in sensitivity, but recover resolution by

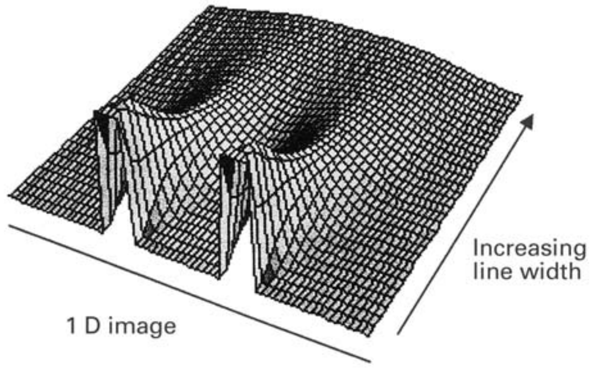



**Figure 2** A comparison of a series of 1D images as a function of the NMR line width with a constant gradient. The NMR line width is varied from 0.1% to 100% of the separation between the two bands in the image. As the resolution degrades so also does the sensitivity. Coherent averaging recovers both by artificially narrowing the NMR resonance. The solid state and coherent averaging line widths can differ by a factor of $10^4$.

using the fact that the gradient and internal Hamiltonians commute,

3. finally, the broad line width does not need to be accepted, it can be refocused through the use of coherent averaging resulting in a simultaneous increase in both resolution and sensitivity.

### Wide-line/strong gradients

Wide-line methods accept the limitation to resolution imposed by the sample's line width and use strong gradients to achieve high spatial encoding. The most successful of these is stray field imaging (STRAFI) where the sample is rotated and translated through the strong magnetic field gradient that is found at the fringe of any superconducting solenoid magnet. The imaging process is incoherent since only a single sensitive slice of the sample is observed at any given time, and the reconstruction process is based on the filtered back projection algorithm. The STRAFI method is described in detail in the article on MRI using stray fields.

Although the STRAFI method gives up the sensitivity advantage from the multiplexing inherent in coherent imaging approaches, a smaller multiplexing advantage can be achieved since there is no need to wait for spin lattice relaxation between the measurement from one slice to the next. In addition, the gradient-imposed noise bandwidth of coherent imaging can be partially avoided through pulsed spin-locked detection, essentially a narrow-band excitation and filtering that locally has good sensitivity.

STRAFI, like most wide-line methods, is typically free of distortions that can plague coherent averaging based imaging methods, but may introduce its own artifacts if the mechanics for sample motion are not excellent. In addition, the method is inherently three-dimensional and so the imaging time is very long.

A challenge of stray field (and wide-line methods) is that the types of contrast that can be introduced are rather limited, and usually based on rotating frame relaxation times. Contrast arising from the commonly observed spectroscopic parameters, particularly the chemical shift cannot be achieved.

### Constant time methods

As pointed out above, the gradient evolution and that from internal interactions commute at high fields, and so these act independently on the spin system. Emid and Creyghton pointed out that since these commute and since the gradient interaction is under experimental control, it is possible to set up the $k$-space sampling in such a fashion that the data are taken at a constant time following the excitation pulse (see **Figure 3**). The internal interactions lead to their normal complex spin evolution, but this is identical from measurement to measurement and therefore can be removed from the convolution: the only thing that changes is the gradient; Equation [4] may thus be rewritten as,

$$\mathrm{image}(x, y, z) = \mathrm{fid}(t_\mathrm{s})\rho(x, y, z) \qquad [5]$$

where the value of the normalized free induction decay at the sampling time, $\mathrm{fid}(t_\mathrm{s})$, is a complex number with magnitude less than unity. Notice that since the sampling time is kept constant, the free induction decay is not transformed into a spectrum by the inverse Fourier transform.

In constant-time methods the line width of the NMR resonance no longer blurs the image and the image resolution is solely determined by the sampling conditions for the coverage of $k$-space. The line width does, however, limit the sensitivity, and so

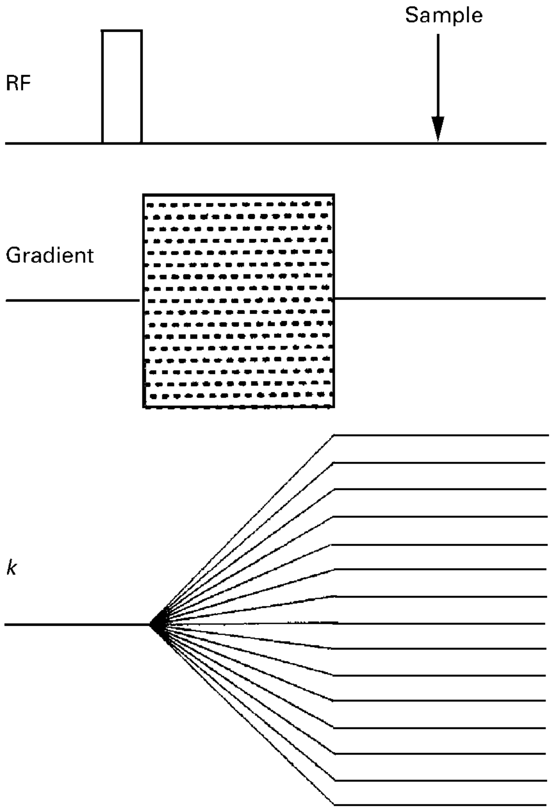



**Figure 3** The pulse sequence and $k$-vectors for constant-time imaging. Notice that all of the data are collected at a constant time following the excitation pulse and yet the $k$-vector is still modulated from point-to-point by systematically varying the strength of the gradient.

constant-time methods make the normal trade-off of sensitivity for resolution. At the same time, however, they achieve a robust experimental setup that is easy to apply.

With the typical 100 G cm$^{-1}$ gradients available from commercial small scale liquid state imaging equipment, constant-time methods can be used to observe rigid solids, but at very low *S/N* ratios and at low spatial resolution.

Constant-time methods are also used in liquid state imaging and in NMR microscopy to avoid distortions from chemical shifts and local variations in the bulk magnetic susceptibility.

### Coherent averaging

One of the most significant advances in the development of high resolution NMR methods for solids is Waugh's demonstration that the complex spin evolution in strongly coupled rigid solids can be refocused. This result is most easily understood in an interaction frame picture that depends on the time independence of the dipolar Hamiltonian in a rigid solid, rather than focusing on the complex spin dynamics. The interaction frame Hamiltonian can be periodically modulated through mechanical rotation (to modulate the geometric dependence), or through a series of RF pulses (to modulate the spin dependence), and in such a fashion that the time-averaged value is zero. Provided that the NMR response is sampled within this frame then the interaction appears to vanish. Coherent in this case corresponds to the structure of the modulation that is responsible for line-narrowing, and not a spatial phase as in the imaging case. Coherent differentiates the steady experimentally controlled, driven modulation from a natural stochastic, incoherent process.

For imaging, coherent averaging means is that we are not left to suffer the blurring from the broad resonance of a typical solid, but we can 'artificially' narrow this to achieve both higher sensitivity and resolution. Indeed, the first solid state images that were acquired by Mansfield and co-workers in 1973 used coherent averaging precisely for this purpose. Coherent-averaging methods have been developed to average dipolar couplings, chemical shifts, susceptibility shifts, and all of these simultaneously.

Coherent averaging for imaging has taken four approaches:

- Magic angle sample-spinning – which has the advantage of permitting contrast based on the isotropic chemical shift. MAS by itself, however, rarely removes the dipolar broadening efficiently.
- Multiple-pulse sequences based on solid echoes – which are very efficient at refocusing all internal interactions and can be combined with pulsed gradients to avoid image distortions.
- Multiple-pulse sequences based on magic echoes – are very similar to the solid echo-based methods, but have the advantage of longer windows for the gradients.
- Multiple-pulse sequences based on off-resonance excitation – where the effective interaction frame rotates at the magic angle and dipolar interactions can be made to vanish. Imaging in this interaction frame requires a combination of matched RF and d.c. gradients.

Magic angle sample spinning (MAS) is conceptually the simplest of these. Andrew demonstrated that a pair of dipolarly coupled spins are effectively decoupled from one another if the sample is spun at the 'magic-angle', the [111] direction – or the bisector of a cube. This rotation effectively removes the dipolar broadening from the system, and has the advantage of also removing the anisotropic chemical shift while leaving the isotropic shift for contrast generation. The challenge, of course, is that as the sample is rotating the imaging experiment must keep up with the sample and a variety of means have been engineered to synchronize the imaging experiment to the frame of reference rotating with the sample (see **Figure 4**). Once in this frame, the imaging methods are very similar to liquid state methods and the solid state aspects can, for the most part, be ignored.

Unfortunately, MAS alone, even at 20 kHz rotation rates, does not produce narrowing of the line width of a dipolarly coupled rigid solid. The challenge here is that the state of the nuclear spins are modulated by energy conserving mutual flip/flops (one spin flips to up while a neighbour flops to down) at a correlation time that is much faster than the modulation frequencies reached through sample rotation. So, MAS works well in special cases but in general must be combined with a multiple-pulse method.

Multiple-pulse methods achieve line narrowing by modulating the spin states quickly and symmetrically (see **Figure 5**). The promise of multiple-pulse approaches to solid-state imaging is both high resolution and good sensitivity, and that these can be combined with great flexibility in terms of contrast and image geometry. The challenge for multiple-pulse imaging is that the gradient evolution no longer commutes with the internal Hamiltonians as they are modulated, and so the spin dynamics of the gradient and internal Hamiltonians are no longer independent of each other. This can lead to pronounced image distortions, where the residual NMR linewidth depends on the gradient-induced frequency shift. The cleanest way to avoid such distortions is to apply the gradient

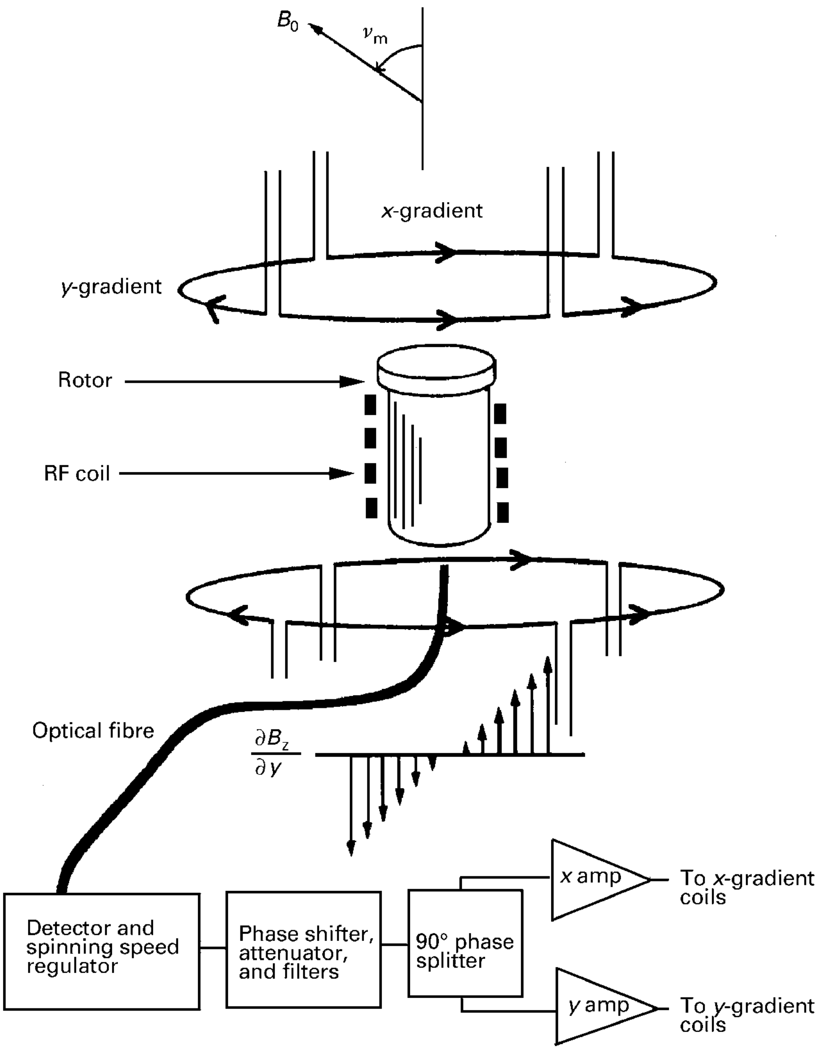



**Figure 4** The experimental setup for MAS imaging. The sample rotates about the magic angle at a frequency of about 10 kHz, and the current through the gradient coils is modulated so that the gradient fields move with the sample. $\nu_m$ is the magic angle, 54.7°.

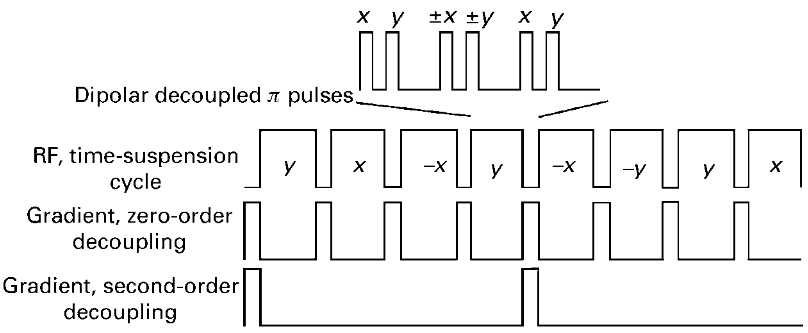



**Figure 5** The pulse sequence for a representative multiple-pulse, coherent averaging imaging measurement. The series of RF pulses lead to a refocusing of the dipolar and chemical shift interactions over the full cycle, and the gradients are added in selected windows. Even though the gradient and internal Hamiltonians do not commute, by restricting the gradients to selected windows an image which is substantially free of distortions can be collected. By placing the gradient pulses between each dipolar decoupled π pulse a larger effective gradient is achieved, but at a cost of some distortions, these are avoided by placing the gradients more sparsely where the gradient RF interaction is avoided, to second order.

in selected, symmetry related windows, which, however, requires specialized hardware to deliver strong fast gradient pulses.

Multiple-pulse imaging permits a surprisingly wide range of image contrast even though the internal Hamiltonian is suppressed during the imaging sequence. This, of course, removes the possibility of building contrast into the image during data sampling, and so most solid-state imaging methods are preceded with a contrast-generating sequence and perhaps slice selection if a two-dimensional image is desired.

## Resolution and field of view

Resolution in NMR imaging of solids is primarily limited by the low sensitivity of NMR. Typical detection limits are of the order of $10^{15}$ spins in the solid state, so that minimum voxels are approximately $10^4$ μm$^3$ (or a cube 20 μm on edge). Along any one dimension the resolution can be somewhat higher, but the volume remains approximately constant. The various approaches to solid state imaging present additional limitations to the resolution.

STRAFI methods rely on sample translation and so the precision and reproducibility of the mechanical arrangement also limits the resolution, here to approximately 10 $\mu$m. STRAFI methods are also set up to provide isotropic voxels in 3D-imaging, although much work has been done with 1D-STRAFI measurements where the question has been carefully tailored so they can be answered by studying profiles of the sample.

Constant-time methods trade-off significant sensitivity for resolution at a rate determined by the gradient strength. These methods typically work best for large objects where lower resolution is required and the gradient evolution time can be kept short.

Magic angle spinning-based imaging methods are very close to liquid state studies and the resolution limits normal in microscopy (better than 10 μm) apply here, however, spinning speeds are not sufficient to study most rigid solids. For rigid materials multiple-pulse coherent averaging is preferred and here the limitation relates to the challenges of sampling in the presence of multiple-pulse and pulsed gradients. In practical terms this permits about 128 volume elements to be acquired along a given axis and so the resolution is tied to the sample size.

This last point is more general, given sampling time and the lower efficiencies of larger detection coils imaging methods are typically limited to 512 elements along a side (or less). So the highest resolutions are achieved with small samples.

A related issue is the allowable size of the field of view. Solid-state imaging methods typically impose mechanical constraints on the sample size, such as to fit within the strong gradient in STRAFI (about 15 mm), or within the rotor for MAS methods (about 5 mm). Multiple-pulse methods work best in very high RF fields, and so the sample is limited to fitting within a relatively small RF coil (about 10 mm). Some single-sided imaging methods have been explored and are certainly feasible for soft materials, but are less well developed for rigid solids.

## Availability of instruments and methods

To date, NMR imaging of rigid solids is still mainly limited to experts who typically build a significant portion of their own equipment. Constant-time methods are the most widely available and can be used for preliminary studies, but the low sensitivity make many studies prohibitively long. STRAFI probes are commercially available, but here also the time to acquire an image is long and the restricted means of generating contrast limits applications. Recently, magic angle sample spinning probes with a single magnetic field gradient along the spinner axis have become commercially available (for NMR spectroscopic studies of semisolids) but probes with 3D gradient sets must still be home-made. All multiple-pulse coherent averaging approaches to imaging are carried out in homebuilt probes, and used with or without an associated fast gradient pulser. A large part of the challenge in employing multiple-pulse methods is the required stability and precision of the RF which typically demands the attention of an experienced spectroscopist to set up.

## Applications

Most of the effort in NMR imaging of rigid solids has been directed towards developing robust, high quality imaging methods, and few applications have been investigated in any detail. However, there are now a variety of good methods and it is an appropriate time for applications.

Proposed applications include the study of structure and morphology in synthetic polymers, an example being the selective imaging of one component of a multicomponent blend (see **Figure 6**). Such studies can be extended to the mapping of rigid versus mobile segments in a processed material (see **Figure 7**). There are a wealth of questions in materials processing and ageing that can be investigated by

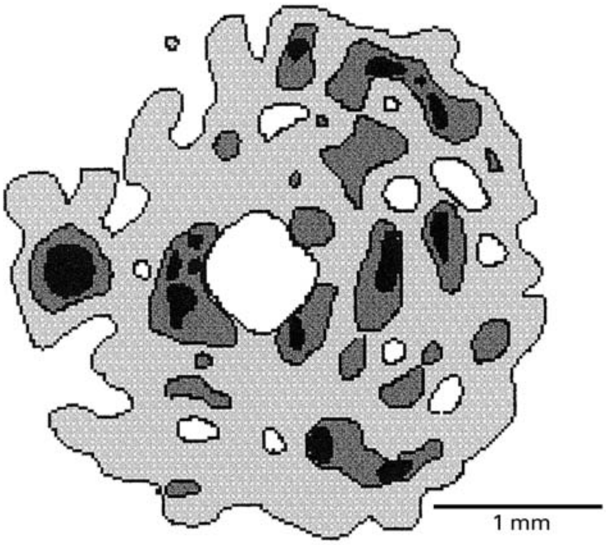



**Figure 6** MAS image of the polybutadiene fraction of a polybutadiene/polystyrene blend cast as a film from toluene. MAS alone at 5 kHz spinning is not sufficient to narrow the polystyrene NMR resonance. The in-plane resolution is $50 \times 50\ \mu m^2$. Figure reproduced with permission from Cory DG, de Boer JC and Veeman WS (1989) Magic angle spinning $^1$H NMR imaging of polybutadiene/polystyrene blends. *Macromolecules* **22**: 1618.

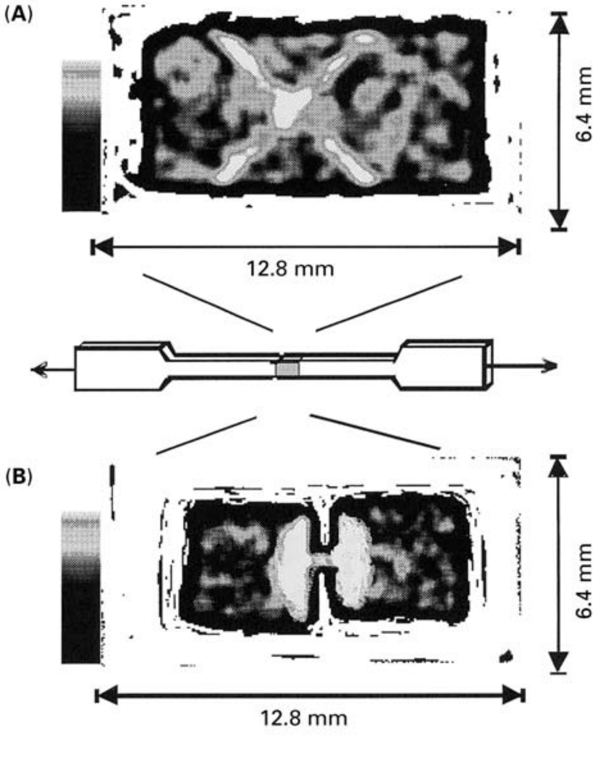



**Figure 7** (A)$T_{2e}$- magic-echo image of a piece of polycarbonate (see schematic) in which a crossed shearband has been created. (B) An image of a sample that has been stretched nearly to the breaking point: The regions with high intensity correspond to areas of low molecular mobility. Figure reproduced with permission from Traub B, Hafner S, Maring D and Spiess HW (1998) In: Blümler P, Blümich B, Botto B and Fukushima E, *Spatially Resolved Magnetic Resonance*, Weinheim: Wiley-VCH, p 191.

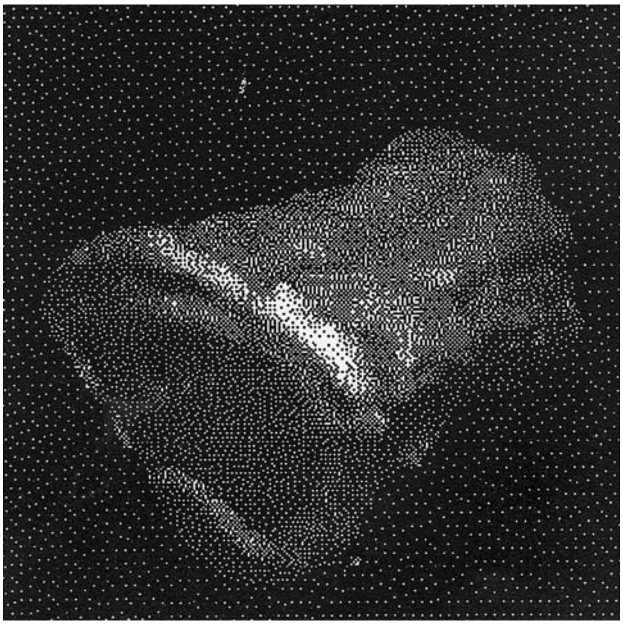


**Figure 8** Three-dimensional constant-time image of a bovine femur showing both the compact bone and the interior bone marrow. The image resolution is $1.1 \times 1.1 \times 4.4\ mm^3$ and the image was acquired over a period of 5 h. Figure reproduced with permission from Balcon BJ (1998) In: Blümler P, Blümich B, Botto B, and Fukushima E (eds) *Spatially Resolved Magnetic Resonance.* Weinheim: Wiley-VCH, p. 84.

such methods, even though the resolution of the image is on the order of tens of micrometres.

There has also been interest in biomedical studies of bone, bone cements and the function of synthetic organs. An example of an image of bovine bone is shown in **Figure 8**. Progress has also been made in developing methods of following bone repair *in vivo*.

## List of symbols

$B$ = magnetic field strength; $k$ = wave number; $s$ = signal; $t$ = time; $\gamma$ = gyromagnetic ratio; $\lambda$ = spatial period; $\nu$ = frequency; $\rho$ = spin density.

*See also:* **High Resolution Solid State NMR, $^1$H, $^{19}$F; MRI Instrumentation; MRI of Oil/Water in Rocks; MRI Theory; MRI Using Stray Fields; NMR Microscopy; NMR of Solids; NMR Principles.**

## Further reading


Blümler P, Blümich B, Botto B and Fukushima E (1998) *Spatially Resolved Magnetic Resonance*. Weinheim: Wiley.

Blümler P and Blümich B (1994) Solid state NMR. *NMR Basic Principles and Progress* 30: 209.

Callaghan PT (1991) *Principles of NMR Microscopy*. Oxford: Oxford University Press.

Cory DG (1992) Solid state NMR imaging. *Annual Reports on NMR Spectroscopy*, **24**: 87–180.
Kimmich R (1997), *NMR-Tomography, Diffusometry, Relaxometry*. Berlin: Springer-Verlag.
Mansfield P and Grannell PK (1975) "Diffraction" and microscopy in solids and liquids by NMR. *Physical Review B* **12**: 3618–3634.

# Rotational Resonance in Solid State NMR

*See* **Solid State NMR, Rotational Resonance.**

# Rotational Spectroscopy, Theory

**Iain R McNab**, University of Newcastle upon Tyne, UK
Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Theory**

## Synopsis

The rotational spectra of gas phase molecules is explained using the quantum theory of angular momentum. The spectra of simple rigid molecules are explained on the basis of selection rules and energy level expressions which depend upon the moments of inertia of the molecules. We consider how the theory needs to be extended to include effects of nonrigidity, and also small effects due to interactions between the angular momentum of the nuclear framework, electron spins, and nuclear spins. A calculation of relative intensities using spherical tensor techniques is given for a diatomic molecule using Hund's case *a*.

## Introduction

The rotational spectroscopy of molecules typically (but by no means always) occurs in the microwave region of the spectrum. In solids, molecules are usually not free to rotate, and in liquids collisions normally render absorption featureless; we therefore consider only the rotational spectroscopy of gaseous molecules.

Spectroscopic assignments are often discussed in terms of selection rules which determine whether or not it is possible for a transition in a molecule to be caused by a particular interaction, such as the electric dipole interaction. The only true selection rules are for total angular momentum ($F$) and parity (+ or –). For complex molecules with many different angular momenta, it is more profitable to discuss spectra in terms of relative intensities rather than 'allowed transitions'. We can readily calculate the relative intensities for many types of transitions in different molecules using the quantum theory of angular momentum.

Rotational effects (fine structure) are also seen in high resolution electronic and vibrational spectra; additional structure due to the interaction of electrons with nuclear electric and magnetic moments may also be observed and is called hyperfine structure. The simplest rotational spectra are associated with diatomic molecules with no electronic orbital or spin angular momentum (i.e. singlet sigma states) and these are considered first.

The theory of rotational spectroscopy depends upon an understanding of the quantum mechanics of angular momentum. Useful results from the quantum theory of angular momentum are given in **Table 1** and useful results in spherical tensor notation are given in **Table 2**.

A full understanding of a rotational spectrum is achieved when we can calculate the rotational energy levels of the molecules and then calculate the

frequencies and intensities of the transitions between them. The cheapness and availability of computers and subroutines for diagonalizing matrices means that most spectra are now assigned on the basis of calculated spectra iterating towards convergence with measured line positions and intensities; we discuss how this is accomplished.

**Table 1** Useful results from the quantum theory of angular momentum

$\boldsymbol{J}^2|JKM\rangle = J(J+1)|JKM\rangle\hbar^2$
$J_c|JKM\rangle = K|JKM\rangle\hbar$
$J_z|JKM\rangle = M|JKM\rangle\hbar$
$J^{\pm}|JKM\rangle = J_x \pm iJ_y)|JKM\rangle = [J \mp M)(J \pm M+1)]^{1/2}|JKM \pm 1\rangle\hbar$

The lower result shows the operation of the shift (or ladder) operator in the space fixed axis system using the normal phase convention.

Angular momentum relationships are based on the commutation relationships

$[J_x, J_y] = i\hbar J_z$
$[J^2, J_z] = 0$

Addition of two angular momenta:

$\boldsymbol{J} = \boldsymbol{J}_1 + \boldsymbol{J}_2$

and the allowed values of the quantum number $J$ are given by

$$J = J_1 + J_2, J_1 + J_2 - 1, \ldots |J_1 - J_2| \qquad [1]$$

Molecule fixed angular momenta are problematic, because the commutation relationships between different components are reversed. There are many ways of taking account of this, but the easiest is to calculate everything in space fixed axes and convert (where necessary) to molecule fixed axes by using Equations [13] and [14].

**Table 2** Useful relationships for the use of spherical tensors in the calculation of matrix elements of angular momentum operators (courtesy of Prof. JM Brown, PTCL, Oxford University)

The three components of a first rank spherical tensor are defined as a linear combination of Cartesian components:

$$\begin{aligned} T^1_{p=1}(\boldsymbol{A}) &= -\frac{1}{\sqrt{2}}(A_X + iA_Y) \\ T^1_{p=0}(\boldsymbol{A}) &= A_Z \\ T^1_{p=-1}(\boldsymbol{A}) &= \frac{1}{\sqrt{2}}(A_X - iA_Y) \end{aligned} \qquad [2]$$

where the superscript '$k$' is the rank and the subscript '$p$' is the component which can have any of $2k+1$ values, $k$, $k-1$, $k-2, \ldots -k$. Note that this transformation essentially corresponds to using $J_z$, $J^+$ and $J^-$ in place of $J_z$, $J_x$ and $J_y$.
The scalar product of 2 irreducible spherical tensor operators of rank $k$ is defined by

$$T^k(\boldsymbol{A}).T^k(\boldsymbol{B}) = \sum_p (-1)^p T^k_p(\boldsymbol{A}) T^k_{-p}(\boldsymbol{B}) \qquad [3]$$

The Wigner–Eckart theorem enables the $M_J$ dependence of a matrix element to be factored out as follows:

$$\langle \gamma J M_J | T^k_p(\boldsymbol{A}) | \gamma' J' M'_J \rangle = (-1)^{J-M_J} \begin{pmatrix} J & k & J' \\ -M_J & p & M'_J \end{pmatrix} \langle \gamma J \| T^k(\boldsymbol{A}) \| \gamma' J' \rangle \qquad [4]$$

where $\langle \gamma J \| T^k(\boldsymbol{A}) \| \gamma' J' \rangle$ is defined by the equation and is known as a reduced matrix element (it has no $M$ dependence).
The Wigner 3-j symbol $\begin{pmatrix} J & k & J' \\ -M_J & p & M'_J \end{pmatrix}$ which appears in the definition is itself defined by the coupling equation

$$|J_1 J_2 J M\rangle = \sum_{M_1, M_2} |J_1 M_1\rangle |J_2 M_2\rangle (-1)^{-J_1+J_2-M} (2J+1)^{1/2} \begin{pmatrix} J_1 & J_2 & J \\ M_1 & M_2 & -M \end{pmatrix} \qquad [5]$$

$J_1$ and $J_2$ (with component quantum numbers $M_1$ and $M_2$, respectively) are coupled together to yield the resultant angular momentum $\boldsymbol{J}$ with component $M$. The 3-j symbol is zero unless the bottom row sum equals zero and the three arguments in the top row must add together according to Equation [1] in **Table 1**.

*Continued*

**Table 2** (*Continued*)

Computation of reduced matrix elements $\langle \gamma J \parallel T^k(\boldsymbol{A}) \parallel \gamma' J' \rangle$ is accomplished by choosing values of components which give a simple value for the matrix element $\langle \gamma J M_J | T^k_p(\boldsymbol{A}) | \gamma' J' M'_J \rangle$ and dividing the matrix element by

$$(-1)^{J-M_J} \begin{pmatrix} J & k & J' \\ -M_J & p & M'_J \end{pmatrix}$$

Two common examples are (a) $\langle J\Omega M | J_Z | J'\Omega' M' \rangle = M \langle J\Omega M | J'\Omega' M' \rangle = M \delta_{JJ'} \delta_{MM'} \delta_{\Omega\Omega'}$
hence $\langle J\Omega \parallel T^1(\boldsymbol{J}) \parallel J\Omega \rangle = \delta_{JJ'} \delta_{\Omega\Omega'} [J(J+1)(2J+1)]^{1/2}$
(b) similarly, $\langle S \parallel T^1(\boldsymbol{S}) \parallel S' \rangle = \delta_{SS'} [S(S+1)(2S+1)]^{1/2}$
Matrix elements of the tensor product of two tensor operators: Let $X^k_p(\boldsymbol{A}, \boldsymbol{B})$ be the tensor product of rank $k$ of $T^{k_1}(\boldsymbol{A})$ and $T^{k_2}(\boldsymbol{B})$:

$$X^k_p(\boldsymbol{A}, \boldsymbol{B}) = (-1)^{k_1-k_2+p} [(2k+1)]^{1/2} \sum_{p_1, p_2} \begin{pmatrix} k_1 & k_2 & k \\ P_1 & P_2 & -P \end{pmatrix} T^{k_1}_{p_1}(\boldsymbol{A}) T^{k_2}_{p_2}(\boldsymbol{B}) \quad [6]$$

Then the reduced matrix element of $X^k_p(\boldsymbol{A}, \boldsymbol{B})$ is related to those of $T^{k_1}(\boldsymbol{A})$ and d $T^{k_2}(\boldsymbol{B})$ by the general expression

$$\langle \gamma J \parallel X^k(\boldsymbol{A}, \boldsymbol{B}) \parallel \gamma' J' \rangle = [2k+1]^{1/2} (-1)^{k+J+J'} \sum_{\gamma'', J''} \begin{Bmatrix} k_1 & k_2 & k \\ J' & J & J'' \end{Bmatrix} \langle \gamma J \parallel T^{k_1}(\boldsymbol{A}) \parallel \gamma'' J'' \rangle \langle \gamma'' J'' \parallel T^{k_2}(\boldsymbol{B}) \parallel \gamma' J' \rangle \quad [7]$$

where $\begin{Bmatrix} k_1 & k_2 & k \\ J' & J & J'' \end{Bmatrix}$ is a Wigner 6-j symbol.

The general formula can be simplified if the two operators commute, i.e. if $T^{k_1}(\boldsymbol{A})$ operates only on one part of a coupled system ($\boldsymbol{J}_1$) and $T^{k_2}(\boldsymbol{B})$ on another ($\boldsymbol{J}_2$) then the reduced matrix elements of $X^k_p(\boldsymbol{A}, \boldsymbol{B})$ are given by:

$$\langle \gamma J_1 J_2 J \parallel X^k(\boldsymbol{A}, \boldsymbol{B}) \parallel \gamma' J_1{}' J_2{}' J' \rangle = \sum_{\gamma''} [(2J+1)(2J'+1)(2k+1)]^{1/2}$$
$$\times \begin{Bmatrix} J_1 & J'_1 & k_1 \\ J_2 & J'_2 & k_2 \\ J & J' & k \end{Bmatrix} \langle \gamma J_1 \parallel T^{k_1}(\boldsymbol{A}) \parallel \gamma'' J_1{}' \rangle \langle \gamma'' J_2 \parallel T^{k_2}(\boldsymbol{B}) \parallel \gamma' J'_2 \rangle \quad [8]$$

$\begin{Bmatrix} J_1 & J'_1 & k_1 \\ J_2 & J'_2 & k_2 \\ J & J' & k \end{Bmatrix}$ is a Wigner 9-j symbol.

A simplification of the above expression which is often encountered is for the scalar product ($X^k = X^0$) of two commuting tensor operators:

$$\langle \gamma J_1 J_2 J M | T^k(\boldsymbol{A}).T^k(\boldsymbol{B}) | \gamma' J_1{}' J_2{}' J' M' \rangle = (-1)^{J'_1 + J_2 + J'} \delta_{JJ'} \delta_{MM'}$$
$$\times \begin{Bmatrix} J & J_2 & J_1 \\ k & J_1{}' & J_2{}' \end{Bmatrix} \sum_{\gamma''} \langle \gamma J_1 \parallel T^k(\boldsymbol{A}) \parallel \gamma'' J_1{}' \rangle \langle \gamma'' J_2 \parallel T^k(\boldsymbol{B}) \parallel \gamma' J_2{}' \rangle \quad [9]$$

If a single operator in a coupled scheme of angular momenta acts on only one part ($J_1$ or $J_2$) then we have the reduced matrix element expressions:
(a) for operator $T^{k_1}(\boldsymbol{A})$ operating only on $J_1$:

$$\langle \gamma J_1 J_2 J \parallel T^{K_1}(\boldsymbol{A}) \parallel \gamma' J_1{}' J_2{}' J' \rangle = (-1)^{J_1+J_2+J'+k_1} \delta_{J_2 J_2'} [(2J+1)(2J'+1)]^{1/2} \begin{Bmatrix} J_2 & J_1 & J \\ k_1 & J' & J_1{}' \end{Bmatrix} \langle \gamma J_1 \parallel T^{k_1}(\boldsymbol{A}) \parallel \gamma' J_1{}' \rangle \quad [10]$$

(b) for operator $T^{k_2}(\boldsymbol{B})$ operating only on $J_2$:

$$\langle \gamma J_1 J_2 J \parallel T^{k_2}(\boldsymbol{B}) \parallel \gamma' J_1{}' J_2{}' J' \rangle = (-1)^{J_1{}'+J_2{}'+J+k_2} \delta_{J_1 J_1{}'} [(2J+1)(2J'+1)]^{1/2} \begin{Bmatrix} J_1 & J_2 & J \\ k_2 & J' & J_2{}' \end{Bmatrix} \langle \gamma J_2 \parallel T^{k_2}(\boldsymbol{B}) \parallel \gamma' J_2{}' \rangle \quad [11]$$

*Continued*

**Table 2** (*Continued*)

For the scalar product of two operators which **both** act on part one ($\boldsymbol{J}_1$) of a coupled scheme, we have

$$\langle \gamma J_1 J_2 JM | T^{k_1}(\boldsymbol{A}).T^{k_1}(\boldsymbol{B}) | \gamma' J_1' J_2' J' M' \rangle$$
$$= \delta_{JJ'}\delta_{MM'}\delta_{J_1J_1'}\delta_{J_2J_2'} \sum_{\gamma'',J_1''} (-1)^{J_1-J_1''}(2J_1+1)^{-1} \langle \gamma J_1 \| T^{k_1}(\boldsymbol{A}) \| \gamma'' J_1' \rangle \langle \gamma'' J_1'' \| T^{k_1}(\boldsymbol{B}) \| \gamma' J_1' \rangle \quad [12]$$

For the relation between operators in space-fixed and molecule fixed co-ordinate systems, we use the notation that p gives the sPace fixed coordinate and q gives the moleQule fixed coordinate of a tensor operator. The relationship between the same tensor in the two coordinate systems is given by:

$$T_p^k(\boldsymbol{A}) = \sum_q D_{pq}^k(\alpha,\beta,\gamma)^* T_q^k(\boldsymbol{A}) \quad [13]$$

where $D_{pq}^k$ (α, β, γ)* = $D_{pq}^k$ (ω)* is the complex conjugate of the *k*th rank rotation matrix $D^k(\omega)$. α, β, γ are the three Euler rotation angles that specify the relationship of the molecule fixed coordinate system to the space fixed coordinate system. The phase convention used here is opposite to that of Edmonds. The inverse relationship is

$$T_q^k(\boldsymbol{A}) = \sum_p D_{pq}^k(\omega) T_p^k(\boldsymbol{A}) = \sum_p (-1)^{p-q} D_{-p,-q}^k(\omega)^* T_p^k(\boldsymbol{A}) \quad [14]$$

The $D_{pq}^k$ (ω)* are simply related to the eigenfunctions of the symmetric top:

$$|JKM\rangle = \left[\frac{2J+1}{8\pi^2}\right]^{1/2} D_{MK}^J(\omega)^* \quad [15]$$

We now quote two additional useful relationships for tensor operators which can be quantized in both molecule- and space-fixed coordinate systems:

$$\langle JKM | D_{p,q}^k(\omega)^* | J'K'M' \rangle = (-1)^{K-M}[(2J+1)(2J'+1)]^{1/2} \begin{pmatrix} J & k & J' \\ -K & q & K' \end{pmatrix} \begin{pmatrix} J & k & J' \\ -M & p & M' \end{pmatrix} \quad [16]$$

$$\langle JK \| D_{\bullet,q}^k(\omega)^* \| J'K' \rangle = (-1)^{J-K}[(2J+1)(2J'+1)]^{1/2} \begin{pmatrix} J & k & J' \\ -K & q & K' \end{pmatrix} \quad [17]$$

Note that the Wigner 3-j, 6-j and 9-j symbols are simply numbers when evaluated for particular values of the angular momentum symbols. There are numerous tabulations and programs available which allow their evaluation and the program Mathematica can handle 3-j and 6-j symbols analytically.

## Born–Oppenheimer approximation

The Born–Oppenheimer approximation assumes that the molecular wavefunction can be written in the form $\psi_{\text{total}} = \psi_{\text{electronic}}\psi_{\text{vibration}}\psi_{\text{rotation}}$ and therefore that the energies due to each type of motion are additive $E_{\text{total}} = E_{\text{electronic}} + E_{\text{vibrational}} + E_{\text{rotational}}$. This gives rise to electronic and vibrational quantum numbers which are often a good first order description of the energy levels and wavefunctions of a molecule. The application of the Born–Oppenheimer approximation leads naturally to the rigid rotor description of a molecule – we treat the rotation separately from the vibrational motion of the nuclei, which are considered later as a perturbation.

The rotational motion of a molecule is quantized in units of $\hbar$, and this free rotation can take place about three orthogonal axes which are associated with three distinct moments of inertia, $I_A$, $I_B$, $I_C$. The moments of inertia are, in fact, quantum mechanical expectation values of the moments of inertia in a particular electronic and vibrational state of the molecule. We associate the moments of inertia with three rotational constants, $A$, $B$, $C$, which determine the energy levels of the molecule.

The Hamiltonian for rotation is $H_r = AJ_A^2 + BJ_B^2 + CJ_C^2$. If the molecule has symmetry then we can take advantage of the symmetry to rewrite the Hamiltonian in a form which is easy to solve. The appropriate Hamiltonians and energy level expressions are given in **Table 3**.

## Molecules without electronic or nuclear angular momentum

The simplest spectra belong to molecules with no electronic angular momentum and no nuclear spin angular momenta and we consider these first. The relevant classifications, Hamiltonians and energy level expressions are given in **Table 3**. The quantum mechanical treatment of spherical motion gives rise

**Table 3** Rigid rotor classifications.
We classify rigid rotors according to their moments of inertia about three perpendicular axes. There is a convention (due to Mulliken) that the three moments of inertia are labelled ordered according to size by $I_A \le I_B \le I_C$ and hence C ≤B ≤A

| | *Moments of inertia* | *Hamiltonian and energy levels* |
|---|---|---|
| Spherical tops: | $I_A = I_B = I_C$ | $H_r = B\boldsymbol{J}^2$<br>$E_r = BJ(J+1)$ |
| Linear molecules: | $I_A = 0 \quad I_B = I_C$ | $H_r = B(J_B^2 + J_C^2) = B\boldsymbol{J}^2$<br>$E_r = BJ(J+1)$ |
| Symmetric tops: | | |
| Prolate | $I_A < I_B = I_C$ | $H_r = AJ_A^2 + B(J_B^2 + J_C^2) = BJ^2 + (A-B)J_A^2$<br>$E_r = BJ(J+1) + (A-B)K^2$ |
| Oblate | $I_A = I_B < I_C$ | $H_r = AJ_A^2 + B(J_B^2 + J_C^2) = BJ^2 + (C-B)J_A^2$<br>$E_r = BJ(J+1) + (C-B)K^2$ |
| Asymmetric top | $I_A \ne I_B \ne I_C$ | $H_{\text{rotation}} = AJ_A^2 + BJ_B^2 + C_C^2J$<br>analytic solutions exist for only the first few levels, these solutions are given in **Table 4**. |

to the eigenfunctions of the spherical top $\psi_{\text{rotation}}$ which we denote by the ket $|JKM\rangle$ ; the letters inside the bracKET denote the quantum numbers used to describe the state. For other problems involving rotation it is convenient to use a basis of $|JKM\rangle$ states and express the eigenfunctions of each problem as a linear combination of them, that is:

$$|J\Gamma\rangle = \sum_K a_K^J |JKM\rangle$$

$$H_{\text{rotation}}|J\Gamma\rangle = E|J\Gamma\rangle$$

For a molecule, there are two projection quantum numbers of $\boldsymbol{J}$ which are simultaneously conserved, $M$ and $K$. $M$ is the space fixed projection of $\boldsymbol{J}$ and $K$ is the molecule fixed projection of $\boldsymbol{J}$. Both $M$ and $K$ can take values from $+J$ to $-J$ in steps of one. In the absence of an electric or magnetic field the eigenvalues are independent of $M$ and each eigenfunction gives rise to an eigenvalue which is $2J+1$ degenerate.

### Spherical top molecules

The spherical top has the same simple energy level pattern as a linear molecule. However, a spherical top cannot have a permanent dipole moment, and therefore has no rotational spectrum.

### Linear molecules (including diatomic molecules)

The linear molecule energy levels are given by $E = B.J(J+1)$, where $J$ is the rotational quantum number for rotation perpendicular to the bond, and $B$ is the rotational constant. The energy level structure is given in **Figure 2A**. A typical rotational spectrum (of a diatomic molecule) is shown in **Figure 1**.

To understand the spectrum we must be able to calculate the absorption frequencies of the lines and their relative intensities. The Beer–Lambert law

$$I(\omega) = I_0(\omega)\exp[-\gamma(\omega)l]$$

describes absorption in a thin sample (after length $l$ is traversed) in terms of the absorption coefficient $\gamma(\omega)$ which is given by:

$$\gamma(\omega) = \frac{8\pi^3}{3hc}\omega\left(\frac{N_m}{g_m} - \frac{N_n}{g_n}\right)|\langle J\,\Gamma|\boldsymbol{\mu}|J'\Gamma'\rangle|^2\ S(\omega,\omega_0)$$

where $S(\omega, \omega_0)$ is a normalized line shape function. Evaluation of $\langle J\,\Gamma|\boldsymbol{\mu}|J'\Gamma'\rangle$ yields selection rules for the basis functions (see below). For the linear molecule with no electronic angular momentum and no nuclear spin, we can calculate selection rules which tell us whether a particular transition is allowed or not, by understanding that we are adding a photon angular momentum of one to the angular momentum of the molecule. Therefore, by the formula for addition of angular momenta (see **Table 1**), the change in angular momentum allowed is $\Delta J = 0,\pm1$ and $\Delta J = 0$ cannot occur from $J = 0$. A complete treatment shows also that $\Delta K = 0$, $\Delta M = 0,\pm1$ and the dipole change occurs along the internuclear axis. The possible absorption energies are given by

$$\begin{aligned}\Delta E = hf &= B[J'(J'+1) - J(J+1)]\\ &= B[(J+1)(J+2) - J(J+1)] = 2B(J+1)\end{aligned}$$

So we expect a series of lines separated by $2B$, as shown in **Figure 1**.

The relative intensities of the lines also depend upon the difference in population of the two levels concerned at the temperature considered, $N_m/g_m - N_n/g_n$ ($g_n$ is the degeneracy of level $n$, etc.). For a sample in thermal equilibrium the relative populations are given by the Boltzmann distribution

$$N_m/g_m = (N_n/g_n)\exp[-(E_n - E_m)/kT]$$

and this factor accounts for the envelope of the spectral intensities shown.

In the case of a molecule with two identical nuclei (such as $CO_2$), relative line intensities are affected by the nuclear spins. The nuclear spin governs the intensities through the generalized Pauli exclusion principle; 'all wavefunctions are antisymmetric with respect to the interchange of fermions (half-integer spin particles) and symmetric with respect to the interchange of bosons (integer spin particles)' A general treatment uses the molecular symmetry group. For a linear molecule with zero spin nuclei the antisymmetric rotational levels are missing entirely. In the case of $CO_2$ the $\Sigma_g^+$ electronic state has only even $J$ levels. If only one pair of identical nuclei have nonzero spins, then the ratio of the statistical

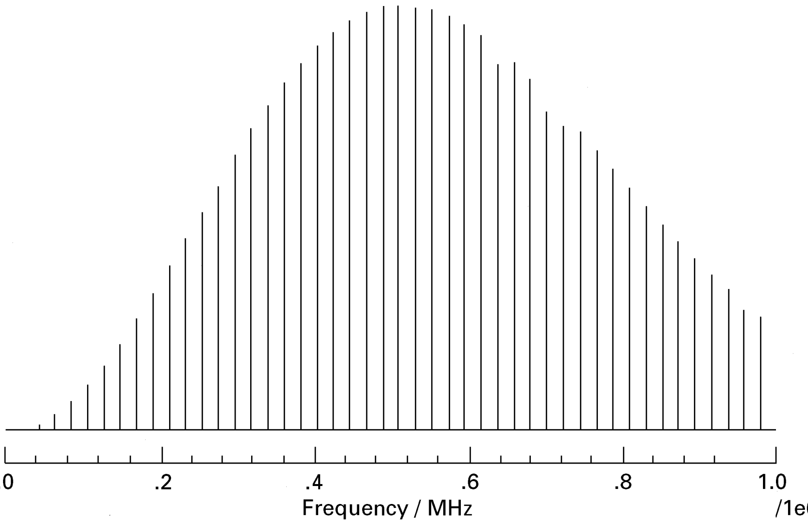



**Figure 1** A simulated spectrum of BrF at 300 K. All simulated spectra were generated using Pgopher, courtesy of Dr. CM Western, University of Bristol.

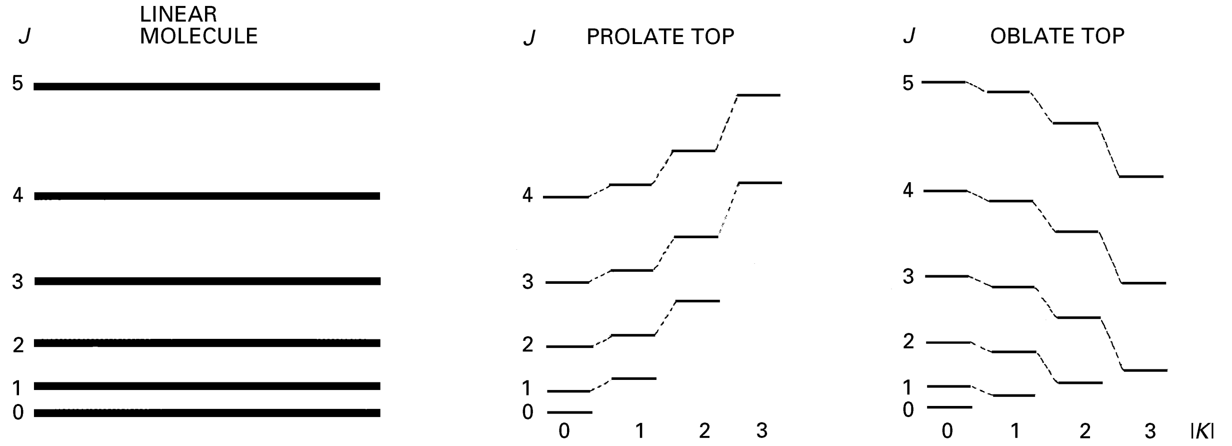



**Figure 2** Energy levels of the linear molecule and prolate and oblate symmetric tops. Modified from Figures 3.1 and 3.2 of Kroto HW (1992) *Molecular Roatation Spectra*, pp 32–33, Canada: Dover Books.

weights of the symmetric and antisymmetric rotational levels is (I+1): I for bosons or I: (I+1) for fermions.

The rotational constant $B$ is easily determined from the spectrum. For a diatomic molecule, $B$ is given by the expression

$$B = \left\langle \frac{\hbar^2}{2\mu R^2} \right\rangle = \frac{\hbar^2}{2\mu} \left\langle \frac{1}{R^2} \right\rangle$$

where $\mu$ is the reduced mass, $R$ is the bond length and the angled brackets show that an average value (expectation value) must be calculated. Back calculation yields the expectation value of $1/R^2$, but inverting this and taking the square root does NOT give the expectation value of $R$, even for a harmonic oscillator! For linear molecules with more than one atom, isotopic substitution can be used to find different $B$ values, and hence the expectation values of $1/R^2$ for several different bonds can be found simultaneously.

### Symmetric top molecules

We consider molecules to be symmetric tops if two of their moments of inertia are the same. It is useful to consider two cases separately, the 'prolate' and 'oblate' symmetric tops. A diagram of the lowest energy levels of both prolate and oblate symmetric tops is given in **Figure 2B**. The selection rules for a symmetric top are the same as for a linear molecule: $\Delta J = 0,\pm 1$ and $\Delta J = 0$ cannot occur from $J = 0$, $\Delta K = 0$, $\Delta M = 0,\pm 1$ and the dipole change occurs along the symmetry axis.

### Asymmetric top molecules

The energy levels of an asymmetric top molecule cannot in general be expressed in a simple algebraic form. Those levels for which such simple solutions exist are given in **Table 4**. The complication arises because the functions $|JKM\rangle$ which we are using as our basis are NOT eigenfunctions for the asymmetric top. However, many asymmetric rotors are nearly a prolate or oblate top, and the perturbations to the energy levels in such cases may be quite small. **Figure 3** shows a schematic correlation diagram between the energy levels in each case. The levels of the asymmetric rotor are usefully LABELLED by the number $\tau = K_A - K_C$ which is NOT a quantum number, but has the useful property that its value increases with increasing energy of the levels. Note that levels of a given $J$ do not cross.

The general problem of finding the allowed energies and transition intensities of the asymmetric rotor is solved by setting up the Hamiltonian matrix and the dipole transition matrix (the 'D' matrix) using the $|JKM\rangle$ functions as a basis. The Hamiltonian matrix is not now diagonal but by diagonalizing the Hamiltonian matrix we find the eigenvalues and eigenvectors of the Hamiltonian that we have used. The eigenvalues can be used to extract the relative intensities of all allowed transitions from the 'D' matrix. The magic of this approach is that we *do not need to know the eigenfunctions themselves*. The diagonalization provides the mixture coefficients for the basis that we have chosen ($a^J_K$) without us having to specify the basis functions apart from their

**Table 4** Energies of the rigid asymmetric rotor. Analytic expressions for the total rotational energy in terms of the rotational constants $A$, $B$ and $C$. As the rotational transitions involving low $J$ can often be observed, these relationships can be of considerable use in making a preliminary determination of molecular constants

| $J_{K_{\text{prolate}},K_{\text{oblate}}}$ | $E(A, B, C)$ |
|---|---|
| $0_{0,0}$ | 0 |
| $1_{1,0}$ | $A+B$ |
| $1_{1,1}$ | $A+C$ |
| $1_{0,1}$ | $B+C$ |
| $2_{2,0}$ | $2A+2B+2C+2[(B-C)^2+(A-C)(A-B)]^{1/2}$ |
| $2_{2,1}$ | $4A+B+C$ |
| $2_{1,1}$ | $A+4B+C$ |
| $2_{1,2}$ | $A+B+4C$ |
| $2_{0,2}$ | $2A+2B+2C-2[(B-C)^2+(A-C)(A-\mathrm{B})]^{1/2}$ |
| $3_{3,0}$ | $5A+5B+2C+2[4(A-B)^2+(A-C)(B-C)]^{1/2}$ |
| $3_{3,1}$ | $5A+2B+5C+2[4(A-C)^2-(A-B)(B-C)]^{1/2}$ |
| $3_{2,1}$ | $2A+5B+5C+2[4(B-C)^2+(A-B)(A-C)]^{1/2}$ |
| $3_{2,2}$ | $4(A+B+C)$ |
| $3_{1,2}$ | $5A+5B+2C-2[4(A-B)^2+(A-C)(B-C)]^{1/2}$ |
| $3_{1,3}$ | $5A+2B+5C-2[4(A-C)^2-(A-B)(B-C)]^{1/2}$ |
| $3_{0,3}$ | $2A+5B+5C-2[4(B-C)^2+(A-B)(A-C)]^{1/2}$ |
| $4_{4,0}$ | no analytic solution |
| $4_{4,1}$ | $10A+5B+5C+2[4(B-C)^2+9(A-C)(A-B)]^{1/2}$ |
| $4_{3,1}$ | $5A+10B+5C+2[4(A-C)^2-9(A-B)(B-C)]^{1/2}$ |
| $4_{3,2}$ | $5A+5B+10C+2[4(A-B)^2+9(A-C)(B-C)]^{1/2}$ |
| $4_{2,2}$ | no analytic solution |
| $4_{2,3}$ | $10A+5B+5C-2[4(B-C)^2+9(A-C)(A-B)]^{1/2}$ |
| $4_{1,3}$ | $5A+10B+5C-2[4(A-C)^2-9(A-B)(B-C)]^{1/2}$ |
| $4_{1,4}$ | $5A+5B+10C-2[4(A-B)^2+9(A-C)(B-C)]^{1/2}$ |
| $4_{0,4}$ | no analytic solution |
| $5_{4,2}$ | $10A+10B+10C+6[(B-C)^2+(A-B)(A-C)]^{1/2}$ |
| $5_{2,4}$ | $10A+10B+10C+6[(B-C)^2+(A-B)(A-C)]^{1/2}$ |

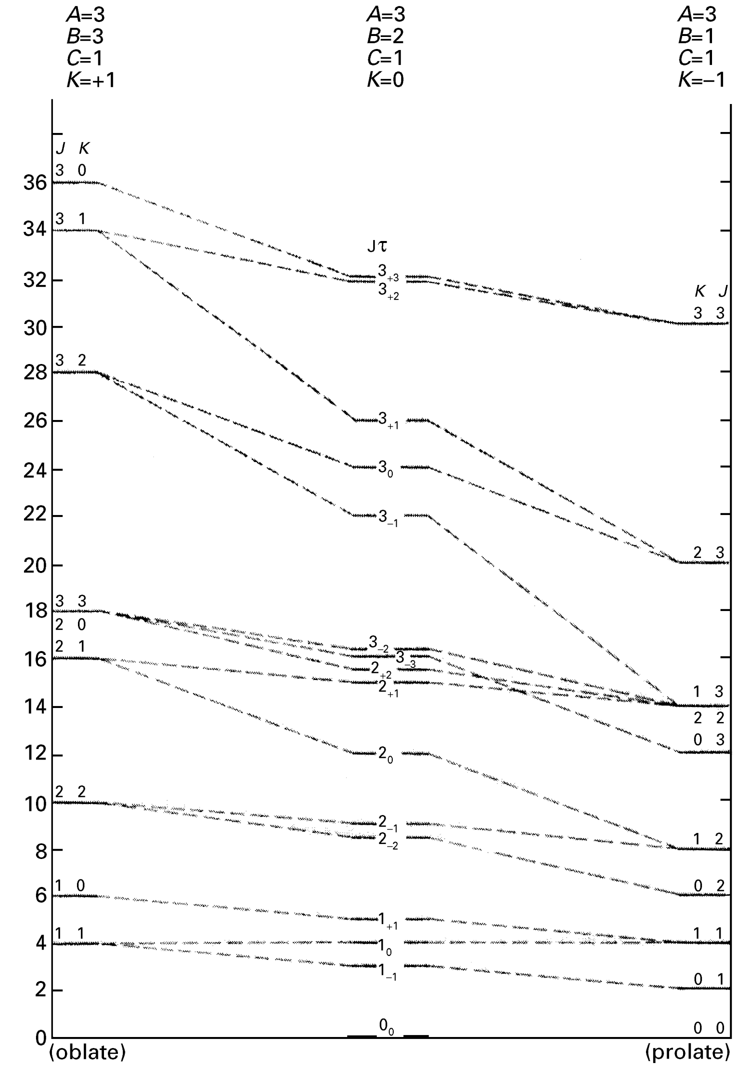



**Figure 3** Correlation between the energy levels of prolate, oblate and asymmetric tops. Reproduced from Figure 2.11 of Gordy W, Smith WV and Trambarulo RF (1953) *Microwave Spectroscopy*, p. 110. New York: Wiley.

properties under various quantum mechanical operators, i.e. their matrix elements. The properties of the basis functions are given in **Table 1** and in spherical tensor form in **Table 2**. For more general problems where the molecule (and therefore the Hamiltonian) includes electron or nuclear spins, it is not possible to write an eigenfunction which describes the property in three-dimensional space, but it is still possible to set up and diagonalize the Hamiltonian and D matrices.

## Vibration–rotation interaction

In some sense we can understand a spectrum if we can assign values of $A$, $B$, $C$ which generate the spectral line positions using the selection rules. The best test of this fit is to examine if the tested and predicted transitions agree with one another. A spectrum has not been correctly fitted unless the constants which have been determined can reproduce the experimental transition frequencies to within the experimental accuracy. From our initial study, we know that $B$ is a measure of $\langle 1/R^2 \rangle$, averaged over the particular electronic and vibrational state involved. In order to fit the experimental data accurately it is usually necessary to consider a more general energy level expression. This is achieved by adding terms to the Hamiltonian which allow for the interaction of vibration and rotation. When a molecule rotates, the centrifugal forces distort the molecule and this changes the moments of inertia. In a linear molecule the effect can be seen as a slight decrease in the spacing between successive rotational lines as $J$ increases. The effect for low $J$ lines is usually less than 1 part in $10^4$, but such changes are observable.

In linear molecules we account for the interaction by changing the rotational Hamiltonian to

$$H_{\text{rotation}} = B\boldsymbol{J}^2 - D\boldsymbol{J}^4 + H\boldsymbol{J}^6 + K\boldsymbol{J}^8 + L\boldsymbol{J}^{10} + \cdots$$

giving rise to the new eigenvalue expressions:

$$E_{\text{rotation}} = B[J(J+1)] - D[J(J+1)]^2 + H[J(J+1)]^3 + K[J(J+1)]^4 + L[J(J+1)]^5 + \cdots$$

As many terms as necessary are used in order to fit the spectrum. Unfortunately such power series expressions are heavily correlated—this manifests itself when say, $B$ and $D$ are determined if a further power is added, the values of $B$, $D$ and $H$ are now determined, but the values of $B$ and $D$ are not the same as before.

For nonlinear molecules, the Hamiltonian including the effects of vibration–rotation interaction is more complicated, but well understood.

## Molecules with electronic angular momentum

In this section we shall only consider the additional complications caused in diatomic molecules but similar considerations apply to more complicated molecules. So far we have considered molecules which have no angular momentum due to their electrons: that is no electronic angular momentum. The addition of spin angular momentum into the problem changes the energy level expressions and causes additional splittings of lines; the ground electronic state of $O_2$ is $^3\Sigma$ and therefore has two unpaired electron spins yielding a total spin of 1. Each rotational level with $J>1$ is now split into three and each line in the spectrum involving $J>1$ is split into six components, three of which are intense (hence the name 'triplet').

Orbital electronic angular momentum gives rise to non sigma electronic states, such as $^2\Pi$ states which have an orbital angular momentum of 1 and a spin angular momentum of $\frac{1}{2}$. NO is a molecule with a $^2\Pi$ ground electronic state and a simulated rotational spectrum of NO is shown in **Figure 4**.

One now has a choice of how to construct a basis set in which to set up the Hamiltonian, and the best basis is the one which most nearly diagonalizes the Hamiltonian matrix. However, nature does not care which basis we choose and a full calculation in any basis yields the same results; which basis is chosen is purely a matter of convenience. In diatomic molecules the choices of basis for the inclusion of spin most normally used were first considered by Hund, and are called Hund's cases *a* and *b* (see **Figure 5** for explanation).

We shall only consider Hund's case *a*. We are creating a basis function by coupling together the angular momenta in a particular order: $\boldsymbol{L}$ and $\boldsymbol{S}$ (the electronic and spin angular momentum of the electrons) are strongly coupled to the internuclear axis, and have projections $\Lambda$ and $\Sigma$ on this axis. The combination of $\Lambda+\Sigma = \Omega$. The angular momentum along the internuclear axis ($\boldsymbol{r}$) is $\Omega\boldsymbol{r}$ which is coupled with the angular momentum of the rotating nuclear

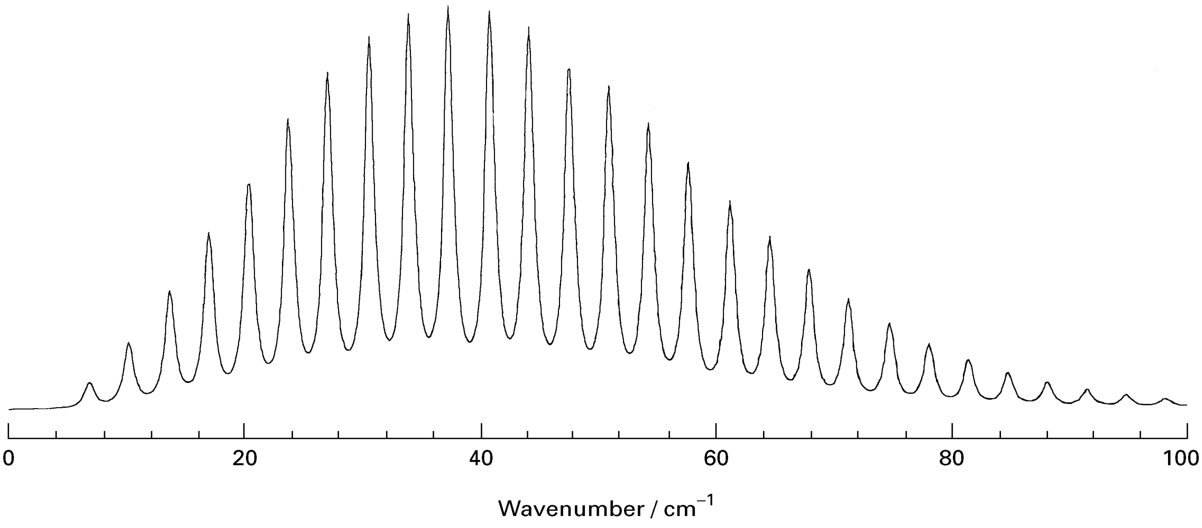



**Figure 4** Simulated rotational spectrum of the NO ground state ($X^2\Pi$) – note that the first rotational line is 'missing' due to the electronic angular momentum $\Lambda = 1$ ($\Pi$-state); such missing lines aid the assignment of electronic states.

framework (***R***) to give the resulting total angular momentum (excluding nuclear spins) of ***J***. In the case of the diatomic molecule, the projection quantum number $K$ is called $\Omega$ in Hund's *case a* and the eigenfunctions are labelled as $|J\Omega M\rangle$ rather than $|JKM\rangle$. The full eigenfunction includes the dependence on electron spin and can be written as the simple product function $|S\Sigma\rangle\,|J\Omega M\rangle = |S\Sigma;\ J\Omega M\rangle$. The fact that we can simply multiply the two functions together in this way (one for electron spin, one for resultant angular momentum) makes calculation of some matrix elements extremely simple. The total eigenfunction, including the electronic and vibrational state involved is written $|v, \Lambda, S\Sigma; JKM\rangle$.

#### Extra terms in the Hamiltonian

The inclusion of electron spin into the problem increases the complexity of the Hamiltonian which must be considered. In general, we must now add the following terms to the rotational Hamiltonian, some of which may be zero in a particular problem

$$\mathrm{H}_{\text{fine structure}} = \mathrm{H}_{\text{spin-spin}} + \mathrm{H}_{\text{spin-orbit}} + \mathrm{H}_{\text{spin-rotation}} + \mathrm{H}_{\text{lambda doubling}}$$

## Hyperfine interactions

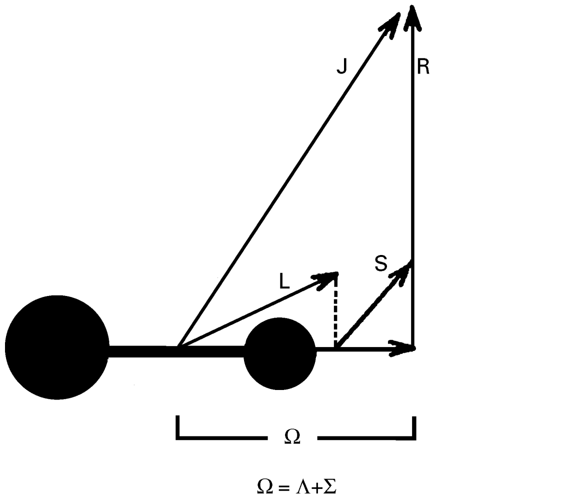
If the nuclei in a molecule have angular momenta $I_n > 1$ then complications in the spectra may arise due to the electric quadrupole field of the nuclei. Such additional structure is commonly observed and arises because there is an interaction between the electric field gradient along the internuclear axis (due to the electrons), the electric quadrupole moment of the nucleus, and the angular momentum of the molecule. The energy change due to the quadrupole moment can be written:

$$E_{\text{quadrupole}} = -eQqf(I, J, F)$$

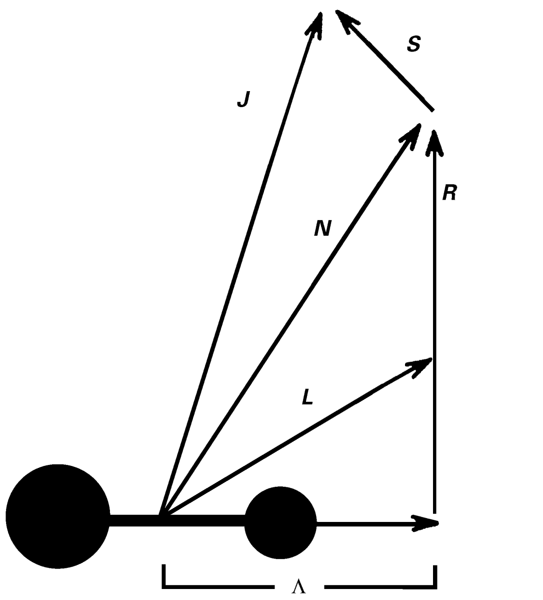
where $-eqQ$ is the nuclear quadrupole moment multiplied by the field gradient and $f\,(I, J, F)$ is an angular momentum coupling coefficient called Casimir's function. The effect of the quadrupole interaction on a spectral line is shown in **Figure 6**.

If a molecule has nonzero spin angular momentum, then magnetic hyperfine interactions may be observed between any nuclei with nonzero spin and between such nuclei and the electron spins. Such interactions cause the zero-field splittings in NMR.

The additional terms which must be considered due to magnetic hyperfine interactions are as follows:




Vector coupling diagram for Hund's case (a).




Vector coupling diagram for Hund's case (b).

**Figure 5** Vector diagrams for Hund's coupling cases *a* and *b*. Reproduced from Figures iii.17 and iii.18 of Carrington A (1973) *Microwave Spectroscopy of Free Radicals,* pp 129–130. New York: Academic.

$$\mathrm{H}_{\text{hyperfine}} = \mathrm{H}_{\text{fermi-contact}} + \mathrm{H}_{\text{spin-dipole, nuclear dipole}} + \mathrm{H}_{\text{nuclear dipole-nuclear dipole}}$$

there are other possible terms, but they have not commonly been necessary in interpreting the spectra which have been observed. Typically, hyperfine interactions have magnitudes which lead to splittings in spectra between 1–1000 MHz.

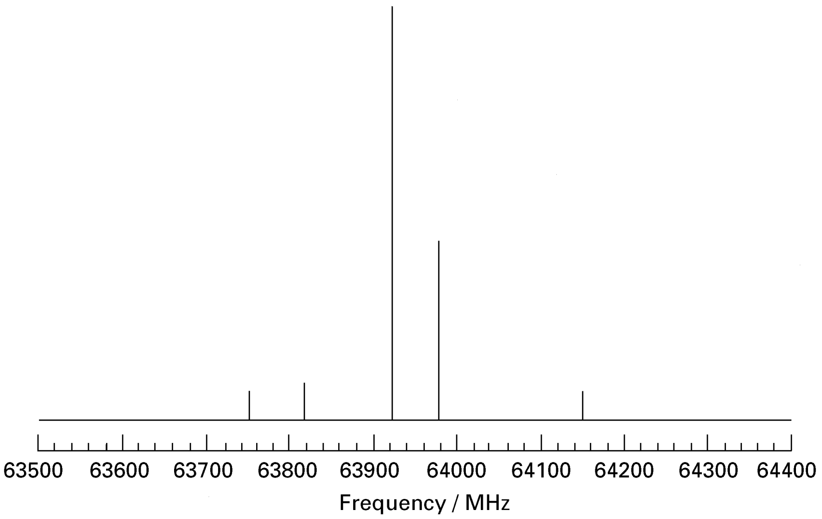



**Figure 6** Additional structure in a single rotational line ($J=3$–2) of BrF due to the quadrupolar constant $eqQ=909.2$ MHz of the bromine nucleus ($I=3/2$).

## Frequencies and Intensities of transitions

To calculate the relative intensity of a particular transition we use the interaction Hamiltonian between the dipole moment of the molecule and the electric vector of the radiation: $H_{\text{interaction}}=-\boldsymbol{E}.\boldsymbol{\mu}$. We need to calculate the matrix elements of this interaction using the eigenfunctions for the states of interest. In the simplest cases we have considered our eigenfunctions are the functions $|JKM\rangle$ and we need to calculate the matrix elements $\langle \nu, \Lambda, S\Sigma; JKM| -\boldsymbol{E}.\boldsymbol{\mu} |\nu, \Lambda, S'\Sigma'; J'K'M'\rangle$; the intensity of the transition is proportional to the square of this matrix element. The procedure is to expand the scalar product in space fixed axes using Equation [3] – both $\boldsymbol{E}$ and $\boldsymbol{\mu}$ are 1st rank tensors (vectors):

$$\boldsymbol{E}.\boldsymbol{\mu} = T^1(\boldsymbol{E}).T^1(\boldsymbol{\mu}) = \sum_p (-1)^p T^1_p(\boldsymbol{E}) T^1_{-p}(\boldsymbol{\mu})$$

and we insert this into the matrix element

$$\langle \nu, \Lambda, S\Sigma; JKM| - \sum_p (-1)^p T^1_p(\boldsymbol{E}) T^1_{-p}(\boldsymbol{\mu}) |\nu, \Lambda, S'\Sigma'; J'K'M'\rangle$$

as the electric field does not operate on the basis states, we can take the electric field outside the matrix element:

$$-\sum_p (-1)^p T^1_p(\boldsymbol{E}) \langle \nu, \Lambda, S\Sigma; JKM| T^1_{-p}(\boldsymbol{\mu}) |\nu, \Lambda, S'\Sigma'; J'K'M'\rangle$$

To evaluate the matrix element we need to transform the tensor for the dipole moment into molecule fixed coordinates, which we can do using Equation [13]

$$T^1_{-p}(\boldsymbol{\mu}) = \sum_q D^1_{-pq}(\omega)^* \, T^1_q(\boldsymbol{\mu})$$

which gives us:

$$-\sum_{p,q} (-1)^p T^1_p(\boldsymbol{E}) \langle \nu, \Lambda, S\Sigma; JKM| D^1_{-p,q}(\omega)^* \, T^1_q(\boldsymbol{\mu}) |\nu, \Lambda, S'\Sigma'; J'K'M'\rangle$$

Now $T^1_q(\mu)$ only operates on $\nu$, $\Lambda$ and $D^1_{-p,q}(\omega)^*$ only operates on $|JKM\rangle$, and neither operator effect $|S\Sigma\rangle$ so we can separate this out to:

$$-\sum_{p,q} (-1)^p T^1_p(\boldsymbol{E}) \langle \nu, \Lambda| T^1_q(\boldsymbol{\mu}) |\nu, \Lambda\rangle \times \langle JKM| D^1_{-p,q}(\omega)^* | J'K'M'\rangle \langle S\Sigma| S'\Sigma'\rangle$$

$T^1_p$ (**E**) is a constant number which tells us the intensity of the light and (through $p$) its polarization; $\langle S\Sigma|S'\Sigma'\rangle$ yields the condition that $S = S'$ and $\Sigma = \Sigma'$ if the expression is nonzero; $\langle \nu, \Lambda\,|\,T^1_q\ (\boldsymbol{\mu})\,|\,\nu, \Lambda\rangle$ is constant for a particular state, and so (finally) it is the matrix elements $\langle JKM\,|\,D^1_{-p,q}(\omega)^*\,|\,J'K'M'\rangle$ (which are the elements of the '$D$-matrix') which determine the relative intensities of the transitions and the selection rules. Using Equation [16] (**Table 2**) we evaluate the $D$-matrix elements to:

$$(-1)^{K-M}[(2J+1)(2J'+1)]^{1/2} \times \begin{pmatrix} J & 1 & J' \\ -K & q & K' \end{pmatrix} \times \begin{pmatrix} J & 1 & J' \\ -M & -p & M' \end{pmatrix}$$

The selection rules arise from the properties of the 3-j symbols which yield $\Delta J = 0,\pm 1$. The selection rule on $K$ depends upon which components of the dipole moment are nonzero but, as $q = -1,0,1$ only, then $\Delta K = 0,\pm 1$. The selection rule on $M$ depends on the polarization of the light. For $z$ polarized light $p = 0$ and $\Delta M = 0$. For nonpolarized light $\Delta M = 0,\pm 1$.

In more complicated cases, where the eigenfunctions are linear combinations of the basis functions, the relative intensities are found by multiplying the $D$ matrix by the eigenfunction coefficients.

The frequencies of transitions are found by setting up the requisite Hamiltonian matrix in the same basis and diagonalizing it, evaluation of the matrix elements of the Hamiltonian matrix is accomplished in the same fashion as for the example above. The eigenvectors and eigenvalues found together with the $D$-matrix enable a complete calculation of the spectrum to be made.

## Summary

We have considered the theory necessary to understand rotational spectroscopy, concentrating on some simple cases. In a first approximation, spectra can be understood on the basis of rigid rotor models. Allowing for the fact that molecules are not rigid complicates spectra and introduces new molecular constants. We have seen that when electronic spin and orbital angular momenta are included in the picture we need to extend our basis sets, and the utility of spherical tensors has been illustrated in calculating expressions which yield the relative intensities of rotational transitions.

## List of symbols

$a_K{}^J$ = mixture coefficients; $B$ = rotational constant; $D$ = dipole transition matrix; $-eqQ$ = nuclear quadrupole moment; $E$ = energy; **E** = electric vector; $f(I,J,F)$ = angular momentum coupling coefficient (Casimir's function); $F$ = total angular momentum quantum number; $g$ = degeneracy; $H$ = Hamiltonian; $I$ = moment of inertia; $I$ = intensity; $I_0$ = incident intensity; $J$ = rotational quantum number; **L**, **S** = electronic and spin angular momenta; $M,K$ = projection quantum numbers; $R$ = bond length; $S(\omega, \omega_0)$ = normalized line shape function; $\gamma(\omega)$ = absorption coefficient; $\mu$ = reduced mass; $\psi$ = wavefunction.

*See also:* **IR Spectroscopy, Theory; Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Spectroscopy of Ions; Symmetry in Spectroscopy, Effects of.**

## Further reading

Bernath PF (1995) *Spectra of Atoms and Molecules*, Chapter 6. New York: Oxford University Press.

Bunker PR (1979) *Molecular Symmetry and Spectroscopy*. London: Academic Press.

Dixon RN (1965) *Spectroscopy and Strucure*, Chapter 3. London: Methuen and Co.

Edmonds AR (1985) *Angular Momentum in Quantum Mechanics*, 2nd edn. Princeton: Princeton University Press. (contains tables of 3-j and 6-j coefficients).

Gordy W and Cook RL (1984), *Microwave Molecular Spectra; Techniques of Chemistry Volume XVIII* 3rd edn. New York: Wiley Interscience.

Herzberg G (1991) *Molecular Spectra and Molecular Structure; Volume I: Spectra of Diatomic Molecules*. Florida: Krieger Publishing Company.

Herzberg G (1991) *Molecular Spectra and Molecular Structure; Volume II: Infrared and Raman Spectra of Polyatomic Molecules*. Florida: Krieger Publishing Company.

Hirota E (1985) *High-Resolution Spectroscopy of Transient Molecules*. Berlin: Springer-Verlag.

Kroto HW (1991) *Molecular Rotation Spectra*. Toronto: Dover.

Townes CH and Schawlow AL (1975) *Microwave Spectroscopy*. New York: Dover.

# S

# $^{29}$Si NMR

**Heinrich C Marsmann**, Universität Paderborn, Germany

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

All group 14 elements – except germanium – share, from the viewpoint of the NMR spectroscopist, one rather important feature: they all have at least one isotope with a spin of $\frac{1}{2}$ that is a minor component of the isotopes of the element. The most important element in this group is, of course, carbon and most people are familiar with obtaining and interpreting 13C NMR spectra. Consequently a comparison with the situation of carbon might be helpful (**Table 1**).

An inspection of **Table 1** shows that $^{29}$Si has a higher share of the isotopic mixture but the absolute value of the magnetic moment is slightly lower than that of $^{13}$C. This leads to a lower resonance frequency. A complication arises from the fact that the spin and magnetic moment are antiparallel, leading to negative sign of $\gamma$.

The chemistry of silicon is very different to that of carbon and in organic compounds there is usually a lower silicon content than carbon. These shortcomings meant that silicon NMR spectroscopy had a slow start. After the first report by Lauterbur and coworkers in 1962 there had been only a few papers per year. But this has now changed dramatically and one collection of $^{29}$Si chemical shifts now contains about 9000 data sets. One area of special growth is $^{29}$Si NMR measurement on solids. Here applications range from crystalline silicates over amorphous silica gels to silicon-containing surfaces and coatings.

**Table 1** Comparison of carbon and silicon isotopes.

| | *Natural abundance (%)* | *Spin* | *γ*$^a$ | *ν*$^b$ (MHz) | *Other isotopes* |
|---|---|---|---|---|---|
| $^{13}$C | 1.108 | 1/2 | 6.7263 | 25.145004 | $^{12}$C (98.89%) |
| $^{29}$Si | 4.70 | 1/2 | –5.3141 | 19.867184 | $^{28}$Si (92.21%)<br>$^{30}$Si (3.09%) |

$^a$ Gyromagnetic ratio.
$^b$ In a magnetic field, where the nuclei of TMS resonate at 100 MHz.

## Measuring $^{29}$Si NMR

### Referencing

The established standard compound to calibrate $^1$H and $^{13}$C spectra is tetramethylsilane $(CH_3)_4Si$ (TMS). As this substance contains silicon also, it is natural to use it as standard for Si NMR. Although TMS is an inert substance, has a low boiling point and a rather short relaxation time, its chemical shift is in the middle of the shift range of other organosilicon compounds and peak misassignments are possible. Two strategies to circumvent this problem have been developed. The first is the use of secondary standards. **Table 2** lists a number of those used in the literature.

$M_8Q_8$, shown in **Table 2**, has a lower symmetry in the solid state and its resonances are split into several lines, but it is the common reference compound used as an internal standard for solid state Si NMR. The second tactic is to use no standard compound in the sample at all. The referencing is done externally relative to a separate sample that contains TMS in the same solvent as is used in the analysis sample.

Early reports of silicon NMR data employ the magnetic field definition of chemical shifts instead of

**Table 2** Common reference compounds for silicon NMR.

| *Formula* | *Abbreviation* | *Shift relative to TMS* (ppm) | | |
|---|---|---|---|---|
| $(CH_3)_4Si$ | TMS | 0.00 | | |
| $[(CH_3)_3Si]_4C$ | | 3.6 | | |
| $[(CH_3)_3Si]_2$ | HMDS | −19.79 | | |
| $[(CH_3)_3Si]_2O$ | $M_2$ | 7.22 | | |
| $[(CH_3)_2SiO]_4$ | $D_4$ | −19.86 | | |
| $(CH_3O)_4Si$ | TMOS | −78.22 | | |
| $(C_2H_5O)_4Si$ | TEOS | −81.65 | | |
| $[(CH_3)_3SiO]_4Si$ | $M_4Q$ | 8.62, −104.08 | | |
| $[(CH_3)_3SiOSiO_{3/2}]_8$ | $M_8Q_8$ | 12.4,[a] – 108.6[a] | | |
| | | 11.77,[b] | 11.72,[b] | 11.51[b] |
| | | −108.36,[b] | −108.64,[b] | |
| | | −109.36,[b] | −109.71[b] | |

[a] In solution.
[b] In the solid state.

the now universally accepted frequency based one, resulting in a reversed sign for chemical shift data.

## Pulse sequences

In principle there are two cases to distinguish: many, especially inorganic, compounds contain only $^{29}$Si as a useful magnetic nucleus. Here, single pulse experiments are applicable only. Because of the rather slow relaxation in such compounds, a 30° pulse is used with a repetition rate of about 20 s.

On the other hand, silicon compounds with organic side-groups usually have resonances split into many lines by spin–spin couplings with the protons. These are normally removed by decoupling. Because of the negative gyromagnetic ratio of $^{29}$Si the nuclear Overhauser effect (NOE) is also negative, leading to negative signals with an enhancement factor of $\eta_0 = -2.52$ if the relaxation is dominated by dipolar interactions with the protons. With few exceptions, the relaxation path is divided between the dipolar term and others – mostly the spin–rotation interaction. This results in much lower enhancement factors and even null signals. This can be overcome by:

1. Doping of the sample with a shiftless relaxation agent such as $Cr(acac)_3$ at a concentration of $\sim 10^{-2}$ mol dm$^{-3}$. However, such a compound might react with the sample and it is also difficult to remove.
2. Inverse gated decoupling. Here the proton decoupler is switched off during a recovery time (three to five times the silicon relaxation time). The advantage of not contaminating the sample is offset by an ineffective use of spectrometer time.
3. The use of pulse programmes such as insensitive nuclei enhanced by polarization transfer (INEPT) or distortionless enhancements by polarization transfer (DEPT). There are two advantages in using these programmes. The first is a signal enhancement by population transfer from the protons that depends on the number of coupling protons. A good review has been given by Blinka *et al.* The second advantage is the removal of the background signal caused by the silicon content of the probe. However, these techniques depend on the proton–silicon coupling constant and as these can cover a considerable range in organosilicon compounds this can lead to unexpected results. An example is shown in **Figure 1** where a $^{29}$Si INEPT spectrum of an octasilsesquioxane is displayed. Because the siloxane skeleton consists of a cube with silicon on the vertices and oxygen on the edges of the cube, the cage can only rotate as a whole. This way the silicon relaxes almost solely by the spin–dipole path. The relaxation times are rather long and the signals very narrow.

The NMR of solids differs from that of liquids insofar as spatial interactions are not averaged out by Brownian motion. The NMR resonances are therefore so broad that chemical shifts and indirect spin–spin couplings are not resolved. There are different ways of handling solid state NMR problems. Here is not the place to cover all of them. The most common presently used method starts with a powder of crystalline or amorphous silicon compound. In the absence of nuclei with a strong magnetic moment such as $^{1}$H or $^{19}$F or ions with unpaired electrons, the source of broadening of the signals comes from the fact that the chemical shift is a tensor. This chemical shift anisotropy (CSA) can be removed by fast spinning (3000–15 000 Hz) around an angle of 54°44′ (magic-angle spinning, MAS). The CSA is then reduced to a series of sidebands spaced at the rotational frequency and with an intensity profile governed by the CSA. By rotating the sample fast enough, these sidebands can be moved out of the region of the chemical shifts of the sample. On the other hand, by rotating slowly, the CSA can be determined. The *direct* spin–spin coupling between $^{29}$Si nuclei is small because the magnetic moment is small and the average distance between $^{29}$Si nuclei is large because the abundance of $^{29}$Si is just 4.7%. For most crystalline silicates, single resonance NMR spectroscopy under MAS is the best choice. Depending on the quality of the crystals, narrow peaks with line widths between 0.1 to 3 ppm can be obtained. In addition to single pulse spectroscopy, incredible natural abundance double quantum transfer experiment (INADEQUATE) or 2D correlation spectroscopy (2D-COSY) can be carried out to determine the connectivity of the silicon atoms.

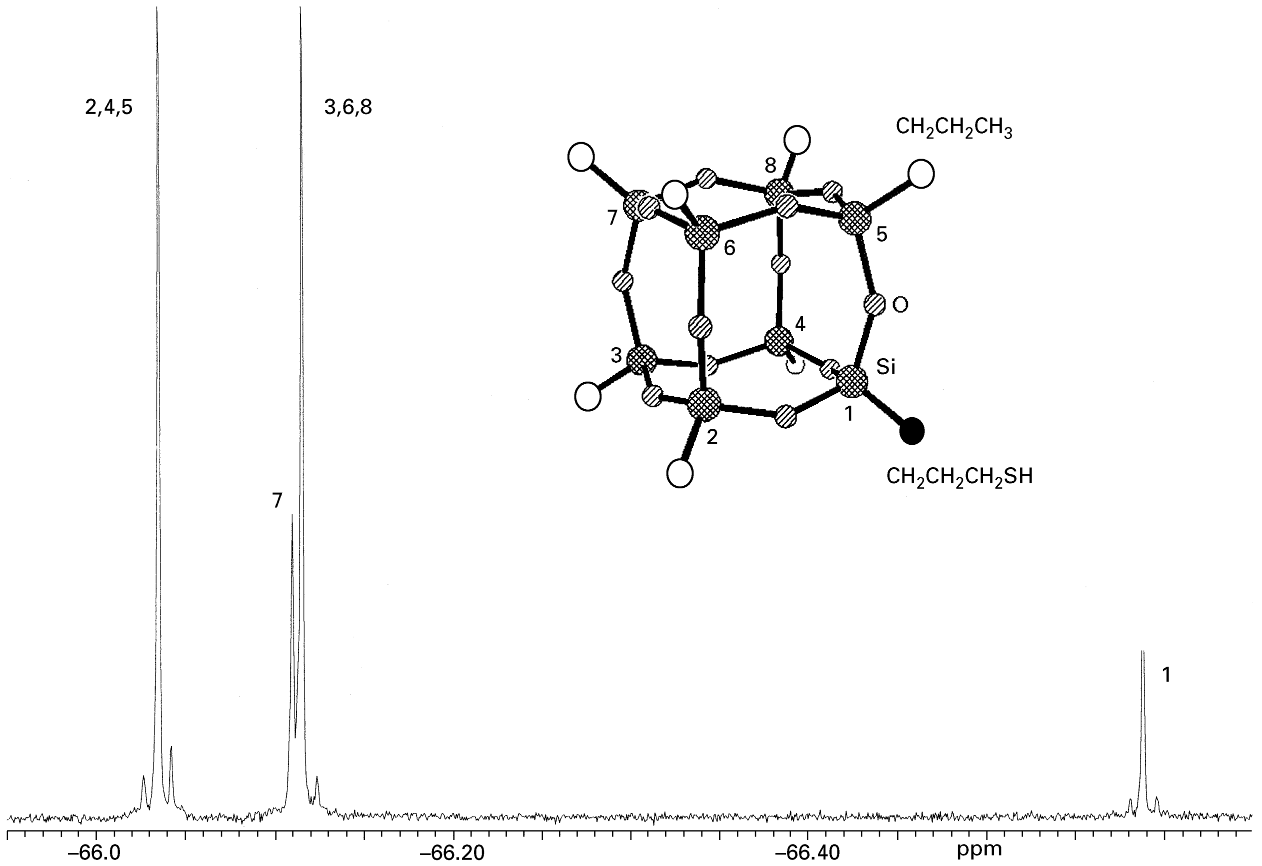


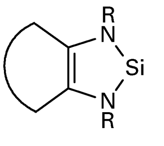

**Figure 1** $^{29}$Si NMR spectrum of (3-mercaptopropyl)hepta(propyl)octasilsesquioxane. The small signals are due to $^{29}$Si–$^{29}$Si coupling.

Organosilicon compounds contain hydrogen and the strong magnetic moment of $^{1}$H leads to very broad lines in the solid state by *direct* proton–silicon spin–spin interaction. This coupling is usually removed by high power decoupling of the protons. The spin coupling with protons offers the opportunity to obtain stronger signals by the cross-polarization (CP) technique. The connection between the proton and silicon spins is possible if both spins are rotating with the same speed around the magnetic field, i.e. $\gamma_H B_{1H} = \gamma_{Si} B_{1Si}$ (spin locking, Hartmann–Hahn condition). The most effective time for this spin locking is determined by the proton–silicon distance. Because of the higher incidence of protons the $^{1}$H relaxation time is usually shorter, so that beside the inherently stronger signals the pulse repetition rate can be faster.

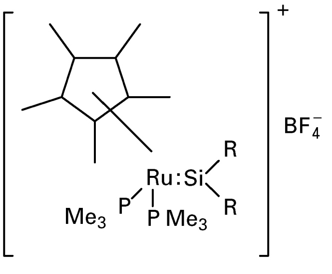
## $^{29}$Si chemical shifts

### General features

In principle, two kinds of silicon compounds are possible. The first kind, derived from divalent silicon [Si(II)] are normally thermodynamically unstable. Some exist at high temperatures or may be trapped using cryogenic temperatures. However, a few are stable enough to be handled at room temperature, but the small number of examples does not allow a general definition of a specific range of chemical shifts. Nevertheless, derivatives with a high coordination number are strongly shielded. Thus, for instance, decamethylsilicocene, $[(CH_3C)_5]_2Si$, has a shift of −398 ppm, the highest shielding measured so far. Another situation where Si(II) seems to be stable is between two nitrogen atoms, as in [1]. The silicon is then rather deshielded with values of the chemical shift between +78.3 and +96.9 ppm.




[1]




[2]

The majority of the $^{29}$Si NMR data reported involve derivatives of Si(IV). The most deshielded silicon (+268.7 ppm) has been reported for [2], the most shielded for $SiI_4$ (–351.7 ppm), but the majority of shifts are found between –200 and 150 ppm. As with the other heavier nuclei, the chemical shift depends primarily on the coordination number of the silicon in such a way that a low number of substituents around the silicon leads to deshielding (e.g. disilylenes: 13 to 175 ppm) and a high coordination number gives high negative numbers (e.g. sixfold coordination: –198 to –135 ppm). The regions of chemical shifts for some important classes of silicon compounds are depicted in **Figure 2**.

There have been several proposals to rationalize silicon chemical shifts. Most of them centre on the charge on the silicon atom. Ernst and co-workers observed that the chemical shifts follow a parabolic curve if plotted against the sum of the electronegativities of the substituents of the silicon. Thus, above the sum of electronegativities of 9.5, a further increase of the electronegativity leads to an increase of the shielding. However, a decrease of shielding is observed if this sum is below 9.5.

Radeglia and co-workers consider that not only the charge on the silicon determines the shielding but also its distribution. This is due to the paramagnetic contribution to the shielding constant. They define a reduced shielding constant $\sigma^*$ on the basis of the valence shell p orbitals of the silicon:

$$\sigma^* = \frac{\sigma_{\mathrm{P}}}{\sigma_{\mathrm{P}_0}} = \frac{P_{\mathrm{u}}\langle r^{-3}\rangle_{\mathrm{P}}}{P_{\mathrm{u}}\langle r^{-3}\rangle_{\mathrm{P}_0}} = P_{\mathrm{u}}^* R^* \qquad [1]$$

where the index 0 pertains to a situation in which all electrons are distributed evenly around the silicon. The terms $P_{\mathrm{u}}^*$ and $R^*$ are then calculated to various degrees of sophistication. One approach uses Slater-type orbitals and Slater rules; $R^*$ is then given by:

$$R^* = \frac{\langle r^{-3}\rangle_{\mathrm{P}}}{\langle r^{-3}\rangle_{\mathrm{P}_0}} = \left(1 + 0.35\frac{f q_{\mathrm{Si}}}{Z_0(\mathrm{Si})}\right)^3 \qquad [2]$$

with $Z_0(\mathrm{Si}) = 4.15$, $f = 0.2135$ as an empirical factor and $q_{\mathrm{Si}}$ stands for the charge on the silicon atom:

$$q_{\mathrm{Si}} = 4 - \sum_{i=\mathrm{A}}^{\mathrm{D}} h_i \qquad [3]$$

The bond polarity $h_i$, calculated from their electronegativities, is summed up over all four substituents (A to D) on the silicon. The term $P_{\mathrm{u}}$ takes elements of the bond order matrix around the silicon atom into account, using the electronegativity as a measure of electron distribution:

$$P_{\mathrm{u}}^* = \frac{P_{\mathrm{u}}}{P_{\mathrm{u}_0}} = \frac{1}{2}\sum_i h_i - \frac{1}{6}\sum\sum_{i>j} h_i h_j \qquad [4]$$

The $\sigma^*$ values so calculated do not give chemical shifts directly but have to be fitted empirically. Nevertheless, it is remarkable that the shape of the curve of $^{29}$Si chemical shifts as a function of the number of substituents (**Figure 3**) is reproduced by such calculations. The literature and a discussion of this procedure and other aspects of silicon chemical shift interpretation has been given by Marsmann (see Further reading section). Empirical aspects of chemical shifts of different classes of some silicon compounds only are given briefly below.

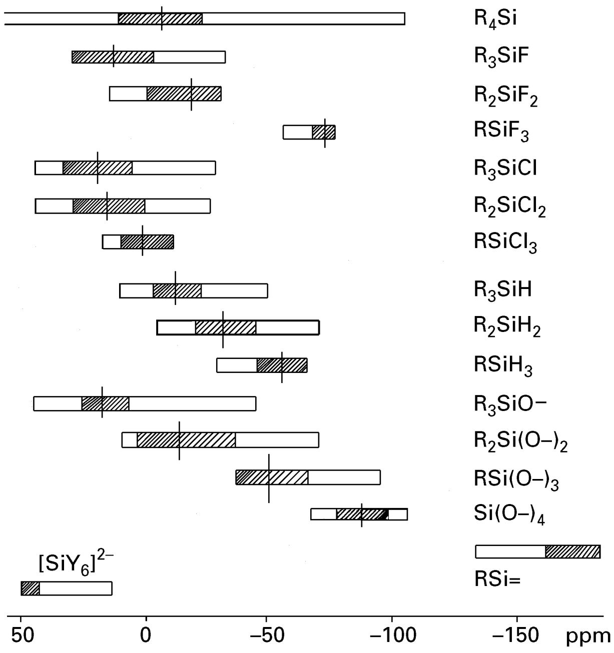



**Figure 2** Ranges of chemical shift of selected classes of silicon compounds. (R is an arbitrary organic ligand, Y = F or OR). See also **Figures 5**, **6** and **8**.

## Organosilanes

Silicon surrounded by four carbon atoms gives resonances between ~+100 and –107 ppm. As expected, shifts of silicon with aliphatic substituents are clustered around 0 ppm. Increased shielding is found if the ligands around the silicon atom have $\pi$ bonds in the $\beta$ position and thus the centre of shifts is moved

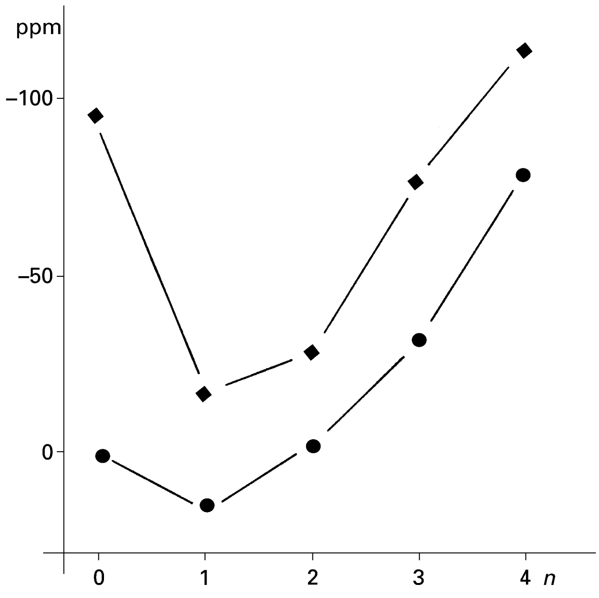
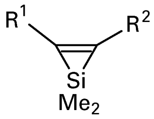
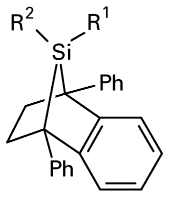



**Figure 3** The effect of successive substitution on $^{29}$Si NMR chemical shifts. (●) $(CH_3)_{4-n}Si(OCH_3)_n$; (◆) $H_{4-n}SiF_n$.

to −35 ppm if silicon is connected to four $sp^2$-hybridized carbons. Very high shielding is observed for silicon in silacyclopropanes (∼−60 to −40 ppm) and silacyclopropenes [3] (−87 to −106 ppm). The influence of $\pi$ orbitals is discussed in the later case. The strong deshielding for the bridge silicon in silanorbornadienes [4] (76–98 ppm) is also attributed to a $\sigma$–$\pi$ interaction.



[3]



[4]

## Siloxanes and silicates

One of the main interests in silicon NMR is because it allows distinction between the building units of a polysiloxane. A sketch of an imaginary polysiloxane to make the concept of building units and their nomenclature clearer is given in **Figure 4**. If the substituent R is not a methyl group, appropriate notation for that substituent is added as a superscript. A collection of chemical shift ranges of such building units has been published by Williams and is shown here as **Figure 5**.

Ring strain has a marked effect on the shifts. This can be seen for example, with dimethylsiloxanes in (**Table 3**). In polysiloxanes with different middle

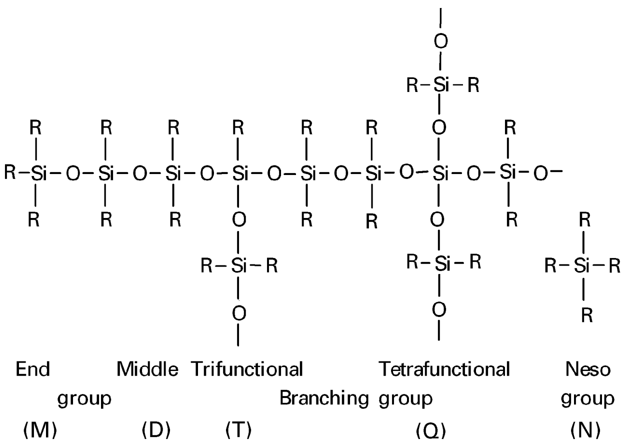



**Figure 4** Building units of a hypothetical polysiloxane.

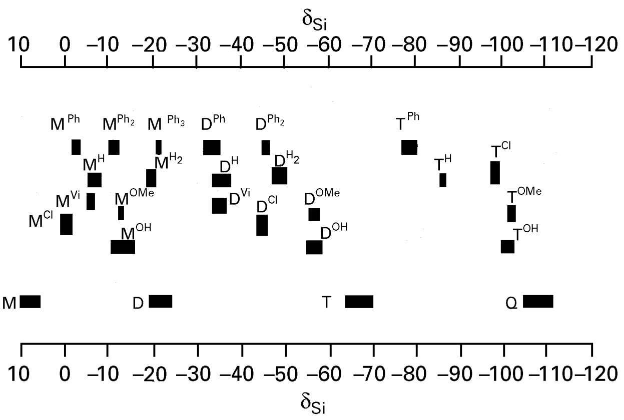


**Figure 5** Shift ranges of building units in polysiloxanes. Reproduced with permission of John Wiley & Sons from Williams EA (1989) NMR spectroscopy of organosilicon compounds. In: Patai S and Rappoport Z (eds) *The Chemistry of Organic Silicon Compounds*. New York: John Wiley & Sons.

groups, it is possible to distinguish the triad and even the pentad structures in the backbone of the polymer.

Approximately the same behaviour, but with a smaller range, is observed if R = OH or $O^-$ in silicates. However, as they are now all Q units, they are distinguished by a superscript $n$ ($n$ = 0–4), indicating the number of non-connecting oxygen atoms around the silicon. For instance, $Q^2$ designates a middle group with two connecting and two non-connecting

**Table 3** $^{29}$Si chemical shifts of dimethylsiloxanes.

| $n$ in $[(CH_3)_2SiO]_n$ | $\delta$ (ppm relative to TMS) |
|---|---|
| 3 | −9.12 |
| 4 | −19.51 |
| 5 | −21.93 |
| 6 | −22.48 |

oxygens. In aqueous solution there are fast equilibria between ionized and unionized forms. This causes a dependency of the chemical shifts of the building units on the pH. As the acidity of the building units differs and the existence of individual silicate ions is pH-dependent, this makes the analysis of such aqueous silicate solutions very difficult and some resonances are still not assigned. To make things a little more manageable, the signal of the monosilicate $Si(OH)_4$ or its corresponding ions are taken as a secondary shift standard. A rough conversion into the TMS scale can be obtained by adding −71.3 ppm. The use of highly $^{29}$Si-enriched material proved to be very successful and a number of individual ions could be identified by selective decoupling and $^{29}$Si, $^{29}$Si COSY experiments. It was found that silicate ions in solution consist mostly of small fused rings. A collection of the shift ranges of acyclic silicic acid ions and esters is given in **Table 4**. Ring strain again leads to deshielding, so that, for example, the signal of the middle groups of trimeric rings can be found close to the end groups and that of their trifunctional branching close to the region of middle groups. The use of Si NMR of solid silicates provides valuable information. The isotropic shifts in silicates follow similar rules to aqueous silicates. The low CSA value of <100 ppm makes the MAS spectra of silicates easy to observe. There are, however, some differences. In crystalline silicates, the shifts also depend on the crystallographic symmetry, which can be lower than in the dissolved state. The most famous example is that $M_8Q_8$, which gives two sharp lines in solution but splits into seven signals in the solid (**Table 2**). The Si–O–X angle dependence of shifts – not observable for solutes – is responsible for the spread of ∼30 ppm for the disilicate from ∼−70 to −100 ppm compared with a region of about 4 ppm for the end groups in aqueous silicates. It is possible to determine the Si–O–X angles on the basis of chemical shifts. One of the triumphs of $^{29}$Si MAS NMR spectroscopy was the determination of the distribution of silicon and aluminium in aluminosilicates, e.g. zeolites. Silicon chemical shifts move ∼5 ppm to higher field if an Si–O–Si bond is exchanged for an Si–O–Al, as can be seen from **Figure 6**. Although it is not possible to disprove a statistical distribution of Si and Al in the framework, it seems that the Al sites are highly ordered and that the Loewenstein rule is observed, i.e. that no direct Al–O–Al connections in aluminosilicates are stable. Another field for $^{29}$Si MAS NMR is the characterization of glasses, silica-gels and their organically modified surfaces which are used in chromatography or catalysts. Here the silicate framework is amorphous, i.e. although each silicon is surrounded by four oxygen atoms there is no long-range order. Compared with the signals of well-crystallized silicates this kind of material gives very broad signals with almost no features. In silica-gels and similar materials the non-connecting oxygens are in the form of OH or OR group. This allows CP MAS NMR spectra to be obtained. **Figure 7** shows a CP MAS spectrum of a silica-gel with an organically modified surface.

**Table 4** Shift ranges (ppm) of some classes oligomeric silanes.

| X | $X_4Si$ | $X_3Si$– | $X_2Si<$ | XSi– (trivalent) | –Si– (tetravalent) |
|---|---|---|---|---|---|
| H | −95.6 | −107 to −67 | −124 to −79 | −161 to −53 | −165 |
| $CH_3$ | 0.0 | −2 to +28 | −49 to −20 | −88 to −78 | −118 to −136 |
| Cl | −18.5 | −10 to +20 | −10 to +15 | −46 to −15 | −79 to −81 |

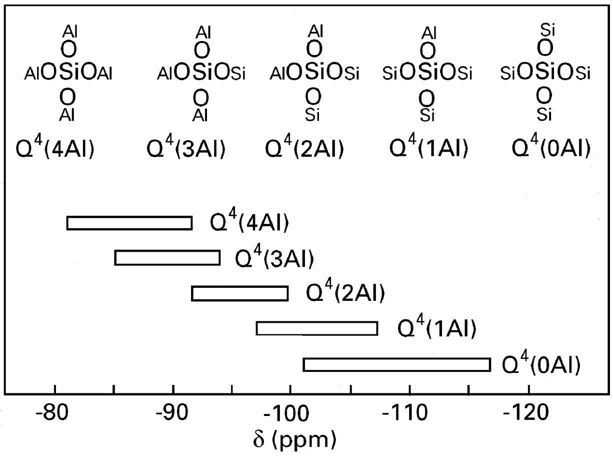



**Figure 6** Chemical shifts of the building units in aluminosilicates. Reproduced with permission of John Wiley & Sons from Engelhardt G and Michel D (1987) *High Resolution Solid-State NMR of Solids and Zeolites*. New York: John Wiley & Sons.

## Trimethylsilyl derivatives

The $(CH_3)_3Si$– moiety is a very useful fragment in several kinds of chemistry. It is used in organic synthesis because it can direct reactions along a particular path or act as a protecting group. In addition, it can be split off by gentle methods. The exchange of the protons in OH or >NH groups for a $(CH_3)_3Si$– group increases the volatility and the solubility in organic solvents. This has many applications in analytical chemistry. Schraml showed that the trimethylsiloxy group is useful as a NMR tag for the analysis of natural products such as sugars, lignins, etc. Therefore, silicon chemical shifts for these types of compounds are very common in the literature. About a quarter of all reports of silicon NMR shifts deal with them. The regions of chemical shifts for trimethylsilylated compounds are given in **Figure 8**.

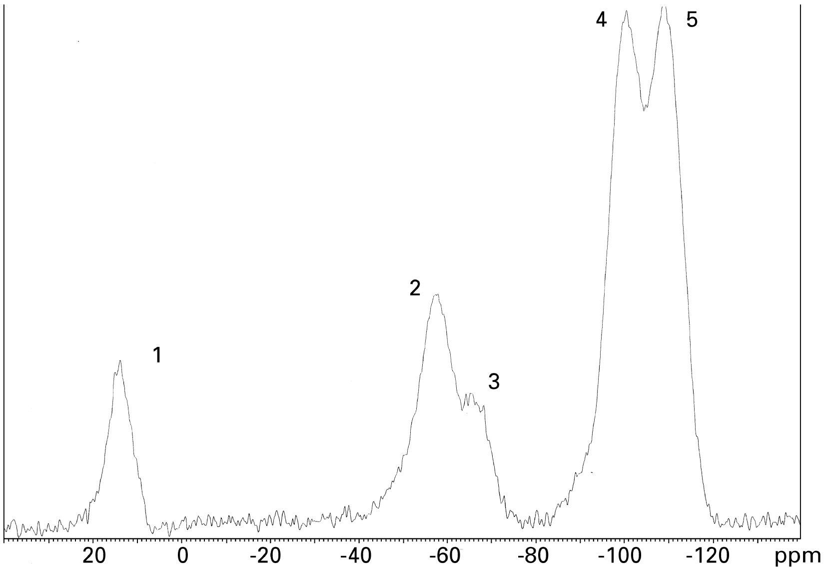



**Figure 7** $^{29}$Si MAS CP spectrum of a surface modified silica-gel. The signal strength is no indication of the concentration of the corresponding building group. Peak 1: ($CH_3SiO$–; peaks 2,3: $T^2$,$T^3$ of the functional group RSi(O–)$_3$; peaks 4,5: $Q^3$,$Q^4$ of the silica-gel skeleton.

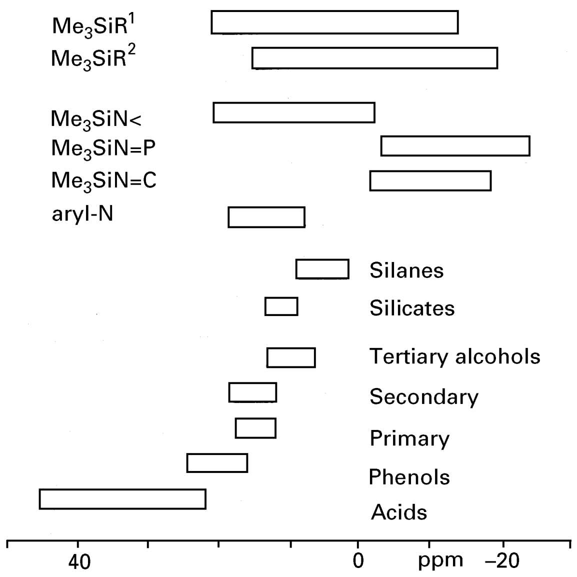



**Figure 8** Regions of chemical shifts for trimethylsilylated compounds ($R^1$ = aliphatic $R^2$ = unsaturated organic residue).

The trimethylsilyl group bound to a carbon atom has relatively clear cut regions. If the carbon is an aliphatic one, most of the resonances appear between −1.6 and +3.8 ppm. If the carbon is substituted with electronegative elements, more positive values are possible. A sp$^2$-hybridized carbon results in shifts to higher fields and the resonances are now clustered around −3.8 ppm. The silicon is more shielded if the carbon is a ring member (∼−11 ppm).

In $(CH_3)_3SiN$– derivatives the spread of chemical shifts is small if the nitrogen is aliphatic (−3.3 to 17.3 ppm), while it is more diverse if the nitrogen has a double bond. If the nitrogen is in an aromatic ring system, a compact region with positive values is discernible (3 to 18 ppm). However, no clear pattern emerges for other environments of this type.

A wealth of data exist for the trimethylsiloxy group. The greatest spread of data is found for the trimethylsilyl esters of acids. There is a linear correlation between the silicon chemical shift and the strength of the acid. The overlapping regions for the shift values of alcohols makes further experiments necessary for an assignment. One procedure is the use of hydrogen–silicon couplings to the non-methyl part of the molecule. A review has been given by Schraml.

## Polysilanes

An extensive review of this area has been presented by Williams. The structural principles of the polysilanes are the same as those of the alkanes. Thus, the rules applied for alkanes were extended to the silanes. This approach after Paul and Grant, was used by West and Stanislawski to derive Equation [5] for dimethylsilanes.

$$\delta_{Si(k)} = B + \sum_{l} A_l n_{kl} \qquad [5]$$

with $A_l$ the chemical shift parameter for the $l$th atom from the $k$th position and $n_{kl}$ the number of atoms present in that position. The term $B$ is constant (8.5 ppm) and $A_l$ is −25.8 ± 0.1 ppm for the $\alpha$ position, +3.9 ± 0.05 ppm for the $\beta$ position, +1.2 ± 0.09 ppm for the $\gamma$ position and +0.2 ± 0.01 ppm for the $\delta$ position relative to the silicon atom considered. This works well for end, middle and trifunctional branching groups but not so well for tetrafunctional branching and otherwise substituted silanes. Hahn used a similar concept to describe the shifts of silanes:

$$\delta_{\mathrm{Si}} = -96.02 - 2.43a - 3.73a^2 + Bb + C \quad [6]$$

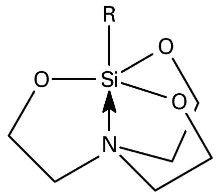
where −96.02 represents the idealized shift of monosilane, $a$ the number of silicon atoms in the $\alpha$ position, $b$ those in the $\beta$ position, and $B$ and $C$ are the incremental shifts produced by the $\beta$ and $\gamma$ silicon atoms.

## Silicon with a coordination number of three

A coordination number of three is found for compounds in which the silicon extends a $p_\pi$–$p_\pi$ bond to a partner. The usual condition is that the two other ligands to the silicon are sterically demanding, such as a phenyl group with substituents in the *ortho* position. The assumption of a $p_\pi$–$p_\pi$ bond in disilylenes is supported strongly by the chemical shift anisotropy data obtained from solid-state silicon NMR measurements which are comparable to that of a $p_\pi$–$p_\pi$ carbon bond. Although little data exist for this class of silicon compound it seems that all of them are deshielded compared with TMS. At present, compounds with a Si=C fragment have shifts between 13 and 144 ppm, where a silicon attached to the carbon leads to the higher numbers. Compounds with a Si=Si moiety resonate between 52 and 65 ppm and in a Si=P fragment the silicon resonances are found between 148 and 176 ppm. The spread of shifts is surprisingly low for simple transition metal derivatives of silylenes. Shift values range from 83.6 ppm for $(t\text{-}C_4H_9S)_2Si\ Fe(CO)_4$ to 142.1 ppm for $ArHSi{=}MnCp(CO)_2$.

## The effect of higher coordination numbers

Although the normal coordination number of silicon is four, the atom still acts as a Lewis acid, especially if the silicon is bound to electronegative ligands and hence coordination numbers of five and six are possible. However, because of equilibria it is not always possible to establish the degree of higher coordination in solution. In this respect, couplings to other magnetic nuclei are of help. Thus for silatranes [5] – the largest group of fivefold coordinated silicon compounds measured so far – the additional bond to nitrogen could be assured by the observation of a small $^{29}$Si–$^{15}$N coupling (0.20–3.37 Hz). In addition, the difference between the shifts in solution and in the solid state is small (−1.0 to −5.1 ppm). It is estimated that the additional bond to the nitrogen results in an increased shielding of ~−10 to −30 ppm. Shifts of silatranes have been observed between −107 and −59 ppm.




[5]

Diols and other oxygen-containing ligands can give rise to sixfold coordination. The mineral thaumasite displays resonances between −182.1 and −177.9 ppm. Complexes with diols have resonance lines between −197 and −135 ppm. Sixfold coordination is discussed here. The effect of putting two additional oxygen atoms in the coordination sphere is ~−100 ppm upfield shift. The result is less pronounced if five oxygen atoms form the first ligand sphere around the silicon. Resonances are found between −131 to −120 ppm, giving an extra shift of ~−50 ppm compared with tetraalkoxysilanes.

Fluorine is also capable of expanding the coordination sphere of silicon. Here the coupling between $^{29}$Si and $^{19}$F can be used to confirm the number of fluorines around the silicon. Fluxional behaviour is often observed. The hexafluorosilicate anion ($SiF_6^{2-}$) resonates at a very high field (−191.7 ppm). More common are fivefold coordinated compounds measured in solution as well as in the solid state. The extra fluorine results in shifts to low frequencies between −28 and −80 ppm; the larger effect is observed in difluorides.

Donating solvents such as hexamethylphosphoramide or dimethylpropylurea give rise to upfield shifts which have been interpreted as caused by the formation of pentacoordinated complexes. The solvents shifts seem to follow Gutmann's donor number for solvents.

### Transition metal derivatives

This has been a fast growing area but, owing to the very diverse situation for a silicon bound to a transition metal ion, only very rough trends can be distinguished so far. The most positive shift values arise from non-classical transition metal complexes such as [2]. The other limit is formed by a number of platinum complexes that display values between –90 to –29 ppm. Within a group of the periodic table there seems to be a trend towards lower frequency on going to the heavier homologues. Some useful observations have been made by Pannell and Bassingdale for more conventional complexes where the silicon is connected to the metal by a single bond. Comparing the shift of the complex to that of a compound where the metal is exchanged by a methyl group, it is found that the silicon shifts are +40 ppm higher in the complex if the metal is Fe or Ru, but –13 to –39 ppm lower for Re.

## Coupling constants involving $^{29}$Si

All magnetic nuclei within a molecule interact, leading to line splittings or sometimes line broadening. There are two cases to consider:

1. The silicon interacts with an abundant isotope, e.g. $^{1}$H, $^{19}$F or $^{31}$P. Then the silicon resonance is split according to the well-known rules for spin–spin interaction.
2. The silicon couples with another dilute spin such as $^{29}$Si itself or, for example, $^{13}$C or $^{195}$Pt. The coupling then gives rise to doublets to the left and right of a strong centre line in the form of satellites.

For a complete collection of coupling constants, the relevant reviews should be consulted (see Marsmann, and Schraml and Bellama in the Further reading section). Some frequently encountered coupling situations for silicon are discussed below.

### Silicon–hydrogen couplings

The coupling with hydrogen is also visible in proton NMR spectrum and this was used even before $^{29}$Si NMR spectroscopy was feasible. Because of the many and complex splittings arising from silicon–proton interactions, the coupling is usually removed by decoupling in the spectra of organosilicon compounds. Of the three possible pathways, the Fermi contact term is the most efficient. Therefore, because of the negative gyromagnetic ratio of $^{29}$Si, the sign of the coupling over one bond is negative, but absolute values only are given in the following. The magnitude of the coupling then depends on the s orbital density between the hydrogen and the silicon. Semiempirical and empirical relationships have been developed by several authors, e.g.

$$^{1}J(\text{Si,H}) = 725\alpha_{\text{Si}}{}^{2} + 15.9 \quad [7]$$

$$^{1}J(\text{Si,H}) = 969.7\alpha_{\text{Si}}{}^{2} + 5.9N_{\text{P}} - 41.0 \quad [8]$$

$$^{1}J(\text{Si,H}) = 810\alpha_{\text{Si}}{}^{2} + 4.3N_{\text{P}} - 1.40N_{\text{Me}} \quad [9]$$

for compounds of the type $H_nSiR_{4-n}$ (R = $CH_3$, $C_6H_5$; $n$ = 1–3). Here $\alpha_{Si}{}^2$ is the square of the s character of the bond used by the silicon, $N_P$ is the number of phenyl groups and $N_{Me}$ is the number of methyl groups bonded to the silicon. On a purely empirical basis, there is a rough correlation between the electronegativity of the other bonding partners of the silicon and the magnitude of the coupling constant $J$ in such a way that highly electronegative substituents favour a large value for $J$(Si,H). The range for the magnitude of coupling constants over one bond, $^{1}J$(Si,H), is between 74.8 Hz ($H_3SiK$) and 371.7 Hz ($HSiF_3$). An exception is the value of 420 Hz for the transition metal complex $H_2Si[Mn(CO)_5]_2\cdot4py$.

Of the couplings over two bonds, $^{2}J$(Si,H), the most important is for the Si-$CH_3$ fragment. Values range from 6.0 to 9.5 Hz in most cases. An exception is $(CH_3)_3SiLi$ which has a value for $^{2}J$(Si,H) of 2.8 or 3.5 Hz.

The large values of $^{2}J$ found if the bonding goes over a transition metal have been interpreted as a sign of an agostic interaction, e.g. constants between 20 and 69 Hz for $^{2}J$(Si,Mn,H).

### Silicon–fluorine couplings

For one-bond couplings, $^{1}J$(Si,F), typical values lie in the 167–488 Hz range, if only compounds with a coordination number of four for the silicon are considered. The decreased share of the s orbital per bond leads to lower values for the coupling in compounds with higher coordination numbers for silicon, e.g. 131–300 Hz for fivefold and 108–182 Hz for sixfold coordinated silicon compounds.

### Silicon–phosphorus couplings

In silylphosphines, it was possible to rationalize the one-bond phosphorus–silicon coupling constants $^{1}J$(P,Si) according to an empirical equation:

$$^{1}J(\text{P},\text{Si}) = 26.3 + 8u + 5v - Zw \quad [10]$$

where 26.3 Hz is the value for the reference compound $H_3SiPH_2$, $u$ the number of silyl groups, $v$ the number of methyl groups substituting P–H and $w$ is the number of methyl groups substituting the Si–H bonds. The empirical factor $Z$ lies between 3.3 and 5. A silicon–phosphorus double bond gives a splitting of 149 to 155 Hz.

A body of data exists for two-bond silicon coupling to phosphorus with an intervening nitrogen. Very low values are observed if the nitrogen is $sp^3$-hybridized, but values of $^2J$(Si,N,P) between 8 and 42 Hz are typical for nitrogen with a double bond.

### Silicon–carbon and silicon–silicon couplings

Both $^{13}C$ and $^{29}Si$ count as magnetically dilute isotopes. Therefore couplings are found only in the form of small satellites with an intensity of ~0.5% ($^{13}C$) or 2.3% ($^{29}Si$) on both sides of a strong central line. The Fermi contact term should be the dominant pathway to transfer magnetization between silicon and carbon. Thus the interaction depends on the s-electron density on both nuclei. The following semiempirical equations have been discussed:

$$^1J(\mathrm{Si,C}) = 555.5\alpha_c{}^2\alpha_{\mathrm{Si}}{}^2 + 18.2 \qquad [11]$$

$$^1J(\mathrm{Si,C}) = -1227.77P_{\mathrm{sSisC}}{}^2 + 26.0 \qquad [12]$$

where $\alpha_C{}^2$ and $\alpha_{Si}{}^2$ are the square of the s character in the bonds that connect the carbon and the silicon and $P_{sSisC}$ is the bond order element pertaining to the Si–C bond.

Values for $^1J$(C,Si) are between 44 and 107 Hz. The high end is characteristic of silicon bonded to substituents of high electronegativity and $sp^2$-hybridized carbon. For a silicon–carbon double bond a value of 83.7 Hz has been reported.

Silicon–silicon coupling constants over a single bond range between 23 and 186 Hz. For Si=Si double bonds, values between 155–158 Hz are characteristic. This is somewhat higher than the values found for diarysilanes (~85 Hz). The $^2J$(Si,Si,Si) in stress-free situations lies in the 3–13 Hz range, but in polycyclic silanes ranges between 14 and 24 Hz. Because ring strain seems to decrease one-bond coupling constants, both types of constants are difficult to distinguish in that class of silanes. Two-bond couplings over oxygen are generally small. $^2J$(Si,O,Si) data are between 0.5 and 14 Hz. The lower end is typical of silsesquioxanes and the larger values are found for Si–O–Si fragments in the skeletons of trimethylsilyl silicates.

## Spin–lattice relaxation

For an isotope with a negative gyromagnetic ratio it is even more important to know of the pathways by which the $^{29}Si$ dissipates its energy from the excited state to the surroundings (lattice). In principle there are five paths, and the total spin–lattice relaxation time can be calculated from the contributions arising from each one:

$$\frac{1}{T_1} = \frac{1}{T_1^{\mathrm{DD}}} + \frac{1}{T_1^{\mathrm{SR}}} + \frac{1}{T_1^{\mathrm{CSA}}} + \frac{1}{T_1^{\mathrm{SC}}} + \frac{1}{T_1^{\mathrm{E}}} \qquad [13]$$

Here $T_1^{DD}$ is the dipolar, $T_1^{SR}$ the spin–rotation, $T_1^{CSA}$ the chemical shift anisotropy, $T_1^{SC}$ the scalar and $T_1^{E}$ the electronic contribution to the spin–lattice relaxation. At the heart of all relaxation paths is the formation of fluctuating magnetic fields.

Owing to the strong magnetic moment of the electron, $T_1^{E}$ is very efficient. The extreme variation of $T_1$ in solids is explained by the effect of paramagnetic ions or molecular oxygen. In solution, shiftless paramagnetic relaxation agents [e.g. $Cr(acac)_3$, $10^{-2}$ molar] are used to shorten $T_1$ to ~10 s. Molecular oxygen also acts as a relaxation agent. It was calculated that for a solution saturated with oxygen at 1 bar, $T_1^{E}$ is between 35 and 57 s. If other relaxation pathways are to be measured or a small line width is desired, paramagnetic impurities have to be removed. Paramagnetic ions that leach from the sample tube are removed by rinsing them with a chelating reagent (e.g. EDTA) before use. Oxygen is removed by bubbling argon through the sample or by vacuum techniques.

In the absence of unpaired electrons, the most efficient pathways for relaxation are the dipolar and the spin–rotation mechanisms. Because the interaction with protons is the most dominant one in this respect, the discussion will be restricted to this aspect. The dipolar term $T_1^{DD}$

$$\frac{1}{T_1^{\mathrm{DD}}} = \hbar^2\tau_c \sum \gamma_{\mathrm{H}}^2\gamma_{\mathrm{Si}}^2 r_{\mathrm{SiH}}^{-6} \qquad [14]$$

then depends on the gyromagnetic ratios of the $^1H$ and $^{29}Si$ isotopes, the distance between them and a

correlation time, $\tau_c$ (Eqn [14]). Owing to the larger radius of silicon compared with that of carbon, dipolar relaxation is less efficient than in $^{13}$C NMR. Another aspect of the longer Si–H bond distance is that hydrogens placed farther apart in the molecule or on solvents can now also contribute to the relaxation process. The correlation time $\tau_c$ is the mean time the molecule, of a part of it, rotates one radian. This process is temperature dependent (**Figure 9**).

If in organosilicon compounds, the signal splittings caused by the coupling with the protons are removed by decoupling, then a modulation of the signal intensity occurs that is connected with the dipolar term. If the silicon relaxes by dipolar interaction with the protons only, then the signal intensity is multiplied by a factor of −1.52 from the nuclear Overhauser effect (NOE):

$$\mathrm{NOE} = \frac{\text{signal intensity with irradiation of the protons}}{\text{signal intensity without irradiation}}$$

$$= 1 + \frac{\gamma_{\mathrm{H}} T_1}{12 \gamma_{\mathrm{Si}} T_1^{\mathrm{DD}}} \qquad [15]$$

The Overhauser enhancement factor $\eta$ is calculated as

$$\mathrm{NOE} = 1 + \eta \qquad [16]$$

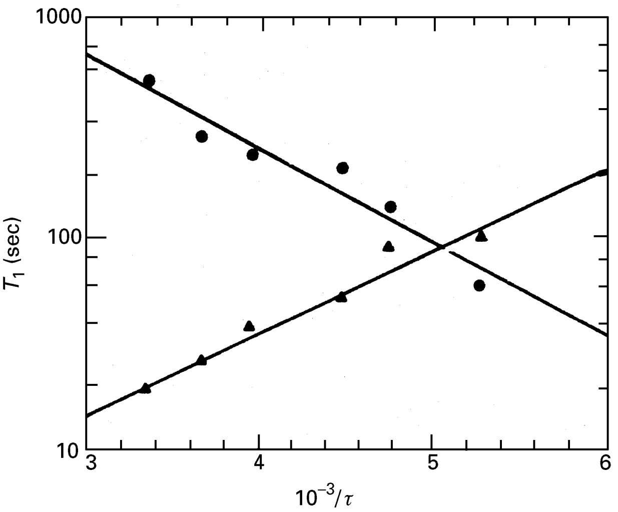



**Figure 9** $^{29}$Si Spin–lattice relaxation in TMS. (●)$T_1^{DD}$; (▲)$T_1^{SR}$; assuming $T_1^{CSA}$>>$T_1^{DD}$. Reproduced with permission of the American Chemical Society from Levy GC, Cargioli JD, Juliano PC, and Mitchell TD (1973) *Journal of the American Chemical Society*. **95**: 3445.

If the dipolar mechanism is the only pathway, then $\eta$ becomes $\eta_0 = -2.52$. Otherwise the share of the dipolar relaxation on the whole relaxation process can be calculated by:

$$T_1^{\mathrm{DD}} = -\frac{2.52 T_1}{\eta} \qquad [17]$$

Therefore, in the unfortunate case of $\eta = -1$ the signal intensity is zero. The other major relaxation mechanism involves spin–rotation interaction. The rotation of a molecule or a part of it generates a ring current through the electrons connected to it. By the tumbling motion of the molecules, fluctuating magnetic fields are produced. The spin–rotation relaxation time $T_1^{\mathrm{SR}}$ is given by

$$\frac{1}{T_1^{\mathrm{SR}}} = \frac{2\pi IkT}{h^2} \frac{2C_\perp^2 + C_\parallel^2}{3} \tau_{\mathrm{j}} \qquad [18]$$

The correlation time $\tau_j$ is the angular momentum correlation time, the time the molecule changes its angular momentum, usually the time between collisions. The terms $C_\parallel$ and $C_\perp$ are components of the spin–rotation interaction tensor and $I$ is the moment of inertia of the molecule. For small step diffusion, $\tau_c$ and $\tau_j$ are connected by the Hubbard relation:

$$\tau_{\mathrm{c}} \tau_{\mathrm{j}} = \frac{I}{6kT} \qquad [19]$$

Because for any given molecule in Equation [17] all data are constant except $T$ and $\tau_j$, this results in a temperature dependence which is the reverse of that for $T_1^{\mathrm{DD}}$ (**Figure 9**). At high temperatures the relaxation process is dominated by the spin–rotation contribution and at low temperatures by the dipolar part. In between is a maximum value for the total relaxation time $T_1$. Another consequence is that the intensity of proton decoupled $^{29}$Si spectra is temperature dependent.

Typical values for the relaxation times for organosiloxanes are between 35 to 205 s. Lower values are found for aqueous solutions of silicates (0.3–4.6 s), higher ones for silicon compounds without protons, e.g. chlorosiloxanes (53–352 s).

## List of symbols

$B$ = magnetic flux density; $J$ = coupling constant; $T_1^{\mathrm{DD}}$ = dipolar spin–lattice relaxation time; $T_1^{\mathrm{CSA}}$ = chemical shift anisotropy relaxation time;

$T_1^E$ = electronic contribution to the spin–lattice relaxation time; $T_1^{SC}$ = scalar relaxation time; $T_1^{SR}$ = spin–rotation relaxation time; $\gamma$ = gyromagnetic ratio; $\eta$ = NOE enhancement factor; $\delta^*$ = reduced shielding constant; $\tau_c$ = correlation time; $\tau_j$ = angular momentum correlation time.

*See also:* **Chemical Shift and Relaxation Reagents in NMR; NMR of Solids; NMR Pulse Sequences; NMR Relaxation Rates; Nuclear Overhauser Effect; Parameters in NMR Spectroscopy, Theory of; Solid State NMR, Methods; Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules.**

## Further reading


Blinka TA, Helmer BJ and West R (1984) Polarization transfer NMR spectroscopy for silicon-29: the INEPT and DEPT techniques. *Advances in Organometallic Chemistry* **23**: 193–218.

Denk M *et al* (1994) *Journal of the American Chemical Society* **116**: 2691–2692.

Engelhardt G and Michel D (1987) *High Resolution Solid-State NMR of Silicates and Zeolites*. Chichester: John Wiley & Sons.

Gehrhus B *et al* (1996) *Journal Organometallic Chemistry* **521**: 211–220.

Harris RK and Knight CTG (1983) *Journal of the Chemical Society, Faraday Transactions* 2, **79**: 1525; 1539.

Horn H-G (1992) Spectroscopic properties of silicon–sulfur-compounds with at least one Si–S-bond – a review. *Journal für Praktische Chemie* **334**: 201–213.

Jutzi P *et al* (1989) *Chem. Ber.* **122**: 1629–1639.

Likiss PD (1992) Transition meta complexes of silylenes, silenes, disilenes and related species. *Chemical Society Reviews* 271.

Marsmann H (1981) *NMR Basic Principles Progress* **17**: 65–235.

Pannell KH and Bassingdale AR (1982) *Journal of Organometallic Chemistry* **229**: 1.

Schraml J (1990) $^{29}$Si NMR spectroscopy of trimethylsilyl tags. *Progress in Nuclear Magnetic Resonance Spectroscopy* **22**: 289–348.

Schraml J and Bellama JM (1976) $^{29}$Si nuclear magnetic resonance. In: *Determination of Organic Structures by Physical Methods*, Vol 6, pp. 203–269. New York: Academic Press.

Uhlig F, Herrmann U and Marsmann H, $^{29}$Si NMR database system. http://oc30.uni-paderborn.de/~chemie/fachgebiete/ac/ak_marsmann or http://platon.chemie.uni-dortmund.de/acii/fuhlig

Williams EA (1983) *Annual Reports on NMR Spectroscopy* **15**: 235.

Williams EA (1989) Spectroscopy of organosilicon compounds. In: Patai S and Rappoport Z (eds) *The Chemistry of Organic Silicon Compounds*, Vol 1, pp 511–554. New York: John Wiley & Sons.

Williams EA and Cargioli JD (1979) *Annual Reports on NMR Spectroscopy* **9**: 221.


# Scandium NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# Scanning Probe Microscopes

**JG Kushmerick** and **PS Weiss**, Pennsylvania State University, University Park, PA, USA

Copyright © 1999 Academic Press

**SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS**

**Methods & Instrumentation**

The atomic resolution and spectroscopic capabilities of scanning probe microscopes (SPMs) have enabled elucidation of the great heterogeneity of surface sites including: defects, step edges, lattice impurities, adsorbates, and grown structures. One or more of these minority sites often function as the active sites for surface processes, and their individual investigation is thus required to gain insight into such processes. Such specific information cannot typically be acquired by spectroscopies that measure ensemble averages of the surface.

The scanning tunnelling microscope (STM) is the most suited, and the most developed of the various SPMs, to perform local spectroscopic measurements. Discussion of STM techniques will constitute the bulk of this article. It also has the most restricted range of accessible substrates in terms of conductivity and roughness. The atomic force microscope (AFM) has limited spectroscopic capabilities but can image a wider range of samples. The near-field scanning optical microscope (NSOM) has excellent spectroscopic, but limited spatial resolution. These latter two SPMs are discussed at the end of this article.

The basic working principles of the STM rely on the quantum mechanical properties of electrons. When an atomically sharp metal probe tip is brought within a few Å of a conducting or semiconducting surface, electrons can tunnel through the energy barrier between the probe tip and surface. By applying a constant DC bias voltage ($V$), a net tunnelling current ($I$) can be induced between the probe tip and the sample under study. Raster scanning the tip across the surface, through the use of piezoelectric transducers while maintaining a constant tunnelling current, images a surface of constant density of electronic states. The resulting image is a convolution of topographic and electronic properties of the sample surface.

The tunnelling current is exponentially dependent on the tip–sample separation and linearly dependent on the densities of tip and sample electronic states. The applied bias voltage determines which electronic states, on the both sides of the tunnelling junction, are being sampled, and thus allows acquisition of *spatially resolved spectra*. **Figure 1** is a schematic of the tunnelling potential barrier. Adjusting the bias voltage probes different electronic states, allowing the local density of states to be mapped as a function of energy. This is the basis of electronic spectroscopy with the STM.

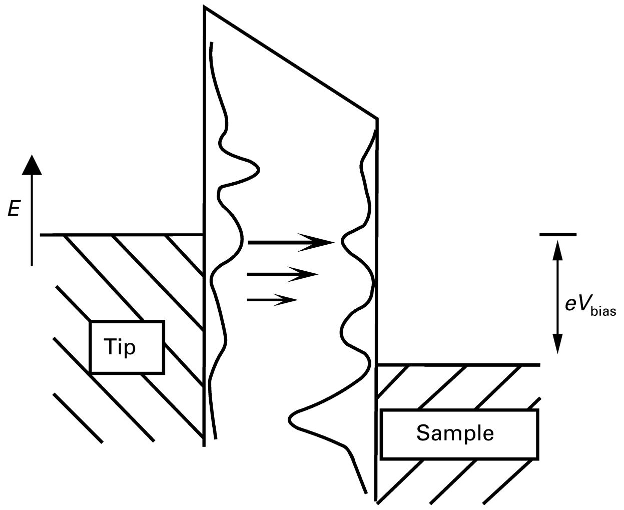



**Figure 1** Energy level diagram for the tip–sample tunnelling gap, depicting electrons tunnelling from tip to sample. The local density of states for both the tip and sample as a function of energy is represented graphically.

## Scanning tunnelling microscope

### Experimental methods

**Voltage-dependent imaging** One straightforward means to gain spectroscopic information from the STM is to acquire multiple images of the same area sequentially at different bias voltage, and thus different energies. These images are then overlaid. Each image shows the relative contributions from features at the particular energy relative to the Fermi level defined by the bias voltage (at energies $eV$, where $V$ is the bias voltage and $e$ is the charge on an electron). Another method is to acquire tunnelling current vs. bias voltage ($I$–$V$) characteristics at every imaged point, or at selected locations.

A dramatic example of this technique is the contrast reversal observed for some semiconductor

surfaces. When voltage-dependent imaging of GaAs(110) is performed, only the Ga atoms are imaged when electrons are tunnelling into surface states, while the As atoms are imaged when electrons are tunnelling from filled surface states to the tip. **Figure 2** is a superposition of two images of the GaAs(110) surface. The orange image was obtained at positive sample bias, thus imaging the Ga atoms, while the green image was obtained with a negative sample bias, showing the As atoms. The cause of this contrast reversal is charge transfer from the Ga to the more electronegative As atoms, which results in the localization of the lowest lying empty states and highest lying filled states on the Ga and As atoms, respectively.

Voltage-dependent imaging can also be applied to understand bonding of adsorbate molecules to surfaces. **Figure 3** shows a $Ni_3$ cluster on $MoS_2$ at a temperature of 4 K, imaged at three different bias voltages. The cluster's contribution to the electronic structure varies dramatically as can be seen at these different energies. At +2 V sample bias (electrons tunnelling into the empty states on the surface) the $Ni_3$ cluster appears as a significant protrusion from the $MoS_2$ surface, indicating that it enhances the local density of empty electronic states at this energy. Similarly, we see that the cluster depletes the local density of filled states at –2 V sample bias (electrons tunnelling from the filled sample electron states to the tip) since the cluster then appears as a depression in the surface. At a sample bias of +1.4 V the cluster is not directly apparent but a diffuse ring ~30 Å in diameter surrounding it is imaged. This ring is outside the atomic positions of the $Ni_3$ cluster imaged at +2 V (~ 16 Å diameter), and results from a perturbation of the $MoS_2$ surface electronic structure by the $Ni_3$ cluster. This ring is believed to be purely electronic in origin with little change in atomic positions of the substrate. Each image shows a different contribution of the $Ni_3$ cluster to the surface electronic structure, demonstrating how voltage-dependent imaging can measure electronic states as a function of both energy and position.

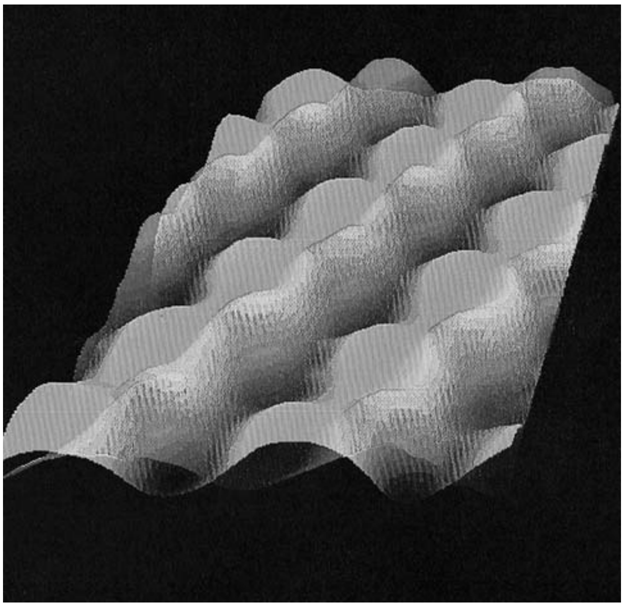


**Figure 2** A composite of two images of the GaAs(110) surface. The orange features obtained at positive sample bias are the Ga atoms, while green features obtained at a negative sample bias are the As atoms. Feenstra RM, unpublished results. (See Colour Plate 56).

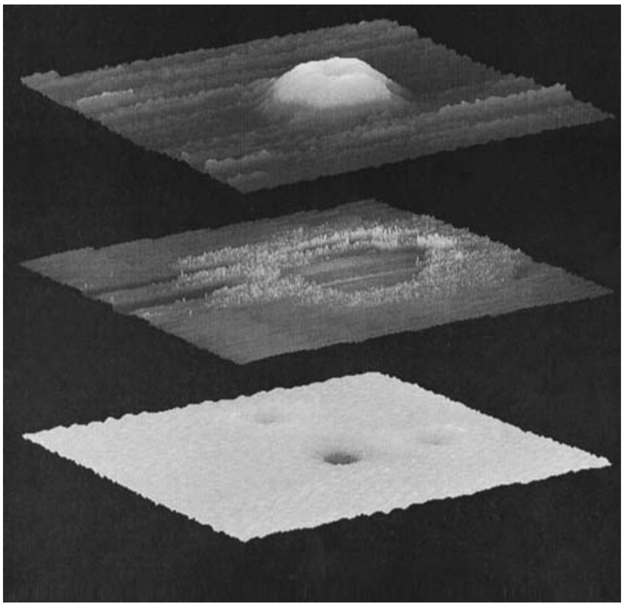


**Figure 3** Three STM images of a $Ni_3$ cluster adsorbed on a $MoS_2$ basal plane at 4 K. All three images show a 60Å × 60Å area and are plotted as three-dimensional representations with the same aspect ratio and with the same angle of view. The images were acquired with sample biases of: +2 V (upper), +1.4 V (middle), and –2 V (lower). Reproduced with the permission of the American Chemical Society from Kushmerick JG and Weiss PS (1998) *Journal of Physical Chemistry B* **102**: 10094–10097. (See Colour Plate 57).

Voltage-dependent imaging can be a powerful technique but it does have some inherent problems. It is necessary to obtain many images in order to map the surface electronic structure as a function of energy. Thermal drift and piezoelectric creep can make overlay and comparison of successive images difficult. The fact that the constant current images acquired are a convolution of geometric and electronic effects further complicates the interpretation of observed features. The latter technique of recording complete or selected sets of *I–V* characteristics,

discussed in the next section, overcomes some of these problems.

**Local current–voltage measurements** Spectroscopic information over a large energy scale can be obtained by acquiring complete *I*–*V* curves at one or many locations. This can be accomplished by releasing feedback control while holding the probe tip steady and measuring the current with respect to applied bias voltage. Spectroscopy with the probe tip held in place allows scanning at both polarities and through zero bias.

Spectra are usually plotted as $(dI/dV)/(I/V)$ vs. $V$ for comparison to other measurements of surface densities of states. This normalizes the spectra, removing the dependence on voltage and the exponential dependence on tip–sample separation. The derivative $dI/dV$ can be numerically calculated, or a lock-in amplifier can be used to measure $dI/dV$ phase sensitively, with a superimposed sinusoidal modulation on the bias voltage.

Synchronizing the feedback blanking and the applied voltage ramp enables acquisition of *I*–*V* curves at specified points in an image, thus mapping the energy and position dependence of the surface electronic structure in one image. **Figure 4** shows spectra acquired over Si atoms at three different locations in the unit cell of the Si(111)-(7 × 7) surface reconstruction. The three spectra show different electronic features due to the local bonding and environment of the Si atom probed.

Collection of data for the entire energy region of interest circumvents the need to construct a spectrum from several images. The dynamic signal range plays a role in determining how large an energy range can be scanned. The current goes to zero as the magnitude of the bias voltage decreases. Thus, features at high and low bias voltages can be hard to resolve in a single spectrum (recall that the preferred display of spectra, $(dI/dV)/(I/V)$ vs. $V$, has tunnelling current in the denominator).

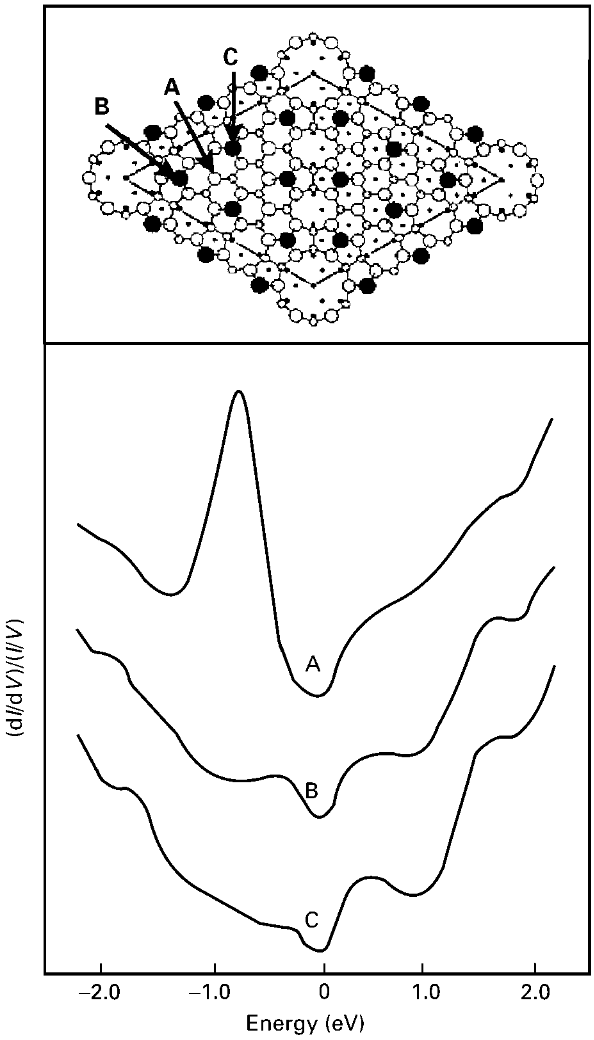



**Figure 4** Atom-resolved spectra of the Si(111)-(7 × 7) surface obtained at positions indicated in schematic. Reproduced with permission of the American Institute of Physics from Wolkow R and Avouris PH (1988) *Physical Review Letters* **60**: 1049–1052.

**Scanning tunnelling spectroscopy** Closely related to voltage-dependent tunnelling is the modulation technique of scanning tunnelling spectroscopy (STS). This technique consists of superimposing a small sinusoidal modulation voltage on the constant DC bias voltage. The modulation frequency is typically chosen to be higher than the cut-off frequency of the feedback loop, resulting in a constant average tunnelling current. If the modulation frequency is too low the control electronics will attempt to compensate by adjusting the gap spacing. Alternatively, control of the tip height can be released during acquisition of spectra. By measuring the in-phase tunnelling current modulation with a lock-in amplifier, $dI/dV$ can be recorded simultaneously with the constant current image. Structure in $dI/dV$ as a function of the applied bias voltage, can be attributed to structure in the local density of surface states.

Application of STS at constant average tunnelling current suffers from two disadvantages. The first is that the magnitude of the $dI/dV$ signal scales as $I/V$, and thus becomes progressively small at low bias voltages. The second is that at low bias voltages the tip–sample separation reduces in order to maintain a constant average tunnelling current. If the density of states is too low the tip will come into contact and damage the surface. This form of STS, like voltage-dependent imaging, requires the acquisition of multiple images to map out electronic structure as a function of energy.

**Barrier height spectroscopy** Lack of direct knowledge of the effective height and width of the tunnelling barrier during spectroscopic measurements makes quantitative understanding of spectra difficult. Properties of the tunnelling barrier can however be investigated, giving information complementary to *I–V* spectroscopy. Barrier height spectroscopy consists of measuring the dependence of the tunnelling current on the tip–sample separation, at constant bias voltage. What is actually measured is the apparent barrier height, which is defined as

$$\Phi_{\mathrm{ap}} = \left[\frac{1}{1.025} \cdot \frac{\mathrm{d}(\ln G_{\mathrm{t}})}{\mathrm{d}z}\right]^2$$

where $G_t$ is the tunnelling conductance. Thus by measuring current as a function of tip–sample separation, $G_t$ and $\Phi$ap can be calculated.

By applying a small modulation to the tip–sample separation ($z$) at constant bias voltage and constant tunnelling current, a lock-in amplifier can be used to measure d$I$/d$z$. The modulation frequency is chosen to be larger than the feedback loop bandwidth but smaller than resonant mechanical frequencies of the microscope. The d$I$/d$z$ signal measured is directly related to the local surface work function and often provides useful contrast.

Another method of measuring the apparent barrier height is to record the tunnelling current directly as a function of tip–sample separation for a constant bias voltage. The tip–sample separation is reduced while the tunnelling current is measured. Although conceptually simple, there are experimental complications that must be taken into account. As the tip–sample separation becomes very small, the attractive forces between them tend to deform the tip and cause the actual separation to be smaller than expected. In fact, while point contact is typically used as the reference for tip–sample separation, this contact is usually realized with a jump to contact from small separation. It can also be difficult to maintain a constant bias voltage, as the tip–sample separation becomes very small, since the junction impedance can become comparable to that of the current preamplifier causing a significant decrease in the actual bias voltage. By measuring the voltage across the junction as well as the tunnelling current this problem can be overcome. **Figure 5** shows tunneling current and bias voltage as a function of tip–sample separation for a Au(110) sample and W tip (most probably covered with Au). The dramatic decrease in bias voltage at small tip–sample separations can be seen.

**Inelastic electron tunnelling spectroscopy** The majority of the tunnelling current consists of elastically tunnelling electrons. Inelastic pathways in which tunnelling electrons excite transitions can be used for recording local spectra. For a vibrational transition of a molecule contained in the junction this can occur above the threshold voltage of $|V| = \hbar\omega/e$ (**Figure 6**), where $\hbar\omega$ is the energy of a molecular

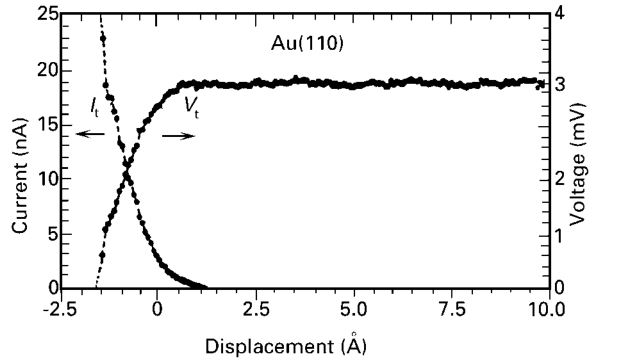



**Figure 5** Tunnelling current ($I_t$), and bias voltage ($V_t$) during tip approach on Au(110). Reproduced with permission of the American Institute of Physics from Olesen L *et al.* (1996) *Physics Review Letters* **76**: 1485–1488.

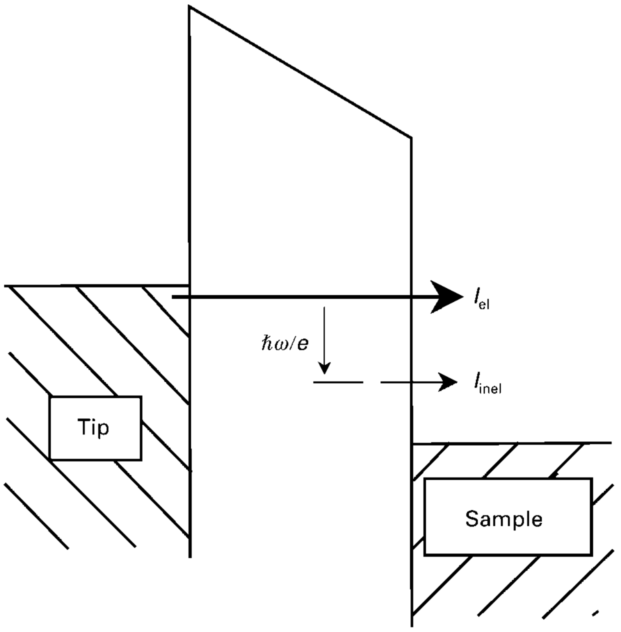



**Figure 6** Energy level diagram for tip–sample tunnelling gap depicting both elastic ($I_{el}$) and inelastic ($I_{inel}$) tunnelling. Above the threshold voltage $|V| = \hbar\omega/e$ tunnelling electrons can excite a vibrational transition, for molecules in the tunnel junction. The densities of states of both tip and sample have been neglected for clarity.

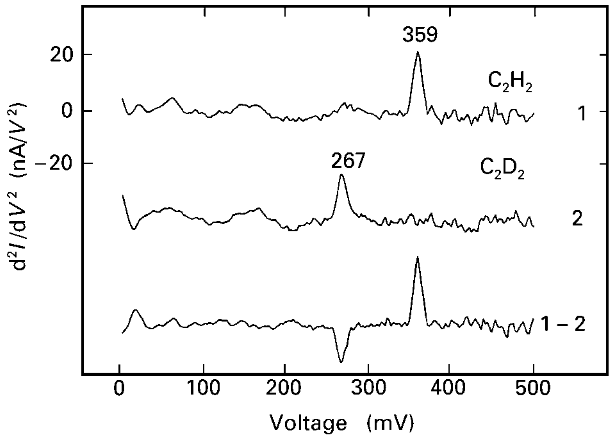



**Figure 7** Inelastic electron tunnelling spectra of $C_2H_2$ (1) $C_2D_2$ (2) taken with the same STM tip and the difference spectra (1 – 2). Reproduced with the permission of the American Association for the Advancement of Science from Stipe BC *et al.* (1998) *Science* **280**: 1732–1735.

vibrational transition, and $e$ is the charge of an electron. Inelastic pathways effectively increase the available number of final states for a tunnelling electron, thus producing a kink in the $I$–$V$ curve at $|V| = \hbar\omega/e$ for each vibrational transition excited. Using a similar modulation technique to that previously described $d^2I/dV^2$ vs. $V$ can be measured directly with a lock-in amplifier (as the second harmonic of the modulation frequency). This has the benefit of transforming the kinks in $I$–$V$ to peaks and dips in $d^2I/dV^2$ vs. $V$, some of which may be assigned to molecular vibrations (of energy $eV$). Vibrational spectra can be obtained in this fashion for molecules in metal–insulator–metal sandwich tunnel junctions at low temperature. Limiting peak widths of the observable features also requires low temperatures as in tunnel diodes and sandwich tunnel junctions. **Figure** 7 shows inelastic electron tunnelling spectra of acetylene ($C_2H_2$) and perdeuterated acetylene ($C_2D_2$), adsorbed on Cu(100) at 8 K, and the difference spectrum. The spectrum of acetylene has a peak at 359 meV corresponding to the C=H stretch. This peak is shifted to 267 meV for the C=D stretch in perdeuterated acetylene. Tuning the bias voltage to the energy of the vibrational mode and monitoring $d^2I/dV^2$ allows vibrational spectroscopic imaging. **Figure** 8 demonstrates vibrational spectroscopic imaging of acetylene and perdeuterated acetylene of Cu(100).

Inelastic electron tunnelling spectroscopy with the STM allows unambiguous chemical identification of surface species, as demonstrated above. Electronic spectroscopy is also capable of differentiating between limited sets of adsorbates but does not as a rule enable such determinations. The vibrational spectra of isolated molecules also shed light on the chemical environment and bonding changes of minority

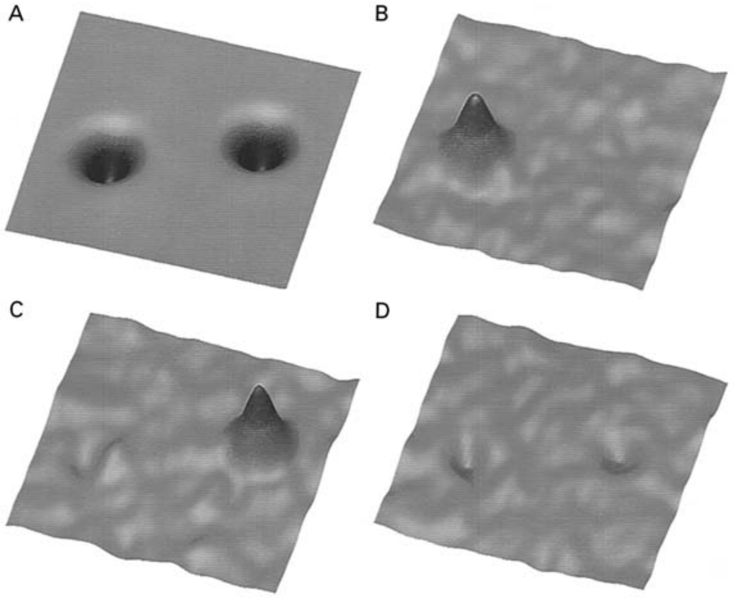



**Figure 8** Vibrational spectroscopic imaging of $C_2H_2$ and $C_2D_2$. (A) Constant current STM image of a $C_2H_2$ molecule (left) and $C_2D_2$ molecule (right). The $d^2I/dV^2$ images of the same area recorded with a bias voltage of (B) 358 mV, (C) 266 mV, and (D) 311 mV, with a 10 mV modulation. All images are 48 Å × 48 Å with 1 nA DC tunnelling current. Reproduced with the permission of the American Association for the Advancement of Science from Stipe BC *et al.* (1998) *Science* **280**: 1732–1735. (See Colour Plate 58).

surface sites, and their critical role in surface processes, such as chemistry, corrosion, and film growth.

**Photon emission** Electrons tunnelling inelastically between tip and surface can induce photon emission from the tunnel junction. On some surfaces, the inelastic tunnelling is enhanced by exciting surface plasmons, which decay by emitting a photon. Measuring emitted photons as a function of applied bias voltage yields information analogous to conventional inverse photon emission experiments. Unlike conventional inverse photon emission experiments occupied and unoccupied electronic states can be probed with photon-emission STM by scanning positive or negative bias voltages, respectively. Dispersing the emission with a spectrometer allows the spectral fingerprint of a feature to be measured. The spatial resolution of the STM allows the emission spectrum of isolated molecules to be recorded.

## Instrumentation

**Microscope** Design considerations for spectroscopy with the STM are the same as for the STM in general. Due to the exponential dependence of tunnelling current on tip–sample separation, vibration isolation is of critical importance. The most demanding of the techniques mentioned above, inelastic electron tunnelling spectroscopy, requires special vibration isolation down to the level of ~ 0.001 Å over a limited bandwidth and operation at extremely low temperatures (ca. 4 K). Various designs for vibration isolation have been implemented, from placing the STM on top of a Viton stack, to mounting the instrument on a pneumatically suspended laser table enclosed in an acoustic isolation chamber. The STM itself should be constructed rigidly so as to yield the highest possible resonance frequencies. Shielding from electronic noise is also important, but is determined primarily by the design of the control electronics as will be discussed below.

Other aspects of the microscope design depend upon the intended experiments. *I–V* spectra can be obtained in air, if the system to be studied is not air-sensitive. Investigation of isolated adsorbates on metal single crystals requires ultrahigh vacuum, to enable sample preparation, and often cryogenic temperatures, to limit thermally activated diffusion. Cryogenic cooling also reduces thermal drift allowing an area to be studied repeatedly.

**Electronics** Low-noise electronics are important to maintain the stability of the probe tip and to avoid coupling of the control electronics to AC voltages powering the electronics or in nearby equipment. Proper shielding and planning can reduce the electronic noise to sufficiently low levels that this is rarely a limiting factor.

Blanking the feedback loop, which is required for many of the spectroscopic techniques discussed, is typically accomplished through the use of a sample-and-hold circuit. In normal operation, the STM maintains a constant tunnelling current by driving the *z*-piezoelectric transducer with the error signal generated from the difference of the tunnelling current (converted to a voltage by the preamplifier) and a reference voltage. A sample-and-hold circuit interrupts the input to the feedback control loop and thus maintains a constant voltage to the piezoelectric transducer controlling tip–sample separation. The applied *z*-piezoelectric voltage can be held constant for up to several seconds with such a circuit. If longer blanking times are required, the use of a digital feedback blanking circuit can hold the voltage constant indefinitely. In addition, nearly constant drift can be corrected in microscopes where hold times are long compared to drift rates.

**Microscope probe tip** Since the observed spectral features are a convolution of both the tip and the sample electronic density of states, understanding the role of the microscope probe tip in determining the observed spectra is critical to allowing interpretation of spectral features. A tip with a constant, preferably flat, density of states is typically desirable so that the sample's electronic structure dominates the spectral features observed. Alternatively a probe tip with a single sharp spectral feature can be useful in obtaining spectra.

Semiconductors have greatly varying densities of states and thus contributions from metal probe tips are less prominent. Metal surface state densities vary to a much smaller degree and are thus comparable to those of the tip states, making electronic spectroscopy of metals more complicated. To enable comparison between spectra obtained at various surface positions it is important that the tip structure, and thus density of states remains constant between measurements. Rearrangement of the tip apex can greatly affect the observed spectra and thus lead to spurious data interpretation.

Heteroatoms adsorbed to the tip can also play a large role in determining the observed spectra. Many studies have shown that the transfer of an adatom or molecule to the STM tip can affect the observed topography, e.g. yielding atomic resolution on an electronically flat close-packed metal substrate. The spectroscopic effects can be even larger. Special care must be taken when probing semiconductor surfaces. Tip–sample contact resulting in some

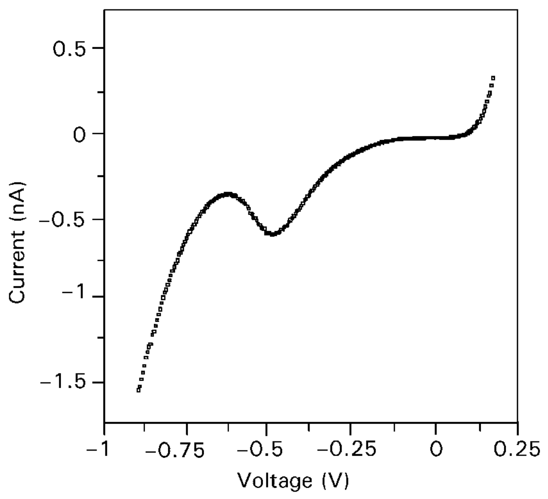



**Figure 9** *I–V* scan of $MoS_2$ with a tungsten tip demonstrating negative differential resistance caused by the presence of $MoS_2$ on the tip. Kushmerick JG and Weiss PS unpublished results.

semiconductor material on the probe tip apex can give rise to odd effects, such as negative differential resistance (NDR). The tunnelling current between tip and sample normally increases with increasing bias voltage. If, however, there is a band gap on both the tip and sample, the tunnelling current can decrease with increasing bias voltage, due to the decrease in overlap of the density of states when the two band gaps are not aligned. **Figure 9** is an *I–V* curve exhibiting NDR of $MoS_2$, with a tungsten tip that came in contact with the surface. By backing away from the surface and field cleaning the tip, *I–V* curves without NDR are again obtained.

Negative differential resistance can also be observed for surfaces with localized trap states. A tunnelling electron can become localized for long times in these surface states, when they are in resonance with the Fermi level of the tip. Electrons so localized electrostatically repel other electrons causing a decrease in tunnelling current, referred to as a coulomb blockade. The voltage at which the NDR occurs is a measure of the energy of the localized trap state.

## Other scanning probe microscopes

### Atomic force microscope

It is possible to perform spectroscopy with SPMs other than the STM. The use of metal-coated cantilevers allows spectroscopic measurements with the AFM. The AFM maps surface topography by monitoring the attractive and/or repulsive probe–surface interactions as a surface is scanned by a cantilever. With a metal-coated cantilever, *I–V* curves can be obtained for surfaces incompatible for STM study, such as metal structures covered with an insulating oxide layer. The insulating film acts as a tunnelling barrier of constant width, allowing spectroscopic measurements analogous to constant-separation *I–V* scans.

Functionalization of a cantilever by deposition of a molecular film allows the chemical forces between molecules to be probed. The chemical force (bonding) between a functionalized cantilever and surface is measured by monitoring cantilever deflection while the sample approaches, makes contact with, then is drawn back from the probe. The deflection of the cantilever is then converted to a force from the cantilever spring constant. **Figure 10** is a plot of force versus cantilever displacement curves for three combinations of tip and sample functionalization. The hysteresis in the curves is a measure of the adhesive interactions between probe

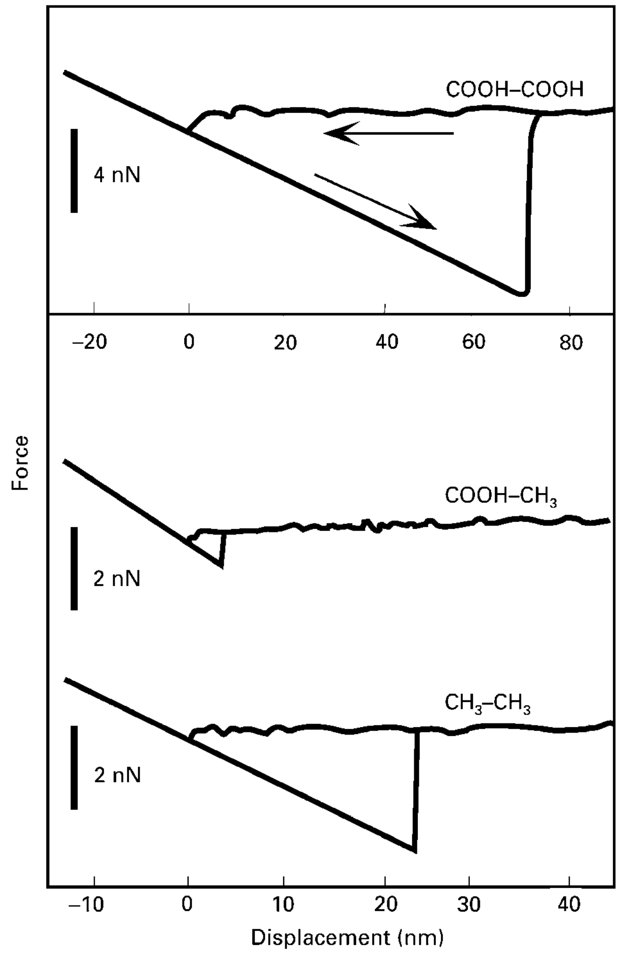



**Figure 10** Force versus displacement curves recorded between functionalized atomic force microscope cantilever probes and surfaces. The adhesive interactions are strong for like–like interactions (COOH–COOH and $CH_3$–$CH_3$) but weak for interaction between unlike functional groups (COOH–$CH_3$). Noy A, Frisbie CD and Lieber CM, unpublished results.

and sample. It can be seen that the strongest adhesion is when both the cantilever and sample are functionalized with the –COOH (hydrophilic) groups. Functionalization of the tip and sample with (hydrophobic) $-CH_3$ groups also yields a strong attractive interaction, but adhesion between –COOH and $-CH_3$ is minimal due to their smaller interactions. Such measurements are sometimes called chemical force microscopy.

### Near-field scanning optical microscope

The NSOM also enables localized spectroscopic measurements. This technique circumvents the diffraction-limited resolution of conventional optical microscopy by scanning a light source or collector in close proximity to the surface of interest. A metal-clad optical fibre typically serves as the probe, allowing light to be either emitted or collected from its apex which is free of any metal cladding. Other probes can also be used. By bringing the probe–sample separation into the near-field regime, the spatial resolution achieved is determined by the size of the unclad probe apex and can go well below the far-field limit of $\lambda/2$, where $\lambda$ is the wavelength of light used. The NSOM can record absorption and fluorescence spectra, as well as measure the refractive index of surface and subsurface species. Spatial resolution of 12 nm has been achieved with visible light. While resolution is lower than that attainable with the STM, the ease of interpretation and familiarity of optical spectra make this technique attractive. Systems particularly suited to study with NSOM include unique biological structures such as the photosynthetic membrane shown in **Figure 11**.

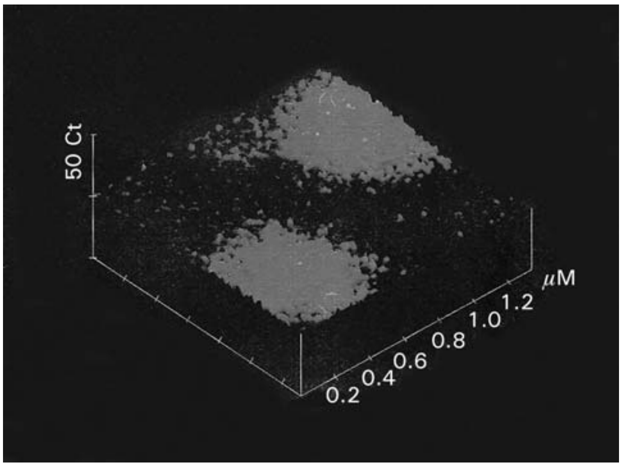



**Figure 11** Fluorescence emission near-field scanning optical microscope image (1.3 μm × 1.3 μm) of a photosynthetic membrane fragment. Reproduced with permission of the American Chemical Society from Dunn RC *et al.* (1994) *Journal of Physical Chemistry* **98**: 3094–3098. (See Colour Plate 59).

If infrared absorption or Raman scattering is used as the contrast mechanism, vibrational spectra of samples can be obtained. The combination of the nanoscale spatial resolution of a scanned probe with the chemical specificity of vibrational spectroscopy allows *in situ* mapping of chemical functional groups with subwavelength spatial resolution. **Figure 12** is a shear force image of a thin polystyrene film along with a representative near-field spectrum of the

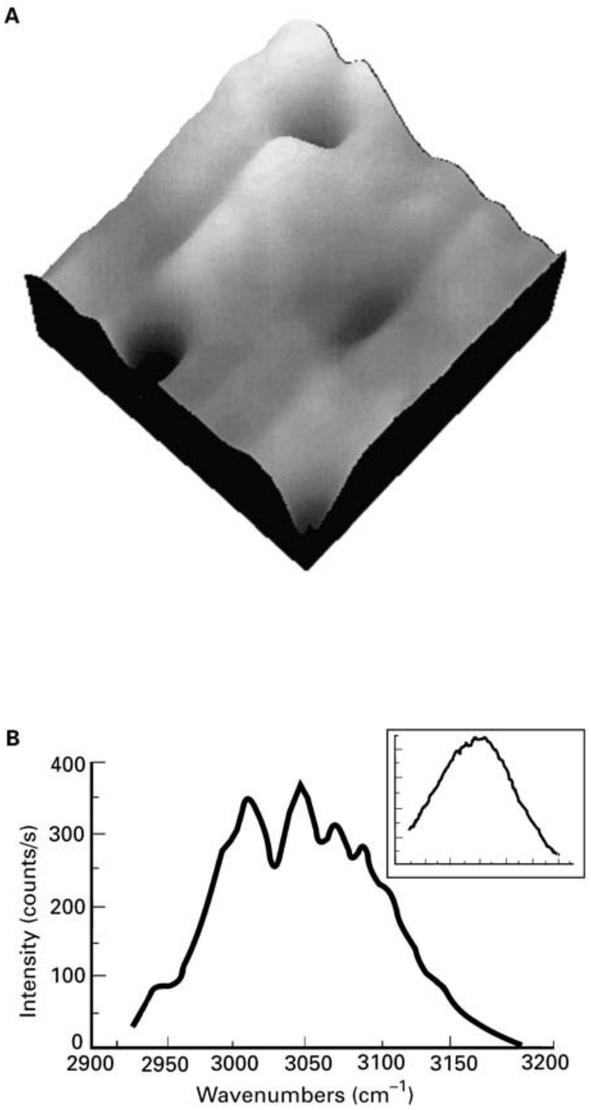



**Figure 12** (A) A 3.1 × 3.1 μm shear force image of a thin polystyrene film deposited on a glass cover slip. The full-scale *z*-range is 62 nm. (B) Near-field IR transmission spectrum of a thin polystyrene film in the aromatic C–H stretching region. The inset is the laser output over the same spectral range in the absence of polystyrene absorption. Reproduced with permission of Stranick SJ, Richter LJ, Cavanagh RR and Michaels C, unpublished results.

aromatic C–H stretch region recorded in the microscope with a 1 second integration time per point.

Also suitable for study are dopants in semiconductors and local structures in semiconductor lasers.

## Overview

Spectroscopies with SPMs are being developed rapidly. The ability to study isolated or small structures of adsorbates has allowed incredible insight into the rich chemistry of surfaces, particularly the defining roles that defect-sites play. The recent demonstrations of single molecule vibrational spectroscopies with the STM and NSOM have further opened up new avenues for investigation.

## List of symbols

Å = Ångstrom; $e$ = charge of an electron; $E$ = energy; $G_t$ = tunnelling conductance; $I$ = tunnelling current; K = Kelvin; $V$ = bias voltage; $z$ = tip–sample separation; $\hbar$ = Plancks constant/$2\pi$; $\Phi_{ap}$ = apparent barrier height; $\omega$ = vibrational frequency.

*See also:* **Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; Inorganic Condensed Matter, Applications of Luminescence Spectroscopy; Scanning Probe Microscopy, Applications; Scanning Probe Microscopy, Theory; Surface Plasmon Resonance, Applications; Surface Plasmon Resonance, Instrumentation; Surface Plasmon Resonance, Theory; Surface Studies By IR Spectroscopy**.

## Further reading

Betzig E and Trautman JK (1992) Near-field optics: Microscopy, spectroscopy, and surface modification beyond the diffraction limit. *Science* **257**: 189–195.

Bonnell DA (ed) (1993) *Scanning Tunneling Microscopy and Spectroscopy: Theory, Techniques, and Applications*. New York: VCH Publishers.

Chen CJ (1993) *Introduction to Scanning Tunneling Microscopy*. NewYork: Oxford University Press.

Feenstra RM (1994) Scanning tunneling spectroscopy. *Surface Science* **299/300**: 965–979.

Sarid D (1991) *Scanning Force Microscopy: With Applications to Electric, Magnetic, and Atomic Forces*. Cambridge: Cambridge University Press.

Stroscio JA and Kaiser WA (1993) *Methods of Experimental Physics: Scanning Tunneling Microscopy*. New York: Academic Press.

Wiesendanger R (1994) *Scanning Probe Microscopy and Spectroscopy: Methods and Applications*. Cambridge: Cambridge University Press.

# Scanning Probe Microscopy, Applications

**CJ Roberts**, **MC Davies**, **SJB Tendler** and **PM Williams**, The University of Nottingham, UK

Copyright © 1999 Academic Press

| SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS |
|---|
| Applications |

The family of scanning probe microscopes (SPMs) have revolutionary imaging capabilities on a range of materials. For example, atomic resolution images of metal and semiconductor surfaces produced by the scanning tunnelling microscope (STM) or images of individual biomolecules in aqueous environments recorded by atomic force microscopy (AFM) are now routine in the literature. Perhaps, less well known, is the even greater potential of these and other probe-based techniques to produce spatially resolved spectroscopic information at the atomic or molecular level. Following a brief introduction to the principal SPMs available, this article will review as comprehensively as possible the wide ranging applications of spectroscopic SPM in semiconductor, material and life sciences.

## Methods and instrumentation

The term SPM encompasses a family of surface-sensitive techniques, each based upon the interrogation at the nanometre level of a specific physical property by a sharp proximal probe. For example the original SPM, the STM measures local conductivity and the AFM local surface hardness. **Figure 1** provides a summary of the operation of the most popular types of SPM. Extensive discussions on the operations of the various forms of SPMs can be found in numerous reviews of the subject. Here we briefly highlight the three SPMs most readily applied to spectroscopic measurements, the STM, the AFM and the near-field scanning optical microscope (NSOM).

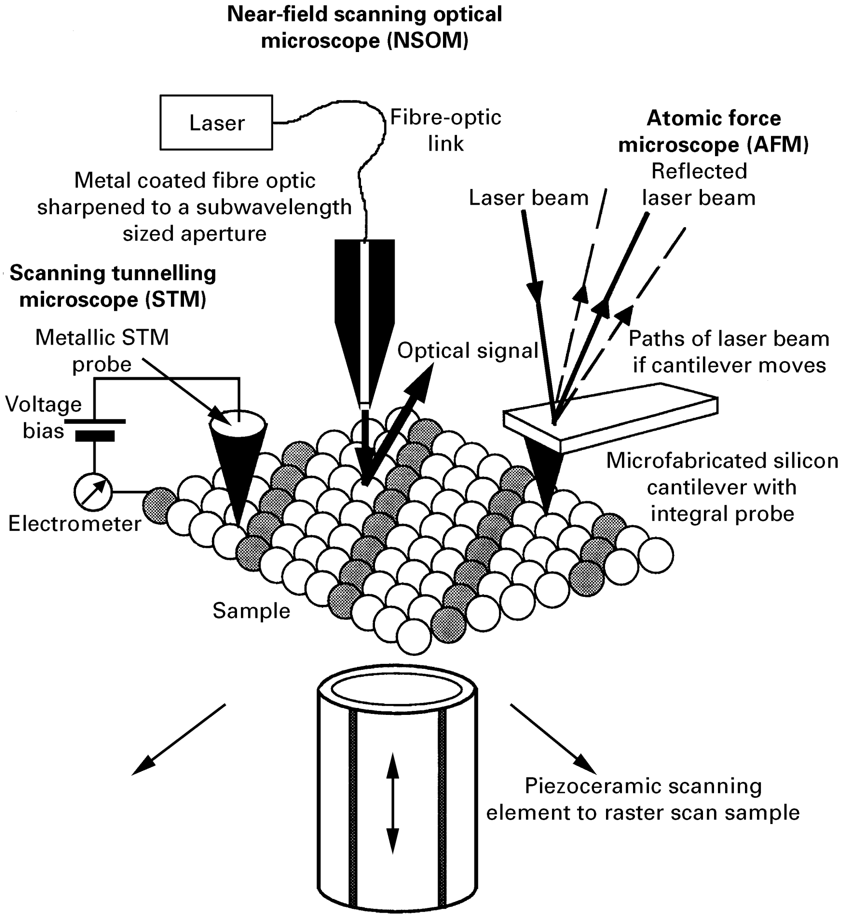



**Figure 1** Schematic representation of the key components of a scanning tunnelling microscope (STM), a near-field scanning optical microscope (NSOM) and an atomic force microscope (AFM). In a STM, a sharp metallic probe is brought into close proximity with a conducting sample. A small bias voltage between probe and sample causes a tunnelling current to flow. This current is recorded as the sample is scanned beneath the probe. In NSOM, the sample is placed in the near-field region of a subwavelength-sized light source. The transmitted or reflected optical signal is used to form an image of the scanning sample. AFM microscopy relies upon the effect of repulsive and attractive forces between the probe and sample to bend a supporting cantilever. The bending of the cantilever, and hence the force, is extracted by monitoring the path of a laser beam reflected from the back of the cantilever.

## STM

It was noted early in the use of STM that the appearance of a surface, particularly at atomic resolution, often changes dramatically with applied sample-tip bias voltage. This phenomenon results since electrons tunnelling between the tip and the surface do so between discrete electronic states and these states change with applied bias. In the extreme case of changes in tip bias polarity, either occupied (tip positive) or empty states (tip negative) in the sample are responsible for image contrast. Hence, it is possible to 'image' the location of specific bonding and antibonding orbitals and, with care, identify surface species by their electronic signature.

The first type of spectroscopy investigated by STM was current ($I$) versus gap distance ($s$). Here the tunnelling current is recorded as the metallic STM probe approaches a sample surface. The local barrier height ($\phi$) of the surface can also be estimated from this $I$–$s$ data using, $\phi = 0.952(\mathrm{d}\ln I/\mathrm{d}s)^2$ where $s$ is in Å (note that $\phi$ is a strong function of tip shape, a generally unknown parameter). Electronic structure has also been studied by ramping the tip bias voltage ($V$) while recording the resultant tunnelling current. This measurement can be carried out at a single point or at each point in a topography scan, hence producing spatially resolved spectroscopic data. Plotting $\mathrm{d}\ln I/\mathrm{d}\ln V$ versus $V$ has been shown to be proportional to the density of states in the low voltage limit ($\phi > V$).

## AFM

Since its inception by Binnig *et al.* in 1986, AFM has become an important and widespread tool for

imaging surface topography with nanometre resolution. AFM is essentially a very sensitive profilometer, measuring surface topography using a sharp stylus, or probe, mounted on a soft spring, or cantilever (**Figure 1**). The ability of AFM to image samples in ambient or aqueous environments is particularly attractive in biomolecular and electrochemical studies.

By exploiting the local nature of an AFM probe and its pico-Newton force sensitivity considerably more information can be extracted from AFM than just surface topography. For example local tribology and force-probe–sample separation spectra. During such force probe–sample measurements the probe is moved towards the surface at constant velocity until it is brought into contact with the sample up to a predetermined point of maximum load. The direction of motion is then reversed and the probe is withdrawn from the sample surface. As the probe is withdrawn from the sample the probe may stick to the surface due to interactions between the probe and the sample. The magnitude of this 'sticking' force and its temporal evolution can reveal details of the type and dynamics of the forces occurring between probe and surface.

## NSOM

NSOM represents one of the most promising optical techniques that aims to overcome the Abbé barrier, and yet retain most of traditional optical microscopy's utility. A NSOM typically illuminates a local area of a sample by transmitting laser light through a subwavelength sized aperture, as defined by the end of a tapered metal-coated optical fibre. An image is then formed by raster-scanning the aperture close to the sample surface and collecting either the transmitted or the reflected light (**Figure 1**). In such a regime, it is the aperture of the NSOM probe that determines the ultimate resolution.

Since NSOM utilizes optically based contrast it has the potential to exploit optical spectroscopies with resolution comparable to the probe dimensions, i.e. tens of nanometres. For example, steady state and time-resolved fluorescence spectroscopy and Raman spectroscopy have been demonstrated with NSOM. Despite some notable successes it is important to note that the combination of the difficulty in the interpretation of NSOM data and the often poor optical efficiency of NSOM probes, presently makes the application of NSOM the most challenging form of SPM.

## Applications of spatially resolved SPM

The breadth of applications of SPM for spectroscopic measurements is considerable. In order to highlight key examples, the discussion is classified by the nature of the sample investigated, ranging from atomic scale studies on semiconductors and metals to the study of biomolecular interactions.

### Semiconductors

The first atomically resolved scanning tunnelling spectroscopy (STS) data were obtained by Hamers *et al.* (1986) on Si(111)-(7 × 7), showing chemically inequivalent atoms within each unit cell. Since this time many semiconductors have been studied by STS studies carried out under ultrahigh vacuum. For example, $dI/dV$ curves recorded from In atoms adsorbed onto the Si(111)-(7 × 7) surface show that covalent bonding between the In and Si surface states saturates the Si intrinsic metallic states. STS studies of Li on *p*-type Si(001) show strong negative differential resistance and the related existence of thermally activated electron traps. Spatially resolved STS has also distinguished inequivalent sites on a roughened Si(001) surface. Spectra recorded from terraces show bonding and antibonding states at +0.5 V and –0.5 V; however, Si atoms recorded from a step show a marked metallic character. It should be noted that a quantitative analysis of such spectra can be problematic due to the generally unknown nature of the tip's electronic structure and the change in shape of the tunnelling barrier with applied tip–sample voltage bias. Nevertheless, such sensitivity has been very successfully exploited for the study of semiconductor surfaces.

Although many methods exist for the optical characterization of semiconductor surfaces, when high spatial resolution is required, for example with an inhomogeneous surface, new techniques are required. Here NSOM has significant potential and has been employed to map photoluminescence intensity simultaneously with topography on quantum well structures and hence local carrier density. The data was shown to be consistent with a diffusion-based model and the existence of short-lifetime carriers at the quantum-well boundary. Low-temperature NSOM at 4.2 K has been used to show a vertical dependence of spectral shape in GaAs quantum dots. NSOM has also been employed to acquire Raman spectra from rubidium-doped regions of $KTiO_2PO_4$ sample, although along acquisition times and very small signal levels have limited progress. Despite this, NSOM Raman has also been successfully

employed to map residual stresses in silicon wafers associated with deformation (**Figure 2**).

## Metals

In comparison to semiconductor surfaces, metals display smaller corrugations in their density of states due to very delocalized bonding. Nevertheless, early STM studies revealed relatively high atomic corrugations on Au(111). It has since been shown that this 'super' resolution results from the nature of the electronic density of states on face-centred cubic metal surfaces.

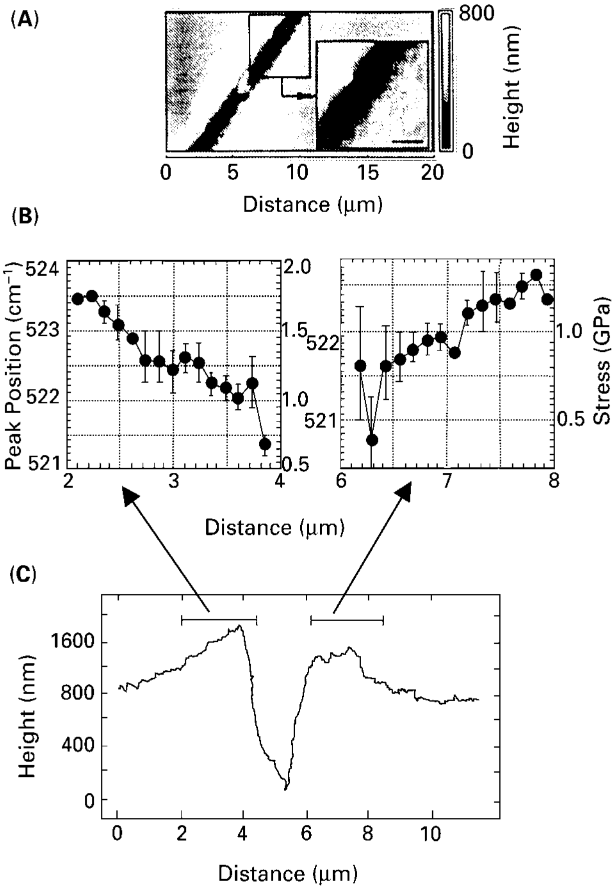



**Figure 2** (A) Topography image recorded using NSOM showing a scratch on a silicon wafer surface. Inset is an enlargement of the area that was Raman mapped. The scale bar is 1 μm. An array of 26 by 21 spectra were recorded with step sizes of 154 nm and 190 nm in the *X* and *Y* directions, respectively. Each spectrum took 60 s to acquire giving a total image acquisition time of just over 9 h. In (B) the value of centre frequency of the silicon band was extracted and is shown as a function of distance across the scratch; the lateral position of the data points is shown on the topographic cross section (C). Reprinted with permission from Webster S, Batcheldes DN and Smith DA (1998) Submicron resolution measurement of stress in silicon by near-field Raman spectroscopy. *Applied Physics Letters* **72**, 1478–1480.

Elemental-specific contrast on metals using STS has been demonstrated for copper on W(110) and Mo(110) surfaces. Resonant tunnelling via surface and image states provide elemental identification. Theoretical treatment of elemental identification of adsorbates on metals indicates that up to 2 Å peaks should be present in images of electropositive elements on Pt(111) and that 0.35 Å depressions would result from electronegative oxygen atoms.

Low-temperature ultrahigh-vacuum STM has been used to perform atomically localized spectroscopic measurements on single Fe atoms adsorbed onto the Pt(111) surface. Using d*I*/d*V* spectra a resonance was found to occur in the adatom local densities of states that is centred 0.5 eV above the Fermi energy. This feature had a width of approximately 0.6 eV, and occurred only when the tip was within angstroms (laterally) of the centre of an Fe adatom. Following on from this work it was found that Fe adatoms strongly scatter metallic surface state electrons and hence are good building blocks for constructing atomic-scale barriers to confine electrons. 'Quantum corral' barriers constructed by individually positioning Fe adatoms using the STM tip reveal, via STS, discrete resonances inside the corrals, consistent with size quantization (**Figure 3**).

## Superconductors

Cryogenic STS is an ideal tool for studying the electronic nature of superconductors and has produced local d*I*/d*V* spectra at the BiO cleavage planes of a bismuth cuprate superconductor at 4.2 K. The spectra confirm a large gap parameter associated with an apparently gapless density of states on the uppermost layer. Spatial variations of the gap parameter on a 100 Å scale were attributed to variations in BiO metallicity with two characteristic d*I*/d*V* spectral shapes over regions of metallic and nonmetallic BiO layers. Low-temperature STS has also been employed to probe the local effects of magnetic impurities on superconductivity. Tunnelling spectra obtained near magnetic adsorbates reveal the presence of excitations within the superconductor's energy gap that can be detected over a few atomic diameters around the impurity at the surface. These excitations are locally asymmetric with respect to tunnelling of electrons and holes. A model calculation based on the Bogoliubov–de Gennes equations can be used to understand the details of the local tunnelling spectra.

## Polymers

New higher throughput NSOM probes have recently permitted the recording of Raman spectra from

(A)

(B)

(C)

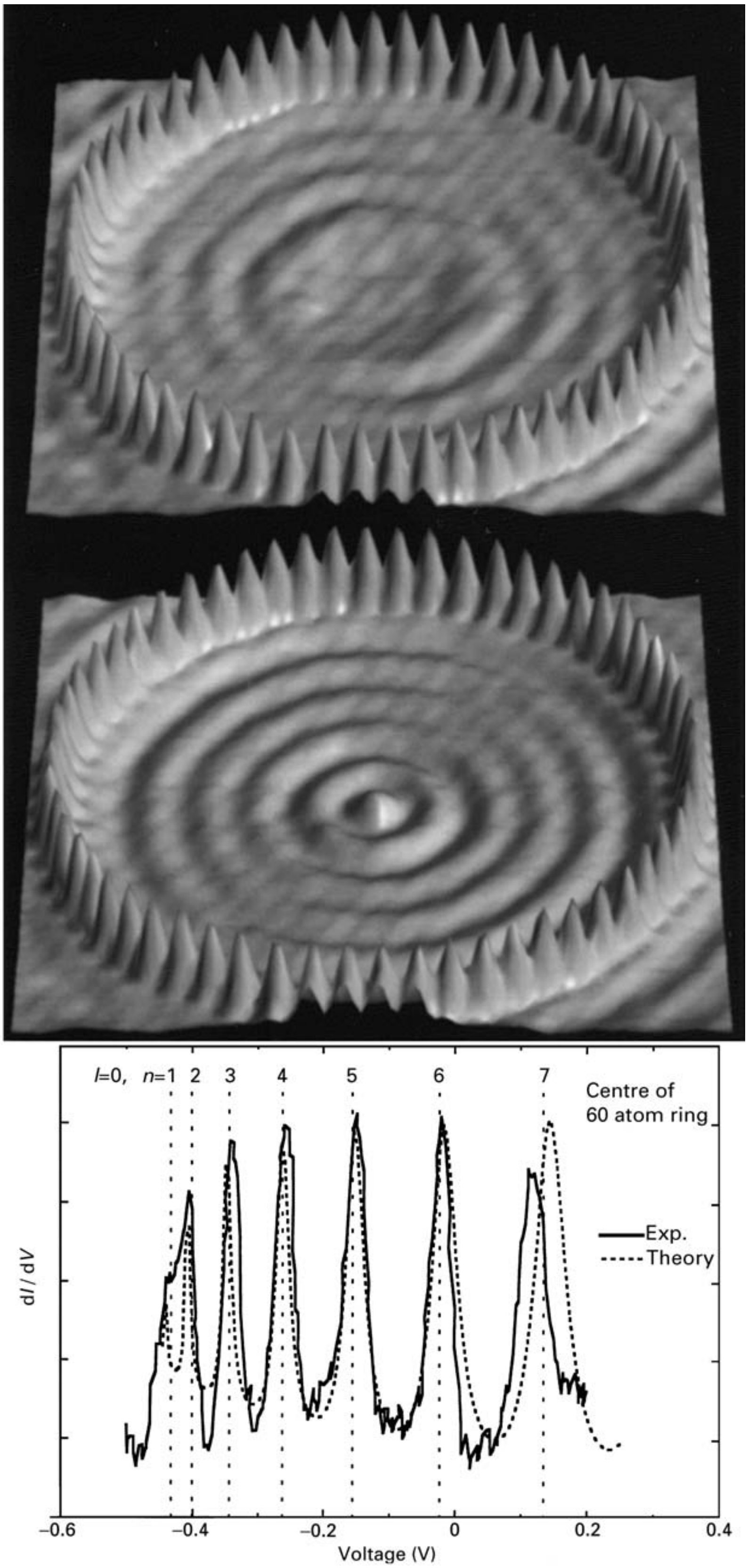



**Figure 3** Perspective views of a 60 atom Fe ring recorded at tunnelling current of 1 nA and tip bias voltages of (A) 10 mV and (B) –10 mV. The quantum interference patterns inside the ring change with energy. The energy dependence of the lowest density of states at the centre of the ring is illustrated by the d*I*/d*V* spectra in (C). The sharp peaks in the spectra indicate sharp resonances in the lowest density of states. These data match theoretical results based upon the particle-in-a-box model very closely. Reprinted from *Physica D* **83**: Crommie MF, Lutz CP, Eigler DM and Heller EJ, Quantum corrals, pp 98–108, 1993 with kind permission of Elsevier Science – NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

polystyrene spheres labelled with different dyes and adsorbed on silver substrates. No significant differences in near-field and far-field Raman spectra were observed with the NSOM data, clearly demonstrating true chemical identification. Nonresonance Raman spectra of polydiacetylene crystals (**Figure 4**) demonstrate the feasibility of acquiring spatially resolved Raman spectra despite very low signal levels.

NSOM has also been used to probe the excitonic transitions in J-aggregates of 1,1′-diethyl-2,2-cyanine iodide grown in poly(vinyl sulfate) thin films. Fluorescence spectra recorded as a function of the NSOM tip position along individual aggregates show only slight variations and are very similar to the bulk aggregate spectrum. The absence of spectral broadening is assigned as evidence for a uniform, well-ordered molecular structure within the aggregates.

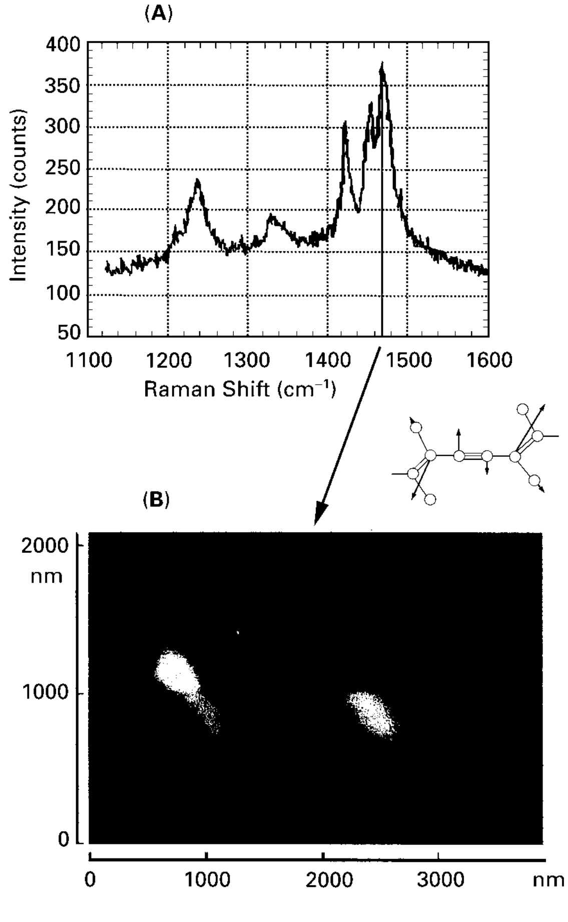



**Figure 4** (A) A pre-resonant near-field Raman spectrum of a polydiacetylene microcrystal taken using 633 nm excitation, approximately 150 nm aperture and a 60 s exposure. (B) A 3 μm × 2 μm near-field Raman image of the polydiacetylene microcrystallites. A Raman spectra as in (A) is obtained at every point in the image, and the intensity of any peak may be chosen to produce a grey scale image, in this case the 1485 $cm^{-1}$ feature. The vibration mode responsible for this line is shown.

## Biological systems

NSOM fluorescence image and spectrographs of recombinant *Escherichia coli* cloned to produce green fluorescence protein (GFP) show a difference in fluorescence distribution within individual bacteria. Fluorescence activity of GFP can thus be used as a convenient indicator of transformation. Improvements in NSOM probe–sample distance control have facilitated the fluorescence imaging of thick biological specimens, such as neurons, astrocytes and mast cells, which also fluoresce in the far-field and hence would normally reduce optical resolution. NSOM has also been used to provide high-resolution information on *in situ* interactions between proteins in biological membranes, in particular human red blood cells invaded by the malaria parasite, *Plasmodium falciparum*. During infection, the parasite expresses proteins that are transported to the cell membrane. Host and parasite proteins were selectively labelled in indirect immunofluorescence antibody assays, and simultaneous NSOM dual-colour excitation fluorescence maps produced.

Presently karyotypes of human metaphase chromosomes are used to detect genetic defects like deletions or translocations, where the chromosomes are treated by the trypsin–Giemsa protocol, to produce a typical banding pattern and imaged by optical microscopy. Because of the diffraction limit in optical microscopy, even the smallest visible band contains around 1 million base pairs. Improved resolution has been demonstrated using fluorescence NSOM on the treated chromosomes compared to conventional light microscopy.

## Single molecule studies

Spatially resolved STS has been used to characterize the electronic structure of $C_{60}$ molecules on a range of substrates including Au(001), Au(110), Au(111) and Al(111). Due to a lattice mismatch between the overlayer $C_{60}$ and the substrate Au(100) surface, a uniaxial stress is applied resulting in several types of oblique lattices and modified electron charge density around the $C_{60}$ molecules. Charge transfer from the substrate to the molecules and intermolecular bonding under stress were observed in STS data. STS also clearly differentiates inequivalent adsorption sites on Au(111) and Al(111). The STM tip has been used to locally excite single $C_{60}$ molecules to luminesce with an emission spot size of 0.4 nm. Fullerenes have been

employed as STM tips, showing improved performance when studying graphite surfaces.

A number of groups have employed NSOM to record fluorescence spectra from single dye molecules. Fluorescence spectra taken in the near-field showed no broadening due to long-range inhomogeneities as are apparent in far-field spectra. Using picosecond light pulses, time-resolved near-field fluorescence images of single sulforhodamine 101 dye molecules and rhodamine 6G have been recorded. Since metal surfaces near radiating dipoles influence fluorescence lifetimes, the fluorescence decay of single molecules is dependent on the relative position of the tip and the molecule. Polarization-sensitive NSOM has been used to resolve mesoscopic spectral inhomogeneities in small crystals of the dye 1,1-diethyl-2,2-cyanine iodide, where the crystals showed strong absorption perpendicular to their long axis and no absorption in the two orthogonal directions.

The sensitivity of fluorescence resonance energy transfer (FRET) has been extended to the single-molecule level by measuring energy transfer between single donor and acceptor fluorophores linked by a short DNA molecule using NSOM. Dual colour images and emission spectra combined with photodestruction dynamics have been used to determine the presence and efficiency of energy transfer. In contrast to ensemble measurements, dynamic events on a molecular scale are observable in single-pair FRET measurements because they are not cancelled out by random averaging.

The ability of AFM to directly measure discrete intermolecular forces as low as 10 pN was highlighted as long ago as 1993. Since then, a number of groups have exploited this ability, using AFM to determine the forces required to separate individual receptor-ligand molecules including avidin-biotin, cell adhesion proteoglycans, antibody–antigen and hydrogen bonding between nucleotide bases. In addition, the potential for mapping surface groups by their functionality using AFM has been exploited to spatially locate the adhesive and frictional interactions between hydrophobic and hydrophilic organic monolayers and biotin–streptavidin (see **Figure 5**).

Recently molecular dynamics has been used to model the disruption of biotin–streptavidin as it occurs in force adhesion experiments and to relate these forces to molecular structure and conformation. The force is calculated from the steepest slope in the free energy profile along the unbinding pathway. Interestingly the calculated rupture forces show a similar spread in values as is found in experimental data, suggesting that this spread is due to heterogeneity in the reaction pathway of biotin–streptavidin.

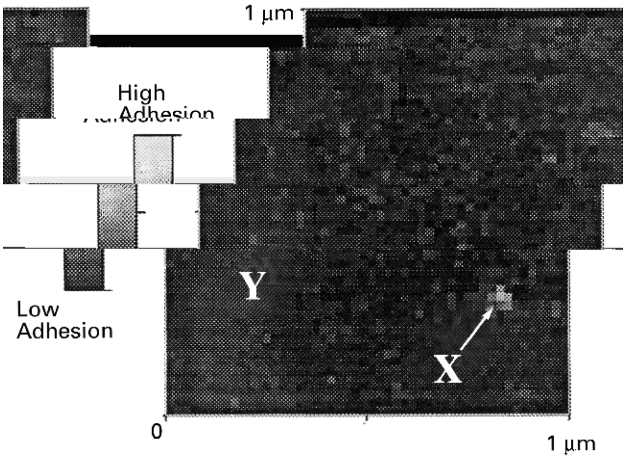



**Figure 5** A spatially resolved force adhesion map recorded using an AFM probe coated in biotin on a 90% biotin-blocked streptavidin surface. Biotin–streptavidin are a ligand–receptor pair with very high affinity ($10^{15}$ $M^{-1}$) often employed as a model system for molecular recognition studies. The contrast on the image corresponds to the amount of adhesion felt by the probe at the surface. Light areas represent high levels. If the biotin-coated probe contacts the surface in a region where free binding sites exist (i.e. streptavidin unblocked by free biotin) then adhesion based on the biotin–streptavidin specific molecular interaction would be expected. Position X in the image corresponds to an area of adhesion and Y a typical area of no interaction. Hence, X marks the spot of an open streptavidin binding pocket. Reproduced with permission of Gordon and Breach Publishers from Roberts CJ, Allen S, Chen X, Davies MC, Tendler SJB and Williams PM (1998). Measurement of intermolecular forces using force microscopy: Breaking individual molecular bonds. *Nanobiology* **4**: 163–175.

AFM has also been used to measure inter- and intra-chain forces in DNA nucleotides. The interaction forces between complementary 20 base pair lengths of single-stranded oligonucleotides ($(ACTG)_5$ and $(CAGT)_5$) and the forces required to stretch and break polydisperse homopolymers of inosine were measured. AFM has also been employed to study the interaction between $d(T)_{20}$ (dT = 2′-deoxyribosylthymine) and poly(dA) oligonucleotides (dA = 2′-deoxyribosyladenine). As before, rupture forces after binding were measured; however, the dependence of the rupture force with time of contact between the probe and sample was also observed. After 30 s contact, only low adhesion was observed. A maximum probe sample adhesion was observed after 2 min. This incubation-time dependence was interpreted as slow reorientation of partially hybridized oligonucleotide strands into more stable structures and was suggested as a means of studying double-helix annealing processes.

Single-molecule force spectroscopy on dextran filaments linked to a gold surface has been carried out using AFM by vertical stretching, the applied force being recorded as a function of the elongation. At low

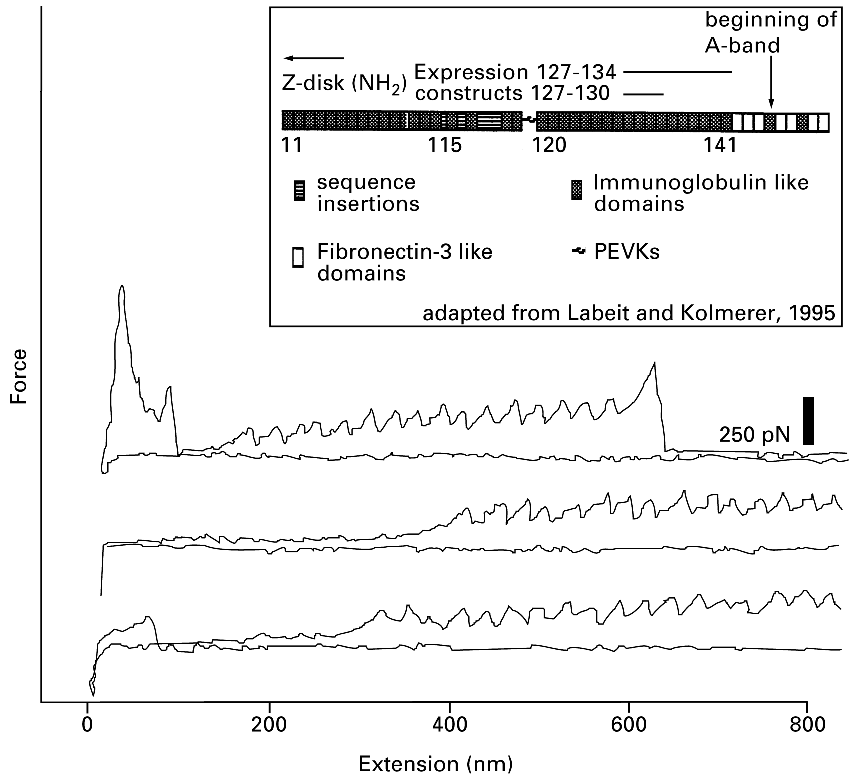



**Figure 6** Three typical force extension curves obtained by stretching titin molecules. The curves show periodic structure on the retract portion of the data consistent with the unfolding of individual titin domains. Reproduced with permission from Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE (1997) Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276,** 1109–1112.

forces the entropic deformation of dextran dominates and can be described by the Langevin function with a 6 Å Kuhn length. At elevated forces the strand elongation was governed by the twist of bond angles. At higher forces the dextran filaments underwent a distinct reversible conformational change. This ability to stretch and relax long-chain molecules has also been exploited to unfold individual domains of the giant sarcomeric protein of striated muscle, titin (**Figure 6**). At large extensions, the restoring force exhibited a sawtooth-like pattern. Measurements on recombinant titin immunoglobulin segments of two different lengths exhibited the same pattern and allowed the discontinuities to be attributed to the unfolding of individual immunoglobulin-like domains. The forces required to unfold individual domains ranged from 150 to 300 pN and depended on the pulling speed. Upon relaxation, refolding of immunoglobulin domains was observed.

## Future trends

The resolution and adaptability of scanning probe techniques are increasingly being exploited to carry out spectroscopy measurements, particularly for electronic and optical surface properties. In addition, new variants of traditional scanning probe spectroscopic methods continue to be developed. For example, STM spectroscopy performed with magnetic probe tips has yielded new information about the spin-resolved nanoelectronic properties of magnetic nanostructures. Also, an adaption of STM to allow the probing of surface acoustic waves is proposed that reaches submicrometre resolution for the quantitative evaluation of elastic constants and studies of nanoscale structures. In AFM, measuring the oscillation amplitude of the probing AFM tip and phase-shift between the cantilever response as a function of the tip–sample distance, allows the analyses of the dynamic interaction of the AFM tip with the sample surface. This has been termed dynamic force spectroscopy and has been proposed as a new method of rapidly mapping probe–sample interactions. Advances in NSOM spectroscopic applications are presently centred on improving the optical efficiency of the probe, either through more precise control of fibre optic probe geometry or through the development of semiconductor-based tips not dissimilar to those in use in AFM.

Other SPM technologies for specific applications are also being developed. For example, magnetic resonance effects have been studied using modified AFMs. The sample is mounted on an AFM cantilever in a magnetic field gradient and exposed to a radiofrequency field which drives the spins into precession. The resultant periodic force is sensed in the normal way from the flexure of the AFM lever. This technique has also been extended to spatially map magnetic resonance data and to detect NMR effects. These data exemplify the reason for the continued rapid growth in SPM applications. The adaptability of the technique to address problems from a range of scientific disciplines and its ability to operate under conditions of vacuum to liquid from 4 K to 1000 K, will ensure this pace of SPM advancement continues for the foreseeable future.

## List of symbols

$I$ = current; $s$ = gap distance (Å); $V$ = voltage; $\phi$ = barrier height.

*See also:* **Scanning Probe Microscopes; Scanning Probe Microscopy, Theory.**

### Acknowledgement

SJBT is a Nuffield Foundation Science Research Fellow.

## Further reading

Allen S, Davies J, Davies MC *et al* (1996) *In situ* observation of streptavidin–biotin binding on an immunoassay well surface using SFM. *FEBS Letters* **390**: 161–164.

Ambrose WP, Goodwin PM, Martin JC and Keller RA (1994) Alteration of single molecule fluorescence lifetime in near-field optical microscopy. *Science* **265**: 364–367.

Betzig E and Chichester RJ (1993) Single molecules observed by near-field scanning optical microscopy. *Science* **262**: 1422–1425.

Binnig G, Quate CF and Gerber C (1986) Atomic force microscope. *Physical Review Letters* **56**: 930–933.

Berndt R, Gaisch R, Gimzewski JK, Reihl B, Schlittler RR, Schneider WD and Tschudy M (1993) Photon emission at molecular resolution induced by a scanning tunnelling microscope. *Science* **262**: 1425–1427.

Crommie MF, Lutz CP and Eigler DM (1993) Spectroscopy of a single atom. *Physical Review B* **48**: 2851–2854.

Crommie MR, Lutz CP, Eigler DM and Heller EJ (1993) Quantum corrals. *Physica D* **83**: 98–108.

Grubmüller H, Heymann B and Tavan P (1996) Ligand binding: molecular mechanics calculation of the streptavidin–biotin rupture force. *Science* **271**: 997–999.

Hallen HD, Larosa AH and Jahncke CL (1995) Near-field scanning optical microscopy and spectroscopy for semiconductor characterization. *Physica Status Solidi A* **152**: 257–268.

Hamers RJ, Tromp RM and Demuth JE (1986) Surface electronic structure of Si(111)-(7×7) resolved in real space. *Physical Review Letters* **56**: 1972–1975.

Hamers RJ (1996) Scanned probed microscopies in chemistry. *Journal of Physical Chemistry* **100**: 13103–13120.

Higgins DA and Barbara PF (1995) Excitonic transitions in J-aggregates probed by NSOM. *Journal of Physical Chemistry* **99**: 3–7.

Hoh JH, Cleveland JP, Prater CB, Revel JP and Hansma PK (1993) Quantized adhesion detected with the atomic force microscope. *Journal of the American Chemical Society* **114**: 4917–4919.

Kazantsev DV, Gippius NA, Oshinivo D and Forchel A (1996) Spectroscopy of GaAs/AlGaAs microstructures with submicron spatial resolution using a near-field scanning optical microscope. *JETP Letters* **63**: 550–554.

Lee GU, Kidwell DA and Colton RJ (1994) Sensing discrete streptavidin–biotin interactions with atomic force microscopy. *Langmiur* **10**: 354–357.

Rief M, Oesterhelt F, Heymann B and Gaub HE (1997) Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science* **275**: 1295–1297.

Rugar D, Yannoni CS and Sidles JA (1992) Mechanical detection of magnetic resonance. *Nature* **360**: 563–566.

Shao Z and Yang J (1995) Progress in high resolution atomic force microscopy in biology. *Quarterly Reviews in Biophysics* **28**: 195–251.

Smith DA, Webster S, Ayad M, Evans SD, Fogherty D and Batchelder D (1995) Development of a scanning near-field optical probe for localized Raman spectroscopy. *Ultramicroscopy* **61**: 247–252.

Wolf EL, Chang A, Rong ZY, Ivanchenko YM and Yu FR (1994) Direct mapping of the superconducting energy gap in single crystal $Bi_2Sr_2CaCuO_{8+X}$. *Journal of Superconductivity* **7**: 355–360.

Zeisel D, Dutoit B, Deckert V, Roth T and Zenobi R (1997) Optical spectroscopy and laser desorption on a nanometer scale. *Analytical Chemistry* **69**: 749–754.

# Scanning Probe Microscopy, Theory

**AJ Fisher**, University College London, UK

Copyright © 1999 Academic Press

**SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS**

**Theory**

## Introduction

The three most important scanning probe techniques are

- Scanning tunnelling microscopy (STM);
- Scanning force microscopy (SFM, also known as atomic force microscopy, AFM);
- Scanning near-field optical microscopy (SNOM).

The three methods give different types of information, and require correspondingly different theoretical treatments. STM probes the electronic states of a surface; SFM probes the force (or force gradient) between a tip and a surface; while SNOM probes the electromagnetic field near a surface.

However, all three techniques share several common features. First, they measure local, not average, surface properties. Any theory must therefore include the local surface properties if it is to be useful. Second, they all lack a simple *inversion theorem*: in no case is it possible to infer directly physical properties of the system from the scanning probe results. Interpretation therefore has to proceed by an indirect 'interpretation cycle':

1. Build a model of the relevant local features (e.g. structure, excitations) of the system under study;
2. Develop a theory of the scanning probe experiment concerned;
3. Combine (1) and (2) to determine the predicted experimental signal from the model adopted;
4. Alter the model if the predictions and the experiment do not match.

In this article we shall examine what type of model of the physical system under study is appropriate under item (1) of the 'interpretation cycle' for each technique, and how a suitable theory of the experiment can be constructed for item (2).

## The scanning tunnelling microscope: electronic spectroscopy

### General considerations

The fundamental physical process in STM is the tunnelling of electrons between the tip and the sample under study, through the barrier formed by the vacuum between them (see **Figure 1**). The 'height' of this barrier in energy is approximately equal to the work functions of the tip or sample material. In the simplest possible one-dimensional model, we assume that the electron potential energy $V$ takes a constant value $V_0$ through the tunnelling gap; the barrier height is therefore $(V_0 - E)$ where $E$ is the electron energy. The electron wavefunctions then decay in the vacuum like $\exp(-\kappa z)$, where $z$ is the coordinate normal to the surface and

$$\frac{\hbar^2\kappa^2}{2m_e} = (V_0 - E) \qquad [1]$$

If $V_0 - E$ is $5\ \text{eV} = 8.01 \times 10^{-19}\ \text{J}$ then $\kappa = 1.15 \times 10^{10}\ \text{m}^{-1}$.

Tunnelling can occur from tip to sample and from sample to tip. If no bias is applied to the system (i.e. if the electrochemical potentials of the electrons in tip and sample far from the junction are equal) the rates of tunnelling in opposite directions are equal, and no net current flows. (Note that if the tip and sample have different work functions, a finite charge transfer will occur at zero bias to establish a dipole layer at the surfaces, and hence an electric field in the vacuum gap; it is the potential difference arising

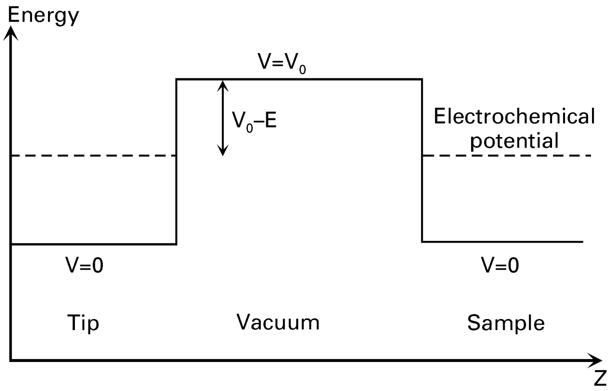



**Figure 1** Schematic diagram of an STM junction at zero bias, illustrating the meaning of the symbols defined in the text.

from this field which equalizes the electrochemical potential in the two materials.)

Suppose now that a finite bias potential $\Delta\Phi$ is applied to the system (see **Figure 2**), of a sign which raises the electrochemical potential for electrons on the left of the junction by $|e|\Delta\Phi$ relative to those on the right. Over a range of energies, electrons are now more likely to tunnel from left to right than *vice-versa*, and a net current flows from right to left. If the difference in chemical potentials is small so that current is dominated by electrons with a single energy $E$, we can use the fact that the current is proportional to the tunnelling probability and hence to the absolute square of the wavefunction to deduce that it will vary with the tip-sample separation $d$ like $\exp(-2\kappa d)$, with $\kappa$ given by Equation [1]. Taking the value of $\kappa$ we estimated earlier, we obtain the often-quoted rule of thumb that the tunnel current should reduce by roughly a factor of ten whenever the tunnel gap is increased by 1 Å = $10^{-10}$ m.

This is an approximate theory of the tunnelling process, but it says nothing about the contrast to be expected when the STM tip is moved across the surface. A better theory must take account of the atomistic structure of the tip and the surface, as well as a better theory of the tunnelling between them. In doing this, it is important to realize that the energy of the tunnelling electrons being used to probe the system is very similar to the energy of electrons in the bonding orbitals holding the atoms together. There is therefore a very close relationship between the tunnelling process, the electronic structure, and the atomic (or chemical) structure of the system.

Step (1) of the 'interpretation cycle' for STM must therefore involve a model of the atomic and electronic structure of the surface, including any adsorbates or surface defects. In practice this is most often obtained numerically using density-functional theory, in which the total energy of the electrons in the system is calculated from the electronic charge density, rather than from the full many-electron wavefunction. The Hartree–Fock method, which employs an approximate form for the many-electron wavefunction which neglects the correlations between the motions of the electrons, is also used. Such calculations are now relatively standard, and many can be found in the literature for surfaces of different types. Step (2) must involve a three-dimensional theory of electron tunnelling between the surface (represented in this way) and the tip; we now turn to this more difficult step.

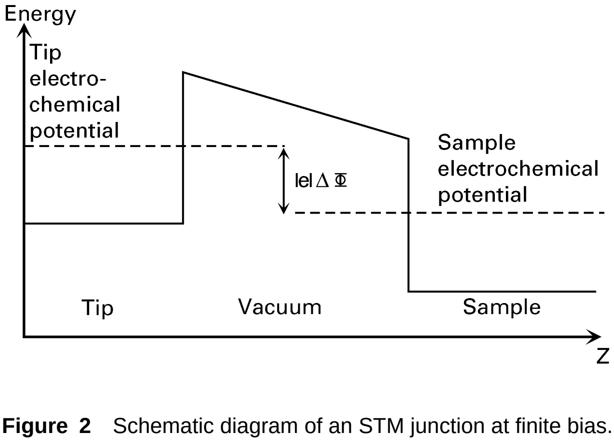



**Figure 2** Schematic diagram of an STM junction at finite bias.

### Perturbation theory

The interpretation of many spectroscopies (for example, optical spectroscopy) proceeds by the identification of a well-defined 'perturbation' which is applied to the system when the experiment is performed. This is both convenient (because the response of the system to the perturbation is not too difficult to evaluate in terms of the matrix elements of the perturbation) and conceptually useful (because it allows a clear separation between the 'system' and the 'probe'.

This is not straightforward in STM. There are two reasons: (i) the tip and the sample may be very close, and hence strongly coupled together, and (ii) even when this is not the case, there are mathematical difficulties in separating the Hamiltonian into parts describing a noninteracting tip and sample, because the kinetic energy operator for the electrons appears in both parts. Nevertheless problem (ii) has been solved, and it has been shown that a sensible perturbation theory can be constructed in which the appropriate matrix element is that of the electron current density operator, evaluated over a surface $S$ separating the tip and the sample (see **Figure 3**). We write

$$M_{ts} = \frac{\hbar^2}{2\mathrm{i}m_\mathrm{e}} \int_S \mathrm{d}^2\boldsymbol{r}\,[\psi_t \nabla \psi_s - \psi_s \nabla \psi_t] \qquad [2]$$

where $\psi_t$ is a one-electron state of the tip (in the absence of the sample) and $\psi_s$ is a state of the sample (in the absence of the tip). Note that in order to derive this result, one has to assume that *both* states are valid solutions of the Schrödinger equation in the neighbourhood of the surface $S$; this implies that the potential for electrons at $S$ must be equal to the vacuum potential. The transition rate for electrons from state $t$ to state $s$ (or *vice-versa*) can then be written

$$W_{ts} = \frac{2\pi}{\hbar} |M_{ts}|^2 \delta(E_t - E_s) \qquad [3]$$

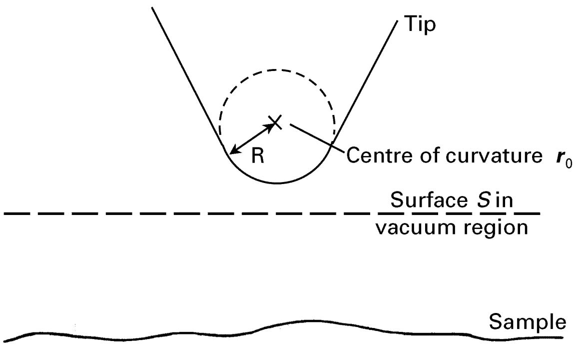



**Figure 3** The important quantities in the Tersoff–Hamann model of STM. The matrix element is evaluated on the surface $S$; the conductance is proportional to the sample density of states at the tip centre of curvature $r_0$.

and the total current from tip to sample as

$$I = \sum_{ts} \frac{2\pi}{\hbar} |M_{ts}|^2 \delta(E_t - E_s) f_t(E_t)[1 - f_s(E_s)] \quad [4]$$

where $f_t(E)$ and $f_s(E)$ are the occupation probabilities for electron states with energy $E$ in the tip and the sample respectively.

The most commonly used model in interpreting STM data is the Tersoff–Hamann model, in which the analysis is carried a step further. It is assumed that the tip wavefunction is an $s$-wave, and decays into the vacuum like

$$\psi_t(\boldsymbol{r}) = \Omega^{-1/2} \kappa R \exp(\kappa R) \frac{\exp(-\kappa|\boldsymbol{r} - \boldsymbol{r}_0|)}{\kappa|\boldsymbol{r} - \boldsymbol{r}_0|} \quad [5]$$

where $\Omega$ is a normalization volume, $\boldsymbol{r}_0$ is the centre of the curvature of the tip, $R$ is the radius of curvature, and $\kappa$ is as defined earlier. In this special case the integral in Equation [2] can be evaluated exactly (under the assumption that $\psi_s$ obeys the free-space Schrödinger equation), and one finds in the limit of small bias that the differential conductance $\sigma$ of the STM is

$$\sigma = \frac{32\pi^3}{\hbar} e^2 N_t(E_F) R^2 \kappa^{-4} \exp(2\kappa R) \sum_s |\psi_s(\boldsymbol{r}_0)|^2 \delta(E_s - E_F) \quad [6]$$

Here $N_t(E_F)$ is the total tip density of states at the Fermi energy. This is a very simple and important result; it tells us that the tunnelling conductance measures the sample density of states at the Fermi energy, evaluated at the centre of curvature of the tip (i.e. some distance outside the sample surface). This is relatively straightforward to calculate, and easy to interpret in simple chemical terms. The model disregards all details of the tip; they are absorbed into the values of the constants $R$ and $\Omega$. The (usually unknown) structure of the end of the tip can therefore be disregarded, at the cost of sacrificing any information about the absolute value of the conductance. It is largely because of these advantages that the Tersoff–Hamann model is so popular.

The approximations leading to Equation [6] are valid only if there is no electric field in the vacuum. Nevertheless, the Tersoff–Hamann model is often used to interpret images taken at finite bias voltage $\Delta\Phi$, or even data from the 'spectroscopic' mode of the STM in which the tip position is held fixed and the bias varied. The density of states involved in Equation [6] is projected onto a 'window' of energies $\Delta E = e\Delta\Phi$, rather than onto a single energy. There is no theoretical justification for this, as the true states of the system are bound to be modified by the addition of such a bias voltage, but it has proved useful as a way of qualitatively rationalizing STM data provided the bias is not too large.

It is possible to extend perturbation theory beyond the Tersoff–Hamann model, for example by including tunnelling to or from states of non-zero angular momentum on the tip, or by using states explicitly calculated from a particular atomistic model to find the matrix element in Equation [2]. However, both these approaches require additional information about the geometry of the tip and the electronic states it supports. This is generally not available from experiment, as a tip will be modified by the forces acting in the course of the experiment (as discussed in more detail below); even if the tip is well-characterized before use (for example, by electron microscopy or field-ion microscopy), this information will become out-of-date once the experiment starts.

Another extension of this type of perturbation theory is to the case where there is some additional electronic order in the tip or the sample – for example, magnetic or superconducting order. In the case of magnetic order one is led to consider separate currents of spin-up and spin-down electrons, proportional to the spin-resolved components of the density of states. For a superconductor, the tunnel current depends on the quasiparticle density of states.

### Beyond perturbation theory

Perturbation theory of this kind leads to an appealing picture of STM. Nevertheless it is not always justified; here we list some of the reasons why it may break down.

- A substantial redistribution of charge and potential takes place, so the effective one-electron Schrödinger equation is altered. This effect has been predicted theoretically when the tip-sample separation drops below about 3 Å; it tends to result in a lowering of the potential energy for an electron in the vacuum and a collapse of the tunnelling barrier.
- The electron tunnelling probability between tip and sample is not small. In practice this occurs only when the electron transport is no longer dominated by tunnelling – either because a physical contact or *nanojunction* has been formed between the two, or because the tunnel barrier has collapsed completely (see above). The signature of this state of affairs is that the STM conductance becomes of the order of the quantum of conductance, $e^2/h$.
- Although small, the tunnelling matrix element through the vacuum is not the smallest energy scale in the problem. This can occur when, for example, a highly insulating molecule is adsorbed on a surface; tunnelling through the molecule can then be just as difficult as tunnelling through the vacuum, so it is not appropriate to treat the vacuum tunnelling as a perturbation.

For all these problems, the theoretical cure is the same: one must perform a single coupled calculation for the electron states in the whole system (tip plus adsorbate – if any – plus sample) under a non-zero bias, allowing a current to flow.

However, this has proved to be very difficult without additional simplifications. In the elastic scattering quantum chemistry (ESQC) method developed by Joachim and Sautet, there is no self-consistency in the Hamiltonian for the electrons and only a relatively small basis set, giving very limited flexibility to the electron wavefunctions. In another approach, pioneered by the group of Tsukada, a more detailed numerical representation of the wavefunction is adopted: the wavefunctions are calculated on a mesh of points and full self-consistency is achieved between the wavefunctions and the electronic potential. The simplification in this case is that the wavefunctions far from the tunnel junction are those of a fictitious ‘jellium’ in which the positive charge of the nuclei is smeared out into a uniform background. In yet a third approach the conductance is calculated in a non-perturbation manner between two localized states, rather than between the true bulk states of the tip and sample.

### Other factors

There are also other factors that are known to be important in STM. One of these is the mechanical interaction between the tip and the sample; the forces that arise can distort the tip, with the result that the displacement of the tip apex is not the same as that recorded from the piezoelectric actuators controlling the tip (see **Figure 4A**). This effect was revealed by careful measurements of the *corrugation* of the STM image of adsorbates on metals as a function of the conductance.

A second effect is that of the tip electric field. This can be very large: fields above $10^9$ V m$^{-1}$ can be obtained when a potential difference of a few volts is

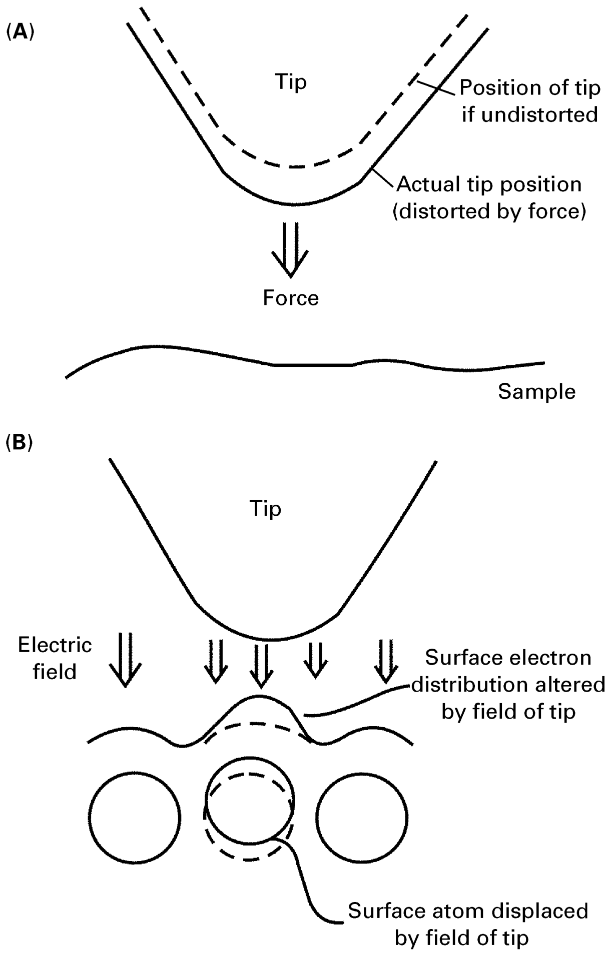



**Figure 4** Two physical phenomena which alter STM images from those predicted by simple theory: (A) Tip–sample forces distort the actual change in separation from that measured by the piezoelectric transducers. (B) Electric fields from the tip cause motion of surface atoms and distortion of surface electron states.

dropped over a narrow tunnelling gap. This can have two results: first it distorts the atomic structure of the surface, causing movements of a few tenths of an angstrom in the surface atoms, and second it distorts the electronic structure, changing the tunnelling probability at different points on the surface (see **Figure 4B**). These effects have been shown to be important in images of the Si(001) surface.

A third complicating effect is the inelastic scattering of electrons from other excitations during tunnelling. These other excitations may be electronic; the most important example is surface plasmons, which may be found in either the tip or the sample. Scattering from surface plasmons produces electromagnetic disturbances near the tunnel gap which can result in the emission of electromagnetic radiation from the STM. The frequency distribution of the emitted light is then characteristic of the surface plasmon spectrum. The distortion of the surface plasmons can also have other, more subtle, effects since it determines whether or not the electron experiences the effect of the 'image interaction' outside a surface. This can in turn have a large effect on the electron potential and the tunnelling current.

Alternatively, the scattering excitations may be atomic vibrations. These may result in phonon-assisted sidebands around resonant tunnelling peaks, corresponding to the absorption or emission of phonons. In extreme cases the transfer of electronic energy to atomic motion may produce atom transfer between tip and sample, or even desorption; this is a form of DIET (desorption induced by electronic transitions) and may be used to break bonds selectively on surfaces.

## The scanning force microscope: force spectroscopy

In order to interpret these experiments one needs to bear in mind the different types of forces that can act between the tip and the sample.

- At large distances the force most commonly present is the Van der Waals force. Between two atoms the Van der Waals force energy decays with separation $z$ according to the well-known $z^{-7}$ law, but for a sphere above a planar surface (one simple model for the tip–surface system) the decay is only as $z^{-2}$. This relatively slow fall-off tells us that in SFM, unlike STM, the large-scale structure of the tip is important.
- If the sample is an insulator, it may be locally charged. The interaction between these local 'patch charges' and the tip also decays like a power law in the tip–sample separation. The patch charges are difficult to control; the highest-resolution SFM results are generally obtained on conducting samples.
- At smaller distances (of the order of 3–5 Å separation) local interactions between the closest atoms of the tip and sample start to become important. These include the onset of covalent bonding, and local electrostatic forces.
- As the tip-sample separation drops below the sum of the atomic radii of the atoms, the Pauli exclusion principle raises the energy of the overlapping electron distributions, producing a repulsive force. If the tip and sample are forced together beyond this point, atomic deformations (first elastic, then plastic) occur.

### Models for the forces

Of these interactions, the Van der Waals attraction and the Pauli repulsion are universal; the presence of the others depends on the nature of the material. The combination of Van der Waals and Pauli interactions is often captured by the simple '6-12' Lennard-Jones interatomic potential

$$V(r) = \epsilon\left[\left(\frac{r}{\sigma}\right)^{-12} - \left(\frac{r}{\sigma}\right)^{-6}\right] \qquad [7]$$

in which the attractive $r^{-6}$ term represents the Van der Waals force and the repulsive $r^{-12}$ term the Pauli force. Simulations of generic interatomic interactions are often performed using this potential, although it cannot be expected to be realistic for anything other than interactions between the simplest rare-gas solids. More realistic calculations include approximate forms for the electrostatic and covalent interactions between the atoms, or (better still) find these forces directly from the electronic structure of the materials involved.

### High-resolution SFM operation

With this in mind, let us examine the most common modes of SFM operation when high-resolution information about the surface is required.

- **Non-contact mode**. In this mode the tip is kept at a distance from the sample in the attractive part of the force–distance curve; usually it is then scanned across the sample, and the tip–sample distance adjusted to keep the cantilever displacement (and hence the force) constant. This procedure keeps the tip in the region where the tip–sample force is (relatively) well understood, but with the price that the force is determined by the cumulative

effect of a large number of atoms – hence the resolution of individual atomic-scale features is seldom possible.

- **Contact mode**. Here, by contrast, the tip is allowed to penetrate into the repulsive regime of **Figure 5**. This has the advantage that one expects a large component of the force to be determined by a relatively small number of atoms near the tip apex, but the disadvantage that the force becomes dependent on complex atomic processes involving the irreversible deformation of the tip–sample junction. Images with apparent atomic resolution can be seen in contact mode on simple crystalline materials such as alkali halides, but the conclusion of careful simulations is that the atomic-scale features are not, in fact, correlated with the positions of atoms in the surface. This theoretical conclusion is reinforced by the failure to resolve atomic defects (known to be present on the surface) in experiments.

One might think that a technique intermediate between contact and non-contact modes could be devised simply by bringing the tip close to the surface, but not in contact with it. In fact this is very difficult because of the 'jump-to-contact' phenomenon: a static tip held above a surface by a SFM cantilever with a given force constant $k_{cant}$ can be stable only as long as the force gradient from the tip–sample interaction is less than $k_{cant}$ (see **Figure 5**). The force gradient of a Van der Waals interaction between a tip and a flat surface diverges as the separation between them is reduced, so this condition is always violated and the tip snaps into contact with the sample. If the tip is pulled off the surface, a similar jump out of contact occurs (although between different values of tip–sample separation).

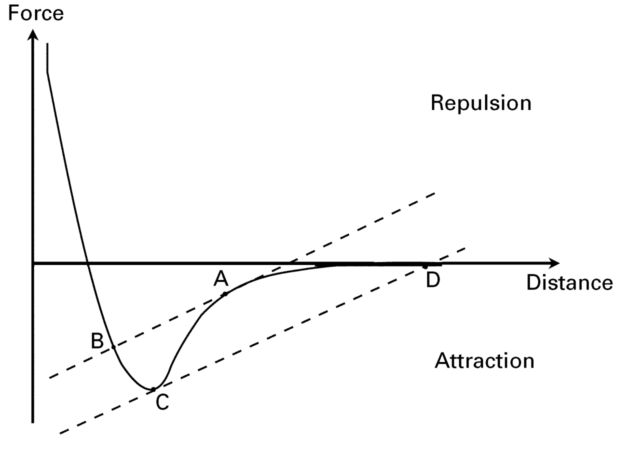



**Figure 5** Schematic force–distance curve for an SFM experiment. On approach, the tip jumps from point A to point B; on retraction, it jumps from C to D. The dotted lines have a slope equal to the cantilever force constant $k_{cant}$.

Since a very interesting range of tip–surface separations is rendered unavailable by the jump to contact, it would be desirable to eliminate it. To date, this has been done in two ways. First, a dynamical approach is used: the cantilever is vibrated above the surface with an amplitude of several hundred angströms, in such a way that its point of closest approach is only a few angströms from the surface. The difference from before is that the tip is accelerating rapidly away from the surface as it approaches; this suppresses the jump to contact. One way of expressing this is to say that the effective cantilever force constant is increased from $k_{cant}$ to $k_{cant} + M_{tip}\omega^2$, where $M_{tip}$ is the total mass of the vibrating tip and $\omega$ is the angular frequency of vibration. The tip is usually scanned while keeping the vibrational period constant; this corresponds approximately to a scan of constant force gradient. Atomic resolution has been obtained using this technique, initially on the Si(111)–$7 \times 7$ surface but now also on others. It seems this resolution can be understood in terms of the interaction between the tip and the surface near the point of closest approach, but the theory is complicated because the vibration of the tip samples all the different regions of the potential surface described above during a cycle, so a unified model containing all of them must be used.

A second approach is to control the force on the tip directly, generally by means of a small magnet mounted on the back. This removes the need to model a complicated tip oscillation, but imposes stringent demands on the response and stability of the electronics controlling the force. Direct measurements of tip–sample potential curves have now been reported using this technique, but comparison with theory is still in its infancy.

### Measurements of elastic properties

If local but not ultra-high-resolution measurements are required to probe the elastic properties of a hard material, there are advantages in using high-frequency measurements.

## The scanning near-field optical microscope-optical spectroscopy

The theory of scanning near-field optical microscopy is somewhat similar to that of STM, with the transport of light (or photons) replacing the transport of electrical current (electrons). Instead of the Schrödinger equation, the Maxwell equations for the electromagnetic field must be solved near the tip and

the sample, taking into account the local electromagnetic properties of each medium. In some respects this is easier, because (in the absence of non-linear media) the Maxwell equations are truly linear and no self-consistency (of the type needed between effective one-electron wavefunctions and the potential) is needed. Also, since the characteristic wavelengths and decay lengths for optical photons are much larger than atomic dimensions, a continuum treatment of the tip and sample materials is almost always sufficient. On the other hand, the Maxwell equations require treatment of two coupled vector fields.

As in the STM case, the equations must in practice be solved numerically. Perturbation theory is seldom employed, and most calculations make a direct solution for the optical modes at a fixed frequency, either by a transfer matrix approach, or by using the Dyson equation to obtain the solution from that of an exactly soluble system (for example, free space).

## Conclusions

STM appears to be the most subtle of the scanning probe methods, relying as it does on quantum-mechanical tunnelling. In fact the theory for this technique is the best developed of all the scanning probe family, but much progress remains to be made in accounting correctly for the nature of the tip and for tip–sample interactions. The theory of near-field optical microscopy is similar in spirit, and in some ways more straightforward. The understanding of SFM data is very incomplete, particularly for experiments with resolution on the atomic scale.

## List of symbols

$d$ = tip–sample separation; $e$ = electronic charge; $f(E)$ = occupation probability for electron state with energy $E$; $\hbar$ = Planck's constant divided by $2\pi$; $k_{cant}$ = force constant of SFM cantilever; $m_e$ = mass of electron; $M_{tip}$ = mass of vibrating tip assembly; $M_{ts}$ = matrix element connecting tip and sample states in STM; $r_0$ = centre of curvature of STM tip; $R$ = radius of curvature; $V$ = potential energy of electron; $W_{ts}$ = transition rate between tip and sample states; $\kappa$ = decay constant for electron wavefunctions $\psi$; $\sigma$ = differential conductance; $\Phi$ = electrostatic potential; $\omega$ = angular frequency of vibration; $\Omega$ = normalization volume for tip wavefunction in the Tersoff–Hamann model.

*See also:* **Scanning Probe Microscopes; Scanning Probe Microscopy, Applications; Surface Plasmon Resonance, Applications; Surface Plasmon Resonance, Instrumentation; Surface Plasmon Resonance, Theory.**

## Further reading

Briggs GAD and Fisher AJ (1999) Molecules on semiconductor surfaces: STM experiment and atomistic theory hand in hand. *Surface Science Reports* **246**: 1–81.

Blöchl PE, Joachim C and Fisher AJ (eds) (1993) *Computation for the Nanoscale*. Dordrecht: Kluwer Academic.

Chen J (1993) *Introduction to Scanning Tunnelling Microscopy*. Oxford: Oxford University Press.

Datta S (1995) *Electronic Transport in Mesoscopic Systems*. Cambridge: Cambridge University Press.

Israelachvili J (1992) *Intermolecular and Surface Forces*, 2nd edn. London: Academic Press.

Sautet P (1997) Images of adsorbates with the scanning tunnelling microscope: theoretical approaches to the contrast mechanism. *Chemical Reviews* **97**: 1097–1116.

Wiesendanger R (1994) *Scanning Probe Microscopy and Spectroscopy – Methods and Applications*. Cambridge: Cambridge University Press.

Wiesendanger R and Güntherodt H-J (eds) (1996) *Scanning Tunnelling Microscopy III*, 2nd edn. Berlin: Springer-Verlag.

# Scanning Tunnelling Microscopes

*See* **Scanning Probe Microscopes.**

# Scanning Tunnelling Microscopy

*See* **Scanning Probe Microscopy, Applications**

# Scanning Tunnelling Microscopy, Theory

*See* **Scanning Probe Microscopy, Theory**.

# Scattering and Particle Sizing, Applications

**F Ross Hallett**, University of Guelph, Ontario, Canada

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
| --- |
| Applications |

The scattering of radiation has been used for many years as a noninvasive technique for the determination of particle sizes in suspension and in aerosols. Two fundamentally different experimental approaches are used. In the first approach, fluctuations in the scattered light intensity arising from the diffusive motion of the particles are monitored and analysed by autocorrelation, a procedure that yields the intensity autocorrelation function, a function that corresponds to the Fourier transform of the frequency spectrum. The intensity fluctuation approach, when applied to light scattering, has been given a wide variety of names, including quasi-elastic light scattering (QELS), intensity fluctuation spectroscopy (IFS), photon correlation spectroscopy (PCS) and the currently favoured dynamic light scattering (DLS). Diffusing wave spectroscopy (DWS) is a special type of DLS technique which provides estimates of particle sizes in concentrated, nearly opaque, suspensions. The second approach involves the measurement of the intensity of light, X-rays or neutrons scattered by the particles as a function of the scattering angle. Depending on the source of radiation these are named static light scattering (SLS), small angle X-ray scattering (SAXS) and small neutron scattering (SANS).

When the scattering particles are small, monodisperse and spherical, size information can be obtained from either SLS or DLS data by fitting to relatively simple mathematical expressions. Generally, however, the particles are polydisperse (i.e. characterized by a size distribution) and sometimes they can be multimodal, and/or nonspherical. These properties of the sample can severely complicate the analysis and recovery of size information can require the use of relatively complex and indirect data analysis methods. This is because the direct analysis falls into a class of mathematical operations known as 'ill conditioned' transforms that are unstable and severely affected by truncation errors and noisy data. Much of the more recent work in light scattering has centred on finding analysis procedures that minimize these problems. The presence of polydispersity introduces a further consideration. Since the amount of light scattered by a particle is proportional to $r^n$, where $r$ is the particle radius and $n$ can be as large as 6, there is the consequence that the larger particles of a distribution can scatter considerably more light than the smaller ones. This leads to $z$-average particle sizes and intensity weighted size distributions in DLS. Conversion of these to number averaged quantities, such as those obtained by analysis of electron microscopic images, requires knowledge of the particle's scattering factor or structure factor and its incorporation into the analysis. These factors can be very simple for very small uniform spheres and

approach unity for point particles. When particular criteria are satisfied, approximately when particle diameters are significantly smaller than a wavelength, Rayleigh–Gans–Debye scattering factors can be used. However, once the particle's diameter is of the same order of magnitude or larger than the wavelength then Mie theory must be used to compute scattering factors. To date, Mie factors can be easily computed only for spherical or coated spherical systems.

## Basic physics of elastic scattering

### The scattering vector

During elastic scattering processes, the incident radiation is redirected at all scattering angles as a result of its encounter with the scatterers. By suitable use of pinholes or slits the light scattered at a particular angle, $\theta$, from the main beam can be detected and analysed (see **Figure 1**). The momentum transfer, which results from such a process, is defined as the vector difference between the incident wave vector, $\boldsymbol{k}_0$, and the scattered wave vector $\boldsymbol{k}_s$. This vector difference, called the scattering vector ($Q$), has a magnitude that can be obtained from the geometry of the scattering arrangement. That is,

$$|\boldsymbol{Q}| = \frac{4\pi n_0}{\lambda}\sin(\theta/2) \qquad [1]$$

where $n_0$ is the refractive index of the medium surrounding the particles and $\lambda$ is the wavelength (in a vacuum) of the incident light. It is clear from Equation [1] that $Q$ has the dimensions of an inverse length. Indeed, $Q^{-1}$ sets the length scale which will be probed by the light scattered at angle $\theta$. At low angles, $Q^{-1}$ is large and light scattered in this direction will contain information only on the gross dynamic and structural properties of the particles, whereas light scattered at a high angle contains this information at a finer scale and might, for example, carry information on the internal structure of the particle. **Figure 2** shows a comparison of scattering techniques, and their respective particle sizing ranges. DLS has a greater range than other scattering techniques because it is concerned with the distance ($2\pi/Q$) that a particle diffuses before dephasing of the scattered light occurs. Note also that due to the relatively short wavelengths of X-rays and neutrons (typically 0.1 to 1 nm as compared with visible light, 400 to 700 nm), $Q^{-1}$ is appropriate for particle sizing only when $\theta$ is very small. As a result, only small angle X-ray or neutron spectrometers are of interest here.

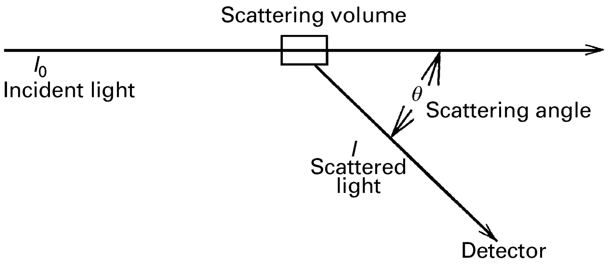



**Figure 1** The scattering geometry.

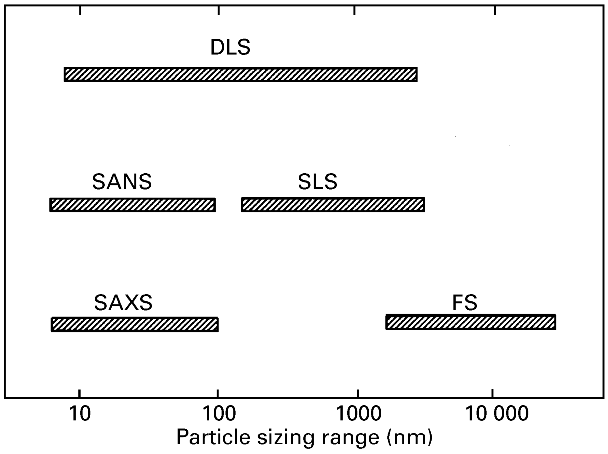



**Figure 2** Particle sizing ranges for different scattering instruments. DLS: dynamic light scattering, SANS: Small-angle neutron scattering, SLS: static light scattering, SAXS: small-angle X-ray scattering, FS: Fraunhofer scattering.

### Dynamic light scattering

DLS is an established technique for measuring the average size and the size distribution of particles in a suspension. The technique has advantages of being relatively fast, noninvasive, and requires minimal sample preparation (as compared with electron microscopy for example). But it does require low particle concentration. As well, dynamic light scattering results are often open to misinterpretation if one is unaware of the optical properties of the sample and the method of data analysis. The following discussion reviews some of the basic concepts of dynamic light scattering and outlines some of the pitfalls which are often encountered in data interpretation. A modification of DLS is DWS, which can be used to obtain approximate size information at high particle concentrations and is described later.

All DLS spectrometers use a laser as a source of coherent light. This means that all the light incident on the sample is in phase. At any instant the scattering

particles will have a particular set of positions within the scattering volume. Because of these positions, the relative phases of the electric fields of the scattered wavelets will differ at the detector due to the differing optical pathlengths that they must travel. The instantaneous electric field at the detector is the superposition of the fields due to all the scattered wavelets and will, at time $t$, have a value $E(t)$. At the time, $t+\tau$, which is a very small delay time, $\tau$, later than $t$, the particles, which are diffusing, will have new positions that are slightly removed from those at the earlier time. Superposition of the slightly shifted scattered wavelets will modify the total electric field at the detector to a new value. $E(t+\tau)$. Therefore, as time progresses, the electric field at the detector will fluctuate as the Brownian random walks that characterize the diffusion of the particles continue.

The main problem in data recovery in the DLS experiment is the extraction of quantitative information from a fluctuating signal. Since detectors are sensitive to light intensity (which is related to the square of the total electric field), small rapidly diffusing particles will yield rapidly fluctuating intensities, whereas larger particles and aggregates generate relatively slow fluctuations. In modern DLS spectrometers the rate of the fluctuations is measured by autocorrelation analysis of a real-time sequence of intensity data that has been digitized into time intervals (sometimes called sampling times) each of which are much shorter than a fluctuation time. The result is an intensity autocorrelation function, $g^{(2)}(\tau)$, where $\tau$ is an instrumentally delay time having a value $\tau = kt$, where $k$ is the channel number of the autocorrelation function and $t$ is equivalently the length of the sampling time or the delay time between one channel of $g^{(2)}(\tau)$ and the next. Ideally, the function (see **Figure 3**) has $\tau = 0$ limit equal to the normalized mean square intensity $\langle I(t)^2\rangle/\langle I(t)\rangle^2 = 2$ and decays to an asymptotic limit (as $\tau$ approaches infinity) equal to the normalized mean intensity squared, $\langle I(t)\rangle^2/\langle I(t)\rangle^2 = 1$. Practically, however, finite pinhole sizes and the detection optics reduce the $\tau = 0$ intercept from its ideal value.

The rate of decay of $g^{(2)}(\tau)$ is indicative of the typical fluctuation time of the scattering signal and of the rate of diffusion of the scatterers. Quantitatively, $g^{(2)}(\tau)$ can be related to the electric field autocorrelation function through the relation

$$g^{(2)}(\tau) = 1+|g^{(1)}(\tau)|^2 \qquad [2]$$

and when the scatterers are spherical and monodisperse, $g^{(1)}(\tau)$ is related to the translational

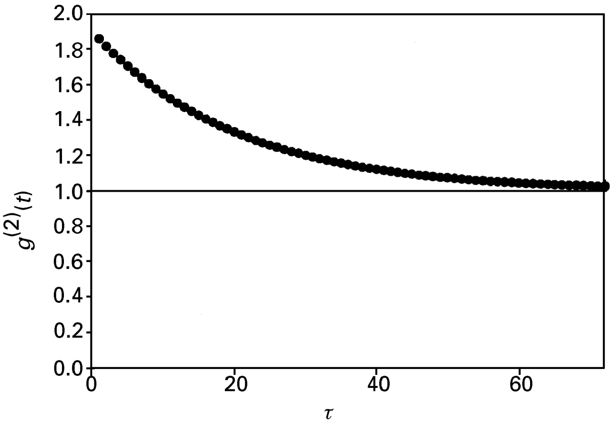



**Figure 3** A normalized intensity autocorrelation function.

diffusion coefficient, $D$, by

$$g^{(1)}(\tau) = \mathrm{e}^{-DQ^2\tau} \qquad [3]$$

Thus, for such a system, diffusion coefficients can rapidly be obtained by simple least-squares fits of Equation [3] to electric field autocorrelation functions recovered from the experimental intensity autocorrelation functions, or by linear fits to the natural logarithm of $g^{(1)}(\tau)$.

The more common situation is one where the solution contains a size distribution of scatterers. In this situation

$$g^{(1)}(\tau) = \int_0^\infty G(\Gamma)\,\mathrm{e}^{-\Gamma\tau}\,\mathrm{d}\Gamma \qquad [4]$$

where $\Gamma = DQ^2$ and $G(\Gamma)$ is the distribution of decay rates that would be obtained from spherical particles whose radius distribution is given by $G(r)$. Since $G(\Gamma)\mathrm{d}\Gamma = G(r)\,\mathrm{d}r$ and the Stokes–Einstein relation is

$$D = \frac{k_\mathrm{B} T}{6\pi\eta r} \qquad [5]$$

where $k_\mathrm{B}$ = Boltzmann's constant; $T$ = temperature K and $\eta$ = coefficient of viscosity then Equation [4] becomes

$$g^{(1)}(\tau) = \int_0^\infty G(r)\,\mathrm{e}^{-\frac{k_\mathrm{B} T}{6\pi\eta r}Q^2\tau}\,\mathrm{d}r \qquad [6]$$

In principle, a complete Laplace inversion of Equation [6] would yield the $G(r)$, the radius distribution of the scattering particles. However, this inversion is termed 'ill conditioned' because of mathematical instability and because unattainably high precision in the experimental data is required. Several alternative methods of various complexities have been developed and brief treatments of two of these are given here. The first and most common method of proceeding is called moment analysis or the method of cumulants. Any monomodal distribution can be described in terms of its moments. The object is to obtain these moments of $G(\Gamma)$ without actually performing the inversion. Specifically one attempts to obtain the first and second moments from which one can obtain the mean value and the variance of $G(\Gamma)$ respectively. In this approach the exponential $e^{-\Gamma\tau}$ in Equation [4] is expanded about the mean value $e^{-\bar{\Gamma}\tau}$. The result, after taking the natural logarithm of both sides, becomes a polynomial of the form

$$\ell n[g^{(1)}(\tau)] = -\bar{\Gamma}\tau + \frac{\mu_2}{2!}\tau^2 - \frac{\mu_3}{3!}\tau^3 + \cdots \qquad [7]$$

Most spectrometers contain fitting routines which output the correlation time $\tau_c = \Gamma^{-1}$, the $z$-average diffusion coefficient and radius, calculated using Equation [5] and $\mu_2$. Since

$$\mu_2 = \overline{\Gamma^2} \qquad [8]$$

the variance of the distribution can be determined from

$$\text{var} = \overline{\Gamma^2} - \bar{\Gamma}^2 \qquad [9]$$

Most modern spectrometers provide a visual display of the log-normal distribution that has the same mean radius and variance as obtained from the moments analysis of the data. This can be somewhat misleading if the true size distribution of the sample is not close to being log-normal and can be a serious misinterpretation if the sample is multimodal. Some spectrometers do, however, provide enhancements of the above procedure if a sample is suspected to be multimodal. Further, since no consideration has been given to the fact that large particles scatter more light than small ones, then all the results of moments analysis are intensity weighted or $z$-averages quantities. The broader the size distribution, the more these quantities can differ from their corresponding number-average quantities.

A variety of mathematically more sophisticated procedures have been developed to approximate the inversion of Equation [6], but without presupposing the form of the size distribution. Most of these approaches are variants of a discrete method in which $g^{(1)}(\tau)$ is fitted to a sum of $m$ exponential functions, each premultiplied by a weighting factor $\omega_m$. The intent is to minimize the sum of squares

$$\left[\left(g^{(1)}(\tau) - \sum^{m} \omega_m \exp(-\Gamma_m \tau)\right)\right]^2 \qquad [10]$$

with respect to the variables $\omega_m$ and $\Gamma_m$. As mentioned earlier, this is notoriously unstable, and if no constraints were applied it is essentially impossible to obtain a unique set of best fit parameters. One of the more common restraints is to specify, in advance of the fit, the range and values of a 'trial' set of $\Gamma_m$ that are expected to span the range of $\Gamma$s corresponding to the distribution of particle sizes present. The smallest and largest values of the trial $\Gamma_m$ (i.e. $m = 1$, and $m = m$) can easily be estimated from the final and the initial slopes of $g^{(1)}(\tau)$. In the technique called exponential sampling, the remaining $\Gamma_m$s are exponentially spaced between these limiting values. Once all the $\Gamma_m$ have been preset, only the $\omega_m$s corresponding to the amplitudes, or relative weights of each of the exponentials remain to be determined by the fit. Even with this simplification there is rarely a single solution with a unique set of $\omega_m$, especially if $m$ is 50 or more. However, by constraining all $\omega_m$ to be positive [by using non-negative-least squares (NNLS) methods] a unique set of $\omega_m$ can usually be found. Since each $\Gamma_m$ has a corresponding $r_m$, plots of $\omega_m$ versus $r_m$ represent the radius distribution of the sample. If the data are of sufficient quality (often runs of several hours are necessary) then reliable and reproducible plots of the radius distribution, $G(r)$, can be obtained.

The amplitudes $\omega_m$ represent the amount of light scattered by each particle size $r_m$ and represent an intensity weighted distribution of radii. To obtain number distributions one must include the relative scattering ability of each size in the distribution. This can be accomplished by including Rayleigh–Gans–Debye or Mie scattering factors in the analysis. The intensity weighted distribution and the number distribution for a set of phospholipid vesicles is shown in **Figure 4**.

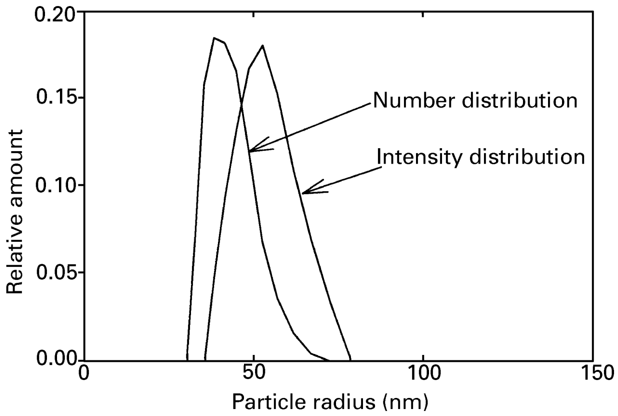



**Figure 4** Intensity and number distributions obtained from the same DLS data from dioleoylphosphatidylglycerol vesicles prepared by extrusion through filters of 50 nm radius pore size.

Although a number of 'user friendly' DLS spectrometers and advanced analytical software packages are available from a number of manufacturers, and size distributions are rapidly and neatly displayed on the computer monitor, it is still important that the user perform a number of checks to ensure that these results are reliable. If one is safely operating in the single scattering regime (no multiple scattering effects), then the particle size distribution should remain the same if the sample concentration is doubled or halved. The presence of multiple scattering leads to multiple phase shifts in the scattered light which, in turn, lead to 'fictitious' particles with sizes smaller than the true size. On the other hand, if sample concentrations are too small, number fluctuations can lead to particles whose apparent size is erroneously large. The presence of dust particles in the sample must be scrupulously avoided for the same reason. Secondly, it is essential that experimental runs be sufficiently long that the measured autocorrelation function is a true statistical representation of the sample. Recovery of size distributions requires extremely accurate data and it is not uncommon that runs of several hours are necessary to achieve the required level of data quality. Better quality data and more data always leads to more confidence in the result. For this reason, new multi-angle DLS experiments have been shown to improve the resolution of measured size distribution over single angle determinations.

Samples must be diluted to the point where they are optically almost clear for proper DLS investigations. As a result standard DLS methods cannot be applied to more concentrated suspensions such as milk or paint unless they are diluted extensively. However, DWS can provide estimates of mean particle sizes in concentrated opaque suspensions. The apparatus used in DWS is the same as for DLS except that the incident light and the scattered light are usually conducted through single fibre optics. DWS depends on multiple scattering being so severe that the incident photons experience a random walk (diffusing wave) in the sample before being finally scattered into the returning fibre optic for detection. For back-scatter detection, the simplest form of the DWS electric field autocorrelation function is

$$g^{(1)}(\tau) = \exp\left[-\gamma(6\tau/\tau_0)^{1/2}\right] \qquad [11]$$

where $\tau_0 = (Dk_0)^{-1}$. The constant $\gamma$ must be determined independently by calibration with known size scatterers or by transmission measurements. Once this is done, however, good estimates of $D$ (and therefore $r$ for spherical systems) have been obtained in a variety of dense scattering media.

## Static light scattering, small angle X-ray and neutron scattering

Even though lasers are still commonly used as the source of incident radiation, static scattering does not require coherence. Therefore, other forms of radiation such as X-ray and neutrons, even when pulsed, are suitable sources. The scattering vector dependence of the intensity, $I(Q)$, of the radiation scattered by particles has the form

$$I(Q) = KNm^2\, P(Q)\, S(Q)\, I_0 \qquad [12]$$

where $I_0$ is the incident radiation intensity, $N$ is the number of particles in the scattering volume, $m$ is the particle mass and $K$ is an instrumental constant. The functions $S(Q)$ and $P(Q)$ are the interparticle scattering factor and the intraparticle scattering factor respectively. The interparticle scattering factor arises from interference between wavelets scattered from different particles. In a suspension this term contains information on the radial distribution function of the scatterers and on their interaction potential. In a more ordered system, where regular spacings are found, this term results in Bragg diffraction peaks when conditions for constructive interference are satisfied. In the case of light scattering, the avoidance of multiple scattering suspensions usually ensures that particle concentrations are sufficiently dilute that interactions are negligible and $S(Q) = 1$. However, since neutrons

and X-rays are more penetrating and are scattered more weakly by particles, they are more commonly used at concentrations where interactions are important and where $S(Q) \neq 1$. However, measurements of $S(Q)$ will not be considered here.

The intraparticle scattering factor $P(Q)$, is related to the structural properties of the scatterer and is the term that contains particle size information. For extremely small noninteracting particles where $r << \lambda$, $P(Q)$ approaches unity, and the scattering is said to be isotropic. In this situation Equation [12] greatly simplifies and $I(Q) \propto m^2$. For uniform spherical or globular particles this means that $I(Q) \propto r^6$. If concentrations can be accurately determined by alternative means, then the radius of an unknown particle can be obtained by simple ratios, provided the unknown and the known particles have identical refractive indices.

In the regime where $Q^2r^2 << 1$, the scattering is weakly dependent on $Q$, and P(Q) can be described by the Guinier approximation.

$$P(Q) = \exp\left(\frac{Q^2 R_g^2}{3}\right) \quad [13]$$

For small angle scattering (e.g. small angle neutron and X-ray scattering) Equation [13] is further approximated by

$$P(Q) = 1 - \frac{Q^2 R_g^2}{3} \quad [14]$$

For larger particles, explicit Rayleigh–Gans–Debye (RGD) expression, for $P(Q)$ are known for several different particle shapes and are available in the literature. These RGD functions are valid if the condition

$$Q^2 r^2 \left(\frac{n^2}{n_0^2} - 1\right) < 1 \quad [15]$$

is satisfied ($n$ is the refractive index of the particle and $n_0$ is the refractive index of the medium). For the case of scattering of vertically polarized incident light from uniform spherical particles, $P(Q)$ has the form

$$P(Q) = \left(\frac{n}{n_0} - 1\right)^2 \left[\frac{3}{u^3}(\sin u - u\cos u)\right]^2 \quad [16]$$

where $u = Qr$.

When particle sizes are too large to satisfy the RGD criterion then Mie theory must be used to evaluate scattering factors. The need for a more sophisticated theory arises because wavefronts of the incident light do not remain planar as they traverse large particles, thereby altering intraparticle interference and the angular dependence of the scattered light. For vertically polarized incident light, the angle dependence of the scattered light is contained in the $S_{11}$ component of the Mueller matrix:

$$S_{11}(\cos\theta, n, n_0, r) = \sum_{j=1}^{\infty} \frac{2j+1}{j(j+1)}(a_j\pi_j + b_j\tau_j) \quad [17]$$

where

$$\pi_j = \frac{P_j^1 \cos(\theta)}{\sin(\theta)} \quad \text{and} \quad \tau_j = \frac{d}{d\theta} P_j^1 \cos(\theta) \quad [18]$$

and $P_j^1\cos(\theta)$ is the associated Legendre polynomial. The scattering coefficients, $a_j$ and $b_j$ are known only for spherical systems such as uniform spheres, hollow shells and coated spheres.

Fraunhofer scattering has become a popular optical sizing tool for particles substantially larger than a micrometre in diameter. In this case each particle behaves similarly to a pinhole aperture, and the low angle scattered light is a superposition of Fraunhofer diffraction patterns. This technique is widely used for the analysis of commercial food-based emulsions and sauces.

Since X-rays and neutrons have relatively low scattering cross-sections for most particulate suspensions, RGD scattering factors apply extremely well in the analysis of SAXS and SANS data, even when the scattering particles are above the size limitations imposed by Equation [15]. Expressions equivalent to Equations [12] and [16] have been written down for X-ray and neutron scattering. In addition to particle size analysis, these techniques, SANS especially, offer the ability to use contrast variation methods to enhance the scattering from the structure or substructure of interest. For SANS studies of water-soluble particles, this can be accomplished by varying the $H_2O$:$D_2O$ ratio in the solvent mixture or by exchanging deuterium for hydrogen in the scatterer. For SAXS the contrast variation can be accomplished if heavy atoms can be strategically attached to the scatterer (**Figure 5**).

Effects due to polydispersity are also important in the analysis of static light scattering data. For a

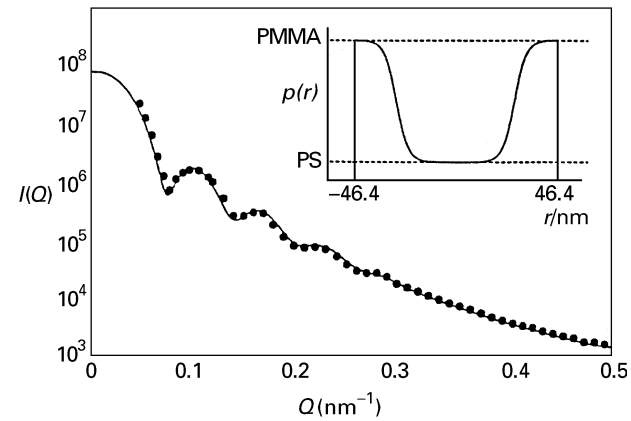



**Figure 5** Experimental SAXS data from a composite latex with a polystyrene core and a poly(methyl methacrylate) shell. The number average diameter is 92.8 nm. The inset displays the electron density cross-section of the particle derived from contrast variation studies. Reproduced with permission of Hüthig & Wepf Verlag from Ballauff MB, Bolze J, Dingenouts N, Hickl P. and Pötschke D (1996) Small-angle X-ray scattering on latexes. *Macromolecular Chemistry and Physics* **197**: 3043–3066.

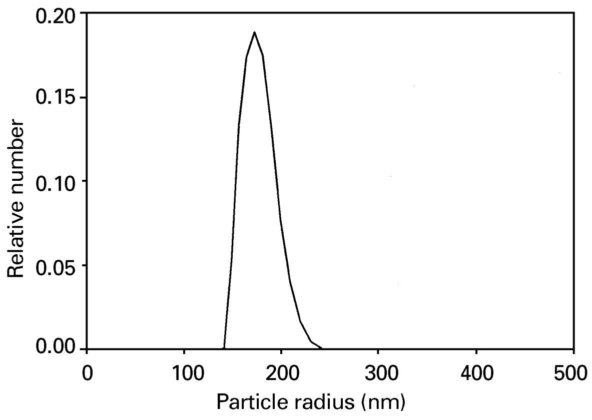



**Figure 6** Number distribution of particle sizes from a sample of microfluidized milk obtained by the inversion of *I* (*Q*) data from SLS.

polydisperse sample, Equation [12] becomes.

$$I(Q) = I_0 KN \int_0^{\infty} M^2(r)P(Q, n, n_0, r)G(r)\,\mathrm{d}r \quad [19]$$

Recovery of the distribution of radii $[G(r)]$ present in the sample requires the inversion of Equation [21] and as for dynamic light scattering (see Eqn [6]) this transform is ill conditioned. However, discrete methods analogous to those used for DLS have been applied successfully, thereby allowing size distributions to be obtained from SLS data. Depending on the particle size range, $P(Q, n, n_0, r)$ can be Guinier, RGD, or Mie, with the latter being the most universal. **Figure 6** demonstrates a size distribution of microfluidized milk obtained from the inversion $I(Q)$ from SLS data according to Equation [19] and assuming the sample behaves as uniform spherical particles. This is corroborated by electron microscopic studies in this case. However, knowledge of the shape and structure (e.g. sphere, coated sphere, rod, Gaussian coil, etc.) is not always available. If the wrong morphology is assumed then the inversion results can be meaningless. In addition, for SLS, $I(Q)$ data from particle sizes below ~ 100 nm in diameter is so close to being isotropic that inversion and recovery of size distributions becomes impossible. The opposite limitation arises in SANS and SAXS. Because of the significantly shorter wavelengths of neutrons and X-rays, the upper limit on diameter for particle sizing for these techniques is typically 100 nm (see **Figure 2**). Finally, as for DLS the need for extremely high quality data is just as vital for the successful recovery of particle size distributions, and all inversion routines work best when the distributions are unimodal and relatively narrow. To improve the recovery of broad size distributions and multimodal distributions, SLS measurements have sometimes been combined with sample fractionation techniques. The most successful of these devices, multiangle light scattering-field flow fractionation (MALS-FFF) spectrometers shows great promise for sizing complex particulate mixtures.

## List of symbols

$E(t)$ = electric field at time $t$; $g^{(1)}(\tau)$ = electric field autocorrelation function; $g^{(2)}(\tau)$ = intensity autocorrelation function; $I_0$, $I$ = intensity of incident and scattered light respectively; $\boldsymbol{k}_0$, $\boldsymbol{k}_s$ = incident and scattered wave vector respectively; $n_0$, $n$ = refractive index of the medium and of the particle respectively; $Q$ = scattering vector; $r$ = particle radius; $R_g$ = radius of gyration; $T$ = temperature (K); $\eta$ = viscosity coefficient; $\theta$ = scattering angle; $\lambda$ = wavelength of incident light; $\tau$ = time delay; $\omega_m$ = weighting factor.

*See also:* **Diffusion Studied Using NMR Spectroscopy; Electromagnetic Radiation; Fourier Transformation and Sampling Theory; Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Light Sources and Optics; Neutron Diffraction, Instrumentation; Scattering Theory.**

## Further reading


Ballauff M, Bolze J, Dingenouts, Hickl P and Pötschke D (1996) Small-angle X-ray scattering on latexes. *Macromolecular Chemistry and Physics* **197**: 3043–3066.

Barnes MD, Lermer N, Whitten WB and Ramsey JM (1997) A CCD based approach to high-precision size and refractive index determination of levitated microdroplets using Fraunhofer diffraction. *Review of Scientific Instruments* **68**: 2287–2291.

Berne B and Pecora R (1976) *Dynamic Light Scattering*. New York: John Wiley & Sons.

Brown W (ed) (1993) *Dynamic Light Scattering*. Oxford: Clarendon Press.

Brown W (ed) (1996) *Static Light Scattering, Principles and Development*. Oxford: Clarendon Press.

Bryant G, Abeynayake C and Thomas JC (1996) Improved particle size distribution measurements using multiangle dynamic light scattering 2. Refinements and applications. *Langmuir* **12**: 6224–6228.

Chu B (1991) *Laser Light Scattering*. Academic Press: Boston.

Dalgleish DG, West SJ and Hallett FR (1997) The characterization of small emulsion droplets made from milk proteins and triglyceride oil. *Colloids Surfaces A: Physiochemical and Engineering Aspects* **123–124**: 145–153.

De Vos C, Deriemaeker L and Finsy R (1996) Quantitative assessment of the conditioning of the inversion of quasi-elastic and static light scattering data for particle size distributions. *Langmuir* **12**: 2630–2636.

Filella M, Zhang J, Newman ME and Buffle J (1997) Analytical applications of photon correlation spectroscopy for size distribution measurements of natural colloidal suspensions: capabilities and limitations. *Colloids Surfaces, A: Physicochemical and Engineering Aspects* **120**: 27–46.

Glatter O and Kratky O (eds) (1982) *Small Angle X-Ray Scattering*. London: Academic Press.

Hallett FR, Craig T, Marsh J and Nickel B (1989) Particle size analysis: number distributions by dynamic light scattering. *Canadian Journal of Spectroscopy* **34**: 63–70.

Maret G (1997) Diffusing-wave spectroscopy. *Current Opinion in Colloid and Interface Science* **2**: 251–257.

Pike ER and McNally B (1997) Theory and design of photon correlation and light-scattering experiments. *Applied Optics* **36**: 7531–7538.

Schmitz KS (1990) *Introduction to Quasielastic Light Scattering by Macromolecules*. New York: Academic Press.

Strawbridge KB and Hallett FR (1994) Size distributions obtained from the inversion of $I(Q)$ using integrated light scattering spectroscopy. *Macromolecules* **27**: 2283–2290.

Wyatt PJ (1998) Submicrometer particle sizing by multiangle light scattering following fractionation *Journal of Colloid and Interface Science* **197**: 9–20.


# Scattering Theory

**Michael Kotlarchyk,** Rochester Institute of Technology, NY, USA

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Theory**

The term 'scattering' refers to an interaction between an incident radiation and a target material resulting in a redirection, and possibly a change in energy, of the incident radiation. In effect, there is an exchange of momentum and energy between the radiation and target. From a fundamental standpoint, the scattering of photons, neutrons, and charged particles from atomic and molecular systems needs to be treated quantum-mechanically. On the other hand, for electromagnetic radiation having wavelengths near or larger than those in the visible region of the spectrum, a classical wave approach is justified as well. In either case, the aim of any useful scattering theory is to provide the calculational framework for predicting the quantity most readily measured in the laboratory, namely the flux or intensity of radiation scattered into a detector. Ultimately, the connection between a scattering measurement and theory is made through a quantity known as the double-differential scattering cross-section.

## Scattering cross-sections

A scattering cross-section, $\sigma$, is a quantity proportional to the rate at which a particular radiation–target interaction occurs. More specifically, if the incoming radiation is considered as being composed of quanta or 'particles' (for example, photons or neutrons), a cross-section is a scattering rate (number of scattering events per unit

time) per unit incident radiation flux, where the latter is the number of incident particles striking the target surface per unit time per unit area. In cases where the radiation is being treated as a continuous classical wave, as in the case of long-wavelength electromagnetic radiation, scattering cross-sections are determined by dividing the power of the scattered wave by the intensity of the incident wave. Dimensionally, a cross-section represents an area, with the basic unit being the barn, which represents an area of $10^{-28}$ m$^2$. A scattering cross-section should not be interpreted as a true geometric cross-sectional area, but as an effective area that is proportional to the probability of interaction between the radiation and target.

In a laboratory setting, one measures differential scattering cross-sections. These are determined by placing a detector at a particular angular position at a substantial distance away from the scattering target. **Figure 1** shows a standard scattering geometry. The incident beam travels in the positive $z$-direction and one considers the radiation scattered into a differential solid angle $\mathrm{d}\Omega$ at polar angle $\theta$ and azimuthal angle $\phi$. One can then measure an angular differential scattering cross-section (in barn steradian$^{-1}$) given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\mathrm{d}R/\mathrm{d}\Omega}{\Phi_{\mathrm{in}}} \qquad [1]$$

where $\mathrm{d}R$ is the rate of scattering into solid angle $\mathrm{d}\Omega$, and $\Phi_{\mathrm{in}}$ is the incident flux. The most fundamental type of cross-section is the double-differential scattering cross-section, $\mathrm{d}^2\sigma/\mathrm{d}\Omega\mathrm{d}E'$. The quantity $[\mathrm{d}^2\sigma/(\mathrm{d}\Omega\,\mathrm{d}E')]\,\mathrm{d}\Omega\,\mathrm{d}E'$ is the number of particles, each with incident energy $E$, scattered (per unit time) into solid angle $\mathrm{d}\Omega$ with energy between $E'$ and $E' + \mathrm{d}E'$,

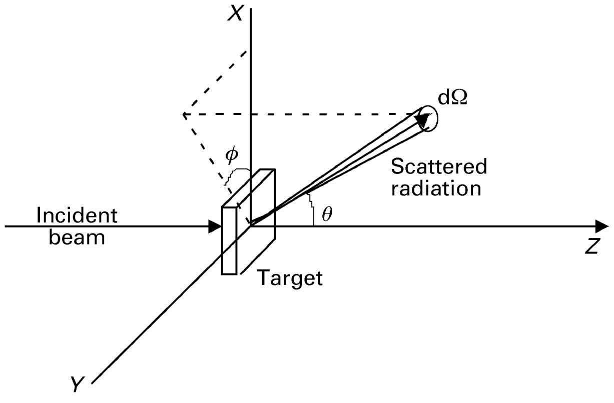



**Figure 1** Standard scattering geometry for measuring differential scattering cross-section.

divided by the flux of the incident beam. Once the double-differential cross-section is derived or measured, $\mathrm{d}\sigma/\mathrm{d}\Omega$ and $\sigma$ can be calculated by integrating over the energy of the scattered radiation and solid angle.

## Development of the double-differential cross-section

From a quantum mechanical standpoint, the calculation of the double-differential cross-section is based on time-dependent perturbation theory, which provides the general framework for calculating transition rates between quantum states.

A scattering event is a type of transition between two states of the combined system of incident particle + scattering medium. Take the system's unperturbed Hamiltonian $H_0$ to be

$$H_0 = H_{\mathrm{S}} + H_{\mathrm{R}} \qquad [2]$$

where $H_{\mathrm{S}}$ and $H_{\mathrm{R}}$ are the Hamiltonians of the scatterer and radiation, respectively, assuming no interaction between the two. The combined system is in an eigenstate of $H_0$ both before and after the scattering interaction takes place. Employing Dirac notation, let the kets $|E_{\mathrm{i}}\rangle$ and $|E_{\mathrm{f}}\rangle$ denote the initial and final energy eigenstates of the scatterer, i.e.

$$H_{\mathrm{S}}\,|E_{\mathrm{i}}\rangle = E_{\mathrm{i}}\,|E_{\mathrm{i}}\rangle \quad \text{and} \quad H_{\mathrm{S}}\,|E_{\mathrm{f}}\rangle = E_{\mathrm{f}}\,|E_{\mathrm{f}}\rangle \qquad [3]$$

If $k$ is the wavenumber ($2\pi$/wavelength) of the incident beam, a quantum of the incident radiation is represented by an eigenstate $|\boldsymbol{k}\lambda\rangle$ of definite momentum $\hbar\boldsymbol{k}$ and polarization (or spin) $\lambda$ (see **Figure 2**). The scattered quantum falls into state $|\boldsymbol{k}'\lambda'\rangle$. These are also energy eigenstates, i.e.

$$H_{\mathrm{R}}\,|\boldsymbol{k}\lambda\rangle = E\,|\boldsymbol{k}\lambda\rangle \quad \text{and} \quad H_{\mathrm{R}}\,|\boldsymbol{k}'\lambda'\rangle = E'\,|\boldsymbol{k}'\lambda'\rangle \qquad [4]$$

Again, $E$ and $E'$ are the energies of the incident and scattered particles. For the composite system, the initial state $|m\rangle$ and final state $|n\rangle$ are given by the products

$$|m\rangle = |\,\boldsymbol{k}\lambda\rangle\,|E_{\mathrm{i}}\rangle \quad \text{and} \quad |n\rangle = |\,\boldsymbol{k}'\lambda'\rangle\,|E_{\mathrm{f}}\rangle \qquad [5]$$

which are eigenstates of $H_0$, so

$$H_0\,|m\rangle = E_m\,|m\rangle \quad \text{and} \quad H_0\,|n\rangle = E_n\,|n\rangle \qquad [6]$$

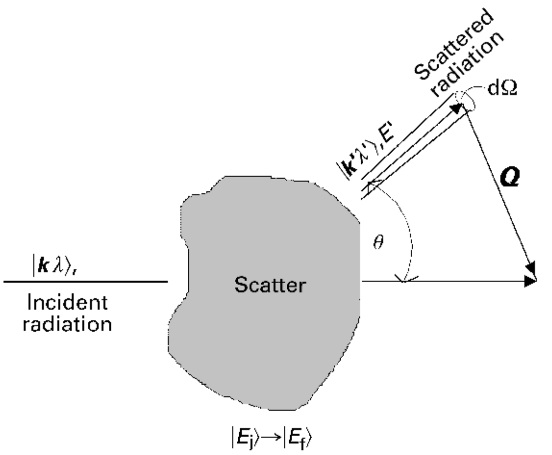



**Figure 2** Setup for deriving a double-differential scattering cross-section

where the initial and final energies of the system are $E_m = E_i + E$ and $E_n = E_f + E'$, respectively.

Time-dependent perturbation theory gives rise to the following general result for the transition probability per unit time between initial state $|m\rangle$ and final state $|n\rangle$:

$$W(m \to n) = \frac{2\pi}{\hbar} |\langle n \mid \mathbf{T} \mid m\rangle|^2 \, g(E_n) \qquad [7]$$

Here, $g(E_n)$ denotes the density of final states for the system, where $g(E_n)\mathrm{d}E_n$ is the number of final states between $E_n$ and $E_n + \mathrm{d}E_n$. $\langle n \mid \mathbf{T} \mid m\rangle$ is called the transition matrix or $\mathbf{T}$-matrix, and it is based on $V$, the interaction potential between the incident radiation and the scatterer, according to the expansion

$$\langle n \mid \mathbf{T} \mid m\rangle = \langle n \mid V \mid m\rangle + \sum_I \frac{\langle n \mid V \mid I\rangle\langle I \mid V \mid m\rangle}{E_m - E_I} + \cdots \qquad [8]$$

Basically, the interaction potential induces scattering events and, hence, transitions between the initial and final states of the combined system. When the first term, i.e. matrix element $\langle n | V | m\rangle$, is the only one responsible for transitions, one says that the transition is first-order in the perturbation, and Equation [7] is referred to as Fermi's golden rule. Higher-order transitions require passing through some intermediate states, denoted here by $I$, between the specified initial and final states.

Once the transition rate is developed for a specific type of interaction, Equation [1] can be used to obtain the angular differential cross-section, $\mathrm{d}\sigma/\mathrm{d}\Omega$. For a target in a known initial state $|E_i\rangle$ and known final state $|E_f\rangle$, one writes

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{if} \mathrm{d}\Omega = \frac{\mathrm{d}R_{if}}{\Phi_{\mathrm{in}}} = \frac{W(E_i, \boldsymbol{k}\lambda \to E_f, \boldsymbol{k}'\lambda')}{\Phi_{\mathrm{in}}} \qquad [9]$$

In this equation, $\Phi_{\mathrm{in}}$ now represents the flux of a single incident particle. It is usually the case, however that the scatterer is not initially in one of its pure eigenstates. Instead, the target is normally in thermal equilibrium at some known temperature $T$. In this case, the scatterer is in a 'mixed quantum state' and it is necessary to perform a summation over possible initial states, $|E_i\rangle$, weighted according to the Maxwell–Boltzmann probability distribution, $P_i = \exp(-E_i/k_B T)/\Sigma_j \exp(-E_j/k_B T)$, where $k_B$ is Boltzmann's constant. The angular differential cross-section then becomes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\sum_i P_i W(E_i, \boldsymbol{k}\lambda \to E_f, \boldsymbol{k}'\lambda')}{\Phi_{\mathrm{in}}\,\mathrm{d}\Omega} \qquad [10]$$

The procedure for obtaining the double-differential scattering cross-section is to now sum over possible final states $|E_f\rangle$ of the scatterer, subject to the condition that the energy of the combined system of radiation + target is conserved. The latter is accomplished by attaching the Dirac delta function, $\delta(E_m - E_n)$, to each term in the sum. Finally, we can introduce

$$\hbar\omega = E - E' \qquad [11]$$

which is an important parameter that represents the energy transferred to the material medium by a quantum of the radiation. Putting these ideas together gives the following expression for the double-differential cross-section:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\omega} = \hbar\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}E'} = \frac{\hbar\sum_{i,f} P_i W(E_i, \boldsymbol{k}\lambda \to E_f, \boldsymbol{k}'\lambda')\delta[\hbar\omega - (E_f - E_i)]}{\Phi_{\mathrm{in}}\,\mathrm{d}\Omega} \qquad [12]$$

## Light scattering – an example

The purpose of this section is to illustrate, for a specific case, the derivation of the angular and

double-differential scattering cross-section just outlined. The vehicle chosen is the scattering of visible light from a collection of $N$ optically isotropic atoms in thermal equilibrium. As will be seen, the calculation culminates in a central quantity known as the dynamic structure factor, which not only appears in light scattering, but in the scattering of X-rays and thermal neutrons as well.

For optical frequencies, where the wavelength of light is large compared with the size of an atom, the interaction occurs through the coupling of the radiation to the electric dipole moment $\boldsymbol{\mu}$ of each atom in the target material. The Hamiltonian of the atomic system in the presence of the electromagnetic (EM) field of the light beam is

$$H = H_{\mathrm{S}} + H_{\mathrm{R}} - \sum_{l} \boldsymbol{\mu}_l \cdot \boldsymbol{E}(\boldsymbol{R}_l) \qquad [13]$$

where $H_S$ is the Hamiltonian of the atomic system in the absence of the EM field, $H_R$ is the Hamiltonian of the pure light field, and the last term represents the sum of the interactions between the dipole moment of the $l$th atom in the system and the light's electric field vector, $\boldsymbol{E}$, at the location $\boldsymbol{R}_l$ of that atom.

The present derivation of the scattering cross-section is based on a non-relativistic quantum electrodynamic approach. In this picture, the modes of the radiation field are quantized and the electric field is treated as a quantum-mechanical operator that annihilates or creates photons populating the various modes. The field operator is given by

$$\boldsymbol{E}(\boldsymbol{R}_l) = \sum_{k\lambda} \mathrm{i}\sqrt{\frac{\hbar c \boldsymbol{k}}{2\varepsilon_0 L^3}}\, \epsilon_{k\lambda}[a_{k\lambda} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l) - a^{\dagger}_{k\lambda} \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l)] \qquad [14]$$

At the initial stage of the calculation it is necessary to imagine that the radiation field is confined to a region of space having volume $L^3$; however, this volume drops out of the calculation shortly. The term $\epsilon_{k\lambda}$ is the unit polarization vector of photon state $|\boldsymbol{k}\lambda\rangle$. $a_{k\lambda}$ and $a^{\dagger}_{k\lambda}$ are sometimes called the photon 'annihilation' and 'creation' operators. $a_{k\lambda}$ has the effect of lowering the number of photons $n_{k\lambda}$ in state $\boldsymbol{k}\lambda$ to $n_{k\lambda}-1$, whereas $a^{\dagger}_{k\lambda}$ raises the number of photons from $n_{k\lambda}$ to $n_{k\lambda}+1$.

The transition rate for the $l$th atom is now given by Equation [7], where the interaction potential associated with the transition matrix is $\boldsymbol{\mu}_l \cdot \boldsymbol{E}(\boldsymbol{R}_l)$. If the atom is initially in state $|A\rangle$, then the combined initial state for the field + atom is $|m\rangle = |1_{k\lambda},0_{k'\lambda'}\rangle|A\rangle$, indicating that there is one photon in incident mode $\boldsymbol{k}\lambda$ and no photons in the scattered mode $\boldsymbol{k}'\lambda'$. Likewise, the final state is $|n\rangle = |0_{k\lambda},1_{k'\lambda'}\rangle|B\rangle$. Because of the action of the operators $a_{k\lambda}$ and $a^{\dagger}_{k\lambda}$, along with the orthogonality property of the photon states, the terms in the perturbation expansion of the transition matrix can only be non-zero if the initial and final states differ by unity in one, and only one, mode of the field. A careful inspection shows that all terms in the perturbation expansion vanish except for the second-order contributions. These are terms that involve intermediate states, $|I\rangle$, of the combined system, which are sometimes referred to as virtual states. For single photon scattering, two types of virtual transitions can occur, as illustrated by the Feynman diagrams in **Figure 3**. In the first type of transition, we have the intermediate state $|I\rangle = |0_{k\lambda},0_{k'\lambda'}\rangle|J\rangle$, where one pictures that an intermediate atomic state $|J\rangle$ is created when the incident $\boldsymbol{k}\lambda$-photon is absorbed; the scattered $\boldsymbol{k}'\lambda'$-photon is then born, leaving the atom in final state $|B\rangle$. In the second type of transition, the intermediate state is represented by $|I\rangle = |1_{k\lambda},1_{k'\lambda'}\rangle|J\rangle$ where one imagines that the scattered $\boldsymbol{k}'\lambda'$-photon appears before the incident $\boldsymbol{k}\lambda$-photon is absorbed.

Because their existence is not real, transitions to and from the virtual states do not need to conserve energy. However, the energies of the initial state and the final state must match exactly, i.e. $E_A + \hbar ck = E_B + \hbar ck'$. Taking this into account, the transition matrix for atom $l$ works out be the sum of two terms

$$\langle n \mid \mathbf{T} \mid m\rangle_l = -\frac{\hbar c}{2\varepsilon_0 L^3}\sqrt{\boldsymbol{k}\boldsymbol{k}'}\exp[\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_l] \times \sum_J \left( \frac{\langle B|\,\mu^l_{k'}\mid J\rangle\langle J\mid \mu^l_{k}\mid A\rangle}{E_A - E_J + \hbar c\boldsymbol{k}} + \frac{\langle B\mid \mu^l_{k}\mid J\rangle\langle J\mid \mu^l_{k'}\mid A\rangle}{E_A - E_J - \hbar c\boldsymbol{k}'} \right) \qquad [15]$$

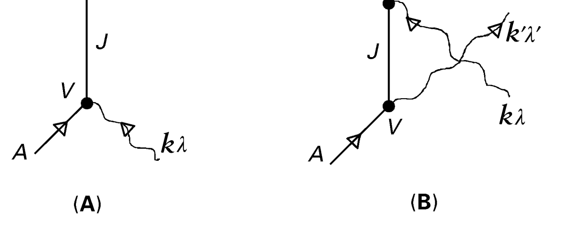



**Figure 3** Feynman diagrams for the two types of virtual transitions associated with the scattering of a photon by an atom

where $\mu^l_k = \epsilon_{k\lambda}\cdot\boldsymbol{\mu}_l$ and $\mu^l_{k} = \epsilon_{k'\lambda'}\cdot\boldsymbol{\mu}_l$. Details leading to this result can be found in the Further reading section.

Carrying out the calculation of the scattering cross-section also requires the incident flux associated with a single photon, which is

$$\Phi_{\text{in}} = c/L^3 \qquad [16]$$

and the density of states factor for a photon (of a specified polarization) scattered into solid angle dΩ about wave vector $\boldsymbol{k}'$. If recoil of the atom is neglected, the expression is given by

$$g(E_n) = \left(\frac{L}{2\pi}\right)^3 \frac{\boldsymbol{k}'^2}{\hbar c}\,\mathrm{d}\Omega \qquad [17]$$

One now has the following expression for the angular differential cross-section for the scattering of light by the $l$th atom:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_l = \frac{\boldsymbol{k}\boldsymbol{k}'^3}{(4\pi\varepsilon_0)^2}\left|\sum_J\left(\frac{\langle B \mid \mu^l_{\boldsymbol{k}'} \mid J\rangle\langle J \mid \mu^l_{\boldsymbol{k}} \mid A\rangle}{E_A - E_J + \hbar c\boldsymbol{k}} + \frac{\langle B \mid \mu^l_{\boldsymbol{k}} \mid J\rangle\langle J \mid \mu^l_{\boldsymbol{k}'} \mid A\rangle}{E_A - E_J - \hbar c\boldsymbol{k}'}\right)\right|^2 \qquad [18]$$

This result is valid for either inelastic Raman scattering or elastic Rayleigh scattering. For the case of Raman scattering, $|A\rangle$ and $|B\rangle$ are different discrete electronic states of the atom, hence the frequency of the scattered photon is either less than or greater than the frequency of the incident radiation. A down-shifted frequency gives rise to the so-called Stokes spectral line, whereas an up-shifted frequency corresponds to the anti-Stokes line. In the case of Rayleigh scattering, the atom returns to its original state, so that $|B\rangle = |A\rangle$ and $\boldsymbol{k}' = \boldsymbol{k}$. The cross-section then becomes

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)^{\text{Rayleigh}}_l = \frac{\boldsymbol{k}^4}{(4\pi\varepsilon_0)^2}\left|\epsilon_{\boldsymbol{k}\lambda}\cdot\tilde{\alpha}_l(\boldsymbol{k})\cdot\epsilon_{\boldsymbol{k}'\lambda'}\right|^2 \qquad [19]$$

where we define the *atomic polarizability tensor* for atom $l$ to be

$$\tilde{\alpha}_l(\boldsymbol{k}) = \sum_J\left(\frac{\langle A \mid \boldsymbol{\mu}_l \mid J\rangle\langle J \mid \boldsymbol{\mu}_l \mid A\rangle}{E_A - E_J + \hbar c\boldsymbol{k}} + \frac{\langle A \mid \boldsymbol{\mu}_l \mid J\rangle\langle J \mid \boldsymbol{\mu}_l \mid A\rangle}{E_A - E_J - \hbar c\boldsymbol{k}}\right) \qquad [20]$$

If the atomic charge distribution is spherically symmetric, then the polarizability is a scalar quantity and

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)^{\text{Rayleigh}}_l = \frac{\boldsymbol{k}^4}{(4\pi\varepsilon_0)^2}\left|\epsilon_{\boldsymbol{k}\lambda}\cdot\epsilon_{\boldsymbol{k}'\lambda'}\right|^2\left|\alpha_l(\boldsymbol{k})\right|^2 \qquad [21]$$

The well-known $k^4$-dependence of the cross-section is responsible for pronounced Rayleigh scattering of light at the short-wavelength end of the visible spectrum.

Now consider the full system of $N$ atoms in thermal equilibrium. Recall from Equation [15] that the transition matrix for a given atom contains a phase factor $e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_l}$ that depends on the position of that atom. At this point, let us introduce the momentum transferred to the material medium as

$$\hbar\boldsymbol{Q} = \hbar\boldsymbol{k} - \hbar\boldsymbol{k}' \qquad [22]$$

$\boldsymbol{Q}$ is usually referred to as the wave vector transfer (see **Figure 2**). The angular cross-section for coherent scattering from all the atoms of the system is obtained by first summing the transition matrix element over $l$, then squaring the result, and finally averaging over the initial energy states associated with the thermal motion of the atoms at equilibrium temperature $T$. In the case of identical atoms, all with the same polarizability $\alpha$, the cross-section reduces to

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right) = N\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_0 S(\boldsymbol{Q}) \qquad [23]$$

where $(\mathrm{d}\sigma/\mathrm{d}\Omega)_0 = [k^4/(4\pi\varepsilon_0)^2]\,|\epsilon_{k\lambda}\cdot\epsilon_{k'\lambda'}|^2|\alpha|^2$ is the Rayleigh scattering cross-section for a single atom

and

$$S(\boldsymbol{Q}) = \frac{1}{N}\sum_i P_i \langle E_i \mid \left| \sum_{l=1}^{N} \exp(\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{R}_l) \right|^2 \mid E_i\rangle$$
$$= \frac{1}{N}\left\langle \sum_{l,l'} \exp\left[-\mathrm{i}\boldsymbol{Q}\cdot(\boldsymbol{R}_l - \boldsymbol{R}_{l'})\right] \right\rangle \qquad [24]$$

is called the 'static structure factor'. The large angled bracket represents an ensemble average at temperature $T$. The function $S(\boldsymbol{Q})$ reflects the spatial correlations between pairs of particles in the scattering medium.

To write the double-differential cross-section, the expression for $(\mathrm{d}\sigma/\mathrm{d}\Omega)$ is summed over the final motional states with the inclusion of an energy-conservation factor, i.e.

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}\omega} = N\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_0 S(\boldsymbol{Q},\omega) \qquad [25]$$

where

$$S(\boldsymbol{Q},\omega) = \frac{1}{N}\sum_{i,f} P_i \left| \langle E_f \mid \sum_{l=1}^{N} \exp(\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{R}_l) \mid E_i\rangle \right|^2 \times \delta\left(\omega - \frac{E_f - E_i}{\hbar}\right) \qquad [26]$$

is the aforementioned dynamic structure factor. By using an integral representation of the delta function

$$\delta\left(\omega - \frac{E_f - E_i}{\hbar}\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}t \exp(-\mathrm{i}\omega t) \times \exp\left[\mathrm{i}(E_f - E_i)t/\hbar\right] \qquad [27]$$

the structure factor can (with some work) be transformed into the following more useful form:

$$S(\boldsymbol{Q},\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}t \exp(-\mathrm{i}\omega t) \times \left\langle \frac{1}{N}\sum_{l,l'} \exp\left[-\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{R}_l(0)\right] \exp\left[\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{R}_{l'}(t)\right] \right\rangle \qquad [28]$$

$S(\boldsymbol{Q},\omega)$ is determined by the spatial and temporal correlations between the various atoms in the scattering medium. Integrating the dynamic structure factor over $\omega$ produces the expression for the static structure factor, $S(\boldsymbol{Q})$.

## Scattering of classical electromagnetic waves

Light scattering can also be treated in the framework of classical EM waves. Since Maxwell's equations accurately describe electromagnetic phenomena almost down to the atomic scale, one might ask whether it is really necessary to go through the detailed quantum treatment previously described. To answer this question, the classical approach is outlined below.

Consider Maxwell's equations for the electric and magnetic field vectors, $\boldsymbol{E}$ and $\boldsymbol{B}$, in a nonmagnetic scattering medium void of free charges and currents. Assigning an $\exp(-\mathrm{i}\omega_0 t)$ time-dependence to all fields, two of the equations (Faraday's law and Ampere's law) become the pair

$$\nabla \times \boldsymbol{E} - \mathrm{i}c\boldsymbol{k}_\mathrm{v}\boldsymbol{B} = 0 \qquad [29]$$

$$\nabla \times \boldsymbol{B} + \mathrm{i}\frac{n^2\boldsymbol{k}_\mathrm{v}}{c}\boldsymbol{E} = 0 \qquad [30]$$

$k_\mathrm{v} = \omega_0/c$ is the wavenumber of the light in vacuum and $n = (\varepsilon/\varepsilon_0)^{1/2}$ is the refractive index of the medium ($\varepsilon$ and $\varepsilon_0$ are the permittivity of the medium and vacuum, respectively.) Eliminating $\boldsymbol{B}$ from the equations results in a single equation for the electric field:

$$\nabla \times \nabla \times \boldsymbol{E} - n^2\boldsymbol{k}_\mathrm{v}^2\boldsymbol{E} = 0 \qquad [31]$$

The triple product becomes $\nabla\times\nabla\times\boldsymbol{E} = \nabla(\nabla\cdot\boldsymbol{E}) - \nabla^2\boldsymbol{E}$. However, Gauss's law for a charge-free region is $\nabla\cdot\boldsymbol{E} = 0$, so the result is the following equation for the field:

$$\nabla^2\boldsymbol{E} + n^2\boldsymbol{k}_\mathrm{v}^2\boldsymbol{E} = 0 \qquad [32]$$

For a homogeneous medium with uniform index of refraction $n = \bar{n}$, the solution to Equation [32] is a propagating plane-wave with modified wavenumber $k = \bar{n}k_\mathrm{v}$. For scattering to occur, there must be local

refractive-index fluctuations present:

$$\Delta n^2(\boldsymbol{r},t) = n^2(\boldsymbol{r},t) - \bar{n}^2 = \frac{\varepsilon(\boldsymbol{r},t)-\bar{\varepsilon}}{\varepsilon_0} = \frac{\Delta\varepsilon(\boldsymbol{r},t)}{\varepsilon_0} \quad [33]$$

Equation [32] then becomes

$$\nabla^2\boldsymbol{E} + \boldsymbol{k}^2\boldsymbol{E} = -\frac{\Delta\varepsilon(\boldsymbol{r},t)}{\varepsilon_0}\boldsymbol{k}_{\mathrm{v}}^2\boldsymbol{E} \quad [34]$$

$\Delta\varepsilon(\boldsymbol{r}, t)/\varepsilon_0$ represents local fluctuations of the relative permittivity (dielectric constant) in the target. The task of classical light-scattering theory has been reduced to solving Equation [34]. The only role of quantum theory, therefore is to calculate the atomic polarizability $\alpha$ – in other words, the microscopic properties of the scattering medium. Once this is achieved, the polarizability can then be related to the dielectric constant through the well-known Clausius–Mossotti relation.

The scattered wave reaching location $\boldsymbol{r}$ is constructed from the Green's tensor $\tilde{G}(\boldsymbol{r}-\boldsymbol{r}')$, which is the solution to Equation [34] with the right-hand side, or source term, replaced by $\tilde{I}\delta(\boldsymbol{r}-\boldsymbol{r}')$ where $\tilde{I}$ is the unit dyad. The delta function represents a point scatterer situated at $\boldsymbol{r}'$. For a field point far from the scatterer, i.e. $r >> r'$, the form of the Green's tensor is

$$\tilde{G}(\boldsymbol{r}-\boldsymbol{r}') = \left(\tilde{I} - \boldsymbol{e}_r\boldsymbol{e}_r\right)\frac{\exp[\mathrm{i}\boldsymbol{k}r]}{4\pi r}\exp[-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}'] \quad [35]$$

where $\boldsymbol{e}_{\mathrm{r}}$ is the unit vector in the direction of $\boldsymbol{r}$, and $\boldsymbol{k}' = k\boldsymbol{e}_{\mathrm{r}}$ denotes the scattered wave vector. The scattered wave, $\boldsymbol{E}_{\mathrm{s}}(\boldsymbol{r}, t)$, corresponds to the inhomogeneous solution to Equation [34]. It is constructed from the Green's tensor to be the following integral over points $\boldsymbol{r}'$ in the scattering volume:

$$\boldsymbol{E}_{\mathrm{s}}(\boldsymbol{r},t) = \int_V \mathrm{d}^3r'\tilde{G}(\boldsymbol{r}-\boldsymbol{r}')\cdot\frac{\Delta\varepsilon(\boldsymbol{r}',t)}{\varepsilon_0}\boldsymbol{k}_{\mathrm{v}}^2\boldsymbol{E}(\boldsymbol{r}',t) \quad [36]$$

If scattering from the medium is sufficiently weak, one is justified in applying the so-called Rayleigh–Gans–Debye approximation. This simply says that the field inside the scatterer may be approximated to be the same as the field of the incident light. That is, one makes the replacement $\boldsymbol{E}(\boldsymbol{r}', t) = \boldsymbol{E}_0\exp[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega_0 t)]$, which leads to

$$\boldsymbol{E}_{\mathrm{s}}(\boldsymbol{r},t) = \frac{\boldsymbol{k}_{\mathrm{v}}^2}{4\pi\varepsilon_0}\frac{\exp[\mathrm{i}\boldsymbol{k}r-\omega_0 t]}{r}(\tilde{I} - \boldsymbol{e}_r\boldsymbol{e}_r) \cdot \boldsymbol{E}_0\int_V \mathrm{d}^3r'\Delta\varepsilon(\boldsymbol{r}',t)\exp(\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{r}') \quad [37]$$

In the case of condensed phases, fluctuations in the dielectric constant may contain a slow time-dependence caused by slowly varying density fluctuations on length-scales of the same order of magnitude as the wavelength of light. The dielectric fluctuations may then be replaced by fluctuations in the number density $\Delta\rho(\boldsymbol{r}', t)$ according to

$$\Delta\varepsilon(\boldsymbol{r}',t) = \left(\frac{\partial\varepsilon}{\partial\rho}\right)_T \Delta\rho(\boldsymbol{r}',t) \quad [38]$$

where $\Delta\rho(\boldsymbol{r}', t) = \rho(\boldsymbol{r}', t) - \bar{\rho}$ with $\bar{\rho}$ being the average density of the material. Furthermore, if one denotes the spatial Fourier transform of $\Delta\rho(\boldsymbol{r}', t)$ by

$$\Delta\rho(\boldsymbol{Q},t) = \int_V \mathrm{d}^3r'\Delta\rho(\boldsymbol{r}',t)\exp(\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{r}') \quad [39]$$

then the scattered field becomes

$$\boldsymbol{E}_{\mathrm{s}}(\boldsymbol{r},t) = \frac{\boldsymbol{k}_{\mathrm{v}}^2}{4\pi\varepsilon_0}\frac{\exp[\mathrm{i}(\boldsymbol{k}r-\omega_0 t)]}{r} \times (\tilde{I} - \boldsymbol{e}_r\boldsymbol{e}_r)\cdot\boldsymbol{E}_0\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T \Delta\rho(\boldsymbol{Q},t) \quad [40]$$

This clearly shows that light scattering arises because of local density fluctuations in the medium.

The angular differential cross-section is calculated by taking $\mathrm{d}P/\mathrm{d}\Omega$, where $\mathrm{d}P$ is the time-averaged power scattered into solid angle $\mathrm{d}\Omega$, and dividing by the intensity of the incident light. This is equivalent to writing $\mathrm{d}\sigma/\mathrm{d}\Omega = r^2\langle|\boldsymbol{E}_{\mathrm{s}}|^2\rangle/|\boldsymbol{E}_0|^2$, so that

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\boldsymbol{k}_{\mathrm{v}}^4}{(4\pi\varepsilon_0)^2}\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T^2 \left|(\tilde{I} - \boldsymbol{e}_r\boldsymbol{e}_r)\cdot\varepsilon\right|^2\langle\Delta\rho(\boldsymbol{Q})\Delta\rho(-\boldsymbol{Q})\rangle \quad [41]$$

where $\epsilon$ is the unit polarization vector of the incident beam. The time average $\langle\Delta\rho(Q)\Delta\rho(-Q)\rangle$ is statistically equivalent to an ensemble average. Since the number density can be expressed as a sum of delta functions, i.e.

$$\rho(\boldsymbol{r}') = \sum_{l=1}^{N} \delta(\boldsymbol{r}' - \boldsymbol{R}_l) \qquad [42]$$

its Fourier transform is simply

$$\rho(\boldsymbol{Q}) = \sum_{l=1}^{N} \exp[\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{R}_l] \qquad [43]$$

so that

$$\Delta\rho(\boldsymbol{Q}) = \sum_{l=1}^{N} \exp[\mathrm{i}\boldsymbol{Q}\cdot\boldsymbol{R}_l] - \bar{\rho}\delta(\boldsymbol{Q}) \qquad [44]$$

Therefore, with the exception of forward scattering along the direction of the incident light ($Q = 0$), the quantity $\Delta\rho(Q)$ may be replaced by $\rho(Q)$, and

$$\frac{1}{N}\langle\rho(\boldsymbol{Q})\,\rho(-\boldsymbol{Q})\rangle = \frac{1}{N}\left\langle\sum_{l,l'} \exp[-\mathrm{i}\boldsymbol{Q}\cdot(\boldsymbol{R}_l - \boldsymbol{R}_{l'})]\right\rangle \qquad [45]$$

The latter expression is identical to that of the static structure factor $S(Q)$ previously identified in the quantum derivation. Thus, the cross-section $d\sigma/d\Omega$, is again given by $N(d\sigma/d\Omega)_0 S(Q)$, except the basic unit of scattering from the system is now

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_0 = \frac{\boldsymbol{k}_\mathrm{v}^4}{(4\pi\varepsilon_0)^2}\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T^2 \left|\left(\tilde{I} - \boldsymbol{e}_r\boldsymbol{e}_r\right)\cdot\epsilon\right|^2 \qquad [46]$$

The double-differential cross-section is developed in the same way as before, and the result is identical to the expression given by Equation [25], which again involves the dynamic structure factor $S(Q, \omega)$. The latter can now also be written in terms of correlations involving Fourier components of the density fluctuations:

$$S(\boldsymbol{Q},\omega) = \frac{1}{2\pi N}\int_{-\infty}^{+\infty} \mathrm{d}t\exp[-\mathrm{i}\omega t] \times \langle\rho(-\boldsymbol{Q},0)\,\rho(\boldsymbol{Q},t)\rangle \qquad [47]$$

## Role of dynamic structure factors in spectroscopy

For light scattering, both the quantum and classical derivations of the double-differential cross-section result in an expression that is the product of $(d\sigma/d\Omega)_0$ and the dynamic structure factor, $S(Q, \omega)$. This decomposition of $d^2\sigma/d\Omega d\omega$ into two factors, the first being the angular cross-section from a single scattering unit and the second representing space and time-dependent structure within the system, is what makes the radiation useful as a spectroscopic tool. This separation occurs whenever the radiation couples weakly to the scattering medium. For example, the scattering of X-rays and thermal neutrons fall into this category. X-rays interact with the atomic electrons, and the basic scattering unit involves the classical electron radius $r_0 = (e^2/mc^2)/4\pi\varepsilon_0$ ($e$ and $m$ are the electron's charge and mass, respectively) and

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_0 = r_0^2\,|\epsilon_{\boldsymbol{k}\lambda}\cdot\epsilon_{\boldsymbol{k}'\lambda'}|^2\left(\frac{\boldsymbol{k}'}{\boldsymbol{k}}\right)^2 \qquad [48]$$

The scattering of thermal neutrons occurs because of the interaction between the neutrons and the atomic nuclei of the target. The basic scattering unit in this case is given by

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_0 = b^2\left(\frac{\boldsymbol{k}'}{\boldsymbol{k}}\right)^2 \qquad [49]$$

where $b$ is called the bound scattering length of an atomic nucleus.

The dynamic structure factor is a function of the two parameters $Q$ and $\omega$. Loosely speaking when the radiation imparts a momentum $\hbar Q$ and an energy $\hbar\omega$ to the system, it, in effect, probes the structure and dynamics of the system with a spatial resolution of

$Q^{-1}$ and a temporal resolution of $\omega^{-1}$. **Table 1** shows the regions of $(Q, \omega)$-space and the corresponding resolutions in real-space and real-time accessible to photon and neutron spectroscopies using current instrumentation and measurement techniques.

One often expresses the dynamic structure factor in the form

$$S(\boldsymbol{Q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}t \exp[-\mathrm{i}\omega t] F(\boldsymbol{Q}, t) \qquad [50]$$

with

$$F(\boldsymbol{Q}, t) = \left\langle \frac{1}{N} \sum_{l,l'} \exp[-\mathrm{i}\boldsymbol{Q} \cdot \boldsymbol{R}_l(0)] \exp[\,\mathrm{i}\boldsymbol{Q} \cdot \boldsymbol{R}_{l'}(t)] \right\rangle \qquad [51]$$

$F(Q, t)$ is known as the intermediate scattering function. The dynamic structure factor is just the temporal Fourier transform of $F(Q, t)$. For scattering measurements that involve the frequency, or energy, domain, one determines $S(Q, \omega)$ directly. Examples include optical mixing spectroscopy in light scattering, as well as crystal spectrometry and time-of-flight measurements in the case of thermal neutron scattering. On the other hand, the intermediate scattering function is directly measured in the time domain. Both photon correlation spectroscopy and neutron spin–echo spectroscopy measure $F(Q, t)$.

Detailed discussions of $S(Q, \omega)$ and $F(Q, t)$ for various target systems appear in the Further reading section. To illustrate the type of information that is contained in these functions, consider the specific case of scattering from a fluid consisting of $N$ identical, independently moving scattering centres (or particles). In this case, the scattering from the various particles adds incoherently and is determined by a simplified self intermediate scattering function governed by the motion of a single particle:

$$F_{\mathrm{S}}(\boldsymbol{Q}, t) = \langle \exp[-\mathrm{i}\boldsymbol{Q} \cdot \boldsymbol{R}(0)] \exp[\mathrm{i}\boldsymbol{Q} \cdot \boldsymbol{R}(t)] \rangle \qquad [52]$$

The brackets represent a thermal average at temperature $T$.

From purely classical considerations, the scattering function always reduces, at least approximately, to the Gaussian form

$$F_{\mathrm{S}}^{\mathrm{cl}}(\boldsymbol{Q}, t) = \exp[-\tfrac{1}{2} Q^2 W(t)] \qquad [53]$$

where $W(t)$ is a width function that corresponds to the mean-square displacement of the scattering particle as a function of time. The specific form of the width function depends on the type of motion exhibited by the particle. $W(t)$ for a particle (mass $M$) in an ideal gas ($V_0 = k_{\mathrm{B}}T/M$) and for a diffusing particle (diffusion coefficient $D$) are given by

$$W(t) = \begin{cases} V_0^2 t^2 & \text{for ideal gas} \\ 2Dt & \text{for diffusing particle} \end{cases} \qquad [54]$$

**Figure 4** is a comparison of the behaviour of these functions, along with $W(t)$ for a particle moving through a typical liquid. Fourier transforming the $F_{\mathrm{S}}^{\mathrm{cl}}(Q, t)$s gives the corresponding forms for the (self) dynamic structure factors:

$$S_{\mathrm{S}}^{\mathrm{cl}}(\boldsymbol{Q}, \omega) = \begin{cases} (2\pi Q^2 V_0^2)^{-1/2} \exp[-\omega^2/2V_0^2 Q^2] \\ \quad \text{for ideal gas} \\ \dfrac{1}{\pi}\left[\dfrac{DQ^2}{\omega^2 + (DQ^2)^2}\right] \\ \quad \text{for diffusing particle} \end{cases} \qquad [55]$$

**Table 1** Accessible regions of $(Q,\omega)$-space and corresponding resolutions in real space and time for various types of spectroscopies[a]

| *Type of spectroscopy* | *Q* *(μm⁻¹)* | *ħω* *(eV)* | *Spatial resolution* *(μm)* | *Temporal resolution* *(s)* |
|---|---|---|---|---|
| Optical mixing | $10^{-1}-10^{1}$ | $10^{-9}-10^{-6}$ | $10^{-1}-10^{1}$ | $10^{-10}-10^{-6}$ |
| Photon correlation | $10^{-1}-10^{1}$ | $10^{-15}-10^{-9}$ | $10^{-1}-10^{1}$ | $10^{-6}-10^{0}$ |
| X-ray | $10^{1}-10^{5}$ | $10^{1}-10^{2}$ | $10^{-5}-10^{-1}$ | $10^{-17}-10^{-16}$ |
| Thermal neutron | $10^{2}-10^{5}$ | $10^{-8}-10^{1}$ | $10^{-5}-10^{-2}$ | $10^{-16}-10^{-6}$ |
| Neutron spin-echo | $10^{1}-10^{2}$ | $10^{-9}-10^{-2}$ | $10^{-2}-10^{-1}$ | $10^{-14}-10^{-7}$ |

[a] Ranges are approximate and to nearest order of magnitude.

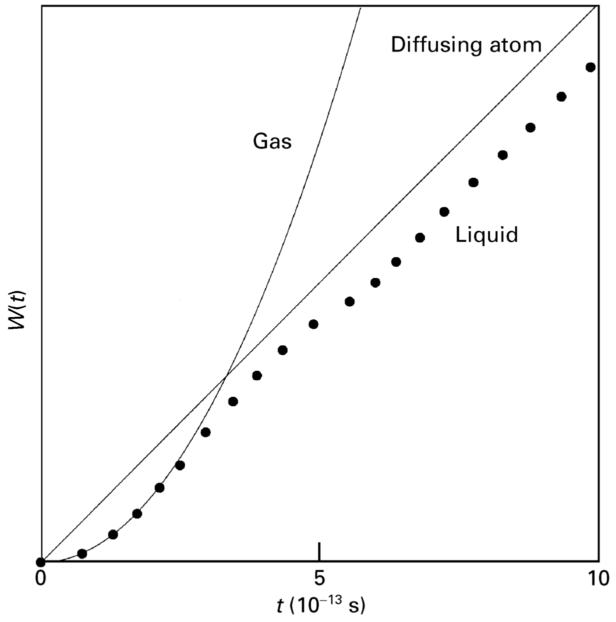



**Figure 4** Mean-square displacement function for a particle in an ideal gas, a diffusing particle, and a particle moving through a typical liquid.

For a fixed value of $Q$, the functions $S_S^{cl}(\mathbf{Q}, \omega)$ give the line shapes for the gas and diffusion cases to be Gaussian and Lorentzian, respectively.

The latter results neglect any quantum corrections to the self intermediate scattering functions. In the case of the ideal gas, the full quantum-mechanical expression for $S_S(\mathbf{Q}, \omega)$ is obtained from the corresponding classical one simply by multiplying by two exponential factors, i.e.

$$S_S(\mathbf{Q}, \omega) = \exp[\hbar\omega/2k_BT]\exp[-\hbar^2Q^2/8Mk_BT] \\ S_S^{cl}(\mathbf{Q}, \omega) \quad [56]$$

The first factor, $\exp(\hbar\omega/2k_BT)$, is known as the 'detailed balance factor' – it produces an asymmetry in the quantum-mechanical structure factor, whereas the classical one is an even function of $\omega$. The second factor, $\exp(-\hbar^2Q^2/8Mk_BT)$, can also be written as $\exp(-E_R/4k_BT)$, where $E_R = \hbar^2Q^2/2M$ is the recoil energy of the target particle. Hence this exponential factor is known as the recoil factor. Equation [56] is exactly true only in the ideal gas case; however, it is also approximately valid for other scattering systems as well.

## List of symbols

$|A\rangle$ = initial state of atom; $a_{k\lambda}$ = photon annihilation operator for mode $k\lambda$; $a_{k\lambda}^\dagger$ = photon creation operator for mode $k\lambda$; $b$ = bound scattering length of atomic nucleus; $\mathbf{B}$ = magnetic field vector; $|B\rangle$ = final state of atom; $c$ = speed of light in vacuum; $D$ = diffusion coefficient; d$P$ = time-averaged power scattered into solid angle d$\Omega$; d$R$ = rate of scattering into solid angle d$\Omega$; $(d\sigma/d\Omega)_l$ = angular differential cross-section for scattering by the $l$th atom; $(d\sigma/d\Omega)_{if}$ = angular differential cross-section for target in specific initial state $|E_i\rangle$ and final state $|E_f\rangle$; $d\sigma/d\Omega$ = angular differential cross-section; $d^2\sigma/d\Omega dE'$ = double differential cross-section; $(d\sigma/d\Omega)_0$ = angular differential cross-section from single scattering unit; $d^2\sigma/d\Omega d\omega$ = $d^2\sigma/d\Omega dE'$ multiplied by $\hbar$; $(d\sigma/d\Omega)_l^{\text{Rayleigh}}$ = angular differential cross-section for Rayleigh scattering by the $l$th atom; d$\Omega$ = differential solid angle; $e$ = electronic charge; $\mathbf{e}_r$ = unit vector in direction of $\mathbf{r}$; $E$ = energy of incident radiation quantum; $\mathbf{E}$ = electric field vector; $\mathbf{E}_0$ = electric field vector for incident wave; $E_A$ = initial energy of atom; $E'$ = energy of scattered radiation quantum; $E_f$ = energy of target medium after scattering; $|E_f\rangle$ = quantum state of target medium after scattering; $E_i$ = energy of target medium before scattering; $E_I$ = energy of intermediate state of radiation/target system; $|E_i\rangle$ = quantum state of target medium before scattering; $E_J$ = energy of intermediate atomic state; $E_m$ = energy of radiation/target system before scattering; $E_n$ = energy of radiation/target system after scattering; $E_R$ = recoil energy of target particle; $\mathbf{E}_s$ = electric field vector for scattered wave; $F(\mathbf{Q}, t)$ = intermediate scattering function; $F_S(\mathbf{Q}, t)$ = self intermediate scattering function; $F_S^{cl}(\mathbf{Q}, \mathrm{t})$ = classical self intermediate scattering function; $g(E_n)$ = density of final states; $\tilde{G}$ = Green's tensor; $\eta = h/2\pi$, where h is Planck's constant; $H_0$ = Hamiltonian of incident radiation + scattering medium, with no interaction present; $H_S$ = Hamiltonian of scattering medium; $H_R$ = Hamiltonian of incident radiation; $i$ = square-root of $-1$; $\tilde{I}$ = unit dyad; $|I\rangle$ = an intermediate state of radiation/target system; $|J\rangle$ = an intermediate atomic state; $k$ = wavenumber of incident radiation; $k_B$ = Boltzmann's constant; $k_v$ = wavenumber of light in vacuum; $\mathbf{k}$ = wavevector of incident radiation; $k'$ = wavenumber of scattered radiation; $\mathbf{k}'$ = wavevector of scattered radiation; $|\mathbf{k}\lambda\rangle$ = quantum state of incident radiation quantum; $|\mathbf{k}'\lambda'\rangle$ = quantum state of scattered radiation quantum; $L^3$ = volume of region in which radiation is confined; $m$ = mass of electron; $|m\rangle$ = quantum state of radiation/target system before scattering; $M$ = mass of scattering particle;

$n$ = index of refraction of medium; $|n\rangle$ = quantum state of radiation/target system after scattering; $n_{k\lambda}$ = number of photons in mode $k\lambda$; $\bar{n}$ = average index of refraction of medium; $N$ = number of atoms; $P_i$ = Maxwell–Boltzmann probability distribution for target in different initial energy states $|E_i\rangle$; $Q$ = wave vector transfer; $\boldsymbol{r}$ = position vector of field-point where radiation is detected; $\boldsymbol{r}'$ = position vector of source-point inside scattering medium; $r_0$ = classical electron radius; $R_l$ = location of $l$th atom; $S(Q)$ = static structure factor; $S_S(Q, \omega)$ = self dynamic structure factor; $S_S^{cl}(Q, \omega)$ = classical self dynamic structure factor; $S(Q, \omega)$ = dynamic structure factor; $t$ = time; $T$ = absolute temperature; $\mathbf{T}$ = transition matrix; $V$ = interaction potential between radiation and scatterer; $V_0$ = speed, $k_B T/M$; $W$ = transition probability per unit time; $W(t)$ = width function or mean-square displacement function; $\tilde{\alpha}_l$ = atomic polarizability tensor for $l$th atom; $\alpha_l$ = scalar polarizability of $l$th atom; $\Delta\rho(Q, t)$ = spatial Fourier transform of local number-density fluctuations in scattering medium; $\Delta n$ = local fluctuation in refractive index; $\Delta\varepsilon$ = local fluctuation in permittivity; $\Delta\rho(\boldsymbol{r}', t)$ = local fluctuation in number density in scattering medium; $\varepsilon$ = electric permittivity of scattering medium; $\varepsilon_0$ = electric permittivity of vacuum; $\epsilon_{k\lambda}$ = unit polarization vector for mode $k\lambda$; $\bar{\varepsilon}$ = average permittivity of medium; $\theta$ = polar angle; $\lambda$ = polarization of incident radiation; $\lambda'$ = polarization of scattered radiation; $\boldsymbol{\mu}_l$ = electric dipole moment of lth atom; $\boldsymbol{\mu}_{k'}^{l}$ = product of $\varepsilon_{k\lambda}$ and $\boldsymbol{\mu}_l$; $\boldsymbol{\mu}_{k'}^{l}$ = product of $\varepsilon_{k'\lambda'}$ and $\boldsymbol{\mu}_l$; $\bar{\rho}$ = average number density in scattering medium; $\rho(\boldsymbol{r}', t)$ = local number density in scattering medium; $\rho(Q, t)$ = spatial Fourier transform of local number density in scattering medium; $\sigma$ = scattering cross-section; $\Phi_{in}$ = incident flux; $\phi$ = azimuthal angle; $\omega$ = angular-frequency shift of radiation due to scattering; $\omega_0$ = angular frequency of light.

*See also:* **Electromagnetic Radiation; Inelastic Neutron Scattering, Applications; Inelastic Neutron Scattering, Instrumentation; Light Sources and Optics; Neutron Diffraction, Theory; Rayleigh Scattering and Raman Spectroscopy, Theory; Scattering and Particle Sizing, Applications; X-Ray Spectroscopy, Theory.**

## Further reading

Berne BJ and Pecora R (1976) *Dynamic Light Scattering*. New York: Plenum Press.

Bohren C and Huffman D (1983) *Absorption and Scattering of Light by Small Particles*. New York: John Wiley & Sons.

Chen SH and Kotlarchyk M (1997) *Interaction of Photons and Neutrons with Matter*. Singapore: World Scientific.

Chu B (1974) *Laser Light Scattering*. San Diego: Academic Press.

Foderero A(1971) *The Elements of Neutron Interaction Theory*. Cambridge, MA: M.I.T. Press.

Heitler W (1954) *The Quantum Theory of Radiation*. London: Oxford University Press.

Kerker M (1969) *The Scattering of Light and Other Electromagnetic Radiation*. New York: Academic Press.

Louisell WH (1973) *Quantum Statistical Properties of Radiation*. New York: John Wiley & Sons.

Lovesey SW (1984) *Theory of Neutron Scattering from Condensed Matter*. Oxford: Clarendon Press.

Mott NF and Massy HSW (1949) *The Theory of Atomic Collisions*, 2nd edn. Oxford: Oxford University Press.

# Sector Mass Spectrometers

**R Bateman**, Micromass, Wythenshaw, Manchester, UK

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| Methods & Instrumentation |

The magnetic sector is the oldest type of mass analyser. Experiments initiated in 1906 by JJ Thompson demonstrated the different deflections of various 'positive' rays in superposed crossed magnetic and electric fields. In 1912, using this equipment, the isotopes of neon were discovered, and over the next decade magnetic sector instruments built separately by Aston and by Dempster were used to identify most of the isotopes and determine their relative abundance. In 1931 Bainbridge added a velocity filter to Dempster's basic design to limit the energy distribution of ions entering the magnet. This refinement improved mass resolution and the relative mass measurement precision of the isotopes. This allowed the 'packing fraction' of the isotopes to be determined.

Industrial interest in mass spectrometry grew in the following decade and in 1942 the first commercial mass spectrometer, a magnetic sector instrument, was built by Consolidated Engineering Corporation and sold to the Atlantic Refining Company for the analysis of gasoline refining streams. Magnetic sector instruments are now only one of several different types of mass spectrometer manufactured commercially. Nevertheless magnetic sector mass spectrometers remain the instrument of choice for a number of important applications, in particular in the areas of target compound trace analysis, accurate mass measurement, isotope ratio measurement and fundamental ion chemistry studies.

## Principal of operation

Ions with mass ($m$) and charge ($ze$) accelerated through an electrical potential difference ($V$) will have velocity ($v$) and kinetic energy (KE) where:

$$\mathrm{KE} = zeV = \tfrac{1}{2}\, mv^2 \qquad [1]$$

Ions with charge ($ze$) moving through a magnetic field ($B$) with velocity ($v$) are subject to the Lorentz force ($F$) orthogonal to the direction of the field and direction of travel. Consequently ions travel with a circular trajectory with radius ($r_\mathrm{m}$) in which the centripetal force is provided by the Lorentz force:

$$F = B\ ze\ v = m(v^2/r_\mathrm{m}) \qquad [2]$$

Eliminating ($v$) from Equations [1 and 2]:

$$m/ze = B^2 r_\mathrm{m}^2/2V \qquad [3]$$

In the SI system of units for $B$ (tesla), $r_\mathrm{m}$ (metres), $V$ (volts) and $m/z$ (daltons per unit of electronic charge) Equation [3] becomes:

$$m/z = 4.786 \times 10^7 \cdot B^2 r_\mathrm{m}^2/V \qquad [4]$$

In its simplest form, the mass spectrometer transmits ions of a particular mass and charge from source to detector via a circular trajectory through a magnetic sector. Here the magnetic sector is a mass filter, and a mass spectrum can be recorded by scanning the magnetic field or accelerating voltage with serial detection of mass peaks. Alternatively, several detectors may simultaneously record several different ion masses, each taking a different trajectory. This principle may be extended to incorporate a continuous detector array in which a complete portion of the mass spectrum is recorded simultaneously. The magnet sector is usually constructed from a laminated iron-cored electromagnet with low inductance coils to allow fast scanning or switching. Superconducting magnets are not appropriate since they are not readily scanned. However, permanent magnets may be used for certain dedicated applications such as isotope ratio determinations.

## Optics

An ion moving in a magnetic field is dispersed with respect to its momentum ($\rho$). The momentum of an

ion is given by:

$$\rho = mv = (2m \cdot \mathrm{KE})^{1/2} \qquad [5]$$

Therefore ions with the same kinetic energy (KE) are, in effect, dispersed with respect to their mass. In addition, the shape of the magnetic sector can be designed to have ion directional focusing properties. A magnetic sector of a particular shape and size will have a particular combination of ion dispersion and directional focusing characteristics.

### Single focusing

An arrangement which combines an ion source with an ion beam width defining slit, a magnetic sector with convergent directional focusing characteristics and an ion 'collector' slit positioned at the image point of the source slit is defined as single focusing. The magnetic sector directional focusing characteristics can be designed to a very high order, but its imaging properties will be limited by any spread in ion energy. Consequently such instruments are used where the energy spread of ions is low, for example with an electron impact ionization source, and where a high resolution is not required (**Figure 1**).

### Stigmatic focusing

An inhomogeneous magnetic field has field components in more than one direction. If the magnetic field is inhomogeneous the focusing in the direction of dispersion is modified, and focusing in the direction orthogonal to the direction of dispersion is introduced. As the focal power in the direction of dispersion reduces, the focal power in the direction orthogonal to that increases, and when the focal power in these two directions become the same the sector has stigmatic focusing. This can provide more efficient ion transmission from source to detector, although higher order focusing terms in the direction of dispersion are usually compromised and ultimate resolution reduced.

A homogeneous magnetic field in which the shape of the sector has been designed to have non-normal ion entry and/or ion exit will also modify focusing in the direction of dispersion and introduce focusing in the orthogonal direction. In a similar way such a magnetic sector can be designed to provide stigmatic focusing (**Figure 2**).

### Mass dispersion and resolution

The mass dispersion coefficient ($D_m$) of a single focusing magnetic sector is proportional to the radius of curvature ($r_m$) of the ion beam trajectory in the magnetic field. The spatial separation ($y$) of two similar masses with mean mass ($m$) and mass difference ($\Delta m$) is related to the mass dispersion coefficient by:

$$y = D_m \cdot \Delta m/m \qquad [6]$$

The ion beam width ($w_b$) at the image position is related to the ion beam width defined by the source slit ($w_s$), the image lateral magnification ($M$) and the sum of the imaging aberration coefficients ($\alpha$) by:

$$w_b = M \cdot w_s + \alpha \cdot r_m \qquad [7]$$

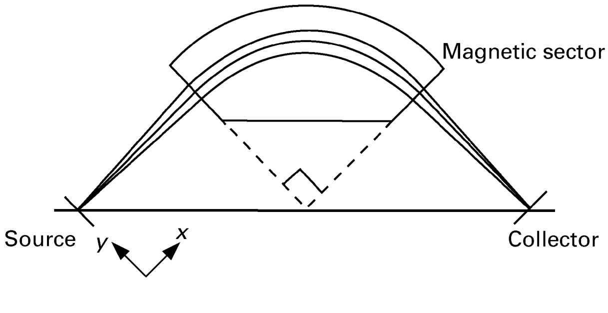



**Figure 1** Single-focusing magnetic sector, with focusing in '*Y*' direction only.

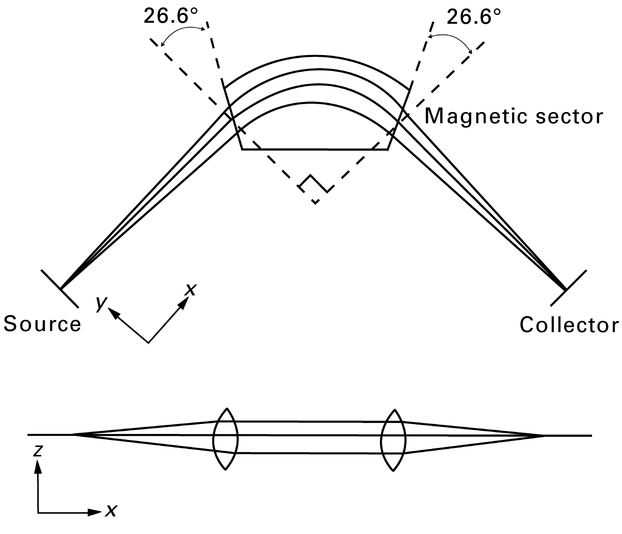



**Figure 2** Single-focusing magnetic sector, with stigmatic focusing in '*Y*' and '*Z*' directions.

The mass resolving power ($m/\Delta m$) for a 'collector' slit width ($w_c$) is given by:

$$m/\Delta m = D_m/(w_b + w_c) = D_m/(M \cdot w_m + w_c + \alpha \cdot r_m) \quad [8]$$

Thus, the mass dispersion coefficient and the slit widths are the most significant parameters in setting the resolution, and the ratio of dispersion to magnification ($D_m/M$) is a key figure of merit in the design of an instrument. However, ultimate resolution is limited by the imaging aberrations.

### Double focusing

It has been pointed out that the magnetic sector will disperse ions with respect to their momentum, and hence with respect to their mass if they are monoenergetic. However, ions will normally have a spread in kinetic energy, depending on the nature of the ion source, and this will broaden the image width. This usually becomes the limiting factor to achieving high resolution.

Momentum dispersion may be considered a combination of mass dispersion and energy dispersion. An electric sector will only disperse ions with respect to their energy, and so if an electric sector is combined with a magnetic sector the overall energy dispersion will be modified. A combination of magnetic sector and electric sector, which is directional focusing and in which the overall energy dispersion is zero, is said to be double focusing. Such a combination does not suffer the same image broadening due to spread in ion kinetic energy, and can achieve much higher resolution.

If the first sector has energy dispersion $D_{e1}$, and the second sector has energy dispersion $D_{e2}$, and image magnification of the second sector is $M_2$, then the overall energy dispersion ($D_e$) is

$$D_e = M_2 \cdot D_{e1} + D_{e2} = 0 \quad [9]$$

If the electric sector precedes the magnetic sector, the double-focusing arrangement is said to be 'forward' geometry, and 'reverse' geometry if the electric sector follows the magnetic. Either arrangement is equally effective as a double focusing optical system. However, the electric sector can provide additional benefits, which are different for each 'geometry'.

**Forward geometry** Here the electric sector, which precedes the magnetic sector, can be used to further advantage if designed to have a lateral magnification less than unity. This will increase the dispersion to magnification ratio ($D_m/M$) and the source slit will be larger than it otherwise would have been to achieve a required resolution. The larger source slit usually results in higher transmission and reduced susceptibility to contamination.

**Reverse geometry** Here the electric sector follows the magnetic sector, and therefore can no longer be used to increase the dispersion-to-magnification ratio. However, the electric sector can now be used as an energy filter to prevent transmission of ions with very different energies – such as the product ions from metastable ion decompositions. In a single-focusing system, or a 'forward' geometry double-focusing arrangement, these product ions appear as defocused artefact peaks in the spectrum when they are generated in the field-free region before the magnetic sector. If the decompositions occur in the magnetic field this usually results in an increase in background noise in the spectrum. In the 'reverse' geometry arrangement these are removed by the electric sector. The electric sector can also be used to further advantage by providing a means of analysing products from metastable ions resulting from ion decompositions in the field-free region between the magnetic and electric sectors. This type of analysis is considered in more detail in the Metastable ion analysis section.

**Split forward and reverse geometry** Here the electric sector is divided into two smaller electric sectors, positioned before and after the magnetic sector. Each electric sector can be smaller than the single electric sector in a 'forward' or 'reverse' geometry, and provided the overall energy dispersion is zero, the arrangement can still be double focusing. This arrangement can benefit both from an increased dispersion-to-magnification ratio, by appropriate design of the first electric sector, and from removal of artefact peaks and background noise – resulting from metastable ion decompositions – by virtue of the second electric sector. Again the final electric sector can provide the means for directly analysing the products from metastable ions dissociating in the field-free region immediately preceding that electric sector (**Figure 3**).

## Parallel detection

Unlike the quadrupole type of mass filter, a magnetic sector may be designed to record the signal from several different masses simultaneously. This is referred to as parallel detection.

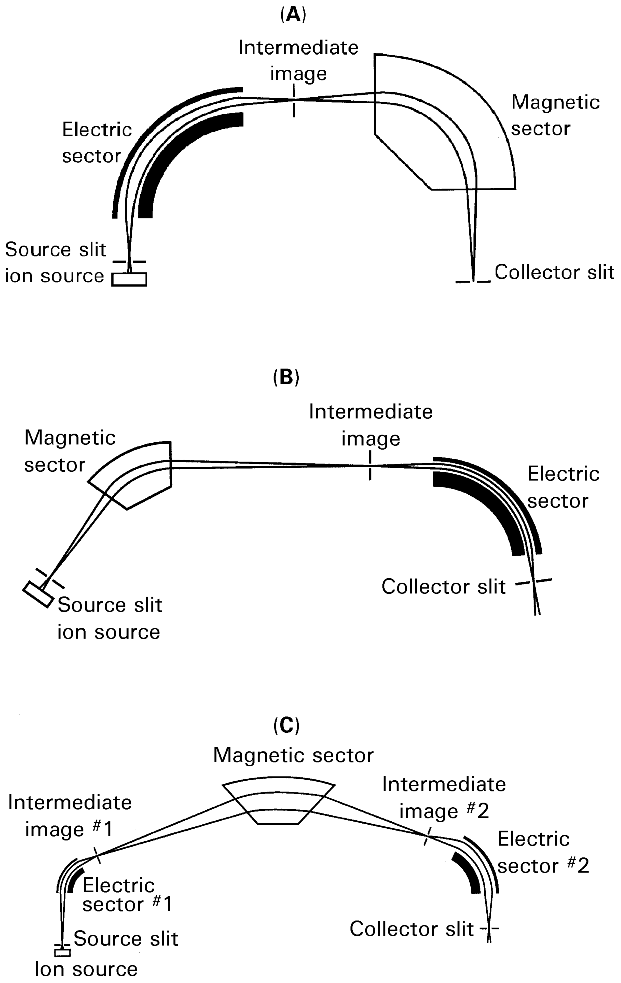



**Figure 3** Modified 'Nier–Johnson'-type double-focusing magnetic sector mass spectrometer (A) with 'forward' geometry, (B) with 'reverse' geometry, (C) with split 'forward' and 'reverse' geometry.

## Multiple collectors

Parallel detection provides a means of accurately recording the relative abundance of two or more different masses, since measurement of the ratio of these peak intensities is not susceptible to fluctuations or drift in the ionization source, or to rapidly changing sample concentration such as encountered in chromatography. Magnetic sector mass spectrometers designed to simultaneously transmit a number of different masses, each with a different trajectory, and incorporating multiple discrete detectors, are used to make the most accurate isotope ratio determinations. Instruments designed specifically for accurate isotope ratio determinations have included combinations of up to 16 Faraday and/or ion counting detectors, and isotope ratios may be determined to a precision of better than 10 ppm using such equipment.

## Continuous array detectors

A magnetic sector mass spectrometer with a single detector may be used to record a mass spectrum by scanning and sequentially detecting mass peaks. The duty cycle for recording each mass in the spectrum is generally poor, and the higher the resolution or the wider the mass range the poorer the duty cycle. An array detector allows simultaneous acquisition over a range of masses thereby improving the duty cycle when used to record a spectrum. Array detectors employing high-density arrays of discrete charge-sensitive detectors or 'single ion' position sensitive detectors are very sensitive, although they are usually limited in size. This imposes a limit on the combination of mass resolution and mass range of simultaneous detection. For example, an array of 2048 discrete detectors can be arranged to simultaneously detect ions over a 10% mass range with a resolution of ~ 4000 (FWHM). A 'single ion' position-sensitive detector simultaneously recording over a similar mass range may achieve approximately twice that resolution, but cannot cope with two ions arriving within one measurement period. This imposes an upper limit on the total ion current onto the detector of between $10^5$ and $10^6$ ions/s, which in turn imposes a limit on the practical mass range for simultaneous detection.

Alternatively, a large-scale photosensitive emulsion plate detector may be used to simultaneously record a large mass range with acceptable mass resolution. Such detectors require the mass spectrometer design to allow simultaneous transmission of a wide range of masses and focusing on to a flat plane. The double focusing arrangement first described by Mattauch and Herzog in 1934 gives first-order double focusing over the entire length of the 'photo-plate'. Mass spectrometers based on this theory have been constructed to successfully record ions simultaneously over a 60:1 mass ratio (**Figure 4**).

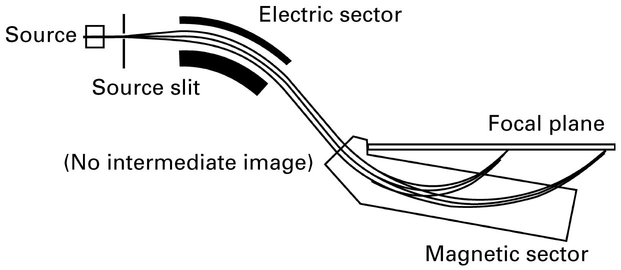



**Figure 4** Modified 'Mattauch–Herzog'-type double-focusing magnetic sector with flat focal plane.

### High resolution

The combination of magnetic and electric sectors to form a double-focusing arrangement includes sufficient degrees of freedom in the choice of design to allow higher order focusing to be achieved. A design constructed by Nier and Roberts in 1951, consisting of a 90° electric sector and 60° magnetic sector, arranged sequentially in a C-shaped geometry, was shown by Nier and Johnson to have first-order energy focusing and second-order directional focusing. Many modern designs of double-focusing instruments, in which all second-order directional- and energy-focusing terms are either zero or near zero, are essentially variations of the 'Nier–Johnson' geometry. Commercially manufactured instruments of this type achieve resolving powers in excess of 150 000 by the 10% valley definition (based on peak width at 5% height) or in excess of 350 000 by the FWHM definition (based on peak width at 50% height). They are used in particular in the petroleum industry for characterization of complex oil mixtures (**Figure 5**).

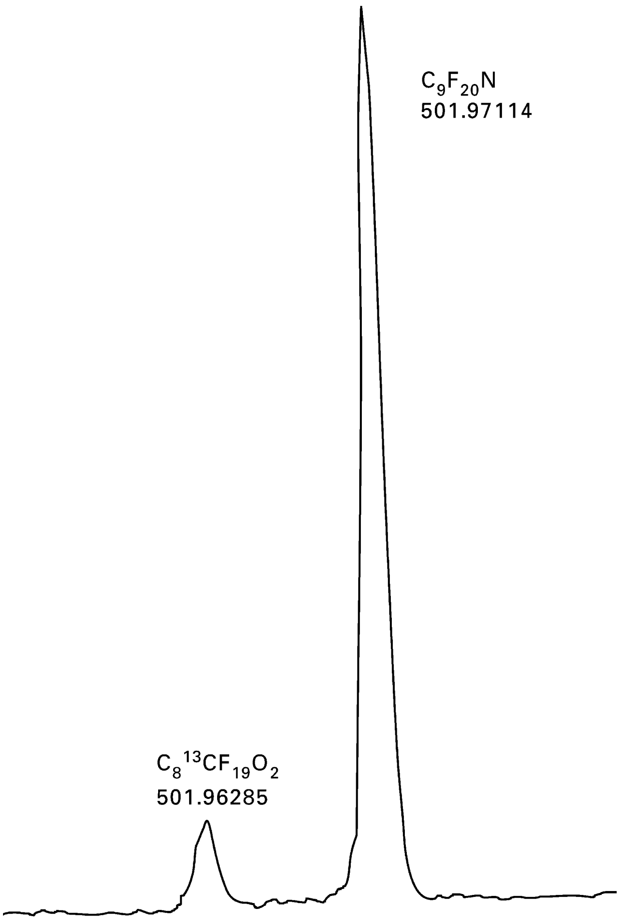



**Figure 5** Doublet from PFTBA and FOMBLIN showing a resolving power of 200 000 (10% valley definition), or 450 000 (full width at half height definition).

## Accurate mass measurement

The capacity for high resolution enables double-focusing magnetic-sector mass spectrometers to be used for accurate mass measurement. Accurate mass is always determined by reference to one or more known masses simultaneously introduced into the mass spectrometer, and there are two different methods in common use.

### Peak matching

This is the simplest and generally most accurate method of measuring the mass of a single peak. Here the magnetic field strength is held constant and the electric sector field and all other voltages are switched together such as to transmit in turn the known and unknown mass peaks. The difference between the two masses, ideally, is small. A narrow voltage scan is applied as each mass is selected to display the peak profile, and the voltages are adjusted until the two peak profiles are exactly superimposed. The unknown mass is calculated from the ratio of the two voltage settings and the known reference mass. A refinement of this method entails the use of two known reference masses, preferably 'bracketing' the unknown mass, and adjusting the switched voltages until all three peak profiles are exactly superimposed. This allows correction for certain instrumental offsets, and a mass accuracy of better than 1 ppm is routinely achievable.

Alternatively voltage scanning through each of the three mass peaks while measuring the peak profiles with a high sampling rate analogue-to-digital converter, and computing the 'centre of mass' of each peak profile, will yield a similar mass measurement accuracy.

### Magnetic scanning

The 'peak matching' method is not appropriate if it is required to quickly measure the mass of a large number of peaks over a wide mass range. Here it is more appropriate to scan the magnetic field instead. However, the magnetic field strength from an iron-cored magnet is difficult to scan precisely. Consequently, a reference mixture is simultaneously introduced into the ion source such as to give a large number of known reference mass peaks at regular intervals throughout the mass range of interest. It is important that these mass peaks do not 'interfere' with any other mass peaks in the spectrum. The

known and unknown masses are 'digitized' and recorded into a data system, and from each peak centre the accurate mass of the unknown peaks may be calculated. Perfluorokerosene (PFK) is commonly used as a 'reference mixture' since its 'electron impact' spectrum exhibits peaks every 12 or 14 daltons from *m*/*z* 31 to beyond *m*/*z* 800. Furthermore, the peaks are mass deficient and may easily be resolved from most other organic compounds. This method can yield an average mass measurement accuracy of about 1 milli-dalton over this mass range.

## High-resolution selective ion recording

A double-focusing magnetic-sector mass spectrometer can be used to select and record the response from target compounds at high resolution and with a high sensitivity. The high resolution enables chemical background masses to be eliminated and consequently allow a lower detection level to be achieved. Selected ion recording provides a much better duty cycle, and therefore sensitivity, than scanning.

The detection and quantification of polychlorinated dibenzo-*p*-dioxins, and in particular the 2,3,7,8-tetrachlorinated dibenzo-*p*-dioxin (2,3,7,8-TCDD), is a major application for double-focusing magnetic-sector mass spectrometers. Despite extensive clean-up procedures, samples still contain compounds such as polychlorinated biphenyls and benzyl phenyl ethers, which have the same nominal masses as the compounds of interest. The sample is 'spiked' with a known amount of the $^{13}C$ isotope labelled form of 2,3,7,8-TCDD, introduced via gas chromatography and recorded by 'high-resolution' mass spectrometry. The measurement is quantified by comparison of the native dioxin response to that from the $^{13}C$-labelled form, and verified by confirmation of the ratio of the major isotopes of both the native and the $^{13}C$-labelled dioxins. At 10 000 resolving power (10% valley definition) the detection level for 2,3,7,8-TCDD is about 1 femtogram, or 3 attomole (**Figure 6**).

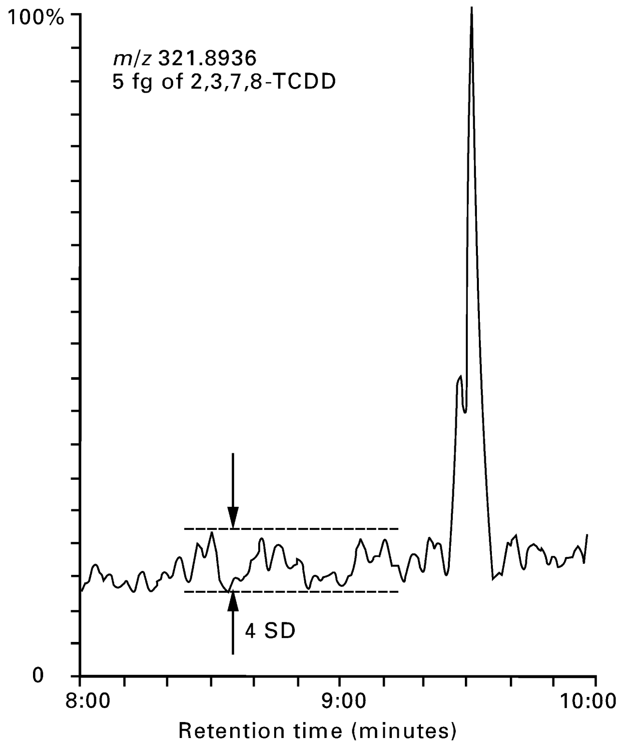



**Figure 6** Detection of 5 femtogram of 2,3,7,8-TCDD with S/N greater than 10:1 from an injection of 1 µL onto a GC capillary column and 'selective ion recording' of four masses plus a 'lock mass' at 10 000 resolving power (10% valley definition).

## Metastable ion analysis and MS/MS

The decomposition of a precursor or 'parent' molecular ion to a product or 'daughter' ion, while in flight through a single focusing magnetic sector mass spectrometer, can give rise to an artefact peak in the mass spectrum. However, such artefact peaks can be instructive in understanding the structure of the precursor ion. Methods have been developed for double-focusing instruments to eliminate such artefact peaks from mass spectra, to specifically record daughter ions from a selected parent ion, and to detect specific parent–daughter ion transitions as a means of detecting target compounds or classes of compounds.

If a parent ion (mass $m_p$) fragments in a field-free region the daughter ion (mass $m_d$) retains the velocity of the parent, only modified by the addition of a velocity component as a result of energy released in the decomposition reaction. Hence, the ratio of the daughter ion kinetic energy to that of the parent is equal to ratio of their masses ($m_d/m_p$). The additional velocity component is randomly distributed and this is seen as a superimposed energy spread.

### Single-focusing magnetic sector

If a parent ion fragments in the flight path between the source and the magnetic field, or the 'first field-free region', the daughter ion has an energy relative to that of the parent in proportion to the ratio of their masses ($m_d/m_p$). Ions originating in the ion source are accelerated to the same kinetic energy. Consequently, after traversing the magnetic sector, the daughter ion appears at a different mass ($m^*$) when viewed relative to the normal spectrum. The

apparent mass is given by:

$$m^* = m_d \cdot (m_d/m_p) \quad [10]$$

The daughter ion peak is often recognizable since it is usually broader as a result of its increased energy spread. The parent and daughter ions can often be deduced from $m^*$ with reasonable confidence if the sample is pure, but not if a mixture. This can provide useful additional information or become a source of noise depending on the circumstances.

### Double-focusing magnetic sector

If a parent ion fragments in the 'first field-free region' of a double-focusing mass spectrometer, the daughter ion, which will have a lower kinetic energy, is unlikely to be transmitted by the electric sector and therefore will not appear in the mass spectrum. If a parent ion fragments in the field-free region preceding the final sector the situation is more complex. The daughter ion can be transmitted through the final (magnetic) sector of a 'forward geometry' arrangement just as it would be in a single-focusing sector design, but will not be transmitted through the final (electric) sector of a 'reverse geometry' or 'split geometry' arrangement. In these latter cases, if the final electric sector only is scanned, the daughter ion will be transmitted when the electric field is at value equal to ($m_d/m_p$) of the value required to transmit 'full energy' ions originating from the source. The result is a spectrum of all the daughter ions of the parent ion transmitted by the magnetic sector. The electric sector scan is a spectrum of ion energies, and therefore will also exhibit the energy distribution of daughter ions, albeit at the cost of low 'mass' resolution. This is known as a 'mass analysed ion kinetic energy spectrum' (MIKES).

Transitions occurring in the 'first field-free region' of a double-focusing mass spectrometer can be studied by scanning the electric sector ($E$) and magnetic sector ($B$) fields synchronously such that the ratio $B/E$ is constant, determined by the value of the selected precursor ion mass ($m_p$). This is known as '$B/E$ linked scanning'. The dissociation energy information is lost as a result of the double-focusing action of the combined electric and magnetic sectors but on the other hand the observed daughter ion 'mass' resolution is increased.

Different types of linked magnetic sector field ($B$) and electric sector field ($E$) scans can yield different information from transitions in the first field-free region, and from the penultimate field-free region where the final sector is an electric sector. These include 'parent ion scans' in which spectra of all the parent ions from which a selected daughter ion ($m_d$) are obtained, and 'constant neutral loss scans' in which spectra of all the transitions in which the same neutral loss ($m_p - m_d$) occur. These are listed in **Table 1**.

**Table 1** Linked magnetic field ($B$) and electric field ($E$) scans for the study of metastable ion decompositions on double-focusing magnetic sector mass spectrometers

| | *'Field-free region' in which decomposition takes place* | |
|---|---|---|
| *Type of analysis* | *First* | *Penultimate (Final sector is electric)* |
| Daughter ions ($m_d$ = constant) | $B/E$ = constant | $B$ = constant ($E$ scan only) |
| Parent ions ($m_p$ = constant) | $B^2/E$ = constant | $B^2E$ = constant |
| Constant neutral loss ($m_p - m_d$ = constant) | $(B/E)^2(1-E')$ = constant | $B^2(1-E')$ = constant |

$E' = E/E_o$, where $E_o$ is the electric field for transmission of undissociated parent ions.

Sector mass spectrometers with two or more sectors, and gas 'cells' in which 'high energy' ion–molecule collisions take place, are used to study ion and neutral chemistry. Tandem sector mass spectrometers with up to six sectors (two magnetic and four electric sectors in EBEEBE sequence) have been constructed to study ion–molecule interactions, multistep dissociations and neutralization–reionization processes.

## List of symbols

$B$ = magnetic field; $D_{e1}$ = energy dispersion (first sector); $D_{e2}$ = energy dispersion (second sector); $D_m$ = mass dispersion coefficient; $E$ = electric sector field; $F$ = Lorentz force; KE = kinetic energy; $m$ = mass; $m_d$ = mass of daughter ion; $m_p$ = mass of parent ion; $m^*$ = apparent mass; $M$ = magnification; $r_m$ = radius; $v$ = velocity; V = potential difference; $w_b$ = ion beam width; $w_c$ = collector slit width; $w_s$ = source slit width; $ze$ = charge; $\alpha$ = imaging aberration coefficient; $\rho$ = momentum.

*See also:* **Ion Dissociation Kinetics, Mass Spectrometry; Ion Imaging Using Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Isotope Ratio Studies Using Mass Spectrometry; Mass Spectrometry, Historical Perspective; Metastable Ions; MS–MS and MS$^n$; Neutralization-Reionization in Mass Spectrometry.**

## Further reading

Beynon JH (1960) *Mass Spectrometry and its Applications to Organic Chemistry*. Amsterdam: Elsevier.
Cooks RG, Beynon JH, Caprioli RM and Lester GR (1973) *Metastable Ions*. Amsterdam: Elsevier.
Duckworth HE (1958) *Mass Spectroscopy*. Cambridge: Cambridge University Press.
Enge H (1967) In: Septier A (ed) *Focusing of Charged Particles*, Chapter 4.2, p 203. New York: Academic Press.
Hintenberger H and Koenig LA (1959) Mass spectrometers and mass spectrographs corrected for image defects. In: Waldron J (ed) *Advances in Mass Spectrometry*, Vol 1, pp 16–35. Oxford: Pergamon Press.
Jennings KR (1983) In: McLafferty FW (ed) *Tandem Mass Spectrometry*, Chapter 9.
McDowell CA (1963) *Mass Spectrometry*, New York: McGraw-Hill.
Milne GW (1971) *Mass Spectrometry: Techniques and Applications*. New York: Wiley Interscience.
Wollnik H (1987) *Optics of Charged Particles*. New York: Academic Press.

# Selenium NMR, Applications

*See* **Heteronuclear NMR Applications (O, S, Se, Te).**

# SIFT Applications in Mass Spectrometry

**David Smith**, Keele University, Stoke-on-Trent, UK
**Patrik Španěl**, Czech Academy of Science, Prague, Czech Republic

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
| --- |
| Applications |

## Introduction

The selected ion flow tube, SIFT, technique is a fast-flow tube/ion-swarm method for the study of the reactions of ions (positive or negative) with atoms and molecules under truly thermalized conditions over a wide range of temperature. It has been extensively used to study ion–molecule kinetics. Its application to atmospheric and interstellar ion chemistry by several eminent groups over a 20-year period has been crucial to the advancement and understanding of these interesting topics. Recently it has been developed as a very sensitive analytical technique for the detection and quantification of trace gases in air and in human breath down to the ppb level and in real time.

## Principle of the SIFT

The SIFT technique was developed because the original fast-flow-tube method, the eminently productive flowing afterglow technique, was not suitable for the study of the reactions considered at that time to be involved in the complex ion chemistry of interstellar clouds. The SIFT (and the flowing afterglow) are flow-tube 'swarm' experiments. The 'ideal' swarm experiment involves the creation of an ensemble of reactant charged particles of number density $n_1$ in an inert buffer gas of number density $n_2$ such that $n_1 << n_2$. Then multiple collisions between the charged particles and the buffer gas ensure the randomization (Maxwellianization) of the charged particle velocities and the relaxation of the charged particle mean energies (which may be high initially) to those appropriate to the buffer gas temperature, $T_g$. Then the introduction of reactant neutral atoms or molecules at a number density, $n_3$, ($n_3 << n_2$) initiates an ion neutral reaction. Thus the rate of change of $n_1$ as a function of $n_3$ can be determined, and the rate of coefficient, $k$, for the reaction can be

calculated since:

$$\frac{\mathrm{d}n_1}{\mathrm{d}t} = -kn_1n_3 \Rightarrow n_1(t) = n_1(0)\exp(-kn_3t)$$
$$\Rightarrow \ln n_1(t) = \ln n_1(0) - kn_3t \quad [1]$$

The $k$ and the ion products of the reaction (determined at the defined temperature $T_g$) are the important parameters needed to model the ion chemistry of ionized media. In a flow system the reaction time $t$ is the length of the reaction zone divided by the ion-flow velocity.

In the SIFT apparatus the ions are created in an ion source which is external to the flow tube. The ions are then extracted from the ion source, selected according to their mass-to-charge ratio using a quadrupole mass filter (see **Figure 1**) and injected into a flowing carrier gas (usually helium at a pressure of 0.5 Torr) via a small orifice (typically ~1 mm diameter). The carrier gas is inhibited from entering the quadrupole mass filter chamber by injecting it into the flow tube through a Venturi-type inlet at near-supersonic velocity in a direction away from the orifice (see **Figure 1**). In this way a swarm of a single-ion species thermalized at the same temperature as the carrier gas are convected along the flow tube (which is usually ~1 m long), sampled by a downstream pinhole orifice, mass analysed and counted by a differentially pumped quadrupole mass spectrometer system.

To study a particular ion–molecule reaction, a reactant gas is introduced at a measured flow rate into the carrier gas. Then by measuring the count rates, $I$, of the reactant and product ions using the downstream mass spectrometer and relating them to the number destiny of the reactant gas in the carrier gas, the $k$ for the reaction (following Eqn [1]) and the ion products can readily be determined (see **Figure 2**). Some SIFT apparatuses can be operated over the wide temperature range 80–600 K. Constructional details for a SIFT are given in some of the cited reviews.

No electrons are present in the carrier gas and therefore the SIFT medium is not a gaseous plasma medium like the flowing afterglow but rather a swarm of positive or negative ions in the carrier gas. Notwithstanding the great success of the flowing afterglow for the study of ion–neutral reactions, it has a serious limitation in that the primary reactant ions may react with their parent (source) gas which is usually present in the carrier gas, and this complicates the interpretation of the mass spectrometric data. This is avoided in the SIFT by the upstream external ion source/mass filter system. So, for example, if methane is introduced into the SIFT ion source, $C^+$, $CH^+$, $:CH_2^+$, $CH_3^+$ or $CH_4^{\bullet+}$ (and even $CH_5^+$) can be

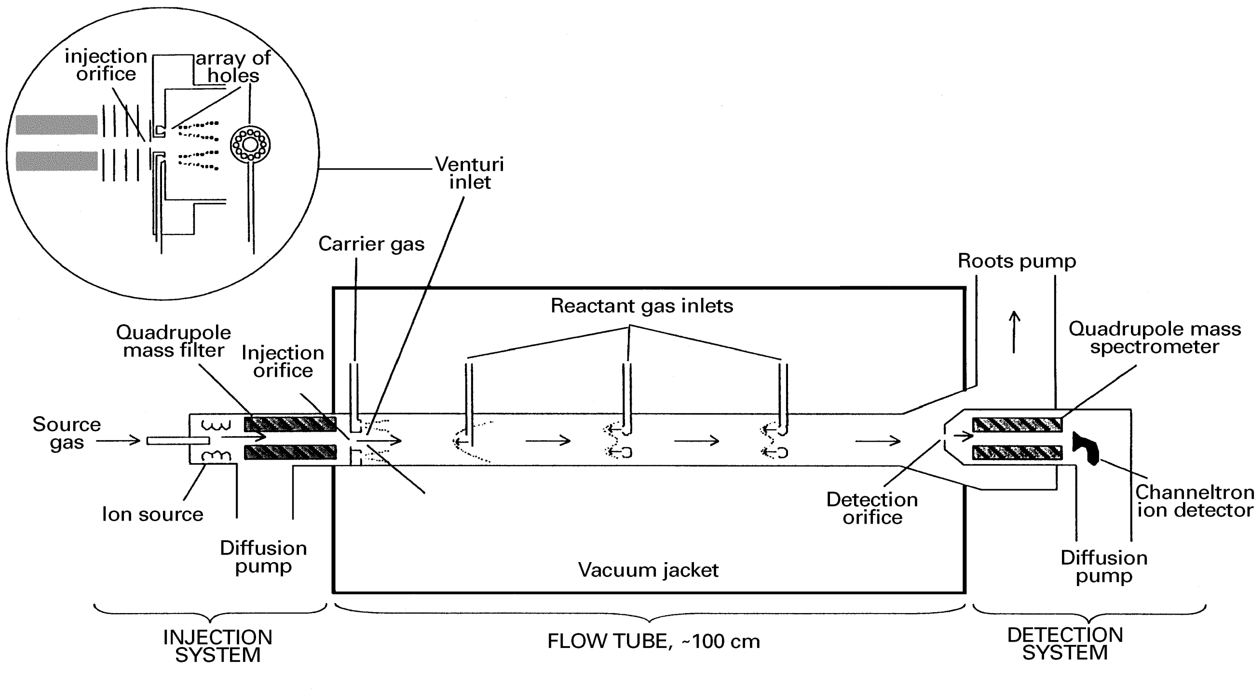



**Figure 1** Schematic representation of a SIFT apparatus. One form of Venturi-type inlet is shown in the upper-left part of the diagram. The vacuum jacket facilitates operation at high and low temperatures.

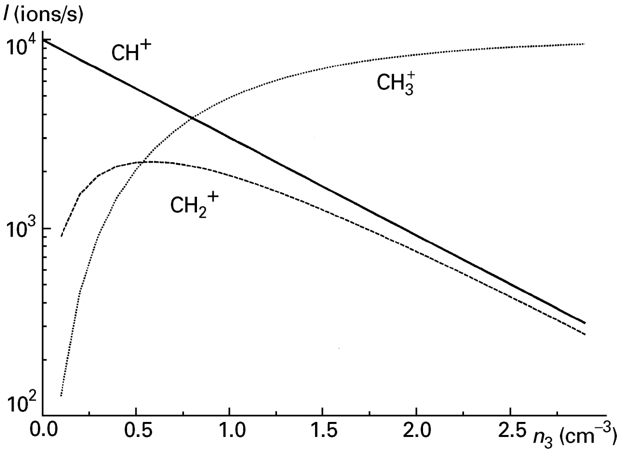



**Figure 2** Dependence of the ion count rate on the $H_2$ number density for the reaction of $CH^+$ with $H_2$. The rate coefficient, *k*, is obtained from the slope of the linear decay plot following Equation [1]. The primary product ion is $CH_2^+$. The $CH_3^+$ ion is produced in the secondary reaction of $CH_2^+$ with $H_2$.

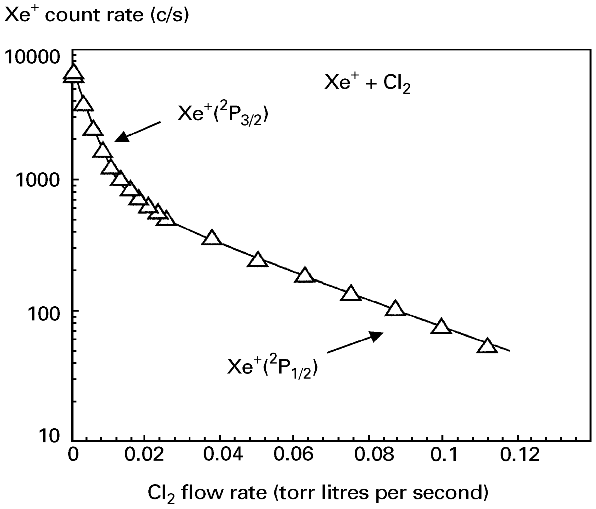



**Figure 3** A typical SIFT decay (c/s = counts per second) obtained when two differently reacting species are present at the same mass. In this example, the excited ion $Xe^+(^2P_{1/2})$ reacts more slowly with $Cl_2$ than the ground state $Xe^+(^2P_{3/2})$ ion.

separately injected into the flowing carrier gas. Significantly, the $CH_4$ source gas is not present in the helium carrier gas and cannot therefore confuse kinetics studies.

The rate coefficients and ion products of the reactions with many gases of practically any positive or negative ion species can be studied using the SIFT, provided that the ions can be extracted from the ion source at a sufficient current (~$10^{-9}$ A is the practical lower limit) and can be injected at a sufficiently low energy to avoid their fragmentation in collisions with the carrier gas. Even weakly bound species such as $H_3O^+(H_2O)_3$ ions have been injected without undue dissociation. Low-pressure and high-pressure election-impact sources and even flowing afterglow sources are routinely used to prepare a wide variety of positive and negative ions. For our recent development of the SIFT for trace gas analysis, a microwave discharge source is used (see below).

A fraction of the ions produced in SIFT ion sources may be electronically and vibrationally excited and after injection survive in the inert carrier gas. Their reactions can then be studied and if their reactivity is different from their ground state analogues the decay curves from which the rate coefficients are derived are nonlinear (see **Figure** 3). However, excited ions can often be relaxed to their ground states by the addition of a suitable 'quenching' gas.

Using SIFT apparatuses, the rate coefficients and products ions for a large number of positive ion–neutral and negative ion–neutral reactions have been studied in several laboratories around the world. They include the bimolecular reactions of doubly charged ions, electronically and vibrationally excited ions and cluster ions, and termolecular association reactions, some over a wide range of temperature. This has led to a greater understanding of the mechanisms, kinetics and energetics of such reactions and also to a clearer understanding of the ion chemistry of ionized media.

## Ion chemistry of the terrestrial atmosphere and interstellar clouds

The SIFT is ideally suited to the study of the ion chemistries of the terrestrial atmosphere, TA, and interstellar clouds, ISC. These are ionized media in which gas phase ion–neutral reactions occur which produce the exotic ions and molecules observed in these regions. The challenge to ion chemists is to identify these reactions, and to this end the SIFT has made a major contribution.

### Atmospheric ion chemistry

In the tenuous upper TA (above 100 km altitude; the ionospheric E- and F-regions) the ion chemistry is simple. Only bimolecular reactions occur between the precursor ions $H^+$, $He^{\bullet+}$, $O^+$, $N^+$, $O_2^{\bullet+}$ and $N_2^{\bullet+}$ (formed by photoionization) and the ambient neutrals O, $O_2$ and $N_2$, e.g.

$$\begin{aligned} He^{\bullet+} + N_2 &\rightarrow N_2^{\bullet+} + He \\ &\rightarrow N^+ + N^{\bullet} + He \end{aligned} \qquad [2]$$

The $k$ and the product ion distributions for such reactions are readily determined using the SIFT. Rocket-borne mass spectrometers have shown that $NO^+$ ions are a major species in the upper TA, even though neutral NO does not exist in measurable concentrations. Flowing afterglow and SIFT experiments have shown that these ions result primarily from the ion–molecule reaction:

$$O^{\bullet+} + N_2 \rightarrow NO^+ + N^\bullet \quad [3]$$

The ion chemistry of the upper TA (summarized in **Figure 4**) proceeds to convert the energetic precursor ions to the less energetic, predominantly diatomic molecular ions, which can be neutralized by the ambient free electrons via the process of dissociative recombination as shown. It was also realized that metastable electronically excited ions of $O^+$ and $O_2^{\bullet+}$ are produced in the upper ionosphere and so it became important to study the reactions of these excited ions. The SIFT is well suited to these studies (excited ions are quenched in flowing afterglow plasmas). The chemistry of these excited species is represented in **Figure 4** by the thick lines.

The ion chemistry of the lower TA (below 100 km) is dominated by termolecular (three-body) reactions of both positive ions and negative ions. At altitudes between about 50 to 90 km, in the ionospheric D-region, most ionizing solar radiations have been filtered out except for $L_\alpha$ and $L_\beta$ radiation, which can selectively ionize NO and $O_2(^1\Delta_g)$ molecules. So the initial ions in the positive ion chemistry are $NO^+$ and $O_2^{\bullet+}$.

SIFT and flowing afterglow studies have shown that both $NO^+$ and $O_2^{\bullet+}$ are relatively unreactive in bimolecular collisions with the major ambient atmospheric neutrals, but they both undergo termolecular (three-body) reactions forming weakly bonded association ions, an important initial reactions in the D-region being:

$$NO^+ + N_2 + N_2 \rightarrow NO^+ \cdot N_2 + N_2 \quad [4]$$

These weakly bonded $NO^+{\cdot}N_2$ ions react efficiently with the ambient $CO_2$ and $H_2O$ molecules in the 'switching reactions':

$$NO^+ \cdot N_2 + CO_2 \rightarrow NO^+ \cdot CO_2 + N_2$$
$$NO^+ \cdot CO_2 + H_2O \rightarrow NO^+ \cdot H_2O + CO_2 \quad [5]$$

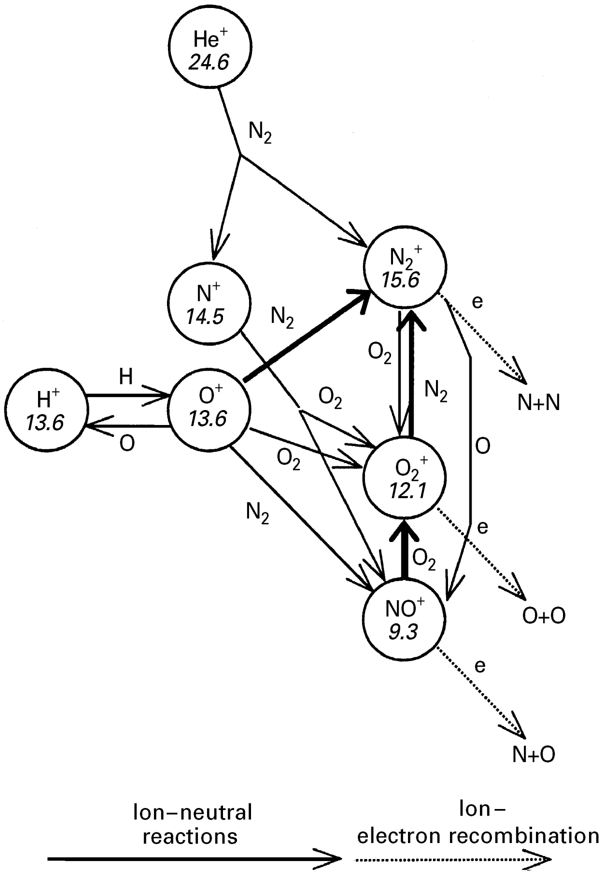



**Figure 4** The ion chemistry of the upper terrestrial atmosphere. Only biomolecular ion–neutral reactions occur, and ion–electron reactions (dissociative recombination) maintain the ionization equilibrium.

Termolecular $H_2O$ addition reactions build up the $NO^+$ hydrates $NO^+(H_2O)_{2,3}$. Finally the $NO^+$ is 'switched out' from the latter ions thus:

$$NO^+(H_2O)_3 + H_2O \rightarrow H_3O^+(H_2O)_2 + HNO_2 \quad [6]$$

Further termolecular reactions produce the hydrated hydronium ions $H_3O^+(H_2O)_n$, which are observed to be the major positive ions in this altitude region.

In the lowest altitude region of the TA the initial positive ions, e.g. $O_2^{\bullet+}$ and $N_2^{\bullet+}$, formed by cosmic ray ionization, are quickly converted to the dimer ions $O_4^{\bullet+}$ and $N_4^{\bullet+}$, which undergo switching reactions with the relatively abundant $H_2O$ and $CO_2$ to finally produce $H_3O^+(H_2O)_n$ (see the left part of **Figure 5**). The chemistry does not stop here, because the many minor reactive species that exist in the lowest part of TA, such as the bases $NH_3$, $CH_3CN$ and $CH_3OH$, undergo ligand switching with the hydrated

hydronium ions producing 'mixed' cluster ions, e.g.

$$H_3O^+(H_2O)_n + CH_3CN \rightarrow H^+(H_2O)_n \cdot CH_3CN + H_2O \quad [7]$$

Switching reactions like [7] can be studied in the SIFT because of the facility to inject cluster ions without undue break-up.

The negative ion chemistry in the lower TA (see the right part of **Figure 5**) is initiated by the electron attachment reactions:

$$\begin{aligned} O_2 + \text{electron} + M &\rightarrow O_2^- + M \\ O_3 + \text{electron} &\rightarrow O^- + O_2 \end{aligned} \quad [8]$$

This chemistry follows a similar pattern to the positive ion chemistry in that the initial ions $O^-$ and $O_2^-$ are converted to $O_3^-$ and $O_4^-$ and these undergo switching reactions with $H_2O$ and $CO_2$ to produce ions such as $O_2^- \cdot H_2O$, $O_2^- \cdot CO_2$ and $CO_3^-$. The further reactions of these ions with other minor neutral

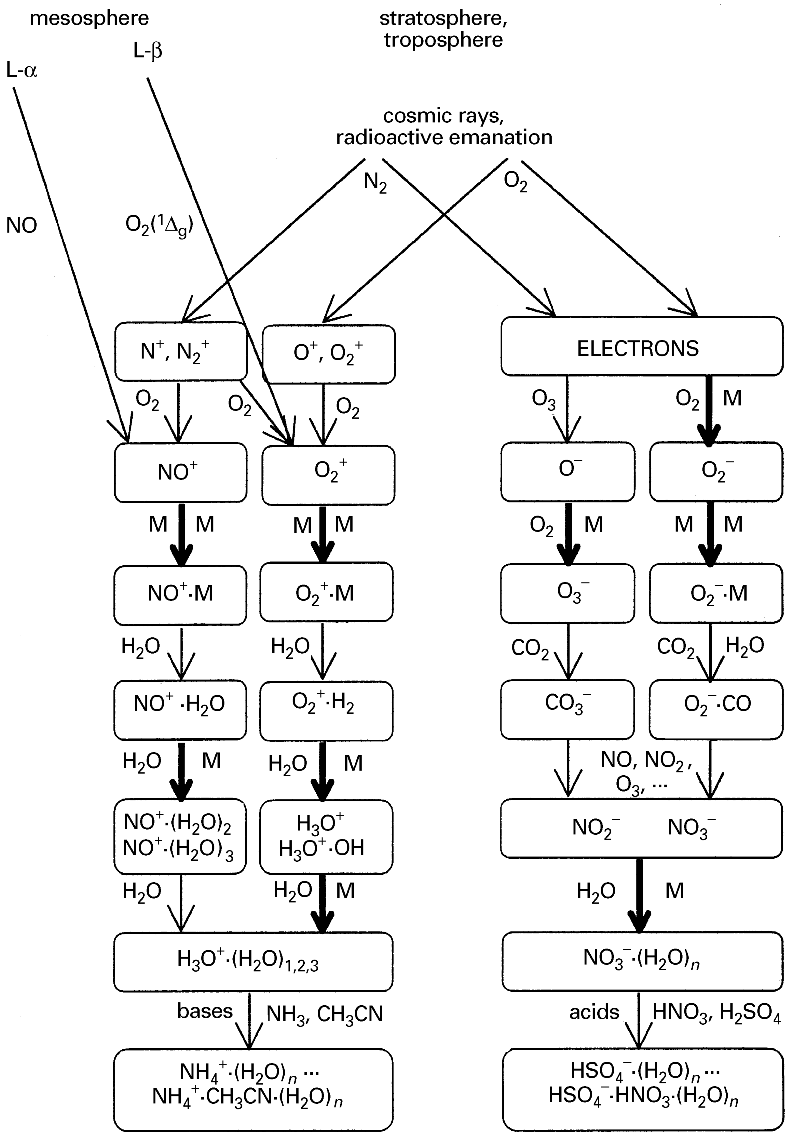



**Figure 5** The ion chemistry of the lower atmosphere. The positive ion chemistry (left column) and the negative ion chemistry (right column) are dominated by termolecular reactions, finally producing the cluster ions in the lower boxes.

species such as $NO_x$ molecules result in the very stable negative ion $NO_3^-$. This ion undergoes ter-molecular association reactions with $H_2O$ producing $NO_3^- \cdot (H_2O)_n$ cluster ions which dominate the ionospheric D-region. At even lower altitudes the production of $NO_3^-$ and its hydrates is followed by their reactions with acids, mostly $HNO_3$ and $H_2SO_4$ (produced from pollutant $NO_x$ and $SO_2$ in the parallel neutral chemistry that is occurring in the lower TA). These reactions produce cluster ions like $NO_3^-(HNO_3)_{2,3}$ and $HSO_4^-(HNO_3)_n$. However, the large ambient concentration of $H_2O$ shifts the equilibrium to the right in the generalized reaction:

$$\begin{aligned} &HSO_4^- \ (H_2O)_{m-1}(\text{acid})_n + H_2O \rightleftharpoons \\ &HSO_4^- \ (H_2O)_m(\text{acid})_{n-1} + \text{acid} \end{aligned} \qquad [9]$$

producing the mixed cluster ions indicated. The SIFT has been of great value in the study of these and many other reactions of cluster ions.

### Interstellar ion chemistry

One of the most interesting events in astronomy during the last few decades has been the discovery of many types of molecules, both ionized and neutral, in the diffuse and dense ISC that pervade the Milky Way galaxy. These ISC are at such enormous distances from the solar system that *in situ* probes, e.g. conventional mass spectrometers, cannot be exploited to analyse their composition. The only tool available to probe their physical conditions and chemical compositions is spectroscopy over the whole spectral range (from radio waves to gamma rays). The diffuse clouds of very low number density (about $10^2$ $cm^{-3}$) consist mainly of H and $H_2$ together with C, N and O atoms which play a central role in the gas phase ion chemistry. The dense clouds of much higher number density ($10^4$–$10^3$ $cm^{-3}$) consist mainly of $H_2$ and He together with minority C, N and O atoms.

In diffuse ISC it is stellar ultraviolet and galactic cosmic rays that largely create the ions, the important initial ions being $C^+$, $H^+$, $H_2^{\bullet+}$ and $He^{\bullet+}$. The dense ISC through which ultraviolet cannot penetrate because of the presence of micrometre-sized 'dust' grains, are ionized by galactic cosmic rays producing $H^+$, and $H_2^{\bullet+}$ and $He^{\bullet+}$. From these initial positive ions begins the gas phase ion chemistries illustrated in **Figures 6** and **7**. The SIFT is especially valuable for the study of interstellar ion chemistry and (together with the ion cyclotron resonance technique) has provided the major contribution to this field.

The primary $H^+$ ions can react with O atoms by 'accidental resonance' charge transfer producing $O^{\bullet+}$ ions, which react rapidly with $H_2$ producing $OH^+$. The $H_2^{\bullet+}$ primary ions react rapidly with $H_2$ producing $H_3^+$, which transfers a proton to O atoms also producing $OH^+$

$$H_3^+ + O \rightarrow OH^+ + H_2 \qquad [10]$$

The $OH^+$ ions react with $H_2$ to give $H_2O^{\bullet+}$, which then reacts with $H_2$ to give $H_3O^+$ in the sequence of H-atom abstraction reactions indicated in **Figure 6**. The closed-shell ion $H_3O^+$ ends this chain, since it does not react with $H_2$, but it does undergo dissociative recombination with electrons, producing OH and $H_2O$ molecules. A similar sequence of H-atom abstraction reactions probably leads to the production of NH, $NH_2$ and $NH_3$ in ISC as is indicated in **Figure 6**. This sequence begins with $N^+$ ions produced from $He^{\bullet+}$ (via

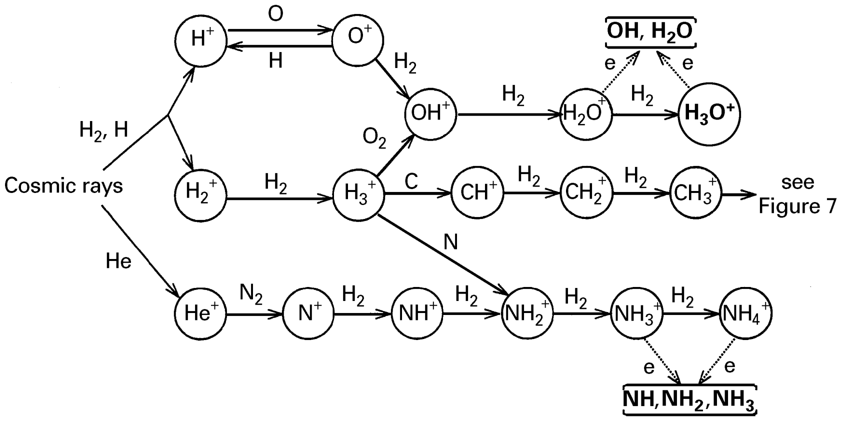



**Figure 6** The initial steps in the ion chemistry of interstellar clouds leading to $H_2O$, $NH_3$ and $CH_4$ (see also **Figure 7**).

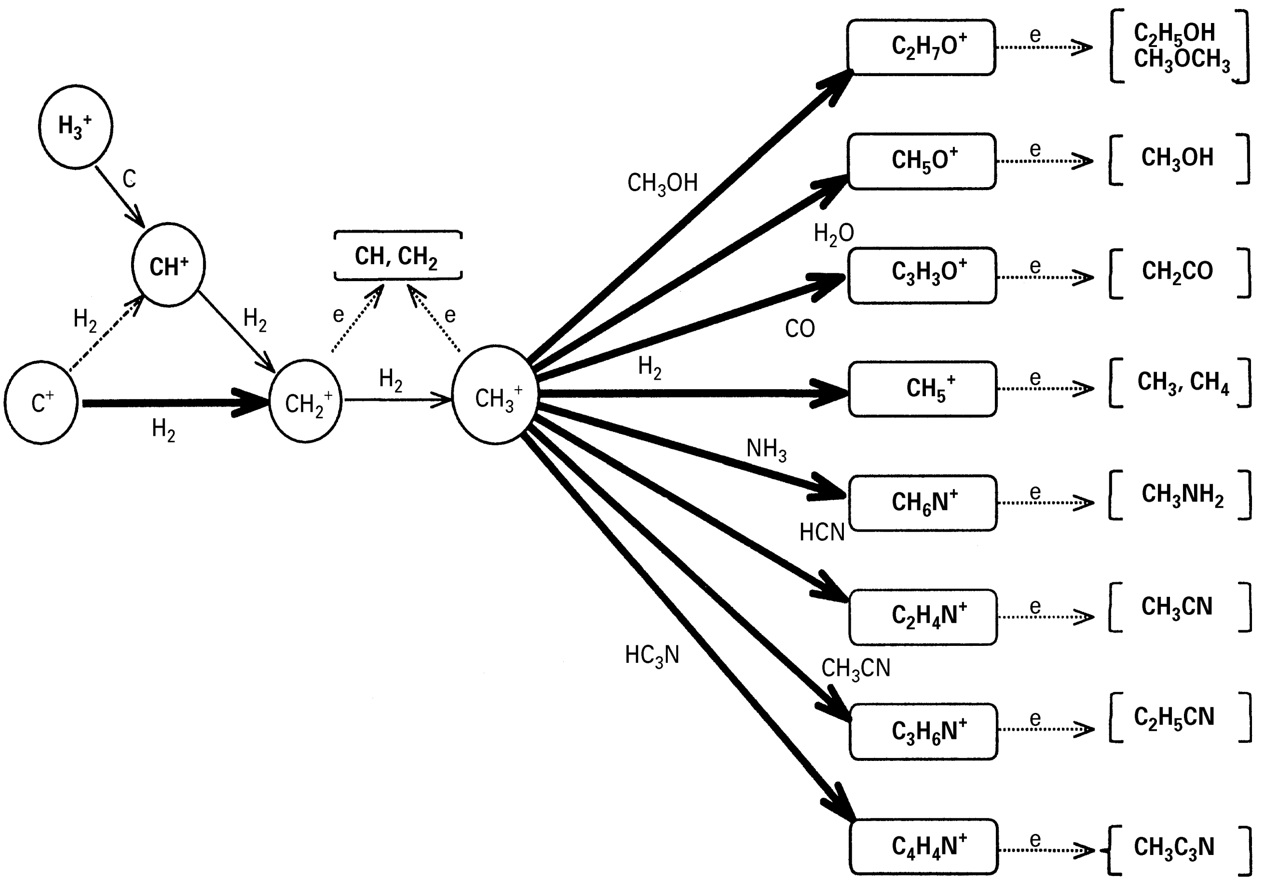



**Figure 7** The production of polyatomic ions and neutral molecules in dense ISC following the radiative association reactions (thick arrows) of $CH_3^+$ with several known interstellar molecules and the subsequent dissociative recombination.

Eqn [2]). The reaction of $H_3^+$ with N atoms is also involved in $NH_3$ production. Similar sequences lead to small hydrocarbon molecules. These sequences of reactions can readily be studied using the selected ion injection facility of the SIFT.

The ion chemistry that results in the polyatomic hydrocarbon molecules observed in dense ISC can begin with slow radiative association reaction (see below) of $C^+$ with $H_2$, producing $CH_2^+$ ions which then react rapidly with $H_2$ to form $CH_3^+$. The last ion can also be produced by the sequence of reactions beginning with the proton transfer reaction of $H_3^+$ with C atoms producing $CH^+$ ions, whence:

$$\begin{aligned} CH^+ + H_2 &\rightarrow CH_2^+ + H \\ CH_2^+ + H_2 &\rightarrow CH_3^+ + H \end{aligned} \qquad [11]$$

SIFT studies have shown that these bimolecular reactions are very fast, but that $CH_3^+$ reacts only slowly with $H_2$ in a termolecular association reaction producing $CH_5^+$ thus:

$$CH_3^+ + H_2 + M \rightarrow CH_5^+ + M \qquad [12]$$

Such termolecular reactions proceed via the formation of loosely bound excited ion, e.g.

$$CH_3^+ + H_2 \rightleftharpoons (CH_5^+)^* \qquad [13]$$

which can be stabilized against decomposition at sufficiently high pressures in collisions with an inert third body M, leaving M internally or kinetically excited thus:

$$(CH_5^+)^* + M \rightarrow CH_5^+ + M^* \qquad [14]$$

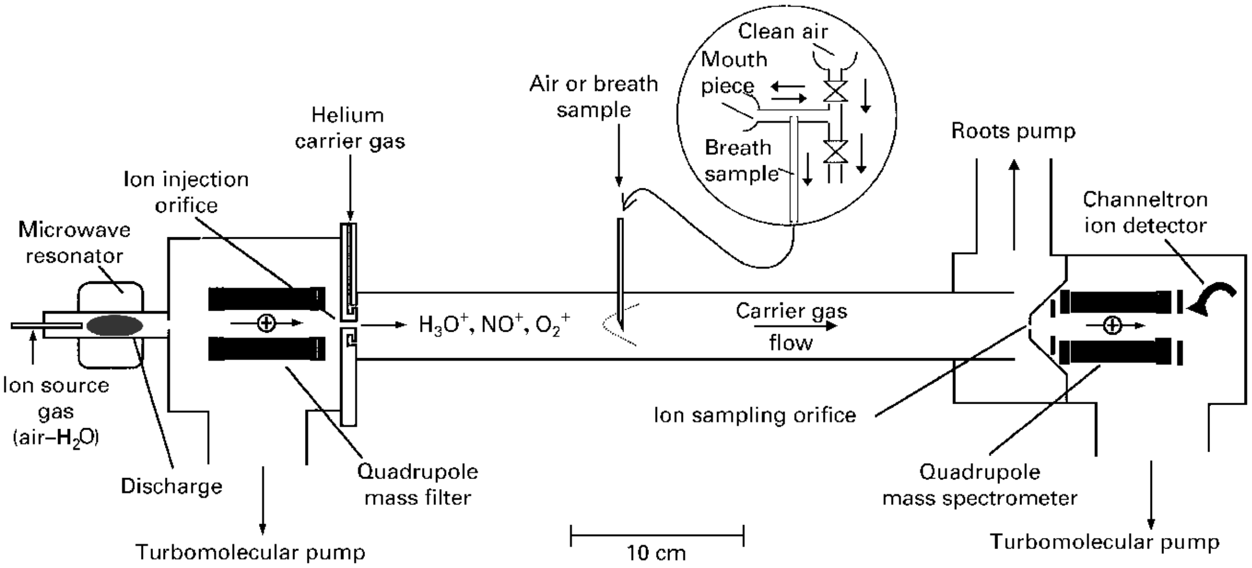



**Figure 8** A SIFT apparatus configured for trace gas analysis, with the stable discharge ion-source for $H_3O^+$, $NO^+$ and $O_2^+$ ions, and a single air/breath sample inlet port.

The kinetics of such termolecular reactions are well understood, and many have been studied using the SIFT. However, they cannot occur in ISC because the pressures in these regions are very low. Instead, the analogous process of radiative association occurs. Continuing with the above example, the bimolecular reactions [13] can occur in ISC, and indeed it is promoted by the low ambient temperatures (as low as 10 K in some clouds), and if the dissociation lifetime of the excited $(CH_5^+)^*$ ion is long enough, then it may emit a photon which will stabilize it against dissociation:

$$(CH_5^+)^* \longrightarrow CH_5^+ + h\nu \qquad [15]$$

Radiative association reactions are very important in ISC chemistry and lead to the production of many of the observed polyatomic molecules in these regions. The $CH_5^+$ ions formed in reaction [15] are neutralized by electrons in ISC to form $CH_4$ (see **Figure** 7).

SIFT studies have shown that $CH_3^+$ ions readily undergo termolecular association reactions with many known interstellar molecular species, including CO, $H_2O$, HCN, $NH_3$, $CH_3OH$ and $CH_3CN$. These ion–molecule associations must surely proceed via radiative association in dense ISC, and in this way complex molecules can be formed, as is indicated in **Figure** 7. Thus these SIFT studies have been crucial in indicating the importance of radiative association reactions in ISC.

Many of the molecular species detected in dense interstellar clouds are seen to contain the rare (heavy) isotopes of some elements (e.g. D, $^{13}C$, $^{18}O$, etc.). However, at first sight the abundance ratios of some molecules containing the rare and common isotopes (e.g. DCN/HCN) were very surprising because they are orders-of-magnitude greater than those expected from their cosmic isotopic ratios. Following extensive SIFT studies, it is now understood that this is due to the phenomenon of 'isotope fractionation' in gas phase ion–molecule reactions. This phenomenon is exemplified by the elementary reaction:

$$D^+ + H_2 \rightleftharpoons H^+ + HD \qquad [16]$$

It is a simple matter to determine the forward (exothermic) and reverse (endothermic) rate coefficients, $k_f$ and $k_r$ for such reactions using the SIFT technique (ideally suited for such studies because separate isotopic ions can be injected and the reaction temperature can be accurately controlled even down to 80 K). For reactions [16] it is observed that $k_f$ increasingly exceeds $k_r$ as the temperature is reduced. This is because the reverse reaction is endothermic by 39.8 meV (3.84 kJ mol$^{-1}$), by virtue of the zero-point-energy difference between $H_2$ and HD, and the ionization energies of H and D differ by 4 meV (0.38 kJ mol$^{-1}$). Hence in cold ISC the reverse reaction is effectively stopped whilst the forward reaction proceeds at the gas kinetic rate. This effectively ensures that much of the deuterium in dense interstellar

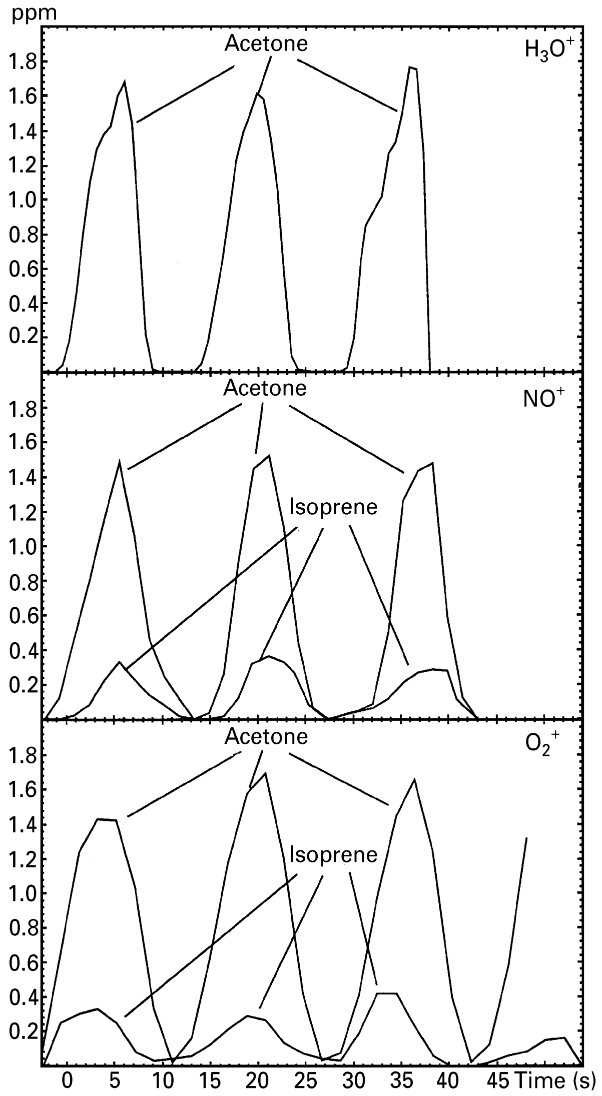



**Figure 9** Concentration time profiles of the trace gases acetone and isoprene in single exhalations of breath obtained with the SIFT using $H_3O^+$, $NO^+$ and $O_2^+$ precursor ions. Concentrations are given in parts per million (ppm) of the breath.

clouds is contained in HD. Also important in ISC are the reactions :

$$H_3^+ + HD \rightleftharpoons H_2D^+ + H_2$$
$$CH_3^+ + HD \rightleftharpoons CH_2D^+ + H_2 \qquad [17]$$

in which D is fractionated into $H_2D^+$ and $CH_2D^+$. Thus the subsequent reactions of these deuterated ions (occurring, of course, in parallel with the $H_3^+$ and $CH_3^+$ reactions) result in the enrichment of deuterium in many interstellar molecules. Similarly, SIFT studies have shown that fractionation of the rare isotopes of heavier elements also occurs, notably of $^{13}C$ into the very abundant interstellar molecule CO via the reaction of $^{13}C^+$ with $^{12}CO$.

## Trace gas analysis; SIFT/MS

The SIFT can be used for accurate determinations of the concentrations $n_M$ of trace gases, M, in an air sample that has been introduced into the carrier gas, if the rate coefficients $k$ are known for the reactions of a chosen injected precursor ion species with the M. Measurements of the count rates at the downstream mass spectrometer system of the precursor ion and each product ion, $I_1$ and $I_p$ respectively, provide values of $n_M$ for each trace gas if $I_p << I_1$. Then following Equation [1]:

$$n_M = \frac{I_p}{I_1 kt} \qquad [18]$$

Since currently $I_1$ may be typically $10^5$ precursor ions per second, a typical $k$ is $10^{-9}$ cm$^3$ s$^{-1}$ and the reaction time $t$ is of the order $10^{-2}$ s. Then for $I_p$ at the low value of one product ion per second (a sensible detection limit), $n_M$ in the carrier gas is $10^6$ cm$^{-3}$. This is a fractional number density of $10^{-10}$ of that of the helium carrier atoms (usually $10^{16}$ cm$^{-3}$). Thus for an air or breath sample introduced into the carrier gas at a relative concentration of 1% of the carrier gas, trace gases at a partial pressure of 10 parts per billion (ppb) can be detected and quantified. Clearly, for large $I_1$, greater air sample concentrations, and longer integration times for $I_p$, the sensitivity can be improved and the detection limit lowered. The experimental configuration for this SIFT analytical method is indicated in **Figure 8**. A transportable version of the SIFT is now available commercially for use *in situ* for breath and environmental analyses.

This analytical method can be easily used for the detection and accurate quantification of several trace gases simultaneously in multicomponent vapour mixtures such as polluted air and human breath. This can be achieved in real time because of the short time response of the method ($t \sim 10^{-2}$ s), by rapidly switching the downstream mass spectrometer between chosen precursor and product ion masses. Thus the time profiles of the concentration of several trace gases in breath can be obtained from single exhalations, as are shown in **Figure 9**. Samples introduced into the apparatus directly from the atmosphere or from containers can be analysed by multiscanning the downstream mass spectrometer. A spectrum obtained by sampling the SIFT laboratory air is shown in **Figure 10**.

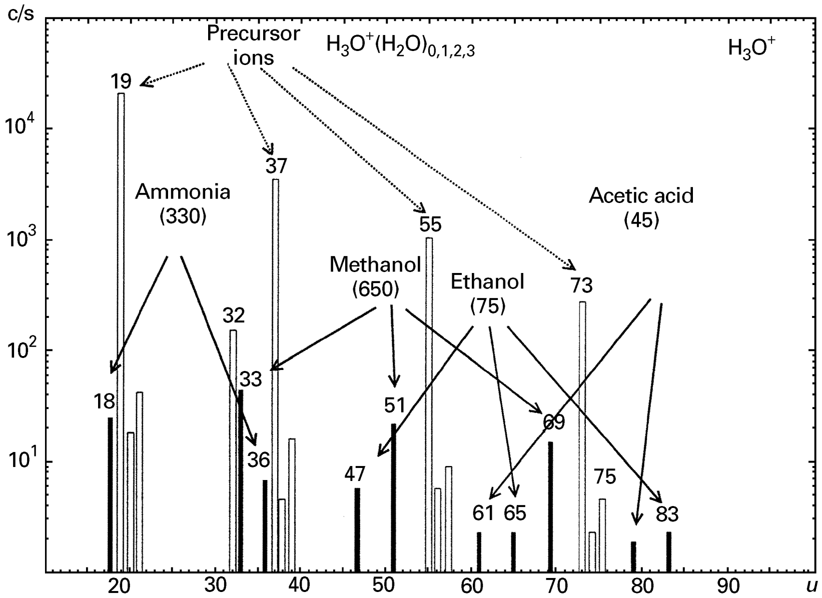



**Figure 10** A mass spectrum obtained using $H_3O^+$, precursor ions following the introduction of laboratory air into the SIFT. Concentrations of the detected species (released from an adjacent laboratory) are given in parts per billion (ppb) in parentheses. *u* = ion mass-to-charge ratio, c/s = counts per second.

A vital point to note is that the mass spectrum of the product ions in much simpler than would be obtained using electron impact ionization of the air or breath sample, because chemical ionization is used. However, only a limited number of precursor ions can be used, which must not react at a significant rate with the major components of air or breath, $N_2$, $O_2$, $CO_2$, $H_2O$ and Ar, but obviously must react efficiently with the trace gases in the sample. In this respect $H_3O^+$, $NO^+$ and $O_2^{\bullet+}$ are the prime candidates, and SIFT studies of a great number of the reactions of these three ions with a wide variety of organic and inorganic compounds have shown that $H_3O^+$ and $NO^+$ are of widest application, with $O_2^{\bullet+}$ being useful for fewer particular trace gases.

Recent SIFT studies have shown that $H_3O^+$ transfers a proton at the gas kinetic (maximum efficiency) rate to all molecules, M, that have proton affinities greater than $H_2O$ molecules, and usually only a single product ion $MH^+$ is formed, which greatly simplifies the analysis of the product ion mass spectrum. However, two products are sometimes observed from these proton transfer reactions, as is the case for some aldehyde reactions, e.g.

$$H_3O^+ + C_4H_9CHO \rightarrow C_4H_9CHOH^+(+H_2O) \quad C_5H_9^+(+2H_2O) \qquad [19]$$

$NO^+$ ions often react with organic molecules via hydride ion ($H^-$) transfer producing $(M{-}H)^+$ ions and HNO molecules e.g.

$$NO^+ + CH_3CHO \rightarrow CH_3CO^+ + HNO \qquad [20]$$

They react with most carboxylic acids and with tertiary alcohols via hydroxide ion ($OH^-$) transfer producing $(M{-}OH)^+$ ions and $HNO_2$ molecules, and undergo rapid association reactions with most ketones producing $NO^+{\cdot}M$ ions.

$O_2^{\bullet+}$ is particularly useful as a precursor ion for species of low proton affinity that do not react with $H_3O^+$ and of high ionization potential that do not react with $NO^+$. The reaction process is invariably charge transfer; it is particularly useful for quantifying NO, $NO_2$ and $NH_3$ in air:

$$\begin{aligned} O_2^{\bullet+} + NO &\rightarrow NO^+ + O_2 \\ O_2^{\bullet+} + NO_2 &\rightarrow NO_2^+ + O_2 \\ O_2^{\bullet+} + NH_3 &\rightarrow NH_3^{\bullet+} + O_2 \end{aligned} \qquad [21]$$

The beauty of this SIFT/MS analytical technique is that all three suitable precursor ion species can be used on the same air or breath sample by switching

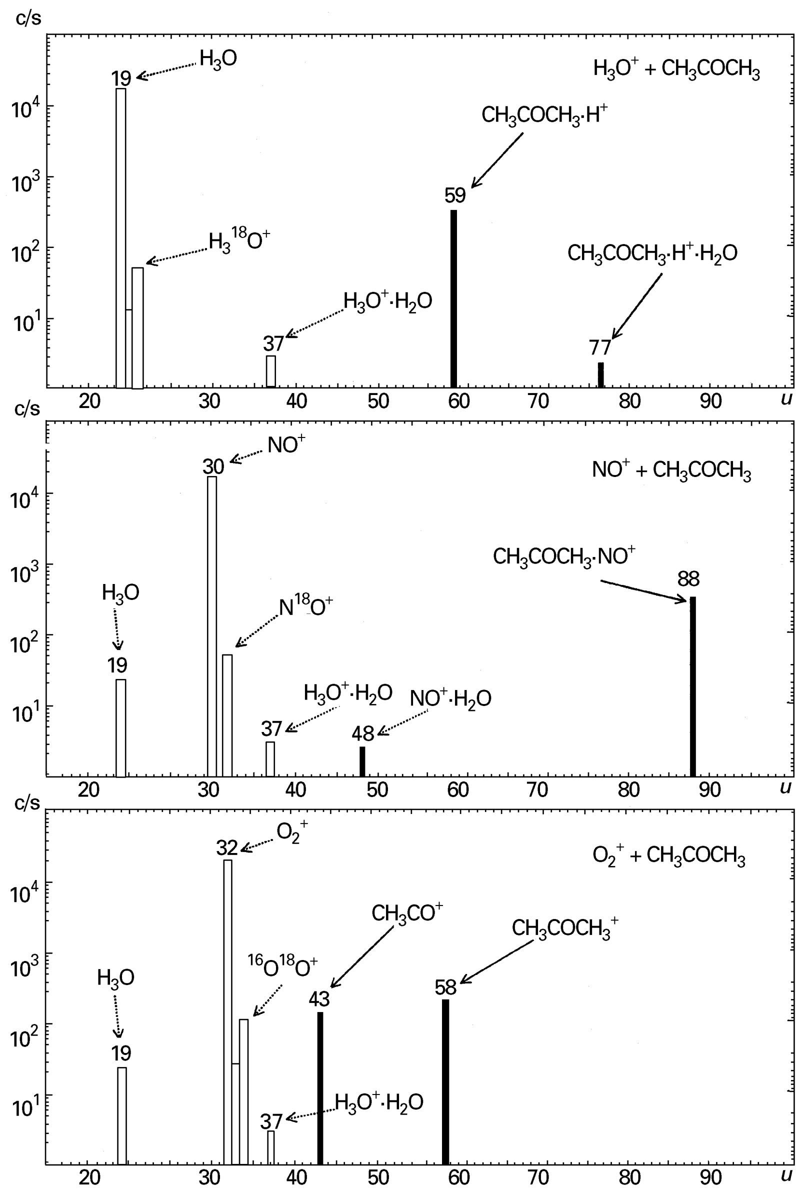



**Figure 11** The SIFT analyses of an air–acetone mixture using $H_3O^+$, $NO^+$ and $O_2^+$ precursor ions injected sequentially into the carrier gas by switching the upstream mass filter (see **Figure 8**).

the upstream mass filter (see **Figure 8**). **Figure 11** shows the results obtained simply for an air–acetone sample. The use of two or three precursor ions ensures that few of the various trace gases in the sample are missed.

An extensive database is required of the rate coefficients and product ions of the reactions of $H_3O^+$, $NO^+$ and $O_2^{\bullet+}$ with many different types of molecules, and this is rapidly being compiled from SIFT experiments. With sufficient knowledge of the ion chemistry, very complex mass spectra can be analysed, such as those shown in **Figure 12**, obtained for a sample of the volatiles from gasoline using both $H_3O^+$ and $O_2^{\bullet+}$ precursor ions. Note the identification of the aromatic and aliphatic hydrocarbons, each compound of mass $M$, being recognized as $MH^+$ ions in the $H_3O^+$ spectrum (proton transfer) and $M^+$ ions in the $O_2^{\bullet+}$ spectrum (charge transfer).

The potential of this SIFT/MS analytical method is enormous for human breath analysis and hence in physiology and biochemistry, for clinical diagnosis and therapeutic monitoring, and for measuring air

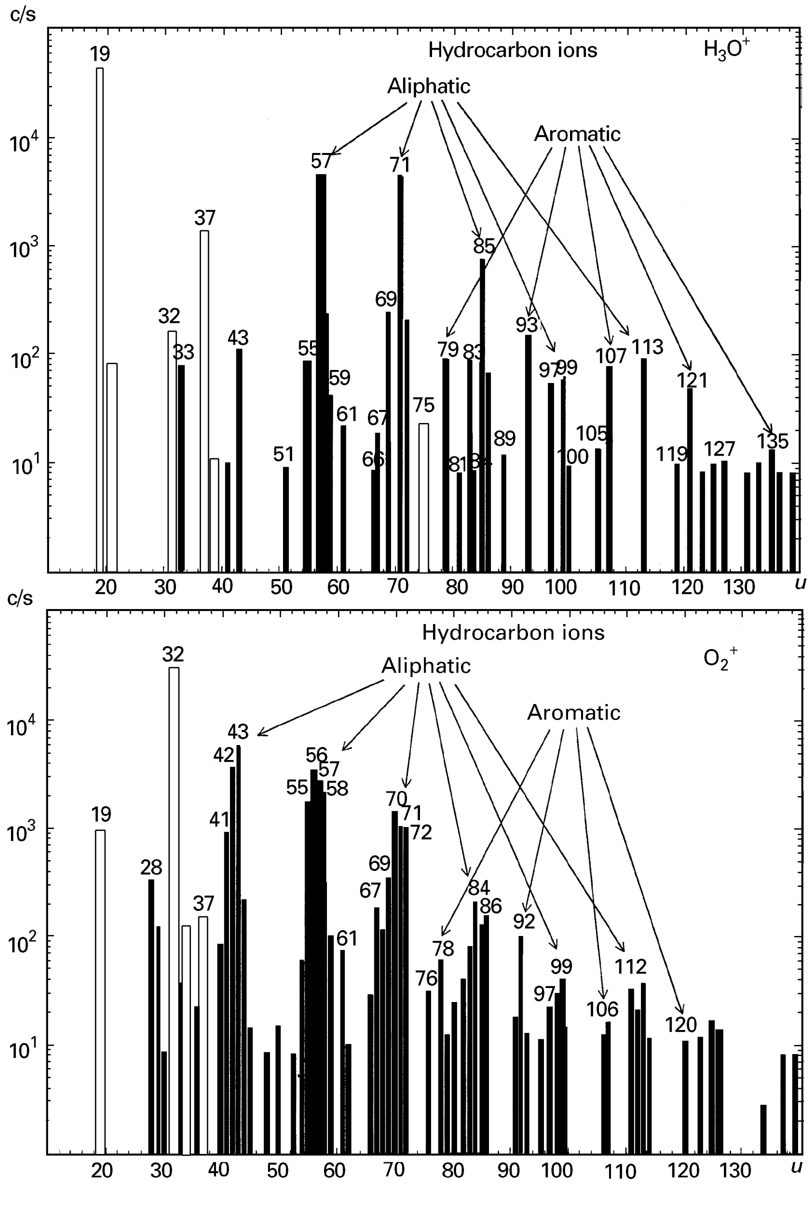



**Figure 12** Mass spectra obtained in the SIFT for an air–gasoline vapour mixture using $H_3O^+$ and $O_2^+$ precursor ions. The component hydrocarbon molecules M, are protonated by $H_3O^+$ producing $MH^+$ ions, whereas charge transfer between $O_2^+$ and the M produces ions like $M^+$ and $(M–H)^+$. *u* = ion mass-to-change ratio, c/s = counts per second.

pollution and analysing food vapour emissions and food flavours. The medical and biochemical value is well illustrated by the analyses of the breath of six healthy individuals following the morning ingestion of a liquid protein meal. The breath concentrations of five species (ammonia, methanol, ethanol, acetone and isoprene) were obtained from only single breath exhalations before and some six hours after the meal. Shown in **Figure 13** are the rise in the ammonia levels as the protein is metabolized and the decrease in the acetone as the body is nourished.

SIFT/MS analyses of the breath of some 30 uraemic patients with end-stage renal failure have shown greatly elevated levels of ammonia compared to healthy subjects. The spectrum obtained of breath of a uraemic patient with diabetes (indicated by the elevated acetone level) who also smokes cigarettes (indicated by the presence of acetonitrile) is shown in **Figure 14**. Many further applications of this new method for trace gas analysis are in train including applications in agriculture (animal welfare) and grassland research.

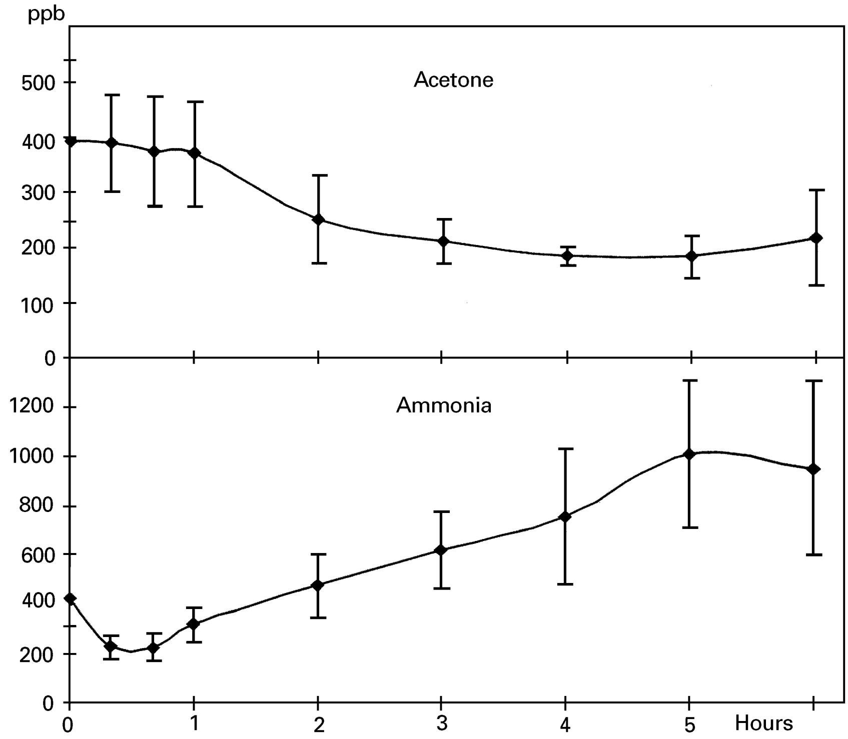



**Figure 13** The mean concentrations of acetone and ammonia (ppb) obtained by the SIFT in single exhalations of breath from six volunteers following a protein meal taken at time 0 (see the text). The vertical bars represent the standard deviations of the six concentrations at each time.

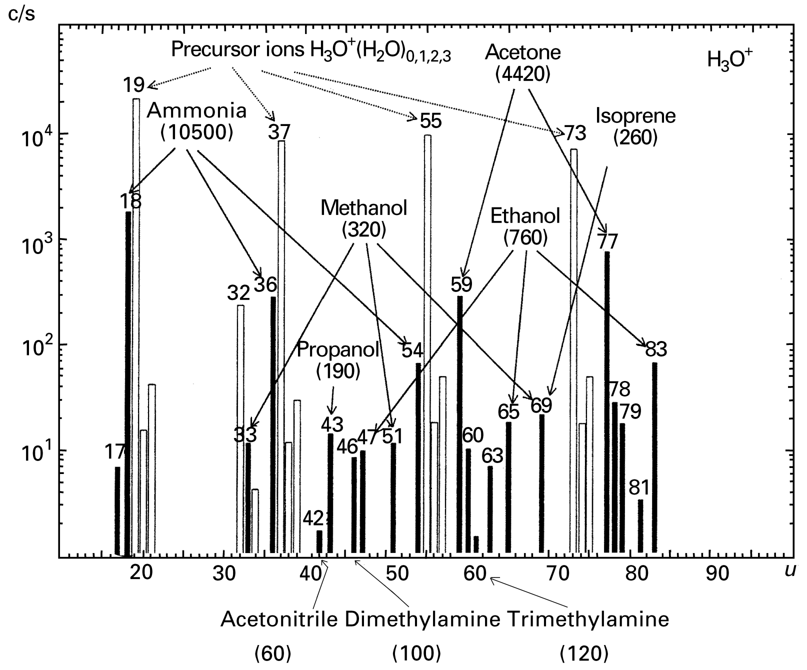



**Figure 14** The SIFT spectrum obtained following the introduction of breath from a diabetic patient with end-stage renal failure who smokes cigarettes (see the text). The concentrations of the major breath gases are given in ppb in parentheses. *u* = mass. c/s = counts per second.

## List of symbols

$h$ = Plank constant; $I$ = count rate; $I_1$ = count rate of precursor ion; $I_p$ = count rate of product ion; $k$ = rate coefficient; $k_f$ = forward rate coefficient; $k_r$ = reverse rate coefficient; $n_M$ = concentration of trace gas M; $n_1$ = number density of reactant charged particles; $n_2$ = number density of buffer gas; $n_3$ = number density of reactant neutral particles; $T_g$ = buffer gas temperature; $t$ = reaction time; $u$ = ion mass-to-charge ratio; $\nu$ = frequency.

*See also:* **Biochemical Applications of Mass Spectrometry; Chemical Ionization in Mass Spectrometry; Cluster Ions Measured Using Mass Spectrometry; Food Science, Applications of Mass Spectrometry; Ion Collision, Theory; Ion Molecule Reactions in Mass Spectrometry; Isotope Ratio Studies Using Mass Spectrometry; Isotopic Labelling in Mass Spectrometry; NQR, Applications; Proton Affinities; Quadrupoles, Use of in Mass Spectrometry.**

## Further reading

Anicich VG (1993) Evaluated bimolecular ion–molecule gas phase kinetics of positive ions for use in modeling planetary atmospheres, cometary comae, and interstellar clouds. *Journal of Physical Chemistry and Reference Data* **22**: 1469–1569.

Duley WW and Williams DA (1984) *Interstellar Chemistry*. London: Academic Press.

Farrar JM and Saunders WH Jr (eds) (1988), *Techniques for the Study of Gas-Phase Ion–Molecule Reactions*, New York: Wiley.

Ferguson EE, Fehsenfeld FC and Albritton DL (1979) Ion chemistry of the Earth's atmosphere. In: Bowers MT, (ed) *Gas Phase Ion Chemistry*, pp 45–81. New York: Academic Press.

Fontijn A and Clyne MAA (eds) (1983) *Reactions of Small Transient Species*. London: Academic Press.

Herbst E and Leung CM (1989) Gas phase production of complex hydrocarbons, cyanopolynes, and related compounds in dense interstellar clouds. *Astrophysical Journal, Supplement Series* **69**: 271–300.

Smith D (1992) The ion chemistry of interstellar clouds. *Chemical Reviews* **92**: 1473–1485.

Smith D and Adams NG (1987) The selected ion flow tube (SIFT): studies of ion–neutral reactions. *Advances in Atomic and Molecular Physics* **24**: 1–49.

Smith D and Spanel P (1995) Swarm techniques. In: Dunning FB and Hulet RG (eds) *Experimental Methods in the Physical Sciences: Atomic, Molecular, and Optical Physics: Charged Particles*, pp 273–298. New York: Academic Press.

Smith D and Spanel P (1996) Ions in the terrestrial atmosphere and in interstellar clouds. *Mass Spectrometry Reviews* **14**: 255–278.

Španěl P, Cocker J, Rajan B and Smith D (1997) Validation of the SIFT technique for trace gas analysis of breath using the syringe injection technique. *Annals of Occupational Hygiene* **41**: 373–382.

Španěl P and Smith D (1996) Application of ion chemistry and the SIFT technique to the quantitative analysis of trace gases in air and on breath. *International Review of Physical Chemistry* **15**: 231–271.

Španěl P and Smith D (1996) Selected ion flow tube: a technique for quantitative trace gas analysis of air and breath. *Medical and Biological Engineering and Computing* **34**: 409–419.

Španěl P and Smith D (1997) SIFT studies of the reactions of $H_3O^+$, $NO^+$ and $O_2^+$ with a series of alcohols. *International Journal of Mass Spectrometry and Ion Processes* **167/168**: 375–388.

Španěl P and Smith D (1999) Selected ion flow tube studies of the reactions of $H_3O^+$, $NO^+$ and $O_2^+$ with some chloroalkanes and chloroalkenes. *International Journal of Mass Spectrometry* **184**: 175–181.

Španěl P and Smith D (1999) Selected ion flow tube-mass spectrometry: Detection and real-time monitoring of flavours released by food products. *Rapid Communications in Mass Spectrometry* **13**: 585–596.

Wayne RP (1985), *Chemistry of Atmospheres*, Oxford: Clarendon.

# Silver NMR, Applications

*See* **Heteronuclear NMR Applications (Y–Cd).**

# Single Photon Emission Computed Tomography

*See* **SPECT, Methods and Instrumentation.**

# Small Molecule Applications of X-Ray Diffraction

**Andrei S Batsanov,** University of Durham, UK

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| Applications |

Single crystal X-ray diffraction is the main source of information on the geometrical structure of molecules and molecular solids, including bond distances (and hence bond orders), bond angles, shapes of coordination polyhedra, conformations of flexible molecules, as well as intermolecular contacts. It can always distinguish between configurational isomers (e.g. *cis* or *trans*), and often optical isomers (enantiomers), too. The enormous and fast-growing mass of X-ray structural data defies any attempt to review it even in the most general way. Numerous compen-diums and reference tables on these data have been published. In fact, X-ray crystallography provides the main bulk of data for any book on structural chemistry and the theory of the chemical bond. More recently, computerized structural databases have given easy access to all available X-ray crystal structures and sophisticated means of extracting from them necessary parameters, performing statistical analysis of the data and finding regularities (structural correlations). On a higher level of precision, it is possible now to determine not only the atomic positions, but also the entire map of electron density in a crystal, visualizing concentrations of bonding electrons and lone electron pairs, determining directly atomic charges, electrostatic potentials, etc. The analysis of the displacement parameters of atoms is also of much importance, both to improve the accuracy of molecular geometry determinations and to understand the dynamic behaviour of crystal structures (**Figure 1**).

## Atomic structure

A crystal structure consists of a periodic pattern of atomic nuclei and a continuous (also periodic) distribution of electron density. As X-rays are scattered by electrons, it is possible to extract structure factors from the reflection intensities, and, by their Fourier transform, to calculate the electron density map, the peaks of which correspond to the centres of atoms. We can then approximate the structure by placing at these points isolated atoms of corresponding elements with ideal spherical symmetry. The positions (coordinates) of these atoms and their displacement parameters (see Atomic displacements below), can be refined by the least-squares technique so as to achieve the best agreement with the observed intensities. Usually, the structure determination ends at this stage. Atomic coordinates are referred to the coordinate axes parallel to the edges of the unit cell and published in fractional form, i.e. as $x/a$, $y/b$ and $z/c$, where $a$, $b$ and $c$ are the unit cell parameters. All kinds of geometrical parameters of the molecule in a crystal can be calculated from these coordinates. Shortest vectors between centres of atoms give bond lengths and, of course, the atomic connectivity, i.e. the order in which the atoms are linked. Usually, the bond order (multiplicity) can be estimated roughly from a bond length in a straightforward way. Angles between bonds indicate the state of hybridization of a given atom. The conformation of a sterically flexible molecule can be described in terms of dihedral angles between planar fragments, or torsion angles. For ring systems, an unequivocal way of describing the conformation is provided by the ring puckering coordinates.

X-ray diffraction provided the first determination ever (in 1912) of the covalent bond length (C–C 1.544 Å in diamond) and is to the present day the overwhelming source of information on 'molecular metrics'. The number of organic and organometallic crystal structures, deposited at the Cambridge Structural Database (CSD) by April 1998, exceeds 182 000 (cf. 67 000 in 1988). Of these, less than 1% were determined by neutron diffraction, all the rest by X-ray diffraction. All other spectroscopic methods that can give exact interatomic distances (gas-phase electron diffraction, microwave spectroscopy, etc.) yielded only hundreds of structure determinations. Furthermore, X-ray diffraction is unique among spectroscopic methods in that the number of experimental data (reflection intensities) exceeds the number of unknown variables (for each atom, three coordinates plus one isotopic or six anisotropic displacement parameters) by an order of magnitude. Currently, the reflections-to-variables ratio of 8 is regarded as the minimum acceptable for a good quality experiment. Therefore in determining the structural formula of a molecule and, to moderate precision, its geometry, the X-ray method needs no preconceived model of the structure, nor prior knowledge of the molecular symmetry. The number of possible reflections being proportional to the unit cell volume, and the latter to the

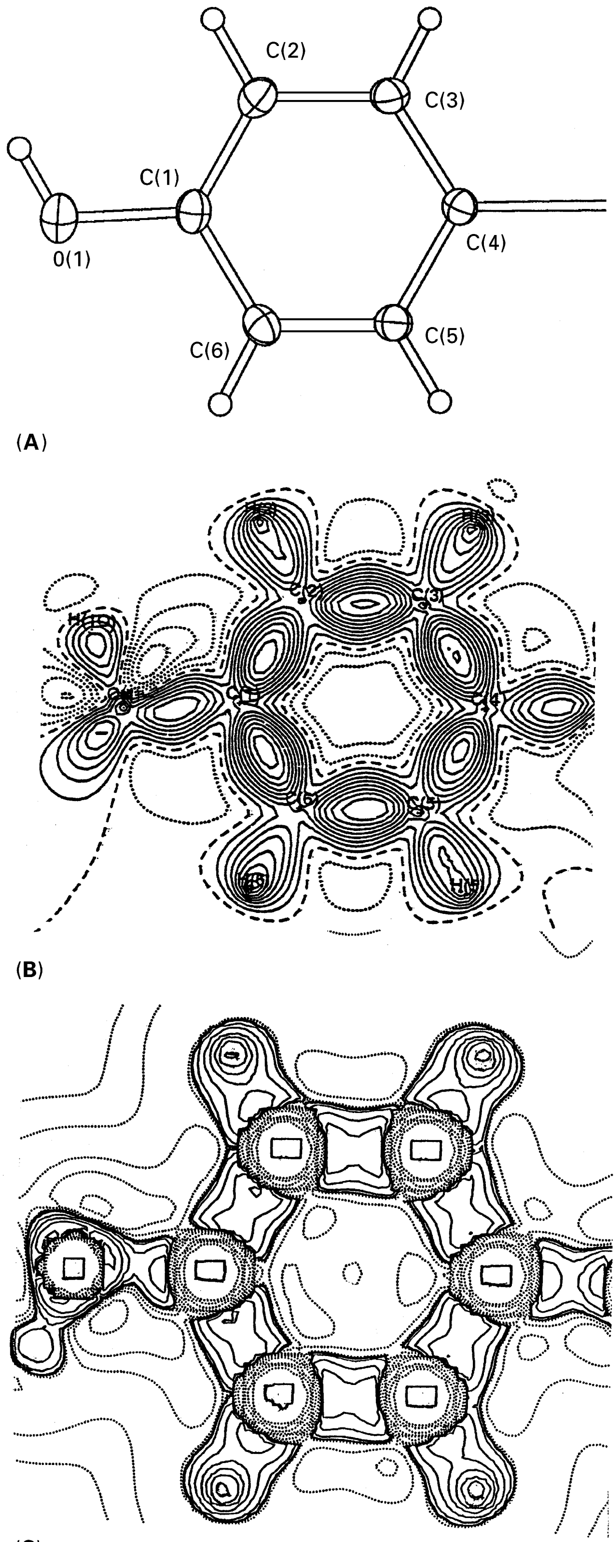



**Figure 1** (A) Molecular structure of the hydroxyphenyl group in the crystal of 4-(methylamido)phenol, showing 50% displacement ellipsoids; maps of the deformation electron density (B) and Laplacian (C) in the same moiety (positive contours solid, negative dashed). Unpublished results of Yufit DS and Muir K, courtesy of the authors.

molecular volume, this favourable situation persists up to molecules of considerable size (200 nonhydrogen atoms and even more).

This makes the X-ray method a very cost-effective tool for identification of newly synthesized compounds or newly isolated natural products: one needs a single crystal of ~0.01 mg to learn the composition and structural formula, usually without destroying the sample itself. In the past, solution of the phase problem was the major difficulty, the only available means (besides trial and error) being the Patterson method, effective only for structures containing at least one heavy atom. The advent of efficient direct methods of phase determination in the 1970s and the enhancement of these methods by phase relations of higher orders in the 1980s (which eliminated the problems for space groups without glide planes and screw axes) reduced the solution of a wholly unknown structure from art to near routine, if a crystal of good quality can be obtained.

A primary (but not unfallable) estimate of the quality of structure determination is R-factor, showing the discrepancy between the experimentally observed structure factors of reflections ($F_o$) and those calculated from the determined structure ($F_c$), $R = \Sigma \big| |F_o| - |F_c| \big| / \Sigma |F_o|$. With modern precision of the data measurement, 'quality' structures have R of 0.02 to 0.06 (for observed reflections). R-factor based on $F^2$ (rather than $F$) became popular recently, being more meaningful for weak reflections. For the same data, R($F^2$) is roughly twice the magnitude of R($F$).

### Bond distances and their precision

It was realized early on that bond lengths in molecular crystals depend mainly on the nature of the elements involved and on the bond order (the number of bonding electron pairs minus the number of antibonding ones), the crystal environment playing only a secondary role. Numerous tables and compendiums of bond length exist. **Tables 1** and **2** list some average ('standard') values, mainly from the most recent and comprehensive reviews by Kennard *et al.* based on the data from the CSD.

The measure of random errors in geometrical parameters, estimated standard deviations (esd), are routinely calculated in least-squares refinements. In good-quality studies of purely organic compounds, they may be as low as one or several thousandths of Å (for bonds not involving hydrogen atoms). Errors being in general in inverse relation to the atomic number, in organometallic structures the esd of the heavy atom coordinates may be as small as 0.0001 Å, but those of the lighter ligand atoms can be of the order of 0.01 Å or more, so the latter define

**Table 1** Selected average bond distances in organic compounds (in Å, $\sigma$ = 0.01–0.02 Å)

| | | | |
|---|---|---|---|
| $RH_2C–CH_2R$ | 1.524 | RHC=CHR | 1.316 |
| $R_2HC–CHR_2$ | 1.542 | $R_2C=CR_2$ | 1.331 |
| $R_3C–CR_3$ | 1.588 | C=C(=C), in allenes | 1.307 |
| $C(sp^2)–C(sp^2)$ in: | | C≡C | 1.181 |
| Biphenyls | 1.490 | C=O in ketones | 1.210 |
| Conjugated dienes | 1.455 | C–N peptide bond | 1.332 |
| C(sp)–C(sp) | 1.377 | C=N | 1.279 |
| C–C in phenyl rings | 1.380 | C≡N | 1.144 |
| $N(sp^3)–N(sp^3)$ | 1.454 | C=S | 1.671 |
| $N(sp^2)–N(sp^2)$ | 1.401 | $R_3P=O$ | 1.489 |
| N=N | 1.240 | $R_3P=S$ | 1.954 |
| RO–OR | 1.469 | Se–Se | 2.340 |
| RS–SR | 2.048 | Si–Si | 2.359 |

| *X* | *$C(sp^3)–X$* | *$C(sp^2)–X$* | *X* | *$C(sp^3)–X$* | *$C(sp^2)–X$* |
|---|---|---|---|---|---|
| F in $RCH_2F$ | 1.399 | 1.340 | $P^+R_3$ | 1.800 | 1.793 |
| Cl in $RCH_2Cl$ | 1.790 | 1.739 | $PR_2$ | 1.855 | 1.836 |
| Br | 1.966 | 1.899 | $P(=O)R_2$ | 1.813 | 1.801 |
| I | 2.162 | 2.095 | $SO_2R$ | 1.786 | 1.763 |
| OH | 1.432 | 1.362 | S(=O)R | 1.809 | 1.790 |
| $N^+(sp^3)$ | 1.499 | 1.465 | SR | 1.819 | 1.773 |
| $N(sp^3)$ | 1.469 | 1.416 | SeR | 1.970 | 1.893 |
| $N(sp^2)$ | 1.454 | 1.355 | $SiR_3$ | 1.863 | 1.868 |

the esd of the metal–ligand *distances*. Differences in bond lengths, not exceeding 3 esd, are regarded as statistically insignificant. However, esd represent only the internal consistency of the data and say nothing of the various systematic errors inherent in the experiment (absorption, extinction, radiation decay of crystal, etc.), or of the model, which has the following limitations.

1. The X-ray diffraction method measures the distances between 'centres-of-gravity' of atoms' electron clouds, which need not necessarily coincide with the atomic nuclei. The difference is most pronounced for H (or D), its sole electron being shifted into the bond region; the X-ray estimates of $C(sp^2)$–H, N–H and O–H bond lengths are 0.93, 0.89 and 0.82 Å, vs. 1.08, 1.01 and 0.97 Å determined by neutron diffraction in the same compounds. The effect is also considerable for sp-hybridized C and N atoms in acetylene and cyano groups.
2. Spherically symmetrical 'isolated' atoms model chemically bonded atoms, which are far from spherical. On refinement, atomic positions shift so as to compensate for the disregarded bonding electrons and/or lone pairs. Thus, a two-coordinate oxygen atom is shifted by 0.007–0.013 Å from its 'neutron' position (**Figure 2**) and C–O bond lengths are overestimated by 0.003–0.005 Å.
3. The X-ray method measures neither instantaneous nor time-average bond lengths, but the distances between mean atomic positions, averaged over all the crystal and all the duration of the experiment. This is much the biggest source of bias (see below).

**Table 2** Selected average metal–ligand distances in organometallic complexes (in Å, $\sigma = 0.03–0.04$ Å)

| | Ligands | | | | | |
|---|---|---|---|---|---|---|
| *Metal* | *CO(terminal)* | *C(aryl)* | *C(alkyl)* | *$C_5H_5$ (centroid of)* | *$PMe_3$* | *Cl* |
| V | 1.95 | 2.11 | | 1.95 | 2.51 | 2.29 |
| Cr | 1.87 | 2.08 | | 1.88 | 2.39 | 2.34 |
| Mn | 1.81 | 2.06 | 2.18 | 1.82 | 2.46 | 2.45 |
| Fe | 1.78 | 2.03 | 2.09 | 1.71 | 2.25 | 2.26 |
| Co | 1.78 | 1.93 | 2.01 | 1.70 | 2.22 | 2.27 |
| Ni | 1.77 | 1.92 | | 1.75 | 2.20 | 2.34 |
| Mo | 1.98 | 2.19 | 2.25 | 2.01 | 2.46 | 2.41 |
| Ru | 1.90 | 2.09 | 2.18 | 1.89 | 2.31 | 2.42 |
| Rh | 1.85 | 2.01 | 2.09 | 1.90 | 2.27 | 2.38 |
| W | 2.00 | | 2.19 | 2.01 | 2.49 | 2.41 |
| Re | 1.94 | | 2.17 | 1.96 | 2.37 | 2.39 |
| Pt | 1.85 | 2.05 | 2.08 | | 2.30 | 2.32 |

## Intermolecular contacts

Distances between atoms not bonded directly to each other, are rationalized in terms of the close packing model. An atom is regarded as a hard sphere of certain radius, van der Waals (vdW) radius, and a molecule as a superposition of such spheres. In a crystal, these bodies are closely packed but cannot dent each other. It is implied that a vdW radius is specific for a given element and remains the same in any contact, so that any A· · ·B contact distance equals the average of the A· · ·A and B· · ·B ones. Many systems of vdW radii have been developed. The pioneering one of Pauling (1939) used anion radii (which are close to the vdW radii) for the lack of direct data. The system of Bondi (1964), the most oft-quoted today, was based on experimental distances in a relatively small number of organic crystals and on some noncrystallographic data (gas kinetic collisions, critical densities, liquid properties).

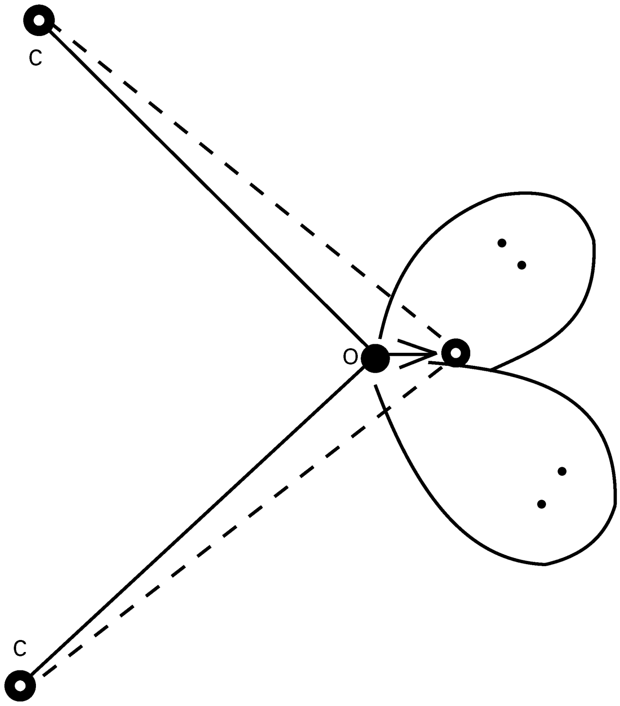



**Figure 2** Ether group geometry determined by neutron (solid) and X-ray (dashed) diffraction; in the latter the O atom is shifted towards its lone-pair electrons.

However, because of the weakness and nonspecific nature of vdW forces, nonbonded distances vary more widely than bond lengths. By a statistical analysis of numerous structures, Rowland and Taylor derived a novel system of vdW radii (**Figure 3**, **Table 3**). Atoms of Cl, Br, I, S, Se and Te, forming one covalent bond, show a highly anisotropic shape in intermolecular contacts (**Figure 4**), which can be described as an ellipsoid compressed along the bond direction, while harder N, O and F atoms remain essentially spherical.

Nonbonded contacts considerably shorter than predicted by the vdW radii, are often indicative of specific interactions between molecules (hydrogen bonds, secondary coordination, charge-transfer, etc.) and are useful in analysing solid-state properties (e.g of organic conductors, so called 'organic metals').

It must be stressed that vdW radii have different meanings in molecular mechanics, where the sum R(A) + R(B) corresponds to the minimum of the potential curve of the *two-atom* A· · ·B interaction. Crystallographic vdW radii were derived to describe the distances of *closest* approach between *polyatomic* molecules, often of complex shapes, and are 8–20% shorter than the equilibrium radii. Only for inert gases, monoatomic in the solid state, do the two systems coincide.

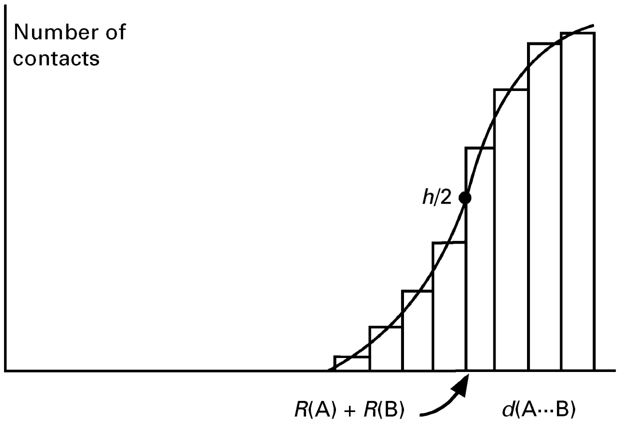



**Figure 3** A typical histogram of nonbonded contact distances; the half-height point A corresponds to the sum of the van der Waals radii (after Rowland and Taylor).

## Structural databases

The abundant data on X-ray crystal structures was made easily accessible with the development of computerized databases. Small-molecule organic and organometallic structures are covered by the Cambridge Structural Database (Cambridge, UK), which lists all compounds containing at least one 'organic' carbon (carbide, carbonate, CN or CO ligand carbon is not regarded as such). For each entry, the crystal lattice parameters and related data, atomic coordinates and (for more recent studies) ADP, bond distances and angles, molecular and structural formula are stored. The database is updated quarterly and widely distributed among academic and nonacademic users worldwide. Its software permits retrieval of entries according to the chemical class, full or partial compound name, composition and chemical connectivity, conditions of determination (pressure, temperature), as well as features of the three-dimensional structure. It also facilitates various statistical and numerical tests on the retrieved data, in order to find regularities of structural parameters and correlations between them.

Crystal structures of minerals and other inorganic substances are covered by the Inorganic Crystal Structure Database which is compiled by the University of Bonn (Germany). It also lists large numbers of molecular structures.

## Absolute structure

Chiral molecules are those not superimposable with their mirror images, such as those containing a tetra-

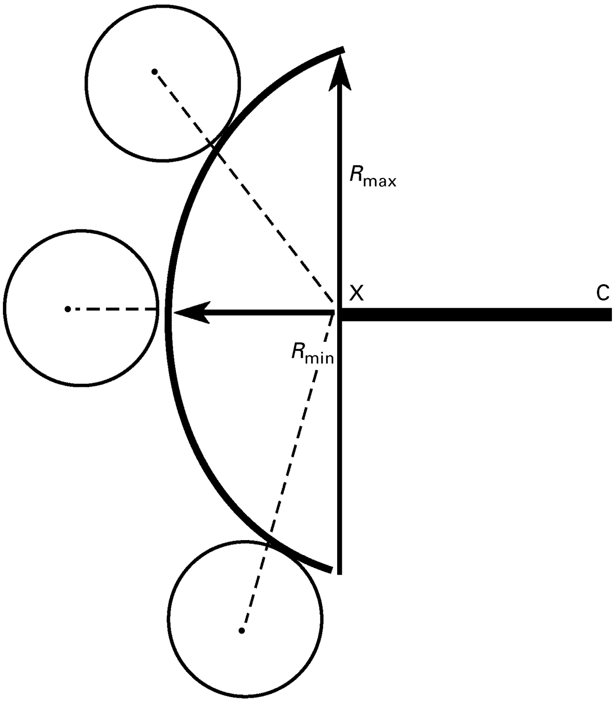



**Figure 4** Anisotropic shape of an atom, forming one covalent bond. Van der Waals radius is at a minimum on the continuation of this bond and maximum normal to it.

**Table 3** Nonbonded contact (or van der Waals) radii for crystals (Å)

| *Element* | *P* | *B* | *ZZ* | *RT* | *NF,max* | *NF, min* |
|---|---|---|---|---|---|---|
| H | 1.2 | 1.20 | 1.16 | 1.10 | 1.26 | 1.00 |
| C | 1.7 | 1.70 | 1.71 | 1.77 | | |
| N | 1.5 | 1.55 | 1.50 | 1.64 | 1.60 | 1.60 |
| O | 1.4 | 1.52 | 1.29 | 1.58 | 1.54 | 1.54 |
| F | 1.35 | 1.47 | 1.40 | 1.46 | 1.38 | 1.30 |
| Cl | 1.8 | 1.75 | 1.90 | 1.76 | 1.78 | 1.58 |
| Br | 1.95 | 1.85 | 1.97 | 1.87 | 1.84 | 1.54 |
| I | 2.15 | 1.98 | 2.14 | 2.03 | 2.13 | 1.76 |
| S | 1.85 | 1.80 | 1.84 | 1.81 | 2.03 | 1.60 |

P, Pauling, 1939–1960; B, Bondi, 1964; ZZ, Zefirov and Zorkii, 1974–1989; RT, Rowland and Taylor, 1996 (see **Figure 3**); NF, Nyburg and Faerman, 1985–1987 (anisotropic system, see **Figure 4**).

hedral carbon (or other) atom with four different substituents (asymmetric centre). They are of great import in biochemistry: natural sugars, amino acids and hence proteins, etc. Molecules of the same chemical composition and molecular geometry, but differing by a mirror reflection (enantiomers, or optical isomers) are indistinguishable in their chemical behaviour (unless another chiral reagent is involved) and in most physical properties. Few physical methods can indicate the difference between enantiomers: optical rotation (ability to rotate the polarization plane of light, in the opposite sense for enantiomers), circular dichroism and the formation of crystals with chiral outer form. None, except X-ray diffraction, can determine the absolute configuration of an asymmetric atom, i.e the real configuration of the substituents in space (*R* or *S*, see **Figure 5** for explanation). The notation introduced by Fischer was based on an arbitrary assignment of the absolute configuration of (+)-glyceraldehyde, to which other optically active compounds could be related by virtue of their chemical origin from one another.

For normal X-ray scattering, centrosymmetrically related reflections with indices *hkl* and $\bar{h}\bar{k}\bar{l}$ always have equal intensities (Friedel's law), thus the diffraction pattern is always centrosymmetric, no matter whether the crystal itself has an inversion centre or not. However, for wavelengths slightly lower than the absorption edge of an atom, a resonant interaction of the X-rays with the inner electrons results in *anomalous scattering*, in which both the intensity and the phase of the radiation is changed. The atomic scattering factor ($f_a$) becomes a complex function:

$$f_a = f + \Delta f' + i\Delta f''$$

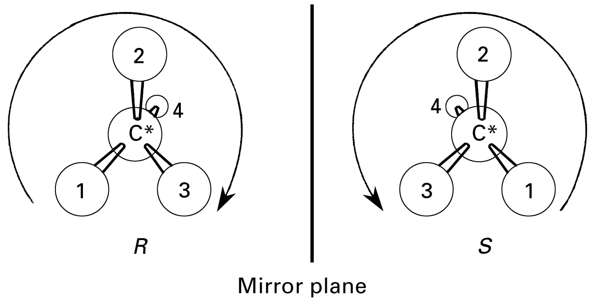



**Figure 5** Notation of the absolute configuration of an asymmetric carbon atom. The substituents are numbered according to certain rules (e.g. beginning with the highest atomic weights, e.g. I > Br > Cl > F); the highest-number substituent pointing away, the numbers of the three others increase clockwise for the *R* configuration and anticlockwise for the *S* one.

where $\Delta f'$ and $\Delta f''$ depend on the wavelength, and for a noncentrosymmetric crystal the *hkl* and $\bar{h}\bar{k}\bar{l}$ reflections differ in intensity. The effect is stronger for heavier atoms and longer wavelengths. Bijvoet (1951) first utilized this anomaly to determine the absolute configuration of the dextrorotatory (+)-tartaric acid in the form of its sodium-rubidium salt, using $ZrK_\alpha$ radiation. (By good luck, this coincided with Fischer's notation.) Today, it is possible to determine reliably the absolute configuration of a medium-sized molecule containing one atom of phosphorus or a heavier element, using $MoK_\alpha$ or $CuK_\alpha$ X-radiation. With the latter wavelength, a careful routine experiment can determine an absolute configuration from the anomalous scattering of oxygen atoms as the heaviest element.

Various methods are used to determine the absolute configuration of the crystal structure, which necessarily defines that of a molecule (although a chiral crystal may consist of nonchiral molecules). The intensities of a few dozens reflections, for which the effect is the strongest, can be calculated together with their inversion equivalent (Friedel pairs) and compared to the observed intensities. Least squares refinements of the 'solved' structure and then of its inverted equivalent, will give a lower *R*-factor for the correct enantiomer. The anomalous scattering corrections themselves ($\Delta f''$) or some coefficients linked to them can be included in the least-squares refinement as a variable. The best method is to refine the 'Flack' parameter during the least-squares procedure. This not only gives the absolute configuration of the crystal, but also (through the esd) a measure of the reliability of the assignment. In addition, it permits the identification and analysis of 'twins by inversion', a common problem faced by physicists trying to grow large enantiomerically pure crystals. Finally, if a structure contains more than one chiral centre, and one of them has a known absolute configuration, those of other(s) can always be determined, even if anomalous scattering is unobservable. For this purpose, an unknown chiral product can be cocrystallized with a known one and the structure of the complex solved by diffraction methods.

## Atomic displacements

Atoms and molecules in crystals are not static, but oscillate at their potential minima and in some cases can jump between different minima. At room temperature, nonhydrogen atoms in a well-ordered organic crystal perform thermal oscillations with mean square amplitudes $\langle u^2 \rangle$ of 0.04–0.05 Å$^2$ (which means the r.m.s amplitudes of ~0.2 Å, comparable

with bond lengths). At 100 K (the practical limit for most cooling devices using liquid nitrogen) $\langle u^2 \rangle$ can be reduced to ~0.015 Å$^2$, but some (zero-point) oscillations persist even at 0 K. Generally, $u$ is in inverse relation to the atomic mass; for outlying atoms of a molecule or some easily rotating substituents, it can be twice the average. Furthermore, atoms can be *disordered*, i.e. have different positions in different asymmetric units of the structure, which are presumed identical by definition of the symmetry of the crystal lattice. Both dynamic and static distortions of the lattice periodicity result in 'smearing' of the electron density, averaged in space over the entire crystal and in time over the entire duration of the experiment. The smearing is described either by atomic thermal parameters, or more precisely atomic displacement parameters (ADP), either in isotropic approximation (spherically averaged for each atom) or in anisotropic approximation, as a symmetrical third-order tensor with six independent parameters. Both representations imply harmonic oscillations; the real vibrations are always anharmonic to some extent, but it is really important to take this effect into account only in high-precision studies of electron density.

Visually, anisotropic ADPs are commonly presented as 'thermal ellipsoids' of a given probability, say, 50% (**Figure 6**). The atom's centre can be found inside the ellipsoid with this probability, and it has *equal* probability of reaching the ellipsoid in any direction. The principal axes of the ellipsoid are proportional to the components of $\langle u^2 \rangle$ in these directions, but other axes are not. The surface visualizing $\langle u^2 \rangle$ components in every direction is not an ellipsoid but a quadric surface of a 'peanut' shape (**Figure 6B**).

## Models of molecular vibrations

The coherent elastic (Bragg) scattering, which alone is measured in the normal X-ray diffraction study, can inform only about the average displacement of every single atom with respect to the crystal lattice. It carries no information whatsoever on how a displacement of one atom relates to that of another atom. In principle, such knowledge can be obtained from inelastic, or thermal diffuse scattering, but the latter is difficult both to measure (its peaks lying under the Bragg ones) and to interpret. Without this knowledge, the same ADPs can be interpreted in different ways, depending on whether the atoms oscillate in phase or out of phase (**Figure** 7).

However, vibrations of a molecular crystal are mainly external modes (oscillations of entire molecules). Intramolecular modes (except some torsional ones), which imply distorting a relatively rigid covalent-bond skeleton, contribute relatively little to the overall amplitudes. From this follows the rigid-bond criterion (Hirshfeld, 1975): if two atoms are linked by a covalent bond, their $\langle u^2 \rangle$ components along the bond direction must be almost equal. For accurately determined structures, this is true within 0.001 Å for X–X bonds (X = C, N, O), 0.003 Å for metal–ligand bonds and 0.005 Å for X–H or X–D bonds. The latter difference is in agreement with the $\langle u^2 \rangle$ of 0.006 Å, calculated from the C–H bond stretching frequency $\nu$ 2700–3300 cm$^{-1}$. Larger differences are indicative of molecular flexibility, as in crystalline Fe(III) complexes (spin-crossover effects), octahedral Cu(II) complexes (dynamic Jahn–Teller and pseudo-Jahn–Teller distortions) and binuclear Mn(II)–Mn(IV) complexes (valence disorder).

In an ideally rigid molecule, the same criterion applies to *all* interatomic vectors. The TLS model,

(A)

(B)

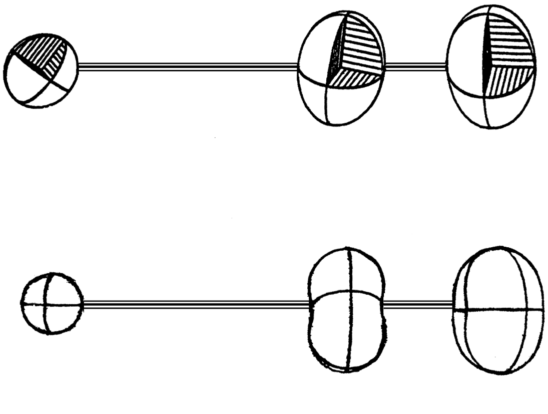
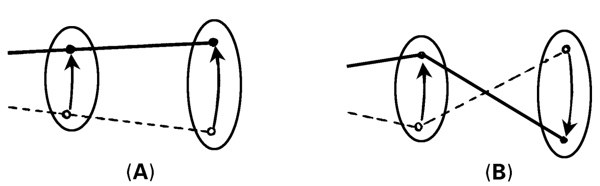


**Figure 6** Atomic displacements in a metal–carbonyl moiety represented by thermal ellipsoids (A) and 'peanut-shaped' surfaces of $(\langle u^2 \rangle)^{1/2}$ (B).



(A) (B)

**Figure 7** In-phase vibrations (A) of a rigid group (CO, CN) and out-of-phase (B) of a flexible group (ethyl) can result in similar ADP ellipsoids.

introduced by Schomaker and Trueblood (1968), describes the motion of such a molecule with three tensors, representing respectively: translations along three principal axes, librations around these axes and screw motions (for noncentrosymmetric cases only). Subtracting the ADPs predicted by this model, from the actual ones, internal nonrigidity of the molecule can be revealed. Torsional motions of a (presumed rigid) group within a molecule can be analysed in the same way, determining torsional amplitudes and, from them, force constants and barriers to internal motion.

### Correcting bond lengths

Rigidly bonded atoms usually perform oscillations along arcs (librations). The electron density thus smeared has its centre-of-gravity (taken for the average atomic position) inside the arc. Hence the X-ray method tends to underestimate bond lengths systematically, the more so (up to 0.03 Å) at higher temperatures (while the real bonds slightly lengthen with temperature). The libration of a C–C bond (~1.5 Å) by 5–6 Å results in a spurious shortening of 0.005 Å, and when by 10 Å, in a shortening of 0.025 Å (**Figure 8**).

If the ADPs from a standard least-squares are consistent with the rigid-body model, bond lengths can

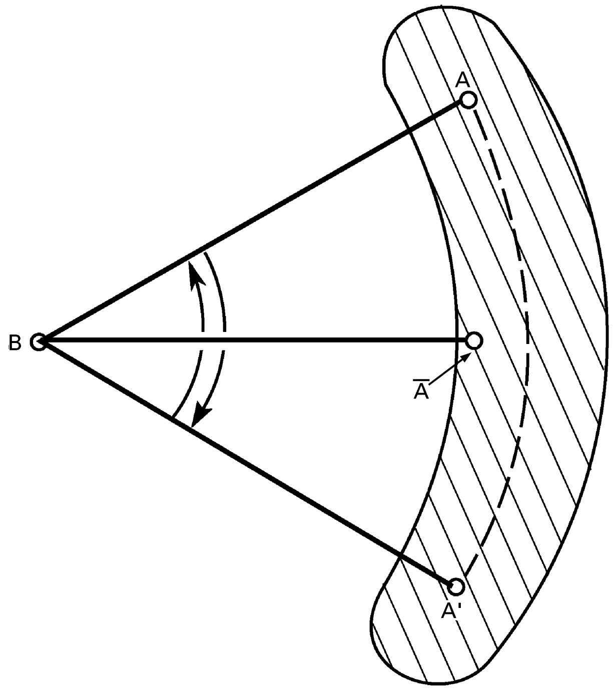



**Figure 8** Systematic underestimation of the length of a librating A–B bond. The terminal atom moves along AA' arc; its average position Ā is found as the centre-of-gravity of the smeared electron density (hatched) and lies within the arc.

be corrected for the spurious shortening. Thus, in a nonrigid molecule of *o*-terphenyl, the mean C–C distance in the phenyl ring is 1.389 Å before correction and 1.398 Å after it, as compared with the $r_g$ = 1.399 ± 0.002 Å from gas-phase electron diffraction. Nevertheless, any thermal motion correction is based on some preconceived idea about the nature of the motion and demands caution. The rigid-bond test can disprove molecular rigidity but not always prove it, as the atoms may vibrate along their connecting vector with the same $\langle u^2 \rangle$, but quite out of phase. The best way to reduce these errors is to diminish the vibrations themselves by cooling the crystal.

### Disorder

From a diffraction study at one temperature one cannot distinguish between dynamic (thermal motion) and static (disorder) displacements. However, if cooling fails to reduce the ADPs, the latter are certainly of static nature, as in the case of ferrocene. The early X-ray studies at room temperature (1951–1956) of the monoclinic modification proved the molecule to possess the crystallographic $C_i$ symmetry, implying the staggered conformation of the rings, in contrast with the eclipsed one in the gas phase. The persistence of a large ADP at 173 K could only be explained as a static disorder, for which various models were suggested. At 164 K, a phase transition occurred into an ordered triclinic structure with four symmetrically independent molecules, all of which adopted nearly (within 10°) eclipsed conformations. A statistical averaging of these molecular orientations results in a picture similar to that observed in the disordered phase. Thus the centrosymmetric conformation is likely to be spurious.

### Erratic ADP

In a least-squares refinement, ADPs tend to become ultimate 'sinks' of all unaccounted for systematic errors, particularly absorption, anisotropic extinction, sometimes incorrect assignment of atomic type or nonstoichiometric occupancy of the position. Erratically elongated ADP ellipsoids, uncorrelated, with those of adjacent atoms, usually betray gross blunders in structure solution, e.g. incorrect crystallographic symmetry or lattice parameters.

## Electron density

The second level of a crystal structure investigation is to go beyond the independent atom model and to resolve more subtle details of the electron density distribution. For such a study one needs to measure

by an order of magnitude more experimental data (measuring all symmetrically equivalents of each reflection rather than just one, increasing the maximum $2\theta$ angle) and to address carefully all possible sources of systematic errors, which can be irrelevant on the atomic approximation level. The mathematical formalisms and software for such studies are still evolving. Only a few hundreds of electron density studies of very uneven quality have been published since 1972 (the earlier works had been attempted with totally inadequate means), of these about 50 were with the multipole technique (see below).

### Deformation maps

The earliest approach was to calculate the total electron density map by a Fourier synthesis and to subtract from its density of isolated atoms, centred at atomic nuclei positions (*promolecule*). These nuclei positions were either determined directly by a neutron diffraction experiment (X–N method) or calculated using only high-order X-ray reflections (X–$X_{ho}$ method), making use of the fact that valence electrons scatter X-rays mainly at lower angles, while the inner electron shells remain practically unperturbed by chemical bonding. The resulting *deformation electron density* maps showed some well-expected features: peaks of electron density in the areas of covalent bonds and lone electron pairs, outward shift of bonding peaks in a cyclopropane ring ('bond bending', **Figure 9**) and so on. However, no peaks appear on bonds involving atoms with the valence shells more than half-filled, e.g. fluorine. In organometallic complexes, only insignificant peaks of electron density were found on the formally quadruple Cr–Cr bond in chromium acetate (0.1 eÅ$^{-3}$) and the single Mn–Mn bond in $Mn_2(CO)_{10}$(0.05 eÅ$^{-3}$), and none at all on the Fe–Fe bond in $[(C_5H_5)Fe(CO)_2]_2$. This reveals the inadequacy of the spherically averaged independent atom as the reference: the $(1s)^2(2s)^2(2p)^5$ atom F has an average of $\frac{5}{3}$ electrons on each p-orbital, which are subtracted from the bonding density. While in fact this atom contributes only one electron for bond formation, the excessive subtraction of $\frac{2}{3}$ *e* more than outweighs the accumulation of electrons in the bond area.

More informative deformation maps could be obtained by subtracting a correctly preoriented independent atom, or a reference atom in a corresponding state of hybridization. Features of one particular bond in a molecule can be most highlighted by subtracting the (calculated by MO methods) electron densities of the two fragments, differing from the molecule by the absence of only this bond.

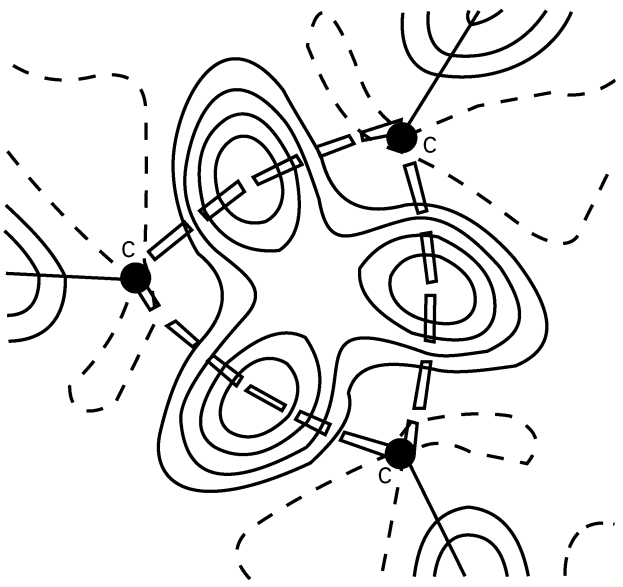



**Figure 9** Deformation electron density map in a cyclopropane ring confirms the concept of 'bent bonds' (dashed lines).

More information can be extracted by topological analysis of the electron density map, particularly using the Laplacian (second gradient) of the electron density:

$$\nabla^2\rho(\mathbf{r}) = \partial^2\rho(\mathbf{r})/\partial x^2 + \partial^2\rho(\mathbf{r})/\partial y^2 + \partial^2\rho(\mathbf{r})/\partial z^2$$

which has maxima in the areas where electron density accumulates (i.e. on chemical bonds) and minima where it depletes.

Determining the phases of structure factors (especially for noncentrosymmetric structures) with the precision sufficient for a deformation density map is an extremely difficult task, as is combining X-ray and neutron data in a sensible way. The irrelevance of the valence electron density for high-order X-ray scattering is also only an approximation (it is correct for the spherically symmetric case only).

### Multipole analysis

More promising is to describe the deformation electron density by a series of spherical harmonic density functions (multipoles), which can be included into least-squares refinement. The inner (core) electron shells of an atom are presumed and the $\kappa$ parameter, which describes the isotropic expansion ($\kappa < 1$) or contraction ($\kappa > 1$) of the valence shell as a whole. Multipole parameters of higher orders describe deviations of the electron density from spherical symmetry. They can be related to the products of atomic

orbitals: monopolar to ss, dipolar to sp, quadrupolar to pp, octupolar to 2p3d and hexadecapolar to 3d3d, and used to analyse atomic orbital populations more directly than density maps.

Studies of octahedral complexes $M(NH_3)_6^{3+}$, $M(CN)_6^{3-}$ (M = Cr, Co) and $Cr(CO)_6$ have confirmed the predictions of the ligand field theory: 66 to 77% of 3d electrons occupy the $t_{2g}$ orbitals (field-stabilized) to the depletion of the destabilized $e_g$ orbitals.

Attempts were made recently to develop a set of deformation parameters, transferable between chemically similar molecules. Their application is much less time-consuming than a full electron density study and may give more realistic least-squares refinement (reduction of displacement parameters) and better predictions of molecular properties than the atomic approximation.

All methods of electron density analysis are hindered by the difficulty of distinguishing between the real (static) deformation and smearing due to thermal motion.

#### Estimates of electric properties

It is possible to estimate directly the charges on atoms and groups of atoms in a crystal, by integrating the electron density over the corresponding areas, if physically sensible borders for them can be suggested. For a monoatomic ion, the dependence of the integral electron density inside a sphere centred on this atom as the function of the radius of this sphere has a distinct minimum, indicating the 'ion boundary'. For the anion in $NH_4Cl$, this gives the radius of 1.75 Å with 17.5 *e* inside the sphere, showing an incomplete charge transfer. In TTF·TCNQ, an organic complex with metallic conductivity, integration over parallelepided-shaped boxes gave the degree of charge transfer of ~0.6 *e* from TTF to TCNQ, in agreement with other estimates. In more complex cases, the physical boundary of an ion can be determined as the surface of zero flux, on which the electron density gradient vanishes.

The electrostatic potential (energy needed to bring a unit positive charge from infinity to a given point) and its first and second derivatives (electric field and its gradient) are of great importance in understanding intermolecular interactions, molecular recognition and reaction mechanisms (as they define the path by which the reagent approaches the substrate). The electrostatic potential can be calculated from the experimental electron density or directly or directly from a high-precision set of structure factors.

## List of symbols

$e$ = electron charge; $f_a$ = atomic scattering factor, comprising the parts independent from ($f$) and dependent on (real $f'$ and imaginary $f''$) the wavelength; $i = \sqrt{-1}$ ; $\mathbf{r}$ = radius-vector; $R$ = discrepancy factor between the observed structure factors of reflections ($F_o$) and those calculated from the structure ($F_c$), $R = \Sigma\,||F_o| - |F_c||/\Sigma|F_o|$; $r_g$ = thermal average internuclear distance; $\langle u^2 \rangle$ = mean square amplitude of oscillations; $\kappa$ = expansion–contraction parameter; $\nu$ = frequency; $\rho$ = electron density; $\sigma$ = estimated standard deviation (esd); $\nabla$ = Laplacian.

*See also:* **Chiroptical Spectroscopy, General Theory; Laboratory Information Management Systems (LIMS); Microwave Spectrometers; Structure Refinement (Solid State Diffraction).**

## Further reading

Allen FH, Bergerhoff G and Sievers R (eds) (1987) *Crystallographic Databases*. Chester: International Union of Crystallography.

Burgi H-B and Dunitz JD (eds) (1994) *Structure Correlation*, Vol 1, 2. Weinheim: VCH.

Coppens P (1997) *X-Ray Charge Densities and Chemical Bonding*. Oxford: Oxford University Press.

Domenicano A and Hargittai I (eds) (1992) *Accurate Molecular Structures, their Determination and Importance*. Oxford: Oxford University Press.

Dunitz JD (1979) *X-Ray Analysis and the Structure of Organic Molecules*. Ithaca: Cornell University Press.

Glusker JP, Lewis M and Rossi M (1994) *Crystal Structure Analysis for Chemists and Biologists*. Weinheim: VCH.

Kitaigorodskii AI (1994) *Molecular Crystals and Molecules*. New York: Academic Press.

# Sodium NMR Spectroscopy

*See* **NMR Spectroscopy of Alkali Metal Nuclei in Solution.**

# Solid State NMR of Macromolecules

*See* **High Resolution Solid State NMR, $^{13}C$.**

# Solid-State NMR Using Quadrupolar Nuclei

**Alejandro C Olivieri**, Universidad Nacional de Rosario, Argentina

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

Compared with solution NMR, the study of solid samples presents certain peculiarities derived from the nuclei being fixed at their lattice positions in a crystal. This rigidity gives rise to the effect of various anisotropic interactions in NMR spectra, the most common of which are given **Table 1**. All these interactions produce orientation-dependent resonances, resulting in characteristic splittings and/or non-Lorentzian broadening of the solid-state NMR signals. The line broadening may range from a few ppm up to thousands of ppm, which helps to explain why much effort has been devoted to the development of line narrowing techniques. Since 1950, for example, Andrew and co-workers have pioneered a method known as magic-angle spinning (MAS), in which the sample is spun inside a suitable rotor, at an angle of 54.7° with respect to the external magnetic field $\boldsymbol{B}_0$ (**Figure 1**). The expected effect of MAS is to average out the orientational dependence of NMR lines, rendering a solution-like spectrum where the relevant

**Table 1** Anisotropic interactions affecting NMR spectra in the solid state

| | Parameter[a] | | | | | |
|---|---|---|---|---|---|---|
| *Interaction* | *G Tensor* | *k* | *P* | *Anisotropy* | *Asymmetry* | *Average value in solution* |
| Chemical shift | $\sigma$ | $\gamma/2\pi$ | $B_0$ | $\Delta\sigma$ | $\eta$ | $\sigma_{iso}$ |
| Direct dipolar coupling | $D$ | $D$ | $S$ | $-3D$ | 0 | 0 |
| Indirect dipolar coupling | $J$ | 1 | $S$ | $\Delta J$ | $\eta_J$[b] | $J_{iso}$ |
| Quadrupole coupling[c] | $q$ | $\frac{e^2Q}{2S(2S-1)h}$ | $P = I = S$ | $3q_{zz}/2$ | $\eta^{\Theta}$ | 0 |

[a] The general form of the Hamiltonian is $h^{-1}H = kIGP$.
[b] The usual assumption is $\eta_J = 0$ and $J$ coaxial with $D$.
[c] The anisotropy is defined as $\Delta q = q_{zz} - (q_{xx} + q_{yy})/2$, and the asymmetry as $3(q_{yy} - q_{xx})/2\Delta q$, where $q_{xx}$, $q_{yy}$ and $q_{zz}$ are the principal components of the $q$ tensor.
[d] $Q$ is the nuclear quadrupole moment and $q_{zz}$ is the maximum value of the field gradient tensor $q$.

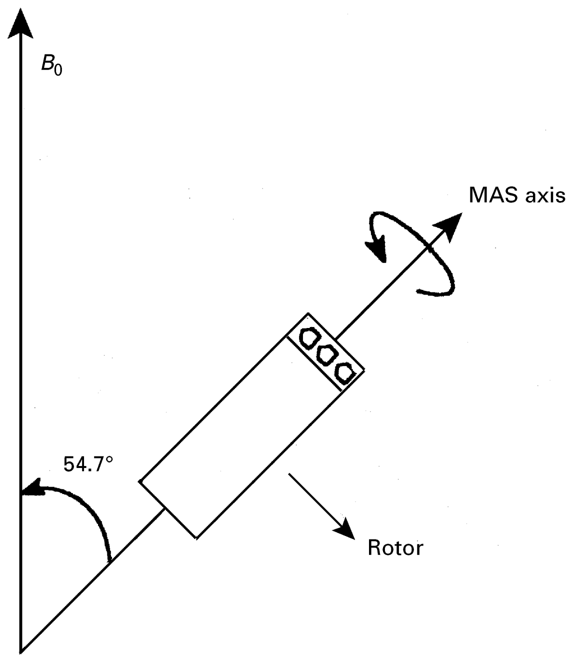



**Figure 1** Scheme showing how MAS is implemented: the sample is placed inside a rotor, which spins in the kHz range about an axis inclined at 54.7° with respect to the external magnetic field $\boldsymbol{B}_0$.

interactions are collapsed to their corresponding isotropic values (**Table 1**). Owing to the successful combination of MAS with sensitivity enhancement pulse sequences (most notably cross-polarization from abundant to dilute spins), solid state NMR has evolved into a technique with sensitivity and resolution comparable to its solution counterpart. It should be borne in mind, however, that MAS is successful provided two conditions are met. On one hand, the dependence of NMR transitions with the orientation should be of the form $(1 - 3\cos^2\theta)$, otherwise split and/or broad lines may result. On the other hand, the frequency of sample spinning should be comparable or higher than the anisotropy. If this latter condition is not attained, the spectrum will not only show an isotropic resonance, but also a number of spinning sidebands separated from the latter by integer multiples of the spinning frequency. The presence of such a sideband manifold may in itself constitute a nuisance, but in certain circumstances it provides a means to measure interesting molecular properties.

The present article has been divided in two main sections, which are briefly described below.

(1) When the observed nucleus is spin-$\frac{1}{2}$ and a neighbouring nucleus is quadrupolar, i.e. spin $>\frac{1}{2}$. In this case, the main interactions suffered by the nucleus are the chemical shift and the dipolar coupling to neighbouring nuclei. One would expect a simple, solution-like spectrum if high-speed MAS was applied. However, a subtle combination of the dipolar coupling with quadrupole interactions at neighbouring nuclei leads to the appearance of additional splittings, which cannot be averaged by MAS. Interesting information concerning molecular and structural properties of solid materials has been gathered from the study of these splittings.
(2) When the observed nucleus is itself quadrupolar. Here the quadrupole interaction is dominant, the others being usually very minor in comparison. Normal MAS speeds are much smaller than typical quadrupole couplings (which lie in the MHz range), producing an enormous number of sidebands, except in the case of the central transition of half-integer spins. In this latter case distinctive line broadening effects appear which cannot be completely removed by MAS. Methods to overcome these difficulties will be reviewed.

## Quadrupole effects on spin-$\frac{1}{2}$ NMR spectra

### Simple MAS spectra

In this section we explain the basis of the success of MAS in achieving spectral narrowing in solid-state NMR. When a single, observed spin-$\frac{1}{2}$ I nucleus is coupled to a single quadrupolar S nucleus, the complete Hamiltonian is:

$$H = H_Z(\mathrm{I}) + H_Z(\mathrm{S}) + H_{\mathrm{CS}}(\mathrm{I}) + H_{\mathrm{CS}}(\mathrm{S}) + H_D(\mathrm{I,S}) + H_J(\mathrm{I,S}) + H_Q(\mathrm{S}) \quad [1]$$

where $H_Z$(I) and $H_Z$(S) are the Zeeman terms for nuclei I and S respectively. They are given by:

$$h^{-1}H_Z(\mathrm{I}) = -\nu_{0\mathrm{I}}I_z \quad [2]$$

$$h^{-1}H_Z(\mathrm{S}) = -\nu_{0\mathrm{S}}S_z \quad [3]$$

where $\nu_{0\mathrm{I}} = \gamma_\mathrm{I}B_0$ and $\nu_{0\mathrm{S}} = \gamma_\mathrm{S}B_0$ are the nominal NMR frequencies. The forms of the other Hamiltonian terms appearing in Equation [1] are summarised in **Table 2**.

The usual assumption in computing NMR spectral resonances is to consider that the Zeeman terms are

**Table 2** Hamiltonian terms affecting an I,S spin system and secular contributions to the NMR spectra.

| *Term* | *Interaction* | *Affected nucleus* | *Secular contribution to NMR spectra*[a] |
|---|---|---|---|
| $h^{-1}H_{cs}(\mathrm{I})$ | Chemical shift of I | I | $-\nu_{0I}[\sigma_{iso}(I) - \Delta\sigma(I)(1-3\cos^2\theta)/3]I_z$ |
| $h^{-1}H_{cs}(\mathrm{S})$ | Chemical shift of S | S | $-\nu_{0S}[\sigma_{iso}(S) - \Delta\sigma(S)(1-3\cos^2\theta)/3]S_z$ |
| $h^{-1}H_D(\mathrm{I,S})$ | Direct dipolar coupling [b,c] | I and S | $D\, I_zS_z(1-3\cos^2\theta)$ |
| $h^{-1}{}_J(\mathrm{I,S})$ | Indirect dipolar coupling [c,d] | I and S | $J_{iso}I_zS_z - (\Delta J/3)I_zS_z(1-3\cos^2\theta)$ |
| $h^{-1}H_Q(\mathrm{S})$ | Quadrupole coupling[e] | S | $\frac{\chi}{4S(2S-1)}[3S_z^2\cos^2\theta - S^2 + \left(\frac{3}{4}\right)(S_+S_- + S_-S_+)\sin^2\theta]$ |

[a] The angle $\theta$ between the main tensor axes and the external magnetic field is characteristic of each interaction; the expressions shown in this table are valid only if the relevant tensors are axially symmetric and coaxial.

[b] $D = (\mu_0/4\pi)(\gamma_I\gamma_S h/4\pi^2 r_{I,S}^3)$ is the dipolar coupling constant ($r_{I,S}$ is the internuclear distance).

[c] The secular contributions for these interactions are valid for heteronuclear spin systems. For homonuclear spin systems, the flip-flop term containing ($I_+S_- + I_-S_+$) is also secular.

[d] The usual assumption of $J$ coaxial with $D$ leads to an effective dipolar constant $D' = D - (\Delta J)/3$.

[e] The quadrupole coupling constant is defined as $\chi = e^2Qq_{zz}/h$.

dominant in Equation [1], and thus that first-order perturbation theory can be applied. Using the information given in **Table 2**, the NMR lines for an observed I nucleus are given by

$$\nu_I = \nu_{0I}[1 - \sigma_{iso}(I)] - m_S J_{iso} + [\Delta\sigma(I)\nu_{0I}/3 - m_S D'](1 - 3\cos^2\ ) \quad [4]$$

where $m_S = -S, -S+1, \ldots, S-1, S$ are the spin components for nucleus $S$. Equation [4] is true provided the relevant tensors $\sigma$, $\boldsymbol{D}$ and $\boldsymbol{J}$ are axially symmetric and coaxial. Notice that the quadrupole interaction plays no role in Equation [4], owing to the fact that the latter does not contain I spin operators (**Table 2**). If MAS is applied at a sufficiently high speed, i.e. when the sample spinning frequency $\nu_r$ is larger than the effective anisotropy ($\Delta\sigma\nu_{0I} - 3m_S D'$), the factor $(1-3\cos^2\theta)$ in Equation [4] becomes $(1-3\cos^2\xi)\,[1-3\cos^2(54.7°)]/2$ (here $\xi$ is the angle between the main axes of the interaction tensors and the sample spinning axis). Since $\cos^2(54.7°) = \frac{1}{3}$, Equation [4] predicts, under MAS, a $J$-coupled multiplet centred at the isotropic I chemical shift:

$$\nu_I = \nu_{0I}[1 - \sigma_{iso}(I)] - m_S J_{iso} \quad [5]$$

i.e. a solution-like spectrum. This constitutes the basis of the successful applications of MAS which led to the development of high-resolution solid-state NMR.

When the spinning speed is low, the result is well known: the isotropic resonance appears flanked by a number of sidebands, located at integer multiples of the spinning frequency. Suitable analysis of the intensities of these sidebands allows of the chemical shift parameters $\Delta\sigma$ and $\eta$ for nucleus $I$.

## Second-order effects

The simple, solution-like Equation [5] was found to apply in most solid-state NMR spectra. However, conference reports in 1979, and papers in the scientific literature soon after, showed both experimentally and theoretically that this was not the case when the quadrupole interaction at a neighbouring S nucleus is comparable to its Zeeman interaction. To interpret these results, first-order theory needs to be corrected with second-order effects. The latter are the result of the interplay of the tensors $\boldsymbol{D}$ and $\boldsymbol{q}$, through non-secular terms of the corresponding $H_D(\mathrm{I,S})$ and $H_Q(\mathrm{S})$ Hamiltonians (specifically, the single quantum transition terms containing the operators $I_\pm S_z$ and $I_zS_\pm$). They lead to the appearance of terms having an orientational dependence other than $(1-3\cos^2\theta)$, which cannot be averaged out even by high-speed MAS. Second-order theory allowed the derivation of the following simple equation:

$$\nu_I = \nu_{0I}[1 - \sigma_{iso}(I)] - m_S J_{iso} + \left(\frac{3D'\chi}{10\nu_{0S}}\right) \times \left[\frac{S(S+1) - 3m_S^2}{S(2S-1)}\right] \quad [6]$$

From the second-order shift (the last term in the right-hand side of Equation [6]), the following six conclusions can be drawn.

(1) Second-order effects scale inversely with the applied magnetic field, in contrast to $J_{iso}$ or chemi-

cal shift effects. Thus, experiments conducted at different fields provide evidence that the observed splittings are indeed due to quadrupole effects.

(2) The second-order shifts depend on $m_S^2$, and therefore their values occur in pairs. When $J_{iso} = 0$ (and $\Delta J$ is also zero), Equation [6] predicts the I line as a 2:1 doublet for $S = 1$, and as a 1:1 doublet for $S = 3/2$ (**Figures 2A** and **2B**). Simple equations for the doublet splittings $s$ (**Figures 2A** and **2B**) exist (Eqn [6]): $s = 9D_\chi/10\nu_{0S}$ for $S = 1$, and for $s = 6D\chi/10\nu_{0S}$ for $S = 3/2$. Notice that for $S = 1$ the doublet is asymmetric (**Figure 2A**), causing $s$ to have a definite sign (the convention is that $s$ is positive if the smallest peak appears at higher frequencies).

For a finite $J_{iso}$, the I line is predicted to be a distorted $J$-multiplet (**Figures 2C** and **2D**). Since the outermost lines of the multiplet shift in opposite direction as compared with the innermost lines, the spectra are 'bunched' at one end, in a manner which also depends on the sign of $\chi$ the quadrupole coupling constant (**Table 2**).

It should be noticed that in all cases the lines are not single peaks but have a distinct powder pattern shape (**Figures 3A** and **4A**).

(3) The average of the second-order shifts over $m_S$ is zero, and hence the isotropic $\sigma_{iso}$(I) is obtained by averaging the multiplet line frequencies.

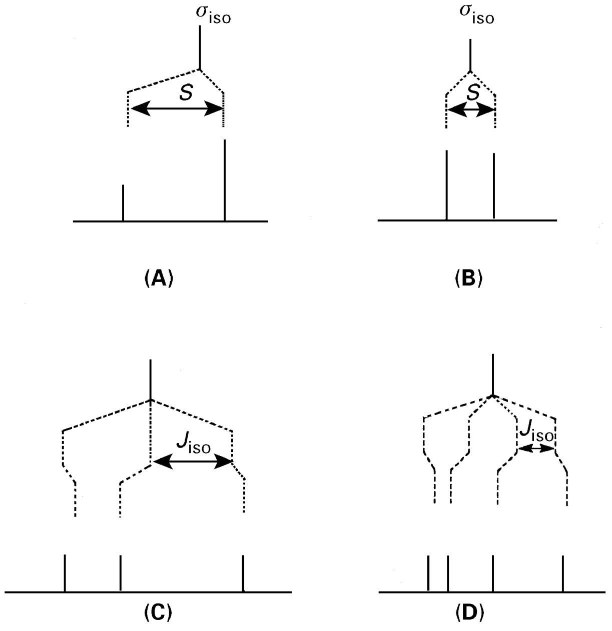



**Figure 2** Spectral appearance of the solid-state NMR spectrum of a spin-$\frac{1}{2}$ nucleus (I), including second-order quandrupole effects from S when: (A) $S = 1$, $J_{iso} = 0$; (B) $S = \frac{3}{2}$, $J_{iso} = 0$; (C) $S = 1$, $J_{iso} \neq 0$; (D) $S = \frac{3}{2}$, $J_{iso} \neq 0$. In all cases Equation [6] applies, with $\chi$ positive. The frequency axes increase from right to left.

(4) The value of $J_{iso}$ is obtained by averaging the multiplet line spacings, or from the central spacing if $S$ is half-integer (**Figures 2C** and **2D**).

(5) The sign of $s$ or the sense of spectral 'bunching' provides experimental access to the sign of $\chi$, which is difficult to obtain from other techniques. An exception is the 1:1 doublet for $S = 3/2$ and $J_{iso} = 0$, in which case the information on the sign of $\chi$ is lost.

(6) If the molecular geometry and the quadrupole parameters are known, the observation of distorted multiplets may allow the determination of the asymmetry in the $\boldsymbol{J}$ tensor $\Delta J$.

When the assumptions of axial symmetry and coaxiality among the tensors are relaxed, the following general equation is obtained:

$$\nu_I = \nu_{0I}[1 - \sigma_{iso}(I)] - m_S J_{iso} + \left(\frac{3D'\chi}{20\nu_{0S}}\right)\left[\frac{S(S+1) - 3m_S^2}{S(2S-1)}\right] \times (3\cos^2\beta^D - 1 + \eta_Q \sin^2\beta^D \cos 2\alpha^D) \quad [7]$$

Equation [7] incorporates not only $\chi$ but also $\eta_Q$, as well as two angles ($\beta^D$ and $\alpha^D$) that fix the mutual orientation of $\boldsymbol{D}$ and $\boldsymbol{q}$. As expected, Equation [6] is a special case of Equation [7] when $\beta^D = 0$, i.e. when $\boldsymbol{D}$ and $\boldsymbol{q}$ are coaxial.

Finally, when the value of $\chi$ is larger than the NMR frequency $\nu_{0S}$ second-order theory breaks down, and one needs to resort to complete full-matrix Hamiltonian calculations, with the results shown in **Figures 3B** and **4B** for $S = 1$ and $S = 3/2$, respectively (in both cases $J_{iso} = 0$). As can be appreciated, for low values of the ratio $(\chi/\nu_{0S})$ the predictions of Equation [6] are in agreement with the full calculations.

It is important to note that the effects described by Equation [6] are only observed in rigid solids. Both in solution and in highly mobile solid phases, random molecular motions average out all anisotropic contributions, leaving only Equation [5] (a further motional effect may be a fast quadrupole relaxation on nucleus S, which would erase the multiplet structure of the I signal).

## Experimental examples and applications

**Tables 3** and **4** summarize nuclear, molecular and structural parameters as well as the spectral appearance for some studied spin pairs giving rise to second-order quadrupole effects on the I line, as described above. In the case of $^{13}C$, $^{14}N$, the ratios $(\chi/\nu_{0S})$ are low and have therefore been studied mainly

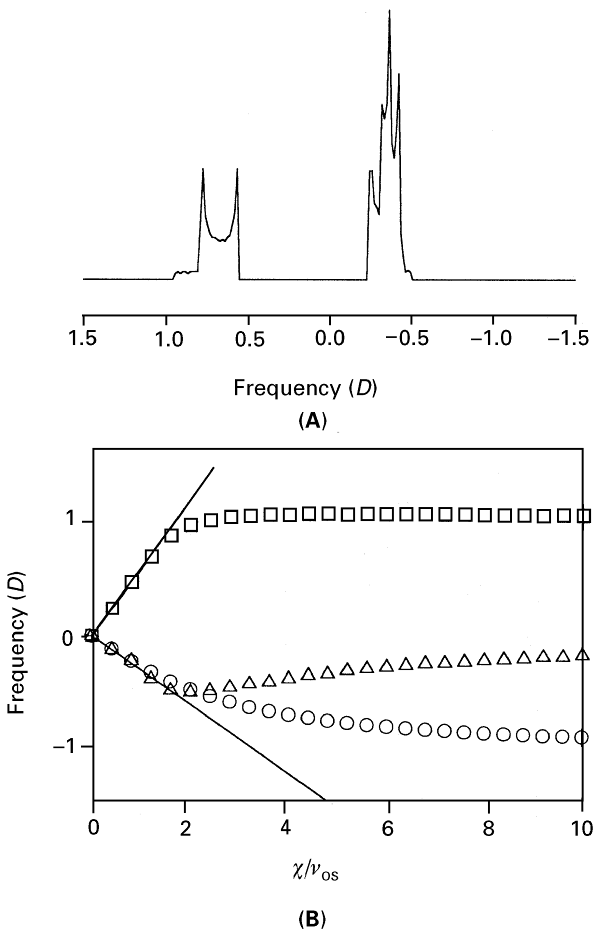



**Figure 3** (A) Powder pattern line shape of an I nucleus coupled to a quadrupolar ($S = 1$) nucleus when $(\chi/\nu_{0S}) = 1$, $J_{iso} = 0$. The frequency axis is in units of the dipolar coupling constant $D$. (B) Frequencies (in units of $D$) of the three lines expected for an I,S pair ($S = 1$) as a function of the ratio $\chi/\nu_{0S}$. The line positions marked with symbols have been obtained by full-matrix Hamiltonian calculations. The solid lines are the values given by Equation [6].

on the basis of the simple Equation [6] (and its extension to non-symmetric $q$ tensors); in most cases $J_{iso} = 0$ and $D' = D$ (see **Table 2**). The theoretical equations have been used to (1) predict the spectral appearance once the molecular geometry and the quadrupole parameters are known, (2) to derive approximate values of $\chi$ (including sign) from the spectra and (3) to aid in spectral assignment, since the affected carbons appear as characteristic doublets.

Most studies on $^{13}C$, $^{14}N$ second-order effects were done using relatively low field solid-state NMR spectrometers. The advent of high-field instruments has displaced this interesting phenomenon to a rather unfortunate second place: at 7.05 T the effects are rarely seen, unless favourable circumstances occur. An interesting example is provided by a recently studied metal cyanide polymer, in which the molecular geometry suggests that all relevant tensors $\sigma$, $\mathbf{D}$ and $q$ are axially symmetric and coaxial. The $^{13}C$ solid-state MAS NMR spectrum at 7.05 T and high spinning speed shows three asymmetric doublets (corresponding to non-equivalent cyanide sites), from which values of $\chi$ in the range −1.9 to −2.5 MHz have been derived (**Figure 5**). The sideband shapes are also doublets with characteristic line shapes, and have been successfully simulated using second-order theory (**Figure 6**). This simulation also provided, as a by-product, $\Delta\sigma = 350$ ppm for the $^{13}C$ chemical shift tensor.

Another spin-1 nucleus causing second-order effects is $^2H$ (**Table 3**). $^{13}C$ MAS NMR spectra of solid deuterated organic molecules show distorted triplets, as expected from Equation [6] when $J_{iso} \neq 0$. In this case, the small value of $\chi$ for $^2H$ is compensated by a large dipolar coupling constant $D$ (**Table 4**).

In the case of $^{13}C$ nuclei coupled to chlorine (**Table 3**), the values of $(\chi/\nu_{0S})$ for $^{35,37}Cl$ are such

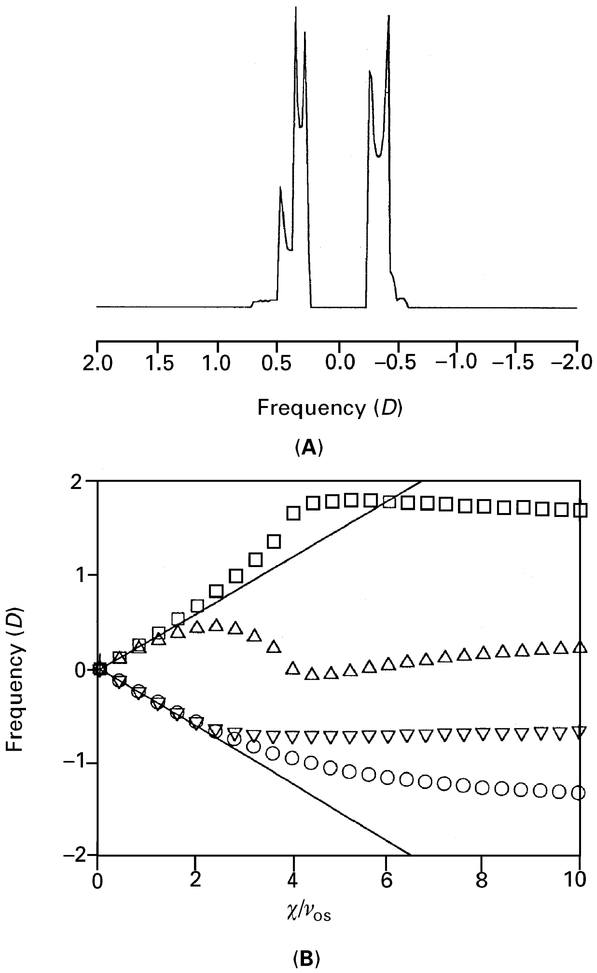



**Figure 4** (A) Powder pattern line shape of an I nucleus coupled to a quadrupolar ($S = 3/2$) nucleus when $\chi/\nu_{0S} = 1$, $J_{iso} = 0$. The frequency axis is in units of the dipolar coupling constant $D$. (B) Frequencies (in units of $D$) of the four lines expected for an I,S pair ($S = 3/2$) as a function of the ratio ($\chi/\nu_{0S}$). The line positions marked with symbols have been obtained by full-matrix Hamiltonian calculations. The solid lines are the values given by Equation [6].

that 1:1 doublets are observed even at high fields, as described by Equation [6] when $J_{iso}$ is negligible. On the other hand, in $^{119}$Sn spectra of chlorostannic compounds, distorted quartets have been observed which allowed the determination of the parameter $\Delta J$ (**Table 4**). Notice that axial symmetry and coaxiality of $\boldsymbol{D}$ and $\boldsymbol{q}$ are plausible assumptions for the X–Cl bond. Since the nuclear properties of both chlorine isotopes are similar (**Table 3**), only average effects are observed in the spectra, except in the special circumstances of very high spectral resolution.

When the quadrupolar nucleus is bromine, $J_{iso}$ effects are important, and $^{13}$C NMR spectra of C–Br carbons appear as asymmetrically distorted quartets (**Figure 7**). Further, the ratio ($\chi/\nu_{0S}$) is in this case large (**Tables 3** and **4**), and second-order theory cannot be applied. Thus, full-matrix calculations were used to account for the observed spectra, as well as a so-called 'inverse' first-order theory, in which the Zeeman term is considered as a small perturbation on the quadrupole Hamiltonian. In any case, knowing both the C–Br distance and Br quadrupole coupling constants, and assuming axial symmetry of all tensors around the C–Br bond, allows one to derive approximate values of $J_{iso}$ and $\Delta J$ from the spectra (**Table 4**). As with chlorine, the separate effects of both bromine isotopes are difficult to distinguish (**Table 3**).

Finally, cases involving $^{31}$P coupled to metals should be mentioned. Spectra which involve coupling of $^{31}$P to several quadrupolar metals, e.g. $^{63,65}$Cu ($S = 3/2$), $^{55}$Mn ($S = 5/2$), $^{59}$Co ($S = 7/2$) and $^{93}$Nb ($S = 9/2$) have been found to consist of dis-torted $J$-multiplets, as expected from Equation [6]. The pair $^{31}$P, $^{63,65}$Cu has been extensively studied in a series of phosphine-Cu(I) complexes. Since the ratio ($\chi/\nu_{0S}$) is low, second-order theory allowed the easy calculation of $\chi$ and $\Delta J$ from the spectra (**Table 4**). It is interesting to note that solid-state NMR is one of the few techniques which allows one to measure $\Delta J$,

**Table 3** Nuclear properties for studied pairs of nuclei as regards second-order effects on spin-½ spectra.

| *I.S.Pair* | $\nu_{0I}$ (MHz)[a] | *Natural abundance of I* (%) | $\nu_{0S}$ (MHz)[a,b] | *S*[b] | *Natural abundance of S*(%)[b] | $10^3$ *Q(S)* (barn) |
|---|---|---|---|---|---|---|
| $^{13}$C,$^{14}$N | 75.43 | 1.11 | 21.67 | 1 | 99.6 | 20.1 |
| $^{13}$C,$^{2}$H | 75.43 | 1.11 | 46.05 | 1 | 0.015[c] | 2.86 |
| $^{13}$C,$^{35,37}$Cl | 75.43 | 1.11 | 29.40 | $\frac{3}{2}$ | 75.5 | −81.1 |
| | | | 24.47 | $\frac{3}{2}$ | 24.5 | −63.9 |
| $^{119}$Sn,$^{35,37}$Cl | 111.82 | 8.58 | [d] | [d] | [d] | [d] |
| $^{13}$C,$^{79,81}$Br | 75.43 | 1.11 | 75.16 | $\frac{3}{2}$ | 50.5 | 331 |
| | | | 81.02 | $\frac{3}{2}$ | 49.5 | 276 |
| $^{31}$P,$^{63,65}$Cu | 121.44 | 100 | 79.52 | $\frac{3}{2}$ | 69.1 | −211 |
| | | | 85.18 | $\frac{3}{2}$ | 30.9 | −195 |

[a] At $B_0$ = 7.05 T, for which $\nu(^1H)$ = 300 MHz.
[b] When two isotopes occur, the first entry corresponds to the nuclear properties for the lighter isotope.
[c] Enriched samples.
[d] See the $^{13}$C,$^{35,37}$Cl case.

**Table 4** Typical values of I,S distances, scalar, dipolar and quadrupole coupling constants, and spectral appearance of I spectra owing to second-order quadrupole effects.

| *I,S pair* | $r_{I,S}$ (pm) | *D* (kHz) | $\Delta J$ (kHz) | $J_{iso}$ (kHz) | *Range of* $\lvert\chi\rvert$ (MHz) | *Spectral appearance of I* |
|---|---|---|---|---|---|---|
| $^{13}$C,$^{14}$N | 110–150 | 0.6–1.6 | $\Delta J \ll D$ | ~0 | 0.5–5 | 2:1 Doublet |
| $^{13}$C,$^{2}$H | 100 | 3.6 | $\Delta J \ll D$ | 0.02 | 0.1–0.3 | Distorted triplet |
| $^{13}$C,$^{35,37}$Cl | 170–180 | 0.5–0.6 | $\Delta J \ll D$ | ~0 | 60–80 | 1:1 Doublet |
| $^{119}$Sn,$^{35,37}$Cl | 220–240 | 0.3–0.4 | −0.4 to −0.8 | 0.2–0.4 | 60–80 | Distorted quartet |
| $^{13}$C,$^{79,81}$Br | 180–190 | 1.2–1.3 | ~0.5 | 0.1–0.2 | 450–500 | Distorted quartet |
| $^{13}$P,$^{63,65}$Cu | 220–240 | 0.8–1.0 | ~0.6 | 1–2 | 10–100 | Distorted quartet |

a parameter of somewhat elusive experimental accessibility.

## Self-decoupling

As discussed above, second-order effects are only observed in rigid solid samples. Random molecular motions in solution or in highly mobile solids produce two phenomena: (1) anisotropic dipolar and quadrupolar interactions are averaged to zero (**Table 1**), and (2) fast longitudinal relaxation is induced on the quadrupolar nucleus S, leading to the collapse of all coupling interactions (both dipolar and scalar). The latter result is known as self-decoupling, and is responsible, for example, for why $^{13}$C nuclei in solution do not normally appear as *J*-coupled when bonded to $^{14}$N or $^{35,37}$Cl nuclei. An interesting situation arises when the solid-state motion is anisotropic: the relevant interactions do not completely disappear, but are scaled down, depending on the extent of the motion. Only self-decoupling would be able to erase the expected splittings in this case.

Appropriate examples are provided by sodium chloroacetates. Both $ClCH_2COONa$ and $Cl_2CHOONa$ show the $^{13}$C–Cl signal as the expected

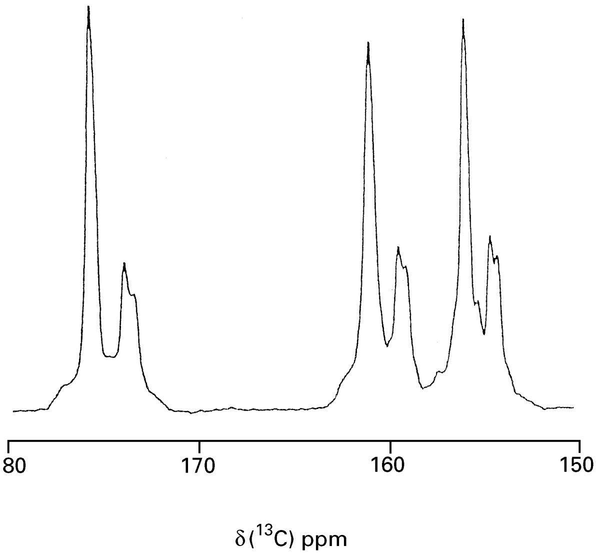



**Figure 5** Solid-state $^{13}$C MAS spectrum of a sample of the polymer $[\{(CH_3)_3Pb\}_4Ru(CN)_6]_\infty$ obtained at a nominal frequency of 75.43 MHz ($B_0$ = 7.05 T) and a spinning speed of 4.3 kHz, by summing all relevant sidebands. There are three different cyanide sites in the solid, each giving a characteristic (negative) splitting *s*.

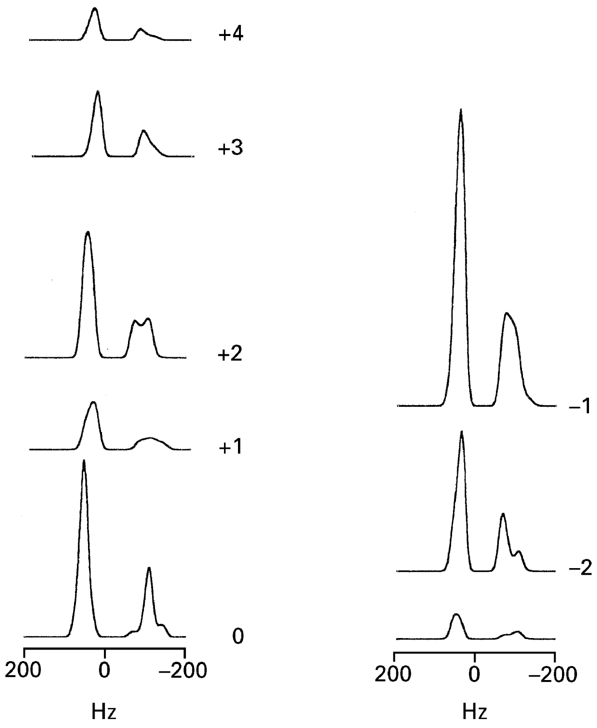



**Figure 6** Simulated shapes of the isotropic line and sidebands (as numbered) for the sample of **Figure 5**. The simulations were done assuming $\chi(^{14}N) = -2.3$ MHz and $\Delta\sigma(^{13}C) = 350$ ppm.

1:1 doublet and 1:2:1 triplet, respectively (provided second-order theory applies) (**Figures 8A** and **8B**). However, $Cl_3CCOONa$ shows a narrow single peak even at low temperatures (**Figure 8C**), owing to a fast rotation of the $Cl_3C$ group around the C–C bond. That the collapse of the expected 1:3:3:1 quartet in the latter case is due to self-decoupling was confirmed by independently measuring the $^{35}Cl$ longitudinal relaxation times in all three salts by using nuclear quadrupole resonance. As expected, the $^{35}Cl$ $T_1$ in $Cl_3CCOONa$ is significantly shorter than in other two compounds.

## Quadrupole effects on NMR spectra from nuclei with spin > $\frac{1}{2}$

### MAS spectra

From the information given in **Table 2**, the following expression for the S NMR signal can be obtained when the observed nucleus is itself quadrupolar:

$$\nu_S = \nu_{0S}[1 - \sigma_{iso}(S)] - m_I J_{iso} + \left[\frac{3\chi(m_S + \frac{1}{2})}{4S(2S-1)} - \frac{1}{3}\Delta\sigma(S)\nu_{0S}\right](1 - 3\cos^2\theta) \quad [8]$$

In general, the spectra are dominated by the quadrupole interaction (i.e. the term containing $\chi$ in Eqn [8]). Although the relevant term in Equation [8] also contains the usual factor $(1 - 3\cos^2\theta)$, the ability of MAS to average out the quadrupole effects depends critically on the spinning speed. Typical values of $\chi$ lie in the MHz range; thus, at experimentally accessible spinning speeds (which rarely exceed tens of kHz) the spectra will have the signal intensity distributed over an enormous number of very weak sidebands. There is an exception to this rule: when $m_S = -\frac{1}{2}$ in Equation [8], the first-order quadrupole effect is zero, and MAS should yield simple spectra. Thus, efforts have been directed to the study of the central transition $(-\frac{1}{2}, +\frac{1}{2})$ in half-integer spin systems. Even when the first-order effect is zero for the latter transition, second-order quadrupolar effects remain which are not completely removed by MAS. The expression for the NMR central transition of half-integer quadrupolar nuclei with axially symmetric $q$, when high-speed MAS is applied, is

$$\nu_S = \nu_{0S}[1 - \sigma_{iso}(S)] - m_I J_{iso} + \nu_{0S}\Delta\sigma_{qs} + \frac{1}{10}\left(\frac{\chi^2}{\nu_{0S}}\right)\frac{9[4S(S+1) - 3]}{[16S(2S-1)]^2} \times (35\cos^4\xi - 30\cos^2\xi + 3) \quad [9]$$

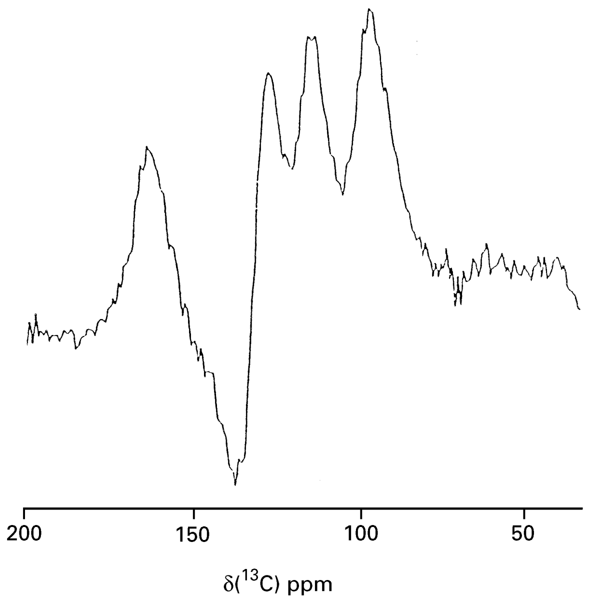



**Figure 7** Quaternary-only solid-state $^{13}C$ NMR spectrum of 1,4-dibromobenzene at 50.33 MHz (4.7 T), showing the C–Br signal as a distorted quartet produced by interaction with $^{79,81}Br$. The negative peak at ~135 ppm is an artifact of the pulse sequence.

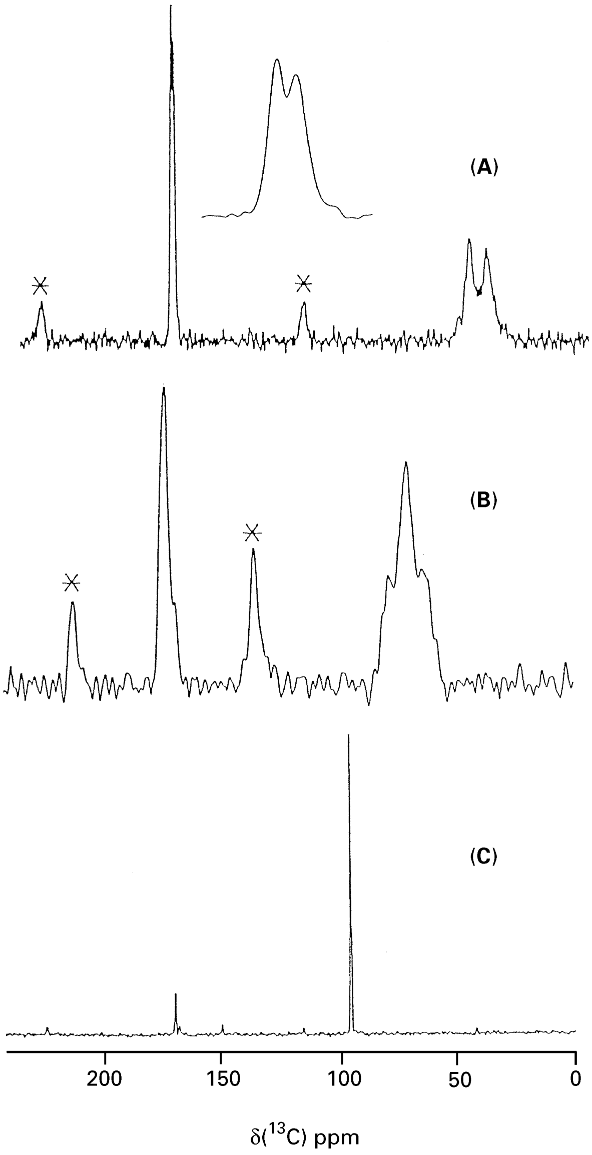



**Figure 8** Solid-state $^{13}C$ NMR spectra of sodium chloroacetates at 75.4 MHz (7.05 T). (A) $ClCH_2COONa$ at room temperature (the insert shows an expansion of the carboxyl region; notice that the effect of coupling to Cl is not limited to directly bonded carbons), (B) $Cl_2CHCOONa$ at 158 K and (C) $Cl_3CCOONa$ at 163 K. Asterisks denote spinning sidebands.

where

$$\nu_{0S}\,\Delta\sigma_{qs} = \frac{3}{10}\left(\frac{\chi^2}{\nu_{0S}}\right)\frac{4S(S+1)-3}{[4S(2S-1)]^2}$$

is the so-called isotropic second-order quadrupolar shift and $\xi$ is the angle between $q_{zz}$ and the rotor axis. Equation [9] will conceivably lead to three effects (**Figure 9**): (1) broad powder pattern line shapes, (2) a shift $\Delta\sigma_{qs}$ of the centre-of-gravity of the spectrum with respect to the isotropic chemical shift, and (3) the occurrence of sidebands (notice that, according to **Table 5**, $\chi^2/\nu_{0S} \ll \chi$ and hence typical spinning speeds will lead to a reasonably low number of sidebands).

In general, appropriate spectral simulations based on Equation [9] are required to retrieve the relevant

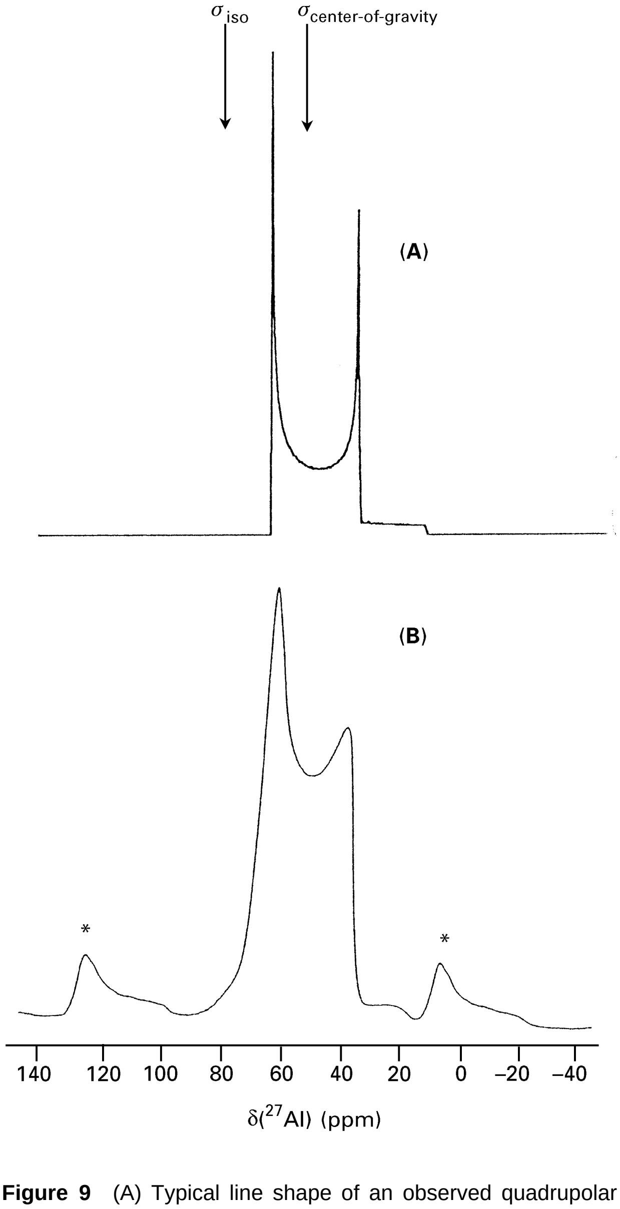



**Figure 9** (A) Typical line shape of an observed quadrupolar nucleus S, showing the second-order quadrupole shift $\Delta\sigma_{qs}$, and the relative position of the centre-of-gravity with respect to the isotropic chemical shift $\sigma_{iso}$. (B) $^{27}Al$ solid-state MAS NMR spectrum of $Sr_8(AlO_2)_{12}{\cdot}Se_2$ at 78.15 MHz (7.05 T). Asterisks denote sidebands. Reproduced with permission of Elsevier Science Publishers from Weller MT, Brenchley ME, Apperley DC and Davies NA (1994) Correlations between $^{27}Al$ magic-angle spinning nuclear magnetic resonance spectra and the coordination geometry of framework aluminates. *Solid State Nuclear Magnetic Resonance* **3**: 103–106.

**Table 5** Nuclear and field gradient properties for some quadrupolar nuclei

| Nucleus | $\nu_{0S}$ (MHz)[a] | S | Natural abundance (%) | $10^3Q(S)$ (barn) | Range of $\|\chi\|$ (MHz) |
|---|---|---|---|---|---|
| $^{23}$Na | 79.35 | $\frac{3}{2}$ | 100 | 100.6 | 1–4 |
| $^{27}$Al | 78.17 | $\frac{5}{2}$ | 100 | 140.3 | 0.5–1 |
| $^{11}$B | 96.25 | $\frac{3}{2}$ | 80.42 | 40.59 | 2–5 |
| $^{17}$O | 40.27 | $\frac{5}{2}$ | 0.037[b] | −25.58 | 1–5 |
| $^{51}$V | 78.86 | $\frac{7}{2}$ | 99.76 | −52 | 0.5–10 |
| $^{7}$Li | 116.59 | $\frac{3}{2}$ | 92.58 | −40.1 | 0.03–0.05 |

[a] At $B_0$ = 7.05 T, for which $\nu(^1H)$ = 300 MHz.
[b] Enriched samples.

information concerning $\chi$ (and $\eta_Q$ in non-symmetric cases), $\sigma_{iso}$ and the chemical shift parameters $\Delta\sigma(S)$ (and $\eta$). Notice that the broadness of the lines may lead to substantial overlap, thereby complicating the spectral interpretation. (**Figure 10**) shows the theoretical effect in a spectrum with two overlapping lines, for typical parameters of $^{17}$O in minerals. Complete separation of peaks is not achieved even at very high magnetic fields.

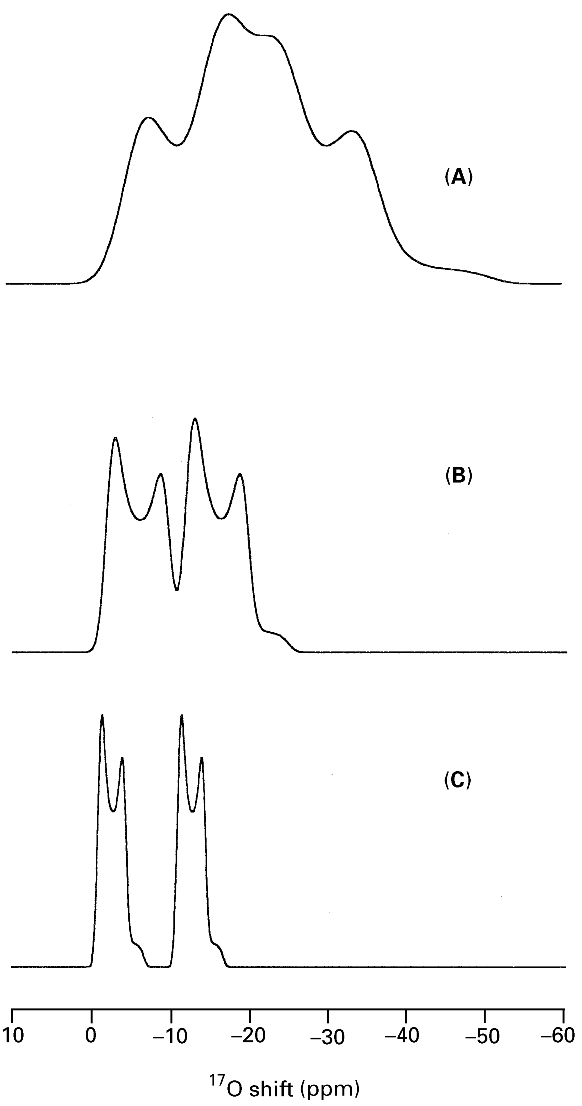



**Figure 10** Second-order quadrupolar spectra expected for two $^{17}$O lines ($S$ = 5/2) located at values of $\sigma_{iso}$ of 0 and 10 ppm, assuming $\chi$ = 2 MHz and axial symmetry for both sites. The external magnetic fields are: (A) 7.05, (B) 11.75 and (C) 17.62 T, corresponding to $^1$H NMR frequencies of 300, 500 and 750 MHz, respectively. The line shapes have been convoluted with a Gaussian broadening.

### Double rotation NMR and other techniques

A significant increase in the fundamental knowledge of quadrupole effects in NMR spectroscopy has taken place in the last decade. Various techniques have been developed to diminish the effects produced by the orientational dependence of Equation [9], or to separate chemical shift from pure quadrupole effects. The most successful seems to be double-rotation (DOR), in which the sample spins simultaneously around two different 'magic' angles: 54.7 and 30.6°. The latter angle has the property that, for $\xi = 30.6°$, the factor $[35\cos^4\xi - 30\cos^2\xi + 3]$ in Equation [9] is zero, leaving a simple spectrum where only the chemical shift and the isotropic second-order quadrupole shift remain. Further distinction of these shifts can be made by recording spectra at different magnetic fields. **Figure 11** shows an example of the dramatic reduction in line width which is attained by application of DOR to a solid sample.

Other relevant methods are (1) quadrupole nutation spectroscopy, a two-dimensional technique which allows the projection of the conventional spectrum in one dimension, and only quadrupolar information in the second frequency axis, (2) dynamic-angle spinning (DAS), in which the sample is spun sequentially rather than simultaneously (as in DOR) about two different 'magic' axes and (3) satellite transition spectroscopy (SATRAS), which monitors NMR transitions other than the central $(-\frac{1}{2},+\frac{1}{2})$ under high-speed MAS, allowing the measurements of both the correct isotropic chemical shift and the quadrupole coupling in a single experiment.

### Studied nuclei

The interest in studying quadrupolar nuclei with NMR is to combine useful chemical shift correlations with information concerning the quadrupole coupling constant. The value of $\chi$ depends on the nucleus itself (though the quadrupole moment $Q$), and on the maximum electric field gradient $q_{zz}$. The latter is a function of both the symmetry and density of the

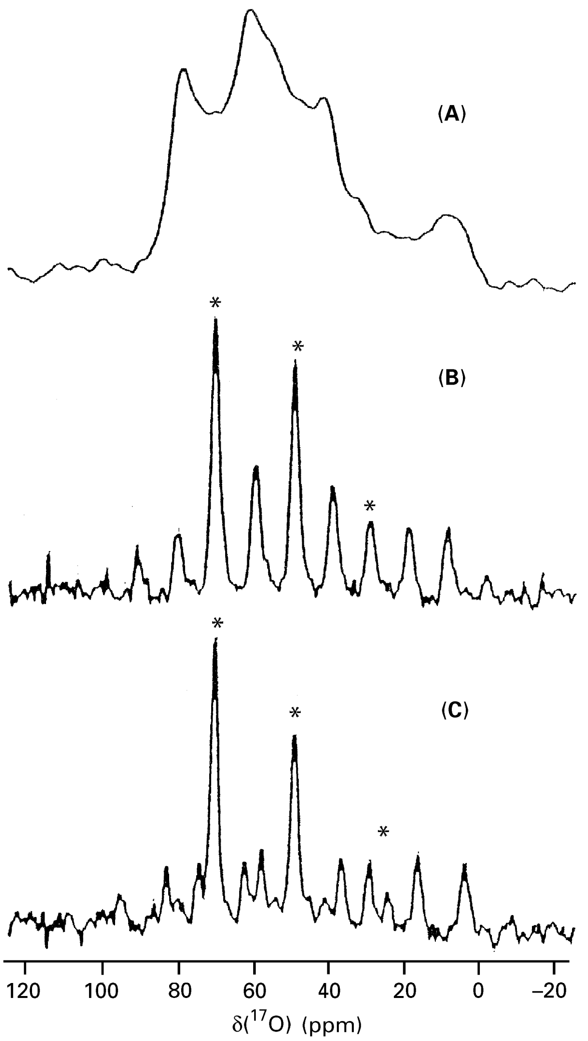



**Figure 11** Solid-state $^{17}O$ NMR spectra of a sample of the mineral diopside, $CaMgSi_2O_6$: (A) using only high-speed MAS, (B) using DOR with a speed of 540 Hz around the second magic-angle of 30.6°, and (C) as in (B), but with a speed of 680 Hz. There are three different oxygen sites in the crystal structure, corresponding to the three $^{17}O$ lines marked with asterisks. The latter are identified in spectra (B) and (C) since their positions are not affected by the spinning speed. Reproduced with permission of Macmillan Magazines Ltd. from Chmelka BF, Mueller KT, Pines A, Stebbins J, Wi Y and Zwaziger JW (1989) Oxygen-17 NMR in solids by dynamic-angle spinning and double rotation. *Nature* **339:** 42–43.

electron distribution, i.e. on the molecular structure. For reasons discussed above, the focus has been restricted on the central transition of half-integer nuclei (which make up almost one-third of all NMR active nuclei). In this regard, $^{23}Na$ and $^{27}Al$ are the most studied, but interest has also been paid to $^{11}B$, $^{17}O$, $^{51}V$ and $^{7}Li$ (see **Table 5** for nuclear and field gradient properties).

A great deal of attention has been paid to $^{27}Al$, owing to its importance in the preparation and dealumination of zeolites, and in the study of inorganic materials and catalysts. Since the values of $\chi$ are relatively low (**Table 5**), the second-order quadrupole shift is small and therefore relatively narrow signals are obtained. Furthermore, the chemical shift range spanned by distinctly coordinated aluminium sites is large (**Table 6**), allowing not only the distinction of Al sites, but also environments within like sites which may differ in bond distances and angles or in hydrogen bonding. In fact, useful correlations between $^{27}Al$ chemical shift or quadrupole parameters and structural features have been found, such as (1) $\sigma_{iso}(^{27}Al)$ varies linearly with Al–O–Al angles in aluminates and with Al–O–Si angles in aluminosilicates, and (2) $\chi(^{27}Al)$ is linearly correlated with the distortion $|\psi| = \sum|\tan(\alpha_i) - 109.48°|$ of the $AlO_4$ tetrahedron.

By studying other quadrupolar nuclei, interesting structural details have been obtained on a number of inorganic solids of undoubted technical importance, such minerals, zeolites, catalysts, ceramics, glasses and cements (**Table 6**). Owing to the technological significance of the studied materials, it is likely that this area of solid-state NMR will experience great progress in the years ahead.

## List of symbols

$B_0$ = magnetic flux density; $D$ = dipolar coupling constant; $D'$ = effective dipolar coupling constant; $h$ = Planck's constant; $I$ = spin-$\frac{1}{2}$ nucleus; $J$ = coupling constant; $q$ = field gradient tensor; $Q$ = nuclear quadrupole moment; $s'$ = doublet splitting;

**Table 6** Chemical shift range in solid materials studied by solid-state NMR of quadrupolar nuclei

| *Nucleus* | *Chemical shift range* (ppm) | *Usual reference* | *Studied materials* |
|---|---|---|---|
| $^{23}Na$ | −50–50 | NaCl (1 M) | Na oxides and salts, zeolites |
| $^{27}Al$ | 150–200 (tetracoord.)<br>20–80 (pentacoord.)<br>−20–50 (hexacoord.) | $AlCl_3$ (1 M) | Zeolites, molecular sieves, catalysts, Al oxides, aluminates, Al glasses and ceramics |
| $^{11}B$ | −2–2 (tetracoord.)<br>10–30 (tricoord.) | $BF_3{\cdot}Et_2O$ | B glasses, oxides, borates |
| $^{17}O$ | 0–500 | $H_2O$ | Oxides, oxoanions, metal carbonyls |
| $^{51}V$ | −100–1000 | $VOCl_3$ | V oxides, vanadia catalysts, vanadates |
| $^{7}Li$ | −5–5 | LiCl (1 M) | Li glasses, oxides |

S = quadrupolar nucleus; $\gamma$ = magnetogyric ratio; $\theta$ = angle between main tensor axes; $\delta$ = chemical shift; $\tau_1$ = relaxation time; $\xi$ = angle between main axes of interaction tensors and sample spinning axis; $\chi$ = quadrapole coupling constant.

*See also:* **High Resolution Solid State NMR, $^{13}$C; NMR in Anisotropic Systems, Theory; NMR of Solids; Solid State NMR, Methods; Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules.**

## Further reading


Abragam A (1989) *Principles of Nuclear Magnetism*. Oxford: Oxford University Press.

Alarcón SH, Olivieri AC, Carss SA and Harris RK (1994) Effects of $^{35}$Cl/$^{37}$Cl, $^{13}$C residual dipolar coupling on the variable-temperature $^{13}$C CP/MAS NMR spectra of solid, chlorinated sodium acetates. *Angewandte Chemie, International Edition in English*: **33**: 1624–1625.

Davies NA, Harris RK and Olivieri AC (1996) The effects of interplay between quadrupolar, dipolar and shielding tensors on magic-angle spinning NMR spectra: shapes of spinning sidebands. *Molecular Physics* **87**: 669–677.

Fyfe CA (1983) *Solid State NMR for Chemists*. Ontario: CFC Press.

Grondona P and Olivieri AC (1993) Quadrupole effects in solid-state NMR spectra of spin-$\frac{1}{2}$ nuclei: a perturbation approach. *Concepts in Magnetic Resonance* **5**: 319–339.

Harris RK (1986) *Nuclear Magnetic Resonance Spectroscopy. A Physicochemical View*. New York: Longman Scientific & Technical.

Harris RK (1996) Nuclear spin properties & notation. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance* Vol 5, pp 3301–3314. Chichester: Wiley.

Harris RK and Olivieri AC (1992) Quadrupole effects transferred to spin-$\frac{1}{2}$ magic-angle spinning spectra of solids. *Progress in Nuclear Magnetic Resonance Spectroscopy* **24**: 435–456.

Lucken EAC (1969) *Nuclear Quadrupole Coupling Constants*. London: Academic Press.

Mason J (ed) (1987) *Multinuclear NMR*. New York: Plenum Press.

Mehring M (1983) *High Resolution NMR in Solids*, 2nd edn. Berlin: Springer-Verlag.

Wasylishen RE and Fyfe CA (1982) High resolution NMR of solids. *Annual Reports in NMR Spectroscopy* **12**: 1–80.

# Solid State NMR, Methods

**JW Zwanziger**, Indiana University, Bloomington, USA
**HW Spiess**, Max-Planck-Institut für Polymerforschung, Mainz, Germany

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Methods & Instrumentation

As in NMR of liquid samples, solid state NMR probes the magnetic interactions of atomic nuclei. These interactions yield detailed information about the local structure and dynamics of the sample, including the bonding types and geometry, the site–site connectivity patterns, and the spatial characteristics and timescales of atomic and molecular motions. All kinds of solids can be studied with NMR, including single crystals and powders, disordered materials such as glass and rubber, and metals and superconductors. Although not as high as in liquid state NMR, spectral resolution is still extraordinary (parts per million or better) but sensitivity is not. Sample volumes of order 100 μL are typical.

The magnetic interactions probed in solid state NMR include those studied in the liquid state, beginning with the Zeeman interaction between the nuclear spin and the applied magnetic field. This induces precession at the Larmor frequency $\omega_0$, which is defined by the nucleus and the strength of the external field. Fields as high as 17.5 T are in use, yielding proton Larmor frequencies of 750 MHz. Internal interactions observed include the chemical shift, and, in favourable cases, scalar couplings. Additionally, magnetic dipole and electric quadrupole interactions, which are observable only indirectly in liquid state NMR spectra, can be detected as frequency shifts in the solid state. In metals, the major spectral observable is the Knight shift. Because all these interactions depend sensitively on the local bonding geometry, they can be used to measure dynamic properties of the sample, either directly through spectral changes as a function of experimental parameters, or indirectly through the nuclear spin relaxation time.

The primary difference between solid state and liquid state NMR is one of timescale. The atomic dynamics of the sample define a natural internal timescale, denoted $\tau$. The motion of interest might be, for example, the rotational tumbling of molecules in a liquid, the reorientation of segments in a polymer, or the hopping of ions in a solid electrolyte. Clearly $\tau$ can range from picoseconds to seconds or more. As the observed nucleus moves to different locations or orientations, its NMR spectrum changes. This occurs both because the different sites may differ chemically, and also because the observable interactions are orientation-dependent, the sense of orientation being defined by the external magnetic field. The different orientations define a range of frequencies $\Delta\omega$ centred on the Larmor frequency. If $\Delta\omega\tau \ll 1$, then a given nucleus samples many local environments during the NMR experiment. All interactions that depend on molecular orientation are in this way averaged to zero, and the spectrum is ‘liquid-like’. ‘Solid-like’ spectra result when $\Delta\omega\tau \gg 1$, for then the anisotropic portions of the interactions remain. These include for example the above-mentioned magnetic dipole and electric quadrupole terms. Typical examples of both regimes are shown in **Figure 1**.

**Figure 1** also shows the principal difficulty encountered in solid state NMR spectra: the additional information provided by the anisotropic interactions can seriously congest the spectrum, making interpretation difficult. Our aim in this article is to outline the current principal methods by which solid state NMR spectra can be acquired in interpretable form. Rather than giving an exhaustive account of the current developments in the field, we present the most important techniques in the context of the physical and chemical problems that they can help to solve.

## Resolving chemically distinct sites

The most frequent application of solid state NMR, as in the liquid state, is resolution of chemically distinct sites in a material. However, as **Figure 1** shows, the anisotropy observed in solid spectra typically create so much spectral congestion that assignment is difficult. Moreover, as **Figure 1** also shows, the anisotropically broadened lines exhibit a variety of step and singularity features which, while informative in their own right, further obstruct a rapid assessment of the types and relative concentrations of distinct sites. The most important method for improving resolution in solid state NMR is magic angle spinning (MAS). This method, so-called because the sample is rotated about an axis inclined at the magic angle of 54.74° with respect to the

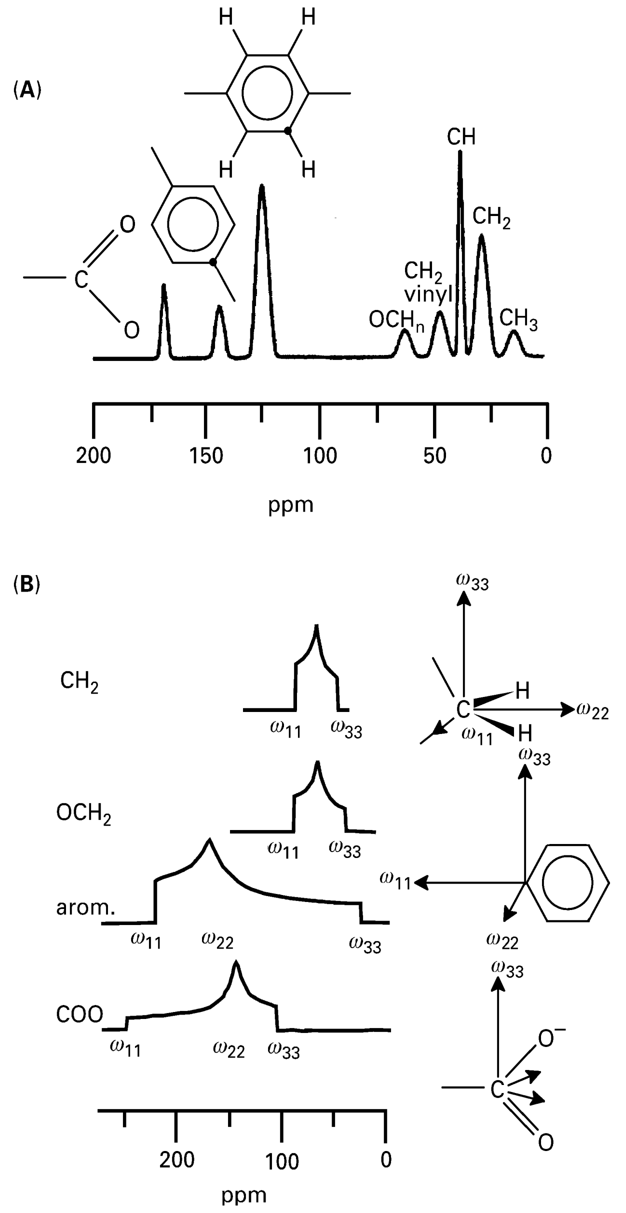



**Figure 1** (A) Typical chemical shifts of carbon in different functional groups. The line widths indicate the ranges observed; in a liquid sample, the actual line width will typically be much smaller. (B) Chemical shift powder patterns of carbons in the same functional groups. Such shapes are observed in powdered solids. The complex shapes, with steps and singularities, arise from the nontrivial orientation dependence of the chemical shift interaction, which is averaged to zero in a liquid but is observable in solids. The powder patterns are shown separately here for convenience; in real samples they overlap, making interpretation difficult. This problem is addressed by techniques such as magic angle spinning. Figure adapted from Schmidt-Rohr K and Spiess HW (1994) *Multidimensional Solid-State NMR and Polymers*. London: Academic Press, 1994.

magnetic field, can enhance the resolution by more than 2 orders of magnitude. It arises because the dominant anisotropies transform under rotations as 2nd rank spherical tensors, i.e. like d-orbitals. Recall that the $d_z{}^2$ orbital has an angular node; this is in fact at 54.74°. Thus an interaction which transforms in the same way can be averaged to zero by spinning about an axis located at this node.

To achieve effective averaging, the rotation frequency $\omega_r$ in MAS must be of the order of, or greater than, the spread of interaction frequencies: $\omega_r/\Delta\omega \gtrsim 1$. The resulting resolution enhancement is dramatic, as shown in **Figure 2**.

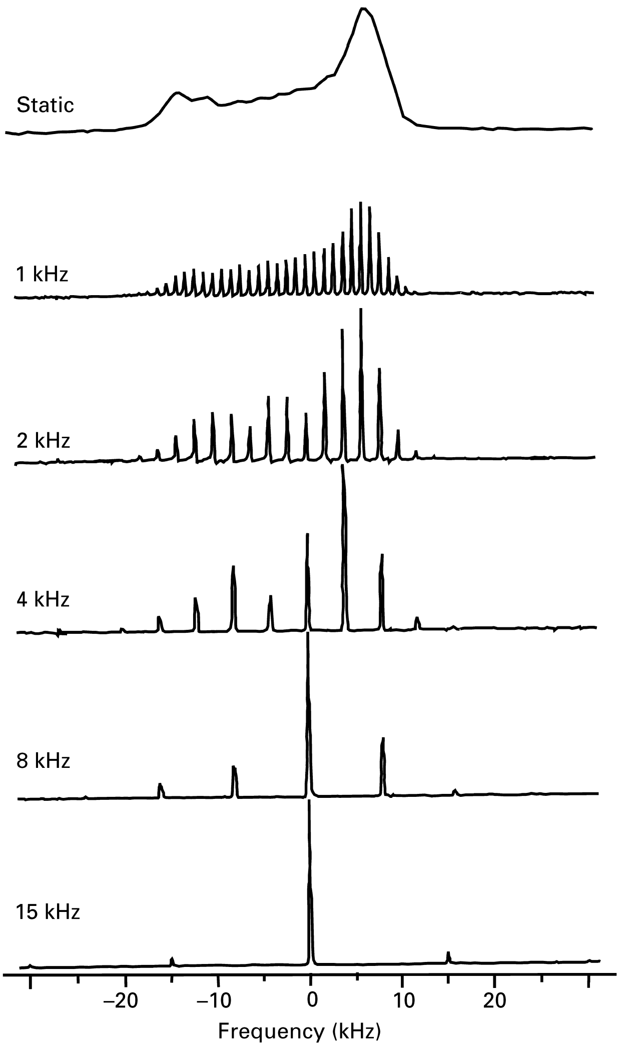



**Figure 2** The effect of magic angle spinning. The figure shows $^{31}P$ spectra of $Na_4P_2O_7{\bullet}10H_2O$, as a function of rotor frequency. Note the extreme line-narrowing achieved, while still in a powdered solid. This illustrates that the symmetry of the chemical shift interaction is such that the full isotropic averaging of the liquid state is more than necessary to suppress the anisotropy; rotation about a single axis is in this case sufficient. The small splittings show that crystallographically, as well as magnetically, different sites can be resolved. Figure adapted from Schnell I, Diploma Thesis, Johannes-Gutenberg-Universität Mainz, 1996; see also Kubo A and McDowell CA (1990) *Journal of Chemical Physics* **92**: 7156.

For nuclei like $^{13}C$, $^{31}P$, and $^{29}Si$, which have modest chemical shift ranges, spinning frequencies of 5–10 kHz are often sufficient, and are well within range of typical commercial MAS probes. $^{1}H$ spectroscopy is particularly challenging in solid state NMR, in contrast to liquids, because of the strong magnetic dipole coupling between protons. This interaction gives a proton line width typically in the range of 20–50 kHz. Commercial MAS NMR probes are now available with spinning frequencies as high as 35 kHz, and so in many cases even $^{1}H$ solid state NMR can be accomplished with the MAS technique.

While MAS can provide significant resolution enhancement, it enhances sensitivity only insofar as the signal from broad resonances is concentrated into narrower resonances. For naturally low-abundance nuclei like $^{13}C$ (1% naturally occurring), this increase may be insufficient. For dilute spins in the presence of an abundant species with good sensitivity (such as protons), i.e. nearly all organic solids, double resonance methods may be used to achieve an additional gain in sensitivity. Coupled with MAS, these techniques are collectively referred to as CP-MAS (CP = cross-polarization). In CP-MAS, magnetization is first excited using the abundant species (typically $^{1}H$), and then transferred to the dilute species ($^{13}C$ or $^{15}N$, say) by simultaneously irradiating both nuclei, at their respective Larmor frequencies. Then, the dilute spin is detected, often with decoupling of the abundant spin. All this is carried out in the presence of MAS, to obtain good resolution of the resulting spectrum. The theoretical sensitivity gain is the ratio of the Larmor frequencies of the two species, for example, 4 for the $^{1}H$–$^{13}C$ pair. In practice, protons often have significantly shorter relaxation times than their CP partners, so this method also allows for shorter recycle delays in pulsed NMR, and thus more rapid acquisition of the spectrum. A disadvantage of CP-MAS is that the CP efficiency is a function of proximity of the abundant and dilute species, so that CP-MAS spectra cannot be assumed to be quantitative reflections of the abundances of the resolved sites.

While MAS and CP-MAS are often sufficient to resolve chemical sites for nuclei like $^{13}C$, $^{31}P$ and $^{29}Si$, this is not the case for other nuclei such as $^{27}Al$, $^{17}O$ and $^{11}B$. The reason is that the first group of nuclei have spin $\frac{1}{2}$, while the second have higher spin. Nuclei with spin greater than $\frac{1}{2}$ are subject to electric quadrupole effects, in addition to chemical shift, and these effects present a significant additional source of line-broadening. Because the quadrupole anisotropy in the presence of a strong magnetic field does not transform simply like a 2nd rank tensor, it cannot be removed completely by MAS alone. During the last 10 years significant progress has been made in devising methods to average quadrupole interactions in addition to chemical shift anisotropy. Double rotation (DOR) and dynamic angle spinning (DAS) both make use of spinning the sample about a time-dependent axis. DOR is a direct extension of MAS, and makes use of a complex rotor-within-a-rotor device. In DAS, the sample spinning axis is hopped between two angles during the experiment. Both DOR and DAS require mechanically sophisticated probes. A third method, multiple-quantum magic angle spinning (MQ-MAS), uses just a MAS probe, but a complex pulse sequence to excite and detect triple and higher order coherences during the MAS experiment. It is mechanically the simplest to implement, but uniform excitation of different chemical sites is difficult with existing pulse sequences. Despite the limitations for each method mentioned above, they have yielded impressive resolution advances for quadrupolar nuclei (**Figure 3**), similar to what is achieved with MAS for nuclei such as $^{13}C$.

The methods described above yield greatly enhanced resolution, and as such help to determine the types and amounts of distinct sites in material. This resolution gain comes at the price of discarding the information available from the interaction anisotropies. This information typically relates to the local site symmetry, for example, distortions in the bond angles, number of nearest neighbours, and so forth. Both high-resolution and interaction anisotropies may be obtained by taking advantage of a second spectral dimension, using so-called 'separation of interactions' experiments. An example is shown in **Figure 4**, where it is seen that the anisotropically broadened resonances are sorted according to their isotropic shift. In this way each resonance may be examined in isolation, and the anisotropy parameters determined with no congestion from neighbouring bands. There are many ways to implement such experiments; an elegant approach for simple chemical shift correlations is to spin the sample at an angle other than the magic angle during the first part of the experiment, followed by a hop to the magic angle and subsequent signal acquisition. In this way the anisotropic interactions 'label' the detected signal, and a double Fourier transform gives the type of spectra shown in **Figure 4**.

## Determining the connectivity between sites

Once the types of sites in the material have been determined, using for example the techniques discussed

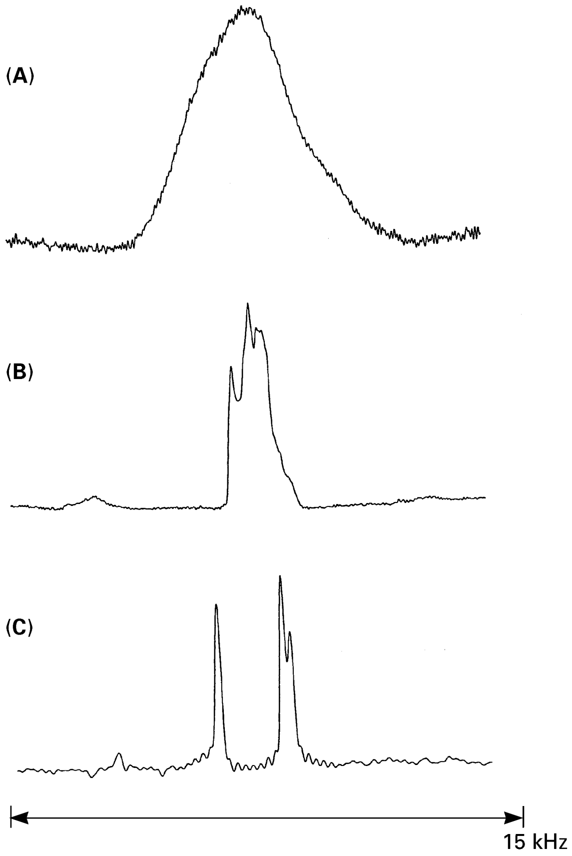



**Figure 3** $^{87}Rb$ NMR spectra of $RbNO_3$. (A) The static spectrum, (B) The effect of magic angle spinning. The symmetry of strong quadrupole interactions is such that spinning about a single axis alone is not sufficient to remove all the anisotropy broadening. (C) Results of multiple-quantum magic angle spinning, one of several methods currently available to obtain high-resolution spectra of quadrupolar nuclei like $^{87}Rb$. In this spectrum, the three crystallographically distinct rubidium sites are resolved, and the total line-narrowing is comparable to that achieved by MAS alone for spin-$\frac{1}{2}$ nuclei like $^{13}C$ and $^{31}P$ (see **Figure 2**). Figure adapted from Brown S, D. Phil. Thesis, Oxford University, 1998, and Brown S and Wimperis S (1997) *Journal of Magnetic Resonance* **128**: 42–61.

above, the second step in determining the material structure can be considered, namely what the connectivity between these sites is. In liquid state NMR connectivities are primarily determined through scalar couplings, a through-bond interaction mediated by the electrons. However, this interaction is very small compared to the anisotropies of the magnetic interactions, and so is hard to probe in solids in any but the most well-ordered samples. On the other hand, magnetic dipole interactions, i.e. the through-space effects of the nuclear magnetic moments on each other, can be substantial. This interaction varies with distance as $r^{-3}$, and so is of particular use in determining local structure. As noted above, dipole couplings are averaged to zero in ‘liquid-like’ spectra, but in solids they have easily observable effects on the resonance line shapes. Such interactions are also observed *indirectly* in liquids, through the nuclear Overhauser effect where they appear as second-order interactions and thereby survive the averaging due to the molecular motion. Knowledge of the through-space connectivities does not give directly a map of the bonding network but, when combined with knowledge of the material composition and chemistry, can yield much about the bonding pattern.

For example, the CP-MAS experiment described above can already be used to obtain some degree of through-space information. The magnetization transfer, from $^1H$ to $^{13}C$ say, is mediated by the magnetic dipole interactions. By varying the duration of the transfer time (usually called the contact pulse), sites can be distinguished by their transfer efficiency. For example, primary and secondary carbons can be selectively excited, relative to tertiary and quaternary carbons due to their greater proximity to protons. This is done simply by using a short contact pulse. Much more elaborate spectral editing schemes yield more accurate results, and are more flexible.

While variants of CP-MAS are particularly suitable for exploring proximities in heteronuclear systems with protons as one partner, other experiments can be performed to probe other heteronuclear systems, and homonuclear couplings. All make use of the dipole coupling as the mechanism for encoding distance information.

In general, one wants to combine the distance measurement with some kind of resolution enhancement, in order to determine which sites are close to which. This is not always possible. For example, in many inhomogeneous solids, only one (broad) resonance will be observed, even with MAS or similar techniques. A well-studied example is sodium in a glass. Because techniques like MAS average the dipole coupling to zero, if they do not provide sufficient resolution enhancement, they should not be used. Then, one studies a static sample. The spatial distribution of species in a static sample can be estimated, by measuring the decay properties of spin echoes.

In spin-echo experiments, an excitation pulse is followed at some time $\tau$ later by a refocussing pulse. At time $\tau$, after the second pulse, an echo will typically form. The typical use of this experiment for measuring distances is to estimate the so-called second moment ($M_2$) of the resonance line. Interactions on a local scale, e.g. the chemical shift and quadrupole interactions, and interactions involving isolated pairs of spins, can be refocussed by using suitably chosen pulses. However, dipole coupling to a bath of partners cannot. Therefore, the echo cannot be

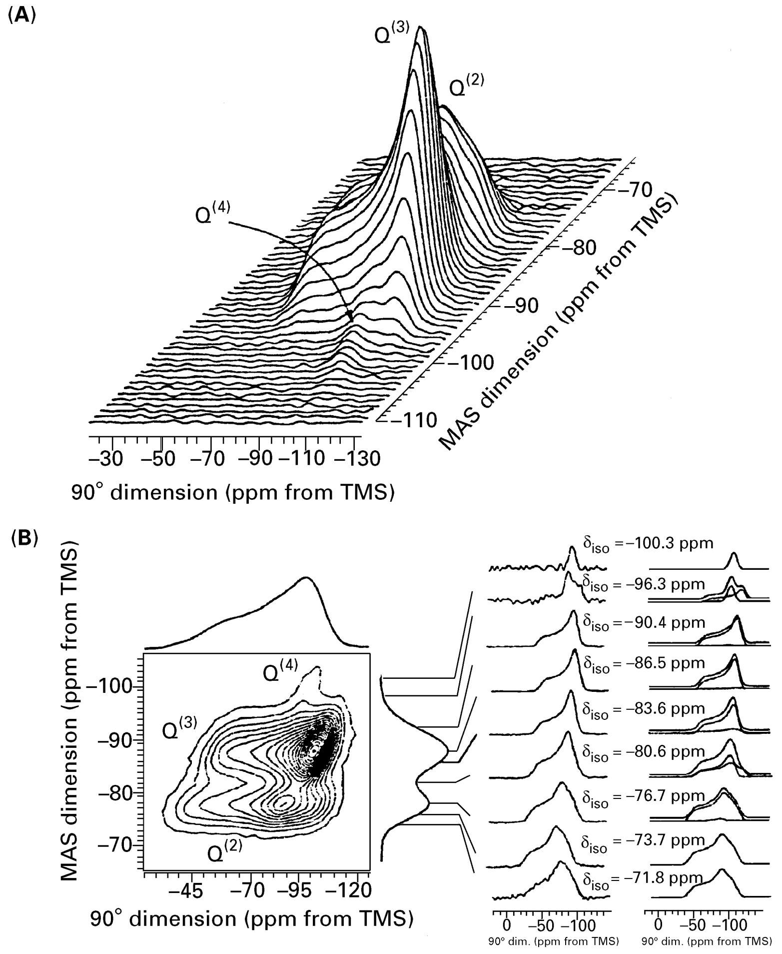



**Figure 4** A separation of interactions-type spectrum of $^{29}Si$ in a glass. (A) By spinning the sample off the magic angle during the first part of the experiment, and on the magic angle in the second, a two-dimensional spectrum is generated that has a high-resolution dimension correlated with the anisotropies of the individual sites. (B) Here, in a glass, each site itself shows a distribution of environments, which can be mapped out quantitatively by taking slices through the two-dimensional spectrum. In this way, bond angle distributions for example, even in complex materials, can be determined, often with superior precision as compared to diffraction-based methods. Figure adapted from Zhang P *et al.* (1996) *Journal of Non-Crystalline Solids* **204**: 294–300.

refocussed indefinitely, but only up to a characteristic time which is a measure of properties of the bath of nuclei coupled to the studied species. The decay constant of the spin-echo envelope is proportional to $M_2$, which is given essentially by summing over $r_{ij}^{-6}$, where the $r_{ij}$ are internuclear distances. Because of the exponent –6, this experiment gives short-range information. It is valuable in assessing qualitative features of the distribution of species in inhomogeneous materials. An important extension of this experiment is called SEDOR, for spin-echo double resonance, in which an additional refocussing pulse is applied to a second nuclear species, and the echo behaviour with and without this secondary pulse are compared. In this way the mixing of different species in an inhomogeneous solid may be assessed.

Similar to SEDOR, but appropriate for isolated pairs of spins, is the rotational echo double resonance method, or REDOR. In REDOR, the combined dynamics of the isolated two-spin system and the sample rotation serve to generate echos at the rotor period. These echoes can be dephased by application of a pulse to one of the coupled partners. The amount of dephasing caused by this additional pulse is a measure of the coupling strength, and hence proximity, of the spins in the pair. This experiment is most applicable to doubly labelled samples, e.g. biopolymers enriched at selected sites with $^{13}C$ and $^{15}N$.

When the spectrum of the material consists of resolved resonances, much more detailed information on the nuclear distances can be derived than is possible with the spin-echo techniques outlined above. Magnetic dipole coupling is still the interaction to probe but, when the various sites are resolved, experiments can be used that give signals *only* if two distinct sites are near enough to each other to have a

significant interaction. Clearly this yields much more detailed information than when the sites serve primarily to generate a background bath. Several types of signals can be generated and measured in this context, but the most precise are the so-called double-quantum coherences. Isolated nuclei (here we have in mind only spin-$\frac{1}{2}$, such as $^{1}$H or $^{13}$C) do not have enough energy levels to support quantum number changes greater than unity, and therefore also cannot support coherences greater than unity. If two such spins are coupled, however, the composite system can support 2-quantum coherence. Experiments can be designed that are selective only for 2-quantum coherence, thus yielding a connectivity map of the sites that are close enough spatially to couple in this way.

Generating such a connectivity map for a solid requires additionally a resolution-enhancement technique, such as MAS. Such techniques, as discussed previously, suppress precisely the spin–spin interactions that the connectivity map is meant to reveal. Therefore, to combine multiple-quantum experiments with MAS, a pulse sequence which counteracts the averaging effect of MAS, thereby restoring the dipole coupling, must be implemented during excitation and reconversion of the 2-quantum coherence. A variety of such dipolar recoupling sequences currently exist, of varying levels of performance and complexity.

The resulting two-dimensional spectra give a connectivity map that can be traced in much the same way as is routinely done for liquid samples (**Figure** 5).

It must be remembered, of course, that the signals observed reflect coupling through space, not through chemical bonds, so additional information about the chemistry must be used to interpret such spectra. Nevertheless, this method is fast becoming routine, as it requires only standard solids NMR instrumentation.

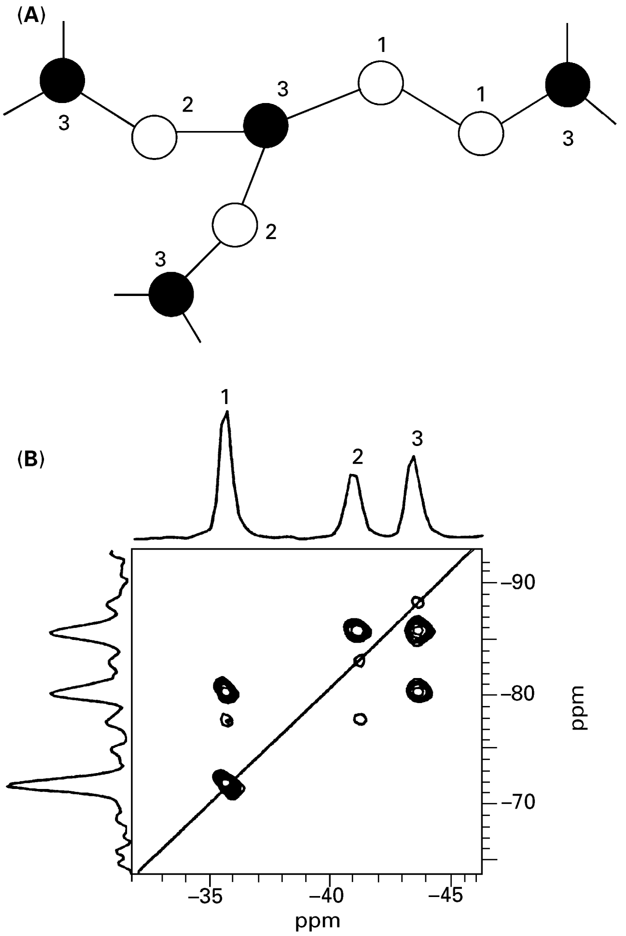



**Figure 5** Double-quantum correlation spectrum of $^{31}$P in a solid phosphate. At top (A) is the molecular fragment derived from assignment of the spectrum (oxygen atoms not shown), and includes phosphate chain branch points (site 3), branch connections (site 2) and chain phosphates (site 1). The spectrum (B) gives signals symmetric across the diagonal, for pairs of sites that are close enough in space to be coupled. Thus from such a spectrum the spatial proximity of resolved sites can be traced, as is done routinely in liquid NMR using scalar couplings (a through-bond interaction). Figure adapted from Feike M *et al.* (1996) *Journal of the American Chemical Society* **118**: 9631–9634.

## Dynamics in solids

The dynamics of atoms in solids may be probed directly, through their effects on the NMR spectra, and indirectly, through the nuclear spin relaxation. Because the NMR signal is observed only after the nuclear magnetization has been perturbed from its equilibrium state, relaxation is a standard feature of all NMR experiments. Two primary relaxation processes are usually identifiable. The first is the relaxation of the total magnetization back to its thermal equilibrium value; this occurs on a timescale denoted $T_1$. The second is the timescale for relaxation of quantum coherences in the spins, and is denoted $T_2$. In liquids, due to the strong decoupling resulting from the molecular motion, these two processes occur on similar timescales. In solids, however, they are usually very different, with $T_2$ ranging typically from $10^{-4}$ to $10^{-2}$ s, and $T_1$ from $10^{-3}$ to $10^{3}$ s. Relaxation occurs because fluctuations in the surroundings of a spin induce transitions within the spin quantum states. Therefore, measurements of relaxation times are indirect probes of the dynamics in the solid. However, identifying what sort of fluctuation is operative is usually very difficult, unless one type of interaction is clearly dominant (conduction electrons in a metal or superconductor is a good example). Otherwise, the best that can be done is to estimate the temperature and magnetic field dependence of the relaxation to be expected from candidate

fluctuation modes, and to compare the predictions with the data.

An easier qualitative assessment of dynamics can often be obtained from resonance line shapes. As noted above, the key distinction between solid-like and liquid-like NMR spectra is the timescale of the atomic motions, compared to the frequency spread of the detected interactions. As the rate of a dynamic process increases, say as a function of temperature, it can be followed through changes in the line widths of the nuclei involved. These changes can be substantial, as a resonance goes from solid-like at low temperatures to liquid-like at high temperatures, with the line width decreasing by orders of magnitude.

The line width strategy is particularly effective if no additional line-narrowing is needed to interpret the low-temperature spectra; however, it is often the case that these spectra will be so congested that additional techniques such as MAS must be applied. In that case, the line-narrowing caused by heating the sample is much less dramatic. In this case, two-dimensional spectroscopy can again be very helpful. For organic solids and polymers, the wideline separation of interactions (WISE) experiment is a convenient qualitative measure of the relative site dynamics. This experiment combines the good resolution found in $^{13}C$ spectra under MAS, with the strong inter- nuclear coupling of protons. The latter feature makes proton spectra particularly good indicators of motional narrowing due to dynamics: broad proton resonances (30–50 kHz) are seen in static samples, and narrow (< 1 kHz) for mobile sites. The WISE experiment works by adding an additional evolution time to the CP-MAS sequence, between the proton excitation and the contact pulse to the carbons, and using relatively slow sample spinning. In this way, a two-dimensional spectrum is obtained, with proton resonances sorted by the carbon sites to which they are bonded (**Figure 6**). One can immediately see, therefore, which carbon sites are mobile (narrow associated proton resonances) and which are static (broad proton resonances).

More detailed information on dynamics is available from so-called exchange experiments. This class of two-dimensional technique provides a correlation between spectral components, which exchange during a mixing period. The exchange may occur because of a chemical transformation during the mixing time, resulting in a new frequency due to a new chemical environment, or because of site reorientation. Since the nuclear spin interactions are orientation-dependent, if the molecular unit changes its orientation during the mixing time, the involved nuclear spins will exhibit altered NMR frequencies, which can be correlated to their initial values.

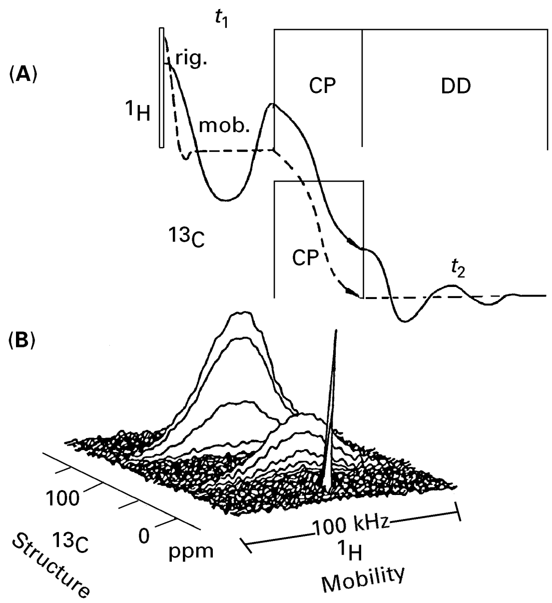



**Figure 6** Two-dimensional wideline separation of interactions (WISE) spectrum of polystyrene-poly(dimethyl siloxane) diblock copolymer (PS-b-PDMS). In the first part of the experiment (A) proton magnetization is allowed to evolve, and then transferred to carbon in the second part of the experiment. In this way the proton spectra of individual carbon sites are sorted by the shift of each site, and the result is a wide-line proton spectrum in one dimension, and a high-resolution MAS carbon spectrum in the other. In this example, (B), the PDMS is seen to be quite mobile: it gives the carbon signal near 0 ppm, and the proton spectrum for this site is very sharp, indicating significant motional narrowing. The PS peaks, on the other hand, give very broad proton resonances, showing that the PS part of this block copolymer is essentially static at this temperature. With this experiment, quick qualitative assessments of relative local mobility in organic solids and polymers can be made. Figure adapted from Schmidt-Rohr K and Spiess HW (1994) *Multidimensional Solid-State NMR and Polymers*. London: Academic Press.

Because the orientation dependences of NMR interactions are well known, it is often a straightforward matter to relate the observed exchange spectrum to the underlying molecular motion that gave rise to it. In this way, very detailed information on microscopic molecular dynamics can be obtained.

Deuterium NMR spectra provide particularly clear examples of the above approach. For the deuterium nucleus, the quadrupole interaction is dominant, by far and in a C–D bond, is aligned with the C–D bond itself. Therefore, changes in time of the deuterium quadrupole orientation give a direct reflection of the orientational dynamics of the C–D bond itself. **Figure 7** shows the two-dimensional exchange spectrum of deuterated dimethyl sulfone ($(CD_3)_2SO_2$). The strong diagonal ridge is a typical deuterium NMR spectrum, and arises from molecules that did not exchange during the mixing time. The pattern of ellipses off the diagonal arises due to

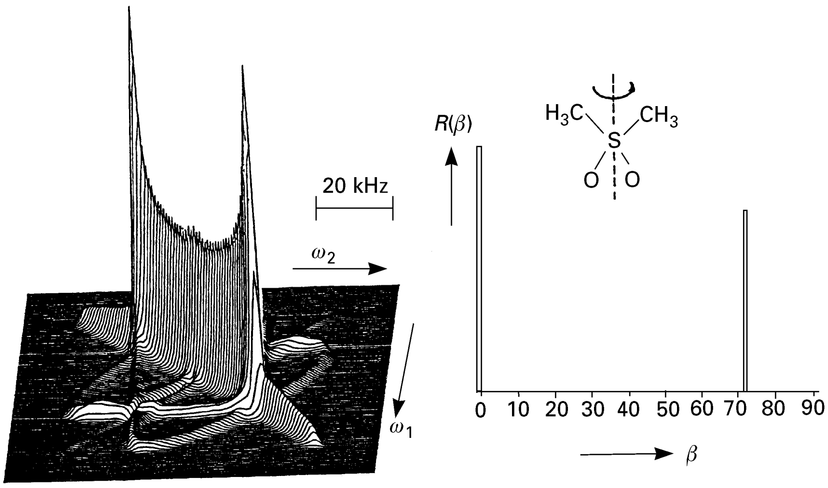



**Figure 7** Two-dimensional $^2$H exchange spectrum of deuterated dimethylsulfoxide. The strong diagonal ridge reflects molecules that have not reoriented during the mixing time, while the pattern of ellipses off the diagonal shows those that have. The ellipses arise due to the orientational dependence of the $^2$H quadrupole interaction, the dominant anisotropy here. The distribution of jump angles is shown on the right, and sharply peaked at zero (static molecules) and 72°, the included angle of the C–D bonds as the entire molecule executes hops about its symmetry axis. With this type of experiment, slow to moderate dynamics of molecules and polymers can be followed in atomic-level detail. Figure adapted from Schmidt-Rohr K and Spiess HW (1994) *Multidimensional Solid-State NMR and Polymers*. London: Academic Press.

deuterium nuclei with one orientation, and hence one frequency, at the start of the experiment, and a second orientation, hence frequency, after the mixing time. This pattern is consistent with 180° jumps of the molecules about their symmetry axis. The exchange experiment can be applied to other nuclei as well, such as $^{13}$C, although it can be harder to relate the spin interaction orientation to a molecular frame of reference.

## Summary

In this article we have attempted to provide a brief overview of modern techniques and their applications in solid state NMR. Far from being exhaustive, we hope instead to have informed the reader about the types of problems that can be investigated fruitfully with this approach, using what have become standard methods. There are many other more specialized techniques, suitable for particular problems, which are described in the current literature. The following bibliography is meant to provide a starting point for newcomers to the field. The 'Further reading' section provides entry into the technical primary literature.

## List of symbols

$M_2$ = second moment; $r_{ij}$ = internuclear distances; $T_1$ = relaxation time of total magnetization; $T_2$ = relaxation time of quantum coherences; $\tau$ = timescale; $\omega_0$ = Larmor frequency; $\omega_r$ = rotation frequency.

*See also:* **$^{13}$C NMR Methods; Chemical Exchange Effects in NMR; Chiroptical Spectroscopy, Orientated Molecules and Anisotropic Systems; Heteronuclear NMR Applications (Ge, Sn, Pb); Heteronuclear NMR Applications (O, S, Se, Te); Liquid Crystals and Liquid Crystal Solutions Studied By NMR; Magnetic Resonance, Historical Perspective; NMR Data Processing; NMR in Anisotropic Systems, Theory; NMR Relaxation Rates; NMR Spectroscopy of Alkali Metal Nuclei in Solution; $^{31}$P NMR; Parameters in NMR Spectroscopy, Theory of; Product Operator Formalism in NMR; Relaxometers; Xenon NMR Spectroscopy.**

## Further reading

Abragam A (1961) *Principles of Nuclear Magnetism*. Oxford: Clarendon Press.

Blümich B (ed) (1994) Solid State NMR I–IV, vols. 30–33 of *NMR Basic Principles and Progress*. Diehl P, Fluck E, Günther H, Kosfeld R, and Seelig J (eds) Berlin: Springer-Verlag.

Fukushima E and Roeder SBW (1981) *Experimental Pulse NMR: A Nuts and Bolts Approach*. London: Addison-Wesley.

Fyfe C (1983) *Solid State NMR for Chemists*. Guelph: CFC Press.
Harris RK and Grant DM (1996) *Encyclopedia of Nuclear Magnetic Resonance*. Chichester: Wiley.
Mehring M (1983) *High Resolution NMR in Solids*, 2nd edn. Berlin: Springer-Verlag.
Slichter CP (1983) *Principles of Magnetic Resonance*, 3rd edn. Springer: Berlin.
Schmidt-Rohr K and Spiess HW (1994) *Multidimensional Solid-State NMR and Polymers*. London: Academic Press.
Traficante DD (ed) *Concepts in Magnetic Resonance, An Educational Journal*. New York: Wiley.

# Solid-State NMR, Rotational Resonance

**David L Bryce** and **Roderick E Wasylishen**,
Dalhousie University, Halifax, Nova Scotia, Canada

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

One of the primary goals of solid-state NMR spectroscopists has been to develop techniques that yield NMR spectra of solid samples with resolution approaching that observed for samples in isotropic liquids. Rapidly spinning samples about an axis inclined at the magic angle ($arc\cos(1/\surd 3) = 54.7356°$) relative to the applied static magnetic field has been found to be highly effective in this regard. In addition, high-power decoupling of abundant spins (e.g. $^1H$) eliminates heteronuclear spin–spin coupling interactions (direct dipolar and indirect *J*-coupling) involving the abundant spins when dilute spins are examined. The availability of commercial NMR instrumentation that permits users to apply these two techniques has contributed to spin-$\frac{1}{2}$ NMR becoming a routine method for examining a wide range of solid materials. Finally, cross-polarization (CP) from abundant spins to dilute spins has been important in improving the sensitivity of the dilute-spin NMR experiment.

Ironically, it is sometimes desirable to selectively reintroduce interactions which are effectively averaged in the magic-angle-spinning (MAS) experiment. Dipolar coupling, for instance, may be recovered in the form of the direct dipolar coupling constant ($R_{DD}$) between isolated spin pairs. The value of $R_{DD}$ is of interest due to its simple relationship with the distance separating two spins, $r_{12}$ (Eqn [1])

$$R_{DD} = \frac{\mu_0}{4\pi}\frac{\gamma_1\gamma_2}{\langle r_{12}^3\rangle}\frac{\hbar}{2\pi} \qquad [1]$$

where $\mu_0$ is the permeability of free space, and $\gamma_i$ are the magnetogyric ratios of the nuclei under consideration.

Rotational resonance (RR) is a MAS NMR technique which selectively restores the dipolar interaction between a homonuclear spin pair, thus allowing the determination of the dipolar coupling constant, $R_{DD}$, and hence, the internuclear distance. Historically, the RR phenomenon was discovered by Andrew and co-workers in a $^{31}P$ NMR study of phosphorus pentachloride, which consists of $PCl_4^+$ and $PCl_6^-$ units in the solid state. This group noticed that when the rate of sample spinning matched the difference in resonance frequencies of the nonequivalent phosphorus centres, their peaks broadened and the rate of cross-relaxation was enhanced.

It is now known that if the RR condition is satisfied, direct dipolar coupling is restored selectively to a homonuclear spin pair. That is, if the sample spinning rate is adjusted to a frequency, $\nu_r$, such that

$$n\nu_r = (\nu_1^{iso} - \nu_2^{iso}) \qquad [2]$$

where *n* is an integer, generally 1–3, and $\nu_1^{iso}$ and $\nu_2^{iso}$ are the isotropic resonant frequencies of spins 1 and 2 respectively, then the two nuclei are said to be in RR. As a result, dipolar coupling between the nuclei is restored (via the 'flip-flop' term in the dipolar Hamiltonian), and 'line broadening' of the resonances at $\nu_1^{iso}$ and $\nu_2^{iso}$ is observed (see **Figure 1**). Additionally, a rapid oscillatory exchange of Zeeman

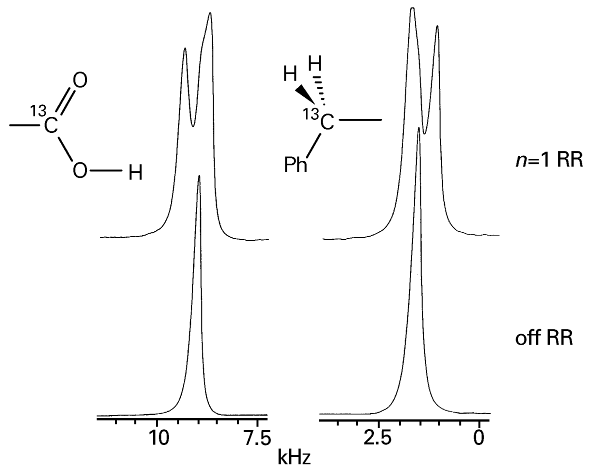



**Figure 1** Effect of RR on the $^{13}C$ NMR line shape of $Ph^{13}CH_2{}^{13}COOH$. Spectra acquired at 4.7 T (50.3 MHz). Top spectrum at $n = 1$ RR, $\nu_{rot} = 7207$ Hz. Bottom spectrum off RR, $\nu_{rot} = 10\,000$ Hz.

magnetization occurs. In fact, it is this exchange of magnetization rather than the line shape which is usually monitored in order to determine the dipolar coupling constant. In order to generate an exchange curve which may be analysed and simulated, one of the two resonances involved must be inverted selectively; the intensity difference of the peaks is then monitored as function of time.

Obviously, it is highly desirable to develop techniques capable of recovering weak dipolar coupling constants from high-resolution NMR spectra obtained under MAS conditions. The focus of the present discussion will be to provide an overview of the basic RR scheme. First, the theory of RR will be outlined, followed by a discussion of the most important experimental techniques employed to measure dipolar coupling constants under conditions of RR. Finally, some examples that illustrate the applications and limitation of the techniques will be described.

## Theory

### Restoring the dipolar interaction: a theoretical approach

The most important interaction in NMR results from the application of a large external magnetic field, $\boldsymbol{B}_0$, to the sample. Termed the Zeeman interaction, its effect on the normally degenerate nuclear spin energy levels is to cause them to split. The Zeeman levels are perturbed by local fields generated by the motion of electrons in the vicinity of the nucleus. These induced local fields are proportional to the applied field. In frequency units, the Hamiltonian operator which accounts for both the Zeeman interaction and this chemical shielding (CS) interaction is

$$h^{-1}\hat{\mathscr{H}}_{\mathrm{Z,\,CS}} = -\sum_i \nu_i \hat{I}_{zi} \qquad [3]$$

where

$$\nu_i = \frac{\gamma_i B_0(1-\sigma_i^{\mathrm{iso}})}{2\pi} \qquad [4]$$

and $\sigma^{\mathrm{iso}}$ is the isotropic chemical shielding constant.

The interaction of interest in RR is the dipolar interaction, an orientationally dependent through-space spin–spin coupling, which leads to a perturbation of the CS-perturbed Zeeman energy levels. For a homonuclear two-spin system, the truncated dipolar Hamiltonian operator is given by the following:

$$\begin{aligned} \hbar^{-1}\hat{\mathscr{H}}_{\mathrm{DD}} = &-\left(\frac{\mu_0}{4\pi}\right)\gamma^2\hbar\langle r_{12}^{-3}\rangle \\ &\times\left[\hat{I}_{z1}\hat{I}_{z2}-\frac{1}{4}\left(\hat{I}_{1+}\hat{I}_{2-}+\hat{I}_{1-}\hat{I}_{2+}\right)\right] \\ &\times\left(3\cos^2\theta-1\right) \end{aligned} \qquad [5]$$

Here, $\hat{I}_+$ and $\hat{I}_-$ are the raising and lowering operators and $\theta$ is the angle between the applied magnetic field and the internuclear vector, $r_{12}$. The factor containing the raising and lowering operators is sometimes referred to as the 'flip-flop' term.

The final interaction that must be considered is the indirect spin–spin coupling interaction, which is mediated by the intervening electrons. The indirect spin–spin Hamiltonian, $\hat{\mathscr{H}}_{\mathrm{J}}$, is often ignored because it is frequently considerably smaller than $\hat{\mathscr{H}}_{\mathrm{DD}}$.

Up until this point, we have implicitly assumed time independence of the interactions and their corresponding Hamiltonian operators. This assumption is valid for a rigid stationary sample. However, when the sample is spun rapidly, each of the internal Hamiltonians becomes time-dependent. For example, $\hat{\mathscr{H}}_{\mathrm{Z,CS}}$ becomes time-dependent when there is chemical shielding anisotropy due to the fact that the orientations of the chemical shielding tensors relative to the applied magnetic field change as the sample

rotates. Then Equation [3] becomes

$$\hbar^{-1}\mathscr{H}_{\mathrm{Z,\,CS}}(t) = -\sum_i \gamma_i B_0\big[\big(1-\sigma_i^{\mathrm{iso}}\big) + \big(\sigma_{i,zz}-\sigma_i^{\mathrm{iso}}\big)\xi(t)\big]\hat{I}_{zi} \qquad [6]$$

where $(\sigma_{i,zz} - \sigma_i^{\mathrm{iso}})$ is a measure of the orientation dependence of the chemical shielding and $\xi(t)$ represents the time dependence of the interaction:

$$\xi(t) = C_1\cos(\omega_{\mathrm{r}}t) + S_1\sin(\omega_{\mathrm{r}}t) + C_2\cos(2\omega_{\mathrm{r}}t) + S_2\sin(2\omega_{\mathrm{r}}t) \qquad [7]$$

Here, $C_1$, $C_2$, $S_1$, and $S_2$ are constants that depend on the nature of the interaction (i.e. CS, dipolar) and $\omega_{\mathrm{r}}$ is the rotor angular frequency.

Summing the CS-perturbed Zeeman Hamiltonian given in Equation [6] with the time-dependent dipolar Hamiltonian gives the total Hamiltonian

$$\hbar^{-1}\mathscr{H}(t) = -\sum_i \gamma_i B_0\big[\big(1-\sigma_i^{\mathrm{iso}}\big) + \big(\sigma_{i,zz}-\sigma_i^{\mathrm{iso}}\big)\xi(t)\big]\hat{I}_{zi} - \left(\frac{\mu_0}{4\pi}\right)\gamma^2\hbar\langle r_{12}^{-3}\rangle \times \xi(t)\left[\hat{I}_{z1}\hat{I}_{z2} - \frac{1}{4}\big(\hat{I}_{1+}\hat{I}_{2-} + \hat{I}_{1-}\hat{I}_{2+}\big)\right] \qquad [8]$$

The parameter $\xi(t)$ completely describes the time-dependence of a rotating solid.

In order to understand some of the essential features of the RR experiment, it is convenient to assume negligible chemical shielding anisotropy. Under these conditions, $\omega_1$ and $\omega_2$, the CS-perturbed Zeeman angular frequencies, are independent of time. In addition, it is convenient to use the spherical tensor notation to describe the direct dipolar interaction. Thus, the total Hamiltonian is

$$\hbar^{-1}\mathscr{H}(t) = \omega_1\hat{I}_{z1} + \omega_2\hat{I}_{z2} + A_{\mathrm{d}}(t)T_{20} \qquad [9]$$

where

$$A_{\mathrm{d}}(t) = -\sqrt{6}(2\pi R_{\mathrm{DD}})\sum_{m=-2}^{2} d_m \exp(-\mathrm{i}\,m\omega_{\mathrm{r}}t) \qquad [10]$$

Here, $A_{\mathrm{d}}(t)$ represents the spatial dependence of the dipolar Hamiltonian, and $d_m$ are Fourier components of the dipole–dipole coupling which depend on Euler angles defining the crystallite orientation with respect to the rotor frame. They are time-independent. The spin part of the truncated dipolar Hamiltonian is

$$T_{20} = \frac{2}{\sqrt{6}}\left[\hat{I}_{z1}\hat{I}_{z2} - \frac{1}{4}\big(\hat{I}_{1+}\hat{I}_{2-} + \hat{I}_{1-}\hat{I}_{2+}\big)\right] \qquad [11]$$

To transform the total Hamiltonian into the doubly rotating frame of reference defined by the Zeeman interactions, the propagator is

$$U^{\mathrm{int}} = \exp\big[\mathrm{i}\big(\omega_1\hat{I}_{z1} + \omega_2\hat{I}_{z2}\big)t\big] \qquad [12]$$

The Zeeman terms and the $\hat{I}_z$ terms of the dipolar Hamiltonian are unaffected by this rotation since it is about the $z$-axis, and so the desired transformation is:

$$T_{20}^{\mathrm{int}} = (U^{\mathrm{int}})T_{20}(U^{\mathrm{int}})^{-1} \qquad [13]$$

The result of the transformation gives a periodic interaction frame dipolar Hamiltonian,

$$\hbar^{-1}\hat{\mathscr{H}}_{\mathrm{DD}}^{\mathrm{int}}(t) = -2(2\pi R_{\mathrm{DD}})\sum_{m=-2}^{2} d_m \left\{\hat{I}_{z1}\hat{I}_{z2}\exp(-\mathrm{i}m\omega_{\mathrm{r}}t) - \frac{1}{4}\left\{\begin{array}{c}\hat{I}_{1+}\hat{I}_{2-}\exp[\mathrm{i}(\omega_\Delta^{\mathrm{iso}} - m\omega_{\mathrm{r}}t)] \\ +\hat{I}_{1-}\hat{I}_{2+}\exp[-\mathrm{i}(\omega_\Delta^{\mathrm{iso}} + m\omega_{\mathrm{r}}t)]\end{array}\right\}\right\} \qquad [14]$$

where $\omega_\Delta^{\mathrm{iso}} = \omega_1 - \omega_2$. If we re-express the rotational resonance condition [2] in angular frequency units as $n\omega_{\mathrm{r}} = \omega_\Delta^{\mathrm{iso}}$, and if $|R_{\mathrm{DD}}\tau_{\mathrm{r}}| \ll 1$, where $\tau_{\mathrm{r}} = \nu_{\mathrm{r}}^{-1}$ is the rotor period, then the time-independent terms vanish and the time average of Equation [14] over one rotor period is

$$\overline{\hbar^{-1}\hat{\mathscr{H}}_{\mathrm{DD}}} = \frac{1}{2}(2\pi R_{\mathrm{DD}})\big(d_m\hat{I}_{1+}\hat{I}_{2-} + d_{-m}\hat{I}_{1-}\hat{I}_{2+}\big) \qquad [15]$$

where $m = \pm 1, \pm 2$. The result of this exercise is that at rotational resonance, parts of the 'flip-flop' term do not average to zero and will therefore contribute to the MAS NMR spectrum.

Some qualitative results of an approximate theoretical treatment of rotational resonance are useful to examine. For $n = 1$ RR, the splitting of each peak is given by $R_{DD}/(2\sqrt{2})$, or ~0.35 $R_{DD}$. For the $n = 2$ case, the splitting is $R_{DD}/4$. The splitting decreases as the order of the RR increases. More rigorous treatments also indicate that the splitting decreases as the chemical shielding anisotropy increases. It is important to note that the observed line widths for homonuclear spin systems are not strictly independent of spinning speed; for a spin pair with differing isotropic chemical shifts, the line widths take on a $\omega_r^{-2}$ dependence at high spinning speeds.

When the RR condition is satisfied, a rapid exchange of Zeeman magnetization occurs in addition to dipolar broadening. We will not present a complete theoretical description of the origins of this exchange, but rather present some of the important approximate results of such a treatment. It is convenient to define

$$\Lambda^2 = \left(T_2^{ZQ}\right)^{-2} - 4\left|\overline{\omega_B^{(n)}}\right|^2 \qquad [16]$$

where $T_2^{ZQ}$ is the zero-quantum relaxation time constant, and $\omega_B^{(n)}$ are the resonant Fourier components associated with the flip-flop term of the dipolar Hamiltonian for RR of order $n$. To monitor the exchange of magnetization, one plots $\langle \hat{I}_{z1} - \hat{I}_{z2}\rangle$ as a function of time. In the limit of very fast dephasing where $T_2^{ZQ}$ is relatively short and thus $\Lambda^2 \gg 0$, the decay of magnetization is exponential:

$$\left\langle \hat{I}_{z1} - \hat{I}_{z2}\right\rangle(t) \approx \exp\left(-T_2^{ZQ}\left|\overline{\omega_B^{(n)}}\right|^2 t\right) \qquad [17]$$

In the case of very slow dephasing where $T_2^{ZQ}$ is relatively long and $\Lambda^2 \ll 0$, the exchange of magnetization oscillates as it decays:

$$\left\langle \hat{I}_{z1} - \hat{I}_{z2}\right\rangle(t) \approx \exp\left(-\frac{t}{2T_2^{ZQ}}\right)\cos\left(\left|\overline{\omega_B^{(n)}}\right| t\right) \qquad [18]$$

In practice, the parameters which influence the observed magnetization exchange curve include $R_{DD}$, $T_2^{ZQ}$, the magnitude of the principal components of the chemical shielding tensors, the relative orientation of the CS tensors with respect to $r_{12}$, and the $J$-coupling constant.

### A pictorial representation of rotational resonance and the exchange of Zeeman magnetization

At this point, it is instructive to provide a qualitative picture of the rotational resonance phenomenon. The energy level diagram for an isolated homonuclear two-spin system where the two nuclei have resonance frequencies $\nu_1^{iso}$ and $\nu_2^{iso}$ is shown in **Figure 2**. Transitions 1 and 2 correspond to the two isotropic peaks, which would be observed in a MAS NMR spectrum. The difference between the isotropic chemical shifts (or, alternatively, the energies) is, according to the diagram, equivalent to the angular frequency $\omega_\Delta^{iso}$. As shown earlier, rotational resonance occurs when an integer multiple of the spinning frequency is equivalent to $\omega_\Delta^{iso}$. In terms of the diagram, it is convenient to think of the mechanical rotation of the sample as supplying the necessary energy for zero-quantum coherence between the two intermediate energy levels. The fact that these two states are linked by mechanical rotation ensures that the dipolar interaction will be recoupled, and that exchange of Zeeman magnetization will occur rapidly.

**Figure 3** illustrates the exchange experiment, in which one of the transitions is selectively inverted, thus creating a nonequilibrium situation in which spins must relax so that the equilibrium Boltzmann populations are re-established. If we consider the diagram on the left to reflect the excess populations in arbitrary units as determined by the Boltzmann distribution, a selective inversion of transition 1 will result in the population distribution shown on the right. Transition 1 is inverted while the intensity of transition 2 remains unperturbed. Techniques for accomplishing this experimentally will be discussed in the next section. Once the inversion has been carried out, the diagram on the right shows a difference of five population units between the two intermediate energy levels. Rotational resonance provides the zero-quantum coherence necessary for an exchange

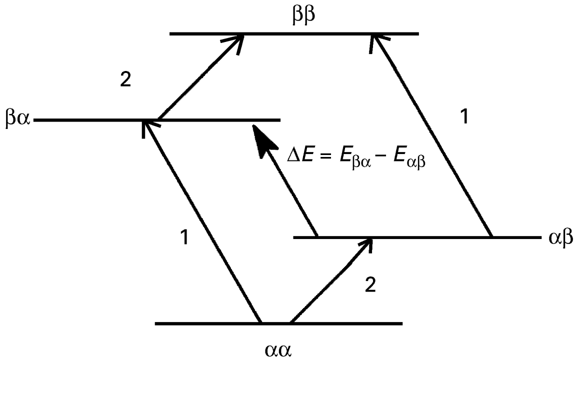



**Figure 2** Simplified energy level diagram for two spin-½ nuclei with different isotropic chemical shifts. The two transitions are labelled '1' and '2', and their difference is greatly exaggerated. The energy of the zero-quantum transition is indicated, which corresponds to the mechanical energy supplied at RR. Here, $J$-coupling is ignored and dipolar coupling is not shown.

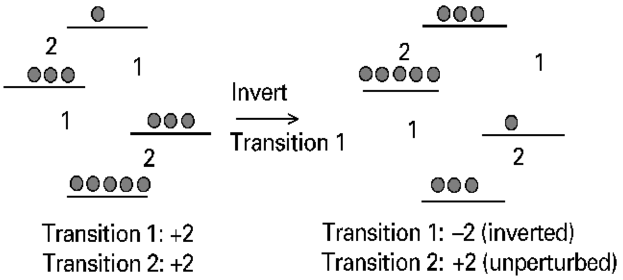



**Figure 3** Energy level and population distribution diagrams for two spin-½ nuclei. The circles indicate the excess population in arbitrary units relative to the least populated level. On the left, an equilibrium Boltzmann-type distribution is represented. Both transitions would show a signal of relative intensity +2. Upon inversion of transition 1 (at right) the populations related to this transition are switched, while the net difference in population for transition 2 is unchanged. Transition 1 would now show an inverted signal with relative intensity −2.

of Zeeman magnetization between these two levels. The zero-quantum relaxation which dampens this exchange is described by the time constant $T_2^{ZQ}$.

## Experimental techniques

### Pulse sequences and cross-polarization

The basic rotational resonance experiment can be as simple a single-pulse excitation, with the rate of MAS adjusted to satisfy the RR condition. A $\pi/2$ pulse followed by acquisition of the free induction decay (FID) will generate a spectrum with significant broadening of the two resonances concerned. Many typical applications involve $^{13}C$ in the presence of $^{1}H$, and benefit from standard CP techniques. **Figure 1** shows an example of the line broadening observed for the $n = 1$ RR condition for the $^{13}C$–$^{13}C$ spin pair in $Ph^{13}CH_2{}^{13}COOH$, with CP.

A typical pulse sequence for carrying out the RR experiment with selective inversion of a particular transition and CP is shown in **Figure 4**. Note that a flipback pulse is applied to the rare spin channel to store the magnetization along the $z$ axis before carrying out the inversion. The efficiency of CP becomes sensitive to the spinning rate, particularly as $\omega_r$ increases. One technique which attempts to circumvent this problem is known as variable amplitude cross-polarization (VACP), where the spin-locking pulses vary in amplitude.

The goal of the RR experiment is the extraction of the homonuclear dipolar coupling constant, $R_{DD}$. This can be done by carrying out lineshape simulations. However, in general this is not done because a Zeeman magnetization exchange experiment is more

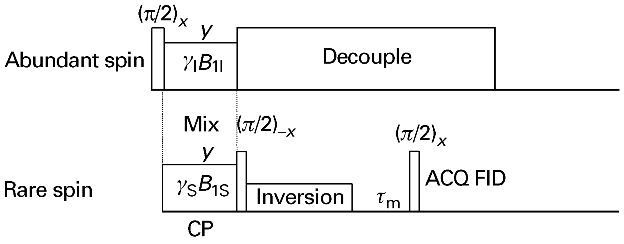



**Figure 4** Pulse sequence for carrying out the RR experiment (see text). In the case of the magnetization exchange experiment, CP of the rare spins is followed by a flipback pulse on the rare spin channel, selective inversion of a particular resonance, and a variable delay before acquisition.

informative, and more sensitive to the magnitude of $R_{DD}$, as will be shown.

### The exchange experiment

To generate an exchange curve, one of the two resonances involved must be inverted selectively. By whichever technique a selective inversion is carried out, it is important that the other resonances not be perturbed. The most frequently used inversion techniques in RR experiments are a long, soft pulse or an asynchronous DANTE (delays alternating with nutation for tailored excitation) sequence.

In cases where the CS anisotropy at one or both of the sites is comparable to the isotropic chemical shift difference between them, difficulties arise in carrying out the inversion with selectivity. Total sideband suppression pulse sequences combined with their time-reversed counterparts may be used to overcome the difficulties associated with large chemical shift anisotropies. Regardless of what technique is used to establish the initial condition of maximum polarization difference, the next step in the experiment is to allow the exchange of Zeeman magnetization for a variable time, $\tau_m$ (see **Figure 4**), before applying a $\pi/2$ acquisition pulse. The equilibration of magnetization between the two sets of spins is described by the approximate Equation [17] or [18], depending on the system.

## Applications and limitations

As mentioned previously, the primary goal of the rotational resonance experiment is to determine the dipolar coupling constant, $R_{DD}$, from which the internuclear distance, $r$, may be calculated. Carbon–carbon separations as large as 6.8 Å have been successfully determined, which corresponds to measuring a coupling as small as 24 Hz. Occasionally, dihedral angle measurements have also been carried

out using RR. At higher order rotational resonances (i.e. $n = 3$ or $n = 4$), where CS anisotropy is more likely to be comparable to $\nu_r$, the lineshapes and exchange curves are more sensitive to the orientation of the chemical shielding tensors. In general, when simulations (of either line shapes or exchange curves) are performed, they depend on $R_{DD}$ and, to varying degrees, on the magnitudes of the principal components of the chemical shielding tensors, their orientations with respect to the internuclear vector and with respect to each other, the magnitude of the $J$-coupling, and the zero-quantum transverse relaxation time constant, $T_2^{ZQ}$.

### Lineshape simulations

In order to effectively determine a dipolar coupling constant based on a lineshape simulation, the chemical shielding tensors and their orientations must be known, as well as the $J$-coupling constant, and $T_2^{ZQ}$. To determine the principal components of the chemical shielding tensors, a MAS NMR spectrum acquired on a singly-labelled compound in the slow-spinning regime may be used to emulate the powder pattern provided the isolated spin approximation is valid. In some cases it is also possible to determine the CS tensor components from a spectrum of the stationary sample. Determining the *orientations* of the CS tensors is a more involved process, although in some cases careful assumptions and clues from local symmetry may be helpful. In practice, a value for $T_2^{ZQ}$ is usually estimated from the observed line widths off the RR condition.

$$(T_2^{ZQ})^{-1} = \pi[\nu_{1/2}(1) + \nu_{1/2}(2)] \quad [19]$$

In many cases, not all these parameters are known for the specific spin system under investigation. Therefore, two techniques that may be employed when a lineshape simulation is desirable are (i) a simulation based on known chemical shielding tensors ($\sigma$), $J$-coupling constants, and $T_2^{ZQ}$ values, where only $R_{DD}$ is varied; (ii) use of a calibration with respect to similar compounds, where $R_{DD}$ (or $r$ itself) can be extracted analytically from the observed splitting of a resonance.

For example, the calibration method (ii) has been employed for a series of $^{13}$C-labelled retinals containing vinylic and methyl carbons, shown in **Figure 5**. The three isotopomers were $^{13}$C-labelled at the (10,20), (11,20), and (12,20) positions. It must be emphasized that the required parameters ($\sigma$, $J$, $T_2^{ZQ}$ and $r$) were known independently from X-ray and previous NMR studies. $T_2^{ZQ}$ was estimated using

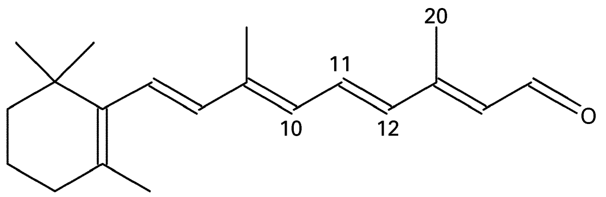



**Figure 5** Structure of the retinal studied using RR, with the labelled carbons indicated. See text for details. (Reprinted with permission of the American Chemical Society from Verdegem PJE, Helmle M, Lugtenburg J and de Groot HJM (1997) (*Journal of the American Chemical Society*, **119**: 169–174).

Equation [19]. The goal of the calibration was to be able to employ a simple, analytic equation relating $r$ to the observed broadening of the vinylic peaks at RR. To accomplish this, the 'ideal' splitting presented above, $R_{DD}/(2\sqrt{2})$, was plotted against the observed splitting, $\Delta\omega$. Simulations showed that $\Delta\omega$ could be reliably reproduced, independent of the actual shape of the line. The resulting equation,

$$\frac{R_{DD}}{\sqrt{8}} = 1.15\left(\frac{\Delta\omega}{2\pi}\right) + 7 \ \text{(Hz)} \quad [20]$$

shows that the approximate theory fits well with experimental results in this case, and allows for a very straightforward determination of $r$ from the observed splitting. The major advantage of using line shape simulations to extract the dipolar coupling constant, in general, is that the spectrometer time involved is less than that for the corresponding magnetization exchange experiment. For molecules similar to the retinal in **Figure 5**, where $r$ is unknown, the analytical empirical Equation [20] can provide the information after a simple 1D NMR experiment.

In spite of the results of the preceding example, the lineshape simulation method has rarely been used in practice, mainly because the $R_{DD}$ values are too small to result in splittings. In such cases, the exchange curve method discussed below is the standard technique for extracting the dipolar coupling constant under RR conditions.

### Exchange curves and simulations

By far the most common method for deriving structural information under RR conditions is through the analysis and simulation of a magnetization exchange curve. Once a suitable state of polarization difference has been achieved between the two sets of spins, 1 and 2, the delay time, $\tau_m$, is varied before applying a $\pi/2$ observe pulse and acquiring the spectrum (see **Figure 4**). Separate NMR experiments

must be performed to generate each point on a magnetization exchange plot.

In order to extract the dipolar coupling constant or structural information, the observed magnetization decay must be simulated. Qualitative and relative distance information is more readily available than quantitative information since the exchange curve depends on the same parameters that the RR line shape depends on. Two common procedures for extracting $R_{\mathrm{eff}}$ are: (i) comparison of the exchange curve with a series of exchange curves of model compounds for which $r$ is known, and (ii) complete simulation of the exchange curve, where $\sigma$, $J$, and $T_2^{\mathrm{ZQ}}$ are known (or estimated).

The magnetization due to naturally abundant NMR-active spins in the sample must be considered. This is done by subtracting the natural-abundance spectrum from that of the labelled sample. Failure to make such a correction could lead to an overestimation of $r$ and an underestimation of $T_2^{\mathrm{ZQ}}$.

A recent example of the application of RR to a structural problem will serve to illustrate its utility as a comparative tool. **Figure 6** shows two peptide fragments in different conformations. This compound models the peptide A$\beta$1-42, a constituent of the amyloid plaques characteristic of Alzheimer's disease. Rotational resonance MAS NMR was used in a qualitative fashion by Costa and co-workers to determine whether the amide conformation in the solid state was '*cis*' or '*trans*'. From previous experiments, model compounds served to give the chemical shielding tensor orientations of the carbonyl carbon. The orientation of the internuclear vector connecting the two labelled carbon atoms with respect to the chemical shielding tensor of the carbonyl carbon is drastically different for the two conformers shown in **Figure 6**. Note that in the *trans* conformation, the internuclear vector lies nearly along the $\sigma_{11}$ component, while in the *cis* conformation it lies nearly along $\sigma_{22}$. The dipolar coupling for the two conformations should, however, be nearly identical. Hence, the variable of interest in this experiment is the CS tensor orientation. Experimental Zeeman magnetization exchange curves were generated and matched to simulated curves (**Figure 7**). It was found that theory matched experiment only when a *trans* geometry was assumed. The $n = 2$ RR experiment was used in this case because at higher spinning speed (i.e. $n = 1$), the orientations of the CS tensors become less influential in determining the course of the magnetization exchange.

This example shows that RR experiments can be used for more than simply extracting the dipolar coupling constant and determining an accurate value for $r_{12}$. In fact, the basic RR technique is probably

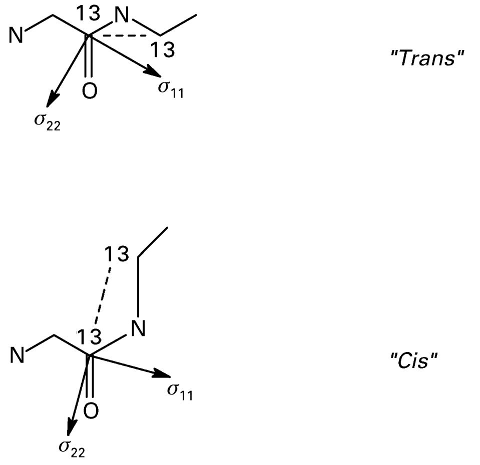



**Figure 6** Fragments of the peptide $\beta$34-42 showing the *cis* and *trans* conformations. Also indicated is the orientation of the carbonyl carbon chemical shielding tensor, with $\sigma_{33}$ perpendicular to the plane. Note the different orientations of the C–C internuclear vector with respect to the CS tensor components. Reprinted with permission of the American Chemical Society from Costa PR, Kocisko DA, Sun BQ, Lansbury PT Jr and Griffin RG (1997) *Journal of the American Chemical Society*, **119**: 10487–10493)

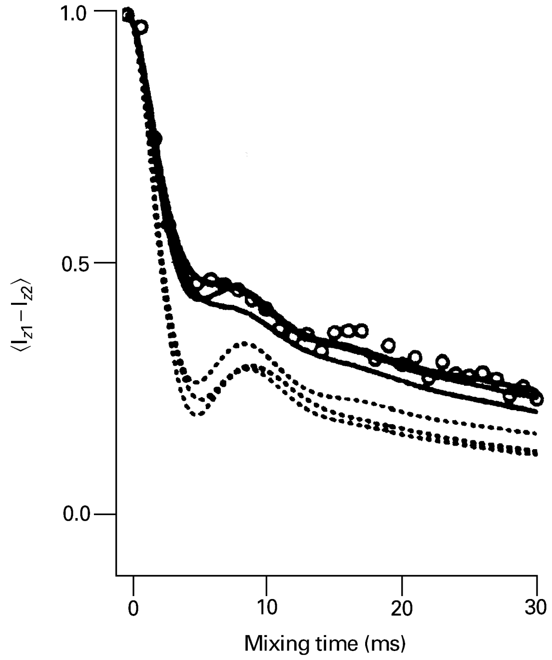



**Figure 7** Zeeman magnetization exchange plot for the peptide $\beta$34-42 fragments shown in **Figure 6**. The open circles are experimentally determined data points; the solid lines result from simulations assuming a *trans* geometry; the dotted lines result from simulations assuming a *cis* geometry. Reproduced with permission of the American Chemical Society from Costa PR, Kocisko DA, Sun BQ, Lansbury PT Jr and Griffin RG (1997) *Journal of the American Chemical Society* **119**: 10487–10493.

**Table 1** A summary of some homonuclear dipolar recoupling techniques.

| *Name* | *Acronym* | *Principle* |
|---|---|---|
| Rotational resonance | RR | Recouples when the difference in chemical shift frequencies is an integer multiple of the MAS speed. |
| Dipolar recovery at the magic angle | DRAMA | In its simplest form, a pair of *x* and –x $\pi$ pulses separated by a delay, $\tau$, results in an observable dipolar broadening (Tycko R and Dabbagh G (1991) Double-quantum filtering in magic-angle-spinning NMR spectroscopy: an approach to spectral simplification and molecular structure determination. *Journal of the American Chemical Society* **113***:* 9444–9448). |
| Simple excitation for the dephasing of the rotational-echo amplitudes | SEDRA | Synchronously applied pulses lead to signal dephasing for dipolar coupled spins (Gullion T and Vega S (1992) A simple magic angle spinning NMR experiment for the dephasing of rotational echoes of dipolar coupled homonuclear spin pairs. *Chemical Physics Letters* **194:** 423–428). |
| Radio frequency driven dipolar recoupling | RFDR | Rotor-synchronized $\pi$-pulses reintroduces flip-flop term (Bennett AE, Ok JH, Griffin RG and Vega S (1992) Chemical shift correlation spectroscopy in rotating solids: radio frequency-driven dipolar recoupling and longitudinal exchange. *Journal of Chemical Physics* **96:** 8624–8627). |
| Unified spin echo and magic echo | USEME | Spin-echo and magic-echo sequences are applied to recover the dipolar interaction (Fujiwara T, Ramamoorthy A, Nagayama K, Hioka K and Fujito T (1993) Dipolar HOHAHA under MAS conditions for solid-state NMR. *Chemical Physics Letters* **212:** 81–84). |
| Combines rotation with nutation | CROWN | Dipolar dephasing occurs due to applied RF pulses (Joers JM, Rosanske R, Gullion T and Garbow JR (1994) Detection of dipolar interactions by CROWN NMR. *Journal of Magnetic Resonance* **A106:** 123–126). |
| Double quantum homonuclear rotary resonance | 2Q1-HORROR | RF field applied at half the rotation frequency in conjunction with RF pulses (Nielsen NC, Bildsøe H, Jakobsen HJ and Levitt MH (1994) Double-quantum homonuclear rotary resonance: efficient dipolar recovery in magic-angle-spinning nuclear magnetic resonance. *Journal of Chemical Physics* **101(3):** 1805–1812). |
| Melding of spin-locking dipolar recovery at the magic angle | MELODRAMA | Rotor-synchronized 90° phase shifts of the applied spin-locking field (Sun B-Q, Costa PR, Kocisko D, Lansbury PT Jr and Griffin RG (1995) Internuclear distance measurements in solid state nuclear magnetic resonance: Dipolar recoupling via rotor synchronized spin locking. *Journal of Chemical Physics* **102:** 702–707). |
| Rotational resonance in the tilted rotating frame | R2TR | Application of an RF field allows selective recoupling when the chemical shift difference is small (Takegoshi K, Nomura K and Terao T (1995) Rotational resonance in the tilted rotating frame. *Chemical Physics Letters* **232:** 424–428). |
| Sevenfold symmetric radiofrequency pulse sequence | C7 | Seven phase-shifted RF pulse cycles lead to dipolar recoupling (Lee YK, Kurur ND, Helmle M, Johannessen OG, Nielsen NC and Levitt MH (1995) Efficient dipolar recoupling in the NMR of rotating solids. A sevenfold symmetric radiofrequency pulse sequence. *Chemical Physics Letters* **242:** 304–309). |
| Dipolar recoupling with a windowless multipulse irradiation | DRAWS | Windowless DRAMA sequence (Gregory DM, Wolfe GM, Jarvie TP, Sheils JC and Drobny GP (1996) Double-quantum filtering in magic-angle-spinning NMR spectroscopy applied to DNA oligomers. *Molecular Physics* **89(6):** 1835–1850). |
| Rotational resonance tickling | $R^2T$ | Ramped RF field during the variable delay removes the $T_2^{ZQ}$ dependence (Costa PR, Sun B and Griffin RG (1997) Rotational resonance tickling: accurate internuclear distance measurement in solids. *Journal of the American Chemical Society* **119:** 10821–10830). |
| Adiabatic passage rotational resonance | APRR | MAS speed varied during CP mixing to achieve more complete polarization transfer (Verel R, Baldus M, Nijman M, van Os JWM and Meier BH (1997) Adiabatic homonuclear polarization transfer in magic-angle-spinning solid-state NMR. *Chemical Physics Letters* **280**: 31–39). |
| Supercycled POST-C5 | SPC-5 | Fivefold symmetric pulse sequence leads to homonuclear dipolar recoupling (Hohwy M, Rienstra CM, Jaroniec CP, Griffin RG (1999) *Journal of Chemical Physics* **110**: 7983–7992). |

better suited to qualitative distance measurements such as in the example given.

It is necessary to make a general comment regarding the influence of molecular motion on the measurement of dipolar coupling constants. In the context of solid-state NMR, it is not $r_{12}$ which is directly measured, but rather the dipolar coupling constant. Molecular librations and vibrations will cause a certain degree of averaging of the dipolar interaction and thus $R_{DD}$. The net result of the

motional averaging of the dipolar coupling is that the calculated distances, $r$, will be too large.

Finally, it is important to recognize that the dipolar coupling constant measured in any NMR experiment also has, in principle, a contribution from the anisotropy in the indirect spin–spin coupling, $\Delta J$. That is, only an effective dipolar coupling constant, $R_{\mathrm{eff}}$ can be measured where

$$R_{\mathrm{eff}} = R_{\mathrm{DD}} - \Delta J/3 \quad [21]$$

The last term in equation [21], $\Delta J/3$, is generally ignored.

### Other homonuclear recoupling methods

Restoring the dipolar coupling between both heteronuclear and homonuclear spin pairs is of great interest. Rotational resonance applies strictly to homonuclear spin pairs, and **Table 1** provides a brief overview of some of the other techniques available for recovering the dipolar coupling and extracting $R_{\mathrm{eff}}$ for homonuclear spin pair from high-resolution MAS spectra.

## Conclusions

At present, RR is best suited for use as a qualitative probe into molecular structure rather than a quantitative one. In most experiments which have been done using the basic RR technique, the internuclear distances were known beforehand as a result of other investigations. Further developments of related RR techniques (such as rotational resonance tickling) may prove to be more useful in obtaining quantitative results. Still, the standard RR experiment is an excellent one for confirming distances between homonuclear spin pairs in a proposed structure.

## List of symbols

$A_{\mathrm{d}}(t)$ = spatial dependence of the dipolar Hamiltonian; $B_0$ = external applied magnetic field; $d_m$ = Fourier components of the dipole–dipole coupling; $h$ = Planck constant; $\hbar$ = Planck constant divided by $2\pi$; $\bar{\hat{\mathcal{H}}}_{\mathrm{DD}}$ = average dipolar Hamiltonian; $\hat{\mathcal{H}}_{\mathrm{DD}}$ = direct dipolar Hamiltonian operator; $\hat{\mathcal{H}}_{\mathrm{J}}$ = indirect spin–spin coupling Hamiltonian; $\hat{\mathcal{H}}_{\mathrm{Z,CS}}$ = chemical shielding perturbed Zeeman Hamiltonian operator; $\hat{I}_-$ = lowering operator; $\hat{I}_+$ = raising operator; $\hat{I}_i$ = spin angular momentum operator for spin $i$; $\hat{I}_{zi}$ = $z$-component of the spin angular momentum operator for spin $i$; $J$ = indirect spin–spin coupling constant; $n$ = order of the rotational resonance; $r_{12}$ = distance between spins 1 and 2; internuclear vector; $R_{\mathrm{DD}}$ = direct dipolar coupling constant (in Hz); $R_{\mathrm{eff}}$ = observed dipolar coupling constant; $t$ = time; $T_{20}$ = spin term in the spherical tensor representation of the dipolar Hamiltonian; $T_2^{\mathrm{ZQ}}$ = zero-quantum relaxation time constant; $U$ = propagator; $\gamma_i$ = magnetogyric ratio of spin $i$; $\Delta J$ = anisotropy of the indirect spin–spin interaction; $\theta$ = angle between the applied field and the internuclear vector; $\Lambda^2$ = dephasing parameter; $\mu_0$ = permeability of free space; $\nu_{\mathrm{r}}$ = rotor frequency in Hz; $\nu_i$, $\nu_i^{\mathrm{iso}}$ = isotropic resonant frequency of nucleus $i$ (in Hz); $\nu_{\mathrm{rot}}$ = rotor frequency (in Hz); $\nu_{1/2}$ = line width at half-height (in Hz); $\xi(t)$ = time-dependence of the NMR interactions as a result of sample rotation; $\sigma_{\mathrm{i}}$ = chemical shielding tensor of spin $i$; $\sigma_{ii}$ = principal component of the chemical shielding tensor ($i$ = 1, 2, 3); $\sigma^{\mathrm{iso}}$ = isotropic chemical shielding constant; $\tau_{\mathrm{m}}$ = variable mixing time; $\omega_B^{(n)}$ = resonant Fourier components; $\omega_i$ = CS-perturbed Zeeman angular frequency of spin $i$ (in rad $s^{-1}$); $\omega_{\mathrm{r}}$ = rotor frequency (in rad $s^{-1}$); $\omega_\Delta^{\mathrm{iso}}$ = difference in isotropic angular frequencies of spins 1 and 2.

*See also:* **Chemical Exchange Effects in NMR; High Resolution Solid State NMR, $^{13}C$; High Resolution Solid State NMR, $^{1}H$, $^{19}F$; NMR in Anisotropic Systems, Theory; NMR of Solids; NMR Pulse Sequences; NMR Relaxation Rates; Solid State NMR, Methods.**

## Further reading

Andrew ER, Bradbury A, Eades RG and Wynn VT (1963) Nuclear cross-relaxation induced by specimen rotation. *Physics Letters* **4**: 99–100.

Garbow JR and Gullion T (1995) Measurement of internuclear distances in biological solids by magic-angle-spinning $^{13}C$ NMR. In Beckmann N (ed), *Carbon-13 NMR Spectroscopy of Biological Systems*, pp. 65–115. New York: Academic Press.

Griffiths JM and Griffin RG (1993) Nuclear magnetic resonance methods for measuring dipolar couplings in rotating solids. *Analytica Chimica Acta* **283**: 1081–1101.

Peersen OB and Smith SO (1993) Rotational resonance NMR of biological membranes. *Concepts in Magnetic Resonance* **5**: 303–317.

Raleigh DP, Levitt MH and Griffin RG (1988) Rotational resonance in solid-state NMR. *Chemical Physics Letters* **146**: 71–76.

Smith SO (1993) Magic angle spinning NMR methods for internuclear distance measurements. *Current Opinion in Structural Biology* **3**: 755–759.

Smith SO (1996) Magic angle spinning NMR as a tool for structural studies of membrane proteins. *Magnetic Resonance Review* **17**: 1–26.

Webb GA, Recent advances in solid-state NMR are reviewed annually in: *Nuclear Magnetic Resonance: Specialist Periodical Reports*. Cambridge: The Royal Society of Chemistry.

# Solvent Suppression Methods in NMR Spectroscopy

**Maili Liu** and **Xi-an Mao**, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan, PR China

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Methods & Instrumentation**

## Introduction

In order to get useful information from NMR spectroscopy of biofluids or biomolecules (proteins, DNA, RNA, carbohydrates, amino acids), it is often necessary to measure $^1$H NMR spectra in aqueous solutions. In these samples, the concentration of solvent water protons is about 110 M and is about $10^5$ times higher than that of the molecules of interest which are usually in the mM concentration range or less. Such a huge excess of water spins can cause many problems for NMR measurements. Firstly, the receiver gain of the spectrometer must be set to a low value to avoid the water signal overloading the receiver. In this circumstance, the analogue-to-digital converter (ADC) will be filled by the water resonance and many of the small signals from the molecules of interest will be below 1 bit of the ADC resolution and hence will not be digitized adequately. Secondly, the water resonance may obscure many solute peaks, resulting in the loss of molecular structural information that could make the spectrum useless. Thirdly, the strong water resonance can cause radiation damping, which provides another relaxation mechanism and shortens the relaxation time of water and hence broadens the water peak. Therefore for $^1$H NMR spectroscopy to be useful in aqueous solution, it is clearly necessary to attenuate the water signal.

There has been a continued interest in developing new methods for solvent suppression in NMR spectroscopy of biomolecules and biofluids and many methods have been proposed. These fall into five categories (1) presaturation, (2) nonexcitation, (3) pulsed field gradient (PFG)-based methods, (4) filtering methods and (5) post-acquisition data processing. Among these methods, soft pulse presaturation is the most frequently used and it is also easy to be incorporated into one- (1D) and multidimensional (*n*D) pulse sequences. The other methods can also be found in some applications, such as the study of exchangeable protons. In the past decade, the use of PFG to enhance the suppression efficiency and to develop new methods has been a popular area of study in NMR spectroscopy. Another advantage of using PFGs is the reduction of radiation damping. The important criteria for a good suppression method are the efficiency, selectivity and phase and baseline properties of the resulting spectrum.

Owing to the limitation of space, this article focuses on the introduction of general principles of the solvent suppression methods, and emphasizes some of the important developments. The reader is encouraged to refer to the original literature and recent reviews for the fundamentals and details of the methods.

## Solvent presaturation

The presaturation method (PR) normally consists of a low-power, soft (long or continuous-wave) pulse at the solvent resonance. It is the simplest and the most widely used method for solvent suppression. The method can be found in a large number of 1D- and *n*D-NMR pulse sequences for measuring the $^1$H NMR spectra of biofluids or biomolecules in aqueous solutions. The general 1D presaturation pulse sequence is shown in **Figure 1A**, where $PR_x$ is the saturation pulse applied along the *x*-axis of the rotating frame. The duration of the PR pulse is normally equal to the preacquisition or relaxation delay. It has been found that a long PR pulse with a fixed phase can lock part of the water magnetization along the direction of the RF field. The locked resonance can be reduced by a so-called phase-shifted PR method as shown in **Figure 1B**. The pulse sequence of the phase-shifted presaturation method has two PR pulses with a duration ratio of 9:1 ($PR_x$:$PR_y$) and a $\pi/2$ phase shift. It had been reported that in conventional 2D COSY [homonuclear chemical shift correlation spectroscopy] and NOESY [2D NOE spectroscopy] experiments on 2 mM lysozyme in a 90% $H_2O$–10% $D_2O$ solution, a water signal suppression of a factor of $10^6$ was achieved using the phase-shift PR method. **Figure 1C** shows a pulse sequence known as NOESYPRESAT, which can be considered as a combination of the phase-shifted PR

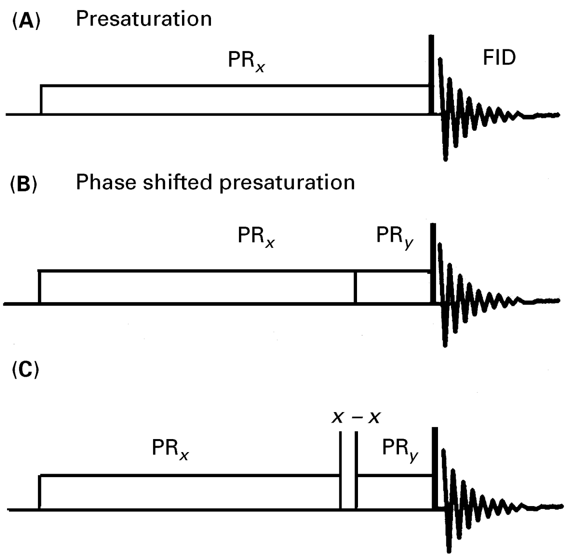



**Figure 1** Water suppression pulse sequences, where the bar symbol represents a 90° pulse: (A) presaturation, (B) phase-shifted presaturation (the ratio of the duration of $PR_x$ and $PR_y$ pulses is 9:1), (C) NOESYPRESAT.

and conventional NOESY approaches and which can further improve the spectral phasing and baseline. It also provides a more effective suppression in the wide base of the water resonance.

In HPLC–NMR or other applications with solvent mixtures, it is often necessary to suppress two solvent peaks at the same time. This is normally achieved by applying soft pulses or continuous irradiation at both solvent resonances or by fast switching the frequency offset between the two signals. With advances in NMR software, it is now possible to automatically search the solvent frequencies and to define suppression pulse offsets accordingly without any reduction in the suppression efficiency.

The presaturation method can cause a loss of signal intensity of exchangeable protons. This is the result of saturation transfer from the solvent resonance that is caused by the chemical exchange during the period of the saturation pulse. Generally, care must be taken when using the method to study systems containing labile protons. On the other hand, since the amount of the saturation transfer can be controlled by the (saturation) pulse length, the approach provides a facile method for the assignment of labile protons and for the study of their exchange rate with water.

## Solvent resonance nonexcitation

A well-established nonexcitation method is the 'Jump-Return (JR)' sequence. It is also known as the [1, −1] method. The sequence consists of a pair of 90° hard pulses that are separated by a delay $\tau$ and have a $\pi$ phase shift,

$$90°_x - \tau - 90°_{-x} - \text{FID} \qquad [1]$$

where FID = free-induction decay. The principle of the sequence can be described using the product operator approach. The first $90°_x$ RF pulse generates transverse magnetization of $-I_y$. During the delay period $\tau$, the transverse magnetization evolves at the relative frequency ($\omega$) with respect to the transmitter

$$-I_y \xrightarrow{\omega\tau} -I_y \cos(\omega\tau) + I_x \sin(\omega\tau) \qquad [2]$$

The last $90°_{-x}$ puts the $-I_y$ term on the right side of Equation [2] back along the $z$-axis of the rotating frame and the $I_x$ term remains unaffected. The sequence thus gives rise approximately to a sine-shaped excitation profile of $\sin(\omega\tau)$. The excitation bandwidth ($SW_b$) depends on the delay $\tau$ and is $SW_b = 1/\tau$. The maximum magnitude of the transverse magnetization can be obtained at the relative frequencies of $1/4\tau$ and $-1/4\tau$, respectively, and the two excitation bands have a 180° phase shift between them. The on-resonance solvent magnetization will be put along the $z$-axis and remains unchanged but the signals at the relative frequencies of $\pm 1/2\tau$ will be inverted. By increasing the number of the RF pulses from two, it is possible to make the excitation bands flatter and the suppression region narrower. One of the improved versions of this pulse sequence is the [1–3–3–1] approach. A disadvantage of the general method is that it is complicated by a phase roll over the spectrum and by the effects of radiation damping. However, the saturation transfer effect remains at a minimum in these methods because of the very short overall duration, typically a few milliseconds, of the pulse sequences.

One other variation of the nonexcitation approaches is the combination of selective and nonselective subsequences. It consists of a soft pulse of angle $\beta$ and a hard pulse of angle $-\beta$. The response to the soft pulse is limited to a narrow range around the solvent resonance, which is cancelled subsequently by the hard pulse. The sequence thus nearly provides a flat and phased excitation profile with a gap at the solvent frequency. On the other hand, the inherent asymmetry of the method makes it sensitive to hardware imperfections.

## Solvent suppression using pulsed field gradients

Because of the advances in probe technology, PFGS have been widely used in high-resolution NMR spectroscopy for coherence selection, magnetization destruction and molecular diffusion coefficient measurement. Many of the water suppression techniques and conventional pulse sequences can be, and have been, modified using PFGs to improve suppression and to increase the spectral quality. It is rare to find a newly proposed pulse sequence not including PFGs. Reviews on the usage of PFGs in water suppression and the other specific topics are available in the recent literature.

A PFG pulse applied along the magnetic field ($z$-axis) direction causes the transverse magnetization (coherence) to rotate with an additional phase of

$$\phi(r) = n\gamma\delta G(r) \quad [3]$$

where $n$ is the coherence quantum order, $\gamma$ is the gyromagnetic ratio of the spin, $r$ is the position of the spin in the gradient direction (the $z$-axis in this case), and $G(r)$ and $\delta$ are the gradient strength and duration, respectively. This position-dependent dephasing can be reversed by applying a PFG pulse of the same strength $[\delta G(r)]$ but in the opposite direction. The dephasing and rephasing properties of the PFG pulses can then be used for solvent resonance suppression.

A simple and efficient pulse sequence using PFGs to dephase the water resonance is the RAW (randomization approach to water suppression, **Figure 2A**) method, in which the transverse magnetization of solvent generated by a selective 90° pulse is destroyed immediately by the following strong PFG pulse. The selective pulse can be a Gaussian-shaped soft pulse to provide a narrow-band excitation and to minimize the off-resonance excitation. Other selective pulses could be used as well. The scheme of ‘sel. 90°-$G_z$’ can be used preceding a preparation pulse in most 2D experiments for solvent suppression. To prevent the longitudinal recovery of the water resonance during the PFG pulse, the $G_z$ pulse can be replaced by a scheme of a composite 180° pulse sandwiched by a pair of bipolar PFG pulses, -$G_z$-90°$_x$-180°$_y$-90°$_x$-$G_{-z}$-.

The WATERGATE (water suppression by gradient-tailored excitation) method has proved popular recently because of its high efficiency and short duration compared with the methods using PR or selective nonexcitation. The method resembles a spin-

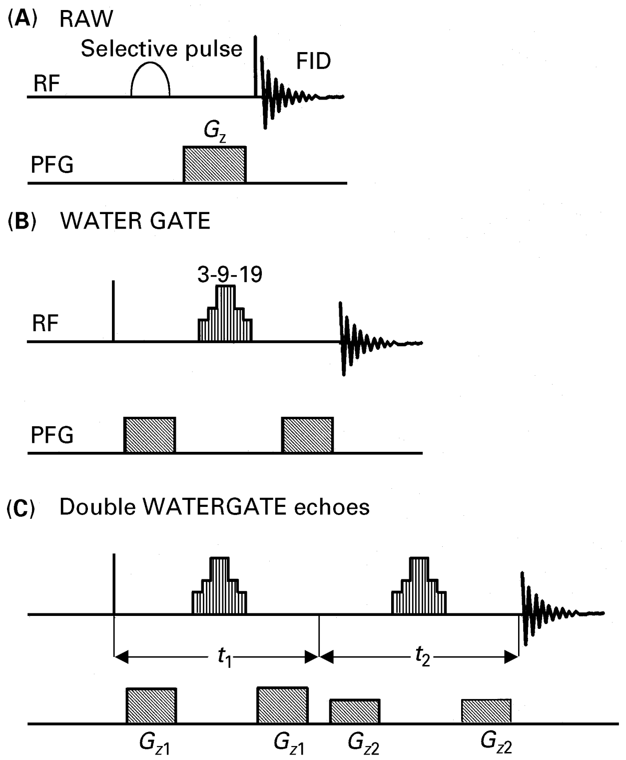



**Figure 2** Water suppression pulse sequences using PFG to selectively dephase the solvent resonance, where the bar and open symbols represent 90° pulses. (A) Randomization approach to water suppression (RAW) sequence. (B) WATERGATE (the composition of the 3-9-19 pulse train and its variations are listed in **Table 1**). (C) Double WATERGATE echo method. It is recommended that different echo times ($t_1 \neq t_2$) and different gradient strengths ($G_{z1} \neq G_{z2}$) are used.

echo sequence with a selective refocusing pulse flanked by two symmetrical gradient pulses as shown in **Figure 2B**. Transverse coherences are dephased by the first gradient and can be rephased by the second gradient, provided they experience a 180° rotation by the selective pulses or the pulse train, denoted by W. This can be a 180° hard pulse sandwiched by a pair of 90° selective pulses. More commonly, it uses a frequency-selective pulse train of $3\alpha$-$\tau$-$9\alpha$-$\tau$-$19\alpha$-$\tau$-$19\alpha$-$\tau$-$9\alpha$-$\tau$-$3\alpha$ (denoted by 3-9-19 or W3 for short), where $62\alpha = 180°$ and $\tau$ is a short delay that is used to control the null-inversion points ($\pm 1/\tau$ Hz). When either kind of selective refocusing pulse is used, the spectral resonances experience a 180° rotation and will be rephased by the second gradient pulse, whilst the net flip-angle at the water resonance frequency approaches zero and thus the water signal will be dephased by both gradient pulses. When selective pulses are used, the duration of the pulse is about 10 ms if a narrower suppression bandwidth is desired. WATERGATE with a W3 pulse train normally takes less than 5 ms, and the saturation of

exchangeable proton resonances is not too serious. Another advantage of using the W3 pulse train is that the null-points can be easily modified for off-resonance water suppression. The disadvantage of using the hard pulse train is that the peak elimination region is wider than when using presaturation and thus any resonances close to the solvent peak will be suppressed. New pulse trains with a narrower noninversion region have been introduced recently. The improvement is achieved by using four (W4) or five (W5) pairs of hard pulses in the pulse train instead of three pairs as in the original W3 sequence. The experimental results (**Figure 3**) indicate that when more element pulses are used in the pulse train, the non inversion region becomes narrower, and the spectral profile becomes wider and flatter, but this is balanced by some sacrifice of suppression efficiency. The composition of different hard pulse trains and parameters are listed in **Table 1**.

The suppression efficiency may be improved by using the double gradient-echo method **Figure 2C**. Although the method provides much better suppression and phase properties, the intensities of any peaks near the solvent will be reduced because the profile of suppression has a squared form.

For comparison of PFG and PR based suppression methods, **Figures 4A** and **4B** show 500 MHz $^1$H NMR spectra of human blood plasma measured using the pulse sequences of WATERGATE (**Figure 2C**) and of NOESYPRESAT (**Figure 1C**) at 30°C, respectively. The low-field region was enlarged and plotted as an inset for both spectra. The experiments were carried out under identical conditions with the exception of the different pulse sequences. The pulse train of W5 was used as the refocusing pulse for the WATERGATE and the inversion bandwidth was set to 3000 Hz ($\tau$ = 1/3000 s). A 2 s $PR_x$ and a 100 ms $PR_y$ low-power ($\gamma B_1$ = 60 Hz; where $B_1$ is the RF magnetic field) pulse were used for solvent saturation in NOESYPRESAT. Both methods provide a high efficiency of solvent suppression; however, the resonances of labile protons (marked by arrows) were observable in **Figure 4A** but suppressed in **Figure 4B** by the saturation transfer effects.

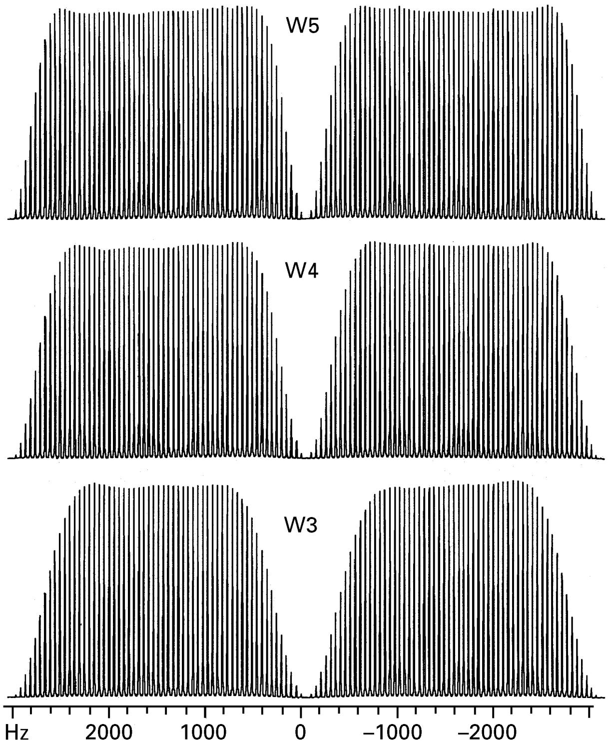



**Figure 3** Experimental excitation profiles of the NMR pulse sequences using W3, W4 and W5 methods. The experiments were carried out at 500 MHz. The bandwidth was set nominally to 3000 Hz ($\tau$ = 1/3000 s).

## Solvent suppression using relaxation, diffusion or multiple quantum filters

Spin and molecular properties that distinguish water from solute signals could be used for the water resonance suppression. The most frequently used properties are the longitudinal relaxation time ($T_1$, also known as the spin–lattice relaxation time), the molecular diffusion coefficient ($D$) and double quantum or multiple quantum coherence filters.

**Table 1** WATERGATE pulse sequence parameters

| *Pulse type* | *Pulse train composition* (deg)[b] | *Phasing for null-points*[c] *at* $k/\tau$ | *Phasing for nullpoints*[c] *at* $(2k+1)/\tau$ |
|---|---|---|---|
| W1 | 90, 90 | $\phi$-$\theta$ [d] | $\phi$-$\phi$ |
| W2 | 45, 135, 135, 45 | $(\phi)_2$-$(\theta)_2$ | $\phi$-$(\theta)_2$-$\phi$ |
| W3[a] | 20.8, 62.2, 131.6, 131.6, 62.2, 20.8 | $(\phi)_3$-$(\theta)_3$ | $\phi$-$\theta$-$(\phi)_2$-$\theta$-$\phi$ |
| W4 | 10.4, 29.4, 60.5, 132.8, 60.5, 29.4, 10.4 | $(\phi)_4$-$(\theta)_4$ | $\phi$-$\theta$-$\phi$-$(\theta)_2$-$\phi$-$\theta$-$\phi$ |
| W5 | 7.8, 18.5, 37.2, 70, 134.2, 134.2, 70, 37.2, 18.5, 7.8 | $(\phi)_5$-$(\theta)_5$ | $\phi$-$\theta$-$\phi$-$\theta$-$(\phi)_2$-$\theta$-$\phi$-$\theta$-$\phi$ |

[a] Original pulse train for WATERGATE sequence.
[b] Each pulse element is separated by a period $\tau$.
[c] $k = 0, 1, 2, 3, \ldots$
[d] $\theta = \phi + \pi$.

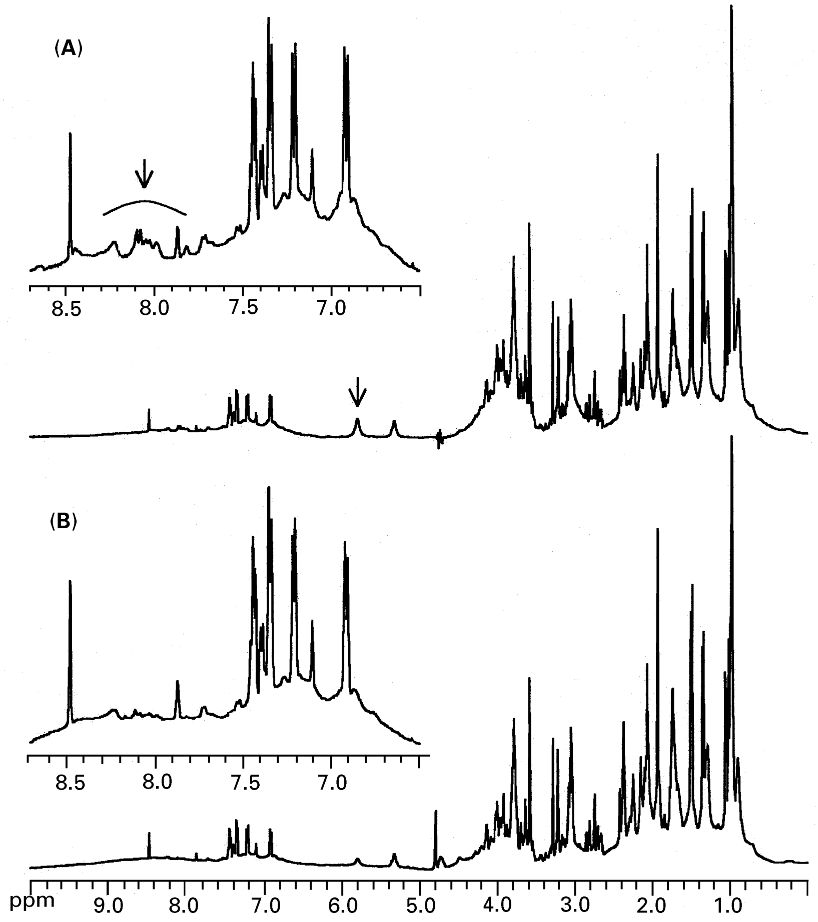



**Figure 4** 500 MHz $^1H$ NMR spectra of human blood plasma obtained using pulse sequences of (A) the double WATERGATE echo method and of (B) NOESYPRESAT under identical conditions. The low-field region is expanded and plotted as the inset. The labile proton resonances (marked by arrows) are observable in (A) but suppressed in (B) by saturation transfer effects.

In biological samples, the longitudinal relaxation times of spins in a protein or other large molecule are often in the region of tens or hundreds of milliseconds, while those of solvent water are 2 to 3 s. Such a large difference in relaxation times makes it possible to suppress the water signal using an inversion recovery scheme (**Figure 5A**). A $\pi$ pulse inverses the magnetization of both solvent and solutes. The remaining transverse magnetization is dephased by the PFG pulse, which also blocks the relaxation pathway associated with the radiation damping effect. The observation pulse is applied when the longitudinal magnetization of water becomes null, after a time $T_1^w \ln 2$, where $T_1^w$ is the longitudinal relaxation time of water. The spins in larger molecules with smaller $T_1$s relax much faster and get closer to their equilibrium magnitude at the time of $T_1^w \ln 2$. For small solute molecules with similar longitudinal relaxation times, the $\pi$ pulse can be replaced by a selective pulse applied at the solvent resonance. Another variation is the use of a series of selective pulses with small flip angles, each of the selective pulses being followed by a PFG pulse. The major advantage of this $T_1$ filter method is its higher selectivity since it provides less attenuation of the resonances close to that of the solvent.

It is also possible to attenuate the water resonance in biofluids and other aqueous samples by the addition of a reagent that causes a significant reduction of the water $T_2$. The broadened water peak can then be attenuated by measuring the spectrum using a spin-echo pulse sequence. This can be achieved for biofluid samples by adding substances containing exchangeable protons that cause water to exchange at an intermediate rate and induce a line broadening. Commonly, guanidinium chloride is used. Alternatively, addition of a paramagnetic ion would also be effective, although care must be taken not to broaden the solute resonances in this case.

**Figure 5B** shows the diffusion coefficient ($D$) filtering method. The method utilizes the difference in molecular mobility between water and solutes such

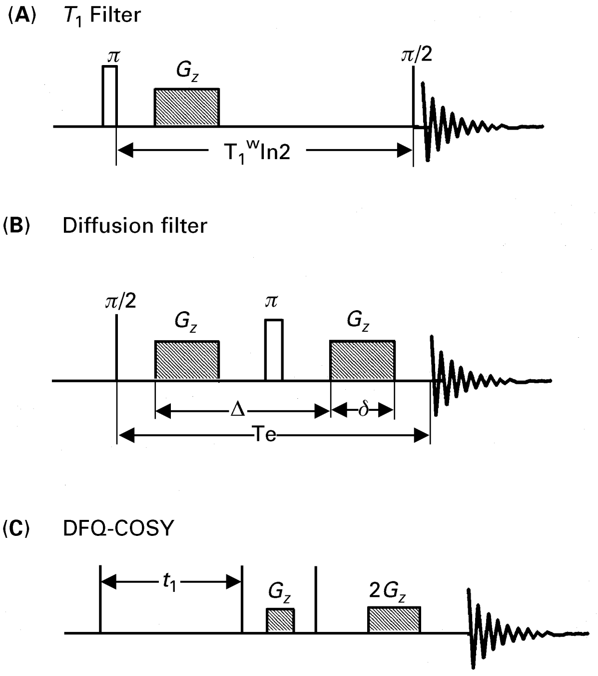



**Figure 5** Solvent suppression pulse sequences based on filtering methods. Method (A) uses a $T_1$ filter to discriminate resonances of solvent and solutes. The difference in molecular diffusion coefficients is used in method (B). $T_e$ is the spin-echo time. (C) Double-quantum filtering COSY, which uses the fact that there is no *J*-coupling between the two equivalent protons in water molecules and thus it cannot be excited to higher quantum coherence. The PFG pulses in (A) and (B) are used to attenuate radiation damping effects and dephase any transverse magnetization. They are used for the desired coherence selection in (C).

as macromolecules. In this experiment, the signal magnitude ($M$) is attenuated according to

$$M = M_0 \exp\left[-D(\gamma\delta\Delta G_z)^2(\Delta - \delta/3)\right] \qquad [4]$$

where $\Delta$ is the time interval between the leading edges of the two PFG pulses, and $G_z$ and $\delta$ are the PFG pulse strength and duration, respectively. The diffusion coefficient of water is more than 10 times that of most larger biological molecules, and thus the water signal will be attenuated, resulting in a spectrum of larger or less mobile molecules. Since $D$ is a molecular property, the attenuation to different spins in a molecule will be the same. This should be useful for quantitative measurement. The diffusion filtering method has now been implemented into most conventional 2D pulse sequences used for diffusion coefficient measurement and for the editing of complex NMR spectra.

The multiple-quantum filtering (MQF) method is an efficient approach to suppression of any resonance that does not experience a specific defined quantum order during the pulse sequence. Since water protons give rise to a single peak in an NMR spectrum and cannot be excited to a quantum order higher than unity, the water resonance cannot pass a MQF and will be eliminated from the spectrum. However, when the MQF is built up based on $^1$H–$^1$H spin coupling, the resulting spectrum is expected to have a dispersive line shape on the coupling splittings, and thus the homonuclear MQF is commonly used for 2D NMR experiments. For example, the pulse sequence of double quantum filtered (DQF) COSY is shown in **Figure 5C**, in which PFG pulses are used to enhance the selectivity. A heteronuclear MQF NMR spectrum often has in-phase line shape, but generally the lower natural abundance of heteronuclei will reduce the sensitivity.

## Postacquisition data processing

As discussed in previous sections, a variety of solvent suppression methods are available. However, even when a high-quality method is used, it is quite common for there to be phase or baseline distortion (or both) in the resulting spectrum, especially in the region of the solvent resonance. In other cases, although the water peak is suppressed, it still remains as the largest peak in the spectrum. This will cause serious ridges (known as '$t_1$ noise') in the $F_1$ dimension of multidimensional NMR experiments and will strongly affect the cross peaks close to it. These disadvantages can be overcome by postacquisition data processing. It has been a standard approach to eliminate a (solvent) signal using a frequency filtering method on most modern NMR spectrometers. Some NMR machine manufacturers also provide software packages of linear prediction and maximum entropy to enhance the quality of NMR spectra and to extract more information from the spectra. However, most off-line data processing methods are focused on automatic baseline correction and convolution, and filtration or subtraction of solvent signal from time- or frequency-domain data sets. The water signal is commonly considered as being a Lorentzian line shape. In some cases it can be treated as a Gaussian function or a weighted Lorentzian–Gaussian function, and can be subtracted from the spectrum or, more commonly, from the time-domain data.

## Radiation damping effects on water suppression

When water is used as solvent in $^1$H NMR spectroscopy, radiation damping is not avoidable, but its effects can be minimized if a proper solvent

suppression method is applied. So far radiation damping has been regarded as a negative effect on $^1$H NMR experiments when water is used as the solvent, although some positive use of radiation damping has been proposed. Radiation damping is a dynamic process similar to the longitudinal relaxation, both leading the magnetization toward the equilibrium state. Physically, radiation damping is caused by the interaction of the FID current with the magnetization itself, and is characterized by a time constant $T_{rd}$, defined by $2(\mu_0\gamma\eta QM_0)^{-1}$, where $\mu_0$ is the vacuum permeability, $\eta$ is the filling factor of the sample in the probe, $Q$ is the quality factor of the detection coil. The large values of $\eta$ and $Q$ of the probe and huge magnetization of water make the radiation damping time very short (in the region of milliseconds), much shorter than its true relaxation time $T_1$ at low concentration (in the region of seconds). As a result, radiation damping interferes with water suppression in many ways.

Among the four experimental strategies discussed above, some could be seriously affected by radiation damping, while others may not be. Because the RF irradiation tends to null the total magnetization, radiation-damping effects will be correspondingly removed. It has been shown that during continuous-wave irradiation, the decay of the magnetization is dependent on $T_1$, $T_2$ and the inhomogeneous contributions from both the static field $B_0$ and the RF field $B_1$, but is independent of radiation damping. The PR method has proven to be a reliable method for water suppression, but with the disadvantage of saturation transfer for exchanging systems. As for the PFG method, if only $z$-gradients are used, radiation damping could occur. Since $z$-gradients can destroy the transverse magnetization only, the remaining longitudinal magnetization, if it is in the $-z$ direction and is still very strong, can easily evolve into a transverse magnetization under the influence of a radiation damping field. As a result, there will be a strong water signal in the spectrum. Care should also be taken when the refocusing PFG is used, since after the refocusing, the transverse magnetization will behave as if PFG were not applied. It should be pointed out that only the positive longitudinal magnetization does not lead to radiation damping. Thus, in order to prevent radiation damping from occurring, PFG experiments should be carefully designed.

If neither PR nor PFG is used, radiation damping may bring about serious problems in water suppression. Because the water magnetization is not attenuated by hard pulses, radiation damping would make the [1, –1] sequence, the inversion-recovery sequence and the DQF-COSY sequence useless, as far as water suppression is concerned. For the simple inversion-recovery method for example, after the spin inversion, the water magnetization returns to the $z$-direction more quickly than the spins in large molecules because of the short radiation damping time. Therefore, it is not possible to use the simple inversion-recovery sequence to measure the relaxation times, nor as a relaxation filter, unless PFG is utilized. In fact, the [1, –1] method, the inversion-recovery method and MQF method have been improved significantly by PFG modifications.

## Summary

Solvent suppression has been one of the rich areas in biological NMR research. The driving force comes from *in vivo* and *in vitro* proton NMR spectroscopy, protein structure determination, metabolic and toxicological studies of biofluids, and medical and functional magnetic resonance imaging. New methods could emerge from the following two fundamental techniques:

1. development of fast switching $B_0$ and $B_1$ gradient coils. A high-quality gradient facility is essential for all types of NMR experiment, including solvent suppression. The use of switching gradients means that transverse magnetization is rapidly dephased, and less signal attenuation caused by diffusion and magnetization transfer results. A properly self-shielded gradient coil can minimize the eddy-current effect which can cause phase distortion throughout the spectrum. The use of magic angle PFGs has proved to be very efficient in the suppression of the solvent resonance and for reducing artefacts in 2D experiments.
2. development of precise and flexible phase- and amplitude-controlled excitation pulses. This facilitates the design of pulse sequences that could produce an exact excitation profile for the solvent peak. Since HPLC–NMR has become more and more routine, the development of double and multiple solvent-peak suppression methods will become important.

Whatever the fundamental technique a new method is based on, it should provide excellent baseline and phase properties and high suppression efficiency, have a short duration to avoid suppression transfer and be easy to be implemented into 1D and $n$D pulse sequences.

## List of symbols

$B_1$ = *RF* magnetic field; $B_0$ = static magnetic field; $D$ = diffusion coefficient; $G(r)$ = gradient strength;

$I_x$ and $I_y$ = transverse magnetizations; $M$ = signal magnitude; $n$ = coherence quantum order; $Q$ = quality factor; $r$ = spin position; $T_1$ = longitudinal relaxation time; $T_1^w$ = water longitudinal relaxation time; $T_{rd}$ = radiation damping time constant; $\beta$ = pulse angle; $\gamma$ = gyromagnetic ratio; $\delta$ = gradient pulse duration; $\Delta$ = time interval; $\eta$ = sample filling factor; $\mu_0$ = vacuum permeability; $\tau$ = pulse delay period; $\omega$ = relative frequency.

*See also:* **Biofluids Studied By NMR; Chromatography-NMR, Applications; Diffusion Studied Using NMR Spectroscopy; Magnetic Field Gradients in High Resolution NMR; NMR Data Processing; NMR Principles; NMR Pulse Sequences; NMR Relaxation Rates; NMR Spectrometers; Product Operator Formalism in NMR; Proteins Studied Using NMR Spectroscopy; Two-Dimensional NMR, Methods.**

## Further reading

Altieri AS, Miller KE and Byrd RA (1996) A comparison of water suppression techniques using pulsed field gradients for high-resolution NMR of biomolecules. *Magnetic Resonance Review* **17**: 27–82.

Gueron M, Plateau P and Decorps D (1991) Solvent signal suppression in NMR. *Progress in NMR Spectroscopy* **23**: 135– 209.

Gueron M and Plateau P (1996) Water signal suppression in NMR of biomolecules. In: Grant DM and Harris RK (eds), *Encyclopedia of Nuclear Magnetic Resonance*, pp 4931–4942. Chichester: Wiley.

Mao X-A and Ye C-H (1997) Understanding radiation damping in a simple way. *Concepts in Magnetic Resonance* **9**: 173–187.

Moonen CTW and Van Zijl PC (1996) Water suppression in proton MRS of humans and animals. In: Grant DM and Harris RK (eds), *Encyclopedia of Nuclear Magnetic Resonance*, pp 4943–4954. Chichester: Wiley.

# Sonically Induced NMR Methods

**John Homer**, Aston University, Birmingham, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Methods & Instrumentation**

Conventionally, nuclear magnetic resonance (NMR) spectroscopy may be viewed as proceeding by photon-stimulated transitions between energy levels of certain nuclei that are quantized due to the influence of a strong homogeneous polarizing magnetic field. However, the required quanta ($h\nu$) of energy necessary to induce transitions need not be restricted in origin to electromagnetic radiation, but may be derived from the phonon, which is the acoustic analogue of the photon. An acoustic wave, which is manifest as sinusoidal pressure variations with a characteristic frequency $\nu(\omega/2\pi)$, can be considered to be composed of a beam of phonons each carrying energy quanta of $h\nu$. The pressure wave and phonon approaches may be invoked to explain various acoustic phenomena. For example, phonons can be considered to be capable of facilitating both experimentally stimulated and naturally occurring nuclear relaxation transitions in solids. On the other hand, the wave approach proves convenient for describing nuclear relaxation processes, and also cavitation, in liquids. Ultrasound can be used experimentally to stimulate NMR transitions, modify nuclear relaxation processes, narrow the resonance bands derived from solids and modify the spectra of solutes dissolved in liquid crystals.

## Principles

### Phonons and nuclear spin-lattice relaxation

Before considering the various possibilities of experimentally manipulating nuclear spin systems with sound it is beneficial to summarize the role of phonons in naturally occurring nuclear relaxation processes.

Debye's theory of the specific heats of solids depends on the existence of a high number of standing, high frequency, elastic waves that are associated with thermal lattice vibrations. Central to his approach is the proposal that in a solid the phonon spectral density ($\rho$) increases continuously, and with a direct dependence on the square of the frequency ($\omega$) to a cutoff frequency ($\Omega$, at about $10^{13}$ Hz) above which the phonon density vanishes: for a solid continuum containing $N$ atoms in a sample of volume $V$, the proportionality constant is $6V/v^3$, where $v$ is the velocity of propagation. At a typical nuclear

Larmor frequency ($\nu_0$) of around $10^9$ Hz there will be a significant phonon spectral density that could facilitate nuclear spin-lattice relaxation. It emerges that the transition probability for such a direct process is very low. However, phonons at frequencies other than $\nu_0$ can contribute to relaxation through indirect processes. Of those possible, one can involve two phonons that could be involved sequentially in absorption or emission, but the probability of this happening is extremely low simply because the numbers of pairs of phonons having energies that combine to match the required transition energy is very low. An alternative indirect process is the so-called Raman process in which a phonon of frequency $\nu$ interacts with the nuclear spins to cause either absorption or emission with the accompanying emergent phonons having frequencies of $\nu-\nu_0$ and $\nu+\nu_0$, respectively. As absorption can only occur when $\nu$ lies between $\nu_0$ and $\Omega$ this process is very much less efficient than emission for which $\nu$ can have any value up to $\Omega$: despite other processes being capable, in principle, of affecting relaxation, the indirect Raman emission process can, therefore, be viewed as being responsible for spin-lattice relaxation in solid samples.

In the case of liquid lattices, the difficulty in adequately characterizing their structures renders them unsuitable for treatment by quantum mechanics. Accordingly, liquid lattices are often treated classically by considering the effects of molecular rotations and translations, with characteristic correlation times, $\tau$, on time-dependent magnetic and electric fields that may influence relaxation processes. Accordingly, it becomes convenient to depart from the phonon description of sound and adopt a classical view of this as the sinusoidal propagation of a pressure wave through a medium.

### Ultrasound

Sound having frequencies above about 18 kHz is traditionally called ultrasound, and in addition to its frequency is characterized by its intensity and velocity of propagation through a medium. As a matter of convenience ultrasound is usually categorized as diagnostic ultrasound (as used for imaging, with frequencies in the megahertz region) or power ultrasound (as used for cleaning, welding etc., with frequencies in the kilohertz range).

The passage of ultrasound through liquids can result in some spectacular phenomena, particularly those resulting from acoustic cavitation. Simplistically, when an acoustic wave passes through a liquid it produces compression and rarefaction of the liquid on successive cycles. On the rarefaction cycle the liquid experiences reduced pressure and, provided a suitable nucleation centre is available, a small cavity will form in the liquid. As migration of entrapped vapour into the cavity depends on the surface area of the interface of the cavity with the liquid this is greater on the rarefaction (expansion) cycle and so the cavity will grow on successive rarefaction cycles. The cavity will either become stable or, over a few acoustic cycles, unstable. In the latter case, the cavity collapses catastrophically with the generation of extreme local temperatures and pressures and the emission of shock waves. If this process occurs near a solid surface, so that there is an imbalance in the force field, a microjet of liquid, starting from the region of the cavity that is most remote from the solid surface, is ejected along the internal symmetry axis of the cavity and strikes the solid surface at velocities in the order of hundreds of m $s^{-1}$. Evidently, ultrasound can be used below the cavitational threshold to pressure modulate the molecular motion in liquids or above the cavitational threshold to subject both liquids and solids to shock waves, liquids to extreme local temperatures and pressures, and solids to the violent impact of microjets.

## Experimentation

Although ultrasound can be generated in a variety of ways, it often proves most convenient to derive it from piezoelectric transducers by applying alternating voltages across opposite silvered faces of discs of the piezoelectric materials (for example, lead zirconium titanate) that have been suitably cut to generate either longitudinal or transverse waves from their surfaces. An alternating signal, applied at the natural resonance frequency of the piezoelectric crystal, causes the surfaces of the crystal to expand and contract at the resonance frequency so that corresponding sinusoidally varying pressure waves are generated and propagated in the desired direction through chosen media.

For many NMR purposes, acoustic waves are often transmitted to liquids using metal (such as titanium) horns that are coupled to one of the piezoelectric crystal faces, or more simply by enabling intimate contact between the crystal and the liquid: to provide acoustic transmission through solids the transducer can be bonded directly to an optically flat surface of the solid.

In the first low frequency (20 kHz) ultrasound/NMR experiments, the ultrasound was delivered to samples in a conventional iron magnet NMR spectrometer using apparatus similar to that shown in **Figure 1**. The titanium alloy horn used was sufficiently long (77 cm) to enable the piezoelectric device

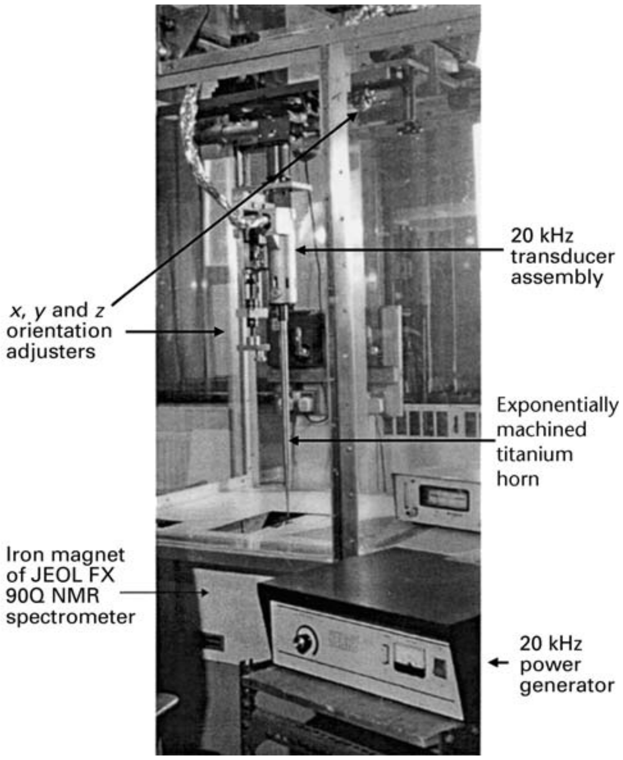



**Figure 1** Early 20 kHz SINNMR apparatus.

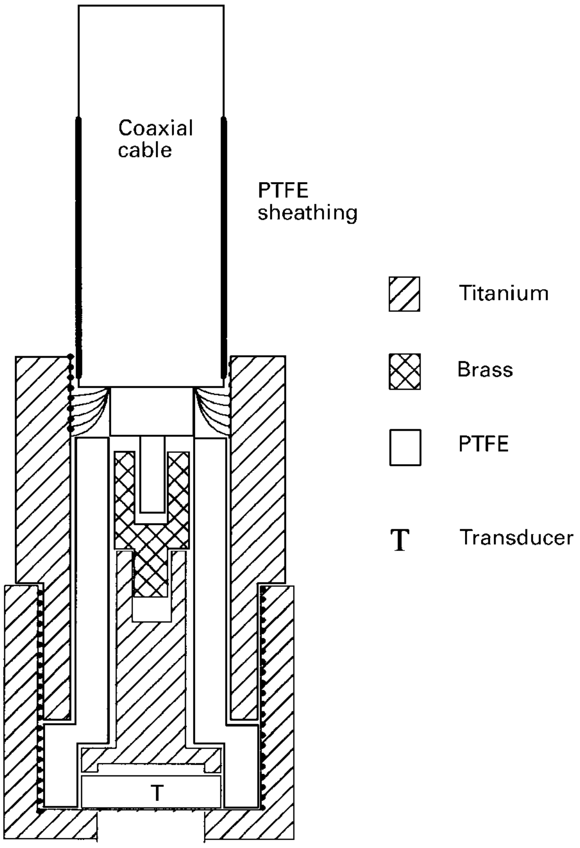



**Figure 2** Schematic diagram of a transducer assembly capable of generating high intensity ultrasound in the megahertz region. Reproduced with permission from RL Weekes, Ph.D. thesis, Aston University, 1998.

to be remote from the NMR detector region as the latter is sensitive to pick-up of extraneous a.c. signals. The horn was machined to provide exponential reduction in its diameter both to provide a coupling tip capable of fitting inside a conventional NMR tube and also to provide mechanical amplification of the ultrasound. Evidently, this equipment with its physically large horn is not suited to insertion down the bore of cryomagnets. Consequently, devices have been produced that are particularly suited to operation with MHz ultrasound and which can be easily inserted into the top of NMR sample tubes in cryomagnets: such a device is illustrated schematically in **Figure 2**. Essentially, this facilitates electrical contact with the piezoelectric transducer by way of compressional contacts and enables the latter to be brought into intimate contact with liquid samples about 1 cm above the active NMR coil region where little pick-up by the latter from the former is experienced: very recently, similar devices to that shown in **Figure 2** have been constructed that allow acoustic irradiation of NMR samples from underneath. By using this general approach NMR/ultrasound experiments have become possible at acoustic frequencies up to 10 MHz. It is worth noting that many high frequency piezoelectric transducer discs have an impedance that can be matched to the output of readily available, and relatively inexpensive, transceiver devices that can, therefore, be used to drive the transducers without extensive modification.

Naturally, when contemplating high acoustic frequency/NMR experiments in cryomagnets it was a matter of priority to resolve two questions. First, does the introduction of ultrasound into the bore of a cryomagnet cause the latter to quench? Extensive experiments have shown that even with acoustic intensities approaching 500 W $cm^{-2}$ the magnets do not quench, although naturally caution is recommended when undertaking such experiments. Second, what is the limit of acoustic frequency at which cavitation can be induced? Using the self-indicating colour-sensitive Weissler reaction as dosimeter it has been shown that cavitation can readily be achieved at frequencies up to 10 MHz, without the involvement of undertones of the transducer natural frequency.

## Applications

By extrapolation of the role of lattice phonons in nuclear relaxation processes in solids it is not a great step to appreciate that the application of ultrasound to both solids and liquids may be used to manipulate phenomena of interest to NMR spectroscopists. Although not yet an area of considerable activity the relatively few examples of the combined use of NMR and ultrasound that will now be described indicate that further such studies will prove profitable.

### Acoustic nuclear magnetic resonance (ANMR)

If a solid is irradiated with ultrasound it produces a phonon spectral density that is much larger than the spectral density arising from natural lattice vibrations at the irradiation frequency. If acoustic irradiation is at the Larmor frequency of dipolar nuclei, the rate of stimulated transitions is increased by a factor of about $10^{11}$, but this is probably insufficient to make the detection of acoustically stimulated transitions detectable. However, for quadrupolar nuclei the ultrasonic transition rates can be increased by a further factor of about $10^4$ and enable the observation of net acoustic energy by ANMR despite the competition from relaxation transitions. The selection rules for allowed transitions due to quadrupolar coupling are $\Delta m = \pm 1$ (but not $m = -\frac{1}{2} \leftrightarrow +\frac{1}{2}$) and ±2. Accordingly, ANMR (which is not susceptible to skin depth problems) can be detected at both $\nu_0$ and $2\nu_0$. Although ANMR can, with difficulty, be detected directly it is usual to detect its effect through the additional saturation of NMR signals that are detected normally through stimulation by RF irradiation.

Although there are many examples of ANMR experiments on solid samples there has been considerable debate as to whether similar experiments are possible using liquid samples. This debate appears to have been resolved by relatively recent work on $^{14}N$ ($\nu_0$ = 6.42 MHz at a magnetic field of 2.1 T) ANMR saturation experiments on acetonitrile and *N*,*N*-dimethylformamide using acoustic frequencies ranging from about 1 MHz to 10 MHz. Only at an acoustic frequency corresponding to the nuclear Larmor frequency was saturation of the $^{14}N$ signal observed, using an acoustic intensity of about 2.5 W cm$^{-2}$.

### Acoustically induced nuclear relaxation

If solids are irradiated with ultrasound having a frequency below the nuclear Larmor frequency, reference to the discussion of the Raman phonon relaxation process indicates that relaxation emission transitions should be favoured and that spin-lattice relaxation times might be reduced. Correspondingly, the acoustic modulation of normal molecular motion in liquids might result in the reduction of $T_1$. The earliest indication that $T_1$ can be reduced by the application of ultrasound derived from work on an aqueous colloidal sol of $As_2S_3$ when, in the 1960s, a reduction in $T_1$ was noted. Since that observation, detailed investigations have been undertaken on the effects of ultrasound, at various frequencies and intensities, on the values of $T_1$ for $^1H$, $^{13}C$ and $^{14}N$ in several liquids and liquid mixtures. Importantly, it has been established that the normal values of $T_1$ in liquids can be reduced by irradiation with ultrasound. Although the acoustic frequency used (1–6 MHz) appeared to have little effect on the observations, it was found that as the acoustic intensity was increased the value of $T_1$ decreased by up to 60% of its natural value, and that the extent of the decrease appeared to correlate roughly with the molecular environment of the nucleus studied. As, for the small molecules studied, the nuclei were all in their extreme narrowing limit (short correlation times) a possible explanation for the reduction in the values of $T_1$ is that their correlation times were increased by the application of ultrasound. This is not inconsistent with the rarefaction–compression effects of the acoustic pressure wave, which may be considered to impose on the molecules a motion that corresponds to a dominant translational correlation time of the order of the inverse of the acoustic frequency (~$10^{-6}$ s). It was also observed that as the intensity of the ultrasound was increased further the values of $T_1$ increased from the minimum value achieved. Although several explanations of this increase in $T_1$ are possible, it is most likely that it arose as a result of the rather crude apparatus used, causing heating of the samples and a normal increase in $T_1$. Recently, further investigations of the acoustic reduction in the values of $T_1$ have been conducted using improved apparatus. The now reproducible results show that $T_1$ for liquid samples can indeed be progressively reduced, to a limiting value, by the systematic increase in intensity of the acoustic field. **Figure 3** shows typical plots of the signal-to-noise ratio of the $^{13}C$ quaternary carbon of 1,3,5-trimethylbenzene in cyclohexane with increasing acoustic intensity at 2 MHz and reflects the progressive reduction of $T_1$. If the explanation for the reduction in the values of $T_1$ for nuclei in the liquid phase is in fact that the ultrasound imposes a dominant translational correlation time on small molecules, an exciting possibility, currently being investigated, is that the choice of a suitable acoustic frequency could be used to modify the correlation times of large biomolecules and hence

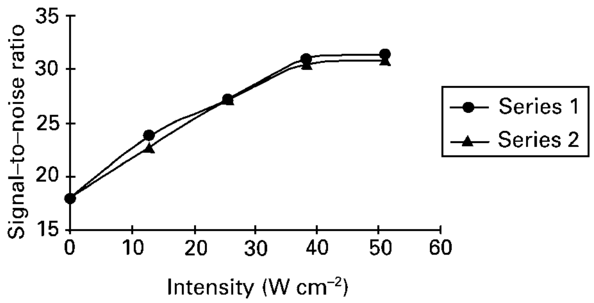



**Figure 3** Dependence on 2 MHz ultrasound intensity of the $^{13}C$ signal-to-noise ratio for the quaternary carbon of 1,3,5-trimethylbenzene in a 1:1 molar mixture with cyclohexane. Reproduced with permission from AL Weekes, Ph.D. thesis, University of Aston, 1998.

reduce the associated values of $T_1$ and speed up their study by NMR. Such an approach, however, has accompanying problems, not least of which is the possibility that the application of ultrasound may cause conformational changes to the macromolecules studied. Similar changes have been observed during studies of *N*,*N*-dimethylacetamide where increasing intensity of 20 kHz ultrasound was found to induce free rotation about the N–C=O single bond and cause averaging of the two *N*-methyl $^1H$ chemical shifts to a single value.

It has also been demonstrated by other workers that ultrasound can reduce $T_1$ for a gadolinium chloride solution, and they, like the originators of the technique, have suggested that the approach might find use in magnetic resonance imaging: this exciting possibility remains to be investigated.

Although less work has been done on the acoustic reduction of $T_1$ in solids than in liquids, it has been established that, as for liquids, the natural values of $T_1$ in solids can be reduced by the application of ultrasound. By coupling 20 kHz ultrasound to a sample of trisodium phosphate dodecahydrate in an open mesh nylon sack immersed in a liquid, the normal value of the $^{31}P$ $T_1$ was reduced from 7.1 s (obtained from MAS NMR measurements) to 2.1 s. Subsequently, similar reductions (by a factor of about two) have been observed for the values of $T_1$ for $^{13}C$ in diamonds to which high frequency piezoelectric transducers were bonded directly.

### Ultrasound and the NMR of liquid crystals

Due to the anisotropy in the molecular magnetic and electrical properties of liquid crystals they can, when in their nematic mesophase, be orientated by the application of external magnetic and electric fields. In the context of NMR this enables liquid crystals containing low concentrations of dissolved solutes to be orientated by the magnet polarizing field. One beneficial consequence of this is that the solutes themselves become orientated and yield NMR spectra which show dipolar spin coupling splittings and which are quite different from their usual isotropic liquid state spectra. These, for example, enable the solute molecules structures to be determined. If a liquid crystal, in its nematic mesophase, is located in a magnetic field alone the molecular director adopts a reasonably well defined orientation with respect to the direction of the applied field. If, in a constant magnetic field, the orientation of the liquid crystal director can be changed, the appearance of the spectra of dissolved solutes should be changed. The possibility of using ultrasound to manipulate the director orientation of both thermotropic and lyotropic liquid crystals in appropriate NMR magnets has been investigated. The $^2H$ spectrum of benzene-$d_6$ dissolved in the nematic liquid crystal ZLI-1167, that normally aligns with its director perpendicular to the directions of the applied magnetic fields, was studied at about 30°C below the clearing temperature, both with and without the application of ultrasound along the bore of a cryomagnet. In the absence of ultrasound, the normal $^2H$ spectrum, composed of a pair of sharp quadrupolar split resonances, was observed. When the sample was irradiated with 2 MHz ultrasound the $^2H$ resonances broadened as shown in **Figure 4**. A temperature gradient within the sample should result in tails between the inner edges of the two resonances the minimum separation between which corresponds to the sample being close to the clearing temperature. On the other hand, a gradient in the director orientation should result in outer trailing edges to the resonances, as observed. Analyses of several such spectra led to the conclusion that the ultrasound, possibly through acoustic streaming, can result in a dispersion of the normal director orientation with a maximum induced change of about 20° from the normal direction.

### SINNMR spectroscopy

It is well known that the width of the naturally broad resonances arising from static solids can be reduced by the coherent averaging processes that result from rapidly spinning samples about so-called magic angles. For dipolar nuclei the magic angle is set at 54° 44′ in the MAS NMR technique, while for quadrupolar nuclei an additional magic angle is used in the double orientation rotor, or DOR, method in order to minimize second order quadrupolar broadening effects. This is not necessary for liquids where the normally rapid and random molecular motion

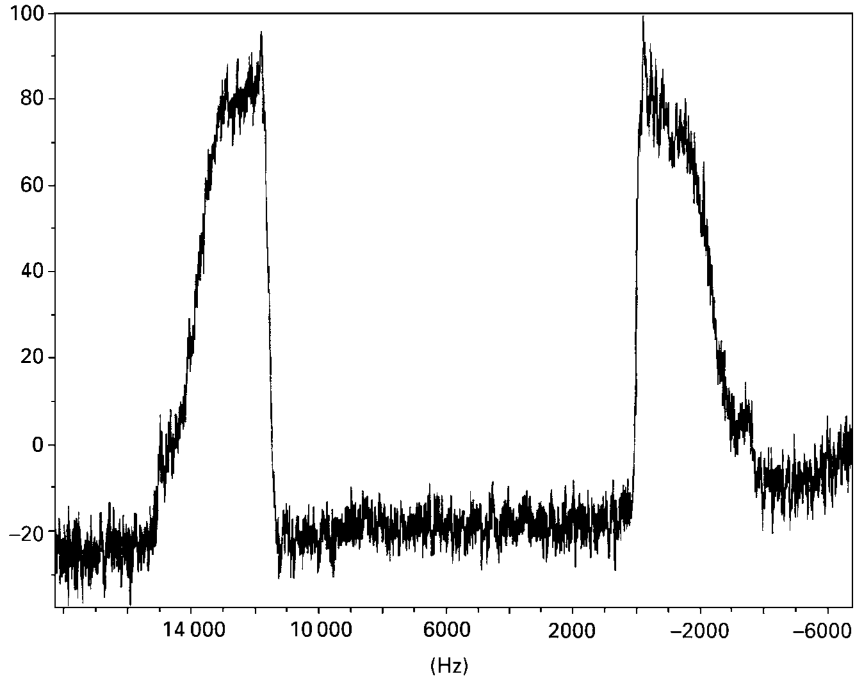



**Figure 4** $^2$H NMR spectrum of benzene-$d_6$ in liquid crystal ZLI-1167 irradiated with 2 MHz ultrasound at ~ 20 W cm$^{-2}$. Reproduced with permission from SA Reynolds, Ph.D. thesis, Aston University, 1997.

leads naturally to narrow lines with isotropic characteristics. Evidently, if particulate matter could be made to mimic the motional characteristics of (large) molecules in the liquid phase, and undergo rapid random motion, it should be possible to narrow the resonances from the solids through an incoherent motional process. If this situation could be achieved without spinning the sample, the production of spinning sidebands, that confuse MAS NMR and DOR spectra, could be avoided.

There are several possible ways of inducing the necessary incoherent motion of particles to facilitate their resonance line narrowing. An obvious way is to produce a fluidized bed in the NMR sample tube, but this has been tried without success. An alternative is to utilize the effects of Brownian motion where molecular bombardment of fine particles in suspension can cause their incoherent motion. The latter approach (ultrafine particle NMR) has been demonstrated using very small particles (nm size) that were perceived to be necessary to respond appropriately to Brownian motion. The necessity to use extremely small particles for this type of experiment is open to question because many experiments in the author's laboratory have shown that micrometre sized particles, suspended in density matched liquids, respond to Brownian motion and yield resonances that are significantly reduced relative to those from static solids.

Another way of inducing appropriate incoherent motion of suspended particles, and producing narrow resonances from solids, is to irradiate the suspension with ultrasound. Whilst this idea relies on the consequences of several phenomena such as streaming, cavitation and shock waves, it may be considered initially from the viewpoint of early theoretical treatments of a single particle subject to an acoustic field. These showed that, for a physically anisotropic particle, the application of an acoustic field will drive the particle to one of three equilibrium orientations of which that with the long particle axis parallel to the direction of the acoustic field is the most stable. Having achieved the most stable equilibrium configuration any motional perturbation from this will be followed by a very rapid return to the equilibrium orientation, i.e. the particle will rotate rapidly through some relatively small angle in a more or less coherent fashion. Obviously this is not what is required to narrow the resonance lines arising from the solid. However, when an assembly of many particles is subject to acoustic irradiation the effects of cavitation producing microjet impact on particles will cause them to rotate, as will shock waves. The sequential effects of these phenomena, together with the effects of induced interparticle collisions, can be adequate to cause rapid incoherent rotation of the particles. These principles are implicit in the sonically induced narrowing of the NMR spectra of solids (SINNMR) technique that was first demonstrated in 1991.

There are so many interdependent parameters (such as support liquid density and viscosity, particle

size, and acoustic frequency and intensity) that govern the success of the SINNMR experiment that it can by no means yet be considered a routine analytical tool. Nevertheless, continuing extensive investigations are slowly revealing the key features of the experiment and some of the findings are worthy of particular note.

The technique was placed on a fairly reproducible basis, using 20 kHz ultrasound delivered via a long titanium alloy acoustic horn to suspensions contained in a normal NMR high resolution tube in an iron magnet based spectrometer. Work on resin-coated (to prevent chemical reaction) 'particles' of aluminium and its alloys resulted in the production of $^{27}$Al SINNMR spectra which revealed resonances of full width at half maximum height (FWHM) as low as 350 Hz: this compares most favourably with the FWHM of about 700 Hz obtained using MAS NMR and the FWHM of about 9000 Hz for a static sample of aluminium. The fact that the SINNMR resonances showed Knight shifts typical of the metallic species appears to provide an unequivocal proof of the validity of the SINNMR experiment. Subsequent studies of the $^{23}$Na and $^{31}$P spectra of trisodium phosphate dodecahydrate provided valuable insight into the SINNMR experiment. $^{31}$P measurements of relaxation times using both MAS NMR and SINNMR revealed that the acoustically induced particle correlation times are of the order of $10^{-7}$ s, which is quite fast enough to cause the motional narrowing of the NMR spectra of the suspended solids. Interestingly, it was shown that the correlation times of the particles reduced as their size increased in support media of decreasing density and viscosity. Somewhat disappointingly, however, these detailed studies revealed that under the conditions employed only about 2% of the solid sample participated in the SINNMR narrowing. Such a low efficiency of SINNMR would appear to render it a poor competitor to MAS NMR and DOR. Nevertheless, the potential value of SINNMR has been established through successful narrowing of $^{11}$B, $^{27}$Al, $^{29}$Si and $^{23}$Na resonances in a range of materials, including glasses. In view of this a concerted thrust has been initiated to both improve the sensitivity of the technique and reduce its complexity.

The most recent work on SINNMR has resulted in the development of the first dedicated acoustic/NMR probehead. This accepts special SINNMR sample tubes that contain a piezoelectric transducer (interchangeable so that a range of acoustic frequencies from 1 to 10 MHz can be used) at its base. This configuration permits the ultrasonic irradiation of particles in less dense support liquids so that the particles can be levitated by the acoustic field and induced to undergo incoherent motion in the NMR coil region. The use of ultrasound in the MHz range should produce smaller cavities and facilitate the study of much smaller particles than those used in the 20 kHz experiments, because the latter generates cavities of about 90 μm and microjetting from these is only effective with particles whose size is greater than the cavity dimensions. The reproducibility and efficiency of the SINNMR experiment has been improved using this dedicated apparatus and the preliminary results are most encouraging: typical SINNMR spectra are shown in **Figure 5**. When designing the dedicated SINNMR probehead and vessel mentioned above, particular attention was devoted to avoiding significant heating of the sample. As a result of this the apparatus can be used, not only for SINNMR studies, but for the reduction in the values of $T_1$ for liquid samples. This is now routinely possible for small molecules and the possibility of inducing similar changes in macromolecules is now the subject of intensive investigation.

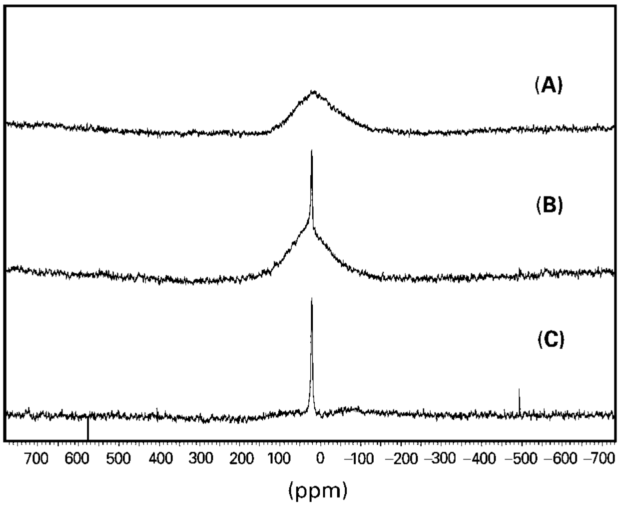



**Figure 5** The $^{23}$Na NMR spectra of particulate trisodium phosphate dodecahydrate suspended in a bromoform/chloroform mixture (A) without acoustic irradiation and (B) and (C) irradiated with 2 MHz ultrasound at ~30 and 50 W cm$^{-2}$ respectively. Reproduced with permission from AL Weekes, Ph.D. thesis, Aston University, 1998.

## List of symbols

$N$ = number of atoms; $T_1$ = nuclear spin–lattice relaxation time; $v$ = velocity of propagation; $V$ = sample volume; $\Delta m$ = change in magnetic quantum number: transition selection rule; $\nu_0$ = Larmor frequency; $\nu(\omega/2\pi)$ = characteristic frequency; $\rho$ = phonon spectral density; $\omega$ = frequency; $\Omega$ = cut-off frequency.

*See also:* **Heteronuclear NMR Applications (As, Sb, Bi); Heteronuclear NMR Applications (B, Al, Ga, In, Tl); Heteronuclear NMR Applications (Ge, Sn, Pb); Heteronuclear NMR Applications (La–Hg); Heteronuclear NMR Applications (O, S, Se, Te); Heteronuclear NMR Applications (Sc–Zn); Heteronuclear NMR Applications (Y–Cd); Liquid Crystals and Liquid Crystal Solutions Studied By NMR; NMR Principles; NMR Relaxation Rates; NMR of Solids; Solid State NMR, Methods.**

## Further reading

Abragam A (1989) *Principles of Nuclear Magnetism*. New York: Oxford University Press.

Beyer RT and Letcher SV (1969) *Physical Ultrasonics*. London: Academic Press.

Emsley JW and Lindon JC (1975) *NMR Spectroscopy in Liquid Crystal Solvents*, Oxford: Pergamon Press.

Homer J (1996) Ultrasonic irradiation and NMR. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, Vol 8, pp 4882–4891. Chichester: Wiley.

Homer J, Patel SU and Howard MJ (1992) NMR With Ultrasound, *Current Trends in Sonochemistry*, Cambridge: Royal Society of Chemistry Special Publication No. 116.

Homer J, Paniwnyk L and Palfreyman SA (1996) Nuclear Magnetic Resonance Spectroscopy Combined with Ultrasound. *Advances in Sonochemistry* 4: 75–99.

Suslick KS and Doktycz SJ (1990) Effects of Ultrasound on Surfaces and Solids, *Advances in Sonochemistry* **1**: 197–230.

# SPECT, Methods and Instrumentation

**John C Lindon**, Imperial College of Science, Technology and Medicine, London, UK

Copyright © 1999 Academic Press

**SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS**
**Methods & Instrumentation**

## Introduction

The SPECT technique is part of the armoury of nuclear medicine for the diagnosis of pathological conditions. Unlike other imaging techniques such as X-ray tomography or even some uses of magnetic resonance imaging (MRI), this technique can be used to identify functional abnormalities rather than anatomical disturbances.

SPECT involves the injection into the body of a radioactive pharmaceutical product (a radionuclide) such as technetium-99m or thallium-201 which decays with the emission of gamma rays. Usually the radiopharmaceutical is a protein with the radioactive atom attached and the molecule is designed to have the desired absorption properties for the tissue to be imaged. Some are accumulated into heart muscle and are used for cardiac imaging, whilst others penetrate the brain and are used for studying brain function. Yet others can be targeted to the lungs. Thus a healthy tissue will take up a known amount of the SPECT agent and this then appears as a bright area in a SPECT image. If a tissue is abnormal it is possible that uptake of the radiopharmaceutical will be amplified or depressed according to circumstances and then this will appear as a more intense spot or a dark area, respectively, on the image. This can be interpreted by a nuclear medicine expert in terms of the suspected pathological state.

Like X-ray tomography, SPECT imaging requires the rotation of a photon detector array around the body to acquire data from multiple angles. Using this technique, the position and concentration of the radionuclide distribution can be determined. Because the emission sources (in this case injected radiopharmaceuticals) are inside the body, this is more difficult than for X-ray tomography, where the source position and strength are known because the X-ray source is outside the body. It is necessary to compensate for the attenuation experienced by emission photons from injected tracers in the body and contemporary SPECT machines use mathematical reconstruction algorithms to generate the image taking this into account.

SPECT imaging has lower attainable resolution and sensitivity then positron emission tomography (PET). The radionuclides that are used for SPECT imaging emit a single gamma ray photon (usually about

140 keV), whereas in PET the positron emission results in two high-energy gamma ray 511 keV photons.

## Methods and instrumentation

The technique requires the detection of the gamma rays emitted by the distribution of the radiopharmaceutical in the body. These gamma rays have to be collimated for detection by a gamma ray camera. The collimators contain thousands of parallel channels made of lead with square, circular or hexagonal cross-sections through which the gamma rays pass. These typically weigh about 25 kg and are about 5 cm thick with a length and breadth of about 40 by 20 cm. Models for special high-energy studies can be much more substantial and can weigh up to 100 kg. Such a collimator is termed a parallel-hole collimator and has a resolution which increases with distance from the gamma ray source. The resolution can be altered by using channels of different sizes. By going to smaller channels there is a trade-off in sensitivity. This has to be borne in mind during patient studies as it has an effect on scan times. Other types of collimator have been developed and these include converging hole collimators. Collimators are positioned above a very delicate single crystal of sodium iodide which is the heart of a gamma camera. This type of collimator/camera arrangement is called an Anger camera after its inventor.

The gamma rays emitted by the radiopharmaceutical in the body can be scattered by electrons within molecules in the body. This is known as Compton scattering and some such scattered photons are thus lost to the Anger camera because of the deflections caused. Second, the gamma rays can cause a photoelectron effect within an atom in the body (promotion of an electron to a higher orbital or even release of the electron) and again this gamma photon will be lost to the detection process.

Usually, Compton scattering is the most probable cause of attenuation in a SPECT image. Conversely, it is possible for a Compton scattered photon to be deflected into the Anger camera's field of view and in this case there is no information available on where the gamma photon originated and hence no spatial information on the location of the radiopharmaceutical. This process leads to loss of image contrast. A typical Anger camera equipped with a low-enegy collimator detects only about 1 in $10^4$ gamma photons emitted by the radiopharmaceutical and a modern Anger camera has an intrinsic resolution of between 3 and 9 mm.

A gamma ray which passes through the collimator assembly will hit the sodium iodide crystal and generate a light photon which interacts with a grid of photomultiplier tubes behind and which collect the light for further processing. SPECT images are produced from these light signals. Sensitivity has been improved by the introduction of multi-camera SPECT systems. A triple-camera SPECT system equipped with high-resolution parallel-hole collimators can produce a resolution of 4–7 mm. Finally, other types of collimator such as the so-called pinhole type have been designed for imaging small organs such as the thyroid gland or limb extremities and for studies on laboratory animals.

The signals for the detected photons are reconstructed into an image using algorithms originally based on those used for X-ray tomography but which allowed for photon attenuation and Compton scattering. A typical method would be the filtered back-projection approach.

A number of experimental parameters have to be optimized in order to obtain the best SPECT image. These include attenuation, scatter, linearity of detector response, spatial resolution of the collimator and camera, system sensitivity, minimization of mechanical movements, image slice thickness, reconstruction matrix size and filter methods, sampling intervals and system deadtime. In a hospital, calibrating and monitoring these functions are usually performed by a Certified Nuclear Medicine Technician or a medical physicist.

## Applications

SPECT is used routinely to help diagnose and stage the development of tumours and to pinpoint stroke, liver disease, lung disease and many other physiological and functional abnormalities. Although SPECT imaging resolution is not as high as that of PET, the availability of new SPECT radiopharmaceuticals, particularly for the brain and head, and the practical and economic aspects of SPECT instrumentation make this mode of emission tomography particularly attractive for clinical studies of the brain.

*See also:* **MRI Theory; PET, Methods and Instrumentation; PET, Theory; Zero Kinetic Energy Photoelectron Spectroscopy, Theory.**

## Further reading

Brooks DJ (1997) PET and SPECT studies in Parkinson's disease. *Baillieres Clinical Neurology* **6**: 69–87.

Corbett JR and Ficaro EP (1999) Clinical review of attenuation-corrected cardiac SPECT. *Journal of Nuclear Cardiology* **6**: 54–68.

Dilworth JR and Parrott SJ (1998) The biomedical chemistry of technetium and rhenium. *Chemical Society Reviews* **27**: 43–55.
Germano G (1998) Automatic analysis of ventricular function by nuclear imaging. *Current Opinion in Cardiology* **13**: 425–429.
Hom RK and Katzenellenbogen JA (1997) Technetium-99m-labeled receptor-specific small-molecule radiopharmaceuticals: recent developments and encouraging results. *Nuclear Medicine and Biology* **24**: 485–498.
Krausz Y, Bonne O, Marciano R, Yaffe S, Lerer B and Chisin R (1996) Brain SPECT imaging of neuropsychiatric disorders. *European Journal of Radiology* **21**: 183–187.
Kuikka JT, Britton KE, Chengazi VU and Savolainen S (1998) Future developments in nuclear medicine instrumentation: a review. *Nuclear Medicine Communication* **19**: 3–12.
Powsner RA, O'Tuama LA, Jabre A and Melhem ER (1998) SPECT imaging in cerebral vasospasm following subarachnoid hemorrhage. *Journal of Nuclear Medicine* **39**: 765–769.
Ryding E (1996) SPECT measurements of brain function in dementia; a review. *Acta Neurologica Scandinavica* **94**: 54–58.

# Spectroelectrochemistry, Applications

**RJ Mortimer**, Loughborough University, UK

Copyright © 1999 Academic Press.

**ELECTRONIC SPECTROSCOPY**
**Applications**

## Introduction

Spectroelectrochemistry encompasses a group of techniques that allow simultaneous acquisition of spectroscopic and electrochemical information *in situ* in an electrochemical cell. Electrochemical reactions can be initiated by applying potentials to the working electrode, and the processes that occur are then monitored by both electrochemical and spectroscopic techniques. Electronic (UV-visible) transmission and reflectance spectroelectrochemistry has proved to be an effective approach for studying the redox chemistry of organic, inorganic and biological molecules, for investigating reaction kinetics and mechanisms, and for exploring electrode surface phenomena. In this article a selection of representative examples are presented, the emphasis being on the applications of transmission electronic (UV-visible) spectroelectrochemistry to the study of redox reactions and homogeneous chemical reactions initiated electrochemically within the boundaries of the diffusion layer at the electrode–electrolyte interface.

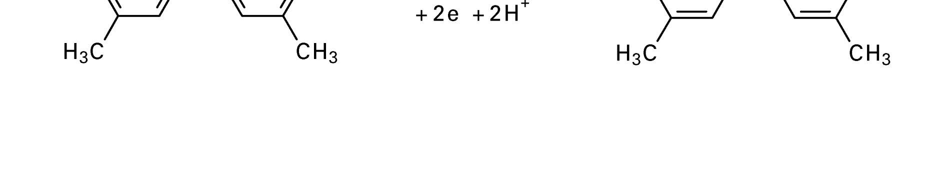
## Organic systems

Many organic systems exhibit redox states with distinct electronic (UV-visible) absorption spectra and are therefore amenable to study with spectroelectrochemical techniques.

### *o*-Tolidine

The technique of transmission spectroelectrochemistry, using an optically transparent electrode (OTE), was first demonstrated in 1964 using *o*-tolidine, a colourless compound that reversibly undergoes a 2-electron oxidation in acidic solution to form an intensely yellow coloured species (Eqn [1]). This system soon became a standard for testing spectroelectrochemical cells and new techniques.

**Figure 1** shows absorbance spectra, for a series of applied potentials, recorded in an electrochemical cell employing an optically transparent thin-layer electrode (OTTLE). Curve a was recorded after application of +0.800 V vs saturated calomel electrode (SCE), which under thin-layer electrode




[1]

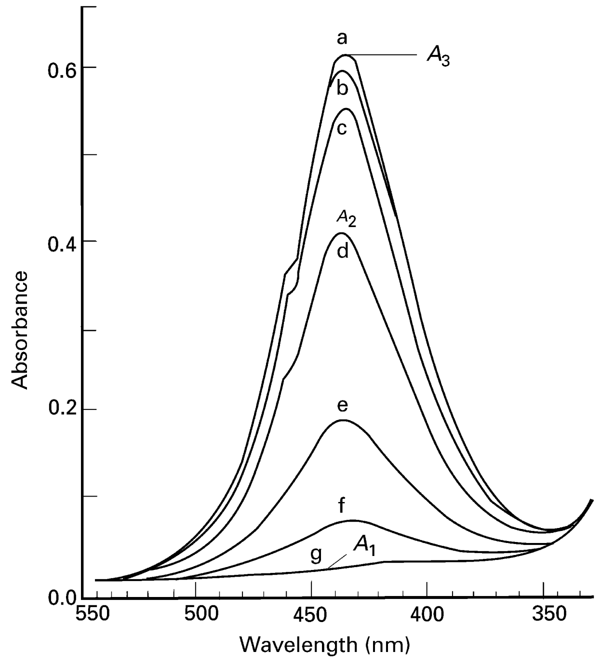



**Figure 1** Thin-layer spectra of 0.97 mM *o*-tolidine, 0.5 M ethanoic acid, 1.0 M $HClO_4$ for different values of $E_{applied}$. Cell thickness 0.017 cm. Potential vs SCE: (a) 0.800 V, (b) 0.660 V, (c) 0.640 V, (d) 0.620 V, (e) 0.580 V, (f) 0.600 V, (g) 0.400 V. Reprinted with permission from DeAngelis TP and Heineman WR (1976) *Journal of Chemical Education* **53**: 594–597. © 1976 Division of Chemical Education, American Chemical Society.

conditions causes complete electrolytic oxidation of *o*-tolidine to the yellow form ([O]/[R] > 1000, where O represents the oxidized form and R the reduced form). Curve g was recorded after application of +0.400 V, causing complete electrolytic reduction ([O]/[R] < 0.001), with the intermediate spectra corresponding to intermediate values of $E_{applied}$. The absorbance at 438 nm reflects the amount of *o*-tolidine in the oxidized form, which can be calculated from the Beer–Lambert law.

Determination of $E^0$, the reversible electrode potential, and $n$, the number of electrons in the *o*-tolidine redox reaction, can be determined from the sequence of spectropotentiostatic measurements (**Figure 1**). For a reversible system,

$$\mathrm{O} + ne \rightleftharpoons \mathrm{R} \qquad [2]$$

the [O]/[R] ratio at the electrode surface is controlled by the applied potential according to the Nernst equation:

$$E_{\text{applied}} = E^0 + \frac{RT}{nF}\ln\left(\frac{[\mathrm{O}]}{[\mathrm{R}]}\right)_{\text{surface}} \qquad [3]$$

In an OTTLE cell, on application of a new potential, the concentrations of O and R in solution are quickly adjusted to the same values as those existing at the electrode surface. Thus, at equilibrium:

$$\left(\frac{[\mathrm{O}]}{[\mathrm{R}]}\right)_{\text{surface}} = \left(\frac{[\mathrm{O}]}{[\mathrm{R}]}\right)_{\text{solution}} \qquad [4]$$

The Nernst equation in a thin-layer cell can then be written as:

$$E_{\text{applied}} = E^0 + \frac{RT}{nF}\ln\left(\frac{[\mathrm{O}]}{[\mathrm{R}]}\right) \qquad [5]$$

For the *o*-tolidine spectra, 438 nm is used as the monitoring wavelength and the ratio [O]/[R] is determined from the Beer–Lambert law:

$$\frac{[\mathrm{O}]}{[\mathrm{R}]} = \frac{(A_2 - A_1)/\Delta\varepsilon b}{(A_3 - A_2)/\Delta\varepsilon b} = \frac{A_2 - A_1}{A_3 - A_2} \qquad [6]$$

where $A_1$ is the absorbance of the reduced form, $A_3$ is the absorbance of the oxidized form, $A_2$ is the absorbance obtained at an intermediate applied potential, $\Delta\varepsilon$ is the difference in molar absorptivity between O and R at 438 nm and $b$ is the light path length in the thin-layer cell. Thus the Nernst equation can be expressed as

$$E_{\text{applied}} = E^0 + \frac{RT}{nF}\ln\frac{(A_2 - A_1)}{(A_3 - A_2)} \qquad [7]$$

**Figure 2** gives $E_{applied}$ vs log([O]/[R]) for the data from **Figure 1**. The plot is linear as predicted from Equation [7], the slope being 0.031 V, which corresponds to an $n$ value of 1.92, with an intercept of 0.612 V vs SCE.

## Methyl viologen

Mechanistic information is often available from spectroelectrochemical measurements. To illustrate the acquisition of semiquantitative information using a rapid scan spectrometer (RSS), the reduction of methyl viologen (the 1,1′-dimethyl-4,4′-bipyridilium dication) under semi-infinite linear diffusion conditions is presented. Methyl viologen ($MV^{2+}$) undergoes two consecutive one-electron reductions to the radical cation ($MV^{\bullet+}$) and neutral species ($MV^0$) in an EE mechanism. In acetonitrile at an OTE coated with a

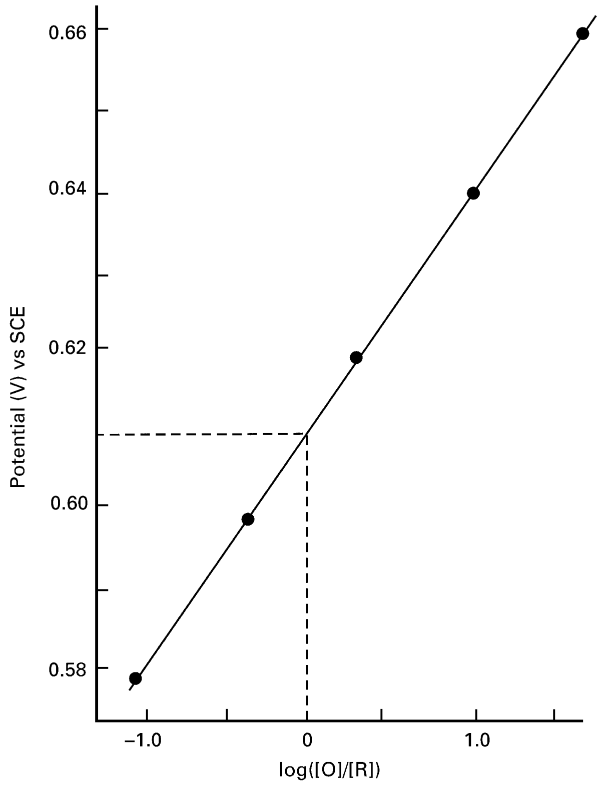
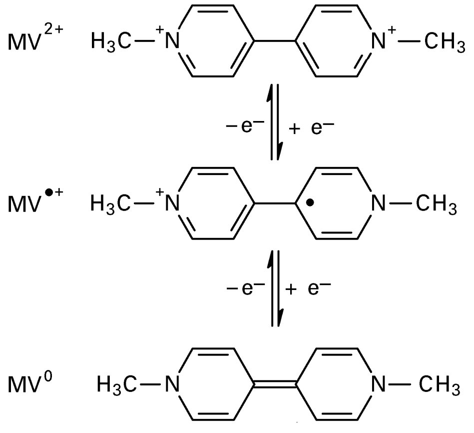



**Figure 2** Plot of $E_{applied}$ vs log([O]/[R]) from spectra in **Figure 1**. Reprinted with permission from DeAngelis TP and Heineman WR (1976) *Journal of Chemical Education* **53**: 594–597. © 1976 Division of Chemical Education, American Chemical Society.

tin oxide film, both waves appear reversible with peak potentials $(E_p)_1 = -0.36$ V and $(E_p)_2 = -0.76$ V vs Ag/AgCl (Eqn [8]).




[8]

If the electrode is stepped some 0.200 V more negative than $(E_p)_1$ during a chronoamperometric experiment, absorbance spectra taken by RSS show two absorbance bands at $\lambda_{max}$ equal to 390 and 602 nm. Interestingly, if the experiment is repeated with the potential stepped slightly beyond $(E_p)_2$, the spectra taken are qualitatively identical to those obtained at $(E_p)_1$. This can be interpreted as due to the equilibrium between the three methyl viologen redox species in the diffusion layer, which greatly favours the radical ion $MV^{\bullet+}$, since $K_{eq} >> 10^3$ for the reaction:

$$MV^{2+} + MV^0 \underset{k_b}{\overset{k_f}{\rightleftharpoons}} 2\,MV^{\bullet+} \qquad [9]$$

Analysis of the spectroelectrochemical working curves for this mechanism shows that when the radical ion is being monitored spectrally, the slopes of the $A$ vs $t^{1/2}$ plots obtained by chronoamperometric reductions at potentials of the first and second waves, respectively, should be in a ratio of 1:1.20. This ratio assumes that the electrode reaction at both waves occurs at the diffusion-controlled rates, and that the three species are in thermodynamic equilibrium in the diffusion layer. The ratio for methyl viologen is 1.21 at $\lambda_{max} = 620$ nm. The larger ratio of 1.79 at $\lambda_{max} = 390$ nm is believed to be caused by band overlap from $MV^0$, which absorbs near the 390 nm band of $MV^{\bullet+}$.

If the chronoamperometric electrolysis is continued beyond several seconds, the rate of growth of the absorbance at the shorter wavelength of 390 nm decreases considerably owing to the formation of a dimer that absorbs near the longer-wavelength band.

### Pyrene reduction

Reduction of the polycyclic aromatic pyrene serves as another excellent example of an EE mechanism where follow-up chemical reactions complicate the overall mechanism (**Figure 3**).

The one-electron reduction to the radical anion produces a ground doublet state with allowed transitions expected in the visible region of the spectrum. Spectra taken by RSS during chronoamperometric reduction at a potential 0.200 V more negative than $E_p$ of the first wave ($E_p = -2.06$ V vs SCE in acetonitrile–TEAP (tetraethylammonium perchlorate) showed only a major band with a wavelength maximum at 492 nm in the visible region (**Figures 3A** and **3C**). Reduction at $E_p$ of the second wave produced spectra with wavelength maxima at 455 and 520–530 nm (**Figure 3A**, curve b). These maxima are similar to those for a spectrum obtained from the chemical reduction of pyrene and attributed to the dianion, except that the long-wavelength band at 602 nm reported earlier is absent. There is doubt, however, that this spectrum is the dianion because the second wave is irreversible; a new oxidation

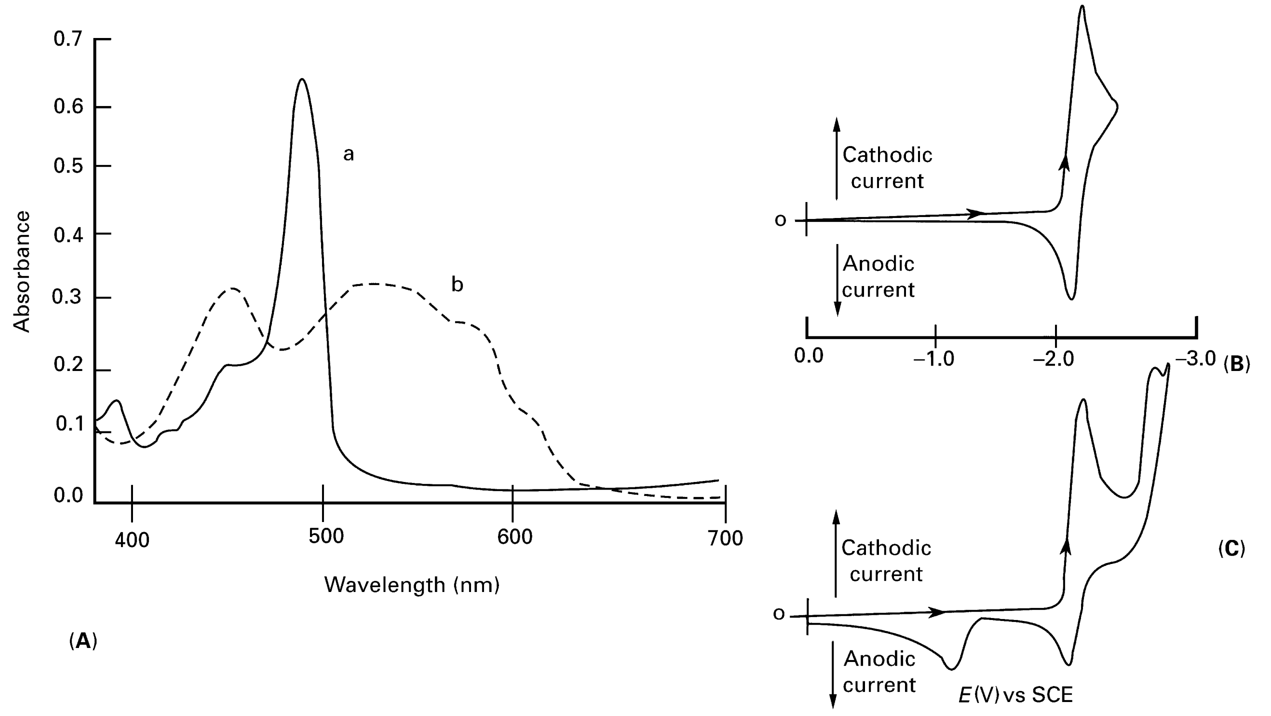



**Figure 3** Spectra and cyclic voltammograms for the reduction of pyrene. (A) Curve a, spectrum of monoanion radical; $\lambda_{max}$ 492, 446, 385 nm; curve b, spectrum obtained from reduction at the second wave for pyrene (see (C)); $\lambda_{max}$ 455 and 520–530 nm. (B) Cyclic voltammogram for the reduction of pyrene to the monoanion radical in acetonitrile–TEAP at a tin oxide OTE. (C) Cyclic voltammogram for the reduction of pyrene. After reduction at the second wave, a new oxidation wave more positive than for the oxidation of the radical appears. Reprinted by courtesy of Marcel-Dekker, Inc. from Kuwana T and Winograd N (1974) Spectroelectrochemistry at optically transparent electrodes. I. Electrodes under semi-infinite diffusion conditions. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 7, pp 1–78. New York: Marcel-Dekker.

wave more positive in potential than the wave for the oxidation of the radical anion appears (**Figure 3C**); and no spectrum due to the free radical appears during chronoamperometric reduction at $E_p$ of the second wave. In any EE mechanism where the waves are sufficiently separated that the equilibrium constant for the disproportionation reaction is large, the equilibrium between the three species (pyrene, radical anion and dianion) would favour the presence of the radical anion in the diffusion layer. The supposed absence of rapid electron exchange between pyrene and dianion to form the radical anion suggests an EEC mechanism in which the dianion undergoes a fast homogeneous chemical reaction to a species more stable than the radical. A likely candidate is the monoanion formed through protonation.

## Inorganic systems

There is a wide range of inorganic systems amenable to study by the spectroelectrochemical approach. In particular, transition metal complexes, with their rich redox state-dependent electronic spectra, have been intensively studied.

### Hexacyanoferrate(III/II)

The hexacyanoferrate(III/II) (ferricyanide/ferrocyanide) system in aqueous solution is a well known electrochemically reversible redox couple (Eqn [10]).

$$[Fe^{III}(CN)_6]^{3-} + e \rightleftharpoons [Fe^{II}(CN)_6]^{4-} \qquad [10]$$

Furthermore, as the hexacyanoferate(III) ion is brilliant yellow in colour and the hexacyanoferrate(II) ion is only very pale yellow, this redox couple is particularly suited as a model system for electronic (UV-visible) absorbance spectroelectrochemical studies.

**Figure 4** shows UV-visible absorption spectra recorded in a spectropotentiostatic experiment in an OTTLE cell on reduction of hexacyanoferrate(III) at a sequence of applied potentials. Curve a is at +0.50 V vs SCE reference electrode, where the redox system is in the oxidized state ($[Fe^{III}(CN)_6]^{3-}$/$[Fe^{III}(CN)_6]^{4-}$ > 1000). Curve h is at +0.00 V vs SCE, where the redox system is in the reduced state

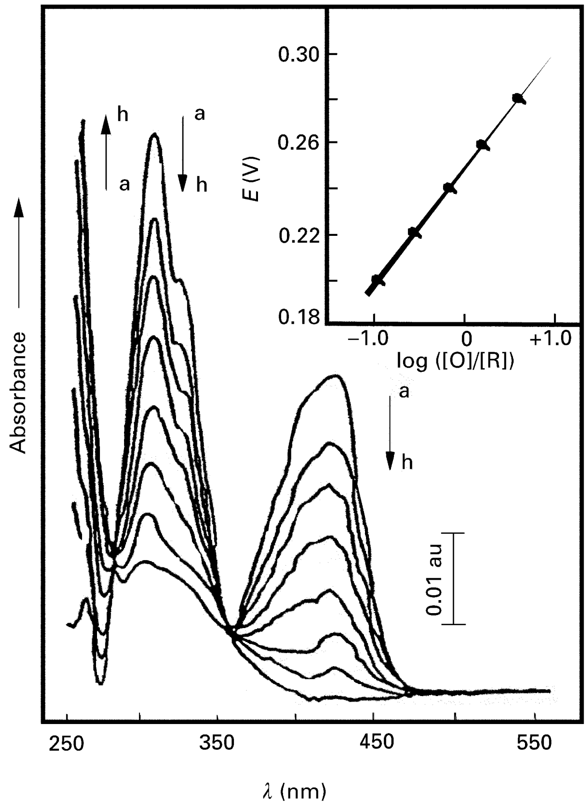
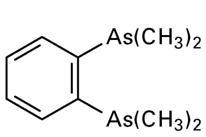



**Figure 4** *In situ* UV-visible absorption spectra of 2.0 mM $K_3Fe(CN)_6$ in aqueous 1M KCl at a sequence of applied potentials vs Ag/AgCl: (a) 0.50 V, (b) 0.28 V, (c) 0.26 V, (d) 0.24 V, (e) 0.22 V, (f) 0.20 V, (g) 0.17 V and (h) 0.00 V. Inset shows the plot of $E_{applied}$ vs log ([O]/[R]): ● at 312 nm and ▲ at 420 nm. Reprinted from Niu J and Dong S (1995) *Electrochimica Acta* **40**: 823–828, © 1995, with permission from Elsevier Science.

($[Fe^{III}(CN)_6]^{4-}/[Fe^{III}(CN)_6]^{3-} > 1000$), while the intermediate spectra correspond to intermediate values of applied potentials. The inset plot in **Figure 4** demonstrates the reversibility of this system in accordance with Equation [7].

### $[Tc^{III}(diars)_2Cl_2]^+$

The complex $[Tc^{III}(diars)_2Cl_2]^+$ (diars = [1]) provides another example of a reversible redox couple for which the spectropotentiostatic method has been applied.




[1] Diars

**Figure 5** shows a thin-layer cyclic voltammogram for this system and **Figure 6** gives a series of spectra for a spectropotentiostatic experiment. Each spectrum was recorded 5 min after potential application so that $([O]/[R])_{solution}$ is at equilibrium with the electrode potential. Spectrum h is the oxidized form, whereas spectrum a is the reduced form. A Nernst plot from the spectra in **Figure 6** is shown in **Figure 7** ($E^0 = -0.091$ V vs SSCE, $n = 0.99$).

### Polypyridylruthenium(II) complexes

The prospect of developing new materials of relevance to the emerging field of molecular electronics, modelling electron-transfer processes in biological systems and producing new electroactive and photoactive catalysts has led in recent years to considerable interest in transition metal polypyridyl complexes. Two recent examples of the application of the OTTLE spectroelectrochemical technique to the study of these fascinating systems are described here.

**Identification of mixed-valence states in polynuclear polypyridylruthenium(II) complexes** Mixed-valence complexes provide an ideal way of studying electron transfer – the most fundamental process in chemistry – under controlled conditions. Polynuclear complexes

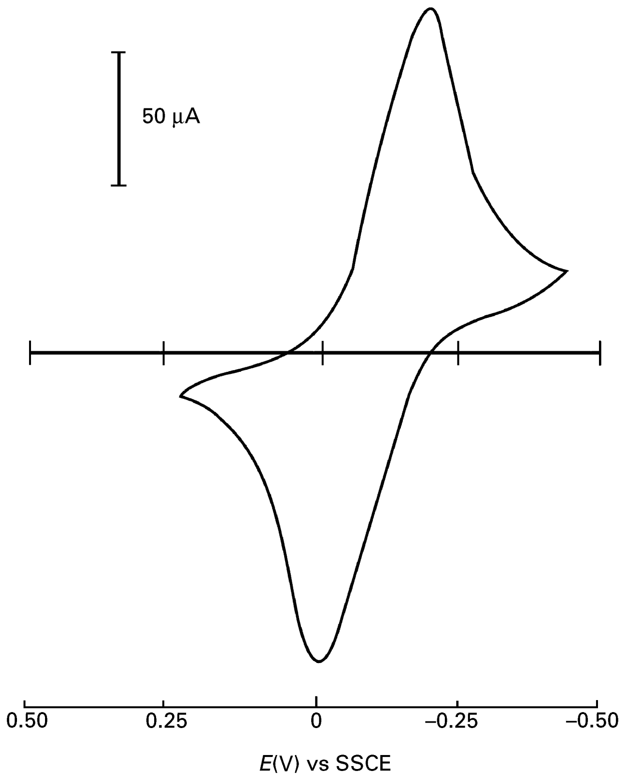



**Figure 5** Thin-layer cyclic voltammogram at 2 mV $s^{-1}$ of 0.87 mM $[Tc^{III}(diars)_2Cl_2]^+$, 0.5 M TEAP in DMF. (SSCE = Sodium chloride saturated calomel electrode.) Reprinted with permission from Hurst RW, Heineman WR and Deutsch E (1981) *Inorganic Chemistry* **20**: 3298–3303. © 1981 American Chemical Society.

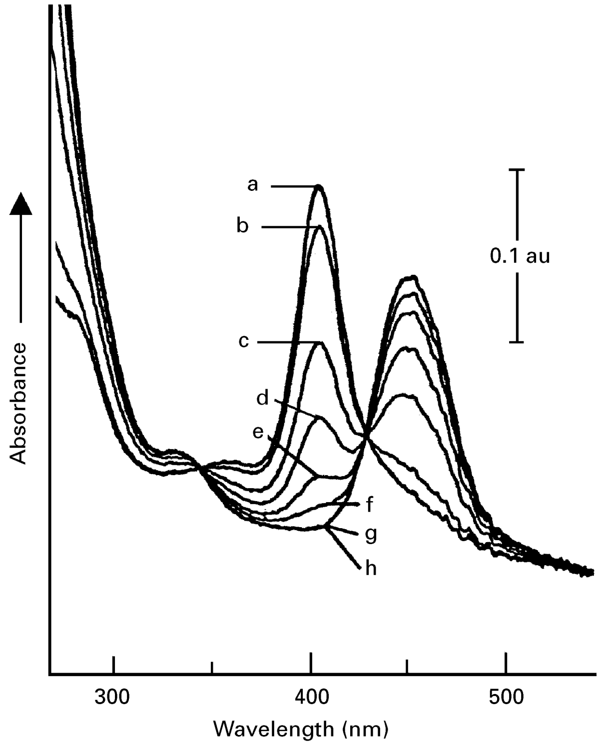



**Figure 6** Spectra recorded during an OTTLE spectropotentiostatic experiment on 0.87 mM $[Tc^{III}(diars)_2Cl_2]^+$, 0.5 M TEAP in DMF. Applied potentials vs SSCE: (a) −0.250 V; (b) −0.150 V; (c) −0.100 V; (d) −0.075 V; (e) −0.050 V; (f) −0.025 V; (g) 0.100 V; (h) 0.250 V. Reprinted with permission from Hurst RW, Heineman WR and Deutsch E (1981) *Inorganic Chemistry* **20**: 3298–3303. © 1981 American Chemical Society.

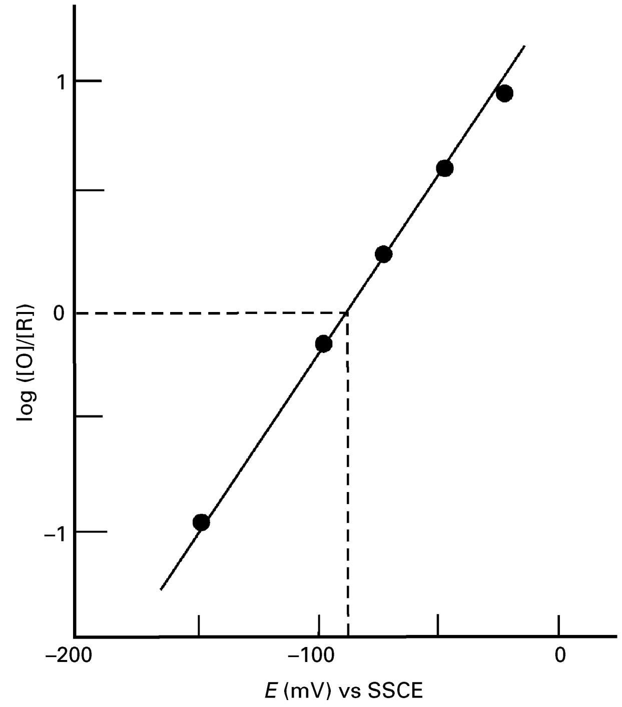



**Figure 7** Nernst plot for spectropotentiostatic experiment on 0.87 mM $[Tc^{III}(diars)_2Cl_2]^+$, 0.5 M TEAP in DMF. Data at 403 nm from **Figure 6** are used. Reprinted by courtesy of Marcel-Dekker, Inc. from Heineman WR, Hawkridge FM and Blount HN (1984) Spectroelectrochemistry at optically transparent electrodes. II. Electrodes under thin-layer and semi-infinite diffusion conditions and indirect coulometric titrations. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 13, pp 1–113. New York: Marcel-Dekker.

containing polypyridylruthenium(II) moieties are of particular interest for the study of mixed valency because of their kinetic inertness in both the +II and +III oxidation states, generally reversible electrochemical behaviour, and good π-donor ability which allows interaction with bridging ligand orbitals. Spectroelectrochemical measurements can be used to probe electrogenerated mixed-valence states in such complexes. A recent example (**Table 1** and **Figure 8**) is the controlled-potential oxidation of the [2,2] species of the complex $[\{Ru(bipy)_2\}_2(\mu\text{-}OMe)_2][PF_6]_2$ in an OTTLE cell. Oxidation of the [2,2] species to the mixed-valence [2,3] state results in the collapse of the metal-to-ligand charge transfer (m.l.c.t) bands at 589 and 364 nm and the generation of a new transition at ~1800 nm ($\varepsilon = 5000$ dm$^3$ mol$^{-1}$ cm$^{-1}$), which disappears on further oxidation to the $Ru^{III}$ state. The observations that this transition is not solvatochromic and that the half-width of the peak is much narrower than the value predicted from Hush theory for vectorial intervalence charge-transfer bands both point to a class III (Robin and Day fully delocalized) mixed-valence state.

**Electronic properties of hydroquinone-containing ruthenium polypyridyl complexes** Ruthenium polypyridyl complexes bound to hydroquinone/quinone moieties are expected to yield information on the behaviour of hydroquinone-type compounds in biological processes. Furthermore, ruthenium(II)–hydroquinone complexes involving O and N bonds are likely to absorb well into the visible region and therefore have potential as dyes in sensitized solar cells. A recent example in the application of spectroelectrochemistry to the study of hydroquinone-containing ruthenium polypyridyl complexes is the oxidation of $[Ru(bipy)_2(HL^0)]^+$ ($H_2L^0$ = 1,4-dihydroxy-2,3-bis(pyrazol-1-yl)benzene) (**Figure 9**).

The spectral changes associated with the first two-electron oxidation step are reversible, and unstable long-lived intermediates are not present, as indicated by the clear isobestic points at 327, 398, 446 and 614 nm (**Figure 9**). After the first two-electron oxidation the m.l.c.t. band at 490 nm blue shifts to approximately 416 nm, and a new feature appears at 700 nm for $[(Ru(bipy)_2(HL^0)]^{2+}$. The presence of

**Table 1** Electronic spectral data for the dinuclear complex $[\{Ru(bipy)_2\}_2(\mu\text{-}OMe)_2][PF_6]_2$ in $CH_2Cl_2$ at 240 K

| *Oxidation state* | $\lambda_{max}$ (nm) ($10^{-3}$ $\varepsilon$($dm^3$ $mol^{-1}$ $cm^{-1}$)) |
|---|---|
| [2,2] | 572 (12), 420 (sh), 359 (15), 293 (79), 242 (58) |
| [2,3] | 1 800 (5), 480 (9), 340 (12), 292 (94), 242 (57) |
| [3,3] | 580 (6), 380 (sh), 248 (64) |

Reprinted with permission from Bardwell DA, Horsburgh L, Jeffrey JC *et al* (1996) *Journal of the Chemical Society, Dalton Transactions*, 2527.

significant absorption features between 400 and 500 nm in the spectrum of the oxidized compound suggests that in the complex the metal centre is still in the ruthenium(II) state, consistent with interpretation from electrochemical data. The oxidized complex is therefore most likely the analogous ruthenium(II)–quinone species. After oxidation of the hydroquinone to quinone, the $Ru^{II} \rightarrow bipy(\pi_2^*)$ m.l.c.t. shifts to the blue as a result of the stabilization of the $t_{2g}$ level when the $\sigma$-donating ability of the ligand is decreased. Further oxidation results in the irreversible loss of the intense feature between 700 and 800 nm and of the band at 416 nm and the generation of a yellow complex likely to be a

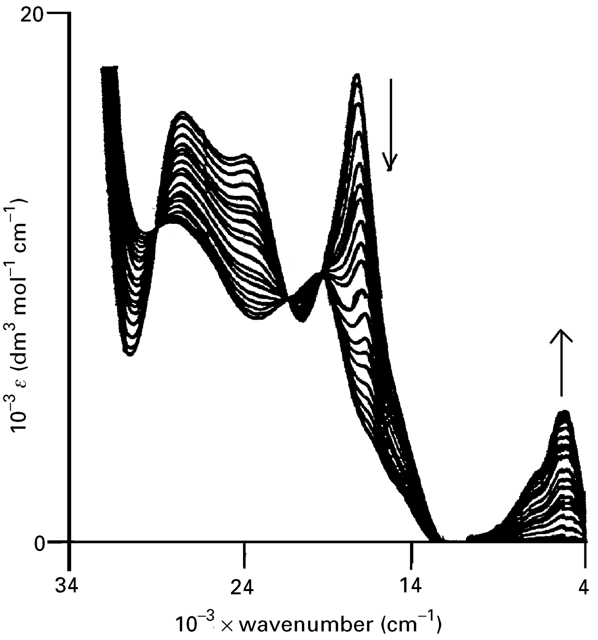



**Figure 8** Successive electronic spectra of the dinuclear complex $[\{Ru(bipy)_2\}_2(\mu\text{-}OMe)_2][PF_6]_2$ in propylene carbonate at 240 K recorded during electrochemical oxidation to the mixed-valence $Ru^{II}Ru^{III}$ state, showing the disappearance of the $Ru^{II} \rightarrow$ m.l.c.t. bands and the appearance of the near-IR band. Reprinted with permission from Bardwell DA, Horsburgh L, Jeffrey JC *et al* (1996) *Journal of the Chemical Society, Dalton Transactions*, 2527–2531.

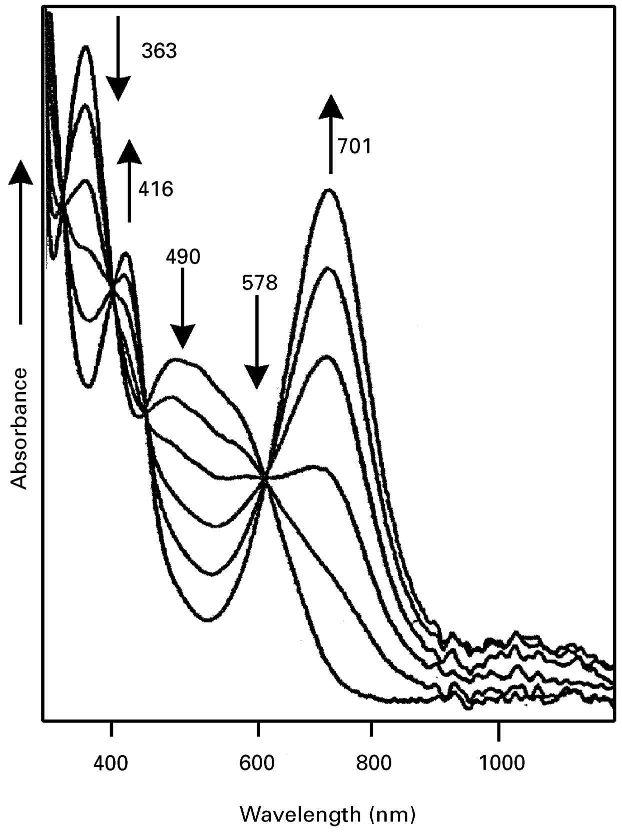



**Figure 9** Spectroelectrochemical oxidation of $[Ru(bipy)_2(HL^0)]^+$ ($H_2L^0$ = 1,4-dihydroxy-2,3-bis(pyrazol-1-yl)benzene) as a function of time between 0 and 20 min. Reprinted with permission from Keyes TE, Jayaweera PM, McGarvey JJ and Vos JG (1997) *Journal of the Chemical Society, Dalton Transactions*, 1627–1632.

complex in which the pyrazole is bound to the ruthenium in a monodentate fashion.

## Biological systems

Numerous biological redox systems have been studied by the spectroelectrochemical approach, including cytochromes, myoglobin, photosynthetic electron transport components, spinach ferrodoxin, blue copper proteins, retinal, and vitamin $B_{12}$ and its analogues. Two classic examples are presented here.

### Vitamin $B_{12}$

Vitamin $B_{12}$ (cyanocob(III)alamin) is an example of a quasi-reversible redox system that exhibits slow heterogeneous electron-transfer kinetics. Cyclic voltammetry alone suggests that the reduction of vitamin $B_{12}$ is a single two-electron process at $E_{pc} = -0.93$ V vs SCE to the Co(I) redox state (**Figure 10A**). However, thin-layer spectroelectrochemistry using a

Hg–Au minigrid OTTLE in a spectropotentiostatic mode reveals that reduction takes place via two consecutive one-electron steps (**Figures 11** and **12**). **Figure 11** shows thin-layer spectra for the reduction to $B_{12r}$, which occurs in the potential range −0.580 to −0.750 V, and **Figure 12** shows the spectral changes for the further reduction to $B_{12s}$, which occurs in the range −0.770 to −0.950 V. Nernst plots for these two reduction processes (using Eqn [7] above) give values of $E_1 = -0.655$ V, $n = 1$ and $E_2 = -0.880$ V, $n = 1$, respectively. The two one-electron reduction processes are clearly shown by the plot of absorbance at 363 nm vs potential in **Figure 10B**, the first one-electron reduction occurring in a region with no apparent cathodic current (**Figure 10A**).

**Cytochrome *c*** Often biological macromolecules will not undergo direct heterogeneous electron transfer with an electrode. Instead, mediator titrants are used that exchange electrons heterogeneously with the electrode and homogeneously with the macro-

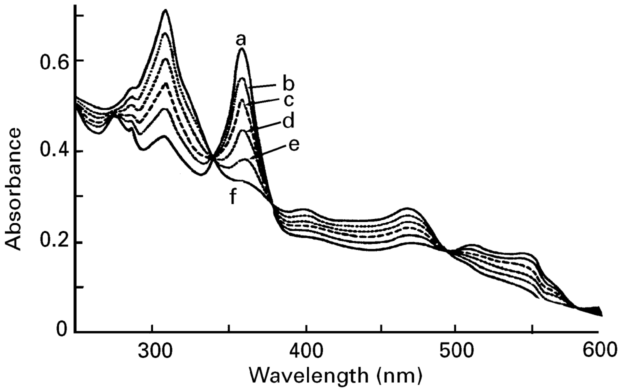



**Figure 11** Thin-layer spectra for reduction of vitamin $B_{12}$ to $B_{12r}$ in a solution of 1 mM vitamin $B_{12}$, Britton–Robinson buffer pH 6.86, 0.5 M $Na_2SO_4$. To obtain the spectra, the potential was stepped in 0.5 mV increments and maintained at each step for 3–5 min until spectral changes ceased. Applied potentials vs SCE: (a) −0.550 V; (b) −0.630 V; (c) −0.660 V; (d) −0.690 V; (e) −0.720 V; (f) −0.770 V. Reprinted by courtesy of Marcel-Dekker, Inc. from Heineman WR, Hawkridge FM and Blount HN (1984) Spectroelectrochemistry at optically transparent electrodes. II. Electrodes under thin-layer and semi-infinite diffusion conditions and indirect coulometric titrations. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 13, pp 1–113. New York: Marcel-Dekker.

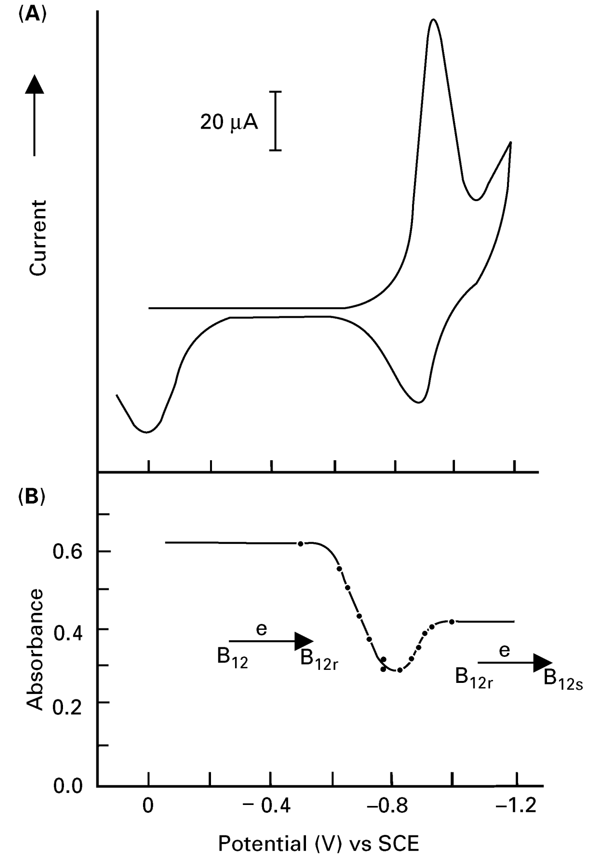



**Figure 10** (A) Thin-layer cyclic voltammogram of 1 mM vitamin $B_{12}$, Britton–Robinson buffer pH 6.86, 0.5 M $Na_2So_4$. (B) Plot of absorbance at 368 nm vs potential, recorded at effectively ~ 0.003 mV $s^{-1}$, from spectra in **Figures 11** and **12**. Reprinted by courtesy of Marcel-Dekker, Inc. from Heineman WR, Hawkridge FM and Blount HN (1984) Spectroelectrochemistry at optically transparent electrodes. II. Electrodes under thin-layer and semi-infinite diffusion conditions and indirect coulometric titrations. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 13, pp 1–113. New York: Marcel-Dekker.

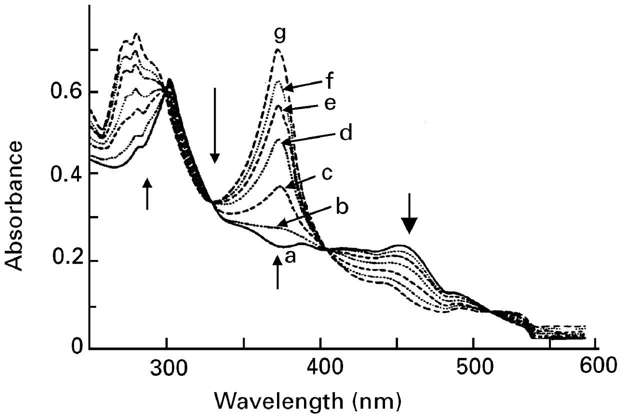



**Figure 12** Thin-layer spectra for reduction of vitamin $B_{12r}$ to $B_{12s}$ in a solution initially of 1 mM vitamin $B_{12}$, Britton–Robinson buffer pH 6.86, 0.5 M $Na_2SO_4$. To obtain the spectra, the potential was stepped in 0.5 mV increments and maintained at each step for 3–5 min until spectral changes ceased. Applied potentials vs SCE: (a) −0.770 V; (b) −0.820 V; (c) −0.860 V; (d) −0.880 V; (e) −0.900 V; (f) −0.920 V; (g) −1.000 V. Reprinted with permission from Rubinson KA, Itabashi E and Mark Jr HB (1982) *Inorganic Chemistry* **21**: 3771–3773. © 1982 American Chemical Society.

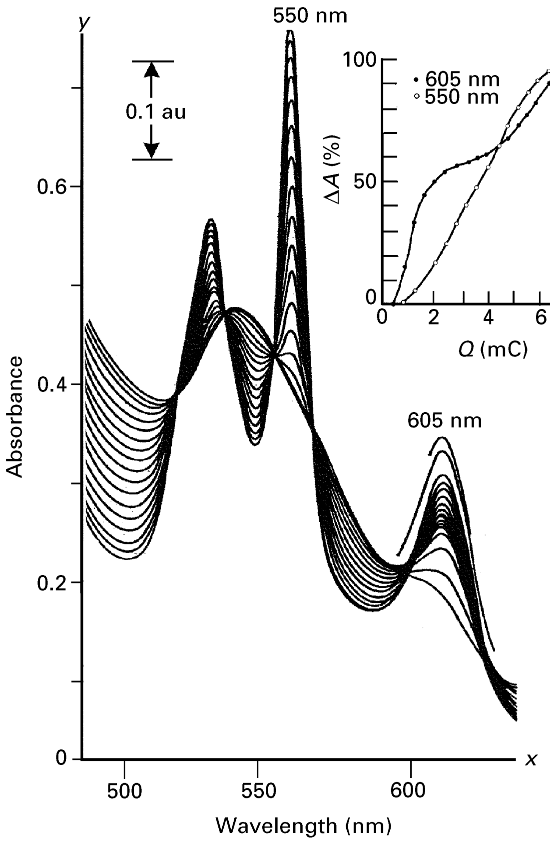



**Figure 13** Spectrocoulometric titration of cytochrome *c* (17.5 μM) and cytochrome *c* oxidase (6.3 μM) by reduction with electrogenerated methyl viologen radical cation ($MV^{\bullet+}$) at a $SnO_2$ OTE. Each spectrum was recorded after $5 \times 10^{-9}$ equivalents of charge (0.5 mC) were passed. Spectra correspond to titration from totally oxidized to totally reduced forms. The final two spectra around 605 nm were recorded after excess $MV^{\bullet+}$ was present. Inset shows titration curves at 550 and 605 nm. Reprinted with permission from Heineman WR, Kuwana T and Hartzell CR (1973) *Biochemical and Biophysical Research Communications* **50**: 892–900.

molecules. **Figure 13** gives spectra obtained for the reduction of a mixture of the haem proteins cytochrome *c* and cytochrome *c* oxidase, both initially in the fully oxidized state.

Each spectrum was recorded after the coulometric addition of $5 \times 10^{-9}$ equivalents of reductant, the methyl viologen radical cation ($MV^{\bullet+}$) electrogenerated at a $SnO_2$ OTE. The reaction sequence is an EC catalytic regeneration mechanism:

$$\text{Electrode} : MV^{2+} + e \rightleftharpoons MV^{\bullet+}$$
$$\text{Solution} : MV^{\bullet+} + \text{cyt } c \text{ (ox)} \rightleftharpoons MV^{2+} + \text{cyt } c \text{ (red)} \quad [11]$$

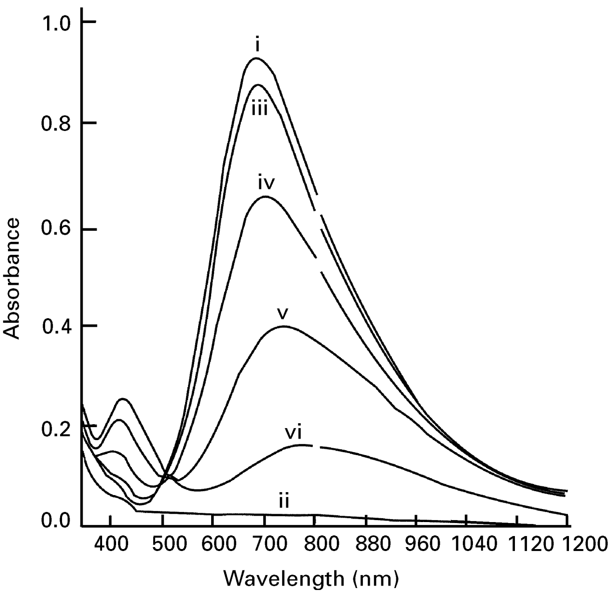



**Figure 14** Spectra of iron hexacyanoferrate films on ITO-coated glass at various potentials [(i) +0.50 V (PB, blue); (i) –0.20 V (PW, transparent); (iii) +0.80 V (PG, green); (iv) +0.85 V (PG, green); (v) +0.90 V (PG, green); (vi) +1.20 V (PX, yellow)] vs SCE with 0.2 M KCl + 0.01M HCl as supporting electrolyte. Reproduced with permission from Mortimer RJ and Rosseinsky DR (1984) *Journal of the Chemical Society*, Dalton Transactions, 2059–2061.

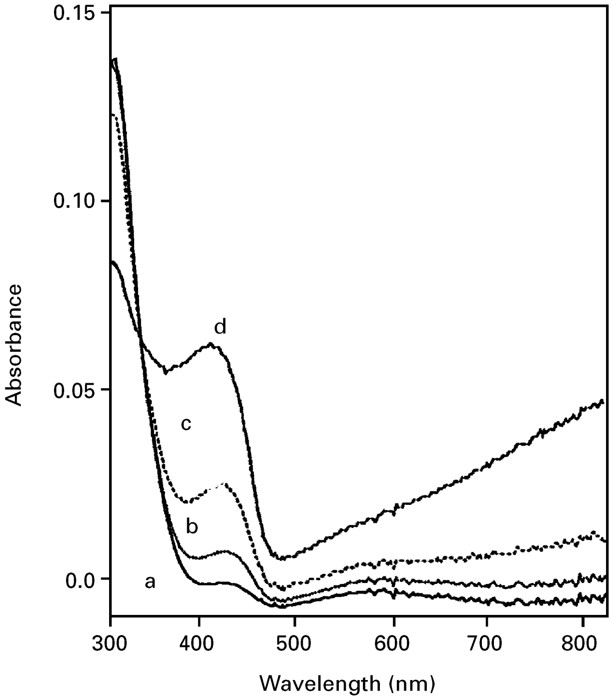



**Figure 15** Spectra of poly(*m*-toluidine) films on ITO in 1M hydrochloric acid at (a) –0.20 V, (b) +0.10 V, (c) +0.20 V, (d) +0.30 V vs SCE. Reproduced with permission from Mortimer RJ (1995) *Journal of Materials Chemistry* **5**: 969–973.

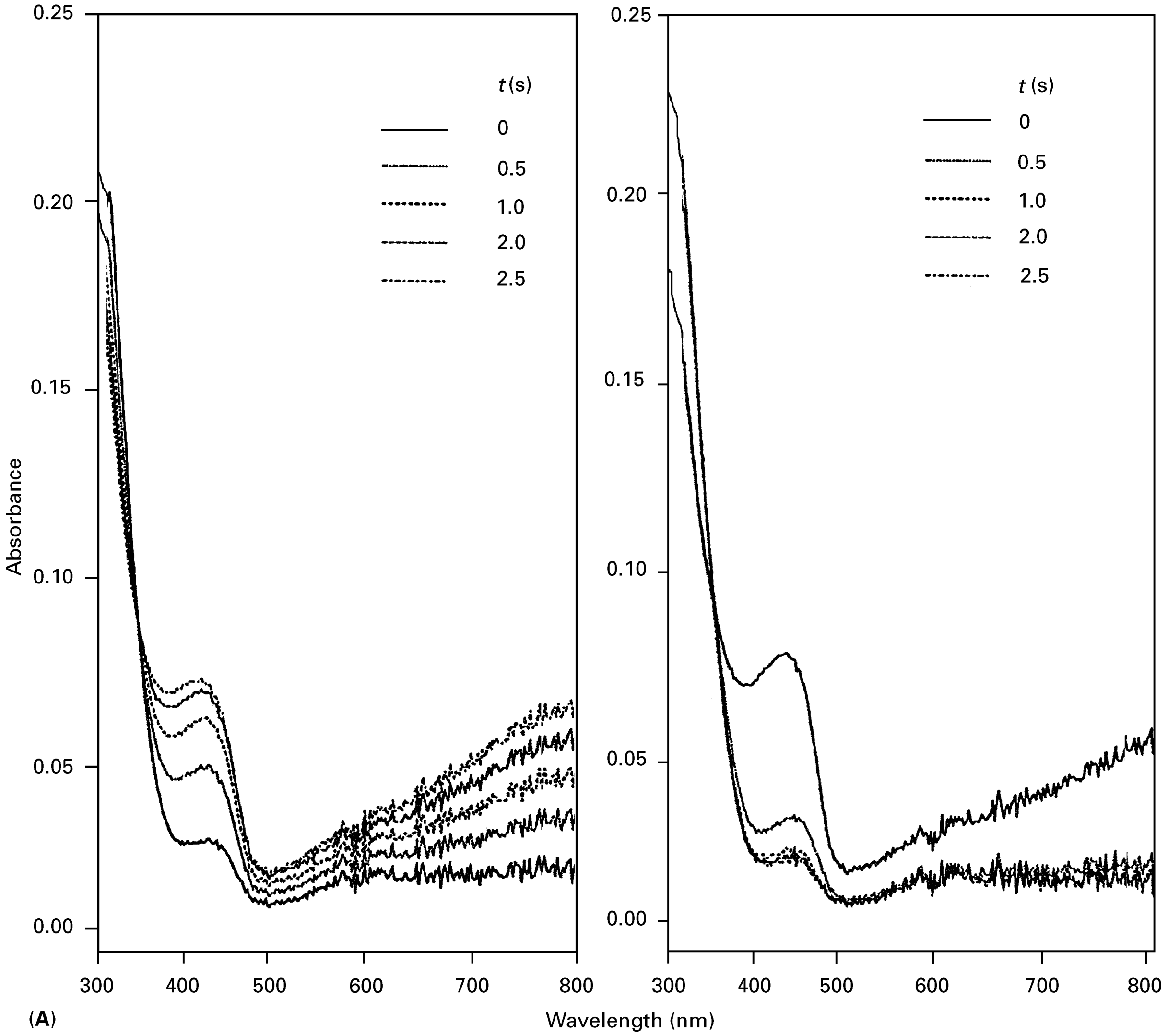



**Figure 16** Spectra recorded at times indicated after potential switching of poly(*m*-toluidine) films on ITO in 1 M hydrochloric acid (A) Potential step −0.20 to +0.40 V vs SCE. (B) Potential step +0.40 to −0.20 V vs SCE. Reproduced with permission from Mortimer RJ (1995) *Journal of Materials Chemistry* **5**: 969–973.

In solution, one $MV^{\bullet +}$ species can reduce a single haem site in cytochrome *c* or one of two in the oxidase. The absorbance increase (**Figure 13**) at 605 nm corresponds to the reduction of the two haem components of cytochrome *c* oxidase; the increase at 550 nm corresponds to the reduction of the haem in cytochrome *c*. Study of plots of absorbance change vs coulometric charge (see inset of **Figure 13**) indicate that $MV^{\bullet +}$ initially reduces one of the haem groups in cytochrome *c* oxidase, then the haem in cytochrome *c*, before it reduces the second haem of the oxidase.

## Modified electrodes

Immobilization of chemical microstructures onto electrode surfaces has been a major growth area in electrochemistry in recent years. Compared to conventional electrodes, greater control of electrode characteristics and reactivity is achieved on surface modification. Potential applications of such systems include the development of electrocatalytic systems with high chemical selectivity and activity, coatings on semiconducting electrodes with photosensitizing and anticorrosive properties, electrochromic displays, microelectrochemical devices for the field of molecular electronics and electrochemical sensors with high selectivity and sensitivity.

Spectroelectrochemical measurements, both *ex situ* and *in situ*, are frequently used in the characterization of modified electrodes. In the case of *in situ* spectroelectrochemical measurements, the modified electrode can be considered to be analogous to an OTTLE, the redox active layer being physically or

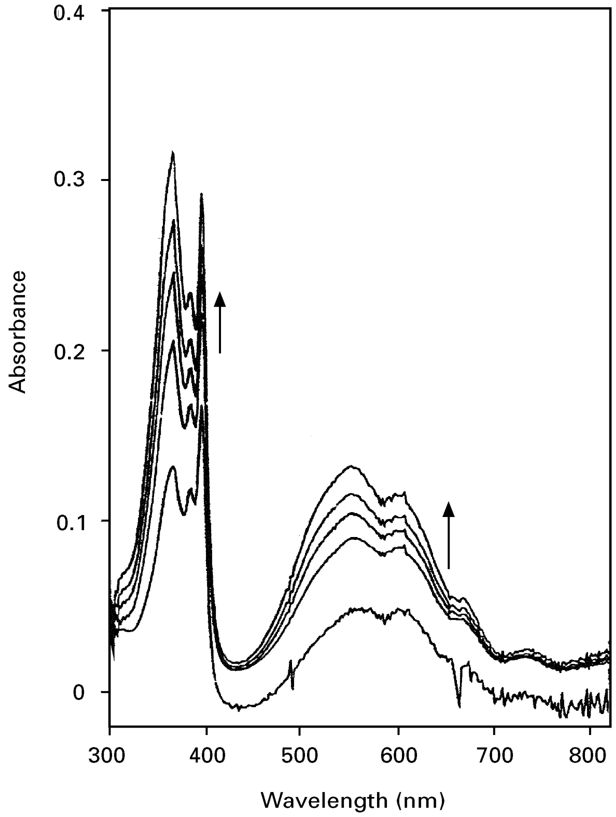



**Figure 17** Spectra recorded at –0.90 V vs SCE during the 2nd, 4th, 6th, 8th and 10th cyclic voltammograms for an ITO/Nafion electrode in 0.1 mM 1,1′-dimethyl -4,4′-bipyridilium dichloride +0.2 M KCl (pH 5.5). The vertical arrows indicate absorbance increase with scan number. For a comparable experiment in the absence of Nafion, the maximum absorbance was <0.01. Reproduced with permission from Mortimer RJ and Dillingham JL (1997) *Journal of the Electrochemical Society* **144**: 1549–1553.

chemically confined to the electrode surface. Electronic spectroelectrochemistry sees significant use in the study of electrodes modified with electrochromic surface films, for which some examples are given below.

## Characterization of electrochromic materials

Chemical species that can be electrochemically switched between different colours are said to be electrochromic. Electrochromism results from the generation of different visible-region electronic absorption bands on switching between redox states. The colour change is commonly between a transparent ('bleached') state and a coloured state, or between two coloured states. In cases where more than two redox states are electrochemically available, the electrochromic material may exhibit several colours and be termed polyelectrochromic. Likely applications of electrochromic materials include their use in controllable light-reflective or light-transmissive devices for optical information and storage, antiglare car rear-view mirrors, sunglasses, protective eyewear for the military, controllable aircraft canopies, glare-reduction systems for offices, and 'smart windows' for use in cars and in buildings.

**Prussian blue** Prussian blue (PB; iron(III) hexacyanoferrate(II)) thin films can be switched to Prussian white (PW) on electrochemical reduction and to Prussian yellow (PX) on oxidation via the partially oxidized Prussian green (PG). For all these electrochromic redox reactions, there is concomitant ion ingress/egress in the films for electroneutrality. The spectra of PX, PG, PB and PW are shown in **Figure 14**, together with two intermediate states between blue and green. The intense blue colour in the $[Fe^{III}Fe^{II}(CN)_6]^-$ chromophore of PB is due to an intervalence charge-transfer (CT) absorption band centred at 690 nm. The yellow absorption band in PX corresponds with that of $[Fe^{III}Fe^{III}(CN)_6]$ in solution, both maxima ($\lambda_{max} = 425$ nm) coinciding with the (weaker) $[Fe^{III}(CN)_6]^{3-}$ absorption maximum. On increase from +0.50 V vs SCE to more oxidizing potentials, the original PB peak shifts continuously to longer wavelengths with diminishing absorption, while the peak at 425 nm steadily increases, owing to the increasing $[Fe^{III}Fe^{III}(CN)_6]$ absorption. The reduction of PB to PW is by contrast abrupt, with transformation to all PW or all PB without pause, depending on the potential that is set. In the cyclic voltammogram of a PB-modified electrode, the broad peak for PB $\rightleftharpoons$ PX in contrast with the sharp PB $\rightleftharpoons$ PW transition emphasizes the range of compositions involved. This difference in behaviour, supported by ellipsometric measurements, indicates continuous mixed-valence compositions over the blue-to-yellow range in contrast with the presumably immiscible PB and PW, which clearly transform one into the other without intermediacy of composition.

**Conducting polymers** Chemical or electrochemical oxidation of numerous resonance-stabilized aromatic molecules including pyrrole, thiophene, aniline, furan, carbazole, azulene and indole produces electronically conducting polymers. In their oxidized forms, such conducting polymers are 'doped' with counteranions (*p*-doping) and possess a delocalized $\pi$ electron band structure, the energy gap between the highest occupied $\pi$ electron band (valence band) and the lowest unoccupied band (the conduction band) determining the intrinsic optical properties of these materials. The doping process (oxidation) introduces polarons (in polypyrrole, for example, these are radical cations delocalized over

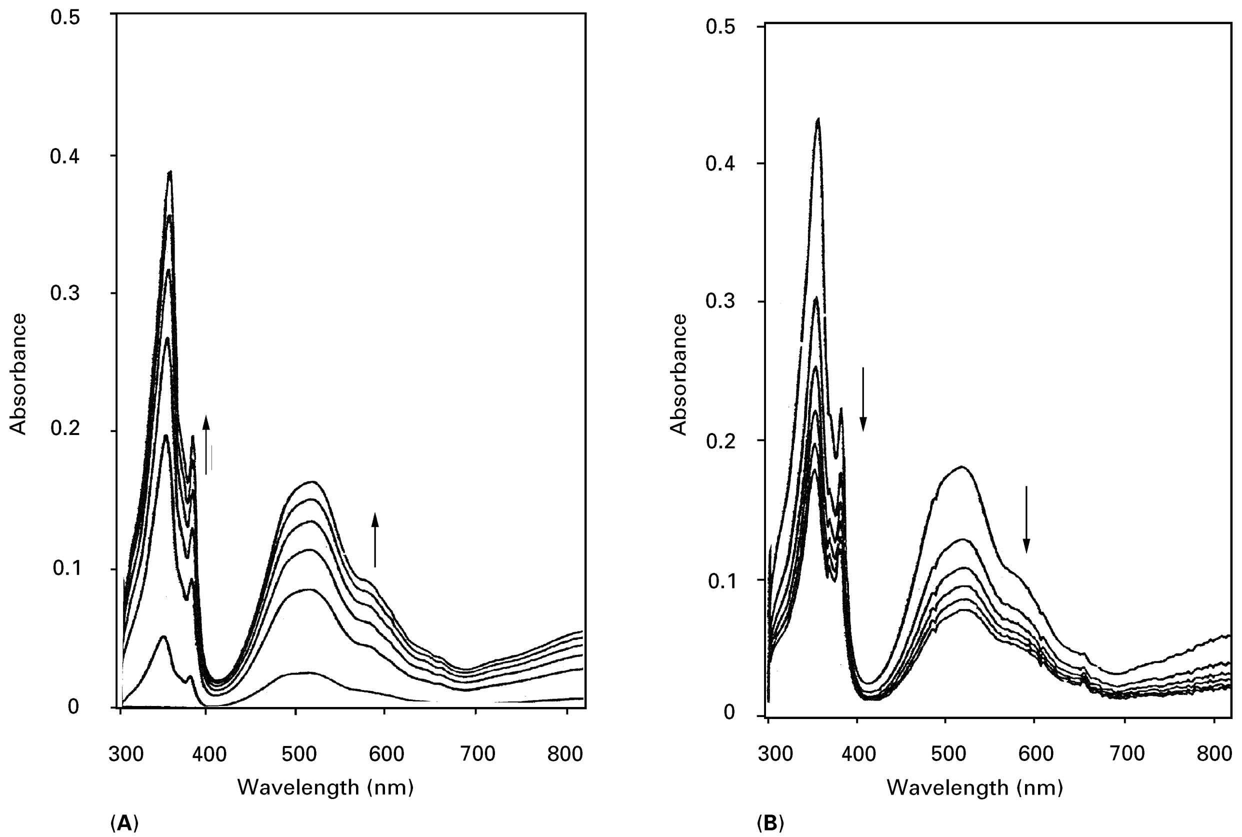



**Figure 18** (A) Spectra recorded at $t = 0$, 10, 20, 30, 40 and 50 s in response to a potential step from +1.00 to –0.90 V vs SCE for an ITO/Nafion/1,1′-di-*n*-hexyl-4,4′-bipyridilium electrode in 0.1 mM 1,1′-di-*n*-hexyl-4,4′-bipyridilium dibromide +0.2 M KCl (pH 5.5). The vertical arrows indicate absorbance increase with time. (B) Spectra recorded at $t = 0$, 10, 20, 30, 40 and 50 s in response to a potential step from –0.90 to +1.00 V vs SCE for an ITO/Nafion/1,1′-di-*n*-hexyl-4,4′-bipyridilium electrode in 0.1 mM 1,1′-di-*n*-hexyl-4,4′-bipyridilium dibromide +0.2 M KCl (pH 5.5). The vertical arrows indicate absorbance decrease with time. Reproduced with permission from Mortimer RJ and Dillingham JL (1997) *Journal of the Electrochemical Society* **144**: 1549–1553.

ca. four monomer units), which are the major charge-carriers. Reduction of conducting polymers with concurrent counteranion exit removes the electronic conjugation, to give the 'undoped' (neutral) electrically insulating form.

All conducting polymers are potentially electrochromic in thin-film form, redox switching giving rise to new optical absorption bands in accompaniment with transfer of electrons/counteranions. Good examples are the polymers of aniline, *o*-toluidine and *m*-toluidine, which are easily prepared as thin films by electrochemical oxidation from aqueous acid solutions of the appropriate monomer. The electrical and electrochromic properties of such polyanilines depend not only on oxidation state but also on the protonation state, and polyanilines are in fact polyelectrochromic (transparent yellow to green to dark blue to black), the yellow–green transition being durable to repetitive colour switching. Spectra for a poly(*m*-toluidine)-modified electrode are illustrated in **Figure 15**.

The two low-wavelength spectral bands observed are assigned to an aromatic $\pi - \pi^*$ transition (≤330 nm) related to the extent of conjugation between the adjacent rings in the polymer chain, and to radical cations formed in the polymer matrix (≤440 nm). With increase in applied potential, the ≤330 nm band absorbance decreases and the ≤440 nm increases (**Figure 15**), the isobestic point indicating that the two species have the same chemical stoichiometry with differences only in electrons. Beyond +0.30 V, the conducting region is entered; the ≤440 nm band decreases as a broad free carrier electron band ~800 nm is introduced. Response times for the yellow–green transition following a potential step can be determined using a diode array spectrophotometer (**Figure 16**).

**Viologens in Nafion** In addition to being important mediator titrants, 1,1′-disubstituted-4,4′-bipyridiliums (viologens) are a major group of electrochromic materials. Electrochromism occurs in bipyridiliums

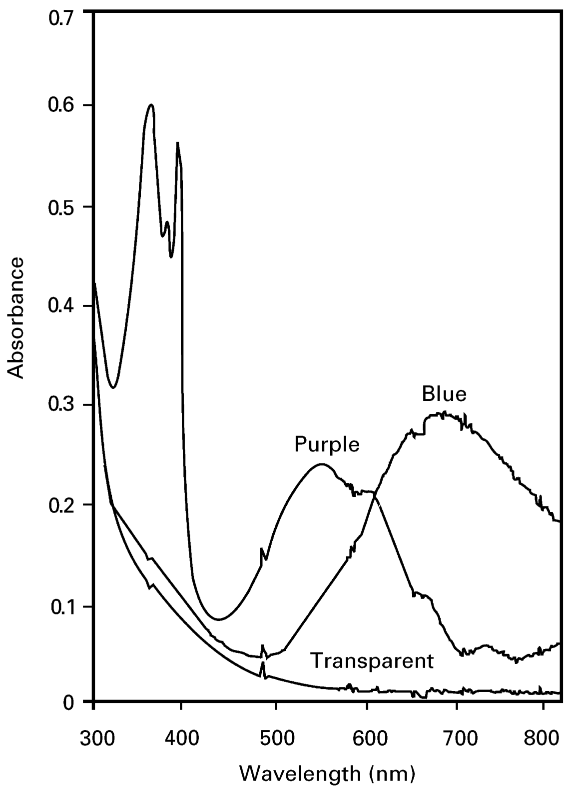



**Figure 19** Spectra recorded at +0.50 V (blue), –0.20 V (transparent) and –0.90 V (purple) vs SCE for an ITO/PB/Nafion/methyl viologen electrode in 0.1 mM 1,1′-dimethyl-4,4′-bipyridilium dichloride +0.2 M KCl (pH 5.5). Reproduced with permission from Mortimer RJ and Dillingham JL (1997) *Journal of the Electrochemical Society* **144**: 1549–1553.

because, in contrast to the bipyridilium dications, the radical cations formed on electroreduction have a delocalized positive charge, coloration arising from an intramolecular electronic transition. Suitable choice of nitrogen substituents to attain the appropriate molecular orbital energy levels can, in principle, allow colour choice of the radical cation. For short alkyl chain length, 1,1′-dialkyl-4,4′-bipyridiliums, both the dication and radical-cation states, are soluble in water and any electrochromic device (ECD) using such bipyridiliums would have the limitation of a low write–erase efficiency. One solution to this problem involves electrostatic binding of bipyridilium dications into anionic polyelectrolyte films. When a Nafion-modified electrode is immersed in an aqueous solution of 1,1′-dialkyl-4,4′-bipyridilium (alkyl = methyl, ethyl, *n*-propyl, *n*-butyl, *n*-pentyl, *n*-hexyl), the 1,1′-dialkyl-4,4′-bipyridilium accumulates in the anionic polyelectrolyte such that its concentration is considerably higher than that in the bulk solution. **Figure 17** illustrates the uptake of 1,1′-dimethyl-4,4′-bipyridilium into a Nafion film monitored by *in situ* spectral measurements. The coloured form in each case is purple, from the presence of monomeric (blue, $\lambda_{max}$ 600 nm) and dimeric (red, $\lambda_{max}$ 500 nm) viologen radical cations. For such viologen-incorporated Nafion films, the electrochromic response times are in excess of 60 s for both coloration and bleaching and independent of viologen size. **Figure 18** shows absorbance spectra measured every 10 s in response to a potential step between the oxidized and reduced forms, for the case of the 1,1′-di-*n*-hexyl-4,4′-bipyridilium system. The longer response time for the oxidation (bleaching) reflects the slower diffusion of radical-cation dimers through the Nafion film compared to that of the monomeric radical cations.

Five-colour polyelectrochromicity, by application of an outer Nafion layer (with subsequent electrostatic incorporation of the methyl viologen system) to an inner layer of PB, is possible (**Figure 19**). The transparent/purple viologen dication/radical cation electrochromicity operates in the potential region where the PB is in its (reduced) transparent state; the bilayer electrode system thus exhibits yellow/green/blue/transparent/purple colours.

## List of symbols

$A$ = absorbance; $b$ = path length; e = electron; $E$ = electrode potential; $E^0$ = reversible electrode potential; $E_p$ = peak potential; $E_{pa}$ = anodic peak potential; $E_{pc}$ = cathodic peak potential; $F$ = Faraday constant (96 485 C mol$^{-1}$); $k$ = rate constant; $K_{eq}$ = equilibrium constant; $n$ = number of electrons; O = oxidized form; R = reduced form; $R$ = gas constant (8.315 J K$^{-1}$ mol$^{-1}$); $t$ = time; $T$ = temperature in kelvin; $\varepsilon$ = molar extinction coefficient; $\lambda_{max}$ = wavelength of maximum absorbance.

*See also:* **Colorimetry, Theory**; **Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy**; **Ellipsometry**; **Spectroelectrochemistry, Methods and Instrumentation**.

## Further reading

Heineman WR and Jensen WB (1989) Spectroelectrochemistry using transparent electrodes – an anecdotal history of the early years. *ACS Symposium Series* **390**: 442–457.

Heineman WR and Kissinger PT (1984) Large-amplitude controlled-potential techniques. In: Kissinger PT and Heineman WR (eds) *Laboratory Techniques in Electroanalytical Chemistry*, Chapter 3. New York: Marcel Dekker.

Heineman WR, Hawkridge FM and Blount HN (1984) Spectroelectrochemistry at optically transparent electrodes. II. Electrodes under thin-layer and semi-infinite

diffusion conditions and indirect coulometric titrations. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 13, pp 1–113. New York: Marcel-Dekker.

Kuwana T and Heineman WR (1976) Study of electrogenerated reactants using optically transparent electrodes. *Accounts of Chemical Research* 7: 241–248.

Kuwana T and Winograd N (1974) Spectroelectrochemistry at optically transparent electrodes. I. Electrodes under semi-infinite diffusion conditions. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 7, pp 1–78. New York: Marcel-Dekker.

Mortimer RJ (1994) Dynamic processes in polymer modified electrodes. In: Compton RG and Hancock G (eds) *Research in Chemical Kinetics*, Vol 2, pp 261–311. Amsterdam: Elsevier.

Mortimer RJ (1997) Electrochromic materials. *Chemical Society Reviews* **26**: 147–156.

Murray RW (1984) Chemically modified electrodes. In: Bard AJ (ed) *Electroanalytical Chemistry. A series of Advances*, Vol 13, pp 191–3\68. New York: Marcel-Dekker.

Niu J and Dong S (1996) Transmission spectroelectrochemistry. *Reviews in Analytical Chemistry* **15**: 1–171.

# Spectroelectrochemistry, Methods and Instrumentation

**Roger J Mortimer**, Loughborough University, UK

Copyright © 1999 Academic Press

**ELECTRONIC SPECTROSCOPY**
**Methods & Instrumentation**

## Introduction

Spectroelectrochemistry encompasses a group of techniques that allow simultaneous acquisition of electrochemical and spectroscopic information *in situ* in an electrochemical cell. A wide range of spectroscopic techniques may be combined with electrochemistry, including electronic (UV-visible) absorption and reflectance spectroscopy, luminescence spectroscopy, infrared and Raman spectroscopies, electron spin resonance spectroscopy and ellipsometry. Molecular properties such as molar absorption coefficients, vibrational absorption frequencies and electronic or magnetic resonance frequencies, in addition to electrical parameters such as current, voltage or charge, are now being used routinely for the study of electron transfer reaction pathways and the fundamental molecular states at interfaces. In this article the principles and practice of electronic spectroelectrochemistry are introduced.

## Cell design

In electronic UV-visible spectroelectrochemistry an optical beam traverses an optically transparent electrode (OTE) in one of the three configurations shown in **Figure 1**.

In transmission spectroelectrochemistry the optical beam is directed perpendicularly through the OTE and the adjacent solution either under semi-infinite linear diffusion conditions (**Figure 1A**) or in a thin-layer cell where diffusion is restricted (**Figure 1B**). Internal reflection spectroscopy (IRS) involves introducing the optical beam through the rear side of an OTE at an angle greater than the critical angle so that the beam is totally reflected (**Figure 1C**). In IRS spectral changes are observable owing to the small penetration of the electric field vector into the solution. Both transmission and reflection spectroscopy have been coupled with numerous electrochemical excitation signals to generate a variety of spectroelectrochemical techniques.

### Optically transparent electrodes

As the working electrode in a spectroelectrochemical experiment, an OTE needs to have both wide optical and potential 'windows', a sufficiently low resistance for good electrode potential control, good stability and surface reproducibility. **Table 1** summarizes the optical and resistance data of some OTEs which are typically thin conducting films on substrate surfaces or minigrids (electroformed mesh).

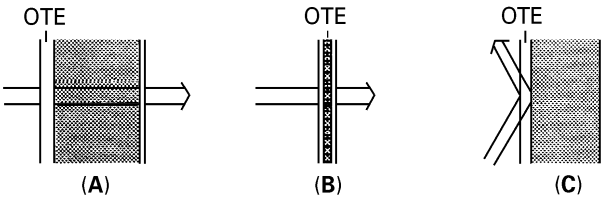



**Figure 1** Schematic diagram of spectroelectrochemical techniques at an optically transparent electrode (OTE). (A) Transmission spectroelectrochemistry; (B) transmission spectroelectrochemistry with an optically transparent thin-layer electrode (OTTLE) cell; (C) internal reflection spectroscopy (IRS). Reprinted by courtesy of Marcel Dekker, Inc. from Heineman WR, Hawkridge FM and Blount HN (1984) Spectroelectrochemistry at optically transparent electrodes. II. Electrodes under thin-layer and semi-infinite diffusion conditions and indirect coulometric titrations. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 13, pp 1–113. New York: Marcel-Dekker.

**Thin conducting films** Any conductor becomes transparent if it is made sufficiently thin, and in OTE design a compromise is usually made between resistance and transmission values (**Table 1**). Most OTEs are prepared by vapour deposition or cold sputtering of a thin film of metal such as platinum or gold or a 'doped' oxide such as tin oxide (Nesa) or indium oxide (Nesatron) on a transparent substrate such as glass, quartz or plastic. OTEs based on thin conducting films were first introduced in 1964 with the use of an antimony-doped tin oxide film on a glass substrate (Nesa glass) in a transmission spectroelectrochemical study of the electrooxidation of *o*-tolidine. Tin-doped indium oxide (ITO) is presently the most common conducting film used for UV-visible spectroelectrochemical studies.

The 'doped' semiconductor oxides, as used in liquid-crystal displays, are particularly attractive owing to their wide potential window (+1.2 to −0.6 V vs saturated calomel electrode (SCE) between solvent oxidation and reduction. Furthermore, the absence of surface oxidation/reduction currents (since these surfaces are already oxidized) is another advantage over platinum and gold OTEs. **Figure 2** shows typical spectra of n-type tin oxide films on various substrates. The optical absorption by the free carriers in the 'doped' tin oxide is in the infrared region, giving transparency in the visible region. Film thickness can be accurately calculated from the interference patterns that are observed in the spectra.

**Minigrids** Minigrid electrodes consist of electroformed wire mesh (40–800 wires cm$^{-1}$) of Au, Pt, Ni, Ag or Cu. The light is transmitted through the microscopic holes between the wires of the minigrid, which in operation functions as a planar electrode after electrolysis has proceeded for sufficient time that the diffusion layer depth becomes large compared to the wire and hole dimensions. The minigrid transmittance varies from 22% to 82%, depending upon the number of wires per centimetre. Since light passes through the holes in the minigrid, the optical window is essentially unlimited. The electrochemical

**Table 1** Optical transmission and electrochemical data on various optically transparent electrodes (OTEs)

| *Type of OTE* | *Transmission range* | *Resistance* ($\Omega$ sq$^{-1}$) |
|---|---|---|
| Pt film (vapour deposited) | 220–near IR, 10–40% | 15–25 |
| Hg–Pt film (electrode-posited Hg) | 220–near IR, 10–30% | 10–25 |
| Au film (vapour deposited) | 220–near IR, 10–80% | 5–20 |
| Sb-doped indium oxide (Nesa) | 360–near IR on glass, 70–85% | |
| | | 5–20 |
| | 240–near IR on quartz, 50–85% | |
| Sn-doped indium oxide (Nesatron) | As Sb-doped indium oxide | 5–20 |
| Au, Hg–Ni and Hg–Au minigrids | UV–visible–IR, 22–80% | < 0.1 |

Reprinted with permission from Kuwana T and Heineman WR (1976) Study of electrogenerated reactants using optically transparent electrodes. *Accounts of Chemical Research* **7**: 241–248 © 1976 American Chemical Society.

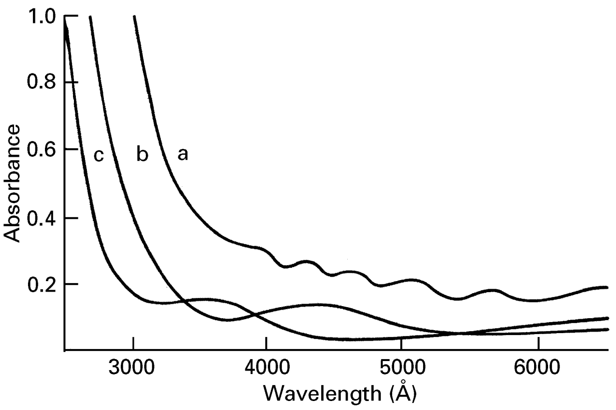



**Figure 2** Transmission spectra of $SnO_2$ coatings on various substrates: (curve a) glass, 3 $\Omega$ sq$^{-1}$; (curve b) Vicor, 6 $\Omega$ sq$^{-1}$; (curve c) quartz, 20 $\Omega$ sq$^{-1}$. Reprinted by courtesy of Marcel Dekker, Inc. from Kuwana T and Winograd N (1974) Spectroelectrochemistry at optically transparent electrodes. I. Electrodes under semi-infinite diffusion conditions. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 7, pp 1–78. New York: Marcel-Dekker.

properties of the gold minigrid are similar to those of the gold film OTE. The negative potential limit of gold and nickel minigrids can be extended by around 0.4 V by deposition of a thin mercury film, which has a high overpotential to hydrogen evolution.

## Transmission spectroelectrochemical cells

Two types of cell geometry can be defined on the basis of the electrolyte solution thickness adjacent to the electrode (**Figure 1A** and **1B**).

**Semi-infinite linear diffusion conditions** The rate of an electrochemical process depends not only on electrode kinetics but also on the transport of species to/from the bulk solution. Mass transport can occur by diffusion, convection or migration. Generally, in a spectroelectrochemical experiment, conditions are chosen in which migration and convection effects are negligible. The solution of diffusion equations, that is the discovery of an equation for the calculation of oxidized form [O] and reduced form [R] concentrations as functions of distance from electrode and time, requires boundary conditions to be assumed. Usually the electrochemical cell is so large relative to the length of the diffusion path that effects at walls of the cell are not felt at the electrode. For semi-infinite linear diffusion boundary conditions, one can assume that at large distances from the electrode the concentration reaches a constant value.

**Semi-infinite linear diffusion spectroelectrochemical cells** In the semi-infinite linear diffusion spectroelectrochemical cell geometry (**Figure 1A**), the cell is analogous to a conventional electrochemical cell, the electrode being in contact with solution much thicker than the diffusion layer adjacent to the electrode. Semi-infinite linear diffusion spectroelectrochemical cell design requires that electrolysis products generated at the counterelectrode should not interfere with the absorbance measurement and that complete deoxygenation should be easily achieved. **Figure 3** shows the classic sandwich-cell design, with a thin film OTE as working electrode. The reference electrode and counterelectrode and the side arms for degassing are positioned so that the cell may be placed with the surface of the OTE in a horizontal plane. A Luggin capillary places the reference electrode near to the surface of the OTE for minimization of solution resistance in the control of the working electrode potential.

These cells are normally used for experiments (chronoamperometry and chronocoulometry) in which large-amplitude steps are applied in order to carry out an electrolysis in the diffusion region. For

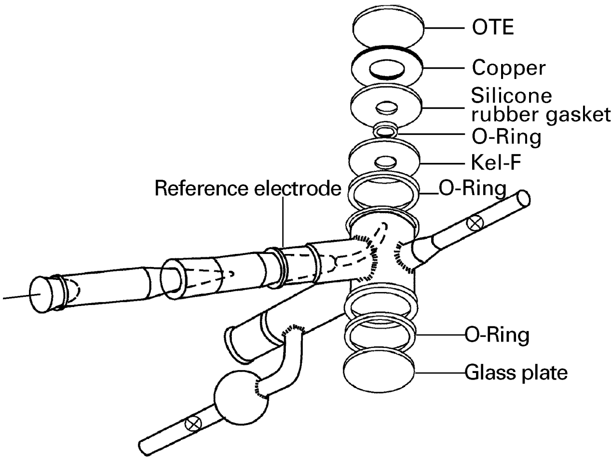



**Figure 3** Sandwich cell for transmission measurement under semi-infinite linear diffusion conditions. Reprinted by courtesy of Marcel-Dekker, Inc. from Kuwana T and Winograd N (1974) Spectroelectrochemistry at optically transparent electrodes. I. Electrodes under semi-infinite diffusion conditions. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 7, pp 1–78. New York: Marcel-Dekker.

the case of a general reduction reaction,

$$\mathrm{O} + n\mathrm{e} \rightleftharpoons \mathrm{R} \qquad [1]$$

the absorbance change can be described by considering a segment of solution of thickness d$x$ and cross-sectional area $A_{\mathrm{elec}}$ (**Figure 4**). If species R of molar absorption coefficient $\varepsilon_{\mathrm{R}}$, is the only species absorbing at the monitored wavelength, then the differential absorbance upon passage of light through this segment is

$$d\mathrm{A} = \varepsilon_{\mathrm{R}}[\mathrm{R}](x,t)\,\mathrm{d}x \qquad [2]$$

The total absorbance is then

$$A = \varepsilon_{\mathrm{R}} \int_0^{\infty} [\mathrm{R}](x,t)\,\mathrm{d}x \qquad [3]$$

If R is a stable species, the integral in Equation [3] is the total amount of R produced per unit area and is equal to $Q/nFA_{\mathrm{elec}}$, where $Q$ is the charge passed in electrolysis, $n$ is the number of electrons and $F$ is the Faraday constant. Since $Q$ is given by the integrated Cottrell equation, which describes the

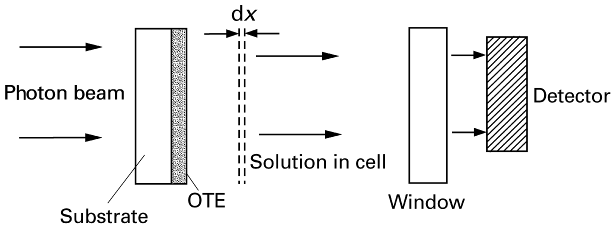



**Figure 4** Schematic view of the experimental arrangement for transmission spectroelectrochemistry.

chronoamperometric response,

$$Q = \frac{2nFA_{\text{elec}}[\text{O}]D_{\text{O}}^{1/2}t^{1/2}}{\pi^{1/2}} \qquad [4]$$

we have

$$A = \frac{2\varepsilon_{\text{R}}[\text{O}]D_{\text{O}}^{1/2}t^{1/2}}{\pi^{1/2}} \qquad [5]$$

which shows that the absorbance should be linear with $t^{1/2}$. Analysis of the slopes of $A$ vs $t^{1/2}$ plots are useful in mechanistic studies where coupled homogeneous reactions follow the initial electrode reaction.

**Optically transparent thin-layer electrochemical cells** The optically transparent thin-layer electrode (OTTLE) cell, first reported in 1967, consists of a sandwich structure with a minigrid OTE working electrode (**Figure** 5). The assembly is placed in a cup of solution containing both the counterelectrode and reference electrode and is filled either by capillary action or by applying nitrogen pressure to give a thin (<0.2 mm) solution layer confined next to the OTE. These cells can easily be constructed in the laboratory using ordinary microscopy slides, 100 μm Teflon adhesive tape spacers, a minigrid and epoxy resin.

A large ohmic potential drop is often present in OTTLE cells owing to the nonuniform current distribution within the thin-layer cavity caused by the large distance between the working electrode and counterelectrode. This is not a problem, as experiments generally involve exhaustive electrolyses, where any ohmic drop can be 'out-waited'. The OTTLE cell design enables the techniques of thin-layer electrochemistry, cyclic voltammetry, controlled potential coulometry and UV-visible spectroscopy to be performed in one unified experiment using a small quantity of solution.

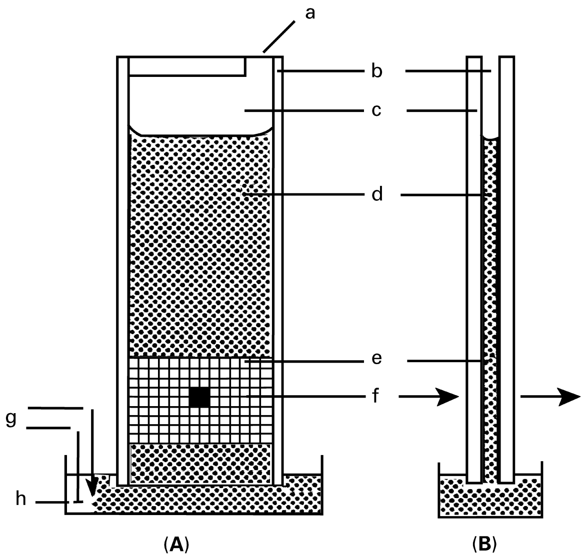



**Figure 5** Optically transparent thin-layer electrode (OTTLE) cell: (A) front view; (B) side view. (a), Point of suction application to change solution; (b) Teflon tape spacers; (c) microscope slides (1 × 3 in.); (d) solution; (e) transparent gold minigrid electrode; (f) optical path of spectrometer; (g) reference and counter electrodes; (h) solution cup. Epoxy resin holds the cell together. Reprinted with permission from DeAngelis TP and Heineman WR (1976) *Journal of Chemical Education* **53**: 594–597. © 1976 American Chemical Society.

**Spectropotentiostatic measurements with an OTTLE cell** Determination of $E^0$, the reversible electrode potential, and $n$, the number of electrons in a redox reactions, can be performed in a sequence of spectropotentiostatic measurements with an OTTLE cell. Generally, for the reversible system given in Equation [1], the [O]/[R] ratio at the electrode surface is controlled by the applied potential according to the Nernst equation:

$$E_{\text{applied}} = E^0 + \frac{RT}{nF}\ln\left(\frac{[\text{O}]}{[\text{R}]}\right)_{\text{surface}} \qquad [6]$$

In an OTTLE cell, on application of a new potential, the concentration of O and R in solution are quickly adjusted to the same values as those existing at the electrode surface, giving at equilibrium:

$$\left(\frac{[\text{O}]}{[\text{R}]}\right)_{\text{surface}} = \left(\frac{[\text{O}]}{[\text{R}]}\right)_{\text{solution}} \qquad [7]$$

The Nernst equation is a thin-layer cell can then be generally expressed as

$$E_{\text{applied}} = E^0 + \frac{RT}{nF}\ln\left(\frac{[\text{O}]}{[\text{R}]}\right) \quad [8]$$

In order to obtain $E^0$ and $n$ values, a series of potentials are sequentially applied to the thin-layer cell containing a test solution. The redox couple is incrementally converted from one oxidation state into another by the applied potentials, resulting in different ratios of [O]/[R] that can be determined spectrally. Each applied potential is maintained until electrolysis ceases, so that the equilibrium value of [O]/[R] is established as defined by the Nernst equation. The $E^0$ and $n$ values can then be obtained by a Nernst plot made by the values of $E_{\text{applied}}$and the corresponding ratios of [O]/[R]. In practice, the range of applied potentials is selected to span $E^0$ of the redox couple, so that spectra for complete reduction/ oxidation and intermediate values of [O]/[R] can be obtained. Selecting a wavelength (usually the maximum wavelength) of O as the monitoring wavelength, for example, the ratio [O]/[R] determined by recording the *in situ* absorbance changes at a certain applied potential is produced using the Beer–Lambert law,

$$\frac{[\text{O}]}{[\text{R}]} = \frac{(A_\text{i} - A_\text{R})/\Delta\varepsilon b}{(A_\text{O} - A_\text{i})/\Delta\varepsilon b} = \frac{(A_\text{i} - A_\text{R})}{(A_\text{O} - A_\text{i})} \quad [9]$$

where $A_\text{R}$ is the absorbance of the reduced from, $A_\text{O}$ is the absorbance of the oxidized form, $A_\text{i}$ is the absorbance obtained at an intermediate applied potential, $\Delta\varepsilon$ is the difference in molar absorptivity between O and R at the selected wavelength and $b$ is the light path length in the thin-layer cell. Substituting Equation [9] into Equation [8], the Nernst equation expressed by absorbance is obtained:

$$E_{\text{applied}} = E^0 + \frac{RT}{nF}\ln\frac{(A_\text{i} - A_\text{R})}{(A_\text{O} - A_\text{i})} \quad [10]$$

The intercept and slope of the straight line of $E_{\text{applied}}$ vs $\ln(A_\text{i} - A_\text{R})/(A_\text{O} - A_\text{i})$ then give $E^0$ and $n$, respectively. Values of $E^0$ and $n$ for numerous reversible redox couples in aqueous, nonaqueous and molten salt solvents have been determined from Nernst plots of such spectropotentiostatic experiments.

### Spectroelectrochemical cells for modified electrode studies

Immobilization of chemical microstructures onto electrode surfaces has been a major growth area in electrochemistry in recent years. In their characterization using *in situ* UV-visible absorption spectroelectrochemical measurements, a modified electrode can be considered to be analogous to an OTTLE, the redox active layer being physically or chemically confined to the electrode surface. Spectroelectrochemical cell design is often simple, a rectangle of OTE being mounted transverse to optical beam direction in a conventional 1 cm cuvette. The counterelectrode, placed opposite the working electrode, is typically a loop of platinum wire through which the light beam can pass. A machined polytetraethylene lid with appropriate holes is used to hold the Luggin capillary from the reference electrode above the light path and the working electrode and counterelectrode in place.

### Internal reflection spectroelectrochemical cells

A typical cell configuration is shown **Figure 6**. This cell has most of the electrode area masked and

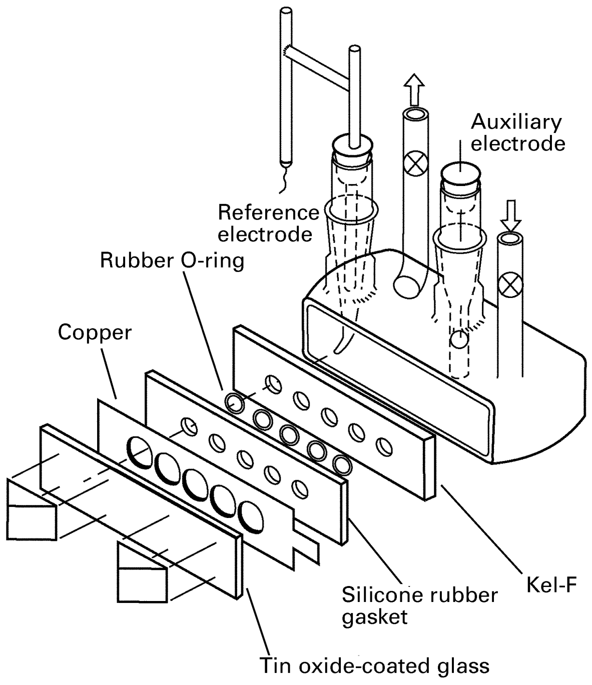



**Figure 6** Internal reflection spectroscopy (IRS) cell and various attachments. Reprinted by courtesy of Marcel Dekker, Inc. from Kuwana T and Winograd N (1974) Spectroelectrochemistry at optically transparent electrodes. I. Electrodes under semi-infinite diffusion conditions. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 7, pp 1–78. New York: Marcel-Dekker.

exposes only the region where the light beam is incident to the electrode–solution interface. The cell shown allows for five reflections, but the number can be selected for the requirements of the experiment.

## Electrochemical control and techniques

Modern electrochemical research encompasses a wide range of techniques, including those based on potential sweep, potential or current step, use of hydrodynamic electrodes, impedance measurements and electrolysis. Electrode processes are investigated at the working electrode, under potentiostatic or galvanostatic control, with a counterelectrode being used to complete the electrical circuit. Historically, in controlled potential experiments, the counterelectrode was also the reference electrode, with the double function of passing current and acting as a reference potential for controlling the potential of the working electrode. Nowadays, three-electrode systems are routinely used where the current passes from the working electrode to a counterelectrode (of larger area than the working electrode), the separate reference electrode serving purely as a reference potential and not passing current. Voltammetric, step and coulometric methods are the most frequently used electrochemical techniques in UV-visible spectroelectrochemistry.

### Voltammetric methods

Electrochemical techniques in which a potential is imposed upon an electrochemical cell and the resulting current is measured are termed voltammetric methods. Numerous methods have been developed, with variation in the type of potential waveform impressed on the cell, the type of electrode used and the state of the solution in the cell (quiescent or flowing). Voltammetry has proved to be very useful for analysing dilute solutions, both quantitatively and qualitatively, for inorganic, organic and biological components, measuring thermodynamic parameters for redox systems, and studying the kinetics of coupled homogeneous chemical reactions. In spectroelectrochemistry, linear-sweep and cyclic voltammetric methods are generally used in quiescent solutions. In linear-sweep voltammetry, for the general redox reaction in Equation [1] above, the potential of the working electrode is swept from a value $E_1$, at which O cannot undergo reduction, to a potential $E_2$, at which the electron transfer is driven rapidly. In cyclic voltammetry, a triangular waveform is applied; once the potential reaches $E_2$, the direction of the sweep is reversed and the electrode potential is scanned back to $E_1$.

**Potentiostatic control** The heart of modern electrochemical instrumentation is the potentiostat, which has control of the voltage across the working electrode–counterelectrode pair; it adjusts this voltage in order to maintain the potential difference between the working and reference electrodes (which it senses through a high-impedance feedback loop) in accord with the programme supplied by a function generator. The instrumentation requirements for thin-layer spectroelectrochemistry are not as exacting or expensive as some other types of spectroelectrochemistry. Since the large ohmic potential drop precludes very fast measurements, relatively inexpensive (slow) potentiostats with a good digital voltmeter are usually adequate.

### Step techniques

In step techniques, a potential or current step is instantaneously applied to the working electrode. Following this perturbation to the electrochemical system, current, charge or potential is monitored versus time.

**Chronoamperometry** Chronoamperometry involves the study of the variation of the current response with time under potentiostatic control. Generally the working electrode is stepped from a potential at which there is no electrode reaction to one corresponding to the mass-transport-limited current, and the resulting current–time transient is recorded. In double-step chronoamperometry, a second step inverts the electrode reaction and this method is useful in analysing cases where the product of the initial electrode reaction is consumed in solution by a coupled homogeneous chemical reaction.

**Chronocoulometry** Chronocoulometry is similar to chronoamperometry except that the current is integrated and the variation of charge with time is studied. The advantages of integration are that the signal increases with time, facilitating measurements towards the end of the transient, when the current is almost zero; integration is effective in reducing signal noise; it is relatively easy to separate the capacitative charge from the faradaic charge.

In an OTTLE cell, coulometry is generally performed by application of a potential that causes complete electrolysis of the electroactive species. Electronic integration of the resulting current gives the total charge consumed by the electrode process, which can be related to the number of moles and electrons involved in the redox reaction by Faraday's law. It is important to carry out a second experiment

in the absence of the electroactive species, to allow subtraction of the 'blank' charge for charging of the electrode/electrolyte interface and any background redox reactions involving the solvent and electrode.

**Chronopotentiometry** Generally constant-current chronopotentiometry is employed, in which the constant current applied to the working electrode causes the electroactive species to be reduced at a constant rate. The potential of the electrode moves to values characteristic of the redox couple and varies with time as the [O]/[R] concentration ratio changes at the electrode surface. In the case of a reduction, after the concentration of O drops to zero at the electrode surface, the flux of O becomes insufficient to accept all the electrons being forced across the electrode–electrolyte interface. The potential, at this transition time, then rapidly changes towards more negative values until a new, second reduction can start.

## Spectroscopic measurements

UV-visible spectral measurements under electrochemical control can often be made using a conventional spectrometer. For thin-layer spectroelectrochemistry experiments, a sufficiently large sample spectrometer compartment is required to accommodate the OTTLE cell.

### Rapid scan and diode array spectrometers

For kinetic measurements, analysis of complete time-resolved spectra is possible using a rapid-scan spectrometer (RSS) interfaced with a microcomputer. With a RSS it is possible to record a 1000-point 450 nm wide spectrum in the range 240–800 nm in about 5 ms, although signal averaging is generally necessary to obtain the required sensitivity. Although RSS instruments have been employed extensively and with great success, the instrumental design is now obsolete and usually instruments are used that employ diode arrays in combination with a polychromator. In these 'optical multichannel analysers', the whole spectrum is spatially dispersed by a polychromator and then imaged onto the detector, which consists of an array of tiny photodiodes.

### Reflectance spectroscopy measurements

A schematic diagram of the typical apparatus required for reflectance spectroscopy is given in **Figure 7**. The optical components consist of a highly stabilized intense light source, frequently a mercury or mercury/xenon arc, a monochromator, a polarizer, the electrochemical system, a photodetector (photomultiplier or photodiode) and appropriate collimating and focusing lenses. The electrode potential is periodically modulated by either a square or a sinusoidal waveform and the small changes in reflectivity so caused are detected with a lock-in amplifier. For kinetic measurements where the shape of the reflectance–time transient is required, the lock-in amplifier is replaced with a signal averager. The requirements for the cell and electrode design are identical to those of ellipsometry.

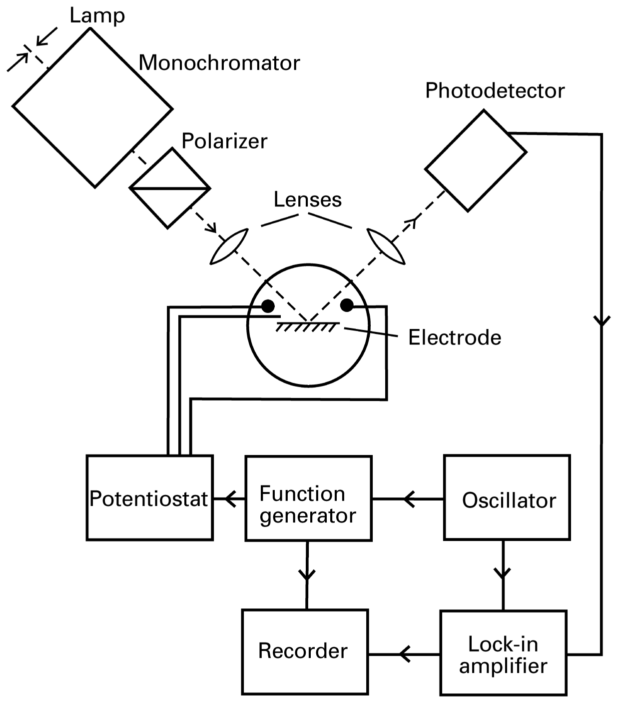



**Figure 7** Block diagram of typical apparatus for reflectance spectroscopy Reprinted with permission from Robinson J (1984) Spectroelectrochemistry. *Electrochemistry Specialist Periodical Reports.* Vol 9, pp 101–161. London: Royal Society of Chemistry.

## List of symbols

$A$ = absorbance; $A_{elec}$ = electrode area; $b$ = path length; $D$ = diffusion coefficient; e = electron; $E$ = electrode potential; $E^0$ = reversible electrode potential; $F$ = Faraday constant (96 485 C mol$^{-1}$); $n$ = number of electrons; O = oxidized form; $Q$ = charge; $R$ = gas constant (8.315 J K$^{-1}$mol$^{-1}$); R = reduced form; $t$ = time; $T$ = temperature in kelvin; $x$ = distance from electrode.

*See also:* **Colorimetry, Theory; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy; Ellipsometry; Light Sources and Optics; Spectroelectrochemistry, Applications.**

## Further reading

Bard AJ and Faulkner LR (1980) *Electrochemical Methods: Fundamental and Applications*, Chapter 14. New York: Wiley.

Fisher AC (1996) *Electrode Dynamics*. Oxford: Oxford University Press.

Heineman WR and Jensen WB (1989) Spectroelectrochemistry using transparent electrodes –an anecdotal history of the early years. *ACS Symposium Series* **390**: 442–457.

Heineman WR and Kissinger PT (1984). In: Kissinger PT and Heineman WR (eds) *Laboratory Techniques in Electroanalytical Chemistry*, Chapter 3. New York: Marcel Dekker.

Heineman WR, Hawkridge FM and Blount HN (1984) Spectroelectrochemistry at optically transparent electrodes. II. Electrodes under thin-layer and semi-infinite diffusion conditions and indirect coulometric titrations. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 13, pp 1–113. New York: Marcel Dekker.

Kolb DM (1988) UV-visible reflectance spectroscopy. In: Gale RJ (ed) *Spectroelectrochemistry*, Chapter 4. New York: Plenum Press.

Kuwana T and Heineman WR (1976) Study of electrogenerated reactants using optically transparent electrodes. *Accounts of Chemical Research* **7**: 241–248.

Kuwana T and Winograd N (1974) Spectroelectrochemistry at optically transparent electrodes. I. Electrodes under semi-infinite diffusion conditions. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 7, pp 1–78. New York: Marcel Dekker.

Mortimer RJ (1994) Dynamic processes in polymer modified electrodes. In: Compton RG and Hancock G (eds) *Research in Chemical Kinetics*, Vol 2, pp 261–311. Amsterdam: Elsevier.

Murray RW (1984) Chemically modified electrodes. In: Bard AJ (ed) *Electroanalytical Chemistry. A Series of Advances*, Vol 13, pp 191–368. New York: Marcel-Dekker.

Niu J and Dong S (1996) Transmission spectroelectrochemistry. *Reviews in Analytical Chemistry* **15**: 1–171.

Robinson J (1984) Spectroelectrochemistry. *Electrochemistry Specialist Periodical Reports*, Vol 9, pp 101–161. London: Royal Society of Chemistry.

Sawyer DT, Sobkowiak A and Roberts JL Jr (1995) *Experimental Electrochemistry for Chemists*, 2nd edn, pp 284–286. New York: Wiley.

# Spectroscopy of Ions

**John P Maier**, University of Basel, Switzerland

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Ions and radicals are transient species which are not readily accessible to conventional techniques for spectroscopic characterization. There are essentially three problems to be overcome–the production in sufficient concentration, the availability of a sensitive technique enabling their IR or electronic spectra to be recorded and the ability to identify the observed spectral features. The involvement of mass-selection not only leads to the solution of the last problem, but enables methods based on particle detection – fragment ions, electrons and photons – to be incorporated. The aim of the spectroscopic studies is, on the one hand, to provide a fingerprint of the species by its vibrational or electronic spectrum, enabling its identification in various terrestrial and space environments, and on the other hand, the spectroscopic analysis leads to information on geometric structures, force fields and fundamental interactions.

Ions are known to be important entities in space, for example in comets and interstellar clouds and their electromagnetic spectrum is the means not only to identify them but also to gain physicochemical data on their surroundings. The same applies in the nonintrusive monitoring in the laboratory, such as of plasmas and flames. It is the purpose of this article to give some examples of the studies on mass-selected species, both of their electronic and IR spectra. This has been achieved in the 1990s in experiments measuring the electronic absorption spectra of mass-selected cations, anions and neutral radicals in an inert neon matrix at ≈ 5 K and more recently in the gas phase for anions following electron photodetachment. IR spectroscopy of ionic complexes, prototypes of fundamental interactions in chemical and biochemical phenomena, and intermediates of ion–molecule reactions, can also be achieved with a mass-selected technique involving vibrational predissociation spectroscopy in an ion trap.

## Absorption spectroscopy in neon matrices

### Technique

This method combines mass spectrometry and matrix isolation. The study of transients trapped in the neutral environment of a rare gas is an established spectroscopic approach. It suffers, however, from selectivity; though sufficient concentrations of elusive radicals and ions can be obtained in rare gas matrices, spectral overlap as a result of the simultaneous presence of many such species usually restricts the interpretation. This is overcome by growing matrices with mass-selected ion beams.

**Figure 1** shows the essential features of the instrument developed. A number of ion sources have been used, hot-cathode discharge for anions and cations as well as a cesium sputter one for anions. These have been designed to generate copious amounts of specific ions. The extracted ions are passed through an electrostatic deflector into a quadrupole mass spectrometer before codeposition with excess neon to form a matrix at ≈ 5 K. Several stages of differential pumping are incorporated in the instrument to isolate the ion source from the cryosurface. Mass-selected ion currents in the nA range are usually required for a successful measurement and the kinetic energies of the ions are chosen to be in the 50–150 eV range. The deposition takes about 2 h resulting in a thin neon matrix (≈ 150 μm) over about 1 $cm^2$ area. A typical ion density is $10^{15}$ to $10^{16}$ $cm^{-3}$. The neutrality in the matrix is assured by the presence of counter-ions generated *in situ* either from impurities present or via the species deposited themselves.

The absorption spectrum of the mass-selected molecules is measured by a technique which enables the

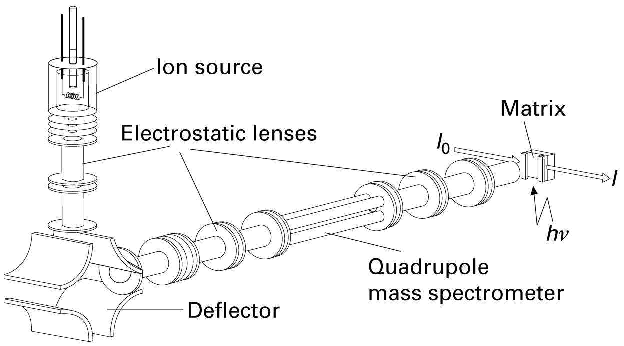



**Figure 1** Apparatus used for the measurement of absorption spectra of mass-selected cations, anions and neutral radicals in neon matrices.

whole length of the thin matrix to be interrogated. The species and neon are condensed on a copper substrate and the light is passed through slits into the side of the matrix and traverses it parallel to the metal surface. By this means, absorption path lengths of 1–2 cm are achieved. For measurements in the IR, only a reflection configuration can be used, leading to two orders of magnitude lower sensitivity.

A major aim of this technique is to identify and locate the characteristic electronic transition of carbon chain-like species. This is for two reasons: (1) to be able to consider their relevance in astrophysical phenomena in view of their spectral features and (2) to plan gas-phase experiments with this knowledge.

## Electronic spectra of cations

As an example, the polyacetylene cations $HC_nH^+$ are considered. They are readily produced in a hot-cathode discharge source fed with a mixture of acetylene diluted with helium. Ion–molecule reactions lead to polymerization. When a particular species is selected according to the number of carbon atoms in the chain, with currents in the nA range, and codeposited with excess neon, the electronic absorption spectra can be measured. **Figure 2** shows the characteristic strong transitions for the species with an even number of carbon atoms. The polyacetylene cations have an open-shell electron configuration with X $^2\Pi$ ground states and the band systems apparent correspond to $\pi^*$–$\pi$ electron excitation, i.e. to electronic transitions of $^2\Pi$–X $^2\Pi$ symmetry.

In the spectra observed, the first band at lower energy is the most intense and the origin transition. The peaks lying to shorter wavelength involve the excitation of vibrational modes in the upper electronic state. Under the conditions of the experiment, rotational motion is eliminated because the species are held rigid in the surrounding neon lattice. The low ambient temperature of 5 K constrains the population to the vibrations $v = 0$ level in the ground electronic state from which all transitions then originate. In addition, the geometry change on $\pi$-electron excitation is relatively small; hence the origin bands dominate and the relative simplicity of the spectra (**Figure 2**). The vibrational pattern and the types of normal modes excited indicates that these are transitions of linear (or quasilinear) polyacetylene chains.

The electronic transition shifts towards the IR by regular increments (≈ 100 nm) for each acetylene unit. This is a characteristic feature of carbon chain molecules and can be simply modelled by an electron

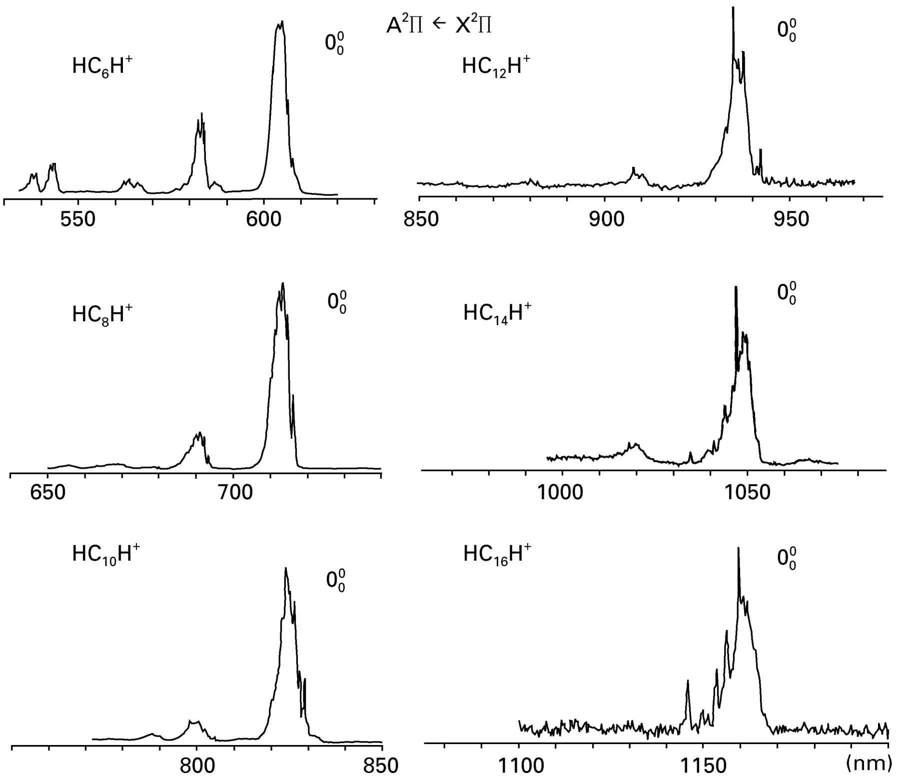



**Figure 2** Electronic absorption spectra of mass-selected polyacetylene cations in 5 K neon matrices.

in a box treatment. In addition to the inverse energy dependence on the length of the chain, the oscillator strength of these transitions grows. This is a contributing factor for the successful detection of the longer species in the laboratory even though the attainable ion current of the mass-selected species is decreasing. It is also one of the attractive features for consideration of such species in astrophysical phenomena. Finally, with the location of the transition in a neon matrix, the corresponding measurements in the gas phase have become a realistic proposition. Typical shifts for the electronic transitions of the cations (**Figure 2**) should be in the 100–200 cm$^{-1}$ range to the blue on passing from the neon matrix to the gas phase.

## Electronic spectra of anions

In an analogous way to the measurement of the electronic absorption spectra of cations, those of anions can be obtained. A sputter source has been used to generate the pure carbon anions. In this, a graphite rod is bombarded with cesium ions. The carbon species are formed by sputtering and gas phase processes and the ions produced are extracted for mass-selection. Sufficient ion concentrations have been attained for the spectroscopic studies of carbon anions in the $C_4^-$ to $C_{20}^-$ range. The carbon species are unusual among anions in that they have large electron detachment energies (3–5 eV) and possess one or more bound excited electronic states. This is illustrated in **Figure 3** by the observed electronic spectra of the $C_{2n+1}^-$ anions detected after mass-selected deposition.

Up to four band systems are discernible (for $n = 4$, 5) and these are the various $^2\Pi - X\,^2\Pi$ transitions arising by $\pi - \pi$ electron excitation. The arrow placed on the wavelength scale (**Figure 3**) indicates the electron detachment threshold in the gas phase. Thus, the highest lying excited electronic state is stabilized with respect to the isolated state as a result of solvation by the neon atoms. The vibrational structure is again relatively simple, indicating a chain structure for the carbon skeleton. In addition to the frequencies of the fundamentals, which can be inferred from the electronic spectra—these are mainly the totally symmetric stretching modes and some bending ones excited in double quanta—antisymmetric, IR active modes have also been observed with this approach. This followed the recording of the infrared spectrum after the mass-selected deposition by a reflection arrangement.

## Electronic spectra of mass-selected neutral radicals

The technique has been extended to the spectroscopic study of mass-selected neutral species. In the case of

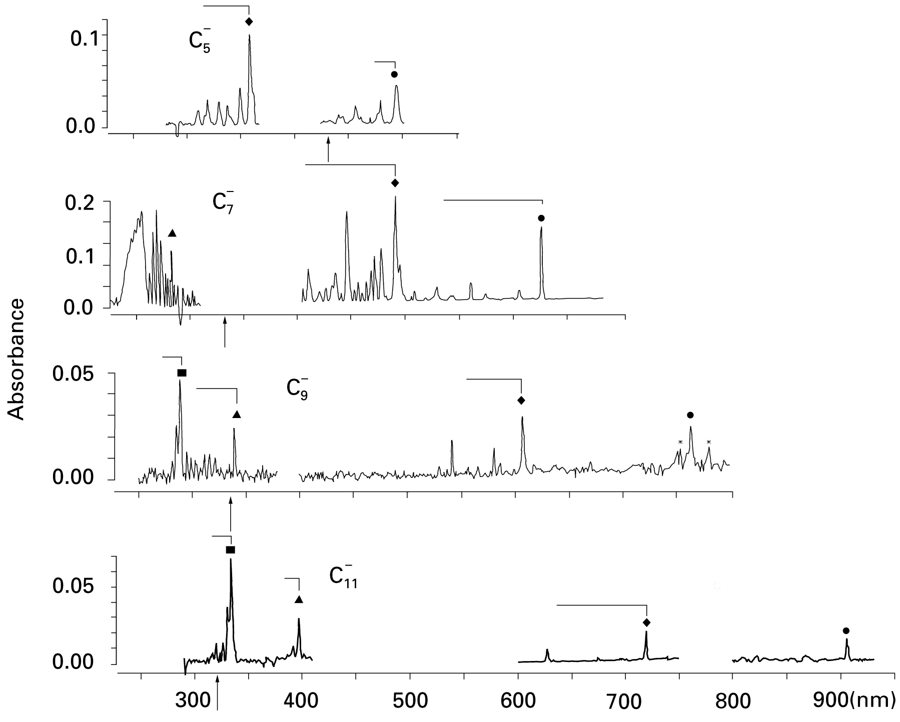



**Figure 3** Absorption spectra of mass-selected carbon anions in neon matrices at 5 K showing several $^2\Pi - X\,^2\Pi$ electronic transitions.

the carbon molecules this is best accomplished by selection of the corresponding anions and subsequent detachment of the electron either during or after growth of the matrix using a broad band photon source. By this means the long sought electronic spectra of the linear carbon chains $C_{2n+1}$ ($n$ = 2–7) could be identified. Some of these are shown in **Figure 4**. The electronic transition corresponds to $\pi* - \pi$ excitation. The carbon chains with odd numbers of atoms are closed–shell species and the symmetry of the observed band systems is $^1\Sigma_u^+$–X $^1\Sigma_g^+$. In contrast, the chains with even numbers of carbon atoms are paramagnetic with a X $^3\Sigma_g^-$ ground state. Their transitions have been detected for $C_{2n}$ ($n$ = 2–5) by this method. The band systems of the linear forms of the larger species are not apparent in the spectra, providing direct experimental evidence for a change of geometrical structure above $C_{10}$. In contrast, the linear chains $C_{2n+1}$ persist beyond $n = 8$. Thus, the potential of the technique of combining mass and matrix-isolation spectroscopies for the characterization of neutral species in addition to the ionic ones is clear.

## Electronic spectra of anions in the gas-phase

The electronic spectra of carbon chain anions have now also been measured in the gas phase incorporating a mass-spectrometric technique. Because the energy region of the electronic transition is known from the measurements on the mass-selected species in neon matrices (see above), the gas-phase study is made that much easier. The sensitive approach adopted involves selection of the respective anion, laser excitation and photodetachment. A schematic outline of the apparatus is given in **Figure 5**.

The anions are conveniently prepared in a pulsed d.c. discharge source containing a few per cent of acetylene in argon expanded in a supersonic free jet with a backing pressure of about 9 bar. The argon carrier gas ensures clustering and cooling of the formed anions. The beam is passed through a skimmer into a time-of-flight mass spectrometer. The ion of interest, e.g. $C_7^-$, is identified by its flight time after which the excitation and photodetachment takes place. In the case of $C_7^-$, the A $^2\Pi$ – X $^2\Pi$ transition (cf. **Figure 3**) is sought. A tunable laser is scanned in this wavelength region while a second laser photon of fixed energy (532 nm) causes the photodetachment whenever the first photon is in resonance with the electronic transition. The total energy of the two photons absorbed exceeds the electron affinity of $C_7$.

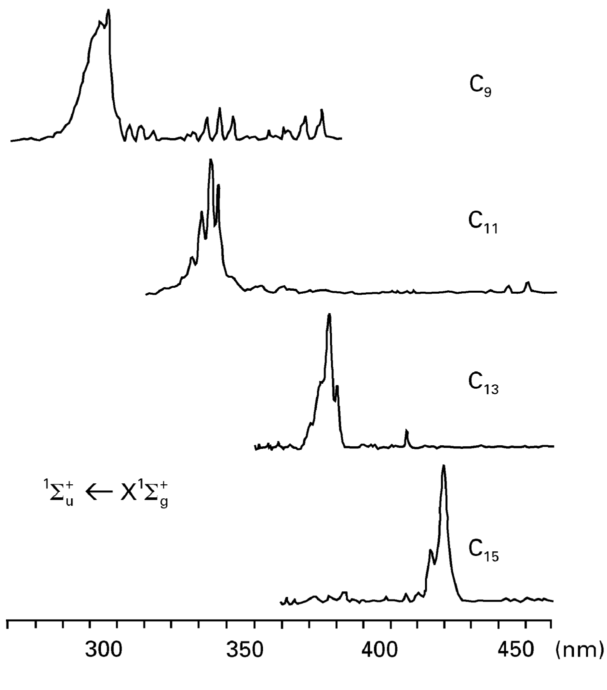



**Figure 4** Electronic absorption spectra ($^1\Sigma_u^+ - X^1\Sigma_g^+$) of neutral carbon chains observed after mass-selected codeposition of their anions with neon at 5 K and detachment of the electrons by photon illumination.

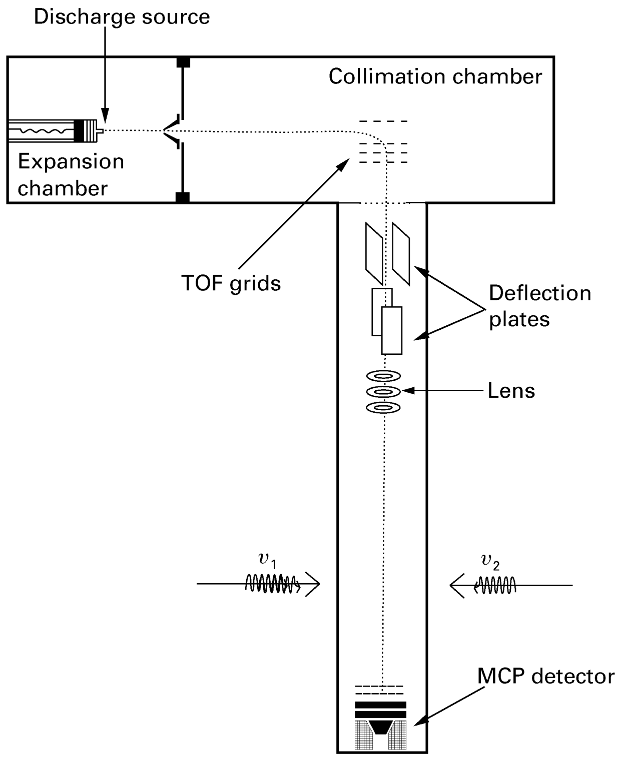



**Figure 5** Experimental arrangement used to observe the electronic transitions of mass-selected carbon anions in the gas-phase.

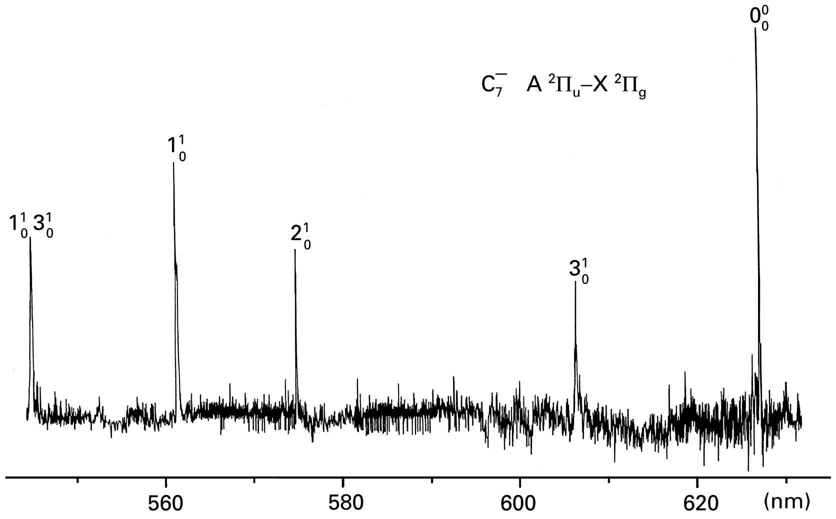



**Figure 6** The A $^2\Pi_u$ – X $^2\Pi_g$ electronic transition of $C_7^-$ measured in the gas phase using a two-colour photon excitation and electron detachment approach.

The detection is usually of the neutral mass-selected species ($C_7$).

In **Figure 6** is seen the recorded spectrum of the A $^2\Pi$ – X $^2\Pi$ transition of $C_7^-$. The striking feature is the narrow line width of the bands, attributed primarily due to the production of cold anions (20–40 K) in the supersonic discharge source. The analysis of the vibrational structure is straightforward – it is consistent with a transition of a molecule linear in both the ground and excited electronic state. The excitation of the three totally symmetric stretching modes in progressions and as combinations in the upper electronic state is apparent. The absence of bands corresponding to bending modes is associated with a relatively rigid structure of $C_7^-$. Similar measurements in the gas phase have now been realized for carbon anions in the $C_5^-$ to $C_{11}^-$ range. Such data then allow a direct comparison with astronomical observations, for example, on the diffuse interstellar bands, a long-standing puzzle.

## Infrared spectra of ionic complexes

### Approach

The aim is to measure IR spectra in the gas phase of ionic complexes. As illustration, the $H_2$–$HCO^+$ species is considered. Such ionic complexes are intermediates in ion–molecule reactions and have been detected by mass spectrometry in the earth's stratosphere. The forces involved in their binding are ion(induced) dipole, and are involved in biological phenomena. The binding energies are generally intermediate between those of van der Waals species and

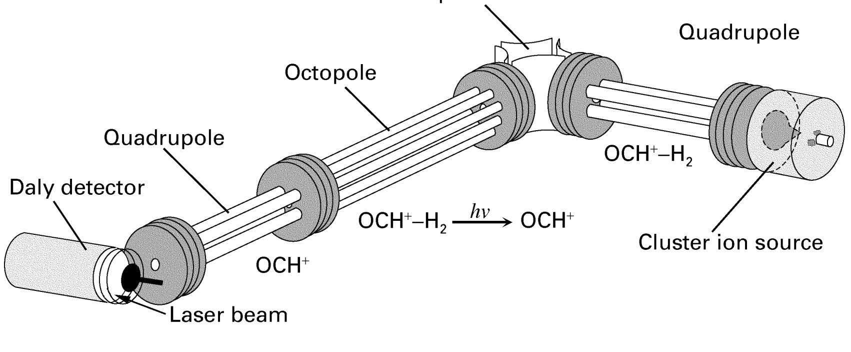



**Figure 7** The setup of the instrument used to measure infrared spectra of ionic complexes via vibrational predissociation spectroscopy.

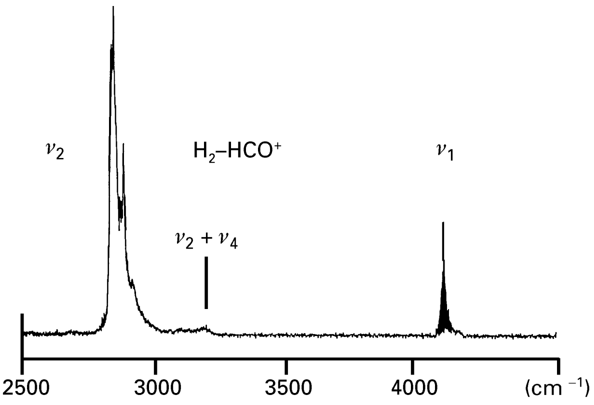



**Figure 8** Infrared spectrum of the mass-selected $H_2$–$HCO^+$ ionic complex recorded with the apparatus shown in **Figure 7**.

those involving hydrogen bonds. In the case of $H_2$–$HCO^+$, the binding energy is ≈ 1400 cm$^{-1}$, whereas it is merely 150 cm$^{-1}$ for He-$HCO^+$, another of the species studied.

There are three concepts in the experiment, the schematic outline of which is given in **Figure 7**. The first is the production and selection of the complexes. This can be achieved only at low temperatures and consequently a supersonic expansion combined with electron impact ionization has been used. To produce $H_2$–$HCO^+$, a 15:1 mixture of $H_2$/CO at a total pressure of 4–5 bar is passed through a pulsed orifice and the complex is formed in the expansion reaction where 70 eV electron impact produces the ions. The latter are extracted via a skimmer into a quadrupole mass spectrometer and the chosen ion is injected into an octopole, and confined for the desired time. The second part involves the excitation of a vibrational transition of the ionic complex. Radiation from a tunable infrared laser passes down the middle of the octopole and intercepts the species. The photon energy is chosen so that it exceeds the binding energy of the ionic complex. When the infrared transition is excited, in due course the complex dissociates. Because the vibrational predissociation is slow for such ionic complexes – the transfer of energy from the mode excited is inefficient, e.g. $\nu_1$(H–H) in $H_2$–$HCO^+$, into the fragmentation channel ($HCO^+$ + $H_2$) – there is usually no spectral broadening.

The fact that fragment ions are produced leads to the third underlying principle of the experiment, namely the sensitive detection of the absorption process by counting the resulting fragment ions. Thus, in the octopole both the ionic complex ($H_2$–$HCO^+$) as well as the product ion ($HCO^+$) are constrained, but the final quadrupole (**Figure 7**) is tuned to transmit only the fragment ion. The IR spectrum of an ionic complex is observed by monitoring the intensity of the fragment ions as a function of the laser frequency.

### Vibrational spectrum of $H_2$–$HCO^+$

**Figure 8** shows an IR spectrum for the complex in the 2500–4500 cm$^{-1}$ region. The two strong bands are the $\nu_1$ (H–H stretch) and $\nu_2$ (C–H stretch) fundamentals as well as combination bands involving the bending levels (in the spectrum only the $\nu_2 + \nu_4$ band is indicated—others lie adjacent to the $\nu_2$ peak). Both $\nu_1$ and $\nu_2$ frequencies are lower than the values for the isolated units indicating a weakening of the H–H

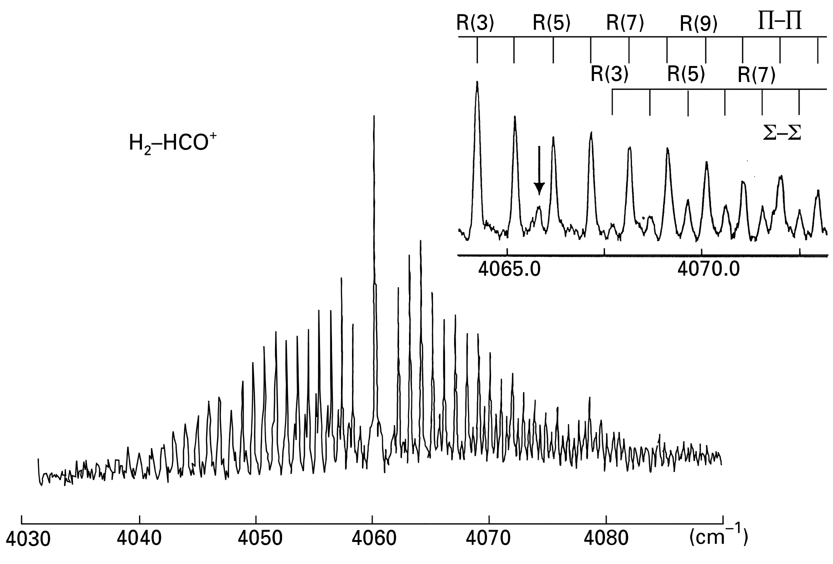



**Figure 9** Rotationally resolved $\nu_1$ band (H–H stretch) of the $H_2$–$HCO^+$ ionic complex.

and H–C bonds in the ionic complex. Thus it can be seen that this approach enables vibrational spectra of mass-selected ionic complexes to be recorded, like the spectra of stable molecules, although the concentrations of the ionic species are minute in comparison.

When the resolution is increased (0.02 $cm^{-1}$), rotational structure on some of the bands becomes apparent. This is particularly striking for the $\nu_1$ band (4060 $cm^{-1}$) of the H–H stretching motion (**Figure 9**). The rotational structure, with Σ–Σ and Π–Π components, is consistent with a T-shaped semirigid symmetric top. Assuming that the $HCO^+$ unit has similar bond distances as in the free species, then the rotational constant evaluated from the analysis of the rotational structure implies a $H_2$ to $HCO^+$ distance of ≈ 175 pm.

Similar studies have now been accomplished on a number of ionic complexes ranging from R–$HCO^+$, R–$N_2H^+$ to species such as R–$NH_4^+$, with R = He, Ne, Ar and $H_2$. This leads to an insight and understanding of the interactions occurring between neutral species and ions as well as of their geometrical structure.

## Solvation of ionic cores

As the measurement of the IR spectra is based on mass-selected species, the number of ligands surrounding the ionic core can readily be varied. This is illustrated in **Figure 10** for the case of the $HCO^+$ surrounded by an increasing number of argon atoms where the changes of the $\nu_1$ transition energy are followed. The IR spectrum can be recorded for each mass-selected entity by monitoring the dominant photofragmentation product ion. The characteristic shifts and patterns in the IR spectrum lead to models of solvation and representation of structural changes.

The addition of the first argon atom results in a shift of 274 $cm^{-1}$ in the $\nu_1$ frequency (H–C stretch) relative to that in the free $HCO^+$ ion. This large change indicates that the argon atom is bound at the hydrogen end. The analysis of the rotational structure points to a linear, rigid geometry. The argon atom is located about 213 pm from the ion core. As further argon atoms are added, the $\nu_1$ frequency shifts in the opposite direction, towards higher energy, in a fairly systematic fashion; by the time 12 argon atoms are solvating the ionic core, the red shift relative to the free $HCO^+$ fundamental is back to 156 $cm^{-1}$. As can be seen from **Figure 10**, the incremental shifts are largest for 2–5 argon atoms, then there occurs a small red shift, while in the range 6–12 atoms the blue shift is rather small. When 13 argon atoms are attached, the $\nu_1$ band splits into two components.

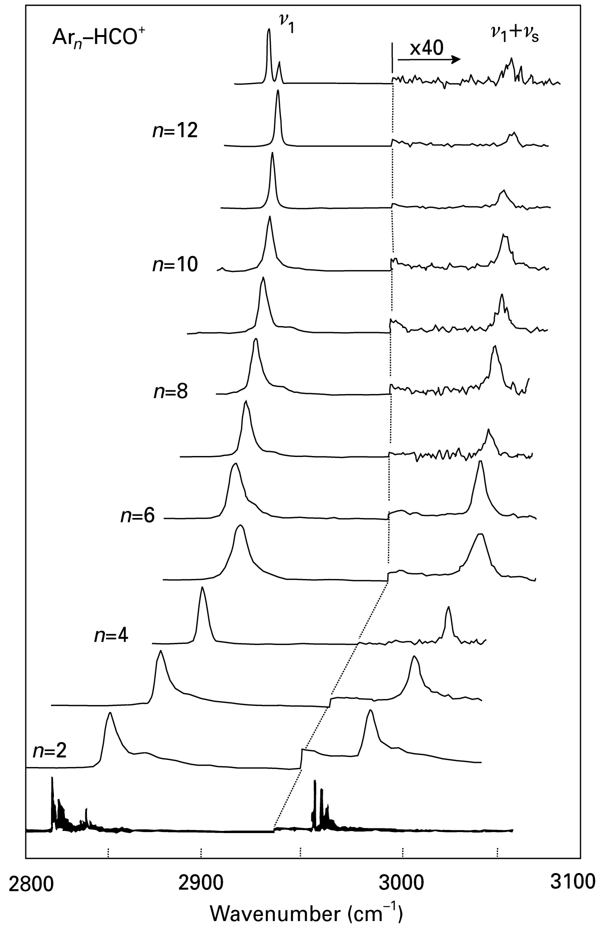



**Figure 10** Infrared predissociation spectra of the $HCO^+$ ionic core solvated by specific number of argon atoms.

These spectroscopic signatures can be rationalized in terms of a simple structural model. The first argon atom occupies a linear proton-bound position. The next 2–5 atoms form the first primary solvation ring at one end. The addition of 5–11 argon atoms forms the second ring at the other end. The marked splitting with 13 atoms, as well as a distinct drop in the binding energy, is associated with the beginning of a second solvation shell and the presence of at least two isomers.

*See also:* **Cluster Ions Measured Using Mass Spectrometry; Ion Dissociation Kinetics, Mass Spectrometry; Ion Energetics in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Ion Structures in Mass Spectrometry; Multiphoton Excitation in Mass Spectrometry; Photoionization and Photodissociation Methods in Mass Spectrometry.**

## Further reading

Bieske EJ and Maier JP (1993). Spectroscopic studies of ionic complexes and clusters. *Chemical Reviews* **93**: 2603–2621.
Herbst E (1990) The chemistry of interstellar space. *Angewandte Chemie, International Edition* (English) **29**: 595–608.
Kroto HW (1981) The spectra of interstellar molecules. *International Reviews in Physical Chemistry* **1**: 309–376.
Maier JP (ed) (1989) *Ion and Cluster Ion Spectroscopy and Structure*. Amsterdam, Elsevier.
Maier JP (1992) Electronic spectroscopy – ions in space and on earth. *Chemistry in Britain* 437–440.
Maier JP (1997) Electronic spectroscopy of carbon chains. *Chemical Society Reviews* **26**: 21–28.
Weltner W and Van Zee RJ (1989) Carbon molecules, ions and clusters. *Chemical Reviews* **89**: 1713–1747.

# Spin Trapping and Spin Labelling Studied Using EPR Spectroscopy

**Carmen M Arroyo**, US Army Medical Research Institute of Chemical Defense, Aberdeen Proving Ground, MD, USA

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## The EPR/spin label technique

Spin labelling is used to monitor physical, biophysical or biochemical properties of substances, proteins, lipids, and cell membranes. This is achieved by introducing a persistent free radical (such as a nitroxide) into the system, which is sensitive to this physical surrounding. Such spin probes are ideally suited to investigate the dynamic aspects of molecular interactions. A spin probe study should proceed in two steps. The first is an analysis of the EPR spectra, yielding physical parameters about the orientation and the motion of the spin probe in the host system. The second step is the interpretation of these parameters in terms of relevant molecular models for the formulation. The section on spin labels will use as an illustrative example the results obtained when various perfluorinated polyether (PEPE) materials containing reactive antivesicant agents were labelled with stearic acid spin probes. These spin-labelled ointment formulations were exposed to different concentrations of vesicant (blistering) agents. The attachment of the spin label at a particular site on an ointment formulation provides a unique means of probing the physical environment in the formulation. If the probe molecule is placed at or near the reactive component of the formulation, then the magnetic resonance experiments may be used to determine the penetration profile. However, there have been many examples of spin labelling in other fields, including liquid crystals and cell membranes.

## EPR/spin label technique for antivesicant agents

Initial screening was performed using the electron paramagnetic resonance (EPR)/spin-label technique by inserting stable organic nitroxide free radicals, known as spin labels, into a heterogeneous complex system consisting of polytetrafluoroethylene (PTFE) particles and fluorinated oils of the proposed formulation. The EPR spectra of the spin labels are very sensitive to the rate at which the label is able to reorient after a magnetic field ($\boldsymbol{B}_0$) is applied. Thus, knowledge of this functional dependence allows the evaluation of the degree of mobility and permeability permitted in the environment of the label. Ointment formulations of reactive topical skin protectants (rTSPs) or topical skin protectants (TSPs) based on PFPE (i.e. fomblin®) were prepared and spin labelled. Four N-oxyl-4-4'-dimethyloxazolidine derivatives of stearic acid, 5-NS, 7-NS, 12-NS and 16-NS, were used as spin probes. The spin-labelled vehicle, fomblin® and the vehicle containing chloroamide [chlorinated glycoluril; 1,3,4,6-tetrachloro-7,8-diphenyl-2,5-diiminoglycoluril (S-330), an antivesicant] were exposed to various concentrations of 2-chloroethyl ethyl sulfide (half-sulfur mustard). The physical

insight into the molecular motion is defined by the order parameter ($S$) which depends on the frequency and amplitude. The EPR experiments detect only the motion and orientation of the –N•–O group, so the order parameter is related to the $x$, $y$, $z$ coordinates. Therefore, $S$ depends on the depth of penetration of the paramagnetic group into the vehicle (fomblin®) and on the chemical composition of the reactive antivesicant under investigation. The net change of the viscosity of the vehicle and the chemical composition were seen to affect the penetration profile.

The value of S obtained from the EPR data provides an analytical tool for comparison of the formulation of topical skin protectants and reactive topical skin protectants. The changes in the interior of this heterogeneous system of the labelled-TSP or labelled-rTSP (controls) were compared with exposed labelled-TSP/or rTSP. The spin-labelled formulations were exposed to the vesicant agents in a dose/time-dependent manner. The results show that the EPR/spin-labelling technique provides an analytical tool for determining the resistance of rTSP to the breakthrough and fluxes of vesicant agents.

The formulated candidates were labelled with stearic acid spin probes. Four kinds of N-oxyl-4,4′-dimethyloxazolidine derivatives of stearic acid (**Figure 1**) were used as spin probes. The nitroxide group (–N•–O) is attached at various positions along the fatty acid chain to situate the nitroxide groups at different depths in the hydrophobic interior of this heterogeneous complex system of TSP or rTSP. These probes are amphiphilic with polar regions which anchor the probes at the hydrophilic interface and the large, rigid steroid frameworks embedded in the lipid chain region. The largest hyperfine splitting (32 G) occurs when the magnetic field is parallel to the long axis of the fatty acid (and hence perpendicular to the plane of a well-ordered multilayer system). These steroid probes lack the rigidity of the group to which the doxyl moiety is attached. It is the behaviour of the lipid to which the –N•–O group is attached that is important, so the order matrix is usually transformed to the molecular coordinate system, with $z$ as the long axis of the spin label. It is in the molecular coordinate system that the $S$-matrix has axial symmetry. The outermost EPR lines move in and the line widths become narrower as the doxyl group moves from C-5 to the C-7, C-12 and C-16 positions.

The mixture of the formulation labelled with the spin probe was placed in a tissue cell sample holder (illustrated in **Figure 2**) for EPR measurement at room temperature. Once the EPR spectra of the control samples were recorded and characterized, titration experiments were performed on the same

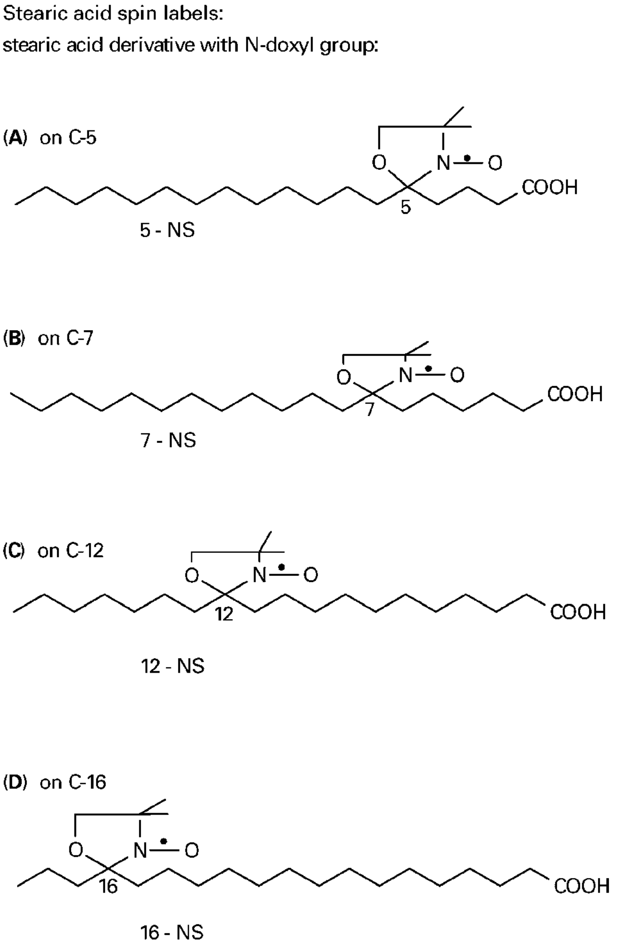



**Figure 1** Chemical structure of the stearic acid spin labels.

control with the particular vesicant agent under study. The vesicant agent, an oily liquid, was carefully distributed over the surface area using a plastic spatula. This dispersion process was performed without disturbing the surface of the labelled formulation. The EPR quartz flat cell was maintained at room temperature in a chemical fume hood for an hour to allow venting of the volatile agent and, after one hour, the EPR spectra were monitored as described for the control samples.

The EPR of the labelled-formulation control exhibits two low-field peaks as illustrated in **Figure 3**, which shows the EPR spectrum of a vehicle (fomblin®-RT-15) incubated with 5-doxyl stearic acid. The spectrum resembles that of an immobilized spin probe. The value of $S$ of the strongly immobilized probe can be calculated from the spectra using the equation:

$$S = \frac{(T_{\parallel} - T_{\perp})}{3[(T_{zz} - a^{1})/2]}$$

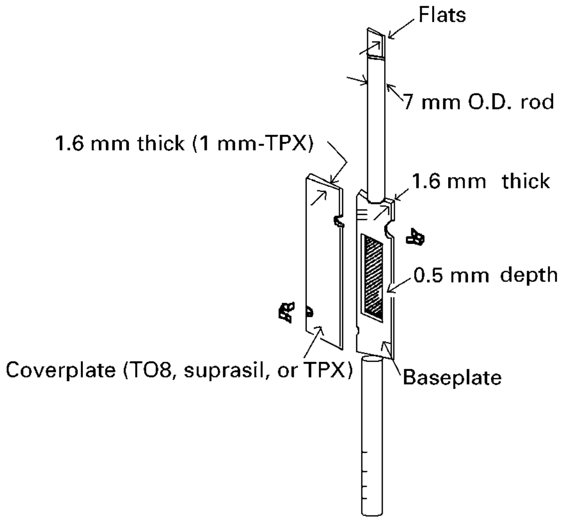



**Figure 2** EPR cavity cell for semisolid samples.

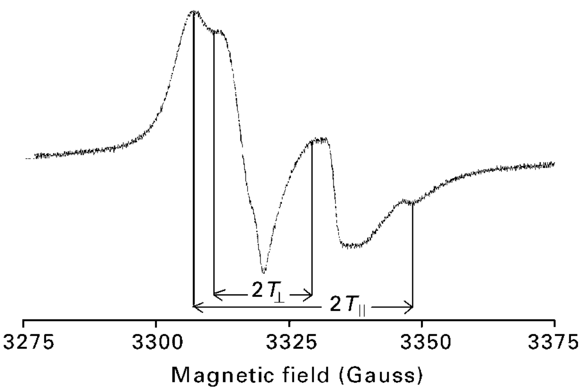



**Figure 3** EPR spectrum of a formulation incubated with 5-doxyl stearic acid. The spectrum is typical of an immobilized spin-probe where $T_{\|}$ is the outer hyperfine splitting ($A_{\|}$) and $T_{\perp}$ is the inner hyperfine splitting ($A_{\perp}$).

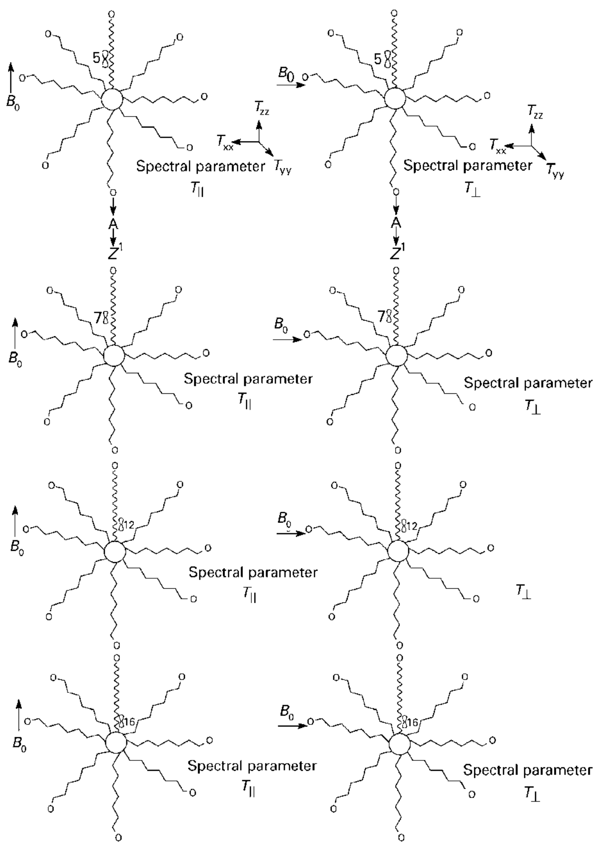
where $T_{\|}$ is the outer hyperfine splitting, $T_{\perp}$ is the inner hyperfine splitting, $a^1$ is $(T_{\|}+2T_{\perp})/3$, and the tensor $T_{zz} = 32$ G. A change in $S$ could be interpreted as representing a net change in the breakthrough and flux of the vesicant agent.

Some general principles should be followed in such formulation studies. (1) The use of experimentally determined 'low' probe–lipid ratios is a requirement. A sufficient quantity of the spin probe should be incorporated into the formulation to permit the recording of an EPR trace with a reasonable signal-to-noise level. (2) The spectral parameters $T_{\perp}$ ($A_{\perp}$), $T_{\|}$ ($A_{\|}$) and $S$ derived from the EPR spectra should be plotted as a function of the vesicant concentration under study. If titration experiments indicate optimal conditions of probe–lipid ratios, then $S$ may be used as a measure of formulation fluidity. (3) It must, however, be remembered that $S$ is a function of both the motion of the probe and the polarity of the environment of the probe. (4) The probe spectral perturbation could be monitored by initially loading the formulation with probe and continuously measuring the EPR spectra to identify 'low' probe concentrations for formulation systems that destroy the EPR signal of the spin label with time. The decrease in probe concentration could be estimated from a double integration of the EPR trace. The decrease in the various spectral parameters could be plotted as a function of time.

The –N$^{\bullet}$–O moiety is attached directly to the hydrocarbon, as illustrated in **Scheme 1**. The probe is incorporated into the lipid portion of the perfluorinated grease and produces an EPR spectrum of an immobilized spin-label probe. An idealized cross section of this heterogeneous system is illustrated in the schematic representation I. The fatty spin-label (A) intercalates with the long hydrophobic chain normal to the plane of the matrix. The rigid spirane structure of the doxyl group places the $z$-axis (arrow) of the nitroxide parallel to the extended lipid chain. Thus, the direction of maximum splitting will be observed when the magnetic field ($\boldsymbol{B}_0$) is normal to the plane of the matrix (i.e. along the arrow A); the splitting becomes progressively smaller as the matrix is rotated in the magnetic field, and the minimum splitting occurs when the matrix is parallel to the magnetic field.

In principle, the polarity profile could be determined by measuring the three-hyperfine splitting constants $T_{xx}$, $T_{yy}$ and $T_{zz}$ for the –N$^{\bullet}$–O group at various positions along the lipid chain. In practice, none of these parameters can be measured directly because of the partial motion average of the electron–nuclear dipolar interaction. An indirect method of obtaining the polarity profile is to estimate $T_{\perp}$ and $T_{\|}$ with the spin labels in randomly oriented samples. The –N$^{\bullet}$–O group at the C-5 position in the lipid matrix has been reported to be in an environment that is more polar than the C-12 and C-16 positions which are in a more hydrocarbon-like environment (**Scheme 1**).

Humidity is a very important factor in such studies since vesicant agents hydrolyse at a relatively rapid rate. The polarity profile can be studied using lipid films supported on glass wool in a hydration chamber. These samples can be equilibrated at 100% relative humidity or dehydrated over phosphorus pentoxide ($P_2O_5$); the removal of water would effectively abolish the polarity profile. The main point is




**Scheme 1** A fatty acid spin probe incorporated into a very complex heterogeneous system that consists of a particular probe (0.2–10 μm) and a lipid matrix. The axis of motional averaging $Z^1$ is the normal on the matrix surface. The arrow (i.e. along the arrow A) indicates the direction of the nitroxide $z'$-axis.

that the defined parameter *S* provides an operationally relative change at some depth in this complex heterogeneous system. Ion transport across lipid films can be studied by measuring the maximum and minimum splitting for the (–N•–O) group at various positions along the lipid chain. In addition, a penetration profile into the lipid matrices can be measured.

The order parameter *S* depends on the depth of penetration of the paramagnetic group (–N•–O) into the homogeneous hydrophobic vehicle, the viscosity of the fluorochemical material and the chemical properties of the vesicant agent under investigation. Therefore, the net change in viscosity of the vehicle and the chemical composition affect the penetration profile. The probability that the paramagnetic group of the probes will be located at different depths in this complex heterogeneous system provides a unique tool to observe the permeation process at a molecular level. In addition, if we assume that the labelled and the unlabelled lipids (base oil of the formulation) intercalate to different depths in the lipid matrix, then to account for the large effect at the C-5 position, the vesicant agent must penetrate to at least the C-2 position to change the environment of the formulation. This penetration profile can be improved by using spin labels bound in the polar head group region, by filling in all points from C-3 to C-16, and by performing X-ray diffraction experiments on the same samples to obtain a more accurate estimate of the overall matrix thickness. Nevertheless, the essential features are clear, stearic acid spin-labels provide a means of estimating system heterogeneity, because they are randomly distributed in the plane of the matrix, position themselves so that the α-carbon of the fatty acyl chain is aligned with the head group region of the lipid matrix, and faithfully reflect the degree of orientational order and dynamics of the lipid molecules. The observed decay of *S* can be related to the permeability properties of the system, because it is known that the EPR signal of the nitroxide group disappears almost instantaneously when the paramagnetic moiety is located at the polar interface of a lipid matrix. This observation is reinforced by the results obtained with the stearic acid spin probes (**Figure 4**). **Figure 4** shows representative EPR spectra of fomblin® HC/04 (relative molecular mass = 1500; viscosity index: 60) labelled with 12- and 7-doxyl stearic acids in the presence of a chlorinated glycoluril (S-330) control and exposed to different concentrations of half-mustard. It is well known that compounds containing a chloroamide group are generally preferred. S-330 provides an efficient antivesicant, which is characterized by its high efficiency and long protective times. One possible mechanism is that S-330 produces sufficient chloride ions to prevent the formation of the sulfonium ion intermediate, which is the active form of sulfur mustard. The graph in **Figure 4** illustrates the decay of the EPR integral signal (signal intensity in relative units) as a function of H-MG concentration. The decay of the EPR signal was dependent on the depth of penetration of the paramagnetic group into this heterogeneous matrix. The increase of H-MG concentration affects the penetration profile as shown in **Figure 4** for the four different spin-probes.

Spin probes only indirectly reflect the ordering of the host system. What information about the molecular architecture of this heterogeneous system can be deduced from these experiments? These labelled-formulations exhibit very little anisotropy (i.e. their magnitude and sign depend on the orientation of the radical with respect to the applied magnetic field) when hydrated with distilled water. Addition of nonelectrolytes such as sucrose has no effect, but the addition of salt results in a large increase in EPR spectral anisotropy. This effect is thus one of charge,

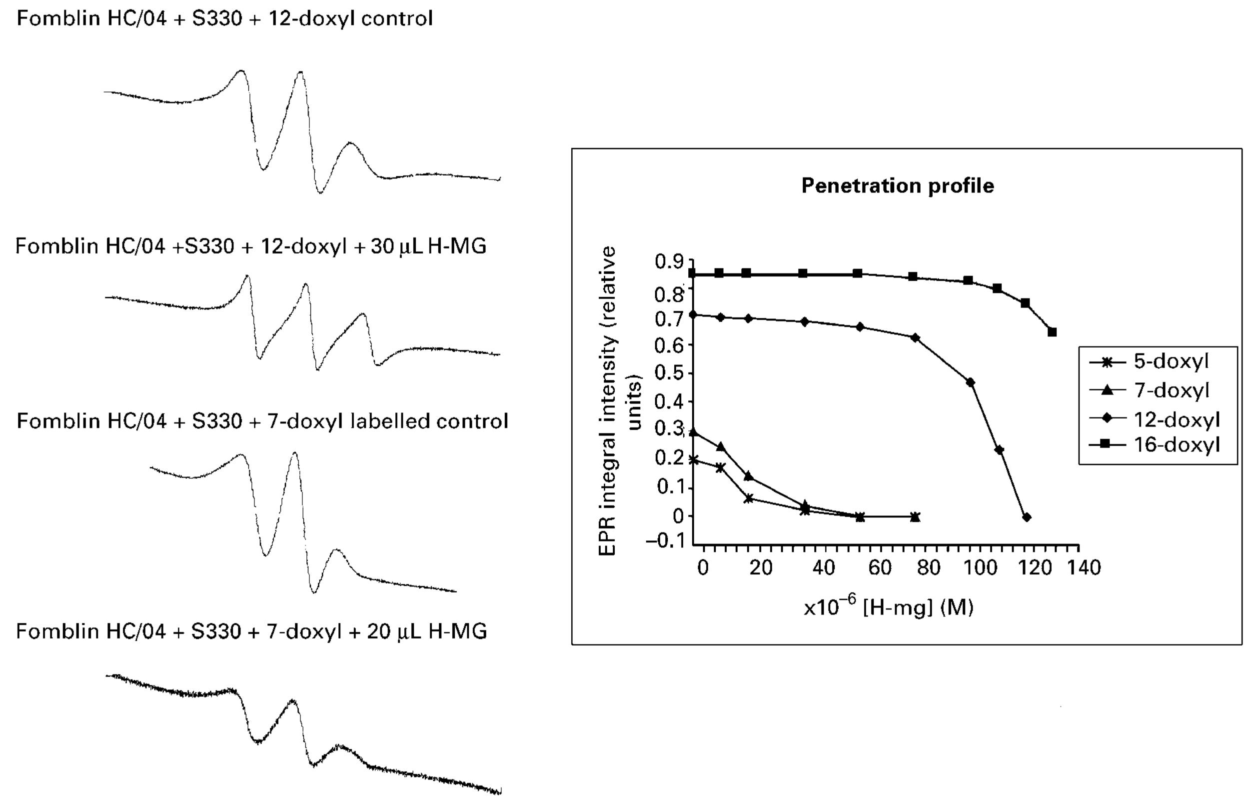



**Figure 4** EPR spectra of a labelled reactive topical skin protectant with 12- and 7-doxyl in selected controls and exposed samples. The graph shows a plot of EPR signal integral (relative units) obtained from the labelled formulation as a function of H-MG concentrations for spin-label probes, 5-, 7-, 12-, and 16-doxyl stearic acids.

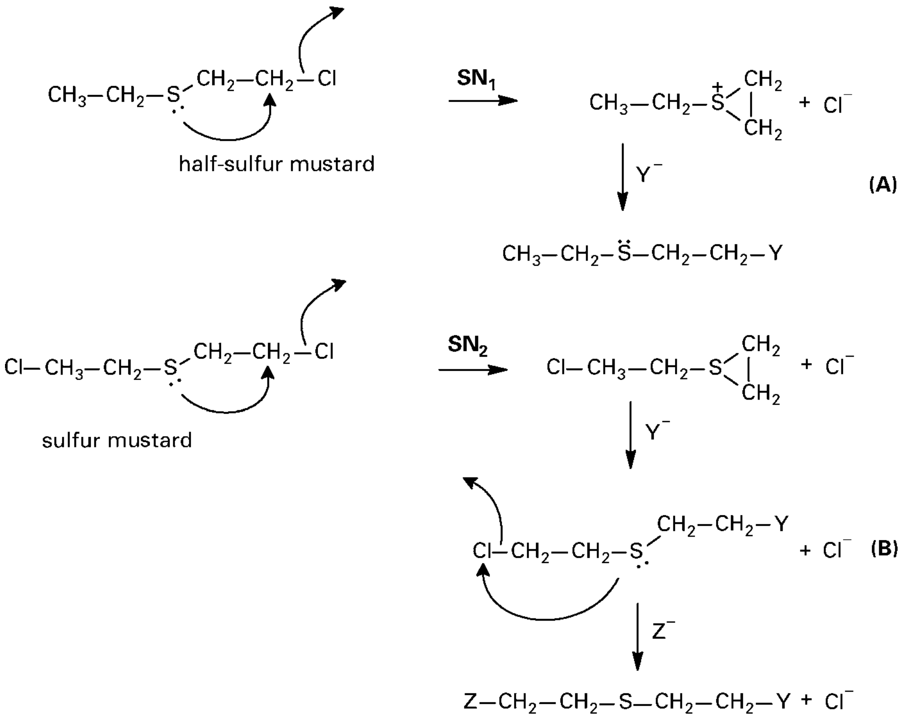
not of osmotic or vapour pressure. It has been shown to be owing to the cation and to depend on the cation valence. The order of effectiveness in promoting spectral anisotropy is $NaCl = KCl = LiCl < MgCl_2$, whereas the chloride, thiocyanate and sulfate salts of the same cation were equal in effectiveness. The cations act by reducing the surface charge density of the heterogeneous complex system, which consists of small particles and the lipid matrix. With the reduction in repulsion between groups, the ointment formulation can contract and achieve a higher degree of order. A net charge of the groups leads to molecular repulsion, an expansion of the formulation, and a decrease in EPR spectral anisotropy.

Vesicant agents of the 'mustard' type present some special features in their reaction mechanism(s). It is known that the mechanism involves, as a first step, the formation of a cyclic sulfonium ion and the release of chloride ($Cl^-$). This mechanism is illustrated for the cases of half-mustard and sulfur mustard in **Scheme 2**. In addition, these vesicant agents are persistent agents, depending on pH and moisture. The mustard is hydrolysed to form HCl and thiodiglycol. Ointment formulation-incorporated spin probes are sensitive to the polarity and fluidity of their local environment. Cation binding, pH alterations, and the action of many substances perturb labelled formulations to yield characteristic changes in their respective EPR spectra. An adequate interpretation of EPR spectral changes in terms of the structure of the host matrix requires that alterations in fluidity and/or polarity be distinguished from changes in probe–probe interaction (e.g. dipole–dipole and electron–electron exchange broadening).

The spectral alterations noted upon addition of vesicant agents to a labelled formulation are caused by changes in (1) motion of the probe, (2) the polarity of the environment of the probe, (3) alteration in fluidity, and/or (4) the permeability profile. If the vesicant agent induces changes that involve radical interactions, the magnitude of the spectral alteration that depends on probe concentration will disappear.

## EPR/spin trapping

EPR spin trapping techniques have successfully been applied to determine and identify free radical intermediates in biology and toxicology. Spin trapping allows one to determine if short-lived free radicals




**Scheme 2**

are involved as reaction intermediates by scavenging the reactive radical to produce more stable nitroxide radicals. This technique involves reaction of the initially-generated radical (itself either too short-lived or of too low a concentration to directly detect) with an added organic compound, known as spin trap, to generate stable radical adducts from whose EPR spectra information about the original radical may be obtained.

Two kinds of spin traps have been developed, nitrone and nitroso compounds. **Figure 5** illustrates the most common commercially available traps. Nitrones are the spin traps of choice for the study of oxygen-centred radicals. The most popular nitrone traps identified in **Figure 5** have a β-hydrogen that can provide considerable information about the radical trapped. However, some information is lost using these nitrones because the trapped radical adds to a carbon adjacent to the nitrogen. Nitroso compounds can provide more information than nitrones as the radical to be trapped adds to the nitroso nitrogen, and, therefore, more information about the hyperfine splitting parameters is obtained.

## *In vivo* spin trapping of oxygen-centred radicals

Nitrones have emerged as the most popular spin traps for biological applications, and out of several nitrone spin traps, the cyclic 5,5-dimethyl-1-pyrroline *N*-oxide (DMPO) has received most attention, since it yields distinct and characteristic adducts with superoxide radical anion ($O_2^{\bullet -}$) and hydroxyl radical (•OH). The use of DMPO as a probe for oxyradical generation in biology systems is not without limitations as high concentrations of DMPO have been suggested to have serious toxic effects on biological tissue. A low concentration of DMPO (10 mM) was used to detect free radical generation in hearts with ischaemia/reperfusion insult. **Figure 6** shows a scheme of a typical spin trapping experiment in an isolated heart. In the effluent immediately after reperfusion, DMPO–OOH, the superoxide spin adduct of DMPO, was obtained. DMPO at a 10 mM concentration range did not interfere with the left ventricular (LV) function during the control perfusion period. Enzyme leakage from hearts also supported non-toxicity findings of DMPO at 10 mM, confirming that the DMPO superoxide adduct is genuine evidence of the generation of superoxide upon reperfusion and is not an artificial generation owing to the cytotoxicity of DMPO.

Furthermore, application of the spin trapping technique in intact animals require an understanding of the stability of the spin traps and the spin adducts *in vivo*. A new class of α-phosphorus-containing DMPO analogues, 5-(diethoxyphosphoryl)-5-methyl-1-pyrroline *N*-oxide (DEPMPO) has been

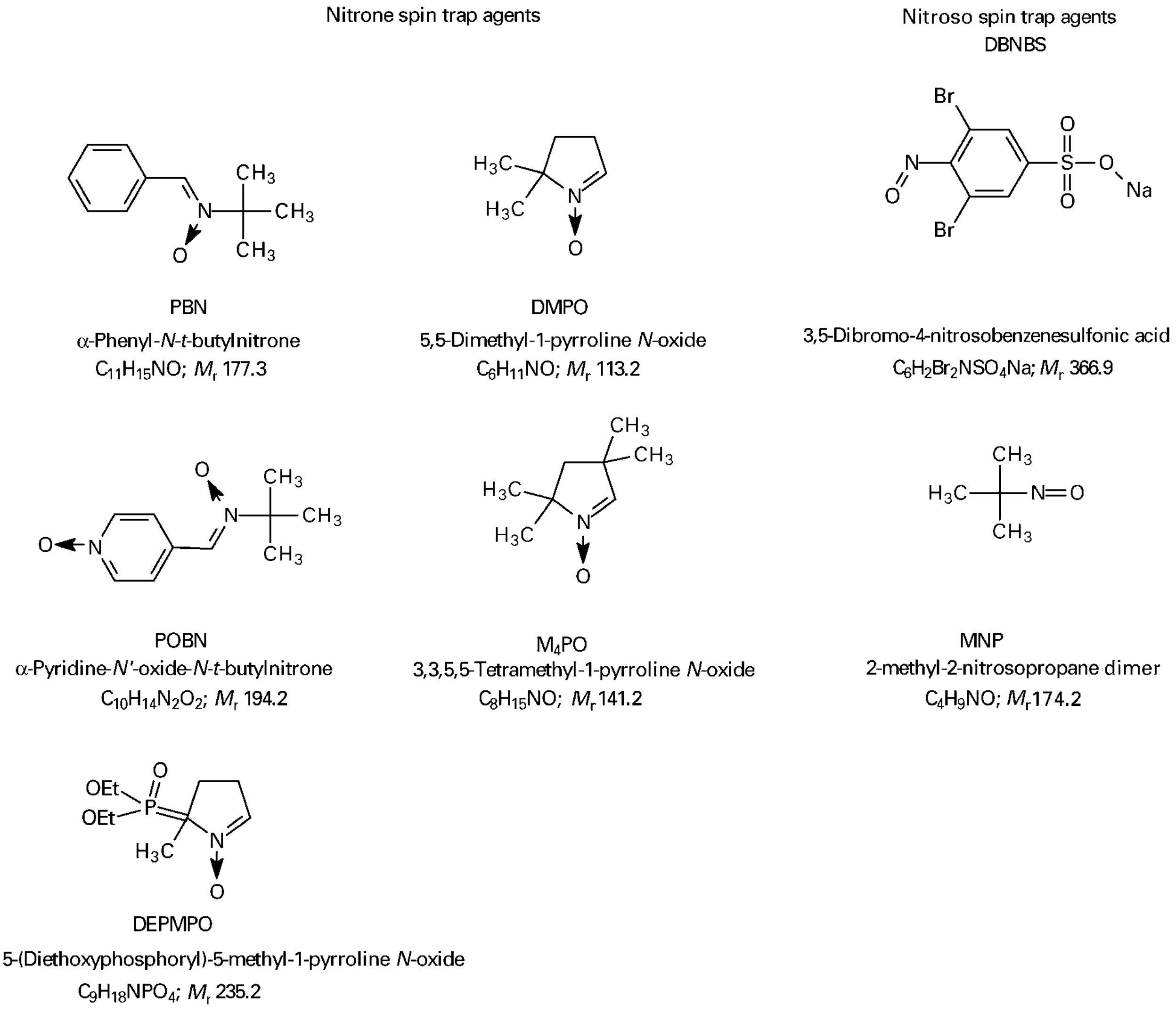



**Figure 5** Names, structural formulae and relative molecular mass of the most common spin trapping agents commercially available.

synthesized and characterized for the generation of superoxide during the reperfusion of ischaemic isolated hearts. DEPMPO can trap and form a stable adduct for both •OH as DEPMPO–OH, and $O_2^{\bullet-}$ as DEPMPO–OOH, giving EPR spectra that are characteristic of each. Thus, unlike DMPO, DEPMPO can be used to distinguish between superoxide dependent and independent mechanisms that lead to the hydroxyl radical. DEPMPO is a good candidate for trapping radicals in functioning biological systems, and represents an improvement over the commonly used trap DMPO.

A very important aspect of spin trapping is to positively identifying the radicals under study. This assignment requires a knowledge of how certain features of the structure of the trapped radical influence the EPR spectrum of the spin adduct. Sometimes it is very difficult to determine unambiguously the precise structure of the spin adduct from the EPR signal obtained. Isotopic substitution EPR experiments are recommended in an attempt to identify the observed adducts. The strategy is that the unpaired electron in a radical interacts with the nucleus of the atom it orbits, and the spin of the nucleus determines the number of lines or peaks in the spectrum. For example, $^{13}C$ has a nuclear spin of $\frac{1}{2}$ while $^{12}C$ has no spin. An unpaired electron, which is associated with atoms having no spin, will exhibit an EPR spectrum containing only a single line. The spin of the nucleus influences the resonance of the unpaired electron so that the EPR resonance splits into two or more lines. The number of EPR resonance observed is equal to $2I + 1$, where $I$ is the nuclear spin. A practical

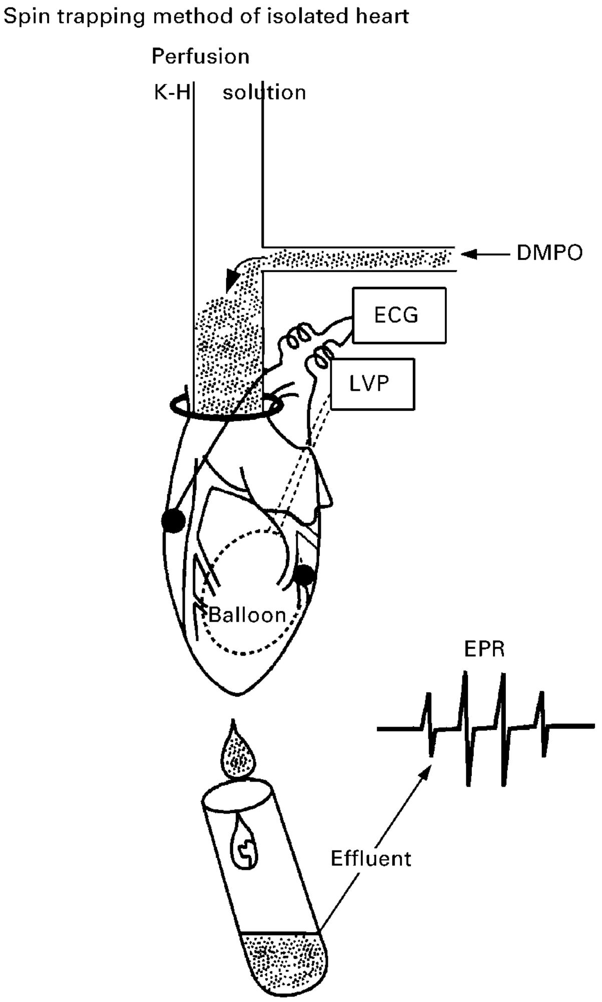



**Figure 6** Scheme of a typical spin trapping experiment in an isolated heart. Coronary effluents are collected and immediately frozen in liquid $N_2$ to prevent spin adduct decay. Frozen samples are thawed just before EPR measurement and EPR spectra are recorded using 160 or 250 μL flat cells. Typical spectrometer conditions are as follows: magnetic field 3350 G, power 8 mW, response 0.3 s, modulation amplitude 1.0–1.25 G, receiver gain $6.5 \times 10^3$, sweep time 2 min, modulation frequency 100 kHz and temperature 23–28°C.

example using this concept follows. When human monocyte cells were exposed to 2-chloroethyl ethyl sulfide (H-MG) in the presence of the nitroso spin trap MNP, an EPR signal consisting mainly of a primary triplet was observed (**Figure 7A**). The complexity of the EPR spectrum suggested the trapping of a hydrogen-abstraction radical of polypeptide molecules. To verify the assignment, $^{13}$C-labelled (at the $C_2$ amino acids) of human tumour necrosis factor-alpha (TNF-α) were reacted with H-MG in the presence of MNP. The resulting EPR spectrum (**Figure 7B**) was detected and simulated using five different hyperfine coupling constants, indicating the presence of a MNP-$^{13}$C-centred adduct. Therefore, the, spin trapping technique, properly applied, leaves no doubt about the nature of the radical or the intensity and duration of its production in biological systems.

## EPR/Spin trapping in toxicology

The involvement of reactive intermediate species (RIS) in chemical agent toxicity could be also investigated using EPR/spin trapping techniques. For example, toxic oedemagenic gases, such as phosgene ($OCCl_2$), bis(trifluoromethyl)disulfide (TFD) and perfluoroisobutylene (PFIB) can be studied by spin trapping. Two types of apparatus that have proved useful for these experiments are shown in **Figure 8**. They consist of a 'tube' which connects via a 7/25 tapered ground-glass joint to a Wilmad EPR 'flat cell' (**Figure 8A**) and a 'U' tube which also with a tapered ground-glass joint to a Wilmad 'flat cell' (**Figure 8B**). In a typical experiment, one positions the tube or the 'U' tube vertically and a solution of the spin trap is placed in one chamber. The chambers are stoppered with rubber septa through which long (~#8) syringe needles are inserted. The gas radical producers are then passed through the spin trap solution for 15–20 min. For deoxygenating purposes, bubbling purified nitrogen or argon gas through the solution is sufficient. The excess of gas can escape through a small syringe needle, as indicated in **Figure 8**. When gassing is complete the system is stoppered and the contents of the tubes and sample cell are thoroughly mixed and shaken into the EPR flat cell, which is inserted into the microwave cavity of the EPR spectrometer. Relatively simple modifications of this basic experimental design allow the study of other toxic chemicals. Highly reactive intermediate species of phosgene have been identified by adding $^{13}$C-labelled phosgene into a PBN solution in benzene (**Figure 9**).

EPR/spin trapping can be used in the study of reactive intermediate species in a wide range of biological and toxicological systems. This technique allows the identification of radical species [and hence their mechanism(s) of production], and the indirect and real-time observation and quantification of radical reactions. To summarize, spin trapping is a powerful technique for the indirect EPR observation of many reactive free radicals. More ideas for dealing with this technique are well documented and discussed in reviews by Janzen, Anderson Evans, Mason, Thornalley and Buettner.

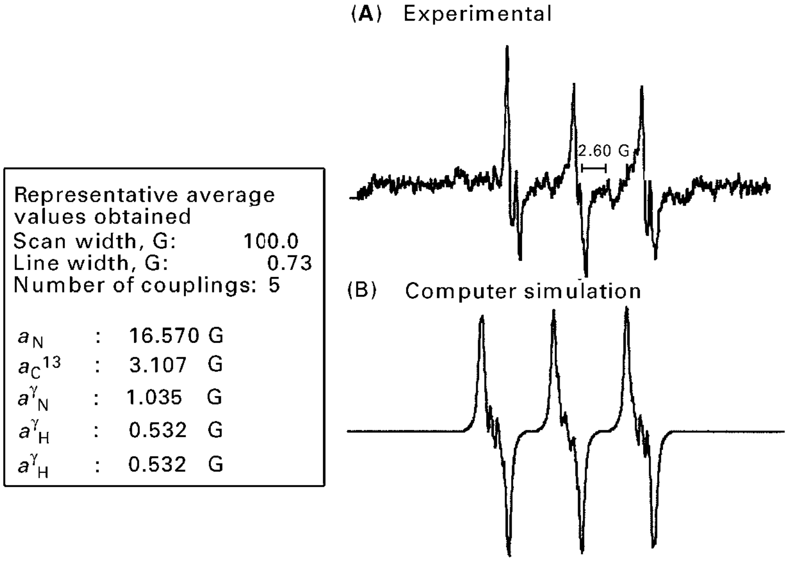



**Figure 7** (A) EPR spectrum of MNP-adducts observed when THP-1 (monocyte cells) suspensions were exposed to H-MG ($2 \times 10^{-4}$M) in the presence of MNP (7 mg). The magnetic field was set at 3350 G, microwave power 10 mW, modulation amplitude 8 G, microwave frequency 9.474 GHz. (B) EPR spectrum of the MNP adduct of $^{13}$C-labelled serine $^{13}C_2$ amino acid generated by H-MG via a hydrogen atom abstraction mechanism. Computer simulation using the EPR parameters is given in the box.

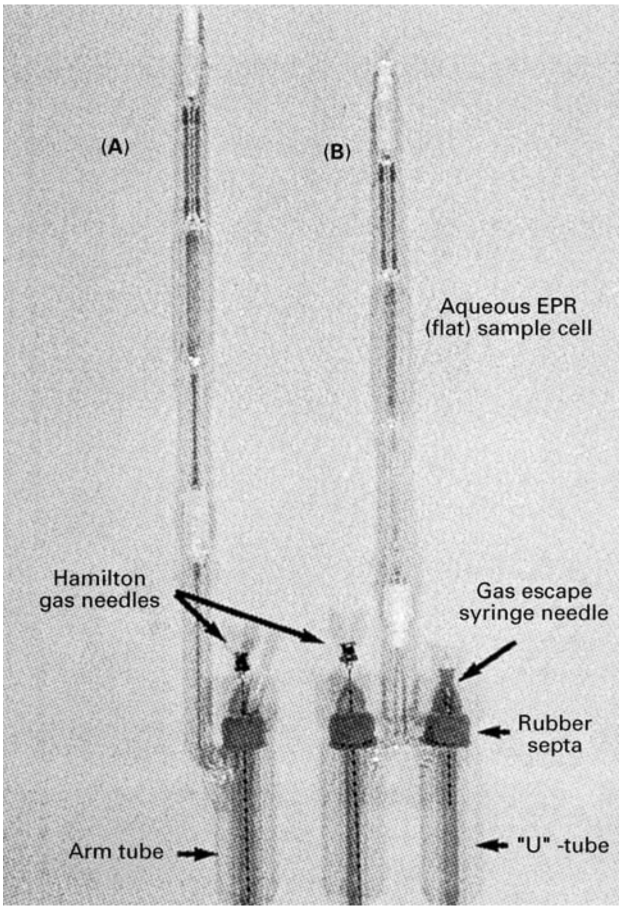



**Figure 8** Typical apparatus for gas aqueous or dielectric solvent EPR/spin trapping experiments.

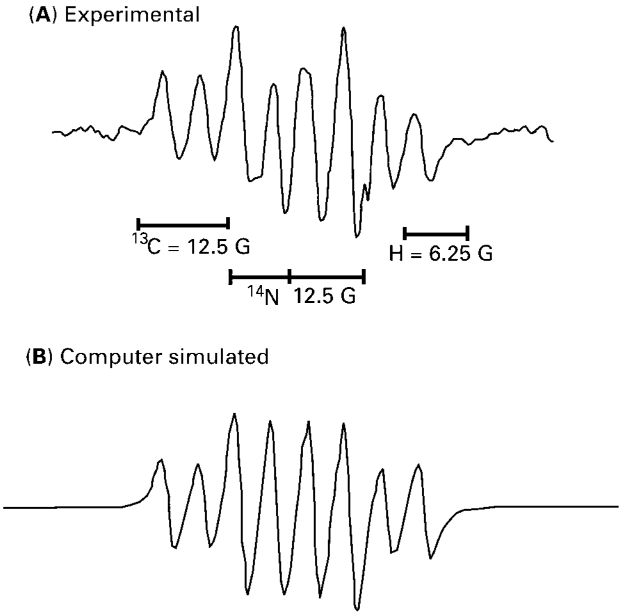



**Figure 9** (A) Spin adducts formation of PBN-$^{13}$COCl in benzene. Receiver gain $1.25 \times 10^5$; modulation amplitude 1.0 G. (B) Computer simulation of (A) using $a_N$ = 12.4 G; $a_H$ = 6.25 G and $a_{13C}$ = 12.5 G with a line width of 2.5 G. Insert: spectral parameters for $^{13}$C-phosgene obtained from the observed nitrone spin adducts.

## List of symbols

$B_0$ = magnetic flux density; $I$ = nuclear spin; $S$ = order parameter; $T_\parallel$ = outer hyperfine splitting; $T_\perp$ = inner hyperfine splitting.

*See also:* **Chemical Applications of EPR; EPR, Methods; EPR Spectroscopy, Theory.**

## Further reading

Anderson Evans C (1979) Spin trapping. *Aldrichimica Acta* **12**: 23–29.

Arroyo CM and Janny SJ (1995) EPR/Spin label technique as an analytical tool for determining the resistance of reactive topical skin protectant (rTSPs) to the breakthrough of vesicant agents. *Journal of Pharmacological and Toxicological Methods* **33**: 109–112.

Arroyo CM, Von Tersch RL and Broomfield CA (1995) Activation of alpha-human tumour necrosis factor (TNF-α) by human monocytes (THP-1) exposed to 2-chloroethyl sulphide (H-MG). *Human & Experimental Toxicology* **14**: 547–553.

Arroyo CM and Keeler JR (1997) Edemagenic gases cause lung toxicity by generating reactive intermediate species. In: Baskin SI and Salem H (eds) *Oxidants, Antioxidants, and Free Radicals,* Vol 17, pp 291–314. Washington, DC: Taylor & Francis.

Buettner GR (1987) Spin trapping: ESR parameters of spin adducts. *Free Radical Biology Medicine* **2**: 259–303.

Berliner LJ (1976) *Spin Labelling – Theory and Application.* New York: Academic Press.

Frejaville C, Karoui H, Tuccio, Le Moigne F, Culcasi M, Pietri, Lauricella R and Tordo P (1995) 5-(Diethoxyphosphoryl)-5-methyl-1-pyrroline N-oxide: A new efficient phosphorylated nitrone for the *in vitro* and *in vivo* spin trapping of oxygen-centered radicals. *Journal of Medicinal Chemistry* **38(2)**: 258–265.

Janzen EG (1980) A critical review of spin trapping biological systems. In: Pryor WA (ed) *Free Radicals in Biology*, Vol IV, pp 115–154. New York: Academic Press.

Mason RP, Stolze K and Morehouse KM (1987) Electron spin resonance studies of the free radical metabolites of toxic chemicals. *British Journal of Cancer* **8**: 163–171.

Mason RP, Hanna PM, Burkitt MJ and Kadiiska MB (1994) Detection of oxygen-derived radicals in biological systems using electron spin resonance. *Environmental Health Perspective* **102**(10): 33–36.

McConnell HM and McFarland BG (1970) Physics and chemistry of spin labels. *Quarterly Reviews of Biophysics* **3(1)**: 91–136.

Pantini G and Antonini A (1988) Perfluoropolyethers for cosmetics. *Drugs & Cosmetic Industry*, September 1988.

Speck JC (1959) Polychloro-7,8-disubstituted-2,5-diiminoglycoluril for use as an antivesicant. *United States Patent Office Patented*, No 2 885 305, May 5, 1959.

Thornalley PJ (1986) Theory and biological applications of the electron spin resonance technique of spin trapping. *Life Chemistry Reports* **4**: 57–112.

# Stark Methods in Spectroscopy, Applications

*See* **Zeeman and Stark Methods in Spectroscopy, Applications.**

# Stark Methods in Spectroscopy, Instrumentation

*See* **Zeeman and Stark Methods in Spectroscopy, Instrumentation.**

# Stars, Spectroscopy of

**AGGM Tielens**, Rijks Universiteit, Groningen, The Netherlands

Copyright © 1999 Academic Press

ELECTRONIC SPECTROSCOPY
Applications

## Introduction

Spectroscopy is the key to unlocking the information in starlight. Stellar spectra show a variety of absorption lines which allow a rapid classification of stars in a spectral sequence. This sequence reflects the variations in physical conditions (density, temperature, pressure, size, luminosity) between different stars. The strength of stellar absorption lines relative to the continuum can also be used in a simple way to determine the abundances of the elements in the stellar photosphere and thereby to probe the chemical evolution of the galaxy. Further, the precise wavelength position of spectral lines is a measure of the dynamics of stars and this has been used in recent years to establish the presence of a massive black hole in the centre of our galaxy and the presence of planets around other stars than the Sun.

## Stellar classification

All stellar spectra show absorption lines due to a variety of species. For the Sun, these were first discovered by Joseph von Fraunhofer in the early 1800s. A sample of stellar spectra is shown in **Figure 1**. The patterns in these lines allow stellar spectra to be grouped in a classification scheme. Depending on the spectral characteristics, stars are designated by a letter from the sequence O, B, A, F, G, K and M. This spectral sequence is summarized in **Table 1**. Since temperature controls ionization and excitation of the atoms, this spectral classification basically reflects a temperature sequence. This temperature sequence is also obvious from the colours of stars.

The strength of an absorption line is a measure of the opacity in the line compared with that in the continuum. For A stars, which have surface temperatures around 10 000 K, the $n = 2$ level of hydrogen can be excited and, because H is by far the most abundant element in (almost) all stars, the opacity and hence the visual region of the spectrum are dominated by absorption out of these levels giving rise to prominent Balmer lines (**Figure 1**). For higher stellar surface temperatures (e.g. O stars), the fraction of the H atoms excited to the $n = 3$ level – which provide the continuum opacity in the visible – increases more rapidly than that in the $n = 2$ level. As a result, the strength of the lines relative to the continuum decreases (**Figure 1** and **Table 1**). For the hottest stars, H is completely ionized and Thompson scattering by free electrons now dominates the continuum opacity. Lines of helium, which has a higher ionization potential, are still present. Of course, these He lines, which originate from levels much higher in energy than the ionization potential of hydrogen, require high stellar surface temperatures for their excitation. For stars much cooler than A, the $H^-$ ion provides the continuum opacity. Because the population of the levels leading to Balmer absorption of hydrogen becomes very small in such cool stars, the strength of the H lines decreases relative to the continuum. Various trace elements with lower energy

**Table 1** Spectral types

| *Spectral class*[a] | $T_{eff}$[b] (K) | *Colour* | *Spectral characteristics* |
|---|---|---|---|
| O | >30 000 | Bluish white | Relatively few lines. $He^+$ lines dominate |
| B | 10 000–30 000 | Bluish white | More lines; neutral He lines dominate; hydrogen Balmer lines developing |
| A | 7 500–10 000 | White | Very strong hydrogen Balmer lines, decreasing later; $Ca^+$ line appears |
| F | 6 000–7 500 | White | Hydrogen Balmer lines and ionized metal lines declining; neutral metal lines increasing |
| G | 5 000–6 000 | Yellow | Many metal lines; lines of $Ca^+$ strong; neutral metal lines continue to increase |
| K | 3 500–5 000 | Reddish | Molecular bands appear; neutral metal lines dominate |
| M | 2 500–3 500 | Red | Neutral metal lines strong; molecular bands dominate |

[a] Each spectral class is subdivided into subclasses ranging from 0 to 9 with 0 the hottest and 9 the coolest type in the class.
[b] Approximate photospheric temperature range for main sequence stars.

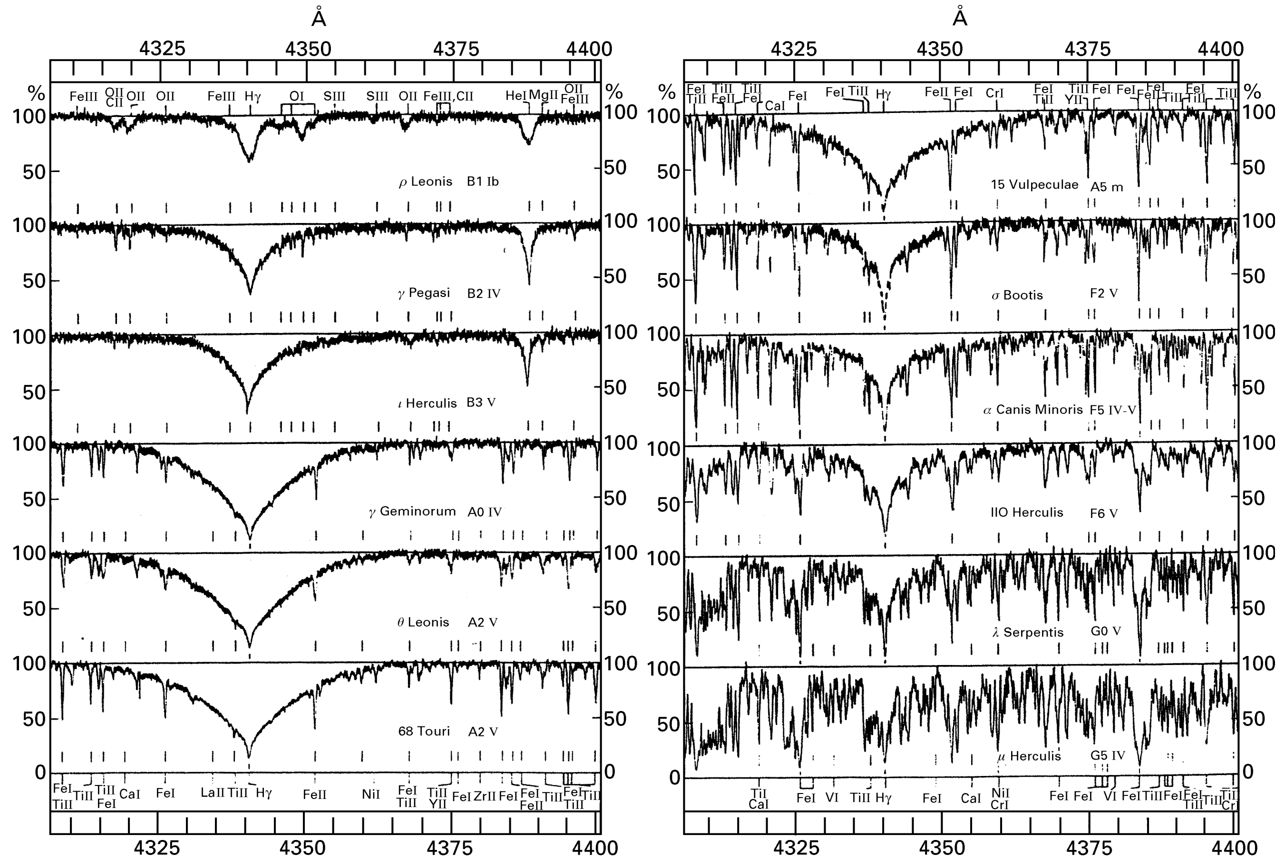



**Figure 1** Line intensities in the 4300–4400 Å region of spectra of representative stars of spectral types B–G. The more prominent lines are labelled at the top and bottom. Note how the strength of the Balmer Hγ line increases in strength relative to the continuum from top to bottom in the left-hand column and then decreases again in the right-hand column. At the same time, metal lines increase in strength, first from ionic and then from neutral species. Finally, note the difference in line width between different luminosity classes (see top left-hand side).

levels now take over the spectrum (**Figure 1** and **Table 1**). For the coolest K and M stars, most of the metals are neutral and molecules can survive in the stellar photosphere. Lines from these species will dominate the spectral appearance.

In addition to this spectral class, stars are also characterized by a luminosity parameter. This luminosity classification is made on the basis of the width of spectral lines. **Table 2** summarizes this classification. The width of spectral lines increases as the gas pressure increases. This so-called pressure broadening is due to the perturbation of atomic energy levels by other, nearby species. The physically largest stars have the lowest surface densities and pressures. Lines from these stars are therefore broader than from smaller stars (**Figure 1** and **Table 2**). This difference in size, which results in a difference in stellar luminosity, has led to the naming scheme from supergiants to dwarfs.

Stellar classification of a spectrum is actually based upon the intensity ratio of pairs of lines which are sensitive to temperature or luminosity. The lines used depend on the appropriate spectral type. This provides a straightforward way to classify stellar spectra and to determine rapidly the physical conditions in the stellar photosphere, including density, temperature, pressure, luminosity and size.

**Table 2** Luminosity classes[a]

| *Luminosity class* | *Star type* |
|---|---|
| I | Supergiants |
| II | Bright giants |
| III | Giants |
| IV | Subgiants |
| V | Main sequence[b] |
| VI | Subdwarfs |
| VII | White dwarfs |

[a] Along this sequence, the width of spectral lines increases from supergiants to white dwarfs.
[b] Main sequence stars are also known as dwarfs.

While these spectral and luminosity classes can be used to classify most stars, there is in addition a

bewildering collection of stars with spectra which deviate in various respects. Generally, these variations reflect differences in the elemental abundances in the photospheres of these stars. Most stars have abundances very similar to the Sun. However, deep in the interior of each star, nucleosynthesis converts hydrogen and helium into heavier elements such as carbon and nitrogen. In some stars, these freshly synthesized elements can be exposed in the stellar photosphere either owing to the effects of extensive mixing of deeper layers with the surface or because much of the stellar envelope has been lost in a stellar wind. **Table 3** contains a sample of such special stars and their spectroscopic characteristics.

The carbon stars are rare variable giants with temperatures similar to those of classes G and M. Their spectra are characterized by strong lines from carbon-bearing molecules (CN, $C_2$). These stars form when nucleosynthetically processed material, which is enriched in carbon owing to helium burning, is dredged up from the interior. When the abundance of carbon becomes larger than that of oxygen in the photosphere, all the oxygen is locked up in carbon monoxide and carbonaceous molecules rather than oxides dominate the composition. The spectra of these stars also show lines due to technetium (Tc), which is radioactive with a half-life of about 1 million years. Clearly, this element was recently formed in the interior and dredged up to the surface. Hence the presence of Tc in stellar spectra directly attests to the importance of nucleosynthesis in stellar interiors. So-called S stars are thought to be an intermediate stage in the stellar evolution from M to C stars where the abundances of carbon and oxygen approximately balance. ZrO bands now dominate the spectra (**Table 3**). Wolf Rayet stars are very luminous and hot stars with weak hydrogen lines and very strong helium lines. Their spectra show very wide emission features due to ionized He, C, N and O, originating in a wind from the star. Depending on type, they have excess carbon (WC) or excess nitrogen (WN) in their photosphere (**Table 3**). Some O, B and A stars (hot emission line stars) show hydrogen emission lines originating in a stellar wind.

**Table 3** Additional spectral types

| *Name* | *Spectral class* | *Spectral characteristics* |
|---|---|---|
| Carbon stars | C | Strong CN bands and $C_2$ bands |
| Heavy metal stars | S | ZrO bands |
| Wolf Rayet stars | WN | $N^{2+}$ and $N^{3+}$ emission lines |
| | WC | Ionized carbon and oxygen lines |
| Hot emission line stars | Of, Be, Ae | Bright hydrogen emission lines |

## Elemental abundances

The determination of stellar abundances is one of the main applications of stellar spectroscopy. The strength of photospheric absorption lines can provide information on the relative abundances of the elements in the photosphere of the star. Generally, this is done by measuring the strength of spectral lines relative to the continuum, the so-called equivalent width. The relationship between the equivalent width and the number of absorbers is called the curve of growth in stellar spectroscopy. For weak lines, the equivalent width is directly proportional to the number of absorbers. When the intrinsic strength of a line is larger or the number of absorbers is larger, the centre of the line saturates and the equivalent width becomes almost independent of the number of absorbing particles. For very strong lines or very large number of absorbers, absorption in the wings of the line become important and the equivalent width of the line will increase proportionally to the square root of the number of absorbers present. Stellar spectra contain many lines of a given element with known intrinsic strength. These can be used to construct an empirical curve of growth for that element. Comparison of such curves of growth for different elements yields then the relative elemental abundances. **Figure 2** shows an example for iron and titanium lines in the Sun. Similarly, we can compare the curves of growth for other stars with that for the Sun and determine elemental abundances for these stars relative to solar.

This semiempirical method is fairly straightforward but does assume that all the lines and the continuum involved are formed in the same region. Moreover, these line formation regions would have to have similar physical conditions in all stars. This

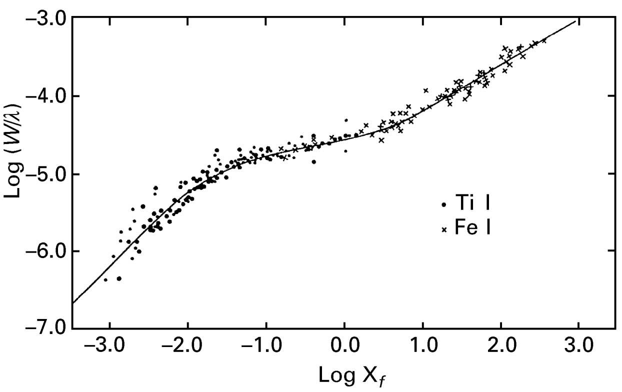



**Figure 2** Empirical curve of growth for solar Fe I and Ti I lines. The *y*-axis is the equivalent width (line strength relative to the continuum) and the *x*-axis is based on the oscillator strength of the transition.

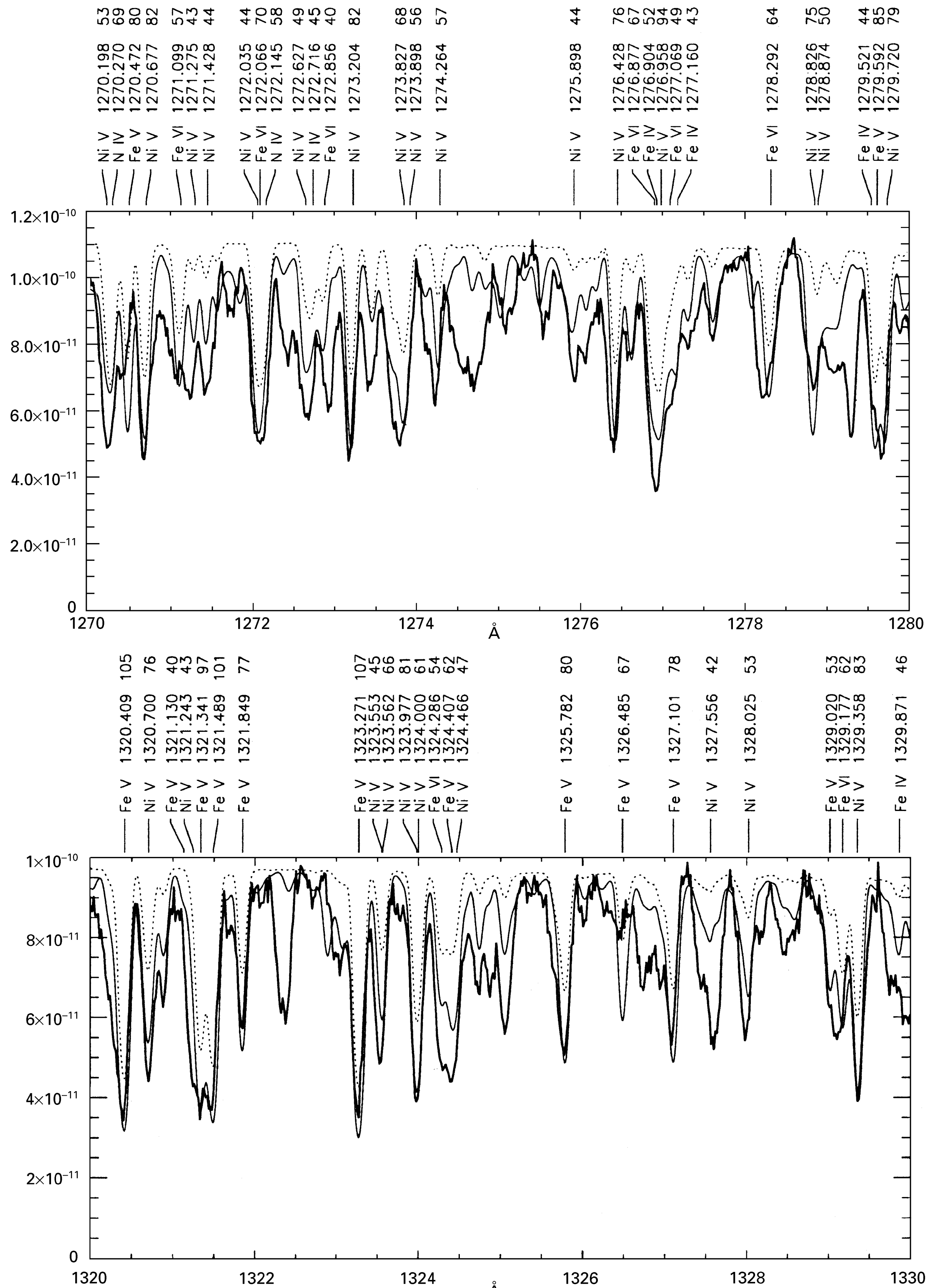


**Figure 3** Observations of the star BD+75°325′; (thick solid line) obtained by the Goddard high-resolution spectrometer on board the Hubble space telescope, showing numerous iron and nickel lines in various ionization stages. These data are compared with model atmosphere spectra. The thin solid line is the best fit model for an iron abundance of $4 \times 10^{-4}$ and the dotted line is for solar abundances ($4 \times 10^{-5}$). A solar iron-to-nickel ratio has been assumed in both models. Reproduced with permission from Lanz T, Hubeny I and Heap SR (1997) *Astrophysical Journal* **485**: 843.

is not always justified. Furthermore, cool stars have very crowded spectral regions where line overlap is a severe problem. In these cases, detailed modelling is a prerequisite for the determination of accurate stellar abundances. Sophisticated techniques have been developed which model in detail the physical structure of the stellar photosphere and its interaction with light. These models solve the equation of statistical equilibrium, regulating the individual level populations, the equation for hydrostatic equilibrium, governing the stellar pressure structure, and the radiative transfer equations describing the absorption and emission of light in the stellar photosphere. Comparison of models calculated for a variety of abundances with the observations allows the determination of the elemental abundances. In general, good agreement between models and observations can be obtained (**Figure 3**).

Analyses of this kind have shown that nearly all stars have very similar elemental compositions. The spectra of some stars, however, reveal much lower abundances than in the Sun. These so-called subdwarfs are metal-poor by factors up to 500. These compositional variations are correlated with the mass, age and dynamics of the stars within the Milky Way. The stars that formed first have the lowest elemental abundances. As those stars evolved – the more massive ones more rapidly than the less massive ones – they 'polluted' the interstellar medium with the nucleosynthetic products formed in their interiors either through a gentle wind (low-mass stars) or through a violent supernova explosion (massive stars). In this way, later generations of stars are formed from gas with higher elemental abundances. It is this elemental enrichment that drives the evolution of the Milky Way and other galaxies.

## Stellar dynamics

In addition to spectral and luminosity classification and abundance determination, stellar spectra also provide the radial motion of the absorbing gas through the Doppler shift. High-resolution spectra can thus provide information on stellar outflows in K and M giants, carbon stars, Of, Be and Ae emission line stars and Wolf Rayet stars. In general, the velocity information available in stellar spectra can be used to probe the dynamics of stars in the galaxy. For example, this technique has been used to trace the dynamics of stars in the centre of our galaxy. The derived velocity law implies a supermassive object in the centre of the galaxy with $3 \times 10^6$ solar masses. This provides strong evidence for the presence of a massive black hole in the centre of the Milky Way.

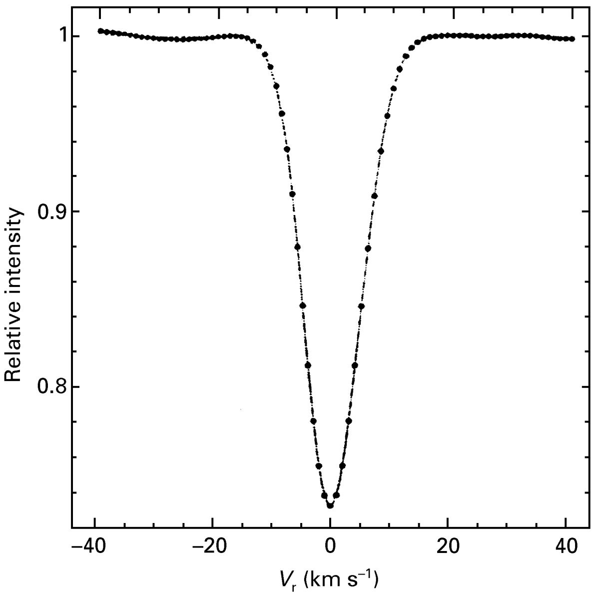



**Figure 4** Typical cross-correlation function used to measure the radial velocity. This figure represents the mean of the spectral lines of the star 51 Peg. The position of the Gaussian function fitted (solid line) is a precise measurement of the Doppler shift to an accuracy of about 15 m s$^{-1}$. The width of the cross-correlation function reflects the star's rotational velocity. Reproduced with permission of Macmillan Magazines Ltd. from Mayor M and Queloz D (1995) *Nature* **378**: 355.

By monitoring radial velocities over a long period, stellar spectroscopy can also be used to search for stellar companions, be they binary stars, brown dwarfs (stellar-like companions which are not massive enough to start hydrogen burning in their interiors) or planets. In recent years, new techniques have been developed to search for planets orbiting solar-type stars using Doppler shifts. In order to eliminate systematic wavelength shifts, the spectra are calibrated either by passing the stellar light through iodine gas or by using stable, fibre-fed spectrometers with simultaneous Th–Ar wavelength calibration. In long-term monitoring programmes, radial velocities of stars can then be measured with an accuracy of about 15 m s$^{-1}$, using a cross-correlation technique which concentrates the Doppler information for some 5000 stellar absorption lines (**Figure 4**). For comparison, the wobble introduced in the Sun's radial motion by Jupiter is about 13 m s$^{-1}$. The handful of planetary companions found so far in this way have masses in the range 0.5–10 Jupiter masses and orbits with solar system dimensions. In this way, stellar spectroscopy has allowed us to complete the 'Copernican' revolution.

*See also:* **Atomic Spectroscopy, Historical Perspective; Cosmochemical Applications Using Mass**

**Spectrometry; Environmental and Agricultural Applications of Atomic Spectroscopy; Interstellar Molecules, Spectroscopy of.**

## Further reading

Gustafsson B (1989) Chemical analyses of cool stars. *Annual Review of Astronomy and Astrophysics* **27**: 701–756.

Jaschek C and Jaschek M (1987) *The Classification of Stars*. Cambridge: Cambridge University Press.

Kudritzki RP and Hummer DG (1990) Quantitative spectroscopy of hot stars. *Annual Review of Astronomy and Astrophysics* **28**: 303–345.

Mihalas D (1970) *Stellar Atmospheres*. San Francisco: Freeman.

Yamashita Y, Nariai K and Norimoto Y (1977) *An Atlas of Representative Stellar Spectra*. Tokyo: University of Tokyo Press.

# Statistical Theory of Mass Spectra

**JC Lorquet**, Université de Liège, Belgium

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Theory**

## The model

Ionization via 70 eV electronic impact brings about a large number of Franck–Condon transitions to many electronic states of the ion. The molecular ion then undergoes internal conversions to its lowest electronic state owing to the existence of very fast and efficient radiationless transitions which themselves result from the presence of numerous crossings between potential energy surfaces. **Figure 1** shows a calculated set of potential energy curves for the $F_2^{\bullet+}$ molecular ion.

This picture gives some idea of the complexity to be expected, for example, in the case of the isoelectronic ion $CH_3CH_3^{\bullet+}$. For polyatomic molecular ions, the pattern of surface crossings is extremely complicated because the density of electronic states is much higher than that detected by photoelectron spectroscopy and because of the number of nuclear degrees of freedom. In a polyatomic system, the sequence of radiationless transitions results from the presence of conical intersections and is usually over after about $10^{-13}$ s. Once in the electronic ground state, vibrational energy is assumed to be redistributed (randomized) throughout the different vibrational degrees of freedom on a timescale short with respect to the reaction lifetime. When it fragments, the vibrationally excited molecular ion has forgotten its

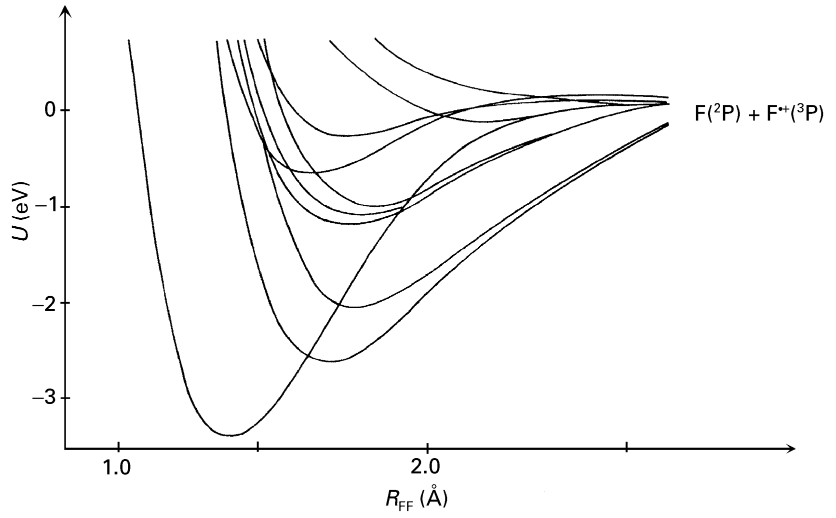



**Figure 1** Potential energy curves for the $F_2^{\bullet+}$ ion.

initial conditions. It undergoes a series of competing, consecutive, unimolecular reactions that can be described by a statistical theory. The theory is known under two acronyms, viz., RRKM (after Rice, Ramsperger, Kassel and Marcus) and QET (for quasi-equilibrium theory, as suggested by Rosenstock, Wahrhaftig and Eyring).

Probably the most convincing argument for the validity of a statistical approach comes from the independence of the fragmentation pattern with respect to the way energy is delivered. Except in rare cases, there is no correlation between the breakdown diagram determined from photoion–photoelectron coincidence (PIPECO) experiments and the photoelectron spectrum. In addition, when the molecular ion (e.g., $C_2H_4^{\bullet+}$) can be prepared by a collision process (e.g., $C_2H_2^{\bullet+} + H_2$ or $C^+ + CH_4$), the fragmentation ratios (e.g., $C_2H_4^{\bullet+}$: $C_2H_3^+$: $C_2H_2^{\bullet+}$) at a given internal energy are usually found to be equal within the experimental error to those obtained when the system is prepared by photon impact (see **Figure 2**).

These observations imply that the dissociation rate constants are smooth, monotonically increasing functions of the internal energy $E$ of the decaying molecular ion alone, irrespective of the way energy is delivered to the molecule, and do not depend on the electronic and vibrational quantum numbers that specify a particular molecular state. The criterion of validity of a statistical treatment (viz., that all of the properties of the system be completely characterized by a single parameter, its internal energy) is thus fulfilled. Exceptions to this pattern of behaviour exist, but they are rare. Many authors have attempted to develop a 'mode-selective chemistry', i.e. to promote reaction involving a chosen bond by excitation of that bond alone. These efforts have met with very limited success for unimolecular reactions, except in the case of van der Waals complexes where the coupling between the high-frequency intramolecular modes and the low-frequency intermolecular mode is weak.

## Phase space

In principle, the evolution of a molecular ion on its lowest potential energy surface is governed by the equations of classical (or quantum) mechanics. Consider a molecular ion made up of $N$ atoms. Its evolution is known when we know the values of the $3N$ coordinates and of the $3N$ conjugated momenta (related to generalized velocities) as a function of time. This information can be graphically represented as a trajectory in a $6N$-dimensional hyperspace, called the phase space.

Randomness is not accounted for by the deterministic classical equations. Indeed, at low internal energies, the nuclear trajectories appear regular (quasi-periodic). The molecular ion can then be described as a set of ($3N$–6) harmonic or anharmonic but separable oscillators to which, if relevant, one, two or three overall rotations have to be added. Only a small fraction of the available phase space is then visited. However, at higher internal energies, the usual model of a collection of ($3N$–6) independent oscillators breaks down. The nuclei then carry out extremely complicated trajectories and the fraction of phase space visited increases dramatically. If, for all trajectories characterized by a given energy, this fraction reaches a value of 100%, the system is declared 'ergodic' and statistical mechanics then generates exact equations. In technical terms, one has reached a state of microcanonical equilibrium. Even when the limit of 100% is not reached, useful equations can be derived from a statistical treatment, and this is the origin of the denomination 'quasi-equilibrium theory' adopted by Rosenstock and colleagues.

There are essentially two explanations that account for the success of the statistical approach.

### Efficient intramolecular vibrational energy redistribution (IVR)

The potential energy surface of the ground electronic state of the molecular ion is assumed to be so anharmonic that the various vibrational degrees of freedom are strongly coupled. IVR is then expected to be rapid compared with the timescale of the reaction. When it follows photon excitation or electron impact, IVR is a sequential process. First, after a time of the order of $10^{-13}$ s, energy is redistributed among the optically active modes. The second step (for which a timescale of the order of about $10^{-11}$ s can be proposed) consists of an energy exchange from symmetric to antisymmetric modes by anharmonic coupling (e.g. Fermi resonances). Efficient IVR implies that each individual nuclear trajectory has visited most parts of the available phase space before dissociating. This case is represented schematically in **Figure 3A**. The isolated molecular ion is then expected to reach spontaneously the ergodic limit under collision-free conditions. IVR is known to be inefficient in the case of van der Waals complexes (and probably also in the case of weakly bonded species) because of the disparity between the high intramolecular frequencies and the low intermolecular ones.

### Randomness of the initial conditions

For thermal reactions, the success of the statistical approach can be explained by invoking the

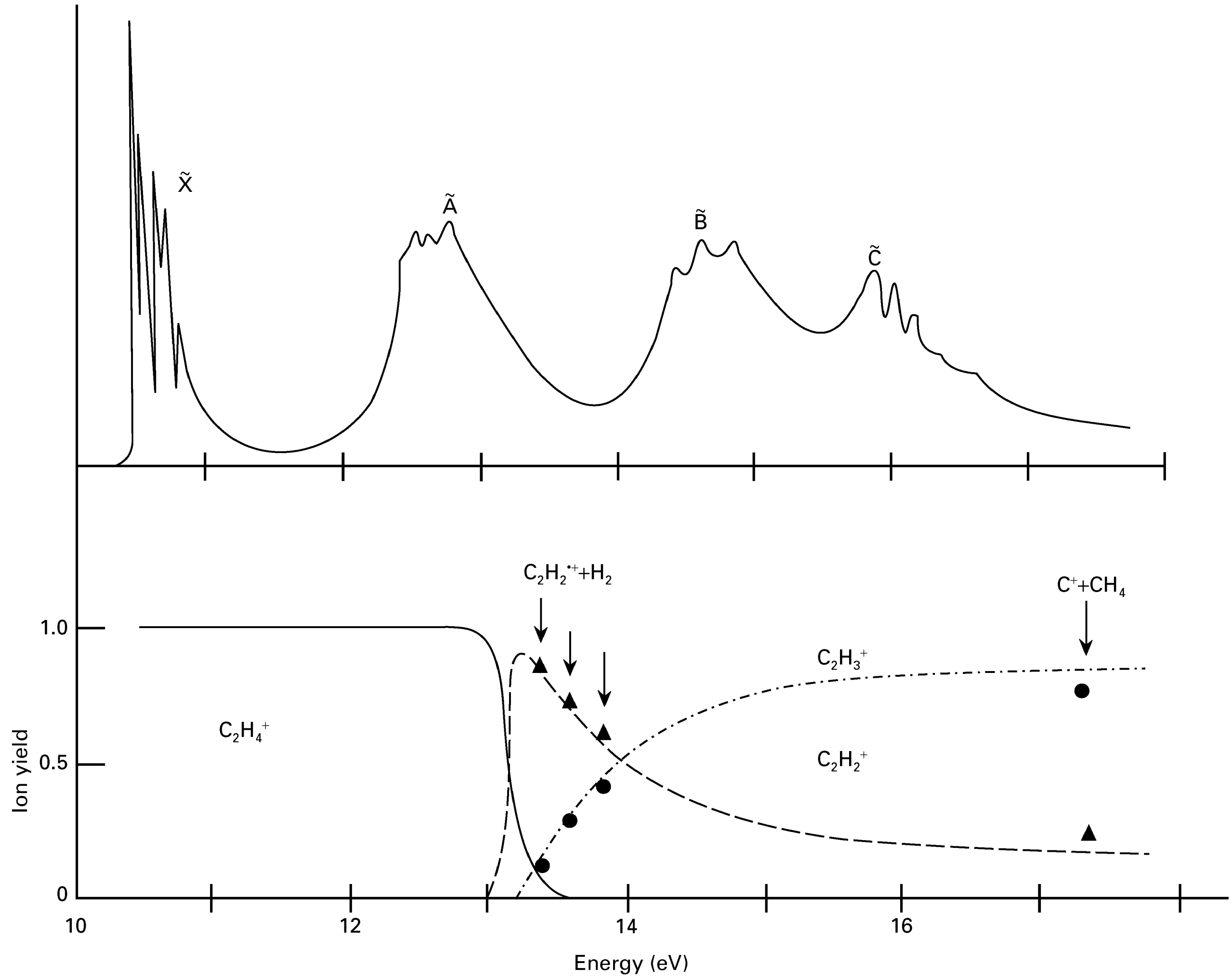



**Figure 2** Top: photoelectron spectrum of ethylene. Bottom: breakdown graph of ethylene determined by photoion–photoelectron coincidence measurements. Triangles and dots: yields of $C_2H_2^{\bullet+}$ and $C_2H_3^+$ ions, respectively, obtained by collision experiments. The ground and excited electronic states of the $C_2H_4^{\bullet+}$ ion are denoted by X̃, Ã, B̃ and C̃.

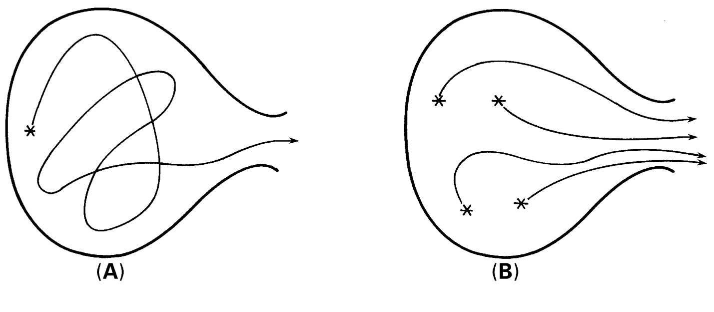



**Figure 3** (A) A single ergodic trajectory visits a substantial part of phase space. (B) A swarm of nonergodic trajectories originating from widely different initial conditions fills up the entire phase space.

impossibility of specifying the initial conditions if energy is delivered by molecular collisions. As shown in **Figure 3B**, even if each individual trajectory does not meet the requirements of ergodicity, an average over a large number of randomly distributed initial conditions will lead to a quasi-uniform sampling of the available phase space. For mass spectrometric experiments, this argument is apparently irrelevant. However, it should be noted that the initial conditions are also poorly defined, even if the experiment is carried out under collision-free conditions. As mentioned in the first section, the initial conditions are

determined by the conversion of electronic into vibrational energy via a cascade of radiationless transitions to the ground state of the molecular ion. As suggested by **Figure 1**, the pattern of surface crossings can be extremely complicated in a polyatomic system. Each process ends up at a different phase-space cell of the reacting ground-state ion. In addition, autoionization (a process in the course of which internal energy of the ionic core is transferred to a Rydberg electron, which is then ejected) generates a molecular ion characterized by its particular initial distribution of the vibrational energy. All of these processes lead to a pattern where the initial conditions are scattered at various positions in phase space. The averaging over these widely different initial conditions ensures the validity of a statistical approach.

## The transition state

The concept of transition state plays an essential role in the theory because it provides a simple, compact, but approximate description of the dynamics of the reaction. It can be defined in various ways. Essentially, one assumes that there exists some critical configuration of the reacting molecule, called the transition state, such that once the molecule has reached this configuration, it will irreversibly proceed on to products. (The term activated complex is sometimes used, but it is best to restrict its use to bimolecular reactions.)

The transition state, defined as it is as a point of no return, is often associated with the top of a potential energy barrier; that is, it is defined as a saddle point with a negative curvature of its potential energy surface along the reaction coordinate and hence with a single imaginary frequency. However, many molecular ions dissociate with no reverse activation energy barrier.

The transition state is not a stable molecule but a fictitious entity obtained after removal of the reaction coordinate, thus having one degree of freedom less than the reactant. Its best definition is that of a dividing surface in phase space. Reactant and products are assumed to be separated by a surface whose crossing is assumed to be irreversible (i.e., having crossed it once, the trajectories terminate as products without recrossing the surface backwards). The reaction is modelled as a flux through this dividing surface and its rate is measured by counting the number of times the surface is crossed per second (see **Figure 4**).

## The RRKM-QET equation

Consider a microcanonical ensemble of molecular ions with an energy between $E$ and $E + \delta E$. The total number of quantum states available to members of the ensemble is denoted $\rho(E)\delta E$, where the function $\rho(E)$ is called the density of states. Of these, a certain fraction corresponds to transition states. The rate constant $k(E)$ is proportional to the ratio of these two quantities. Many unknown parameters defining the transition state cancel out in the final expression, which reads simply

$$k(E) = \sigma \int_0^{E-E_0} \rho^{\ddagger}(E')\,\mathrm{d}E'/h\rho(E) = \sigma N^{\ddagger}(E-E_0)/h\rho(E) \qquad [1]$$

where $N^{\ddagger}$ denotes the number of accessible quantum states of the transition state, i.e. those having a potential energy $E_0$ and a translational kinetic energy $\varepsilon$ in the reaction coordinate, with the remainder $(E - E_0 - \varepsilon)$ in the bound degrees of freedom, and $\sigma$ is the reaction path degeneracy, i.e. the number of equivalent paths leading from the reactant to the products. (The quantities affixed with a double dagger refer to the transition state, which has one degree of freedom less than the reactant.)

At threshold (i.e. when $E = E_0$), $N^{\ddagger}(0) = 1$. Thus, the minimum rate is in principle equal to $1/h\,\rho(E_0)$. However, when the internal energy $E$ is equal to, or even is slightly less than the barrier, then the reaction proceeds via tunnelling and a transmission probability $\kappa$ is then inserted into Equation [1]. Isotope effects are then to be expected. When the shape of the barrier can be described by an inverted parabola, a simple expression can be derived for the transmission

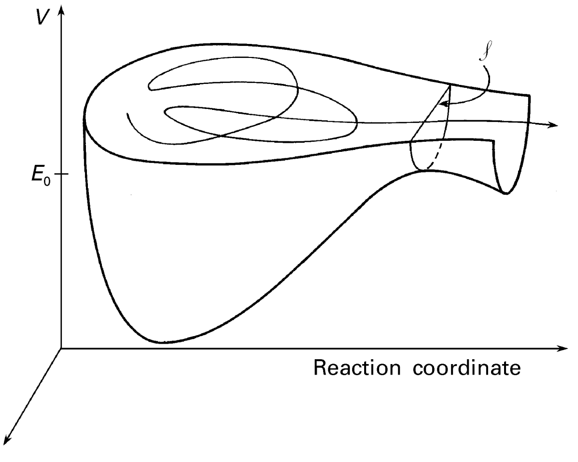



**Figure 4** Potential-energy surface for a unimolecular fragmentation, with the dividing surface s separating the reactant from the products.

coefficient:

$$\kappa(E) = \frac{\exp[2\pi(E - E_0)/h\nu^{\ddagger}]}{1 + \exp[2\pi(E - E_0)/h\nu^{\ddagger}]} \quad [2]$$

where $\nu^{\ddagger}$ is the modulus of the imaginary frequency of the saddle point. Expressions for more complicated, unsymmetrical barriers are also available.

## Practical calculations

There exists a computer algorithm (developed by Beyer and Swinehart) for counting all the possible combinations of the vibrational quantum numbers that are consistent with a specified value of the internal energy, even for a collection of anharmonic (but independent) oscillators.

Alternatively, analytical expressions for the energy-level densities can be obtained by inverse Laplace transformation of the corresponding partition function. This works well for the rotational degrees of freedom. For example, for a single rotor having a symmetry number equal to $\sigma$, one has simply (except at very low energies):

$$\rho(E) = \sigma^{-1}(BE)^{-1/2} \quad [3]$$

where $B$ is the rotational constant.

However, for a collection of oscillators, the set of energy levels is not dense enough and the energy-level density has to be numerically calculated by the steepest-descent method implemented by Forst. Nevertheless, an approximate closed-form expression involving an empirical correction for the effect of the zero-point vibrational energy has been developed by Whitten and Rabinovitch. For a system of $n$ oscillators having frequencies $\nu_i$, one has:

$$\rho(E) = \frac{(E + aE_z)^{n-1}}{(n-1)!\prod_{i=1}^{n} h\nu_i} \quad [4]$$

where $E_Z$ is the total zero-point energy and $a$ is an empirical parameter that usually ranges between 0.7 and 0.98. This formula leads to satisfactory results for the calculation of the denominator of Equation [1]. For its numerator, a direct count of the sum of states $N^{\ddagger}(E - E_o)$ is preferable, because at lower energies the vibrational levels are usually too widely spaced for a classical or semiclassical approximation to work.

The numerical values of the frequencies and moments of inertia needed for an actual calculation can be obtained from commonly used quantum-chemistry programs. This is the only possibility for the frequencies of the transition state.

The role played by the overall rotations raises a difficult problem. The constraint of angular momentum conservation prevents certain rotational degrees from getting involved in randomization. Very often, the transition state is approximated as a symmetric top, with two equal moments of inertia. It is then usually assumed – possibly incorrectly – that the degenerate two-dimensional external rotation is unable to couple significantly with the vibrational degrees of freedom, whereas the rotation about the reaction coordinate is able to exchange energy with them. However, although unavailable for randomization, the rotational energy stored in the degenerate external two-dimensional rotation modifies the radial potential and gives rise to an effective potential characterized by a centrifugal barrier.

## More elaborate statistical models

It has been found useful to introduce new concepts and methods, particularly in the study of reactions for which no potential barrier is encountered along the reaction coordinate, a case often encountered in the dissociation of molecular ions. In such a case, the conservation of the orbital angular momentum during the reaction process has important consequences.

### Microcanonical variational transition state theory (VTST)

If the transition state is defined as a structure denoting unstable equilibrium between the reactant and the products, then any point of its phase space having a nonzero velocity in the forward direction will react. However, this criterion is oversimplified: the condition is necessary but not sufficient. After having left the transition state, some trajectories may return to cross it again. Hence, the rate constant calculated by the transition state theory is an overestimate of the exact value. (This conclusion has been challenged, however, when energy is selectively delivered to the system, as in laser light excitation.) Leaving the controversial cases aside for deeper scrutiny, it follows that the best choice for the transition state is the one that minimizes the calculated rate constant. Therefore, in VTST, the numerator of Equation [1] is written as $N^{\ddagger}[E - V_{\text{eff}}(R^{\ddagger})]$, where $V_{\text{eff}}(R)$ denotes the effective potential for motion along the reaction coordinate $R$ (usually, the sum of the actual and a centrifugal potential), and where $R^{\ddagger}$

denotes the position of the dividing surface that minimizes $N^{\ddagger}$.

### Transition state switching

The rate constant does not necessarily admit a unique minimum. Thus, several transition states may simultaneously exist along the reaction coordinate with characteristics depending on the energy and angular momentum. For example, a tight transition state may be followed by an orbiting complex. The former determines the magnitude of the rate constant, whereas the latter controls the translational, rotational and vibrational energy distributions of the products.

### Bottlenecks

This term is often encountered, but is used with two different connotations. First, it may describe an impediment to IVR (resulting, for example, from the presence of a heavy atom or from a disparity in the frequencies). Alternatively, it may describe a throttling of the reactive flux (i.e., a new kind of transition state) due, for example, to a region of strong curvature of the reaction path resulting, for example, from a detour around a conical intersection between two potential-energy surfaces.

### The statistical adiabatic channel model (SACM)

Troe and Quack have defined reaction channels as a set of potential-energy curves obtained by connecting the vibration–rotational states of the reactant to those of the pair of products along the reaction coordinate. (This procedure is referred to as an adiabatic correlation.) Barriers arise when imposing conservation of the orbital angular momentum during the reaction process to the transitional modes (i.e. the vibrational bending modes that correlate with rotations or orbital motion of the fragments). A channel is said to be open when its potential energy curve never exceeds the available internal energy $E$. No reference to a transition state is made in this theory. The rate constant now depends on two parameters, viz., the internal energy $E$ and the total angular momentum $J$. Its expression is remarkably similar to the RRKM-QET equation:

$$k(E, J) = \frac{N^*(E, J)}{h\rho(E, J)} \qquad [5]$$

where $N^*(E, J)$ is now defined as the total number of open channels.

### The orbiting transition state phase space theory (OTS/PST)

The unimolecular reactions of polyatomic ions are assumed to be governed by the long-range part of the potential. The principle of microscopic reversibility has been used by Klots to express the rate constant, not in terms of the properties of a transition state, but in terms of the cross section for association of fragments. The latter is then evaluated by the Langevin theory, which is based on the consideration of a long-range attractive potential (e.g. $-\alpha q^2/2R^4$). Alternatively, an orbiting transition state located at a centrifugal barrier has been reintroduced by Chesnavich and Bowers. In both cases, the density of states is no longer evaluated as a simple convolution between the vibrational and rotational parts, but by introducing restrictions resulting from the conservation of angular momentum. This theory often provides too large estimates for the rate constant.

### Effective number of states

Information theory shows that if the assumption of complete energy randomization breaks down, then Equation [1] remains valid provided it is interpreted as a ratio between an *effective* number of states divided by an *effective* density of states. Both the numerator and the denominator are then reduced (but not necessarily to the same extent) by a so-called entropy deficiency factor. A mechanism of cancellation of errors arises, which accounts for the success of the simple theory.

### Nonadiabatic reactions

Nonadiabatic reactions (i.e. those involving a change in the electronic state or structure) occur more frequently in mass spectrometry than is commonly thought. Surprisingly enough, statistical methods are found to be useful even in these cases.

## Kinetic energy release distributions (KERDs)

The internal energy in excess of the thermodynamic dissociation threshold is partitioned among the translational, rotational and vibrational degrees of freedom of the pair of fragments. In the case of molecular ions, the translational kinetic energy release distribution (KERD) can be measured and its study is a precious source of information on the dynamics of the reaction. Indeed, it is a stronghold of the more elaborate statistical theories presented in the previous paragraph because the usual RRKM-QET theory is in principle not appropriate for such a

study. The reason is that transition state theory concentrates on the behaviour of the system up to the dividing surface, whereas the product energy disposal is determined by the potential felt by the fragments as they separate. However, Marcus, Wardlaw and Klippenstein have proposed an extension that describes the evolution of the transitional modes along the reaction coordinate and that, at the same time, conserves angular momentum.

Phase space orbiting transition state theory works much better for the calculation of KERDs than for that of the rate constants, thereby demonstrating that the former are controlled by the long-range part of the potential whereas the latter are governed by its shorter range. In addition, Klots has introduced a set of effective temperatures to parametrize the observed distributions. The SACM has also demonstrated its usefulness in the case of weakly bonded species.

The maximum entropy method and the associated surprisal theory are an outgrowth of information theory. They involve a comparison between the actual shape of the KERD and the hypothetical, most statistical, so-called 'prior distribution'. Two precious pieces of information can be derived from this comparison: (i) an identification of the constraint that operates on the dynamics and prevents it from being statistical and (ii) the magnitude of the entropy deficiency which can be related to the fraction of phase space effectively sampled by the transition state. Values of 75–80% have been obtained in the case of the halogenobenzene ions.

## Concluding remarks

It has often been naggingly remarked that the RRKM-QET theory can fit anything and predict nothing. To counter this criticism, many authors have multiplied skilful consistency checks (study of isotope effects, preparation of the ion via a bimolecular reaction or via charge reversal in addition to electron or photon impact, time-resolved studies all the way from the millisecond to the nanosecond timescales, etc.) and have removed arbitrariness via *ab initio* calculations of frequencies. However, it should be realized that ability to fit the experiments by no means implies that the theory is exact and that its basic assumptions (full energy randomization and existence of a good transition state) are fulfilled. It has been seen that Equation [1] cannot be grossly in error because of a mechanism of cancellation of errors. In contradistinction, KERDs (for which the cancellation of errors does not work because they basically depend on the numerator only) provide a much better way to test the validity of the assumption of complete energy randomization than a study of fitted rate constants.

## List of symbols

$B$ = rotational constant; $E$ = internal energy; $E_Z$ = zero-point energy; $h$ = Planck constant; $J$ = total angular momentum; $k(E)$ = rate constant; $N^\ddagger$ = number of accessible states of the transition state; $R^\ddagger$ = position of dividing surface that minimizes $N^\ddagger$; $V_{\text{eff}}(R)$ = effective potential for motion along reaction coordinate $R$; $\kappa$ = transmission probability; $\nu^\ddagger$ = modulus of imaginary frequency of saddle point; $\nu_i$ = oscillator frequency; $\rho(E)$ = density of states; $\sigma$ = symmetry number.

*See also:* **Fragmentation in Mass Spectrometry; Ion Dissociation Kinetics, Mass Spectrometry; Metastable Ions; Photoelectron-Photoion Coincidence Methods in Mass Spectrometry (PEPICO).**

## Further reading

Baer T and Hase WH (1996) *Unimolecular Reaction Dynamics. Theory and Experiments.* Oxford: Oxford University Press.

Baer T (1996) The calculation of unimolecular decay rates with RRKM and *ab initio* methods. In: Baer T, Ng CY and Powis I (eds) *The Structure, Energetics and Dynamics of Organic Ions*, pp 125–166. Chichester: Wiley.

Baer T and Mayer PM (1997) Statistical Rice–Ramsperger–Kassel–Marcus quasiequilibrium theory calculations in mass spectrometry. *Journal of the American Society for Mass Spectrometry* 8: 103–115.

Chesnavich WJ and Bowers MT (1979) Statistical methods in reaction dynamics. In: Bowers MT (ed) *Gas Phase Ion Chemistry*, Vol 1, pp 119–151. New York: Academic Press.

Derrick PJ and Donchi KF (1983) Mass spectrometry. In: Bamford CH and Tipper CFH (eds) *Comprehensive Chemical Kinetics*, Vol 24, pp 53–247. Amsterdam: Elsevier.

Gilbert RG and Smith SC (1990) *Theory of Unimolecular and Recombination Reactions.* Oxford: Blackwell.

Illenberger E and Momigny J (1992) *Gaseous Molecular Ions.* Darmstadt: Steinkopff.

Lifshitz C (1989) Recent developments in applications of RRKM-QET. *Advances in Mass Spectrometry* **11A**: 713–729.

Lifshitz C (1992) Recent developments in applications of RRKM-QET. *Advances in Mass Spectrometry* **12**: 315–337.

Lorquet JC (1994) Whither the statistical theory of mass spectra? *Mass Spectrometry Reviews* **13**: 233–257.

Lorquet JC (1996) Non-adiabatic processes in ionic dissociation dynamics. In: Baer T, Ng CY and Powis I (eds) *The Structure, Energetics and Dynamics of Organic Ions*, pp 167–196. Chichester: Wiley.

Wardlaw DM and Marcus RA (1988) On the statistical theory of unimolecular processes. *Advances in Chemical Physics* 70: 231–263.

# Stereochemistry Studied Using Mass Spectrometry

**Asher Mandelbaum**, Technion–Israel Institute of Technology, Haifa, Israel

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Applications**

Stereochemical effects in mass spectrometry have been used for configurational assignment in numerous organic systems. The unsurpassed sensitivity of mass spectrometry and the possibility of interfacing the mass spectrometer with a variety of separating devices (most commonly gas chromatography–mass spectrometry (GCMS) and liquid chromatography–mass spectrometry (LCMS)), make it a most useful tool for structural assignment of minor constituents of complex mixtures, including stereochemical information. The mass spectral stereochemical effects may also be an important and useful tool in structural studies of gas-phase ions and in mechanistic investigations of their fragmentation processes.

Mass spectral stereochemical effects occur in the molecular radical cations $M^{\bullet+}$, obtained usually on electron ionization (EI), in protonated molecules $MH^+$ upon chemical ionization (CI) and other soft ionization techniques (fast atom bombardment (FAB), electrospray, thermospray), and, to a lesser extent, in negative ions. These effects are observed in the normal mass spectra and also by tandem mass spectrometric techniques (MS/MS).

The different mass spectral behaviour of stereoisomers (this term will not include enantiomers in this article) may be due either to the different thermochemical nature of the chemistry of their gas-phase ions, or it may result from the different kinetics of their reactions. These two aspects will be dealt with in the following sections.

## Thermochemical considerations

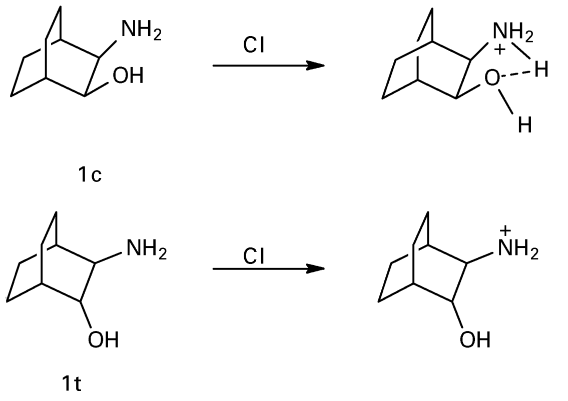
### Electron ionization

Small, often insignificant or within the experimental error, differences (in most reports below 0.1 eV) have been observed in the ionization energies of stereoisomers in a number of systems. The appearance energies of some specific fragment ions often show a more pronounced dependence on the configuration of stereoisomers. Lower energies have been reported for $[M–R]^+$ ions obtained from stereoisomers with higher enthalpies of formation in various dialkylcycloalkanes and in other more complex heterocyclic and polycyclic systems. The difference between the appearance energies could be quantitatively correlated with the difference in the enthalpies of formation of the stereoisomers in several systems. The lower appearance energies of the thermochemically less stable stereoisomers have been attributed to the release of steric strain in the course of the fragmentation.

Since the early days of mass spectrometry it has been proposed, that the release of steric strain results, in many cases, in lower abundances of the molecular ions and in higher abundances of fragments in the EI mass spectra of the thermochemically less stable isomers in numerous systems. The differences between the mass spectra of stereoisomers in those systems are usually not large, and there are numerous exceptions which cast doubt on the reliability of this approach in real problems of configurational assignments.

### Chemical ionization

In contrast to the ionization energies, considerably different proton affinities and gas-phase basicities have been reported for a number of stereoisomeric difunctional compounds, which have an impact on their behaviour upon chemical ionization. For example, the gas-phase basicities of *cis*- and *trans*-2-amino-3-hydroxybicyclo[2.2.2]octanes 1c and 1t (**Scheme 1**) are 926 and 909 kJ $mol^{-1}$, respectively.




**Scheme 1**

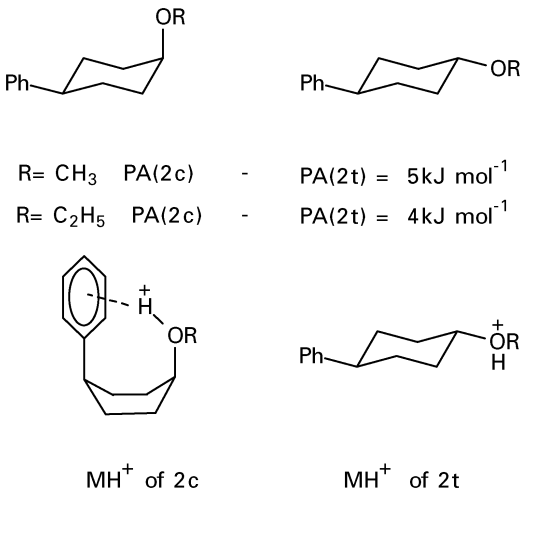
The difference is attributed to the different interfunctional distance in the two stereoisomers (2.31 and 3.50 Å respectively) resulting from their different dihedral angles about the HOC–$CNH_2$ bond (30° and 150°). The effective intramolecular hydrogen bond (often termed proton bridging), which is possible in the $MH^+$ ion of the *cis*-amino alcohol 1c (**Scheme 1**), results in the higher gas-phase basicity of this stereoisomer.

A smaller difference (4–5 kJ $mol^{-1}$) has been reported between the stereoisomeric *cis*- and *trans*-4-phenyl-1-alkoxycyclohexanols 2c and 2t (**Scheme 2**), and it has been attributed to proton bridging between the alkoxylic oxygen atom and the phenyl ring, which is possible only in the $MH^+$ ions of 2c (**Scheme 2**).

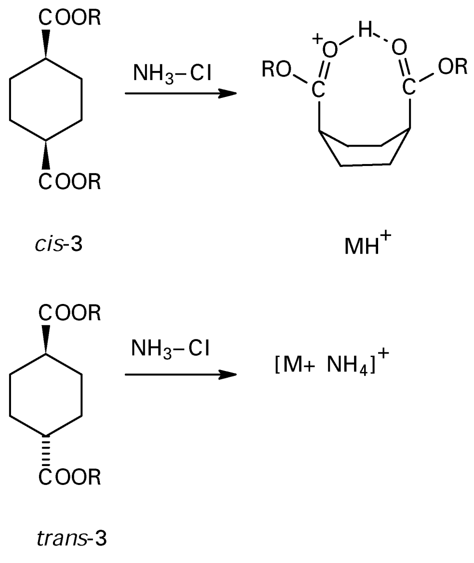
Easy distinction between stereoisomers containing two or more basic sites, may be achieved upon CI using selected reagent gases, based on the remarkable difference in their proton affinities and gas-phase basicities, due to the effectiveness of proton bridging. Thus, diesters with remote alkoxycarbonyl groups (e.g. fumarates, mesaconates, *trans*-1,3 and 1,4-cyclohexane dicarboxylates) give rise to low abundance $MH^+$ ions and to abundant $[M+NH_4]^+$ adduct ions upon $NH_3$-CI, because of the higher proton affinity of ammonia. On the other hand, the *cis* analogues with adjacent ester groups (e.g. maleates, citraconates, *cis*-1,3- and 1,4-cyclohexane dicarboxylates), which may undergo intramolecular proton bridging on protonation, afford abundant $MH^+$ ions under $NH_3$-CI (**Scheme 3**).

Stabilization of the $MH^+$ ions by intramolecular hydrogen bonding may be used as a simple tool in the configurational assignment of stereoisomeric diols, diethers and other difunctional analogues. Stereoisomers with adjacent basic functions afford abundant proton-bridged $MH^+$ ions under CI conditions. Counterparts with remote functions, which cannot be stabilized by proton bridging, undergo fragmentation processes resulting in less abundant $MH^+$ ions in their CI mass spectra. Typical examples are shown in **Scheme 4**, and many others have been reported.

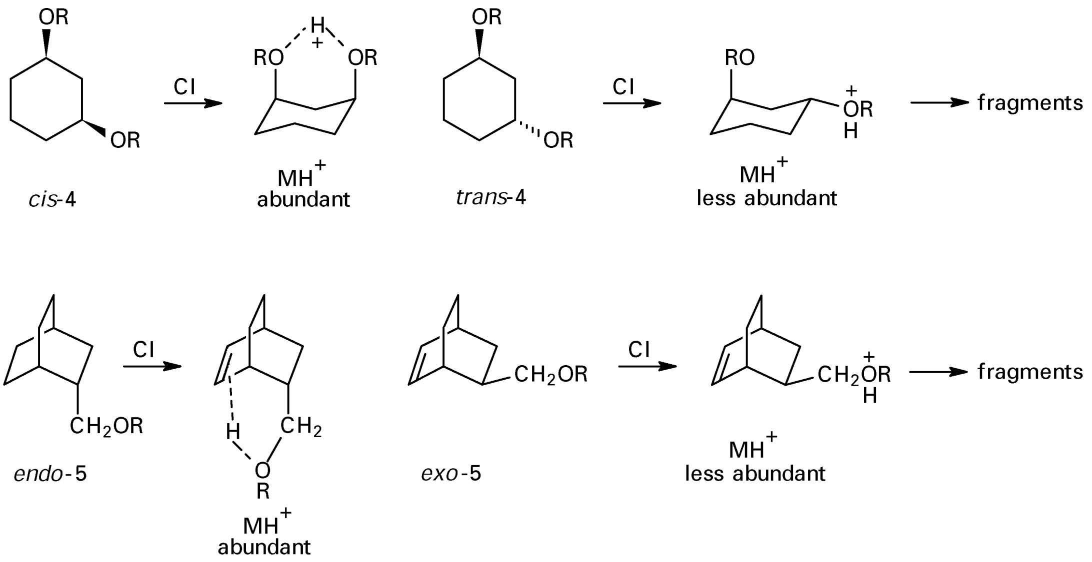
The above stabilization effect is limited to systems such as 4 and 5 (**Scheme 4**), containing at least one functional group (e.g. OH, OR) which undergoes ready dissociation upon CI. Other systems, e.g. diesters, with functional groups that are stable under CI conditions, behave in a different manner. Protonated esters of alkanoic or cycloalkanoic acids exhibit a high degree of stability upon isobutane-CI. Protonation occurs at the carbonylic oxygen atom,




**Scheme 2**

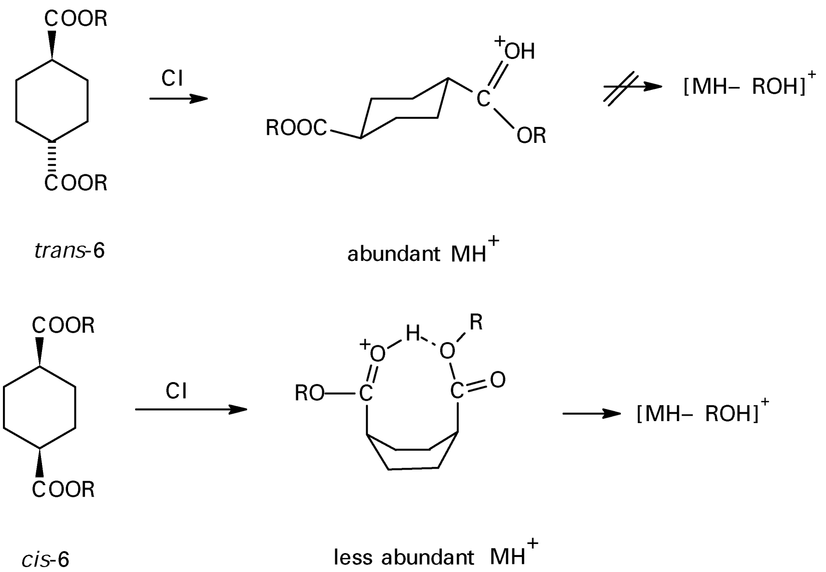



**Scheme 3**




Scheme 4

and the 1,3-proton transfer to the alkoxy oxygen, that would enable alcohol elimination, is a symmetry forbidden process. The nonbridged $MH^+$ ions of diesters, with the two remote noninteracting alkoxycarbonyl groups (e.g. *trans*-6) (**Scheme 5**), are stable (and consequently highly abundant) under isobutane-CI. On the other hand, stereoisomers with adjacent ester groups (e.g. *cis*-6) (**Scheme 5**), undergo facile proton transfer between the two alkoxycarbonyls *via* proton-bridged intermediates, resulting the efficient elimination of ROH (**Scheme 5**). In the latter systems, the thermochemically stabilized proton-bridged species are kinetically destabilized, resulting in less abundant $MH^+$ ions.




Scheme 5

## Kinetic considerations

A wide variety of fragmentation processes of the $M^{\bullet+}$ and $MH^+$ ions obtained upon EI and CI involve formation of new bonds. Such unimolecular rearrangements take place via cyclic transition states. Molecular ions of stereoisomers often differ in the accessibility of the transition structures for fragmentation processes, which involve bond formation or concerted rupture of several bonds. Relatively large differences in the energy of activation of such processes may result in extreme cases in total suppression of particular pathways in one of the stereoisomers. A large number of systems have been reported that show stereospecific fragmentation behaviour upon EI and CI. A limited number of typical examples will be given in the following.

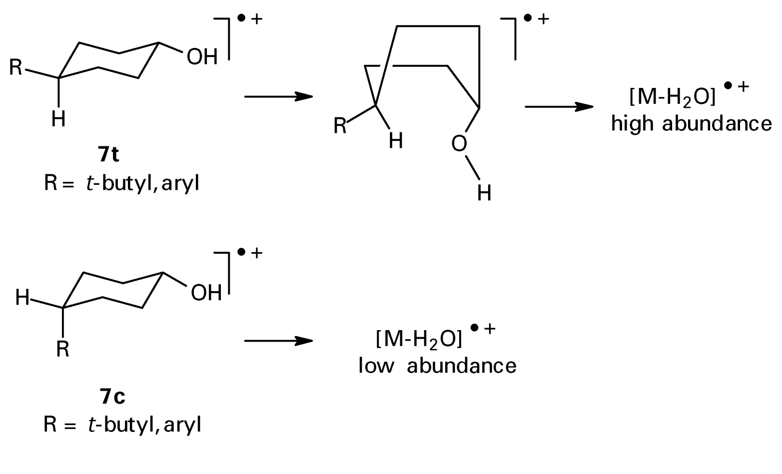
### Hydrogen transfer

**Elimination of $H_2O$ from cycloalkanols and related processes** One of the early cases of stereospecific fragmentation processes of gas-phase ions was the EI-induced dehydration of stereoisomeric 4-sustituted cyclohexanols. *Trans*-4-*t*-butyl- and -4-arylcyclohexanols 7t (**Scheme 6**) afford highly abundant $[M-H_2O]^{\bullet+}$ ions, which are much less abundant in the mass spectra of the *cis* isomers 7c (**Scheme 6**). Similar stereospecificity has been also observed in the elimination of methanol and acetic acid from the corresponding methyl ethers and acetates. Deuterium labelling studies have shown that the H-atom from position 4 is abstracted in the course of the elimination of $H_2O$ from the *trans*-alcohols, but not to an appreciable extent in the case of the *cis* isomers. The low energy H-transfer from the tertiary (and benzylic when R=aryl) position 4 to the hydroxy group is the stereospecific step in this fragmentation. The cyclic transition state of this H-transfer is attainable only in 7t (**Scheme 6**) in boat conformation. The high stereospecificity of this process indicates retention of the original structure in the molecular radical cation.

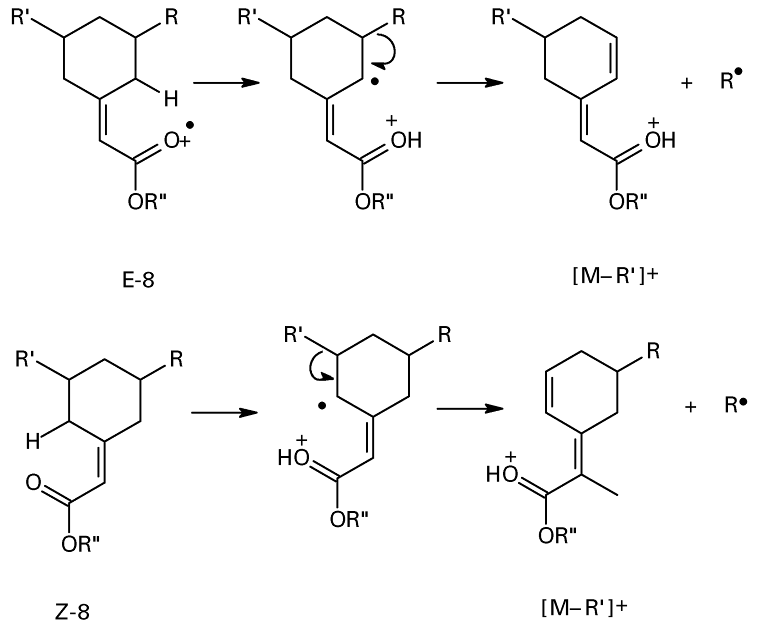
A similar stereospecific behaviour has also been observed in the stereoisomeric 3-arylcyclohexanols and in their methyl ethers and acetates, but not in the 2-aryl analogues, which undergo cleavage of the $C_1$–$C_2$ bond prior to fragmentation, resulting in the loss of stereochemical information. Many additional systems have been shown to undergo EI-induced stereospecific eliminations of water and other neutral molecules, which may be applied in structural studies.

It is noteworthy that stereoisomeric cycloalkanols and their derivatives exhibit a lower degree of stereospecificity in their dissociation upon CI. Protonation occurs at the oxygen atom in these materials, and the concurrent elimination, that takes place by a simple cleavage of the C–O bond, is governed by thermochemical factors.

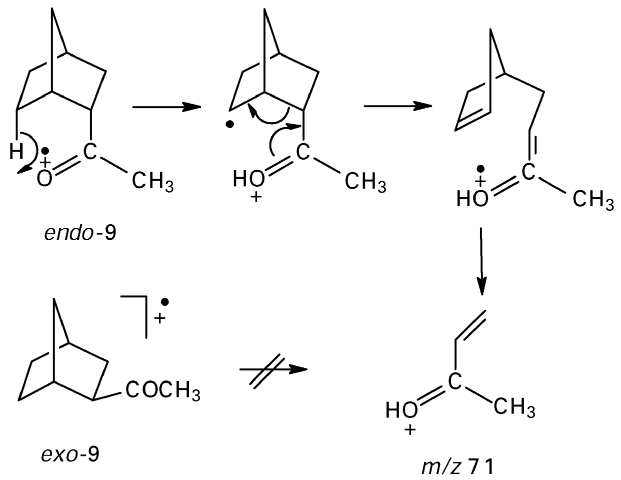
**Cycloalkylidene acetates** The geometrically isomeric *E*- and *Z*- disubstituted cycloalkylidene acetates (e.g. E-8 and Z-8) exhibit a highly specific EI-induced loss of one of the two substituents R and R′ at the two homoallylic positions, as shown in **Scheme 7**. The specific H-transfer from the allylic position adjacent to the carbonyl toward the carbonylic oxygen atom has been proposed as explanation for the stereospecificity of this process. These results enable easy distinction and configurational assignment of stereoisomers in analogous systems




**Scheme 6**




Scheme 7

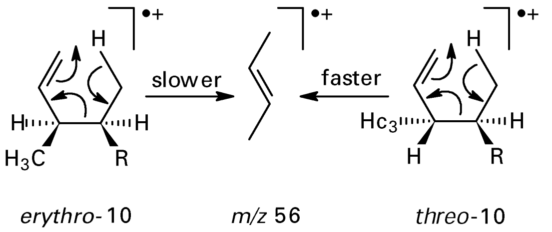



Scheme 8

(including acyclic analogues), which is not straightforward by other spectroscopic methods. They also indicate that the rotation about the double bond in the $M^{\bullet+}$ ions must be slower than the H-transfer in this process.

**McLafferty rearrangement and related processes** The stereospecificity of the McLafferty rearrangement is exemplified by the different behaviour of *endo*- and *exo*-acetylnorbornanes *endo*-10 and *exo*-10 upon EI, shown in **Scheme 8**. The *m/z* 71 ion is abundant in the EI-mass spectrum of *endo*-10, but absent in that of *exo*-10. McLafferty rearrangement is the initial step in the formation of the *m/z* 71 ion with the involvement of the H-atom from position 6, which is in the proper distance from the carbonylic oxygen only in the *endo*- isomer.

Steric interactions in the cyclic transition structures for a hydrogen transfer result in different abundance of the butene radical cation fragment, obtained from the diastereoisomeric unsaturated alcohols *erythro*-10 and *threo*-10. The higher energy of the transition state of *erythro*-10, due to the steric interaction of the methyl and R groups, results in a lower abundance of the fragment, as compared with the stereoisomeric *threo*-9 (see **Scheme 9**).



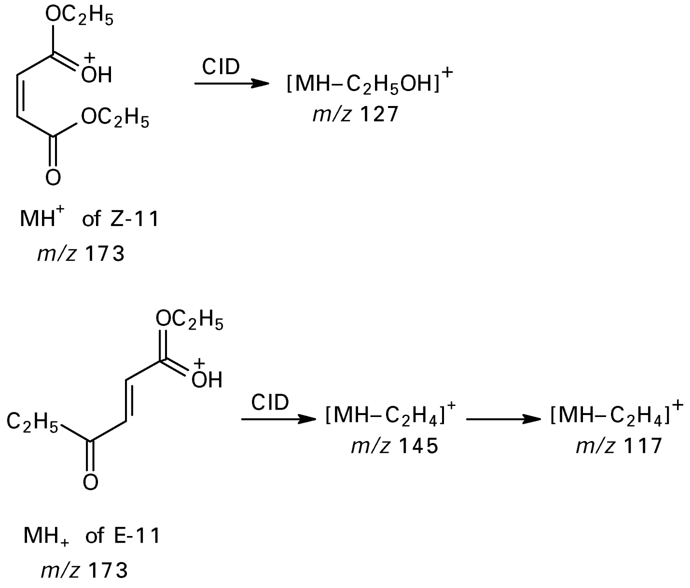

**Scheme 9**

**Alcohol elimination from MH+ ions diesters upon CI and CID** It has been previously mentioned that MH+ ions of diesters, with the two remote noninteracting alkoxycarbonyl groups, are stable and consequently highly abundant under isobutane-CI. On the other hand, stereoisomers with adjacent ester groups undergo a facile proton transfer between the two alkoxycarbonyls via proton bridged intermediates, resulting in the efficient elimination of ROH (**Scheme 5**). This behaviour enables easy distinction between stereiosomeric diesters by CI mass spectrometry, and also by CID measurements of MH+ ions. CID mass spectra of the *m/z* 173 MH+ ions of diethyl maleate Z-11 and fumarate E-11 exhibit an entirely different behaviour (**Scheme 10**): that of Z-11 shows an abundant *m/z* 127 [MH–EtOH]+ ion which is absent in the spectrum of E-11, while that of E-11 exhibits abundant *m/z* 145 $[MH–C_2H_4]^+$ and *m/z* 117 $[MH–2C_2H_4]^+$ ions, which do not appear in the CID spectrum of Z-11.

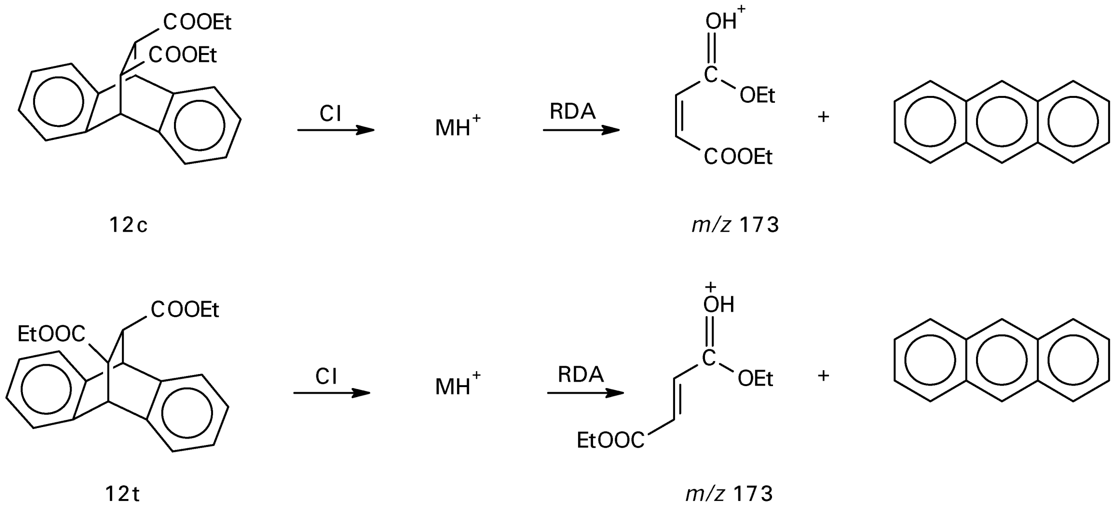
These distinctive features of the CID spectra also enable structural assignments and quantitative relative abundance estimates of protonated maleate and fumarate ions, which are formed by mass spectral fragmentation of higher systems. For example, this method was used to determine that the retro-Diels–Alder (RDA) fragmentation of *cis*- and *trans*-2,3-diethoxycarbonyl-5,6,7,8-dibenzobicyclo[2.2.2.]octanes 12 under i-$C_4H_{10}$-CI and $CH_4$-CI conditions is highly stereospecific, giving rise to protonated diethyl maleate and fumarate, respectively (see **Scheme 11**). This behaviour is consistent with a single-step concerted mechanism, analogous to the ground state RDA process occurring in neutral molecules in the condensed phase. On the other hand, analogous dissociation is nonstereospecific in *endo*-, *exo*- and *trans*-2,3-diethoxycarbonyl-5,6-benzobicyclo[2.2.2]-octanes 13, indicating involvement of a step-wise mechanism in this system (**Scheme 12**).

## Retro-Diels–Alder (RDA) fragmentation

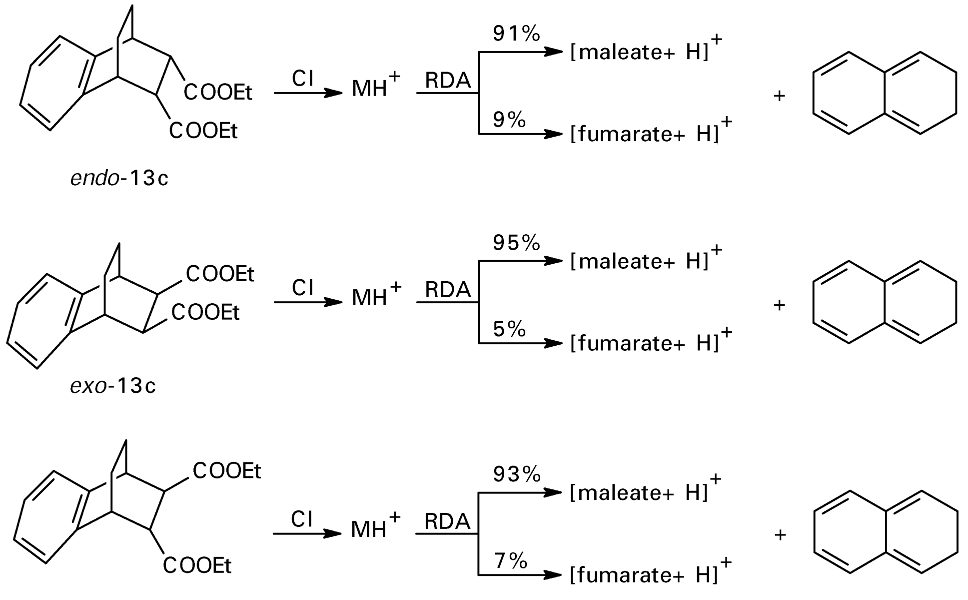
The retro-Diels–Alder (RDA) fragmentation is highly stereospecific in a variety of bi-, tri-, tetra- and pentacyclic systems. Two examples are shown in **Scheme 13**. The diene radical cations are the most abundant species observed in the EI mass spectra of the *cis* isomers 14c and 15c, and practically absent in the *trans* counterparts 14t and 15t. The very high efficiency of this process in the *cis* isomers and the high stability of the molecular ions of the *trans* analogues suggest that the RDA dissociation of gas-phase ions takes place by a concerted mechanism, which exhibits symmetry conservation characteristics similar to those of the ground-state RDA process in neutral species.




**Scheme 10**




**Scheme 11**

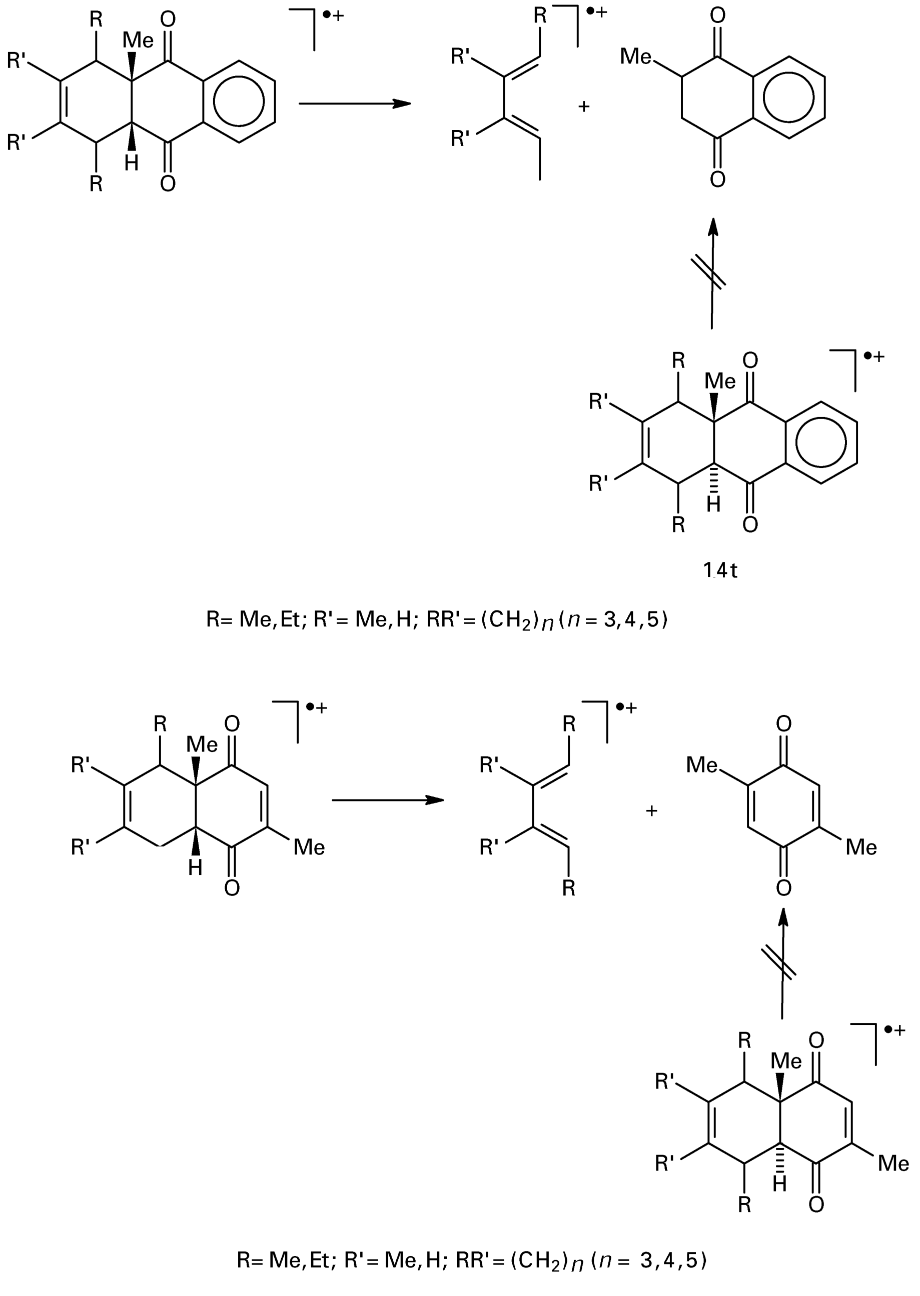



**Scheme 12**

Similar behaviour has been also observed upon CI. Protonated dienes are formed only from *cis* isomers. Protonated dienophiles have been also observed in certain cases under CI, again only in *cis* isomers.

In contrast with the above behaviour, there have been reports on the occurrence of EI-induced RDA fragmentation in compounds with a *trans* junction of the cyclohexane and the adjacent rings (e.g. 16 and 17, **Scheme 14**). These results are indicative of a step-wise dissociation in these systems, in which the allylic bonds cleaved in the course of this process are either highly (as in 16) or lightly (as in 17) substituted. It has been proposed that the nonstereospecific step-wise behaviour of systems such as 16, results from the low energy requirement for the cleavage of the fully substituted allylic 9–14 bond, which presumably is the initial step of the step-wise RDA dissociation. In systems such as 17, the low substitution pattern of the bonds cleaved in the course of the RDA process may result in a relatively high energy of activation of the concerted pathway and in the consequent preference of the step-wise mechanism.

The partial stereospecificity of the RDA process in system 18 (**Scheme 15**) has been the subject of a thorough examination. Theoretical calculations and critical energy measurements led to the conclusion

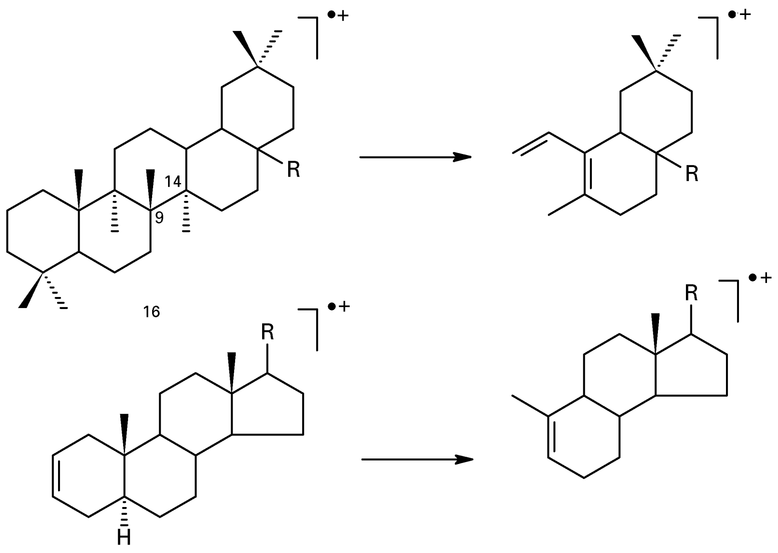



**Scheme 13**

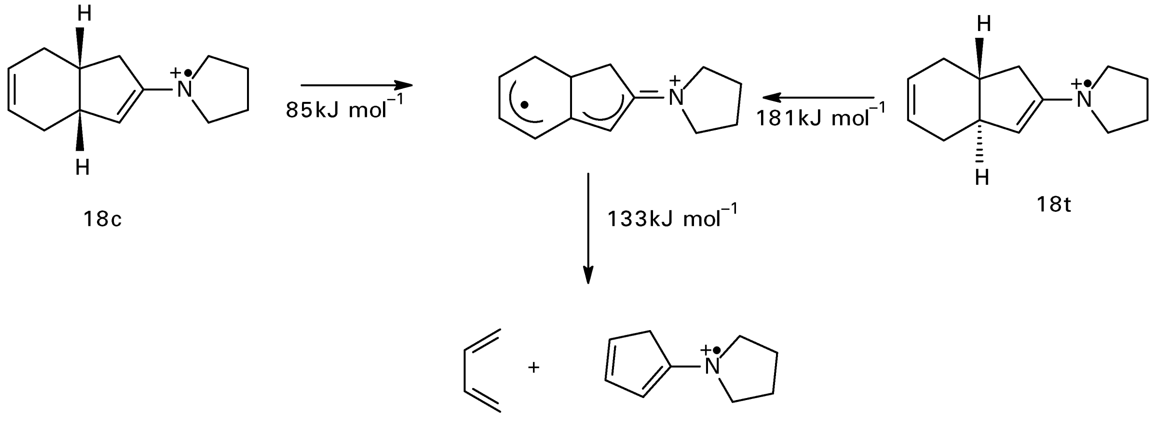
that the RDA fragmentation of both stereoisomers is a step-wise process involving a common distonic intermediate, but the energy barriers are different for the stereoisomers (**Scheme 15**).

### Anchimeric assistance

In many systems, variations in the abundances of certain fragment ions in the mass spectra of stereoisomers have been ascribed to anchimeric assistance. The loss of bromine from the molecular ions of stereoisomeric 1,2-dibromocyclopentanes and dibromocyclohexanes 19 (**Scheme 16**) is strongly affected by their configuration. The ion abundance ratio $[M–Br]^+/M^{\bullet+}$ is higher for the *trans* isomer by a factor of 10 in the dibromocyclopentane and 42 in the dibromocyclohexane system at 70 eV, and the difference between the stereoisomers increases at lower ionization energies. A similar effect has also been observed in acyclic diastereoisomeric vicinal dibromoalkanes. Anchimeric assistance of one bromine atom in the expulsion of the other from the molecular ion is proposed as the explanation of the stereospecificity of this process



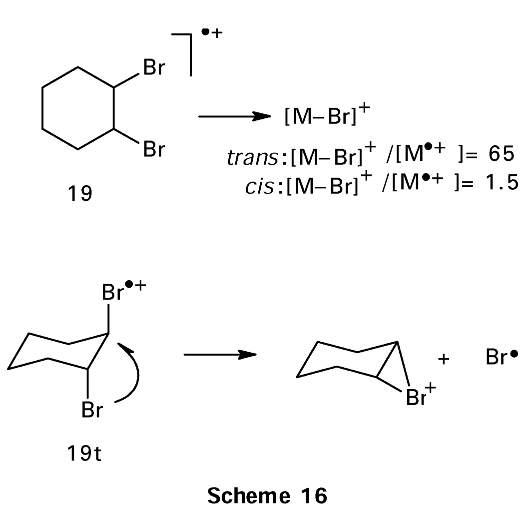

Scheme 14



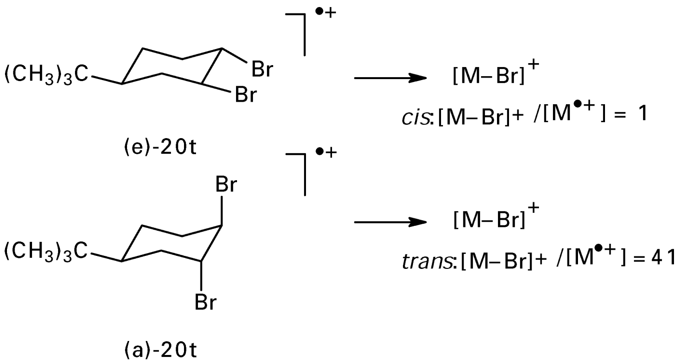

Scheme 15

(**Scheme 16**). This proposed mechanism finds support in the stereospecific behaviour of the two *trans*-1,2-dibromo-4-*t*-butylcyclo-hexanes (a)-20t and (e)-20t (**Scheme 17**). The stereoisomer (a)-20t, with the two axial Br atoms in the antiperiplanar conformation, gives rise to a much more abundant $[M–Br]^+$ ion than the diequatorial analogue (e)-20t.

The greater extent (by a factor of 2–10) of elimination of acetic acid from the $MH^+$ ions of *trans*-1,2- and 1,3-diacetoxycyclopentanes and -cyclohexanes upon $CH_4$-CI and isobutane-CI, as compared with the *cis* isomers, has been also interpreted in terms of anchimeric assistance. The carbonylic oxygen of the nonprotonated acetoxy group in the *trans* isomers (e.g. 21t in **Scheme 18**) assists in the elimination of




Scheme 16

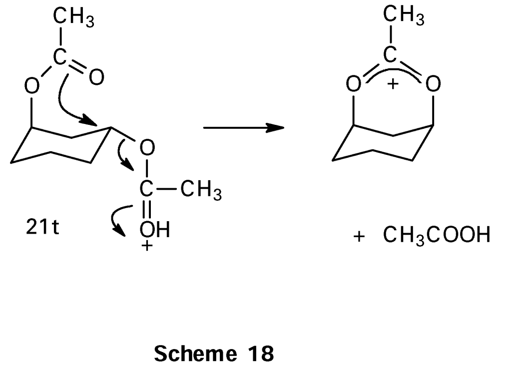



Scheme 17

$CH_3COOH$, affording the stabilized cyclic fragment-ion.

Numerous additional examples of a distinctive mass spectral behaviour of stereoisomers, which have been attributed to anchimeric assistance both under EI and CI conditions, have been proposed in the literature. In some cases, particularly upon CI, additional factors may be responsible for the distinctive behaviour. For instance, the higher abundance of the $[MH–CH_3COOH]^+$ ion in the CI mass spectra of the *trans*-diacetate 21t could also be attributed (at least in part) to the greater stability of the internally hydrogen bonded $MH^+$ ion of the *cis*-isomer.

Direct evidence for the operation of anchimeric assistance has been found in the CI behaviour of stereoisomeric 1,4-dialkoxycyclohexanes 22 (**Scheme 19**). The *trans*-diethers 22t afford very abundant $[MH–ROH]^+$ ions upon CI in contrast to the corresponding *cis* counterparts, suggesting anchimeric assistance in the elimination of alcohol from the $MH^+$ ions of the *trans*-diethers. Collision induced dissociation (CID) measurements of the $[MH–ROH]^+$ ions, obtained from various suitably deuterium labelled stereoisomeric 1-ethoxy-4-methoxy-cyclohexanes, indicated formation of symmetrical bicyclic ethyl and methyl oxonium ions by an anchimerically assisted alcohol elimination from the *trans*-diethers (the elimination of methanol is shown in **Scheme 19**). On the other hand, the CID measurements show that the *cis* isomers afford isomeric $[MH–ROH]^+$ ions, in which positions 2 and 3 (as well as 1 and 4, and 5 and 6) are not equivalent. These two results, namely the symmetrical structure and the high abundance of the $[MH–ROH]^+$ ions in the CI mass spectrum of the *trans*-diether 22t, in contrast to the non-symmetrical monocyclic structure and low abundance of these ions in the *cis* counterpart, are suggested as direct evidence for anchimeric assistance in the gas-phase ion dissociation process in that system. *Ab initio* calculations support the anchimerically assisted elimination observed in 22t. The energy difference between the anchimerically assisted and non-assisted elimination mechanisms in this system is not large (~ 2–3 kcal $mol^{-1}$).

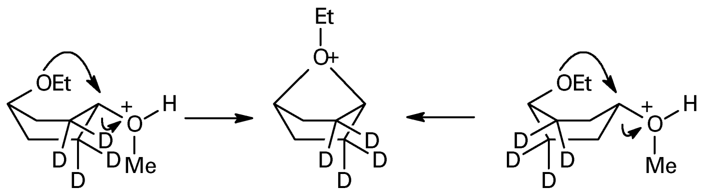



Scheme 18

### Stereoelectronic effects

One of the early reported cases of distinctive mass spectra of stereoisomers was the EI-induced behaviour of deacetylcyclindrocarpol 23a and of its epimer at C-19, 23b (**Scheme 20**). The more pronounced loss of the hydrogen atom from position 19 of the molecular ion of 23b, as compared with that of 23a (10.6% versus 1.7%), was attributed to the antiperiplanar relationship of the 19-C–19-H bond and the p-orbital of the adjacent nitrogen atom in 23b, in contrast to the epimer 23a.

The pronounced different behaviour of the stereoisomeric bicyclic carbamates 24a and 24b (**Scheme 21**) also indicates the occurrence of a stereoelectronic effect. The *trans* isomer 24a, with the axial methoxycarbonylmethyl group, affords the much more abundant $[M–CH_2COOCH_3]^+$ ion



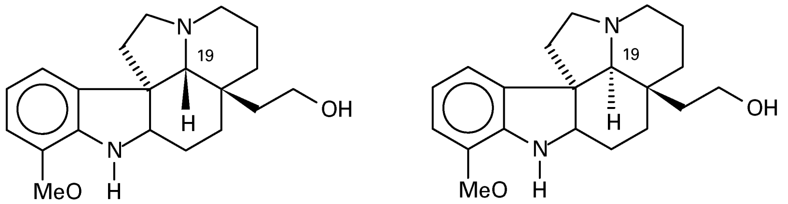

**Scheme 19**

possibly due to the stereoelectronic assistance of the π-orbital at the adjacent carbamate group. Similar stereoelectronic effects, resulting in distinctive bond dissociation processes of stereoisomers, have been observed in several other heterocyclic systems.

## Ion–molecule reactions

Ion–molecule reactions, studied using the ion cyclotron resonance (ICR) technique, show considerable stereospecificity in a number of systems, which enable distinction between stereoisomers. An early example of such a process is the gas-phase acetylation of *endo*- and *exo*-norborneol, shown in **Scheme 22**. The fourfold lower reactivity of the *endo*-isomer *endo*-25, as compared with the epimeric *exo*-25, is attributed to the steric hindrance in the approach of the bulky triacetyl cation to this stereoisomer.

Similar stereoselective behaviour has also been observed in the analogous gas-phase acetylation of *cis*- and *trans*-1-decalones and of the acetates of *cis*- and



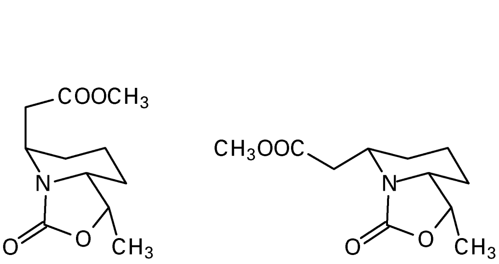

23a

[Ma–H]$^+$ 1.7%

23b

[M–H]$^+$ 10.6%

**Scheme 20**

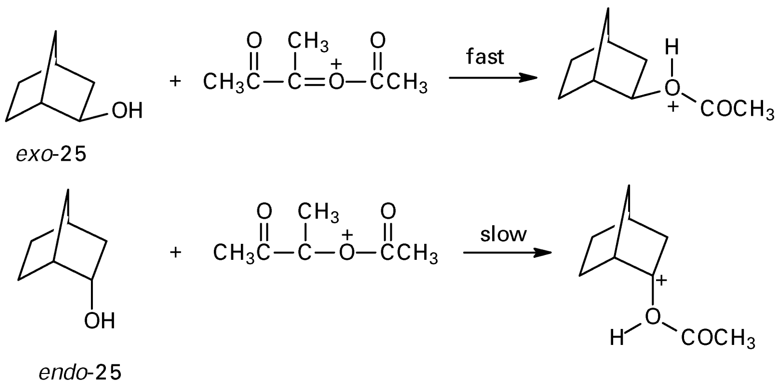



24a

24b

$\frac{[M-CH_2COOCH_3]^+}{M^{\bullet+}}$ 137 6.3

**Scheme 21**




Scheme 22

*trans*-4-*t*-butylcyclohexanols. The reaction is faster with the less hindered *trans* isomers in both cases.

## Enantiomers: chiral recognition

Unimolecular fragmentation processes studied by MS are of achiral nature, and consequently insensitive to chirality differences. A key procedure for chiral recognition using MS is to add a chiral component to the process and detect the possibly different diastereoisomeric interactions under a variety of conditions (CI, FAB, electrospray ionization (ESI), Fourier transform ion cyclotron resonance (FTICR), CID).

The formation and dissociation behaviour of the gas-phase protonated dimers (or higher clusters) of dialkyl tartarates under CI conditions is one of the early and most widely studied examples of chiral recognition by mass spectrometry. The protonated dimer consisting of two enantiomeric molecules and that consisting of two molecules of the same enantiomer are diastereoisomeric species, and as such they may exhibit distinctive behaviour, usually of a quantitative nature. Isotope labelling of one of the enantiomers is necessary in order to observe the different behaviour of the two protonated dimers.

Chemical ionization using chiral reagent gases (e.g. 1-amino-2-propanol or 2-amino-1-propanol) has been shown to induce distinctive behaviour between enantiomers. Enantiomers could also be differentiated via $S_N2$ reactivity with chiral reagents in ion–molecule reactions. The resulting diastereoisomeric products were distinguished by tandem mass spectral techniques (MS/MS).

Host–guest interactions using chiral hosts have been relatively widely investigated under a variety of mass spectral conditions as tools for chiral recognition in numerous systems. The use of chiral matrices in secondary ion mass spectral measurements (FAB ionization) is a promising way for distinguishing between enantiomers.

## Conformational studies of biopolymers

Electrospray ionization (ESI) has become a powerful method for the investigation of thermally labile polar biopolymers. The attachment of a large number of protons to the basic sites of proteins in the course of the ESI experiment affords a spectrum of multiply charged ions in the gas phase, which allows mass analysis using mass spectrometers with relatively low mass-to-charge ranges. With the steadily increasing use of this technique it has become apparent that more information may be obtained from the results of ESI analysis than just the molecular weight of the protein.

Careful measurements of ESI mass spectra of proteins show that changes in the solvent or in the pH of the examined solution may have an effect on the charge-state distributions. These different distributions have been interpreted as being the result of differences in the conformation of the proteins in the examined solutions. For example, ESI spectra of lysozyme, obtained from a 100% aqueous solution at pH 5.0 and from a solution containing 50% acetonitrile and 1% formic acid, show charge state maxima at 9+ and 12+, respectively. These different charge-state distributions are interpreted in terms of the degree of folding of the proteins in the original solutions. The protein molecules are unfolded to a greater extent in the organic than in the aqueous solution.

The unfolding of the molecules exposes sites or protonation, which were buried in the folded native

conformation, resulting in the higher charge states in the ESI mass spectrum obtained from the 50% acetonitrile–1 % formic acid solution.

Hydrogen/deuterium exchange measurements with the aid of ESI mass spectrometry have also been used successfully to explore conformational changes of proteins in solution. Elements of secondary structure that involve internal hydrogen bonding in the core of the protein are protected against hydrogen exchange, whereas regions exposed to the solvent undergo hydrogen exchange more readily.

*See also:* **Chemical Ionization in Mass Spectrometry; Fast Atom Bombardment Ionization in Mass Spectrometry; Fragmentation in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Ion Structures in Mass Spectrometry; MS–MS and MS$^n$; Peptides and Proteins Studied Using Mass Spectrometry; Proton Affinities.**

## Further reading

Green MM (1976) Mass spectrometry and the stereochemistry of organic molecules. *Topics in Stereochemistry* **9**: 35–110.

Harrison AG (1992) *Chemical Ionization Mass Spectrometry*, 2nd edn, pp 172–185. Boca Raton, Florida: CRC Press.

Mandelbaum A (1977) Application of mass spectrometry to stereochemical problems. In Kagan H (ed) *Stereochemistry*, vol. 1, pp 137–180. Stuttgart: Georg Thieme Publishing.

Mandelbaum A (1983) Stereochemical effects in mass spectrometry. *Mass Spectrometry Reviews* **2**: 223–284.

Meyerson S and Weitkamp AW (1968) Stereoisomeric effects on mass spectra. *Organic Mass Spectrometry* **1**: 659–668.

Robinson CV (1996) Protein secondary structure investigated by electrospray ionization. In: Chapman JR (ed.) *Protein and Peptide Analysis by Mass Spectrometry*, pp 129–139, Totowa, New Jersey: Humana Press.

Splitter J and Turecek F (1994) *Application of Mass Spectrometry to Stereochemical Problems*. New York: VCH.

Sawada M (1997) Chiral recognition detected by fast atom bombardment mass spectrometry. *Mass Spectrometry Reviews* **16**: 73–90.

Turecek F (1991) Stereoelectric effects in mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **108**: 137–164.

Turecek F (1987) Stereochemistry of organic ions in the gas phase. *Collection Czech. Chemical Communications* **52**: 1928–1984.

# Stray Magnetic Fields, Use of in MRI

*See* **MRI Using Stray Fields.**

# Structural Chemistry Using NMR Spectroscopy, Inorganic Molecules

**GE Hawkes**, Queen Mary and Westfield College, London, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

A very wide range of experiments is available to the NMR spectroscopist for inorganic structural analysis. The full armoury of single- and multi-dimensional NMR techniques used for the structural analysis of organic compounds may be used for inorganic systems, with the added dimension of the multinuclear approach. X-ray structural analysis of crystalline compounds has long been fundamental to inorganic structure determination and it is now clear that solid-state NMR, particularly with the magic angle spinning (MAS) technique, is capable of providing complementary data making the synergistic combination of X-ray and MAS-NMR methods very powerful indeed. In reviewing applications of NMR methods to inorganics, it is important to consider the insight that NMR provides into dynamic aspects of the molecular structures; intermolecular exchange processes in solution, and intramolecular fluxionality in both solution and solid state can be delineated, and very often thermodynamic and kinetic parameters derived which can be interpreted in mechanistic terms. It is often the case, particularly for transition metal compounds, that the molecule is paramagnetic. While the unpaired electron(s) can cause undesirable effects on the NMR spectra, e.g. excessive broadening of resonances, it may often be the case that the paramagnetic *shift* at ligand resonances may be useful in resolving erstwhile overlapping signals, or in providing information on the distribution of unpaired spin density throughout the molecule. In addition to the intrinsic information from such paramagnetic molecules, the addition of a paramagnetic compound (e.g. a lanthanide chelate) to a solution of a diamagnetic compound can induce paramagnetic shifts and changes in line widths in resonances of the latter compound (shift and relaxation agents). These changes can be valuable in providing structural information or in enhancing contrast in magnetic resonance images.

## The multinuclear approach

In any structural problem, the first step is to decide which NMR experiments will provide the most definitive information in the shortest time. Usually, although not exclusively, solution-state spectra are measured, and if the inorganic compound includes an organic ligand, then $^{1}$H and $^{13}$C spectra are essential. The choice of additional spectra from other nuclei must be directed using the same criteria – information content and economy of time. In this it is necessary to know the inherent sensitivity of the isotope, and the value of its spin quantum number ($I$). Isotopes with $I = 1/2$ generally give rise to narrower resonances ('high-resolution' nuclei) whereas those with $I > 1/2$ often yield much broader resonances (quadrupolar nuclei) and spectral information (chemical shifts and scalar couplings) may be obscured. Some NMR-active isotopes of potential interest to the inorganic chemist are listed in **Table 1**, and this list is by no means exhaustive since there are considerably more than 100 NMR-active stable isotopes across the periodic table.

Careful consideration should be given to the ease of observation of the NMR spectrum of a given isotope. Generally, the ease of observation will increase with increase in the magnetic field strength of the spectrometer, but it is possible that at the higher end of the available magnetic field strength range (≥ 14 T), the resonances of certain nuclei, particularly the heavier metals such as platinum, mercury or lead, may become broadened by the influence of chemical shift anisotropy (see below), thereby abrogating the beneficial effects of the higher field. Another consideration is the combination of low natural abundance and small nuclear magnetic moment (low resonance frequency) giving a low intrinsic receptivity, as for $^{57}$Fe. In such cases the situation may be partially alleviated by resorting to isotopic enrichment, as illustrated in **Figure 1**, which shows two $^{57}$Fe resonances from a 20 mM solution of the superstructured haem model compound $^{57}$Fe$^{II}$PocPiv(1,2-diMeIm) (CO) [1] (94.5% enriched in $^{57}$Fe). Even at this high level of enrichment and high magnetic field, the spectrum still required about 20 h of instrument time for the spectral accumulation. The spectrum does illustrate the extreme sensitivity of the $^{57}$Fe chemical shift to fairly

**Table 1** NMR properties of some isotopes for the inorganic chemist

| *Isotope* | *I* | *Abundance*[a] | *Receptivity*[b] | *Q* ($10^{-28}$ m$^2$)[c] | *Ξ* (MHz)[d] | *Reference*[e] |
|---|---|---|---|---|---|---|
| $^{1}$H | 1/2 | 99.985 | $5.7\times10^{3}$ | | 100.0 | $SiMe_4$ |
| $^{2}$H | 1 | 0.015 | $8.2\times10^{-4}$ | $2.7\times10^{-3}$ | 15.4 | |
| $^{6}$Li | 1 | 7.4 | 3.6 | $-8\times10^{-4}$ | 14.7 | $Li^+$aq |
| $^{7}$Li | 3/2 | 92.6 | $1.5\times10^{3}$ | $-4.5\times10^{-2}$ | 38.9 | $Li^+$aq |
| $^{11}$B | 3/2 | 80.4 | $7.5\times10^{2}$ | $3.6\times10^{-2}$ | 32.1 | $BF_3{\cdot}Et_2O$ |
| $^{13}$C | 1/2 | 1.1 | 1.00 | | 25.1 | $SiMe_4$ |
| $^{17}$O | 5/2 | 0.037 | $6.1\times10^{-2}$ | $-2.6\times10^{-2}$ | 13.6 | $H_2O$ |
| $^{19}$F | 1/2 | 100 | $4.7\times10^{3}$ | | 94.1 | $CCl_3F$ |
| $^{23}$Na | 3/2 | 100 | $5.3\times10^{3}$ | 0.12 | 26.5 | $Na^+$aq |
| $^{27}$Al | 5/2 | 100 | $1.2\times10^{3}$ | 0.15 | 26.1 | $Al(H_2O)_6^{+}$ |
| $^{29}$Si | 1/2 | 4.7 | 2.1 | | 19.9 | $SiMe_4$ |
| $^{31}$P | 1/2 | 100 | $3.8\times10^{2}$ | | 40.5 | 85% $H_3PO_4$ |
| $^{51}$V | 7/2 | 99.8 | $2.2\times10^{3}$ | 0.3 | 26.3 | $VOCl_3$ |
| $^{57}$Fe | 1/2 | 2.2 | $4.2\times10^{-3}$ | | 3.2 | $Fe(CO)_5$ |
| $^{59}$Co | 7/2 | 100 | $1.6\times10^{3}$ | 0.4 | 23.6 | $Co(CN)_6^{3-}$ |
| $^{71}$Ga | 3/2 | 39.6 | $3.2\times10^{2}$ | 0.11 | 30.5 | $Ga(H_20)_6^{3+}$ |
| $^{103}$Rh | 1/2 | 100 | 0.18 | | 3.2 | |
| $^{109}$Ag | 1/2 | 38.2 | 0.28 | | 4.7 | $Ag^+$aq |
| $^{113}$Cd | 1/2 | 12.3 | 7.6 | | 22.2 | $CdMe_2$ |
| $^{119}$Sn | 1/2 | 8.6 | 25.2 | | 37.3 | $SnMe_4$ |
| $^{139}$La | 7/2 | 99.9 | $3.4\times10^{2}$ | 0.2 | 14.1 | $La^{3+}$aq |
| $^{183}$W | 1/2 | 14.4 | $5.9\times10^{-2}$ | | 4.2 | $WO_4^{2-}$aq |
| $^{195}$Pt | 1/2 | 33.8 | 19.1 | | 21.4 | $Pt(CN)_6^{2-}$ |
| $^{199}$Hg | 1/2 | 16.8 | 5.4 | | 17.9 | $HgMe_2$ |
| $^{207}$Pb | 1/2 | 22.6 | 11.8 | | 20.9 | $PbMe_4$ |

[a] The natural abundance of the isotope.
[b] A rough guide to the ease of observation of the NMR spectrum, relative to $^{13}$C.
[c] Approximate values for the quadrupole moment.
[d] The resonance frequency in a magnetic field strength that gives the $^{1}$H resonance of $SiMe_4$ at exactly 100 MHz.
[e] Commonly accepted chemical shift reference standard material.

remote structural effects since the two $^{57}$Fe resonances are believed to arise from the presence of the two atropisomers (α and β) due to restricted rotation of the pivaloylamido picket. In considering NMR spectra of the quadrupolar nuclei, in addition to the question of sensitivity, there is the question of resolution of chemical shifts, since chemical shift differences within a spectrum may be obscured by relatively large line widths ($W_Q$) exhibited by the resonances due to quadrupolar relaxation. As a very rough guide to the line width problem, the line width may increase with the square of the quadrupole moment, hence less broadening is expected in spectra of $^{2}$H or $^{6}$Li:

$$W_Q \propto \left(\frac{e^2 q_{zz} Q}{h}\right)^2 \left(1 + \frac{1}{3}\eta^2\right) \qquad [1]$$

where $e$ is the electronic charge, $h$ Planck's constant, $q_{zz}$ the largest component of the electric field gradient and $\eta$ the asymmetry parameter for $q$. However, since the electric field gradient at the nucleus, caused by the surrounding electron distribution, is also important, this may counter the effect of a larger quadrupole moment as for the $^{51}$V spectrum of the product of partial hydrolysis of $VO(NO_3)_3$ (**Figure 2**, which also includes the $^{17}$O spectrum). $^{17}$O enrichment can often be achieved starting with the relatively inexpensive source $H_2{}^{17}O$, and the relatively narrow lines often exhibited by $^{17}$O resonances can provide a wealth of structural data as shown in **Figure 3** for the aqueous isopolytungstate solution (enriched to 5% $^{17}$O), where the $^{17}$O chemical shifts are sensitive to a variety of structural features, particularly the metal–oxygen bond lengths.

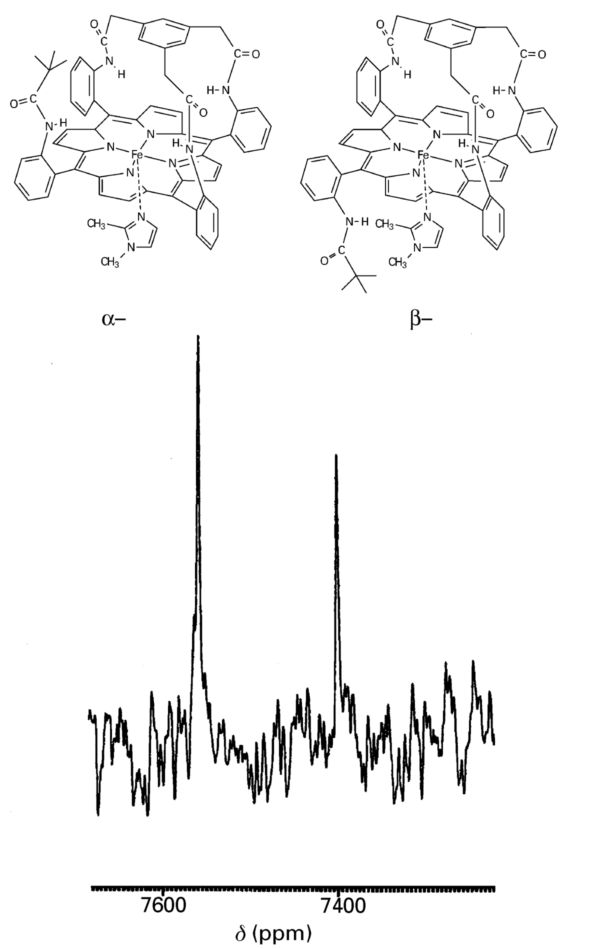



**Figure 1** 19.58 MHz $^{57}$Fe NMR spectrum of the $^{57}Fe^{II}$ PocPiv(1,2-diMeIm)(CO) adduct in $CD_2Cl_2$ solution at 298 K. Reproduced with permission from Gerothanassis IP, Kalodimos CG, Hawkes GE and Haycock PR, *Journal of Magnetic Resonance* 1998, **131**: 163–165.

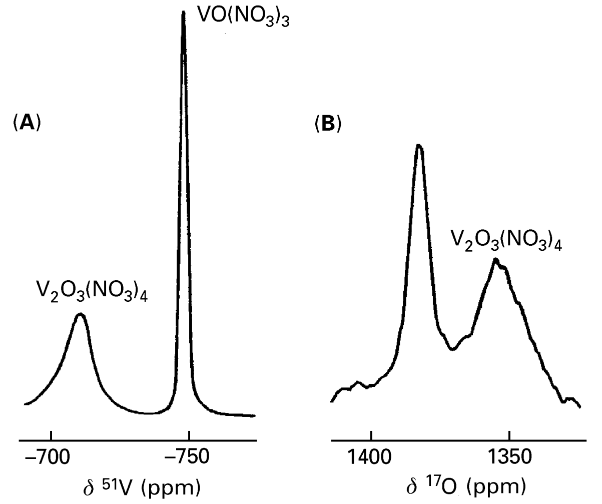



**Figure 2** NMR spectra at 294 K of the $VO(NO_3)_3$ – $H_2O$ (mole ratio 1 : 0.3) system in $MeNO_2$: (A) 105.1 MHz $^{51}$V spectrum and (B) 54.2 MHz $^{17}$O spectrum. Reproduced with permission from Hibbert RC, Logan N and Howarth AW, *Journal of the Chemical Society, Dalton Transactions* 1986, 369–372.

## The NMR parameters

### Chemical shifts

For samples of inorganic molecules in solution, each chemically distinct site for an atom in the molecule will result in a distinct *isotropic* chemical shift for its nucleus in the NMR spectrum. What is important from the structural point of view is that these chemical shifts are resolved in the spectrum, and this in turn will be determined by the sensitivity of the chemical shift to structural changes. Some nuclei are more sensitive than others, and this is usually represented by the reported chemical shift range of the nucleus. $^1$H chemical shifts in inorganic compounds typically span a range ~20 ppm, whereas heavier isotopes often exhibit much greater chemical shift ranges, e.g. hundreds of ppm for $^{17}$O, $^{19}$F, $^{29}$Si and $^{31}$P, and this may run to thousands of ppm for heavy metals such as $^{195}$Pt or $^{199}$Hg. Each isotropic chemical shift is in fact an average of three principal chemical shift values. The chemical shift is determined by the interaction of the electron distribution with the spectrometer magnetic field, hence a nucleus in an asymmetric electronic environment (the general case) will experience a change in chemical shift with the orientation of the molecule in the magnetic field. The chemical shift is thus represented as a second-rank tensor ($3 \times 3$ values) and it is possible to find a molecule-fixed Cartesian coordinate system which diagonalizes this tensor to give the three principal components ($\delta_{11}$, $\delta_{22}$, $\delta_{33}$). In any solution sample the molecules undergo rapid random motion, including rotation, and as a result these components average to a single isotropic chemical shift ($\delta_{iso}$):

$$\delta_{iso} = (\delta_{11} + \delta_{22} + \delta_{33})/3 \qquad [2]$$

Although it is not possible to obtain values for the independent components from solution-state spectra, it is possible in certain cases to obtain some information such as the chemical shift anisotropy $\Delta\delta$ (see below):

$$\Delta\delta = \delta_{33} - (\delta_{11} + \delta_{22})/2 \qquad [3]$$

where $\delta_{33}$ is the component furthest removed from the average $\delta_{iso}$.

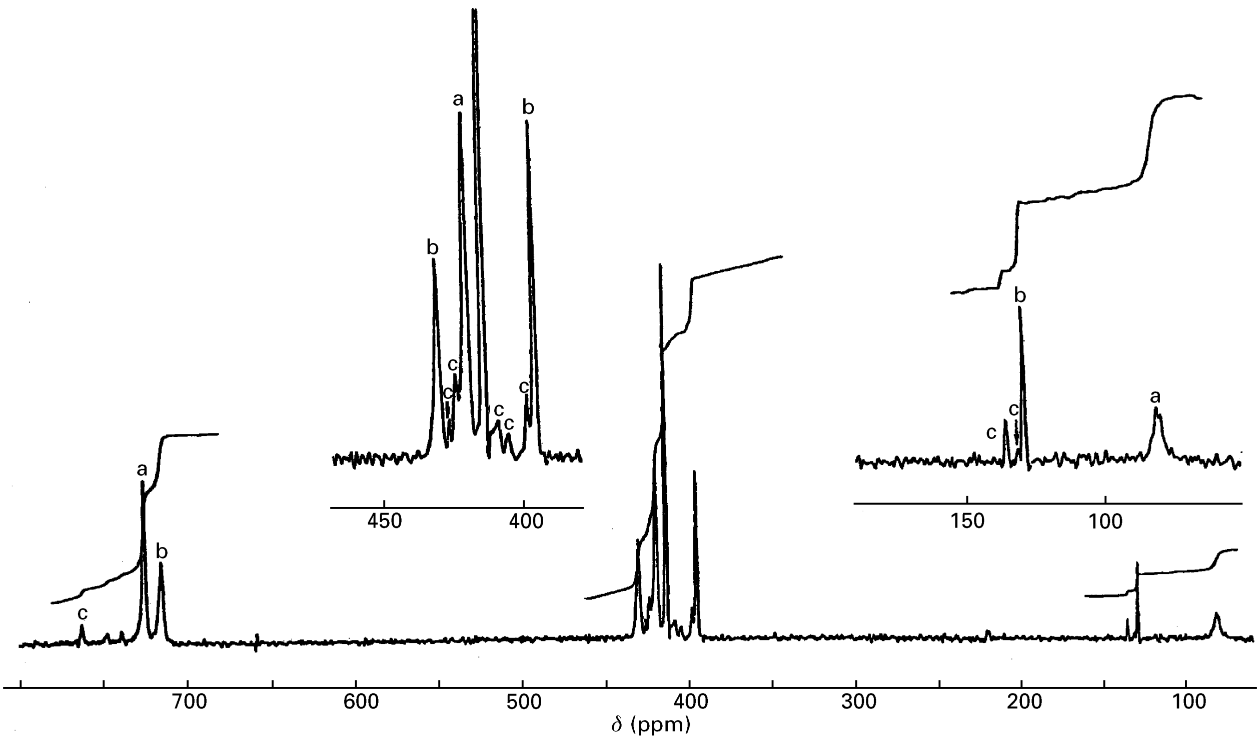



**Figure 3** 54.2 MHz $^{17}O$ NMR spectrum of isopolytungstate at 353 K, pH 1.1. Species: a, α-$[H_2W_{12}O_{40}]^{6-}$, metatungstate; b, α-$[HW_{12}O_{40}]^{7-}$; c, $\psi'$-metatungstate; probably $\beta$-$[HW_{12}O_{40}]^{7-}$. Reproduced with permission from Hastings JJ and Howarth OW, *Journal of the Chemical Society, Dalton Transactions* 1992, 209–215.

MAS-NMR spectra of the solid state can provide values for the components $\delta_{ii}$ as shown in **Figure 4**. The *static* spectrum (**Figure 4A**) of the powder sample is the superposition of different chemical shifts resulting from all possible orientations of the molecules in the discrete particles of the sample, and the components $\delta_{ii}$ are as indicated. Usually such broad lines would mask any resolution of distinct chemical shifts. The MAS spectra (**Figures 4B–D**) consist of a centre-band resonance at the isotropic chemical shift ($\delta_{iso}$) and side bands spaced at the rotation frequency, and offer two advantages over static spectra. The first advantage is that within the centre band there may be chemical shift resolution (in this case there is only one $^{31}P$ environment and therefore only one isotropic chemical shift) and the second is that the total integrated intensity of the spectrum is within the relatively sharp lines and so the sensitivity of the observation is greatly enhanced in comparison with that of the static spectrum. The intensity pattern of the spinning side bands roughly follows the static spectrum, and these intensities can be used to obtain values for the components $\delta_{ii}$. To illustrate the utility of such measurements, it has been shown that the $^{31}P$ isotropic shifts and chemical shift tensor components ($\delta_{ii}$) from a series of phosphido-bridged iron complexes $Fe_2(CO)_6(\mu\text{-}X)(\mu\text{-}PPh_2)$ gave an excellent correlation with the crystallographically determined Fe–P–Fe bond angles. In a related study on iron complexes with asymmetrical bridging carbonyl ligands Fe···CO–Fe the $^{13}C$ chemical shift anisotropy and the component $\delta_{33}$ (associated with the C–O bond axis) both correlated with the difference in the two Fe–C distances. The isotropic $^{13}C$ chemical shift is not a reliable indicator of the metal–carbonyl group bonding, and typically cannot be used to distinguish unequivocally between terminal and bridging carbonyl groups (e.g. M–CO *vs* M–CO–M). However, as shown in **Table 2**, the chemical shift anisotropy values are distinctive of the carbonyl group bonding.

**Table 2** $^{13}C$ chemical shift anisotropy values for terminal and bridging carbonyl groups

| | Δδ (ppm) | | |
|---|---|---|---|
| | *Co terminal* | *$\mu_2$-CO* | *$\mu_3$-CO* |
| $(C_5H_5)_2Fe_2(CO)_4$ | 444 | 138 | – |
| $Rh_6(CO)_{16}$ | 390 | – | 194 |

Data reproduced with permission from Gleeson JW and Vaughan RW, *Journal of Chemical Physics* 1983, **78**: 5384–5392.

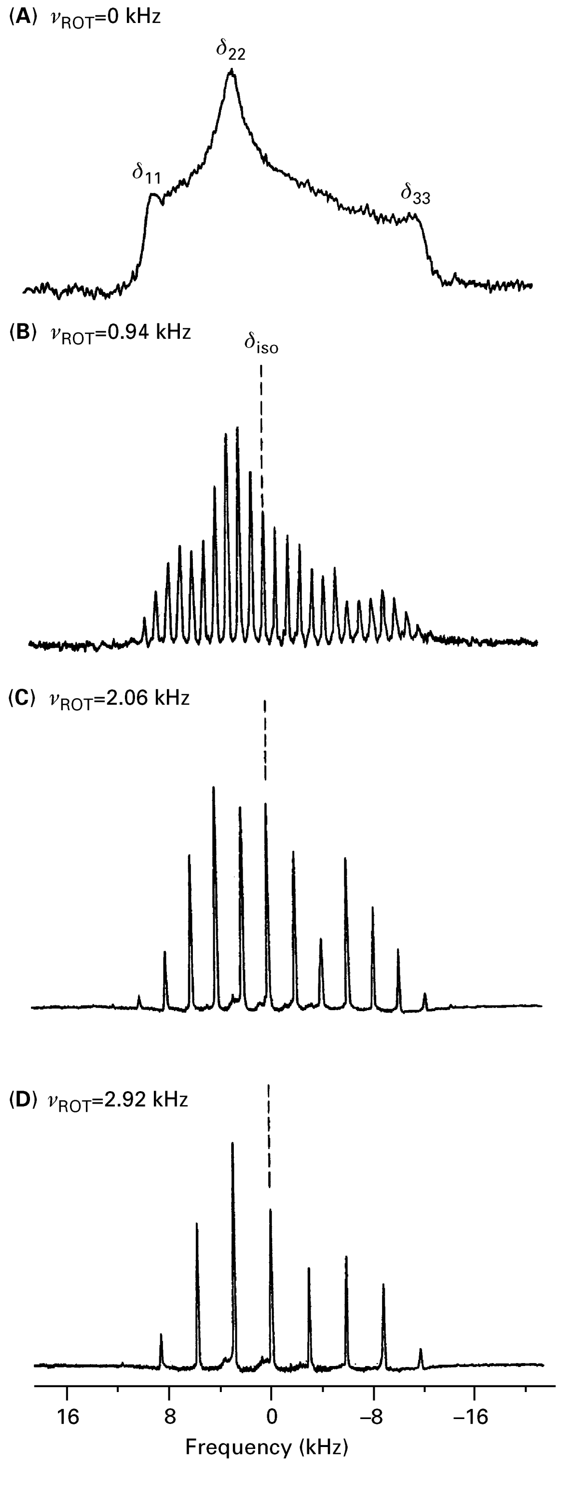



**Figure 4** 119.05 MHz solid-state $^{31}P$ NMR spectrum of diethyl phosphate spinning at the magic angle. $\nu_{ROT}$ indicates the spinning frequency. Reproduced with permission from Herzfeld J and Berger AE, *Journal of Chemical Physics* 1980, **73**: 6021–6030.

## Coupling constants

The two important parameters to consider are the internuclear scalar couplings and the internuclear dipolar couplings. Scalar or spin–spin coupling constants are observed in both solution- and solid-state spectra and are usually considered to be transmitted *via* bonding electrons. The magnitude of the scalar couplings varies dramatically with the isotopes concerned, the nature and number of intervening bonds, coordination number, oxidation state, etc. Interproton couplings are typically small (< 20 Hz) but couplings between heavier isotopes may be fairly large, up to about 10 000 Hz, for example, for the one-bond $^{31}P$–$^{199}Hg$ coupling in $Ph_3PHgX_2$, X = $OCOCH_3$ or $OCOCF_3$. The splitting patterns induced by the scalar couplings are used to determine the number of interacting nuclei and the magnitudes of the couplings may be interpreted in terms of bonding and conformation. Internuclear dipolar couplings are direct through space interactions between magnetic nuclei and values may be as large as 50 000 Hz between protons, being dependent upon the inverse third power of the separation ($r^{-3}$). The splitting caused by this effect is dependent on the orientation of the internuclear vector in the spectrometer magnetic field and for molecules in solution undergoing rapid molecular tumbling the effect averages to zero. Therefore, in solution spectra there is no obvious direct effect on one-dimensional spectra due to dipolar couplings. However, they are responsible for indirect effects, such as nuclear relaxation and the nuclear Overhauser effect (see below). In solid-state NMR, dipolar interactions provide one mechanism for line broadening, particularly with hydrogen present in the molecule. MAS (e.g. using 4 mm o.d. rotors at 15 kHz) is used to reduce the dipolar interactions and for the observation of nuclei other than hydrogen this is often used in conjunction with high-power $^1H$ decoupling. For observation of solid-state $^1H$ spectra it is often necessary to use the combination of MAS with a suitable pulse sequence (CRAMPS; combined rotation and multiple pulse spectroscopy). If the dipolar interactions are relatively weak then MAS alone may be sufficient to allow resolution of chemical shifts, as shown in **Figure 5** for some metallo-hydride complexes. While much attention has been focused on methods for reducing or eliminating the effects of the dipolar interactions in solid-state NMR, the presence of the dipolar interaction could be useful in showing the spatial proximity of atoms in a structure. Several multiple-pulse two-dimensional experiments have been proposed, including trains of pulses, synchronized in time with the MAS rotation period. These sequences refocus the dipolar interaction, as illustrated in **Figure 6** for the $^{31}P$ double-quantum–single-quantum correlation spectrum of a polycrystalline powder sample of $Cd_3(PO_4)_2$. There are six crystallographically distinct

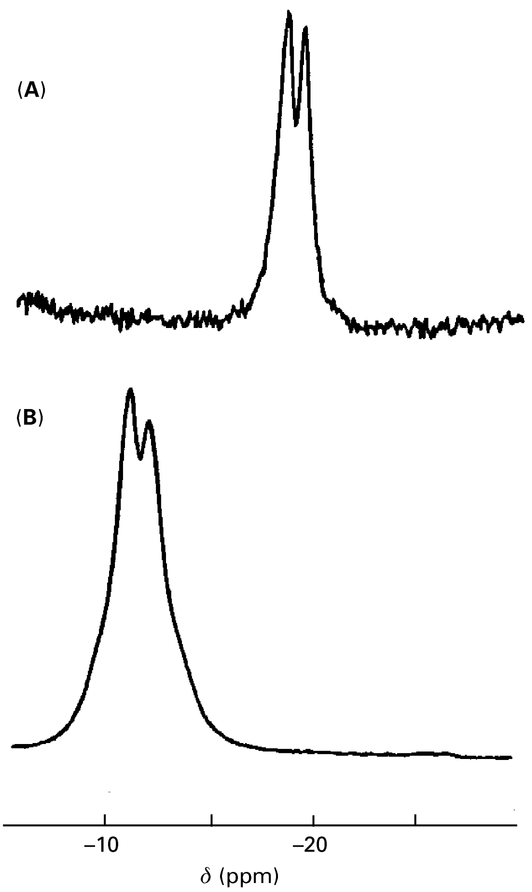



**Figure 5** 300 MHz solid-state $^1H$ MAS-NMR spectra: (A) $H_2Os_3(CO)_{10}$, MAS rate 8.1 kHz; (B) $H_2FeRu_3(CO)_{13}$, MAS rate 9.5 kHz. Reproduced with permission from Aime S, Barrie PJ, Brougham DF, Gobetto R and Hawkes GE, *Inorganic Chemistry* 1995, **34**: 3557–3559.

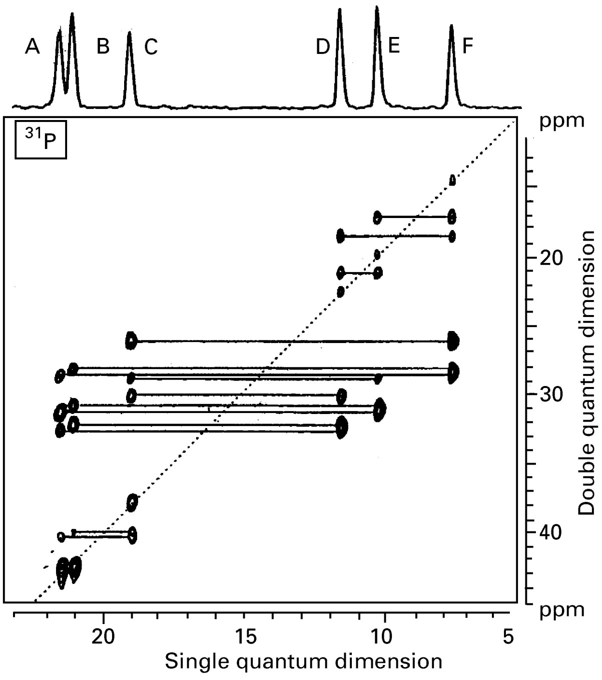



**Figure 6** 202.5 MHz solid-state $^{31}P$ MAS-NMR spectrum of $Cd_3(PO_4)_2$. The two-dimensional spectrum shows the single-quantum–double-quantum dipolar correlations Reproduced with permission from Dollase WA, Fecke M, Förster H, Schaller T, Schell I, Sebald A and Stevernagel S, *Journal of the American Chemical Society* 1997, **119**: 3807–3810.

phosphorus sites in the structure and six resolved $^{31}P$ resonances in the one-dimensional spectrum. The contours link pairs of phosphorus sites which have a measurable dipolar interaction (are in spatial proximity) and the more intense correlations indicate greater proximity in the structure.

### Relaxation times

The principal relaxation times measurable for resolved resonances from solution state samples are the spin–lattice relaxation time ($T_1$) and the spin–spin relaxation time ($T_2$). Spin echo methods may be used to measure $T_2$ values and these are often useful in defining chemical exchange rate processes. $T_1$ values are readily obtained from the inversion–recovery experiment and can be directly used to provide structural information. For nuclei with spin $I = 1/2$ in diamagnetic molecules in solution there are two principal mechanisms which contribute to the rate of the relaxation ($T_1^{-1}$), namely the dipole–dipole interaction and the chemical shift anisotropy mechanism. The dipole–dipole interaction occurs between magnetic nuclei which are in close spatial proximity in the molecule and if the population distribution of nuclei across the energy levels of one site is disturbed away from its equilibrium value (usually the populations are equalized by a second radiofrequency field) then this is reflected as a change in the intensity of the resonance from the other site. This is the so-called nuclear Overhauser effect (NOE) and is widely used as a structural tool; particularly the interproton NOE is used in a qualitative or quantitative manner to estimate distances between protons in biomolecules, and thereby serve to define conformation. Such interproton NOEs will similarly be useful for determination of structure in organometallic species, and in addition the structural inorganic chemist will be able to utilize the homonuclear NOE between other nuclei which are present at a high level of abundance, for example the $^{31}P$–$^{31}P$ NOE might be particularly useful. The heteronuclear NOE might also be expected to provide useful conformational information where sensitivity permits, and $^{13}C$–$^{1}H$ and $^{31}P$–$^{1}H$ are obvious candidates, and $^{6}Li$–$^{1}H$ has also been used. The relaxation rate of a nucleus due to its dipole–dipole interaction ($T_1^{-1}$) with a neighbouring nucleus at a distance $r$ is proportional to $r^{-6}$.

This distance dependence has been put to a number of uses, in particular in structural studies of molecular hydrogen complexes. If hydrogen is bound as two distinct M–H groups then the separation between the hydrogens will be greater than if the hydrogen is bound as molecular hydrogen M–($H_2$). Therefore, for the molecular hydrogen case the hydrogens will experience a stronger mutual dipole–dipole interaction and the rate of spin–lattice relaxation will be greater. Ideally the experiment should be calibrated by measuring the relaxation rate for a pair of nuclei at a known separation in the same molecule, perhaps for nuclei in an organic ligand.

The rate of nuclear spin–lattice relaxation due to the chemical shift anisotropy mechanism $(T_1^{CSA})^{-1}$ depends upon the magnitude of the chemical shift anisotropy ($\Delta\delta$) and the strength of the spectrometer magnetic field ($B_0$) squared. This dependence of the rate on $B_0^2$ provides a means of estimating a value for $\Delta\delta$ if the mechanism is important (significant value for $\Delta\delta$) and if the measurement of the rate of relaxation can be made at several different field strengths ($B_0$). This may be applied to $^{13}$C relaxation for the carbonyl groups of organometallic complexes (see **Table 2**). A related study on metal carbonyl complexes made use of the quadrupolar relaxation rate shown by those nuclei with spin $I > 1/2$, here $^{17}$O with $I = 5/2$. The relaxation rate of such quadrupolar nuclei is dominated by the contribution, $(T_1^{Q})^{-1}$, from the quadrupolar mechanism, and in favourable cases it is possible to determine the quadrupole coupling constant (QCC; *cf.* Eqn [1]). **Table 3** shows values for the $^{17}$O QCC for the carbonyl groups of some metallo-carbonyl complexes, and again the derived parameter is seen to be diagnostic of the type of carbonyl group.

## Dynamic processes

NMR spectroscopy is a most powerful method for the investigation of dynamic processes occurring at the molecular level. In particular, both intramolecular and intermolecular chemical exchange processes in solution may be investigated and for inorganic compounds these processes include ligand exchange, conformational changes, rearrangements and fluxional processes. There are various NMR parameters which may be used to monitor the dynamic process, and the particular set of NMR experiments to be used may depend in part on the order of magnitude for the rate coefficient of the process. The rate process may be termed either 'slow' or 'fast', but these labels really depend on the NMR parameter being used to monitor the process. For many years the most common method to study rate processes both qualitatively and quantitatively has been the band shape method. Here the NMR parameters may be chemical shifts (measured in frequency units) and/or coupling constants. For a two-site exchange process A ↔ X the spectrum will be affected when the rate is within the limits

$$T_2 \geq k_r^{-1} \geq \left(\tfrac{1}{2}\pi^2\Delta\nu^2 T_2\right)^{-1} \quad [4]$$

where $\Delta\nu$ is the difference in resonance frequency (between sites A and X) or the coupling constant being averaged. Exchange processes in the fast exchange limit ($k_r \geq 10^4$ Hz) may contribute to the rotating frame relaxation time ($T_{1\rho}$), and measurement of $T_{1\rho}$ as a function of the strength of the spin-lock field can give a value for the rate coefficient. More recently, magnetization transfer experiments, both one- and two-dimensional, have been used to explore multi-site slow exchange situations. In these experiments the population distribution between the nuclear energy levels for one or more of the sites is disturbed from the equilibrium. This can be by equalization of the populations (cf. saturation as described above for the NOE) or by inversion of the populations by a selective 180° radiofrequency pulse for the 1D experiment or a non-selective pulse for the 2D experiment. The chemical exchange can then transmit the disturbance throughout the exchanging system. However, since spin–lattice relaxation is always occurring in order to restore the equilibrium nuclear distribution, then this method is applicable when the rate coefficient $k_r \geq T_1^{-1}$. In the two-dimensional experiment, which is exactly the same as the 2D NOESY experiment, the advantage is that it is possible to obtain a very clear picture of the magnetization transfer pathways in addition to being able to quantify the rate coefficients. This is illustrated in **Figure 7**, where the off-diagonal contours link chemical shifts of pairs of slowly exchanging methyl groups among the six distinct methyls of the 2,4,6-tris(3,5-dimethylpyrazol-1-yl)pyrimidine (tdmpzp) ligands.

**Table 3** $^{17}$O quadrupole coupling constant values for terminal and bridging carbonyl groups

| | *QCC* (MHz) | | |
|---|---|---|---|
| | *Co terminal* | $\mu_2$-*CO* | $\mu_3$-*CO* |
| $(C_5H_5)_2Fe_2(CO)_4$ | 1.47 | 3.3 | – |
| $Rh_6(CO)_{16}$ | 2.02 | – | 0.09 |

Data reproduced with permission from Hawkes GE and Randall EW, *Journal of Magnetic Resonance* 1986, **68**: 597–599.

## Sensitivity enhancement by polarization transfer

### One-dimensional experiments

The solution-state polarization transfer experiments described here all depend upon the existence of a resolved scalar coupling between a sensitive nucleus (e.g. $^{1}H$, $^{31}P$) and an insensitive nucleus (e.g. $^{57}Fe$, $^{109}Ag$). The one-dimensional experiments are based upon the so-called INEPT or DEPT pulse sequences and involve the initial creation of anti-phase magnetization for the more sensitive nucleus; this is effectively the selective inversion (of populations) for part of the multiplet of the sensitive nucleus. This has the effect of enhancing the population differences across the energy levels of the coupled, less sensitive nucleus, thus making the observed resonances more intense. Polarization (population differences) has thus been transferred from the more sensitive to the less sensitive nucleus. For a single acquisition the sensitivity improvement for spin $I = 1/2$ nuclei is of the order of the ratio of the resonance frequencies; hence using the $^{31}P$ polarization to drive the $^{109}Ag$ populations results in a sensitivity enhancement factor ~ 8.6 compared with 'single pulse' observation of the $^{109}Ag$ spectrum. There is a second benefit to using the polarization transfer sequences in that typically the spectrum must be accumulated over a period of time and each individual acquisition sequence should be separated by a 'relaxation delay'. The intensity of the observed resonance is derived from the nonequilibrium populations of the more sensitive nucleus, and it is often the case that relaxation times for $^{1}H$ and $^{31}P$ are shorter than for the metal nuclei, and therefore the accumulation sequence with polarization transfer can be repeated with greater frequency than the 'single pulse' observation. An example of the sensitivity enhancement is shown in **Figure 8** where the $^{109}Ag-\{^{31}P\}$ INEPT experiment provides considerably improved signal/noise ratio over the normal acquisition, with about one sixth of the number of scans accumulated with INEPT. There is one disadvantage common to all 1D and 2D polarization transfer experiments and this is the need to have prior knowledge of the magnitude of the scalar coupling constant ($J$, Hz) since there are delays in the various pulse sequences which are related to $J$. It is

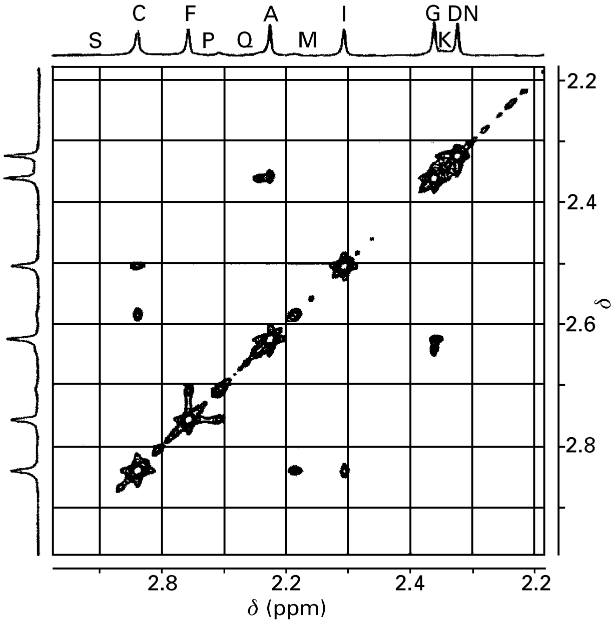



**Figure 7** Methyl region of the 400 MHz $^{1}H$ two-dimensional EXSY NMR spectrum of $[ReBr(CO)_3(tdmpzp)]$ in $CDCl_2CDCl_2$ solution at 296 K. Reproduced with permission from Gelling A, Noble DR, Orrell KG, Osborne AG and Šik V, *Journal of the Chemical Society, Dalton Transactions* 1996, 3065–3070.

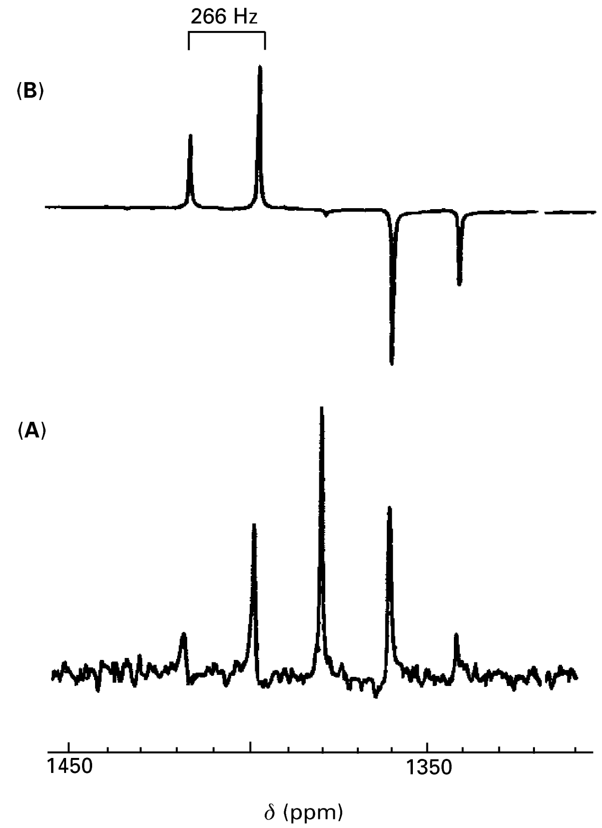



**Figure 8** 13.97 MHz $^{109}Ag$ NMR spectra of $[Ag(dppe)_2]NO_3$ in $CDCl_3$ solution at 300 K: (A) single pulse acquisition (accumulation of 12 111 scans); (B) $^{109}Ag-\{^{31}P\}$ INEPT experiment (accumulation of 2048 scans). Reproduced with permission from Berners Price SJ, Brevard C, Pagelot A and Sadler PJ, *Inorganic Chemistry* 1985, **24**: 4278–4281.

often possible to obtain a value for the coupling constant from the $^{1}H$ (or $^{31}P$) spectrum, but in other cases a guess must be made for *J*, and the experiment possibly repeated for a range of assumed *J* values.

### Two-dimensional experiments

Two-dimensional experiments provide heteronuclear shift correlations with the spectrum detected at the frequency of the more sensitive nucleus, and involve a double polarization transfer from sensitive to less sensitive to sensitive nucleus. The sensitivity improvement over the direct 'single pulse' observation of the less sensitive nucleus is even more dramatic than for the one-dimensional methods; here it is ~ $R^{5/2}$, where *R* is the ratio of the resonance frequencies, and for the $^{31}P$–$^{109}Ag$ example used above the sensitivity improvement is ~218. These experiments are often used to facilitate the observation of the spectrum of the less sensitive nucleus. Several experiments may be used and, when the one bond coupling is known the choice is between HMQC (heteronuclear multiple quantum coherence) and HSQC (heteronuclear single quantum coherence). There are advantages and disadvantages to both experiments; for example, the HMQC pulse sequence has fewer pulses than the HSQC, making the former experiment less susceptible to instrumental imperfections. However the resulting HMQC two-dimensional plot includes homonuclear coupling for the more sensitive nucleus (e.g. $^{1}H$ or $^{31}P$), reducing the intensities of the correlation peaks, and the line widths are determined by the relaxation rate of the multiple quantum coherence which may be faster than that of the single quantum coherence leading to broader lines in the HMQC plot. An example is the $^{109}Ag$–$^{31}P$ correlation shown in **Figure 9**.

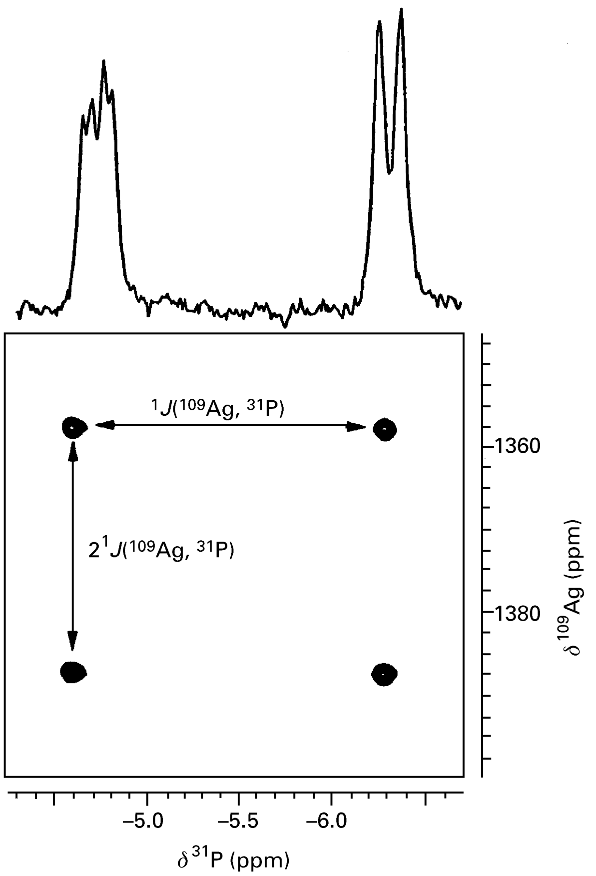



**Figure 9** A $^{109}Ag$–$^{31}P$ HSQC correlation experiment with inverse ($^{31}P$) detection ($^{109}Ag$ at 23.3 MHz, $^{31}P$ at 202 MHz) of a silver–chiral ferrocene complex in the presence of an excess of the isonitrile $CNCH_2(CO_2Me)$. This shows a single $^{109}Ag$ resonance. Reproduced with permission from Lianza F, Macchioni A, Pregosin P and Rüegger H, *Inorganic Chemistry* 1994, **33**: 4999–5002.

## Paramagnetic systems

The incidence of paramagnetism in inorganic molecules and materials is fairly common, and is due to the presence of unpaired electrons, usually associated with a metal centre. This paramagnetism will often lead to large chemical shifts of the NMR-active nuclei in the sample and may also induce severe line broadening of the resonances. Certainly paramagnetic metal centres in a variety of biomolecules may result in spectra dispersed on the chemical shift scale over hundreds of ppm, compared with tens of ppm for diamagnetic analogues. Such effects are fairly common, for example, in the $^{1}H$ and $^{13}C$ NMR spectra of a wide range of natural haem systems containing iron, and in model haem and porphyrin systems wherein the paramagnetic shifts may be related, in part, to the distribution of inpaired electron spin density throughout the molecule. The paramagnetism of inorganic complexes may be used to good effect in other areas. Lanthanide metals complexed with a range of organic ligands have been used for a number of years as shift or relaxation reagents. The shift (relaxation) reagent, when added to a solution of a diamagnetic compound, may form a weak complex with the diamagnetic molecule and result in paramagnetic changes in the chemical shifts (relaxation rates) of the nuclei in the substrate. The magnitudes of these changes in shift or relaxation rate depend upon the geometry of the weak complex and so may be analysed to give information about the structure of the diamagnetic compound. A second area of application of paramagnetic organometallic complexes is as contrast agents for use in MRI experiments. When the complex is introduced *in vivo*,

if there is a differential distribution of the agent between normal tissue and a lesion, then the agent will induce a differential in the relaxation rates of the water protons in these regions. Since the contrast in the MR image can be tuned to the relaxation properties of the water protons, then enhanced contrast in the image is obtained.

## List of symbols

$B_0$ = magnetic field strength; $e$ = electronic charge; $h$ = Planck's constant; $I$ = spin quantum number; $J$ = coupling constant; $k_r$ = rate coefficient; $q_{zz}$ = largest component of electric field gradient; $Q$ = quadrupole moment; $r$ = distance between nuclei; $R$ = ratio of resonance frequencies; $T_1$ = spin–lattice relaxation time; $T_2$ = spin–spin relaxation time; $T_{1\rho}$ = rotating frame relaxation time; $W_Q$ = line width; $\delta_{ii}$ = component of chemical shift; $\delta_{iso}$ = isotropic chemical shift; $\Delta\delta$ = chemical shift anisotropy; $\nu$ = resonance frequency; $\eta$ = asymmetry parameter; $\Xi$ = resonance frequency in a magnetic field strength that gives the $^1$H resonance of $SiMe_4$ at exactly 100 MHz.

*See also:* **Chemical Exchange Effects in NMR; Chemical Shift and Relaxation Reagents in NMR; Halogen NMR Spectroscopy (excluding $^{19}$F); Heteronuclear NMR Applications (As, Sb, Bi); Heteronuclear NMR Applications (B, Al, Ga, In, Tl); Heteronuclear NMR Applications (Ge, Sn, Pb); Heteronuclear NMR Applications (La–Hg); Heteronuclear NMR Applications (O, S, Se, Te); Heteronuclear NMR Applications (Sc–Zn); Heteronuclear NMR Applications (Y–Cd); High Resolution Solid State NMR, $^{13}$C; High Resolution Solid State NMR, $^1$H, $^{19}$F; Inorganic Compounds and Minerals Studied Using X-ray Diffraction; NMR of Solids; NMR Relaxation Rates; NMR Spectroscopy of Alkali Metal Nuclei in Solution; Nuclear Overhauser Effect; $^{31}$P NMR; $^{29}$Si NMR; Solid State NMR, Methods; Solid State NMR, Rotational Resonance; Solid State NMR Using Quadrupolar Nuclei; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Two-Dimensional NMR Methods.**

## Further reading

Aime S, Botta M, Fasano M and Terreno E (1998) Lanthanide(III) chelates for NMR biomedical applications. *Chemical Society Reviews* **27**: 19–29.

Bertini I and Luchinat C (1986) *NMR of Paramagnetic Molecules in Biological Systems*. Menlo Park, CA: Benjamin/Cummings.

Brey WS (ed) (1988) *Pulse Methods in 1D and 2D Liquid-phase NMR*. New work: Academic Press.

Gielen G, Willem R and Wrackmeyer B (eds) (1996) *Advanced Applications of NMR to Organometallic Chemistry*. Chichester: Wiley.

Mann BE (1974) $^{13}$C NMR chemical shifts and coupling constants of organometallic compounds. *Advances in Organometallic Chemistry* **12**: 135–213.

Mann BE (1991) The Cinderella nuclei. *Annual Reports on Nuclear Magnetic Resonance Spectroscopy* **23**: 141–207.

Mason J and Jameson C (eds) (1987) *Multinuclear NMR*. New York: Plenum Press.

Orrell KG, Šik V and Stephenson D (1990) Quantitative investigations of molecular stereodynamics by 1D and 2D NMR methods. *Progress in Nuclear Magnetic Resonance Spectroscopy* **22**: 141–208.

Orrell KG (1999) Dynamic NMR spectroscopy in inorganic and organometallic chemistry. *Annual Reports on Nuclear Magnetic Resonance Spectroscopy* **37**: 1–74.

Pregosin PS (ed) (1991) *Transition Metal Nuclear Magnetic Resonance*. Amsterdam: Elsevier.

Sandström J (1982) *Dynamic NMR Spectroscopy*. London: Academic Press.

Sievers RE (ed) (1973) *Nuclear Magnetic Resonance Shift Reagents*. New York: Academic Press.

Willem R (1988) 2D NMR applied to stereochemical problems. *Progress in Nuclear Magnetic Resonance Spectroscopy* **20**: 1–94.

# Structural Chemistry Using NMR Spectroscopy, Organic Molecules

**Cynthia K McClure,** Montana State University, Bozeman, MT, USA

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

Nuclear magnetic resonance spectroscopy is one of the most powerful tools that chemists use to determine the structure of compounds. Generally, NMR spectroscopy is the technique that most chemists, especially organic chemists, use first and routinely in structural analysis. In organic compounds, this non-destructive spectroscopic analysis can reveal the number of carbon and proton atoms and their connectivities, the conformations of the molecules, as well as relative and absolute stereochemistries, for example. The recent advent of pulsed field gradient (PFG) technology for NMR spectrometers has allowed the routine acquisition of sophisticated one-dimensional (1D) and two-dimensional (2D) NMR spectra in relatively short periods of time on complex organic molecules. This in turn has revolutionized organic structure determination such that deducing the three-dimensional structure of compounds takes a fraction of the time it used to. Mention of relevant 2D experiments that can aid in structure determination will be made in the appropriate sections herein.

This article is geared toward the analyses of small organic compounds, and will cover the following topics: practical tips in sample preparation; basic principles of one-dimensional $^1$H and $^{13}$C NMR spectroscopy and their use in organic structure determination, including chemical shifts, coupling constants and stereochemical analyses; and the application of more sophisticated 1D and 2D experiments to structure elucidation. Examples of structural analyses of organic compounds via NMR methods are ubiquitous in the literature such that it is impractical to mention more than just a few of them here. Therefore, the reader is encouraged to peruse the organic chemistry literature to find structural analyses of the specific types of organic compounds of interest. This article will deal mainly with generalities of organic compound structure elucidation, although several relevant examples will be presented.

## General practical considerations

Deuterated solvents are utilized with FT NMR spectrometers to provide an internal lock signal to compensate for drift in the magnetic field during the experiment. The more common solvents used for organic compounds are $CDCl_3$, $CD_3CN$, $CD_3OD$, acetone-$d_6$, benzene-$d_6$, DMSO-$d_6$ and $D_2O$. Since all deuterated solvents contain some protonated impurities (e.g. $CHCl_3$ in $CDCl_3$), one should choose a solvent that will not interfere with the NMR peaks of interest from the sample. Tetramethylsilane (TMS) is usually added to the sample as an internal standard for both proton and carbon spectra, being set at 0.0 ppm in both cases. However, the small protonated solvent impurities also make good standards as the chemical shifts of these peaks are published in many texts and are reported relative to TMS.

Protons provide the highest sensitivity for NMR observations, and therefore only small quantities of sample are needed (1–10 mg in 0.5 mL of solvent for an FT instrument). $^{13}$C NMR has a much lower sensitivity than proton NMR due to the low natural abundance of $^{13}$C (1.1%) compared with $^1$H (100%), and the fact that the energy splitting and hence the resonance frequency for carbon is approximately one quarter that of proton. Thus, for a spectrometer whose $^1$H frequency is 300 MHz, the frequency for $^{13}$C is 75.5 MHz. To obtain a carbon NMR spectrum in a timely manner, one needs to use either more sample than for a $^1$H NMR spectrum (>20 mg of a compound with MW≈150–300 g mol$^{-1}$), or a higher field spectrometer.

## $^1$H NMR

As mentioned earlier, $^1$H NMR is a very valuable method for obtaining information regarding the molecular structure of organic compounds with any number of protons. The electronic environment, as well as near neighbours and stereochemistry, can be determined by analysing the chemical shifts and spin–spin couplings of protons. The relative number of protons can be determined by direct integration of the areas under the peaks (multiplets), as the number of protons is directly proportional to the area under the peaks produced by those protons. To obtain

accurate integrations, however, the relaxation delay needs to be at least 5 times the longest $T_1$ in the sample.

### Proton chemical shift

Chemical shifts are diagnostic of the electronic environment around the nucleus in question. Withdrawal of electron density from around the nucleus will deshield the nucleus, causing it to resonate at a lower field (higher frequency or chemical shift). Higher electron density around a nucleus results in shielding of the nucleus and resonance at higher field (lower frequency or chemical shift (δ)). Therefore, basic details of the molecular structure can be gleaned from analysis of the chemical shifts of the nuclei. Factors that affect the electron density around the proton in question include the amount of substitution on the carbon (i.e. methyl, methylene, methine), the inductive effect of nearby electronegative or electropositive groups, hybridization, conjugation interactions through π bonds, and anisotropic (ring current) effects. Tables of proton chemical shifts can be found in various texts, such as those listed in Further reading.

As alkyl substitution increases on the carbon that possesses the proton(s) in question, the deshielding increases due to the higher electronegativity of carbon compared with hydrogen (e.g. $CHR_3$ > $CH_2R_2$ > $CH_3R$), producing a downfield shift of the resonances (methine most downfield, methyl most upfield). The deshielding effect of electron-withdrawing groups depends directly upon the electronegativity of these groups, and upon whether their effects are inductive (less effective) or through resonance (more effective). This deshielding effect falls off rapidly with increasing number of bonds between the observed proton and the electronegative group. One can, therefore, estimate chemical shifts of alkyl protons by analysing the amount of carbon substitution and the effects of nearby electron-withdrawing groups. A fairly accurate calculation of chemical shifts for methylene protons attached to two functional groups (X–$CH_2$–Y) is possible by using Shoolery's rule, where the shielding constants for the substituents, $\Delta_i$, are added to the chemical shift for methane. Tables of these shielding constants can be found in most texts on NMR spectroscopy.

To some extent, hybridization also influences the electron density around the proton in question by electronegativity effects. With increasing *s* character in a C–H bond, the electrons are held closer to the carbon nucleus. The protons consequently experience less electron density and are, therefore, more deshielded. This reasoning applies very well to protons attached to $sp^3$ rather than $sp^2$ carbons. For sp (acetylenic) protons, however, anisotropic effects are the dominating factors.

Electron-donating or electron-withdrawing groups directly attached to aromatic or alkene $sp^2$ carbons greatly affect the chemical shifts of aromatic or vinyl protons via π bond interactions (resonance). Thus, vinyl protons on the β-carbon of an α,β-unsaturated carbonyl system are further downfield (more deshielded) than the proton on the α-carbon due to resonance, and the opposite holds true for the β-proton(s) of a vinyl ether, as shown in **Figure 1**. In aromatic systems, electron-withdrawing groups deshield the protons ortho and para to it relative to unsubstituted benzene, while a group that is electron-donating by resonance will shield the ortho and para protons such that they resonate at a field higher than unsubstituted benzene (δ7.27). Empirical methods for estimating the chemical shifts of protons on substituted alkenes and benzene rings have been developed (see Further reading). It should be realized, however, that the anisotropies of aromatic and alkenyl systems are also responsible for the larger than expected downfield shifts of the protons. The large downfield shift of aldehyde protons (~δ9.5) is due in large part to the anisotropic shielding/deshielding effect (called the cone of shielding/deshielding in carbonyls), as seen in alkenes and aromatic compounds.

Shielding and deshielding effects via anisotropy caused by ring currents can also affect protons not directly attached to the alkene, alkyne, carbonyl or aromatic systems. A good example of this in shown in **Figure 2**. The calculated chemical shift of the methine proton $H_a$ in the absence of any ring current effects is δ4.40, while the observed chemical shift is

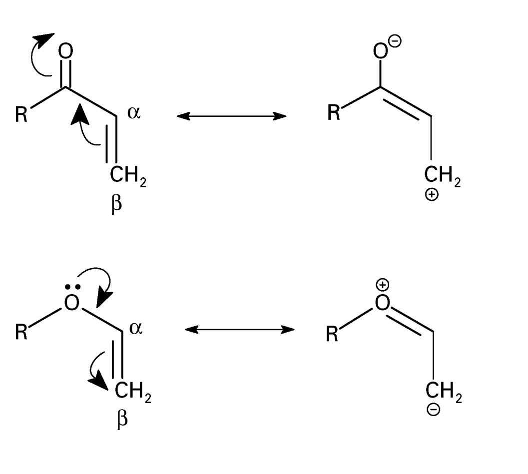



**Figure 1** Shielding and deshielding effects due to resonance.

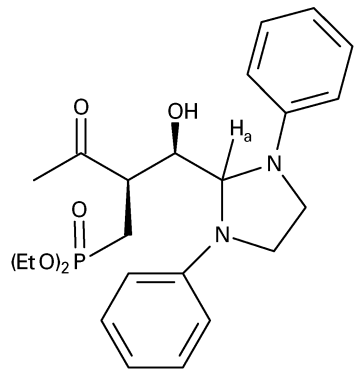



**Figure 2** Proton $H_a$ is deshielded by ~1 ppm due to the ring current of the nearby phenyl group.

δ5.44. The low energy conformation of the molecule (from molecular modelling) has one of the phenyl rings very near the proton $H_a$. Therefore, it appears that the ring current of this phenyl group is deshielding this proton by ~1 ppm.

In organic molecules possessing protons with very similar chemical shifts, such as steroids or carbohydrates, it would be advantageous to be able to simplify the spectrum by eliminating all spin–spin splittings, thereby allowing the determination of resonance frequencies by only chemical shift effects. An improved method has been developed to produce this type of 'chemical-shift spectrum', and is illustrated in **Figure 3**. Overlapping resonances are resolved into singlets, and this allows for a more straightforward structural assignment of the resonances. Near neighbours, coupling constants and relative stereochemistries can be determined by other spectral editing techniques and experiments (see below).

## Through-bond coupling: determination of near neighbours and stereochemistry

The analysis of through-bond spin–spin coupling (scalar or *J* coupling) allows for ready determination of the number of neighbouring protons, as well as the relative stereochemistry in certain cases. See the texts listed in Further reading for more in-depth discussions of spin–spin coupling. In short, spin–spin couplings occur between magnetically nonequivalent nuclei (here, protons) through intervening bonding electrons, and decreases with increasing number of intervening bonds. Protons that are chemically equivalent (interchangeable by a symmetry operation) are magnetically equivalent if they exhibit identical coupling to any other nucleus not in that set. However, protons with the same chemical shift do not split each other even when the coupling constant

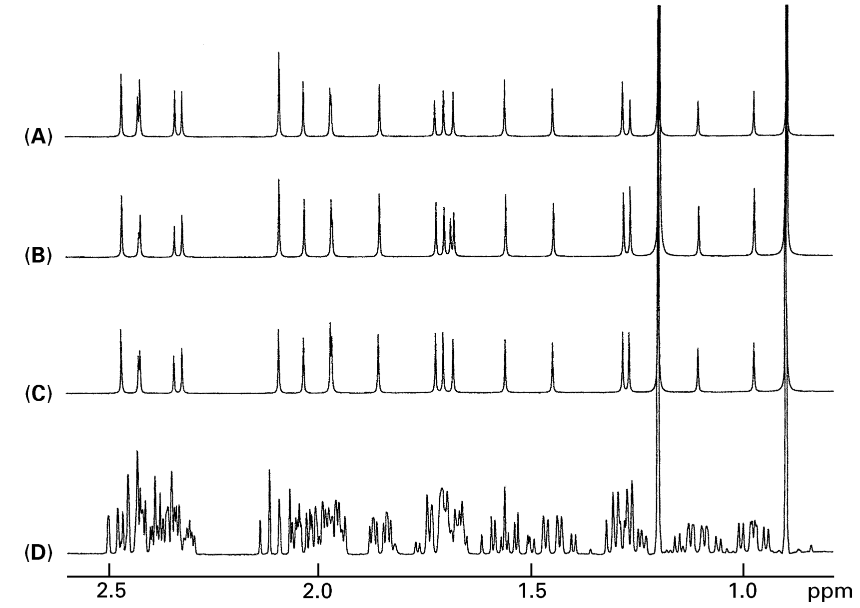



**Figure 3** Chemical shift spectra of 4-androsten-3,17-dione obtained from (a) the reflected *J* spectrum; (b) the purged *J* spectrum (the additional response near δ1.7 is from the residual water signal); and (c) the *z*-filtered *J* spectrum. The conventional $^1H$ spectrum is shown in (d). Reprinted with permission from Simova S, Sengstschmid H and Freeman R (1997) Proton chemical-shift spectra. *Journal of Magnetic Resonance* **124**: 104–121.

between them is non-zero. Rapid rotation about a C–C single bond, such as with a $CH_3$ group, results in an average environment for each methyl proton, and hence, equivalence.

Interacting protons with very different chemical shifts are weakly coupled if the difference in chemical shifts between the coupled protons, $\Delta\delta$, is large compared to the coupling constant $J$, i.e. $\Delta\delta/J > 10$. The multiplets resulting from this weak coupling are considered 'first-order' patterns, and can be interpreted easily. The multiplicity is governed by the $(2nI + 1)$ rule, where $n$ is the number of *magnetically equivalent* coupled protons and $I$ is the spin of the nucleus. In first-order systems, the multiplicities and peak intensities of coupled protons can be predicted using Pascal's triangle. For example, a proton split by two magnetically equivalent neighbours will be a triplet with peak intensities of 1:2:1. The frequency difference between the lines of the multiplet is the coupling constant, $J$, reported in Hz, and is invariant with changes in the strength of the magnetic field. The recent greater accessibility to higher NMR field strengths has enabled the interpretation of most proton spectra as first-order. In symmetrical spin systems, these simple rules do not apply and a more rigorous analysis is needed.

The Pople spin notation system is generally utilized to indicate the degree of difference among nuclei. Thus, in a two spin system, AX indicates a molecule with two nuclei where the chemical shift difference is much larger than the coupling between them (weakly coupled system, first-order analysis possible), whereas AB indicates a molecule containing two strongly coupled nuclei with similar chemical shifts. An $A_2BB'$ notation indicates a set of two equivalent nuclei (A) interacting with two nuclei (B, B') that are chemically, but not magnetically, equivalent.

For proton NMR, the most diagnostic couplings are 2-bond ($^2J$, geminal), 3-bond ($^3J$, vicinal), and 4-bond ($^4J$, W-type) couplings. Geminal couplings can be quite large, but may not be evident due to the symmetry associated with the carbon and protons in question. As mentioned above, the lack of geminal coupling is due to the identical chemical shifts of the protons involved. Vicinal 3-bond proton–proton couplings tend to be the most useful when determining stereochemistry, although coupling beyond three bonds can be important in systems with ring strain (small rings, bridged systems) or bond delocalization, as in aromatic and allylic systems.

For simple organic molecules, pattern recognition of multiplets can simplify structure determination. For example, the presence of an upfield triplet due to three protons and a more downfield quartet due to two protons with the same coupling constant is most probably due to an ethyl group ($X–CH_2CH_3$). Therefore, it is useful to look for common patterns. Many preliminary assignments can be made in a standard 1D $^1H$ spectrum due to the reciprocity of coupling constants ($J_{AB} = J_{BA}$) With more complex patterns due to coupling to several magnetically non-equivalent protons, interpretation can be done via first-order analysis only if no two of the spins within an interacting multispin system have $\Delta\delta/J \leq 6$. Multiplets such as doublet of doublets (dd), doublet of triplets (dt), triplet of doublets (td), doublet of quartets (dq), doublet of doublets of doublets (ddd), etc., can usually be analysed by first-order techniques, especially if the spectrum was run at a fairly high magnetic field. A very useful and practical guide to first-order multiplet analysis that utilizes either a systematic analysis of line spacings or inverted splitting trees to determine the couplings is listed in Further reading.

Measurement of coupling constants can usually be done directly from the 1D spectrum with well resolved multiplets, or aided by simple 1D homonuclear decoupling experiments where irradiation of one of the *weakly* coupled nuclei (i.e. nuclei with very different chemical shifts) simplifies a multiplet by eliminating that spin–spin interaction. Several two dimensional techniques also help to determine the coupling network, and determine and assign coupling constants. A 2D COSY (correlated spectroscopy) spectrum is a homonuclear experiment, and provides a map of the proton–proton $J$-coupling network in the molecule. The spectrum contains a set of auto-correlated peaks along the diagonal ($\omega_1 = \omega_2$), which is the original spectrum. For those spins that exchange magnetization due to $J$-coupling, $\omega_2 \neq \omega_1$ and off-diagonal peaks appear. The diagonal peaks that correspond to $J$-coupled spins are connected by symmetrical pairs of off-diagonal peaks. In general, strongly coupled protons are handled better in a COSY experiment than with conventional 1D homonuclear decoupling. However, in molecules that have overlapping resonances, it can be difficult to accurately assign the cross peaks. New computer programs are being developed to provide automated processing and assignment of the data (see Further reading). Further simplification of the spectrum can be attained by utilizing a DQF-COSY (double quantum filtered COSY) experiment, where singlets are essentially eliminated from the spectrum. Coupling constants can be attained from a COSY spectrum, but it is not a trivial process. The $J$ values measured from a COSY spectrum also tend to be slightly larger than the actual $J$ value.

The two-dimensional experiment, homonuclear *J*-resolved spectroscopy, is utilized to accurately measure the scalar coupling constants. This method can readily resolve overlapping signals, as well as strongly coupled systems. From the contour plot of the spectrum, multiplets are resolved along the *y*-axis (Hz), and the coupling constants are read directly along this axis. Projection of the multiplet onto the *x*-axis (δ-axis) provides a single resonance line for each distinct spin system without the effects of coupling (i.e. is proton-decoupled), and accurate values of δ (ppm) can be attained.

### Stereochemical assignments

Accurate stereochemical assignments are generally only possible in rigid or ring systems where free rotation about carbon–carbon bonds is hindered or not possible. As mentioned above, vicinal, three-bond couplings ($^3J$) can be quite diagnostic of the stereochemical relationship between the coupling protons. The Karplus equation (Eqn [1]) can predict the vicinal coupling constant $^3J_{H–C–C–H}$ with reasonable accuracy if the H–C–C–H dihedral angle is known. Thus, dihedral angles near 0° or 180° have the largest coupling constants, while a dihedral angle of 90° has a coupling constant near 0 Hz.

$$\begin{aligned} ^3J &= A\cos^2\phi + C \quad (\phi = 0–90°) \\ ^3J &= A'\cos^2\phi + C' \quad (\phi = 90–180°) \end{aligned} \qquad [1]$$

Use of this relationship in alkenes and ring (or bridged) systems works very well to predict stereochemistry. For alkenes, *trans* coupling is in the range of 12–18 Hz, and *cis* coupling is 6–12 Hz. See the texts listed in Further reading for tables listing coupling constants in various alkenyl systems. *Trans*, diaxial protons on a six-membered carbocyclic ring have a dihedral angle of ~180° and a *J* value of 6–14 Hz (typically 8–10 Hz), whereas protons oriented axial–equatorial or equatorial–equatorial with dihedral angles near 60° have coupling constants of 0–5 Hz (typically 2–3 Hz). In five-membered rings, vicinal *trans* and *cis* coupling constants may be similar in magnitude due to the reduced conformational flexibility of the ring relative to six-membered rings. Five-membered rings will adopt a twist conformation to relieve eclipsing interactions. For example, in the oxazolidinones in **Figure 4**, it was found that $^3J_{4\text{-}5\ cis}$ = 7.2 Hz, and $^3J_{4–5\ trans}$ = 5.9 Hz.

Where no coupling is possible because of a quaternary centre or if the coupling constants fall on the borderline between two possible orientations nuclear Overhauser effect (NOE) measurements may need to be taken in order to definitively establish the relative stereochemistry. The use of pulsed field gradients (PFGs) in the acquisition of NOE enhancement spectra (called GOESY) now allows one to avoid the need to compute difference NOE spectra, as was done in the past. With NOE difference spectra, it was hard to avoid subtraction artifacts, and thus difficult to obtain accurate NOE values, especially small (<1%) enhancements. Utilizing PFGs, the only resonances now seen in the spectrum are those from spins which are cross-relaxing with the irradiated spin.

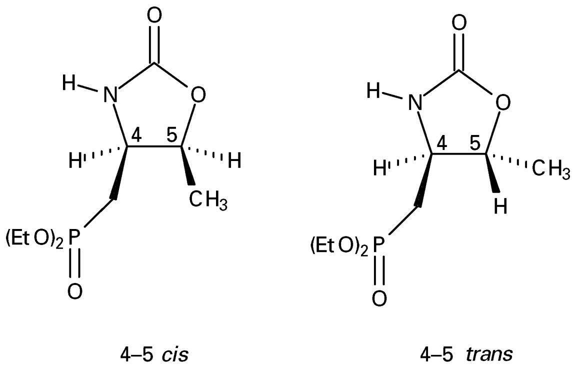



**Figure 4** Oxazolidinones exhibiting coupling constants of $^3J_{4–5\ cis}$ = 7.2 Hz and $^3J_{4–5\ trans}$ = 5.9 Hz.

The stereochemical assignments for the oxazolidinones in **Figure 4**, were confirmed by the NOE enhancements of 13% between H4 and H5 in the *cis* isomer, and only 2% between H4 and H5 in the *trans* isomer. There was also an NOE enhancement of 5% between the $CH_2$ α to the phosphonate and H5 in the *trans* isomer. In the examples shown in **Figure 5**, the vicinal coupling constants were all small (0–3 Hz), thus indicating that all the methine protons were in axial–equatorial or equatorial–equatorial relationships. The NOE measurements (GOESY experiment) indicated in **Figure 5** enabled final assignments of the relative stereochemistries.

## $^{13}$C NMR

Information regarding the number and types of carbons in an organic compound can be provided by carbon NMR spectroscopy. Since the chemical shift range is greater for carbon than for proton, a greater dispersion of signals is seen. Different functional groups that contain at least one carbon, such as

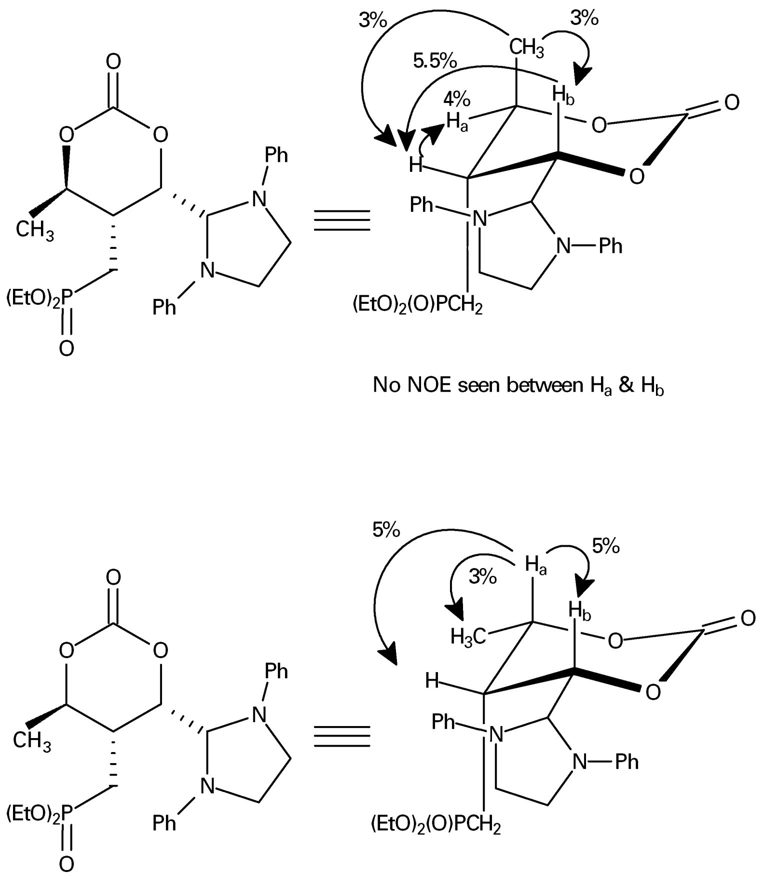



**Figure 5** NOE enhancements (GOESY experiments) of two diastereomeric cyclic carbonates to confirm relative stereochemistries.

ketone, ester and amide carbonyls, alkenes, alkanes, alkynes, nitriles, imines, etc., can generally be readily distinguished by $^{13}C$ NMR spectroscopy. The chemical shifts of aromatic and alkene carbons, however, are in the same chemical shift range, and at times cannot be differentiated. This is in contrast to aromatic and alkene protons which exhibit different chemical shift ranges.

In general, the factors that affect the chemical shifts of carbons are the same as for protons (i.e. electron density around the nucleus in question, and anisotropy effects). Carbon chemical shifts can be readily calculated from tables of shift effects found in many texts. However, unlike protons attached to $sp^2$ carbons, $sp^3$ carbons attached to $sp^2$ carbons exhibit only a small shift difference. There are also few good substituent parameters available for calculating the chemical shifts of alkene carbons bearing polar groups, unlike the calculation of $^1H$ NMR chemical shifts near polar groups. However, in systems where resonance is present, some predictions can be made of relative shift differences in the carbons (see **Figure 1**).

Carbon–proton connectivities can be determined using several methods. The number of protons directly attached to the carbon in question will split the carbon resonance according to the $2nI + 1$ rule seen in proton NMR. There tends to be, however, much overlap of the multiplets in fully proton-coupled carbon spectra, sometimes such that it is very difficult to distinguish between the various multiplets. Routine carbon spectra are therefore measured fully proton decoupled for simplicity. Information regarding the exact number of protons attached to the carbons can be acquired from APT, DEPT or INEPT experiments. In APT spectra, the carbons bearing an odd number of protons (CH, $CH_3$) can be distinguished from carbons with no or two attached protons (quaternary C, $CH_2$). DEPT and INEPT experiments can distinguish between all four types of carbons (primary, secondary, tertiary and quaternary). Heteronuclear 2D $J$-resolved spectroscopy can also be used to obtain the multiplicities of the carbons, as well as $^1J_{C-H}$.

A complete mapping of the protons to the carbons they are attached to is possible via a HETCOR

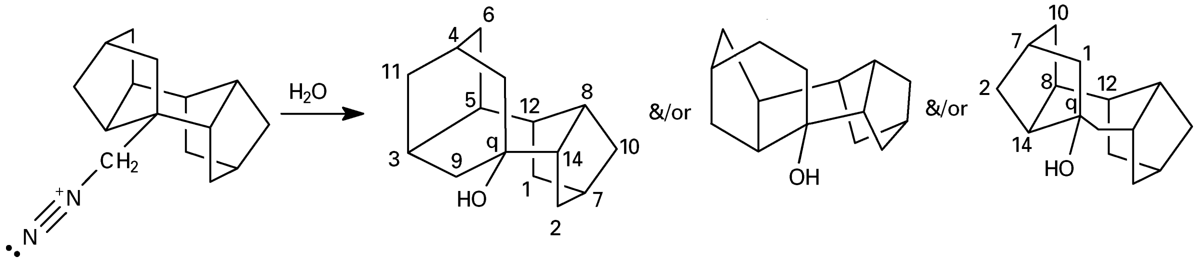
(heteronuclear chemical shift correlation) experiment. This method correlates the peaks of the proton spectrum of a compound with the peaks of its carbon spectrum. A contour plot of the spectrum has a cross peak at the intersection of the vertical line drawn from a carbon resonance (plotted along the *x*-axis) with the horizontal line drawn from a proton peak (plotted along the *y*-axis). However, this method is relatively insensitive as it suffers from the low natural abundance of $^{13}C$ atoms in the molecule. Utilization of the inverse detection method, HMQC (heteronuclear correlation through multiple quantum coherence), can alleviate this problem as $^{13}C$ responses are observed in the $^{1}H$ spectrum. Normally, $^{1}H$–$^{13}C$ coupling information is included in the $^{1}H$ dimension, although proton decoupling from carbon is possible. Quaternary carbons, however, will not be present in a HMQC spectrum. A method to aid in assignments of quaternary carbons, as well as carbon and proton connectivities, is the HMBC (heteronuclear multiple bond correlation) experiment. In these spectra, cross peaks are observed connecting the $^{13}C$ signals to $^{1}H$ signals two or more bonds away.

The INADEQUATE experiment is designed to map out the entire carbon skeleton of a molecule by providing carbon–carbon connectivities, and offers great possibilities for organic structure determination. However, it suffers severely from the low natural abundance of covalently bound $^{13}C$–$^{13}C$ pairs. Several groups have recently offered modifications of the pulse sequence to try to overcome this limitation. It remains to be seen, however, if these new programmes will produce the higher sensitivities needed for this experiment to become a routine analytical procedure.

Even without the newer modifications to the pulse sequence, the INADEQUATE experiment can be an invaluable method for structure determination in some cases, as illustrated in the example in **Figure 6**. The rearrangement of the cage hydrocarbon diazonium ion [1] in the presence of water could lead to three possible alcohol products [2], [3] and [4]. Only two alcohol products were obtained from this reaction in isolated yields of 38% and 7%. The structures of these two products could only be determined by the 2D INADEQUATE experiment due to the hydrocarbon nature of their structures, and therefore, the lack of any distinguishing details in the proton spectra. Each compound also had the same number of methylene, methine and quaternary carbons, thus precluding the utilization of structure determination by simple 1D $^{13}C$ spectroscopy.

The INADEQUATE and APT spectra for the major product are shown in **Figure 7**. In an INADEQUATE spectrum, the pairs of adjacent carbons, and hence the connectivity, can be mapped out similarly to a COSY spectrum. The major difference here is that the original spectrum is not on the diagonal in an INADEQUATE spectrum (as in a COSY spectrum), but is in the *x*-axis direction (= normal $^{13}C$ frequencies) along the line $\nu_1 = 0$ (residual single quantum signals). The *y*-axis is the frequency $\nu_1$, the double quantum frequency that is the sum of the frequencies of the two coupled nuclei referenced to a transmitter frequency at zero. The peaks arising from two coupled nuclei (here adjacent carbons) with shifts $\nu_a$ and $\nu_b$ have coordinates of $((\nu_a + \nu_b), X)$, where $X$ is the frequency of the carbon in a single quantum coherence spectrum (1D spectrum). At the double quantum frequency ($\nu_1$) for each pair of adjacent carbons, doublets will occur at the coordinates of $((\nu_a + \nu_b), \nu_a)$ and $((\nu_a + \nu_b), \nu_b)$. Thus, the midpoint between each pair of signals lies on a line with slope of 2, and helps to distinguish the real peaks from artifacts. The (C7–C1) – (C1–C7), (C14–C8) – (C8–C14) and (C3–C5) – (C5–C3) peak pairs are illustrated on the spectrum (B) in **Figure 7**.

In the example illustrated in **Figure 6**, it may be noted that compound [3] would require the grouping




[1] [2] [3] [4]

**Figure 6** Possible products from the rearrangement of structure [1] in water. Bratis AD, Bruch MD and Murray RK Jr unpublished results.

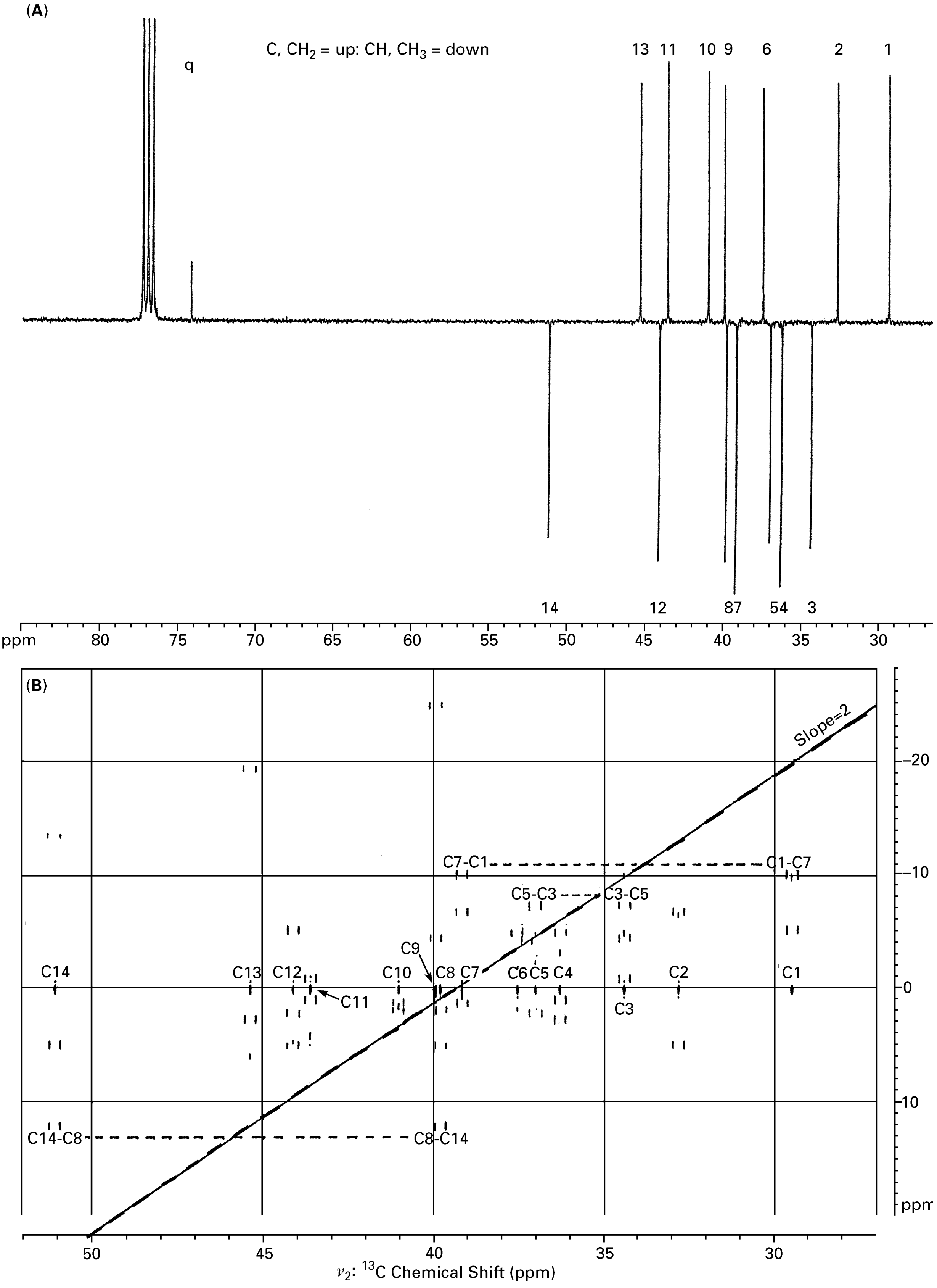
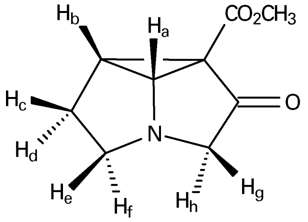



**Figure 7** (A) APT $^{13}C$ spectrum of compound [2]; (B) INADEQUATE spectrum of compound [2], with the connectivities of C1 to C7, (see **Figure 6**) C8 to C14 and C3 to C5 shown. Bratis AD, Bruch MD and Murray RK Jr unpublished results.

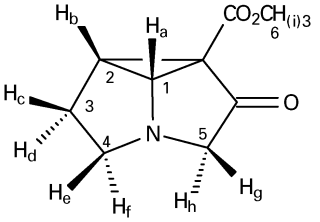
**Table 1** Comparison of calculated proton distances, dihedral angles and estimated *J* couplings, with *J* couplings and NOE results for compound [6]




| | *Mechanics* | | | *Experimental* | |
|---|---|---|---|---|---|
| *Protons* | *Distance* (Å) | *Angle* (°) | *J coupling* (Hz, estimated) | *J coupling* (Hz) | *NOE* |
| $H_a$–$H_b$ | 2.71 | 7.1 | 8 | 4.9 | yes |
| $H_b$–$H_c$ | 2.47 | 13 | 7.5 | 6.5 | yes |
| $H_b$–$H_d$ | 2.92 | 108 | 1 | 0 | ND |
| $H_c$–$H_d$ | 1.81 | 109 | 10–20 | 14.0 | yes |
| $H_c$–$H_e$ | 2.37 | 32 | 6 | 7.0 | yes |
| $H_c$–$H_f$ | 2.77 | 87 | <1 | 0 | ND |
| $H_d$–$H_e$ | 3.09 | 154 | 8 | NA | ND |
| $H_d$–$H_f$ | 2.48 | 35 | 6 | 6.2 | ND |
| $H_e$–$H_f$ | 1.79 | 107 | 10–20 | 12.3 | yes |
| $H_g$–$H_h$ | 1.8 | 108 | 10–20 | 18.0 | yes |

NA = not available; ND = no NOE detected.

**Table 2** HETCOR data for compound [6]

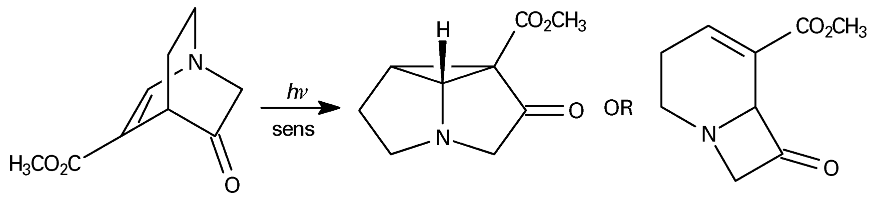



| *Carbon* | *$^{13}$C δ* | *Proton(s)* | *$^1$H δ* |
|---|---|---|---|
| C-1 | 63.7 | $H_a$ | 4.23 |
| C-2 | 43.7 | $H_b$ | 3.06 |
| C-3 | 26.5 | $H_{c,d}$ | 2.25, 1.99 |
| C-4 | 64.3 | $H_{e,f}$ | 3.32, 2.79 |
| C-5 | 68.8 | $H_{g,h}$ | 3.64, 2.84 |
| C-6 | 53.1 | $H_i$ | 3.75 |

$-CH_2CH_2-$ to be present, which is not seen in the INADEQUATE spectrum. For the major compound to be structure [4], the connectivity of carbon 1 would have to be to carbon 7 and the quaternary carbon, q. The connectivity found for carbon 1 was to carbon 7 and the methine carbon 12. Therefore, only structure [2] fully supports all the spectral data for the major isolated product from this rearrangement.

## Putting it all together

The following is an outline of the basic procedure that a practising organic chemist follows when deducing the structure of an organic compound via NMR spectroscopy. First, standard 1D $^1$H and $^{13}$C spectra are acquired and analysed. If needed, a 'chemical shift spectrum' can provide a straightforward assignment of the δ value of all the resonances, including overlapping multiplets. Proton spin–spin couplings and near neighbours are determined either directly from the 1D $^1$H spectrum, or assisted by homonuclear decoupling experiments. If these experiments are not conclusive due to overlapping resonances, changing the solvent or utilization of a shift reagent can on occasion resolve the overlapping multiplets. A map of the *J*-coupling network in the molecule is available through 2D COSY or TOCSY experiments. New programs to assist in assignment of the cross-peaks in complicated COSY or TOCSY spectra are now becoming available. Utilization of the DQF-COSY experiment eliminates all the singlets from the spectrum. The homonuclear 2D *J*-resolved spectrum allows for the separation of overlapping resonances and, therefore, accurate measurement of the coupling constants and chemical shifts. A number of $^{13}$C experiments, such as APT, DEPT, INEPT, INADEQUATE, and heteronuclear 2D *J*-resolved experiments, can be run to assist in determining proton–carbon attachment, carbon–




[5] [6] Oxa-di-pi-methane rearrangement product [7] 1,3-Acyl shift product

**Figure 8** Possible products from the photolysis of compound [5]. Kiessling AJ and McClure CK unpublished results.

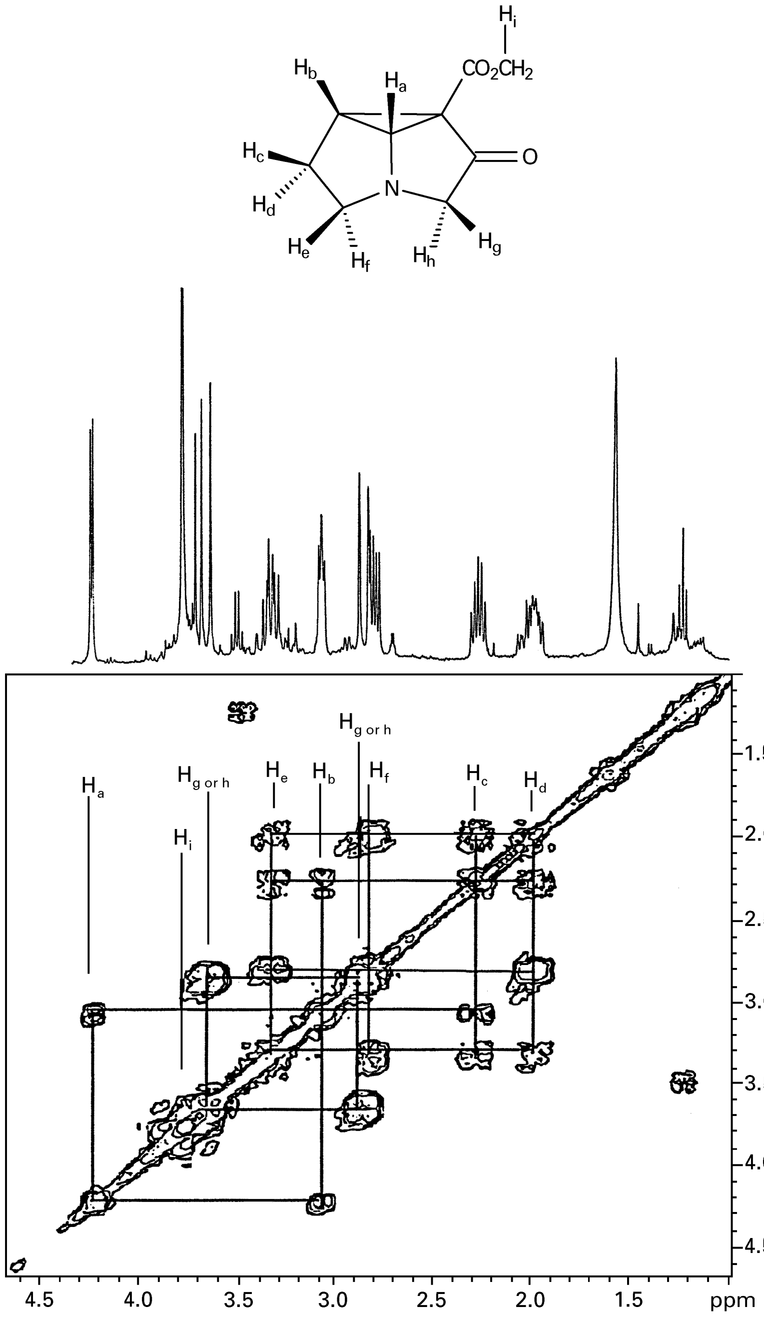



**Figure 9** COSY spectrum of compound [6].

carbon attachments for carbon skeleton determination, and values of $^{1}J_{C\text{–}H}$. A 2D HETCOR spectrum will confirm the proton–carbon connectivities. NOE, GOESY and ROESY experiments can assist in determining stereochemical information where coupling constants cannot.

The following is an example that illustrates how several of the experiments listed above, as well as molecular mechanics calculations, can be used to deduce the structure of the compound produced by a photochemical rearrangement. The photochemical irradiation of the bicyclic compound [5] was run under sensitized conditions (acetophenone). Two possible products, [6] and [7] are theoretically possible, and are shown in **Figure 8**. The 1,3-acyl shift product [7] normally arises from photolysis under non-sensitized conditions, but can be formed from certain compounds under sensitized photolysis conditions. The oxa-di-pi-methane rearrangement product [6] was the desired compound. From 1D $^{1}$H and $^{13}$C NMR spectra, the only product isolated from the photochemical rearrangement did not appear to

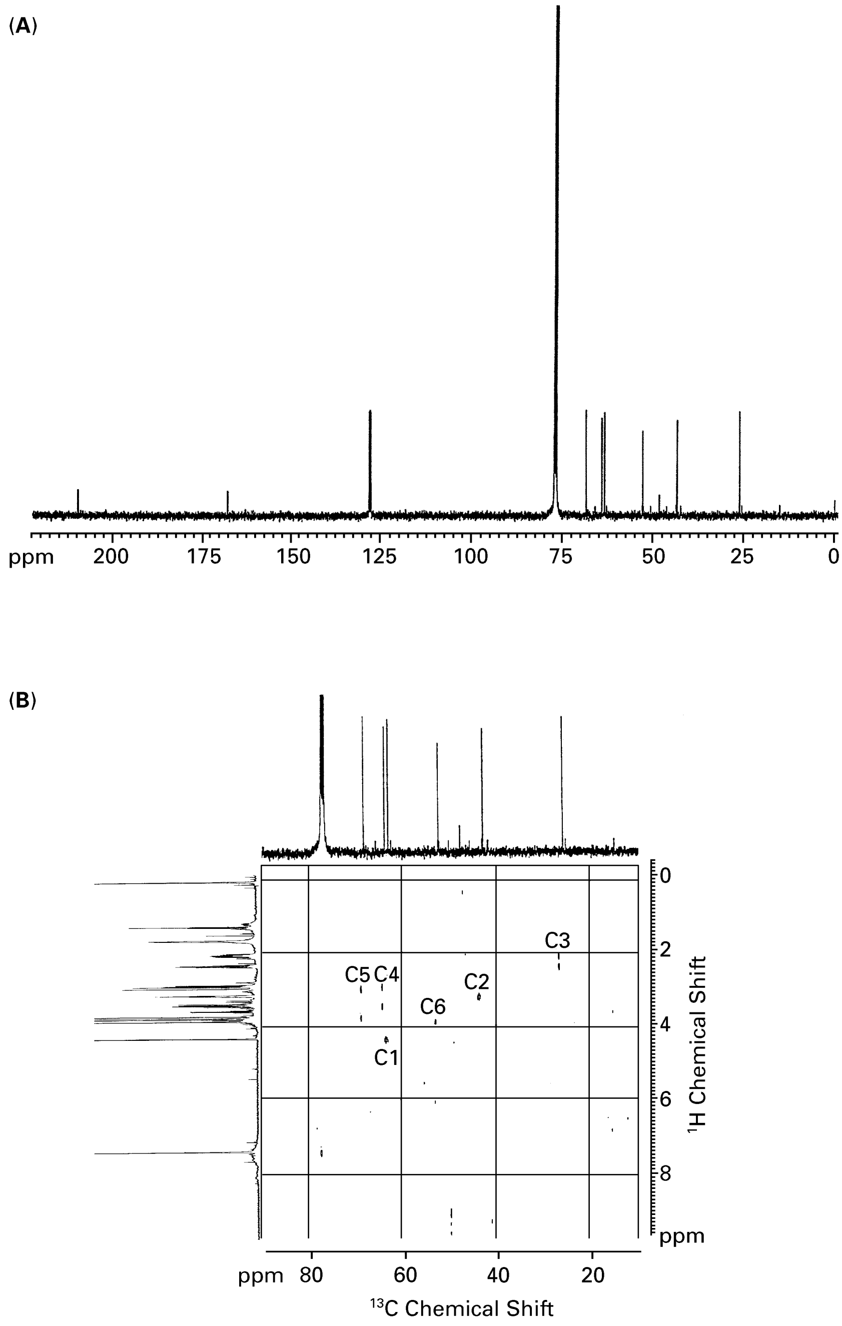



**Figure 10** (A) $^{13}$C spectrum of compound [6] (see **Figure 2**) in $CDCl_3$/benzene-$d_6$; (B) HETCOR spectrum of compound [6].

contain an olefin, as would be seen in the α,β-unsaturated ester [7]. Thus, the rearrangement most likely went via the oxa-di-pi-methane rearrangement, and not by the 1,3-acyl shift mechanism.

Further proof that the structure of the photoproduct was indeed [6] is as follows. From the standard COSY spectrum (**Figure 9**), all the proton–proton coupling networks could be established. The proton responsible for the peak at δ4.22 (d, $J = 4.8$ Hz) was coupled only to a proton at δ3.06 (dd, $J = 6.5$, 4.9 Hz), which in turn was coupled to only one other proton at δ2.25. Of the protons $H_a$–$H_f$, only proton $H_a$ was expected to be coupled to only one other proton, namely $H_b$. The chemical shift of δ4.22 was also reasonable for $H_a$. The multiplet at δ3.06 was therefore assigned to $H_b$, and was coupled to one other proton at δ2.25. This other proton could be either $H_c$ or $H_d$.

In order to assist in the assignments, the tricyclic structure [6] was submitted to molecular mechanics calculations to estimate the dihedral angles between the protons, and thus approximate the coupling constants using Equation [1] (see **Table 1**). According to these calculations, $H_b$ and $H_c$ had a dihedral angle of ~13° and thus an estimated coupling constant of 7.5 Hz, while $H_b$ and $H_d$ had a dihedral angle of ~110° and an estimated coupling constant of 1 Hz. The measured $J$ value between signals at δ3.06 and δ2.25 was 6.5 Hz, closely matching the calculated coupling constant between $H_b$ and $H_c$. Therefore, $H_c$ was assigned to the signal at δ2.25. This multiplet (ddd) exhibited three coupling constants of 14.0, 7.0 and 6.9 Hz. The large coupling of 14 Hz would be consistent with geminal coupling to proton $H_d$. The calculations predicted that in addition to $H_b$, $H_c$ would couple to $H_e$ with a dihedral angle of 32° and an estimated coupling constant of 6 Hz. $H_f$ was predicted to be nearly orthogonal to $H_c$, and thus have little or no coupling to $H_c$. From this data, $H_d$ was assigned to the multiplet at δ1.99 and $H_e$ to the multiplet at δ3.32. From the COSY spectrum, the multiplet at δ1.99 was further coupled to the multiplet at δ2.79, which was assigned to $H_f$. The multiplet (dd) at δ2.79 had two coupling constants of 12.3 and 6.2 Hz. The large coupling constant was geminal coupling with $H_e$. The other coupling constant was consistent with the calculation of the dihedral angle of 35° between $H_f$ and $H_d$. The protons $H_g$ and $H_h$ were coupled only to each other, and the exo proton $H_g$ was assigned as the downfield doublet of doublets at δ3.65.

The distances between the protons of the proposed structure were also calculated by molecular mechanics and are summarized in **Table 1**. The photoproduct was submitted to NOE experiments to verify the spatial relationships. The signals assigned to $H_a$, $H_b$ and $H_e$ yielded meaningful data, and the NOE results are shown in **Table 1**. Results of the NOE experiments are in agreement with the proposed structure, where $H_a$, $H_b$, $H_c$ and $H_e$ are shown to be in a *cis* relationship.

The $^{13}C$ and HETCOR spectra in **Figure 10A** and **Figure 10B** respectively, further verified the proposed structure [6]. No carbon signals were detected in the alkene region of the spectrum, consistent with the lack of alkene protons. The only carbonyl peak detected was at δ207.6, consistent with the ketone in [6]. The HETCOR data is summarized in **Table 2**.

*See also:* **$^{13}C$ NMR Methods; $^{13}C$ NMR Parameter Survey; Chemical Exchange Effects in NMR; Chemical Shift and Relaxation Reagents in NMR; Chromatography-NMR, Applications; Enantiomeric Purity Studied Using NMR; Magnetic Field Gradients in High Resolution NMR; NMR Data Processing; NMR Pulse Sequences; Nuclear Overhauser Effect; Structural Chemistry using NMR Spectroscopy, Peptides; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals; Two-Dimensional NMR Methods.**

## Further reading

Bourdonneau M and Ancian B (1998) Rapid-pulsing artifact-free double-quantum-filtered homonuclear spectroscopy. The 2D-INADEQUATE experiment revisited. *Journal of Magnetic Resonance* **132**: 316–327.

Bruch MD (ed) (1996) *NMR Spectroscopy Techniques*, 2nd edn. New York: Marcel Dekker.

Derome AE (1987) *Modern NMR Techniques for Chemistry Research*. Oxford: Pergamon.

Hoye TR, Hanson PR and Vyvyan JR (1994) A practical guide to first-order multiplet analysis in $^1H$ NMR spectroscopy. *Journal of Organic Chemistry* **59**: 4096–4103.

Lambert JB, Shurvell HF, Lightner DA and Cooks RG (1998) *Organic Structural Spectroscopy*. New York: Macmillan.

Sengstschmid H, Heinz S and Freeman R (1998) Automated processing of two-dimensional correlation spectra. *Journal of Magnetic Resonance* **131**: 315–326.

Silverstein RM, Bassler GC and Morrill TC (1998) *Spectrometric Identification of Organic Compounds*, 6th edn. New York: John Wiley.

Simova S, Sengstschmid H and Freeman R (1997) Proton chemical-shift spectra. *Journal of Magnetic Resonance* **124**: 104–121.

Stonehouse J, Adell P, Keeler J and Shaka AJ (1994) Ultrahigh-quality NOE spectra. *Journal of the American Chemical Society* **116**: 6037–6038

# Structural Chemistry Using NMR Spectroscopy, Peptides

**Martin Huenges and Horst Kessler**, Technische Universität München, Garching, Germany

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

## Introduction

The majority of biological processes depend directly on peptides and proteins. The sequence of most peptides and all proteins are encoded genetically and the polypeptides are post-translationally modified, processed and transported to their specific location in the cell. The wide range of possible chemical structures (owing to the combination of functional groups of their amino acid residues), especially in their three-dimensional dynamic conformation, allows peptides and proteins to play many different roles in biological processes, such as hormone/receptor interactions, cellular adhesion and cellular recognition, transport mechanisms between cell compartments or through membranes, and the processing of almost all chemical compounds, including peptides and proteins, to name only a few.

Although the conformation and dynamics of peptides and proteins are encoded in their sequence, they are not yet reliably predictable based on it. The determination of the dynamic 3D structure therefore was, and still is, of utmost importance for the interpretation and artificial modulation of their functions.

The challenges posed by peptides and proteins have strongly stimulated the development of modern NMR spectroscopy. Most of the multidimensional NMR techniques now available designed and applied initially to peptides and proteins.

We will discuss in this article first some general problems which arise in the NMR spectroscopy of peptides. Then, NMR techniques for signal assignment and extraction of conformational parameters will be described, followed by a short excursion into structure determination using NMR parameters. The final part will include the analysis of peptide dynamics based on NMR data.

## General problems with peptides

Peptides are composed of a linear, branched or cyclic array of amino acid residues (**Figure 1**). The peptide chain is defined and numbered from the N to the C terminus. The $\alpha$ carbon atoms of the amino acids are linked by peptide bonds. The bonds to the $\alpha$ carbon atom are described by their bond angles $\phi$(N–C$^\alpha$), $\psi$(C$^\alpha$–CO) and $\chi_1$(C$^\alpha$–C$^\beta$). The usually planar peptide bonds prefer the *trans*-configuration ($\omega$ = 180°) as shown in **Figure 1** for the Phe-Gly and Gly-Val bonds. Only in the case of Xaa-Pro pairs are *cis*- and *trans*-conformations of similar energy. In **Figure 1** the Val-Pro bond is in the *cis*-configuration. The barrier between *cis* and *trans* peptide bonds is between 16 and 20 kcal mol$^{-1}$ which leads to a slow exchange between the conformations on the NMR time-scale at room temperature. At higher temperature the exchange between the two states occurs at an increased rate, leading to a coalescence of the two signal sets in NMR spectra.

Rotations around $\phi$, $\psi$ and $\chi_1$ are fast on the NMR time-scale. At a result it is not straightforward to distinguish between a single preferred conformation and a rapid equilibrium between several conformations (see below). Linear peptides, approximately up to dodecamers, are normally very flexible and do not exist in or prefer a single conformation, although sometimes a slight preference for a distinct set of structures is observed. However, cyclization and/or sterically demanding substitution restrain the conformational space and often allow the identification of a preferred conformation.

The equilibrium between the different conformations of a peptide can be very sensitive to the environment. Hence, different conformations can be found in a single unit cell of a crystal, between different crystals, between crystal and solution as well as between 'free' and receptor bound peptides. 'Free' conformations are embedded in the solvent whose chemical and physical properties can induce drastic changes when different solvents are used. The evidence for such conformational changes often is indirect (solvent induced signal shifts), but a careful analysis of the whole 3D structure can also lead to the detection of such exchange processes (see e.g. antamanide). Direct observation of solvent induced conformational shifts is possible in the case of the *cis*/*trans* isomerization of an alkylated peptide bond,

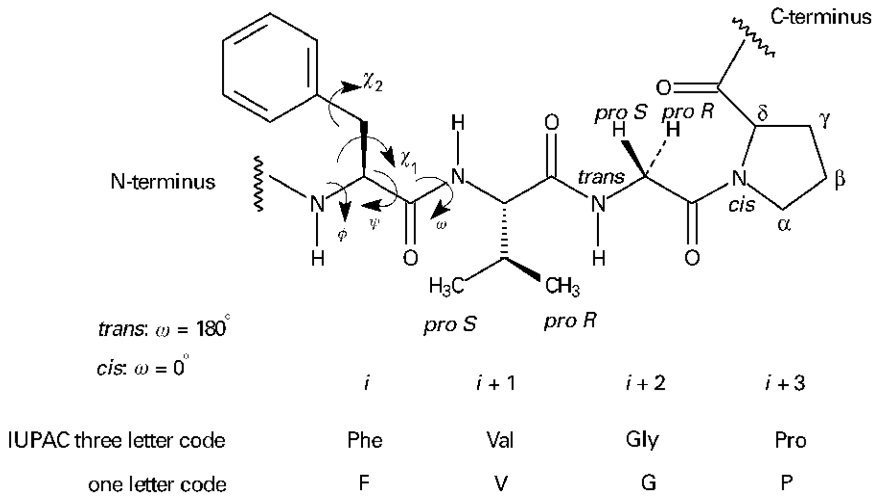



**Figure 1** Schematic representation of the tetrapeptide sequence – FVGP – used to illustrate the nomenclature.

since rotation around this bond is slow enough to be resolved on the NMR time-scale. For example, when cyclosporin A is dissolved in $CDCl_3$, benzene or THF one strongly dominating conformation is observed that contains a $NMeVal^9$-$NMeVal^{10}$ *cis* amide bond. The corresponding conformation with the same amide in the *trans*- conformation is populated at less than 5%. In more polar solvents, such as $CD_3CN$ and MeOH, a number of coexisting conformations are found, whereas a single but very different conformation is found when the peptide is bound to the receptor. However, in cases of peptides with unmodified CONH peptide bonds *cis*-conformations are only very rarely observed.

Generally, it is recommended that proof of conformational homogeneity, i.e. the dominance of a single or a few conformation(s) under given conditions, be obtained before beginning a detailed NMR analysis. Criteria for preferred conformations are:

- Large chemical shift dispersion within the set of $H^N$ and $H^\alpha$ signals.
- Large shift difference between diastereotopic protons such as $H^\alpha$ protons of Gly, or of $H^\beta$ protons in the side-chains of Phe, Tyr, His, Trp, Ser, Cys and the two $\beta$ methyl groups of Val, for example.
- Strong differences in $H^N$–$H^\alpha$ coupling constants of different residues and $H^\alpha$–$H^{\beta proR}$/$H^\alpha$–$H^{\beta proS}$ coupling constants in each side-chain.
- Pronounced differences in NOE intensities.
- Appearance of long-range NOEs between protons of non-neighboured residues.

Only if these criteria are met is it worthwhile initiating a careful conformational analysis. In general, conformational restraints, such as cyclization, binding to a receptor or complexation with metal ions are necessary to fulfil these conditions.

If the NMR data indicate a flexible structure, a structural discussion is not meaningful since the bioactive conformation of interest, i.e. the conformation the peptide bound at the receptor, is selected out of a large ensemble of alternative conformations.

## Assignment of signals

A prerequisite for the extraction of conformational parameters is the assignment of each signal in the spectrum to a specific spin system, and the assignment of these spin systems to specific residues in the peptide chain. The pulse sequences discussed here are shown in **Figure 2**.

### Assignment of spin systems

The COSY (correlation spectroscopy) experiment yields information about connectivity between nuclei. COSY cross peaks can be expected for each resolved scalar coupling between nuclei that are connected by two or three bonds. An unambiguous identification of individual amino acid spin systems can be complicated by the overlap of signals in the vicinity of the diagonal of spectrum, overlap of cross peaks for the long side-chains of Arg, Lys, Pro and Leu and the frequently insufficient signal intensities of resonances that are coupled to many neighbouring nuclei (e.g. $H^\gamma$ of Leu couples to eight vicinal neighbours). A list of other pulse techniques used in structural studies of peptides is given in **Table 1**.

The TOCSY experiment is the most efficient way to obtain complete assignments of spin systems (**Figure 3**). TOCSY is often called HOHAHA.

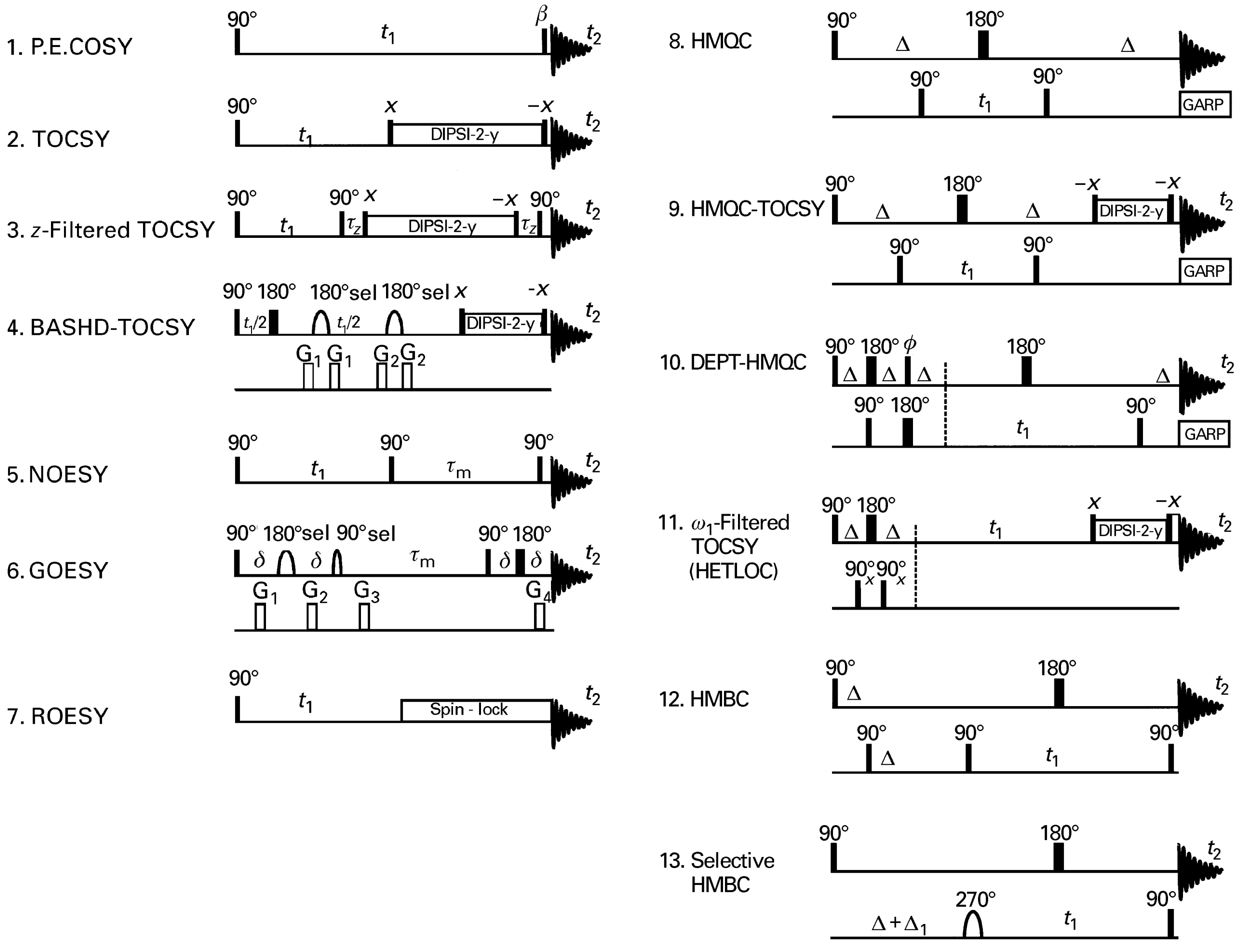



**Figure 2** Pulse sequences used in this article. All experiments can be acquired in the phase-sensitive mode, using either TPPI or the States method. Sequences 12 and 13 are best displayed in the magnitude mode because of phase modulation owing to homonuclear couplings in $F_2$. All inverse correlations, except for sequences 12 and 13, can be preceded by a BIRD (bilinear rotating decoupling) pulse sandwich to allow for a fast repetition rate of scans. For simplicity gradients are not given here in most cases.

The duration of the mixing period determines the efficiency of transfer. If a sufficiently long mixing time is chosen, correlations of the whole proton spin systems are found. TOCSY experiments with short mixing times reveal mainly correlations between directly coupled nuclei.

The spin systems of the naturally occurring amino acid residues can be divided into two groups. While the side-chains of Ala, Arg, Gln, Glu, Gly, Ile, Leu, Lys, Met, Thr and Val exhibit unique spin systems and therefore can be identified with relative ease, the residues in the second group, including Asn, Asp, Cys, His, Phe, Ser, Trp and Tyr, all have similar AMXY proton spin systems. Nevertheless, the aromatic residues of the second group can be unambiguously assigned using $H^{\beta}$–H (ring) NOE cross peaks, whereas the Ser-spin system can be easily distinguished from all other possible AMXY spin systems in COSY experiments because of its weak $H^{\beta}$–$H^{\beta'}$ coupling. Trp is easy to identify via a heteronuclear HMQC experiment owing to its characteristic downfield shift of the $\beta$ carbon signal.

Heteronuclear spectroscopy also reduces problems with signal overlap because of the large chemical shift dispersion of $^{13}C$ nuclei. The most popular heteronuclear correlation experiments for this purpose are HMQC and HMQC-TOCSY experiments (e.g. **Figure 4**). The latter contains information similar to the homonuclear $^{1}H$-TOCSY, but, as for all heteronuclear experiments, its sensitivity is lower owing to the lower gyromagnetic ratio and low natural abundance of the $^{13}C$ nucleus.

HMQC sequences can be combined with DEPT editing, allowing for the editing of multiplicities in heteronuclear correlations. The resulting experiments, such as HDQC (heteronuclear double quantum correlation), HTQC (heteronuclear triple quantum correlation) and HQQC (heteronuclear

**Table 1** Pulse techniques necessary for structural studies of peptides

| *Technique* | *Purpose* |
|---|---|
| P.E.COSY | COSY with simplified multiplet structure. P.E.COSY allows for the accurate measurements of homonuclear coupling constants |
| TOCSY = HOHAHA | Assignment of spin systems. If a long mixing time is used, TOCSY gives total correlation between all nuclei in a spin system |
| *z*-filtered TOCSY | TOCSY sequence that leads to cross peaks with pure phases. *z*-filtered TOCSY needs long measurement times owing to the random variation of the *z*-filter in 6 to 12 steps between 110 µs and 20 ms |
| NOESY | NOESY gives distance information about nuclei that are separated by less than 500 pm in space |
| ROESY | ROESY is ideally suited for the observation of nuclear Overhauser effects for medium-sized peptides at low field strengths |
| HMQC | HMQC correlates the shifts of protons with a directly bound heteronucleus. Very sensitive |
| HMQC-TOCSY | HMQC with subsequent TOCSY transfer to coupled protons |
| DEPT-HMQC | DEPT-edited HMQC, which allows for the distinction of CH, $CH_2$ and $CH_3$ groups. Exclusive selection of these multiplicities is possible with the related HDQC, HTQC and HQQC techniques |
| $\omega_1$- filtered TOCSY = HETLOC | Extraction of coupling constants to proton-bearing heteronuclei. Because magnetization is distributed among a large number of spins, this method is rather insensitive |
| HQQC | Assignment of methyl groups in crowded spectra, when folding is not feasible |
| HMBC | Assignment of carbons and protons and determination of long-range coupling constants |
| Selective HMBC | Useful variation of HMBC. For peptides, the selective pulse is usually applied to the carbonyl carbons |

quadruple quantum correlation, make it possible to exclusively excite CH, $CH_2$ or $CH_3$ groups. This results in a simplified assignment procedure.

An alternative way of overcoming problems with overlapping resonances in crowded spectral regions is to apply band-selective excitations. Band-selective pulses can be used to selectively excite a desired spectral region in one or more dimensions. The reduction of spectral width in one or more dimensions improves the digital resolution attainable in the chosen dimension, and thus helps to reduce ambiguities in the resonance assignment procedure. As a welcome side-effect it also shortens the measuring time of the experiment.

Resolution can further be improved substantially by semi-selective homonuclear decoupling during both the acquisition and the evolution dimensions. This can be achieved in the acquisition dimension by use of homonuclear shaped pulse decoupling in combination with the time-shared decoupling mode during data acquisition and in the evolution dimension by application of a semi-selective refocusing pulse together with a non-selective refocusing pulse in the centre of the evolution period.

An example of an experiment implementing these techniques is the BASHD- (band selective homonuclear decoupled) TOCSY experiment. Band selection in the evolution dimension is achieved by the excitation sculpting method. The key element of this method is a double pulse field gradient spin echo (DPFGSE) that leads to pure phase spectra with flat baselines. This cluster of pulses rephases only the selected magnetization affected by the 180° pulses and avoids any evolution of the *J*-coupling during this period. The combination of selective pulses and pulse field gradients to select the desired coherence pathway results in pure phase spectra largely devoid of artefacts. This principle can also be extended to any existing homonuclear and heteronuclear selective NMR experiment, as demonstrated by semi-selective 2D TOCSY, ROESY, HSQC and HSQC-TOCSY experiments.

## Sequential assignment

Sequential assignment of residues in a peptide chain requires correlation across the peptide bond that separates the proton spin systems of adjacent residues. This sequential information can be provided by dipolar couplings using NOESY or ROESY experiments, or by heteronuclear scalar couplings using HMBC experiments.

When only homonuclear proton experiments can be used (e.g. for reasons of sensitivity), NOESY or ROESY experiments are necessarily the method of choice. Short-range NOE signals, such as those observed between amide and aliphatic protons, are usually also observed for sequentially adjacent residues. Among these the $H^N{}_i$–$H^N{}_{i+1}$ and $H^\alpha{}_i$–$H^N{}_{i+1}$ connectivities are especially important for establishment of the sequence (see **Figure 5**). In practice, however, ambiguities can be encountered in the sequential assignment step owing to overlap of cross peaks, particularly if the peptide contains multiple residues of one type, or if long-range NOE signals are also found in the $H^\alpha$–$H^N$ region of the NOESY or ROESY spectrum, as in the case of folded peptides and small proteins. Sequential assignment of these molecules therefore can be ambiguous and requires a

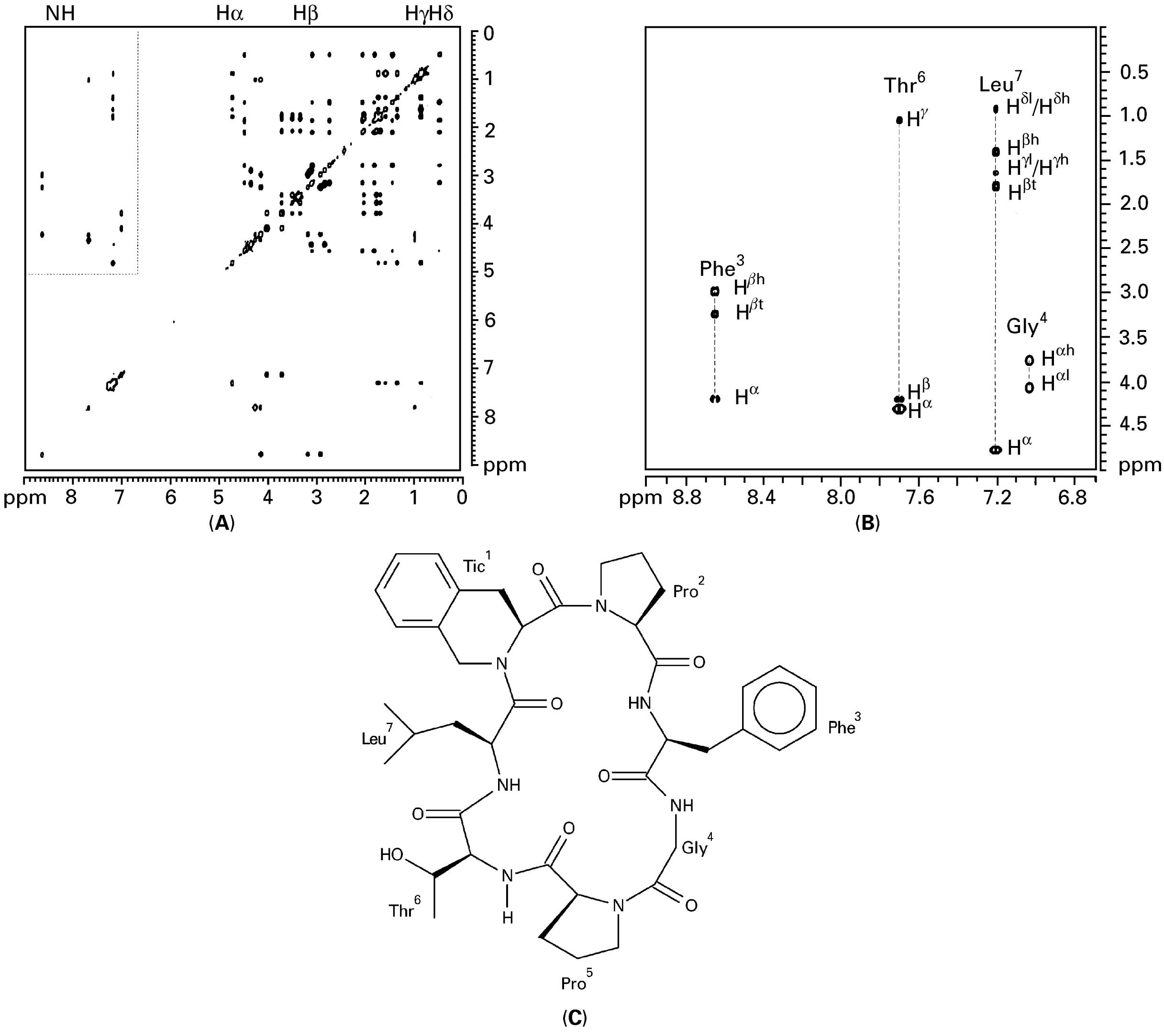



**Figure 3** (A) 500 MHz TOCSY spectrum of 20-mmol-L$^{-1}$ *cyclo* (-Tic-Pro-Phe-Gly-Pro-Pro-Thr-Leu- in DMSO at 300 K with TPPI applied in the $F_1$ dimension. Mixing time was 80 ms. A number of relayed connectivities can be observed. This spectrum is typical for small cyclic peptides. Linear peptides of this size would exhibit much less chemical shift dispersion of $H^N$ and $H^\alpha$ protons. The indicated part of the spectrum is expanded in (B) which is the fingerprint region ($F_2$: amide protons, $F_1$: aliphatic protons); coupled nuclei are connected by dashed lines and assigned to the respective residues. (C) Sequence of *cyclo* (-Tic-Pro-Phe-Gly-Pro-Pro-Thr-Leu-). Tetrahydroisochinolin (Tic) is an unnatural proline like amino acid which lacks the amide proton.

tedious and time-consuming analysis of the NOESY or ROESY spectra. In those cases, a significantly increased resolution in the $H^\alpha$–$H^N$ region of the ROESY spectrum can be achieved with a BASHD-ROESY pulse sequence, incorporating band selection and homonuclear decoupling in the $H^\alpha$ region of the spectra. Band selection in the evolution dimension is performed with the DPFGSE technique as described above.

This NOE-based approach requires previous knowledge of the peptide sequence. The known sequential position of a residue with a unique or characteristic spin system can be used as a first ‘anchor point’, from which sequencing in both directions can be carried out. In cases where the sequence is unknown, a different strategy must be used. In this case, each spin system must be assigned individually to its type of amino acid *before* a sequential assignment can be achieved. Obviously this approach is restricted to relatively small molecules (up to 30 residues).

It is also possible to determine the sequence of a peptide using scalar coupling to the carbonyl carbon. This is best done with standard or selective HMBC experiments, where only carbonyl carbons are excited and indirectly detected during $t_1$. The sequential assignment is then unambiguous and does not require any knowledge about the conformation of the peptide (**Figure 6**). A clear distinction between the

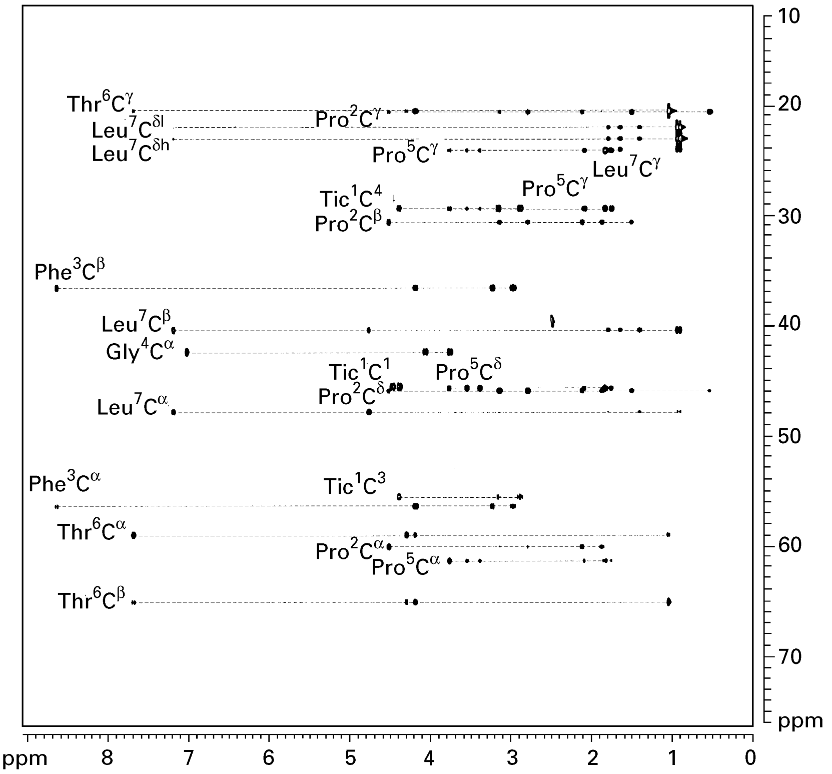



**Figure 4** 500 MHz HMQC-TOCSY spectrum of *cyclo* (-Tic-Pro-Phe-Gly-Pro-Pro-Thr-Leu-) in DMSO at 300 K. Mixing time was 80 ms. In addition to the directly bound protons the entire spin system can be observed. Coupled nuclei are connected by dashed lines and assigned to their respective residues.

proof of the complete assignment and the conformational analysis is then possible. Such a sequential assignment might also be used as an independent proof of the existence of a specific peptide bond, as for example formed by cyclization.

A complete assignment also includes the stereospecific assignment of diastereotopic groups such as methylene protons or geminal methyl groups in Val or Leu. Such additional information can significantly increase the quality of a 3D structure, which is especially important in the case of small peptides, where typically only a small number of long-range NOEs is available for the conformational analysis. Diastereotopic assignment will be discussed in more depth below.

## Extraction of conformationally relevant parameters

### NOE effects

Nuclear cross relaxation in liquids is caused by mutual spin flips in pairs of dipolar-coupled spins, induced by motional processes. Cross-relaxation efficiency depends on the spatial distances between the relaxing nuclei and leads to a transfer of magnetization between the spins. It causes intensity changes known as nuclear Overhauser effects (NOE). NOEs, as well as ROEs, can only be observed between nuclei that are separated in space by less than 500 pm. The NOE can be rationalized as heat flow from a non-equilibrium state to another neighbouring spin. In the NOE (or ROE) experiment such deviation from the Boltzmann equilibrium of spin states is created via specific pulsing and the efficiency of the heat flow to neighbouring nuclei (NOE build-up) is measured via the induced intensity changes of their NMR signals. The efficiency of dipolar relaxation is a function of the field strength (represented by $\omega_0$) and the motion (rate of relaxation referred to the external magnetic field) of the molecule, described by the molecular correlation time $\tau_c$.

The intensity of a cross peak appearing in a NOESY spectrum contains information about the relative distances between the two nuclei that contribute to the cross peak. It can be shown that the

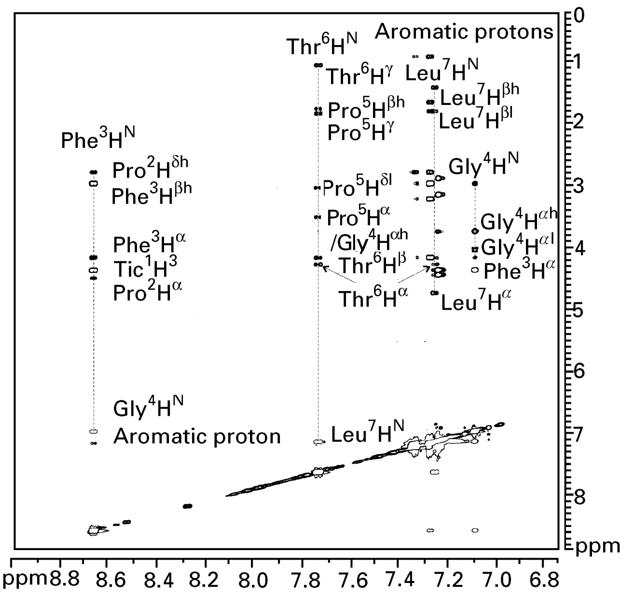



**Figure 5** Part of the 500 MHz ROESY spectrum of *cyclo* (-Tic-Pro-Phe-Gly-Pro-Pro-Thr-Leu-) in DMSO at 300 K. Correlations between the amide protons and $H^\alpha$ protons of one residue and the $H^\alpha$ protons of the preceding residue are essential for the sequential assignment. Cross peaks between neighbouring amide protons are an important source of sequential information. Cross peaks not assigned belong to aromatic protons.

cross-relaxation rate $\sigma$ that determines the NOE transfer is obtained from Equation [1].

$$\sigma = W_2 - W_0 = \frac{\gamma^4 h^2 \tau^c}{4\pi^2 10 r^6}\left(\frac{6}{1+4\omega^2\tau_c^{\,2}} - 1\right) \qquad [1]$$

where $r$ is the internuclear distance and $W_2$ and $W_0$ are the transition probabilities for the double-quantum and zero-quantum transition respectively. At a given $\omega$ the variables that determine the size of $\sigma$ are the correlation time $\tau_c$ and the interproton distance $r^6_{ij}$. It should be noted that for specific combinations of $\tau_c$ and $\omega$ the second term becomes zero or negative. The build up of cross peak intensity in a multispin system is given by $A(\tau_m) = \exp\{-R\tau_m\}$:

$$A_{ij}(\tau_m) = \left\{-R_{ij}\,\tau_m + \frac{1}{2}\sum R_{ik}R_{ij}\,\tau_m^2 + \ldots\right\}A_{ii}(0) \qquad [2]$$

where $A(t_m)$ is the cross peak intensity as a function of the mixing time $\tau_m$, $R$ the relaxation matrix and $R_{ij}$ the relaxation rate between spins $i$ and $j$. For sufficiently short mixing times the quadratic term and those of higher order in $\tau_m$ can be ignored. The cross peak intensity is then directly proportional to the cross-relaxation rate and thus to the inverse sixth power of the distance between the nuclei (when the molecule behaves as a rigid body, $\tau_c$ = constant):

$$A_{\mathrm{NOE}} \propto \sigma_{ij} \propto \frac{1}{r_{ij}^6} \qquad [3]$$

The intensities of NOE effects depend on the size of $W_2$ and $W_0$. As can be seen for Equation [1] the NOE vanishes when $W_2 = W_0$, which occurs approximately when the inverse correlation time $\tau_c^{-1}$ is of the order of the Larmor frequency $\omega_0$.

Similar relationships can be derived for ROESY. In this case, the cross peaks are generated by cross relaxation of transverse magnetization. In the rotating frame, $\sigma$ is given by Equation [4]

$$\sigma = u_2 - u_0 = \frac{\gamma^4 h^2 \tau}{4\pi^2 10 r^6}\left(\frac{3}{1+4\omega^2\tau^2} + 2\right) \qquad [4]$$

where $u_2$ and $u_0$ are the transition probabilities for the double-quantum and zero-quantum transitions in the rotating frame, respectively.

It is important to note that ROE effects, in contrast to NOE effects, are always positive and never vanish.

NOESY as well as ROESY experiments can both provide distance information. However, there are some important differences in their application and in the evaluation of the resulting spectra.

The usefulness of either of these techniques depend strongly on the time-scale of the motional processes that cause the cross relaxation. We have to distinguish three cases: (a) the fast-motion limit (extreme narrowing limit) with a short correlation time $\tau_c \ll \omega_0^{-1}$ (positive NOEs) (**Figure** 7). This applies for small molecules in non-viscous solutions. (b) The slow-motion limit (spin-diffusion limit) with a long correlation time $\tau_c \gg \omega_0^{-1}$ (negative NOEs) which applies to large macromolecules such as proteins at the maximum currently used magnetic field strengths. (c) For intermediate sized molecules only small, or even no NOE effects at all are observed. This is the case for peptides with a relative molecular mass of 500 to 1000 Da at resonance frequencies $\omega_0$ of 300 MHz. Differences of internal mobility, for example via the rotation of side-chains, can then lead to the appearance of both positive and negative NOEs in the same NOESY spectrum, making it impossible to evaluate molecular distances from these data. If only small NOE effects are observed, the ROESY techniques should be used.

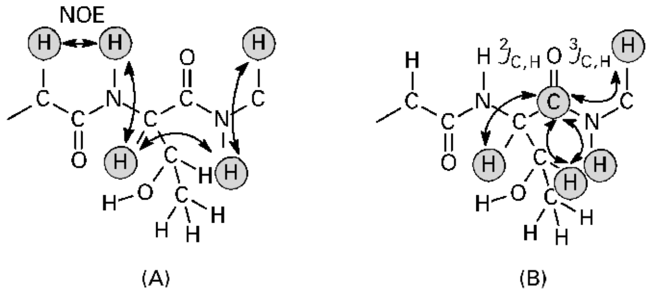



**Figure 6** Polypeptide segment with the possible interresidue connectivities between spin systems. (A) Sequential short distance NOEs for peptides. (B) Observable through-bond couplings, useful for the sequencing of peptides.

Two problems have to be considered in the evaluation of ROESY spectra. First, the offset dependence of the spin-lock field introduces intensity variations into the spectrum. Peak intensities will have to be corrected to take this effect into account when distances are to be calculated. The cross peak intensity as a function of offset from the transmitter (when NOE effects are neglected) is given by Equation [5].

$$A(\gamma B_1) = A\ \sin^2\theta_k\ \sin^2\theta_l \qquad [5]$$

where $\theta_{k,l}$= arctan $(\gamma B_l/\Omega_{k,l})$ and $\Omega_{k,l}$ are the offsets of spins $k$ and $l$ from the transmitter. The use of the compensated ROESY sequence leads to a higher intensity for peaks at the edge of the spectrum compared with the standard ROESY. In this case the peak intensity is given by Equation [6].

$$A(\gamma B_1) = A\ \sin\theta_k\ \sin\theta_l \qquad [6]$$

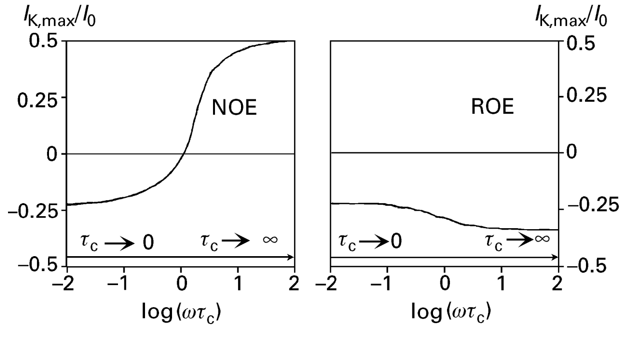



**Figure 7** Dependence of the maximum NOE and ROE cross peak intensities $I_{k,max}$ (standardized on a diagonal signal $I_0$) on $\omega$ and $\tau_c$ for very short mixing times.

Second, undesired TOCSY peaks appear because some nuclei that are spin coupled experience similar fields during the application of the spin-lock and fulfil the Hartmann–Hahn condition. Since the TOCSY peaks are phase shifted by 180° with respect to the ROESY peaks, they can easily be recognized. However, the superposition of contributions from direct and indirect transfer results in a decrease of cross peak intensity and therefore in distances which are too long. When only lower boundaries are used as restraints in MD calculations this would lead to lower restraints and a less well-defined structure but would not induce wrong results. In addition, different internal correlation times, such as the above-mentioned different flexibility of the molecule have a smaller influence in ROESY than in NOESY spectra.

NOESY spectra are preferred in the slow-motion limit but never near the transition from positive to negative NOEs ($W_2 \approx W_0$, $\sigma \to 0$) because the different internal mobility induces larger errors in distances. In such cases, lowering the temperature (to slow down molecular rotation) is recommended.

## Evaluation of NOESY and ROESY spectra

Dipolar cross-relaxation rates, and thus distances, can be determined through NOESY or ROESY experiments using various approaches. The measurement of build-up rates involves the recording of several NOESY spectra with different mixing times. To ensure equal conditions, the measurements should be made in succession. The integrals of cross peaks are determined, and the volumes are plotted as a function of mixing time:

$$A(\tau_m) = 1 - \exp(-\sigma\tau_m) \qquad [7]$$

The derivative of Equation [7] at an extrapolated mixing time of zero yields the rate of build-up of the cross peak. This initial build-up rate is directly

proportional to the cross-relaxation rate.

$$\mathrm{d}A\frac{(\tau_{\mathrm{m}}=0)}{\mathrm{d}\tau_{\mathrm{m}}}=\sigma \qquad [8]$$

The simplest and most common approach is the measurement of a single NOESY or ROESY spectrum with a short mixing time. At short mixing times the NOE build-up is in the linear range. Under this condition, it can be assumed that only direct enhancements $\sigma_{ij}$ contribute to the cross peak intensity $a_{ij}(\tau_{\mathrm{m}})$. The evaluation of a single NOESY spectrum can be done by either integration of the cross peaks or in a more qualitative manner by visual inspection of the spectrum. The second approach is often used in the case of NOESY spectra of proteins where an insufficient signal-to-noise ratio and extensive overlap prevents the accurate integration of cross peaks. In large molecules 'spin diffusion', i.e. a rapid flow of magnetization from one nucleus via another nucleus to a third one, is the most severe problem. Only very short mixing times can be used and a complete treatment of the relaxation matrices is recommended. In the first approximation the NOEs are only used qualitatively in proteins: cross peaks are then classified according to several semi-quantitative categories – usually strong, medium, weak – which correspond to distance ranges. This approach is not recommended for peptides since the integration of cross peaks should lead to considerably more accurate distances and spin diffusion is not so efficient in peptides as it is in large molecules. The so-called ISPA (isolated spin pair approximation) is closer to reality for molecules which are close to the $\tau_{\mathrm{c}} \approx \omega_0$ condition.

Owing to the fact that absolute values of correlation times are usually not available, interproton distances cannot be directly calculated. Distances are instead obtained by calibration of the cross peak intensities against an internal distance standard, usually the distance between diastereotopic geminal protons (178 pm) or aromatic protons of Tyr (242 pm). Assuming isotropic tumbling and rigid-body model for all parts of the molecule, Equation [9] is then used to calculate all interproton distances:

$$r_{ij}=r_{\mathrm{ref}}\left(\frac{a_{\mathrm{ref}}}{a_{ij}}\right)^{-6} \qquad [9]$$

Usually, the NOE enhancements for the structural and conformational analysis of peptides are extracted from 2D NOESY spectra. The GOESY experiment, a 1D version of the NOESY experiment, uses selective excitation of separated signals and yields accurate measurements even for tiny enhancements. The DPFGSE NOE technique (see above) achieves better sensitivity by not discarding one of the coherence transfer pathways (in contrast to the GOESY technique), while spectra have the same characteristics as GOESY spectra. Therefore, the DPFGSE NOE technique is to be preferred.

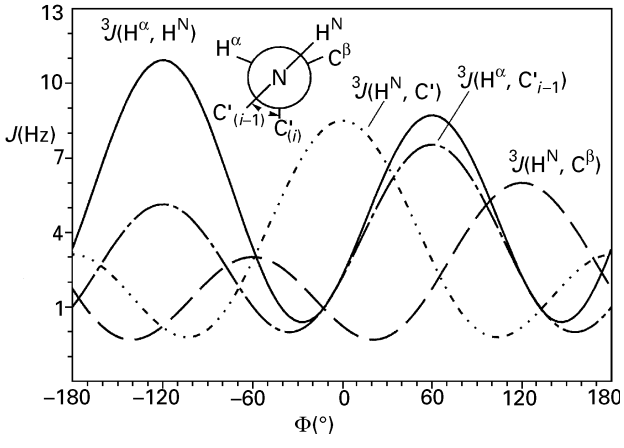



**Figure 8** The Karplus curve for four coupling constants about the $\phi$ dihedral angle of an L-amino acid. There are four possible dihedral angles for a given coupling constant. Utilizing a combination of all four coupling constants it is usually possible to narrow down the choice to a single angle. Reproduced with permission of Wiley-VCH from Eberstadt *et al* (1995) *Angewandte Chemie*, *International Edition in English* **34**: 1671–1695.

### Determination of coupling constants

Many $J$ coupling constants illustrate a clear dependency on dihedral angles and therefore are an important source of conformational information. This relationship is particularly distinct for $^3J$ couplings. The model for the relationship between bond angles and the coupling constant most often used is that proposed by Karplus (**Figure 8**).

$$^3J=A\ \cos^2\theta+B\ \cos\theta+C \qquad [10]$$

The equation holds for almost all coupling constants (**Table 2**). Only the coefficients $A$, $B$ and $C$ have to be adjusted, depending on the type of the two coupled nuclei and their environment. However, even if the coefficients have been determined, the multiple angles that fulfil Equation [10] (up to four values can be obtained) remain a problem.

Despite these problems, the application of coupling constants, in addition to proton–proton distances, in modern conformational analysis of peptides is indispensable. In principle, the coupling constants shown in **Table 2** can be used to determine the $\phi$, $\psi$ and $\chi_1$ angles.

For the determination of homonuclear or heteronuclear three-bond coupling constants, three fundamentally different approaches are used: (1) direct measurements of splittings caused by $J$-couplings, (2) the so-called E.COSY-signal patterns in which the splittings can be measured as shift differences of signals and (3) special experiments which lead to modulation of signal intensities via $J$ (**Table 3**). A frequent requirement for the latter is a $^{15}N$-enriched sample, e.g. for the $J$-modulated ($^{15}N$, $^{1}H$)-COSY experiment or the HNHA experiment. Whereas labelled compounds are routinely used for proteins and nucleic acids they are expensive for peptides and therefore rarely used.

**Determination of coupling constants from the shape of the signal.** All these techniques have to correct or compensate for the partial overlap of multiplet lines by either using additional parameters which depend on the shape of the signals or by fitting a model to the overlapped experimental signal. However, this procedure requires that the line width is still smaller than the coupling constant, and furthermore that the signals have a good signal-to-noise ratio.

The determination of coupling constants from an antiphase (e.g. COSY) cross peak yields values which are inherently too large owing to the reciprocal signal cancellation of the antiphase pattern. Kim and Prestegard have described an especially simple method for the determination of coupling constants for AX spectra. The splitting of the maxima in the absorptive and in the dispersive signal is measured and the $J$ coupling is calculated via an extensive cubic equation (**Figure 9**). Only two calculations are required for the COSY spectra in which the phase are 90° shifted. This procedure is especially useful for the determination of coupling constants that cannot easily be extracted from peaks in E.COSY spectra (see below), e.g. for $H^N$–$H^\alpha$ cross peaks. No additional spectrum has to be recorded.

The simulation of the line shape using a linear combination of reference signals is based on the simulation of a complex nonresolved multiplet, starting out from a library of experimental multiplets. This technique has found widespread application in the determination of heteronuclear long-range coupling constants from HMBC cross peaks of peptide samples with $^{13}C$ in natural abundance. The major advantage is the relatively high sensitivity of the HMBC experiment because of inverse detection, as well as the fact that the identity of heteronuclear couplings directly ensues from the assignment of the 2D cross signals. In principle, long-range coupling constants can be determined directly from the cross peaks which represent the active coupling. However, since the long-range heteronuclear coupling constants are approximately of the same magnitude as $^{2}J_{H,H}$ and $^{3}J_{H,H}$ coupling constants (1–10 Hz), the rather small heteronuclear antiphase coupling constant $^{n}J_{C,H}$ cannot be read directly because of overlapping and reciprocal cancellation of the numerous multiplet lines. Keeler *et al.* have developed an elegant procedure which allows the determination of the heteronuclear long-range coupling

**Table 2** Coupling constants used to determine $\phi$, $\psi$ and $\chi$, angles in peptides

| *Angle* | *Coupling constant* |
|---|---|
| $\phi$ | $^{3}J_{HN,H^\alpha}$, $^{3}J_{C^\beta,HN}$ $^{3}J_{H^\alpha,CO(i-1)}$ |
| $\psi$ | $^{3}J_{H^\alpha N(Ii-1)}$ |
| $\phi$ and $\psi$ | $^{1}J_{C^\alpha,H^\alpha}$ |
| $\chi_1$ | $^{3}J_{H^\alpha,H^\beta}$, $J_{CO,H^\beta}$, $J_{N,H^\beta}$ |

**Table 3** Experiments that are used for determining the most important coupling constants in peptides

| *Coupling* | *Technique* |
|---|---|
| $H^N$–$H^\alpha$ | Direct reading or Kim–Prestegard method |
| $H^\alpha$–$H^\beta$ | E.COSY techniques |
| Couplings within the proline pyrrolidine ring | E.COSY techniques |
| $^{13}CO$–$H^\beta$ | HMBC (qualitative), Keeler method |
| $H^N$–$C^\beta$ | HETLOC |

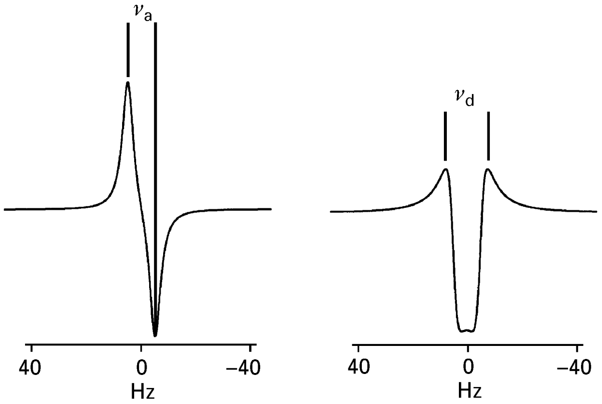



**Figure 9** Determination of the separation of the signal maxima of an antiphase doublet for the absorptive ($\nu_\alpha$) and dispersive ($\nu_d$) component. Based on these two parameters the coupling constant $J$ can be calculated. Reproduced with permission of Wiley-VCH from Eberstadt *et al* (1995) *Angewandte Chemie, International Edition in English* **34**: 1671–1695.

constant from the line shape even in such cases (**Figure 10**). The difference between a cross section in a HMBC spectrum and the corresponding, reconstructed proton multiplet (e.g. from a 1D spectrum) is the heteronuclear coupling constant of interest. In practice, the proton multiplet will show some overlap in the 1D spectrum, and therefore a TOCSY spectrum with a pure phase (e.g. with a *z*-filter) has to be recorded. The homonuclear reference multiplet now differs from the HMBC multiplet only by the absence of the heteronuclear antiphase coupling and the signal amplitude. This can be simulated by scaling of the reference signal in the time domain with the term $\sin(\pi J_{trial} t)$. The amplitude and $J_{trial}$ can now be varied by a nonlinear optimization until the deviation between the two spectra reaches a minimum. At this point $J_{trial}$ should be equal to the coupling constant of interest.

This method has the advantage that a large number of coupling constants can be determined from a single HMBC spectrum, which usually contains a large number of long-range correlations. However, the quality of the heteronuclear spectrum and especially the signal-to-noise ratio is of crucial importance for the convergence of the optimization. Therefore, the method is time consuming with respect to both recording of the spectra as well as their processing, but it yields accurate values for the coupling constants between heteronuclei.

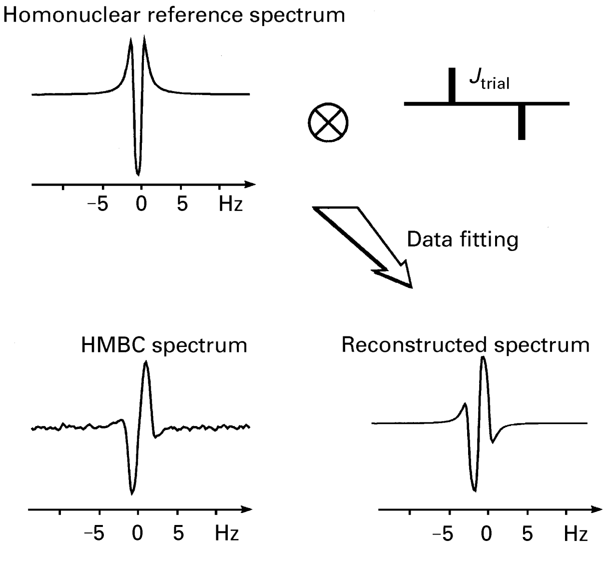


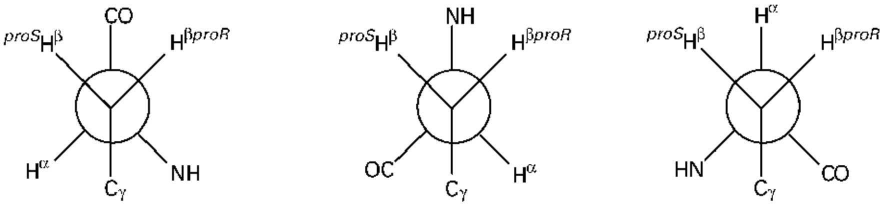

**Figure 10** Schematic of the Titman–Keeler method for the evaluation of heteronuclear coupling constants from HMBC spectra. The synthetic spectrum is computed from the homonuclear reference spectrum and a chosen heteronuclear coupling constant $J_{trial}$. This spectrum is then compared with the actual HMBC spectrum and $J_{trial}$ is iteratively varied until a good fit is obtained. Reproduced with permission of Wiley-VCH from Kessler H and Seip S (1994) In: Croasmun WR and Carlson RMK (eds) *Two-Dimensional NMR Spectroscopy*, pp 619–654.

Assuming that the staggered rotamers are predominantly populated (**Figure 11**), qualitative considerations together with accurately determined homonuclear coupling constants are often sufficient for the diastereotopic assignment of methylene protons. The $\chi_1$ angle can be set to −60° if the two $^3J_{H^\alpha,H^\beta}$ and $^3J_{H^\alpha,H^{\beta'}}$ coupling constants are small (both ~3 Hz). If one strong and one weak coupling is observed, $\chi_1$ can be either 60 or 180°. To differentiate these two cases, stereospecific assignment of the $H^\beta$ protons is required. This is possible with the aid of qualitative heteronuclear *J*-couplings (e.g. between $^{13}CO$ and $H^\beta$ or $^{15}N$ and $H^\beta$) and NOE or ROE cross peak intensities to the different $H^\beta$ protons.

**Determination of coupling constants by the E.COSY principle** The E.COSY (exclusive correlation spectroscopy) principle yields a simplified cross peak multiplet, since only the 'connected transitions' are excited. This means that the signal intensity in an A, M cross peak from a three-spin system AMX can only be found in those parts of the multiplet pattern where the spin states of the third nucleus X have been conserved. To obtain such an E.COSY pattern, a mixing of spin states of the X nucleus (e.g. by the application of a non-sensitive 90° pulse) must be avoided. The coupling between M and X can then be extracted from the passive coupling of the A, M cross peak as the shift of two in-phase multiplets, which are separated in the indirect dimension and, therefore, have no interfering influence on each other. The only requirement is that the splitting in $F_1$ (the coupling between A and M) must be larger than the line width (**Figure 12**).

The E.COSY technique with the highest sensitivity is P.E.COSY (primitive E.COSY), where the retention of the spin states of the passive spin is achieved using a small flip angle of the mixing pulse and subtraction of the dispersive diagonal via a reference spectrum. The resulting cross peaks contain strong signal intensities for connected transitions but vanishing intensities for non-connected transitions.

In heteronuclear spectroscopy E.COSY patterns can be easily obtained if no 90° pulse is applied to the heteronucleus (i.e. the states $\alpha$ and $\beta$ are not mixed). $\omega_1$-Hetero-filtered (HETLOC) experiments are the method of choice for the determination of long-range coupling constants between protons and $^{13}C$ or $^{15}N$ nuclei in natural abundance that carry a directly connected proton (**Figure 13**). For the determination of the $\chi_1$ angle the coupling $^3J_{H^\beta}$ can be determined by using a heteronuclear $\omega_1$-half-filter (X-half-filter) at the beginning of a NOESY or TOCSY experiment (before the $t_1$ delay). Only protons which are directly coupled to the magnetically active




| [1] | [2] | [3] |
|---|---|---|
| $\chi_1 = -60°$ | $\chi_1 = 180°$ | $\chi_1 = +60°$ |
| $^3J(H^\alpha, H^{\beta proR})$ large | $^3J(H^\alpha, H^{\beta proR})$ small | $^3J(H^\alpha, H^{\beta proR})$ small |
| $^3J(H^\alpha, H^{\beta proS})$ small | $^3J(H^\alpha, H^{\beta proS})$ large | $^3J(H^\alpha, H^{\beta proS})$ small |
| NOE $(H^N, H^{\beta proR})$ strong | NOE $(H^N, H^{\beta proR})$ strong | NOE$(H^N, H^{\beta proR})$ weak |
| NOE $(H^N, H^{\beta proS})$ weak | NOE $(H^N, H^{\beta proS})$ strong | NOE $(H^N, H^{\beta proS})$ strong |
| $^3J(CO, H^{\beta proR})$ small | $^3J(CO, H^{\beta proR})$ large | $^3J(CO, H^{\beta proR})$ small |
| $^3J(CO, H^{\beta proS})$ small | $^3J(CO, H^{\beta proS})$ small | $^3J(CO, H^{\beta proS})$ large |

**Figure 11** The three staggered conformers about the $\chi_1$ angle of residues with $\beta$-methylene protons (see text for details).

heteronucleus ($^{13}C$ or $^{15}N$) are selected, while $^{12}CH$ or $^{14}NH$ protons are suppressed. Obviously, no heteronuclear decoupling can be performed during the acquisition. The delay $\Delta$, is adjusted according to the value of the heteronuclear coupling constant for the $\omega_1$-half-filter ($\Delta = 1/2J_{H,X}$) resulting in in-phase magnetization at the end of the two delays. The subsequent TOCSY sequence affects only the $^{13}C$-coupled protons and transfers the magnetization through the entire spin system. In $\omega_1$ the signals are split by the large value of the $^1J_{X,H}$ coupling (e.g. $^1J_{N,H} = 90$ Hz) in $\omega_2$ by the desired long-range heteronuclear coupling constant. However, by this method heteronuclear $^nJ_{H,X}$ couplings can only be determined for heteroatoms which bear a directly bound proton. The latter causes the required large splitting in $\omega_1$ via $^1J_{X,H}$. Fortunately, using the sensitivity of an inverse detection experiment many heteronuclear coupling constants can be determined from a single spectrum for molecules in natural isotopic abundances.

## Structure determination of peptides

The utilization of NMR data for the determination of the three-dimensional structure of peptides involves the use of computer simulations. The methods can be broken down into two general categories: molecular mechanics/dynamics (MM or MD) and distance geometry (DG) calculations.

MM and MD use a force field to describe the molecule and estimate the potential energy of the given conformation. The standard force field contains a term for distortion of bond lengths, bond angles and dihedral angles plus non-bonded terms for Coulombic interactions and a Lennard-Jones description of the attraction/repulsion of atoms. The application of experimental restraints is achieved by simply introducing an additional term, a so-called penalty function. This penalty function serves to minimize the differences between the calculated values and experimental data.

The second general approach, DG calculations, utilizes a description of a molecule based solely on distances. Bond lengths, bond angles and torsion angles are converted into ranges of allowed distances according to the molecular constitution. Distances which satisfy these ranges are chosen randomly to create a distance matrix. The diagonalization of this matrix then produces Cartesian coordinates. The experimental distances are compared with the distances generated from DG calculations based on the covalent structure; if the experimental distances are tighter, they replace those on the consideration of the covalent geometry.

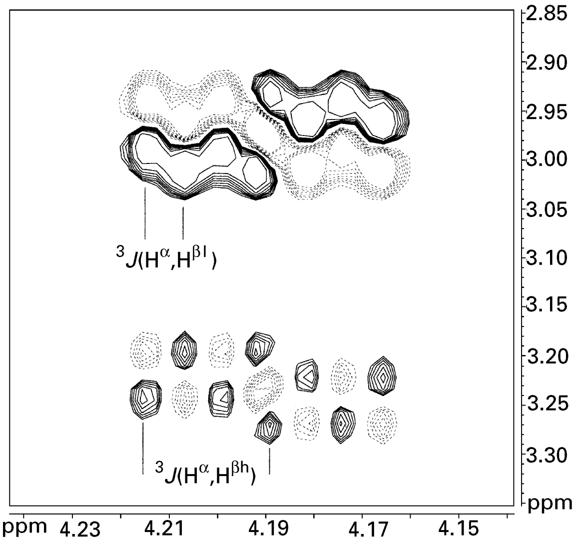



**Figure 12** $H^{\alpha}$, $H^{\beta}$ region of the 500 MHz P.E.COSY spectrum of cyclo(-Tic-Pro-Phe-Gly-Pro-Pro-Thr-Leu-) in DMSO at 300 K. The enlarged view is of the $Phe^{3}$ $H^{\alpha}$, $H^{\beta}$ cross peaks. The displacement of the multiplet patterns can be used to determine the passive couplings with high accuracy. $H^{\beta l}$ means the low field $\beta$ proton and $H^{\beta h}$ the high field proton.

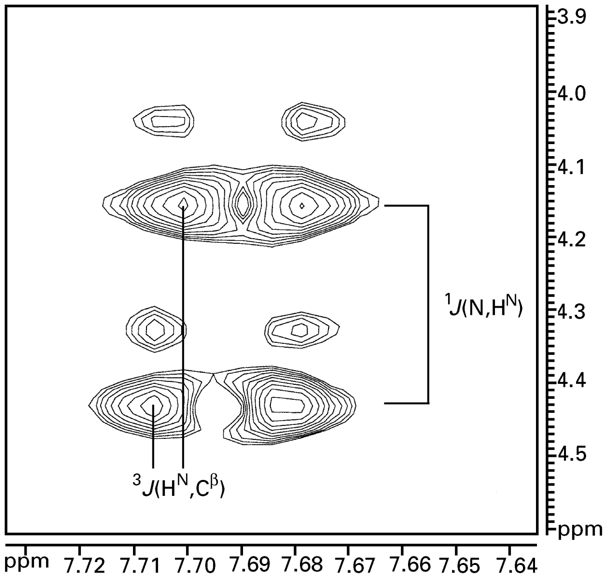



**Figure 13** Part of the 500 MHz HETLOC ($\omega_1$-filtered TOCSY) spectrum of *cyclo*(-Tic-Pro-Phe-Gly-Pro-Pro-Thr-Leu-) in DMSO at 300 K. The enlarged view is of the $Thr^{7}$ $H^{N}$, $H^{\beta}$ cross peak. The coupling constant is extracted from the separation of the cross peak components in the better resolved acquisition dimension ($F^{2}$).

The most important parameters for the determination of a three-dimensional structure are NOE-derived distances, bond angles derived from *J*-coupling constants and temperature dependences of $H^{N}$ proton chemical shifts. Distances from NOESY or ROESY spectra are directly used as restraints for the calculations. Recently, coupling constants have also been used in computational structure refinement. The penalty function employed in this case is directly based on the Karplus equation. If more than one coupling for a single dihedral angle is available the restraints from coupling constants are quite useful at the level of structure refinement. A limited number of experimental NOE values for peptides normally requires special care. In small peptides more or less all of the protons are on the surface, in contact with the solvent, and thus there are only distances to one side whereas in proteins there are many protons in the core region that are completely surrounded by other protons. Hence, for peptides it is indispensable to collect as much experimental data as possible. This means that accurate, quantitative NOE data as well as correct treatment of *J* coupling constants are required.

The direct application of the temperature coefficients in structure refinement is problematic, since, while a temperature coefficient may indicate that an $H^{N}$ proton is shielded from the solvent (for example by being involved in a hydrogen bond), it does not allow for identification of the acceptor of that hydrogen bond. Therefore, temperature coefficients, unfortunately, are frequently used only to confirm the final structure. The analysis of the radial distribution function (rdf) of the solvent around an amide proton shows, in our experience, distinct peaks when this proton is solvent exposed. The size and sharpness of these rdfs correlate directly with the size of the temperature gradient.

Conformational analysis of a peptide normally begins with the simple assumption of only a single dominating conformation using restrained MD under vacuum, beginning with various starting structures (to prevent structural bias, it is, however, recommended to begin with DG calculations to create the first crude starting structures for the MD). The resulting MD structure is further refined by recalculations of the molecule within an explicit solvent box. This can contain $H_2O$, DMSO, $CHCl_3$, $CH_3OH$ or others, depending on the solvent used for the measurement. The best procedure uses a truncated octahedron to allow periodic boundary conditions for an almost spherical box but also cubic boxes may be used. The quality of the final structure is finally checked by a long trajectory MD calculation (100 ps or more), without all experimental restraints but within the solvent. If all restraints are fulfilled (within about 10 pm), and the same result is always obtained regardless of the selected starting structure, it can be concluded that a ‘single’

satisfying conformation exists (a 'single conformation' may still include some flexibility, of the order of ±20° for torsions). There might be other conformations which also fulfil the experimental data, especially when the system is underdetermined owing to an insufficient number of constraints, but this should be apparent if different starting structures lead to different results.

If one part of the molecule turns out to be well defined while another part shows larger deviations from the experimental constraints, it can be assumed that the latter part is undergoing intramolecular motion that is fast on the chemical shift time-scale. Short distances $r$ contribute most strongly to the observed NOE intensities (because of the $r^6$ dependence), and hence not all distance constraints derived from NOEs can be fulfilled at the same time if an equilibrium among several conformations exists. Given that there is a sufficient number of restraining parameters one can try to analyse this equilibrium by using time-dependent NOEs and/or time-dependent coupling constants, making the assumption that the constraints are not fulfilled at each simulation step, but rather only over a whole trajectory. This allows for analysis not only of the flexibility but also of the detailed nature of the molecular processes involved.

In addition, ensemble calculations may be used to analyse flexible structures. However, these calculations are time consuming, difficult to analyse and cannot directly include solvents. Hence, this procedure is only used in rare cases.

Side-chain mobility is analysed mainly by assuming a rapid equilibrium between three staggered conformations. The populations of the conformations can then be derived from the homonuclear and heteronuclear coupling constants using Pachler's equation. The observed coupling $J_{obs}$ then results as the average over all three rotamers, $J_i$, weighted with their respective population $P_i$.

$$J_{obs} = \sum_i P_i J_i \quad (i = 1-3) \qquad [11]$$

$$P_{[1]} = \frac{J(H^{\alpha}, H^{\beta proR}) - J_{sc}}{J_{ap} - J_{sc}} \qquad [12]$$

$$P_{[2]} = \frac{J(H^{\beta proR}, CO) - J_{sc}}{J_{ap} - J_{sc}} \qquad [13]$$

$$P_{[2]} = \frac{J(H^{\alpha}, H^{\beta proS}) - J_{sc}}{J_{ap} - J_{sc}} \qquad [14]$$

$$P_{[3]} = \frac{J(H^{\beta proS}, CO) - J_{sc}}{J_{ap} - J_{sc}} \qquad [15]$$

For homonuclear couplings, the antiperiplanar coupling $J_{ap}$ is 13.6 Hz and the synclinal coupling $J_{sc}$ is 2.6 Hz while for heteronuclear $^1$H–$^{13}$C couplings $J_{ap}$ is 8.5 Hz and $J_{sc}$ is 1.4 Hz. It should be noted that homonuclear coupling constants and NOE effects alone do not always yield an unambiguous diastereotopic assignment of the $\beta$-methylene protons. This means that incorrect values for the dominant bond angle ($\chi_1$) may be obtained. This is especially the case if MD calculations are performed only under vacuum. In such cases, a globular structure is often predicted for the molecule, which only opens into a realistic conformation when the solvent is explicitly included in the calculations. Especially as peptides have a large surface which can interact with the solvent, it is essential to perform MD calculations in explicit solvents.

## Relaxation parameters and molecular dynamics

NMR spectroscopy is uniquely capable of comprehensively characterizing the internal motions of peptides in solution at the atomic level over time-scales ranging from picoseconds to hours. NMR techniques used for the study of dynamics include relaxation rate measurements, dynamic NMR and line shape analysis, magnetization transfer experiments, NOESY and ROSEY and amide proton exchange measurements.

For diamagnetic peptides in isotropic solvents, the primary mechanism of nuclear magnetic relaxation of protonated $^{13}$C nuclei and of $^{15}$N nuclei at natural abundance is the dipolar interaction with the directly bound protons. At high magnetic fields, chemical shift anisotropy (CSA) also contributes to the relaxation of the heteronuclei. The rates of these relaxation processes are governed by both the internal motions and the overall rotational motion of the molecule. Consequently, characterization of $^{13}$C and $^{15}$N heteronuclear relaxation can provide information about internal dynamics of peptides on time-scales faster than the rotational correlation time.

The $T_{1\rho}$ times of protons have been measured to study conformational exchange on the microsecond to millisecond time-scale. However, the complex interaction with surrounding protons, which is strongly dependent on the molecular geometry, may lead to artefacts in the interpretation of the data.

The overall tumbling rate, an important parameter for NMR spectroscopy, can best be determined by measuring $T_1$ relaxation times. To determine the overall correlation time $\tau_c$ at least two different field strengths are required. In most cases, the function ($T_1$ versus $\tau_c$) allows for two possible values of $\tau_c$. Usually the true $\tau_c$ can be selected based either on reasonable estimates as a function of relative molecular masses or after consideration of additional relaxation rates such as $T_2$ relaxation times or heteronuclear NOEs (**Figure 14**).

Specific models for internal motions can be used to interpret heteronuclear relaxation, such as restricted diffusion and site-jump models. However, model-free formal methods are preferable, at least for the initial analysis, since available experimental data generally are insufficient to completely characterize complex internal motions or to uniquely determine a specific motional model. The model-free approach of Lipari and Szabo for the analysis of relaxation data has been used for proteins and even for peptides. It attempts to reproduce relaxation rates by a weighted product of spectral density functions with different correlation times $\tau_i$. The weighting factors are identified as order parameters $S_i^2$ for the molecular rotational correlation time $\tau_c$ and optional further local correlation times $\tau_i$. The term $(1-S_i^2)$ would then be proportional to the amplitude of the corresponding internal motion. However, the Lipari–Szabo approach is based on the assumption that molecular and local correlation times are not coupled, i.e. they should be distinct enough (e.g. differing by at least a factor of 10 in time) to allow for this separation. However, in small molecules the rates of these different processes are of the same order of magnitude, and the requirements of the Lipari–Szabo approach may not be fulfilled. Molecular dynamics simulation provide a complementary approach for the interpretation of relaxation measurements.

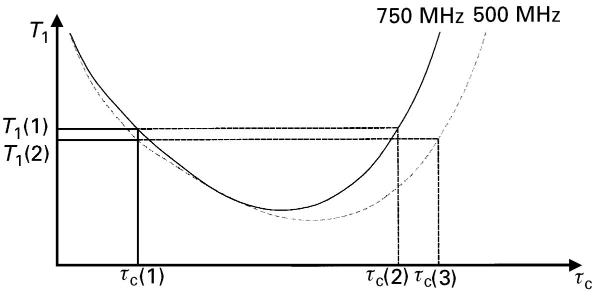



**Figure 14** Dependence of $T_1$ on $\tau_c$ at two different field strengths. The different $T_1$ times [$T_1(1)$ and $T_1(2)$] clearly correspond to $\tau_c(1)$; the alternative values for $\tau_c$ can be ruled out.

## List of symbols

$A$ = cross peak intensity; $B$ = magnetic flux density; $J$ = coupling constant; $P_i$ = population of rotamer $i$; $R_{ij}$ = relaxation rate between spins $i$ and $j$; $r$ = internuclear distance; $u_2$, $u_0$ = probability of double- and zero-quantum transitions, respectively, in the rotating frame; $W_2$, $W_0$ = transition probability for double- and zero-quantum transitions, respectively; $\gamma$ = gyromagnetic ratio, $\sigma$ = cross-relaxation rate; $\tau_1$, $\tau_2$ = correlation times; $\tau_c$ = correlation time; $\tau_m$ = mixing time; $\phi$, $\psi$, $\omega$ = peptide backbone angles; $\chi$ = bond angles of peptide side-chains; $\omega_0$ = Larmor frequency.

*See also:* **NMR Pulse Sequences; Nuclear Overhauser Effect; Proteins Studied Using NMR Spectroscopy; Solvent Suppression Methods in NMR Spectroscopy; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals; Two-Dimensional NMR Methods.**

## Further reading

Eberstadt M, Gemmecker G, Mierke DF and Kessler H (1995) Scalar coupling constants – their analysis and their application for the elucidation of structures. *Angewandte Chemie, International Edition in English* **34**: 1671–1695.

Evans JNS (1995) Biomolecular NMR Spectroscopy. Oxford: Oxford University Press.

Kessler H and Seip S (1994) NMR of Peptides. In: Croasmun WR, and Carlson RMK (eds) *Two-Dimensional NMR Spectroscopy*, pp 619–654. Weinheim: VCH.

Kessler H and Schmitt W (1996) Peptides and polypeptides. In: Grant DM and Harris RK (eds.) *Encyclopedia of Nuclear Magnetic Resonance*, pp 3527 –3537. Chichester: John Wiley & Sons.

Lipari G and Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. *Journal of the American chemical society* **104**: 4546 – 4559.

Neuhaus D and Williamson MP (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. Weinheim: VCH.

Parella T (1996) High quality 1D spectra by implementing pulsed-field gradients as the coherence pathway selection procedure. *Magnetic, Resonance in chemistry* **34**: 329–347.

van Gunsteren WF and Berendsen HJ (1990) Computer simulation of molecular dynamics: methodology, applications and perspectives in chemistry. *Angewandte Chemie, International Edition in English* **29**: 992.

Wüthrich K (1986) *NMR of Proteins and Nucleic Acids*. New York: Wiley.

# Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals

**Alexandros Makriyannis** and **Spiro Pavlopoulos**,
University of Connecticut, Storrs, CT, USA

Copyright © 1999 Academic Press

| MAGNETIC RESONANCE |
| --- |
| **Applications** |

## Scope and applications

Information on the structure of drug molecules and their interactions with their therapeutic sites of action is of critical importance in the design and development of new drugs. Of all the analytical methods, nuclear magnetic resonance spectroscopy (NMR) is the most exquisitely suited to provide such experimental results. The field of NMR is advancing continuously to include new pulse sequences and methods as well as progressively larger field instruments and improved probes. This fast progress in NMR methods and technologies has served to expand dramatically its applications in drug research. Indeed NMR, used jointly with X-ray crystallography and computational/graphical approaches, has revolutionized structure-based drug design.

Currently the availability of a plethora of multidimensional/multinuclear NMR methods allows us to extract information on the structures and dynamic behaviours of a wide range of drug molecules of up to 30 kDa in size. These include the small and medium-sized traditionally used therapeutic drugs, to higher molecular weight peptides, proteins, nucleotides, nucleic acids and polysaccharide biotechnology products.

Progressively, more detailed structural and dynamic information has become available because of our increased ability to measure more effectively the basic NMR parameters used in structural analysis, namely, proton and carbon chemical shifts, coupling constants, relaxation parameters ($T_1$, $T_2$) and the exceedingly valuable nuclear Overhauser effect. Such measurements are, in turn, used to obtain information on the three-dimensional structure of molecules as well as their conformational properties and dynamic behaviour. Additionally, the new solution NMR methods allow researchers to study the interactions of a drug molecule with its site of action on the biopolymer (enzyme, receptor, nucleic acid, etc.). Such studies can lead to insights regarding the bioactive conformation of a flexible drug molecule, which constitutes invaluable information for drug design. Here we shall discuss the most commonly exploited experiments for extracting the individual NMR parameters mentioned above, and also how such parameters are utilized to obtain structural information.

## Conformational analysis of small molecules

Information on the structural properties of small drug molecules in solution can be obtained from a number of NMR parameters including $^1$H and $^{13}$C chemical shifts, $^1$H–$^1$H and $^1$H–$^{13}$C scalar coupling constants, $^1$H nuclear Overhauser effects (NOEs), as well as relaxation measurements. Here, the conformational analysis of CP55,940 (**Figure 1**) is used to illustrate the most common experiments encountered in studying small molecules (<1000 Da) in solution. This synthetic compound is structurally related to $\Delta^9$-tetrahydrocannabinol ($\Delta^9$-THC), a psychoactive component of marijuana, and has received much attention because it was used as the high affinity radioligand during the discovery and characterization of the G-protein coupled cannabinoid receptor (CB1). The elucidation of the conformational properties of this compound and its congeners provides information on the steric requirements for a productive interaction at the cannabinoid receptor active site.

### Double quantum filtered correlation (DQF-COSY) and total correlation (TOCSY) spectroscopy

These experiments provide information on $^1$H chemical shifts and $^1$H–$^1$H scalar coupling. Spectral assignments are made initially by an analysis based on integrated peak areas and chemical shifts in the one-dimensional spectrum. Subsequently, they are speci-fically assigned by analysis of scalar or spin–spin coupling connectivities observed by $^1$H–$^1$H double quantum filtered correlation spectroscopy (DQF-COSY). This is a two-dimensional experiment where the information is spread onto a plane in which the diagonal is equivalent to the one-dimensional spectrum, and the scalar coupling is manifested as an off-diagonal crosspeak between the two

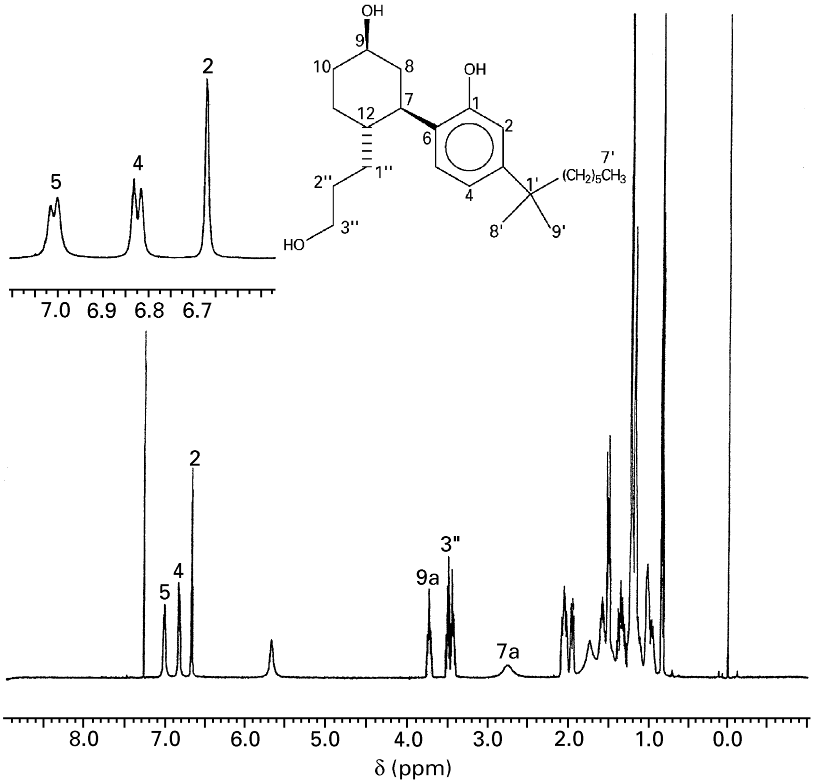



**Figure 1** The one-dimensional ¹H spectrum of CP55,940 with an expanded scale of the aromatic region.

resonances in question. Thus even though a resonance may not be visible in a one-dimensional spectrum due to overlap, its position may be identified from a crosspeak in a two-dimensional experiment. Protons that are part of a spin system often give rise to a pattern of diagonal peak–crosspeak connectivities that can be traced from a 'starting point' in the spin system to an 'end point'. It is the presence of such connectivity patterns that makes this experiment such a powerful tool in assignment.

For example, the $^{1}H$ NMR spectrum of CP55,940 is shown in **Figure 1**. A logical starting point for assignment was the resonance at δ 3.74 that was assigned to H9a because the size of the peak as measured by integration was consistent with one proton and because this aliphatic resonance is expected to be deshielded. In the DQF-COSY spectrum (**Figure 2**) vicinal coupling to H10e, H8e, H10a and H8a is observed. Assignment of H10e and H10a resonances was made with the support of the crosspeak connectivities of H9a with H10a, H10e, H8a and H8e. The DQF-COSY spectrum clearly shows three components (H10e, H8e and H11e) under the multiplet at δ 2.06, in which three related strong geminal $^{2}J$ couplings, H10e/a ($F_1$ = δ 2.09, $F_2$ = δ 1.38), H8e/a ($F_1$ = δ 2.05, $F_2$ = δ 1.53) and H11e/a ($F_1$ = δ 1.98, $F_2$ = δ 1.13) can be discerned. This is a prime example of the improvement in spectral resolution of two-dimensional experiments over one-dimensional experiments.

Complete assignment of a whole spin system may be limited because of severe spectral overlap. To overcome this, DQF-COSY data are often used in conjunction with total correlation spectroscopy (TOSCY). This experiment results in a transfer of magnetization across an entire spin system and consequently crosspeaks may be observed between each resonance of a spin system. Thus it is possible to determine whether a particular overlapped region of the spectrum contains all unidentified members of a chemical spin system.

$^{1}H$ chemical shifts and scalar coupling constants can be measured directly from one-dimensional spectra if the peaks are well resolved, or, if spectra are too complex, they may be measured from DQF-COSY spectra crosspeaks. However, such measurements are often inaccurate and so are used as a basis of simulating the observed 1D spectrum to obtain more accurate values. The measurements are used as a starting point and are systematically altered until the stimulated spectrum best matches the observed spectrum. $^{1}H$–$^{1}H$ scalar coupling constants are especially useful in providing information on the dihedral angle within a HC–CH system and are thus one

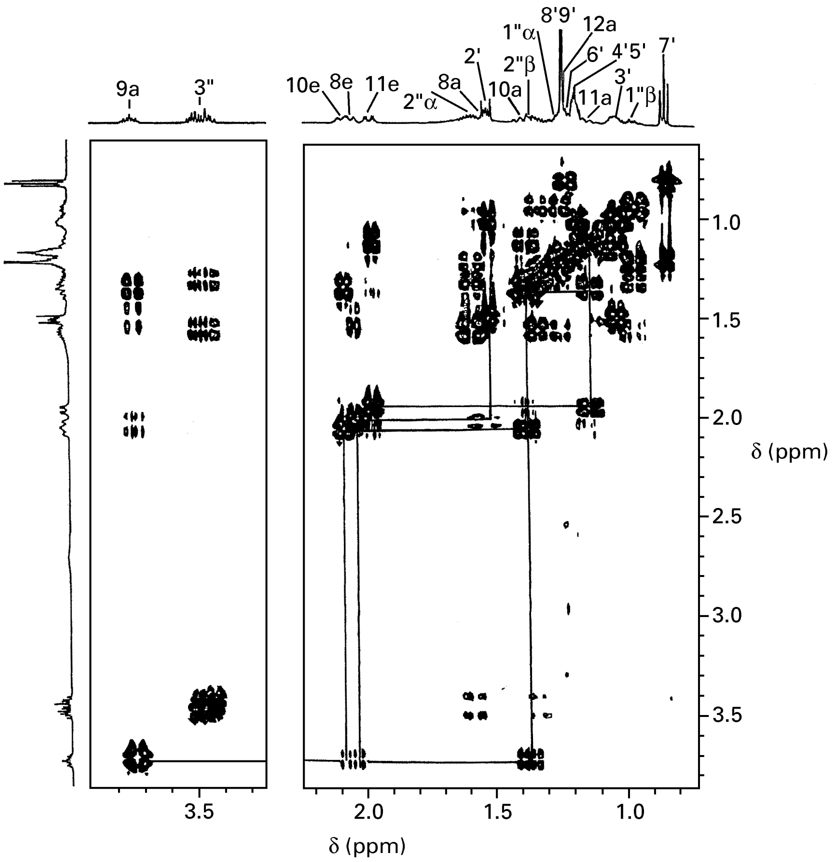



**Figure 2** Contour plots of expanded regions from the DQF-COSY spectrum of CP55,940. The lines highlight the connectivities between the H8, H9a, H10 and H11 resonances.

of the most important sources of conformational information.

#### Nuclear Overhauser effect spectroscopy (NOESY)

The nuclear Overhauser effect (NOE) is another important NMR parameter used in conformational analysis because the magnitude of the NOE is inversely proportional to the sixth power of the interproton distance in space ($I_{NOE} \propto r^{-6}$). NOE spectroscopy (NOESY) is two-dimensional experiment that may be run routinely in which the NOE is manifested as a crosspeak between two resonances indicating that the two protons are near in space.

For example, in the case of CP55,940, two NOE crosspeaks were assigned to the spatial coupling of H5 with H8a and H12a (**Figure 3**). Such crosspeaks are congruent with a conformation in which the planes of the two rings are almost perpendicular, with the Ph–OH oriented toward the α face of the cyclohexyl ring. An NOE crosspeak between the phenolic hydroxyl proton and the adjacent aromatic H2 indicates that these two protons are spatially near each other and thus coupled through a dipole–dipole interaction (**Figure 3**). Such a result indicates that in its preferred conformation, the Ph–OH proton points away from the cyclohexyl ring and towards the H2 proton.

The full analysis of NOESY and DQF-COSY spectra of other analogues, plus computational studies, further showed that this was typical for all congeners of CP55,940 and that the dimethylheptyl chain adopts one of four preferred conformations, in all of which the chain is almost perpendicular to the phenol ring. The most biologically active conformations were such that all hydroxyl groups were oriented towards one face of the cyclohexyl ring system (**Figure 4**), a feature that may be an important requirement for cannabimimetic activity.

## Structure of drug macromolecules

An increasing number of therapeutic drugs are composed of proteins or peptides and knowledge of their three-dimensional structures has helped in the design of structurally modified variants with improved biological activities and pharmacokinetic properties.

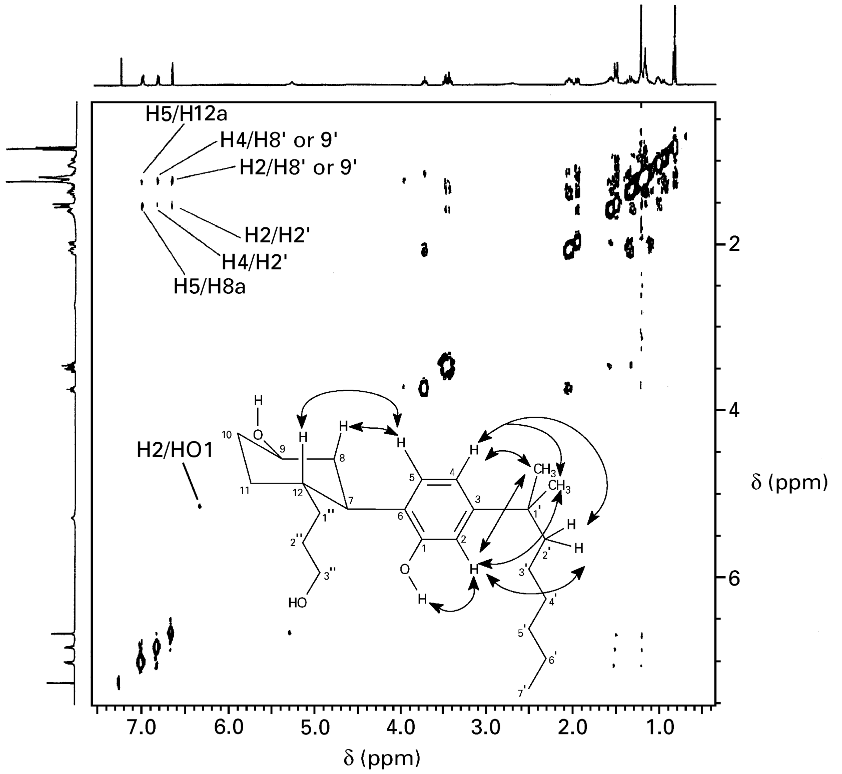



**Figure 3** Contour plot of the 500 MHz CP55,940 NOESY spectrum in $CDCl_3$. The NOE interactions for CP55,940 are indicated with arrows.

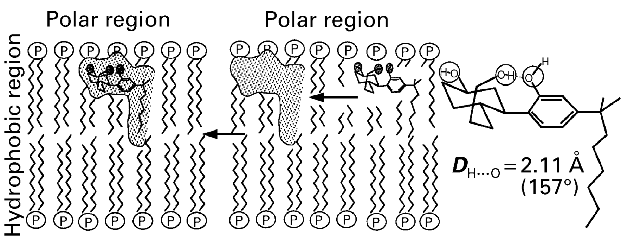



**Figure 4** Representation of the active biological conformation of CP55,940. According to the current hypothesis, the ligand preferentially partitions in the membrane bilayer where it assumes an orientation and location allowing for a productive collision with the active site.

Such a case is insulin, which was first extracted from pancreas tissue, used in a patient in 1922, and its structure first determined in 1972. It has a molecular weight of 5.8 kDa and consists of a 21 amino acid peptide (chain A) that is connected to a 30 amino acid peptide (chain B) by two disulfide bonds. A third intra-subunit disulfide bond exists in chain A.

The structure of insulin has been probed using X-ray crystallography, while NMR spectroscopy was used to determine its three-dimensional structure in solution. As with other similar-sized molecules, standard two-dimensional $^1$H homonuclear experiments can adequately provide such structural information.

The use of two-dimensional experiments such as DQF-COSY, TOCSY and NOESY to assign and determine the three-dimensional conformation of peptides and proteins is well established. Briefly, the method used for the assignment of amino acid chains is not dissimilar from that of assigning small molecules. Each amino acid residue gives rise to a characteristic spin pattern that can be identified using the complementary DQF-COSY and TOCSY spectra, where connectivities between all protons within a spin system are observed in the TOCSY, and connectivities between neighbouring protons are observed in the DQF-COSY. An example of the DQF-COSY spectrum for the des-pentapeptide insulin monomer is shown in **Figure 5**. Of the 46 amino acids in the monomer, non-overlapped spin systems for four valine, two gylcine, one threonine and one alanine residue were distinguished. The latter two were sequence specifically assigned as threonine-8 in the A chain and alanine-14 in the B chain as these were the only alanine and threonine residues in the sequence. In addition, a further nine AMX spin systems could be clearly distinguished that were subsequently assigned to amino acids with the aid of the TOCSY and NOESY spectra. A combination of DQF-COSY

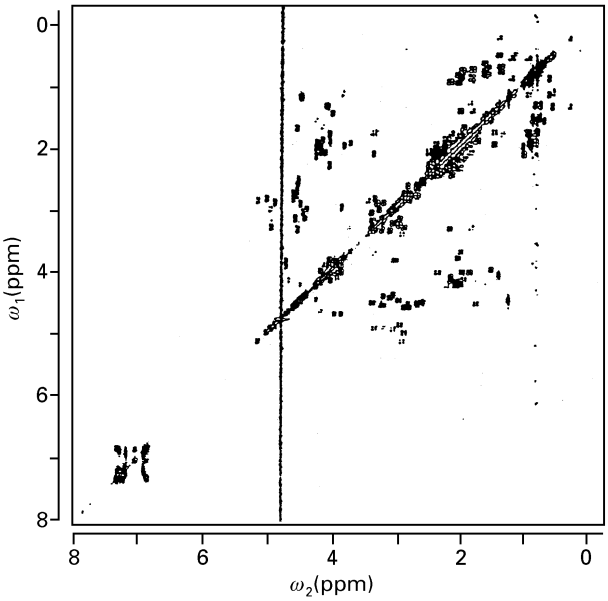



**Figure 5** The 500 MHz $^1$H DQF-COSY spectrum of a shortened analogue of insulin, des-(B26–B30)-pentapeptide insulin in $D_2O$. The spectrum was recorded on a Bruker AM 500 spectrometer interfaced with an ASPECT 3000 computer in a phase-sensitive mode. 2048 complex data points were acquired in the $t_2$ dimension, with a total of 512 free induction decays (FID) collected for transformation in the $t_1$ dimension. Each FID was acquired with a total of 128 scans.

and TOCSY data were required in order to delineate the remaining overlapped amino acid spin systems.

Once the spin systems arising from individual amino acids have been identified, the correct sequence of amino acids is determined from NOE data, acquired using the NOESY experiment. For molecules that fall within a particular molecular weight range (1000–3000 Da), the magnitude of the NOE is small, and a modification that is referred to as rotating Overhauser enhancement spectroscopy (ROESY) must be employed. Typically, a series of NOESY (and/or ROESY) spectra are collected in different solvents such as $D_2O$, $H_2O$ or DMSO, with mixing times ranging from 50 to 600 ms. Different mixing times are required to gain an estimate of the magnitude of the NOE and subsequently an estimate of the distance between the protons. Care must be taken at longer mixing times, as it is possible to observe an indirect magnetization transfer between protons that are further apart than 5 Å via a third proton that is appropriately positioned between them. To avoid this spin diffusion effect, it is preferable to acquire NOESY experiments with the shortest mixing times possible that will result in good quality spectra.

The determination of the amino acid sequence is based on the fact that NOEs will always be observed between particular protons from neighbouring residues regardless of the secondary structure of the protein. For example the proton attached to the α carbon (Hα) of an amino acid will always be within approximately 4.5 Å of the proton attached to the nitrogen (NH) of the neighbouring amino acid that is attached to the carboxyl end. In the NOESY spectrum, the NH resonances tend to occur within a particular spectral region, as do the Hα resonances. Thus, there is a particular region of crosspeaks between them in which a continuous connectivity pattern can be distinguished that begins with an NOE between the first and second residue and ends with an NOE between the terminal residue and the one preceding it (**Figure 6**). Such patterns reveal the sequence of amino acids in the protein. Similar patterns can also be distinguished between other groups of resonances (e.g. correlations between NH protons of neighbouring residues, Hβ protons and NH protons of neighbouring residues) that can be used to confirm the sequence or resolve ambiguities. Breaks in this continuous pattern do occur in cases where resonances overlap or in some instances due to the structure of the amino acid chain, for example when a proline is present. However, once various lengths of polypeptides have been identified it is usually possible to surmise the order in which the various lengths are connected. This is achieved by considering different possibilities until a process of elimination arrives at the arrangement that best fits the data.

Once the spectra have been assigned, data that contain structural information can be extracted. Nuclear Overhauser effects between non-neighbouring residues are the most revealing source of conformational information, and scalar coupling constants are important in providing information on torsion angles. All such conformational parameters are used as constraints in computational calculations that arrive at the three-dimensional conformation of the protein. For small molecules, such as the non-classical cannabinoids, it is possible to infer the conformation from NMR data without the aid of a computer, given that there are a relatively small number of NOEs generated and there are limited conformational possibilities from which to choose. However, for peptides, proteins and other macromolecules such as DNA, where a large number of measured NMR parameters must be taken into consideration, computational methods are an essential tool in the determination of the vast conformational possibilities. The computational techniques utilized seek to systematically adjust the position of all nuclei in the molecule, so that all distances and bond angles derived from NOE data and coupling constants are satisfied. At the same time, the structure must not exceed set

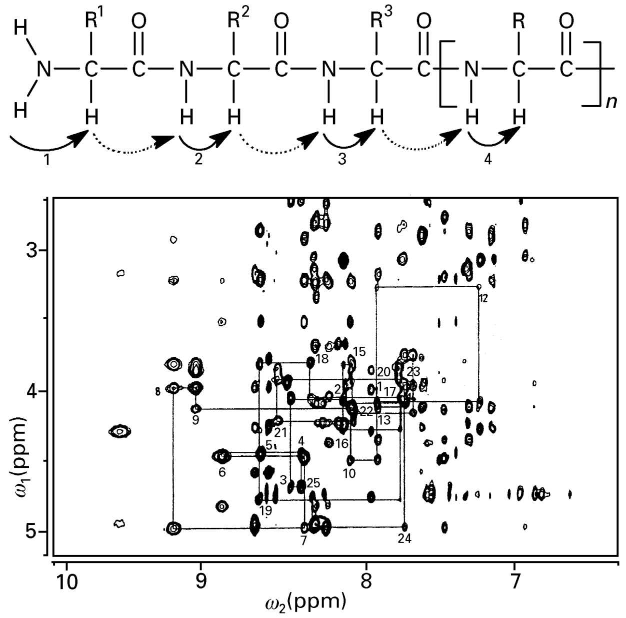



**Figure 6** The fingerprint region of a 500 MHz $^{1}$H NOESY spectrum which contains NOEs between NH and Hα resonances of des-(B26–B30)-pentapeptide insulin. The expected intra-residue and inter-residue NOEs that form the highlighted sequential pattern of B chain backbone NH and Hα resonances are represented as solid and dashed arrows, respectively.

physical limits for bond lengths, bond angles, torsion angles, van der Waals contacts and Coulombic interactions between atoms. The challenge in such methods is to ensure that all possible conformations are sampled while not allowing the molecule to exceed the set physical limits. Various algorithms and methods for calculation exist; however, the most common are the distance geometry and/or restrained molecular dynamics methods.

Using this basic methodology, the three-dimensional conformation of insulin was determined, and a significant amount of structure–activity information was gained by the study of insulin analogues and insulins purified from different sources.

## Structural analysis of drug-binding domains

Macromolecules such as proteins and nucleic acids form the sites at which drugs interact. Knowledge of three-dimensional conformations assists in the design of analogues that are more potent and have improved pharmacokinetic properties. Furthermore, the structural analyses of protein receptors and enzymes adds to the knowledge of biological systems, and therefore assists in identifying novel types of therapeutic agents. Because of their large molecular weights, most macromolecular therapeutic targets cannot be studied using exclusively $^{1}$H homonuclear methods.

The advent of three-dimensional and heteronuclear pulse techniques has greatly expanded the ability to study macromolecules of up to 30 kDa in size. Three-dimensional techniques are a natural progression of the two-dimensional experiments. The pulse sequence is altered so that a vertical domain is introduced and the information is spread into a third dimension, so that the spectrum now is projected into a cube instead of a plane. The diagonal of the cube is equivalent to the one-dimensional spectrum, and crosspeaks that may be overlapped in a two-dimensional spectrum can be resolved in the third dimension.

A case in point is the structure determination of the insulin receptor substrate-1 (IRS-1). Insulin binds to a membrane-bound receptor that is a ligand-activated protein tyrosine kinase. Upon insulin binding there is an autophosphorylation of several tyrosine residues on the cytosolic side of the receptor. This enhances the tyrosine kinase activity of the insulin receptor towards other substrates and is required for signal transduction. A cascade of events is initiated, the first of which is the phosphorylation of IRS-1. This occurs when IRS-1 binds to the insulin receptor via a specific domain of the protein that is termed the phosphotyrosine binding (PTB) domain.

The structure of this domain was determined while interacting with a tyrosine-containing peptide derived from a receptor. As such, this study is also an example of the use of NMR to study the interactions between molecules. The three-dimensional NMR experiment was coupled with heteronuclear techniques to increase the level of resolution of the spectra. For larger proteins such as IRS-1, homonuclear experiments are limited because proton resonances become broader, and the efficiency of magnetization transfer between protons is decreased. With the advent of recombinant DNA technology, this problem can be overcome by producing the proteins to contain isotopic labels, such as $^{13}$C, $^{15}$N or $^{2}$H, either at specific sites or uniformly throughout the protein. The magnetic properties of these nuclei offer significant advantages over protons. Carbon and nitrogen nuclei resonate over a much larger spectral width or range of frequencies, and as such are less likely to suffer from overlap. Also their scalar couplings to each other and to protons are higher in magnitude, and a more efficient magnetization transfer takes place than for $^{1}$H–$^{1}$H couplings. Thus, for macromolecules, heteronuclear experiments that correlate $^{15}$N to $^{13}$C nuclei or to protons offer

significant gains in sensitivity and resolution compared to homonuclear experiments. The most often used heteronuclear experiments are the heteronuclear single quantum coherence experiment (HSQC) and the related heteronuclear multiple quantum coherence experiment. These experiments allow the measurement of one- and two-bond heteronuclear couplings (and homonuclear $^{13}C$–$^{13}C$ couplings). They are most often combined with traditional two-dimensional experiments such as NOESY and TOCSY to yield a three-dimensional experiment. For example, in the case of an HSQC-NOESY spectrum of a protein, two of the axes represent the heteronuclei such as $^{15}N$ and the protons which are directly attached to the nitrogen nuclei, while the third axis contains chemical shifts of protons which share an NOE effect with the amide proton. This offers a significant increase in resolution compared to a traditional two-dimensional NOESY. A large array of these types of three-dimensional, heteronuclear-edited experiments have been designed to extract structural information in various situations.

Using these methods, the structure of the PTB domain of the IRS-1 protein was found to be similar to phosphopeptide-binding regions of several other proteins. Once the structural details were known, the different binding specificities could be compared and rationalized based on the interactions with their substrates.

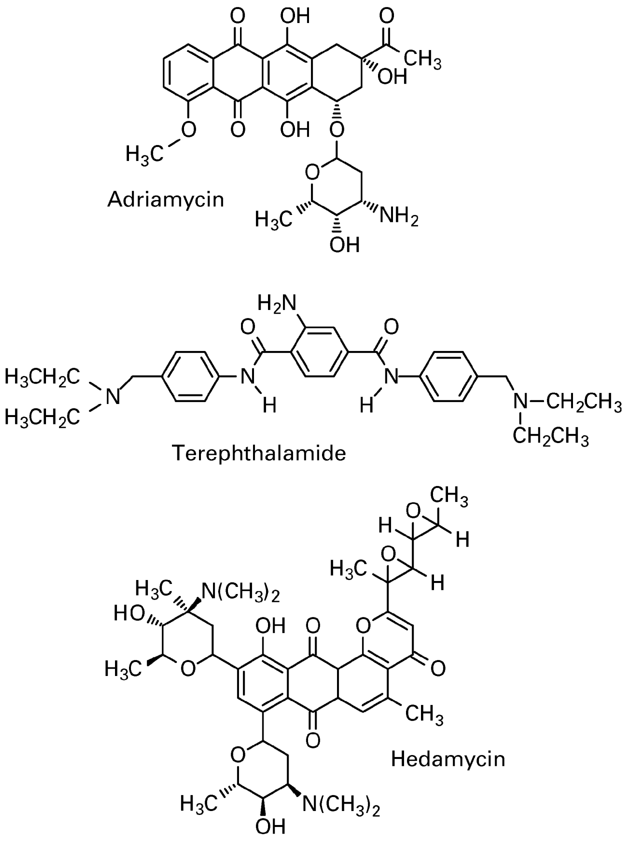



**Figure 7** Chemical structures of intercalating and minor groove DNA binding ligands.

## Drug interactions with macromolecular targets

NMR spectra are capable of supplying information about molecular interactions in solution. When a drug interacts with a receptor in a reversible manner, a number of effects may be observed in the spectra due to the exchange of the molecules between free and bound states. These effects will be discussed in relation to studies of antitumour antibiotics binding to short sequences of oligonucleotides. Compounds such as adriamycin (**Figure 7**) are currently in use as chemotherapeutic agents, and an understanding of factors involved in binding specificity may lead to more effective drugs to combat cancer. Chemical exchange effects have been exploited to great effect in obtaining information about these interactions.

The free and bound states of the molecules represent two different chemical environments in which participating molecules may be found. Thus, the same nucleus of a particular molecule may be characterized by different values for NMR parameters, such as chemical shift, when located in each different environment and give rise to different sets of resonances. The ability to measure the parameters that characterize each environment is dependent on the rate at which the nucleus exchanges between them. The exchange rate therefore has a significant effect on the appearance of NMR spectra, and the exchange rate is dependent upon the affinity of the drug for the receptor. The kinetics of the interaction can be examined by acquiring a series of spectra in which the concentrations of the reactants or the temperature are modulated.

**Figure 8** shows an expansion of the aromatic region of a series of $^{1}H$ NMR spectra of the oligonucleotide duplex d(GGTAATTACC)$_2$ to which has been added increasing amounts of a terephthalamide derivative (**Figure 7**). It is clear that the addition of the ligand causes significant perturbations of the free DNA resonances, and these indicate that the molecules are interacting. At ligand: DNA ratios between 0:1 and 1:1 there is a mixture of DNA molecules that are bound and unbound. The DNA resonances observed are averages of resonances arising from each of these states, and this is characteristic of fast exchange between the bound and unbound state due to a low affinity of the ligand for the binding site. There is a point in the titration in which the DNA

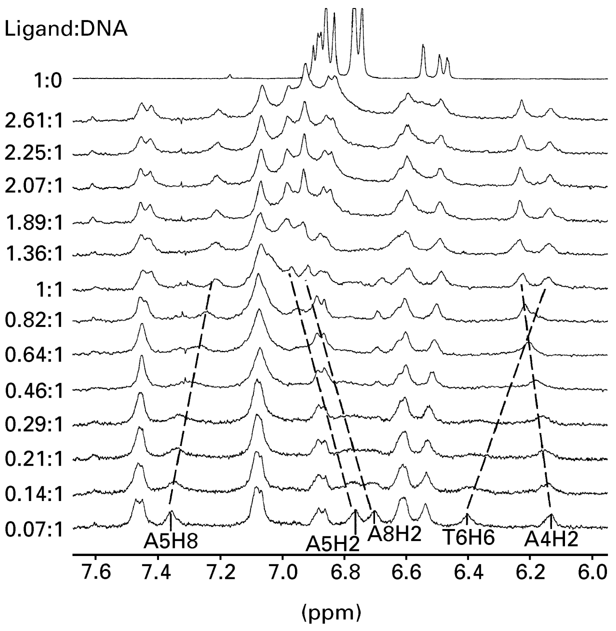



**Figure 8** Expanded aromatic regions from the 300 MHz $^1$H NMR spectra of complexes between a terephthalamide derivative and d(GGTAATTACC)$_2$, recorded at 10 °C. Increasing ligand concentrations cause perturbations to chemical shifts of nuclei located at or near the binding site. All perturbations are upfield except for adenine H2 resonances that are perturbed downfield and are located on the floor of the minor groove.

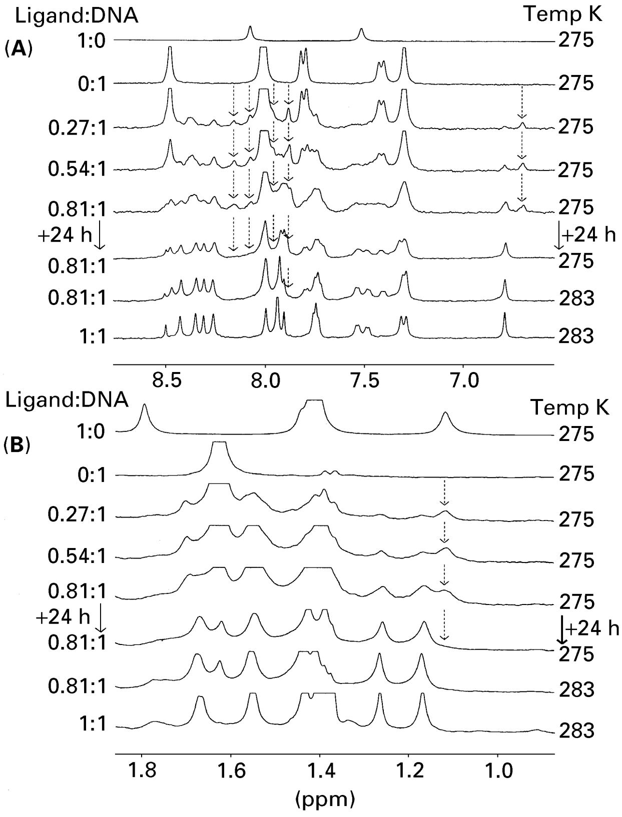



**Figure 9** Expanded regions from the 300 MHz $^1$H NMR spectra of complexes between hedamycin and d(CACGTG)$_2$, showing (A) aromatic resonances and (B) methyl resonances. Spectra up to 0.8:1 ligand:DNA ratio were recorded at 2 °C and the complex allowed to equilibrate for 24 h at this temperature. Subsequent spectra were recorded to 10 °C as the resonances were sharper. The dotted arrows highlight resonances that disappear after the 24 h equilibration period.

resonances are no longer perturbed, and this identifies the point at which all binding sites have been occupied and there are only bound DNA molecules present in solution. By noting the resonances that are most perturbed and the direction of the perturbation (upfield or downfield), it was determined that the ligand was bound at the 'ATTA' binding site. The observed perturbations were consistent with the ligand being inserted edge on into the minor groove. This places protons on the floor of the groove in the same plane as the aromatic drug which are then deshielded due to ring current effects. Protons above and below the plane of the ring experience shielding ring current effects.

An example of a ligand that has a high affinity for the binding site and exhibits slow exchange characteristics is shown in **Figure 9**, where the antitumour antibiotic hedamycin (see **Figure 7**) was titrated into a solution of the oligonucleotide duplex d(CACGTG)$_2$. Upon addition of hedamycin, the free DNA resonances diminish in intensity and new peaks appear in the spectrum. These new peaks do not correspond to chemical shifts of the free duplex or ligand resonances and increase in intensity with increasing ligand concentration. This suggests that they arise from the bound form of these molecules, and that slow exchange conditions exist due to the high affinity of the ligand for the binding site. In this particular case it became apparent that time-dependent changes in the spectra were taking place. Allowing the mixture to equilibrate for 24 h resulted in certain resonances disappearing from the spectrum and sharpening of the remainder of the resonances. Given that hedamycin was subsequently shown to intercalate and alkylate, the transient peaks may represent reaction intermediates prior to alkylation where the chromophore is intercalating reversibly to sites other than the most favoured binding site, and the sharpening of the spectra is caused by alkylation of the DNA by the ligand.

In cases where the binding affinity is high, or the binding is irreversible, a detailed model of the interaction between the molecules can be constructed based on intermolecular NOEs. A NOESY spectrum of the 1:1 complex of hedamycin with d(CACGTG)$_2$ is shown in **Figure 10**. Assignment of the resonances

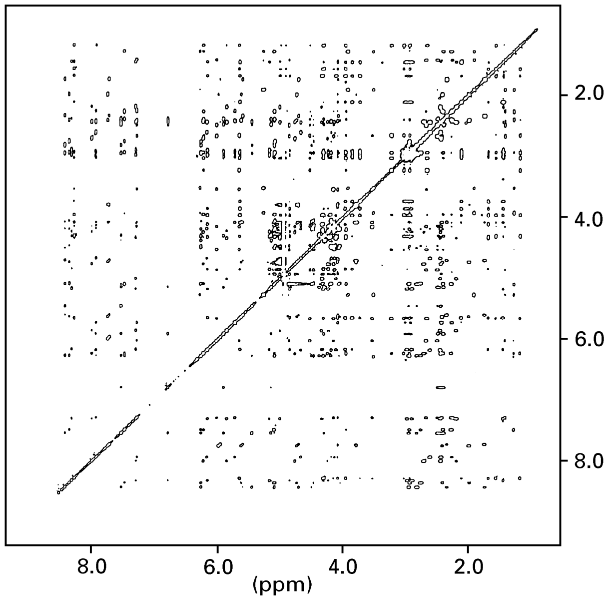



**Figure 10** Contour plot of the 500 MHz $^{1}H$ NOESY spectrum of the 1:1 hedamycin:DNA complex recorded at 10 °C with a mixing time of 300 ms. Spectra were acquired on a Bruker 500AMX spectrometer with 2048 complex data points in the $t_2$ dimension and a total of 512 free induction decays (FID) collected for transformation in the $t_1$ dimension. Each FID was acquired with a total of 128 scans.

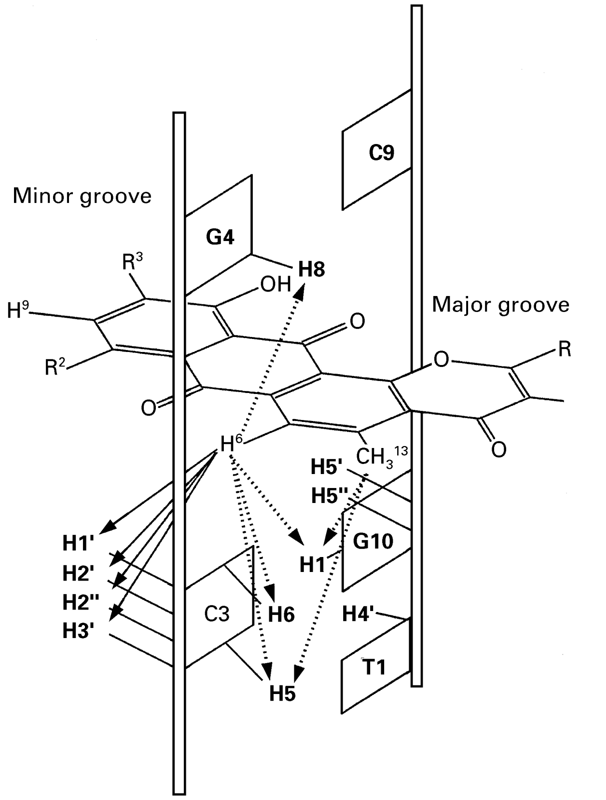



**Figure 11** Schematic representation of NOEs observed between the central GC basepairs and the hedamycin chromophore. Contacts to major groove protons are shown as dotted arrows and contacts to the minor groove are shown as solid arrows.

was achieved using a combination of COSY, TOCSY and NOESY spectra. As in the case of proteins, the NOESY spectra of oligonucleotides yield characteristic crosspeak patterns that allow the sequential assignment of residues. Once the spectrum had been assigned and intramolecular NOEs arising from the DNA duplex and hedamycin had been eliminated, a total of 61 intermolecular NOEs from each of the oligonucleotide strands to protons located on the chromophore, epoxide chain and sugar groups of hedamycin were identified. These intermolecular NOEs identified the binding site and allowed the orientation of the ligand within the binding site to be determined.

For example, a number of intermolecular NOEs, summarized in **Figure 11**, showed that the molecule was intercalated between the central CG basepairs. Similarly, intermolecular contacts showed that the sugar groups of the molecule were located in the minor groove (**Figure 12**) and the epoxide chain in the major groove. Furthermore, the epoxide chain was shown to be in an appropriate location, so that the terminal carbon was capable of alkylating to the N7 of the guanine. Direct evidence for this was observed in that the guanine H8 proton became labile (i.e. disappeared when spectra were acquired in $D_2O$) following complexation.

Another example involving proteins is the interaction of the PTB domain of the IRS-1 protein with a tyrosine phosphorylated peptide in which the peptide was found to bind in a surface-exposed pocket of the PTB domain. More specifically, the peptide was bound along one strand of the β sheet structure of the PTB domain and interacted with an α helix. Hydrogen bonding, van der Waals contacts and hydrophobic interactions stabilized the interaction. The peptide was found to be in a type I β turn, and the N-terminal residues of the peptide were in an extended conformation that formed an additional strand of the PTB domains β sheet.

## Combinatorial methods

Traditional methods of drug discovery involved a search amongst a range of diverse compounds that were derived from nature. This was a long and arduous process in which advances were due more often to serendipity rather than scientific thought and technique. The means to study the three-dimensional conformation of drugs and their receptors opened the door to a more rational approach in which compounds could be synthesized to better interact with a

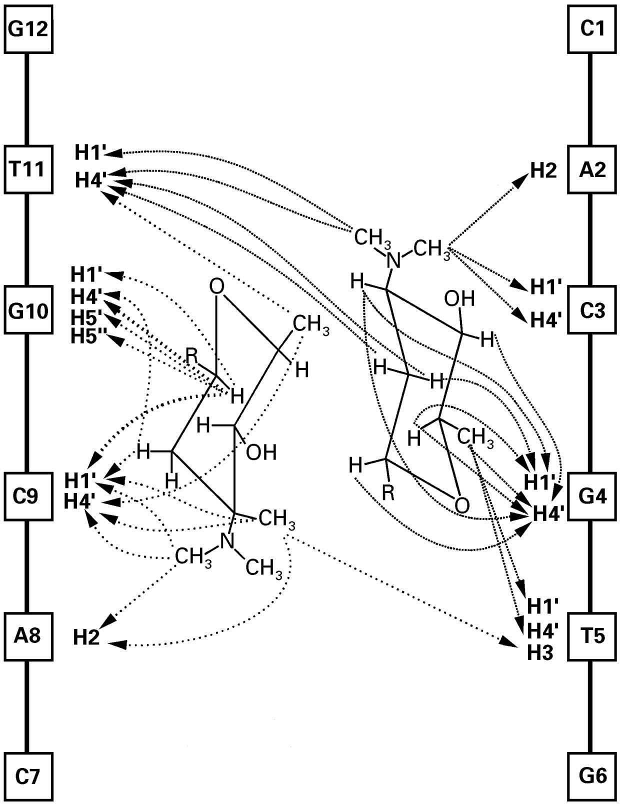



**Figure 12** Schematic representation of contacts between the sugar rings of hedamycin and the DNA duplex. Contacts are observed only to minor groove protons and each ring is associated predominately with one DNA strand.

receptor. Along with the better understanding of drug/receptor interactions came improvements in biochemical techniques to isolate receptors and test compounds for their affinity towards these receptors. The ability to test thousands of compounds in a relatively short amount of time using high throughput screening methods resulted in pressure to produce large numbers of diverse compounds to be tested.

### Analysis of solid-state intermediates

Combinatorial chemistry is a term used to describe the production of a large number (thousands) of diverse compounds, and different methods are employed to achieve this. One such method is solid-phase synthesis where the chemistry occurs on a solid support that may be packed into a column. The solid support may be a material such as crosslinked polystyrene and the compounds to be produced generally consist of a number of basic pharmacologically relevant structures (pharmacophores) that are to be linked in different ways. The method requires that one of the pharmacophores be linked to the solid support, then each subsequent pharmacophore can be delivered to the column with appropriate reagents, allowed to react, and then washed off the column. By varying the order in which the pharmacophores are delivered, different compounds can be produced.

This method relies heavily on the use of NMR to assess the success of each reaction step. As a non-destructive technique, a sample of the resin may be removed, placed in the NMR machine and then returned to the column. However, the fact that the compound is attached to the resin results in broad resonances if solution NMR techniques are utilized. This difficulty is circumvented by using solid-state NMR methods, including magic angle spinning (MAS). These experiments lead to narrow linewidths and high quality spectra, so that even two-dimensional experiments can be performed.

### Combinatorial methods in drug discovery

One of the most interesting applications of NMR in drug research is in the field of high throughput screening. One such recently described approach bridges combinatorial chemistry with biochemical screening and was named ‘SAR by NMR’. The first step in this method is to identify the ability of ligands to bind to a target preparation of the molecular therapeutic target (e.g. an enzyme) in solution. Thus, batches of ligand mixtures (e.g. one batch containing 20 compounds) are allowed to interact with the enzyme preparation in solution while following changes in the protein $^{15}N$ and $^{1}H$ NMR frequencies.

In this manner, batches that produce a response can be identified quickly and the compounds within each batch further tested to identify specific compounds that bind to adjacent but different sites on the protein. These ligands are then further optimized using rational design techniques to improve the binding at each respective site. The second generation ligands are then linked together to produce a compound that has a higher affinity than either of the two lead compounds. Using this method it was possible to identify two ligands that bound with micromolar affinities to the FK506 binding protein, which is involved in immunosuppression when activated. Linking these two individual ligands resulted in a compound that bound to FK506 binding protein with a nanomolar affinity. Preliminary use of this technique, which is applicable only to biomolecules of less than 30 kDa, indicates that it can have wide-ranging usefulness and is a potentially valuable tool in drug research.

## List of symbols

$F_1 = \omega_1 = y$-axis frequency domain obtained by Fourier transformation with respect to $t_1$; $F_2 = \omega_2 = x$-axis frquency domain obtained by Fourier transformation with respect to $t_2$; $t_1$ = evolution time in a 2D pulse sequence; $t_2$ = detection period in a 2D pulse sequence; $T_1$ = spin–lattice (longitudinal) relaxation time; $T_2$ = spin–spin (transverse) relaxation time.

*See also:* **High Resolution Solid State NMR, $^{13}$C; High Resolution Solid State NMR, $^{1}$H, $^{19}$F; Macromolecule-Ligand Binding Studied By NMR; NMR Pulse Sequences; Nuclear Overhauser Effect; Nucleic Acids Studied Using NMR; Proteins Studied Using NMR Spectroscopy; Small Molecule Applications of X-ray Diffraction; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Peptides.**

## Further reading

Anderson RC, Stokes JP and Shapiro MJ (1995) Structure determination in combinatorial chemistry: Utilization of magic angle spinning HMQC and TOCSY NMR spectra in the structure determination of wang-bound lysine. *Tetrahedron Letters* **36**: 5311–5314.

Boelens R, Ganadu ML, Verheyden P and Kaptein R (1990) Two-dimensional NMR studies on des-pentapeptide-insulin. Proton resonance assignments and secondary structure analysis. *European Journal of Biochemistry* **191**: 147–153.

Brange J, Ribel U, Hansen JF *et al.* (1988) Monomeric insulins obtained by protein engineering and their medical applications. *Nature* **333**: 679–682.

Craik DJ (1996) *NMR in Drug Design*, Boca Raton: CRC Press.

Davis SN and Granner DK (1996) Insulin, oral hypoglycemic agents, and the pharmacology of the endocrine pancreas. In: Hardman JG, Limbird LE, Molinoff PB, Ruddon RW and Gilman AG (eds) *The Pharmacological Basis of Therapeutics*, 9th edn, pp 1487–1517. New York: McGraw-Hill.

Pavlopoulos S, Bicknell W, Craik DJ and Wickham G (1996) Structural characterization of the 1:1 adduct formed between the antitumor antibiotic hedamycin and the oligonucleotide duplex d(CACGTG)$_2$ by 2D NMR spectroscopy. *Biochemistry* **35**: 9314–9324.

Shuker SB, Hajduk PJ, Meadows RP and Fesik SW (1996) Discovering high affinity ligands for proteins: SAR by NMR. *Science* **274**: 1531–1534.

Wilson SR (1997) Introduction to combinatorial libraries: Concepts and terms. In: Wilson SR and Czarnick AW (eds) *Combinatorial Chemistry, Synthesis and Applications*, pp 1–23. New York: Wiley Interscience.

Xie XQ, Melvin LS and Makriyannis A (1996) The conformational properties of the highly selective cannabinoid receptor ligand CP-55,940 *The Journal of Biological Chemistry* **271**: 10640–10647.

Zhou MM, Huang B, Olejniczak ET *et al.* (1996) Structural basis for IL-4 receptor phosphopeptide recognition by the IRS-1 PTB domain. *Nature Structural Biology* **3**: 388–393.

# Structure Refinement (Solid State Diffraction)

**Dieter Schwarzenbach**, University of Lausanne, Switzerland
**Howard D Flack**, University of Geneva, Switzerland

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Methods & Instrumentation**

Crystallographic structure refinement is generally understood to be the last step in the determination of a crystal structure by diffraction methods. The usual procedure of a crystal structure analysis includes collection of X-ray or neutron diffraction intensities, data reduction yielding structure factor amplitudes, the solution of the crystallographic phase problem yielding approximate structural parameters and finally refinement of these parameters to obtain a best fit of the observed structure factor amplitudes with the amplitudes calculated from the optimized model. The methods used to accomplish these successive steps depend on the type of compound and crystal. It is convenient to distinguish between small structures containing up to 100 or 200 symmetrically independent atoms, and macromolecular structures. The former are solved with atomic resolution, often routinely and nearly automatically, using standard and well-tested program packages. Modern efficient data collection apparatus employing area detectors and

adequate computing power have enabled small-molecule crystallography to become an analytical technique with turn-around times measured in days or even fractions of a day. Complete small-molecule structure determinations are normally carried out on single crystal data, but the number of successful structure solutions from synchrotron powder diffraction data is steadily increasing. The methodology applied to macromolecular structures has evolved quite independently. Corresponding software packages and algorithms are distinct to a large extent from those applied to small structures, but small-molecule crystallography has started to incorporate some of the techniques applied to macromolecules and the dialogue is open. The present article is devoted mainly to methods of small-molecule structure refinement against single-crystal diffraction data, offered by widely distributed program packages such as SHELX-97, CRYSTALS, XTAL and applied to chemical crystallography.

## Model fitting

The interpretation of measured quantities using a model derived from theory is a universal process in the physical sciences. The model is expressed by mathematical equations and contains constant numbers and parameters whose values are not fixed by the theory. The aim is then to find values for these variable parameters that best reproduce the observed quantities. The quality of the fit is judged according to a criterion that defines the distance between observed and modelled quantities. Model fitting is thus equivalent to the minimization of a multiparameter function expressing a distance criterion. Observations are always carried out to a limited precision, characterized by the standard uncertainty akin to the standard deviation or square root of the variance of a statistically distributed quantity. An important question to be answered is then the derivation of the standard uncertainties of the adjusted parameters of the model. This question is complicated by the fact that models are invariably imperfect: since they are obeyed only approximately, they cannot in principle obtain a perfect fit even for perfectly precise data. A crystallographic example is the use, in all models, of the kinematic scattering theory, which is correct only in the limiting case of vanishing diffracted intensity. Such model deficiencies are often referred to as 'systematic errors'. They may be reduced by performing experiments more closely related to the theoretical model. A further problem derives from applying corrections to the observations. Thus, observed diffraction intensities are corrected for background scattering, Lorentz and polarization factors and absorption effects. The resulting pseudo-observations (squared structure amplitudes on a relative scale) are then used as observations in the model fitting. Their values are in general correlated since an uncertainty in a correction and also systematic errors will affect all of them. This correlation is neglected in nearly all refinement software.

## The crystallographic model

The standard *procrystal model* assumes that the electron distribution in the crystal is very nearly equal to a superposition of previously known rigid atomic density distributions, which are smeared by harmonic lattice vibrations. The structure factor $F(h, k, l)$ with integer Miller indices $h, k, l$ for $N$ atoms per unit cell, located at the positions $x_n$, $y_n$, $z_n$ then becomes

$$F(h,k,l) = \sum_{n=1}^{N} p_n f_n T_n \exp[2\pi i(hx_n + ky_n + lz_n)] \quad [1]$$

The atomic scattering factors $f_n$ are the Fourier transforms of the spherical atomic electron distributions. They are considered as known from quantum-chemical calculations. The site occupation parameters $p_n$ may assume values different from unity if the structure is disordered. The Debye–Waller factors $T_n$ allow for the atomic thermal motions. They are functions of the atomic displacement parameters $U^{i,j}$. Omitting the atom index $n$ and representing the Miller indices and lengths of the reciprocal lattice vectors by $h_i$ and $a_i^*$, respectively:

$$T = \exp\left[-2\pi^2 \sum_{i,j} U^{i,j} a_i^* a_j^* h_i h_j\right] \quad [2]$$

There results nine adjustable parameters per atom, i.e. three positional coordinates and six displacement parameters, and sometimes an additional site occupation parameter, although the point symmetry of an atom in a special position may reduce the number of adjustable parameters. The model value of the net intensity $I(h, k, l)$, Bragg peak minus background, is

$$I(h,k,l) = K\,Lp\,T\,y\,|F(h,k,l)|^2 \quad [3]$$

where $K$ is a scale factor, $L_p$ is the Lorentz-polarization correction, $T$ is the transmission factor allowing for X-ray absorption, and $y$ is an extinction

correction allowing for some deficiencies in the kinematic scattering theory. The observed net intensities are routinely corrected for $Lp$ and somewhat less routinely for $T$. The resulting corrected observations $|F_{obs}|^2$ and the corresponding model quantities $|F_{calc}|^2$ thus become

$$|F_{obs}(h,k,l)|^2 \rightarrow K\,y\,|F(h,k,l)|^2 = |F_{calc}(h,k,l)|^2 \quad [4]$$

where $y$ depends on one or more parameters, and $K$ is also adjustable. Taking the square root of both sides gives the model in terms of structure amplitudes $|F|$. Each observed quantity is accompanied by an estimation of its uncertainty, usually derived from counting statistics and/or the spread of multiple and symmetry-equivalent observations around their average value. The model may of course be augmented at will by additional parameters (twinning ratios, enantiomorph-polarity parameter, charge density or anharmonic motion parameters, etc.), but the above-mentioned are included in every refinement software.

## The principle of least squares

By far the most important distance criterion between observed and modelled quantities is the least-squares deviance

$$D = \sum_{h,k,l} w(h,k,l)\ \{|F_{obs}(h,k,l)|^2 - |F_{calc}(h,k,l)|^2\}^2 \quad [5]$$

or analogously in terms of $|F|$, $w(h, k, l)$ being a weighting factor. The problem of structure determination, including the crystallographic phase problem, is thereby formulated in terms of an optimization criterion: find the minimal deviance $D$ by varying all the parameters of the model defined by Equations [1], [2] and [4]. The solution would be unique and could be easily obtained if the equations for $|F_{calc}(h, k, l)|^2$ describing the model were linear. However, they are highly nonlinear, and a unique solution is not assured (although solutions of ordered structures at atomic resolution are indeed most probably unique). The optimization for nonlinear model equations becomes an iterative process, starting with approximate parameter values obtained by the methods of structure determination. For this reason, the structure solution and the structure refinement steps are usually dealt with separately.

In the following, we denote the $M$ variable parameters of a structure by $p_m$ ($1 \le m \le M$), the $J$ observations $|F_{obs}(h, k, l)|^2$ or $|F_{obs}(h, k, l)|$ by $O_j$ ($1 \le j \le J$) and the corresponding model quantities by $C_j$. The usual iterative Gauss–Newton method for approaching the optimal value of $D$, starting from approximate values $p_m^{(k)}$ for the adjustable parameters obtained in the $(k)$th iteration, is to linearize the model equations for $C_j$. The deviance $D$ (see Eqn [5]) to be minimized with the improved parameter values $p_m^{(k+1)}$ then becomes

$$D = \sum_{j=1}^{J} w_j\ \{O_j - C_j[p_1^{(k+1)} \cdots p_M^{(k+1)}]\}^2 \quad [6]$$

$$C_j[p_1^{(k+1)} \cdots p_M^{(k+1)}] \approx C_j[p_1^{(k)} \cdots p_M^{(k)}] + \sum_{m=1}^{M} (\partial C_j/\partial p_m)_k\ [p_m^{(k+1)} - p_m^{(k)}] \quad [7]$$

The partial derivatives are evaluated with the parameters $p_m^{(k)}$. This leads to the matrix equation $\mathbf{N}[\mathbf{p}^{(k+1)} - \mathbf{p}^{(k)}] = \boldsymbol{\nu}$ with elements $N_{m,n}$ of the *normal matrix* $\mathbf{N}$ and $\nu_m$ of the vector $\boldsymbol{\nu}$.

$$N_{m,n} = \sum_{j=1}^{J} w_j (\partial C_j/\partial p_m)_k (\partial C_j/\partial p_n)_k, \qquad \nu_m = \sum_{j=1}^{J} w_j [O_j - C_j^{(k)}](\partial C_j/\partial p_m)_k \quad [8]$$

The process starts with the parameters $p_m^{(0)}$ obtained with the methods of structure solution. Algorithms for solving Equation [8] are described in the literature. If the number of observations is much larger than the number of parameters, e.g. $J/M \approx 10$, and if the structure is well-ordered, the diagonal terms $N_{m,m}$ are large compared to the off-diagonal terms, and $\mathbf{N}$ is easily invertible. A refinement may be started with the scale factor and the positional parameters. The displacement parameters are then found automatically starting from values with reasonable orders of magnitude or from zero, and convergence is rapid. However, the observations may not be particularly sensitive to certain structural features of interest. The normal matrix then risks being nearly singular, the results of matrix inversion may be dominated by rounding errors, and convergence of the iterative adjustment is not guaranteed and may be

meaningless. Examples of ill-conditioned problems are the parameters of weakly scattering hydrogen atoms in the presence of heavy atoms, or those of split atoms in disordered structures. In such cases, experience shows that it is in the use of a more appropriate model, constraints and restraints (see below), that refinement can be achieved rather than in resort to more advanced numerical algorithms.

The Gauss–Newton algorithm is not the only method capable of minimizing the deviance (Eqn [5]). The deviance of ill-conditioned nonlinear least-squares problems may possess several, often shallow, minima. There is then no guarantee that a minimum found starting from a given trial structure is the lowest one. The method of *simulated annealing* may be suitable for such problems as it permits us to leave a local minimum in search of another. It is frequently used in macromolecular crystallography.

## Estimation of uncertainty

A result of a measurement is not complete without a quantitative statement of its uncertainty. Likewise, the value of a quantity derived from the measurement must be accompanied by its uncertainty. The uncertainty reflects the lack of exact knowledge of the value owing to random and systematic effects including deficiencies in the model. The quantitative measure of uncertainty is called standard uncertainty. It is an estimate of the standard deviation $\sigma$ (i.e. the positive square root of the variance) of the probability density function of the quantity.

The method of least squares allows standard uncertainties to be obtained for all adjusted parameters. For any weights $w_j$, and for any probability density function of the observations, linear least squares is an unbiased estimator if the weights are independent of the observations, and if the model is perfect, i.e. represents physical reality correctly for some values of the parameters. In particular, the *Gauss–Markov theorem* states that minimal variances of the estimates are obtained if the weights are chosen equal to the inverse variances of the observations, $w_j = \sigma_j^{-2}$. For these, and only for these weights, the inverse of the normal matrix $\boldsymbol{N}^{-1}$ is an unbiased estimate of the variance–covariance matrix of the model parameters. In particular, the square roots of the diagonal terms of $\boldsymbol{N}^{-1}$ are the standard uncertainties of the adjusted parameters. We have tacitly assumed in the foregoing that the observations are statistically independent, i.e. that their covariances are zero. The Gauss–Markov theorem applies in fact also to correlated observations. Generally valid expressions for the variance–covariance matrix of the model parameters for weights other than minimum-variance weights are rarely implemented in refinement software. As a rule, crystallographic software implements estimated minimum-variance weights, $w_j = u_j^{-2}$, where $u_j$ is the standard uncertainty of $O_j$. The use of reliable values for these standard uncertainties is of utmost importance. They may be estimated from the spread of multiply measured or symmetry-equivalent Bragg intensities about their average values, the intensity fluctuations of periodically measured check reflections and Poisson statistics of count rates. The latter may be estimated from the observed $O_j$, the model $C_j$ or a combination of these quantities. When other weights, e.g. unit weights, are used, the apparent standard uncertainties of the model parameters obtained from $\boldsymbol{N}^{-1}$ will be meaningless.

The goodness-of-fit $S$ is a measure of the extent to which the calculated model values $C_j$ agree with the observations $O_j$. For miminum-variance weights, $S$ is calculated as

$$S^2 = (J - M)^{-1} D_{\min} = (J - M)^{-1} \sum_{j=1}^{J} u_j^{-2} (O_j - C_j)^2 \qquad [9]$$

where $D_{\min}$ is the deviance at convergence, and the $u_j$ are the standard uncertainties of the $O_j$. $J$ and $M$ have been defined above. The expectation value of $S$ for a perfect model is $\langle S \rangle = 1.0$. If the probability density functions of the observations are Gaussian, $D_{\min}$ is distributed like $\chi^2$ with $J - M$ degrees of freedom. In practice, values for $S$ near unity are rarely obtained even for fairly estimated $u_j$ values. Apart from the possibility of inappropriate measurement procedures, this is explained by imperfections in the model. For example, we may recall that the standard procrystal model neglects the effects of chemical bonding and anharmonic thermal motions, that the kinematic theory of diffraction may be a questionable approximation to reality, and that X-ray absorption effects may have been incompletely accounted for. It is common practice among crystallographers to multiply the standard uncertainties of the parameters obtained from $\boldsymbol{N}^{-1}$ by $S$. Although this is equivalent to the unrealistic assumption that a lack of fit is due entirely to a constantly proportional underestimation of the standard uncertainties of the observations, it does have the advantage of increasing the standard uncertainties of the parameters obtained from $\boldsymbol{N}^{-1}$. Moreover, other elaborate schemes of uncertainty manipulation are in use for attaining $S \approx 1$.

It may be important to stress at this point that standard uncertainties of parameters optimized with crystallographic programs depend on the method of refinement. Thus, if a parameter has been fixed because it was highly correlated with another parameter, thereby leading to an ill-conditioned normal matrix, the resulting standard uncertainties refer to an effective model where the fixed parameter assumes its true value which is previously known. Adjusting correlated parameters by alternatingly fixing one at its current value while refining the other is an unacceptable practice: it masks the fact that the parameters may not be both obtainable from the available data, and their standard uncertainties will be severely underestimated. Another example is block-diagonal refinement: in order to save storage space and computer time, the normal matrix is decomposed into a set of smaller matrices by neglecting off-diagonal terms outside the latter. With increasing computer power, this practice has been progressively abandoned but may still be attractive for large structures. The resulting standard uncertainties may, however, be underestimated, particularly if one or more of the neglected matrix elements assume large values. Wherever possible, the standard uncertainties should be calculated from an unrestricted full normal matrix.

## Maximum likelihood

Maximum likelihood is a more general statement of the optimization problem than least squares. Instead of optimizing a distance criterion such as the deviance $D$, the problem may be stated in the following manner: ‘Given a model and a set of observations, what is the likelihood of observing those particular values, and for what values of the parameters of the model is that likelihood a maximum?’ [*International Tables for Crystallography*, Vol. C (1992) p 605. Dordrecht: Kluwer Academic.] The answer depends on the probability density functions $\rho$ of the observations $O_j$. Assuming the observations to be uncorrelated and the model quantities $C_j$ to be the expectation values of the $O_j$, the probability of observing the set of the $O_j$ is the likelihood $L$,

$$L(p_1 \cdots p_M) = \prod_{j=1}^{J} \rho[O_j - C_j(p_1 \cdots p_M)] \qquad [10]$$

$L$, or more easily the logarithm ln $L$, is maximized by varying the parameters $p_m$. Simple algebra gives for the gradient of ln $L$

$$\partial \ln L/\partial p_m = (1/2)\sum_{j=1}^{J} [\zeta_j^{-1} \mathrm{d}\phi(\zeta_j)/\mathrm{d}\zeta_j]\,[\partial\zeta_j^2/\partial p_m]$$
$$\phi(\zeta_j) = \ln \rho(\zeta_j); \;\; \zeta_j = (O_j - C_j)/\sigma_j \qquad [11]$$

where $\sigma_j$ is the square root of the variance of the function $\rho(\zeta_j)$, or a measure for the width of $\rho(\zeta_j)$. The term $\partial\zeta_j^2/\partial p_m$ is the gradient of Equation [6] with $w_j = \sigma_j^{-2}$. For a Gaussian function, $\rho(\zeta) = (2\pi\sigma^2)^{-1/2}\exp(-\zeta^2/2)$, maximum likelihood is identical with minimum-variance least squares. If $\rho(\zeta)$ is not Gaussian, Equation [11] shows that maximum likelihood is equivalent to minimizing a least-squares deviance, where the minimum-variance weights are multiplied after each cycle by a function of the current deviates $O_j$–$C_j$. For a Cauchy function, $\sigma(\zeta) = \pi\sigma(1 + \zeta^2)^{-1}$, the modified weights are $w_j = [\sigma_j^2 + (O_j - C_j)^2]^{-1}$. Strongly discordant data are thus down-weighted because they are much more probable for a Cauchy function than for a Gaussian.

The probability density functions of the observations are generally unknown, but the Gauss–Markov theorem ensures that least-squares is always an acceptable estimator. However, the results of least squares are strongly influenced by discordant observations, so-called outliers. The *robust-resistant* techniques use weight-modification functions of $O_j$–$C_j$ which progressively down-weight outliers. These functions implicitly define probability functions $\rho$. They may alternatively be interpreted as an appreciation of the reliability of certain measurements. This approaches the frequently used option to simply omit discordant observations because they are judged to be unreliable.

## Constraints

It is usually necessary, and often desirable, to impose relations between the parameters of a model. The problem to be solved is then a minimization of Equation [6] while imposing subsidiary conditions. A *constraint* satisfies such a condition exactly. A *restraint* (soft constraint) is a condition accompanied by a measure of uncertainty or a weight, and accordingly satisfies the condition only approximately.

Most crystallographic refinements require parameter constraints due to site symmetries. For example, the position of an atom on a mirror plane is defined by two independent coordinates only. The displacement tensor of such an atom is defined by four, rather than six, $U^{i,j}$ values, since one of its eigenvectors must be perpendicular to the plane. Modern refinement programs find and apply such constraints automatically for all possible site symmetries. In some space groups the origin of the coordinate system

cannot be fixed with respect to the symmetry elements and thus must be defined by a constraint between positional parameters. Certain programs detect this condition and apply the constraint automatically, others do not. If site occupation parameters are refined, a constraint assuring electroneutrality may be required. The above types of constraint are usually implemented in the software by the method of parameter elimination.

Certain special features to be imposed on a model may be expressed by more complicated constraint equations. We note as an example the assumption of a rigid molecule with prescribed dimensions whose position and orientation are to be refined. The position may be described by the coordinates of the centre of mass and the orientation by three Euler angles with respect to a unitary coordinate system. The atomic coordinates and thus the structure factor, Equation [1], are expressed as functions of these six parameters. The latter may then be adjusted to optimize the deviance. A similar procedure can be used to constrain the atomic displacement parameters of a molecule to rigid-body movements described by a translation tensor, a libration tensor and a translation/libration-correlation tensor (TLS model). This model neglects intramolecular vibrations.

## Restraints

Restraints are used to specify certain features of the structural model which are approximately known. For example, typical values of the bond lengths such as C–C or C–H are well-known; the quality of a structure determination is indeed judged by, among other properties, the plausibility of bond lengths. Bond lengths in well-ordered small-molecule structures are, as a rule, accurately determined from the diffraction data. This is not the case for those macromolecular structure determinations that do not attain atomic resolution. It may not be the case for disordered structures characterized by split-atom positions and superpositions of molecules or structural fragments in various orientations, such as found in structures of fullerenes, $C_{60}$ and $C_{70}$, or in structures containing highly symmetric ions such as $ClO_4^-$. Powder diagrams may also contain insufficient information for the determination of all atomic positions. In such cases, the prior knowledge of bond lengths may permit a successful structure refinement.

Bond lengths are not known with absolute precision, nor does a bond between two atoms assume exactly the same value in all structures. Bond lengths should therefore not be fixed to a particular value by a constraint. Instead, they are introduced in the form of additional observational equations. The distance between atoms *m* and *n* is

$$d_{\text{calc}} = |(x_m - x_n)\boldsymbol{a} + (y_m - y_n)\boldsymbol{b} + (z_m - z_n)\boldsymbol{c}| \approx d_{\text{obs}} \quad [12]$$

where $d_{\text{obs}}$ has an associated standard uncertainty $u$ expressing the confidence to be placed in this information. This pseudo-observation is used like all other observations in Equation [6], [7] and [8], with a weight $w = u^{-2}$. A residual $|d_{\text{obs}} - d_{\text{calc}}|$ of the order of $u$ shows the distance information to be useful. A large residual, on the other hand, indicates a contradiction between distance and diffraction information which must be taken seriously: one of the two types of information or the corresponding weights may be erroneous. Simply increasing the weight of the restraint equation at the expense of the observations may be most inadequate, even if it leads to a closer agreement of the bond length with its expectation.

Other geometrical restraints may be applied to bond angles and torsion angles. Groups of atoms may be restrained to occupy a plane by fixing one or several torsion angles to be 0° or 180°. *Similarity restraints* are used to impose two or more bonds to be of the same length without specifying the value of the bond length. This permits the restraint of the hexagonal symmetry of a phenyl ring, or the icosahedral symmetry of a $C_{60}$ molecule. Restraints are thus far more flexible than rigid-molecule constraints, and may be applied to only part of a molecule. Distance and angle information alone, without diffraction data, has been used to optimize atomic coordinates in framework structures. Examples are refinements of alumosilicate frameworks in zeolites starting with unit cell data as determined from powder diagrams, and assuming diverse space groups belonging to subgroups of the holohedral lattice symmetry (the highest symmetry compatible with the metric of the translation lattice).

The vibrationally rigid bond criterion expresses the expectation that the amplitudes of intramolecular vibrations should be much smaller than those of the translational and librational movements of a molecule. The mean-square displacements of two strongly bonded atoms along their bond should accordingly be approximately equal. Accurate structure determinations confirm this criterion for C–C bonds, but not for metal–ligand bonds in transition-metal complexes. Vibrationally rigid bonds may be imposed by rigid-bond restraints on the anisotropic displacement parameters. If all interatomic distances in a molecule are specified to be rigid, its thermal motions are restrained to rigid-body translation and libration (TLS), except if the molecule is linear or

planar. With respect to a constrained TLS-refinement mentioned above, restraints offer the advantage that partially rigid molecules or structural groups may be defined.

*Shift-limiting restraints* are used to improve the convergence of iterative least-squares cycles. Strongly correlated parameters, such as displacement and site-occupation parameters refined against a low-resolution data set, result in an ill-conditioned pseudosingular normal matrix. The minimum of the deviance with respect to such parameters may be very shallow, successive shifts may erratically oscillate in opposite directions, or the refinement may diverge. A shift-limiting restraint ensures that the shift of the parameter value is of the order of the associated uncertainty $u$, or in terms of Equation [7] $p_m^{(k+1)} - p_m^{(k)} \approx 0$ with weight $w = u^{-2}$. Such a restraint simply adds $w$ to the corresponding diagonal term of the normal matrix. This clearly leads to a better conditioned and more easily invertible matrix. After convergence, all shift-limiting restraints should be removed and a final cycle should be calculated to obtain acceptable standard uncertainties for the parameters. Shift-limiting restraints are a simplified version of the *Levenberg–Marquardt method* in which the diagonal terms of the normal matrix are multiplied by a number which approaches 1.0 at convergence.

In general, any function of the structural parameters may serve as a restraint. An important example in macromolecular crystallography is *molecular dynamics*. In addition to the diffraction data, an empirical potential energy expression $E_{pot}$ is minimized, which contains terms for bond lengths, bond angles, torsion angles and nonbonded interactions. The function minimized is then the sum of the deviance, Equation [5], and the energy, $D + E_{pot}$. Molecular dynamics is a generalization of the distance and angle restraints discussed above where the weights are derived from the expression for the energy.

## Judging the results

The quality of the fit between observed and calculated structure amplitudes is commonly characterized by the agreement factors

$$R1 = \sum_{h,k,l} \Big| |F_{\mathrm{obs}}| - |F_{\mathrm{calc}}| \Big| \; / \sum_{h,k,l} |F_{\mathrm{obs}}|$$

$$wR2 = \left\{ \sum_{h,k,l} w(|F_{\mathrm{obs}}|^2 - |F_{\mathrm{calc}}|^2)^2 \; / \sum_{h,k,l} w|F_{\mathrm{obs}}|^4 \right\}^{1/2} \qquad [13]$$

The conventional index $R1$ usually includes only the stronger reflections; a common selection is $|F_{obs}| > 4u_F$ or $|F_{obs}|^2 > 2u_{F2}$, where $u_F$ and $u_{F2}$ are the standard uncertainties of $|F_{obs}|$ and $|F_{obs}|^2$, respectively. The goodness of fit $S$ defined by Equation [9] is a less appropriate measure since some software adjusts the weights to bring this value close to 1.0. Values of $R1$ up to 0.07 are considered to be acceptable; for very accurate studies, all agreement factors are of the order of 0.02 or lower.

The optimal values of the parameters and their associated standard uncertainties are meaningful only in relation to the model employed in the final least-squares cycle. All constraints and restraints used should be reported. Note that the positions and displacement parameters of hydrogen atoms are often derived from the parameters of the heavier atoms and subsequently suitably restrained or constrained; corresponding bond lengths, such as C–H, are then features of the model and not experimental evidence.

In the final cycle, the best estimate of minimum-variance weights, $w = u^{-2}$, should be used, where $u = u_{F2}$ or $u_F$ for refinements against $|F|^2$ or $|F|$, respectively. The standard uncertainties of the refined parameters obtained from the inverse normal matrix are meaningful only if the standard uncertainties of the observations are meaningful. It should also be borne in mind that even in careful studies, model errors are unavoidable. These include imperfect corrections for absorption and extinction, as well as the standard procrystal approximation neglecting chemical bonding and anharmonic motions. In addition, the uncertainties of molecular parameters such as bond lengths are often computed neglecting correlations between the model parameters. Therefore, published uncertainties of refinement results are approximate and tend to be underestimates. They can be improved only through a careful estimation of $u_{F2}$ or $u_F$ and a reduction of model errors.

## List of symbols

$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ = lattice vectors defining unit cell; $a_i^*$ = length of $i$th reciprocal lattice vector; $C_j$ = $j$th calculated model quantity, either $|F_{calc}(h,k,l)|$ or $|F_{calc}(h,k,l)|^2$; $D$ = deviance, distance criterion in model fitting; $d_{calc}$ = calculated interatomic distance; $d_{obs}$ = observed interatomic distance; $f_n$ = atomic scattering factor of $n$th atom; $F(h, k, l)$ = structure factor with Miller indices $h, k, l$; $|F(h, k, l)|$ = structure amplitude with Miller indices $h, k, l$; $|F_{calc}(h, k, l)|$ = calculated structure amplitude on relative scale; $|F_{obs}(h, k, l)|$ = observed structure amplitude on relative scale; $h, k, l$ = Miller indices; i = square root of –1; $I(h,k,l)$ = intensity of Bragg reflection; $J$ = number of

observations; $K$ = scaling factor; $k$ = index of $k$th refinement interaction (e.g. $p_m^{(k)}$); $L$ = maximum likelihood function; $Lp$ = Lorentz-polarization correction; $M$ = number of adjustable parameters; $\mathbf{N}$ = normal matrix with elements $N_{m,n}$; $O_j$ = $j$th observed quantity, either $|F_{obs}(h,k,l)|$ or $|F_{obs}(h,k,l)|^2$; $p_m$ = $m$th adjustable parameter; $p_n$ = site occupancy factor of $n$th atom; $R1$ = reliability index with respect to $|F(h,k,l)|$; $wR2$ = weighted reliability index with respect to $|F(h,k,l)|^2$; $S$ = goodness of fit, $S = \{D/(J-M)\}^{1/2}$; TLS = rigid-body thermal displacement description; $T$ = transmission factor or Debye–Waller factor; $T_n$ = Debye–Waller factor of $n$th atom; $u$ = standard uncertainty; $U^{i,j}$ = anisotropic atomic displacement tensor element; $\boldsymbol{v}$ = vector of normal equations with elements $v_m$; $w_j$ = weight of $j$th observation; $x_n, y_n, z_n$ = positional coordinates of $n$th atom; $y$ = extinction correction; $\phi$ = logarithm of probability density function; $\rho(\zeta)$ = probability density function; $\sigma$ = standard deviation, square root of variance measure for width; $\zeta$ = random variable with mean 0.0 and variance 1.0.

*See also:* **Powder X-Ray Diffraction, Applications; Small Molecule Applications of X-Ray Diffraction.**

## Further reading


Albinati P, Becker PJ, Boggs PT *et al* (1992) Refinement of structural parameters. In: Wilson AJC (ed) *International Tables for Crystallography*, Vol. C, pp 593–652. Dordrecht: Kluwer Academic.

Brünger AT (1988) Crystallographic refinement by simulated annealing. In: Issacs NW and Taylor MR (eds) *Crystallographic Computing 4*, pp 126–140. Oxford: Oxford University Press.

Brünger AT (1991) A unified approach to crystallographic refinement and molecular replacement. In: Moras D, Podjarny AD and Thierry JC (eds) *Crystallographic Computing 5*, pp 392–408. Oxford: Oxford University Press.

Hall SR, King JDS and Stewart JM (1995) LSLS. In: *Xtal 3.4 User's Manual*. Lamb, Perth: University of Western Australia.

Press WH, Flannery BP, Teukolsky SA and Vetterling WT (1992) *Numerical Recipes: The Art of Scientific Computing*. New York: Cambridge University Press.

Prince E (1994) *Mathematical Techniques in Crystallography and Materials Science*. New York: Springer-Verlag.

Schwarzenbach D, Abrahams SC, Flack HD *et al* (1989) Statistical descriptors in crystallography. Report of the International Union of Crystallography Subcommittee on Statistical Descriptors. *Acta Crystallographica, Section A* **45**: 63–75.

Schwarzenbach D, Abrahams SC, Flack HD, Prince E and Wilson AJC (1995) Statistical descriptors in crystallography. II. Report of a working group on expression of uncertainty in measurement. *Acta Crystallographica, Section A* **51**: 565–569.

Sheldrick GM (1997) SHELXL-97. In: *The SHELX-97 Manual*. Gottingen: University of Göttingen.

Watkin DJ (1994) Lead article: The control of difficult refinements. *Acta Crystallographica, Section A* **50**: 411–437.

Watkin DJ, Prout CK, Carruthers JR and Betteridge PW (1996) *CRYSTALS Issue 10*. Oxford: Chemical Crystallography Laboratory, University of Oxford.


# Sulfur NMR, Applications

*See* **Heteronuclear NMR Applications (O, S, Se, Te).**

# Surface Induced Dissociation in Mass Spectrometry

**SA Miller** and **SL Bernasek**, Princeton University, NJ, USA

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

Analytical tandem mass spectrometry (MS/MS) relies on the ability to activate and dissociate ions in order to identify or obtain structural information about an unknown compound. The most common means of ion activation in tandem mass spectrometry is collision-induced dissociation (CID). CID uses gas phase collisions between the ion and neutral target gas to cause internal excitation of the ion and subsequent dissociation. Surface-induced dissociation (SID) is analogous to the CID experiment except a surface is substituted for the collision gas as shown in **Figure 1**. SID offers several advantages as a means of ion activation in tandem mass spectrometry, including fine control of the internal energy deposition, efficient translational to vibrational energy conversion, large average internal energy deposition, and applicability to both small organic and large biological molecules. Since 1985, low energy polyatomic SID and its accompanying processes have been the subject of much research and continue to gain interest in the mass spectrometry community as an alternative means for ion activation, to study interesting ion/surface reactions, and to cause selective surface modification.

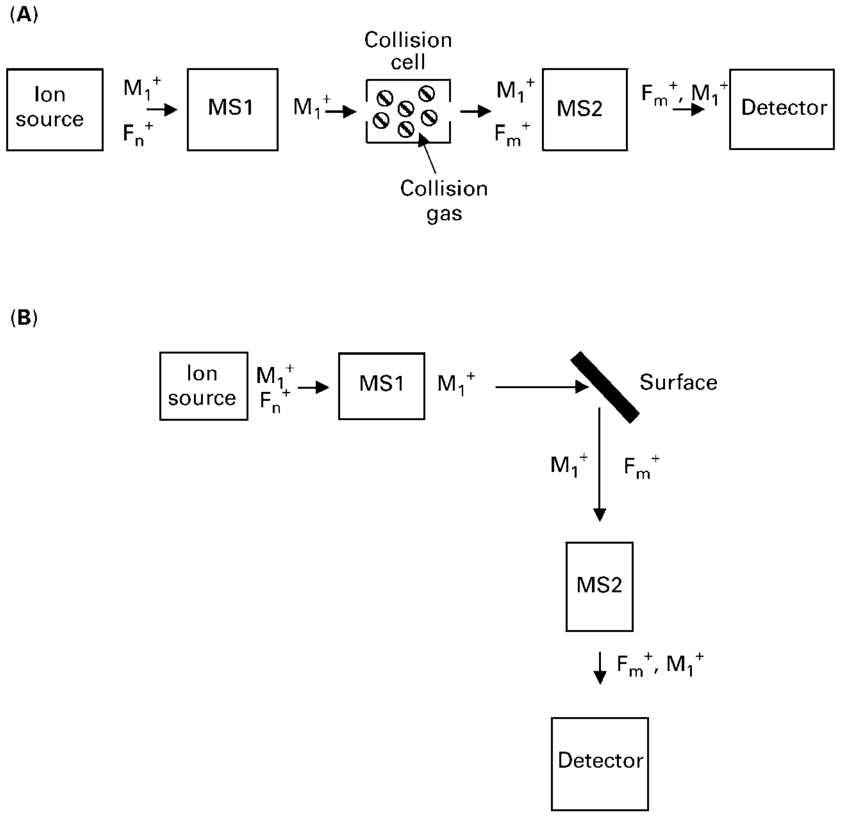
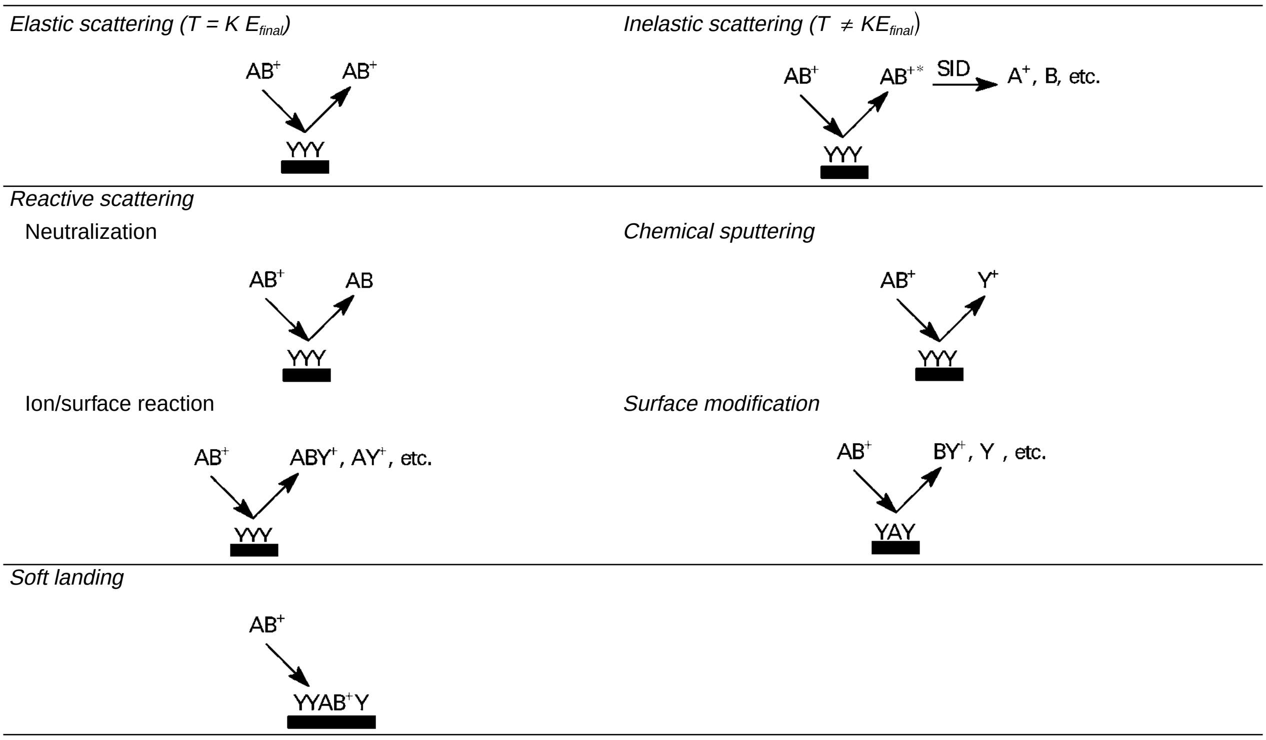



**Figure 1** Simplified schematic of (A) a CID tandem mass spectrometer and (B) a SID tandem mass spectrometer where MS1 and MS2 are the mass analysers, $M_1^+$ is the parent or projectile ion, $F_n^+$ are fragment ions of $M_1^+$ generated in the ion source, and $F_m^+$ are fragment ions formed as a result of (A) CID and (B) SID.

**Table 1** Low energy polyatomic ions/surface collision processes




## Low energy polyatomic ion/surface collisions

When an ion collides with a surface, several processes may occur including elastic, inelastic, and reactive scattering or soft landing. **Table 1** lists possible events which occur for low energy polyatomic ion collisions with organic covered surfaces. Note that 'low energy' refers to the component of kinetic energy normal to the surface which is usually in the range of 1 to 100 eV. The parent or projectile ion, $AB^+$, is accelerated (up to a few hundred eV) towards a substrate which has been chemically modified with an organic overlayer, Y, such as a self-assembled monolayer (SAM) surface (**Figure 2**). Since the projectile ion has a low normal component of collision energy, it interacts only with the outer layer of the adsorbate and not directly with the substrate. The ion will collide with the organic surface and may scatter from the surface without losing any kinetic energy. Generally, the polyatomic ion scatters from the surface with a loss of kinetic energy which is distributed into internal energy of $AB^+$ and of the organic overlayer Y as described in the following equation,

$$Q = T - \mathrm{KE}_{\mathrm{final}} = V + \varepsilon_{\mathrm{surface}} \quad [1]$$

where $Q$ is the inelasticity of the collision event, $T$ and $\mathrm{KE}_{\mathrm{final}}$ are the initial and final kinetic energy of the projectile ion $AB^+$, $V$ is the total vibrational energy of the ion $AB^+$, and $\varepsilon_{\mathrm{surface}}$ is the internal energy deposited in the surface (radiative losses are ignored). If $AB^+$ becomes sufficiently activated, it will dissociate into fragment ions which are characteristic of the structure and chemical composition of the projectile ion.

Several processes are representative of reactive scattering, including neutralization, chemical sputtering, ion/surface reaction, and surface modification. Most reactive scattering processes occur through electron transfer from the surface to the ion. If no other processes occur after the electron transfer, then the only product formed is the neutralized projectile ion, AB, and no ions are detected. Chemical sputtering occurs when sufficient energy is deposited in the surface following the electron transfer and collision event causing adsorbate ions to be liberated from the surface. The incident ion or a fragment ion can incorporate an atom or group of atoms from the surface to form an ion/surface reaction product. Ion/surface reactions may also cause chemical modification of the adsorbate through atom or molecule exchange reactions. One must distinguish between surface modifications that result from damage which occurs to an adsorbate during ion/surface collisions

(for example, chemical sputtering) and a selective chemical modification that results from an ion/surface exchange reaction. Finally, specially chosen projectile ions can be soft landed into the adsorbate, specifically in fluorocarbon and hydrocarbon SAM surfaces, upon collision at very low energies (~1–10 eV). Remarkably, the deposited intact projectile ion remains in the surface as an ion protected by the adsorbate layer. This matrix-isolated ion has been shown to survive for several days in vacuum and a few days in ambient lab air.

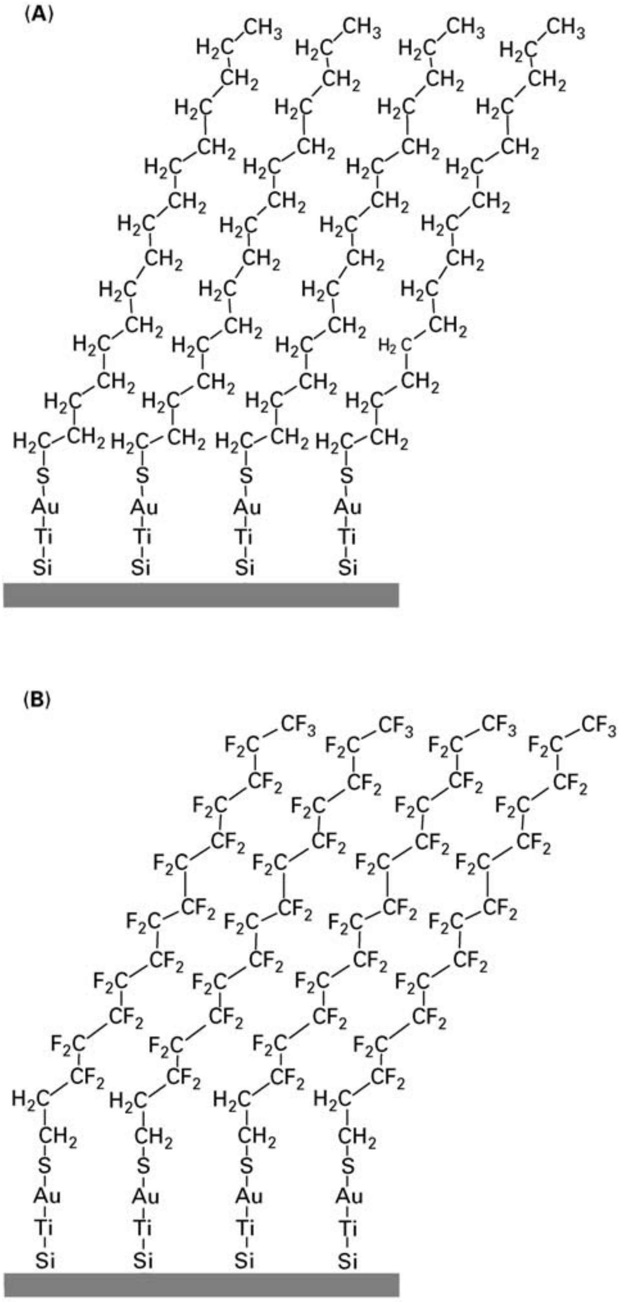



**Figure 2** Example of (A) an alkanethiol self-assembled monolayer surface and (B) a fluorinated alkanethiol self-assembled monolayer surface. Typical substrates are glass or silicon with a thin Ti layer (~100 Å) followed by a layer of gold (~1000 Å). Reproduced with permission from Miller SA, PhD Thesis, Purdue University, 1997.

## Surface-induced dissociation instrumentation

The fundamental requirement for performing an SID experiment is a tandem mass spectrometer. Several tandem mass spectrometers have been developed using various combinations of mass analysers and geometries to perform SID. **Table 2** lists several SID systems (by no means an exhaustive list), which have been organized by on-axis (or in-line) and off-axis geometry. On-axis geometry refers to a tandem mass spectrometer where the exit slit or aperture of the first mass analyser is in-line with the entrance slit or aperture of the second mass analyser. In the off-axis geometry, there is no common axis between the exit and the entrance of the first and second mass analyser. The on-axis and off-axis groups can be further divided into instrument collision geometry, that is, large incident angle/low collision energy and glancing incident angle/high collision energy. The incident angle is defined here as being referenced to the surface normal. Finally, a special group of mass spectrometers which trap ions have also been used to perform SID experiments. Quadrupole ion traps and Fourier transform ion cyclotron resonance mass spectrometers perform the tandem mass spectrometry experiment in a single mass analyser, i.e. the tandem mass spectrometry experiment is performed in one analyser with each MS stage separated in time rather than separated in space as for a two-analyser system.

A simplified example of an SID tandem mass spectrometer is shown in **Figure 1B**. This system utilizes two mass analysers (MS1 and MS2) perpendicular to one another to mass select the projectile ion (MS1) and analyse the scattered product ions (MS2). Ions are generated in the ion source, accelerated, and focused into MS1 where the projectile ion, $M_1^+$, is mass selected. $M_1^+$ is then further accelerated or decelerated (depending on the desired collision energy) and focused onto the surface. The product ions scattered from the surface are accelerated toward MS2 and the SID mass spectrum is then collected by scanning MS2 and measuring the ion current at the detector.

There are several design considerations which must be addressed when constructing an SID

instrument, including vacuum conditions and instrument functionality. The first consideration is determining the vacuum requirement which will be sufficient for performing an SID experiment. For most SID instruments, the vacuum is typically in the range of $10^{-7}$ to $10^{-10}$ torr ($1.3 \times 10^{-5}$ to $1.3 \times 10^{-8}$ Pa) and is dictated by the type of surface to be used in the experiment. If a clean and well characterized single crystal is used as the SID target, then ultrahigh vacuum (UHV) conditions must be met ($<10^{-9}$ torr). For an organic covered surface, such as a self-assembled monolayer surface, less stringent vacuum conditions ($<10^{-6}$ torr) will suffice.

The type of ion source used in an SID experiment depends on the projectile ion under study. For instance, if volatile, small organic compounds are of interest then an electron ionization (EI) and/or chemical ionization (CI) source will be sufficient. However, if one wishes to study biological compounds then a spray or desorption ionization method will be required, such as electrospray (ESI), atmospheric pressure chemical ionization (APCI), desorption chemical ionization (DCI), or matrix-assisted laser desorption ionization (MALDI).

As seen in **Table 2**, several combinations of mass analysers have been used in SID instruments, each with their own advantages and disadvantages. The type(s) of mass analysers used in an SID system depends to some extent on the class of projectile ions to be studied, as with the method of ionization, but more so on the overall performance mass analyser. One must consider the resolution, sensitivity, size, ruggedness, ease-of-use, cost, and the compatibility of one mass analyser with another as well as the information sought in the SID spectra. The authors refer the reader to other articles in this text that discuss advantages and disadvantages of the various mass analysers.

**Table 2** Low energy ion/surface collision instruments.

| |
|---|
| *On-axis* |
| Large incident angle/low collision energy |
| 1. Multisector |
| 2. Tandem quadrupole |
| 3. Hybrid sector/quadrupole |
| Glancing incident angle/high collision energy |
| 1. Hybrid sector/time-of-flight |
| 2. Hybrid sector/quadrupole |
| 3. Multi-Wien filter |
| 4. Multisector |
| *Off-axis* |
| Large incident angle/low collision energy |
| 1. Tandem quadrupole |
| 2. Tandem time-of-flight |
| 3. Hybrid sector/time-of-flight |
| 4. Hybrid sector/quadrupole |
| 5. Hybrid Wien filter/quadrupole |
| *Trapping instruments* |
| 1. Paul quadrupole ion trap |
| 2. Fourier transform ion cyclotron resonance mass spectrometers. |

The extent of fragmentation in an SID experiment is highly dependent on the component of kinetic energy normal to the surface. Thus, having control of the collision energy (kinetic energy of the projectile) is necessary. The range of ion kinetic energy will depend on the collision geometry of the instrument. For example, if glancing incident angles are used then higher overall collision energies will be needed (hundreds of eV to keV), while more normal collisions will require a lower ion kinetic energy (1–100 eV). It has been shown that nearly identical SID spectra can be obtained from both glancing and large incident angle collisions as long as the component of kinetic energy normal to the surface is similar. While the angle of incidence is not that important analytically, subtle fundamental information can be obtained by controlling both the kinetic energy of the projectile and the angle of incidence. More fundamental information about the SID process and other ion/surface collision processes can be determined by measuring the energy of the scattered ions as well as their angular distribution. Hence, it is of interest to control the scattering angle independently of the incident angle and to have the ability to measure the kinetic energy of the scattered ions.

Surface science techniques are useful for monitoring changes in the structure and composition of the adsorbate before, during, and after ion/surface collisions. Common techniques include high resolution electron energy loss spectroscopy (HREELS), low energy electron diffraction (LEED), Auger electron spectroscopy (AES), secondary ion mass spectrometry (SIMS), reflection absorption infrared spectroscopy (RAIRS), X-ray photoelectron spectroscopy (XPS), and ultraviolet photoelectron spectroscopy (UPS). While little ion structural information is gained using these methods, changes in the adsorbate are of interest when studying ion/surface reactions, especially those which lead to surface modification.

A few examples of SID instruments are shown in **Figures 3** to **6** which incorporate some or all of the considerations discussed above. **Figure 3** is an off-axis large incident angle/low collision energy hybrid sector/quadrupole SID instrument of the BEEQ configuration (B = magnetic sector, E = electrostatic

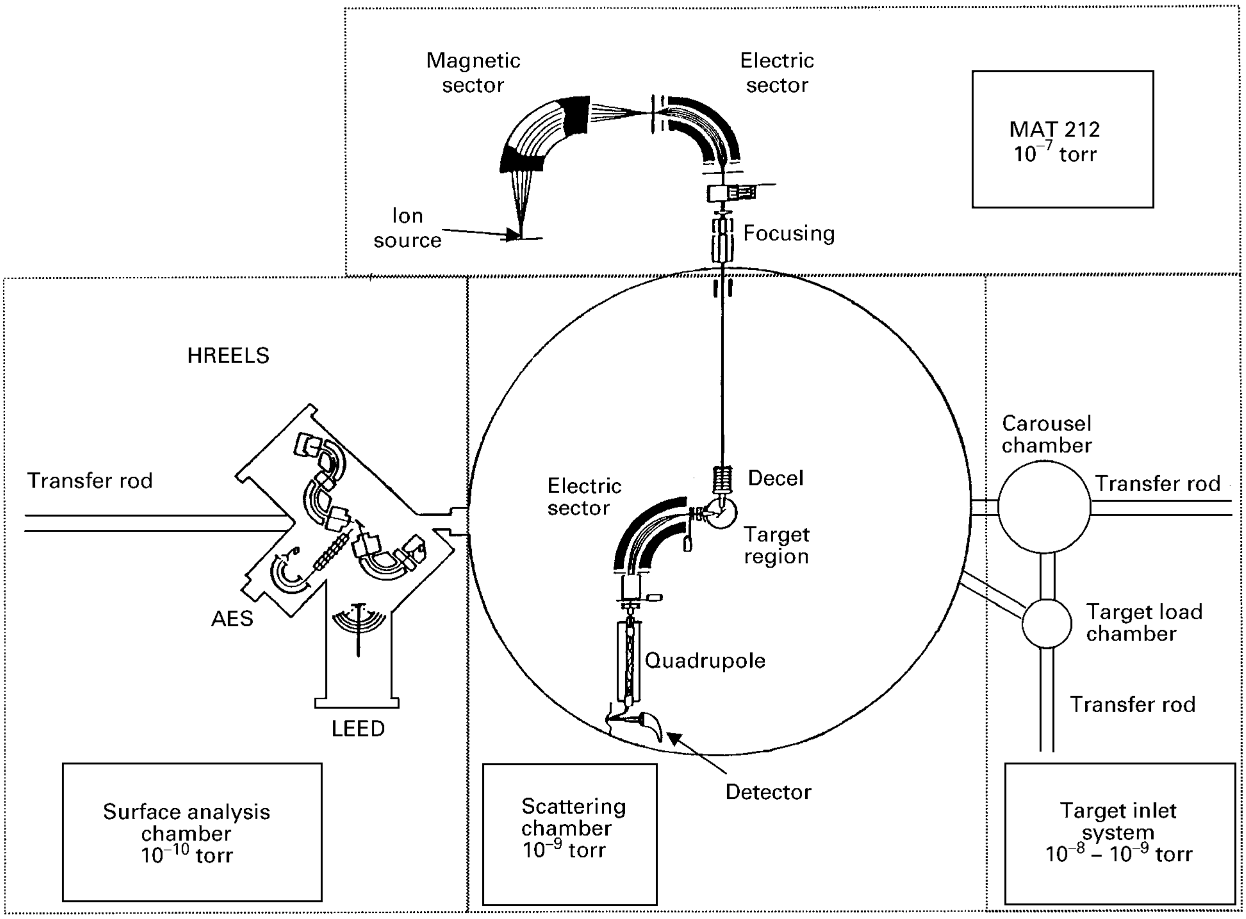



**Figure 3** Overview of a hybrid BEEQ tandem mass spectrometer. Reproduced with permission from Miller SA, PhD Thesis, Purdue University, 1997.

sector, and Q = quadrupole mass filter). This system has the ability to change the scattering angle independently of the incident angle, to measure the kinetic energy of the scattered product ions with the post-collision electrostatic sector, and to do traditional surface analysis in a separate UHV chamber. **Figure 4** presents the SIMION calculated ion trajectories for a custom-built, on-axis SID device used on a commercial four sector tandem mass spectrometer. (Dahl DA and Delmore JE (1989) *SIMION PC/PC2 version 4.02*, Idaho National Engineering Laboratory, EG and G Idaho Inc.) The mass analysed projectile ion beam enters from the left of the device, strikes a target, and the scattered ions are extracted to the right side of the device into the final mass analyser. The SID device can be used in two different modes including a single deflection and dual deflection mode. It is clear in the SIMION calculation that the dual deflection mode (**Figures 4B** and **4C**) produces a more focused ion beam. An advantage of this SID system is that the SID device can be exchanged with a commercial CID device and direct comparisons of SID to high energy CID collisions can be observed. **Figures 5** and **6** show two types of off-axis instruments that maintain large incident angle/low collision energy geometry but use different tandem mass analysers. The tandem time-of-flight system in **Figure 5** has the ability to use several ionization techniques making it amenable to projectile ions ranging from small organics to large biological compounds. This system also employs delayed extraction technology to improve the resolution in the SID product ion mass spectra. A unique feature of the tandem quadrupole system in **Figure 6** is that it is contained on a single 8″ conflat flange and is easily attached to a standard UHV surface analysis system. Most currently used SID systems utilize off-axis large incidence angle/low collision energy geometry due to enhanced overall performance compared to on-axis systems.

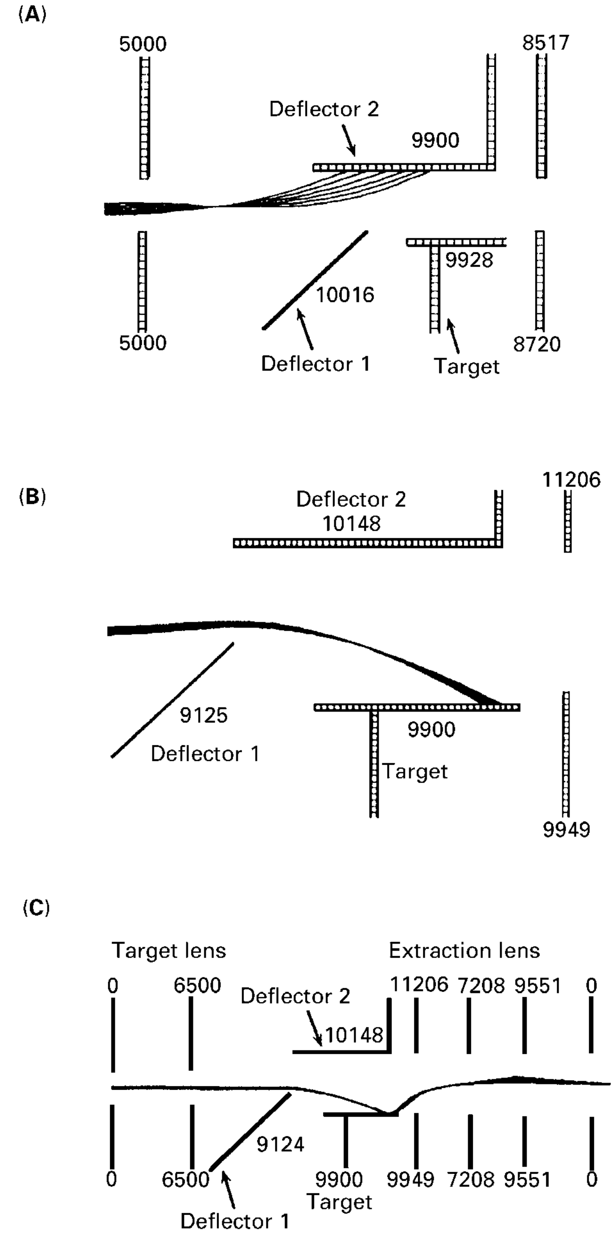



**Figure 4** SIMION ion trajectory calculations of an on-axis SID device with (A) precursor ion deflection in single deflection mode, (B) precursor ion deflection in double deflection mode, and (C) precursor and product ion trajectories in double deflection mode. Numerical labels represent the voltages applied to the electrodes. Reproduced with permission from Durkin DA, Schey KL, (1998) Characterization of a new in-line SID mode and comparison with high energy CID. *International Journal of Mass Spectrometry and Ion Processes* **174**: 63–71.

**Table 3** Characteristics of the SID process.

| | |
|---|---|
| 1. | Internal energy ($V$) deposition is variable depending on the collision energy ($T$) |
| 2. | Internal energy deposition can be made large, more than 10 eV is readily accessible |
| 3. | Internal energy distribution is narrow, typically 4 eV FWHM, and is approximately Gaussian in shape |
| 4. | The efficiency of converting translational energy to internal energy is dependent on the nature of the surface and is approximately 20–30% for fluorocarbon surfaces and 10–20% for hydrocarbon surfaces |
| 5. | The SID efficiency varies over a wide range 1–75% but is typically more than 10% for most polyatomic ion/organic covered surface pairs. |
| 6. | Fragment ions are similar to those noted in CID experiments |
| 7. | SID is proving to be applicable to large biomolecules |

## Surface-induced dissociation

The characteristics associated with the SID process are listed in **Table 3** and emphasize the analytical utility of SID as a method of ion activation. The first characteristic of SID listed in **Table 3** is the ability to control the amount of fragmentation, or internal energy, by adjusting the collision energy. **Figure 7** displays the SID mass spectrum resulting from the collision of p-methyl phenetole molecular ion (136 Th) with a fluorocarbon self-assembled monolayer (F-SAM) surface at two different collision energies. This example illustrates the effect that collision energy has on the extent of fragmentation. At a 15 eV collision energy, only the molecular ion and one fragment ion, *m*/*z* 108 (loss of $C_2H_4$), are observed. When the collision energy is increased to 30 eV, fragmentation occurs giving rise to a high energy fragment ion, *m*/*z* 29 ($C_2H_5^+$ and/or $CHO^+$) as the base peak and *complete loss* of the molecular ion at *m*/*z* 136. The fragment ions in **Figure 7B** are the same as those observed in the electron ionization mass spectrum of p-methyl phenetole but with different ion intensities.

A better fundamental understanding of the SID process is obtained by quantitating the partitioning of the transitional energy of the ion into internal energy of the ion, internal energy of the surface, and the kinetic energy of the scattered ions. In Equation [1], partitioning of the projectile kinetic energy in an ion/surface collision was described. Rearranging this equation to indicate energy conservation,

$$T = \mathrm{KE}_{\mathrm{final}} + V + \varepsilon_{\mathrm{surface}} \qquad [2]$$

the ion transitional energy is the sum of the final kinetic energy of the scattered products, the total internal energy of the projectile ion and the energy absorbed by the surface. Despite the fact that experimentally determined values for $\mathrm{KE}_{\mathrm{final}}$ and $V$

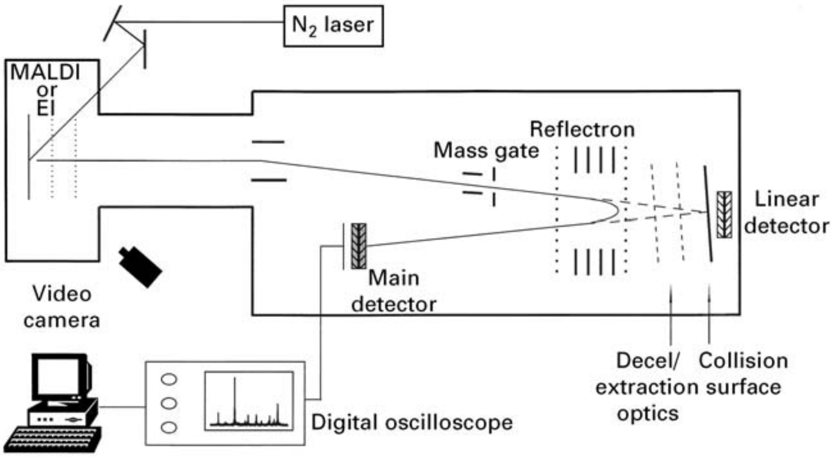



**Figure 5** Overview of a tandem time-of-flight SID instrument. Reproduced with permission from Riederer Jr, DE, PhD, Department of Chemistry, University of Missouri, Columbia.

are dependent on the nature of the surface and projectile ion, general conclusions can be made about the distribution of the projectile ion kinetic energy. For example, the kinetic energy of the scattered product ions represents a small fraction of *T*, typically less than 10%. The value obtained for *V* is approximately 10–30% and has been shown to be linear over a wide range of collision energies. The remaining energy, 60–80%, is absorbed by the surface.

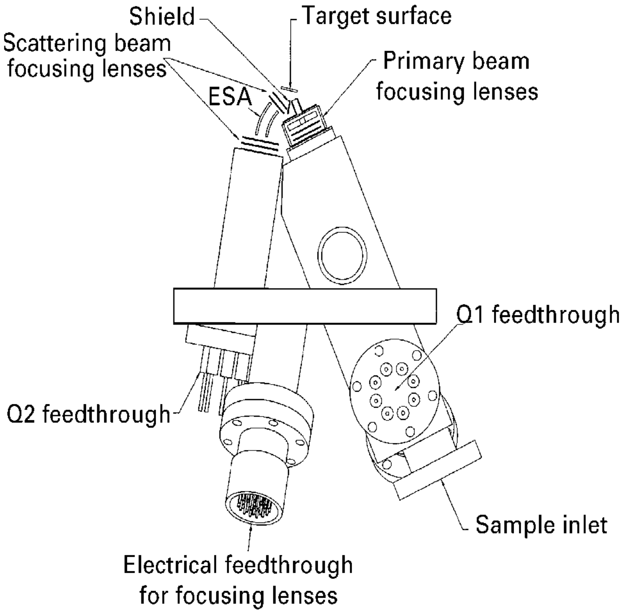



**Figure 6** Overview of a tandem quadrupole SID instrument with a cross-sectional view of the target and ion focusing region. The ion optical element labelled ESA represents an electrostatic analyser. Reproduced with permission from Phelan LM, PhD Thesis, Princeton University, 1998.

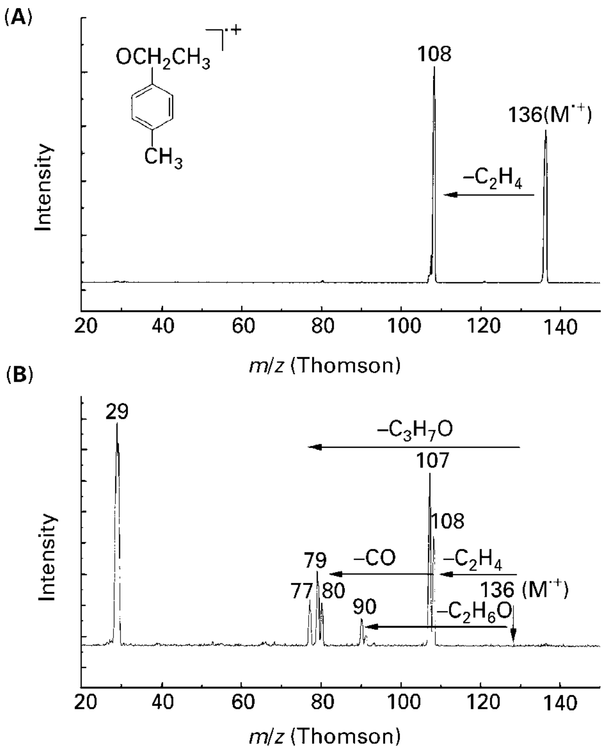



**Figure 7** Surface-induced dissociation ion spectra of the p-methyl phenetole molecular ion (*m*/*z* 136) upon collision with an F-SAM surface at (A) 15 eV and (B) 30 eV. Reproduced with permission from Cooks RG and Miller SA, Collision of ions with surfaces. In: Jennings KR (ed) *NATO ASI Series*, Dordrecht, The Netherlands, Kluwer Academic Publishers, (in press).

### Conversion of ion transitional energy into internal energy (T→V)

A quantitative variable used to judge the effectiveness of a collision ion activation technique is the translational to internal energy conversion efficiency or *T*→*V*%. The internal energy absorbed from a collision event is not single-valued, rather a distribution of internal energies is observed. Internal energy distributions can be measured using a thermometer molecule method which relies on a distinct set of dissociation pathways of known critical energy for a particular ion. The intensities of the fragment ions in the SID mass spectrum of the thermometer ion and the critical energies required to generate each fragment ion are used to calculate the internal energy distribution (P(ε) versus internal energy, ε) deposited in the collision event. Metal carbonyls are often used due to their simple fragmentation pathway by consecutive losses of CO and well known critical energies. **Figure 8A** depicts the SID product ion spectrum resulting from 30 eV collisions of $W(CO)_6^{\bullet +}$ (352 Th) with an 8 μm thick liquid perfluoropolyether (PFPE) surface, $F[CF(CF_3)CF_2O]_{27(avg)}CF_2CF_3$. The internal energy distribution is then calculated from the mass spectrum in **Figure 8A** using the thermometer molecule method and the result is shown in **Figure 8B**. The distribution is approximately Gaussian in shape with a full width at half maximum (FWHM) of 4.4 eV and an average internal energy of 5.5 eV. The *T*→*V* conversion efficiency is calculated to be 18%. Although the values calculated in **Figure 8B** are for one particular ion/surface system, it is a general characteristic of SID to produce internal energy distributions which are approximately Gaussian in shape with FWHMs of approximately 4 eV.

Another method used to determine SID internal energy distributions is the deconvolution method. This method relies on the fact that the mass spectrum, breakdown curve (normalized fragment ion intensities as a function of internal energy), or internal energy distribution can be calculated if two of the three are known. For example, the mass spectrum can be calculated by convoluting the breakdown curve with the internal energy distribution. Since the benzene molecular ion has been well studied and breakdown curves have been measured, the internal energy distribution can be determined by deconvoluting the breakdown curve and the recorded SID mass spectrum of the benzene molecular ion. The deconvolution method has reported slightly higher T→V conversion efficiencies than the thermometer molecule method for similar surfaces. Small variations in the *T*→*V* conversion efficiency for a given surface have been noted with changes in the chemical composition of the projectile ion (see **Table 4**).

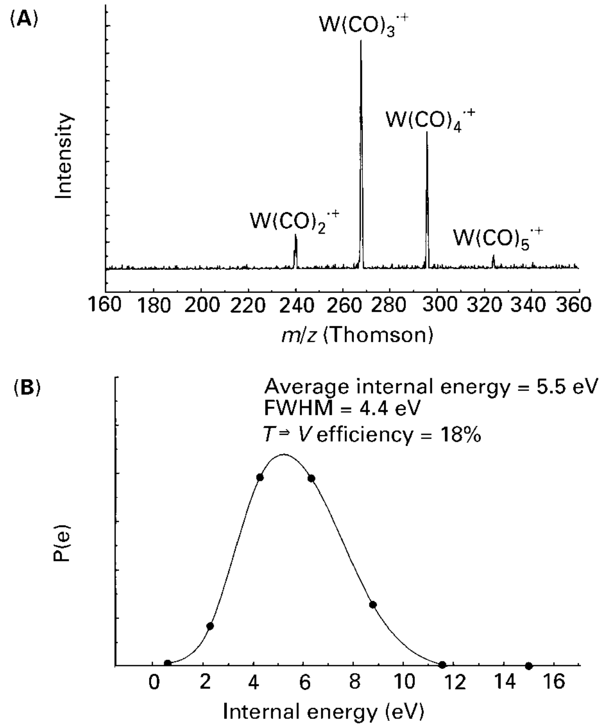



**Figure 8** (A) Surface-induced dissociation mass spectrum produced from 30 eV collisions of $W(CO)_6^{\bullet +}$ (*m/z* 352) with an 8 μm thick liquid perfluoropolyether surface. (B) Calculated internal energy distribution for the spectrum in (A) using the thermometer molecule method. **Figure 8A** reproduced with permission from Pradeep T, Miller SA and Cooks RG (1993) Surface-induced dissociation from a liquid surface *Journal of American Society for Mass Spectrometry* **4**: 769–773.

SID internal energy distributions are very different from those observed in CID. CID internal energy distributions are typically broad in nature and with uncharacteristic shapes. The narrow internal energy distribution found in SID gives a finer control over the energy that is deposited into an ion and thus more control over the extent of fragmentation than CID.

As a result of the high *T*→*V* conversion efficiencies in SID, large amounts of internal energy can be readily deposited into the projectile ion. For example, experiments using the benzene molecular ion have shown the ability of SID to impart more than 20 eV of internal energy. This large internal excitation and the high *T*→*V* conversion efficiency are due to the fact that in SID, the projectile ion cannot 'miss' its target, that is to say, the ion will always experience 'head-on' or small impact parameter collisions, as opposed to CID where ions are more likely

to undergo larger impact parameter collisions causing less excitation.

### Surface effects

Surface effects in ion/surface collision processes are very prominent and continue to be actively investigated. For example, surfaces covered with hydrocarbon pump oil were found to be more effective in reducing neutralization (or increasing secondary ion yield, that is, the ratio of the total secondary ion current to the primary ion current measured at the surface) than clean metal surfaces. In addition, it was also found that the organic nature of the surface affected the energetics of the ion/surface collision, especially the $T \rightarrow V$ conversion efficiency.

**Table 4** Effect of target nature on the estimated translation-to-vibrational energy conversion[a]

| *Ion* | *Target (Co.Energy, eV)* | *$T \rightarrow V$(%)* | |
|---|---|---|---|
| $Cr(CO)_6$ | F-SAM (25) | 19 | |
| $W(CO)_6$ | F-SAM (30) | 18 | Fluorine |
| $W(CO)_6$ | F-SAM (50) | 20 | targets |
| $W(CO)_6$ | PFPE (50) | 18 | |
| Ferrocene | F-SAM (20–50) | 24 (20)[b] | |
| Benzene | F-SAM (10–70) | 28 | |
| $Cr(CO)_6$ | D-SAM (25) | 12 | |
| Ferrocene | D-SAM (20–80) | 15 (13)[b] | |
| $Cr(CO)_6$ | Fc-SAM (25) | 11 | |
| n-Butylbenzene | Fc-SAM(25) | 12 | Hydrocarbon |
| $Cr(CO)_6$ | SS (25) | 11 | like targets |
| $W(CO)_6$ | SS (25–100) | 13 | |
| $Fe(CO)_5$ | SS (25–60) | 13 | |
| $Et_4Si$ | SS (10–100) | 12 | |
| Benzene | H-SAM (10–70) | 17 | |
| $W(CO)_6$ | Alkenyl-termi- ated SAM | 12 | |
| | 30–70 | | Other SAMs |
| Ferrocene | Amino-termi- nated SAM (25) | 13 | |
| Ferrocene | Cyano-termi- nated SAM (25) | 12 | |
| Ferrocene | Si(100) (20–90) | 13 | Single crystal |
| Pyridine | Ag(111) (32) | 20 | targets |

[a] Where PFPE is perfluoropolyether, Fe is ferrocene, D is deuterated, and SS is a stainless steel surface with an adventitious hydrocarbon overlayer.

[b] Denotes that the ($T \rightarrow V$%) has been corrected for initial internal energy from the ionization event.

Prior to using organic covered surfaces, SID was performed on clean metal surfaces. The SID efficiency (ratio of the sum of total fragment ion current to the projectile ion current measured at the surface) of these ion/surface systems was often very poor, <1%. For example, collisions of the ferrocene molecular ion with a clean Si(100) surface at 10 eV result in an SID efficiency of $10^{-2}$. This can be rationalized by examining the thermochemistry of an electron transfer reaction for the projectile ion and the surface. For example, a typical ionization energy for an organic compound is 10 eV and that of a clean metal surface is on the order of 3 or 4 eV (known as the work function of the surface). When the ion approaches a clean metal surface, the enthalpy of the electron transfer reaction (from the surface to the projectile ion) is very exothermic, 6–7 eV, and the neutralization of the projectile ion will be very efficient. However, when an organic layer covers the metal substrate, it changes the thermochemistry of the surface from a work function of a few eV to one that is of the order of the neutralization energy of the projectile ion. This change in thermochemistry reduces the exothermicity and extent of neutralization.

Self-assembled monolayer surfaces (SAMs) of n-alkane thiols on gold have become a surface of choice for studying the SID process. These surfaces offer an array of variability in the chemical functionality of the surface, they are stable in vacuum and air, can be well ordered, characterizable, and are relatively simple to fabricate. The most common of the SAM surfaces used in SID studies are the hydrocarbon SAM (H-SAM) and fluorocarbon SAM (F-SAM) surfaces shown in **Figure 2**. Continuing the thermochemical argument for the electron transfer process above, it would be expected that a fluorocarbon surface would yield higher SID efficiencies due to the larger ionization energy for fluorocarbon molecules, ~13 eV, making the electron transfer reaction endothermic for typical organic projectile ions. Note that although the reaction is endothermic, the translational energy of the projectile ion can be used to drive endothermic reactions. Studies have shown that fluorocarbon surfaces have higher total secondary ion yields than hydrocarbon surfaces, with some reports showing the F-SAM surface to have efficiencies of 55–75% as contrasted with 1–35% for H-SAM surfaces. The variation in secondary ion yields for a given SAM surface correlates with the neutralization energy of the projectile ion, that is, the overall thermochemistry of the electron transfer reaction.

The nature of the surface is an influential factor in the $T \rightarrow V$ conversion efficiency for a particular ion/surface pair. **Table 4** lists several thermometer molecules that have been used to determine the $T \rightarrow V$

conversion efficiency for several hydrocarbon and fluorocarbon surfaces. Fluorocarbon surfaces are more efficient at converting translational energy to internal energy than hydrocarbon surfaces due to a more rigid structure in the F-SAM surface and more massive end-groups. **Figure 9** illustrates the greater effectiveness of the F-SAM surface at dissociating the pyrazine molecular ion than a perdeuterated SAM (D-SAM) surface at the same ion kinetic energy of 25 eV. The F-SAM surface also exhibits a greater signal-to-noise ratio and SID efficiency as a result of reduced neutralization. Note that the pyrazine molecular ion is reactive with the D-SAM surface producing the ion/surface reaction product, $[M+CD_3]^+$, and the ion/surface reaction fragment $[M+D-HCN]^+$. It is interesting to note that the $T \rightarrow V$ conversion efficiency of the liquid PFPE film and an F-SAM surface are almost identical (similar chemical composition) but their physical state of matter is different. There is very little difference in SID and ion/surface reactions between the liquid PFPE and the F-SAM surface for a large variety of projectile ions, supporting the hypothesis that the ion only interacts with the outer layer of the surface.

### Applications

Surface-induced dissociation produces structurally diagnostic ions similar to those found in the analogous gas phase CID experiment. For example, **Figure 10A** shows the SID mass spectrum of the pyrene molecular ion at a collision energy of 100 eV with an F-SAM surface. The spectrum displays significant fragmentation by diagnostic $C_2$ losses. **Figure 10B** displays the CID of pyrene at 100 eV under single collision conditions with Ar and only produces two minor fragment ions. Note that this comparison of SID to CID is made based on specific experimental conditions, namely single collision events for CID. CID is often performed under multiple collision conditions which imparts a larger average internal energy deposition yielding more diagnostic fragment ions. One must be careful when comparing SID and CID since the experimental conditions can greatly bias the observed results. Nonetheless, both ion activation methods provide useful information about ion structure.

Fragment ions produced by SID have been used to distinguish gaseous isomeric ions and to elucidate the structure of biomolecules. The unique characteristics of SID, narrow internal energy distribution and variable control of internal energy, make it a powerful tool for studying isomeric ions. Before low energy ion/surface collisions were used as a means of ion activation, isomeric ion distinction was performed by mass spectrometric techniques such as CID, charge exchange, and kinetic energy release measurements, all of which used high energy collisions. Isomers of $C_5H_6^+$, $C_6H_6^{\bullet+}$, $C_6H_6^{2+}$, $C_2H_3S^+$, $C_2H_4O^{\bullet+}$, and $C_3H_4^{\bullet+}$ have been distinguished by low energy SID. It is worth noting that ion/surface reaction products can serve as a complementary tool when studying isomers in the gas phase due to the often selective nature of ion/surface reactions.

Activation and dissociation of large biomolecule ions is currently the focus of many investigators interested in analytical applications of SID. As noted in the introduction, the need to activate and dissociate increasingly larger molecules is a constant challenge in tandem mass spectrometry. SID offers the advantage of large $T \rightarrow V$ conversion efficiencies and narrow internal energy distributions. While neutralization in SID can pose a problem, improvements are made by choosing a surface which minimizes the extent of neutralization. Thus, fluorocarbon surfaces are often used in SID biomolecule applications. SID mass spectra of peptides yield results consistent with low energy CID, that is, the formation of peptide backbone fragments such as *a*-, *b*-, and *y*-type ions. In addition to backbone cleavages, some peptides exhibit side chain specific cleavage ions, *d*- and *w*-type, in low energy SID experiments which are normally characteristic of high collision energy (keV) CID data. **Figure 11** illustrates the diagnostic fragments resulting from collisions of the doubly protonated melittin ion with an F-SAM surface at 110 eV. The spectrum consists of mainly backbone cleavage fragments (*a*-, *b*-, and *y*-type) with nearly a complete *y*-type series from $y_2$ to $y_{25}$. A complete primary sequence of the melittin ion, with the exception of two residues, can be determined from this single SID mass spectrum. These experiments are also lending insight to the energetics and fragmentation mechanisms of peptides due to the controllable nature of the internal energy deposition characteristic of the SID method.

### Surface-induced dissociation mechanisms

In most cases, polyatomic fragment ions noted in SID are similar to those found in CID and can be explained by unimolecular fragmentation pathways. This general observation has led to the prediction of a two-step SID mechanism involving the collision which causes internal excitation followed by a delayed fragmentation some distance and time away from the collision event. This concept for the mechanism of SID of polyatomic *organic* ions is fairly well accepted, although one recent study shows evidence for dissociation at the surface interface in a prompt

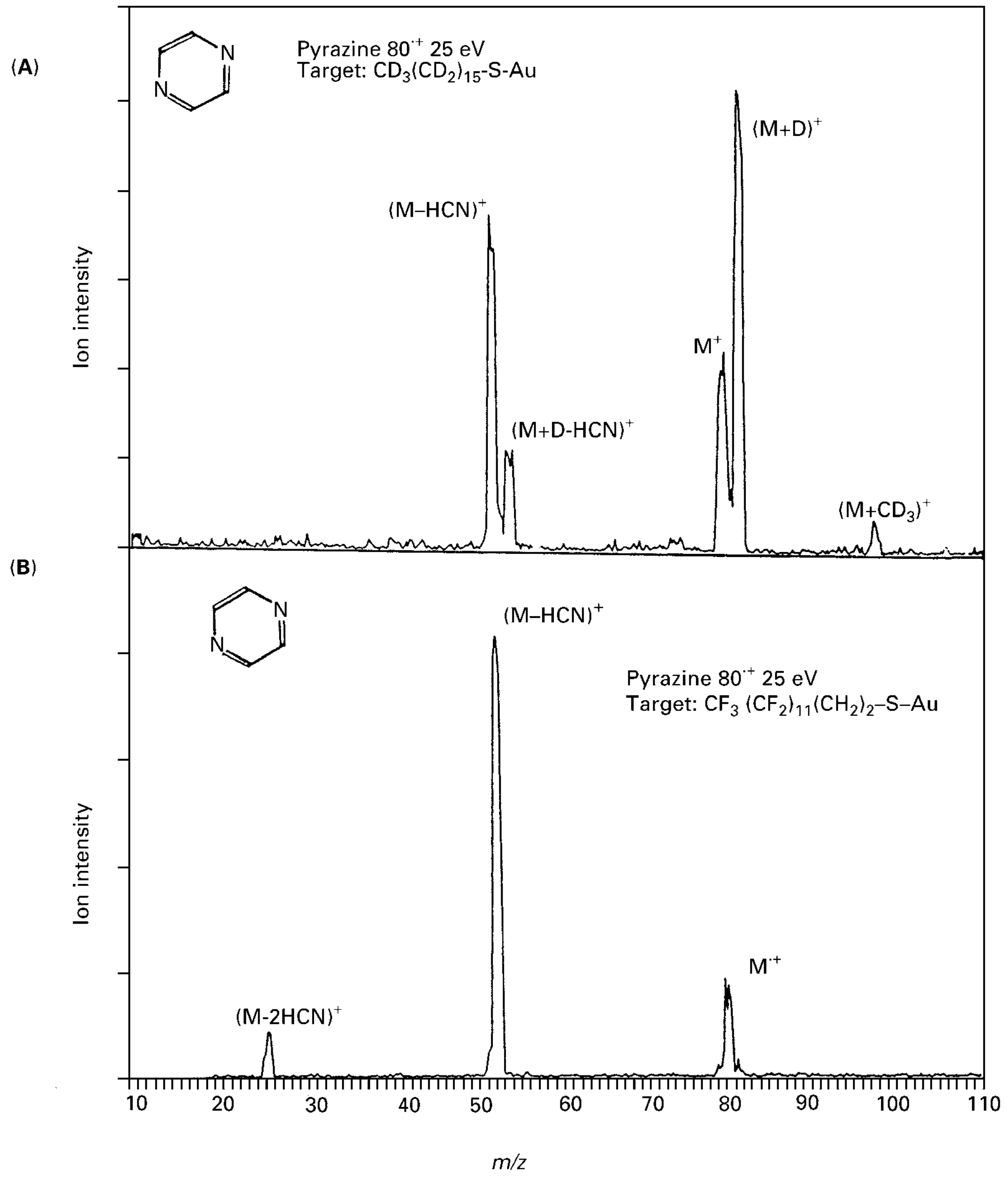



**Figure 9** Product ion spectra resulting from 25 eV collisions of the pyrazine molecular ion (*m/z* 80) with (A) a D-SAM surface and (B) an F-SAM surface. Reproduced with permission from Winger BE, Laue H-J, Horning *et al*, (1992) Hybrid BEEQ tandem mass spectrometer for the study of ion/surface collision processes *Review of Scientific Instruments* **63**: 5613–5625.

or ‘shattering’ mechanism for a specific ion/surface pair.

SID mechanisms are typically interrogated by monitoring the kinetic energy or velocity of the scattered product ions. If delayed unimolecular fragmentation is occurring, then the velocity of the scattered fragment ions will be the same as the scattered projectile ion. If a prompt or shattering mechanism is occurring then the scattered product ions will leave the surface with different velocities or rather the same kinetic energy. At this early stage of investigation, it is not clear if the prompt dissociation mechanism is only specific for certain ion/surface pairs or if both mechanisms commonly occur but to varied extents.

### Other ion/surface processes

The examples discussed thus far have been relatively ‘clean’ examples of SID without the added complexity of peaks resulting from chemical sputtering and ion/surface reactions. There is a growing literature based on observations of chemical sputtering, ion/surface reactions, and ion/surface reactions which

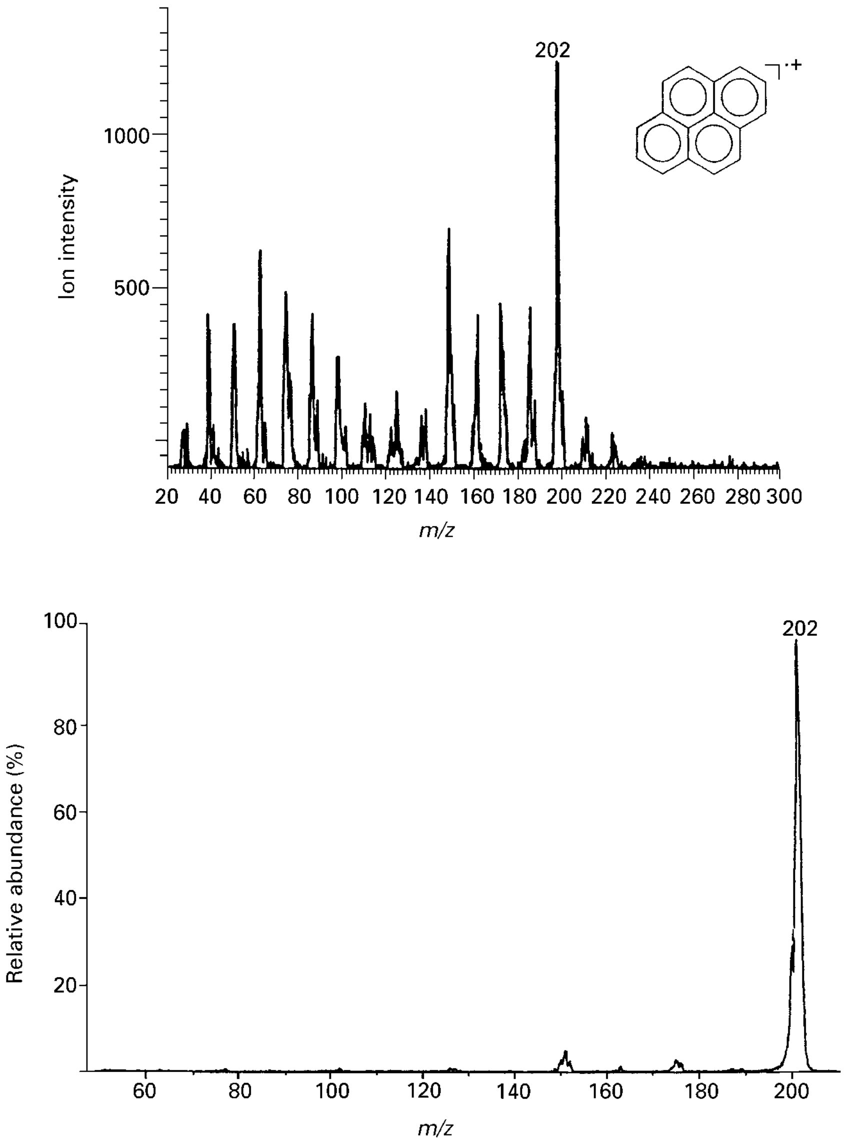



**Figure 10** (A) Surface-induced dissociation spectrum of the pyrene molecular ion (*m/z* 202) with a stainless steel surface at a collision energy of 100 eV. (B) Collision-induced dissociation mass spectrum of the pyrene moelcular ion (*m/z* 202) with Ar under single collision condiitons. Note that the *m/z* axis is not aligned or the same scale between (A) and (B). Adapted with permission fron Riederer Jr DE, PhD Thesis, Purdue University, 1993.

lead to selective surface modification in low energy ion/surface collisions. In order to obtain a better understanding of SID, it is important to have an understanding of all of the ion/surface collision processes shown in **Table 1**.

**Ion/surface reactions** The organic covered surface can be thought of as a chemical reagent in a reaction that uses the projectile ion as the second chemical reagent. **Figure 12** shows the SID mass spectrum for collisions of $(CH_3)_2$ $SiNCS^+$ with an F-SAM surface at 50 eV. Peak assignments can be found in **Table 5**. This spectrum displays a wealth of peaks originating from various ion/surface collision processes including SID, ion/surface reaction, and chemical sputtering. The number and abundance of ion/surface reaction products (especially Si containing product ions) for this ion/surface pair may appear to be remarkable for what would intuitively be thought of as a relatively inert fluorocarbon surface. Indeed, it has been found that the F-SAM surface is very reactive toward organometallic, metallic, and a few organic ions.

Two distinct ion/surface reaction mechanisms have been elucidated. One ion/surface reaction mechanism

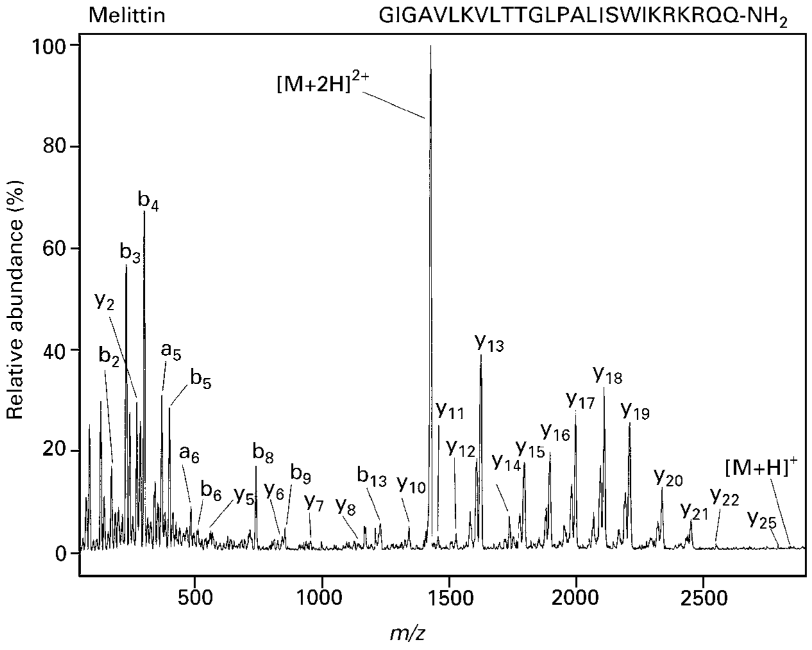



**Figure 11** Surface-induced dissociation spectrum of doubly protonated melittin obtained at a 110 eV collision with a fluorinated SAM surface. Reproduced with permission from Dongré AR, Somogyi Á and Wysocki VH (1996) Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides *Journal of Mass Spectrometry* **31**: 339.

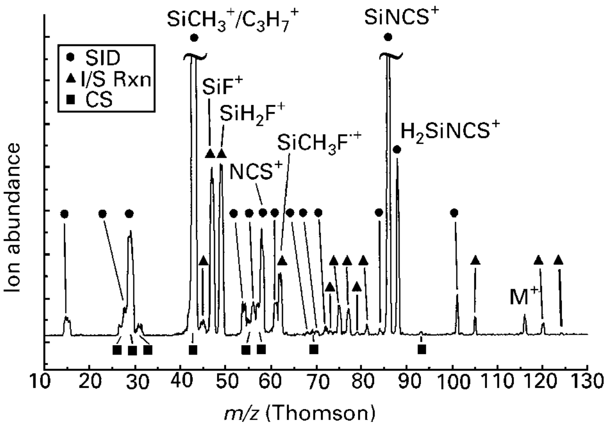



**Figure 12** 50 eV ion/surface collision spectrum of $(CH_3)_2SiNCS^+$ on an F-SAM surface. Note that the ion abundance scale has been expanded. For reference, the $M^+/SiNCS^+$ ratio is 0.05 and the $M^+/SiCH_3^+$ ratio is 0.04. Peak assignments are noted in **Table 5**. Reproduced with permission from Miller SA, Luo H, Jiang X, Rohrs HW and Cooks RG (1997) Ion/surface reactions, surface-induced dissociation, and surface modification resulting from hyperthermal collisions of $OCNCO^+$, $OCNCS^+$, $(CH_3)_2SiNCO^+$, and $(CH_3)_2SiNCS^+$ with a fluorinated self-assembled monolayer. *International Journal of Mass Spectrometry and Ion Processes* **160**: 83–105.

occurs by direct abstraction or oxidative addition of an atom or group of atoms from the surface by the projectile ion. This type of mechanism is most often observed for metal-containing ions and fluorocarbon surfaces. For example, collisions of $W^+$ with an F-SAM surface result in ion/surface reaction products of the form, $WF_n^+$ ($n$ = 1–5). These reaction products have been shown to occur via the direct abstraction process. The second mechanism occurs by a two-step process of (i) an electron transfer from the surface to the projectile ion and (ii) an associative ion/molecule reaction between the neutralized projectile and a chemically sputtered ion from the surface. This mechanism is typically observed for hydrocarbon projectile ions and hydrocarbon surfaces. An example of ion/surface reaction which occurs via an associative ion/molecule reaction is shown in **Figure 13**. The neutralized $d_6$-benzene, $C_6D_6$, reacts with a sputtered proton from the H-SAM surface to form the $[M+H]^+$ product ion. This example also illustrates the advantage of using an isotopically labelled projectile ion. Peak assignments for SID and ion/surface reaction products are less ambiguous than if the $C_6H_6^{\bullet+}$ molecular ion was used as the projectile. It is worth noting that both ion/surface reaction mechanisms have been indirectly supported by the corresponding gas phase ion/molecule experiments.

**Chemical sputtering** Chemical sputtering has been shown to be a very sensitive (submonolayer) surface analysis technique when using appropriate sputtering reagents in low energy ion/surface collisions. The chemical sputtering peaks originating from the

F-SAM surface in **Figure 12** are of low abundance but can be significantly enhanced by using the xenon radical cation which is an efficient charge transfer reagent for fluorocarbon surfaces. For example, **Figure 14** depicts the chemical sputtering spectrum of a fresh F-SAM surface and the liquid PFPE surface. Abundant chemical sputtering peaks representative of the chemical composition of each surface are labelled in the figure. This experiment is also sensitive toward hydrocarbon impurities in the F-SAM surface at *m/z* 27 ($C_2H_3^+$), *m/z* 29 ($C_2H_5^+$), *m/z* 39 ($C_3H_3^+$), *m/z* 41 ($C_3H_5^+$), and *m/z* 43 ($C_3H_7^+$). As noted earlier, the F-SAM and liquid PFPE surface produce very similar SID mass spectra. However, chemical sputtering experiments can differentiate the two, due to the characteristic oxygenated fluorocarbon peak at *m/z* 47, $CFO^+$, in the liquid PFPE spectrum. These experiments also suggest that the liquid/vacuum interface microscopic structure is similar to that of the F-SAM surface and the projectile ion only interacts with the outer portion of the surface at these low energies. Note that chemical sputtering is a low energy (reactive) collision which differs from the more common secondary ion mass spectrometry technique that uses keV ion beam (momentum transfer) collisions to analyse surface composition.

**Surface modification** Selective chemical modification of surfaces is also a consequence of low energy ion/surface collisions. For example, one of the many ion/surface reaction products seen in **Figure 12** is of particular interest, $CH_3FSiNCS^+$ (see also **Table 5**). This reaction product may be formed by the following group transfer reaction,

$$(CH_3)_2SiNCS^+ + F{-}CF_2\text{-(surface)}$$
$$\rightarrow CH_3FSiNCS^+ + CH_3{-}CF_2\text{-(surface)}$$

where F–$CF_2$-(surface) represents the outer $CF_3$ head group on an F-SAM surface. This reaction would yield a chemically modified surface which would be difficult to fabricate using common surface preparation methods. Xenon chemical sputtering has shown that the $(CH_3)_2SiNCS^+$ projectile ion modifies the F-SAM surface by incorporating NCS, methyl, and silyl groups into the fluorocarbon chain. Other examples of chemical surface modification involving group transfer reactions or transhalogenation in low energy ion/surface collisions include the interaction of $OCNCS^+$, $SiCl_4^{\bullet+}$, $CH_2Br_2^{\bullet+}$, and $CH_2Br^+$ with F-SAM surfaces.

**Soft landing** Finally, soft landing of polyatomic ions into self-assembled monolayers can be achieved using appropriate projectile ions with low ion kinetic

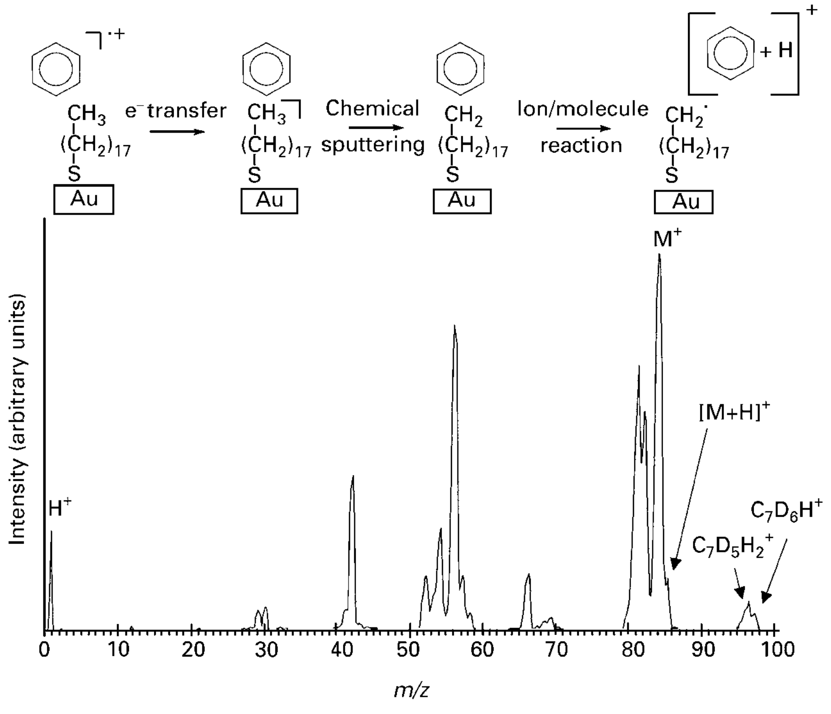



**Figure 13** Product ion spectrum resulting from collisions of the $d_6$-benzene molecular ion (*m/z* 84) with an H-SAM surface at 30 eV. The inset represents the associative ion/surface reaction mechanism for the unlabelled benzene molecular ion.

**Table 5** Peak assignments for $(CH_3)_2SiNCS^+$ from an F-SAM surface at a collision energy of 50 eV.

| *Mass* | *SID* | *Mass* | *Ion/surface reaction* | *Mass* | *Chemical sputtering* |
|---|---|---|---|---|---|
| *m/z* 101 | $CH_3SiNHCS^{\bullet+}$ | *m/z* 124 | $F_2SiNCS^+$ | *m/z* 93 | $C_3F_3^+$ |
| *m/z* 88 | $H_2SiNCS^+$ | *m/z* 120 | $CH_3FSiNCS^+$ | *m/z* 81[a] | $C_2F_3^+$ |
| *m/z* 86 | $SiNCS^+$ | *m/z* 105 | $FSiNCS^+$ | *m/z* 69 | $CF_3^+$ |
| *m/z* 84 | $C_2H_6SiNC^+$ | *m/z* 81[a] | $F_2SiCH_3^+$ | *m/z* 57[a] | $C_4H_9^+$ |
| *m/z* 72 | $C_2H_6SiN^+$ | *m/z* 79 | $F_2SiCH^+$ | *m/z* 55[a] | $C_4H_7^+$ |
| *m/z* 70 | $C_2H_4SiN^+$ | *m/z* 77 | $NCSF^+$ and $FSi(CH_3)_2^+$ | *m/z* 43[a] | $C_3H_7^+$ |
| *m/z* 68 | $C_2H_2SiN^+$ | *m/z* 75 | $FSiC_2H_4^+$ | *m/z* 31 | $CF^+$ |
| *m/z* 61*a* | $HSiS^+$ | *m/z* 73 | $FSiC_2H_2^+$ | *m/z* 29[a] | $C_2H_5^+$ |
| *m/z* 58 | $NCS^+$ and $Si(CH_3)_2^{\bullet+}$ | *m/z* 62 | $FSiCH_3^{\bullet+}$ | *m/z* 27 | $C_2H_3^+$ |
| *m/z* 57[a] | $SiC_2H_5^+$ | *m/z* 61[a] | $FSiCH_2^+$ | | |
| *m/z* 56 | $SiC_2H_4^{\bullet+}$ | *m/z* 49 | $H_2SiF^+$ | | |
| *m/z* 55[a] | $SiC_2H_3^+$ | *m/z* 47 | $SiF^+$ and $FCO^+$ | | |
| *m/z* 54 | $SiC_2H_2^{\bullet+}$ | *m/z* 45 | $NCF^{\bullet+}$ | | |
| *m/z* 43[a] | $SiCH_3^+$ | | | | |
| *m/z* 29[a] | $SiH^+$ | | | | |
| *m/z* 28 | $Si^+$ | | | | |
| *m/z* 15 | $CH_3^+$ | | | | |

[a] Indicates isobaric masses which can have more than one assignment. See **Figure 12** for a scattered product ion spectrum. Reproduced with permission from Miller SA, Luo H, Jiang X, Rohrs HW and Cooks RG (1997) *International Journal of Mass Spectrometry and Ion Processes* **160**: 83–105.

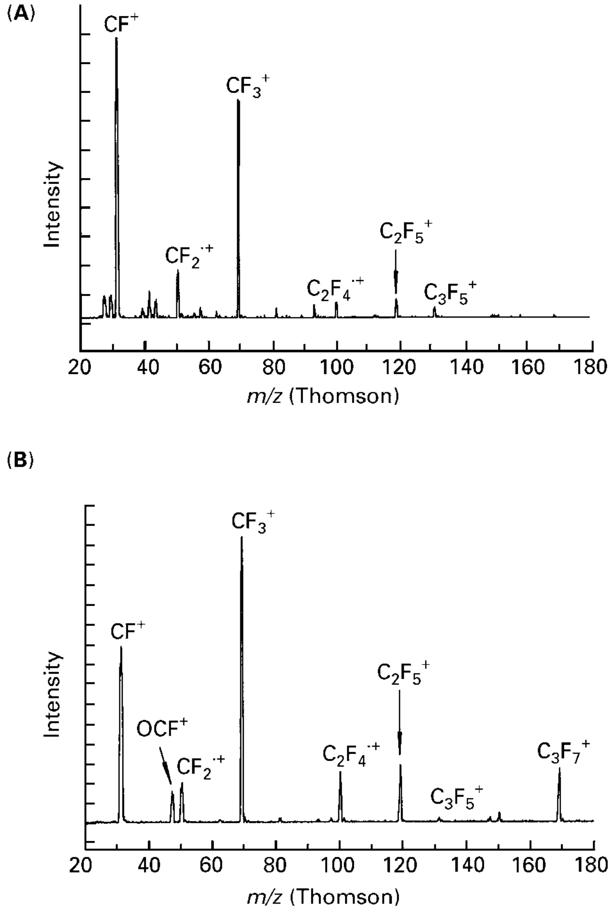



**Figure 14** 90 eV $Xe^{\bullet+}$ chemical sputtering spectra for (A) and F-SAM surface and (B) an 8 μm thick liquid perfluoropolyether surface. Reproduced with permission from Pradeep T, Miller SA, Rohrs HW, Feng B and Cooks RG (1995) Chemical sputtering of F-SAM and liquid PFPE. *Materials Research Society Symposium Proceedings* **380**: 93–98.

energies. The projectile ion used in **Figure 12**, $(CH_3)_2SiNCS^+$, can be soft landed in a F-SAM surface by lowering the ion kinetic energy below 10 eV. The soft-landed ion can be subsequently liberated into the gas phase by $Xe^{\bullet+}$ chemical sputtering. By using chemical sputtering experiments and temperature programmed desorption, the soft-landed species has been shown to survive as an ion in the F-SAM matrix for several days in vacuum and a few days in ambient air. Although these results have only recently been published, this approach may have exciting implications for both fundamental and applied studies.

## List of symbols

$KE_{final}$ = final kinetic energy; $P(\varepsilon)$ = internal energy distribution; $Q$ = inelasticity of collision; $T$ = initial kinetic energy (ion translational energy); $V$ = total vibrational energy; $\varepsilon$ = internal energy; $\varepsilon_{surface}$ = internal energy deposited in surface.

*See also:* **Fragmentation in Mass Spectrometry; Ion Collision Theory; Ion Energetics in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Ion Structures in Mass Spectrometry; Ion Trap Mass Spectrometers; MS–MS and MS*ⁿ*; Quadrupoles, Use of in Mass Spectrometry; Sector Mass Spectrometers; Time of Flight Mass Spectrometers.**

## Further reading

Ada ET, Kornienko O and Hanley L (1998) Chemical modification of polystyrene surfaces by low-energy polyatomic ion beams. *Journal of Physical Chemistry B* 102: 3959.

Busch KL, Glish GL and McLuckey SA (1993) *Mass Spectrometry/Mass Spectrometry: Techniques and*

*Applications of Tandem Mass Spectrometry*. New York: VCH Publishers.
Cooks RG, Ast T, Pradeep T and Wysocki VH (1994) Reactions of ions with organic surfaces. *Acc. Chemistry Research* **27**: 316.
Cooks RG and Miller SA. *Collisions of Ions with Surfaces*. In: Jennings KR (ed) *NATO ASI Series*. Dordrecht: Kluwer (in press).
Cooks RG, Ast T and Mabud MA (1990). Collisions of polyatomic ions with surfaces. *International Journal of Mass Spectrometry and Ion Processes* **100**: 209.
Dongre AR, Somogyi A and Wysocki VH (1996) Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. *Journal of Mass Spectrometry* **31**: 339.
Hanley L (ed) (1998) Special issue on polyatomic ion-surface interactions. *International Journal of Mass Spectrometry and Ion Processes* **174**: 1.
Hayward MJ, Park FDS, Phelan LM, Bernasek SL, Somogyi A and Wysocki VH (1996) Examination of sputtered ion mechanisms leading to the formation of $C_7H_7^+$ during surface induced dissociation (SID) tandem mass spectrometry (MS/MS) of benzene molecular cations. *Journal of the American Chemical Society* **115**: 8375.
Mabud MA, Dekrey MJ and Cooks RG (1985) Surface-induced dissociation of molecular ions. *International Journal of Mass Spectrometry and Ion Processes* **67**: 285.
Miller SA, Luo H, Cooks RG and Pachuta SJ (1997) Soft-landing of polyatomic ions at fluorinated self-assembled monolayer surfaces. *Science* **275**: 1447.
Morris MR, Riederer DE Jr., Winger BE, Cooks RG, Ast T and Chidsey CED (1992) Ion/surface collisions at functionalized self-assembled monolayer surfaces. *International Journal of Mass Spectrometry and Ion Processes* **122**: 181.
Vekey K, Somogyi A and Wysocki VH (1995) Internal energy distribution of benzene molecular ions in surface-induced dissocation. *Journal of Mass Spectrometry* **30**: 212.

# Surface Plasmon Resonance, Applications

**Zdzislaw Salamon** and **Gordon Tollin**,
University of Arizona, Tucson, AZ, USA

Copyright © 1999 Academic Press

| SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS |
|---|
| Applications |

As described in the article on the theory of surface plasmon resonance, surface plasmons create a surface-bound evanescent electromagnetic wave which propagates along the surface of an active medium (usually a thin metallic film), with the electric field intensity maximized at this surface and diminishing exponentially on both sides of the interface. As a consequence of this property, the phenomenon has been utilized extensively in studies of surfaces and of thin dielectric films deposited on the active medium. Although numerous other optical techniques have also been applied to such systems (e.g. ellipsometry, interferometry, spectrophotometry, and microscopy the surface plasmon resonance (SPR) method has some important advantages over all other optical techniques, as follows. The method utilizes a relatively simple optical system, it has a superior sensitivity, and the complete system of measurement is located on the side of the apparatus that is remote from the sample, and thus there is no optical interference from the bulk medium. Furthermore, the surfaces of the sample need no extra treatment to increase reflectivity, because this is achieved by using total reflectance. Additional benefits include the fact that there are three parameters of the resonance that can readily be measured, thereby yielding much more information about the sample and changes within it than the simple interferometric step height used in other sensitive optical techniques. Finally, the recently developed new variant of SPR, which involves both plasmon and waveguide modes producing coupled plasmon waveguide resonance (CPWR), expands spectroscopic sensitivities and capabilities even further, allowing the measurement of anisotropies in both the refractive index and extinction coefficient.

Current interest in the properties of surfaces and surface coatings arises partly from increased applications of thin film devices, including the large field of integrated optics, which uses surface guided waves, and partly from recent developments in biosensors. In addition, many surface phenomena depend on molecular interactions occurring at dielectric interfaces. Surface chemistry applications include charge transfer interactions, acid–base chemistry, chemical

bond formation, and van der Waals forces, all of which participate in the processes of adhesion, catalysis, lubrication, corrosion, contamination and packaging. Surface and interfacial phenomena are also of particular importance in all areas of biology, where molecular phenomena occurring at lipid membrane interfaces or during protein–protein interactions are key events in living systems.

This article focuses on the application of the two major properties of the surface plasmon (SP) evanescent wave (i.e. its identity as a surface-bound and surface-unique electromagnetic phenomenon), to the characterization of the properties of surfaces, interfaces, and thin films. It describes the two principal experimental modes, kinetics and spectroscopy, which are currently used in these SPR applications.

## Principles of SPR applied to thin films

As described in the article on the theory of surface plasmon resonance, plasmons are transverse magnetic (TM; *p*-polarized) evanescent waves generated at the expense of light energy under total internal reflection conditions, which propagate along an active medium surface (usually silver or gold) with their field amplitudes decaying exponentially perpendicular to the metal surface. They obey the well-known dispersion relation, and can only be excited when matching energy and momentum conditions between photons and surface plasmons are fulfilled, as follows:

$$K_{\mathrm{SP}} = K_{\mathrm{ph}} = (\omega/c)\varepsilon_0{}^{1/2}\sin\alpha_0 \qquad [1A]$$

where:

$$K_{\mathrm{SP}} = \omega/c\ (\varepsilon_1\varepsilon_2/\varepsilon_1 + \varepsilon_2)^{1/2} \qquad [1B]$$

is the longitudinal component of the SP wave vector, $K_{\mathrm{ph}}$ is a component of the light wave vector parallel to the active medium surface, $\omega$ is the frequency of the surface plasmon excitation wavelength, $c$ is the velocity of light *in vacuo*, $\varepsilon_0$, $\varepsilon_1$, and $\varepsilon_2$ are the complex dielectric constants for the incident, surface active, and dielectric (or emerging) media, respectively, and $\alpha_0$ is the incident coupling angle.

As can be seen, the resonance condition stated in Equation [1A] can be fulfilled by either changing the incident angle, $\alpha$, at a constant value of photon energy, $h\omega$ (i.e. maintaining a constant value of the light wavelength, $\lambda = \lambda_0$; see **Figures 1A** and **1C**, curve 1), or varying $\lambda$ at a constant value of the incident angle, $\alpha = \alpha_0$ (see **Figures 1B** and **1D**, curve 1). This means that each photon of energy $h\omega$ allows the excitation of exactly one SP mode at a distinctive value of the incident angle, $\alpha$, within a specified metallic layer (silver or gold). Furthermore, any alteration in the optical parameters (described by the complex dielectric constant, $\varepsilon$, in Equation [1B], which contains the refractive index, $r$, and extinction coefficient, $k$) of the metal–emerging medium interface, will affect the $K_{\mathrm{SP}}$ value and therefore change the resonance characteristics, as indicated by Equation (1A). Thus, deposition of any thin dielectric film (sensing layer) on a metallic surface introduces changes in the optical parameters of the metal–emerging medium interface, thereby causing a shift of the SP wave vector to a larger value:

$$K'_{\mathrm{SP}} = K_{\mathrm{SP}} + \Delta K_{\mathrm{SP}} \qquad [2]$$

According to Equation [1A], such a shift of the SP wave vector moves the resonance to either a higher incident angle, $\alpha_1$, at a constant value of exciting light wavelength, $\lambda_0$, or a longer exciting light wavelength, $\lambda_1$, at a constant value of the incident angle $\alpha_0$, as shown in all parts of **Figure 1** (curves 2).

The spectra presented in **Figure 1** demonstrate that the angular (or wavelength) position and shape of the resonance curve is very sensitive to the optical properties of both the metal film and the emergent dielectric medium adjacent to the metal surface. This property is one of three major attributes of SPR which constitute the foundation for the application of surface plasmons (via their interactions with the medium in which they propagate) as an optical probe of the properties of any material in contact with an active layer capable of generating this effect. The other two attributes can be stated as follows: (i) The evanescent electromagnetic field generated by the free-electron oscillations is strongly enhanced as compared with the electromagnetic field of the exciting light, and reaches its maximum at the metal–emerging dielectric medium interface. The enhancement of the field at the metal–dielectric interface magnifies the spectral features of the interface, making optical measurements with the SP electromagnetic waves possible. (ii) The evanescent electromagnetic field generated in the film decays exponentially with penetration distance into the emerging medium, i.e. the depth of penetration into the dielectric material in contact with an active metal layer extends only to a fraction of the light wavelength used to generate the plasmons. This makes the phenomenon sensitive *only* to the metal–dielectric interface region, without any interference from the bulk volume of the dielectric material or any medium that is in contact with it.

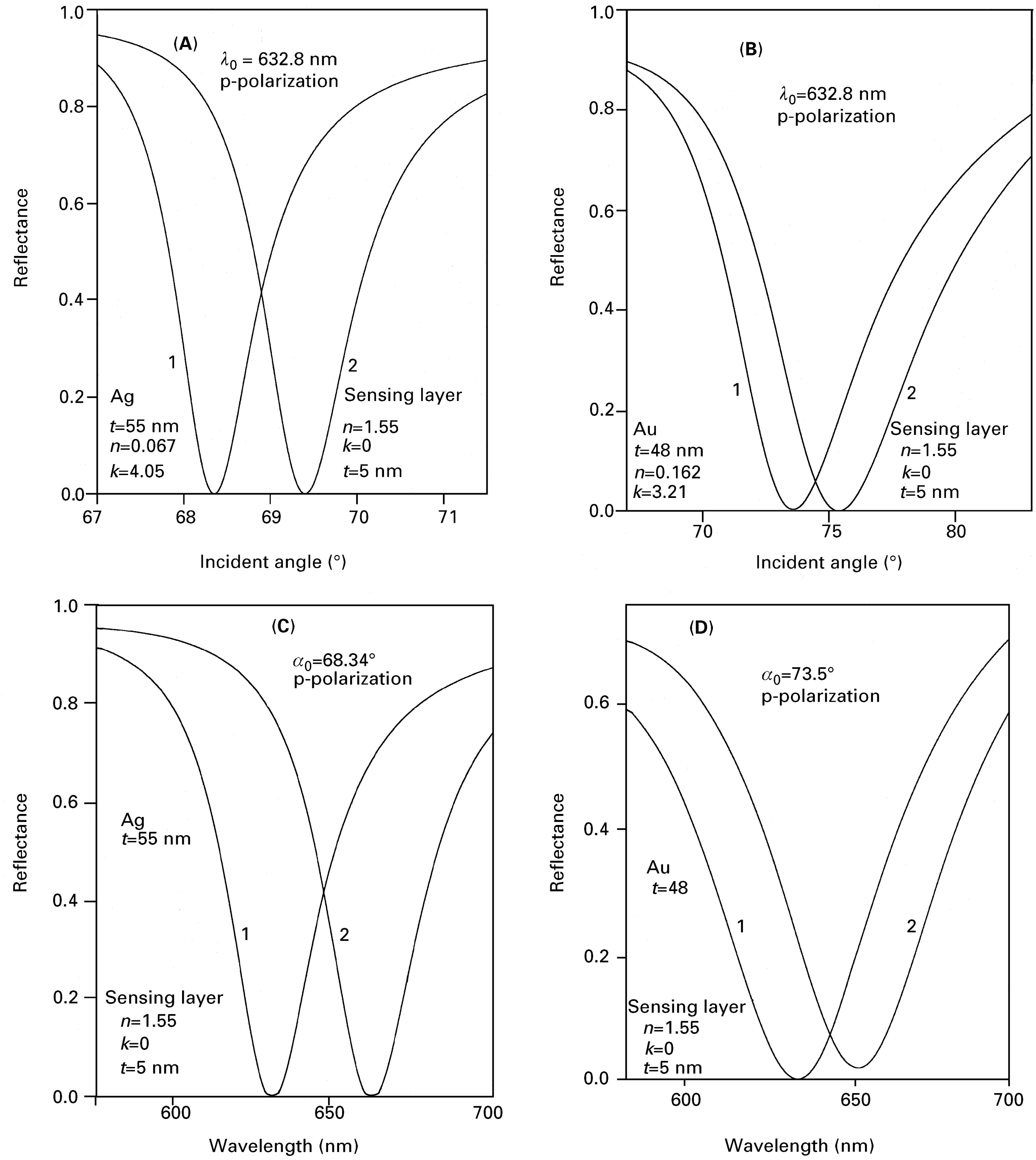



**Figure 1** Influence on the theoretical SPR spectra generated either with silver (panels A and C) or gold (panels B and D) films (curves 1 in all panels), and represented by reflectance vs. either incident angle (panels A and B), or excitation light wavelength (panels C and D), of a thin sensing layer deposited on the metallic surface (curves 2 in all panels). The incident and emerging media are glass ($n_g$ = 1.5151) and water ($n_w$ = 1.33), respectively.

As a consequence of these characteristics, SPR is ideally suited to probe a few nanometres from the metal surface, a distance well below the wavelength of the light used to generate the plasmons. In essence, the surface plasmon technique can be likened to a multiple-beam interferometer, with one narrow reflected fringe, and is therefore capable of similar sensitivity. In contrast, however, it allows optical measurements to be made of dimensions that are one-thousandth of a wavelength or less, which is well below the sensitivity limits of the other optical techniques (e.g. ellipsometry, spectrophotometry,

and various forms of microscopy) used in studies of surfaces and thin films.

## Analysis of SPR spectra

As illustrated above, SPR spectra can be presented either in the form of reflectance ($R$) as a function of incident angle ($\alpha$):

$$R = f(\alpha)_{\lambda_0} \quad [2A]$$

or as a function of the excitation light wavelength ($\lambda$):

$$R = f(\lambda)_{\alpha_0} \quad [2B]$$

Such spectra incorporate information about the following three parameters, refractive index, $n$, extinction coefficient, $k$, and thickness, $t$, which describe the optical properties of both the metal and the dielectric layers. The influence of these parameters on the angular (or wavelength) *position*, the angular (or wavelength) *width*, and the *depth* of the SPR spectrum is completely contained in the characteristic matrix of the thin film assembly, which allows a determination of their values.

### Structural analysis of thin dielectric films

As demonstrated in **Figures 2** and **3**, the above-mentioned three optical parameters of a thin dielectric film deposited on a metal surface are well separated in their effects on the SPR spectra. Therefore, the experimental spectra interpreted in the context of the characteristic matrix of the assembly allow a unique evaluation of $n$, $k$, and $t$. The evaluation procedure is based on fitting a theoretical resonance curve to the experimental one. As also indicated in **Figures 2** and **3**, the CPWR method provides a means for determining the optical parameters using both TM ($p$) and transverse electric (TE) ($s$) polarizations of the excitation light, resulting in two values of the refractive index ($n_p$ and $n_s$) and two values of the extinction coefficient ($k_p$ and $k_s$). These parameters can then be used to calculate the anisotropy of $n$ and $k$, thereby describing the degree of both molecular order (by the anisotropy in $n$) and orientation of chromophoric groups attached to the molecules comprising the thin film (by the anisotropy in $k$). Such information, taken together with the film thickness ($t$), provides insights into the microscopic structure of the film. Furthermore, the optical parameters can also be employed to calculate the mass of a deposited thin layer (see next section for details).

It has to be emphasized that all of these characterizations can be obtained using a single device (in this case, CPWR), containing a metal film covered with a dielectric layer, and using a measurement method that involves only a determination of reflected light intensity under total reflection conditions as a function of either incident angle or light wavelength.

### Determination of thin film mass

As noted above, the values of the $n$, $k$, and $t$ parameters of a thin film layer deposited on the surface of a metal film contain information about the amount of material in the layer. There are two different ways of calculating the adsorbed mass from the refractive index value. The first approach is based on the assumption that the refractive index increment, d$n$/d$C$, is independent of the concentration, $C$, of the adsorbed substance. If this is so, the surface density, $D_s$, i.e. the amount of material per unit surface area, can be evaluated by the following expression:

$$D_s = t\,[(n - n_2)/(\mathrm{d}n/\mathrm{d}C)] \quad [3]$$

where $n$ and $n_2$ are refractive indices of the adsorbed thin film and the emerging medium, respectively, and d$n$/d$C$ is the refractive index increment of the adsorbed substance. Equation [3] can of course only be used if the refractive index increment is known, and depends on the assumption that the value of the increment is constant over the concentration range of the adsorbed material.

The second method of mass calculation is based on the Lorentz–Lorenz relation which can be presented in the most general case, i.e. when the deposited layer contains a mixture of substances, by the following equation:

$$(n^2 - 1)/(n^2 + 2) = A_1N_1 + A_2N_2 + \ldots \quad [4]$$

where $A_i$ and $N_i$ are the molar refractivity and the number of moles of substance per unit volume, respectively. For a pure substance, a mass density, $D$, defined as mass per unit volume of adsorbed material, can be directly related to the refractive index by the following equation:

$$D = MN = M/A\,[(n^2 - 1)/(n^2 + 2)] \quad [5]$$

For an adsorbed layer of thickness $t$ formed from such a pure substance, the above equation can be used to calculate the adsorbed mass ($M$) in μg per cm$^2$, as follows:

$$M = Dt = 0.1\ M/A\ \{t[(n^2 - 1)/(n^2 + 2)]\} \qquad [6]$$

where the thickness is expressed in nanometres. Such a simple mass calculation becomes more complicated when a surface layer is formed from a mixture of substances, as it often is in real measurements. This complexity can, however, still be dealt with, depending upon the specific experimental conditions.

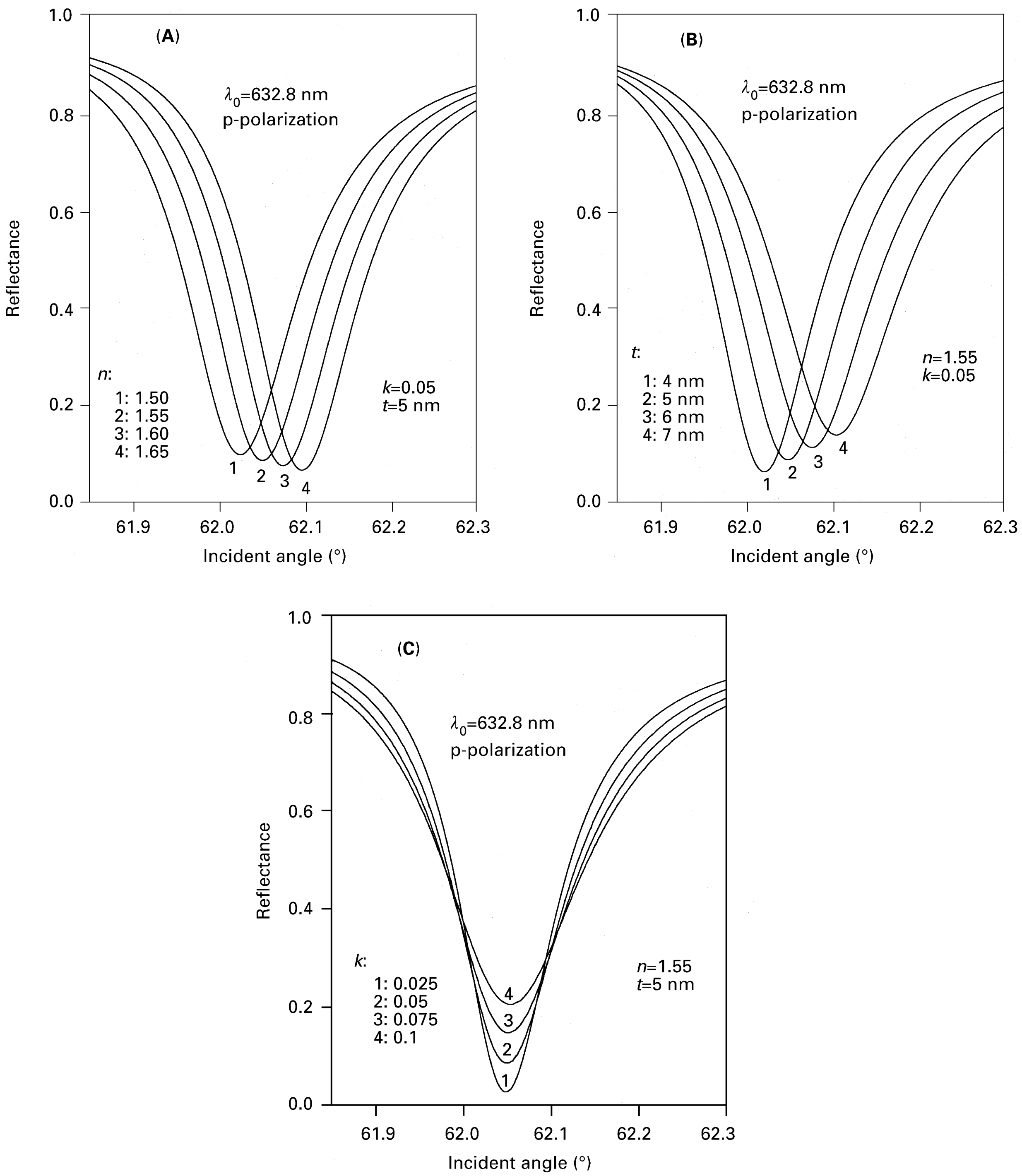



**Figure 2** Changes in the *p*-polarization component of a theoretical SPR spectrum generated with a CPWR device, comprising a silver layer coated with a 460 nm $SiO_2$ film, caused by alterations in either the refractive index (A), the thickness (B), or the extinction coefficient (C), of a light-absorbing dielectric sensing film deposited on the silica film. The incident and emerging media are both as described in **Figure 1**.

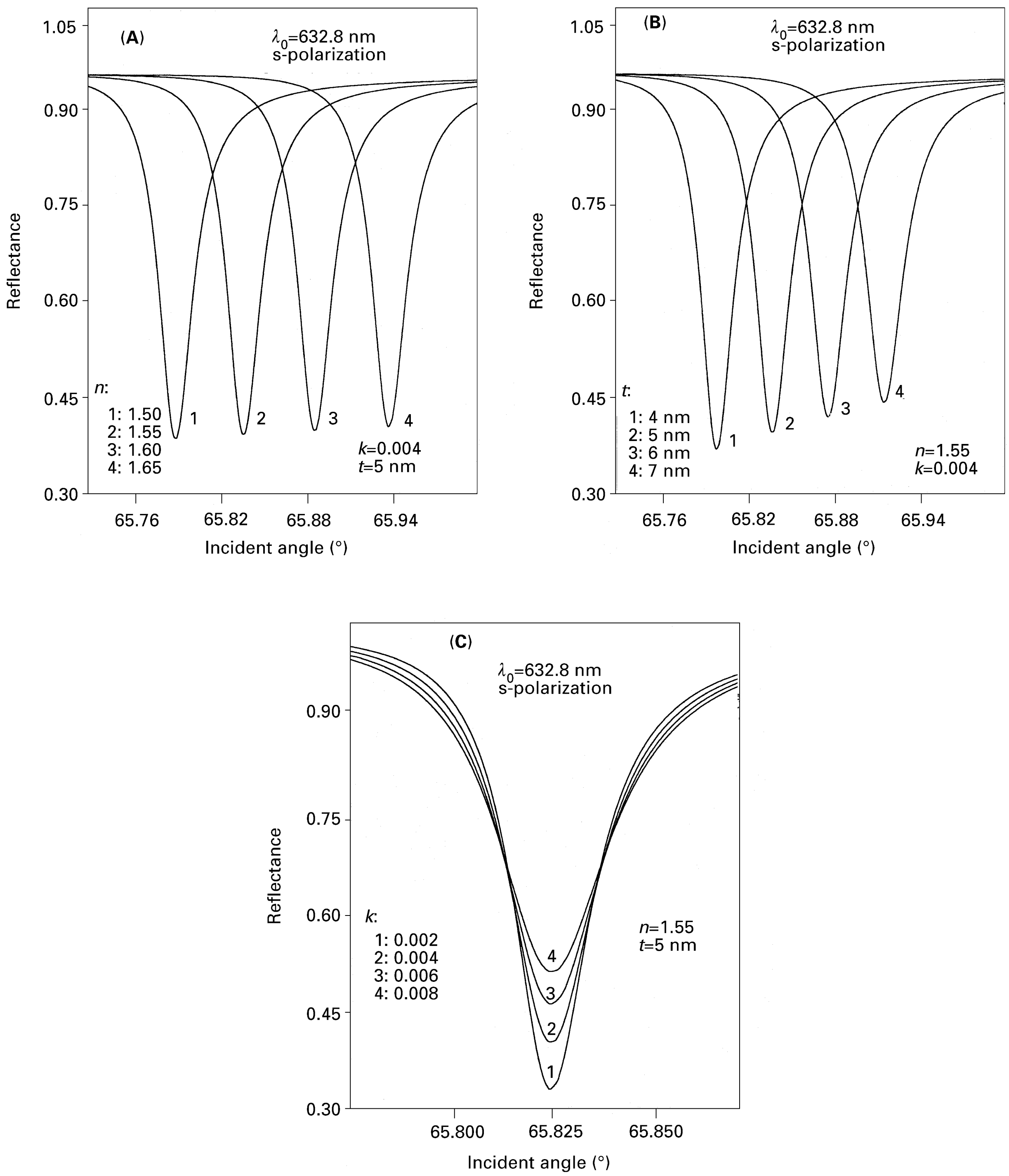



**Figure 3** Changes in the *s*-polarization component of a theoretical SPR spectrum obtained with the CPWR device described in **Figure 2**, by alterations in either the refractive index (A), the thickness (B), or the extinction coefficient (C), of a light-absorbing dielectric sensing film deposited on the silica. The incident and emerging media are the same as in **Figure 1**.

## Applications of the SPR technique

The SPR measurement can be performed, as with other optical measurements, in two modes. Used in a spectroscopic mode, i.e. by measuring and analysing the entire resonance spectrum, SPR provides a tool for the characterization of the microstructural properties of thin films, including thickness, mass distribution (packing density) within the film, and degree of order (orientation) of molecules and of chromophore groups attached to the molecules within the film. In addition, when combined with emission

measurements, i.e. emission excited by the surface plasmon electromagnetic field, it can also provide a *molecular orientation distribution*.

When used in a kinetic mode i.e. by measuring and analysing the changes of reflectance as a function of time at $\alpha = \alpha_0$ = const., and $\lambda = \lambda_0$ = const.:

$$R = f(time)_{\alpha_0,\lambda_0} \qquad [7]$$

this experimental approach provides a means for real-time analysis of molecular interactions.

### Steady-state and kinetic SPR measurements

As a consequence of the above-mentioned characteristics, SPR is ideally suited to studying both structural and mass changes of thin films, including molecular interactions occurring at surfaces and interfaces. These can be examined using either steady-state or time-resolved SPR spectroscopy, and can be applied to a wide range of materials. The time-resolved mode expands the capability of SPR techniques to allow probing of the dynamics of structural and mass alterations of thin films.

Although the phenomenon initially has been utilized extensively by physical scientists in studies of the properties of surfaces and thin films, with a major goal of creating thin film-based opto-electronic devices for optics, spectroscopy, laser optics, solar energy conversion, and space technology, the implementation of this technology to chemical, and especially to biological processes, has proven to be one of the most dynamic and fruitful application areas during recent years. The molecular mechanisms underlying biological and chemical processes are dependent on direct interactions between (bio)molecules. These interactions are often preceded by specific binding between two or more molecules, and for binding to occur the molecules must be able to come close enough to each other to make contact. Such binding may take place between the partners in solution, or with at least one partner attached to a biological surface (e.g. a lipid membrane or a chromosome) or to a physical surface. In situations in which the interactions occur at a surface, there must be a mechanism, such as lateral diffusion in the surface plane, that permits them to come close enough for interaction. These binding processes will not only result in alterations in the mass and composition of the membrane or other surface, but may also cause structural changes within such entities as well.

The simplest case of a biomolecular (or chemical) interaction is the association of two molecules, $X$ and $Y$, to form a complex $XY$. This process is characterized by rate constants for association, $k_{ass}$, and dissociation, $k_{diss}$, and an equilibrium (binding or binding affinity) constant, $K_B = k_{ass}/k_{diss}$. Steady-state SPR measurements, by determining mass density changes occurring in molecular assemblies accompanying binding interactions, allow an evaluation of the binding constants $K_B$. Kinetic SPR measurements enable direct observation in real time of such binding events, resulting in determination of association and dissociation rate constants (as discussed in the next section), as well as the dynamics of binding-induced structural changes.

The ability of SPR to probe both kinetic and thermodynamic processes, as well as to provide microstructural information, make it a very important component of the experimental methodology available to probe molecular interactions occurring at surfaces. Furthermore, it allows some of the limitations of other techniques to be overcome. For example, other methods often require one of the partners to be labelled in some way in order to allow it to be detected. Fluorescent probes, radioactive labels, and attachment of independently detectable molecules (e.g. enzymes) have all been used for this purpose. These suffer from the drawback that they may interfere with the binding of the labelled partner to the unlabelled one, or cause unwanted structural perturbations. SPR observations can be based solely on the dielectric properties of molecules, or their intrinsic light absorption characteristics, and thus require no specific labelling.

### Association and dissociation rate constant measurements

The most straightforward application of SPR technology is the measurement of binding kinetics, i.e. association and dissociation rate constants. This usage is based on the following requirements: First, the binding of one molecule to another, in which one of the molecules is immobilized on the surface of an SPR device, must produce a change in the refractive index, $n$, of the interface which results in an alteration of the SPR spectrum. This condition is only valid when the SPR exciting light wavelength is outside of the absorption spectral region of both interacting molecules, and when the interaction does not induce any structural changes in the interface. Second, the changes in $n$ are assumed to be proportional to the surface concentration of bound molecules, $C$, i.e. $dn/dC$ = constant over the whole concentration range. This assumption can give rise to some error, especially at high concentrations where the proportionality might not be fulfilled. Third, the changes in

$n$ cause *only* a shift of the resonance curve without any alterations in its shape, which is a simplification of the theoretical influence of $n$ on the resonance curve as determined by the characteristic matrix of the thin film assembly. Under these assumptions, changes in the $n$-value can easily be measured by the shift of the SPR resonance curve. This can be greatly simplified by monitoring the reflectance taken at one specific point of the resonance curve as a function of time, as described by Equation [7]. Such a simple optical measurement allows one to directly probe the molecular interaction by monitoring association and dissociation processes in real time.

### Additional applications of SPR

Additional applications of the SPR phenomenon include using the surface plasmon electromagnetic waves to excite emission of surface-bound chromophores, to enhance Raman spectra (surface-enhanced Raman spectroscopy), and as 'surface-bound light' in optical microscopy.

## Sensitivity of various types of surface plasmon resonances

The sensitivity, $S$, of an SPR measurement can be defined as the change in reflectance, measured either at a specified angle, $\alpha_1$, or a specified wavelength, $\lambda_1$ within the range of the resonance curve, divided by the change in one of the optical parameters, i.e.

$$\text{either} \quad S_{\alpha 1} = [\mathrm{d}R_{(\alpha)}/\mathrm{d}n \ \mathrm{d}k \ \mathrm{d}t]_{\alpha_1} \qquad [8A]$$

$$\text{or} \quad S_{\lambda 1} = [\mathrm{d}R_{(\lambda)}/\mathrm{d}n \ \mathrm{d}k \ \mathrm{d}t]_{\lambda_1} \qquad [8B]$$

respectively. As can be seen from the explicit function of $R$ given by the characteristic matrix of a thin film assembly, the reflectance is not only a function of the optical parameters of the sensing layer, but also depends on the optical parameters of the incident and emerging media, as well as those of the metal film that generates the surface plasmons. In addition, the refractive indices and extinction coefficients of these media are related to one another by the complex form of Snell's Law, which complicates the function $R$ even further. In general, however, changes in the experimental value of $R$ are generated by two factors: the shift of the position and the change of the shape of the resonance spectrum. The latter parameter is usually described either by the sharpness of the SPR spectrum, i.e. its half-width, or by the slope of the reflectance function, and characterizes the resolution in either the resonance angle or the resonance wavelength.

Both of these factors affecting sensitivity are dependent upon the field distributions within the resonant structure. The extent of the resonance angle shift is dependent upon the fraction of the resonant mode that lies within the sensing layer; the higher the fraction, the greater the sensitivity. This is why tight confinement of the surface plasmon evanescent wave to the sensing layer, which results in maximizing the fraction of the evanescent field in this region, favours high sensitivity. The resolution in the resonance angle (or wavelength) shift (setting aside instrumental considerations) is determined by the width of the resonance, which is defined by the losses of electromagnetic field energy within the system. The absorption by the metallic layer is the dominant factor in the SPR spectral width. Scattering of the electromagnetic wave causes further losses and increases with increased irregularities in the film structure and surface roughness, and with lower exciting wavelengths.

In summary, the overall sensitivity of the SPR device depends on both the metallic layer material (e.g. gold or silver), and the type of surface plasmon resonance being measured. The increase in the evanescent field is much smaller (about 2-fold) with gold than silver, which translates into about a 4-fold smaller overall sensitivity for an SPR device based on gold. On the other hand, gold layers are usually more stable in practical use than are silver films. The different types of surface plasmon resonances show different distributions of the evanescent electromagnetic field within the resonator device resulting in widely varying sensitivities. Long-range surface plasmon resonance has about 2.5-fold higher overall sensitivity than conventional SPR, whereas CPWR shows an even higher increase in sensitivity: about 3.5-fold (for $p$-polarization) and about 8-fold for $s$-polarization, as compared to conventional SPR. Therefore, the final design of an SPR device is usually a compromise between the different factors influencing overall sensitivity, durability, and any other requirements of the device for a specific practical application.

## List of symbols

$A_i$ = molar refractivity; $c$ = velocity of light; $C$ = concentration of bound molecules; $D$ = mass density; $D_s$ = surface mass density; $k$ = extinction coefficient; $k_{ass}$, $k_{diss}$ = rate constant for association and dissociation; $K_{sp}$ = longitudinal component of the SP wave vector; $K_{ph}$ = photon component of wave vector; $K_B$ = binding affinity equilibrium constant; $M$ = adsorbed mass; $n$ = refractive index;

$N_i$ = number of moles of substance per unit volume; $R$ = reflectance; $S$ = sensitivity; $t$ = thickness; $\alpha_0$ = incident coupling angle; $\varepsilon$ = dielectric constant; $\omega$ = frequency.

*See also:* **Biochemical Applications of Fluorescence Spectroscopy; Biomacromolecular Applications of UV-Visible Absorption Spectroscopy; Chiroptical Spectroscopy, General Theory; Chiroptical Spectroscopy, Orientated Molecules and Anisotropic Systems; Ellipsometry; Surface Plasmon Resonance, Instrumentation; Surface Plasmon Resonance, Theory; Surface-enhanced Raman Scattering (SERS), Applications.**

## Further reading

Garland PB (1996) Optical evanescent wave methods for the study of biomolecular interactions. *Quarterly Reviews of Biophysics* **29**: 91–117.

Harrick NJ (1967) *Internal Reflection Spectroscopy*. New York: Interscience.

Kovacs G (1982) Optical excitation of surface plasmon-polaritons in layered media. In: Boardman AD (ed) *Electromagnetic Surface Modes*, pp. 143–200. New York: Wiley.

Macleod AH (1986) *Thin Film Optical Filters*. Bristol, U.K.: Adam Hilger.

Raether H (1977) Surface plasma oscillations and their applications. *In*: Hass G, Francombe M and Hoffman R (eds) *Physics of Thin Films*, 9, pp. 145–261. New York: Academic Press.

Salamon Z and Tollin G (1998) Surface plasmon spectroscopy: A new biophysical tool for probing membrane structure and function. *In*: Chapman D and Haris P (eds) *Biomembrane Structure*. Amsterdam: IOS Press.

Salamon Z, Macleod AH and Tollin G (1997) Surface plasmon spectroscopy as a tool for investigating the biochemical and biophysical properties of membrane protein systems. II: Applications to biological systems. *Biochimica et Biophysica Acta* **1331**: 131–152.

Salamon Z, Macleod AH and Tollin G (1997) Coupled plasmon-waveguide resonators: A new spectroscopic tool for probing proteolipid film structure and properties. *Biophysical Journal* **73**: 2791–2797.

Salamon Z, Brown MF and Tollin G (1999) Plasmon resonance spectroscopy: probing molecular interactions with membranes. *Trends in Biochemical Sciences* **24**: 213–219.

Wedford K (1991) Surface plasmon-polaritons and their uses. *Optical and Quantum Electronics* **23**: 1–27.

# Surface Plasmon Resonance, Instrumentation

**RPH Kooyman**, University of Twente, Enschede, The Netherlands

Copyright © 1999 Academic Press

**SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS**
**Methods & Instrumentation**

In view of its simple instrumentation and its high surface sensitivity, surface plasmon resonance (SPR) and, more recently, SPR microscopy gains an increasing significance to numerous problems concerned with the study of interactions occurring near to or at surfaces. Applications can be found in the optical behaviour of metals, the study of Langmuir–Blodgett films and self-assembled monolayers, the interactions of proteins with interfaces, or redox reactions at interfaces. A fast-increasing field is the development of sensitive chemooptical sensors, intended to quantitatively and selectively monitor the presence of prespecified and chemical species, which can be in the range from a molecular mass of ~200 Da to 150 kDa. The main reason for the high sensitivity of SPR to surface phenomena should be attributed to the high local electromagnetic field strengths brought about by surface plasmons (SPs).

Several excellent monographs have been published on SPR theory, whereas an overview on applications is the subject of another article in this Encyclopedia. This chapter will be concerned with experimental techniques for a variety of SPR applications.

## Basic requirements

Although SPs can be excited in an electron beam we will only consider the common case where they are generated by means of light.

SPs can be excited in any material where free charge carriers exist; in the majority of practical cases this means that a metal layer deposited on a

dielectric material is used as a medium to carry SPs. It turns out that the particular SP properties strongly depend on the nature of this deposited layer: if the metal is present as an island-like structure with patches of the order of a few tens of a nm, *radiative* plasmons will be excited resulting in large enhancements of the applied (optical) field strengths. This effect is mainly used to enhance surface responses in vibrational spectroscopy, such as Raman and infrared spectroscopy. *Nonradiative* SPs can be excited if the metal layer is applied as a thin (~50 nm) homogeneous layer. These plasmons have the peculiar property that the associated electromagnetic field is evanescent in character, i.e. the wave vector $k_x$ parallel to the interface is (partly) real, whereas that perpendicular to the interface is imaginary. The generation of SPs is an elastic scattering process, implying that both energy and momentum are conserved. Consider the system as depicted in **Figure 1A**, where a metal layer is sandwiched between two media a and p, which is the commonly used Kretschmann configuration. In view of Snell's law, which states that through this whole system $k_x$ remains constant, this has the consequence that SPs can only be excited at the interface a/m when light enters the metal layer through another interface m/p, *under the condition that* $\varepsilon_p > \varepsilon_a$, where $\varepsilon_p$, $\varepsilon_a$ denote the respective dielectric constants. This is further illustrated in **Figure 1B** where the two lines a and p represent the light dispersion relations in the media a and p, respectively, and the curve represents the SP dispersion relation. A further necessary condition is that the exciting light has to be p-polarized, because light polarized along the metal interface cannot exist.

This short theoretical description contains all the information needed for setting up a basic SPR experiment.

## Choice of metal support

An approximate expression for the SPR wave vector can be found in the literature:

$$k_x = (2\pi/\lambda)(\varepsilon_m \cdot \varepsilon_a/(\varepsilon_a + \varepsilon_m)^{1/2}) \quad [1]$$

where $\lambda$ is the wavelength *in vacuo*, and $\varepsilon_m$ is the metal dielectric constant at the wavelength used. Generally, $k_x$ of the incoming light can be adapted to that required by Equation [1] by inclining the beam relative to the interface normal. The following relation holds (see also **Figure 1A**):

$$k_x = (2\pi/\lambda) \cdot \varepsilon_p^{1/2} \cdot \sin\theta \quad [2]$$

However, in practice the maximum angle $\theta$ that can be chosen is around 80°, implying that in a number of cases the dielectric constant of the support, $\varepsilon_p$, has to be selected such that within this material a light beam can have a *practical* $k_x$ matching that of Equation [1]. If medium a is a water solution (refractive index $n_a \sim 1.33$; note that $n = \varepsilon^{½}$) then BK7 glass ($n_p \sim 1.52$) is an appropriate material. For media with higher refractive index SF10 glass ($n_p \sim 1.7$) can be used. (The abbreviations BK7 and SF10 refer to a nomenclature common in lens-making technology.) Convenient shapes of the dielectric support are prisms with apex angle 60 degrees or hemicylinders, which minimize the reflection loss of light when entering the support (cf. **Figure 1A**).

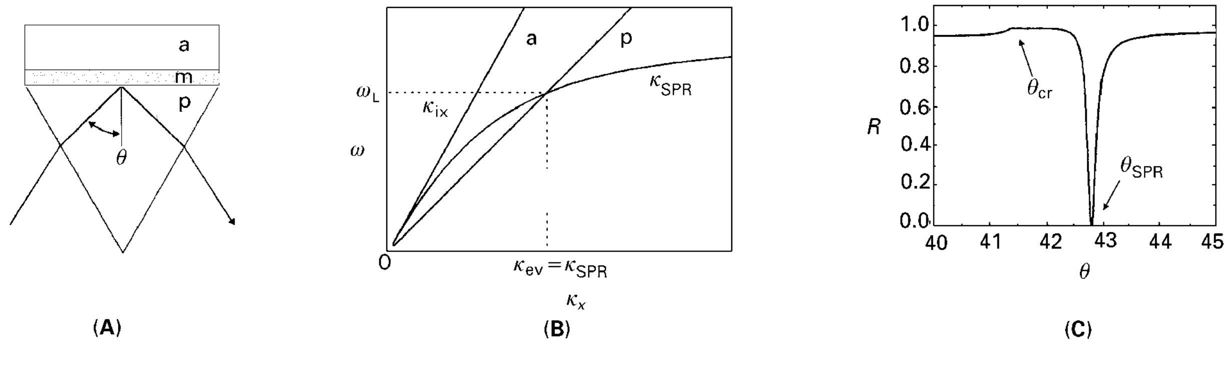



**Figure 1** The SPR experiment. (A) a metal layer m is sandwiched between two dielectric media a and p. The direction *x* is defined parallel to the layer structure; (B) dispersion relation of SPs; see text; (C) a typical SPR curve when medium a has a refractive index $n_a = 1$. The angle $\theta_{cr}$ corresponds to the critical angle for the a/p interface.

### Choice of metal layer

A typical application of SPR spectroscopy is to monitor a layer growth on the metal layer. Such a process can be modelled as a changing $\varepsilon_a$, which experimentally translates into a changing resonance angle $\theta_{SPR}$. **Figure 1C** depicts a reflectance curve obtained from a SPR experiment. From the figure it is clear that in order to obtain maximum sensitivity to variation in $\varepsilon_a$ the slope $dR/d\theta$ has to be as large as possible. For a given $\varepsilon_a$ this depends in a complicated way both on the value of $\varepsilon_m(\lambda)$ and on the thickness of the metal layer. Generally, it can be said that smaller real and imaginary parts, $Re(\varepsilon_m)$ and $Im(\varepsilon_m)$, result in smaller resonance halfwidths, whereas the metal layer thickness is an important parameter determining the minimum reflectance. From **Table 1**, which gives an impression of $\varepsilon_m(\lambda)$ for some metals, it can be concluded that in the visible wavelength range silver is expected to exhibit the narrowest resonances. This is indeed found experimentally; however, in many situations where experiments are performed in a water or air environment, silver tends to undergo unwanted interactions with its environment, making this a less attractive material. In this respect gold is a much more stable material and thus this material has become the standard metal for SPR purposes, despite its resonance width that at $\lambda = 633$ nm is approximately three times larger than that of silver.

For a given metal the optimum layer thickness depends on the wavelength used, as illustrated in **Figure 2**; also the remarks on the importance of the value of $\varepsilon_m(\lambda)$ are apparent from the figure. For any wavelength within the visible range a gold or silver layer thickness can be found corresponding to a vanishing minimum reflectance; for the often used helium-neon laser wavelength $\lambda = 633$ nm this is approximately 47 nm for gold and 49 nm for silver.

**Practical aspects** As already mentioned, in order to be able to excite nonradiative SPs it is important to have available smooth, homogeneous metal layers with a well-defined layer thickness. An appropriate way to prepare these is to sputter or to evaporate the metal at a high rate (~1 nm $s^{-1}$) in a vacuum chamber ($10^{-6}$ mbar) on the substrate of choice. In order to increase the adhesion between metal and underlying dielectric, the support is often precoated with a 2 nm Ti or Cr layer. It is not mandatory to coat directly on the prism; alternatively, a flat substrate, such as a microscope glass cover slip, can be used as a substrate. After coating, this plate is optically connected to the prism using a matching oil; care has to be taken that, in order to avoid spurious reflections, plate, oil and prism have the same refractive index in the wavelength region of interest.

Although a simple incandescent light source can be used for SPR, the use of a small HeNe or diode laser is far more convenient, in view of their well-defined wavelength and high degree of collimation.

**Table 1** Real and imaginary parts of dielectric constant for some metals

| *Metal* | *λ (nm)* | $\varepsilon_{re}$ | $\varepsilon_{im}$ |
|---|---|---|---|
| Aluminium | 600 | –29.8 | 7 |
| | 700 | –46.6 | 22 |
| | 900 | –55.5 | 30 |
| Silver | 500 | –8.23 | 0.3 |
| | 700 | –21.3 | 0.7 |
| | 900 | –38.7 | 1.3 |
| Gold | 600 | –8.37 | 1.2 |
| | 750 | –18.2 | 1.2 |
| | 900 | –28.5 | 1.8 |

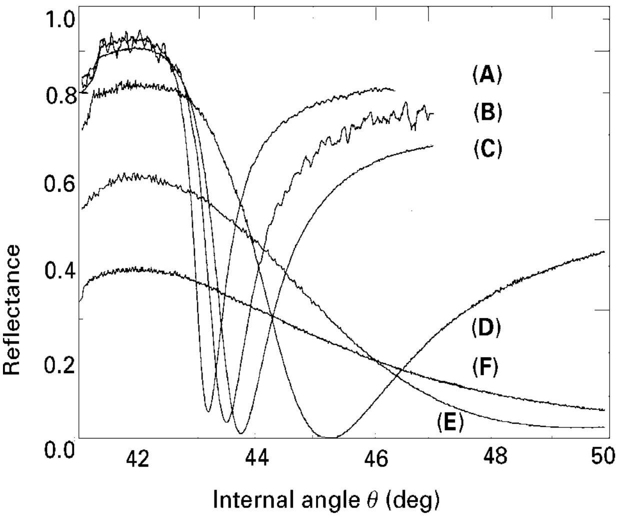



**Figure 2** SPR reflectance curves for a bare gold layer with thickness 44 nm, measured with different wavelengths: (A) 676 nm; (B) 647 nm; (C) 633 nm; (D) 568 nm; (E) 514 nm; (F) 488 nm.

## SPR instrumentation

### Rotation stage

The most straightforward method to measure SPs is depicted in **Figure 3**, where the prism/gold assembly is placed on top of a rotation stage (angular resolution typically 0.01 degree). With such an arrangement, all the features of a SPR curve can be accurately determined. The light detectors consists of simple large area photodiodes whose output is fed into a low-noise current-to-voltage amplifier. The excitation laser beam is polarized such that both s- and p-polarized light enter the prism. The use of a

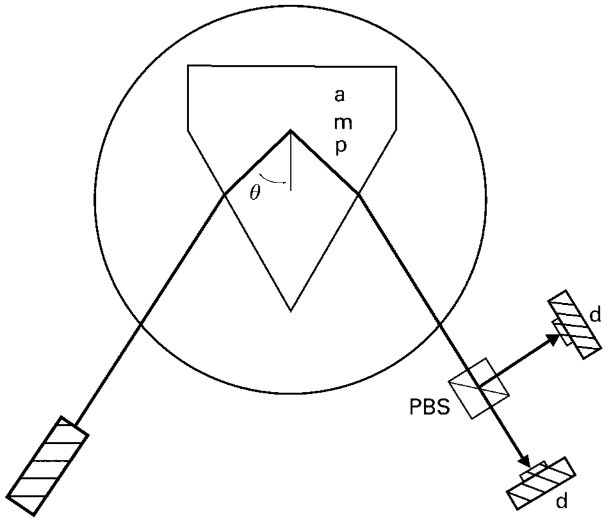



**Figure 3** Rotation stage SPR setup. d: detectors; PBS; polarizing beamsplitter.

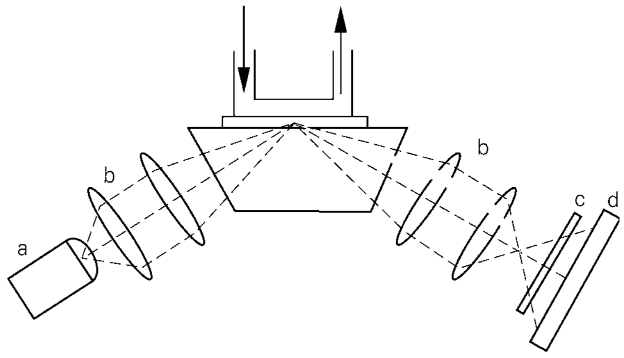



**Figure 4** Use of focused beam. a: laser diode; b: focusing optics; c: neutral density filter; d: diode array. A flow cuvette is placed on top of the metal layer. Reproduced with permission from Sjölander S (1991) *Analytical Chemistry* **63**: 2338. © American Chemical Society.

polarizing beamsplitter serves two purposes: one detector measures only the reflected s-polarized light whose intensity is only affected by laser intensity fluctuations and angular dependent refraction losses, whereas the other detector in addition measures SPR effects. The ratio between the two outputs then provides a signal proportional to the net SPR response. In order to obtain an absolute angular read-out with sufficient accuracy the critical angle can be measured; this angle is very accurately known for a certain $\varepsilon_p/\varepsilon_a$ interface, and is an easily discernible feature in a SPR curve (cf. **Figure 1C**).

The obvious disadvantage of this setup is that it is difficult to monitor fast changes in the SPR curve. To be able to monitor time-dependent changes, the orientation of the rotation stage relative to the light beam is often set such that the reflectance is about halfway between maximum and minimum. A changing $\varepsilon_a$ is then approximately linear in the monitored changed reflectance. However, one has to assume that a change of $\varepsilon_a$ leaves the width of the SPR curve unchanged.

### Use of a focused beam

A more elaborate setup is shown in **Figure 4**. Here, SPs are excited by a focused diode laser beam such that the beam waist is slightly displaced from the top metal/dielectric interface. In this way, a certain distribution of angles is present at the interface. The reflected beam is imaged onto an array consisting of a large number (50–100) of photodiodes. The angular range corresponding to resonance will exhibit a low intensity; consequently, the output of the diode array is a measure of the angular dependent reflectance. Although the distance between individual diodes is relatively large (~ 20 μm) the use of numerical interpolation techniques makes it possible to obtain an angular resolution of ~ 1 millidegree. Because no moving parts are involved, this system is capable of monitoring relatively fast-changing SPR characteristics: the temporal resolution will now be determined by the read-out rate of the diode array. A drawback of this system is that obtaining an *absolute* angular read-out is not as straightforward as in the case of the rotation stage, in view of the limited angular range that is applied to the metal/dielectric interface.

### Photothermal detection

When a nonradiative SP decays it generates heat at the metal/dielectric interface. This observation can be fruitfully exploited to detect the presence of SPs. In a *photoacoustic* approach, the prism/metal assembly is placed in an airtight chamber. SPs are excited with an intensity-modulated light beam; the resulting periodic heat flow from metal to dielectric as a result of plasmon decay causes a periodic pressure variation in the dielectric which can be detected by a microphone in the sample chamber. A significant difference as compared with optical detection is that SPR is now detected as a *maximum* in response. This method is mainly used in fundamental studies where one is interested in SP decay. The method is less suited in situations where the dielectric of interest consists of a fluid, such as water.

In a related approach, the heat flow is detected optically (*photothermal deflection spectroscopy*). A representative setup is shown in **Figure 5**. The prism/metal assembly is placed upon a rotation stage, and SPs are excited in the usual way. Heat

produced by the decaying plasmon results in a gradient of the refractive index immediately above the metal layer. Consequently, the propagation direction of the light beam of the probe laser will be deflected and this can be measured, e.g. by using a position-sensitive detector. By modulating the SPR excitation beam and synchronous detection of the detector output, a differential signal is obtained whose angular dependence is directly related to the SPR curve. Using a HeNe laser as the excitation source an angular resolution of ~0.01 degree can be obtained; contrary to the setups where the reflectance is measured, this resolution can be improved by employing higher laser powers. It has been demonstrated that this deflection method is also useful in water solutions.

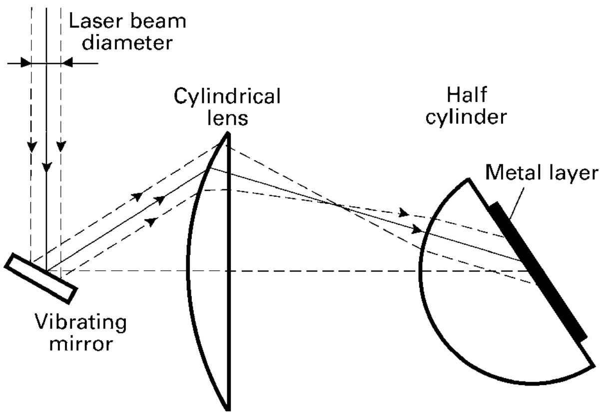



**Figure 6** Vibrating mirror setup. a: vibrating mirror; b: cylindrical lens. Reproduced with permission from Lenferink ATM (1991) *Sensors and Actuators B (Chemical)* **3**: 261. Copyright 1991, with kind permission from Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington, OX5 1GB, UK.

### Sensor configurations

As already mentioned, an important application of SPR is in the field of chemical sensors. Such a SPR sensor system should be simple, compact, and relatively inexpensive, while retaining the angular sensitivity. Some representative SPR sensor systems will now be discussed.

**Vibrating mirror device** In the system as depicted in **Figure 6** the angle under which the light from a laser diode enters the interface is made time-dependent by means of a mirror, vibrating at a frequency of ~50 Hz. If the optical system in the excitation path is designed properly, the light spot during an angular scan of ~5 degrees is stationary on the interface to within 0.2 mm, while the beam divergence is kept within 0.02 degrees. Although this setup can be used to monitor the complete SPR curve, it is dedicated to determine only the angle of minimum reflectance $\theta_{SPR}$, which for the majority of SPR applications is the main parameter of interest. This can be conveniently accomplished as follows: during one cycle of the vibrating mirror the beam traverses the reflectance minimum twice. The time span $\Delta t$ between these two minimum is measured using appropriate electronics. If $\theta_{SPR}$ changes, $\Delta t$ changes accordingly.

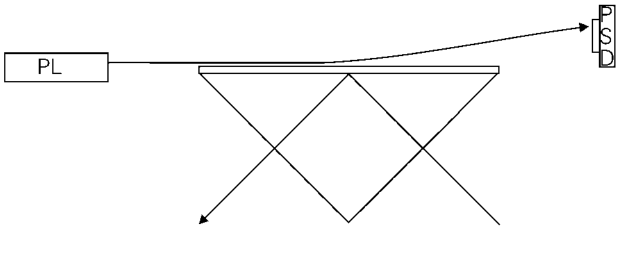



**Figure 5** Basic photothermal detection setup. The position of the deflected beam of a probe laser PL is measured by a position-sensitive detector PSD.

Connected to a personal computer, this provides a versatile means to detect a SPR minimum with an angular resolution of 1 millidegree. The time resolution of such a setup will be determined by the vibration frequency of the mirror. (Note that vibrating mirrors operating at frequencies up to 10 kHz are available). For some applications the difficulty of obtaining an absolute angular read-out will be a disadvantage.

**Use of diffraction-gratings** In the foregoing configurations a SP was excited using a prism. However, plasmons can also directly be produced on a grating surface coated with a thin metal layer. The condition for SPR to occur is determined by the grating periodicity (typical value 2000 lines $mm^{-1}$) and the angle of incidence of the light beam, whereas the value of the reflectance minimum is solely determined by the grating depth (optimum value ~50 nm). Any of the above described setups can be used to detect SPs in this configuration. An advantage of this approach is that gold-coated gratings can be very easily and inexpensively manufactured as disposable replicas, once a holographic master grating has been produced. An important disadvantage is that the light beam has to enter the dielectric/metal interface from the dielectric side, implying that this is only a valuable approach for transparent media.

**Fibre-optic devices** In situations where the sensing surface should be remote from the signal processing equipment (such as in a hostile environment or in a living organism) the use of optical fibres can provide a practical solution. Apart from the trivial use of a fibre to transport the light to and from a separate

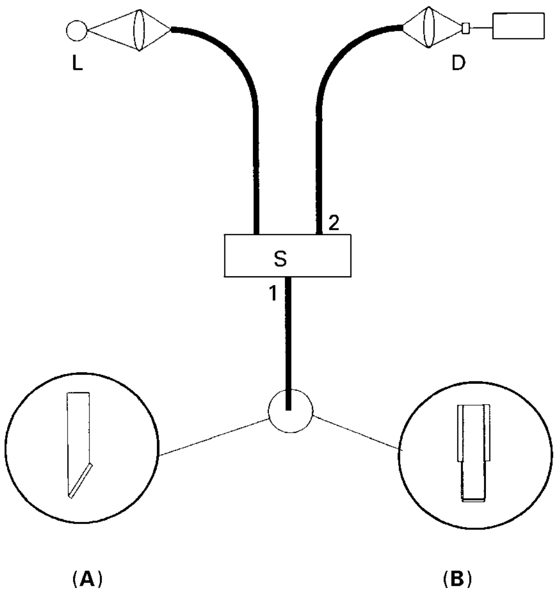



**Figure 7** Use of fibre optics. L: light source; D: detector; S: fibre splitter/combiner; 1, 2: output ports. (A): single mode fibre; (B): multimode fibre (cf. text).

prism/metal assembly, it has also been demonstrated that the fibre can also be used as an *intrinsic* sensing interface, where part of the fibre surface replaces the prism. This concept is schematically shown in **Figure 7**. Light is fed into a fibre connected to an optical splitter/combiner S. The output port 1 of S is connected to a probe fibre, which is decladded at the distal end. After interaction with this end interface (see below) the light is reflected and again enters S, where it is transmitted to output port 2. A fibre connected to this port transports the light to a detector.

Several possibilities exist for the design of the decladded fibre tip. In **Figure 7A** the tip of a monomode fibre is cleaved at a specific angle and is subsequently coated with a metal layer of appropriate thickness. As the wave vector (cf. Eqn [2]) of the propagating light in a monomode fibre is well defined the cleaving angle can be calculated such that excitation of SPs is possible in the $\varepsilon_a$ region of interest. Again, SP excitation is detected by a decreased reflected intensity. In order to obtain optimum response it is advisable to use polarization-preserving fibres. In **Figure 7B** another approach is depicted, involving the use of a multimode fibre. Here, the circumference of the tip is metal coated; additionally a mirror is deposited on the distal end to minimize reflection losses. For a multimode fibre a (discrete) range of wave vectors simultaneously propagates; in order to be able to detect changing characteristics of the metal/dielectric interface it is therefore necessary to use a broadband light source. The SPR information will now be contained in the wavelength-dependent reflectance. An advantage over the monomode configuration is that a broader range of SP wavevectors can be excited.

## SPR microscopy

The foregoing discussion has ignored the possibility to obtain spatially resolved information from a SPR experiment. However, it is known that SPs are collective electron oscillations *with a limited coherence length* $L_x$, implying that two regions in the metal with a mutual distance larger than $L_x$ are capable of supporting SPs which are mutually independent. This phenomenon can be exploited to image structures on top of a metal layer that have a distribution of different $\varepsilon_a$: if the angle of light incidence is such that one particular $\varepsilon_a$ corresponds to resonance, then regions with another $\varepsilon_a$ will exhibit larger reflectance. An example of such an experiment can be seen in **Figure 8**, where the spatially resolved reflectance of an inhomogeneous monomolecular layer, consisting of molecules oriented either tilted or perpendicularly to the surface is depicted.

Of course, in such an experiment one aims to obtain maximum lateral resolution, while retaining the vertical (thickness) resolution. A general rule is that $L_x$ decreases for increasing resonance halfwidths (cf. **Figure 1C**). However, as was pointed out in an

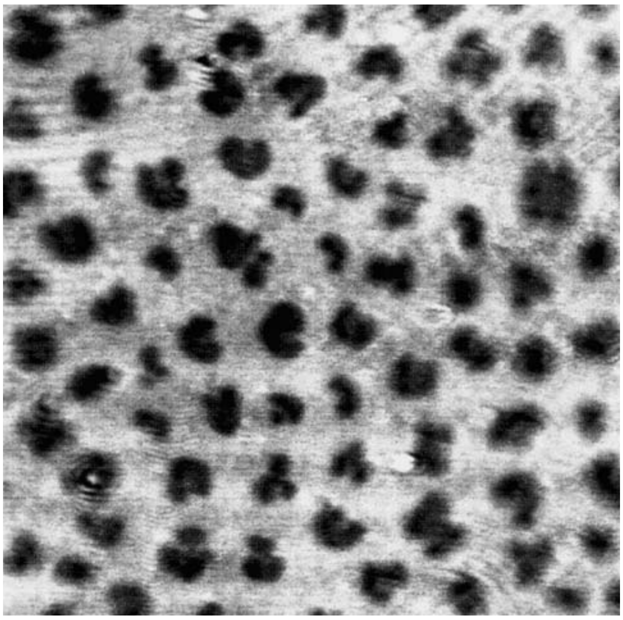


**Figure 8** SP microscopic image of an inhomogeneous monolayer. Dimensions of image 0.2 × 0.2 mm$^2$. Thickness difference of the two types of domains is less than 0.4 nm. Lateral resolution approximately 3 μm.

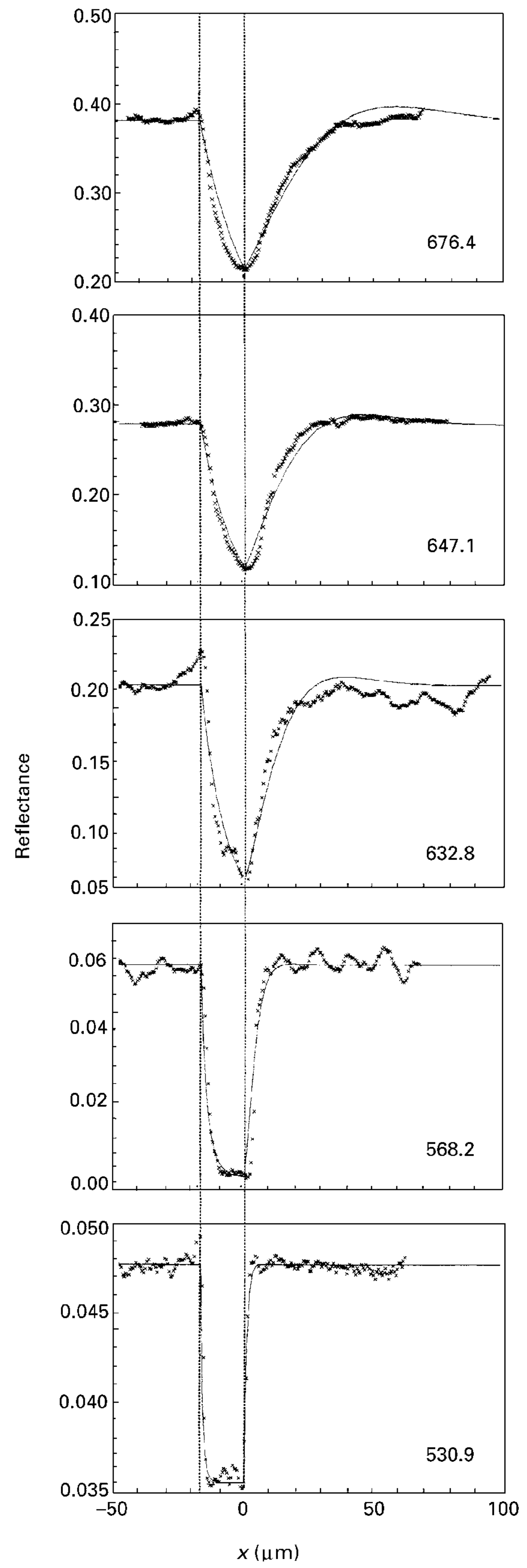



Figure 9 Lateral resolution in SPR microscopy. The vertical lines denote the physical width of a 2.5 nm $SiO_2$ ridge. Number insets correspond to the various wavelengths used.

earlier section, such an increasing SPR halfwidth simultaneously deteriorates vertical resolution. Therefore, for each particular situation a balance has to be sought between these two contradictory conditions. This is exemplified in **Figure 9** where SPR results are shown of a 2.5 nm $SiO_2$ ridge on gold for various wavelengths. It is clearly seen that at the shortest wavelength used the lateral resolution is the highest (~ 2 μm), but the slope is the lowest (compare also **Figure 2**), indicating that it is difficult to obtain sufficient intensity contrast to detect sub nm height differences, which is usually not a problem with standard SPR experiments.

## Instrumentation

The first SPR microscopy experiments were done using a scanning focused beam. However, a straightforward approach, by imaging a collimated beam, proved to be much faster and instrumentally much simpler while at both the lateral and vertical resolution are comparable or better. Therefore, only this last approach will be discussed in more detail.

**Figure 10** gives an overview of a representative setup. A prism/gold layer assembly and an optional cuvet system are placed on a rotation stage (angular increments 0.01 degree). SPs can be excited using light from a HeNe or an argon/krypton ion laser. This last light source has the advantage of providing a large number of high power wavelengths over the whole visible range. After having passed through a Pockels cell and a spatial filter, the light spot

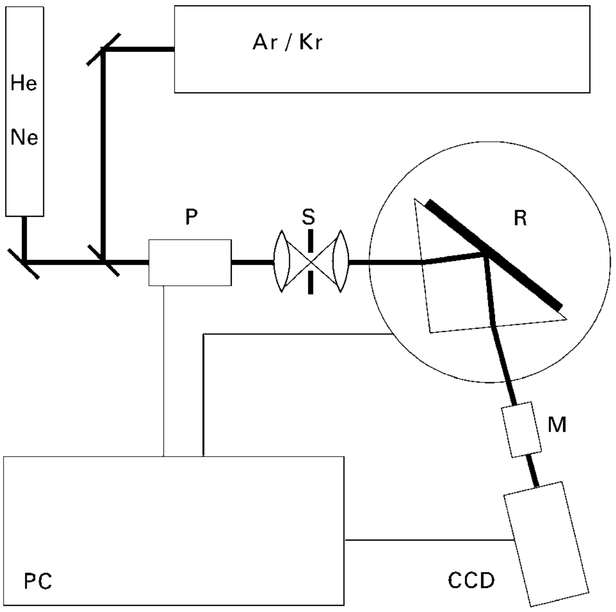



Figure 10 SPR microscope setup. P: Pockels cell; S: spatial filter; R: rotation stage; M: objective; CCD: video camera.

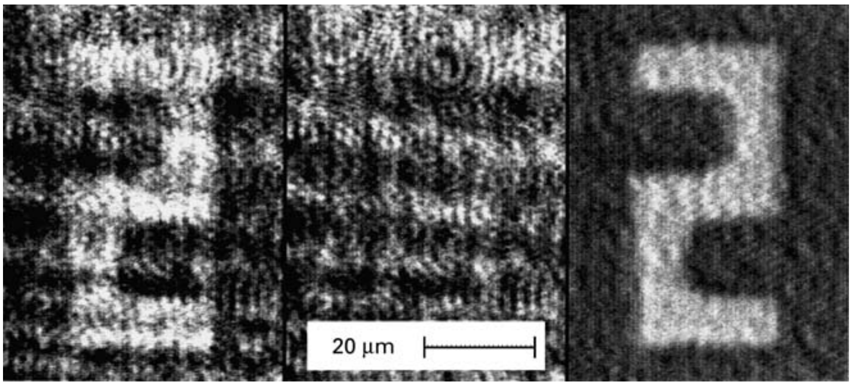



**Figure 11** Improvement of image quality by dividing s- and p-polarized responses. From left to right: p response; s response; p/s response.

entering the prism has an area of ~1 cm$^2$. The intensity profile of the reflected beam is recorded with a microscope consisting of an objective and CCD video camera. It is important that the video camera has a response linear in the light intensity (see below). Depending on the light intensity and required magnification, objectives with focal distances between 5 and 50 mm are used. The Pockels cell and rotation stage are controlled by a microcomputer and the obtained images are also stored in this same computer by using a frame-grabber card.

It also proves possible to use the aforementioned gratings in a SPR microscopic experiment. Rotation of such a plate about the normal is equivalent to changing the angle of incidence of the exciting light. Compared to the standard prism setup this configuration is very compact and can easily be integrated in a conventional light microscope.

#### Improvement of image quality

Apart from the diffraction limit of the use of imaging lens there are a number of factors determining the eventual quality of a SPR microscopy image. As already mentioned, the use of shorter wavelengths generally improves lateral resolution, but simultaneously results in lower intensity contrasts for a given height difference. With such relatively low contrasts the use of a digital frame-grabber (which has a typical dynamic range of $2^8$) can result in *quantization effects during image acquisition*, which become apparent if a low-contrast image is digitally amplified. To avoid this effect an option is to add a number of images with an effective dynamic range substantially larger than that of the frame-grabber used.

Integrating a number of images also results in averaging. This can be particularly important if shot noise is present when working with low light levels. Such an averaging will result in a signal-to-noise (SNR) improvement with a factor of $n^{1/2}$ where $n$ is the number of added images. Another possible way to increase SNR is to increase laser power but this option is of limited value in view of the resulting destructive heating effects on the sample.

Another experimental problem is *lateral nonhomogeneity of the incoming laser beam intensity profile*. This can be partially solved by spatial filtering as indicated in **Figure 10**; however, if there are any unwanted, spurious reflections in the light path between the spatial filter and CCD camera, then again a beam nonhomogeneity will occur owing to the relatively large coherence length of the laser light used. An appropriate method is the use of a Pockels cell with simultaneous digital image processing. Such a cell can be configured such that the application of a voltage results in transmission of either p- or s-polarized light, without moving any part in the light beam. In the SPR microscope the Pockels cell is employed to acquire two images using p- and s-polarized light, respectively. The image with p-polarization contains the SPR microscopic image together with the contrast generated by spatial nonhomogeneities, whereas the image obtained with s-polarization only contains the same unwanted nonhomogeneities. Because the CCD camera output is linear in intensity, the ratio of the two images is a true representation of the SPR-related reflectance variations over the imaged surface. An example of the result of such an approach can be seen in **Figure 11**. Although lateral nonhomogeneities are still visible in the ratioed SPR image, the resolution is approximately 2 μm; for this image the difference in reflectance between the two regions is ~0.01.

#### Combination with other microscopies

Since the birth of scanning probe microscopy several attempts have been undertaken to merge this

approach with SP microscopy. In **Figure 12** a schematic diagram is given where a SPR microscope is combined with a scanning tunnelling microscope (STM). The hemisphere, serving as a prism, is metal-coated, and SPs are excited in the usual way. The metallic STM tip is brought close to the surface (distance ~5 nm). The area of interaction between the tip and the surface is imaged on a photodiode. The tip is allowed to scan laterally over the surface and the photodiode output is monitored simultaneously. It was found that small corrugations on the surface which were detected by the tunnelling current of the STM were equally well monitored by the SPR reflectance. In this way a lateral resolution of ~5 nm could be demonstrated *in a SPR image*. Similar conclusions could be made for a dielectric atomic force microscopy (AFM) tip. Even better results can be obtained if the conically scattered SPR radiation is monitored rather than the specularly reflected light.

A discussion on the contrast mechanism in both setups is far beyond the scope of the present article; a pragmatic remark is that, in most cases, an AFM type tip is preferable in view of the fact that in the latter case no conductivity conditions have to be imposed on the samples of interest.

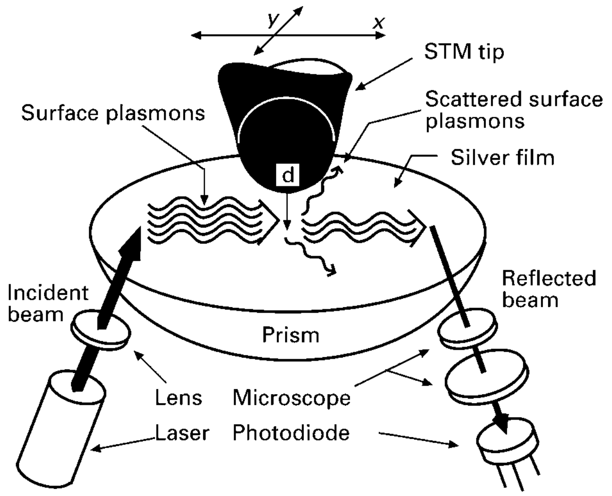



**Figure 12** Scheme of a combined SP microscope and a scanning probe microscope. The tip can either be a STM or an AFM tip. Reproduced with permission from Specht M (1992) *Physical Review Letters* **68**: 476–479. American Physical Society.

## List of symbols

$k_x$ = wave vector; $L_x$ = coherence length; $n$ = refractive index; $n$ = number of added images; $R$ = reflectance; $\Delta t$ = time span; $\varepsilon$ = dielectric constant; $\lambda$ = wavelength *in vacuo*; $\theta$ = angle of light; $\theta_{SPR}$ = angle of minimum reflectance.

*See also:* **Fibre Optic Probes in Optical Spectroscopy: Clinical Applications; Fourier Transformation and Sampling Theory; Light Sources and Optics; Photoacoustic Spectroscopy, Applications; Photoacoustic Spectroscopy, Theory; Scanning Probe Microscopes; Scanning Probe Microscopy, Applications; Scanning Probe Microscopy, Theory; Surface Plasmon Resonance, Applications; Surface Plasmon Resonance, Theory.**

## Further reading

Agranovich VM and Maradudin AA (1982) *Surface Polaritons*. Amsterdam: North-Holland Publishing Company.

Kretschmann E (1971) Die bestimmungen optischer Konstanten von Metallen durch Anregung von Oberflächen plasmaschwingungen. *Zeitschrift Physica* **241**: 313–321.

Lawrence CR and Geddes NJ (1997) SPR for Biosensing. In: Kress-Rogers E (ed) *Handbook of Biosensors and Electronic Noses*. Boca Rata, FL: CRC Press.

Raether H (1988) *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Berlin: Springer Verlag.

Rothenhäusler B and Knoll W (1988) Surface plasmon microscopy. *Nature* **332**: 615–616.

Specht M, Pedarnig JD, Heckl M, and Hansch TW (1992) Scanning plasmon near-field microscope. *Physical Review Letters* 68: 476–478.

# Surface Plasmon Resonance, Theory

**Zdzislaw Salamon** and **Gordon Tollin**,
University of Arizona, Tucson, AZ, USA

Copyright © 1999 Academic Press

| SPATIALLY RESOLVED SPECTROSCOPIC ANALYSIS |
|---|
| Theory |

The physics of surface plasmons propagating along a metal/dielectric interface has been studied intensively, and their fundamental properties have been found to be in good agreement with theoretical concepts based upon the plasma formulation of Maxwell's theory of electromagnetism. The phenomenon has been utilized extensively by physical scientists in studies of the properties of surfaces and thin films. Current interest in the properties of thin surface coatings stems partly from increased applications to thin film devices and, in particular, to recent developments in biosensor devices. This article focuses on the characterization of the surface plasmon resonance phenomenon, with emphasis on the conditions of optical excitation of plasmon resonance and the theoretical analysis of different types of surface resonances.

## Description of surface plasmons

The concept of surface plasmons originates from the plasma formulation of Maxwell's theory, where the free electrons of a metal (or a conductive electron gas) are treated as a high density liquid (plasma). Plasma oscillations in metals are collective longitudinal excitations of the conductive electron gas, and plasmons are the quanta representing these charge-density oscillations. Such oscillations can exist in the bulk media, and can also be localized on an interface between a metallic and a dielectric surface, along which they propagate as waves. In the latter case they are called surface plasmons (SP), or surface polaritons. The spreading electron density fluctuations generate a surface-localized electromagnetic wave which propagates along the plane interface between the metal/dielectric media, with the electric field normal to this interface and vanishing exponentially with penetration distance from the interface. These characteristics of the electromagnetic field are the same as those describing the guided surface waves (also known as evanescent waves) generated optically under total internal reflection conditions when all the incident light is reflected at the boundary of the incident and emerging media. Under such conditions the electric and magnetic fields do not, however, stop abruptly at the boundary. Rather they penetrate a distance into the emerging medium in the form of a surface wave. Although the existence of guided surface electromagnetic waves has been theoretically predicted from Maxwell's equations and investigated during the first decade of the 20th century, it is only since 1960 that they have attracted the interest of the experimentalist and the term 'surface plasmon' has been used. This is partly due to the fact that methods have been developed which enable the optical excitation and detection of surface-bound electromagnetic waves (see below).

It has been shown that Maxwell's equations have solutions resulting in the generation of surface plasmon electromagnetic waves only when the following conditions are fulfilled: (1) one of the adjacent media (i.e. the surface active medium which generates surface plasmon waves) has a negative value for the real part of its complex dielectric constant $\varepsilon$, (2) a component of the wave vector along the interface between these two media, $K$, satisfies an equation which involves the dielectric constants of both media. These conditions are discussed in more detail in the following sections.

## Optical excitation of surface plasmons

The condition that the surface active medium has a negative dielectric constant results in several experimental requirements which have to be fulfilled in order to be able to generate surface plasmon waves. First, not all materials can be utilized as surface active media; gold and silver are the best examples of materials which can support surface plasmons. In addition, the electromagnetic wave in the surface active medium is an evanescent wave under all circumstances. Therefore, in order to satisfy boundary conditions in the surface active and dielectric media, which require that the tangential components of the electrical and magnetical fields be continuous across the boundary (i.e. that a component of the wave vector along the interface must be the same in both media), surface plasmons can only be optically generated by an evanescent wave whose wave vector matches that of the evanescent surface plasmon

electromagnetic wave $K_{sp}$. The latter is given by the following equation:

$$K_{\rm SP} = \omega/c(\varepsilon_1\varepsilon_2/(\varepsilon_1+\varepsilon_2))^{1/2} \qquad [1]$$

where $\omega$ is the frequency of the surface plasmon wave, $c$ is the velocity of light *in vacuo*, and $\varepsilon_1$ and $\varepsilon_2$ are the complex dielectric constants for the surface active and dielectric (emerging) media, respectively. The complex dielectric constant $\varepsilon$ is directly related to the complex index of refraction $N = c/v = n - \mathrm{i}k$, by the following relation:

$$\varepsilon = \varepsilon' + \mathrm{i}\varepsilon'' = N^2 \qquad [2]$$

where the real part of the complex dielectric constant is $\varepsilon' = n^2 - k^2$ and the imaginary part is $\varepsilon'' = 2nk$, $n$ is the refractive index, $k$ is the extinction coefficient, and $v$ is the velocity of light in the dielectric medium.

These constraints require that surface plasmons cannot be directly excited by incident light, and produce a resonance condition for the wave vector of the evanescent wave which excites the plasmons. Furthermore, indirect SP excitation can only be achieved by an evanescent wave generated by p-polarized incident light under total internal reflection conditions. This occurs when a beam of light propagating through an incident medium (e.g. a prism) of higher refractive index ($n_0$) meets an interface with a second (emerging) medium of lower refractive index ($n_2$) at incident angles $\alpha$ larger than the critical angle for total reflection $\alpha_c$, given by $\alpha_c = \sin^{-1}(n_2/n_0)$. Since the prism has a dielectric constant, $\varepsilon_0 = n_0^2$ (i.e. $k_0 = 0$) the following relation must be satisfied: $\varepsilon_0 > \varepsilon_2$ (or $n_0 > n_2$), where $\varepsilon_2$ is the dielectric constant and $n_2$ is the refractive index of the emergent dielectric medium). Although metal-coated diffraction gratings may be used instead of prisms, they require a greater complexity of fabrication without additional benefits, and so they have not been widely used and will not be discussed further.

Despite being totally reflected, the incident beam generates an evanescent electromagnetic field that penetrates a small distance, the order of a wavelength, into the second medium, where it propagates parallel to the plane of the interface. This electromagnetic field can be used to measure the optical properties of interfaces and thin films in various ways. The two main types of applications of optically generated evanescent waves are those based on waveguiding systems, and those used to excite surface plasmon resonance (SPR).

In waveguiding techniques the measured interface (or thin film) is placed in the evanescent region of a guided mode propagating in a dielectric waveguide structure. The optical properties of the interface (or thin film) affect the propagation characteristics of the evanescent surface wave causing changes in the resonant waveguide mode. These changes can be measured by a variety of optical techniques including attenuated total reflection (ATR) spectroscopy, total internal reflectance fluorescence (TIRF), where the evanescent wave is employed to excite fluorescence from molecules in the interface, or interferometry.

In order to generate SPR using an evanescent wave which is produced during an internal total reflection of the light from a prism whose base is coated with a thin metal film, the following two conditions have to be fulfilled. First, a component of the incident light vector parallel to the prism/metal interface, $K_{ph}$ which is described by the following equation:

$$K_{\rm ph} = (\omega/c)\ \varepsilon_0{}^{1/2} \sin\alpha \qquad [3]$$

must be identical with the surface plasmon wave vector, $K_{SP}$ (see Eqn [1]):

$$K_{\rm ph} = K_{\rm SP} \qquad [4]$$

The value of $K_{ph}$ can be adjusted to match that of the surface plasmon wave by changing either $\omega$, i.e. the frequency (or wavelength) of the excitation light, or $\alpha$, i.e. the incident angle (see Eqn [3]). Secondly, because the oscillations of the free electrons in a metal film can only occur along the normal to the plane of the metal surface, only p- (or TM, transverse magnetic) polarization of the incident light is effective in generating surface plasmons. At the resonance condition (Eqn [4]), the incident light is coupled into a SP wave travelling along and bound to the outer active (metal) surface, and the phenomenon is known as surface plasmon resonance (SPR). The SP wave is nonradiative, and can either decay into photons of the same frequency $\omega$ if coupling by the surface roughness takes place, or be converted to heat.

In practical terms there are two configurations, both based on the ATR technique available to optically excite SPR at the metal/dielectric (or emerging medium) interface. In the first, the Kretschmann configuration, the prism is in direct contact with the surface active (metal) medium. In the second, the Otto configuration, the prism is separated by a thin layer of a dielectric (inactive) medium at a distance of approximately one wavelength of excitation light from the metal film. The practical consequences of

using these configurations to excite SPR are described below.

## Analysis of surface plasmon resonances excited by light

Although the SPR phenomenon can be accurately described in physical terms as propagating oscillations of free electrons at a metal surface, there is a simpler and more general approach which has been used to describe light propagation through optically anisotropic layered materials whose properties vary only along the layer normal. This is a standard mathematical tool used to describe the optical properties of multilayered thin-film devices, and the SPR phenomenon can be seen as a straightforward result of the application of such thin-film electromagnetic theory. The analysis applies Maxwell's equations to describe the propagation of a plane electromagnetic wave through a multilayer assembly of thin dielectric films, and is based on the following properties of the structure. The film is considered thin when the phase differences between the various waves in the assembly are constant with time. This condition invariably holds for films which have thicknesses of not more than a few wavelengths. A second requirement is that the thin-film materials are characterized by a complex refractive index, which in the optical region is numerically equal to the optical admittance in free space units. This is defined by the ratio of the total tangential electric ($B$) and magnetic ($C$) field amplitudes of the electromagnetic wave ($Y = C/B$). Additional constraints are that the tangential components of the electric and magnetic field vectors of the electromagnetic wave are continuous across the interface between any two thin films. Also, that in any thin film the amplitude reflection coefficient or reflectivity, sometimes known as the Fresnel reflection coefficient ($r$), defined as the ratio of the amplitudes of the incident and reflected electric field vectors, at any plane within the layer is related to that at the edge of the layer remote from the incident wave, $r_m$, by $r = r_m e^{-2i\delta}$, where $\delta$ is the phase thickness of that part of the layer between the far boundary, m, and the plane in question. Solution of Maxwell's equations describing the propagation of a plane, monochromatic, linearly polarized, and homogeneous electromagnetic field within a multilayer thin-film assembly with the above mentioned attributes results in a relationship which connects the total tangential components of the electric and magnetic field amplitudes at the incident interface with the total tangential components of electric and magnetic field amplitudes which are transmitted through the final interface. This result is of prime importance in describing the optical properties of thin films and forms the basis of almost all calculations. It has the following standard matrix notation, known as the characteristic matrix of the assembly:

$$\begin{bmatrix} B \\ C \end{bmatrix} = \prod_{j=1}^{s} \begin{bmatrix} M_j \end{bmatrix} \begin{bmatrix} 1 \\ y_{j+1} \end{bmatrix} \qquad [5]$$

where $s$ is the number of layers deposited on the incident medium; $y_j$, gives the characteristic oblique admittance of the $j$th layer ($y_{j+1}$ for the emerging medium): $y_j = y_j/\cos\alpha_j = (n - ik)_j/\cos\alpha_j$ for p-polarized incident light, or $y_j = Y_j \cos\alpha_j$, for an s-polarized electromagnetic wave, and $\mathbf{M}_j$ is known as the characteristic matrix of the $j$th thin layer and has the following form:

$$\mathbf{M}_j = \begin{bmatrix} \cos\delta_j & i(\sin\delta_j)/y_j \\ i\, y_j \sin\delta_j & \cos\delta_j \end{bmatrix} \qquad [6]$$

where $\delta_j = (2\pi Y_j/\lambda)t_j \cos\alpha_j$, gives the phase thickness of layer $j$ in the thin-film assembly; $(Y_0 \sin\alpha_0) = (Y_j \sin\alpha_j)$, is the complex Snell's Law which relates $\alpha_j$ to $\alpha_0$, the angle of incidence in the incident medium; $t_j$ is the thickness of the $j$th layer, and $\alpha_0$ and $\alpha_j$ are the incident angles of light of wavelength $\lambda$ for the incident medium (a prism in the SPR system) and the $j$th layer, respectively.

There are two important conclusions which can be deduced from Equation [5]. First, the characteristic matrix of an assembly of $j$ layers is simply the product of the individual matrices taken in the correct order. Second, sufficient information is included in Equations [5] and [6] to allow the full analysis of the electromagnetic field generated at each interface of a multilayer thin-film assembly, thereby yielding transmittance, absorbance, and reflectance for both p- and s-polarizations. Furthermore, the optical admittance presented at the incident interface by the system of layers and emerging medium is the product of the characteristic matrix of the assembly.

The optical admittance parameter has been introduced into thin-film optics with one specific aim, namely to visualize optical phenomena occurring within such systems by means of a graphical representation of the optical events known as the admittance diagram. Although this is one of a class of diagrams known collectively as circle diagrams, it is particularly powerful and attractive and therefore it is used extensively in thin-film optics.

In the case of SPR which is generated optically with the ATR technique, the reflectance of a multilayer system, $R$, defined as the ratio of the energy reflected at the surface of such a structure to the energy which is incident, is an especially important parameter and can be calculated from the following relation:

$$R = rr^* \qquad [7]$$

where:

$$r = (Y_0 B - C)/(Y_0 B + C) = (Y_0 - Y)/(Y_0 + Y) \qquad [7a]$$

is the amplitude reflection coefficient (reflectivity or Fresnel reflection coefficient), and:

$$\begin{aligned} r^* &= \{(Y_0 B - C)/(Y_0 B + C)\}^* \\ &= \{(Y_0 - Y)/(Y_0 + Y)\}^* \end{aligned} \qquad [7b]$$

is the complex conjugate of $r$. $Y_0$ is the admittance of the incident medium (which in the case of the present application to SPR is a non-light-absorbing glass prism, i.e. with $k = 0$, and therefore the $Y_0$ value becomes real and equal to the refractive index of the incident medium, $n_0$).

Equations [5], [6], [7], [7a] and [7b] comprise a full set of mathematical tools to examine optically excited surface plasmon resonance. Such an analysis can be applied to the following three types of resonances (see the next section): (i) conventional surface plasmon resonance (SPR), (ii) the resonance associated with long-range surface plasmons (LRSPR), and (iii) the resonance associated with coupling of plasmon resonances in a thin metal film with waveguide modes in a dielectric overcoating, known as coupled plasmon waveguide resonance (CPWR). Although in all three cases the excitation of surface plasmons is based on coupling light photons to plasmons by the ATR technique, these resonance systems differ in their detailed thin-film structural designs, as described in the next section.

## Variety of surface plasmon resonances

### Conventional surface plasmon resonance

In the most straightforward case, for which the hypotenuse of the prism is coated with a single high-performance metal (Ag or Au) layer (Kretschmann coupling), one can generate surface plasmons on the outer surface of the metal, as indicated in **Figure 1**. This shows that for either a 55 nm Ag or a 48 nm Au layer, as a consequence of an enormous increase in the intensity of the evanescent electromagnetic field, which is produced as a function of either the incident angle, $\alpha$, with $\lambda = \lambda_0 =$ constant (panel A), or the wavelength, $\lambda$, with $\alpha = \alpha_0 =$ constant (panel B). This very large increase of electric field amplitude, as compared to that of the incident light, with the characteristics of a sharp resonance and which can only be obtained with p-polarized excitation light, is a result of a resonant generation of free metal electron oscillations, and is known as conventional surface plasmon resonance (SPR). As anticipated in the previous sections, **Figure 1** demonstrates the dependency of SPR on the optical parameters of the metal films, resulting in a much stronger evanescent electric field obtained with silver than with gold. The easiest means of experimental detection of this phenomenon is by measuring the changes in intensity of the totally reflected light ($R$) which has been used to generate plasmons in the ATR arrangement (**Figure 2**). In both cases, i.e. for $R$ calculated either as a function of $\alpha$ with $\lambda = \lambda_0 =$ constant (**Figure 2A**), or as a function of $\lambda$ with $\alpha = \alpha_0$ = constant (**Figure 2B**), one obtains a resonance curve resembling in shape those shown in both panels of **Figure 1**. As demonstrated above, in conventional SPR the evanescent electromagnetic field reaches its maximum intensity on the outer metal surface and decays very rapidly with distance into the emerging dielectric medium. This effect is demonstrated in **Figure 3A**. In addition, **Figure 3B** illustrates the sensitivity of the SPR phenomenon to the thickness of the metal film.

### Long-range surface plasmon resonance (LRSPR)

The second type of resonance, long-range surface plasmon resonance (LRSPR), is generated in the same way as conventional SPR, but uses a thinner metal layer which is surrounded by dielectric media that are beyond the critical angle so that they support evanescent waves. An example of a calculated LRSPR reflectance versus incident angle curve (**Figure 4A**) demonstrates a very narrow resonance (curve 1) as compared to conventional SPR (curve 2). The distribution of the resonantly generated evanescent electric field intensity along the normal to the film planes, presented in **Figure 4B**, clearly indicates two important differences between this type of resonance and conventional SPR: (1) LRSPR involves two surface bound waves on both the inner

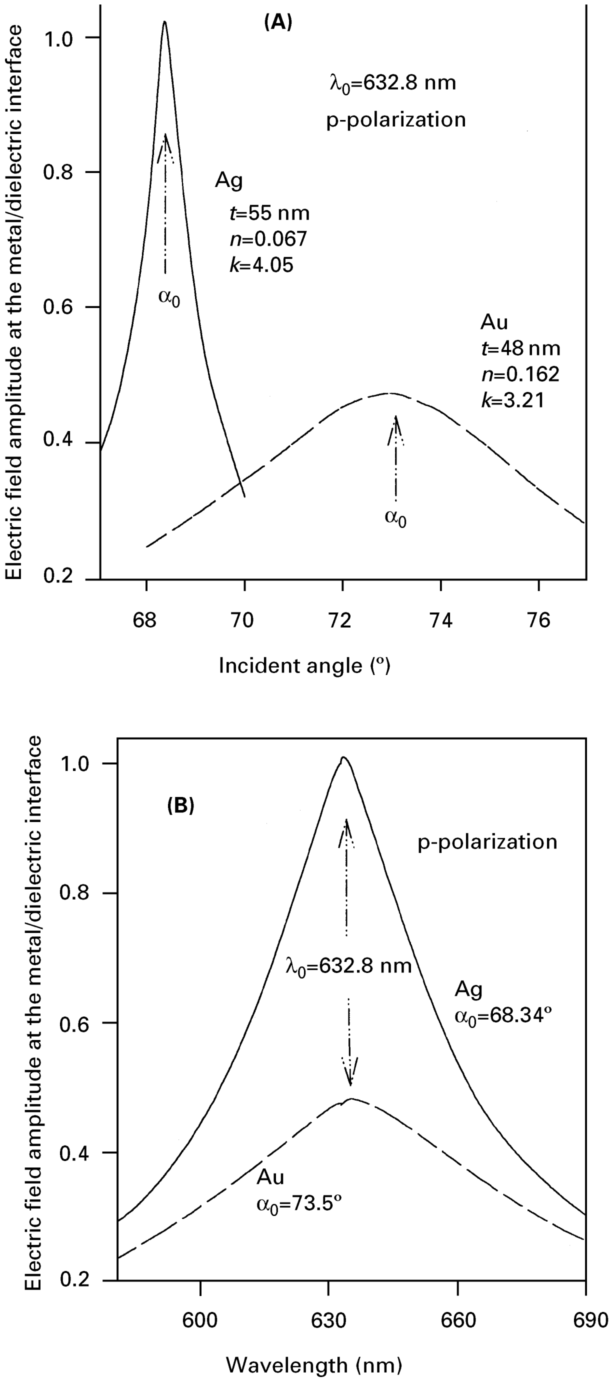



**Figure 1** Resonance spectra of the two most frequently used metal films, i.e. silver (solid line) and gold (dashed line), with the indicated optical parameters, presented as the total evanescent electric field amplitude (normalized at its largest value) generated by surface plasmons on the outer surface of the metal film, versus either the incident angle ($\alpha$, panel A) obtained with p-polarized (transverse magnetic) light of constant wavelength ($\lambda$ = 632.8 nm), or light wavelength (panel B) obtained at $\alpha = \alpha_0$ = constant. $\alpha_0$ is the incident angle at which the resonance excited with light of $\lambda = \lambda_0$ = constant reaches its maximum. The calculation has been done assuming a glass prism as an incident medium ($n_g$ = 1.5151), and water as an emerging medium ($n_w$ = 1.33), both at $\lambda$ = 632.8 nm.

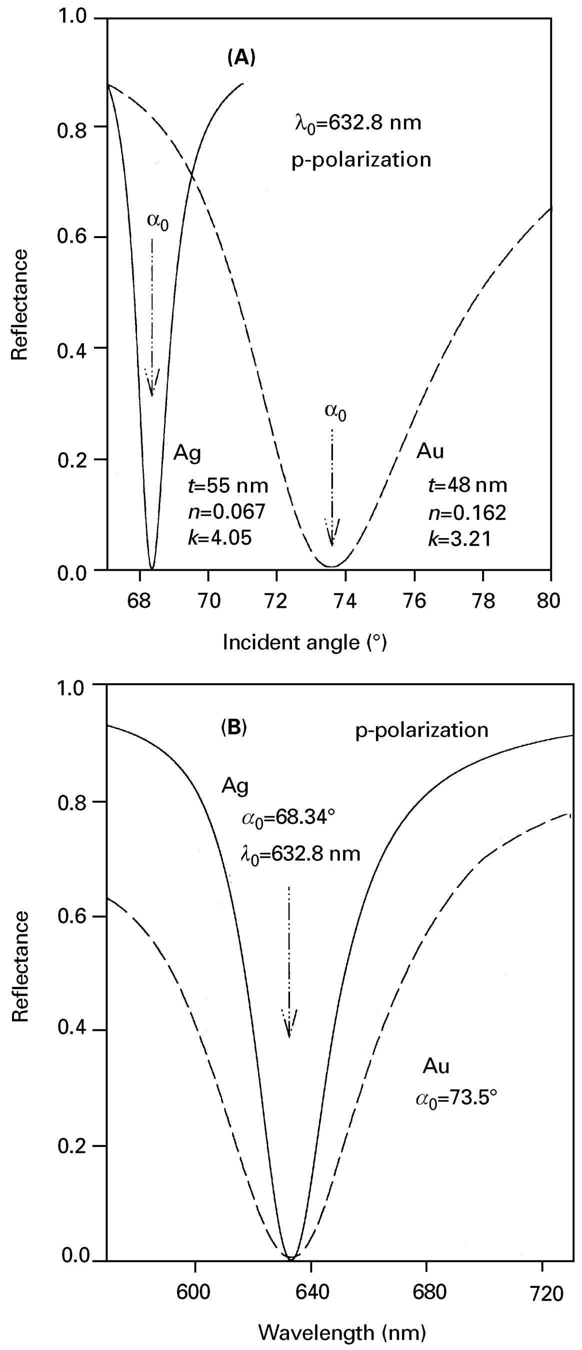



**Figure 2** Reflectance SPR spectra, i.e. reflectance versus either the incident angle ($\alpha$), with a constant value of the wavelength ($\lambda_0$) of the surface plasmon excitation light (panel A), or the wavelength, $\lambda$, at a constant value of $\alpha_0$ (panel B), obtained with silver and gold films. Other experimental conditions and symbols as in **Figure 1**.

and outer surfaces of the metal layer; and (2) it is characterized by a much higher evanescent electric field at the outer metal surface, which results in a sharper resonance curve, as shown in **Figure 4A**.

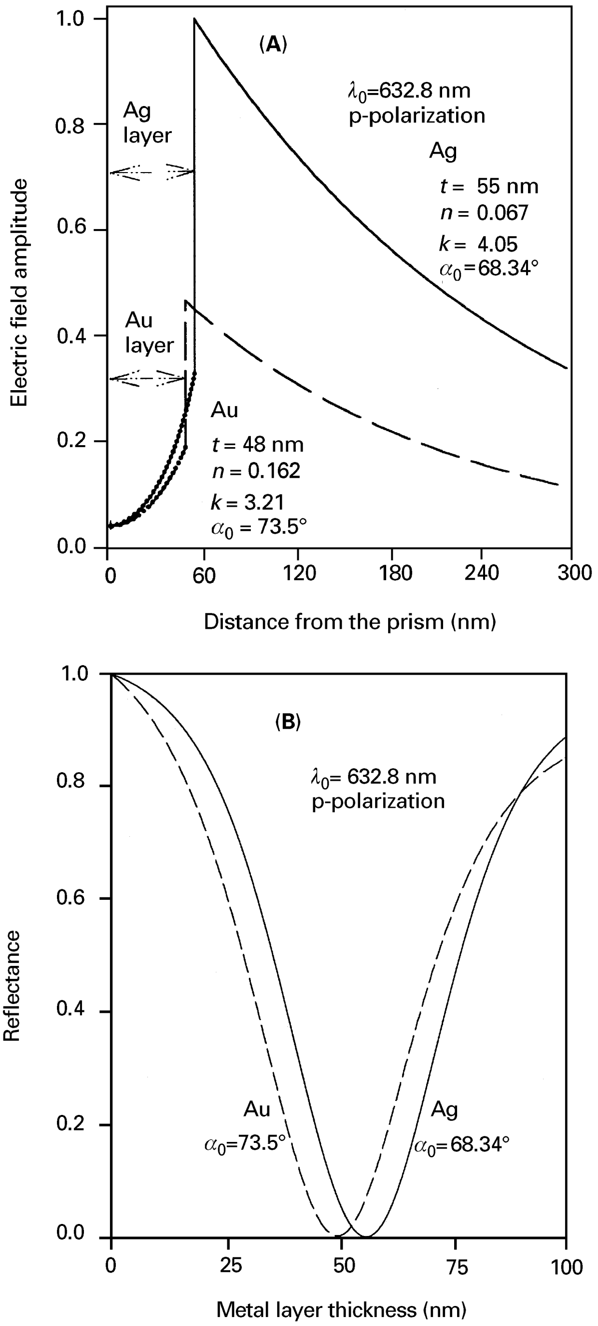



**Figure 3** (A) Calculated amplitude of the evanescent electric field (normalized at its largest value) generated within a metallic film (shown by closed circles), obtained with both silver (solid line) and gold (dashed line) at constant values of $\alpha_0$, as a function of the distance from the glass prism/metal interface. (B) Influence of metal film thickness on the SPR spectra obtained with an excitation wavelength $\lambda_0$ = 632.8 nm for silver (solid line) and gold (dashed line) films. Other conditions as in **Figure 1**.

### Coupled plasmon-waveguide resonance (CPWR)

The third type of surface resonance involves even more complex assemblies in which surface plasmon resonances in a thin metal film are coupled with guided waves in a dielectric overcoating, resulting in excitation of both plasmon and waveguide resonances (CPWR). A coupled plasmon-waveguide resonator contains a metallic layer (the same as in a conventional SPR assembly), which is deposited on either a prism or a grating and is overcoated with either a single dielectric layer or a system of dielectric layers, characterized by appropriate optical parameters so that the assembly is able to generate surface resonances upon excitation by both p- and s-polarized light components (**Figure 5A**). The addition of such a dielectric layer (or layers) to a conventional SPR assembly plays several important roles. First, it functions as an optical amplifier which significantly increases electromagnetic field intensities at the dielectric surface in comparison to conventional SPR, as illustrated by **Figures 5B** and **5C**. This results in an increased sensitivity and spectral resolution (the latter due to decreased resonance linewidths, as shown in **Figure 5A**). Secondly, it enhances spectroscopic capabilities (due to excitation of resonances with both p- and s-polarized light components), which results in the ability to directly measure anisotropies in refractive index and optical absorption coefficient in a thin film adsorbed onto the surface of the overcoating. Thirdly, the dielectric overcoating also serves as a mechanical and chemical shield for the thin metal layer in practical applications.

### Coupled long-range plasmon-waveguide resonance (CLRPWR)

This type of surface resonance can be obtained with a resonator which combines both the long-range surface plasmon and coupled plasmon-waveguide resonators into one device. The resulting resonance spectra are similar in shape and intensity to those obtained with CPWR devices.

## Detection of surface plasmon resonances

As discussed above, SP excitation generates surface electromagnetic modes bound to and propagating along the interface between a metal and a dielectric medium. Although these modes differ considerably from plane electromagnetic waves by having a pronounced dispersion (i.e. energy and momentum are not linearly connected by the speed of light), they demonstrate all other properties common to plane waves such as diffraction and interference. Therefore, the SP modes can, in principle, be detected by the same techniques as for plane electromagnetic

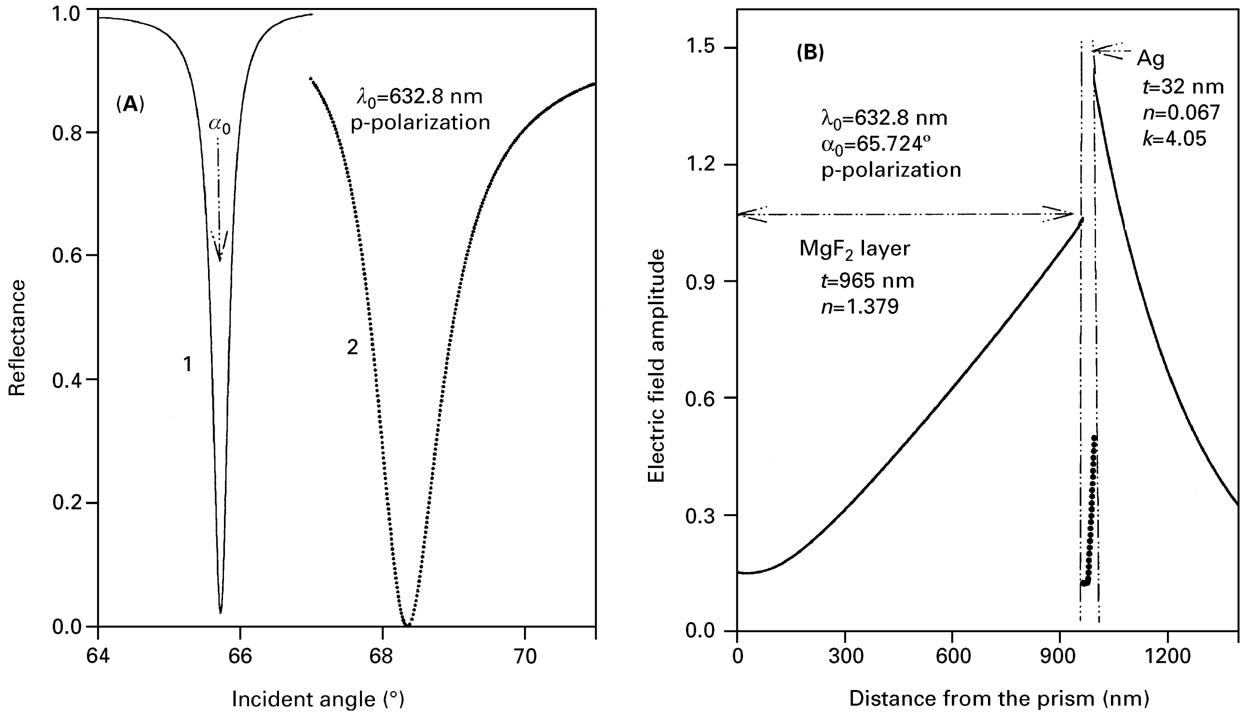



**Figure 4** (A) Calculated long-range SPR spectrum represented by reflectance versus incident angle (curve 1), obtained with p-polarized excitation light ($\lambda_0$ = 632.8 nm) and a glass prism coated with a $MgF_2$ layer (thickness = 965 nm, and refractive index = 1.379) on top of which a 32 nm thick silver film has been deposited. $\alpha_0$ indicates the incident angle at which resonance achieves its maximum. The emerging medium is water with $n_w$ = 1.33. Curve 2 illustrates the conventional SPR spectrum obtained with a 55 nm thick silver layer deposited directly on a glass prism (see **Figure 2A**). (B) Amplitude of the evanescent electric field (normalized to the largest value of the conventional SPR electric field as presented in **Figure 3A**) along the normal to the layer plane, calculated for the thin film design described in panel A and using $\lambda_0$ = 632.8 nm as an excitation wavelength at $\alpha = \alpha_0$ (see panel A).

waves, with two additional constraints, namely, that these are surface bound and are primarily nonradiative modes. Taking these special properties into account, together with the requirements which have to be fulfilled in order to excite SP modes, the most direct means of detecting these modes is by analysing the totally reflected light used to excite SPR.

The reflected light can be examined by analysis of either the transformation of light polarization using ellipsometry techniques, or the alteration of light intensity by applying reflectometry methods. The state of polarization is characterized by the phase, $\delta$, and amplitude, $E_0$, of light polarized parallel (p-polarization) and normal (s-polarization) to the plane of incidence. The difference in polarization state between the incident and reflected light (denoted i and r, respectively) is described by the parameters $\psi$ and $\Delta$, which are defined as follows:

$$\tan\psi = (E^{\mathrm{r}}_{0\mathrm{p}}/E^{\mathrm{r}}_{0\mathrm{s}})/(E^{\mathrm{i}}_{0\mathrm{p}}/E^{\mathrm{i}}_{0\mathrm{s}}) \qquad [8]$$

$$\Delta = (\delta^{\mathrm{r}}_{\mathrm{p}} - \delta^{\mathrm{r}}_{\mathrm{s}}) - (\delta^{\mathrm{i}}_{\mathrm{p}} - \delta^{\mathrm{i}}_{\mathrm{s}}) \qquad [9]$$

where tan $\psi$ and $\Delta$ are the changes in the amplitude ratio and the phase difference, respectively, on reflection. The variables $\psi$ and $\Delta$ often referred to as ellipsometrical angles, are related to the ratio between the overall complex reflection (Fresnel) coefficients of the respective light components, $r_p$, and $r_s$, by the following equation:

$$r_{\mathrm{p}}/r_{\mathrm{s}} = \tan\psi e^{\mathrm{i}\Delta} \qquad [10]$$

The Fresnel coefficients of the interface, $r_p$ and $r_s$, can, with the aid of Maxwell's theory (as shown in the preceding section) be expressed as functions of the wavelength of the light, $\lambda$, the incident angle, $\alpha$, and the optical properties of the reflecting system.

The reflectometry technique, which is based on measurement of reflected light intensity under ATR conditions, is experimentally a much simpler method than ellipsometry. It is therefore used much more frequently, especially in various sensor applications. As shown above, in this methodology, at a particular

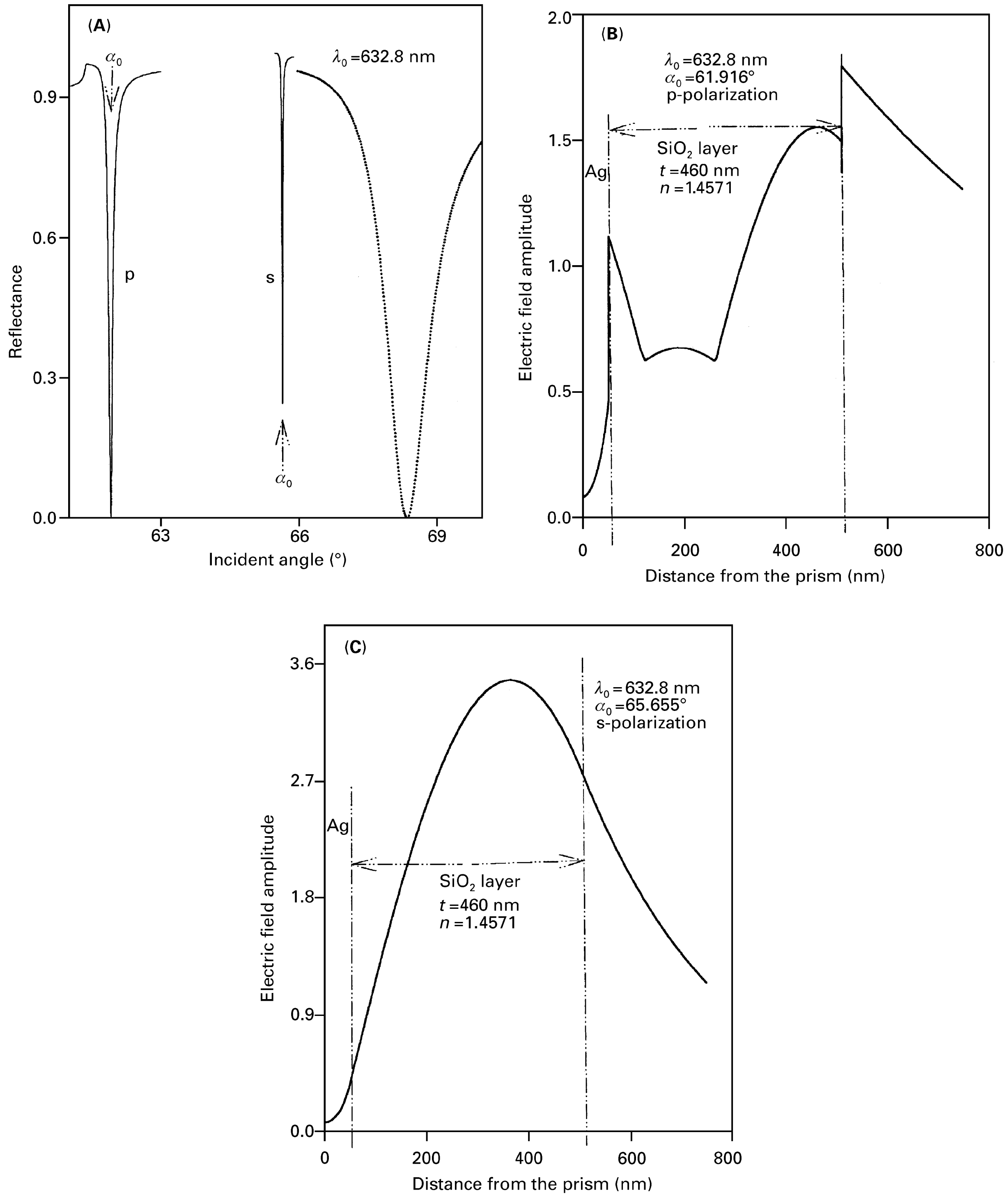



**Figure 5** (A) Coupled plasmon-waveguide resonance (CPWR) spectra presented as reflected light intensity versus incident angle, calculated with $\lambda_0$ = 632.8 nm as an excitation wavelength for both p- and s-polarizations, and a glass prism ($n_g$ = 1.5151) coated with a 55 nm thick silver layer overcoated with a 460 nm $SiO_2$ film. The emerging medium is water ($n_w$ = 1.33). The curve plotted with solid points illustrates the conventional SPR spectrum obtained with a 55 nm thick silver layer (see **Figure 2A**). $\alpha_0$ indicates the incident angle at which the resonance achieves its maximum. (B) and (C) Amplitudes of the evanescent electric fields, as a function of the distance from the glass prism/metal interface, within a silver layer, an $SiO_2$ film, and the emerging medium for p- (panel B) and s- (panel C) polarized light of $\lambda_0$ = 632.8 nm, calculated at $\alpha = \alpha_0$ (see panel A), and normalized to the largest value of the conventional SPR electric field as shown in **Figure 3A**.

incident angle the light wave vector matches the wave vector of the plasmon, fulfilling resonance conditions for plasmon generation (see Eqn [4]). During the resonance interaction, energy is transferred from photons to plasmons, so that the effect of plasmon excitation can be observed as a sharp minimum of the reflectance when either the angle of light incidence is varied at the same light wavelength, or the light wavelength is varied at the same incident angle, thus defining an SPR spectrum. In both instances, this spectrum reflects the resonance in absorption of incident photons.

SPR can also be detected with a fluorescence technique known as total internal reflectance fluorescence (TIRF) employed with waveguided systems. The application of TIRF to monitor SPR is based on the following properties of the surface modes. First, there is a possibility that the nonradiative SP modes can, under specific conditions, decay into light, therefore allowing emission techniques to be utilized to detect the resonance. As noted above, nonradiative surface plasmons can decay into photons of the same frequency if coupling by the surface roughness takes place. The intensity of emitted light images the SPR occurring at silver and gold surfaces producing an emission resonance curve similar to that of the reflectance resonance curve obtained under ATR conditions. Secondly, the surface bound electric field generated by SP modes, which can be much higher than that of the plasmon excitation light (see **Figures 1, 3A, 4B, 5B and 5C**), can be used as an efficient source to excite fluorescence emission of molecules adsorbed at the SPR active surface. This property of the surface plasmon electromagnetic field allows the monitoring of resonance by using fluorescent labelling of molecules adsorbed (or immobilized) on the active surface of the SPR producing medium. In both cases the measurement of fluorescence emission intensity as a function of either incident angle, $\alpha$, with excitation wavelength, $\lambda = \lambda_0 =$ constant, or $\lambda$ with $\alpha = \alpha_0 =$ constant, will generate an excitation emission curve. Such an excitation emission resonance curve will reflect either the SPR absorption resonance spectrum, as usually measured with ATR (see above), or the combination of the absorption spectrum of a fluorescent label and the SPR phenomenon, for measurement of emission intensity versus $\lambda$, at $\alpha = \alpha_0 =$ constant.

## List of symbols

$B$ = tangential electric field amplitude; $c$ = velocity of light in vacuo; $C$ = tangential magnetic field amplitude; $E_0$ = amplitude of light; $k$ = extinction coefficient; $K$ = wave vector; $K_{ph}$ = wave vector of light component parallel to interface; $K_{SP}$ = evanescent wave vector; $\mathbf{M}_j$ = matrix of $j$th layer; $n$ = refractive index; $N$ = complex index of refraction; $r$ = Fresnel reflection coefficient; $r^*$ = complex conjugate of $r$; $R$ = reflectance; $s$ = number of layers; tan $\psi$ = change in amplitude ratio; $t_j$ = thickness of $j$th layer; $v$ = velocity of light in dielectric medium; $y_i$ = oblique admittance; $Y_0$ = admittance of incident medium; $\alpha$ = incident angle; $\alpha_c$ = critical angle for total internal reflection; $\delta$ = phase thickness; $\Delta$ = change in phase difference; $\varepsilon'$ = real part of dielectric constant; $\varepsilon''$ = imaginary part of dielectric constant; $\lambda$ = wavelength; $\omega$ = frequency of surface plasmon wave.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; Ellipsometry; Fluorescent Molecular Probes; Inorganic Condensed Matter, Applications of Luminescence Spectroscopy; Organic Chemistry Applications of Fluorescence Spectroscopy; Surface Plasmon Resonance, Applications; Surface Plasmon Resonance, Instrumentation.**

## Further reading

Garland PB (1996) Optical evanescent wave methods for the study of biomolecular interactions. *Quarterly Reviews of Biophysics* **29**: 91–117.

Harrick NJ (1967) *Internal Reflection Spectroscopy*. New York: Interscience.

Kovacs G (1982) Optical excitation of surface plasmon-polaritons in layered media. In: Broadman AD (ed) *Electromagnetic Surface Modes,* pp 143–200. New York: Wiley.

Macleod AH (1986) *Thin Film Optical Filters*. Bristol: Adam Hilger.

Raether H (1977) Surface plasma oscillations and their applications. In: Hass G, Francombe M and Hoffman R (eds) *Physics of Thin Films*, Vol. 9, pp 145–261. New York: Academic Press.

Salamon Z, Brown MF and Tollin G (1999) Plasmon resonance spectroscopy: probing molecular interactions within membranes. *Trends in Biochemical Sciences* **24**: 213–219.

Salamon Z, Macleod AH and Tollin G (1997) Coupled plasmon-waveguide resonators: A new spectroscopic tool for probing proteolipid film structure and properties. *Biophysical Journal* **73**: 2791–2797.

Salamon Z, Macleod AH and Tollin G (1997) Surface plasmon resonance spectroscopy as a tool for investigating the biochemical and biophysical properties of membrane protein systems. I: Theoretical principles. *Biochimica et Biophysica Acta* **1331**: 117– 129.

Wedford K (1991) Surface plasmon-polaritons and their uses. *Optical and Quantum Electronics* **23**: 1– 27.

# Surface Studies By IR Spectroscopy

**Norman Sheppard**, University of East Anglia, Norwich, UK

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES

Applications

The investigation of surfaces, and of molecular layers adsorbed on surfaces, by electromagnetic radiation has been carried out principally by infrared spectroscopy. This is because of the high sensitivity of present-day Fourier transform infrared (FT-IR) spectrometers; the capability for IR spectroscopy to obtain data from mixed-phase samples with gas/solid, liquid/solid, or gas/liquid interfaces; and because of the availability of very large databases relating the positions and relative strengths of infrared absorptions to structural features of organic and inorganic molecules. As described below, the sampling techniques used differ substantially whether the systems under investigation involve finely divided samples (powders or porous solids) or whether the surface involved is flat. In the early history of the subject, spectral features relating to adsorbed layers and other surface phenomena could only be detected if very high area finely divided samples were used so that the radiation beam could pass through many interfaces. However, since the advent of FT-IR spectrometers, infrared sensitivity has so much improved that nowadays a measurable spectrum can be produced from even a single monolayer on a flat surface.

After reviewing the experimental techniques involved, we survey the principal applications of the infrared method under the headings surface characterization, physical adsorption, and chemisorption and catalysis.

## Experimental techniques

### High-area, finely divided, surfaces

Surfaces, because of their unsaturated surface fields, normally require to be cleaned from contamination derived from the atmospheric environment before systematic research can be carried out on them. For finely divided samples of high area it is adequate to mount them in a high-vacuum enclosure (~$10^{-6}$ mbar) provided with infrared-transparent windows and the means for treating the sample in oxygen or hydrogen at elevated temperatures. The samples themselves are most often studied in transmission, usually in the form of pressed discs derived from powders. These are prepared using a hydraulic press in a manner similar to that used for the standard potassium bromide pressed-disc sampling procedure for IR spectroscopy. The pressure required for coherent disc formation is greater for the commonly studied oxide layers than for the softer potassium bromide, but the discs so prepared remain porous for adsorption studies. Alternatively, a powdered sample can often be made to cohere on an infrared-transparent disc, or on a fine metal grid, through sublimation or by deposition from a solvent.

Finely divided metal samples require that the opaque particles are separated from each other for transmission purposes. Usually this is done by distributing (supporting) them on the surface of an oxide which is transparent over relevant regions of the spectrum. Such samples are prepared by depositing metal salts from solution on the oxide particles followed by evaporation of the solvent; a disc is then pressed from the mixed powder and inserted in the IR vacuum cell; finally, the salt is reduced in hydrogen at appropriate temperatures so as to form metal particles distributed over the surface of the oxide support. Very high area powders of silica and alumina, of areas between 200 and 300 $m^2\ g^{-1}$, are commercially available and are frequently used as metal supports. They have the advantages that they are largely infrared-transparent down to ~1300 or ~1100 $cm^{-1}$ respectively, and hence permit the study of many group-characteristic absorptions from organic adsorbates. Silica is a more catalytically inert support than alumina.

Samples prepared as described above can be good models for working catalysts of either the oxide or metal types and many infrared studies of surface phenomena are undertaken in conjunction with catalytic investigations. Loose powders can alternatively be studied by diffuse reflection, with the advantage for kinetic studies that surface reactions, rather than diffusion processes, are more likely to be rate determining than is the case with the fine-pored pressed discs.

### Low-area flat surfaces

Adsorbents in the form of flat surfaces are of very low area and normally ultra-high vacuum (UHV) conditions (~$10^{-10}$ mbar) are required in order to

preserve them from contamination. Where the substrate is transparent, infrared spectra of the surface layers can be obtained either by transmission or by reflection; when the substrate is non-transmitting, as in the case of metals, then reflection is normally used. Experiments involving flat surfaces allow the application of polarized radiation for the purpose of obtaining information about the orientation of the adsorbed molecules with respect to the surface. For transparent substrates the used of radiation polarized in or perpendicular to the plane of incidence, in combination with the measured angle of incidence, can determine the direction of the dipole change with respect to the surface associated with each group-characteristic vibration. The orientations of even flexible molecules with respect to the surface can be deduced from such measurements.

In the case of metals the effect of the free response of the conduction-electrons to a charge above the surface can be modelled in terms of an image of opposite sign at the same distance below the surface as is shown in **Figure 1**. In the infrared context it is the dipole moment change associated with the vibration that interacts with the radiation. **Figure 1** shows that a component of such a dipole change that is parallel to the surface is cancelled out by its image, whereas a component perpendicular to the surface is doubled in magnitude. Hence only modes with perpendicular dipole components are IR allowed; in general these are those modes of vibration that are symmetrical with respect to all the symmetry elements associated with the surface complex. For example, a CO molecule adsorbed perpendicular with respect to the surface will give absorption bands from the $\nu$CO or $\nu$CM (M = metal) bond-stretching modes but not from the OCM bending modes. Such considerations constitute the metal surface selection rule (MSSR), which is widely used for the determination of molecular orientation or, if this is known, as an aid in the assignment of vibrational modes. For work with metals, reflection–absorption infrared spectroscopy(RAIRS) uses near-grazing incidence in order to maximize the strength of the electric vector of the incident infrared radiation that is perpendicular to the surface.

UHV is normally required when studying low-area flat surfaces (exceptionally this would not be a requirement if the adsorbate, such as a surfactant, is capable of displacing surface impurities) and this requires sophisticated equipment. Also, the high sensitivity needed for the measurement of spectra from single monolayers requires the use of FT-IR spectrometers with selective photoconductive infrared detectors; the mercury/cadmium telluride detector which covers the major range of the spectrum down

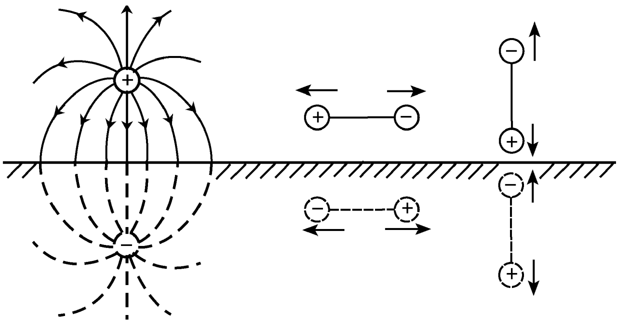


**Figure 1** Charges and their images near metal surfaces; the origin of the metal surface selection rule (MSSR).

to ~700 cm$^{-1}$ is widely used. **Figure 2** illustrates a typical experimental arrangement for RAIRS on a metal surface under UHV conditions. Spectroscopic work carried out on single crystals with known types of adsorption sites, such as are readily available for metals, are of great use in interpreting the more complex spectroscopic phenomena obtained from finely divided samples. Individual particles of the latter can exhibit facets with a variety of atomic arrangements and adsorption sites which can be studied one-at-a-time on single crystals. **Figure 3** shows the different atomic arrangements, and hence adsorption sites, of the (111), (100) and (110) faces of a face-centred cubic metallic lattice. UHV facilities also permit complementary spectroscopic methods involving particles such as electrons (as in high-resolution electron energy loss spectroscopy) or diffraction methods (as in low-energy electron diffraction) to be employed in order to characterize the same system further.

Adsorption on metal electrodes, which can be cleaned in solution by electrode reactions, is also studied by RAIRS. There is added interest in the effects of the variable electrode potential on the spectra and structures of the adsorbed species.

The surfaces of infrared-transparent materials that are available in the form of shaped and polished crystals, such as silicon or germanium, can be studied with good sensitivity by using attenuated total internal reflection (ATR) in conjunction with multiple reflection procedures.

Sum frequency generation (SFG) is a recent spectroscopic development in which two laser beams, one in the visible region and the other of variable frequency in the infrared region, generate infrared-modulated signals in the visible region at the sum of the two frequencies. As the signals come only from the interface and not from the bulk, this technique is being exploited in high-pressure catalyst work and for surfactant research.

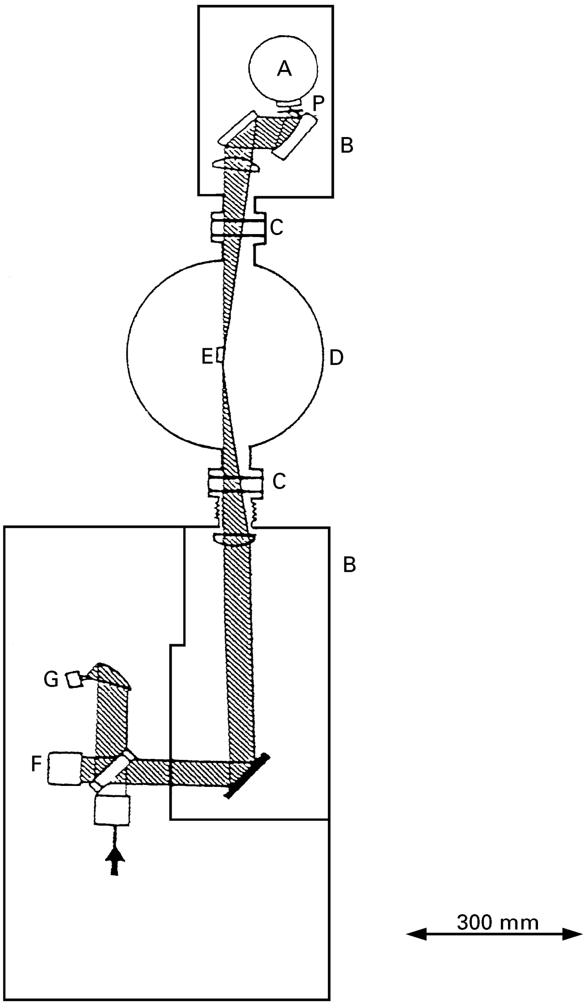



**Figure 2** The optical arrangement of an FT-IR spectrometer for reflection–absorption (RAIRS) work in ultrahigh vacuum (UHV). A, detector; B, KBr lens: C, KBr window; D, UHV chamber; E, sample; F, Michelson interferometer; G, Globar source; P, grid polarizer. Reprinted from Chesters MA (1986) *Journal of Electron Spectroscopy and Related Phenomena* **38**:123 Copyright (1986), with permission from Elsevier Science.

## Surface characterization

The long-range patterns of surface atomic arrangements are principally monitored by low-energy electron diffraction (LEED). Whereas in principle the top layers of a lattice have different frequencies and hence wavenumbers from those of the bulk lattice, the associated absorptions often fall within a spectral region dominated by the latter and are hence difficult to identify. Transition metal oxides are exceptional in that the variable valency associated with the metallic element can lead to the generation of surface M = O groups (M = metal) that give absorptions of notably higher wavenumber than those of the lattice modes.

Infrared contributions to our knowledge of surfaces are mostly short-range in type and involve the identification of different types of site through the adsorption of probe molecules chosen for this purpose. CO is a well known probe of surface sites on metal surfaces. As discussed in the Chemisorption section below, its $\nu$CO bond-stretching vibration has distinct wavenumber ranges for adsorption on linear (on-top), twofold and threefold bridging sites. Although linear and twofold sites can occur on each of the surfaces shown in **Figure 3**, the threefold one is specifically characteristic of (111) surfaces and can be used to identify such facets on metal particles. Distinctions can sometimes be made between twofold CO sites on different facets. The wavenumbers of CO absorptions can also be used to characterize surface cation sites of different charge (different formal oxidation states) on transition metal oxides as shown in **Figure 4** for a partially reduced Ni oxide sample. For hydrocarbon adsorption the characteristic spectrum of ethylidyne ($CH_3C$) also plays a useful role in identifying (111) facets on finely divided metals.

One of the earliest discoveries of surface infrared spectroscopy was that oxide surfaces, such as those of $SiO_2$ or $Al_2O_3$, retain chemisorbed OH groups after the removal of water molecules adsorbed from the atmosphere. These can only be removed by high-temperature treatment and are presumably generated by the reaction of ambient water molecules with otherwise free valencies on the surface of the oxide lattice, according to the reaction $O^{2-} + H_2O \rightarrow 2OH^-$ or its covalent equivalent. In the case of alumina, for example, individual absorptions amongst a multiplicity of OH bond-stretching absorptions can be identified with linear, two- and threefold adsorption sites, for each of two types of surface aluminium atoms which in the bulk lattice have four- or six-fold coordinations, i.e. are in formal IV or VI oxidation states. Silica has only four silicon coordination and correspondingly simpler $\nu$OH spectrum consisting

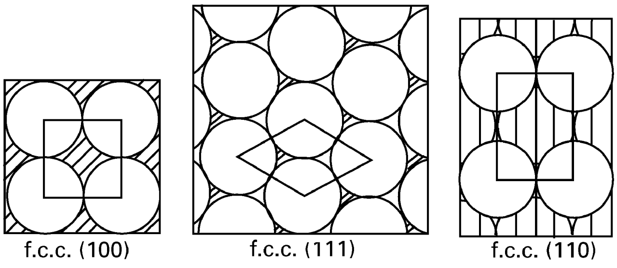



**Figure 3** The arrangements of atoms, and the resulting adsorption sites, on the (100), (111) and (110) surfaces of a face-centred-cubic metal.

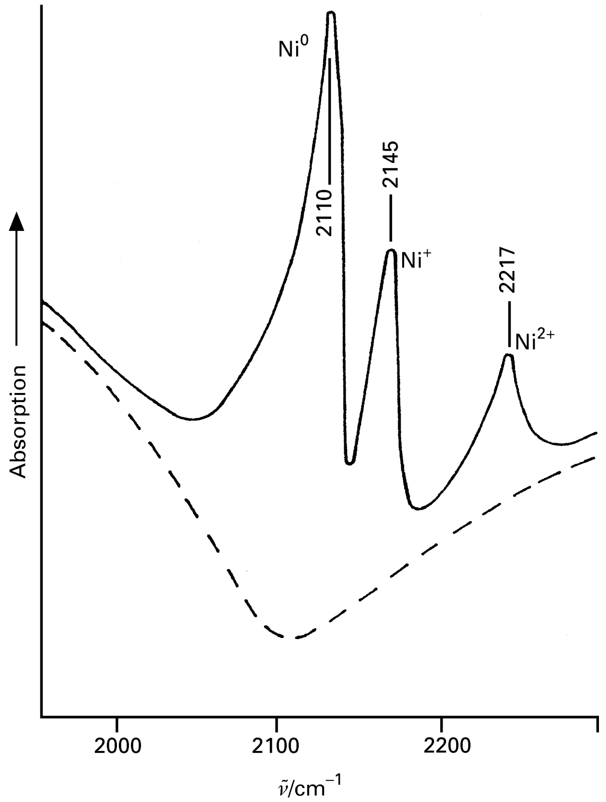



**Figure 4** The IR spectrum of CO adsorbed on partially reduced NiY zeolite showing the oxidation-state dependence of $\nu$CO. Reproduced with permission of John Wiley & Sons Ltd. from Davydov AA (1984) *Infrared Spectroscopy of Adsorbed Species on the Surfaces of Transition Metal Oxides*, Copyright John Wiley & Sons Ltd.

mainly of absorptions of pairs of free and hydrogen-bonded OH groups on adjacent sites.

An acidic oxide such as alumina exhibits relatively inert OH groups, strongly acidic OH groups that are capable of proton donation (Brønsted acidity), plus aluminium ions that act as electron-deficient sites (Lewis acidity). The relative proportions of these on such oxide surfaces are analysed using the infrared spectrum of adsorbed pyridine. The spectra are measured in the 1700–1500 cm$^{-1}$ region. Pyridine hydrogen-bonded to the weaker surface OHs gives a weakly perturbed spectrum; that interacting with the strongly acidic OHs forms the pyridinium ion by proton transfer which has a characteristic additional absorption at 1540 cm$^{-1}$; that interacting with metal-coordination sites shows a change in wavenumber of a skeletal absorption near 1600 cm$^{-1}$.

Basic oxides frequently adsorb carbon dioxide from the atmosphere to give surface carbonates according to the reaction $O^{2-} + CO_2 \rightarrow CO_3^{2-}$, a process readily monitored by infrared spectroscopy. With the well known exceptions of gold and platinum, metal surfaces adsorb oxygen from the atmosphere leading, in the cases of aluminium and iron, to the production of multilayer 'passivating' oxide films or (with the participation of absorbed water) thick films of rust, respectively. Infrared spectroscopy can monitor these surface corrosion processes.

## Physical adsorption

In this section we consider the spectroscopic study of the association of molecules with surfaces by intermolecular forces ranging from van der Waals to strong hydrogen bonding.

**Figure 5** shows the absorption bands in the $\nu$CH bond-stretching region from methane adsorbed on porous silica glass. In the gas phase the triply degenerate $\nu$CH mode at 3019 cm$^{-1}$ is infrared active and is to be identified with the strong band from the adsorbed species at ~3006 cm$^{-1}$. On the surface an additional feature has appeared in the spectrum at 2899 cm$^{-1}$ which is readily identified as the gas-phase forbidden $\nu$CH 'breathing' mode, known from gas-phase Raman spectroscopy to occur at 2917 cm$^{-1}$. The one-sided surface forces have distorted the original tetrahedral shape of the methane molecule so as to cause this mode to become active. The considerable breath of the ~3006 cm$^{-1}$ absorption of the surface species, notably less than that of the gas-phase vibration–rotation band, was interpreted in terms of quasi-free rotation of the molecule about a single axis perpendicular to the silica surface.

The spectrum of **Figure 6** shows the interaction of the very acidic surface OH group on zeolite HY with adsorbed ethene. The low wavenumber, broad profile, and intensification of the shifted $\nu$OH absorption upon ethene adsorption indicate a hydrogen bond of considerable strength, comparable to that between water molecules, even although the bonding is only to the $\pi$-electrons of the adsorbed ethene. This complex

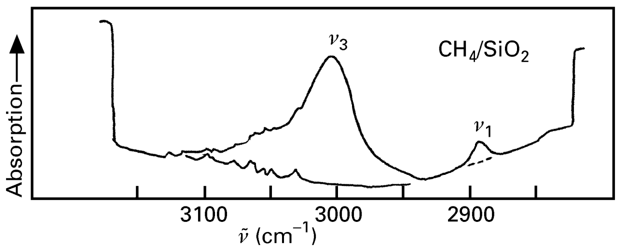



**Figure 5** The IR spectrum of $CH_4$ adsorbed on high-area porous silica glass in the $\nu$CH bond-stretching region showing the presence of a gas-phase forbidden absorption. Reproduced with permission of the Royal Society from Sheppard N and Yates DJC (1956) *Proceedings of the Royal Society of London, Series A* **238**: 69.

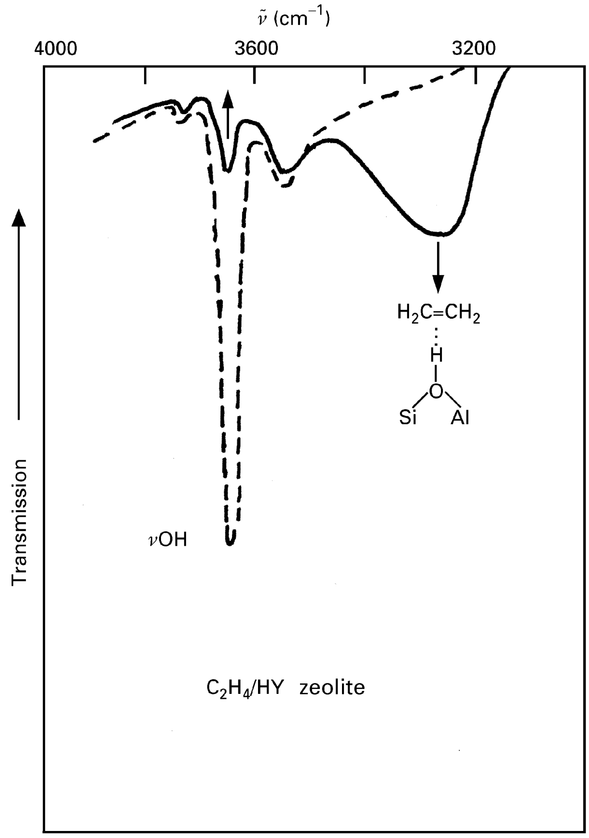



**Figure 6** The IR spectrum of ethene adsorbed on the acid OH groups of HY zeolite. Solid line, ethene adsorbed; dashed line, background. Reproduced with permission of the Royal Society of Chemistry from Liengme BV and Hall WK (1966) *Transactions of the Faraday Society* **62**: 3229.

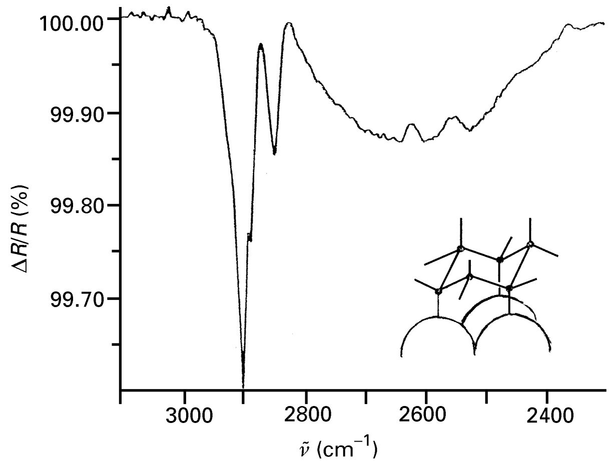



**Figure 7** The IR spectrum of cyclohexane adsorbed on a Pt(111) surface. The broad absorption near 2600 cm$^{-1}$ is from a form of hydrogen bonding between axial CH bonds and surface Pt atoms. Reprinted from Chesters MA and Gardener P (1990) *Spectrochimica Acta, Part A* **46**: 1011, Copyright (1990), with permission from Elsevier Science.

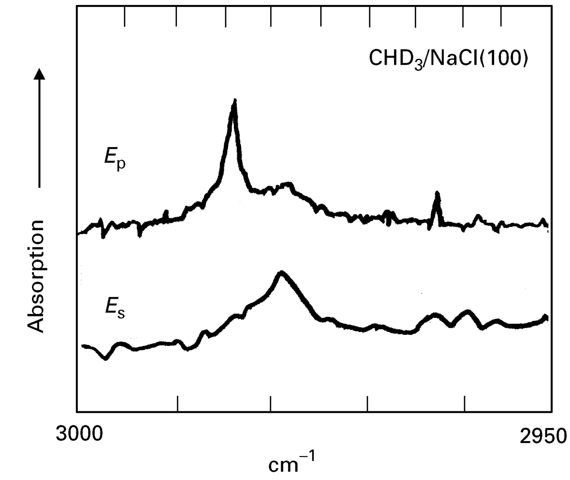



**Figure 8** The IR spectrum from $CHD_3$ adsorbed at 33 K on NaCl(100). $E_p$ and $E_s$ refer to radiation polarized in and perpendicular to the plane of incidence, respectively. Reprinted with permission from Davis KA and Ewing GE (1997) *Journal of Chemical Physics* **107**: 8073. Copyright 1997, American Institute of Physics.

is clearly an intermediate in the higher temperature formation of the carbenium ion $C_2H_5^+$.

Hydrogen bonds are normally considered to form between acidic XH groups and electron-rich bases. However, surface infrared spectroscopy, in conjunction with HREELS, has shown that such bonds can also occur between electron-rich CH bonds of paraffins and electron-deficient sites on metal surfaces. **Figure** 7 shows the spectrum of cyclohexane adsorbed on the (111) surface of platinum. The very broad band centred at ~2620 cm$^{-1}$ is from a proportion of CH bonds of the adsorbed cyclohexane in a hydrogen-bonded type of environment. As the separation between the three parallel axial CH bonds on one side of the cyclohexane molecule is almost exactly the separation between adjacent Pt atoms on a threefold site of the (111) surface, it is clear that the hydrogen bond is of the type CH···Pt.

**Figure** 8 is a spectrum taken at 33 K in the νCH region of $CD_3H$ adsorbed on the (100) face of the face-centred-cubic lattice of NaCl. It is seen that there are two well resolved absorption bands, one sharp and the other broad, resulting from CH bonds that are oriented differently with respect to the surface. One of these occurs with the incident light polarized in the plane of incidence but is eliminated when the light is polarized perpendicular to this; the other is present in both spectra. The former band hence has its νCH vibrational dipole change perpendicular to the surface, whereas the direction of the

latter has both parallel and perpendicular components. Considerations of relative intensities, taking into account the angle of incidence, show that the broader low-wavenumber band is from CH bonds that are at ~70° with respect to the surface. It is hence concluded that the parent $CH_4$ molecules are adsorbed with three of their four CH bonds on the surface.

Even in the absence of the special effects of hydrogen bonding, bandwidths of 10 $cm^{-1}$ or more are common from adsorbed species on polycrystalline substrates owing to interactions with sites that differ in their detailed environments. Absorptions obtained from adsorbed species on single-crystal planes with uniform and well defined sites can, in contrast, be very sharp with bandwidths of less than 1 $cm^{-1}$. In the case of methane itself, and of a number of other molecules such as CO and $CO_2$, the resolution of the spectra on alkali metal halide single-crystal surfaces are such that even the fine-structure splitting caused by the vibrational couplings of more than one molecule in the surface unit cell can readily be resolved.

## Chemisorption and catalysis

The quantitative and energetic aspects of the chemisorption of molecules on surfaces have long been investigated but, until in the 1950s it became possible to obtain infrared spectra, the actual structures of the surface species could only be a matter for speculation. The spectra show that in fact finely divided adsorbents give absorptions from several different surface species, and that the nature of the latter can vary as a function of coverage. Simpler spectra are obtained on single-crystal surfaces of known atomic arrangements. However, even so, the deductions of the structures of the chemisorbed species can be difficult because of uncertainties related to the effects of surface bonding on the spectra of the attached groups. The usual group-characteristic wavenumber ranges can no longer be assumed to be reliable because of the electron-donating or -withdrawing properties of the surface atoms and also, when there is multiple bonding to the surface, because of strains associated with cyclic bonding features. The procedure adopted is to use the spectra to suggest possible alternative structures for the adsorption complexes, and then to look for molecular analogues of known structures whose spectra can be obtained for comparative pattern-recognition purposes. This approach is well exemplified by the results obtained for chemisorption on metal surfaces, an area much studied because of the ready availability of single crystals of metals which can be cut so as to display particular surface planes.

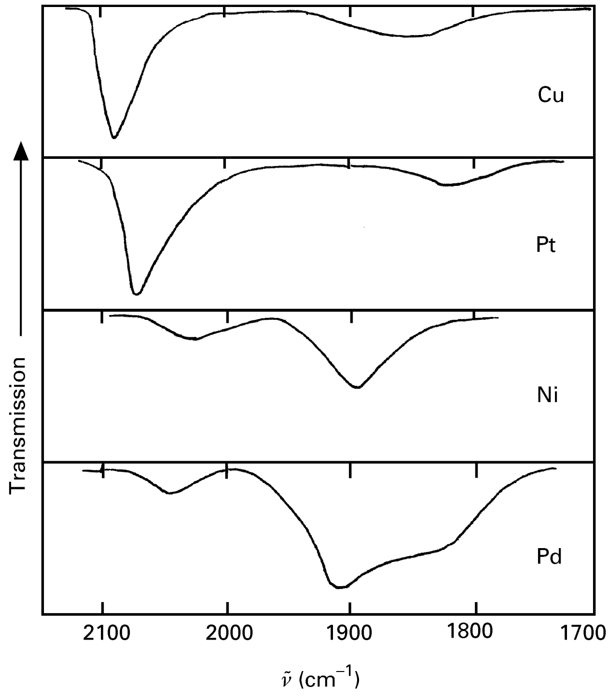



**Figure 9** The IR spectra of CO adsorbed on the silica-supported metals Cu, Pt, Ni and Pd. Absorptions above 2000 $cm^{-1}$ are from linear (on top) CO bonded to one metal atom; those below this value are from CO bridge-bonded to two or three metal atoms. Reprinted with permission from Eischens RP, Pliskin WA and Francis S A (1954) *Journal of Chemical Physics*, **22**: 1786. Copyright 1954 American Institute of Physics.

**Figure 9** shows high-coverage spectra obtained from CO chemisorbed on the silica-supported metals Cu, Pt, Ni and Pd. The several metals show different proportions of absorption bands above and below 2000 $cm^{-1}$ which are characteristic of adsorption on linear (on-top) and bridged sites, respectively. These structural assignments were deduced by comparison with the spectra of metal carbonyls. The spectral ranges attributable to such surface species are as follows: linear, 2120–2000 $cm^{-1}$; twofold bridge, 2000 – ~1870 $cm^{-1}$; and threefold bridge ~1900–1800 $cm^{-1}$. These ranges apply whichever crystal face is involved. Within each range the characteristic absorptions increase in wavenumbers with increasing coverage. This is caused by strong vibrational coupling within the array of parallel molecules on the surface, mostly of a dipolar nature related to the exceptional strength of the $\nu CO$ absorptions. The mixture of linear and bridged CO species found from the spectra from the finely-divided samples is caused by adsorption on different sites, usually different facets, on the metal particles.

**Figure 10** shows $\nu CO$ spectra at full coverage from chemisorption on an Rh(111) single-crystal electrode. These are plotted as a function of the

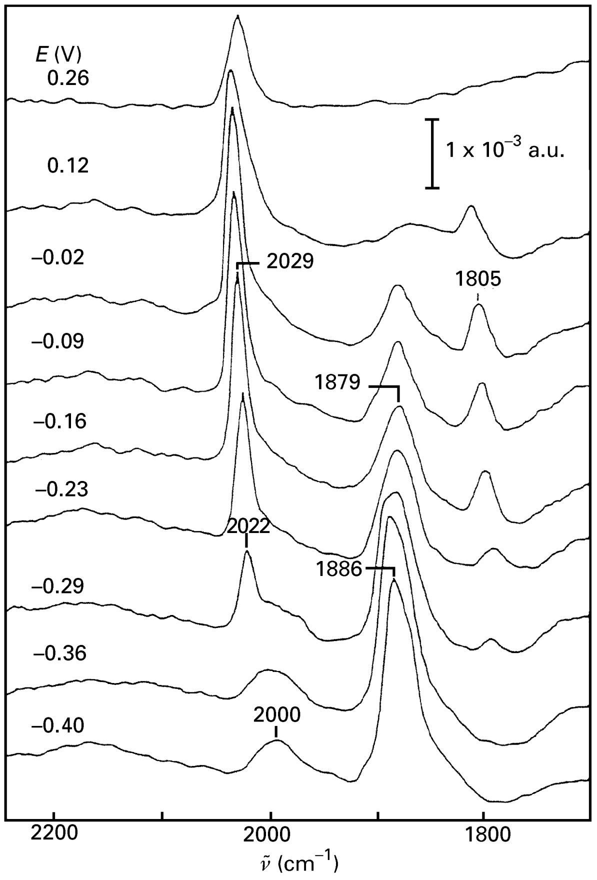



**Figure 10** The IR spectra of CO adsorbed at full coverage on an Rh(111) electrode in 0.1 M $NaClO_4$ at various electrode potentials. Reprinted from Chang SC and Weaver MJ (1990) *Surface Science* **238**: 142, Copyright (1990), with permission from Elsevier Science.

electrode potential (with respect to a standard Ag/AgCl electrode) in 0.1M $NaClO_4$ and show the interest of this additional variable in electrode work. It is seen that at the lowest electrode potential the spectrum is dominated by the absorption at 1886 $cm^{-1}$ from a bridged species but at higher potentials, before desorption sets in, a linear species becomes dominant, absorbing at 2029 $cm^{-1}$. More generally, work on electrodes shows that, at a given coverage, negative potentials favour bridged species over linear species and that the wavenumbers of $\nu$CO absorptions from linear species increase in value with increasingly positive electrode potentials – a milder version of the dependence of $\nu$CO on metal oxidation state reported above.

Hydrocarbons on metal surfaces provide greater challenges in spectral interpretation and we choose the example of ethene chemisorbed on different metal surfaces. Here the relevant model compounds are inorganic binuclear or trinuclear metal clusters with the hydrocarbon ligand of interest and additional CO ligands occupying the positions of the metal atoms of the surface complex. One of the unexpected aspects of the adsorption of ethene is that (111) faces of many metals are covered by the dissociative ethylidyne species $CH_3CM_3$ (M = metal) near room temperature. Its spectrum was attributed to this structure by comparison with the spectrum of the model compound $(CH_3C)Co_3(CO)_9$, considered as a possibility because electron diffraction had shown that the CC bond of the adsorbed species is perpendicular to the surface. This example shows the importance of the metal–surface selection rule (MSSR). For this species, as a ligand or as a surface complex, the modes of vibration are fully separable into those with dipole changes either perpendicular or parallel to the surface (parallel or perpendicular to the CC bond, respectively). Only the former modes are active under the MSSR but both sets are active in the infrared spectrum of the model compound. **Figure 11** compares the infrared spectrum of the ethylidyne species on the Pt(111) surface with that of the model compound. The bands marked with asterisks in the spectra of the model compound (in order of decreasing wavenumber, $\nu CH_3$ symmetrical stretch, $\delta XH_3$ symmetrical bend and $\nu$CC stretch) are those which give dipole changes perpendicular to the surface; the other doubly degenerate modes give dipole changes parallel to the surface ($\nu CH_3$ asymmetric stretch, $\delta CH_3$ asymmetric bend and $CH_3$ rock). The positions of the 'missing' modes of the surface species, indicated by arrows, have been identified in the HREEL spectrum of the same system where the selection rules are more relaxed.

It has been shown by spectroscopy that at low temperatures ethene adsorbs on Pt(111) as the

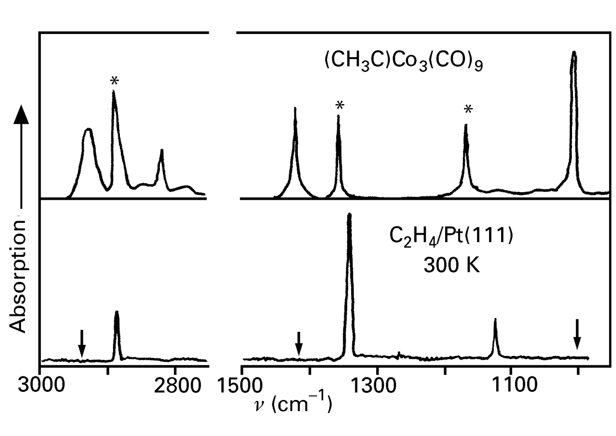



**Figure 11** A comparison of the IR spectrum from ethene adsorbed on Pt(111) at room temperature with that of the model compound $(CH_3C)Co_3(CO)_9$. Asterisks indicate absorptions of the model compound allowed in the spectrum of the adsorbed species by the metal-surface selection rule; arrows indicate other bands observed by HREELS. Courtesy Chesters MA.

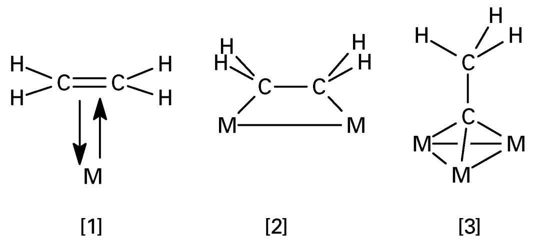



**Figure 12** The structures of the principal adsorbed species from the adsorption of ethene on metal surfaces; [1] $\pi$-complex; [2] the di-$\sigma$ species; [3] ethylidyne ($CH_3C$).

$MCH_2CH_2M$ (di-σ adsorbed) species with a cyclic $C_2M_2$ skeleton and that this transforms into ethylidyne on warming to near room temperature. The (111) faces of other metals, notably Pd and Cu, show low-temperature spectra from another less strongly perturbed $H_2CCH_2$ adsorbed species in which there is bonding from a single metal atom to the $\pi$-electron distribution of the C=C double bond. Its spectrum is closer to that of ethene itself but with those modes which involve CC stretching occurring at lower wavenumbers. The structures of these three species are shown in **Figure 12. Figure 13** shows two spectra from ethene adsorbed on a silica-supported Pt sample, of the type used in catalysis, at 180K and at room temperature. On these are indicated absorptions from the above three species with the MSSR-allowed modes still dominant. It is seen that the spectra from the catalyst sample are comprehensively accounted for in terms of the species that had been identified one-at-time on single-crystal surfaces.

For the purpose of catalysis, the structure of the surface-adsorbed reactant should be sufficiently perturbed in order to promote reactivity, but not so strongly adsorbed that it cannot be removed by reaction. Below the temperature for the onset of catalysis, controlled by the energy of activation, the reactive species will be one or more of the chemisorbed species. The spectra of such species will weaken or disappear when catalysis commences while less reactive species are retained. In the case of ethene hydrogenation over metal catalysts, the order of reactivity in the presence of hydrogen is [1] > [2] > [3], with [3], the ethylidyne species, being very slow to be removed. By room temperature over Pt/ $SiO_2$, when the di-$\sigma$ species has all been converted to ethylidyne, it is clear that the $\pi$-species, [1], is the catalytically active one. On Pt single crystals this mainly occurs on non-close-packed planes, and it may be inferred that catalytic reduction occurs on rougher, non-(111), surfaces of the metal particles. In a similar manner, it has been shown by single-crystal spectroscopy that the reactive species in the reduction of nitrogen to ammonia over the Fe catalyst (the Haber process) is a di-σ species involving the NN molecule chemisorbed to two Fe atoms which dissociates to adsorbed N atoms during catalysis.

The transition metal oxides form the other principal class of catalysts. These differ from the metals in that they have both acid and base sites in the same surface (the metal and oxygen atoms/ions, respectively) and react differently according to which of these properties is dominant. **Figure 14** shows the infrared spectrum from the heterolytic dissociation of hydrogen on polycrystalline ZnO to given surface

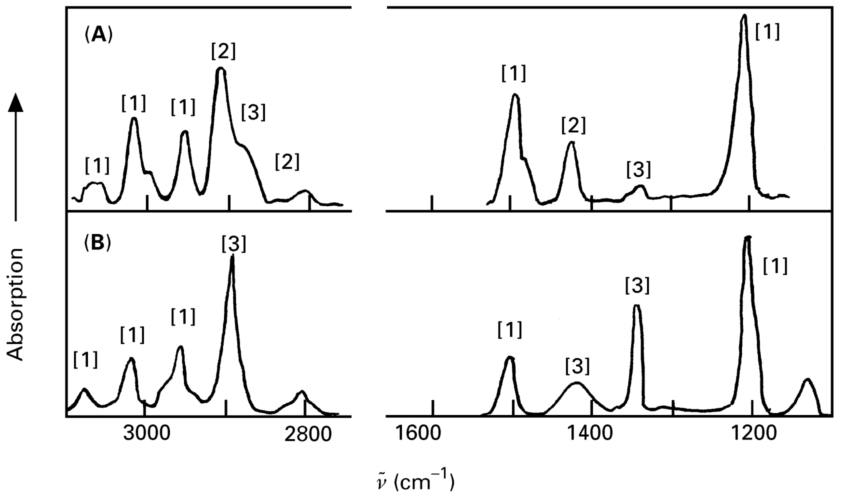



**Figure 13** The IR spectra of ethene adsorbed on silica-supported Pt (A) at 180 K and (B) at room temperature, labelled according to the structural assignments of the absorption bands. Reprinted with permission from Mohsin SB, Trenary M and Robota H *Journal of Physical Chemistry*, (1988) **92**: 5229 and (1991) **95**: 6657. Copyright 1988,1991, American Chemical Society.

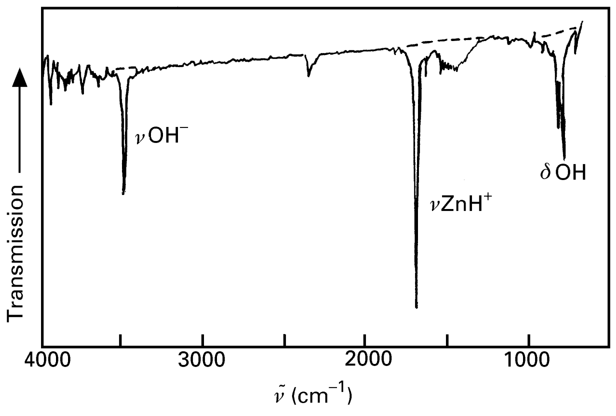



**Figure 14** The IR spectra of hydrogen adsorbed on ZnO. Reproduced with permission of the Royal Society of Chemistry from Hussain G and Sheppard N (1990) *Journal of the Chemical Society, Faraday Transactions* **86**: 1615.

$HZn^+$ and $OH^-$ groups. Oxides have mostly been studied in high-area form to date, including the zeolites whose acidic activity occurs on well defined sites within the pores of the crystalline material. Flat single crystals of oxides are difficult to clean in UHV because of their insulating properties. 'Single-crystal' spectroscopic work on oxides is increasing being carried out on thin films grown epitaxially on metal surfaces.

## Other vibrational spectroscopic techniques for surfaces

Raman spectroscopy provides valuable complementary vibrational information to IR spectroscopy but its applications to adsorbed molecules has been principally limited to the study of finely divided samples for reasons of reduced sensitivity. Exceptionally, flat monolayers of long-chain surfactant have given Raman spectra using multireflection techniques. Surface-enhanced Raman spectroscopy (SERS) gives greatly enhanced sensitivity but only for molecules adsorbed on the roughened surfaces of the coinage metals, particularly silver.

High-resolution electron energy loss spectroscopy (HREELS), with its higher sensitivity but lower resolution, has played a strongly complementary role to IR in the study of molecules adsorbed on single-crystal metal surfaces.

Inelastic neutron scattering (INS) and inelastic electron tunnelling spectroscopy (IETS) have found more limited applications to the study of the adsorption of molecules on high-area surfaces.

*See also:* **ATR and Reflectance IR Spectroscopy, Applications; High Resolution Electron Energy Loss Spectroscopy, Applications; Inelastic Neutron Scattering, Applications; Inelastic Neutron Scattering, Instrumentation; IR Spectroscopy, Theory; Raman and IR Microspectroscopy; Surface-Enhanced Raman Scattering (SERS), Applications.**

## Further reading

Bell AT and Hair ML (1980) *Vibrational Spectroscopies for Adsorbed Species*, ACS Symposium Series 137. Washington, DC: American Chemical Society.

Clark RJH and Hester RE (eds) (1988) *Spectroscopy of Surfaces*, Advances in Spectroscopy, Vol 16. New York: Wiley.

Davydov AA (1984) *Infrared Spectroscopy of Adsorbed Species on the Surfaces of Transition Metal Oxides*. New York: Wiley.

Sheppard N and De La Cruz C (1996, 1998) Vibrational spectra of hydrocarbons adsorbed in metals. *Advances in catalysis*, Part I, **41**: 1–112; Part II, **42**: 181–313.

Sheppard N and Nguyen TT (1978) The vibrational spectra of CO chemisorbed on the surfaces of metal catalysts. In: Clark RJH and Hester RE (eds) *Advances in Infrared and Raman Spectroscopy*, Vol. 5. London: Heyden.

Suëtaka W (1995) *Surface Infrared and Raman Spectroscopy – Methods and Applications*. New York: Plenum Press.

Willis RF (ed.) (1980) *Vibrational Spectra of Adsorbates*, Springer Series in Chemical Physics 15. Berlin: Springer-Verlag.

Yates JT Jr and Madey TE (1987) *Vibrational Spectroscopy of Molecules on Surfaces*. New York: Plenum Press.

# Surface-Enhanced Raman Scattering (SERS), Applications

**WE Smith** and **C Rodger**, University of Strathclyde, Glasgow, UK

Copyright © 1999 Academic Press

VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES
Applications

Surface enhanced Raman scattering (SERS) was first demonstrated by Fleischmann and colleagues in 1974. In a study of the adsorption of pyridine at a silver electrode, they noted that the Raman scattering was considerably stronger when the surface of the electrode was roughened. Jeanmaire and Van Duyne and Albreicht and Creighton reported that the Raman scattering from pyridine adsorbed on a roughened surface was enhanced by a factor of $10^6$ compared to the equivalent concentration of pyridine in solution. This huge increase in signal stimulated a great interest in the technique and it remains one of its main advantages. The technique has been applied in many fields, including surface science, medicinal chemistry and analytical chemistry. Several books and reviews have been written: early developments were surveyed by Furtak and Reyez and Laserna has produced an informative overview indicating the potential to develop a powerful quantitative and qualitative analytical methodology. Chang and Furtak have written a comprehensive book on the subject. Articles directed towards specific applications include one by Vo Dinh targeted at chemical analysis and two by Nabiev and colleagues and Cotton and colleagues targeted at biological and medicinal applications.

## The mechanism of the surface enhancement

The nature of the mechanism that produces SERS is still the subject of debate. Two main mechanisms of enhancement are now most commonly proposed. These are electromagnetic enhancement and charge transfer or chemical enhancement. Electromagnetic enhancement does not require a chemical bond between the adsorbate and the metal surface. It arises from an interaction between surface plasmons on the metal surface and the adsorbed molecule. Chemical or charge transfer enhancement requires a specific bond between the adsorbate and the metal plus energy transfer between the metal and the adsorbate during the Raman scattering process. There is evidence for both mechanisms. The predominant view appears to be that both may occur.

### Electromagnetic enhancement

On smooth surfaces, surface plasmons exist as waves of electrons bound and confined to the metal surface. However, on a roughened metal surface, the plasmons become localized and are no longer confined and the resulting electric field can radiate both in a parallel and in a perpendicular direction. When an incident photon falls on the roughened surface, excitation of the plasmon resonance of the metal may occur, causing the electric field to be increased both parallel and perpendicular to the surface. The adsorbate is bathed in this field and the Raman scattering is amplified. This mechanism has been studied and reviewed by Weitz, Moskovits and Creighton.

Since SERS has been obtained from molecules spaced off the surface, the existence of enhancement from this type of mechanism is well established.

### Charge transfer enhancement

The enhancement from the charge transfer mechanism is believed to result from resonance Raman scattering from new resonant intermediate states created by the bonding of the adsorbate to the metal. The adsorbate molecular orbitals are broadened into resonance by interaction with electrons in the conduction band. Resonance states whose energies lie near the Fermi energy are partially filled, while those lying well below are completely filled. Otto has provided much evidence of the existence of this effect. He showed that there was a specific first layer and has extensively reviewed the field. Campion reported direct experimental evidence linking new features in the electronic spectrum of an adsorbate to SERS, under conditions where electromagnetic enhancements were unimportant. He noted that it was difficult to observe charge transfer only because the electromagnetic effects had to be accounted for and removed. This problem was overcome by conducting SERS on an atomically flat, smooth single-crystal surface where the electromagnetic effects were small and well understood. He adsorbed pyromellitic dianhydride (PMDA) on to copper(III) and observed an enhancement of a factor of 30. In addition, a low-

energy band in the electronic spectrum from the adsorbed PMDA was observed that was absent in the solution PMDA spectrum.

## Selection rules

Selection rules have been derived for electromagnetic SERS enhancement. The advantage of electromagnetic enhancement is that, since no new chemical species is formed on the surface, the selection rules can be based on the properties of the molecular adsorbate rather than on an ill-understood surface complex. In its simplest form, assuming no specific symmetry rules, the most intense bands are those where a polarization of the adsorbate electron cloud is induced perpendicular to the metal surface. However, more detailed selection rules can be obtained when the molecule has symmetry elements. Creighton and Moskovits have independently reviewed the principles.

## Nature of the substrate

The active substrates are usually made from a limited number of metals. Silver, gold and copper are the most commonly used SERS-active metals, although the use of lithium is well established. These substrates were chosen because their surface plasmons exist in or close to the visible region. Ideally, the excitation from the laser should coincide with the plasmon resonance frequency of the particular surface created and conditions such that the efficiency of absorption of the light is reduced and the efficiency of scattering increased. Silver is the most commonly used substrate, although gold is often used particularly in the near infrared. The original experiments used electrochemistry and this is a good method of obtaining a suitable surface. The scale of the roughness required is between about 40 and 250 nm for visible excitation with silver. SERS of pyridine obtained using an electrode setup results in certain bands appearing strongly in the pyridine spectra and the relative intensity and absolute intensity is dependent on voltage. The maximum enhancement is believed to be when the Fermi level matches the energy of the $\pi$ orbital of pyridine. The electrode working surface can be difficult to reproduce and is prone to annealing with time in certain environments. However, sensitive qualitative analysis is feasible.

Colloidal suspensions are particularly attractive as they can be prepared in a one-pot process and are inexpensive. Reliable SERS analysis is possible as a fresh surface is available for each analysis. Many different methods of colloid preparation have been reported. Some groups always use freshly prepared colloid for their experiments, but recent emphasis has been on obtaining reproducible, monodisperse colloid that is stable for several months. In particular, colloid prepared by the citrate reduction of silver can be produced in almost monodisperse form and with a lifetime of up to one year or more. The particle size of these colloids varies. In one standard preparation of citrate-reduced silver colloid, a transmission electron microscopy study indicated that the particles were approximately 36 nm in their longest dimension and were small hexagonal units (**Figure 1**). Photoelectron correlation spectra of the suspension indicated that the average particle size approximated to a sphere was 28 nm.

Metal colloidal particles adsorbed upon or incorporated into porous membranes such as filter papers, gels, beads, polymers, etc. have been developed as SERS-active substrates. Although these substrates are claimed to be reproducible, they are not widely used, probably because they involve complicated preparative procedures and are susceptible to contamination and self-aggregation. Ruled gratings can be used to give good reproducibility and abraded surfaces, although not so reproducible, they are attractive because of their simplicity of preparation. Numerous researchers have reported that immobilization of the colloidal particles as ordered arrays on films gives reproducible and sensitive SERS sensors.

## Surface enhanced resonance Raman scattering

Surface enhanced resonance Raman scattering (SERRS) is obtained by using a molecule with a

**Figure 1** Transmission electron microscopy image of silver colloid, × 250.

chromophore as the adsorbate and tuning the excitation radiation to the frequency of the chromophore. The effect was originally reported by Stacey and Van Dyne in 1983. The enhancement obtained is very much greater than that of either resonance Raman or SERS, enabling very sensitive analysis and low detection limits to be achieved. Although SERRS is best considered as a single process, it arose experimentally from the combination of two previously studied effects, namely resonance scattering and surface enhanced Raman scattering (SERS). It is a unique process and different effects can be obtained depending on the nature of the chromophores used and the choice of laser excitation. **Figure 2** illustrates the main choices. In **Figure 2A**, the molecular chromophore (curve a) is chosen not to coincide with the maximum of the plasmon resonance (curve b). If laser excitation at the molecular absorption maximum is used, the maximum contribution to the overall effect from resonance enhancement would be expected. With the arrangement in **Figure 2A** and with the excitation at the molecular resonance maximum, it has been reported that for azo dyes there is reduced sensitivity to surface enhancement mechanisms, providing a signal that is less sensitive to the nature of the surface and that has a recognizable molecular 'fingerprint' related to the resonance spectrum, making this arrangement better for quantitative analysis.

The second possible arrangement illustrated in **Figure 2A** is where the laser excitation is set off the frequency of the adsorbate resonance and at the maximum of the plasmon resonance (2). For resonance experiments on the molecule alone, this would be described as a preresonant condition and often SERRS undertaken in this way is written as SE(R)RS. More orientation information is to be expected and additional bands have been observed and assigned as due to mechanisms of surface enhancement. However, in this preresonant condition, the selectivity of resonance still applies. Thus, it is possible to pick out individual molecules in the presence of a matrix of interferents, but the effect will now be more dependent on the angle of the adsorbate to the surface. For many surface studies this is a key point and consequently this experimental process may be preferred for surface analysis.

**Figure 2B** gives an alternative case in which the molecular chromophore (curve a) coincides with the surface plasmon maximum (curve b). Similar considerations will apply, but a greater increase in sensitivity is likely at the resonance and plasmon resonance maximum frequency.

Hildebrandt and Stockburger carried out an extensive study on SERRS of Rhodamine 6G in order to explore the enhancement mechanisms involved. They

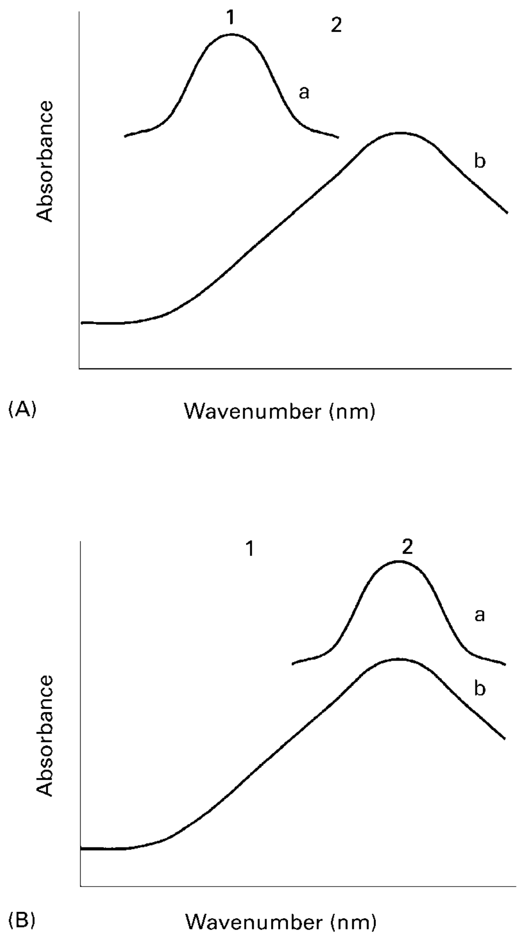



**Figure 2** Illustration of the different arrangements for SERRS: curve a is the molecular absorbance and curve b is the plasmon absorbance. In (A) the molecular and plasmon absorbances do not coincide. Position (1) represents excitation at the molecular maximum and (2) represents excitation at the plasmon maximum. In (B) the molecular and plasmon maxima coincide. Position (1) represents excitation away from molecular and plasmon maxima and (2) represents excitation at the absorbance and plasmon maximum.

reported that two different types of adsorption sites on the colloid surface were responsible for the enhancement experienced: an unspecific adsorption site that had high surface coverage on the colloid surface resulted in an enhancement factor of 3000 and could be explained by a classical electromagnetic mechanism; a specific adsorption site was only activated in the presence of certain anions ($Cl^-$, $I^-$, $Br^-$, $F^-$ and $SO_4^{2-}$). This specific site had a low surface coverage (approximately 3 per colloidal particle); however, an enhancement of $10^6$ was claimed to result. This enhancement was believed to be due to a charge transfer mechanism. This study was continued into the near-infrared region and extended to include gold colloid and gold and silver colloid supported on filter

papers. The enhancement experienced from anion activation with silver colloid was stronger by a factor of 47 in the near-infrared region compared with the visible region. The authors concluded that this phenomenon could be accounted for by the charge transfer transition being shifted towards the red for Rhodamine 6G, increasing the resonance effect in the near-infrared region.

## Advantages and disadvantages of SERS/SERRS

SERS incorporates many of the advantages of Raman spectroscopy in that visible lasers can be used so that flexible sampling is possible, and there is little signal from water so that *in situ* examination of such surfaces as those of colloidal particles in aqueous suspension or of electrode surfaces under solvent can be carried out. The greatest advantages are the sensitivity that can be obtained and the selectivity of the signals. Since SERS will detect compounds down to a level of about $10^{-9}$ M, adsorbates at monolayer coverage or less can be studied easily. Experiments on pyridine are a classic example. At well below monolayer coverage, pyridine is believed to lie with the plane of the ring almost flat on the metal surface. Under these conditions, there is very little intense Raman scattering since the main polarizability changes in the molecule are parallel to the surface. As the surface density of pyridine increases, the molecule is forced into a more vertical position and the signal begins to appear quite rapidly. This forms a good probe of when monolayer coverage occurs. Further, the existence of selection rules means that an indication of the nature of the surface processes can be obtained.

There are a number of key limitations on the method. First, to obtain a large effect, SERS can be used only for adsorbates on a limited number of metal surfaces in correctly prepared (roughened) form. Second, the very large surface enhancement coupled to the need for a specific molecule to be adsorbed on the surface makes the technique prone to interference. Contaminants that give strong surface enhancement can be detected in much lower concentration than the adsorbate studied, leading to problems in identification. The additional complexity that the intensity of the bands depends on only partially understood selection rules and can change depending on the angle of the molecule to the surface and the degree of packing makes it difficult to assign bands. Finally, there is a tendency in SERS for photodecomposition to occur on the surface. Characteristic broad signals that have been reported as being due to specific surface adsorbates are probably pyrolysed species on the surface. Notwithstanding these problems, SERS is unique in providing a fascinating insight into the adsorption mechanisms of molecules on suitable surfaces *in situ*.

The technique of SERRS might be assumed to have some of the same disadvantages as SERS and more limitations, but in fact SERRS is proving to be a much more effective technique for analytical science. The major advantage of SERRS is that, if correctly applied, the chromophore signal dominates. Since related spectra are obtained by resonance from solution, the spectra on the surface can easily be recognized, and since the Raman signal from the chromophore is enhanced more than any other molecule this particular species is very readily identified at the surface. Thus, in contrast to the difficulty in assigning signals in SERS in some cases, the signal assignment in SERRS is often simple and reliable. Further, and rather surprisingly, a fluorescence quenching mechanism occurs on the surface so that both fluorescent and nonfluorescent dyes give good SERRS. Provided the molecule is attached to the surface, there is little fluorescence background. In fact, it is often useful to establish the fluorescence background against the strong SERRS signals in order to measure the degree of adsorption and desorption from the surface. Thus, a wide range of chromophores is available. Further, the technique requires very low laser powers and consequently the photodegradation common in SERS is seldom a problem.

The characteristic spectra routinely observed with SERRS permit the identification of mixtures without the need for preseparation. Munro and colleagues have reported the analysis and characterization of 20 similar monoazo dyes, all of which produced unique characteristic spectra that in turn permitted the simultaneous analysis and detection of five dyes presented in a crude mixture. They addressed the problems associated with reproducibility and have focused much attention on improving and standardizing the production of the silver colloid generally used to obtain SERS. They concluded that, with careful attention to detail, a relative standard deviation (RSD) of 5% was routinely obtained.

## Applications of SERRS and SERS

The advantages reported above have been exploited in numerous research fields, including the following.

*Biochemistry* SERRS methodology can be modified in order to provide a biocompatible environment for biological materials. The identification of water-soluble porphyrins and their photostability and

interaction with roughened metal surfaces have been reported. For copper chlorophyllin, spectra obtained by using different excitation frequencies permitted a better understanding of a complex system. SERRS enabled the identification of novel chromophores in eye lenses without preseparation of the crude mixture. Finally the oxidation and spin state of proteins such as myoglobin and cytochrome P450 can be probed, and the use of a fluorescent low-concentration protein solution to study labelled tyrosines has been reported.

*Medicinal chemistry* SERRS has been used to detect the antitumour drug mitoxantrone and its interaction with DNA *in situ*. The adsorption of the drug complex onto a colloidal surface did not destroy or interfere with the native structure. Therefore, bonding information from the complex was extracted from the SERRS spectra. Selective detection on DNA at ultra-low detection by SERRS has been developed.

*Surface chemistry* SERRS has been used to probe electrode surfaces *in situ* to extract structural information and to provide quantification. The *in situ* SERRS detection of compounds such as 2,4,6-trinitrobenzene sulfonic acid covalently bound to tin oxide was observed when the chemically modified surface was coated with silver. The spectra collected provided rapid and sensitive structural information that was semiquantitative. SERRS has also been used to follow reactions occurring at well below monolayer coverage at roughened metal surfaces.

*Polymer science* The ability to probe surfaces and boundaries using *in situ* SERS has been exploited extensively in polymer chemistry to characterize the surface of polymers for comparison with the bulk properties, and to determine the molecular geometry, orientation of polymer side groups adjacent to the metal surface and information on bonding, for example of polymer–metal composites such as adhesives and coatings.

*Forensic science* Modern Raman spectrometers connected to microscopes enable the examination of small amounts of material such as single fibres. The sensitivity and selectivity of SERRS can be exploited in forensic science by determining the nature of the dye mixture *in situ* from a single fibre, from an ink or from a lipstick smear.

*Corrosion science* Studies of corrosion inhibitors, particularly for copper using SERS of the inhibitor adsorbed on the roughened metal surface, have been used to selectively identify the species. The limitation of requiring a roughened metal surface of a particular metal can be overcome by applying colloid to a smooth nonactive surface, but this field has yet to be exploited.

Practical uses of SERRS have been developed. It has been used to prepare a robust disulfide pH indicator by coupling pH-sensitive dyes – methyl red, cresol violet and 4-pyridinethiol – to cystamine, which adsorbs strongly to the roughened metal surface, forming monolayer coverage of the complex with colloidal silver and allowing strong SERRS to be recorded. Changes in the pH result in changes in the chromophores of the dyes that were easily detected by SERRS. As the SERRS spectrum obtained from the complex was pH sensitive, it was possible to obtain quantitative pH determination.

Another example involves the exploitation of the sensitivity of the technique to analysis of trace amounts of nitrite in fresh and sea waters: sulfanilamide was added to the water sample and reacted with any nitrite present, forming a diazonium salt that was then coupled with ethylenediamine to produce an azo dye that was then detected by SERRS. An enhancement of a factor of $10^9$, a relative standard deviation of 10–15% and a limit of detection of picograms were reported. This method was superior to existing colorimetric and chemiluminescence techniques used to analyse the water samples for nitrite. Ultrasensitive detection of metal ions has been reported. A limit of detection at the nanogram level was claimed. The metal ions nickel or cobalt and a ligand, a mixture of 2-pyridinecarboxyaldehyde and 2-pyridinehydrazone or 1,10-phenolanthroline, form a complex on the roughened metal surface that is then detected by SERRS.

## Single molecule detection

Rhodamine 6G has been used extensively as a model dye to probe the nature of the SERRS effect. It is an extremely strong fluorophore when excited by visible radiation. Hence normal Raman is not observed except with near-infrared excitation. However, when SERRS is used the dye adsorbs strongly to the roughened metal surface and consequently this strong fluorescence is quenched and an extremely strong, enhanced Raman signal is observed. **Figure 3** illustrates the resonance Raman and SERRS spectra collected from Rhodamine 6G. Attomolar levels ($10^{-18}$ M) of detection have been reported for this system, which is approaching single molecule detection. The fluorescence-quenching properties of surface enhancement coupled with the additional sensitivity obtained from SERRS have been exploited by several researchers. Rhodamine 6G adsorbs very effectively on the roughened silver

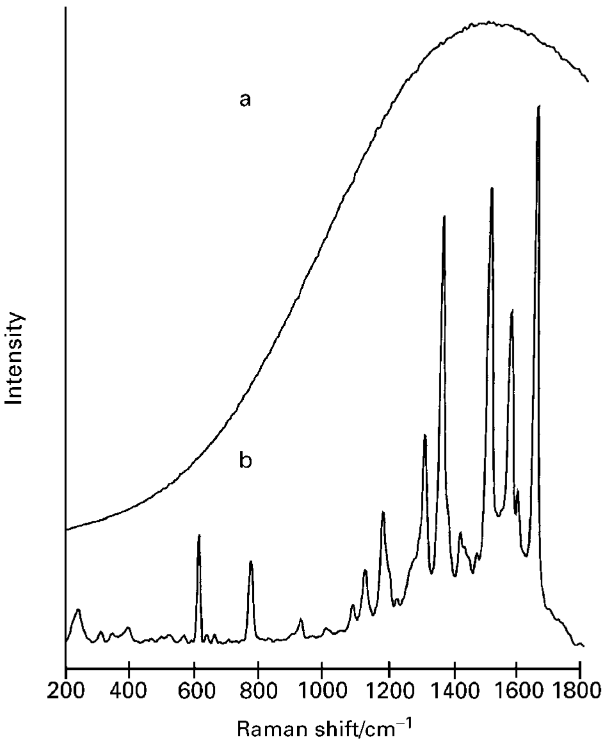



**Figure 3** Curve a: Solution spectrum from a $10^{-6}$ M Rhodamine 6G solution using 514.5 nm excitation, demonstrating the predominance of fluorescence over resonance Raman scattering. Curve b: SERRS spectrum taken from a suspension of aggregated silver colloid to which 150 μL of a $10^{-8}$ M Rhodamine 6G solution has been added using 514.5 nm excitation.

surface. However, the detection of single adsorbates of dopamine or phthalazine on colloidal clusters, with a limit of detection at picogram levels, illustrates that ultrasensitivity of this technique for other adsorbates is possible.

The ability of SERRS to detect one molecule has recently been demonstrated by three groups. Nie has isolated colloidal particles with rhodamine adsorbed onto glass slides and obtained spectra from the individual particles. The particles are preselected for size to ensure that the surface plasmon of the single particle is in the visible region. Kneipp has used near-infrared anti-Stokes scattering and statistical methods to demonstrate that single molecules can be observed in suspension and Graham and colleagues have shown that one molecule of DNA labelled with a covalently attached fluorescein dye can be detected in the interrogation volume of suspended and aggregated colloid.

## Conclusion

In summary, surface enhancement results in a huge enhancement in Raman scattering and the ability to observe Raman signals at very low concentrations. The resulting spectra provide molecular information and are unique to individual molecules. Sample preparation is simple and it is possible to undertake analysis *in situ* under water or in air or in vacuum. Since there are new selection rules and the effect is dependent on the metal used and the degree of surface roughness, there is a wealth of surface information to be obtained from SERS provided the limitations in terms of contamination and photodecomposition are remembered. The main problems are the limited number of surfaces to which the method can be applied and difficulties in interpreting the spectra.

Additional advantages can be obtained from the use of SERRS. It is more sensitive and the resulting spectrum can be related back to the molecular resonance spectrum, making for more confidence in assignments. Fluorescence is quenched, signals from the adsorbate are much more intense than from contaminants and there is less dependence on the exact nature of the surface. This makes for unique applications for SERRS, which include simpler, more sensitive and more selective quantitative analysis and single molecule detection.

*See also:* **Biochemical Applications of Raman Spectroscopy; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy; FT-Raman Spectroscopy Applications; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; MRI of Oil/Water in Rocks; Polymer Applications of IR and Raman Spectroscopy.**

## Further reading

Campion A, Ivanecky JE, Child CM and Foster M (1995) *Journal of the American Chemical Society* **117**: 11 807.

Chang RK and Furtak TE (1982) *Surface Enhanced Raman Scattering*. New York: Plenum Press.

Cotton TM and Chumanov G (1991) *Journal of Raman Spectroscopy* **22**: 729.

Creighton JA (1988) In: Clark RIH and Hester RE (eds) *Spectroscopy of Surfaces*. Chichester: Wiley.

Fleischmann M, Hendra PJ and McQuillan AJ (1974) *Chemical Physics Letters* **26**: 163.

Furtak TE and Reyez J (1980) *Surface Science* **93**: 351.

Laserna JJ (1993) *Analytica Chimica Acta* **283**: 607.

Moskovits M (1982) *Journal of Chemical Physics* **77**: 4408.

Nabiev I, Chourpa I and Manfait M (1994) *Journal of Raman Spectroscopy* **25**: 13.

Vo Dinh T, Alak A and Moody RL (1988) *Spectrochimica Acta* **43B**: 605.

Weitz DA, Moskovits M and Creighton JA (1986) In: Hall RB and Ellis AB (eds) *Chemistry and Structure at Interfaces, New Laser and Optical Techniques*, p 197. Florida: VCH.

# Symmetry in Spectroscopy, Effects of

**SFA Kettle**, University of East Anglia, Norwich, UK

Copyright © 1999 Academic Press

**FUNDAMENTALS OF SPECTROSCOPY**

**Theory**

In this brief article an attempt is made to review those aspects of group theory that may be of concern to a nonspecialist reader. It is written on the assumption that the reader may at some point decide to study a particular aspect of the subject in more detail. However, specialist articles are not always the most accessible and so a particular effort has been made to explain key points in nonmathematical terms. Hopefully, this will provide an insight which will be helpful should the going get tough! For simplicity, reference will always be made to 'molecules'. This should in no case be taken to exclude other species.

Modern spectroscopy makes extensive use of group theory. It does so at several related points. Perhaps most important is in the interpretation of the data that it produces. However, spectral interpretation does not stand alone. In order that the data may be interpreted, the states to which they relate must also be subject to a group theoretical description. In that the states are normally product functions of some sort, the individual functions must themselves have a group theoretical basis. Then, the spectral data correspond to transitions that are made allowed by virtue of some physical process – electric dipole transitions are perhaps the most widely studied but the continuing growth of NMR and EPR methodologies, *inter alia*, is making magnetic dipole transitions of increasing importance; it is also relevant when optical activity is the subject of discussion. The physical process also has to be susceptible to a group theoretical description.

There is an alternative approach to the topic which is of no less validity. This is the study of commuting operators. The statement that two operators commute is equivalent to the statement that you can, in principle at least, make simultaneous measurements of the physical properties associated with each of them. The set of operators that make up the point group of a molecule commute with the Hamiltonian operator and so one can have knowledge of energies which correspond to symmetry-distinct levels. If one can recognize the existence of a valid operator then it can be used in the classification of energy levels; there is no requirement that one can implement in some way the associated physical process. The permutation operator and time-reversal operators fall into this category, but so too do the operations of the point groups of spectroscopic importance. So, it can prove convenient to extend the concept of the group of operations which commute with the Hamiltonian beyond the simple point group operators and to incorporate others with them.

But, to begin at the beginning. To the majority of users the applications of group theory to spectroscopy start with the simple point groups. The symmetry of a molecule is expressed by the statement that there exist within it certain symmetry elements, such as rotation axes or mirror planes, which relate equivalent parts of a molecule to each other. To apply such an approach there is an implicit assumption that the molecule is a rigid body locked into its lowest potential energy arrangement. It is commonplace to state that it is not the symmetry elements that are of importance but, rather, the corresponding operations. More accurate still is to state that it is the corresponding operators that are of importance, since it is they which commute with the Hamiltonian. It is also important to recognize that there is no requirement of a 1:1 relationship between symmetry element and operator. There is only a single identity (leave alone) operator – often denoted E – but for any molecule there is an infinite number of $C_1$ axes ($C_n$ means rotation by $360/n$). There is a 1:2 relationship between $C_n$, $n > 2$, axes and the corresponding $C_n$ operators. Other symmetry elements/operators that will commonly be encountered are mirror planes/mirror plane reflections, denoted $\sigma$. These admit of a classification into three types. First, $\sigma_v$, where the v stands for vertical and indicates that the mirror plane contains the rotational axis of highest $n$ (i.e. if the axis is vertical, so too is the mirror). Second, $\sigma_h$, where the h stands for horizontal and indicates that the mirror plane is perpendicular to the rotational axis of highest $n$ (i.e. if the axis is vertical, the mirror is horizontal). Thirdly, $\sigma_d$, where the d is the first letter of the word diagonal. These mirror planes bisect the angle between two twofold axes and also have an axis of at least twofold symmetry lying in them. Care is needed, however, because the use of the symbol $\sigma_d$ is subject to abuse. If a set of mirror planes is correctly labelled $\sigma_d$ in one point group they may well be labelled the same way in a

subgroup, even though the latter does not contain the two-fold axes of the parent. Other symmetry elements/operators are a centre of symmetry/inversion. These are commonly denoted i (the same symbols are usually used for symmetry element and operator, although different fonts may be used to distinguish them). The final symmetry elements/operators are denoted $S_n$. They correspond to a rotation about $360/n$ followed either by reflection in a plane perpendicular to the axis or by inversion in the centre of mass of the molecule. Either definition may be used; the sets resulting are in a 1:1 correspondence. The former definition is perhaps the more popular but the latter is perhaps the more useful (it is convenient to have a simple way of determining parity in a descent in symmetry from spherical to point group). Each set of symmetry operators comprise a group which is given a unique label, usually a combination in some way of the labels used for the operations that it contains. An exception to this is found for those groups which have a symmetry such that the coordinate axes are all symmetry-related. They have individual names ($I_h$, $O_h$, $T_d$ are the labels of the most important icosahedral, octahedral and tetrahedral groups).

The combination of the group operators gives rise to a group multiplication table and the characters of sets of matrices which multiply isomorphically to the group operators are collected in character tables. These character tables are the starting point for many applications and are widely available, most commonly as appendices in relevant books. A typical character table is given in **Table 1**.

The table of characters is square; each row corresponds to a different irreducible representation of the group. Each irreducible representation has a unique label, given at the left-hand side, such labels are widely used. The label E (with or without suffixes or primes) indicates double degeneracy. Not used here, but the label T (or sometimes F) indicates triple degeneracy. For emphasis, the characters are divided into four sets which will be seen to have a very simple relationship with each other. At the right-hand side of the table are given two columns of basis functions. These are functions which may be used to generate the set of characters in their row but, more usefully, they provide information which can be useful in a spectroscopic analysis. For instance, the electric dipole operators have the same symmetry species as the coordinate axes (here, $A_2'' + E'$). One important use of character tables is to decompose reducible representations into their irreducible components. Reducible representations are produced by most real-life applications of group theory to molecular problems, when a sum of irreducible representa-

**Table 1** Table of characters

| $D_{3h}$ | E | $2C_3$ | $3C_2$ | $\sigma_h$ | $2S_3$ | $3\sigma_d$ | | |
|---|---|---|---|---|---|---|---|---|
| $A_1'$ | 1 | 1 | 1 | 1 | 1 | 1 | | $z^2$: $X^2+y^2$ |
| $A_2'$ | 1 | 1 | –1 | 1 | 1 | –1 | $R_z$ | |
| E' | 2 | –1 | 0 | 2 | –1 | 0 | $(T_x, T_y)$ | (x,y)<br>$(1/\sqrt{2}\,[x^2–y^2],xy)$ |
| $A_1'$ | 1 | 1 | 1 | –1 | –1 | –1 | | |
| $A_2''$ | 1 | 1 | –1 | –1 | –1 | 1 | $T_z$ | z |
| $E''$ | 2 | –1 | 0 | –2 | 1 | 0 | $(R_y, R_x)$ | (zx, yz) |
| ? | 6 | 0 | –2 | 2 | 2 | –2 | | |

**Table 2** Direct product table for the $D_{3h}$ group

| $D_{3h}$ | $A_1'$ | $A_2'$ | $E'$ | $A_1''$ | $A_2''$ | $E''$ |
|---|---|---|---|---|---|---|
| $A_1'$ | $A_1'$ | $A_2'$ | $E'$ | $A_1''$ | $A_2''$ | $E''$ |
| $A_2'$ | $A_2'$ | $A_1'$ | $E'$ | $A_2''$ | $A_1''$ | $E''$ |
| $E'$ | $E'$ | $E'$ | $(A_1'+A_2'+E')$ | $E''$ | $E''$ | $(A_1'+A_2'+E'')$ |
| $A_1''$ | $A_1''$ | $A_2''$ | $E''$ | $A_1'$ | $A_2'$ | $E'$ |
| $A_2''$ | $A_2'$ | $A_1''$ | $E''$ | $A_2'$ | $A_1'$ | $E'$ |
| $E''$ | $E''$ | $E''$ | $(A_2'+A_2''+E'')$ | $E'$ | $E'$ | $(A_1'+A_2'+E')$ |

tions is generated and the details of the sum have to be determined. For instance, immediately beneath the character table above is a reducible representation which has $2A_2' + E' + E''$ components. There are simple systematic ways of decomposing reducible representations based on the orthonormality of the irreducible representations. This can be seen in that the product of pairs of characters of any two different irreducible representations, summed over every operation of the group, sum to zero. On the other hand, the squares of characters of any irreducible representation, summed over all of the operations of the group, sum to the number of operations in the group, 12 in the group above.

One cannot go far in spectroscopy without encountering product functions. The individual component functions of product functions have their symmetry properties described by one of the irreducible representations in the appropriate character table. In order to determine the symmetry species of the product function a table of direct products is needed, which enables this to be determined from the irreducible representations of the component functions. Such direct product tables may be obtained from the corresponding character table. That corresponding to the $D_{3h}$ group above is shown as **Table 2**.

The symmetry of **Table 2** arises, of course, because it results from the multiplication of numbers. Such direct product tables can be used to determine the symmetry species of product functions, irrespective of the number of functions; they can be included in sequence. So, the symmetry of the first overtone or combination bands in vibrational spectroscopy is immediately obtained from a direct product table.

However, care has to be taken in the case of overtones of degenerate vibrations. The case of the first overtone illustrates the problem. For the overtone of a doubly degenerate mode there are just three overtone functions, roughly aa, ab and bb. Four are given in the table above. Similarly, for the overtone of a triply degenerate mode the overtone functions, roughly, are aa, ab, bb, bc, cc and ca, six in total, compared with the nine that appear in a direct product table. The reason is that the direct product can be divided into a symmetric and an antisymmetric part (with respect to interchange of the component functions). For overtones, only the symmetric part is relevant, the antisymmetric functions self-cancel (the functions above are rough because they are not all properly symmetric).

Direct product tables also have importance in determining the selection rules of spectroscopy. The statement that a particular transition is allowed is a statement that a transition moment integral is nonzero. Selection rules are general statements about which of all of the possible transition moment integrals can be nonzero. It is group theory which is used to determine which integrals can be nonzero.

An integral may be regarded as the sum of an infinite number of tiny component contributions, one from each of the infinite number of tiny volume elements of space. But for any given tiny volume of space there are $(n-1)$ symmetry-related tiny volume elements (where the number of symmetry operations in the group is $n$; $n$ is often called the 'order' of the group). When all of the symmetry-related volume elements make equal (in magnitude and sign) contributions to the integral their sum does not vanish and the integral (which includes the contributions from all of such sets) is nonzero. If $n/2$ of the symmetry-related tiny volume elements are positive and $n/2$ are negative then their combined contribution to the integral is zero. Since this pattern is repeated for each and every set of symmetry-related tiny volume elements, the integral itself is zero. But can one make statements about the phase relationships between the symmetry-related members of sets of tiny volume elements?

Character tables such as that above are statements about the phase relationships between functions in the space spanned by the character table. For instance, it is easy to see that any $A_2'$ function has equal positive and negative contributions – multiply each $A_2'$ character by the number of regions of space to which it refers (the numbers at the head of each column) – and so an integral over all space is symmetry-required to be zero. Only integrals of $A_1'$ symmetry (in the $D_{3h}$ point group) can be nonzero. In general, only integrals which transform as the totally symmetric irreducible representation (which always has characters of +1) of a group can be nonzero. Now, this irreducible representation always occurs on the leading diagonal of the table of direct products and nowhere else in this table. This means that it occurs whenever the direct product is one formed between an irreducible representation and itself. It is this matching which is the basis of simple selection rules. For instance, the often-stated requirement for infrared activity – that to be spectrally active a vibrational mode has to transform like a coordinate axis. This arises because in the triple direct product derived from the symmetry species of:

$$\text{(excited state) (operator) (ground state)}$$

the vibrational ground state is totally symmetric – the molecule is assumed to be nonvibrating (actually, it would almost be equally valid to say that it is assumed that only nondegenerate vibrations are excited in the ground state – this is the way the group theory works out). For the corresponding integral to be nonzero, the symmetry species of the excited state (which is the symmetry of the mode excited) has to be the same as that of the operator. It is an electric dipole transition and so the symmetry of the operator is that of an electric dipole – and this is the same as that of a coordinate axis.

All of this has assumed a rigid, isolated, molecular species. If we are interested in molecular rotations the theory clearly needs extension – it covers bulk rotations (these find a place in the list of basis functions in the character table above) but there is nothing evident in the theory which would enable the classification of rotational energy levels. Equally, the zero point vibrational energy – and its destructive effect on the assumed molecular geometry – has conveniently been neglected. In fact, there is a validity in this last step. It arises from the fact that the vibrations of a molecule are such that it explores a multidimensional space (each molecular vibration represents an independent variable and so the greater the number of vibrations, the greater the number of dimensions that the molecule can explore vibrationally). This multidimensional space has an inherent symmetry. Thus, if a molecule has (when rigid) a mirror plane of symmetry, then if it is distorted there will be another configuration which is of equal energy to the first but the mirror image of it. That is, all of the symmetry operations of the rigid molecule have a relevance to the multidimensional potential surface. In fact, if one works through the theory, one ends up with a group which is isomorphic to that of the rigid molecule. It is for this reason that the use of the rigid-molecule group gives valid results – it is

isomorphic to the correct group. This theory might be called the theory of nearly-rigid molecules; molecules that make but small deviations away from an equilibrium geometry. What of the other extreme – molecules which, whilst retaining their molecular identity, are nonetheless internally very mobile. One has to deal with the molecular symmetry group (as opposed to the molecular point group) of a molecule. Consider what has become a classic example, the molecule $CH_3BF_2$, which has a rather free rotation about the C–B bond axis. Consider the molecule in a configuration in which it has no point group symmetry other than the identity. Appropriate is an arrangement in which, viewed down the C–B axis, one F almost eclipses an H. Suppose that the H is slightly to the left. Two arrangements of equal potential energy can be generated by rotating one of the two other hydrogens into the position occupied by the first. Three equienergetic arrangements in all. Three more can be obtained by rotating the second F to take the place of the first. Now we have six equivalent arrangements. Six more can be obtained if we go back to the beginning but now place the H in an equivalent position but now to the right. The (physically feasible) interconversions which relate these 12 arrangements constitute the molecular symmetry groups of $CH_3BF_2$. The definition of 'physical feasibility' is at the heart of the definition of the molecular symmetry group of a molecule and, in contrast to the definition of a molecular point group, has sometimes proved somewhat controversial. In part, at least, this has resulted from attempts to define such groups in a compact way. For instance, inversion of the positions of all particles in the centre of mass of the molecule – a physically highly unfeasible operation – can usefully be included!

Next, the assumption of an isolated molecule. When in solution, a molecule will be subject to an infinity of different solvent environments, unless there is some reason that a particular solvent is 'frozen' around the dissolved species. Such cases are rare. The infinity of environments will cause a broadening of most spectral transitions but there is little more that can be said in general. Much more can be said when the environment is that of a crystalline solid. For a solid, formally, one is not concerned with a point group but, rather, with a space group. These are the space groups of classical crystallography but care has to be taken in that classical crystallography may make use of a nonprimitive unit cell (body or face centred). For spectroscopic purposes it is essential that only primitive unit cells are used (a unit cell, by pure translations generates the entire crystal. If one doubles the size of the unit cell, by body-centring it, for instance, then the number of translation operations is halved). Even so, it is to be noted that there is no unique definition of a unit cell. If the spectroscopy is such that the surface of the crystal is not of importance then the crystal faces are ignored and the crystal is, effectively, infinite. An important simplification arises because for the vast majority of spectroscopic measurements performed on a crystal, the wavelength of the incident radiation is much greater than a typical primitive lattice translation vector. This enables the use of a so-called factor group in which the entire set of translation operations are incorporated into the identity operation. The set of operations which remain are isomorphic to one of the 32 crystallographic point groups; this isomorphism is exploited by the use of the character table of the appropriate one of the 32 crystallographic point groups in the subsequent analysis. Pictorially, one can think of there being a coupling between all of the molecules within a (primitive and, of course, arbitrary) unit cell, but in reality the formal analysis covers all unit cells. The method of analysis can be called the 'unit cell' or 'factor group' models (they differ in approach, not results); the splittings that occur on 'molecular' features are referred to as 'factor group', 'correlation field' or 'Davydov' splittings. If coupling between molecules does not have to be invoked, the molecular environment in the crystal, the site, may well lead to splittings in (molecular) degenerate modes – site symmetries are commonly lower than molecular. The 'site group' model is then appropriate and the splittings are known as 'site splittings'. The irreducible representation of the group of all translations does not enter any of these analyses. However, they become important in vibrational spectroscopy, for example, if anharmonic vibrations are considered (and they well may be, involving for instance, coupling with lattice modes). In such cases the Brillouin zone, representing the $k$ vectors which define the irreducible representations of the group of all translations, has to be invoked.

Related to the spectroscopy of crystals is the spectroscopy of surfaces and, particularly, the spectroscopy of species adsorbed on crystal surfaces. For perfectly conducting metals, there is an important selection rule in that such surfaces image any electric dipole within an adsorbed molecule. When such dipoles are perpendicular to the surface the dipoles reinforce; when they are parallel to the surface they cancel. This gives rise to the so-called 'surface selection rule', that it is only possible to observe by electric dipole spectroscopy those modes which involve dipole moment changes perpendicular to the surface. This requirement can be expressed group theoretically by use of the so-called diperiodic groups in two dimensions.

Often not too far away from an application of group theory in spectroscopy is the fact that functions derived from atomic orbitals are under study. The study of the group theory of spherical symmetry is fascinating, an infinitesimal rotation about any axis being a symmetry operation. The corresponding infinitesimal rotation operators have important commutation relationships which provide a basis for the theory of angular momentum, for instance. The descent from spherical to molecular symmetry usually means a reduction in degeneracy. For a species corresponding to an angular momentum $j$ the character, $\chi$ corresponding to a rotation of $\theta°$ is given by:

$$\chi_j(\theta) = \frac{\sin\frac{1}{2}(2j+1)\theta}{\sin\frac{1}{2}\theta}$$

This equation is equally applicable when $j$ is halfintegral but then it is necessary to work in the so-called double group, in which the factor of $\frac{1}{2}$ in front of the $\theta$ means that a rotation of 720° is required to give the identity. Corresponding to each operation in the 0–360° sector there is another in the 360–720°. A $C_2$ rotation still corresponds to a 180° rotation but it now takes four such operations in succession to generate the identity. Perhaps not surprisingly, $C_2$ in a double group closely resembles $C_4$ in a 'normal' group. So, although the group $C_{2v}$ is Abelian (all characters |1|), the $C_{2v}$ double group has degenerate representations – and, indeed, its character table is isomorphic with that of the $C_{4v}$ 'normal' group.

An important application of group theory in spectroscopy arises from the recognition that direct products are at the heart of the topic. Consider the case where two degenerate irreducible representations combine to give a reducible representation (which they invariably do). Of course, the irreducible representations could describe either operators or functions. The irreducible components of the reducible representation will be linear combinations of functions of the type that we have already met, aa, ab, ac and bc, etc. (these functions have the disadvantage that they occur in a direct product of an irreducible representation with itself; however, we seek to illustrate the principle, not to give a general example). The coefficients which relate the basis functions of the starting irreducible representations to the final combinations of product functions do not depend on the particular problem in hand. The so-called vector coupling or 'Clebsch–Gordon' coefficients are universal and so tables of them are available. However, care is needed in their use. This is because for any degenerate irreducible representation an infinite number of choices of basis function is of equal validity and can be used. But the coefficients just introduced are based on a specific choice of basis functions; these same functions must be used in the application of tables of coupling coefficients or errors will surely result (and sometimes the basis functions chosen are not the most obvious – the reason why will become evident in the next paragraph). There are several important outcomes from the use of coupling coefficients. The existence of coupling coefficients means that the final functions are related. If one final (product) function, which could be the intensity of a spectral transition, were available, either by calculation or by measurement, then the values of all of the others follow, based on a so-called 'reduced matrix element' (which is written mathematically with two bars on each side of the operator). This is the Wigner–Eckart theorem. Related to this is the so-called replacement theorem. If the relative values of a set of integrals are determined using coupling coefficients then the same pattern holds for a different set of integrals involving basis functions from the same irreducible representations. The two sets of integrals are proportional to each other. This can be very useful if, for example, one only has an approximate expression for an operator or wavefunction.

It is evident that the more symmetry, the greater the utility of coupling coefficients. Put another way, the greater the inherent degeneracies, the greater their use. The logical consequence of this is that they will be of greatest use in spherical symmetry (and the descent from spherical symmetry to point group symmetry has been described above). In spherical symmetry, that which we have called vector coupling coefficients become replaced by $3j$ symbols ($j$ as in $j$–$j$ coupling) and higher extensions, the $6j$ and $9j$ symbols. These $nj$ symbols have many advantages. First, they are defined in such a way that they are basis-independent. Secondly, they summarize in compact form what would otherwise be lengthy summations. Those practised in the art use them wherever they can!

*See also:* **Tensor Representations.**

## Further reading

Butler PH (1981) *Point Group Symmetry Applications*. New York: Plenum.

Harter WG (1993) *Principles of Symmetry, Dynamics and Spectroscopy*. New York: Wiley.

Heine V (1960) *Group Theory in Quantum Mechanics*. London: Pergamon.

Piepho SB and Schatz PN (1983) *Group Theory in Spectroscopy*. New York: Wiley.

Tsukerblat BS (1994) *Group Theory in Chemistry and Spectoscopy*. London: Academic Press.

Wolbarst AB (1977) *Symmetry and Quantum Systems*. New York: Van Nostrand.

# Tensor Representations

**Peter Herzig**, Universität Wien, Vienna, Austria
**Rainer Dirl**, TU Wien, Vienna, Austria

Copyright © 1999 Academic Press

FUNDAMENTALS OF SPECTROSCOPY
Theory

## Introduction

Tensor representations, synonymous for product representations and their decomposition into irreducible constituents, are useful concepts for the treatment of several problems in spectroscopy. Important examples are the classification of the electronic states in atoms and the derivation of selection rules for infrared absorption or the vibrational Raman or hyper-Raman effect in crystals. In the first case the goal is to reduce tensors which are defined as products of one-particle wave functions, while in the second case tensors for the dipole moment, the electric susceptibility or the susceptibilities of higher orders have to be reduced according to the irreducible representations of the relevant point groups.

## Basics of group theory

### Group postulates

A set of elements $g_1$, $g_2$,... forms a group **G** if there is a composition law defined for the elements such that the following conditions are satisfied:

$$g_i\, g_j = g_k, \qquad g_k \in \mathbf{G} \tag{1}$$

$$(g_i\, g_j) g_k = g_i (g_j\, g_k) \tag{2}$$

$$g_i\, e = e\, g_i = g_i \tag{3}$$

$$g_i\, g_i^{-1} = g_i^{-1}\, g_i = e \tag{4}$$

Equation [1] represents the binary composition law, Equation [2] states the associative law, Equation [3] shows the existence of an identity element, and Equation [4] shows that to each $g \in \mathbf{G}$ there exists a unique inverse element.

### Conjugate elements and classes

For any given group element $g_i \in \mathbf{G}$ the product $gg_ig^{-1}$ can be formed for any arbitrary $g \in \mathbf{G}$ and is called the *conjugate element* of $g_i$ by $g$. This defines an equivalence relation which subdivides the group **G** into mutually disjoint subsets where the latter, symbolized by $[g_i] = \{gg_ig^{-1} | \, g \in \mathbf{G}\}$, are called classes of conjugate elements.

### Products of groups

**Direct product of groups** Consider two groups **L** and **M** and let $\mathbf{L} \times \mathbf{M}$ be the product set which consists of all ordered pairs $\{(l_j, m_k) | l_j \in \mathbf{L}, m_k \in \mathbf{M}\}$. The set $\mathbf{L} \times \mathbf{M}$ defines a *direct product* group, if its composition law is defined by:

$$(l_j, m_k) * (l_r, m_s) = (l_j l_r, m_k m_s) \tag{5}$$

If **L** and **M** are finite groups, then the order $|\mathbf{L} \times \mathbf{M}|$ of the product group is given by $|\mathbf{L} \times \mathbf{M}| = |\mathbf{L}| \cdot |\mathbf{M}|$, where the symbol $|\mathbf{G}|$ denotes the order of some group **G**.

**Kronecker product of groups** The special case $\mathbf{L} = \mathbf{M} = \mathbf{G}$ allows one to define the product group $\mathbf{G} \times \mathbf{G}$ which contains as a special subgroup $\mathbf{G} \boxtimes \mathbf{G} = \{(g,g) | g \in \mathbf{G}\}$ that is isomorphic to the group **G**. The subgroup $\mathbf{G} \boxtimes \mathbf{G}$ is sometimes called the *Kronecker product* group or likewise *diagonal subgroup* of the direct product group $\mathbf{G} \times \mathbf{G}$.

### Representations of groups

**Matrix representations** If for every element $g_i \in \mathbf{G}$ there is a corresponding element $g_i' \in \mathbf{G}'$, such that if $g_i\, g_j = g_k$ in **G**, it follows for the corresponding product of operations $g_i' g_j' = g_k'$ in $\mathbf{G}'$, then the two groups are said to be *homomorphic*. A *representation* of a group is defined as a group of nonsingular square matrices homomorphic with the group. The number of rows (columns) of the matrices is called the *dimension* of the representation. Consider the product $g_i\, g_j = g_k$ in **G**. The analogous product of representation matrices is written as:

$$G(g_i)\, G(g_j) = G(g_k) \tag{6}$$

**Carrier spaces – linear operator representations** A vector space $V$ is called the carrier (or representation) space for the group **G** if there exists a

homomorphic image of **G** of operators $O(\mathbf{G}) = \{O(g)|g \in \mathbf{G}\}$ that leave the space $V$ invariant. The representation $O(\mathbf{G})$ is called the *linear* operator representation $L(\mathbf{G})$ of **G** in $V$ if the operators $O(g) = L(g)$ are linear ones. If $V$ is a unitary space and each operator $L(g) = U(g)$ leaves any scalar product invariant, then $U(\mathbf{G})$ is called a *unitary* operator representation of **G** in $V$.

For instance, the Hilbert space $L^2(\mathbb{R})^3$ is the carrier space for the 3-dimensional rotation group $\mathbf{O}(3, \mathbb{R}) = \mathbf{SO}(3, \mathbb{R}) \times \mathbf{C}_i$ where $\mathbf{C}_i = \{e, i\}$ and the symbol $i$ denotes the inversion operation $(\vec{x} \longmapsto -\vec{x})$ in $\mathbb{R}^3$, respectively. If, using the *Euler* angles to characterize the group elements $g = (\omega) = (\alpha,\beta,\gamma) \in \mathbf{SO}(3, \mathbb{R})$ then the linear operators

$$U(\omega) = \exp(-\mathrm{i}\alpha L_z)\exp(\mathrm{i}\beta L_y)\exp(-\mathrm{i}\gamma L_z) \quad [7]$$

$$[U(i)f](\vec{x}) = f(-\vec{x}) \quad [8]$$

define a unitary operator representation of $\mathbf{O}(3, \mathbb{R})$ on $L^2(\mathbb{R}^3)$, where $f \in L^2(\mathbb{R}^3)$ represents an arbitrary function of the Hilbert space in question and where the operators $L_x$, $L_y$, $L_z$ denote the Cartesian components of the orbital angular momentum operator.

As another example, the Hilbert space $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$ is the carrier space for the direct product group $\mathbf{O}(3, \mathbb{R}) \times \mathbf{SU}(2)$ where the unitary operators $U(\omega,i^\sigma) = U(\omega)U(i^\sigma)$ with $\sigma = 0,1$ that are defined by Equations [7] and [8] represent the orbital part together with the spin part a unitary operator representation of $\mathbf{O}(3, \mathbb{R}) \times \mathbf{SU}(2)$ on the Hilbert space in question. The spin part reads:

$$V(\omega') = \exp(\mathrm{i}\alpha' S_z)\exp(-\mathrm{i}\beta' S_y)\exp(-\mathrm{i}\gamma' S_z) \quad [9]$$

where the entries $S_x$, $S_y$, $S_z$ denote the Cartesian components of the intrinsic (spin) angular momentum which are proportional to the *Pauli* matrices $\sigma_j$, respectively. Here, it should be noted that due to the fact that **SU**(2) represents the universal covering group of $\mathbf{SO}(3, \mathbb{R})$, the range of variation of the group parameters $(\omega') \in [0, 2\pi) \times [0, \pi) \times [0, 4\pi)$ describing **SU**(2)-elements differs from that of $(\omega) \in [0, 2\pi) \times [0, \pi) \times [0, 2\pi)$ describing $\mathbf{SO}(3, \mathbb{R})$ elements. In fact, $\mathrm{Ker}_\varphi = \{(0,0,0), (0,0,2\pi)\}$ where $\varphi(\mathbf{SU}(2)) = \mathbf{SO}(3, \mathbb{R})$ represents the nontrivial kernel of the homomorphism. Apart from this, we have

$$W(\omega, i^\sigma, \omega') = U(\omega, i^\sigma) \otimes V(\omega') \quad [10]$$

where $W(\omega,i^\sigma,\omega')$ denotes an arbitrary element of the unitary operator representation $U(\mathbf{O}(3, \mathbb{R}) \times \mathbf{SU}(2))$ acting on the Hilbert space $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. In this context it should be noted that $L^2(\mathbb{R}^3)$ is used to describe the quantum mechanical motion of a spinless particle in 3 dimensions, whereas $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$ is used as the carrier space to describe the quantum mechanical motion of a particle with spin $\frac{1}{2}$ in 3 dimensions.

**Basis of a representation** Let $\mathcal{H}$ be the carrier space of the group **G** whose elements $g \in \mathbf{G}$ are represented by unitary operators $U(g) \in U(\mathbf{G})$, which is typical when applying group theoretical methods to quantum mechanical problems. Assume that $\dim \mathcal{H} = n$ and consider a set of $n$ linearly independent functions $\varphi_i$ defined in this configuration space with the property:

$$U(g)\varphi_k = \sum_{i=1}^{n} G(g)_{ik}\, \varphi_i \quad [11]$$

This set of functions defines a basis for the representation $G$. It is clear from the last equation that the representation matrix for some particular group operation is completely defined if a basis is given and vice versa. It is therefore convenient to label the representation according to the basis, e.g. $G_\varphi(g)$ in order to stress the basis dependence of multidimensional matrix representations. Note in particular, if $\{\varphi_j|j = 1, 2, 3, \ldots n\}$ defines an orthonormal basis, then the corresponding matrix representation $G_\varphi(\mathbf{G}) = \{G_\varphi(g)|g \in \mathbf{G}\}$ of **G** must be a unitary one.

**Reducibility of representations, irreducible representations** Multidimensional matrix representations of groups are not unique and, if defined via bases of some carrier spaces, sensitively depend on the chosen basis. Any nonsingular linear transformation of the basis $\{\varphi_j|j = 1, 2, 3, \ldots n\}$ to a new basis say $\{\psi_j|j = 1, 2, 3, \ldots n\}$, leads to the following well-known transformation formulae:

$$\psi_k = \sum_{i=1}^{n} C_{ik}\, \varphi_i \quad [12]$$

$$G_\psi(g) = C^{-1}\, G_\varphi(g) C \quad [13]$$

where $C$ is an invertible matrix with coefficients $C_{ij}$ defining a similarity transformation. Note in passing, if $\{\varphi_j\}$ and its counterpart $\{\psi_j\}$ are orthonormal bases of the **G**-invariant carrier space $\mathcal{H}$, then $C$

can be chosen unitary without any loss of generality and the matrix representations $G_\varphi(\mathbf{G})$ and $G_\psi(\mathbf{G})$ are unitary too.

A representation is *reducible* if there is a matrix $C$ which converts all matrices of a representation into the same block-diagonal form. When such a matrix is found the reducible representation is decomposed into a number of representations of dimensions smaller than the original dimension. Representations found in the reduction process that cannot be reduced further are called *irreducible representations*. There are a number of theorems for irreducible representations valid in particular for *finite* groups which shall be given without proof:

1. The number of irreducible representations which are nonequivalent, i.e. not related by a similarity transformation, is equal to the number of classes in the group.
2. The sum over the squares of the dimensions of the irreducible representations is equal to the number of elements in the group $\mathbf{G}$ (*order* of the group, $|\mathbf{G}|$).
3. There is an orthogonality relation between the different irreducible representations of a group (denoted by a Greek letter as a left superscript):

$$\sum_g {}^{\alpha}G(g)^*_{ij}\, {}^{\beta}G(g)_{kl} = (|\, \mathbf{G}\, |\, /\, |{}^{\alpha}G\, |)\delta_{\alpha\beta}\, \delta_{ik}\, \delta_{jl} \qquad [14]$$

where $|{}^{\alpha}G|$ is the dimension of the irreducible representation $\alpha$.

**Characters** Let the matrices ${}^{\alpha}G(g)$ form a (reducible or irreducible) representation of a group. The trace of matrix ${}^{\alpha}G(g)$ is called the *character* of the operation $g_i$ in the representation ${}^{\alpha}G(g)$:

$${}^{\alpha}\chi(g) = \sum_i {}^{\alpha}G(g)_{ii} \qquad [15]$$

An orthogonality relation corresponding to Equation [14] also exists for the characters of irreducible representations:

$$\sum_g {}^{\alpha}\chi(g)^{*\beta}\chi(g) = |\, \mathbf{G}\, |\, \delta_{\alpha\beta} \qquad [16]$$

**Symmetrization of states** One of the most important applications of group theoretical methods in quantum mechanics consists of the task to construct so-called *symmetrized* states which by definition not only form an orthonormalized basis of $\mathbf{G}$-invariant carrier space but also have a peculiar transformation law with respect to the group $\mathbf{G}$. Let $\mathcal{H}$ be a $\mathbf{G}$-invariant carrier space which for the sake of simplicity implies that $U(\mathbf{G})$ forms a unitary operator representation on $\mathcal{H}$ and that there exists an orthonormalized basis $\Phi = \{{}^{\alpha}\phi^s_j | \alpha \in A_G, s = 1, 2, \ldots m_\alpha, j = 1, 2, \ldots |{}^{\alpha}G|\}$ with the following properties:

$$\left\langle {}^{\alpha}\phi^s_j, {}^{\beta}\phi^t_k \right\rangle = \delta_{\alpha\beta}\, \delta_{st}\, \delta_{jk} \qquad [17]$$

$$U(g)\, {}^{\alpha}\phi^s_j = \sum_{k=1}^{|{}^{\alpha}G|} {}^{\alpha}G(g)_{kj}\, {}^{\alpha}\phi^s_k \qquad [18]$$

Here it should be remarked that the symbol $\langle \phi, \psi \rangle$ denotes the scalar product in $\mathcal{H}$ and that $A_G$ symbolizes the index set of the labels which characterize the irreducible $\mathbf{G}$-representations, and that the index $s = 1, 2, \ldots m_\alpha$ indicates the frequency of the irreducible representations ${}^{\alpha}G$ in $\mathcal{H}$, respectively.

Usually, the symmetrization of states is carried out by means of a specific projection method which mainly relies upon the properties of the underlying *group algebra*. Without going into any details, we merely state the form and properties of the so-called *units* of the group algebra. The units of the group algebra are represented in the Hilbert space $\mathcal{H}$ by the following operators:

$$\mathsf{P}^{\alpha}_{jk} = \frac{|{}^{\alpha}G|}{|\mathbf{G}|} \sum_g {}^{\alpha}G(g)^*_{jk}\, U(g) \qquad [19]$$

whose properties are well known. Note that $\mathsf{P}^{\alpha}_{jj}$ are projection operators, whereas $\mathsf{P}^{\alpha}_{jk}$ with $j \neq k$ are shift operators. For further details the reader is referred to the relevant literature. However in this context it is worth emphasizing that there exist some modifications of this method. In particular, it is possible to construct complete sets of commuting operators instead of the units of the group algebra. This method deserves extra attention since it circumvents the representation dependence of the units by solving simultaneous eigenvalue equations.

**Spherical harmonics – irreducible SO(3, ℝ)-representations** The most prominent functions of mathematical physics with applications in many areas of physics and related disciplines are the so-called *spherical harmonics* which form a basis of the Hilbert space $L^2(S)$ where the symbol $S$ denotes the unit sphere. Spherical harmonics are eigenfunctions of the angular momentum operators $L^2$ and $L_z$, respectively.

For integer values of $l$ the spherical harmonics take the following form where, in particular, the Condon–Shortley phase convention has been adopted:

$$Y_l^m(\theta,\phi) = N_{lm} P_{lm}(\theta)\exp(\mathrm{i}m\phi) \quad [20]$$

$$N_{lm} = \mathrm{i}^{m+|m|}\left\{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}\right\}^{1/2} \quad [21]$$

$$P_{lm}(\theta) = \frac{1}{2^l l!}\sin^{|m|}\theta\frac{\mathrm{d}^{l+|m|}(\cos^2\theta-1)^l}{\mathrm{d}(\cos\theta)^{l+|m|}} \quad [22]$$

The group elements $(\omega) \in \mathbf{SO}(3, \mathbb{R})$ are represented in the specific Hilbert space $L^2(S)$ in principle by the same unitary operators as given in Equation [7] though the corresponding Hilbert spaces are different. The matrix elements of the $(2l+1)$-dimensional irreducible representations of $\mathbf{SO}(3, \mathbb{R})$ are given by:

$$U(\omega)Y_{lm} = \sum_{m'} R^l_{m'm}(\omega)Y_{lm'} \quad [23]$$

$$R^l_{m'm}(\omega) = \exp\{-\mathrm{i}(m'\alpha+m\gamma)\}d^l_{m'm}(\beta) \quad [24]$$

$$d^l_{m'm}(\beta) = \{(l+m')!(l-m')!(l+m)!(l-m)!\}^{1/2} \times\sum_k(-1)^{k+m'-m} \times\frac{\cos^{2l-m'+m-2k}(\beta/2)\sin^{2l+m'-m}(\beta/2)}{(l-m'-k)!(l+m-k)!(m'-m+k)!k!} \quad [25]$$

Equation [24] is also valid for the half-integral values $j$ so that direct products of representations needed for the coupling of arbitrary angular momenta can be calculated. However, one should be aware that for half-integral values $j$ the spherical harmonics due to their definition cannot serve as basis for the corresponding $\mathbf{SU}(2)$-representations.

## Direct products of representations

Let $\mathcal{H}^{\alpha\beta}_{12} = \mathcal{H}^\alpha_1 \otimes \mathcal{H}^\beta_2$ be the carrier space of the direct product group $\mathbf{G}\times\mathbf{G}$ where the latter contains $\mathbf{G}\boxtimes\mathbf{G}$ as a diagonal subgroup. Let us assume that $\mathcal{H}^\alpha_1$ is the carrier space for the irreducible representations ${}^\alpha G$ and $\mathcal{H}^\beta_2$ is the carrier space for the irreducible representations ${}^\beta G$, respectively. In other words, we assume symmetrized basis functions ${}^\alpha\varphi_i$, $i=1, 2, \ldots, |{}^\alpha G|$ of $\mathcal{H}^\alpha_1$ and equivalently symmetrized basis functions ${}^\beta\psi_j$, $j=1, 2, \ldots, |{}^\beta G|$ of $\mathcal{H}^\beta_2$, respectively. The group elements $(g_1, g_2) \in \mathbf{G}\times\mathbf{G}$ are represented by the unitary operators $U(g_1,g_2) = U_1(g_1)\otimes U_2(g_2)$ where $U_1(g_1)$ acts nontrivially on $\mathcal{H}_1$ only and similarly $U_2(g_2)$ on $\mathcal{H}_2$, respectively. Now, if all possible products

$$\Psi^{\alpha,\beta}_{ij} = {}^\alpha\varphi_i \otimes {}^\beta\psi_j \quad [26]$$

are formed, then an orthonormal basis of the product space $\mathcal{H}^{\alpha\beta}_{12}$ is defined, since the factor bases $\{{}^\alpha\varphi_i\}$ and $\{{}^\beta\psi_j\}$ are assumed to be symmetrized ones. This allows one to define for the Kronecker product group $\mathbf{G}\boxtimes\mathbf{G}$ a unitary $|{}^\alpha G|\times|{}^\beta G|$-dimensional matrix representation (in general reducible) which is of the following form:

$$U(g,g)\,\Psi^{\alpha,\beta}_{ij} = \sum_{rs} G^{\alpha,\beta}(g)_{rs,ij}\Psi^{\alpha,\beta}_{rs} \quad [27]$$

$$G^{\alpha,\beta}(g) = {}^\alpha G(g)\otimes{}^\beta G(g) \quad [28]$$

This representation is called the *direct product* of the representations ${}^\alpha G$ and ${}^\beta G$. It can be decomposed into the so-called *direct sum* of its irreducible constituents, say ${}^\sigma G$, of $\mathbf{G}$ written symbolically as:

$$G^{\alpha,\beta}(g) = \sum_\sigma \oplus\, a_{\alpha\beta;\sigma}\,{}^\sigma G(g) \quad [29]$$

$$a_{\alpha\beta;\sigma} = |\mathbf{G}|^{-1}\sum_g {}^{\alpha\times\beta}\chi(g)\,{}^\sigma\chi(g)^* = |\mathbf{G}|^{-1}\sum_g {}^\alpha\chi(g){}^\beta\chi(g)\,{}^\sigma\chi(g)^* \quad [30]$$

The coefficients $a_{\alpha\beta;\sigma}$ indicate how often the irreducible representation ${}^\sigma G$ is contained in the direct product (*multiplicity* or *frequency*).

**Clebsch–Gordan coefficients** The canonical basis for the irreducible representation ${}^\sigma G$ in Equation [29] is given by the set of functions ${}^\sigma\Phi^q_k$, $k=1, 2, \ldots$. $|{}^\sigma G|$, where the index $q=1, 2, \ldots, a_\sigma$ indicates how many times ${}^\sigma G$ is contained in ${}^\alpha G\otimes{}^\beta G$. The functions ${}^\sigma\Phi^q_k$ can be expressed as linear combinations of the products ${}^\alpha\varphi_i\otimes{}^\beta\psi_j = \Psi^{\alpha\beta}_{ij}$:

$${}^\sigma\Phi^q_k = \sum_{ij}\Psi^{\alpha\beta}_{ij}\langle\alpha i,\beta j\,|\,\sigma qk\rangle \quad [31]$$

The linear-combination coefficients $\langle\alpha i, \beta j|\sigma qk\rangle$ are commonly called *Clebsch–Gordan coefficients*. The Clebsch–Gordan coefficients form a unitary matrix

$C^{\alpha\beta}$ whose columns are given by the following expressions and are satisfying specific transformation laws which allow on their systematic computation:

$$\left\{\vec{C}^{\alpha\beta}_{\sigma qk}\right\}_{ij} = \langle \alpha i, \beta j | \sigma qk \rangle \quad [32]$$

$$G^{\alpha,\beta}(g)\vec{C}^{\alpha,\beta}_{\sigma qk} = \sum_{m} {}^{\sigma}G(g)_{mk}\vec{C}^{\alpha\beta}_{\sigma qm} \quad [33]$$

Here, it is worth emphasizing that Clebsh–Gordan matrices $C^{\alpha\beta}$ are nonsymmetrically indexed, since their rows are labelled by the pairs $(i, j)$ whereas their columns are labelled by the triplets $(\sigma, q, k)$, respectively. The inverse form of Equation [31] is given by the following expression:

$$\Psi^{\alpha\beta}_{ij} = \sum_{\sigma qk} {}^{\sigma}\Phi^{q}_{k} \langle \sigma qk | \alpha i, \beta j \rangle \quad [34]$$

## 3*nj* Symbols and atomic spectra

### Clebsch–Gordan coefficients and Wigner's 3*j* symbols: coupling of two angular momenta

We now consider the coupling of two angular momenta. For this purpose we have to specify Equation [31] to the case **G** = **SU**(2) which is here written as:

$$|(j_1 j_2) jm\rangle = \sum_{m_1, m_2} |j_1 m_1 j_2 m_2\rangle \langle j_1 m_1 j_2 m_2 | jm \rangle \quad [35]$$

First, it should be noted that the states $|j_1, m_1\rangle = \phi^{\lambda_1}_{j_1,m_1}$ and similarly $|j_2, m_2\rangle = \phi^{\lambda_2}_{j_2,m_2}$ are written in a shorthand notation where the principal quantum numbers $\lambda_1$, $\lambda_2$ are suppressed. Accordingly, the irreducible representation labels are $j_1$ and $j_2$ for the uncoupled representations and $j$ for the coupled representations. The index $m$ $(-j \leq m \leq j)$ numbers the different functions of the basis. For the products of the uncoupled functions $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ we simply write $|j_1 m_1 j_2 m_2\rangle$ for short. The multiplicity index of Equation [31] can be dropped here, because in the special unitary group SU(2) each irreducible representation occurs exactly once in the direct product. It should be noticed that a variety of different symbols are in use in the literature for the Clebsch–Gordan coefficients and different phase conventions exist. Because of

$$m_1 + m_2 = m \quad [36]$$

$$j_1 + j_2 \geq j \geq |j_1 - j_2| \quad [37]$$

the sum (Eqn [35]) is only formally a double sum since the condition (Eqn [36]) must hold and for given values $j_1$ and $j_2$ the range of variation of $j$-values is determined by the triangle inequality (Eqn [37]) respectively.

Finally, we emphasize that the Clebsch–Gordan coefficients for **SU**(2) are directly related to the Wigner's 3$j$ symbols where the latter have the advantage of being more symmetric than the former.

$$\langle j_1 m_1 j_2 m_2 | jm \rangle = (2j+1)^{1/2}(-1)^{-j_1+j_2-m} \times \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{pmatrix} \quad [38]$$

### Calculation and symmetry properties of the 3*j* symbols

The following general formula for the calculation of 3$j$ symbols is due to Racah:

$$\begin{aligned} &\begin{pmatrix} a & b & c \\ \alpha & \beta & \gamma \end{pmatrix} \\ &= (-1)^{a-b-\gamma} F(abc)\{(a+\alpha)!(a-\alpha)!(b+\beta)! \\ &\times (b-\beta)!(c+\gamma)!(c-\gamma)!\}^{1/2} \\ &\times \sum_{k} (-1)^{k}\{k!(a+b-c-k)!(a-\alpha-k)! \\ &\times (b+\beta-k)!(-b+c+\alpha+k)! \\ &(-a+c-\beta+k)!\}^{-1} \end{aligned} \quad [39]$$

$$F(abc) = \left\{ \frac{(-a+b+c)!((a-b+c)!(a+b-c)!}{(a+b+c+1)!} \right\}^{1/2} \quad [40]$$

where the summation over $k$ extends over all values for which all the factorials are defined. It is valid for $\alpha + \beta + \gamma = 0$ and $\Delta(abc) = 1$ (the latter being a compact notation for the triangle condition (Eqn [37]) otherwise the 3$j$ symbols vanish by definition. Much simpler formulae exist for some of the special cases.

The 3$j$ symbols have a high degree of symmetry. Many (but not all) of these symmetries can be seen by using a highly redundant notation which has been introduced initially by Regge:

$$\begin{pmatrix} a & b & c \\ \alpha & \beta & \gamma \end{pmatrix} = \begin{bmatrix} -a+b+c & a-b+c & a+b-c \\ a+\alpha & b+\beta & c+\gamma \\ a-\alpha & b-\beta & c-\gamma \end{bmatrix} \quad [41]$$

The following operations performed on the square-bracket Regge symbol define 72 symmetries of the $3j$ symbol:

(i) Permutations of the columns. The symbol is multiplied by $(-1)^{a+b+c}$ if the permutation is odd.
(ii) Permutations of the rows. As in (i) the symbol is multiplied by $(-1)^{a+b+c}$ if the permutation is odd.
(iii) Transposition about the main diagonal (like the transposition of an ordinary matrix).

### Coupling of three angular momenta

If three angular momenta $\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3$ have to be coupled to the total angular momentum $\mathbf{J}$ there are different ways leading to bases related to each other by a unitary transformation. One possibility is to start from the uncoupled state $|j_1m_1\, j_2m_2\, j_3m_3\rangle$ and to perform first of all the coupling $\mathbf{J}_1 + \mathbf{J}_2 = \mathbf{J}_{12}$ and then $\mathbf{J}_{12} + \mathbf{J}_3 = \mathbf{J}$ leading to a coupled state denoted as $|((j_1j_2)j_{12}j_3)\,jm\rangle$. Another possibility is the coupling scheme $\mathbf{J}_2 + \mathbf{J}_3 = \mathbf{J}_{23} = \mathbf{J}_1 + \mathbf{J}_{23} = \mathbf{J}$ with the corresponding wavefunction $|(j_1(j_2j_3)j_{23})jm\rangle$. With these definitions the following relations between the two types of state functions hold:

$$|((j_1\,j_2)\,j_{12}\,j_3)jm\rangle = \sum_{j_{23}} |(j_1\,(j_2\,j_3)\,j_{23})\,jm\rangle\, V^{(j)}_{j_{23}j_{12}} \qquad [42]$$

$$\begin{Bmatrix} a & b & c \\ d & e & f \end{Bmatrix} = F(abc)F(aef)F(dbf)F(dec) \times \sum_k \left\{ \frac{(-1)^k (k+1)!}{(a+b+d+e-k)!(b+c+e+f-k)!(a+c+d+f-k)!} \times \frac{1}{(k-a-b-c)!(k-a-e-f)!(k-b-d-f)!(k-d-e-c)!} \right\} \qquad [47]$$

$$V^{(j)}_{j_{23}\,j_{12}} = \langle (j_1\,(j_2)\,j_3)\,j_{23})\,jm\,|\,((j_1\,j_2)\,j_{12}\,j_3)jm\rangle \qquad [43]$$

The coefficients $V^{(j)}_{j_{23}\,j_{12}}$ which only depend on the values of the $j$s but are independent of $m$, are called *recoupling coefficients* and can be expressed in terms of Clebsch–Gordan coefficients:

$$V^{(j)}_{j_{23}\,j_{12}} = \sum_{\text{all } m_i} \langle j_2m_2j_3m_3\,|\,j_{23}m_{23}\rangle\langle j_1m_1j_{23}m_{23}\,|\,jm\rangle \times \langle j_1m_1j_2m_2\,|\,j_{12}m_{12}\rangle\langle j_{12}m_{12}j_3m_3\,|\,jm\rangle \qquad [44]$$

Analogous to the introduction of the $3j$ symbols (Eqn [38]) one defines $6j$ symbols:

$$V^{(j)}_{j_{23}\,j_{12}} = \{(2j_{23}+1)(2j_{12}+1)\}^{1/2}(-1)^{j_1+j_2+j_3+j} \times \begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix} \qquad [45]$$

Equation [41] can thus be written as:

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} = \sum_{\substack{\text{all } m_i \\ \text{all } \mu_i}} (-1)^s \left[ \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & l_2 & l_3 \\ m_1 & \mu_2 & -\mu_3 \end{pmatrix} \times \begin{pmatrix} l_1 & j_2 & l_3 \\ -\mu_1 & m_2 & \mu_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & j_3 \\ \mu_1 & -\mu_2 & m_3 \end{pmatrix} \right] \qquad [46]$$

where $s = l_1 + l_2 + l_3 + \mu_1 + \mu_2 + \mu_3$.

### Calculation and symmetry properties of the 6*j*-symbols

With the abbreviation (Eqn [40]) Racah's formula for the $6j$ coefficients reads:

The $6j$ symbols are invariant under interchange of the columns. A further invariance exists with respect to the interchange of any two numbers from the top row with their counterparts in the bottom row, e.g.:

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} = \begin{Bmatrix} j_1 & l_2 & l_3 \\ l_1 & j_2 & j_3 \end{Bmatrix} \qquad [48]$$

The $6j$ symbols are different from zero only if the following four triangle conditions are fulfilled:

$$\Delta(j_1j_2j_3) = \Delta(j_1l_2l_3) = \Delta(l_1j_2l_3) = \Delta(l_1l_2j_3) = 1 \qquad [49]$$

and if the sum of the side lengths of these triangles are integers. Further symmetries of the 6*j* symbols have been found but they are not discussed here.

### 9*j* and 12*j* symbols

The 9*j* symbols are required for the coupling of four angular momenta. An example is the coupling of two orbital and two spin angular momenta either within the Russel–Saunders (*LS*) or the *jj* coupling scheme. The 9*j* symbols can also be used as recoupling coefficients for the transformation from one of the two schemes into the other. They can be expressed in terms of 3*j* or 6*j* symbols. Analogously, the coupling between five angular momenta can be described by the 12*j* symbols for which formulae in terms of 6*j* and 9*j* symbols exist.

## Property tensors and vibrational spectra of crystals

### Property tensors

Like other physical objects, as, for example, elementary particles, atoms and molecules, crystals are characterized by their symmetry which, as has been known now for more than a century, has a determining influence on their physical properties. The underlying symmetry principle, often called 'Neumann–Minnigerode–Curie principle' can, for our purposes, be written as:

$$G_{\text{object}} \subseteq G_{\text{property}} \qquad [50]$$

where $G_{\text{object}}$ is the symmetry group of the object and $G_{\text{property}}$ the symmetry group of the physical property. Therefore the symmetry operations for the crystal are also valid symmetry operations for its physical properties.

On the other hand, crystals are media that are anisotropic, which means that the application of certain causes onto the crystal (e.g. by an electric field) leads to a response or an effect (like the induced polarization) which depends on the orientation of the crystal. Both quantities, the electric field **E** and the electric polarization **P**, are vectors which, in general, point in different directions. For sufficiently low electric fields this leads to a linear relationship between the electric field and the electric polarization, which can be written as tensor equations as follows:

$$\mathbf{P} = \chi\mathbf{E} \qquad [51]$$

$$\begin{aligned} P_x &= \chi_{xx}\, E_x + \chi_{xy}\, E_y + \chi_{xz}\, E_z \\ P_y &= \chi_{yx}\, E_x + \chi_{yy}\, E_y + \chi_{yz}\, E_z \\ P_z &= \chi_{zx}\, E_x + \chi_{zy}\, E_y + \chi_{zz}\, E_z \end{aligned} \qquad [52]$$

$$P_i = \chi_{ij} E_j \qquad [53]$$

Note that Equation [51] represents formally the tensorial relationship, while Equation [52] expresses this relation by the (Cartesian) components of **P** and **E** and some coefficients $\chi_{ij}$, whereas Equation [53] states this relation by using Einstein's summation convention. Here, the coefficients $\chi_{ij}$ are the components of the electric susceptibility tensor which is a tensor of rank 2. The tensor $\chi$ is an example of what is usually called a *property tensor* or *matter tensor*. Strictly speaking, property tensors describe physical properties of the static crystal which belong to the totally symmetric irreducible representation of the relevant point group. Properties, however, that depend on vibrations of the crystal lattice are described by tensors which belong to the different irreducible representations. The corresponding tensors are then often designated as *tensorial covariants*.

If the electric field is not low enough for a linear relationship (Eqn [53]) to hold higher-order terms become important:

$$P_i = \chi^{(1)}_{ij}\, E_j + \frac{1}{2}\chi^{(2)}_{ijk}\, E_k E_j + \frac{1}{6}\chi^{(3)}_{ijkl}\, E_l E_k E_j \; + \cdots \qquad [54]$$

While $\chi^{(1)}$ is important for the Raman effect, the electric susceptibilities of higher order ($\chi^{(2)}$, $\chi^{(3)}$, etc.) are tensors of ranks 3, 4, etc., and are necessary for the description of the hyper-Raman effect.

All tensors can be classified by their parity, i.e. by their behaviour under the symmetry operation of the inversion of the space. Tensors (like vectors) are said to have even parity if they remain unchanged under space inversion and have odd parity if they change sign. In a tensor equation such as Equation [54], the parities on both sides of the equations must be the same. Thus, since **E** and **P** have odd parity, the parity of the susceptibility tensor must be even. Instead of using the term 'parity' one usually designates tensors as *polar tensors* (tensors of even rank having even parity or tensors of odd rank having odd parity) or *axial tensors* (tensors of odd rank having even parity or tensors of even rank having odd parity).

A further criterion for the classification of tensors is their behaviour under time reversal. This is important when one is interested in magnetic properties, for instance in the Raman effect of magnetic crystals. All tensors we are dealing with in this section are assumed to be invariant under time reversal. The latter are also known in the literature as *i*-tensors.

In addition to the above-mentioned classification, a property tensor may have an *intrinsic symmetry*, i.e. a symmetry with respect to the interchange of certain indices which is determined by the symmetry properties of the tensors for the cause and for the effect.

### Transformation properties of tensors of rank [*m*]

Here we consider the *Euclidean* vector space $\mathbb{E}^3 \simeq \mathbb{R}^3$ and assume that the components of an arbitrary vector $\mathbb{R}^3$ are its Cartesian components with respect to a fixed orthonormal basis where the former are denoted by $x_1 = x$, $x_2 = y$, $x_3 = z$, respectively. Accordingly a tensor of rank [$m$] is defined as follows:

$$\left\{\mathsf{T}^{[m]}\right\}_{i_1,i_2,\ldots i_m} = x_{i_1} x_{i_2} \cdots x_{i_m} \quad [55]$$

$$\mathsf{T}^{[m]'} = \mathrm{g}\ \mathsf{T}^{[m]} \quad [56]$$

$$\mathrm{g}\mathsf{T}^{[m]} = \mathbf{M}^{[m]}(g)\ \mathsf{T}^{[m]} \quad [57]$$

$$\mathbf{M}^{[m]}(g) = M(g) \otimes M(g) \otimes \cdots \otimes M(g) \quad [58]$$

$$\left\{\mathsf{T}^{[m]'}\right\}_{i_1,i_2,\ldots i_m} = M_{i_1 j_1}(g) M_{i_2 j_2} \cdots M_{i_m j_m} \left\{\mathsf{T}^{[m]}\right\}_{j_1,j_2,\ldots j_m} \quad [59]$$

Here the entries $\{\mathsf{T}^{[m]}\}_{i_1,i_2,\ldots i_m}$ denote the Cartesian components of the tensor of rank [$m$] where $i_j = 1, 2, 3$ has to be taken into account. Equations [56] to [58] describe the transformation properties of an arbitrary tensor of rank [$m$] with respect to the orthogonal group $\mathbf{O}(3, \mathbb{R})$, respectively. The matrix group $\mathbf{M}^{[m]}$ is the $m$-fold tensor representation of $\mathbf{O}(3, \mathbb{R})$, where $M(g)$ is a real 3-dimensional matrix representation of $g \in \mathbf{O}(3, \mathbb{R})$, which implies that $\mathbf{M}^{[m]}$ defines a real $3^m$-dimensional $\mathbf{O}(3, \mathbb{R})$-representation. Moreover, note that in Equation [59] Einstein's summation convention has been used. Finally, one has to distinguish carefully between *polar* and *axial* tensors of rank [$m$] since their inherent transformation properties with respect to $\mathbf{O}(3, \mathbb{R})$, being stated below, are significantly different.

$$\mathrm{g}\ \mathsf{T}^{[m]}_{\mathrm{polar}} = \mathbf{M}^{[m]}(g)\ \mathsf{T}^{[m]}_{\mathrm{polar}} \quad [60]$$

$$\mathrm{g}\ \mathsf{T}^{[m]}_{\mathrm{axial}} = (\Delta(g))^m\ \mathbf{M}^{[m]}(g)\ \mathsf{T}^{[m]}_{\mathrm{axial}} \quad [61]$$

$$\Delta(g) = \det M(g) \quad [62]$$

Here it is important to notice that $\Delta(g) = \det M(g) = +1$ for proper rotations and $\Delta(\mathrm{g}) = \det M(\mathrm{g}) = -1$ for improper rotations, like for the inversion or reflections. Thus, the factor $(\Delta(g))^m$ occurring in Equation [61] is the $m$-th power of this phase factor and confirms what has been stated in the preceding section as regards the transformation properties of tensors with respect to the space inversion.

### Calculation of property tensors (tensorial covariants)

Property tensors of arbitrary rank [$m$] can be constructed by symmetry adapting their Cartesian components along the lines of the well-known projection method where the group $\mathbf{G}$ is assumed to be a finite subgroup of $\mathbf{O}(3, \mathbb{R})$. In other words, here the symmetry adaptation of the tensors is carried out for point groups and some nontrivial extensions like for magnetic point groups, which we discuss later. The procedure can be divided into several steps. First, the projection operators are defined, and secondly they are applied to the Cartesian tensors, and finally the shift operators are applied to the symmetrized tensors to obtain the corresponding partner tensors. The procedure is summarized as follows:

$$^{[m]}\mathsf{P}^{\gamma}_{ij} = |^{\gamma}G \,||\, \mathbf{G}\,|^{-1} \sum_{g} {}^{\gamma}G(g)^{*}_{ij}\ \mathrm{g} \quad [63]$$

$$^{[m]}\mathsf{X}^{\gamma}_{i} = {}^{[m]}\mathsf{P}^{\gamma}_{ii}\ \mathsf{T}^{[m]} \quad [64]$$

$$^{[m]}\mathsf{X}^{\gamma}_{j} = {}^{[m]}\mathsf{P}^{\gamma}_{ji}\ \mathsf{T}^{[m]} \quad [65]$$

Several remarks are necessary. The units $^{[m]}\mathsf{P}^{\gamma}_{ij}$ of the corresponding group algebra are provided with the superscript [$m$] in order to indicate that their representation is realized in the space of tensors of rank [$m$], respectively. The first step is to use Equation [64] in order to generate just one tensor for the $|^{\gamma}G|$-dimensional irreducible representation. In fact, the number of free parameters must coincide with the

frequency of ${}^{\gamma}G$ in the reducible tensor representation $\mathbf{M}^{[m]}$. The other tensors of a multidimensional tensor basis are obtained from Equation [64] by applying the appropriate shift operators. The given formulae are valid for the calculation of polar and axial tensors, since one merely has to use either Equation [60] for polar tensors or Equation [61] together with Equation [62] for axial tensors, respectively. For the sake of clarity, the explicit forms of the calculated tensors the following notation is used in order to define tensors of rank 2 and rank 3 in matrix form:

$$\mathsf{T}^{[2]} = \begin{bmatrix} xx & xy & xz \\ yx & yy & yz \\ zx & zy & zz \end{bmatrix} \qquad [66]$$

$$\mathsf{T}^{[3]} = \begin{bmatrix} xxx & xxy & xxz \\ xyx & xyy & xyz \\ xzx & xzy & xzz \\ - & - & - \\ yxx & yxy & yxz \\ yyx & yyy & yyz \\ yzx & yzy & yzz \\ - & - & - \\ zxx & zxy & zxz \\ zyx & zyy & zyz \\ zzx & zzy & zzz \end{bmatrix} \qquad [67]$$

Since the components of property tensors are quantities that can be measured experimentally, they have to be real. The results obtained from Equations [64] and [65] are real, if real rotation matrices and real irreducible representations are used. The latter can always be achieved for the multidimensional representations. However, there are 10 crystallographic point groups which have pairs of one-dimensional irreducible representations complex conjugate to each other. In a case like this, one first computes pairs of conjugate complex tensors and then forms two real linear combinations for each conjugate pair as illustrated below for $\mathbf{C}_3$ (or 3 in Hermann–Mauguin notation).

**Polar tensors of rank 2 for group $\mathbf{C}_3$** The characters for $\mathbf{C}_3$ are shown in **Table 1** and the required rotation matrix for the operation $C_3^+$ can be taken from **Table 4** (the matrix for $C_3^-$ is the square of the matrix for $C_3^+$). With these data the polar tensors of rank 2 can be calculated. Using the abbreviations $a = xx + yy$, $b = xy - yx$, $c = zz$, $d = \frac{1}{2}(xx - yy)$, $e = \frac{1}{2}(xy - yx)$, $f = xz$, $g = yz$, $h = zx$ and $j = zy$, the resulting tensors that are complex for the representations ${}^1E$ and ${}^2E$ are displayed in **Table 2**. Adding the two tensors for ${}^1E$ and ${}^2E$ gives the first real tensor and multiplying the one for ${}^1E$ with –i and the one for ${}^2E$ with i and adding them together gives the second real tensor for the 'physically irreducible representation' $E$, as it is sometimes called (see **Table 3**).

**A note on magnetic point groups** In order to describe consistently some properties of magnetic systems that are related to their symmetry, the concept of ordinary point groups was extended, to so-called *magnetic* point groups. In order to cope these types of problems a new operation has been introduced designated as *antisymmetry operation*. The latter does not affect the space coordinates but only reverses the sign of the magnetic moment at each point in space. Instead of ordinary point groups one has to deal with magnetic point groups (colour groups) in which some of the ordinary point-group operations appear in combination with the antisymmetry operation. To summarize, one can also construct symmetry adapted tensors for magnetic point groups, where the crucial difference consists of a

**Table 1** Character table for the group $\mathbf{C}_3$

| $\mathbf{C}_3$ (Eqn [3]) | $E$ | $C_3^+$ | $C_3^-$ |
|---|---|---|---|
| $A$ | 1 | 1 | 1 |
| ${}^1E$ | 1 | $\varepsilon^*$ | $\varepsilon$ |
| ${}^2E$ | 1 | $\varepsilon$ | $\varepsilon^*$ |

$\varepsilon = \exp(2\pi i/3)$

**Table 2** Tensors of rank 2 for the group $\mathbf{C}_3$

| $A$ | ${}^1E$ | ${}^2E$ |
|---|---|---|
| $\begin{bmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & c \end{bmatrix}$ | $\frac{1}{2}\begin{bmatrix} d+ie & e-id & f-ig \\ e-id & -d-ie & g+if \\ h-ij & j+ih & 0 \end{bmatrix}$ | $\frac{1}{2}\begin{bmatrix} d-ie & e+id & f+ig \\ e+id & -d-ie & g-if \\ h+ij & j-ih & 0 \end{bmatrix}$ |

**Table 3** Real Raman tensors for the group $\mathbf{C}_3$

| $A$ | $E$ | |
|---|---|---|
| $\begin{bmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & c \end{bmatrix}$ | $\begin{bmatrix} d & e & f \\ e & -d & g \\ h & j & 0 \end{bmatrix}$ | $\begin{bmatrix} e & -d & -g \\ -d & -e & f \\ -j & h & 0 \end{bmatrix}$ |

**Table 4** Generators for the groups $\mathbf{D}_3$,$\mathbf{D}_{3v}$,$\mathbf{D}_{3d}$

| $C_{3z}^+$ | $C_{2x}$ | $\sigma_x$ | $i$ |
|---|---|---|---|
| $\begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ | $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ |

certain modification of the transformation formulae (Eqns [60] to [62]).

**Raman and hyper-Raman tensors for Laue class $\bar{3}$m** Here, the Raman and the hyper-Raman tensors for trigonal lattices shall be given. The property tensor for the Raman effect is the derivative of the polarizability with respect to the normal coordinate and is a polar second rank *i*-tensor which is symmetric if the resonant Raman effect or a degenerate ground state are excluded while the corresponding tensor for the hyper-Raman effect is a polar *i*-tensor of rank three the internal symmetry of which is dependent on the experimental conditions.

The matrix generators for the groups belonging to this Laue class, namely $\mathbf{D}_3$ (32), $\mathbf{C}_{3v}$ ($3m$) and $\mathbf{D}_{3d}$ ($\bar{3}m$) are shown in **Table 4** and the representation matrices (in real form) for the generators in **Table 5** (the information for $\mathbf{D}_{3d}$ can easily be generated by forming the direct product $\mathbf{D}_{3d} = \mathbf{D}_3 \times \mathbf{C}_i$ where $\mathbf{C}_i = \{E, i\}$).

The obtained tensors (here given without any intrinsic symmetry) are displayed in **Table 6** where for the tensors of rank 2 the two numbers on the right-hand side of the matrices give the numbers of independent components of the tensor without any intrinsic symmetry and for the symmetric tensor, respectively. Summing these numbers for all irreducible representations must give 9 (the number of components of a general tensor of rank 2 without intrinsic symmetry) and 6 (the number of components of a symmetric tensor of rank 2). For the tensors of rank 3 the numbers of independent components without intrinsic symmetry, for intrinsic symmetry with respect to the interchange of two indices and for the totally symmetric tensor are indicated in a similar fashion. The corresponding sums over the irreducible representations must yield 27, 18 and 10 in this case taking into account the required intrinsic symmetries. In **Table 7** the Raman and hyper-Raman tensors for the Laue class $\bar{3}m$ are displayed.

**Table 5** Irreducible representations for the generators

| $\mathbf{D}_3$ | $\mathbf{C}_{3z}^{+}$ | $\mathbf{C}_{2x}$ | $\mathbf{C}_{3v}$ | $\mathbf{C}_{3z}^{+}$ | $\sigma_x$ |
|---|---|---|---|---|---|
| $A_1$ | 1 | 1 | $A_1$ | 1 | 1 |
| $A_2$ | 1 | –1 | $A_2$ | 1 | –1 |
| $E$ | $\frac{1}{2}\begin{bmatrix}-1 & -\sqrt{3}\\ \sqrt{3} & -1\end{bmatrix}$ | $\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$ | $E$ | $\frac{1}{2}\begin{bmatrix}-1 & -\sqrt{3}\\ \sqrt{3} & -1\end{bmatrix}$ | $\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$ |

**Table 6** Tensors

$$\mathbf{P}_2 = \begin{bmatrix} a & 0 & 0\\ 0 & a & 0\\ 0 & 0 & b \end{bmatrix}\begin{matrix}2\\2\end{matrix} \qquad \mathbf{Q}_2 = \begin{bmatrix} 0 & c & 0\\ -c & 0 & 0\\ 0 & 0 & 0 \end{bmatrix}\begin{matrix}1\\0\end{matrix}$$

$$\mathbf{R}_2^{(1)} = \begin{bmatrix} d & 0 & 0\\ 0 & -d & e\\ 0 & f & 0 \end{bmatrix}\begin{matrix}3\\2\end{matrix} \qquad \mathbf{R}_2^{(2)} = \begin{bmatrix} 0 & -d & -e\\ -d & 0 & 0\\ -f & 0 & 0 \end{bmatrix}\begin{matrix}3\\2\end{matrix}$$

$$\mathbf{P}_3 = \begin{bmatrix} a & 0 & 0\\ 0 & -a & b\\ 0 & c & 0\\ - & - & -\\ 0 & -a & -b\\ -a & 0 & 0\\ -c & 0 & 0\\ - & - & -\\ 0 & d & 0\\ -d & 0 & 0\\ 0 & 0 & 0 \end{bmatrix}\begin{matrix}4\\2\\1\end{matrix} \qquad \mathbf{Q}_3 = \begin{bmatrix} 0 & e & f\\ e & 0 & 0\\ g & 0 & 0\\ - & - & -\\ e & 0 & 0\\ 0 & -e & f\\ 0 & g & 0\\ - & - & -\\ h & 0 & 0\\ 0 & h & 0\\ 0 & 0 & i \end{bmatrix}\begin{matrix}5\\4\\3\end{matrix}$$

$$\mathbf{R}_3^{(1)} = \begin{bmatrix} j+k+l & 0 & 0\\ 0 & j & m\\ 0 & n & p\\ - & - & -\\ 0 & k & m\\ l & 0 & 0\\ n & 0 & 0\\ - & - & -\\ 0 & q & r\\ q & 0 & 0\\ s & 0 & 0 \end{bmatrix}\begin{matrix}9\\6\\3\end{matrix} \qquad \mathbf{R}_3^{(2)} = \begin{bmatrix} 0 & l & m\\ k & 0 & 0\\ n & 0 & 0\\ - & - & -\\ j & 0 & 0\\ 0 & j+k+l & -m\\ 0 & -n & p\\ - & - & -\\ q & 0 & 0\\ 0 & -q & r\\ 0 & s & 0 \end{bmatrix}\begin{matrix}9\\6\\3\end{matrix}$$

**Table 7** Raman (RT) and hyper-Raman tensors (HRT) for Laue class

| *Point group* | *Rep.* | *RT* | *Rep.* | *HRT* |
|---|---|---|---|---|
| $\mathbf{D}_{3d}$ ($\bar{3}m$) | $A_{1g}$ | $\mathbf{P}_2$ | $A_{1u}$ | $\mathbf{P}_3$ |
| | $A_{2g}$ | $\mathbf{Q}_2$ | $A_{2u}$ | $\mathbf{Q}_3$ |
| | $E_g$ | $\mathbf{R}_2^{(1)}$, $\mathbf{R}_2^{(2)}$ | $E_u$ | $\mathbf{R}_3^{(1)}$, $\mathbf{R}_3^{(2)}$ |
| $\mathbf{D}_3$ (32) | $A_1$ | $\mathbf{P}_2$ | $A_1$ | $\mathbf{P}_3$ |
| | $A_2$ | $\mathbf{Q}_2$ | $A_2$ | $\mathbf{Q}_3$ |
| | $E$ | $\mathbf{R}_2^{(1)}$, $\mathbf{R}_2^{(2)}$ | $E$ | $\mathbf{R}_3^{(1)}$, $\mathbf{R}_3^{(2)}$ |
| $\mathbf{C}_{3v}$ ($3m$) | $A_1$ | $\mathbf{P}_2$ | $A_1$ | $\mathbf{Q}_3$ |
| | $A_2$ | $\mathbf{Q}_2$ | $A_2$ | $\mathbf{P}_3$ |
| | $E$ | $\mathbf{R}_2^{(2)}$, $\mathbf{R}_2^{(1)}$ | $E$ | $\mathbf{R}_3^{(1)}$, $\mathbf{R}_3^{(2)}$ |

## Selection rules

In order to determine whether a transition between given initial and final states is allowed or forbidden, one only has to know whether the following scalar product is zero or not:

$$\langle \Psi_i^{\alpha}, O_k^{\gamma} \Psi_j^{\beta} \rangle = \int_{\mathbb{R}^3} \mathrm{d}^3 x \, \Psi_i^{\alpha}(\vec{x})^* O_k^{\gamma} \Psi_j^{\beta}(\vec{x}) \qquad [68]$$

where the operator $O_k^{\gamma}$ is the electric dipole moment, the electric susceptibility or the electric susceptibility

of second order, depending on whether infrared absorption, Raman or hyper-Raman scattering is to be investigated and $\Psi_i^\alpha$ and $\Psi_j^\beta$ are the initial and final state, respectively. The integral (Eqn [68]) is different from zero only if the totally symmetric irreducible representation is contained in the direct product representation ${}^\alpha G^* \otimes {}^\beta G \otimes {}^\gamma G$. According to Equation [30] this is the case if the corresponding frequency number

$$a_{\gamma\beta:\alpha} = | G |^{-1} \sum_g {}^\alpha\chi(g)^{*\gamma}\chi(g)^\beta\chi(g) \qquad [69]$$

is greater than zero. Often, $\Psi_j^\alpha$ belongs to the totally symmetric irreducible representation in which case its characters are all equal to 1.

In order to obtain the characters for the (reducible) representation to which the operator $O_k^\gamma$ belongs one only has to consider matrices of the form:

$$M(g) = \begin{bmatrix} \cos\phi_g & -\sin\phi_g & 0 \\ \sin\phi_g & \cos\phi_g & 0 \\ 0 & 0 & \pm 1 \end{bmatrix} \qquad [70]$$

where the positive sign holds for proper operations (pure rotations about the $z$ axis) and the negative sign for improper operations (rotations combined with reflections about the $xy$ plane).

For infrared absorption $O_k^\gamma$ is the operator for the electric dipole moment which is a polar vector transforming like the Cartesian coordinates $x$, $y$, $z$. The corresponding character is given by the trace of the matrix (Eqn [70]).

$$\chi_{\mathrm{IR}}(g) = 2\cos\phi_g \pm 1 \qquad [71]$$

For the Raman effect the operator is that for the electric polarizability which is a polar tensor of rank 2. Without any intrinsic symmetry the character is the square of the trace of the matrix (Eqn [70]). However, since we assume a symmetric tensor here, the character becomes:

$$\chi_{\mathrm{R}}(g) = 2\cos\phi_g(2\cos\phi_g \pm 1) \qquad [72]$$

For the hyper-Raman effect we only consider the case of a totally symmetric polar tensor of rank 3. Under this assumption the character is:

$$\chi_{\mathrm{HR}}(g) = 2\cos\phi_g(4\cos^2\phi_g \pm 2\cos\phi_g - 1) \qquad [73]$$

With this information and only using group character tables one can easily find out to which symmetry species $\psi_i^\alpha$ may belong in order that a transition is observed.

There is a general rule that follows from the symmetry considerations above, called *mutual exclusion rule*: ‘In groups having a centre of symmetry the modes active in Raman scattering are inactive in infrared absorption and vice versa’. An explanation for this behaviour is that the property tensor for infrared absorption is a polar vector and for Raman scattering a polar tensor of rank 2. While the parity under space inversion of the former is odd and thus belongs to an *ungerade* representation of the point group, it is even for the latter and belongs to a *gerade* representation.

In Raman spectroscopy the relative intensity of the scattered light is straightforwardly calculated from the appropriate Raman tensor as follows: If $\mathbf{v}_i$ is a unit vector defining the polarization of the incident laser radiation and $\mathbf{v}_s$ a unit vector characterizing the polarization of the scattered light, then the following proportionality for the intensity $I$ of the total scattered radiation holds:

$$I \sim | \mathbf{v}_i \cdot \boldsymbol{\chi} \cdot \mathbf{v}_s |^2 \qquad [74]$$

where the inner product has to be formed between the three tensor quantities on the right-hand side. With a suitable experimental arrangement the components of the Raman tensor can thus be measured independently.

*See also:* **Atomic Absorption, Theory; IR Spectroscopy, Theory; Nonlinear Raman Spectroscopy, Applications; Nonlinear Raman Spectroscopy, Instruments; Nonlinear Raman Spectroscopy, Theory; Rotational Spectroscopy, Theory; Symmetry in Spectroscopy, Effects of; Vibrational, Rotational and Raman Spectroscopy, Historical Perspective.**

## Further reading

Altmann SL and Herzig P (1994) *Point-Group Theory Tables.* Oxford: Clarendon Press.

Brandmüller J, Illig D and Herzig P (1999) *Symmetry and Physical Properties of Matter. Rank 1, 2, 3 and 4 property tensors for the irreducible representations of the classical and magnetic, crystallographic and non-crystallographic point groups.* IVSLA Series, Vol. 2, Amsterdam: IOS Press.

Brandmüller J and Winter FX (1985) Influence of symmetry on the static and dynamic properties of crystals. Calculation of sets of the Cartesian irreducible tensors

for the crystallographic point groups. *Zeitschrift für Kristallographie* **172**: 191–232.
Chaichian M and Hagedorn R (1998) *Symmetries in Quantum Mechanics. From Angular Momentum to Supersymmetry*. Bristol and Philadelphia: Institute of Physics Publishing.
Chen JQ (1989) *Group Representation Theory for Physicists*. Singapore: World Scientific.
Claus R, Merten L and Brandmüller J (1975) Light Scattering by Phonon-Polaritons. *Springer Tracts in Modern Physics* **75**: 1–237.
Condon EU and Odabaşı H (1980) *Atomic Structure*. Cambridge: Cambridge University Press.
Joshua SJ (1991) *Symmetry Principles and Magnetic Symmetry in Solid State Physics*. Bristol, Philadelphia and New York: Adam Hilger.
Poulet H and Mathieu JP (1976) *Vibration Spectra and Symmetry of Crystals*. New York: Gordon and Breach.
Rotenberg M, Bivins R, Metropolis N and Wooten JK (1959) *The 3-j and 6-j symbols*. Cambridge, Massachusetts: MIT Press.
Wigner EP (1959) *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*. New York: Academic Press.

# Thallium NMR, Applications

*See* **Heteronuclear NMR Applications (B, Al, Ga, In, Tl).**

# Thermospray Ionization in Mass Spectrometry

**WMA Niessen**, hyphen MassSpec Consultancy, Leiden, The Netherlands

Copyright © 1999 Academic Press

| MASS SPECTROMETRY |
|---|
| **Methods & Instrumentation** |

## Introduction

Thermospray ionization is a soft ionization technique, applicable with a thermospray interface for combining liquid chromatography and mass spectrometry (LC-MS). The thermospray interface was developed by Vestal and co-workers at the University of Houston (TX) and subsequently commercialized by Vestal in the company Vestec (Houston, TX). The thermospray interface was the first LC-MS interface, where the analyte ionization is an integral part of the introduction of the column effluent into the mass spectrometer. Between 1987 and 1992, thermospray interfacing and ionization was the most widely used strategy for LC-MS coupling. After 1992, its use diminished in favour of interfacing strategies based on atmospheric-pressure ionization, i.e. electrospray and atmospheric-pressure chemical ionization (APCI).

In a thermospray interface, a jet of vapour and small droplets is formed by heating the column effluent of an LC column or any other continuous liquid stream in a heated vaporizer tube. Nebulization takes place as a result of the disruption of the liquid by the expanding vapour that is formed upon evaporation of part of the liquid in the tube. A considerable amount of heat is transferred to the solvent prior to the onset of the partial inside-tube evaporation. This assists in the desolvation of the droplets in the lower pressure region. By applying efficient pumping directly at the ion source, up to 2 mL min$^{-1}$ of aqueous solvents can be introduced into the MS vacuum system. The ionization of the analytes takes place by mixed mechanisms based on gas-phase ion–molecule

reactions and ion evaporation processes. The reagent gas for ionization can be made either in a conventional way using energetic electrons from a filament or discharge electrode, or in a process called thermospray buffer ionization, where the volatile buffer dissolved in the eluent is involved. Thermospray interfacing and ionization as well as its applications in LC-MS have been reviewed in two excellent review papers by Arpino and in two extensive book chapters.

## Thermospray interface

The thermospray interface is the result of a long-term research project between 1978 and 1984, aimed at the development of an LC-MS interface which is compatible with a flow-rate of 1 mL min$^{-1}$ of an aqueous mobile phase and is capable of providing both electron ionization and chemical ionization mass spectra. Initially, the two most important research topics were related to the ability to (1) achieve a very rapid heating and subsequent vaporization of the column effluent, and to (2) achieve sufficient vacuum conditions to successfully perform analyte ionization and mass analysis, while introducing large amounts of liquid vapour, i.e. the equivalent of 1 mL min$^{-1}$ aqueous solvents. The developments with respect to the various heating systems investigated for the rapid vaporization of the mobile phase from the LC column are summarized in **Table 1**. The highly complex setup of the first prototype, featuring laser vaporization of the mobile phase and an extensive vacuum system containing orthogonal quadrupole analysers, was simplified in subsequent interface designs. The heat required for the evaporation of the 1 mL min$^{-1}$ aqueous mobile phase could, instead of using an expensive laser system, also be provided with an electrically heated vaporizer capillary (see **Table 1**). In addition, the vacuum system could be significantly simplified by connecting a mechanical rotary pump directly to the ion source block. The highly directed flow of the liquid vapour jet from the thermospray vaporizer considerably enhances the pumping efficiency of this pump.

**Table 1** Characteristics of various heating systems investigated in the development of the thermospray interface

| *Heat supply* | *Heated length* (mm) | *Energy flux* (W cm$^{-2}$) | *Total power* (W) |
|---|---|---|---|
| $CO_2$ laser beam focused on liquid jet | 0.3 | 30 000 | 25 |
| Hydrogen flames to heat a copper cylinder at the capillary exit | 3 | 5 000 | 50 |
| Indirect electrically heated capillary | 30 | 700 | 100 |
| Direct electrically heated capillary | 300 | 70 | 150 |

A commercial thermospray interface consists of a direct-electrically heated vaporizer type, mostly fitted into a spray probe, a heated source block featuring a filament, a discharge electrode, a repeller electrode and an ion-sampling cone to the mass analyser, and the exhaust pump outlet. A schematic diagram of a typical thermospray system is shown in **Figure 1**. The temperature of the vaporizer tube must be accurately controlled in order to ascertain the partial liquid evaporation, required for successful thermospray ionization. Automatic compensation for changes in the solvent composition during gradient elution should be incorporated. A liquid nitrogen trap is positioned in the exhaust line between the source block and the rotary pump in order to avoid contamination of the pump oil by solvent used in LC-MS. Obviously, minor instrumental differences with respect to vaporizer design, temperature control and source block design are present between the various commercial systems. Two types of thermospray vaporizers have been in use, i.e. the Vestec-type vaporizer where the temperature control is based on measuring temperatures both at the stem near the solvent entrance and at the tip, where the nebulization is complete, and the Finnigan-type vaporizer where a thermocouple is spot-welded close to the inlet side at approximately one-quarter of the heated length.

A schematic representation of the thermospray nebulization process is shown in **Figure 2**. Initially, in the first part of the vaporizer tube, the liquid is heated until, at a certain stage, the onset of vaporization takes place. The vaporization process will start at the heated capillary walls and results in tearing of the liquid: bubbles are formed within the liquid.

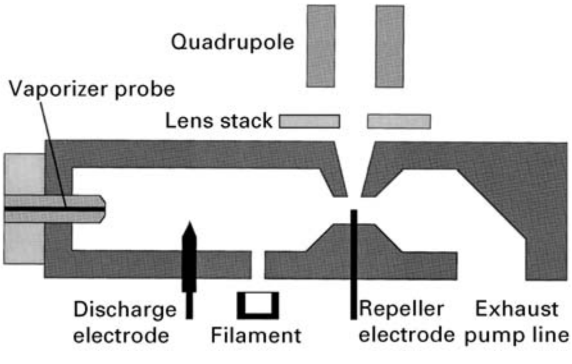



**Figure 1** Schematic diagram of a thermospray interface and ion source.

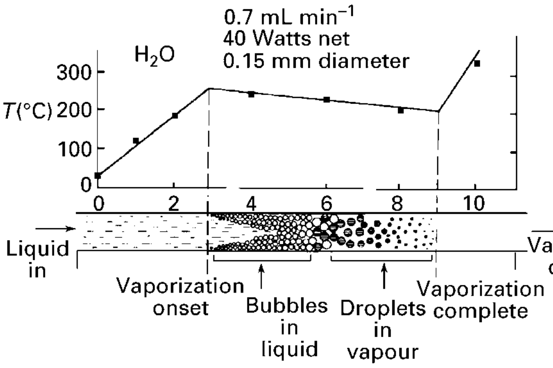



**Figure 2** Schematic representation of the thermospray vaporization process. Reprinted with permission from Vestal ML and Fergusson GJ (1985) *Analytical Chemistry* **57**: 2373–2378, © 1985, American Chemical Society.

Upon continuing vaporization, the stage of bubbles in the liquid transforms to liquid droplets in a vapour. The temperature measured at the vaporizer tube wall over the length of the capillary is also shown in **Figure 2**. When complete solvent evaporation inside the tube would be achieved, a sharp increase of the capillary wall temperature would be observed, where the vapour is heated. However, optimum ionization conditions are achieved at nearly complete inside-tube vaporization. From this description of the nebulization process, it may be concluded that the contact time between the liquid and the analyte molecules dissolved in the liquid and the hot surface of the capillary is relatively short. This limits the extent of thermal decomposition of labile analytes.

Most thermospray interfaces have been fitted onto (triple) quadrupole mass analysers, although thermospray interfaces for magnetic sector instruments were commercially available as well.

## Thermospray ionization modes

The thermospray interface can be used in various modes of ionization, depending on the settings of experimental parameters and the choice of the solvent composition. The thermospray nebulization process provides for a rapid and efficient means to partially evaporate the solvent mixtures introduced into the system by means of the production of small heated droplets. For clarity of the discussion, four ionization modes are distinguished here, i.e. two electron-initiated modes and two liquid-based ionization modes.

The two electron-initiated ionization modes are filament-on and discharge-on ionization. In these modes, the thermospray interface is used as a solvent introduction device, providing nebulization and soft transfer of analytes from the liquid to the gas phase. High-energy electrons are generated by means of a heated filament or at a corona discharge electrode. These electrons produce molecular ions of solvent molecules in the high-pressure (typically 1 kPa) source. In a series of ion–molecule reactions, the solvent molecular ions are converted to solvent-based reagent gas ions, i.e. protonated molecules and clusters, similar to the processes in a chemical ionization source. Protonated $[M+H]^+$ or deprotonated $[M-H]^-$ analyte molecules are produced in positive-ion and negative-ion mode, respectively, as a result of gas-phase proton-transfer reactions, while various other even-electron ionic species (adducts such as the ammoniated molecule $[M+NH_4]^+$, $[M+CH_3OH+H]^+$, or $[M+CH_3COO]^-$) may be produced as well.

The two liquid-based ionization modes are based on ion evaporation processes, initially proposed by Iribarne and Thomson. The mechanism can be summarized as follows: During thermospray nebulization, a superheated mist carried in a supersonic vapour jet is generated. Nonvolatile molecules are preferentially retained in the droplets, which are charged due to the statistical random sampling of the buffer ions in solution. As a result of continuous solvent evaporation from the droplets and repeated droplet breakdown by Rayleigh instabilities, a high local field strength is generated allowing charged species to desorb or evaporate from the droplets. These charged species comprise analyte molecules, present as preformed ions in solution, and buffer-solvent cluster ions, that rapidly equilibrate with the vapour in the ion source. Ion–molecule reactions may occur between the ions and neutrals in the source. A schematic illustration of the thermospray ionization mechanism is provided in **Figure 3**.

The ion evaporation mechanism in thermospray ionization has been criticized, by, among others, Röllgen and co-workers, who propose an alternative model, i.e. the charge residue or soft desolvation model. According to this model, the preformed ion of the nonvolatile analyte molecule is kept in a droplet, which decreases in size due to continuous solvent evaporation and repetitive Rayleigh instabilities until the droplet has become so small that it can be considered as a solvated ion. The ionization is thus the result of soft desolvation of the preformed analyte ions. Interestingly, the discussion between these two mechanisms reappears in the discussion on electrospray ionization, although in a slightly different manner.

Irrespective of the exact mechanism, ion evaporation or soft desolvation, it is important to pursue the generation of preformed ions in solution, i.e. by adjusting pH, and to reduce the influence of

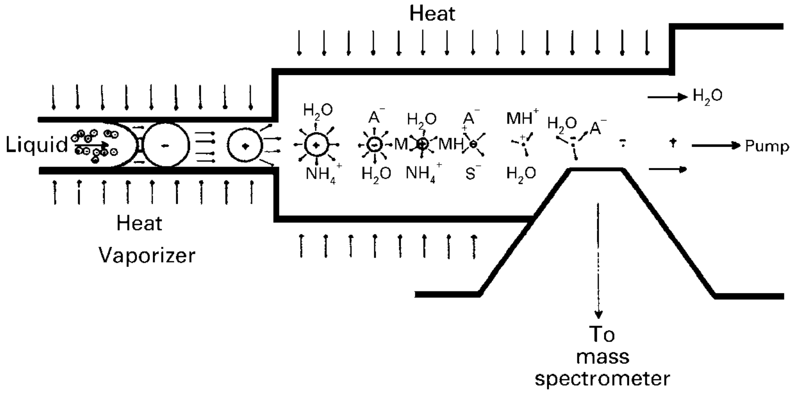



**Figure 3** Schematic illustration of the liquid-based thermospray ionization modes. Reprinted with permission from Vestal ML (1983) *International Journal of Mass Spectrometry and Ion Physics* **46**: 193–196, © 1983, Elsevier Science.

competitive ions, i.e. to keep the ionic strength of other ions as low as possible. Surprisingly, the latter is not a common practice in thermospray ionization. The most widely applied liquid-based thermospray ionization mode appears to perform best in the presence of rather high (0.05–0.2 mol $L^{-1}$) concentrations of ammonium acetate or formate. Under these conditions, the evaporation of solvated ammonium ions generally will be more effective than the ion evaporation of preformed analyte molecules, especially because the latter are present in significantly lower concentrations. As a result, gas-phase ion–molecule reactions between ion-evaporated ammonium ions and neutral analyte molecules will significantly contribute to the ionization yield in most thermospray applications. Therefore, two liquid-based ionization strategies are indicated and discriminated here, i.e. one based on ion evaporation of preformed analyte ions in solution (thermospray ion evaporation mode), and one based on ion evaporation of solvent buffer ions followed by gas-phase ion–molecule reactions with neutral analyte species, efficiently transferred to the gas phase by means of nebulization and subsequent droplet evaporation, to produce protonated or deprotonated analyte ions (thermospray buffer ionization mode).

It must be emphasized that the resulting ions from either mechanism are the same. Therefore, it appears difficult to discriminate between the various mechanisms, especially because a mixed ionization mode, where various processes contribute to the final mass spectrum, is most likely.

Although the ion evaporation mechanism is the most popular view on the thermospray ionization mechanism, the ionization characteristics under typical operating conditions, and for most analytes, are best understood in terms of chemical ionization. Although for particular compounds differences between the filament-on and discharge-on modes were observed, in general these two modes can be treated in the same way.

In both modes, analyte ionization is due to a gas-phase ion–molecule reaction between analyte molecules and reagent gas ions. The latter are generated from the solvent vapour in the high-pressure ion source by means of energetic electrons, basically similar to the generation of any other reagent gas in a source for chemical ionization.

The reagent gas composition is determined by the composition of the solvent mixture or mobile phase introduced. The reagent gas mass spectrum is often quite complex, containing several solvent cluster ions, but is generally dominated by the ionic species derived from the component in the solvent mixture with the highest proton affinity (in positive-ion mode) or the lowest gas-phase acidity (in negative-ion mode), although concentration effects may play a role as well. Proton affinities (gas-phase basicity) of common mobile-phase constituents are given in **Table 2**, while gas-phase acidities are given in **Table 3**. From these data it can be concluded that in positive-ion mode the reagent gas due to a 50:50 water-methanol mixture is dominated by methanol-related ions, e.g. at *m/z* 33, 65 and 97 due to $[(CH_3OH)_n + H]^+$ with $n = 1$, 2 and 3, respectively. After addition of ammonium acetate to this solvent mixture, the most abundant reagent gas ions are ammonium related ions, e.g. at *m/z* 50, 78 and 110, due to $[NH_4.CH_3OH]^+$, $[NH_4.CH_3COOH]^+$ and $[NH_4.CH_3OH.CH_3COOH]^+$, respectively.

In positive-ion mode, analyte ionization to a protonated molecule may take place when the proton affinity of the analyte exceeds that of the reagent gas. Typical values of proton affinities for a number of

monofunctional analytes are given in **Table 2**. For multifunctional analytes, the proton affinity is roughly determined by the proton affinity of the functional group with the highest proton affinity. From **Table 2**, it may be concluded that the solvent mixture without ammonium acetate has a wider applicability range. In practice, however, ammonium acetate is added to the mobile phase in over 80% of the applications with filament-on or discharge-on modes, partly because of the need to apply a buffer in order to achieve reproducible retention times in LC. Next to protonated molecules $[M+H]^+$ a variety of adduct ions may be generated. When the proton affinity of the analyte is within ~30 kJ $mol^{-1}$ of that of the reagent gas, adduct ions, e.g. $[M+NH_4]^+$, may be found. Furthermore, a series of solvent cluster ions may be observed, generally with low intensity. Maeder elaborately studied the various ions observed in thermospray ionizations and proposed a general formula:

$$[\mathrm{M} + \mathrm{A} + x\mathrm{B} - y\mathrm{C}]^+$$

where M is the analyte molecule, A is the attached cation, e.g. the proton or ammonium ion, B is an attached solvent molecule, C is an eliminated molecule, e.g. water, and *x* and *y* take integer values of 0, 1, 2, … . The presence of adduct ions next to the protonated molecule may be useful to ascertain the molecular-mass determination.

In the negative-ion mode, analyte ionization to a deprotonated molecule $[M–H]^-$ may take place when the gas phase acidity of the reagent gas exceeds that of the analyte, while similarly adduct ions may be observed as well. Typical values of the gas-phase acidity for a number of monofunctional analytes are given in **Table 3**.

These same ionization rules can be applied to predict the ionization behaviour of compounds in thermospray buffer ionization, where the analyte ionization is primarily dependent on the gas-phase ion–molecule reaction with ion-evaporated buffer ions, i.e. ammonium and acetate or formate.

In all instances, soft ionization of the analyte molecules is achieved, i.e. generally little fragmentation is observed. Obviously, there are a number of parameters other than the solvent composition that determine the ionization behaviour, e.g. analyte properties, temperatures, pressure. The temperature plays an important role because of its many influences on the ionization behaviour, but also on the production of ions due to thermal decomposition of thermolabile analytes. In most cases, thermal decomposition of analytes already takes place in the vaporizer tube, and the mixture of analyte-related molecules is subsequently ionized. The mass spectrum appears to show fragmentation, although some of the fragment ions observed cannot be explained from a mass spectrometric point-of-view, but rather are due to hydrolysis and subsequent ionization of the hydrolysis product.

**Table 2** Proton affinities of some common mobile-phase constituents ($PA_A$) and of typical compound classes with one functional group ($PA_M$) in kJ $mol^{-1}$

| *Reagent gas* | $PA_A$ (kJ $mol^{-1}$) | *Compound class* | $PA_M$ (kJ $mol^{-1}$) |
|---|---|---|---|
| Methane | 536 | Ethers, esters, ketones | 630–670 |
| Water | 723 | Polycyclic aromatic | 710–800 |
| Methanol | 773 | Hydrocarbons | |
| Acetonitrile | 797 | Carboxylic acids | <800 |
| Ammonia | 857 | Carbohydrates | 710–840 |
| Methylamine | 894 | Alcohols | 750–840 |
| Pyridine | 921 | Thio | 750–880 |
| Dimethylamine | 922 | Amine, nitro | 840–1000 |
| | | Peptides | 880–1000 |

**Table 3** Gas-phase acidities ($\Delta H_{acid}$) of some common mobile-phase constituents and of typical compound classes with one functional group in kJ $mol^{-1}$

| *Reagent gas* | $\Delta H_{acid}$ (kJ $mol^{-1}$) | *Compound class* | $\Delta H_{acid}$ (kJ $mol^{-1}$) |
|---|---|---|---|
| Ammonia | 1657 | Benzyl alcohol | 1662 |
| Water | 1607 | Toluene | 1588 |
| Methanol | 1589 | Alkyl alcohols | 1560–1590 |
| Acetonitrile | 1528 | Ketones, aldehydes | 1530–1550 |
| Acetic acid | 1429 | Anilines | 1510–1540 |
| Formic acid | 1415 | Thiols | 1485–1510 |
| Fluoroacetic acid | 1394 | Phenols | 1400–1470 |
| Trifluoroacetic acid | 1323 | Benzoic acid | 1420 |

The general lack of fragmentation under thermospray conditions has led to the more extensive application of MS/MS instrumentation as well as to research into the possibilities of collision-induced dissociation of ions in the ion source by means of high voltages on the repeller electrode. The latter showed nice perspectives in fundamental studies, with mass spectra quite similar to those observed in MS/MS but they proved to lead to a signal reduction that was too large for successful use in real-life applications.

## Operation and optimization

The thermospray interface for LC-MS is generally considered as a difficult interface. This is due to the fact that for a proper operation the careful

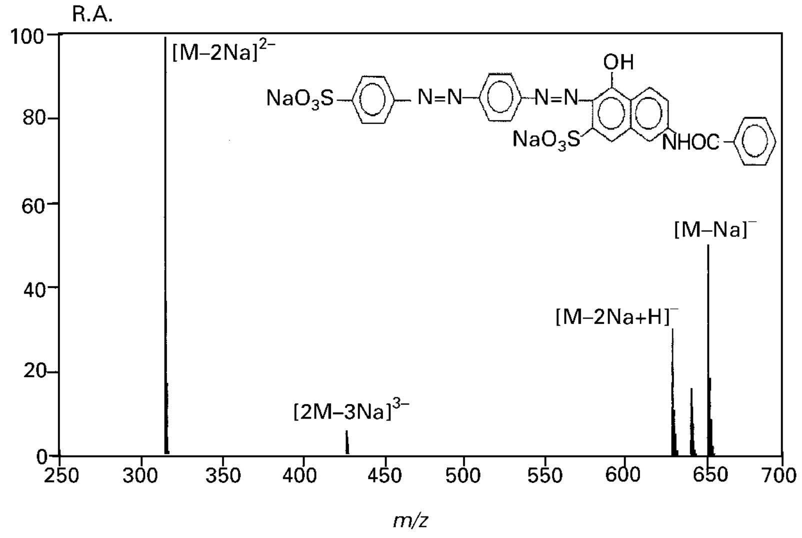



**Figure 4** Negative-ion thermospray mass spectrum of the disulfonated azo dye. Direct Red 81 (mobile phase contains 10 mmol $L^{-1}$ ammonium acetate).

optimization of a variety of mostly interrelated experimental parameters is required.

The performance of the interface is to a large extent determined by the quality of the spray, which in turn depends on the quality of the vaporizer, the solvent composition and the temperature control at the vaporizer. The temperature at the vaporizer depends on the type: with a Vestec-type vaporizer the stem and tip temperature are typically set at ~120°C and 220°C, respectively, while the vaporizer temperature of a Finnigan-type vaporizer is typically set at ~100°C.

Because the thermospray interface contains a dedicated ion source block, tuning and calibration of the source and analyser parameters is obligatory. Calibration and tuning cannot be performed with common calibrants like perfluorokerosene. Diluted solutions of polyethylene glycols are used in most cases, although a tuning and calibration based on clusters of ammonium acetate, ammonium trifluoroacetate and even simply water was proposed and used as well.

After tuning and calibration, the proper functioning of the interface can further be investigated by the injection of a number of standard compounds, e.g. adenosine and tertiary amines, as well as the compound(s) of interest. Subsequently, the system can be optimized to achieve the highest response or the best signal-to-noise ratio. Parameters to be studied are: the ammonium acetate concentration, the concentrations of the organic modifier and possible other mobile phase additives, the flow-rate, the optimum compound-dependent vaporizer temperature, the source block temperature, the repeller potential, as well as the ionization mode. A lower flow-rate generally requires a lower vaporizer temperature, as does a higher content of the organic modifier. However, at a modifier content exceeding 40%, the thermospray buffer ionization mode is generally ineffective. Although the vaporizer temperature should be set in such a way that ~95% solvent vaporization inside the vaporizer is achieved, which in principle is primarily dependent on the flow-rate and the solvent composition, fine-tuning of the vaporizer temperature for a particular application may provide significant improvement of the performance. The analyte-related optimum of the vaporizer temperature may be sharp.

## Applications

Thermospray ionization was especially applied between 1987 and 1992 in combination with LC-MS for a wide variety of compound classes, e.g. pesticides and herbicides, drugs and metabolites, alkaloids, glycosides and several other natural products, as well as peptides.

There are many studies available concerning the characterization of interface and ionization performance for the thermospray LC-MS analysis of pesticides, herbicides and insecticides, the improvement of detection limits and information content of the mass spectra. Compound classes most frequently studied are the carbamates, organophosphorous pesticides, triazine and phenylurea herbicides, chlorinated phenoxy acetic acids, and sulphonylureas.

Analytical strategies for the analysis of pesticides and herbicides in environmental samples, e.g. surface

and tap water, are based on combined solid-phase extraction (SPE), LC separation and subsequent thermospray MS detection, often in a completely automated online system. Specific strategies have been developed for multiresidue screening as well as quantitative determination of pesticides and herbicides from specific compound classes. More recently, there is a growing interest in the determination and identification of pesticide and herbicide degradation products.

Environmental applications of LC-MS, not only pesticides and herbicides, but dyes, shellfish toxins, surfactants, organotin and other environmental contamination were recently reviewed in a multi-authored book, edited by Barceló.

Thermospray LC-MS has also found frequent application in the qualitative and quantitative analysis of drugs and their metabolites in biological fluids, like plasma and urine, and tissue extracts.

In the drug development area, thermospray ionization has found application in open-access approaches, as the technique allows the rapid determination of the molecular mass of a synthesized product without the need to optimize too many experimental variables. In this type of work, the thermospray interface is simply applied as an easy access to the MS.

Thermospray LC-MS, especially in combination with MS/MS was successfully applied in the characterization of drug metabolites. Metabolite screening strategies based on precursor-ion or neutral-loss scan modes in MS/MS were also proposed for the detection of both Phase I and Phase II metabolites. For the Phase II metabolites, neutral loss scan with 176 or 80 Da losses for glucuronide and sulfate conjugates, respectively, were proposed. However, this approach is successful for only some Phase II metabolites, because it was found that the Phase II metabolites often undergo thermally induced ammoniolysis, resulting in mass spectra of the aglycones.

The thermospray interface was the first LC-MS system that allowed reliable quantitative bioanalysis for a wide variety of compounds. Numerous examples were published in the literature. An excellent example is the automated analysis of bambuterol. The automated system, described by Lindberg and coworkers, contained a series of feedback steps in order to assure the various components of the system were operating properly during overnight, unattended analysis and to avoid the loss of valuable sample material. The same approach was applied to the quantitative bioanalysis of cortisol and related steroid compounds. In order to enhance the response of cortisol in thermospray ionization, the compound was derivatized to the 21-acetate using acetic anhydride. This is a viable approach to slightly increase the proton affinity of an analyte to obtain improved ionization characteristics in thermospray buffer ionization.

Thermospray LC-MS was also frequently applied in the analysis of natural products, e.g in extracts from plants or cell cultures, of (modified) nucleosides, endogenous compounds such as prostaglandins, and some peptides. However, later it was demonstrated that alternative LC-MS strategies, e.g. based on electrospray ionization, were far more effective in the MS analysis of peptides.

## Conclusion and perspectives

For a number of years (1987–1992), thermospray LC-MS was the most frequently applied interface for LC-MS. It has demonstrated its applicability in both qualitative and quantitative analysis in various application areas. With the advent of the more robust LC-MS interfaces, based on atmospheric-pressure ionization, the use of thermospray interfacing and ionization rapidly decreased. The newer technology pointed out the limitations of the thermospray system, e.g. in the analysis of thermolabile compounds, ionic compounds, high molecular-mass compounds, as well as in robustness and user-friendliness. Therefore, thermospray as an ionization and interface technique for LC-MS is now history. Thermospray nebulization will continue to be used, e.g. in nebulization for ICP-MS.

*See also:* **Chromatography-MS, Methods; Ionization Theory.**

## Further reading

Arpino PJ (1990) Combined liquid chromatography mass spectrometry. Part II. Techniques and mechanisms of thermospray. *Mass Spectrometry Review* **9**: 631–669.

Arpino PJ (1992) Combined liquid chromatography mass spectrometry. Part III. Applications of thermospray. *Mass Spectrometry Review* **11**: 3–40.

Barceló D (ed) (1996) *Applications of LC-MS in Environmental Chemistry*. Amsterdam: Elsevier.

Blakley CR and Vestal ML (1983) Thermospray interface for liquid chromatography/mass spectrometry. *Analytical Chemistry* **55**: 750–754.

Blakley CR, McAdams MJ and Vestal ML (1978) Crossed-beam liquid chromatograph–mass spectrometer combination. *Journal of Chromatography* **158**: 261–276.

Conver TS, Shawn T, Yang J and Koropchak JA (1997) New developments in thermospray sample introduction for atomic spectrometry. *Spectrochimica Acta B* **52**: 1087–1104.

Iribarne JV and Thomson BA (1976) On the evaporation of small ions from charged droplets. *Journal of Chemical Physics* **64**: 2287–2294.
Iribarne JV and Thomson BA (1979) Field induced ion evaporation from liquid surfaces at atmospheric pressure. *Journal of Chemical Physics* **71**: 4451–4463.
Lindberg C, Paulson J and Blomqvist A (1991) Evaluation of an automated thermospray liquid chromatography – mass spectrometry system for quantitative use in bioanalytical chemistry. *Journal of Chromatography* **554**: 215–226.
Niessen WMA (1998) *Liquid Chromatography – Mass Spectrometry*, 2nd edn. New York: Marcel Dekker.
Röllgen FW, Nehring H and Giessmann U (1989) Mechanisms of field induced desolvation of ions from liquids. In Hedin A, Sundqvist BUR and Benninghoven A (eds), *Ion Formation from Organic Solids (IFOS V)*, pp 155–160. New York: Wiley.
Yergey AL, Edmonds CG, Lewis IAS and Vestal ML (1990) *Liquid Chromatography/Mass Spectrometry, Techniques and Applications*, pp 31–85. New York: Plenum Press.

# Time of Flight Mass Spectrometers

**KG Standing** and **W Ens**, University of Manitoba, Winnipeg, Manitoba, Canada

Copyright © 1999 Academic Press

**MASS SPECTROMETRY**
**Methods & Instrumentation**

The time of flight (TOF) mass spectrometer is perhaps the simplest type of mass analyser, at least in principle. The kinetic energy of an ion of mass $m$ is given by $mv^2/2$, so a measurement of its speed $v$ by timing the flight of the ion over a given path determines the mass when the kinetic energy is known, or when the spectrometer has been suitably calibrated. Such an instrument was first proposed (in 1946) to take advantage of the improvements in timing circuits developed in the Second World War and in succeeding years it developed a reputation as a device with fast response but low resolution when used with an electron impact ion source. The Bendix Corporation manufactured a commercial TOF instrument that achieved considerable popularity, but later the technique fell into disuse when quadrupole mass filters became common. This situation has changed dramatically in recent years, and the field is now one of the most active areas in mass spectrometry. This has come about partly because of improvements in electronics, but mainly because of the development of new methods of ionization, particularly matrix-assisted laser desorption/ionization (MALDI). In addition, interest has shifted to the measurement of compounds of larger mass, for which TOF methods are especially well suited.

## Ionization methods

To define the 'start signal' for a TOF measurement, it is necessary to produce the ions as a series of short bursts. In some methods of ion production, this is achieved naturally, since the source itself is intrinsically pulsed. An early example of such a source is plasma desorption mass spectrometry (PDMS). In this technique, ions are produced by bombardment of the sample with particles of MeV energies, usually fission fragments, and the pulses are formed by individual bombarding particles arriving at the target. More recently, much of the activity in the field has stemmed from the widespread use of MALDI, as remarked above. MALDI is also an intrinsically pulsed source, where the ions are produced by irradiation of a sample with a beam from a pulsed laser. The laser provides the start pulse, so the coupling to a TOF instrument is a natural one.

On the other hand, ions produced in a continuous beam must be formed into pulses by an appropriate device, so an additional complication is introduced. Examples are electron ionization, secondary-ion mass spectrometry (SIMS), and most recently electrospray ionization (ESI). As mentioned above, the earliest commercial TOF instruments used electron-impact ionization, and the difficulty in producing short ion bursts with this method was mainly responsible for their limitations in mass resolution. However, new technology has enabled dramatically improved performance for continuous sources, particularly ESI, and TOF has been gaining popularity for such sources as well.

## A simple model

An idealized TOF instrument is illustrated in **Figure 1**. Here particle or laser bombardment

desorbs an ion of charge $+q$ and mass $m$ at time $t = 0$ from a sample deposited on the target, a plane conducting surface ($z = 0$) that is held at potential $+V$. A parallel grid at $z = s$ is kept at ground potential so that there is a uniform electrostatic field directed along the $z$ axis in the 'source region' between the target and the grid. If the ion starts out with zero velocity at the target, it is accelerated by the electric field and arrives at the grid with kinetic energy $\frac{1}{2}mv_z^2 = qV$ or velocity $v_z = \sqrt{(2qV)/m}$; its average velocity in the source region is half this value. The ion then passes through the ideal grid and travels with constant velocity through a drift region to the plane surface of a detector at $z = s + d$. Thus the total time of flight $t$ is the sum of the time spent in the source region and the time spent in the drift region; i.e. $t = (2s + d)\sqrt{m/(2qV)}$. Measurement of the time of flight with a fast 'clock' determines the mass $m$, since the other parameters in the equation are known. Note that the time of flight is proportional to $\sqrt{m/q}$.

This geometry is close to that introduced by Macfarlane and Torgerson in their pioneering studies in PDMS in 1976, which marked a major step forward in the development of TOF techniques. As in the simple model, ions were ejected from an equipotential surface (by fission fragment bombardment in this case), so the spatial spreads that had limited the performance of earlier instru-ments were removed.

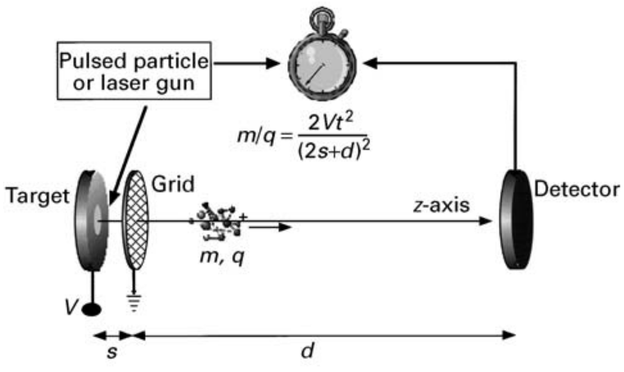



**Figure 1** Schematic diagram of a simple idealized time-of-flight mass spectrometer.

## Features of TOF measurements

The advantages of the technique in the ideal case can be seen from the simplified description above:

- The mass range of the analyser is unlimited, since the clock can simply be allowed to run until the ion of interest arrives at the detector. The only limits on the mass range are imposed by the ion source and the detector.
- Parallel detection of all the ions over the complete mass range is straightforward, because the mass is determined by the measured arrival time at the detector, and the arrival times of all ions can be recorded.
- Defining slits are unnecessary.
- Because of the previous points, sensitivity is high and the instrument has a fast time response.

In the past, the main disadvantage of TOF systems has been poor mass resolution, because of the difficulty of producing an ion beam consisting of very short bursts, and because of the inevitable departures from the ideal case. However, TOF resolution has been substantially increased recently by various technical improvements, as described below. Consequently, TOF instruments now often provide the optimum combination of resolution and sensitivity, particularly in cases where the whole mass spectrum is required. In contrast to the parallel detection capability of TOF instruments, most other types of mass spectrometer operate as mass filters, in which the mass spectrum is obtained by scanning through the mass range, one mass at a time. Thus, in these instruments there is a reciprocal relationship between mass resolution and sensitivity; resolution must be sacrificed to obtain high sensitivity.

## Compensation for ion energy spreads with an electrostatic reflector

The most obvious defect of the simple model given above is its failure to take account of the initial energy that the ion possesses as it leaves the target. Variations in the initial energy may give rise to a considerable spread of times of flight, and thus to a deterioration in resolution. However, a modification to the instrument that alleviates this problem is the introduction of a reflector or ion mirror, as first proposed by Mamyrin (who called it a 'reflectron'). The simplest case is illustrated in **Figure 2**, where ions on a plane surface (the 'object plane') just outside the source region have a distribution of velocities. Ions travel freely from this surface until they enter a uniform retarding electrostatic field (an ion mirror). Like projectiles in the earth's gravitational field, ions follow parabolic paths within the mirror and leave it with the ion velocity component parallel to the mirror axis ($v_z$) reversed. The ion then travels freely to the detector. For $L = L_1 + L_2$, the time spent in free flight is $L/v_z$, and

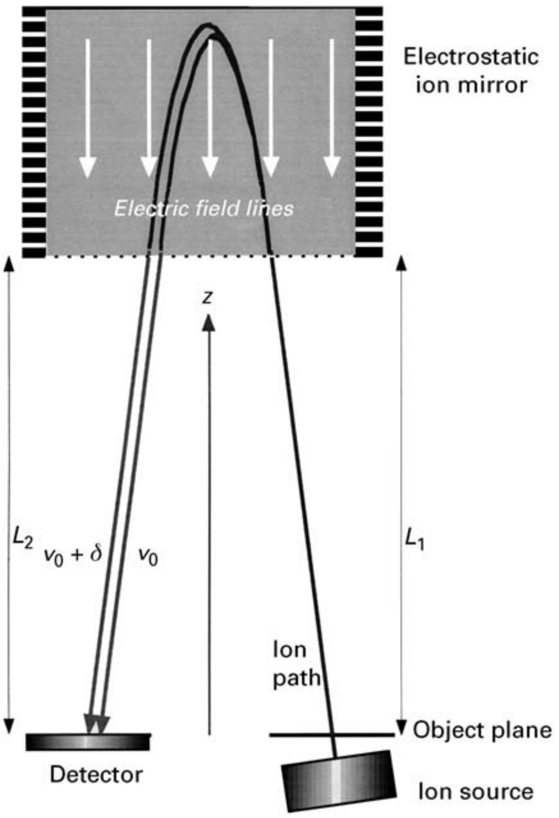



**Figure 2** An illustration of the principle of ion mirror to compensate for velocity spreads. Two ion paths are shown for axial velocities $v_z = v_0$ and $v_z = v_0 + \delta$. The ion with higher velocity spends less time in the field-free region but more time in the mirror.

the time spent in the mirror is $2mv_z/qE$, where $E$ is the magnitude of the retarding electric field. If $v_z = v_0 + \delta$, we can expand as a function of $(\delta/v_0)$ to give a total time of flight:

$$t = \left(\frac{2mv_0}{qE} + \frac{L}{v_0}\right) + \left(\frac{\delta}{v_0}\right)\left(\frac{2mv_0}{qE} - \frac{L}{v_0}\right) + \cdots \quad [1]$$

Setting $2mv_0/qE = L/v_0$, or $E = 2mv_0^2/qL$ removes the first-order term in $\delta/v_0$. Thus the reflector eliminates the effect of a velocity variation $\delta$ to first order. Under this condition, the total time of flight $t = 2L/v_0$, so the ion spends equal amounts of time in the mirror and in free flight. Higher-order terms can be removed by the use of more complicated electric fields, for example by the two-stage mirror described by Mamyrin, but any advantage in doing so is often lost because the resolution may deteriorate because of other effects, particularly in the acceleration region, which must be considered separately.

## Space and velocity focusing in the acceleration region

As pointed out above, variations in ion velocity on the object plane can be corrected to a large extent by the use of a reflector. However, effects during ion production or acceleration may give not only a spread in velocity on this plane but also a spread in time. The latter effects are not corrected by the mirror.

Some of these phenomena were discussed in a classic paper of 1955 by Wiley and McLaren in connection with their TOF studies of ions produced by electron ionization, which in general may have both an initial spatial spread and an initial velocity spread. They considered an ion source with two stages of acceleration in series, each with a uniform longitudinal electric field. They pointed out first that ions with a pure spatial spread in the acceleration region are subject to 'space focusing'; that is, there is some plane beyond the grid that ions of the same mass will reach at approximately the same time. This is because ions initially close to the grid acquire less energy in the acceleration region than more distant ions and therefore are overtaken by the latter after travelling some distance determined by their original position and the accelerating electric fields. In the special case of a single uniform field, the focal plane is a short distance $D = 2s_0$ beyond the grid if the ions originate an average distance $s_0$ inside it. If a two-stage acceleration region is used, the position of the focal plane can be adjusted by changing the ratio of the electric fields in the two regions. When the ions arrive at this plane, they will have a velocity spread because of the differing amounts of energy gained during acceleration, so the original spatial spread has been exchanged for a velocity spread.

This technique does not give perfect spatial focusing, even in the ideal case, but the usual limitation results from an initial velocity spread in addition to the spatial spread. The worst case involves two ions at the same $z$ position but with velocities in opposite directions when the extraction pulse is applied. The ion in the negative $z$ direction must first be turned around, and the two ions will arrive at the focal plane separated by this 'turn-around time'.

If the ions are produced in pulses (e.g. if the electron gun in an electron ionization source is pulsed), then some velocity compensation in the above situation is possible by the use of 'time-lag focusing', also proposed by Wiley and McLaren. In this technique, now usually called delayed extraction, a delay is introduced between ion production and the application of the accelerating fields, during which the ions drift freely. For

simplicity, consider ions starting with zero spatial spread, i.e. with a pure velocity spread. When the accelerating field is applied, the ions will be separated in space according to their velocity. Those ions with higher initial $v_z$, will be closer to the end of the acceleration region, and will receive a smaller accelerating impulse. If the time delay and the amplitude of the accelerating voltage are adjusted properly, ions of the same mass will arrive at a focal plane at approximately the same time.

In the general case there may be both spatial and velocity spreads in the initial ion distribution, so some compromise is necessary to give optimum focusing. However, two currently popular ionization methods, ESI and MALDI, approximate the two limiting cases described above. Ions suitably injected from an ESI source have an appreciable spatial spread, but a very small velocity spread (see below). A pure velocity spread is approximated by the geometry normally used for MALDI, since the MALDI ions are ejected from an equipotential target by a very short laser pulse. In a simple linear TOF spectrometer, resolution can be optimized in both cases by using a two-stage acceleration region and setting the accelerating fields so that the focal plane coincides with the plane of the detector.

## A modern TOF geometry

The best performance of TOF instruments is now achieved by a combination of an electrostatic reflector and the Wiley–McLaren focusing techniques, and this combination is the basis for most high-performance TOF systems.

By themselves, the Wiley–McLaren focusing methods are limited because the narrowest time distributions are achieved for short flight paths, but good mass resolution requires long flight paths to provide time dispersion between ions of different masses. The electrostatic mirror provides energy focusing but does not compensate for time spreads in the source region, so by itself it also offers limited improvement. However, the two methods are highly effective when used in combination. The ions from the source are focused into a flat ion packet near the source and coincident with the object plane of the mirror (see **Figure 2**). The mirror then images the ion packet onto the detector, greatly increasing the time of flight and therefore the time dispersion between species, without appreciably increasing the time spread. The ions at the first focal plane have a considerable velocity spread as mentioned above, but the mirror compensates to first order for velocity spreads in its object plane.

## TOF measurements with a continuous beam

As remarked above, a continuous beam must be formed into pulses before it is introduced into the TOF spectrometer. This requirement is an extra complication that is not present if the beam is intrinsically pulsed. However, there are several cases in which mass analysis of ions produced in a continuous beam can benefit considerably from the features of TOF instruments. For example, electrospray ionization has been the most successful technique for producing ions from intact noncovalent complexes, but these ions are often formed with high mass-to-charge ratios, beyond the range of quadrupole mass filters; TOF imposes no limit on the mass-to-charge ratio ($m/z$) range. A second example involves coupling of separation techniques such as high-performance liquid chromatography with mass spectrometry. The sensitivity and fast time response of TOF instruments are well suited for such an application, but most separation techniques produce a continuous output. The same can be said of coupling TOF instruments with other types of mass spectrometer in order to perform MS/MS measurements as discussed below. For these reasons, there is clearly a need for an effective method for coupling continuous sources to TOF instruments.

A continuous beam can be injected into a TOF spectrometer in the longitudinal geometry of **Figure 1**, but only with very low efficiency. A more practical arrangement is illustrated in **Figure 3**; here a continuous beam of ions enters the TOF instrument perpendicular to its axis with low velocity and is injected into the flight path by the electrical pulses indicated in the figure. This technique takes advantage of the relative insensitivity of TOF measurements to spatial spreads in a plane perpendicular to the TOF axis. Such ‘orthogonal injection’ geometries were first introduced in the early 1960s, but acquired particular relevance when used with an electrospray source and an ion mirror by Dodonov. Limits on the injection efficiency and the resolution are set by the spatial and velocity spreads of the injected beam as described above. The properties of these instruments are therefore considerably improved if they are preceded by an ion guide running at relatively high pressure (up to ~10 Pa) to provide collisional translational cooling of the ions before they enter the TOF spectrometer. In this way, a beam is produced with a small energy spread, limited by thermal velocities, allowing effective spatial focusing as described above.

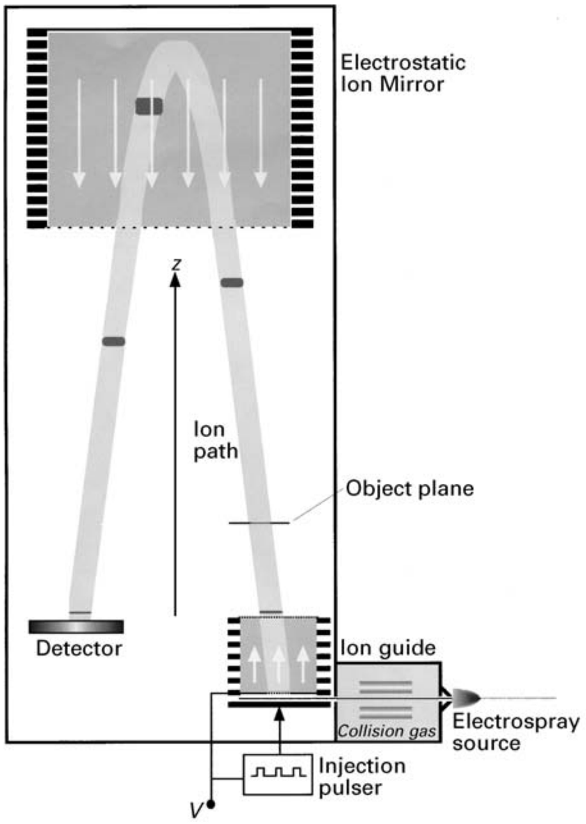



**Figure 3** A schematic diagram of an orthogonal-injection TOF instrument with an electrospray ionization source. Collisional cooling is used in a quadrupole ion guide to produce a beam with a small energy spread, and a small cross section. The pressure in the ion guide is typically tens of millitorr; the main TOF chamber is under high vacuum, typically $10^{-7}$ torr. Ions are pulsed into the spectrometer at a repetition rate of several kilohertz; packets of ions for one mass are shown at several positions along the ion path.

## Daughter-ion measurements in TOF spectrometers

The study of the products of ion breakup can often give useful information about the molecular structure of the parent ion. In the simplest measurement of this type, a metastable parent ion may suffer unimolecular decay as it passes down the flight tube (so-called 'post-source decay'). The velocities of the daughter ions or neutrals can be determined from their times of flight, but such a measurement in a linear TOF spectrometer yields little information, since the *velocities* of the decay products are approximately the same as the velocity of the parent ion, as a result of conservation of momentum. Thus a daughter ion cannot be distinguished by its time of flight in a simple linear instrument from the parent or from other daughters.

On the other hand, the kinetic energy of the parent ion is divided among its decay products, so daughter ions will have *energies* determined by their masses. In contrast to the situation in a linear TOF spectrometer, the total time of flight in a reflecting instrument depends on the ion's energy as well as its velocity, because the ratio of ion energy to charge determines the distance the ion penetrates the mirror and thus the time spent there. The mirror does not correct for the velocity spreads of the daughter ions as well as it does for the parent (assuming that the electric field in the mirror is set to the value appropriate for the parent ion). However, for optimum resolution the problem can be minimized by examining the daughter-ion spectrum in segments, with the mirror field set to an appropriate value for each segment. Alternatively, a mirror with a nonlinear electric field can be used.

Usually there are a number of different parent ions extracted from the sample, each giving rise to a daughter-ion spectrum. It is therefore necessary to separate these in order to identify the particular parent ion giving rise to an observed decay. This is normally done by deflecting all ions except the desired one out of the flight path by some form of 'ion gate', and examining the decay products of the parent ions one by one.

The post-source decay technique suffers from some limitations, such as the limited selectivity obtainable on the parent ion, modest mass accuracy of the daughter ions (compared to the accuracy achieved for parent ions), and the frequently incomplete information obtained from the metastable decay spectrum. These factors have stimulated the development of various tandem instruments, as discussed below.

## Tandem instruments

Tandem mass spectrometers provide a more flexible means of studying molecular structure. Such devices combine a mass filter (MS1) to select the parent ion, a gas cell for collision-induced dissociation, and a second mass spectrometer (MS2) to analyse the daughter ions produced from breakup of the parent. A particularly suitable combination is a quadrupole mass filter (Q) as MS1, a quadrupole ion guide (q) as the collision cell (excited by RF only), and a reflecting TOF instrument as MS2. MS/MS experiments are often limited by the amount of sample available, and by the time available to obtain a spectrum, so the sensitivity and rapid response of TOF offers a significant advantage over scanning instruments for MS2. Because of the parallel detection, the sensitivity can be maximized without reducing mass resolution. On the other hand, MS1 is simply used as a mass filter, so parallel detection offers no advantage, and a quadrupole mass filter is a good choice because it can be efficiently coupled to the collision cell. The

QqTOF geometry thus offers an improved alternative to the popular triple-quadrupole instrument.

## Detection methods

Detection in most types of mass spectrometer depends on measurement of the electrons or the secondary ions ejected from a surface by the impact of the ions of interest, usually after electron multiplication. A TOF detector must obviously have a fast time response. It must also have a large and flat active area, since the cross section of the ion beam at the detector is relatively large (usually several centimetres). Finally, to exploit the high mass range that TOF is capable of, the detector must be sensitive to ions with high *m*/*z*, which for a given energy have relatively low velocity.

The first two of the above demands are met effectively with microchannel plate (MCP) electron multipliers. These are flat arrays (of various dimensions) of micrometre-sized channels, each one acting as a very fast electron multiplier when a voltage gradient exists along it. Most high-resolution TOF instruments now use a microchannel plate for the first element of the detector.

The secondary emission coefficient decreases rapidly for decreasing velocity, and therefore for increasing *m*/*z*. For this reason, the ions are usually accelerated to relatively high energies before they are incident on the detector. Most MALDI/TOF instruments use 30 kV acceleration or more. Even so, for singly charged ions larger than about 50 kDa, the electron emission coefficient is considerably less than unity. Detection by electron emission still appears to be feasible in the high mass range because of the large number of ions desorbed in MALDI, although more efficient detection is possible by using a detector designed to take advantage of secondary ion emission at the cost of some loss of resolution.

The problem of detecting high masses does not occur when an ESI source is used, because it produces ions with much higher charge states, and therefore much higher energy (and lower *m*/*z*) for the same molecular mass. Intact molecular ions from brome mosaic virus with mass larger than 4.6 MDa (*m*/*z* ~25 000) have been observed by ESI/TOF using only 5 kV acceleration.

## List of symbols

$E$ = electric field strength of reflector; $L_1$, $L_2$ = field-free path lengths (see **Figure 2**); $m$ = ion mass; $q$ = ion charge; $t$ = time; time of flight; $v$ = ion velocity; $V$ = electric potential; $z$ = $z$ coordinate; ionic charge; $\delta$ = variation in ion velocity.

*See also:* **Fragmentation in Mass Spectrometry; Ion Molecule Reactions in Mass Spectrometry; Ion Structures in Mass Spectrometry; Ionization Theory; Laser Applications in Electronic Spectroscopy; Metastable Ions; Plasma Desorption Ionization in Mass Spectrometry; Quadrupoles, Use of in Mass Spectrometry.**

## Further reading

Chernushevich IV, Ens W and Standing KG (1999) Orthogonal-injection TOFMS for analyzing biomolecules. *Analytical Chemistry* **71**.

Cotter RJ (1997) *Time-of-Flight Mass Spectrometry*. Washington DC: American Chemical Society.

Dodonov AF, Chernushevich IV and Laiko VV (1994) Electrospray ionization on a reflecting time-of-flight mass spectrometer). In Cotter RJ (ed) (1994)*Time-of-flight Mass Spectrometry*, pp 108–123. Washington DC: American Chemical Society.

Hillenkamp F, Karas M, Beavis RC and Chait BT (1991) Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Analytical Chemistry* **63**: 1193A–1203A.

Macfarlane RD and Torgerson DF (1976) Californium-252 plasma desorption mass spectroscopy. *Science* **191**: 920–925.

Mamyrin BA, Karataev VI, Shmikk DV and Zagulin VA (1973) The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Soviet Physics – JETP* **37**: 45–48.

Spengler B (1997) Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. *Journal of Mass Spectrometry* **32**: 1019–1036.

Stephens WE (1946) A pulsed mass spectrometer with time dispersion. *Physical Review* **69**: 691.

Wiley WC and McLaren IH (1955) Time-of-flight mass spectrometer with improved resolution. *Review of Scientific Instruments* **26**: 1150–1157.

# Tin NMR, Applications

*See* **Heteronuclear NMR Applications (Ge, Sn, Pb).**

# Titanium NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# Tritium NMR, Applications

**John R Jones**, University of Surrey, Guildford, UK

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Applications**

It is undoubtedly true that if tritium were not radioactive it would be one of the most widely used of all NMR nuclei. Its favourable properties – an isotope of hydrogen, one of the most important of all the elements, with a nuclear spin $I = \frac{1}{2}$ and the most sensitive of all NMR nuclei – counts but little with those who quake at the mention of the word 'radioactivity', let alone think of spinning radioactive samples. However, there are those, increasing in number, who believe that tritium is the most favoured of all the nuclei, combining the advantages of favourable radioactive properties (weak $\beta^-$ emitter ($E_{avg} \sim 6$ keV), convenient half-life (12.3 years), ready detection by liquid scintillation counting with good efficiencies (typically >50%) with these positive NMR characteristics. Furthermore the technology for synthesizing and handling tritiated compounds has been in place for many years whilst the development of spectrometers operating at ever increasing fields means that less tritium is required for NMR detection. In addition there is virtually no natural abundance tritium concentration, unlike the situation that exists for stable isotopes, so that the dynamic range is enormous. It is this factor above all others that will lead to an expansion in the use of tritium and tritium NMR spectroscopy in the life sciences. Recent publications show that such possibilities are being increasingly appreciated.

## Properties of the nucleus

As well as having a nuclear spin $I = \frac{1}{2}$ tritium has a high nuclear magnetic moment which is responsible for the magnetogyric constant being larger than for any other nucleus, as also is its sensitivity to NMR detection, 21% higher than that for $^1$H. At 11.7 T, at which field the $^1$H NMR frequency is 500 MHz, the $^3$H NMR frequency will be 533.3 MHz.

## Sample preparation and spectrum measurement

Before embarking on any $^3$H NMR work the personnel must become designated radiation workers, have the appropriate radiochemical facilities and become familiar with tritiation procedures. In this respect it is frequently useful for initial training to be given in appropriate deuteration studies although the corresponding tritium work will usually be carried out on a much smaller scale and the purification procedures will depend greatly on appropriate radio-chromatographic methods, as distinct from chromatographic methods. With appropriate rules and regulations in place a radiochemical laboratory need not be any more hazardous than an ordinary chemistry laboratory, particularly if the rule 'that only the minimum amount of radioactivity consistent with the requirements of the project' is used.

There are two separate units of radioactivity in use, the first being the curie (Ci) which is defined as an activity of $3.7 \times 10^{10}$ disintegrations per second. This is a large unit, hence the frequent use of smaller subunits, the millicurie ($10^{-3}$ Ci) and the microcurie ($10^{-6}$ Ci). The second, and more recently introduced unit, is the becquerel (Bq). At one disintegration per second this is an extremely small amount of radioactivity. The conversions are

$$1 \text{ Bq} = 2.703 \times 10^{-11} \text{ Ci } (27.03 \text{ p Ci}),$$
$$1 \text{ Ci} = 3.7 \times 10^{10} \text{ Bq or } 37 \text{ GBq}$$

Although there are a large number of methods available for preparing tritiated compounds the most widely used stem from the following categories:

- catalytic hydrogenation of an unsaturated precursor using $^3H_2$ gas;
- catalytic halogen–tritium replacement reactions;
- hydrogen isotope exchange reactions catalysed by acids, bases or metals;
- reductions using reagents such as sodium borotritide;
- methylation reactions using reagents such as tritiated methyl iodide.

Recently microwaves have been used to greatly accelerate the rates of many of these reactions whilst the development of microwave-enhanced solid state tritiation procedures offers considerable potential.

For $^3$H NMR analysis 1–10 mCi of material dissolved in ~50–100 μl of a deuterated solvent is usually sufficient to obtain a spectrum of good signal-to-noise in a matter of 1–10 h, depending on whether the radioactivity is located at one site (a specifically labelled compound) or in several positions (a general labelled compound) – this assumes a spectrometer operating at 300 MHz for $^1$H and 320 MHz for $^3$H. For reasons of safety the radioactive samples are placed in narrow cylindrical tubes, sealed at the top, which themselves are placed in standard NMR tubes – this double containment procedure, initially introduced when much higher levels of radioactivity were required, provides a measure of safety as well as reassurance. Experience shows that $^3$H NMR spectra are of two kinds. Firstly there are those in which the specific activity of the compound is less than ~1 Ci mmol$^{-1}$ so that $^3$H–$^3$H couplings are absent and the $^3$H NMR ($^1$H decoupled) spectra consist of a series of single lines, which on integration give the relative incorporation of $^3$H at each site. Nuclear Overhauser effects (NOEs) are small and differential effects even smaller so that there is no need to obtain NOE-supressed spectra. It should also be mentioned that there is no need to synthesize a tritiated organic standard – all the $^3$H chemical shifts are obtained via the $^1$H chemical shifts and the Larmor frequency ratio.

For compounds at high specific activity, e.g. prepared by the addition of $^3H_2$ gas across the unsaturated group of a precursor, there will be tritium–tritium couplings, the magnitude of which are similar to those of hydrogen, i.e. $J(^1\text{H}–^3\text{H}) = J(^1\text{H}–^1\text{H}) \times 1.066$. Small isotope effects are present but these do not complicate the interpretation of the spectra, on the contrary they can aid the analysis of the relative proportions of isotopomers present, e.g. RC $^3H_3$, RC $^1H^3H_2$ and RC $^1H_2{}^3H$. Tritium couplings to boron, deuterium, carbon, fluorine and phosphorus are also in agreement with theory.

## Applications of tritium NMR

### New tritium labelling reagents

The development of $^3$H NMR spectroscopy has made possible many new applications and in the process has stimulated research into the development of new labelling reagents and hence new/improved labelling procedures. One such area is that of tritide reagents. Essentially carrier-free LiB $^3H_4$ can now be obtained via the two-step sequence:

$$\text{BuLi} + {}^3\text{H}_2 \rightarrow \text{Li}^3\text{H} + \text{Bu}^3\text{H}$$
$$4\text{Li}^3\text{H} + \text{BBr}_3 \rightarrow \text{LiB}^3\text{H}_4 + 3\text{LiBr}$$

Similarly, carrier-free sodium triethylborotritide, a useful reagent for the stereo- and regiospecific reduction of carbonyl-containing compounds, can be synthesized in the following manner:

$$\text{BuLi} + \text{NaOBu}^t + {}^3\text{H}_2 \rightarrow \text{Na}^3\text{H} + \text{LiOBu}^t + \text{Bu}^3\text{H}$$
$$\text{Na}^3\text{H} + \text{Et}_3\text{B} \rightarrow \text{NaEt}_3\text{B}^3\text{H}$$

Tri-*n*-butyltin and lithium tri-*s*-butylborotritide are other useful reagents. Increasingly sophisticated tritium labelling technology is being developed as an alternative to the more traditional hydrogenation and catalytic dehalogenation reactions. The procedures will find wide application in the tritiation of molecules of biological importance. Thus *N*-tritioacetoxyphthalimide, a new high specific activity tritioacetylating reagent, has been used to label a number of acetylenes, ketones and alcohols whilst radical-induced tritiodeoxygenation reactions can lead to the synthesis of important heterocyclic compounds.

### New more selective tritiation procedures

Hydrogen isotope exchange reactions are widely used not only to study reaction mechanisms but also for labelling compounds with either deuterium or tritium. The reactions may be catalysed by acids or bases under both homogeneous or heterogeneous conditions and frequently lead to generally labelled compounds. The same is true for transition metals. Recently considerable effort has been directed at developing more selective procedures – homogeneous rhodium trichloride has been shown to be very

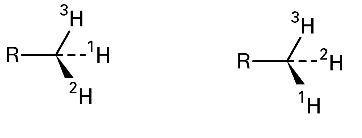
effective in introducing both deuterium and tritium into the *ortho*-aromatic positions of a wide range of pharmaceutically important compounds.

The well-known iridium catalyst $[Ir(COD)(Cy_3P)(Py)]PF_6$, where COD = 1,5-cyclooctadiene and Py = pyridine, demonstrates excellent regioselectivity in isotopic exchange reactions of acetanilides and other substituted aromatic substrates. $^3H$ NMR spectroscopy is invaluable in identifying the site(s) of tritium incorporation – there are many instances where the broad signals in the corresponding $^2H$ NMR spectra are much less informative.

Another iridium catalyst that exhibits good regioselectivity in hydrogen isotope exchange reactions is the complex $[IrH_2(acetone)_2(PPh_3)_2]BF_4$. As in the previous studies the transient existence of a metallocyclic intermediate is indicated. Considerable interest has also been shown in the development of the high temperature solid state catalytic isotopic exchange (HSCIE) method developed by Myasoedev and colleagues. Although labelling is uniform in most instances, some regiocontrol can be exerted by careful temperature control.

## Chiral methyl, stereochemistry and biosynthesis

The analysis of stereochemical problems in both chemistry and biochemistry has benefited greatly from the use of compounds that contain a methyl group with one atom each of $^1H$, $^2H$ and $^3H$. Such compounds exist as a pair of enantiomers, identified by *R* and *S*, and early work in this area will always be associated with the names of Arigoni and Cornforth. Recently a very efficient five-stage synthesis of chiral acetate has been reported (in which the penultimate reaction uses supertritide) with an enantiomeric purity of 100%.




(*S*)-Enantiomer (*R*)-Enantiomer

In the past the determination of whether an unknown sample contained an excess of an (*R*)- or (*S*)-configured chiral methyl group relied on using a reaction in which one hydrogen is removed to generate a methylene group in which tritium is now unevenly distributed between the two methylene hydrogens. The condensation of acetyl coenzyme A with glyoxylic acid catalysed by the enzyme malate synthase, which exhibits a primary kinetic isotope effect $k_H/k_D$ of 3.8, was the chosen reaction. Analysis of the tritium distribution, together with a knowledge of $k_H/k_D$ and the steric course of the reaction, yields the required information – the configuration of the original chiral methyl group and an estimate of the enantiomeric excess.

$^3H$ NMR spectroscopy can provide the necessary information directly; whether $^3H$ has $^1H$ or $^2H$ as a neighbour can be determined directly from the $^1H$–$^3H$ coupling and the $^2H$ isotope shift on the $^3H$ signal. The only problem with the $^3H$ NMR method is that it requires a few mCi of tritiated material, at least with current-day NMR spectrometers. With improvements in spectrometer design and the absence of 'natural abundance' tritium signals this may not always be the case. As it is, the method is direct, does not require any knowledge of the primary isotope effect and no chemical degradations are required.

There are many examples of enzymatic methyltransfer reactions in biochemistry to which the chiral methyl/$^3H$ NMR technology can be applied. One such example involves the important biological methyl donor *S*-adenosylmethionine. Combined with other studies the results show that the transfer of a methyl group to a variety of different nucleophiles all operate with inversion of methyl group configuration.

## Substrate–receptor interactions

Most NMR studies in this area have used $^{13}C$- or $^{15}N$-labelled ligands, the synthesis of which is frequently more demanding than is the case for $^3H$ ligands. Furthermore, the sensitivity of both $^{13}C$ and $^{15}N$ nuclei to NMR detection is considerably less favourable than is the case for tritium. It is therefore somewhat surprising and at the same time disappointing that there are still relatively few examples of protein–ligand interaction studies based on the use of $^3H$-labelled ligands. In an early study $^3H$ NMR spectroscopy was used to monitor the anomeric binding specificity of α- and β-maltodextrins binding to a maltose-binding protein whilst in another study $^3H$ NMR spectroscopy was used to measure the dynamic properties of tosyl groups in specifically $^3H$-labelled tosylchymotrypsin. Preliminary details of a $^3H$ NMR binding study of a tritium-labelled phospholipase $A_2$ inhibitor to bovine pancreatic $PLA_2$ suggest that the tritium atoms are located within the hydrophobic pocket of the protein. In a more extensive study a number of high specific activity tritiated folic acids and methotrexates have been prepared and their complexes with *Lactobacillus casei* dihydrofolate reductase (DHFR) investigated. The $^3H$ NMR results confirm the

presence of three pH-dependent different conformational forms in the complex DHFR·NADP$^{+}$·folate, whereas both the binary and ternary methotrexate complexes (DHFR·MTX, DHFR·NADP$^{+}$·MTX) were shown to exist as a single conformational state.

An interesting $^{3}$H NMR study of the complex formed by [4-$^{3}$H]benzenesulfonamide and human carbonic anhydrase 1 reveals details that are widely different from those obtained when using a fluorinated inhibitor, highlighting the dangers of using fluorine as a 'substitute' isotope for one of the hydrogen isotopes.

#### Macromolecules

The methods that have been developed for tritiating small organic molecules do not lend themselves very readily to the tritiation of large macromolecules such as proteins although there are a few examples where Myasoedov's HSCIE procedure proved successful. It is not surprising therefore that very little work has been reported on, for example, the $^{3}$H NMR of proteins. The polymer area, however, has seen more activity, mainly because it has been much easier to tritiate such compounds – hydrogenation with $^{3}H_2$ gas of a suitable monomer followed by polymerization leads to a specifically tritiated product.

Many polymers are difficult to solubilize and the question has been asked several times whether in view of its good NMR characteristics it is possible to obtain satisfactory solid state spectra. The potential problems have recently been overcome partly by the development of zirconia rotors and partly by enclosing the tritium probe in a Perspex shield so that, in the event of an accident, radioactivity would be retained on a suitable filter. Magic-angle spinning at 17 kHz rotation provides spectra with line widths at half-height of the order of 120 Hz. This has been achieved without $^{1}$H decoupling, this being a more difficult task than for solution studies. It is too early to say at this stage whether $^{3}$H NMR spectroscopy of solids will develop into as widely used a technique as $^{13}$C NMR spectroscopy. The main factor will undoubtedly be how far the current-day improvements in NMR sensitivity can be extended.

*See also:* **Biochemical Applications of Mass Spectrometry; Enantiomeric Purity Studied Using NMR; Isotopic Labelling in Mass Spectrometry; Labelling Studies in Biochemistry Using NMR; Macromolecule–Ligand Interactions Studied By NMR; Microwave Spectrometers; Solid State NMR, Methods; Stereochemistry Studied Using Mass Spectrometry; Structural Chemistry Using NMR Spectroscopy, Organic Molecules.**

## Further reading

Andres H, Morimoto H and Williams PG (1990) Preparation and use of $LiEt_3BT$ and $LiAlT_4$ at maximum specific activity. *Journal of the Chemical Society, Chemical Communications* 627.

Evans EA, Warrell DC, Elvidge JA and Jones JR (1985) *Handbook of Tritium NMR Spectroscopy and Applications*. Chichester: Wiley.

Floss HG and Lee S (1993) Chiral methyl groups: small is beautiful. *Accounts of Chemical Research* **26**: 116–122.

Junk T and Catallo WJ (1997) Hydrogen isotope exchange reactions involving C–H(D,T) bonds. *Chemical Society Reviews*, 401–406.

Kubinec MG and Williams PG (1996). Tritium NMR. In: Grant DM and Harris RK (eds) *Encyclopedia of NMR*, Vol 8, pp 4819–4839. Chichester: Wiley.

Saljoughian M, Morimoto H, Williams PG, Than C and Seligman SN (1996) *Journal of Organic Chemistry* **61**: 9625–9628.

# Tungsten NMR, Applications

*See* **Heteronuclear NMR Applications (La–Hg).**

# Two-Dimensional NMR, Methods

**Peter L Rinaldi**, University of Akron, OH, USA

Copyright © 1999 Academic Press

**MAGNETIC RESONANCE**
**Methods & Instrumentation**

## Introduction

While NMR has been a valuable tool for scientists who must understand the structures, reactions and dynamics of molecules, there have been two major advances since the 1960s, which, more than any other contributions, have kept this an exciting and rapidly evolving field. The first of these was the introduction of the Fourier transform NMR technique by Ernst and Anderson in 1966. This development helped to reduce problems associated with the biggest limitation of NMR, its poor sensitivity. It also set the stage for a second important development. The dispersion of NMR signals, and thus the complexity of molecules which can be studied is related to the magnetic field strength of the instrument. At a time when scientists were preparing evermore complicated structures, the incremental increases in the magnetic field strengths of commercially available instruments were growing smaller. However, the proposal of Jeener in 1971 and the first demonstration, by Muller, Kumar and Ernst in 1975, of multidimensional NMR spectroscopy resulted in a quantum leap in the capabilities of and the prospects for NMR. By dispersing the resonances into a second frequency dimension additional spectral dispersion could be achieved. The dispersion from a 2D experiment performed on a 1980 vintage 200 MHz spectrometer can match that obtained in the 1D spectrum from modern 800 MHz spectrometers. In 2D spectroscopy, the spectral dispersion increases as the square of the magnetic field strength. Furthermore, 2D experiments can have the unique characteristic of providing structural information based on the correlation of the frequencies at which peaks occur. This article deals with the background and practical aspects of obtaining 2D NMR data.

There are quite a few variations of 2D NMR experiments in which properties such as retention time (in liquid chromatography-NMR, LC-NMR), distances (imaging) or diffusion coefficients (diffusion ordered spectroscopy) are the variables along one or more axes in the spectra.

However, discussions in this article will be restricted to experiments in which two frequencies, related to NMR parameters, are plotted along the two axes of the spectra. Other forms of 2D NMR are discussed in other parts of this work. While this article can be read alone, it is useful to refer to other articles for details of various techniques (e.g. weighting, zero filling, sampling rates, complex versus real Fourier transforms, linear prediction, etc.) which are discussed here as they pertain to 2D NMR.

## Fourier transform NMR spectroscopy

**Figure 1A** shows the time domain signal, called the free induction decay (FID), obtained by measuring the response of nuclear spins to an RF pulse. The FID is the sum of many exponentially decaying cosine waves, one for each resolvable singlet in the spectrum. In the example shown in this time domain spectrum, a single frequency is observed; from measuring its period, the frequency can be determined. A typical FID will contain the sum of many oscillating signal components, making it impossible to identify individual frequency components by visual inspection of the time-domain signal. By converting the time domain signal into a frequency domain signal, using a mathematical process called the Fourier transformation, a readily interpretable spectrum (**Figure 1B**) with peaks at discrete frequencies (one for each cosine wave in the original FID) can be obtained.

Each point in the time domain spectrum contains information about every frequency in the frequency domain spectrum. In a typical 1D spectrum up to 100k points are collected in the time domain, thus information about each peak is measured 100k times. The laws of signal averaging tell us that the signal (S) from $n$ measurements increases linearly ($n \times$ S), but that the noise (N) from $n$ measurements

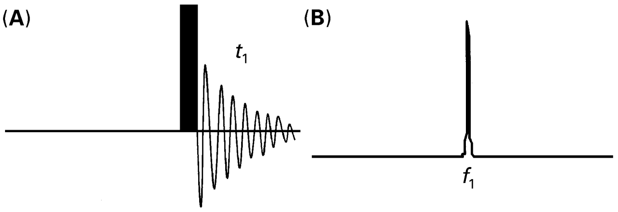



**Figure 1** One-dimensional Fourier transform NMR data: (A) time domain free induction decay (FID) detected after a radiofrequency pulse; (B) frequency domain spectrum after Fourier transformation of the signal in (A).

increases as $n^{½}$ ($n^{½} \times$ N). Therefore, the signal-to-noise ratio (S/N) in the final spectrum improves as $n \times$ S ($n^{½} \times$ N) = $n^{½} \times$ S/N as long as the signal is present throughout the FID. Consequently, the S/N in the final spectrum will be $(10^5)^{½}$ ~300-fold better than that obtained in a single scanned spectrum. This improvement is known as the Felgett advantage. It is described here because it has some important consequences when $n$-dimensional experiments are performed. In practice, S/N gains in 1D NMR are lower than those predicted by the Felgett advantage, because the intensities of the signals decay exponentially during the signal acquisition period. However, in multidimensional NMR, short evolution and acquisition times are used to minimize the size of the data sets. Consequently, very little signal decay occurs and S/N improvement are close to those expected from theory.

## Fundamentals of 2D NMR

### General sequence for collection of 2D NMR spectra

**Figure 2** contains a diagram of a 2D NMR pulse sequence called the NOESY (nuclear Overhauser enhancement spectroscopy) experiment. (NMR spectroscopists have been very liberal in their methods for selecting acronyms to name their experiments). This pulse sequence contains the four basic elements which are common to 2D NMR experiments: preparation, evolution ($t_1$) mixing and detection ($t_2$) times. The filled rectangular boxes represent 90° pulses, which are applied at the $^1$H resonance frequency in this experiment. In general, these pulses can be at a variety of flip angles and can be applied at a variety of frequencies, depending upon the requirements of the experiment and the information desired.

The preparation period is used to put the nuclear spins into the initial state required by the experiment being performed. In this particular sequence the preparation period is a relaxation delay to allow the spins to return to their equilibrium Boltzmann distribution among the energy levels. In some sequences, the preparation period might also contain a coherence transfer step (e.g. by an INEPT-type polarization transfer pulse sequence) to move NMR signal components from one nucleus to another in preparation for the evolution period. The evolution period is used to encode frequency information in the indirectly detected ($t_1$) dimension. The mixing period, which is present in some pulse sequences, is used to transfer magnetization from one nucleus (whose chemical shift information is encoded during $t_1$) to a second nucleus for detection during the acquisition period, $t_2$.

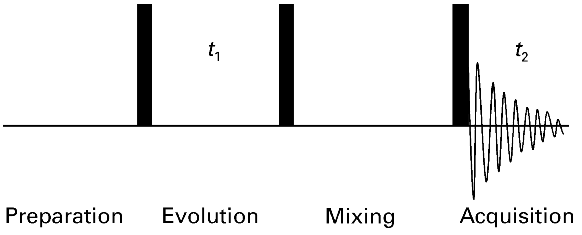



**Figure 2** NOESY 2D NMR pulse sequence.

The NOESY sequence contains a delay during the evolution period to encode $^1$H chemical shifts; however, some pulse sequences contain 180° refocusing pulses to remove chemical shift modulation or coupling to a second nucleus (if the pulse is at the frequency of that second nucleus), or combinations of pulses to remove selected signals or coupling interactions.

The key to the success of 2D NMR is the collection of a series of FIDs, while progressively incrementing the value of $t_1$. At the end of data collection a set of 100–1000 FIDs is obtained (the number of FIDs collected depends upon the desired resolution and spectral window in the $t_1$ dimension) as shown in **Figure 3A**. The intensities of these FIDs are modulated based on the length of the $t_1$ period and the precession of the coherence during $t_1$. If each of these FIDs is Fourier transformed (with respect to $t_2$), a series of spectra is obtained as shown in **Figure 3B**. Each spectrum contains signals which correspond to those found in the normal 1D spectrum of the nucleus which is detected. The intensity of a signal at a specific chemical shift varies from one spectrum to the next. Its intensity is modulated by the NMR interaction ($J$-coupling, chemical shift, multiple quantum coherence, etc.) which is in effect during $t_1$, and by the duration of the $t_1$ period. By plotting the intensities of the two peaks in **Figure 3B** as a function of $t_1$, the curves in **Figure 3C** are obtained. The modulation frequencies of these two curves are different because the detected signals in $t_2$ originate from different coherences which have different precession frequencies in $t_1$.

The intensity variations in these curves are reminiscent of the 1D FIDs. The obvious thing to do with these signals is to transpose the data matrix, and, at each frequency of $f_2$, Fourier transform the data with respect to $t_1$. The result is a spectrum with signal intensity variations as a function of two frequencies as shown in **Figure 3D**. The frequencies plotted along the $f_2$ dimension correspond to those which are detected during $t_2$ (i.e. $^1$H chemical shift and $J$-coupling if the sequence in **Figure 2** is used). The frequency

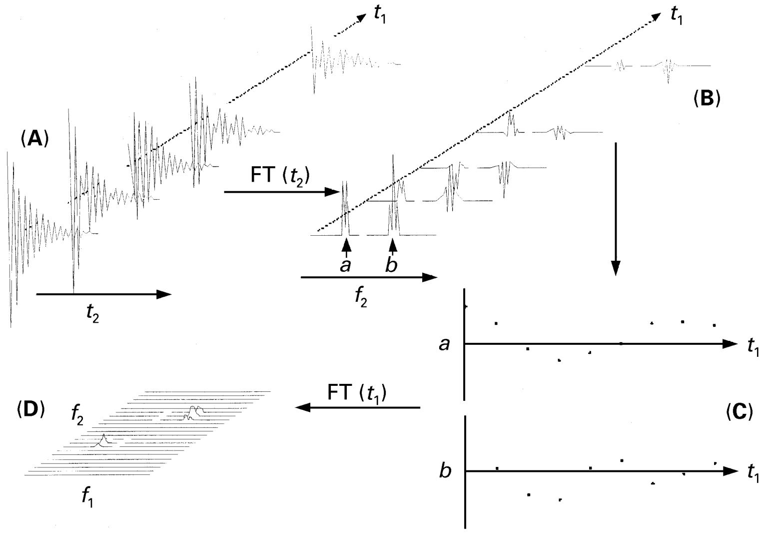



**Figure 3** Schematic illustration of the process used to produce a 2D NMR spectrum.

plotted along the $f_1$ dimension corresponds to the precession properties of the coherences which are selected by the pulse sequence.

## Presentation of 2D NMR data

**Figure 4A** shows a stacked plot of the COSY (correlation spectroscopy) spectrum from ethanol. Detailed fine structure is not resolved in this spectrum because of the greatly reduced digital resolution compared with that obtained in 1D NMR. This reduced digital resolution is not greatly detrimental, and is necessary to keep the 2D data files to a manageable size (see below).

The stacked plot (**Figure 4A**) is not the most desirable way to present the data because it involves a lot of plotting time and background peaks are often hidden by those in the foreground. The preferred method of presenting 2D data is in the form of contours as shown in **Figure 4B**. For those unfamiliar with the generation of contour maps, planes are set at a range of intensity values above a user determined threshold in the spectrum. The contour map is generated by the intersection of the peaks with these

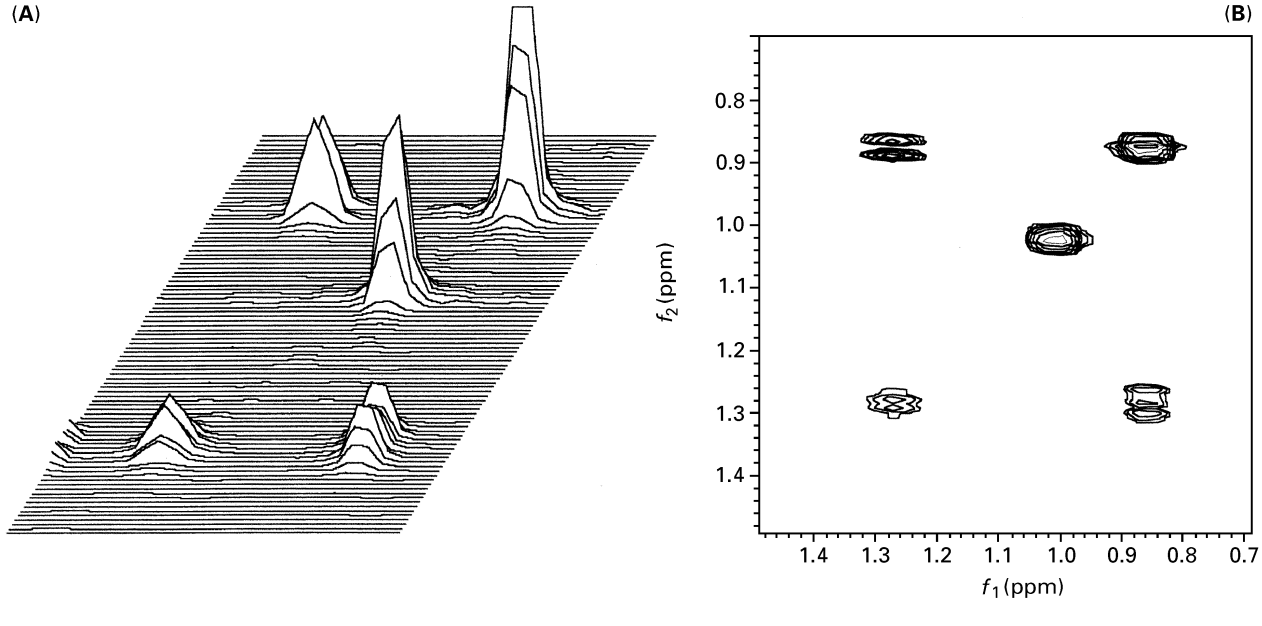



**Figure 4** COSY 2D NMR spectrum of ethanol: (A) stacked plot; (B) contour plot with contours plotted at intensities of $2^n$.

planes. The more intense peaks will intersect a larger number of planes, therefore peaks in **Figure 4B** which are defined by a larger number of contours are more intense than those defined by a small number of contours. In this display mode, peaks are not obscured and the printout is generated fairly rapidly. While it does take a significant amount of computer power to calculate a contour map, modern computers are capable of doing so in much less time than it takes to transmit the data to most plotters. Commercial software packages for manipulating NMR data permit the adjustment of the number and spacing of the contours so as to best display all the peaks in the spectrum. In cases where all the signal intensities are of the same order of magnitude, contour spacing can be small so that a large number of contours accurately defines the peak shapes. In cases where there is a considerable dynamic range of peak heights, contour spacing can be large to prevent the generation of a large number of contours around intense peaks.

## Classes of 2D NMR experiments

Part of the power of 2D NMR comes from its ability to provide tremendous spectral dispersion; however, the structural information present from the correlation of frequencies is equally important. Organic chemists have been performing elegant syntheses for many decades. They choose a target molecule for preparation, and based on their knowledge of chemical reactions select the proper reagents from their stockroom to carry out the chemical transformations necessary to obtain the desired product. Since the development of multidimensional NMR similar possibilities exist for studying molecular structure, reactivity and dynamics. The NMR spectroscopist first defines the nature of the information needed to solve a particular problem. It is then possible to go to the 'NMR stockroom', and choose from a variety of 'NMR reagents', which include pulses, delays, frequencies, RF phases, RF amplitudes, magnetic field gradients, etc. Using the right combination of these reagents, a spectrum can be produced with selected signals which contain the needed information, while removing other undesired signals which interfere with observation of the interesting signals and/or the interpretation of the data.

As an example, **Figure 5** shows a COSY pulse sequence (**Figure 5A**) and the 2D COSY spectrum of heptan-3-one with the 1D $^1H$ spectrum plotted across the top (**Figure 5B**). The 2D spectrum contains a series of peaks along the diagonal whose positions in $f_1$ and $f_2$ correspond to the positions of the peaks in the 1D spectrum. Off the diagonal, cross peaks exist which correlate the frequencies of proton pairs which are coupled to each other. In COSY

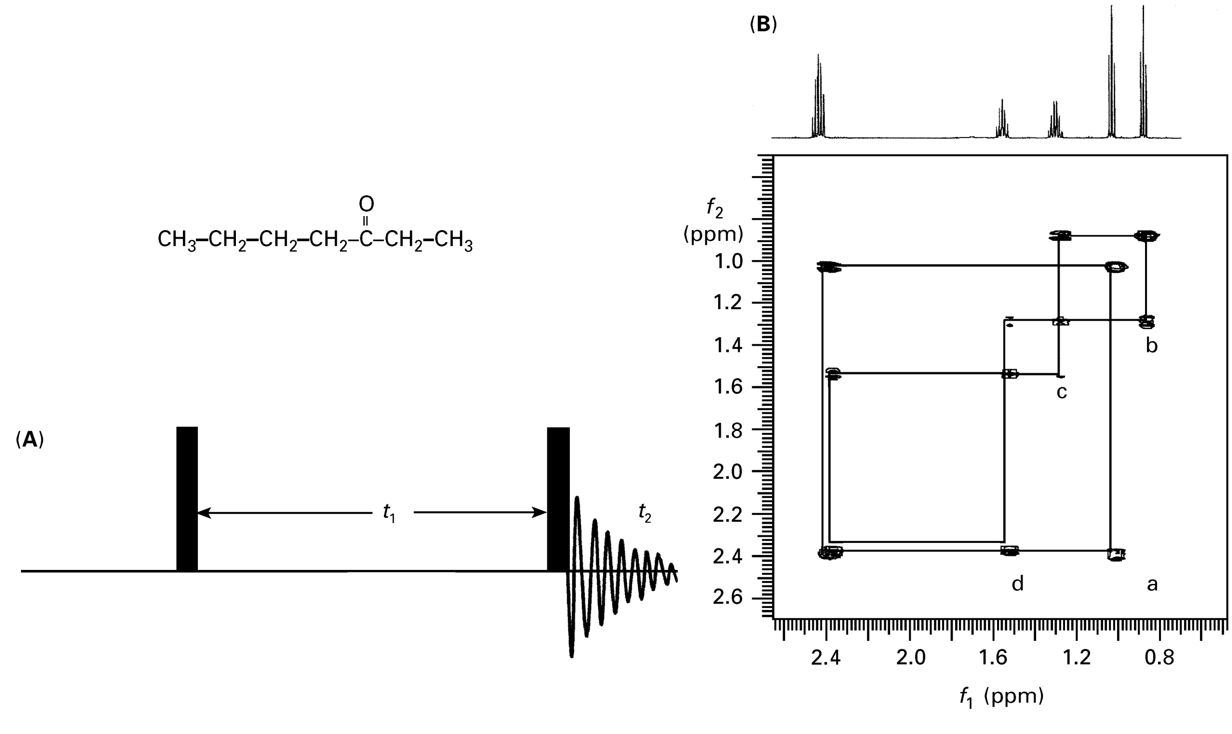



**Figure 5** (A) COSY pulse sequence; and (B) COSY spectrum of heptan-3-one with its 1D $^1H$ spectrum plotted across the top.

spectra, these correlations indicate protons which are on adjacent carbons (or non-equivalent protons on the same carbon). Separate sets of cross peaks are observed for the ethyl (a) and butyl (b, c and d) fragments of the molecule. Because the protons on these two fragments are more than three bonds away from each other, there is no *J* coupling between protons on the two groups. Consequently, none of the resonances from protons on the butyl fragment contain correlations to the resonances of protons on the ethyl fragment.

In general, the NMR parameter plotted along the $f_2$ dimension is related to the signal detected during the $t_2$ time period. The NMR parameter plotted along the $f_1$ dimension is determined by the precession frequencies of the NMR coherences during the $t_1$ period. A specific sequence of pulses can be used to place the NMR coherence on selected nuclei (e.g. $^1H$ or $^{13}C$) to encode their NMR properties during $t_1$; a second series of pulses and delays are then used to transfer that coherence to the detected nucleus based on an NMR interaction (e.g. by *J* coupling, dipolar coupling or internuclear relaxation). Cross peaks in the 2D NMR spectrum identify pairs of nuclei which share this interaction. Some common 2D NMR experiments are shown in **Table 1**, along with the NMR parameters plotted along the $f_2$ and $f_1$ dimensions, the interaction which produces the correlations, the structure information obtained from the spectrum, typical experiment times and some comments.

The experiments can arbitrarily be classified in four groups: homonuclear chemical shift correlation, heteronuclear chemical shift correlation, *J*-resolved, and multiple quantum experiments. The first five experiments in **Table 1** are homonuclear chemical shift correlation experiments. In one subset of homonuclear chemical shift correlation experiments, COSY and TOCSY-type experiments, the same chemical shifts (usually those of $^1H$) are plotted along the $f_1$ and $f_2$ axes. The 2D spectrum contains peaks along a diagonal at the intersections of the chemical shifts of each nucleus. If there is *J*-coupling between two nuclei, then off-diagonal cross peaks connect the diagonal peaks to form a box as in the COSY spectrum of heptan-3-one described above. A second subset of homonuclear chemical shift correlation experiments (NOESY and ROESY) have an appearance identical to that of COSY-type experiments, but with off-diagonal cross peaks that indicate the proximity of two nuclei in space (usually the nuclei must be within 5 Å to produce NOESY cross peaks).

The second group of experiments produces 2D spectra with the chemical shifts of different nuclei along the two axes (e.g. $^1H$ along the $f_1$ axis and $^{13}C$ along the $f_2$ axis). A single cross peak is observed for each coupling interaction in HETCOR, COLOC, HMQC, HSQC and HMBC experiments. The latter three experiments are sometimes put in their own classification, and are called indirect detection experiments. HETCOR-type experiments, which involve detection of the $^{13}C$ signal during $t_2$, were the first

**Table 1** Some common 2D NMR experiments and related information

| *Experiment name* | *$f_2$* | *$f_1$* | *NMR interaction* | *Structure information* | *Time (h)[a]/sample quantity (mg)* | *Comments* |
|---|---|---|---|---|---|---|
| COSY | $\delta_H$ | $\delta_H$ | $^2J_{HH}$ & $^3J_{HH}$ | H–C–H & H–C–C–H | 0.25/1 | Easy |
| Relayed COSY | $\delta_H$ | $\delta_H$ | $^2J_{HH}$ & $^3J_{HH}$ | H's in a spin system | 0.25/1 | Easy |
| TOCSY | $\delta_H$ | $\delta_H$ | $^2J_{HH}$ & $^3J_{HH}$ | H's in a spin system | 0.25/1 | Easy |
| NOESY | $\delta_H$ | $\delta_H$ | H–H dipole–dipole | $r_{HH}$, conformation | 4–12/5 | Usually 10–100 times weaker than COSY |
| ROESY | $\delta_H$ | $\delta_H$ | H–H dipole–dipole | $r_{HH}$, conformation | 4–12/5 | Usually 10–100 times weaker than COSY |
| HETCOR | $\delta_C$ | $\delta_H$ | $^1J_{CH}$ | C–H | 2–8/10 | $^{13}C$ detected |
| Long-Range HETCOR | $\delta_C$ | $\delta_H$ | $^2J_{CH}$ & $^3J_{CH}$ | C–C–H & C–C–C–H | 2–8/10 | $^{13}C$ detected |
| COLOC | $\delta_C$ | $\delta_H$ | $^2J_{CH}$ & $^3J_{CH}$ | C–C–H & C–C–C–H | 2–8/10 | $^{13}C$ detected |
| HMQC/HSQC | $\delta_H$ | $\delta_C$ | $^1J_{CH}$ | C–H | 1/5 | $^1H$ detected |
| HMBC | $\delta_H$ | $\delta_C$ | $^2J_{CH}$ & $^3J_{CH}$ | C–C–H & C–C–C–H | 1–4/5 | $^1H$ detected |
| HOESY | $\delta_C$ | $\delta_H$ | C–H dipole–dipole | $r_{CH}$ | 12/50 | Extremely difficult |
| Homonuclear 2D-J | $\delta_H$ | $J_{HH}$ | $^2J_{HH}$ & $^3J_{HH}$ | Conformation | 0.25/1 | Easy |
| Heteronuclear 2D-J | $\delta_C$ | $J_{CH}$ | $^2J_{CH}$ & $^3J_{CH}$ | Conformation & no. of attached H | 2–8/10 | Moderate |
| 2D-INADEQUATE | $\delta_C$ | $\delta_{Ca}+\delta_{Cb}$ | $^1J_{CC}$ | $^{13}C_a$–$^{13}C_b$ | 12–16/100 | 1 in $10^4$ Molecules extremely difficult |

[a] Typical experimental times for molecule with $M_r = 500$ and experiments performed on a 300–400 MHz spectrometer.

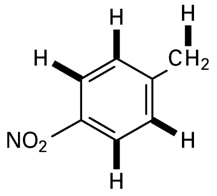
commonly used experiments to provide $^1H$–$^{13}C$ correlations. Later, after the performance of NMR instruments improved, the more sensitive and more challenging $^1H$–$^{13}C$ correlation experiments involving detection of the $^1H$ signal during $t_2$ became popular. In these experiments, the chemical shift of the nucleus X (usually $^{13}C$) is indirectly detected in the $t_1$ dimension. Perhaps, if HMQC-type experiments were popular first, HETCOR-type experiments would now be called indirect detection experiments. The HOESY experiment is the heteronuclear version of the NOESY experiment, and contains cross peaks between the resonances of dissimilar nuclei if there is an NOE interaction between those nuclei.

The third class are *J*-resolved experiments. These produce spectra with the peaks at the frequencies along $f_2$, corresponding to the resonances observed in the 1D spectrum of the detected nucleus. In homonuclear 2D *J*-spectroscopy, the peaks are dispersed into the $f_1$ dimension based on homonuclear *J*-coupling. In heteronuclear 2D J-spectroscopy, the peaks are dispersed into the $f_1$ dimension based on heteronuclear *J*-coupling (e.g. detection of $^{13}C$ in $f_2$ and at the shift of each $^{13}C$ a multiplet, resulting from all of the resolved $J_{CH}$-couplings, is observed in $f_1$).

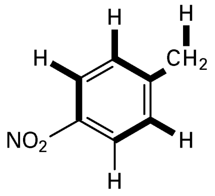
The fourth class of experiments is multiple quantum spectroscopy such as 2D-INADEQUATE. In these experiments, homonuclear shifts (such as those of $^{13}C$) are plotted along the $f_2$ dimension. If two or more nuclei are *J*-coupled to each other, then they can be made to share a common multiple quantum precession frequency during the $t_1$ period. In the 2D-INADEQUATE experiment if $C_A$ and $C_B$ are coupled to each other, then the signals from each of these components will precess at a common double quantum frequency ($\nu_A + \nu_B$) in the $f_1$ dimension.

Usually, two or more experiments are run, where the correlations in each experiment provide a set of structure fragments. The combined fragments can then be fit together like the pieces of a puzzle, and, in most instances, the right combination of multidimensional experiments can provide complete information about the structure of an unknown molecule.

As an example, if HMQC and HMBC spectra were obtained from *p*-nitrotoluene, the HMQC spectrum would provide C–H connectivities, illustrated by the highlighted bonds in structure [1]; HMBC would provide information which relates the $^1H$ shifts with $^{13}C$ shifts of atoms two and three bonds away. Some of these correlations are illustrated in structure [2]. The combined information from the two experiments provides a complete structure of the molecule. While a complete description of these experiments is beyond the scope of this article, some




[1]




[2]

comments on experimental characteristics are worth noting. Experiments which involve $^1H$ detection are generally much more sensitive than those which involve $^{13}C$ detection, largely due to the higher $\gamma$ of $^1H$. Even though HETCOR and HMQC provide similar frequency correlations and identical structure information, the former involves $^{13}C$ detection and generally requires ~30 times more sample to produce a spectrum of the same quality.

Although many of the experiments use similar interactions to provide correlations, experiments which use smaller, long-range *J*-couplings require longer delays (usually ~$1/2J$) than experiments which use large one-bond *J*-coupling. During these longer delays, relaxation effects reduce the intensities of the signals which are finally detected during $t_2$. Consequently, experiments like HMBC produce spectra with poorer S/N than its counterpart, HMQC.

While the entries in the second and third columns of **Table 1** all refer to $^1H$ and $^{13}C$, other combinations of NMR active nuclei can be used to perform most of these experiments. For example, the experiments in the first 5 rows of **Table 1** are $^1H$–$^1H$ homonuclear correlation experiments. These experiments will work just as well with $^{19}F$–$^{19}F$ homonuclear correlation experiments if there are a number of mutually coupled fluorine atoms in the structure to be studied. Likewise, $^{15}N$ could be substituted for $^{13}C$ in HMQC and HMBC experiments.

## Experimental aspects of 2D NMR

### Acquisition conditions

**Instrument requirements** Most instruments which have been installed in the 1990s are capable of performing all of the experiments shown in **Table 1**. The collection of 2D NMR spectra requires a stable instrument and a stable instrument environment. The exact requirements become more stringent at higher magnetic fields. For example, 600–800 MHz spectrometers generally require room temperature fluctuations less than ± 0.5°, draughts should be minimized, and the magnet should be mounted on vibration isolation pads. In some instances it might

be necessary to mount other mechanical equipment near the instrument (near is not used in an absolute sense since some buildings are more efficient at transmitting vibrations throughout the structure than others) on its own vibration isolation pads. All of the experiments shown in **Table 1** can be performed on standard two-channel (i.e. $^1$H and X channels) spectrometers which have been installed since 1990, although HMQC and HMBC experiments might require special accessories in order to run these experiments conveniently.

**Spectral resolution and data size** As mentioned above, a number of separate 1D FIDs are collected, each with a different value for $t_1$. In 1D NMR spectroscopy, 50–100k data points are collected and Fourier transformed to provide a spectrum. The exact number of points depends on the spectral window, expected line widths, and the desired digital resolution (usually 0.1–0.5 Hz per point) in the 1D spectrum. If this digital resolution were maintained in both dimensions of a 2D experiment, the file size could grow to many Gbytes and would be difficult to manipulate and store. Consequently, short cuts are used to minimize the sizes of 2D data files. The first of these shortcuts is to minimize the spectral windows to include only those regions which are expected to contain peaks of interest. For example, in a COSY experiment which contains cross peaks between the resonances of coupled protons, the spectral window is narrowed to exclude singlets and solvent resonances. It is usually worthwhile to collect 1D spectra which correspond to the windows in the two dimensions before attempting to run the 2D experiment. However, this may not be possible in some circumstances, e.g. if sample quantity is limited and the $f_1$ dimension is the $^{13}$C chemical shift in an HMQC experiment.

The second shortcut is to drastically reduce the digital resolution in the 2D spectrum; typically the data is collected to provide 2–4 Hz per point digital resolution in the final spectrum. Typical $t_2$ times are 0.05–0.2 s, an order of magnitude smaller than those used in 1D NMR spectroscopy. The use of longer acquisition times has very little effect on data collection times. If a 1s relaxation delay is used, increasing $t_2$ from 0.05 to 0.2 s results in 15% longer experiment time and provides a four-fold increase in digital resolution in $f_2$ (and a four-fold increase in the size of the data file). The same is not true in the $t_1$ dimension.

Typically, 100–500 separate FIDs are collected to produce a 2D spectrum. To obtain a four-fold increase in digital resolution in $f_1$, four times as many FIDs must be collected, increasing the experiment time by more than four-fold (the $t_1$ period for the last FID will be significantly longer than for the first increment where $t_1 = 0$). If four transients are averaged for each of the 100 FIDs, the S/N in the resulting 2D spectrum would be comparable to that obtained in the corresponding 1D version of the experiment (i.e. only the first $t_1$ increment is collected) obtained by averaging 400 transients. However, in many cases this additional sensitivity is not required and the 2D experiment is longer than is required for signal detection.

**Phase cycling for artefact suppression and coherence selection** There are a number of artefacts which can appear in 2D spectra, including peaks and ridges at the transmitter frequencies in $f_1$ and $f_2$, and mirror images of the real peaks on the opposite side of the spectrum. These can arise from a number of sources, including the fact that some spins experience imperfect pulses. Even with a properly functioning instrument, nuclei whose resonances lie near the edge of the spectral window can experience a significantly different flip angle compared with those nuclei whose resonances lie near the transmitter, as a consequence of resonance offset effects. Additionally, even those nuclei whose resonances fall near the transmitter experience a gradation of flip angles, depending on the position of the nuclei relative to the probe's transmitter coil (i.e. the nuclei in the portion of the sample near the middle of the tube experience a larger flip angle than nuclei in those portions of the sample near the top and bottom of the tube). These artefacts are reduced by using composite pulses in place of simple 180° pulses. To further reduce artefacts, the phases of 180° pulses are shifted by 180° in alternative transients; and the phases of 90° pulses are usually incremented by 90° in a sequence of four transients. A sequence with both a 90° and a 180° pulse requires the averaging of eight transients to obtain a spectrum resulting from all permutation of the two phase cycles. In experiments with many pulses, the number of transients required to cycle the phases of all pulses becomes extremely large (spectra in which the number of transients per FID is 64–256 are typical). Usually, artefacts from imperfections in one pulse are more prominent that those arising from imperfection in the other pulses in the sequence. In those cases, the phases which remove the most severe artefacts are cycled first. When setting up an experiment, it is necessary to know the number of transients needed to complete this minimum phase cycling, and to set up the experiment to collect an integral multiple of this number of transients.

Some of the 2D spectra result from cancellation experiments (i.e. coherence selection). The HMQC is

an excellent example of experiments in this class. As described above, HMQC provides a 2D spectrum correlating the shifts of $^1$H and directly bound $^{13}$C nuclei. In the $^1$H spectrum of heptan-3-one, the peaks which are normally observed are those from $^1$H bound to $^{12}$C (99% of the protons, **Figure 6A**); however, if the vertical scale of the spectrum is increased 100-fold a set of satellite resonances from $^1$H atoms bound to $^{13}$C (1.1% of the signal, **Figure 6B**) are observed. To selectively detect the desired component from $^1$H atoms bound to $^{13}$C, the pulse sequence in **Figure 7A** is used. If the phase, $\phi_1$, of the 90° $^{13}$C pulse is applied along the $+x$- and $-x$-axes on odd and even transients, respectively, the sign of the undesired signals from $^1$H bound to $^{12}$C are unaffected; however, the phases of the desired signals from $^1$H bound to $^{13}$C are altered in odd and even transients. If the FIDs from odd transients (**Figure 6C**) are subtracted from those in even transients (**Figure 6D**) (by altering the phase of the receiver $\phi_3$) the undesired signals cancel and the desired signals add (**Figure 6E**). Detection of the desired signals requires observation of small differences between two large signals. Minor variations in the state of the instrument or its environment will result in imperfect cancellation (as evident by the large residual centre signal in **Figure 6E**) and will produce large noise ridges which obscure the resonances of interest. With a 64-cycle sequence, the residual centre peak can be significantly reduced; however, once an adequate S/N is achieved after a single transient, the experiment must still be run 64 times longer just to complete the phase cycling necessary to remove the artefacts. Furthermore, even when limited sample quantities result in the need to perform signal averaging, the residual signal intensity varies randomly from one pair of transients to the next, and adds like noise. The result is a ridge of noise along $f_1$ at the $f_2$ frequency of intense signals. This noise ridge often called $t_1$-noise, limits the ability to detect weak resonances in the spectrum.

### Pulsed field gradients for coherence selection

Pulsed field gradients (PFGs), also known as gradient enhanced spectroscopy (GES), can be used to achieve coherence selection and minimize the need for extensive phase cycling. In PFG spectroscopy a large $z$-gradient is introduced along the sample's vertical axis (magnitude ~0.1–0.5 T m$^{-1}$ and duration ~1 ms); additional PFGs are introduced later in the sequence to selectively refocus the coherence components of interest and continue to destroy coherence components which are undesired. The spectrum in **Figure 6F** was obtained by collecting a single HMQC transient with the aid of PFG coherence selection. The residual centre peak from $^1$H–$^{12}$C fragments is completely suppressed in one transient. The first obvious advantage of PFG spectroscopy is that excellent coherence selection is obtained in a single transient. Under these circumstances, the number of transients per FID is

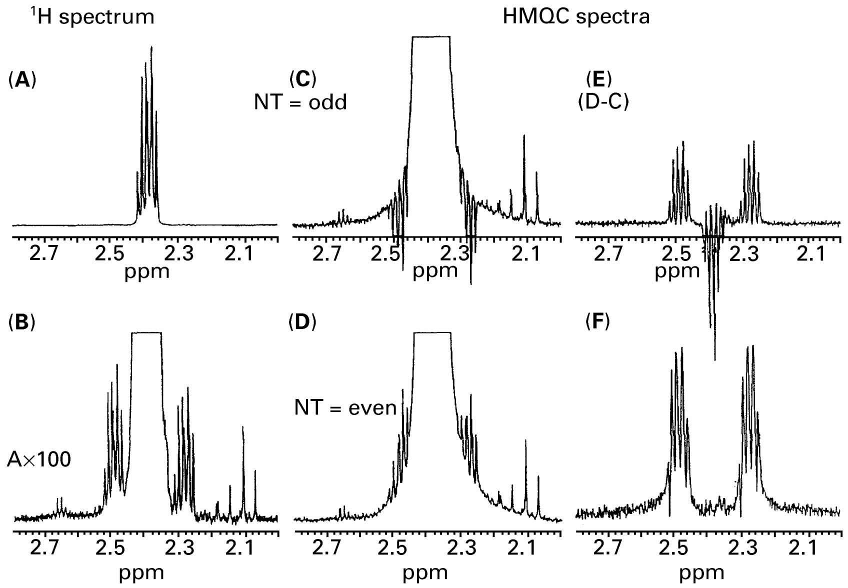



**Figure 6** Methylene regions between 2.2 and 2.5 ppm from the proton spectra of heptan-3-one. (A and B) normal $^1$H spectra; (C and D) HMQC spectra; (A) $^1$H spectrum; (B) 100× vertical amplification of (A); (C and D) spectra obtained from collecting a single HMQC transient with phase cycling for odd and even transients, respectively; (E) is spectrum in (D) minus spectrum in (C); and (F) single transient from HMQC spectrum obtained with PFG coherence selection.

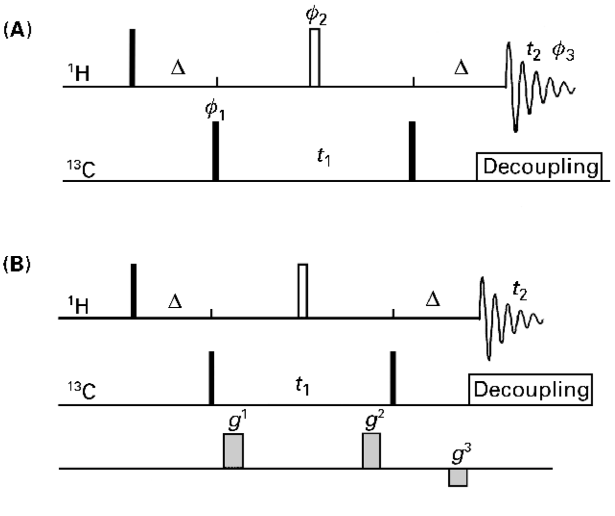



**Figure 7** (A) HMQC pulse sequence with phase cycling for coherence selection; (B) PFG-HMQC pulse sequence. In these diagrams, the filled rectangles are 90° pulses, the unfilled rectangles represent 180° pulses and the shaded rectangles represent field gradient pulses.

determined by sensitivity requirements and not the need to complete a phase cycle. Additionally, in cancellation experiments, the suppression level is less sensitive to instrument instabilities; and because the large undesired signal component never reaches the receiver, instrument gain settings can be optimized for detection of the weak signals of interest.

**Quadrature detection in $f_1$ and $f_2$** When collecting 1D spectra on modern instruments two detection channels are present which independently measure the signals 90° out of phase with respect to one another (**Figure 8A** below). Two FIDs, a real component and an imaginary component (which is 90° out of phase with respect to the real component of the signal), are saved; frequency information is obtained from the real component, and phase information (i.e. whether the signal is to the left or right of the transmitter) is present in the imaginary component. A complex Fourier transformation produces a spectrum that shows peaks with the proper relationship with respect to the transmitter, depending on the relative phase of the imaginary component in **Figure 8A**.

In 2D NMR, it is not possible to use a second detector in the $f_1$ dimension. There are alternatives which provide the equivalent of phase sensitive detection in the $f_1$ dimension; two of these are the States method and time proportional phase incrementation (TPPI).

In the TPPI method a single data set with 512 $t_1$ increments is collected. In each successive $t_1$ increment the phase of the 90° pulse at the end of the $t_1$ period is incremented by 90° with respect to the phase of the corresponding pulse in the previous $t_1$ increment. (An equivalent experiment can be performed in which the phases of the pulses before the $t_1$ period are shifted by 90°). This is equivalent to changing the reference frame in $f_1$ so that the transmitter in the $t_1$ dimension appears to be shifted to one edge of the spectrum. After performing a real Fourier transformation, all peaks will appear to be shifted to one side of the transmitter in $f_1$. The main disadvantage of this technique is that phase distortions can appear for resonances in strongly coupled spin systems.

To obtain true quadrature detection two sets of data with real and imaginary components in $t_1$ must be obtained. In the States method, two set of 2D FIDs are collected and saved. Both sets of data might contain 256 FIDs (2 × 256) and the $t_1$ delays in the corresponding FIDs in the two data sets are identical. Their only difference is that the second set of FIDs is collected with the phase of the 90° pulse immediately after the $t_1$ period shifted by 90° compared with the phase of the same pulse in the first set of FIDs. The first set of FIDs contains the frequency

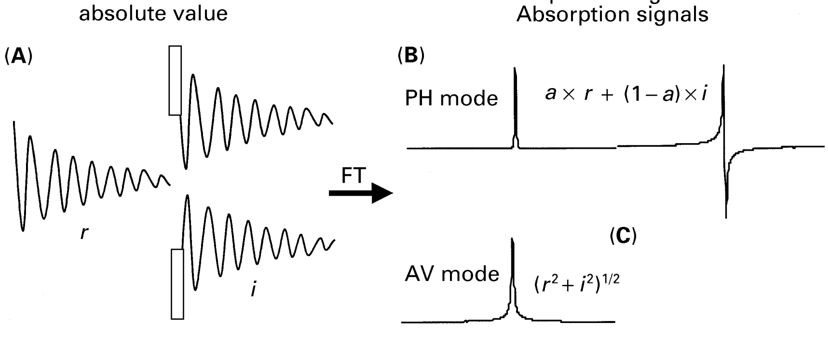



**Figure 8** (A) Schematic illustration of real ($r$) and imaginary ($i$) components of the spectrum before FT ; (B) after FT the data can be represented in phase sensitive (PH) mode, or (C) an absolute value (AV) mode spectrum can be calculated.

modulation information in $t_1$ and the second set of FIDs contains the phase information in $t_1$, similar to quadrature detection in 1D NMR. After each of the FIDs in the two data sets are Fourier transformed, corresponding points from the two data sets are paired to form complex points in $t_1$ and a complex Fourier transformation is performed with respect to $t_1$. This latter method provides data sets which are identical in size (and digital resolution) to those obtained from the TPPI method with equivalent digital resolution. The States method of phase sensitive detection is usually preferred because artefacts are less problematical.

**Establishing a steady state** Ideally, a relaxation delay of 3–5 × $T_1$ should precede the cycle of pulses used to collect each transient. However, this would make experiments impractically long. Normally, 1–2 s relaxation delays are used, even though $T_1$ values, might be 10 s or more. For larger molecules with shorter $T_1$ values, relaxation delays of 100–500 ms are used. Under these conditions, incomplete relaxation occurs. Consequently, the intensities of the signals in the first $t_1$ increment are artificially enhanced relative to the same signals in later increments. This means that the first point in $t_1$ is offset, leading to a large zero frequency offset in the baseline at slices corresponding to the resonance frequencies of all peaks in $t_2$. To eradicate this problem, it is customary to perform 8–32 dummy scans, which are discarded, before collection of the data to be saved. During the dummy scans 'steady state' magnetization is established before data collection commences.

### Data processing

**Zero filling** Most of the effort needed to produce quality 2D spectra from a working spectrometer occurs after the data is collected. With the many data processing techniques available, even poor data can frequently be worked up to produce a useful spectrum. Two basic operations which are common to 1D NMR are zero filling and mathematical weighting of the data to improve resolution or S/N. In 2D NMR it is almost always desirable to add zeros in the $t_1$ and $t_2$ dimensions so that the linear dimensions in each of these directions is 2–4 times larger than the data collected. For example, if the States method was used to collect 2 × 256 FIDs with 1024 points (in $t_2$) per FID, it is useful to zero fill the data so that Fourier transformation is performed on a 2 × 1024 × 2048 matrix. The final displayed spectrum after phasing will be a 1024 × 1024 matrix. It is useful to optimize other processing conditions described below on the smaller data set without zero filling to shorten processing times for the many iterations needed. After this is accomplished, zero filling should be used to produce the best quality spectrum for final display.

**Absolute value versus phase sensitive display** The selection between absolute value (AV) and phase sensitive (PH) display modes is governed by the nature of the data collected. The 1D NMR spectra obtained from modern instruments are always phase sensitive as illustrated in **Figure 8**. As described above, two FIDs are collected, a real (*r*) FID and an imaginary (*i*) FID with a 90° phase shift between them (**Figure 8A**). The phase of the imaginary component can be +90° or −90° with respect to the real component. When a complex FT is performed, the phase information [i.e. whether the imaginary component is a +sin(ω*t*) or a −sin(ω*t*) function] determines the direction of the frequency offset relative to the transmitter. After FT and phasing, if the PH display mode is chosen, two components of the data are obtained (**Figure 8B**). The real component is a pure absorption mode signal which is usually displayed, and the imaginary component is a dispersion spectrum which is either hidden from the operator or disposed. In cases where the spectrum is difficult to phase, the real and imaginary components are combined point by point to generate an AV mode signal as illustrated in **Figure 8C**. This display mode has the advantage that phasing is not required. However, because the resulting spectrum contains both absorptive and dispersive components the peaks are much broader. For this reason, the AV mode display is rarely used in 1D NMR.

In 2D NMR, the use of AV mode display is more common. If phase sensitive detection is not used and/or if the peaks in the spectrum are phase modulated, then it becomes necessary to use an AV mode display. In addition, in 2D NMR there are twice as many phase parameters to adjust, making the phase correction procedure somewhat cumbersome if the right experiment delays were not used to obtain the data. Under these circumstances cross peaks in the spectrum are dispersive rather than pure absorption, and it is necessary to display the data in an AV display mode.

**Weighting** In 1D NMR the selection of weighting functions is based on the desired tradeoff between resolution and S/N ratio. In 2D NMR, the selection of weighting functions is based on the nature of the experiment, the display mode (AV versus PH), the desired resolution and the desired S/N. An entire article could be written to describe the various weighting functions which have been developed. Rather than discuss all of these, it is useful to break them up into several groups based on the general effect they

have on the spectrum and show a limited number of representative weighting functions. These are shown in **Figure 9** along with two types of FIDs which are commonly obtained in 2D NMR.

The groups shown in **Figure 9B** and **9C** are used to provide a smooth decay at the end of a truncated FID to minimize truncation artefacts in the spectrum. These functions are generally used when the FID is similar in shape to the one shown in **Figure 9D**. The use of an exponential decay function (**Figure 9A**) is not common in 2D NMR because it broadens the base of the peak and results in long ridges/tails at the base of intense signals. These tails would obscure other weak cross peaks that fall at the same frequency. It also produces severe line broadening under conditions which provide noticeable smoothing of the FID. If smoothing is desired, the functions in **Figure 9B** and **9C** accomplish this with minimal perturbation of the first half of the signal.

The group of weighting functions (**Figures 8E–G**) is used to provide resolution enhancement when the data must be displayed in AV mode. In addition, some 2D experiments such as HMBC produce FIDs which are echos, like the one shown in **Figure 9G**. Under these circumstances it is usually desirable to match the weighting function to the echo (by adjusting parameters which control the width and displacement of the maximum of the weighting function) so that their maxima coincide and their initial buildup and later decay rates are matched. It is common to use different types of weighting functions in the $t_1$ and $t_2$ dimensions.

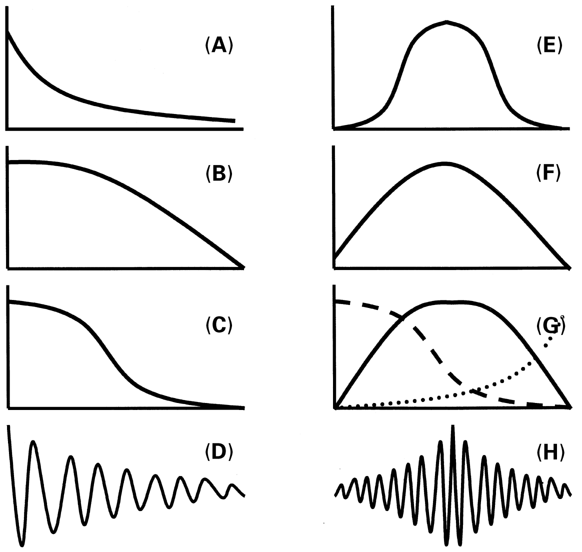



**Figure 9** General shapes of various weighting function used to massage NMR data and typical FIDs: (A) exponential decay, (B) shifted sine function, (C) shifted Gaussian function, (D) typical FID in which acquisition has been truncated before the signal decays, (E) Gaussian function, (F) sine function (G) product of exponentially increasing and shifted Gaussian functions, and (H) echo signal FID.

**Linear prediction** Limited access to instrument time and the volume of data which must be collected require that short cuts, which adversely affect the appearance of the spectrum, must always be taken. Digital signal processing techniques can significantly enhance the appearance of a spectrum without increasing data collection times. In some cases, when instrument time is at a very high premium, it might even be desirable to deliberately reduce the experiment time below the minimum needed for a reasonable spectrum, knowing that processing techniques can be used to compensate for the lost data. Mathematically, it is possible to use the behaviour of a function during time $t$, during which a measurement is made, to predict the behaviour of the function if the measurement time had been extended by time $t'$ (**Figure 10A**). As of writing, in multidimensional NMR, linear prediction is the most used and most useful of these mathematical methods. Essentially, the oscillatory behaviour of the signal intensity as a function of $t_1$ (at a specific $f_2$) is fitted to the sum of a series of cosine waves. Since the number of peaks present at a single $f_2$ in a 2D spectrum is relatively small, the sum of a relatively small number of frequency components is sufficient to stimulate the behaviour of the FID in $t_1$. The function can then be used to artificially synthesize values for the FID in $t_1$ as if a much larger number of $t_1$ increments had been collected. Usually the data is increased to 2–4 times the original size, and zero filling is applied to double the length of the data (e.g. if $2 \times 256$ $t_1$ increments were collected, linear prediction would be used to forward extend the data to $2 \times 1024$ and zero filling could be used to further increase the size in $t_1$ to $2 \times 2048$). This permits improvements in resolution comparable with what would be achieved from an experiment that is up to four times longer than the actual experiment time.

Linear prediction can also be used to remove experimental artefacts from data. For example, the intensity of a single FID in the middle of a 2D experiment (**Figure 10B**) could be distorted if the field were perturbed for some reason. If steady pulses were not applied at the beginning of the experiment, the first few points in $t_1$ might be more intense than they should be (**Figure 10C**). In the former case, with linear prediction, the behaviour of the FID on the either side of the distorted point could be used to approximate the correct value of the distorted point.

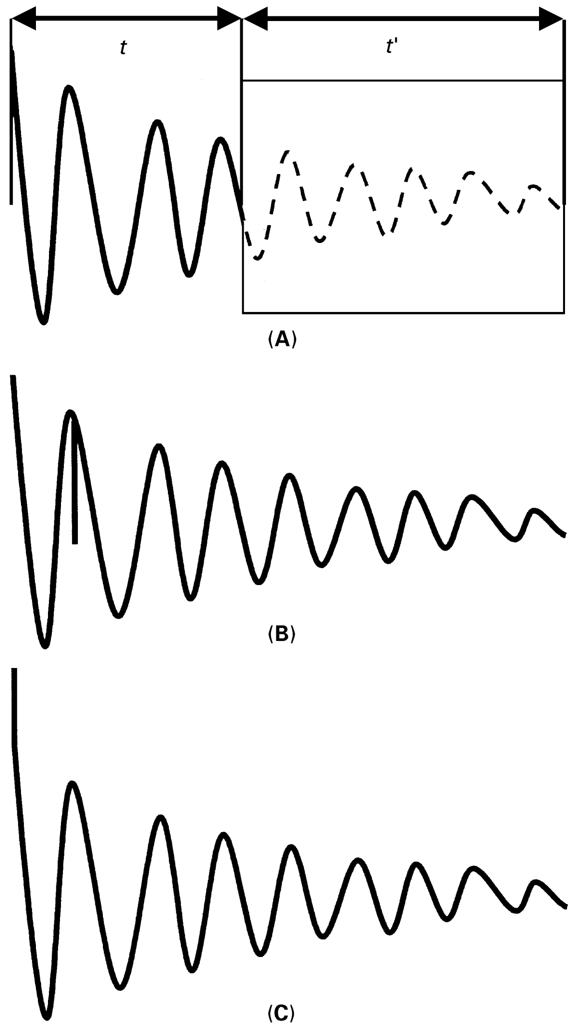



**Figure 10** (A) Time domain NMR signal (—) detected (---) calculated using linear prediction. (B) FID with a distorted point in the middle. (C) FID with the first few points distorted by pulse breakthrough.

In the latter case, if the first two points are distorted, the behaviour of the function for points 3–10 could be used to back predict the proper value of the points in the beginning of the FID.

## List of symbols

$t_1$ = evolution time; $t_2$ = detection time; $T_1$ = relaxation time; $\gamma$ = gyromagnetic ratio.

*See also:* **Diffusion Studied Using NMR Spectroscopy; Macromolecule–Ligand Interactions Studied By NMR; Magnetic Field Gradients in High Resolution NMR; NMR Data Processing; NMR Pulse Sequences; Nuclear Overhauser Effect; Nucleic Acids Studied Using NMR; Product Operator Formalism in NMR; Proteins Studied Using NMR Spectroscopy; Solid State NMR, Methods; Structural Chemistry Using NMR Spectroscopy, Organic Molecules; Structural Chemistry Using NMR Spectroscopy, Peptides; Structural Chemistry Using NMR Spectroscopy, Pharmaceuticals.**

## Further reading

Aue WP, Bartholdi E and Ernst RR (1976) Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *Journal of Chemical Physics* **64**: 2229.

Berger S (1997) NMR techniques employing selective radio frequency pulses in combination with pulsed field gradients. *Progress in NMR Spectroscopy* **30**: 137.

Bovey FA and Mirau PA (1996) *NMR of Polymers*. Academic Press.

Cavanagh J, Fairbrother WJ, Palmer III AG and Skelton NJ (1996) *Protein NMR Spectroscopy Principle and Practice*. Academic Press.

Clore GM and Gronenborn AM (1991) Application of three- and four-dimensional heteronuclear NMR spectroscopy to protein structure determination. *Progress in NMR Spectroscopy* **26**: 43.

Croasmun WR and Carlson RMK (1994) *Two-Dimensional NMR Spectroscopy Applications for Chemists and Biochemists*. New York: VCH.

Ernst RR and Anderson WA (1966) Application of Fourier transform spectroscopy to magnetic resonance. *Review of Scientific Instruments* **37**: 93.

Freeman RA (1997) *A Handbook of Nuclear Magnetic Resonance* 2nd edn. Essex: Longman.

Griffiths PR (1978) *Transform Techniques in Chemistry*. New York: Plenum Press.

Jeener J (1971) *Abstracts AMPERE International Summer School*, Basko Polje, Yugoslavia.

Martin GE and Zektzer AS (1988) *Two-Dimensional NMR Methods for Establishing Molecular Connectivity*. New York: VCH.

Muller L, Kumar A and Ernst RR (1975) Two-dimensional carbon-13 NMR spectroscopy. *Journal of Chemical Physics* **63**: 5490.

Schmidt-Rohr K and Spiess HW (1994) *Multidimensional Solid State NMR and Polymers*. New York: Academic Press.

# U

## UV Spectroscopy of Biomacromolecules

*See* **Biomacromolecular Applications of UV-Visible Absorption Spectroscopy.**

## UV Spectroscopy of Dyes and Indicators

*See* **Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy.**

## UV-Visible Absorption and Fluorescence Spectrometers

**GE Tranter**, GlaxoWellcome Medicines Research, Stevenage, Herts, UK

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
|---|
| Methods & Instrumentation |

UV-visible absorption and fluorescence (together with phosphorescence) spectrometers work with light in the wavelength region extending from the far ultraviolet (175 nm) to beyond the red end of visible light (900 nm). Often, the highest specification absorption instruments are able to extend their range into the near-infrared region (NIR). Many of the elements of UV-visible absorption spectrometers likewise appear in the analogous fluorescence spectrometers and can be conveniently described alongside each other.

### UV-visible absorption instruments

Typically, in a scanning UV-visible absorption spectrometer, light from a suitable source in transmitted through a monochromator (or filters in a low specification instrument) to yield light of the desired wavelength. This is then passed through the sample and thence to a detector (**Figure 1A**). As the monochromator in scanned through its wavelength range so a spectrum is measured from the detector's response. To improve spectral resolution and accuracy, the highest specification instruments have two, or even three, monochromators in series.

In contrast, diode array based instruments, popular as relatively low cost spectrometers, have a reverse arrangement by having the dispersion of wavelengths post-sample by a dispersive optic (e.g. a diffraction grating) which irradiates the diode array detector with the spectrum across its elements. Whilst commercially available diode array instruments are invariably of lower optical quality than the best scanning instruments, and are limited by their spectral resolution through the number of elements in their array, they do offer the advantage of rapid spectral acquisition as the complete

spectrum across the wavelength range may be acquired almost simultaneously.

A similar 'reverse optics' arrangement, with a post-sample monochromator followed by a detector has infrequently been incorporated into scanning instruments, albeit rarely on a commercial level (**Figure 1B**). Reverse optics instruments, whether diode array or scanning, can be more prone to inducing sample photodecomposition and other photoreactive phenomena, as the full intensity of the light source, unattentuated by a monochromator, is incident on the sample.

As the light source and all of the optical components of an instrument will have a wavelength dependence it is necessary to acquire a 'background' spectrum in the absence of a sample with which to correct an acquired sample spectrum. If one wishes to correct simultaneously for the optical properties of a moiety in the sample, such as a solvent is solution studies or the vessel in which it is contained, then this 'reference' may replace the sample when acquiring the background. Although the background and sample spectra have to be acquired separately (although ideally at proximate times) in a 'single beam' instrument, many spectrometers have a 'dual beam' configuration that enables both to be acquired together (**Figure 1C**). In this case, the light is split into two equivalent beams prior to the sample. One beam passes through the sample, as in the single beam case, whereas the other passes through the reference. In most current instruments the two beams are then recombined onto a single optical path to the same detector. The resulting two signals are distinguished by alternately obscuring the beams through the use of 'choppers' (typically a rotating disk with apertures), rotating mirrors (which may be used to generate the two beams alternately) or a combination of the two. Nonetheless, for studies of the highest precision, a sequential background should be acquired to ensure complete correction for the sample optical path, which will inevitably differ slightly from that of the reference path in a dual beam arrangement.

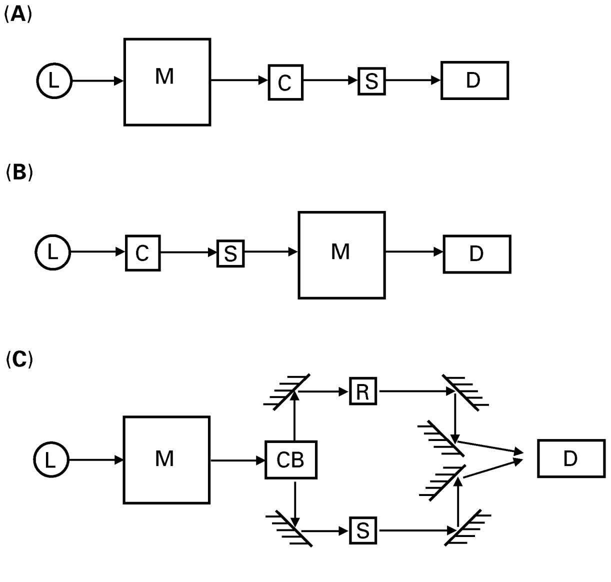



**Figure 1** Conceptual diagrams of absorption spectrometer configurations: (A) single beam (B) single beam reverse optics and (C) dual beam. L = light source(s), M = monochromator(s), C = chopper, B = beam splitter, R = reference sample, S = test sample and D = detector.

The use of oscillating signals brings a further benefit in the ability to use AC rather than DC detection circuitry, where phase lock amplification can be used to best advantage, particularly if a dark period during which light is completely obscured from the detector is also introduced to provide a zero transmission level. This advantage is equally applicable in single beam configurations and therefore is to be found in virtually all scanning instruments. The drawback of employing oscillating signals is that rapid phenomena such as in kinetic studies may be immeasurable on the oscillation time-scale. In these cases a dual beam configuration with two separate detectors and continual illumination may be more appropriate.

Instruments are generally limited by the maximum absorption (measured in absorption units, AU) they can measure due to stray light (see later) and signal-to-noise concerns. Typical limits are, in practice, up to 1 AU for diode arrays and single monochromators, up to 2.5 AU for routine double monochromators and up to 4 AU for the highest specification double monochromator instruments. A once popular technique, now reappearing in the latest instrumentation, for aiding the precise absorbance determination of highly absorbing samples is to introduce an attentuator into the reference beam in a double beam configuration. In so doing, the relative absorbance of the sample to the reference in reduced to around 1 AU, so achieving near optimal conditions regarding signal-to-noise. However, stray light will continue to exert its effects.

As with all instrumentation, proper care and maintenance are essential for correct functioning. Notwithstanding this, UV-visible absorption spectrometers are very reliable if a few precautions are taken with their operation. In particular, instruments should sited in a vibration, vapour and dust free environment with low humidity and minimal temperature variation. For use in the UV region it is necessary to purge the instrument with dry evaporated nitrogen to exclude oxygen, which absorbs below 200 nm and will otherwise generate reactive ozone that damages the optical components. The discipline of continuously purging with nitrogen, whatever the wavelength

region, protects the optical components from airborne contaminants and prolongs their life.

### Light sources

The key requirements for a light source in UV-visible instruments are an adequate coverage of the spectral range with sufficient intensity together with stability of output. Deuterium arcs and tungsten filament lamps are frequently employed in tandem for the lower (180–350 nm) and higher (330–900 nm) wavelength region, respectively. Alternatives for the higher wavelength region include quartz halogen lamps, which are essentially a tungsten filament in a halogen atmosphere within a quartz envelope. These operate at higher temperatures than traditional tungsten filaments, the halogen prolonging the life of the lamp through reacting with vaporized tungsten to minimize blackening of the envelope.

For the complete wavelength region (175–1000 nm) high intensity xenon arcs may be employed. The output of such arcs, particularly those above 100 W, are of great value in fluorescence instruments where high intensity is a prerequisite, but are usually considered unnecessarily powerful for absorption spectrometers, particularly given their heat generation. Other alternatives, currently for specialist use, include tuneable lasers (which may obviate the need for monochromators) and pulsed arc lamps (which give a broad band of intense radiation, like steady-state arcs, but with much reduced heat generation).

### Monochromators

The simplest of wavelength selectors is a set of optical filters in a movable mount such that an individual filter may be placed into the optical path. However, they are wholly inadequate for spectroscopy except for the most basic of instruments. Nonetheless, they may be fruitfully employed to prefilter light prior to a monochromator in order to reduce stray light and attenuate the incident radiation of the optical elements.

The simplest of monochromators consist of a rotatable diffraction grating or prism together with a number of mirrors to guide the beam from the entrance to the exit slit/aperture. As the grating or prism is rotated, so the wavelength of light issuing from the exit slit varies. The size of the slits through which the light is constrained, coupled with the dispersion of the monochromator's optics, determines the 'spectral bandwidth' (SBW) of the light produced. Light nominally of wavelength $\lambda$ is better considered as having a distribution of wavelengths, the width of the distribution about $\lambda$ being given by the SBW. Typically, single monochromators may achieve SBWs of 5–10 nm, whereas for accurate spectroscopy SBWs of the order of 1 nm or less may be appropriate. By coupling monochromators in series the SBW may be reduced to a more suitable figure and likewise improve stray light performance (see later), albeit with a commensurate loss in light intensity. Alternatively, SBW may be reduced by simply increasing the dimensions of a single monochromator such that the physical linear dispersion becomes larger. However, this can lead to impracticalities and problems with focusing of the beam. Whatever the design, the optimal configuration of components and the use of baffles to deflect and absorb unwanted light is of critical importance.

Diffraction gratings have the advantage of ease of manufacture and relatively constant dispersion with wavelength (i.e. the wavelengths are evenly spread out by the grating). Generally holographic gratings have a higher precision than ruled gratings, although the latter may give greater light throughput away from their central wavelength. With diffraction gratings the slits may be kept of fixed size to achieve a given SBW independent of the wavelength. In contrast, the dispersion of a prism is highly wavelength dependent, requiring the coupling of the wavelength drive (prism rotation) to the slit mechanisms. However, prisms do have useful transmission and polarization qualities that are of value in other optical instrumentation such as circular dichroism spectrometers.

### Sample compartment

Solution studies are typically carried out using fused quartz rectangular cuvettes (cells) which are, for historical reasons, of 1 cm pathlength. However, many alternative variants are available, including those of cylindrical construction, thermostatted, flow cells, micro cells and a plethora of specialist types. In particular, pathlengths from 0.001 cm to 10 cm are readily available from commercial sources, allowing the study of a wide range of sample concentrations and quantities whilst avoiding problems of excessive or too little absorption. Whatever the cell construction, for reliable results cells must be located in a fixed, reproducible, orientation and position in the sample compartment. Many instrument manufacturers provide a wide range of attachments for controlling the sample, such as thermostatted cell holders, stirrers, sippers and cell autochangers.

Likewise, there are a vast array of attachments for solid samples, whether for transmission studies or for surface reflectance investigations. In particular, many attachments provide either a method of integrating total reflectance (as in integrating spheres to surround the sample) or to orientate the surface with respect to the light beam and detector so as to probe specular and diffuse reflectance.

Recent advances have been made in employing optical fibres to allow the study of samples remote from the spectrometer. Essentially a fibre optic redirects the light from the sample compartment to the external sample, with another returning the resultant beam back to the detector. By these means absorption can be monitored by either transmission or reflectance, using remote cells, surface probes and submersible probes.

### Detectors

The detector in a standard UV-visible absorption spectrometer is most frequently a photomultiplier tube (PMT) or a silicon diode, the latter being extended to an array in diode array instruments. PMTs have a greater sensitivity and are employed in the most demanding research grade instruments. However, silicon diode devices are considerably smaller, cheaper and do not require the high voltages necessary with PMTs. Both PMTs and silicon diodes have wavelength dependencies which may also dictate their specific use. More recent detectors include charge coupled devices (CCD) and photomultiplier arrays, which will no doubt become more commonplace in the future.

### Calibration

The two axes of an absorption spectrum, namely the wavelength (or correspondingly the energy, frequency or wavenumber) and the absorption (or transmission or intensity), dictate that these two scales of an instrument be calibrated. The wavelength scale calibration is typically accomplished by the use of either a series of line emissions from discharge lamps, the precise wavelengths of which have been tabulated, or through standard filters with known absorption spectra. For general convenience the filter method is the one of choice. The most common filters used are those of holmium oxide or didymium (a mixture of neodymium and praesodymium) oxide in glass. However, these can be difficult to produce consistently and can show variations of some ±4 nm in peak positions at the long wavelength end of their useful range (240–685 nm). Therefore, for accurate calibration, it is necessary to use filters provided with a table of determined peak positions from a reputable source such as a national laboratory for standards. As an alternative to glasses, solutions containing lanthanide ions have proved useful, with less variation but a corresponding decrease in convenience.

Historically, potassium dichromate has been the most extensively employed standard for calibrating the absorbance scale. However, in solution numerous species exist in a series of complex equilibria that are sensitive to pH, concentration and other environmental factors. Consequently, various organic and inorganic compounds have been investigated as alternatives, although many have other complicating features for calibration to the highest accuracy. Solutions also hold the problems of being a test not just of the instrument, but of the laboratory skills of the scientist preparing them. At present, the most practical method of calibration is through the use of neutral density filters, whose absorbance has been established by a reputable source.

### Stray light

One of the main reasons for an apparent deviation from the Beer–Lambert law for absorption, excluding chemical phenomena specific to a sample, is the effect of stray light. In an ideal spectrometer, only light of the correct wavelength (within the spectral bandwidth window) that has impinged upon the sample would reach the detector and be monitored. Any additional sources of light detected in a real spectrometer may be thought of as 'stray light'. Broadly, there are five potential sources of stray light: (i) sample fluorescence/phosphorescence/luminescence etc, (ii) ambient light leakage into the instrument, (iii) transmission of light not through or from (in the case of reflectance) the sample, (iv) imperfections in the monochromator and light source and (v) imperfections in the detector optics.

The first of these, emission by the sample, when it does occur is invariably weak and would only cause problems in the most precise studies or extreme cases. As a molecular phenomenon specific to the sample it is not within the realms of 'instrumental' stray light and must be considered on a case-by-case basis.

The second two sources are manifestations of poor instrumental design; instruments should be light tight and the sample should be sufficiently masked in a blackened compartment to ensure that only light impinging on the sample reaches the detector. This latter condition is sometimes unfortunately overlooked by instrument manufacturers, who may, for example, introduce reflective components in the sample compartment, or cell holders that do not fully mask the cell to within its useable aperture and beyond the dimensions of the light beam.

Finally, the last two sources are, to some degree, unavoidable instrumental stray light. Nonetheless, they can be minimized through careful design and maintenance. Imperfections in the optical surfaces and compromises in the positioning of components in the monochromators, and elsewhere, give rise to unwanted reflections or dispersion. In particular, diffraction gratings are not perfect and furthermore, even in ideal circumstances, they generate repetitions of the wavelength range. Thus the choice of optimal component

configurations, light baffles and component quality is crucial to the stray light performance.

Reverse optics instruments may similarly exhibit stray light introduced at the detector, post sample. In particular, diode array instruments may suffer through internal reflections in the optical surface covering the array, leading to apparent illumination of the incorrect array elements.

### Polarization

All the optical components, particularly the diffraction grating or prism and the light source, cause the light beam to the polarized. For the study of isotropic samples, with no preferred orientation, via transmission methods this is of little consequence. This is not the case for non-isotropic materials, such as crystals and ordered solids, or reflection measurements where one encounters linear dichroic effects. To avoid polarization artefacts it is necessary to insert depolarizing optics at the appropriate positions in the optical path. However, care must be taken to choose a depolarizer that truly depolarizes at each wavelength, rather than one that simply gives a different, but specific, polarization at each individual wavelength (these are intended for use with white light applications). Suitable depolarizers are often based upon multiple scattering (*ala* frosted glass) which, in turn, may give additional stray light concerns.

## Fluorescence spectrometers

Flourescence spectrometers can be divided into either lifetime or steady-state instruments, depending on whether they resolve the temporal behaviour of the emission (or more correctly the excited state), or not, respectively. In both cases there are strong similarities with single beam absorption instruments. Thus, much of the preceding sections is equally relevant to them. However, the levels of photons detected in fluorescence (or equally phosphorescence) are typically much lower than those in absorbance: in the former one is detecting the few photons that are emitted by the sample, in the latter one is detecting those of the light source attenuated by the number absorbed by the sample. As a consequence certain features are optimized differently for fluorescence.

Firstly, fluorescence is detected orthogonal to the direction of the excitation beam incident on the sample (**Figure 2**), so as to delineate the emission photons from those of the excitation beam and minimize those from Rayleigh and Raman scattering, although these always provide a residual level. Hence, for solution studies, special fluorescence cells are required that have orthogonal faces optically transparent and flat. Due to the low levels of photons to be detected it is extremely important to exclude all sources of ambient light from the instrument.

To distinguish the wavelength dependencies of a sample's excitation and emission spectra, monochromators are placed in both the excitation and the emission optical paths. Again, the emission side monochromator and detector may be replaced with a fixed dispersive element (e.g. a diffraction grating) and a diode array. Likewise, in very basic instruments, filters may be substituted for monochromators. For instruments operating at a single excitation wavelength laser sources can be used to good effect.

The selection of excitation wavelength and detected emission wavelength may be independently controlled. Thus the excitation wavelength may be fixed and the emission wavelength scanned to give the emission spectrum, or vice versa to give the excitation spectrum. On many of the higher specification instruments it is possible to automatically scan both emission and excitation wavelengths to give an excitation–emission 2D map.

As the fluorescence is directly proportional to the number of photons absorbed by the sample (in the absence of inner-filter/self-shadowing effects of excessive absorption), it is advantageous to employ very high intensity light sources; xenon arcs are highly suitable. Additional excitation intensity may be achieved by greater spectral bandwidths employed on the excitation side, although this may

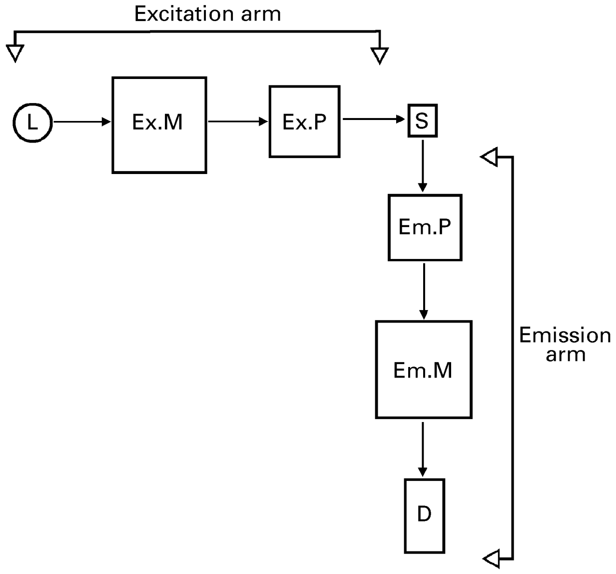



**Figure 2** Conceptual diagram of a fluorescence spectrometer. L = light source, Ex.M = excitation monochromator(s), Ex.P = excitation polarizer, S = sample, Em.P = emission polarizer, Em.M = emission monochromator(s) and D = detector.

compromise the spectral resolution of the results. With the higher levels of light impinging on a sample, unwanted photoreactions can be a problem, as can heat generation in the sample, which accelerates all reactions. Alternatives are pulsed sources, which provide broad band radiation over short periods of time and thus may minimize some of the problems of steady-state arcs. Furthermore, in lifetime instruments they provide a means of determining the time lag between absorption and subsequent emission of photons by the sample – the emission lifetime. In this respect pulsed lasers and flash lamps such as hydrogen arcs are popular.

Similarly, the number of photons detected can be increased by modestly increasing the spectral bandwidth on the emission side. However, again this will have a corresponding effect on spectral resolution. In the most sensitive of instruments the detector, invariably a photomultiplier tube, is cooled to reduce noise and thus improve the signal-to-noise levels.

### Calibration

Many fluorescence studies are carried out without recourse to correction for instrumental response variations with wavelength, or even with time. For some investigations this is adequate, but the reasons for not calibrating stem primarily from its difficulty rather than its irrelevance; for absolute measurements it is essential.

To correct excitation spectra it is necessary to determine the wavelength dependence of the excitation side of the instrument, with the emission side fixed. A common method is to employ a 'quantum counter'; essentially a compound whose absorption spectrum (at an appropriate concentration) is such that more than 99% of all the exciting photons are absorbed over a sufficiently wide wavelength range and whose emission spectrum and quantum yield are independent of the excitation wavelength over this range. The most frequently used quantum counter is rhodamine B in glycerol or ethylene glycol (at 3–8 g $L^{-1}$). This solution exhibits constant (to within 2%) fluorescence efficiency at 610–620 nm when excited in the range 350–600 nm, and only ±5% variation for excitation between 250–350 nm. Measurement of the apparent excitation spectrum of such a sample, monitored at an emission wavelength between 610 and 620 nm, allows direct determination of the wavelength dependence of the excitation side of the instrument.

The effectiveness of this method has led to instruments in which a quantum counter is incorporated into the design by diverting a portion of the excitation light to a separate quantum counter and detector. Whilst such a system importantly allows for correction of any temporal variations whilst measurements on samples is ongoing, it does introduce a difference in the optical path from that of the true exitation beam, with a potential inaccuracy.

Determination of the wavelength dependence of the emission side of an instrument is more problematical. Ideally, light from a 'standard lamp', i.e. one whose calibrated spectral distribution is known, is directly introduced into the emission optical path from the sample compartment. Measurement of the apparent emission spectrum and comparison with the known true distribution of the lamp gives the wavelength characteristics of the emission side. One practical variation of this method is to employ the light from the excitation side of the instrument, for which the wavelength characteristics have already been determined via, say, a quantum counter. In order to direct this light into the emission optical path it is necessary to place a reference scatterer into the sample compartment. Such scatterers must have no appreciable wavelength dependence over the wavelength range of interest (mirrors, whilst achieving the redirection, have wavelength dependencies that make them poor choices). Common choices are flat cakes of magnesium oxide (MgO) or barium sulfate with potassium sulfate binder ($BaSO_4$ in $K_2SO_4$), which can be mounted in the sample position at 45°C to both the excitation and the emission optical paths.

As an alternative to standard lamps and their derivatives, there are numerous compounds whose absolute fluorescence spectra have been documented and may be employed to deduce the emission wavelength characteristics of the instrument. Nonetheless, the calibration of the excitation arm is still an essential procedure.

### Fluorescence anisotropy and polarization

As in absorption spectroscopy, instrumental polarization effects can yield unwanted artefacts and therefore it is appropriate to introduce depolarizers into the optical path before and after the sample if the aim is to monitor the true unpolarized fluorescence spectrum. Unfortunately, this will reduce the light levels commensurately and is therefore frequently not pursued. Other methods, involving the use of polarizers set at 'magic angles' to minimize some unwanted polarisation effects have been devised, but are even less frequently employed.

However, deliberate polarization can be used to great advantage in probing the environment and motion of fluorescent molecules and groups in larger macromolecules. In this case, rotatable plane polarizers are inserted in the optical path just before and after the sample; spectra are acquired in the four possible

combinations of the polarizers, each in the horizontal or vertical orientation relative to the 'horizontal' plane defined by the excitation and emission optical paths. By comparison of the four spectra the mobility of the fluorescent group during the lifetime of the exicted state can be deduced.

### Fluorescence lifetime instruments

Lifetime instruments share most of the optical arrangement of steady-state instruments. Indeed there are commercial instruments that combine both into one versatile spectrometer. The essential optical difference is in the use of intense pulsed light sources, with an emission pulse width typically of the order of 1 ns or less. By coupling the detector and light source trigger with sophisticated electronics and post acquisition processing it is possible to correlate the time between the absorption and subsequent emission of photons by the sample. Essentially, the excitation pulse corresponds to the absorption profile in time. In time-correlated single photon counting methods the delay for the first photon to be subsequently detected is then recorded. This is repeated many thousands of times to give a statistical distribution from which the absorption time profile can be deconvoluted. Alternatively, if the lifetime is sufficiently long, as in phosphorescence, then the complete decay curve of emitted photons from a single exitation pulse can be directly monitored – the pulse excitation method. Finally, rather than employing a flash lamp to provide a pulse of excitation, the intensity of continuous excitation can be modulated and the phase lag of the resulting oscillations on emission intensity observed – the phase resolved method. Whichever method is adopted, in turn the data can be analysed in terms of the fluorescence or phosphorescence lifetimes of the molecular species involved.

## Imaging instruments

The advent of imaging detectors, such as CCD cameras and more recently photomultiplier arrays, has prompted the development of monochromators that are able to spectrally disperse the individual pixels of an image, whilst preserving the spatial integrity of the image. Consequently, absorption and fluorescence instruments are beginning to be developed that are able to produce a spectroscopic image of a sample, each pixel of the image being a complete spectrum. It is apparent that such instruments will find growing use in the investigation of inhomogeneous material for which traditional methods are only able to give spatially averaged results.

*See also:* **Biochemical Applications of Fluorescence Spectroscopy; Biomacromolecular Applications of UV-Visible Absorption Spectroscopy; Dyes and Indicators, Use of UV-Visible Absorption Spectroscopy; Inorganic Condensed Matter, Applications of Luminescence Spectroscopy; Light Sources and Optics; Organic Chemistry Applications of Fluorescence Spectroscopy; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Burgess C and Knowles A (eds) (1981) *Techniques in Visible and Ultraviolet Spectrometry*, Vol 1, Standards in absorption spectrometry. London: Chapman & Hall.

Miller JN (ed) (1981) *Techniques in Visible and Ultraviolet Spectrometry*, Vol 2, Standards in fluorescence spectrometry. London: Chapman & Hall.

Knowles A and Burgress C (eds) (1984) *Techniques in Visible and Ultraviolet Spectrometry*, Vol 3, Practical absorption spectrometry. London: Chapman & Hall.

Clark BJ, Frost T and Russell MA (eds) (1993) *Techniques in Visible and Ultraviolet Spectrometry*, Vol 4, UV spectroscopy. London: Chapman & Hall.

Mattis DA and Bashford CL (eds) (1987) *Spectrophotometry and Spectrofluorimetry, a Practical Approach*. Oxford: IRL.

Ingle JD Jr and Crouch SR (1988) *Spectrochemical Analysis*. Englewood Cliffs, NJ: Prentice Hall.

Silverstein RM, Bassler CG and Morrill TC (1974) Spectrometric Identification of Organic Compounds, 3rd edn. New York: Wiley.

# V

## Vanadium NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

## Vibrational CD Spectrometers

**Laurence A Nafie**, Syracuse University, NY, USA

Copyright © 1999 Academic Press

**VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES**
**Methods & Instrumentation**

### Introduction

Vibrational circular dichroism (VCD) is defined as circular dichroism (CD) in vibrational transitions in molecules. These transitions typically occur in the infrared (IR) region of the spectrum and hence a VCD spectrometer is an infrared spectrometer that can measure the circular dichroism associated with infrared vibrational absorption bands. CD is defined as the difference in the absorption of a sample for left versus right circularly polarized radiation. This difference is zero unless the sample possesses molecular chirality, either through its constituent chiral molecules or through a chiral spatial arrangement of non-chiral molecules.

A molecule, or an arrangement of molecules, is chiral if it is not superimposable on its mirror image. A chiral molecule possesses a handedness and can exist in either one form, an enantiomer, or its mirror image, the opposite enantiomer. A sample of chiral molecules can have varying chiral purity, referred to as enantiomeric excess. The percent enantiomeric excess (%ee) is defined as the percent excess of one enantiomer relative to the total sample. The %ee of a pure sample of only one enantiomer is 100%. If a sample is composed of an equal mixture of both enantiomers, the %ee is 0% and the sample is called racemic. Racemic samples of chiral molecules exhibit no CD spectra, or any other form of natural optical activity.

The first measurements of VCD were achieved nearly 25 years ago. The early instruments used for these measurements were relatively crude by today's standards, but they demonstrated that VCD was a natural phenomenon that could be used to study in more detail the structure and dynamics of chiral molecules. Subsequent improvements in VCD spectrometers included extending the wavelength of coverage from the region of hydrogen-stretching modes into the mid-infrared region where a greater variety of vibrational transitions could be studied. It also included the implementation of Fourier-transform (FT) methods for VCD measurement. This was a particularly important advance, since virtually all modern, commercially available infrared absorption spectrometers are now FT-IR spectrometers. With the advent of FT-VCD, it became possible to construct an efficient VCD spectrometer starting from a commercially available FT-IR spectrometer.

Within the past few years, accessory modules for the measurement of FT-VCD have become available from the manufacturers of several FT-IR

spectrometers. In one case, that of the Bomem Chiral*ir*, the first stand-alone, factory-aligned FT-VCD spectrometer has become commercially available, opening the way for widespread applications of VCD spectroscopy.

The principal applications of VCD spectroscopy include measurements of the conformation, absolute configuration and enantiomeric excess of chiral molecules. Most of the molecules of interest for study with VCD are biological in origin. Many are molecules of pharmaceutical interest. The unique application of VCD is its ability to determine absolute configuration in conjunction with *ab initio* calculations. Remarkably close matches have been achieved between experimental VCD spectra and the corresponding spectra calculated from first principles using quantum-mechanical calculations.

## General measurement principles

The measurement of VCD is quite simple in concept. A sample is placed in the VCD spectrometer and the polarization of the IR radiation passing through the sample is switched between left and right circularly polarized states. If the sample is chiral, a small difference in the intensity of the IR beam for left and right circularly polarized IR radiation occurs and is measured at the detector. **Figure 1** illustrates the definition of VCD with an energy-level diagram for molecular transitions between the zeroth and first vibrational sublevels of the ground electronic state, *g*0 and *g*1). The decadic absorbance, or IR intensity, of the sample for wavenumber frequency $\bar{\nu}$ (equal to the frequency of the radiation divided by the speed of light) is defined as

$$A(\bar{\nu}) = -\log_{10}[I(\bar{\nu})/I_0(\bar{\nu})] \qquad [1]$$

where $I(\bar{\nu})$ and $I_0(\bar{\nu})$ are the single-beam intensities at the detector with and without the sample present, respectively. The circular-polarization differential absorbance, or VCD intensity, is defined as

$$\Delta A(\bar{\nu}) = A_L(\bar{\nu}) - A_R(\bar{\nu}) \qquad [2]$$

The basic measurement layout is illustrated in **Figure 2**. Here radiation from an IR source is dispersed either by a diffraction grating or a Fourier-transform interferometer so that different wavelengths of the radiation can be distinguished. An infrared optical filter is placed in the beam to restrict

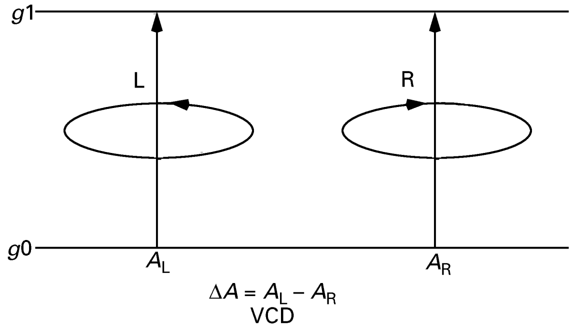



**Figure 1** Energy-level diagram illustrating the definition of VCD as the difference in the absorbance of a molecule for left *versus* right circularly polarized IR radiation in a vibrational transition between states *g*0 and *g*1 in the ground electronic state.

the measurement to the spectral region of interest. This is followed by a linear polarizer to define a single state of polarization of the infrared beam. A photoelastic modulator (PEM) then modulates the polarization state of the beam between left (L) and right (R) circularly polarized states at a frequency in the tens of kilohertz range. Immediately afterwards, the sample is placed in the beam. The VCD of the sample creates an intensity modulation of the IR beam at the polarization-modulation frequency. The radiation is then focused on a detector after which further manipulations are carried out electronically to produce the final VCD spectrum. The detector signal is first amplified using a preamp and then divided into two pathways. One leads directly to the IR spectrum as in an ordinary infrared spectrometer. The other pathway is to a lock-in amplifier, referenced to the PEM, that demodulates the high-frequency polarization-modulation component of the signal and leads to the VCD spectrum as described in more detail below.

The precise way in which the detector signal is processed electronically depends on the kind of VCD spectrometer used. In the following sections, three different VCD spectrometer designs are discussed. The simplest of these is the dispersive VCD spectrometer, and we will use this design to illustrate the basic concepts associated with the electronic processing of VCD spectra. The two subsequent cases involving Fourier-transform VCD spectrometers are more complex, but share the same underlying conceptual basis as the dispersive VCD spectrometer.

## Dispersive VCD spectrometers

In a dispersive VCD spectrometer, the IR source in **Figure 2** consists of a thermal or arc source of infrared radiation, a light chopper, and a grating monochromator. The infrared source of radiation is first

focused on the entrance slit of the monochromator where it is spatially dispersed by the grating. A narrow band of wavelengths (wavenumber frequencies) emerges from the exit slit of the monochromator, and the spectrum is collected by turning the grating and scanning each point in the spectrum sequentially. Successive scans of the monochromator can be averaged to improve the signal-to-noise ratio.

The signal from the detector consists of two components. One is referred to as $I_{DC}$, which represents the IR single-beam transmission of the sample. In a dispersive VCD spectrometer, $I_{DC}$ is modulated at the frequency of the light chopper and carries the information needed for the ordinary IR spectrum as indicated in **Figure 2**. The other component is $I_{AC}$, and it is modulated at the polarization-modulation frequency of the PEM, as well as the frequency of the light chopper. In terms of the transmission intensities at the detector for left and right circularly polarized radiation, $I_L$ and $I_R$, these two components of the detector signal are given by

$$I_{DC}(\bar{\nu}) = (1/2)[I_R(\bar{\nu}) + I_L(\bar{\nu})] \qquad [3]$$

$$I_{AC}(\bar{\nu}) = (1/2)[I_R(\bar{\nu}) - I_L(\bar{\nu})] \sin \alpha_M(\bar{\nu}) \qquad [4]$$

where $I_{AC}$ depends on the sine of the retardation angle of the PEM, which in turn varies sinusoidally at the PEM frequency $\bar{\nu}_M$:

$$\alpha_M(\bar{\nu}) = \alpha_M^0(\bar{\nu}) \sin 2\pi\nu_M t \qquad [5]$$

After some algebra, it can be shown that the ratio of the two intensity components in Equations [3] and [4] is proportional to the VCD intensity as

$$I_{AC}(\bar{\nu})/I_{DC}(\bar{\nu}) = 2J_1[\alpha_M^0(\bar{\nu})]1.15\Delta A(\bar{\nu}) \qquad [6]$$

where $J_1[\alpha_M^0(\bar{\nu})]$ is the first-order Bessel function and is a measure of the efficiency of the PEM setting for the wavenumber frequency specified.

In order to calibrate VCD measurements, one substitutes the sample with a multiple waveplate followed by a linear polarizer. The fast and slow axes of the multiple waveplate are aligned with the axes of the PEM, and the polarizer is set at 45 degrees from these axes. There are four positions of the multiple waveplate and the polarizer relative to the PEM, and these generate a family of four calibration curves. It can be shown that the intersections of these pseudo-VCD curves have the values

$$[I_{AC}(\bar{\nu})/I_{DC}(\bar{\nu})]_{CAL} = \pm 2J_1[\alpha_M^0(\bar{\nu})] \qquad [7]$$

Connecting all the positive intersection points, one obtains a spectral curve, which when divided into Equation [6] allows the isolation of the calibrated VCD intensity spectrum $\Delta A(\bar{\nu})$.

Examples of dispersive VCD and IR spectra for three closely related chiral molecules are presented in **Figure 3**. These spectra are illustrative of a number of basic concepts. All three sets of spectra are recorded in the region of carbon–hydrogen

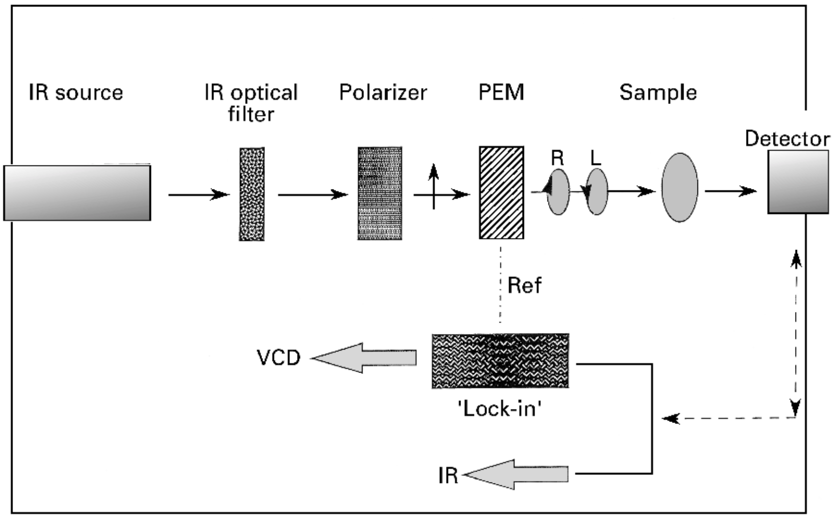



**Figure 2** Diagram illustrating the basic optical layout and electronic pathways for the measurement of VCD. The diagram is applicable to both dispersive VCD spectrometers and FT-IR spectrometers that use a photoelastic modulator (PEM) as the source of the polarization modulation of the light beam between left (L) and right (R) circular states.

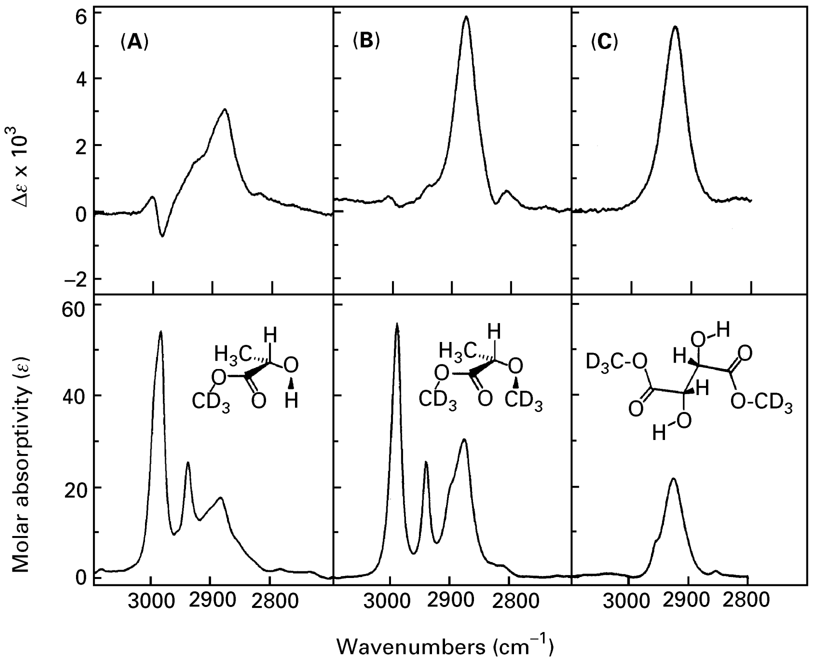


Figure 3 Dispersive IR and VCD spectra in the region of CH stretching vibrations for the molecules (A) (*S*)-methyl-$d_3$ lactate, (B) (*S*)-methyl-$d_3$ 2-(methoxy-$d_3$)-propionate, and (C) di(methyl-$d_3$) D-tartrate illustrating the large positive VCD associated with the methine CH stretching mode in these molecules. The experimental conditions were 0.005 M or 0.01 M solutions in $CCl_4$ in a 1.00 cm fixed-pathlength cell. The resolution of the VCD spectra is 16 $cm^{-1}$ and of the IR spectra is 4 $cm^{-1}$.

stretching vibrations. The VCD instrument was constructed in our laboratory at Syracuse University starting in 1975 and optimized over the years to include various kinds of improvements including computer control for automatic control and signal averaging. The source used was a xenon arc lamp and the detector was a liquid-nitrogen-cooled InSb detector, used for higher-frequency vibrations above 2000 $cm^{-1}$. The intensity scale is in molar absorptivity in units of $M^{-1}$ $cm^{-1}$ where the absorbance has been divided by the concentration in moles per litre and the pathlength in cm. The magnitudes of the VCD spectra are approximately four orders of magnitude smaller than the IR absorbance spectra. The VCD spectra all have a bias toward positive VCD intensity. The source of this bias is demonstrated by the series of three spectra, and it is shown to be the lone methine CH stretching mode. In **Figure** 3A, for (*S*)-methyl-$d_3$ lactate, four CH fundamental modes are present, two for the antisymmetric methyl stretching modes near 3000 $cm^{-1}$, one for the symmetric methyl stretching mode, with an additional Fermi component at lower frequency, near 2940 $cm^{-1}$, and the lone methine stretch near 2880 $cm^{-1}$. Converting this molecule to the deuteriomethoxy analogue, (*S*)-methyl-$d_3$ 2-(methoxy-$d_3$)-propionate, further enhances the lone methine stretching mode relative to the methyl modes as shown in **Figure** 3B. The interfering methyl group is eliminated in the case of **Figure** 3C for di(methyl-$d_3$) D-tartrate where the VCD of the methine can be observed free of other fundamental vibrational modes. The sign of the methine VCD is a marker for the absolute configuration of these and related molecules. The magnitude of the methine VCD is sensitive to the conformation of the molecule in the vicinity of the chiral centre, which for these relatively small molecules is essentially the whole molecule.

Although dispersive VCD spectrometers were the original kind of VCD instrument, they still retain some advantages over the newer Fourier-transform instruments. A relative advantage is present if only a limited spectral range is of interest. In that case a strong source and a narrow filter permit transmission intensities that are higher than could be maintained over a broader spectrum without saturating the detector.

## Fourier transform VCD spectrometers

The first measurements of VCD using a Fourier-transform (FT-IR) spectrometer were published in 1979. The basic idea is to substitute the combination of light chopper and monochromator with an FT-IR spectrometer. In an FT-IR spectrometer, all wave-

lengths of the spectrum of interest are measured at once. The frequencies are distinguished from one another by the interferometer at the heart of the instrument. The infrared light from the source is divided by amplitude at a beamsplitter where one beam is sent to a fixed mirror and the other to a mirror that can change position. The two beams recombine at the beamsplitter and interfere with one another depending on the phase difference of the two light paths. Shorter wavelengths go in and out of phase more rapidly than longer wavelengths and, hence, the different wavelengths can be distinguished from one another by their interference rate or Fourier frequency. The Fourier interference frequency is analogous to the light chopper in the dispersive VCD instrument, but in the case of an FT-VCD instrument, each wavelength has its own 'chopper' frequency. No other changes are needed in the optical setup of the FT-VCD instrument, and hence **Figure 2** is applicable to this instrumental layout as well as that of the dispersive VCD instrument.

The intensity measured by the detector as a function of the moving mirror position, $\delta$, is called an interferogram. The interferogram is a sum of all the intensities of the spectrum at each wavenumber frequency times their Fourier amplitude. Again, there are two intensity components at the detector. One is the ordinary interferogram associated with the single-beam transmission spectrum, and the other is the VCD interferogram that is modulated at the PEM frequency. These component interferograms are given by

$$I_{\mathrm{DC}}(\delta) = \int_0^\infty I_{\mathrm{DC}}(\bar{\nu}) \cos[2\pi\bar{\nu}\delta - \theta_{\mathrm{DC}}(\bar{\nu})] \mathrm{d}\bar{\nu} \qquad [8]$$

$$I_{\mathrm{AC}}(\delta) = \int_0^\infty I_{\mathrm{AC}}(\bar{\nu}) \exp(-2V\bar{\nu}\tau) \cos[2\pi\bar{\nu}\delta - \theta_{\mathrm{AC}}(\bar{\nu})] \mathrm{d}\bar{\nu} \qquad [9]$$

where $V$ is the Fourier frequency and $\tau$ is the time constant of the PEM lock-in amplifier. The ordinary IR interferogram in Equation [8] contains a phase function, $\theta_{\mathrm{DC}}(\bar{\nu})$, that must be determined before the interferogram can be Fourier transformed to yield $I_{\mathrm{DC}}(\bar{\nu})$. This is evaluated by standard techniques. The VCD interferogram contains its own phase function and this phase is transferred from another VCD interferogram associated with a spectrum of only positive VCD intensities so that standard phase-correction algorithms can be used. Equation [9] also contains an exponential-decay function that decreases with higher wavenumber frequency. This function

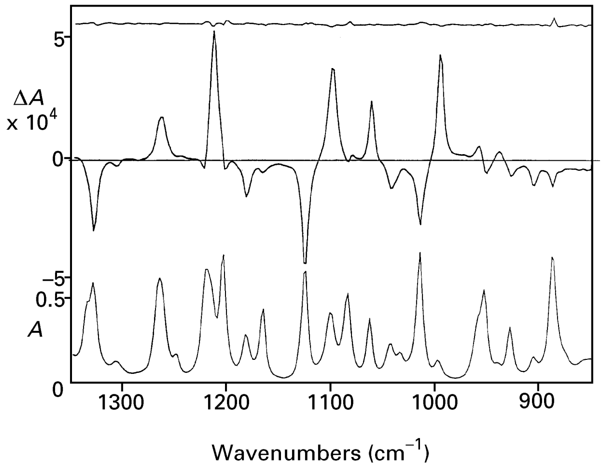



**Figure 4** FT-IR and FT-VCD spectra of (–)-$\alpha$-pinene in the mid-IR region. The experimental conditions were neat liquid with a pathlength of 75 μm a resolution of 4 $cm^{-1}$ and a collection time of 20 min per enantiomer. The final VCD spectrum was obtained from the subtraction of the VCD of the (+)-enantiomer from that of the (–)-enantiomer.

represents the effect of the time constant of the lock-in amplifier used to demodulate the VCD interferogram from the PEM modulation frequency.

Once Equations [8] and [9] have been Fourier transformed, Equations [6] and [7] can be used to isolate the VCD spectrum although both ratios now also include the exponential function of the lock-in time constant. However, this function vanishes when the calibration curve is divided into the ratio of the AC and DC intensities and does not enter the final VCD spectrum.

An example of an FT-VCD spectrum is presented in **Figure 4**. The IR and VCD spectra of (–)-$\alpha$-pinene are between 1350 and 850 $cm^{-1}$. These spectra were measured on the Chiral*ir* VCD spectrometer from Bomem/BioTools. It employs a SiC glower source and a liquid-nitrogen-cooled HgCdTe (MCT) detector. Again we see the VCD spectrum is displayed on an intensity scale that is approximately four orders of magnitude smaller than the corresponding IR spectrum. Good correspondence is present between the peaks in the IR and VCD spectra, although some overlapping of bands is present. It is easy to see that some IR bands are positive and some are negative. According to the definition of VCD, the positive bands absorb left circularly polarized light more strongly than right circularly polarized light. There is no particular correlation between strong IR bands and strong VCD bands. The spectrum illustrates the relative strength of FT-VCD to measure spectra over a wide spectral range at high resolution in a relatively short period of time.

A second example of an FT-VCD spectrum is provided in **Figure 5**. The sample in this case is (+)-camphor and the spectral region is the higher-frequency range of CH-stretching vibrations near 3000 cm$^{-1}$. This spectrum was obtained using a step-scan VCD spectrometer based on an IFS 55 of Bruker Instruments and a VCD accessory bench aligned and optimized in our laboratory at Syracuse University. Step-scan operation offers the advantage of eliminating the decreasing exponential time-constant function associated with the VCD interferogram that disadvantages the higher-frequency region of vibrational transitions. Here a tungsten light source was used in conjunction with an InSb detector. This VCD spectrum is of much higher quality than the corresponding spectrum obtained with a dispersive VCD spectrometer. Step-scan FT-VCD measurement have been carried out as well in the OH and NH stretching regions between 3000 and 3700 cm$^{-1}$. In some respects collecting a step-scan VCD spectrum is similar to collecting a dispersive VCD spectrum. In each case the spectrum is scanned and averaged a limited number of times, typically two to four times, and a relatively long time constant can be employed with the PEM lock – since one is not trying to protect a band of Fourier frequencies. The principal difference between the two kinds of measurements is that the light level is not diminished by the reduction of slitwidth if higher resolution is desired, and all the light is used in the FT measurement rather than leaving most of it on the inside of the monochromator as in the case of the dispersive VCD measurement.

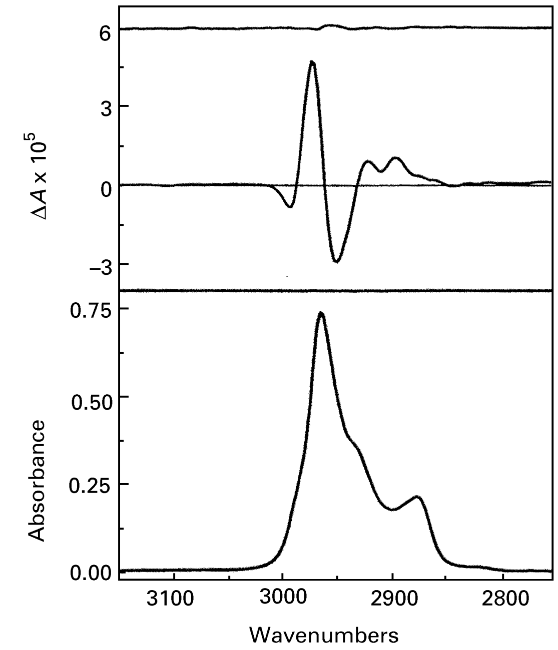



**Figure 5** FT-IR and step-scan FT-VCD spectra of R-(+)-camphor in the CH-stretching region. The experimental conditions were a 0.6 M solution in $CCl_4$, a pathlength of 43 µm, a resolution of 16 cm$^{-1}$ and a collection time of 2 h per enantiomer. The final VCD spectrum was obtained from the subtraction of the VCD of the (–)-enantiomer from that of the (+)-enantiomer.

## Polarization-division FT-VCD spectrometers

In 1989, a new kind of FT-VCD measurement was demonstrated, originally called polarization-modulation interferometry (PMI) and more recently called polarization-division interferometry (PDI). In this approach a polarizing beamsplitter is substituted for the normal amplitude-division beamsplittter. If a linearly polarized infrared beam, with a direction of polarization at 45 degrees relative to the polarization direction of the beamsplitter, is directed to this beamsplitter, then the beam is split into two orthogonally polarized beams. Upon recombination at the beamsplitter, the two beams combine but they do not interfere. The result of the movement of the moving mirror associated with one of the beams is that the polarization state of each wavelength of light cycles continuously through 360 degrees of relative phase retardation at its own Fourier frequency. The cycle starting with vertically polarized radiation is vertical linear, right circular, horizontal linear, left circular and back to vertical linear. In the ideal case, there is no intensity modulation, only polarization modulation. In order to measure a conventional FT-IR absorption spectrum, one can insert a polarizer, say in the vertical position, and the beam is converted from polarization modulation to intensity modulation. Maximum intensity occurs when the beam is vertical linear and minimum when it is horizontal linear. Without the polarizer present, the interferometer is sensitive to linear dichroism in the sample oriented vertically or horizontally at the same Fourier phase as the absorption spectrum (cosine transform) and is sensitive to circular dichroism (VCD) out of phase relative to the absorption spectrum (sine transform). **Figure 6** illustrates the polarization cycles and the mode of operation of this instrument for absorption, linear dichroism and circular dichroism measurements.

The advantage of PDI-FT spectrometers is their independence of PEMs. A PEM has a limited range of wavelength coverage, and there are no PEMs commercially available that operate into the far IR. Yet, there are polarizing beamsplitters that have good efficiency in the far IR, and hence PDI-FT-VCD is the likely approach to extend VCD into this region of the spectrum. To date, the performance of

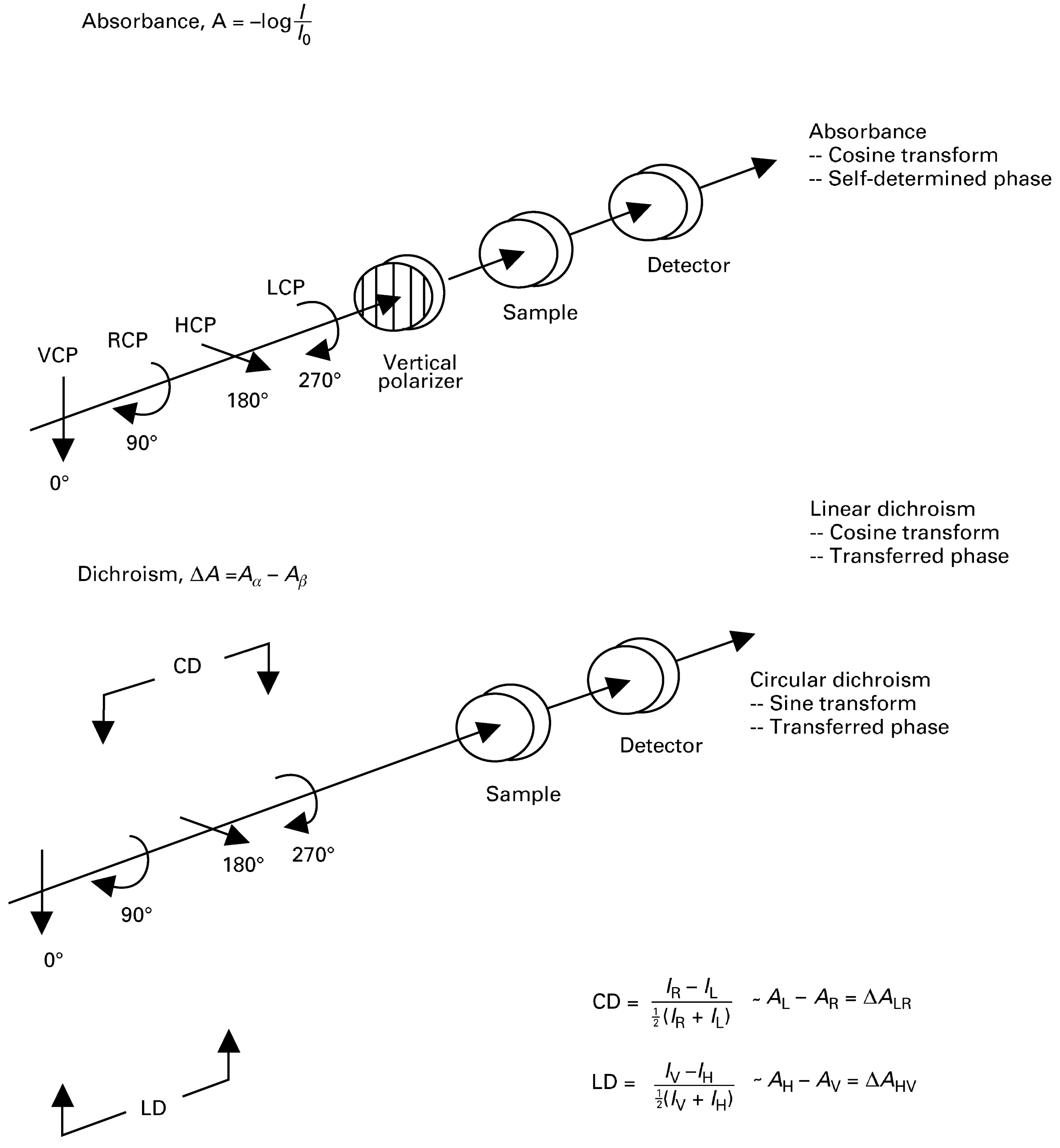



**Figure 6** Diagram illustrating the polarization sequence and measurement setups for FT-IR, FT-VCD and FT-VLD with a PDI-FT spectrometer. For absorption measurements, the placement of a vertical polarizer converts the train of polarization modulation to intensity modulation. The polarizer is removed for VCD and VLD measurements as illustrated.

PDI-FT-VCD spectrometers is somewhat below that of PEM-FT-VCD spectrometers in the region where they have been directly compared, the mid-IR region. Recently, an FT-VCD instrument was described that possessed both PDI capability and conventional polarization-modulation capability using a PEM. Referred to as double polarization modulation interferometry, this technique offers advantages of signal intensity compared to other single polarization modulation FT-VCD spectrometers.

## Artifact suppression

VCD intensities are smaller than IR intensities by approximately four orders of magnitude. As a result they are subject to interference from optical

imperfections in the instrument itself. The manifestations of these imperfections, which differ from instrument to instrument, are called artifacts. Artifacts arise from the combination of birefringence in the optics and a sensitivity to different states of linear polarization by the detector or by optical reflection surfaces. The birefringence arises from strain in windows and lenses. Birefringence alters the polarization state of the light and disturbs the balance between left and right circularly polarized light in the spectrometer. A PEM is an oscillating birefringent plate and its action creates the oscillating left and right circular polarization states in the first place. Stray birefringence in the optics, including within the PEMs, further alters the polarization states in undesirable, unknown ways. Once the symmetry between the left and right circular polarization states is broken, the instrument possesses some, perhaps small, degree of linear polarization modulation at the PEM frequency. If the detector or surface of some other optical element responds differentially to the linear polarization modulation, an artifact intensity is created that coexists with the VCD intensity.

There are two kinds of artifacts. One is independent of the sample and exists as a common background spectrum. It can be recorded in the absence of a sample or with any racemic or non-chiral sample, such as a solvent. Once recorded, it may be subtracted automatically from all future VCD spectra to remove this background signal from the measurement. The second kind of artifact is one that varies with the absorption spectrum of the sample or solvent. This is more difficult to remove. In fact, the only way currently known that it can be removed completely from a measurement is by subtraction of the VCD spectrum of the racemic mixture or the opposite enantiomer of the chiral sample. It is important that the racemic or enantiomer VCD measurement be made under the same conditions as the desired chiral measurement, namely, the same pathlength, concentration and cell position. From the standpoint of signal quality, it is more effective to record the VCD spectrum of the opposite enantiomer rather than the racemic mixture since in the former case, the subtraction adds additional VCD information while at the same time cancelling the common artifact spectrum. Unfortunately, a sample of the opposite enantiomer or the racemic mixture is not always available. For this reason, great care needs to be exercised to reduce the occurrence of both kinds of artifacts in the initial optical alignment of the VCD spectrometer.

In practice, it is found that reducing the constant background artifact also reduces the severity of the absorption-dependent artifact. It has also been found that using lenses instead of mirrors after the first polarizer in the optical train can reduce the background artifact. Mirrors possess both birefringence and sensitivity to different states of linear polarization. When used off-axis, as is usually the case, these effects are enhanced on the IR beam. Lenses, on the other hand, can be used on-axis and exhibit lower artifact-inducing effects. In addition to using lenses, the optical alignment should be purely axial and cylindrically symmetric so that a particular direction in space, beyond the direction of beam propagation, is not favoured in the alignment. The final alignment can be achieved by minor adjustment of the optics on a trial and error basis until a good instrument baseline is reached. If the baseline is relatively flat and close to zero across the spectrum, it is usually found that absorption-dependent artifacts are not a serious problem.

## Absolute VCD intensity

An important aspect of instrumentation performance is absolute intensity calibration. The intensity-calibration procedure described above using the multiple waveplate and second polarizer has been the method of choice for the calibration of VCD spectra for many years. Nevertheless, the technique is prone to variation depending on the accuracy and care taken in the calibration measurement. If the multiple waveplate or the second polarizer is not positioned at the optimum angular orientation, a calibration spectrum is obtained that is not correct. Usually, the calibration spectrum is too small and the resulting calibrated VCD spectrum has intensities that are too large. Another common source of error is the aperture of the beam used in the calibration measurement relative to that used for the VCD measurement. The degree of polarization modulation in a PEM varies with aperture, decreasing from its centre. The calibration measurement determines the $J_1$ function of the PEM averaged over the beam profile, and the correct calibration is obtained only if the aperture of the VCD measurement matches the aperture of the calibration measurement.

In an effort to establish an intensity standard for VCD measurements, a number of laboratories have undertaken the measurement of the mid-infrared VCD spectrum of (−)-$\alpha$-pinene. The results from several laboratories have been obtained to date and the results have been plotted in molar absorptivity units. In **Figure 7**, we present the absolute VCD measurements from three locations. The results are still preliminary, and though they show a variation of the order of 10%, these intensities and others appear to

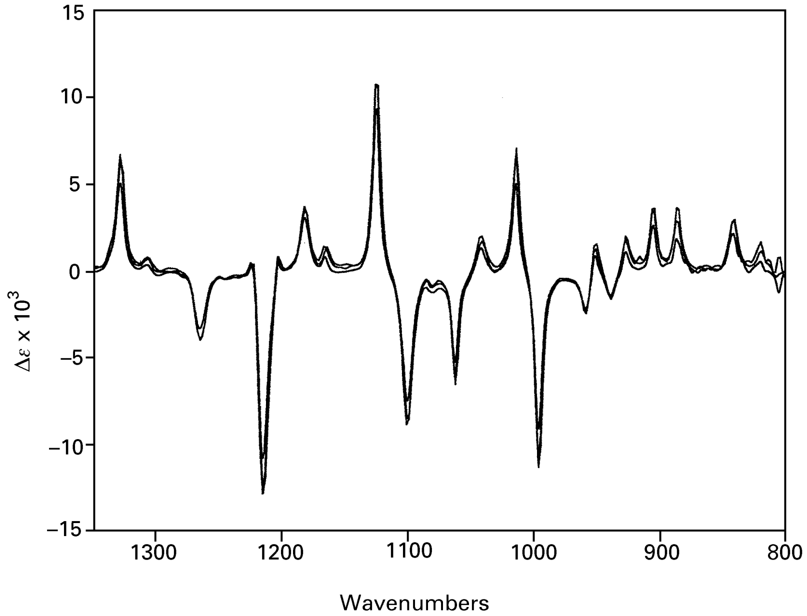



**Figure 7** Measurements of the VCD spectrum of neat (+)-$\alpha$-pinene from three different laboratories illustrating both the variation that can arise in the measurement of absolute intensities and the convergence of the measurements to a relatively narrow range of values.

be converging on a particular set of values. It is hoped that a set of accepted values with a small range of uncertainty will be available in the near future. With a set of absolute intensities in hand, the calibration of a VCD spectrometer could then be carried out by comparison with a standard set of spectra rather than by a calibration measurement subject to the operational errors discussed above.

## Areas of application

There are three principal areas of application of VCD spectroscopy. The first, and simplest, is to use VCD spectra to measure the optical purity in terms of %ee of a sample or series of samples. The second application is to determine the absolute configuration of the molecules in the sample, and the third is to determine the solution-state conformations of molecules present in the sample.

## Determination of enantiomeric excess

In the case of optical-purity measurements, the determination of %ee is based on the fact that the VCD intensity varies linearly from 100% to 0% with the %ee. The VCD spectrum obtains its maximum value for a chirally pure sample of a single enantiomer; it falls to half its value for a %ee value of 50% and vanishes for the racemic solution where the %ee is 0%. This linear relationship is demonstrated in **Figure 8** where the VCD spectra of (−)-$\alpha$-pinene for three different values of %ee are plotted for the region from 1150 to 1075 $cm^{-1}$. Here it is clear that the VCD in both bands decreases in value as the %ee is lowered from 100% to 95% to 90%. A partial least-squares analysis of the entire VCD spectrum of (−)-$\alpha$-pinene from 1350 to 900 $cm^{-1}$ for a wide range of %ee values leads to a degree of precision in predicting the %ee from the VCD spectrum of less than 1%. Similar accuracies have been achieved for other molecules. The only prior requirement for the determination of %ee is a high-quality VCD reference spectrum of a sample with known optical purity. From such a VCD spectrum, the VCD intensities for a pure sample at 100 %ee can be determined and all subsequent unknowns can be referenced to this measurement.

There are several advantages of VCD for the determination of optical purity that are not available in the more traditional methods. First, compared to the measurement of optical rotation, VCD intensity can usually be observed in most molecules at approximately the same level of intensity for the strongest bands in the spectrum. This intensity is approximately four orders of magnitude smaller than the IR absorbance spectrum. On the other hand, values of optical rotation can vary widely.

VCD intensities are not temperature sensitive. VCD spectra are composed of many bands and

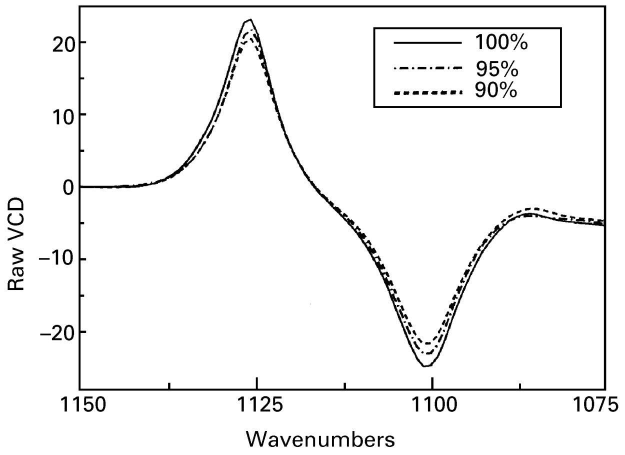



**Figure 8** Three VCD spectra of (−)-$\alpha$-pinene for decreasing values of optical purity, 100%, 95% and 90% enantiomeric excess. The experimental conditions were 70 μm pathlength, 8 $cm^{-1}$ resolution and 2 h of collection time for each sample.

many spectral points, each one of which is a determinant of the %ee relative to the same point in another sample. Averaging over all the points in the spectrum weighted by their importance to the spectrum, leads to an accurate overall determination. By contrast, optical rotation is a single, temperature-sensitive measurement. The multi-spectral aspect of VCD allows accurate results to be obtained even if there is more noise in a VCD spectrum than in a single optical-rotation measurement.

Compared to chiral chromatography, VCD spectra can be obtained without a physical separation of the two enantiomers. In some cases, enantiomeric pairs of molecules cannot be separated sufficiently with a column, and in this case VCD can be useful. Chiral columns are also expensive to develop and operate, and VCD can be used for routine measurement in a time that is less than that usually required for a physical separation and optical-purity measurement.

The prospects are bright for further improvements of VCD to determine the optical purity of samples. The use of VCD for this purpose is still in an early stage of development, and it is likely that improved measurement and analysis techniques will increase the accuracy of VCD %ee determinations to well below 1% for most samples.

## Determination of absolute configuration

A powerful application of VCD is the determination of the absolute configuration of chiral molecules. VCD is more effective than electronic CD (ECD) in this respect for at least two reasons. One is the number of transitions and the richness of a VCD spectrum compared to an ECD spectrum. There are many more possible transitions to consider in looking for ways to connect the CD spectrum to the absolute configuration. The second reason is that it is easier to calculate IR absorption and VCD spectra than it is to calculate UV absorption and ECD spectra. The latter require accurate descriptions of excited electronic state wavefunctions whereas a vibrational spectrum can be calculated only on the basis of the ground electronic state and its response to nuclear motion. This information is readily available when the equilibrium ground-state geometry is optimized and its molecular force field is determined.

There is widespread interest in the capability of VCD to determine absolute configurations of molecules because the method is free of the need to obtain a single crystal for X-ray diffraction measurements, the standard approach to the determination of absolute configuration of a chiral molecule. VCD measurements are typically carried out in solution or with neat liquids. Since many molecules are difficult to crystallize, VCD promises to be an effective way to determine stereo-specific structures in the absence of the availability of crystals.

Over the past several years, several laboratories have demonstrated that the absolute configuration of a molecule can be determined *de novo* without reference to any other measurement or information base. The method involves carrying out an *ab initio* calculation of the VCD of a particular enantiomer of the chiral molecule. This serves as a theoretical prediction of its VCD spectrum. Next, the VCD spectrum of the molecule is measured and compared to the theoretical prediction. If the sign pattern agrees, the absolute configuration is confirmed to be the one calculated. If the signs are opposite, then the absolute configuration of the molecule in the sample is opposite to the configuration used in the calculation.

An example of the determination of absolute configuration is shown in **Figure 9**. Here the FT-VCD spectrum of (*S*)-methyl lactate in the mid-infrared region is compared to the corresponding *ab initio* calculation using density-functional theory (DFT) and magnetic-field perturbation (MFP) at the 6-31G* basis set level. There are no adjustable parameters to the calculation, and the calculated intensities have been plotted using band shapes similar to the experimentally measured IR and VCD spectra. There is excellent agreement in sign and intensity for both the VCD and IR spectra and there is no doubt that the experimental spectrum was the *S*-enantiomer of this molecule and not the *R*-enantiomer. The theoretical programs used to carry out these calculations are

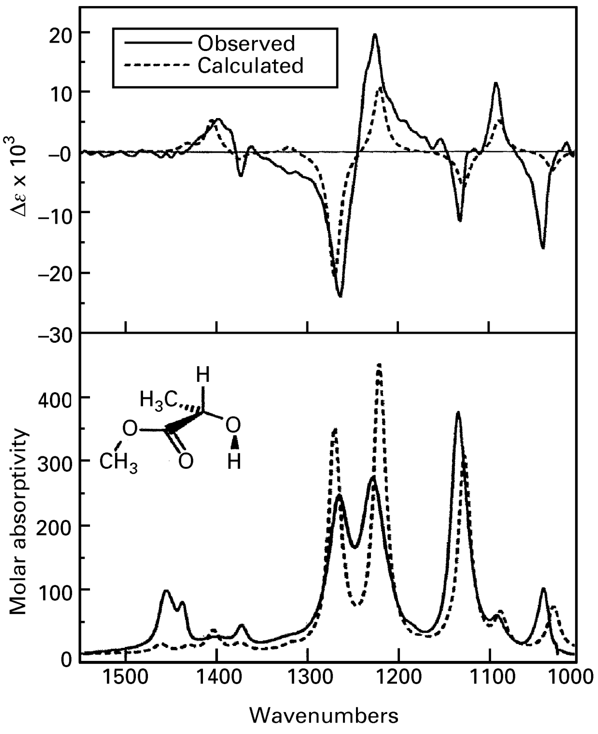



**Figure 9** Comparison of the experimental and theoretical IR and VCD spectra of (*S*)-methyl lactate. The calculation was carried out using density-functional theory and the magnetic-field perturbation theory of VCD intensities. No adjustable parameters were used for the theoretical calculation other than choosing the bandshape for the vibrational transitions.

available commercially with the Gaussian 98 set of quantum-chemistry programs.

In addition to this new powerful approach to the determination of absolute configuration by VCD, it is also possible to gain information about the absolute configuration of members of a family of structurally related molecules by empirical correlation. Among the many bands present in a VCD spectrum, there are often one or more that serve as reliable markers of absolute configuration for a particular chiral centre. One of the best known examples, illustrated in **Figure 3**, is the methine CH stretch of amino acids and hydroxy acids. The *S*-enantiomer (L-amino acid) always exhibits a positive VCD band for reasons that are currently being explored using detailed quantum-chemistry calculations. Other examples of markers of absolute configuration abound, and such markers can typically be found for any set of structurally related molecules.

## Stereo-conformational analysis

The final area of the application of VCD spectroscopy to be discussed here is stereo-conformational analysis. This is the most sophisticated level of VCD application. Here one is concerned about determining the conformation of chiral molecules in solution. This is the principal research application of VCD, and VCD has been used for this purpose for many years. Nearly all classes of chiral molecules have been explored using VCD. Most of these are molecules of biological significance such as amino acids, peptides, sugars, proteins, nucleic acids, and most classes of pharmaceutical molecules. Many excellent reviews have been written on the application of VCD to the study of these molecules. Most recently it has been demonstrated that quantum calculations can be used to determine the presence and relative population of various solution-state conformers of chiral molecules. Theoretically predicted VCD spectra of the most stable conformers of the chiral molecule are compared to the experimental VCD spectrum. A best fit is then sought starting from a Boltzmann distribution of the theoretically determined spectra. Deviations from the VCD spectrum predicted by the Boltzmann distribution are explained in terms of the influence of the solvent on the stability of the various conformers. In this way, new information is obtained about the solution-state structures that are present under particular conditions of solvent, temperature and concentration.

## Future performance

The design and performance of VCD spectrometers have been advanced dramatically over the 25 years since the first measurements of VCD. Improvements are continuing today and the first commercially available VCD instrument was introduced only two years ago. Therefore, there is reason to expect advances to continue in the coming years making VCD of higher and higher quality accessible to those concerned about stereochemistry and molecular structure. With the advent of commercially available theoretical programs for the accurate simulation of VCD spectra, VCD spectroscopy is emerging as a powerful new tool for understanding the absolute structure and dynamics of chiral molecules in solution.

## List of symbols

A = absorbance; $g$ = ground electronic state vibrational sublevel; $I$ = intensity; $J_1$ = first-order Bessel function; $V$ = Fourier frequency; $\alpha_M$ = retardation angle; $\delta$ = mirror position; $\theta$ = phase function; $\nu$ = wavenumber frequency; $\tau$ = time constant of lock-in amplifier.

*See also:* **Biochemical Applications of Raman Spectroscopy; Biomacromolecular Applications of Circular Dichroism and ORD; Chiroptical Spectroscopy, Oriented Molecules and Anisotropic Systems; Chiroptical Spectroscopy, General Theory; ORD and Polarimetry Instruments; Raman Optical Activity, Applications; Raman Optical Activity, Spectrometers; Raman Spectrometers; Vibrational CD, Applications; Vibrational CD, Theory.**

## Further reading

Ashvar CS, Stephens PJ, Eggimann T and Wieser H (1998) Vibrational circular dichroism spectroscopy of chiral pheromones: frontalin (1,5-dimethyl-6,8-dioxabicyclo [3.2.1]octane). *Tetrahedron: Asymmetry.* **9**: 1107–1110.

Devlin FJ and Stephens PJ (1997) *Ab Initio* prediction of vibrational absorption and circular dichroism spectra of chiral natural products using density functional theory: alpha-pinene. *Journal of Physical Chemistry* A. **101**: 9912–9924.

Freedman TB, Long F, Citra M and Nafie LA (1999) Hydrogen stretching vibrational circular dichroism spectroscopy: absolute configuration and solution conformation of selected pharmaceutical molecules. *Enantiomer* (in press).

Gigante DMP, Long F, Bodack L *et al* (1998) Hydrogen stretching vibrational circular dichroism in methyl lactate and related molecules. *Journal of Physical Chemistry A.* (submitted for publication).

Holzwarth G, Hsu EC, Mosher HS, Faulkner TR and Moscowitz A (1974) Infrared circular dichroism of carbon–hydrogen and carbon–deuterium stretching modes. Observations. *Journal of the American Chemical Society* **96**: 251–252.

Keiderling TA (1990) Vibrational circular dichroism. Comparison of technique and practical considerations. In: Ferraro JR and Krishnans K (eds) *Practical Fourier Transform Infrared Spectroscopy. Industrial and Laboratory Chemical Analysis*, pp 203–284. San Diego: Academic Press.

Lipp ED, Zimba CG and Nafie LA (1982) Vibrational circular dichroism in the mid-infrared using Fourier transform spectroscopy. *Chemical Physics Letters* **90**: 1–5.

Long F, Freedman TB, Tague TJ and Nafie LA (1997) Step-scan Fourier transform vibrational circular dichroism measurements in the vibrational region above 2000 $cm^{-1}$. *Applied Spectroscopy* **51**: 508–511.

McCann JL, Rauk A and Wieser H (1998) A conformational study of (1S,2R,5S)-(+)-menthol using vibrational circular dichroism spectroscopy. *Canadian Journal of Chemistry* **76**: 274–283.

Nafie LA (1996) Vibrational optical activity. *Applied Spectroscopy* **50** (5): 14A–26A.

Nafie LA (1997) Infrared and Raman vibrational optical activity: theoretical and experimental aspects: *Annual Review of Physical Chemistry* **48**: 357–386.

Nafie LA and Freedman TB (1998) Vibrational circular dichroism: an incisive tool for stereochemical applications. *Enantiomer* **3**: 283–297.

Nafie LA, Cheng JC and Stephens PJ (1975) Vibrational circular dichroism of 2,2,2-trifluoro-1-phenylethanol. *Journal of the American Chemical Society* **97**: 3842.

Nafie LA, Diem M and Vidrine DW (1979) Fourier transform infrared vibrational circular dichroism. *Journal of the American Chemical Society* **101**: 496–498.

Nafie LA, Lipp ED and Zimba CG (1981) Fourier transform infrared circular dichroism: a double modulation approach. In: Sakals J (ed) *Proceedings of the 1981 International Conference on Fourier Transform Infrared Spectroscopy*, pp 457–468. SPIE.

Nafie LA and Vidrine DW (1982) Double modulation Fourier transform spectroscopy. In: Ferraro JR and Basiles LJ (eds) *Fourier Transform Infrared Spectroscopy*, pp 83–123. New York: Academic Press.

Nafie LA (1988) Polarization modulation FTIR spectroscopy. In: Mackenzies MW (ed) *Advances in Applied FTIR Spectroscopy*, pp 67–104. New York: John Wiley & Sons.

Polavarapu PL and Deng ZY (1996) Measurement of vibrational circular-dichroism below ~ 600 $cm^{-1}$ — progress towards meeting the challenge. *Applied Spectroscopy* **50**: 686–692.

Polavarapu PL (1997) Double Polarization modulation interferometry. *Applied Spectroscopy* **51**: 770–777.

Ragunathan N, Lee N-S, Freedman TB, Nafie LA, Tripp C and Buijs H (1990) Measurement of vibrational circular dichroism using a polarizing Michelson interferometer. *Applied Spectroscopy* **44**: 5–7.

Su CN, Heintz VJ and Keiderling TA (1980) Vibrational circular dichroism in the mid-infrared. *Chemical Physics Letters* **73**: 157–159.

# Vibrational CD, Applications

**Günter Georg Hoffmann**, Hoffmann Datentechnik, Oberhausen, Germany

Copyright © 1999 Academic Press



## Introduction

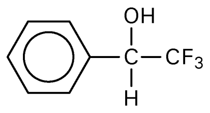
Only a few methods are available for the determination of the absolute configuration of chiral molecules. The most common are synthesis by chemical degradation from a molecule with reliably known stereochemistry, the X-ray method of Bijvoet, or the measurement of electronic circular dichroism (ECD). The chemical method has often been used in the history of chemistry when no other methods were available, but it is too time consuming to be generally applicable. The X-ray method is surely the most important, as it gives starting points for all other methods, but it cannot be applied to molecules that are not crystallizable and it is often too time consuming and comparatively demanding. The measurement of ECD requires a suitable chromophore. If none of these methods is practicable, the measurement of vibrational circular dichroism is a good choice. Since vibrational circular dichroism (VCD) and Raman optical activity (ROA) are complementary techniques (a stereochemical problem that cannot be solved by one technique can most probably be solved by the other), the interested reader should also consult appropriate articles of ROA.

The history of VCD is marked by instrumental advances: the early instruments were useful only in the near IR, first for measuring O–H and N–H stretching vibrations and their overtones, advancing to the C–H stretching region, later covering the C=O stretching vibration, and finally intruding into the fingerprint region. Advances in theory are also clearly discernible: first, simple coupled oscillator models were compared to the experimental spectra, then semiempirical calculations were made, later Hartree–Fock (HF) *ab initio* calculations, and now mainly density functional theory (DFT) calculations.

## Stereochemistry of small chiral molecules

In principle, the stereochemistry of a molecule can be determined by comparing the sign of a single band of one enantiomer with the calculated sign of that band. Unfortunately, the calculations are still not accurate enough for this method. One has to compare the enantiomer's spectrum in a larger spectral region with the calculated spectra of both configurations and then look for the best fit. In the early days of VCD spectroscopy, when the theory of VCD was not well developed, it was possible to derive 'chirality rules' for some classes of compounds.

The first publication of a single molecule VCD effect appeared in 1974. (*S*)-(+)- and (*R*)-(−)-2,2,2-tri-fluorophenylethanol [1] and (*R*)-(−)-neopentyl-1-*d*-chloride were examined using the C–H stretching vibration and its respective C–D analogue. As the signal-to-noise ratio of these first spectra was very low, the former compound was re-examined the following year by another group.



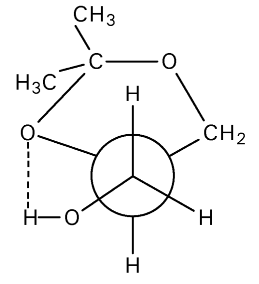

[1]

NMR data together with VCD data (O–H and C=O regions) were used to access the conformation of dimethyl tartrate and (2*S*)-(−)-malic acid dimethyl ester. A later investigation on the conformations of tartaric acid and its esters used *ab initio* calculations to find that the *trans*-COOH conformation with hydrogen bonding was the most stable. These results fit well with the VCD intensities in the C*–O stretching region of the examined compounds if charge flow along the C*–C* bond is assumed.

Using an empirical force field of the Urey–Bradley type, the infrared Raman, VCD and ROA spectra of chlorofluoroacetic acid and its anion were readily interpreted.

In an investigation of the O–H stretching vibration of the (*R*) enantiomer of 2,2-dimethyl-1,3-dioxoalane-4-methanol [2], the conformer containing an intramolecular hydrogen bridge showed a positive VCD effect at 3600 $cm^{-1}$. The free form of the alcohol showed no measurable VCD band.

Spectra of substituted allenes showed a correlation of the sign of the VCD for the asymmetrical stretching vibration of the C=C=C moiety (≈ 1950 $cm^{-1}$) with the absolute configuration. Thus for an (*S*) configuration the VCD was positive. Judging from two

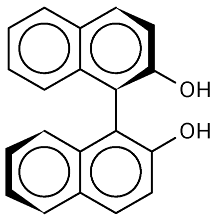



[2]

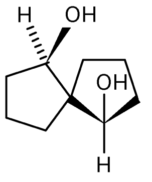
1-halo-3-*t*-butyl allenes, such a correlation also seems to exist in the C(X)–H stretching mode (≈ 3050 $cm^{-1}$).

Anisotropies from $+1.5 \times 10^{-4}$ to $+4.5 \times 10^{-4}$ were found for the VCD in the methine stretching mode of hydroxyacid methyl esters, whereas dimethyl-$d_6$-2,3-*O*-benzylidine-C-$d_1$-L-tartrate and (*S*)-methyl-2-chloropropionate showed only small VCD signals.

(*R*)-2,2′-Dihydroxy-1,1′-binaphthyl [3] was examined in the O–H stretching region and from 950 to 1700 $cm^{-1}$. For (1*R*, 5*R*, 6*R*)-(–)-spiro[4.4]nonane-1,6-diol [4], the theoretical VCD spectra were produced using vibronic coupling theory at the 6-31G level. A comparison of the crystal structure of the ketal of the compound with optically pure (+)-(5α-cholestan-3-one confirmed the results of the VCD determination.




[3]




[4]

Isoflurane [5] and desflurane [6] are relatively new fluorine-contaning anaesthetics. Unlike older anaesthetics such as diethyl ether or chloroform, they are chiral molecules; VCD spectra can be taken and, by comparison with theoretical spectra, their absolute configuration can be determined. This is especially

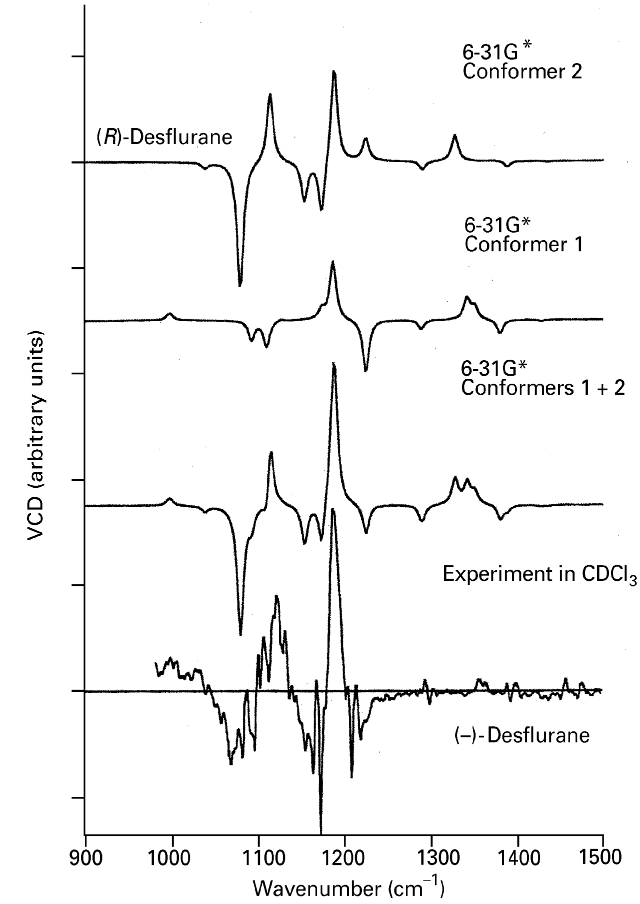
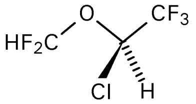


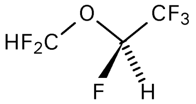

**Figure 1** VCD of desflurane. The experimental spectrum shows the (–)-enantiomer (corrected, as the original label incorrectly read (+) due to a confusion); 0.2 M solution in $CCl_3$). The theoretical calculations were done on the (*R*) configuration. Reprinted with permission from Polavarapu PL, Cholli AL and Vernice G (1992) Determination of absolute configurations and predominant conformations of general inhalation anesthetics: desflurane. *Journal of Pharmaceutical Sciences* **82**: 791–793. © 1992 American Chemical Society.

valuable as the enantiomers have different biological activities; for example the (+)-isomer of isoflurane is nearly twice as effective in activating the potassium current as the (–)-isomer. Experimental and theoretical spectra are shown for desflurane in **Figure 1**. The same assignment has been made for both compounds: the (*R*) configuration for the (–)-isomer and accordingly the (*S*) configuration for the (+)-isomer. For each isomer of desflurane, two dominant conformations were found. In a reinvestigation of the compound using DFT methods (see **Table 1**) and large basis sets, the configurational assignment was confirmed, but three different conformers contributing to the experimental spectrum have been proposed. A study on a third volatile anaesthetic used the same high level of theory: for (–)-1,2,2,2-tetrafluoroethyl methyl ether the (*R*) configuration was been derived

from the VCD spectrum and the *trans*-conformer was found to be dominant in $CCl_4$ solution.



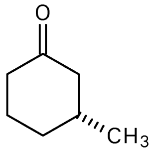
[5]



[6]

The enzymatic synthesis of (2*R*)-(+)-($^2$H) cyclohexanone and *trans*-(2,6-$^2H_2$)cyclohexanone has been reported together with the CD spectrum and the VCD spectrum in the C–H and C–D region.

For (3*R*)-(+)-methylcyclohexanone [7] in the C–H stretching and deformation region, no temperature dependence of the VCD was found. This leads to the conclusion that only one conformation is present in solution. Four of its chiral deuterated isotopomers were also examined in the C–H and the C–D regions. In the first report of VCD in the $CH_2$ bending region, 3-methylcyclohexanone and (+)-*trans*-1,2-cyclopropane dicarboxylic acid dichloride were investigated. The spectra of (*R*)-(+)-3-methylcyclohexanone and (*R*)-(+)-3-methylcyclopentatone showed negative bands in the region of the overtones ($\nu = 4$) of the C–H stretch, which had a distinctly larger rotatory strength in the case of the cyclopentanone derivative. The very first investigation in the poorly accessible 370–620 cm$^{-1}$ region was performed on (*R*)-3-methylcyclohexanone. The result was compared successfully with *ab initio* calculations.



[7]

Small rings (three- and four-membered) have been of great interest among the VCD spectroscopists. The rings have a rigid structure and the compounds are small enough to allow the theoretical spectra to be calculated in reasonable time. Optically active cyclopropanes were studied in the C–H stretch and $CH_2$ bending regions, and deuterated compounds in the respective regions. Spectra could readily be interpreted using the FPC model. The C=O region of a dimethyl ester and the C≡N stretching region of a dinitrile were also investigated. Here the coupled oscillator model could be used with advantage. The model failed for deformational modes.

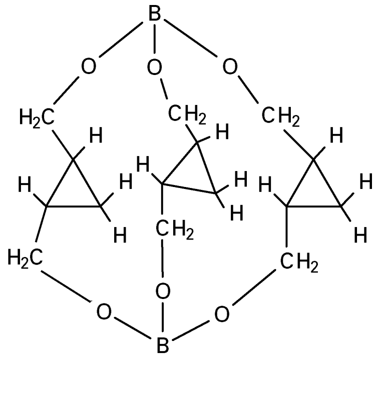
**Table 1** Models/methods cited

| | |
|---|---|
| FPC | Fixed partial charge |
| LMO | Localized molecular orbital |
| MFP | Magnetic field perturbation |
| APT | Atomic polar tensor |
| VCT | Vibronic coupling theory |
| EXC | Excitation scheme |
| HF | Hartree–Fock |
| DFT | Density functional theory |

In the VCD spectra of *trans*-2-phenylcyclopropane carboxylic acid and similar compounds, the symmetrical stretching vibration of the methylene group of the cyclopropane ring always showed a negative sign for the (1*R*, 2*R*) configuration.

With (*S*)-(+)-(1,2-$^2H_2$)cyclopropane in the gaseous phase, the region above 2000 cm$^{-1}$ could only be resolved to 7.2 cm$^{-1}$, but from 900 to 1500 cm$^{-1}$ a resolution of 1 cm$^{-1}$ was reached for the first time, allowing the observation of P, Q, and R branches in the VCD spectrum.

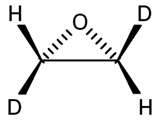
The crystal structure of the triply bridged diborate ester tris(*trans*-1,2,-cyclopropanediyldimethylene) diborate [8] has been determined and its VCD spectrum has been measured from 3150 to 2750 cm$^{-1}$ and from 1500 to 950 cm$^{-1}$.



[8]

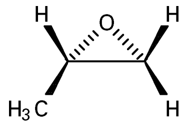
The symmetry of oxirane is lowered from $C_{2v}$ to $C_2$ by partial deuteration. The resulting (*S*,*S*)-(2,3-$^2H_2$) oxirane [9] exhibits two modes in the C–H as

well as two in the C–D stretching region of the infrared spectrum, corresponding to two couplets in the VCD spectrum. Their intensities are affected by a ring current mechanism (C–H) and a Fermi resonance (C–D). The molecule has also been investigated in the gaseous phase.



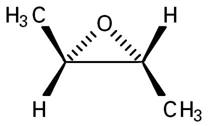

[9]

The VCD spectrum of (*S*)-(–)-epoxypropane [10] in the liquid and in the gaseous phase shows the splitting of the degenerate vibrational modes of the methyl group. Its analysis verified the VCD theory of the perturbed vibrational degenerate modes. Using a resolution of 1 $cm^{-1}$, the CD in the rotational–vibrational spectrum of (*R*)-(+)-methyloxirane has been measured with the result that the Q branch in some bands has the opposite sign to the R and P branches. This can be explained if methyloxirane (in spite of its chirality) is an approximate symmetrical top.




[10]

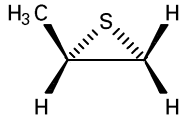
The absolute configuration of *trans*-2,3-dimethyloxirane [11] (2*R*,3*R* for the (+)-enantiomer) has been derived from the VCD spectra and *ab initio* calculations and is consistent with that determined by complexation chromatography.




[11]

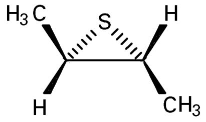
The vibrational circular dichroism of both enantiomers of methyloxirane [12] has been measured in $CCl_4$, in $CS_2$, and in the gaseous phase. The experimental spectra have been compared with a wide variety of theoretical calculations.

An extensive analysis of the experimental VCD of *trans*-2,3-dimethyloxirane [13] and its 2,3-$d_2$-isotopomer was published for 850–1650 $cm^{-1}$.



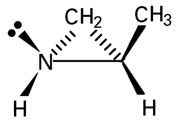

[12]

Comparing the experimental spectra with various calculations, the best results were obtained using the VCT model and the basis set 6–31G$^{*(0,3)}$.




[13]

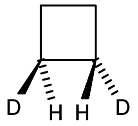
(2*R*)-2-Methylaziridine exists in solution as a mixture of the invertomers. According to *ab initio* calculations the ratio of (1*R*,2*R*)-2-methylaziridine (*trans*) [14] to (1*S*,2*R*)-2-methylaziridine (*cis*) should be 70 to 30. The experimental VCD spectrum is dominated by the effects of the *trans* isomer, as this is not only in excess but also shows greater rotatory strengths.




[14]

The heterocycles 1,2,- and 2,3,-dimethylaziridine were measured from 1500 to 1800 $cm^{-1}$ and the experimental VCD spectra were compared with theoretical calculations using the VCT model.

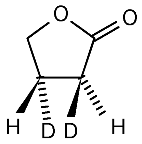
Comparison of the experimental spectra of *trans*-1,2-dideuteriocyclobutane [15] with the FPC as well as with the LMO model shows the former to be considerably more reliable.




[15]

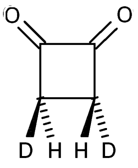
The synthesis, normal coordinate analysis and VCD spectrum have been reported for (2*S*,3*S*)-dideuteriobutyrolactone [16]. Comparison of the latter

with *ab initio* calculations using the MFP method and basis set 6–31G** yielded good qualitative agreement.



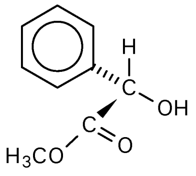

[16]

In a study on the VCD of (3*R*,4*R*)-dideuterio-cyclobutane-1,2-dione [17] the experimental spectra were compared with the calculated rotatory strengths using the MFP model, with good agreement.




[17]

In the region of methyl deformational modes and in the C–H stretching region, the VCD of $\alpha$-phenylethylamine, $\alpha$-phenylethanol, $\alpha$-phenylethylisocyanate, *p*-bromophenylethylamine and (*S*)-methyl mandelate [18] was examined. The bands near 1450cm$^{-1}$ were explained by interaction of the $CH_3$ deformational mode with an energetically neighbouring phenyl vibration.




[18]

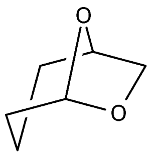
With 1-phenylethanol, 1-phenylethanethiol, 1-chloro-1-phenylethane, D-$\alpha$-phenylglycine-*N*-$d_3$ and (*S*)-methyl mandelate, the VCD of the methine stretching vibration is enhanced by ring currents. For methyl mandelate, a very large value of $\Delta\varepsilon = 5 \times 10^{-3}$ is found for the O–H stretching vibration.

$\alpha$-Phenylethylamine, $\alpha$-phenylethyl alcohol, $\alpha$-phenylethyl isocyanate and methyl mandelate were measured in the 1625–860 cm$^{-1}$ region. The VCD effects, which occur at 1368 and 1182 cm$^{-1}$ with phenylethylamine, were correlated with the stereochemistry of the molecules.

A simple chirality rule was derived for six phenylcarbinols: orienting the fourth substituent to the back and arranging the remaining three substituents in the order OH–Ph–H clockwise, one finds a negative VCD band at 1200 cm$^{-1}$; orienting them counterclockwise results in a positive effect.

Deuterated phenylethanes were observed in the region from 3100 to 2000 cm$^{-1}$. All aliphatic C–H and C–D stretching vibrations could be assigned.

The measurement and the theoretical calculation of the VCD spectrum of 6,8-dioxabicyclo[3.2.1]octane [19] was presented together with a detailed *ab initio* normal coordinate analysis using the APT and the FPC models. In another study, mono- and dimethyl derivatives of 6,8-dioxabicyclo[3.2.1]octane were treated the same way, but compared with calculations of higher accuracy (see also the pheromone [34]). The determination of absolute configuration by VCD has been made for some *exo*-7-derivatives of 5-methyl-6,8-dioxabicylco[3.2.1]octane (R=H, OH, Br or $CH_3$). Using recurring patterns in the 1100–1400 cm$^{-1}$ region, the chiral unit – C* ($CH_2$R)X–, with X=O or S was detected in rings of different size. The signs of these patterns corresponds to absolute configuration.




[19]

## Crystals

The first sample in which VCD was detected unambiguously was a thin slice of a crystal of $\alpha$-nickel sulfate hexahydrate. The compound crystallizes as tetragonal bipyramids in the narrow temperature range 31.5–53.3°C. The same paper, which was published in 1973, also reported the VCD of $\alpha$-$ZnSeO_4 \cdot 6H_2O$. As nickel sulfate is an achiral molecule, the VCD bands can be ascribed to vibrations of the chiral array of water molecules. Five main bands were found: at 2300 cm$^{-1}$ ($\nu_2$ + librations), at 4000 cm$^{-1}$ ($\nu_3$ or $\nu_2$ + librations), at 4200 cm$^{-1}$, at 4350 cm$^{-1}$, and the first part of a strong negative band at 5100 cm$^{-1}$ ($\nu_2 + \nu_3$), which was expected to be symmetrical. $\alpha$-Nickel sulfate has been reinvestigated by the author's group and the latter band showed a sawtooth shape with the steeper descent

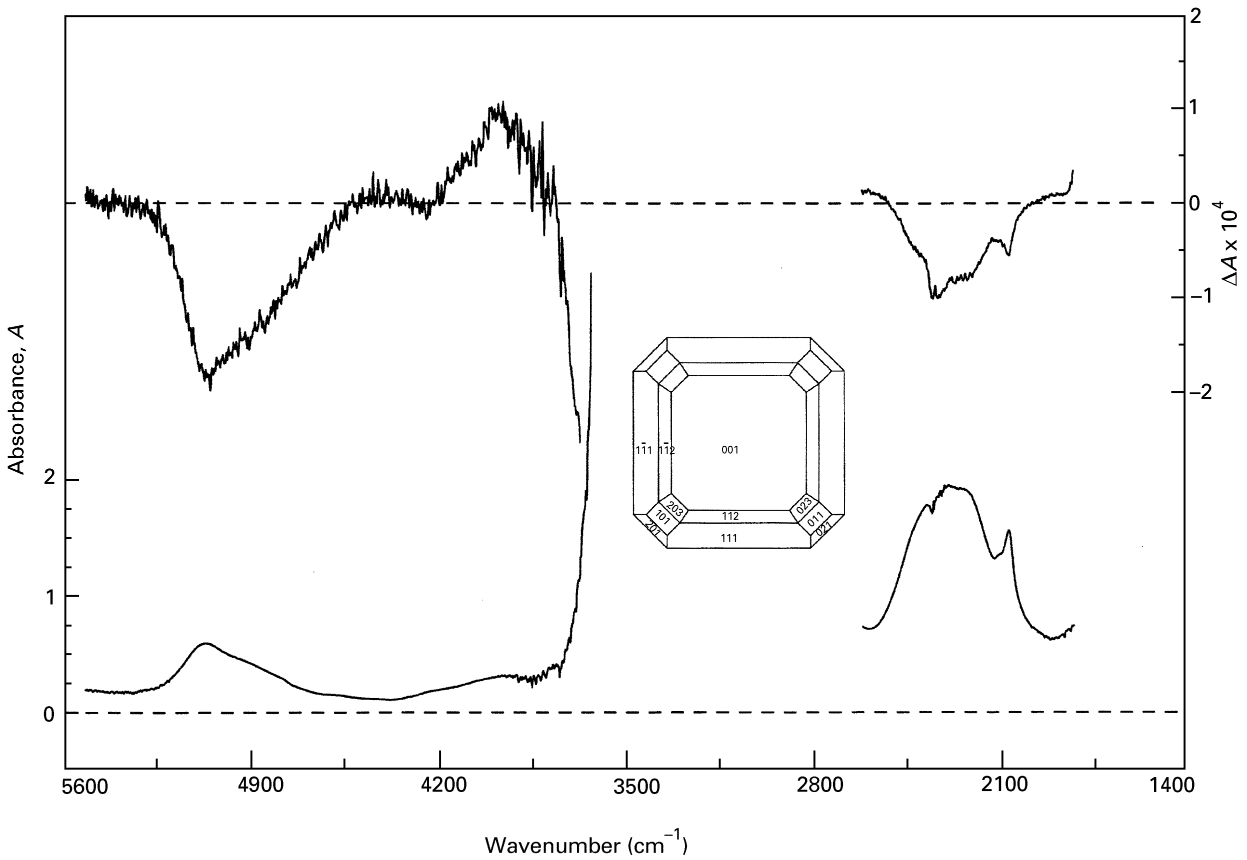



**Figure 2** VCD (upper) and absorbance (lower) spectra of a thin crystal slice of $\alpha$-nickel sulfate hexahydrate ($d$ = 63 µm, cleaved parallel (001) as shown in inset). No spectrum could be taken of bands with absorbance > 2.

on the side of shorter wavelength (**Figure 2**). We also found a new band at 2070 cm$^{-1}$, which, owing to lower resolution, had formerly only been detected in the absorption spectrum and had been attributed to the first overtone of the $\nu_3$ vibration of the $SO_4^{2-}$ ion (species $F_2$ at 1104 cm$^{-1}$).

## Liquid crystals

Studies of the optical activity of liquid crystalline solutions in the infrared region were published one year before the publication of single-molecule VCD effects. Using a solution of 2 mol% *d*-carvone in a liquid-crystalline eutectic of the isomeric *N*-oxides of *p*-methoxy-*p*′-*n*-butylazobenzols, huge effects were observed from the liquid crystal forced into a helical arrangement (cholesteric state) by the chiral solute. The liquid crystal acts as a molecular amplifier and reliably allows the determination of absolute configuration using only tiny amounts of substance. Later the same year, the VCD of a solution of 2% (–)-menthol in *N*-(*p*-methoxybenzylidene)butylaniline was published. Again the effects were extraordinarily large.

Infrared circular dichroism can also be measured using an ATR arrangement consisting of a wire polarizer followed by a KRS5 half-cylinder as ATR element. The spectrum of a 1% solution of cholesteryl chloride in the liquid crystal ZLI-887 was recorded from 4000 to 400 cm$^{-1}$ and compared favourably with a VCD spectrum recorded by the common-beam technique.

## Organometallic compounds

Relatively few papers have been published on the VCD of organometallic complexes. This may be because the common organic ligands are achiral and chirality has to be sought in the arrangement of the ligands around the central atom. The very first spectra of organometallic compounds were of electron transitions of $Pr^{3+}$–tartrate complexes down to 2000 cm$^{-1}$. The next published spectra of real vibrational transitions featured the C–H region only. These studies were on tris(3-trifluoromethylhydroxymethylene-*d*-camphorato) complexes of europium and praseodymium. Studies on complexes with amino acids, ethylenediamine and acetylacetate followed.

The two diastereomeric complexes Δ- and Λ-bis(acetylacetonato)(L-alaninato)cobalt(III) give rise to VCD spectra that can be explained using the degenerate coupled oscillator model (antisymmetric C=O stretching at 1522 $cm^{-1}$) and the ring current mechanism (N–H stretching). The appearance of the latter is illustrated by the fact that while they give nearly equal absorption spectra, the VCD of the Δ-isomer is nearly ten times larger than that of the Λ-isomer.

The spectral data (C–H stretching) of five copper complexes of amino acids and (Δ)*α*′-tris(L-alaninato)cobalt(III) have been obtained. As for the parent amino acids at pH values favouring hydrogen bonds, one detects an enhancement of the methine band by ring current effects. This is even larger as a result of the better closure of the ring by the transition metal ion.

Complexes of the trivalent cobalt and chromium with ethylenediamine have also been examined by VCD. Again substantial enhancements of the VCD effects by ring currents are found for the N–H and C–H stretching vibrations.

The peptide *cyclo*-(Pro-Gly-)$_3$ forms complexes with different alkali and alkaline earth metals that show spectra with sensitivity to the conformation of the peptide; the arrangement of the carbonyl groups is especially of interest. The solvent plays the most important role in the development of conformation. The size of the ion-binding cavity formed by the carbonyl groups and the size and charge of the cation are only of secondary importance.

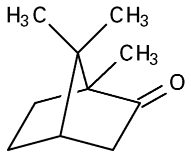
In studies of the interaction of two deoxyribo-oligonucleotides with divalent manganese ions, the resulting changes in the VCD spectra of d(GC)$_{20}$•d(GC)$_{20}$ and d(ATGCATGCAT)•d(ATGCATGCAT) were interpreted in terms of structural changes.

## Biochemical applications

### Terpenes

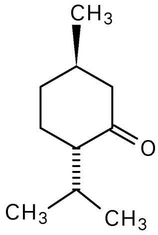
The first VCD investigation on camphor [20] observed the C–H stretching in the principal region as well as in the first overtone and combination regions; the first overtone of the C=O stretching was also measured. Later reports showed VCD also in the mid IR and presented calculational results.

Inherently dissymmetric chromophores, meaning groups that do not gain their chirality only from the influence of their neighbourhood, have always been of interest to investigators studying optical activity, including VCD spectroscopists. The sequence of signs could be correctly predicted for 15 different molecules (including cyclohexanones, menthone



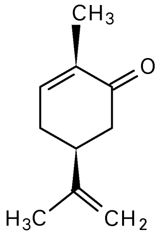

[20]

[21], pinenes, pulegone, and other natural products) featuring the $CH_2$–$C_2H$–C*H moiety.




[21]

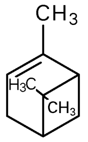
Using a FT-VCD spectrometer, the spectra of (+)-3-methylcyclohexanone, (+)-carvone [22] and (–)-*α*-pinene [23] were observed in the mid-infrared region; a higher signal-to-noise ratio and twice as great anisotropy were obtained than with dispersive instruments. Matrix-isolated molecules feature even larger anisotropies. Accessing the band at about 2920 $cm^{-1}$, one finds values of $5.4 \times 10^{-4}$ for (–)-*α*-pinene and $-6.5 \times 10^{-4}$ for (–)-*β*-pinene, which really are record values (excepting the huge value of 0.02 for methaemoglobin azide).




[22]

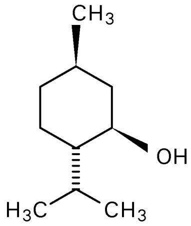
For the six monoterpenes (*S*)-(–)-limonene, (*R*)-(+)-limonene, (*S*)-(–)-perillyl alcohol, (*S*)-(–)-perillaldehyde, (*R*)-(+)-*p*-menth-1-ene and (*R*,*R*)-(+)-*p*-menth-1-en-9-ol, the VCD spectra of the second, third, and fourth overtones of the C–H stretching vibration have been published. The observed couplets can be attributed to a coupled vibration of the $CH_2CH_2C$*H fragment.

Other terpenes studied subsequently in the mid IR include nopinone and (–)-borneol, and a detailed




[23]

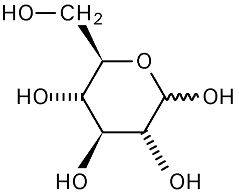
study has focused on the conformers of (+)-menthol [24].




[24]

#### Carbohydrates

Sugars are very good candidates for the measurement of VCD as the more common ECD is dependent on a chromophore, which is almost always absent in this class of natural compounds. The examination of the VCD spectra of six common sugars revealed a chirality rule for the 1150 $cm^{-1}$ band in deuterated dimethyl sulfoxide. Later the FT-VCD spectra of the carbohydrates D-fucose, D-arabinose, D-ribose, D-galactose and D-glucose [25] and their isotopomers deuterated at the hydroxyl group were examined in the same solvent. Some useful correlations between structure and spectra are found, but also some deviations.



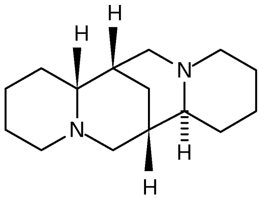

[25]

Cyclodextrins are water-soluble cyclic oligomers of glucose, the most common of which are $\alpha$-, $\beta$- and $\gamma$-cyclodextrin with six, seven or eight glucose moieties, respectively. Owing to their conical shape with a hydrophobic interior and a hydrophilic exterior, they form water-soluble complexes with inorganic or organic compounds. Comparison of the VCD spectra of the $\alpha$- and $\beta$-cyclodextrins with hydroxyl-deuterated $\alpha$-cyclodextrin, cyclodextrin–copper complexes and cyclodextrin inclusion complexes with methyl orange, methyloxirane, *n*-propanol and substituted cyclohexanones sensitively monitors structural changes in dimethyl sulfoxide-$d_6$.

#### Alkaloids

The spectrum of calycanthin in the C–H and N–H stretching regions can be interpreted as due to the coupling of the chromophore with the substituents. This is in contrast to the common coupling of the two chromophores in chiral dimers, which is commonly used to explain the electronic CD.

VCD investigation of a $CCl_4$ solution of (–)-sparteine [26] – the alkaloid from lupin beans – and comparison of the experimental results with calculations using the new EXC theory gave adequate agreement for such a large molecule.



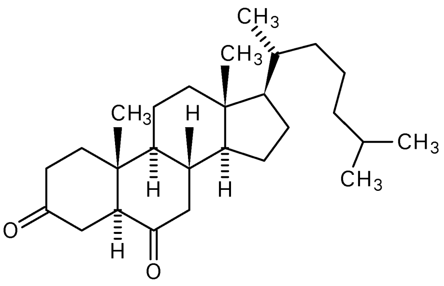

[26]

Vibrational CD in the O–H and N–H stretching bands of the anticancer chemotherapeutic agent taxol and two of its side-chain derivatives has been measured and compared with calculations on taxol fragments.

#### Steroids and their precursors

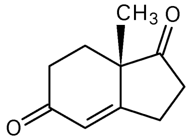
The simple coupled oscillator model, which can readily be applied to large molecules with two identical oscillators, originates from electronic CD. An example of the applicability of the model is given by steroids carrying two carbonyl functionalities. Even this simple model gives good results for the closely related steroids 3,6-dioxo-5$\alpha$-cholestane [27], 3,6-dioxo-5$\beta$-cholic acid methyl ester, 3,7-dioxo-5$\beta$-cholic acid methyl ester, 7,12-dioxo-5$\beta$-cholic acid, 3$\alpha$-hydroxy-7,12-dioxo-5$\beta$-cholic acid, and 3-oxo-5$\beta$-cholic acid with only one exception (the 3,7-dioxo derivative).

(+)-5,6,7,8-Tetrahydro-8-methylindane-1,5-dione [28] is an important precursor in the synthesis of estrone. The signs of its experimental VCD spectrum in the 1400–850 $cm^{-1}$ region can be reproduced adequately even using the small basis set 6–31G. In a later paper the spectra of the target molecule, estrone

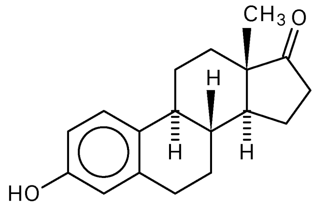


[27]

[29], were calculated with larger basis sets using HF and DFT methods.



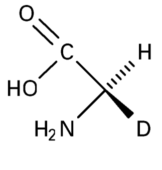
[28]



[29]

## Amino acids

The simplest amino acid studied is (*S*)-(–)-glycine-$C_\alpha$-$d_1$ [30]. Its weak VCD in the methine stretching at 2990 cm$^{-1}$ was studied together with those of L-alanine and L-proline, which in contrast to [30] show a positive effect.



[30]

Nineteen different amino acids were examined using the C–H stretching vibrational region. Aided by these VCD spectra, shown explicitly only for L-valine-N-$d_3$, a chirality rule was deduced for the chiral methine. If it is supposed that the amino acids studied form an intramolecular ring in aqueous solution, the VCD of the C*–H stretching vibration will be enhanced so much by ring currents (**Figure 3**), that it will obscure all other vibrations in this region. A positive effect with a value of more than $10^{-4}$ cm$^{-1}$ L mol$^{-1}$ then indicates an L-amino acid.

The vibrational CD spectra of some L-amino acids have been recorded as a function of pH. A large positive bias has been found for the C–H stretching region at neutral or high pH, whereas at low pH the bias is absent and only very small VCD signals are observed. Again the large bias was attributed to ring currents that are possible in some conformations. Another study examined alanine and its deuterated isotopomers.

Seven *N*-acyl-*N*′-alkylamide derivatives of different amino acids were measured in $CCl_4$ and $CHCl_3$ using the spectral region 3600–3200 cm$^{-1}$. The local conformation of the amide moiety was determined as well as the hydrogen bridge bonds using the VCD spectra. Other derivatives studied include *N*-*t*-BOC-alanine and *N*-*t*-BOC-proline (BOC = butoxycarbonyl).

## Peptides and proteins

The determination of absolute configuration is not of importance in the study of peptides and proteins. Among the peptides that have been studied are polyalanines, polyprolines, polylysines, polytryosines and poly(γ-benzyl-L-glutamate), as well as gramicidin S and other cyclic peptides; proteins examined include α-chymotrypsin, cytochrome *c*, haemoglobin, myoglobin, ribonuclease S and triose-phosphate isomerase. VCD spectroscopy is applied with advantage to access the secondary and tertiary structures of these biopolymers. Only a few typical examples are given here from studies published during the 1990s.

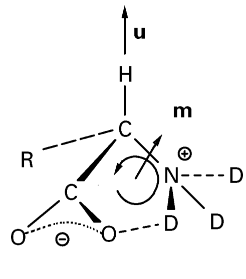
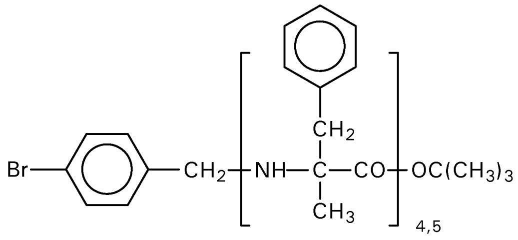


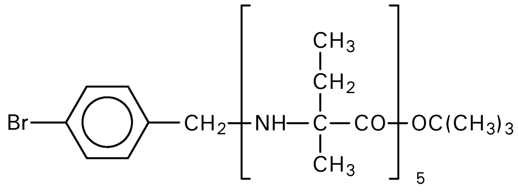
**Figure 3** Ring current mechanism (positive VCD) in the C–H stretching vibration of an L-amino acid. Redrawn from Freedman TB, Balukjian GA and Nafie LA (1985) Enhanced vibrational circular dichroism via vibrationally generated electronic ring currents. *Journal of the American Chemical Society* **107**: 6213–6222.




[31]

Using *ab initio* calculations of the model dipeptide $CH_3$–CONH–$CH_2$–CONH–$CH_3$, the VCD of the four most common secondary structures of proteins were calculated and compared successfully with experimental spectra of albumin, concanavalin A, $(Aib)_2Leu(Aib)_5$ and poly(L-lysine) (with Aib = $\alpha$-amino isobutyric acid). The main structure of these proteins is the $\alpha$-helical, $\beta$-sheet, $3_{10}$-helical and poly(L-proline)II conformation.

The favoured screw sense of homo-oligopeptides of $\alpha$-methylated phenylalanine and isovaline has been studied using *p*-BrBz-[D-($\alpha$Me)Phe]$_{4,5}$-OBu$^t$ [31] and *p*-BrBz-[D-Iva]$_5$-OBu$^t$ [32] in $CDCl_3$ solution. Analysis of their VCD spectra shows that the first two compounds are folded in a right-handed $3_{10}$-helix, whereas the last pentapeptide forms a left-handed helix.




[32]

The calcium-binding milk protein $\alpha$-lactalbumin and lysozyme from hen egg white show very different VCD spectra, though X-ray analysis reveals that the three-dimensional structures in the crystalline state are very similar. If one adds propanol to an aqueous solution of $\alpha$-lactalbumin, the helical regions become enlarged and the spectra become similar.

### Nucleotides and nucleic acids

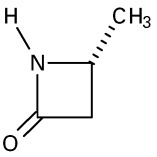
Base-sequence-characteristic bands have been found in the VCD spectra of six different octadeoxynucleotides in buffered $D_2O$. These bands belong to the C=O and C–C stretching regions and do not have a counterpart in the absorption spectrum.

Polyribonucleic acids can be measured in aqueous solution using the windows at 1750–1550 cm $^{-1}$. In contrast to the monomers, which do not show VCD in this region, the dimers and higher polymers show bisignate VCD bands. The spectra have been calculated for the dimers ApA and CpC using the coupled oscillator model.

### Other compounds

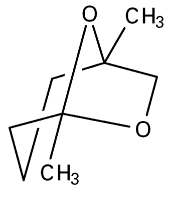
A study on the pharmaceutically applied ephedrines and pseudoephedrines derived valuable stereochemical information from the VCD spectra.

The model $\beta$-lactams 3-methyl- and 4-methylazetidine-2-one [33] readily form dimers in solution, as was clearly observed from the experimental VCD spectra and corresponding *ab initio* calculations.




[33]

A very interesting field of research in the biological area is the chemistry of pheromones. These chemicals strongly attract animals of the same species but of opposite sex. Stereochemistry is essential for the effectiveness of these substances. As pheromones are often applied in the struggle against insect pests, methods are needed to test the chirality of the natural and synthetic pheromones. For frontalin [34] the pheromone of the southern pine beetle (*Dendroctonus frontalis*), the VCD and absorption spectra of two different conformers have been calculated and compared with the experimental spectrum (**Figure** 4). In conformation *a* the six membered ring is in the chair conformation, whereas in conformation *b* it is in the boat conformation. The (1*R*,5*S*) configuration and the energetically more stable *a* conformation were assigned to the (+)-isomer on this basis.




[34]

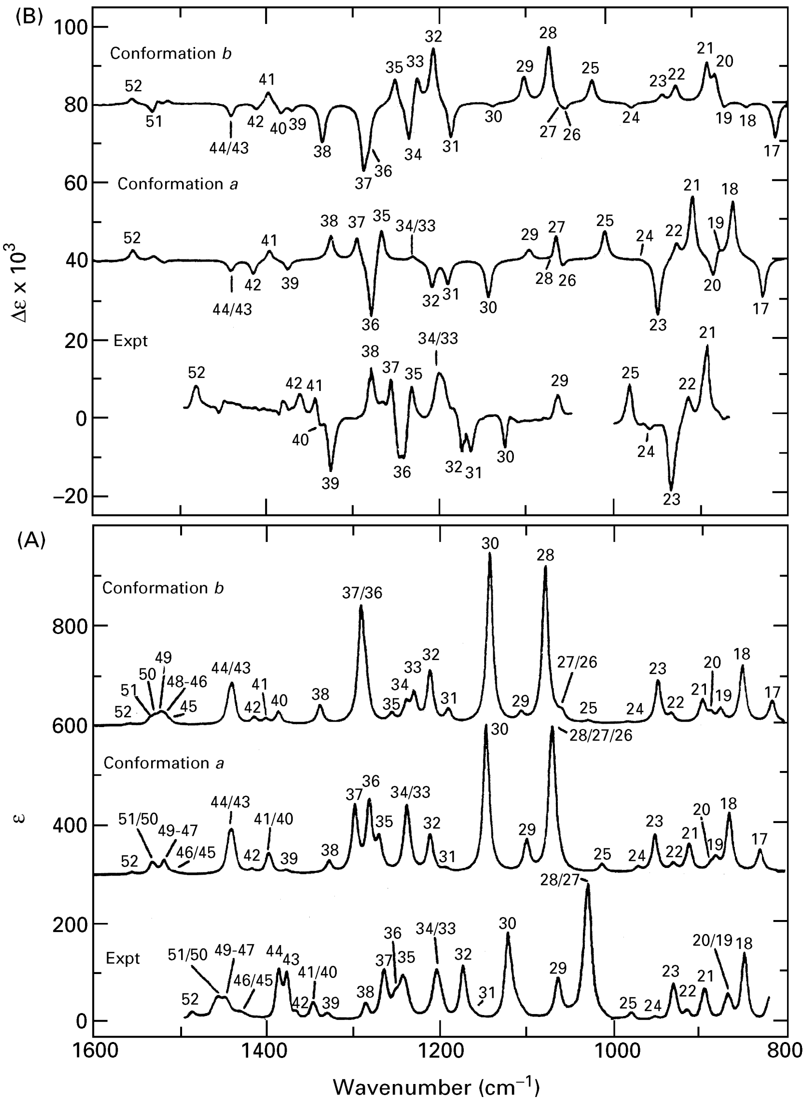



**Figure 4** Experimental (in $CCl_4$) and theoretical (*ab initio* DFT, B3LYP/6-31G*) spectra of (1*R*, 5*S*)-(+)-frontalin: (A) absorption spectra, (B) VCD spectra. Reprinted with permission from Ashvar CS, Stephens PJ, Eggimann T and Wieser H (1998) Vibrational circular dichroism spectroscopy of chiral pheromones: frontalin (1,5-dimethyl-6,8-dioxabicyclo[3.2.1]octane). *Tetrahedron: Asymmetry* **9**: 1107–1110. © 1998 Elsevier Science B.V.

## Synthetic polymers

Depending on the method of polymerization, a synthetic chiral polymer can be obtained from methyl methacrylate that has more or less extended isotactic regions. The VCD spectrum of the compounds is substantially more sensitive to its stereochemistry than is the normal IR spectrum.

Menthyl vinyl ether polymers of the diastereomeric menthols (+)-menthol, (+)-isomenthol and (+)-neomenthol have been synthesized and studied. While the menthyl and the neomenthyl derivatives both showed enhanced VCD features compared to the corresponding monomer, the VCD of the isomethyl derivative was found to stay virtually the same. Poly(menthyl vinyl ether) was studied in greater detail in later work.

## Other applications

### Magnetic VCD

If an achiral substance is put into a strong magnetic field, a VCD spectrum can be taken. Important information about molecules, such as the *g* value, can be obtained in this way.

### Chiral detection

A very sensitive detector using the VCD of the O–H stretching vibration can be constructed using a solid-state laser that is circularly polarized by a photoelastic modulator. Using this, 2,2,2-Trifluoro-1-(9-anthryl)ethanol and benzoin have been separated on the microgram scale by column chromatography on a chiral stationary phase.

### Kinetics

The signal-to-noise ratio of modern VCD spectrometers is now high enough to allow them to follow the course of a chemical reaction involving chiral molecules. Thus, the ratio of isomerization to stereomutation of 1.4±0.4 at 420°C was derived from study of the thermolysis of (1*R*,2*R*)-1-2-dideuteriocyclobutane, and the reaction constants of the racemization and isomerization of (2*S*,3*S*)-cyclopropane-1-$^{13}$C-1,2,3-$d_3$ were obtained at 407.0°C.

*See also:* **Biochemical Applications of Raman Spectroscopy; Biomacromolecular Applications of Circular Dichroism and ORD; Carbohydrates Studied by NMR; Circularly Polarized Luminescence and Fluorescence Detected Circular Dichroism; Induced Circular Dichroism; Magnetic Circular Dichroism, Theory; Nucleic Acids and Nucleotides Studied Using Mass Spectrometry; Organometallics Studied Using Mass Spectrometry; Polymer Applications of IR and Raman Spectroscopy; Proteins Studied Using NMR Spectroscopy; Vibrational CD Spectrometers; Vibrational CD, Theory.**

## Further reading

Ashvar CS, Devlin FJ, Stephens PJ, Bak KL, Eggimann T and Wieser H (1998) Vibrational absorption and circular dichroism of mono- and dimethyl derivatives of 6,8-dioxabicyclo[3.2.1]octane. *Journal of Physical Chemistry A* **102**: 6842–6857.

Bose PK and Polavarapu PL (1999) Vibrational circular dichroism of cyclodextrin complexes. *Journal of the American Chemical Society* **121**.

Hoffmann GG (1995) Vibrational optical activity (VOA). In: Schrader B (ed) *Infrared and Raman Spectroscopy – Methods and Applications*, pp 543–572. Weinheim: VCH.

Keiderling TA (1994) Vibrational circular dichroism spectroscopy of peptides and proteins. In: Nakanishi K, Berova N and Woody RW (eds) *Circular Dichroism – Principles and Applications*, pp 597–521. New York: VCH.

Keiderling TA (1996) Vibrational circular dichroism – applications to conformational analysis of biomolecules. In: Fasman GD (ed) *Circular Dichroism and the Conformational Analysis of Biomolecules*, pp 555–598. New York: VCH.

McCann JL, Rauk A and Wieser H (1998) A conformational study of (1*S*,2*R*,5*S*)-(+)-menthol using vibrational circular dichroism spectroscopy. *Canadian Journal of Chemistry* **76**: 274–283.

Nafie LA (1996) Vibrational optical activity. *Applied Spectroscopy* **50**: 14A–26A.

Nafie LA (1997) Infrared and Raman optical activity: theoretical and experimental aspects. *Annual Review of Physical Chemistry* **48**: 357–386.

Polavarapu PL (1998) *Vibrational Spectra: Principles and Applications with Emphasis on Optical Activity.* Amsterdam: Elsevier.

Rauk A and Freedmann TB (1994) Chiroptical techniques and their relationship to biological molecules, big or small. *International Journal of Quantum Chemistry* **28**: 315–338.

# Vibrational CD, Theory

**Philip J Stephens**, University of Southern California, Los Angeles, CA, USA

Copyright © 1999 Academic Press



## Introduction

Circular dichroism (CD) can be observed in the vibrational transitions of chiral molecules: vibrational circular dichroism (VCD). An example of a VCD spectrum is shown in **Figure 1**, together with the corresponding unpolarized absorption spectrum. The sample is a 0.6 M solution of (1*R*,4*R*)-(+)-camphor in $CCl_4$. Here, we discuss the theoretical analysis of VCD spectra. The current state-of-the-art is illustrated in **Figure 1**, where VCD and absorption spectra of camphor, predicted within the harmonic approximation (HA) using *ab initio* density functional theory (DFT), are shown.

## Theory

We restrict our discussion to the case of isotropic dilute solutions of randomly oriented molecules e.g. liquid solutions or amorphous solid solutions. (In practice, the vast majority of VCD experiments are carried out using liquids at room temperature.) Beer's Law applies:

$$A = \varepsilon c\ell \qquad [1]$$

$$\Delta A \equiv A_{\mathrm{L}} - A_{\mathrm{R}} = (\varepsilon_{\mathrm{L}} - \varepsilon_{\mathrm{R}})c\ell \equiv (\Delta\varepsilon)c\ell \qquad [2]$$

where $A$ = absorbance, L and R denote left and right circularly polarized light, $\Delta A$ = circular dichroism, $\varepsilon$ = molar extinction coefficient, $c$ = concentration (mol L$^{-1}$) and $l$ = pathlength (cm). The unpolarized absorption is

$$A \equiv \tfrac{1}{2}(A_{\mathrm{L}} + A_{\mathrm{R}}) = \tfrac{1}{2}(\varepsilon_{\mathrm{L}} + \varepsilon_{\mathrm{R}})\, c\ell \equiv \bar{\varepsilon} c\ell \qquad [3]$$

Semi-classical treatment of the interaction of molecules with electromagnetic waves leads to equations for $\varepsilon$ and $\Delta\varepsilon$ in terms of molecular properties:

$$\bar{\varepsilon}(\nu) = \frac{8\pi^3 N\nu}{(2.303)3000hc} \sum_{g,k} \alpha_g D_{gk} f_{gk}(\nu_{gk}, \nu) \qquad [4]$$

$$\Delta\varepsilon(\nu) = \frac{32\pi^3 N\nu}{(2.303)3000hc} \sum_{g,k} \alpha_g R_{gk} f_{gk}(\nu_{gk}, \nu) \qquad [5]$$

$$D_{gk} = |\langle g|\boldsymbol{\mu}_{\mathrm{el}}|k\rangle|^2 \qquad [6]$$

$$R_{gk} = \mathrm{Im}[\langle g|\boldsymbol{\mu}_{\mathrm{el}}|k\rangle \bullet \langle k|\boldsymbol{\mu}_{\mathrm{mag}}|g\rangle] \qquad [7]$$

where $g \rightarrow k$ is a molecular excitation of frequency $\nu_{gk}$, $\alpha_g$ is the fraction of molecules in state $g$, and $f(\nu_{gk}, \nu)$ is a normalized line shape function (e.g. Lorentzian). $D_{gk}$ and $R_{gk}$ are the dipole strength and rotational strength of the excitation $g \rightarrow k$. $\boldsymbol{\mu}_{\mathrm{el}}$ and $\boldsymbol{\mu}_{\mathrm{mag}}$ are the electric and magnetic dipole moment operators:

$$\boldsymbol{\mu}_{\mathrm{el}} = -\sum_i e\boldsymbol{r}_i + \sum_\lambda (Z_\lambda e)\boldsymbol{R}_\lambda \equiv \boldsymbol{\mu}^{\mathrm{e}}_{\mathrm{el}} + \boldsymbol{\mu}^{\mathrm{n}}_{\mathrm{el}} \qquad [8]$$

$$\boldsymbol{\mu}_{\mathrm{mag}} = -\sum_i \frac{e}{2mc}\left(\boldsymbol{r}_i \times \boldsymbol{p}_i\right) + \sum_\lambda \left(\frac{Z_\lambda e}{2M_\lambda c}\right) \times \left(\boldsymbol{R}_\lambda \times \boldsymbol{P}_\lambda\right) \equiv \boldsymbol{\mu}^{\mathrm{e}}_{\mathrm{mag}} + \boldsymbol{\mu}^{\mathrm{n}}_{\mathrm{mag}} \qquad [9]$$

Here, $-e$ and $Z_\lambda e$, $\boldsymbol{r}_i$, and $\boldsymbol{R}_\lambda$, $\boldsymbol{p}_i$ and $\boldsymbol{P}_\lambda$ are the charge, position and momentum of electron $i$ and nucleus $\lambda$ respectively. Equations [4] and [5] do not include the effects of the condensed-phase medium either on the molecular properties $\alpha_g$, $D_{gk}$, $R_{gk}$ and $\nu_{gk}$ or on the electromagnetic fields of the radiation: 'solvent effects'.

In the case of vibrational transitions, $g$ and $k$ are vibrational levels of the ground electronic state, $G$. Within the Born–Oppenheimer (BO) approximation:

$$\Psi_g(\boldsymbol{r}, \boldsymbol{R}) = \psi_G(\boldsymbol{r}, \boldsymbol{R})\chi_{Gg}(\boldsymbol{R}) \qquad [10]$$

$$\Psi_k(\boldsymbol{r}, \boldsymbol{R}) = \psi_G(\boldsymbol{r}, \boldsymbol{R})\chi_{Gk}(\boldsymbol{R}) \qquad [11]$$

where

$$H_{el}(\boldsymbol{r}, \boldsymbol{R})\, \psi_G(\boldsymbol{r}, \boldsymbol{R}) = W_G(\boldsymbol{R})\, \psi_G(\boldsymbol{r}, \boldsymbol{R}) \qquad [12]$$

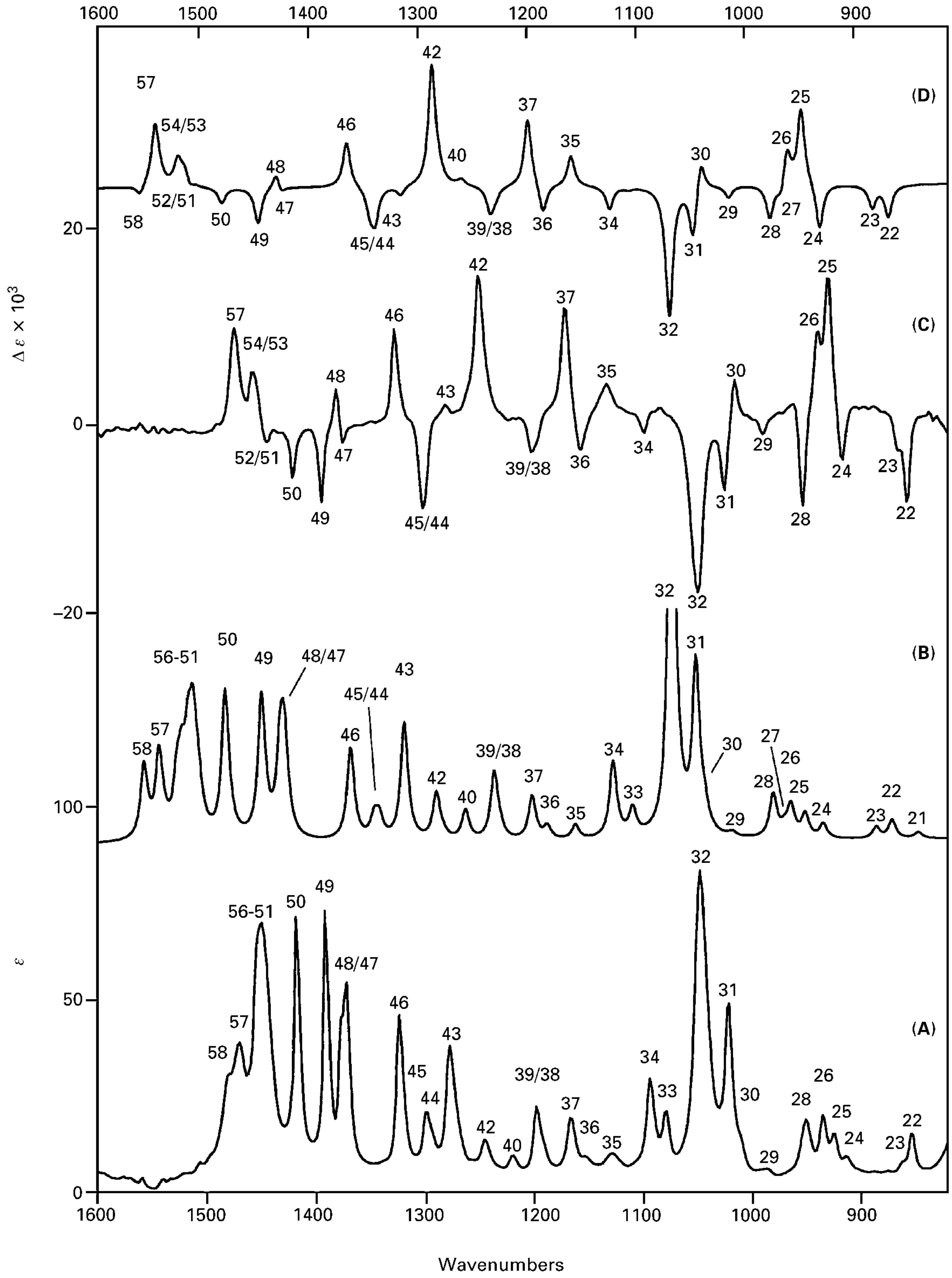



**Figure 1** Experimental (A,C) and calculated (B,D) absorption (A, B) and VCD (C, D) spectra of (1*R*, 4*R*)-(+)-camphor. The resolution of the experimental spectra was 4 $cm^{-1}$. Calculated spectra were obtained using DFT, B3PW91 and 6-31G*. Band shapes are Lorentzian (half width at half height 4 $cm^{-1}$). Fundamental modes are numbered.

$$[W_G(\boldsymbol{R}) + T_n(\boldsymbol{R})]\chi_{Gv}(\boldsymbol{R}) = E_v\chi_{Gv}(\boldsymbol{R}) \quad [13]$$

***r*** and ***R*** denote electronic and nuclear coordinates respectively. $H_{\text{el}}$ is the adiabatic 'electronic Hamiltonian':

$$H_{\text{el}} = T_{\text{e}} + V_{\text{ee}} + V_{\text{en}} + V_{\text{nn}} \quad [14]$$

comprising the electronic kinetic energy and the Coulombic interactions of electrons and nuclei. $\Psi_G$ and $W_G$ are the wavefunction and energy of the ground electronic state. $\chi_{Gv}$ and $E_v$ are the wavefunction and energy of the vibrational level $v$ arising from vibrational motion on the potential energy surface (PES) $W_G(\boldsymbol{R})$.

For simplicity, we restrict discussion now by assuming that only the lowest vibrational level is populated and that the PES, $W_G$, is harmonic:

$$\begin{aligned} W_G &= W_G^o + \tfrac{1}{2}\sum_{\lambda\alpha\lambda'\alpha'}\left(\frac{\partial^2 W_G}{\partial X_{\lambda\alpha}\partial X_{\lambda'\alpha'}}\right)_o X_{\lambda\alpha}X_{\lambda'\alpha'} \\ &= W_G^o + \tfrac{1}{2}\sum_i k_i Q_i^2 \end{aligned} \quad [15]$$

where $W_G^0$ is the energy of $G$ at equilibrium, $\boldsymbol{R} = \boldsymbol{R}_0$; $X_{\lambda\alpha}$ is the displacement of nucleus $\lambda$ ($\lambda = 1 \ldots N$) along Cartesian axis $\alpha$ ($\alpha = x,y,z$); $Q_i$ are normal coordinates, simultaneously diagonalizing the nuclear kinetic energy:

$$T_n = \tfrac{1}{2}\sum_i Q_i^2 \quad [16]$$

The force constants, $k_i$, determine the normal mode frequencies:

$$\nu_i = \frac{1}{2\pi}\sqrt{k_i} \quad [17]$$

The vibrational states of this harmonic PES are of energy

$$E(v_1, v_2 \ldots v_{3N}) = \sum_i \left(v_i + \tfrac{1}{2}\right) h\nu_i \quad [18]$$

$$(v_i = 0, 1, 2 \ldots)$$

For six modes, corresponding to translational and rotational motions, $k_i$ and $\nu_i$ are zero.

Within the HA, electric dipole transition moments are

$$\begin{aligned} \langle g|\boldsymbol{\mu}_{\text{el}}|k\rangle &\equiv \langle\psi_G\chi_{Gg}|\boldsymbol{\mu}_{\text{el}}|\psi_G\chi_{Gk}\rangle \\ &= \langle\chi_{Gg}|\langle\psi_G|\boldsymbol{\mu}_{\text{el}}|\psi_G\rangle|\chi_{Gk}\rangle \end{aligned} \quad [19]$$

which, on expanding $\langle\psi_G|\boldsymbol{\mu}_{\text{el}}|\psi_G\rangle \equiv \boldsymbol{\mu}_{\text{el}}^G$

$$\boldsymbol{\mu}_{\text{el}}^G = \left(\boldsymbol{\mu}_{\text{el}}^G\right)_o + \sum_i \left(\frac{\partial\boldsymbol{\mu}_{\text{el}}^G}{\partial Q_i}\right)_o Q_i + \ldots \quad [20]$$

leads to non-zero transition moments from the vibrational ground state (all $v_i = 0$) only for fundamental transitions involving one mode alone, i.e. to the states $v_i = 1$, $v_j = 0$ ($j \neq i$). The transition moment for the fundamental in mode $i$ is

$$\langle 0|\boldsymbol{\mu}_{\text{el}}|1\rangle_i = \left(\frac{\partial\boldsymbol{\mu}_{\text{el}}^G}{\partial Q_i}\right)_o \left(\frac{\hbar}{4\pi\nu_i}\right)^{1/2} \quad [21]$$

Equation [21] can be rewritten in terms of derivatives of the molecular electric dipole moment $\boldsymbol{\mu}_{\text{el}}^G$, with respect to the Cartesian displacement coordinates, $X_{\lambda\alpha}$. With

$$X_{\lambda\alpha} = \sum_i S_{\lambda\alpha,i}\, Q_i \quad [22]$$

Equation [21] becomes

$$\langle 0|(\boldsymbol{\mu}_{\text{el}})_\beta|1\rangle_i = \left(\frac{\hbar}{4\pi\nu_i}\right)^{1/2}\sum_{\lambda\alpha} S_{\lambda\alpha,i}\, P_{\alpha\beta}^\lambda \quad [23]$$

where

$$P_{\alpha\beta}^\lambda \equiv \left(\frac{\partial\left(\boldsymbol{\mu}_{\text{el}}^G\right)_\beta}{\partial X_{\lambda\alpha}}\right)_o \quad [24]$$

The second-rank molecular tensors, $P_{\alpha\beta}^\lambda$, are termed atomic polar tensors (APTs). Separating electronic and nuclear parts:

$$\begin{aligned} P_{\alpha\beta}^\lambda &= E_{\alpha\beta}^\lambda + N_{\alpha\beta}^\lambda \\ E_{\alpha\beta}^\lambda &= \left[\frac{\partial}{\partial X_{\lambda\alpha}}\left\langle\psi_G\middle|(\boldsymbol{\mu}_{\text{el}}^{\text{e}})_\beta\middle|\psi_G\right\rangle\right]_0 \\ N_{\alpha\beta}^\lambda &= (Z_\lambda e)\delta_{\alpha\beta} \end{aligned} \quad [25]$$

We can further write

$$E^{\lambda}_{\alpha\beta} = 2\left\langle \left(\frac{\partial\psi_G}{\partial X_{\lambda\alpha}}\right)_0 \middle| (\mu^{c}_{el})_\beta \middle| \psi^0_G \right\rangle \quad [26]$$

The dipole strength of the fundamental excitation of mode $i$ is then

$$D_i = \left(\frac{\hbar}{4\pi\nu_i}\right)\left|\left(\frac{\partial\left(\mu^G_{el}\right)}{\partial Q_i}\right)_0\right|^{X}$$
$$= \left(\frac{\hbar}{4\pi\nu_i}\right)\sum_\beta \sum_{\lambda\alpha,\lambda'\alpha'}\left[S_{\lambda\alpha,i}P^{\lambda}_{\alpha\beta}\right]\left[S_{\lambda'\alpha',i}P^{\lambda'}_{\alpha'\beta}\right] \quad [27]$$

The formulation of magnetic dipole transition moments is unfortunately less straightforward. Compare the electronic contributions to the electric and magnetic dipole moments of $G$:

$$\left(\mu^G_{el}\right)^{c} = \langle\psi_G|\mu^{c}_{el}|\psi_G\rangle \quad [28]$$

$$\left(\mu^G_{mag}\right)^{e} = \langle\psi_G|\mu^{c}_{mag}|\psi_G\rangle \quad [29]$$

Considering only non-degenerate electronic ground states (in practice very few chiral molecules are exceptions) $\mu^G_{el}$ and $\mu^G_{mag}$ are qualitatively different because $\langle\Psi_G|\mu^e_{mag}|\Psi_G\rangle = 0$ at all molecular geometries. That is, electrons make zero contribution to the adiabatic magnetic dipole moment. It follows that, in the case of magnetic dipole transition moments, the BO approximation leads to a non-physical result. The treatment of magnetic dipole transition moments requires more accurate vibronic wavefunctions. The vibrational states $g$ and $k$ must be written

$$\Psi_g = \psi_G\chi_{Gg} + \sum_{E\neq G}(\psi_E\chi_{Ee})c_{Gg,Ee}$$

$$\Psi_k = \psi_G\chi_{Gk} + \sum_{E\neq G}(\psi_E\chi_{Ee})c_{Gk,Ee} \quad [30]$$

allowing for the admixture of BO functions of excited electronic states $E$ into the ground state. This in turn permits non-zero vibrational transition moments of $\mu^e_{mag}$ to be obtained; simply put, electronic magnetic dipole transition moments are 'stolen' by mixing of BO states. The reader is referred to the literature for the details. The final result is that

$$\left\langle 0|(\mu_{mag})_\beta|1\right\rangle_i = \left(4\pi\hbar^3\nu_i\right)^{1/2}\sum_{\lambda\alpha} S_{\lambda\alpha,i}M^{\lambda}_{\alpha\beta} \quad [31]$$

where

$$M^{\lambda}_{\alpha\beta} = I^{\lambda}_{\alpha\beta} + J^{\lambda}_{\alpha\beta}$$
$$I^{\lambda}_{\alpha\beta} = \left\langle \left(\frac{\partial\psi_G}{\partial X_{\lambda\alpha}}\right)_0 \middle| \left(\frac{\partial\psi_G}{\partial H_\beta}\right)_0 \right\rangle$$
$$J^{\lambda}_{\alpha\beta} = \frac{i}{4\hbar c}\sum_\gamma (Z_\lambda e)R^0_{\lambda\gamma}\varepsilon_{\alpha\beta\gamma} \quad [32]$$

The tensors $M^{\lambda}_{\alpha\beta}$ are termed atomic axial tensors (AATs); $I^{\lambda}_{\alpha\beta}$ and $J^{\lambda}_{\alpha\beta}$ are the electronic and nuclear components. Here, $(\partial\psi_G/\partial X_{\lambda\alpha})_0$ is the same derivative which occurred already in Equation [26]. The electronic AAT, $I^{\lambda}_{\alpha\beta}$, is the overlap integral with the derivative $(\partial\psi_G/\partial H_\beta)_0$. This latter is defined via

$$H(H_\beta) = H_{el} + H'(H_\beta)$$
$$H'(H_\beta) = -\left(\mu^{e}_{mag}\right)_\beta H_\beta$$
$$H(H_\beta)\psi_G(H_\beta) = W_G(H_\beta)\psi_G(H_\beta) \quad [33]$$

That is: $\psi_G(\boldsymbol{H})$ is the wavefunction of $G$ in the presence of a uniform external magnetic field, $\boldsymbol{H}$, approximating the perturbation by the linear magnetic dipole interaction $H'(\boldsymbol{H})$.

The rotational strength of the fundamental excitation of mode $i$ is then

$$R_i = \hbar^2\sum_\beta \sum_{\lambda\alpha,\lambda'\alpha'}\left[S_{\lambda\alpha,i}P^{\lambda}_{\alpha\beta}\right]\left[S_{\lambda'\alpha',i}M^{\lambda'}_{\alpha'\beta}\right] \quad [34]$$

## Computation

Within the HA, the prediction of a vibrational absorption spectrum amounts to the calculation of the harmonic normal mode frequencies, $\nu_i$, and dipole strengths, $D_i$. The frequencies are obtained from the harmonic force field (HFF). With respect to Cartesian displacement coordinates, this is the Hessian $(\partial^2 W_G/\partial X_{\lambda\alpha}\partial X_{\lambda'\alpha'})_0$. Diagonalization (after mass-weighting) yields the force constants $k_i$; the frequencies, $\nu_i$; and the normal coordinates, $Q_i$, i.e. the

transformation matrices, $S_{\lambda\alpha,i}$. The dipole strengths depend in addition on the APTs; these require calculation of $(\partial\psi_G/\partial X_{\lambda\alpha})_0$.

The prediction of a VCD spectrum amounts likewise to the calculation of the harmonic frequencies and rotational strengths, $R_i$. All of the quantities required in predicting the absorption spectrum are again needed; in addition, the AATs must be calculated. Since $(\partial\psi_G/\partial X_{\lambda\alpha})_0$ is already required for the APTs, the AATs require additionally only $(\partial\psi_G/\partial H_\beta)_0$.

In sum: the prediction of both absorption and VCD spectra requires (i) $(\partial^2 W_G/\partial X_{\lambda\alpha}\partial X_{\lambda'\alpha'})_0$; (ii) $(\partial\psi_G/\partial X_{\lambda\alpha})_0$; (iii) $(\partial\psi_G/\partial H_\beta)_0$. The prediction of the VCD spectrum requires relatively little more than is needed for the absorption spectrum: specifically, $(\partial\psi_G/\partial H_\beta)_0$.

The calculation of molecular properties can be carried out at three distinct levels: (i) *ab initio*, (ii) semi-empirical, (iii) empirical. *Ab initio* methods have increased enormously in accuracy and efficiency in the last two decades and are the focus of our discussion here. *Ab initio* methods have developed in two directions: first, the level of approximation has become increasingly sophisticated and, hence, accurate. The earliest *ab initio* calculations used the Hartree-Fock/self-consistent field (HF/SCF) methodology, which is the simplest to implement. Subsequently, such methods as Møller-Plesset perturbation theory, multi-configuration self-consistent field theory (MCSCF) and coupled-cluster theory have been developed and implemented. Relatively recently, density functional theory (DFT) has become very popular, since it yields an accuracy much greater than that of HF/SCF while requiring relatively little additional computational effort.

The second dimension in which *ab initio* theory has progressed is that of derivative techniques. Many molecular properties of interest–including, as shown above, the HFF, APTs and AATs–can be expressed in terms of derivatives of energies and wavefunctions with respect to perturbations. Such derivatives can be evaluated using either numerical or analytical methods. For example, the energy gradients $(\partial W_G/\partial X_{\lambda\alpha})_0$ can be evaluated either by calculating $W_G$ at $\boldsymbol{R}_0$ and $\boldsymbol{R}_0 + X_{\lambda\alpha}$ and using

$$\left(\frac{\partial W_G}{\partial X_{\lambda\alpha}}\right)_0 \approx \frac{W_G\left(\boldsymbol{R}_0 + X_{\lambda\alpha}\right) - W_G\left(\boldsymbol{R}_0\right)}{X_{\lambda\alpha}} \quad [35]$$

or by formulating an equation for $(\partial W_G/\partial X_{\lambda\alpha})_0$ and then carrying out direct evaluation. Similarly, a Hessian matrix can be obtained by finite-differences of gradients or analytically. Analytical derivative methods are much more efficient. Much of the recent expansion in usage of *ab initio* quantum chemistry has resulted from advances in formulating and implementing analytical derivative techniques for an increasing diversity of molecular properties at an increasing number of theoretical levels.

At the present time, the simultaneous calculation of HFFs, APTs and AATs using analytical derivative *ab initio* methods has been implemented in three program packages: CADPAC, DALTON and GAUSSIAN. The levels of implementation are:

| | |
|---|---|
| CADPAC | HF/SCF |
| DALTON | HF/SCF and MCSCF |
| GAUSSIAN | HF/SCF and DFT. |

The accuracies of these methods are:

HF/SCF < MCSCF << DFT.

The computational effort is:

HF/SCF < DFT << MCSCF.

The ratio of accuracy to effort is:

DFT >> HF/SCF > MCSCF.

Thus, DFT is currently the most cost-effective methodology available.

An additional variable in *ab initio* calculations is the basis set. Two choices are to be made: (i) perturbation-independent or perturbation-dependent; (ii) size and composition. In calculating derivatives with respect to nuclear displacements, $X_{\lambda\alpha}$, one can adopt basis functions which either (a) are not or (b) are functions of nuclear position. The latter add computational complexity but vastly improve convergence of properties with increasing basis set size (i.e. decrease the errors associated with the use of basis sets of finite size). Modern computational packages use only nuclear-position-dependent basis sets. In the same way, derivatives with respect to magnetic fields can use basis functions which either (a) are not or (b) are functions of magnetic field. The standard choice for the latter are so-called London orbitals or gauge-invariant atomic orbitals (GIAOs). The use of GIAOs vastly reduces basis set error and is increasingly *de rigueur* in computation of magnetic properties (e.g. NMR shielding tensors). With regard to the implementation of AATs in CADPAC, DALTON and GAUSSIAN, we should add that DALTON and GAUSSIAN use GIAOs, while CADPAC does not.

With respect to basis set size we can simply note that (a) accuracy increases with increasing basis set size; (b) the rate of increase in accuracy is rapid at small sizes and less rapid at large sizes; (c) with respect to the ratio of accuracy to size, such basis sets

as 6-31G* are generally regarded as optimal for calculations on organic molecules.

Finally, in DFT calculations there is the question of the density functional. The accuracy of DFT calculations varies greatly with the choice of functional. The exact functional gives exact results. Very crude functionals give very inaccurate results. Functionals used in the recent past can be grouped into three classes: (a) local; (b) non-local/gradient-corrected; (c) hybrid. Overall, the relative accuracy is

$$\text{local} < \text{non-local} < \text{hybrid}.$$

At this time, hybrid functionals are generally regarded as state-of-the-art. There are many: the original is 'B3PW91'; a popular-choice is 'B3LYP'. Undoubtedly, current functionals will be soon replaced by yet more accurate functionals.

## Implementation

In **Figure 1** we compare predicted absorption and VCD spectra for camphor to experiment. Predictions are based on DFT calculations using the B3PW91 functional and the 6-31G* basis set. AATs are calculated using GIAOs. The calculations were carried out using the GAUSSIAN program. Spectra are simulated from frequencies, dipole strengths and rotational strengths assuming Lorentzian bandshapes of constant width (half width at half height 4 $cm^{-1}$).

Focussing first on the absorption spectrum, we observe an excellent one-to-one correspondence between predicted and experimental spectra. That is: we can assign the bands of the experimental spectrum to fundamental excitations in such a way that the pattern of frequencies and intensities is extremely similar to that predicted. One notices an overall shift of the predicted spectrum to higher frequency. This shift is attributable both to error in the calculated harmonic frequencies and to anharmonicity, which uniformly lowers experimental frequencies with respect to harmonic frequencies. Calculations on very small molecules, where harmonic frequencies are known, indicate that the two contributions are of the same sign and comparable in magnitude. It is also to be noted that almost all bands of the experimental spectrum can be assigned to fundamental transitions. Conversely, the number of overtone and combination bands observable is very small. It is clear that the HA is a very good approximation and that the neglect of anharmonicity in predicting spectra is not a serious deficiency.

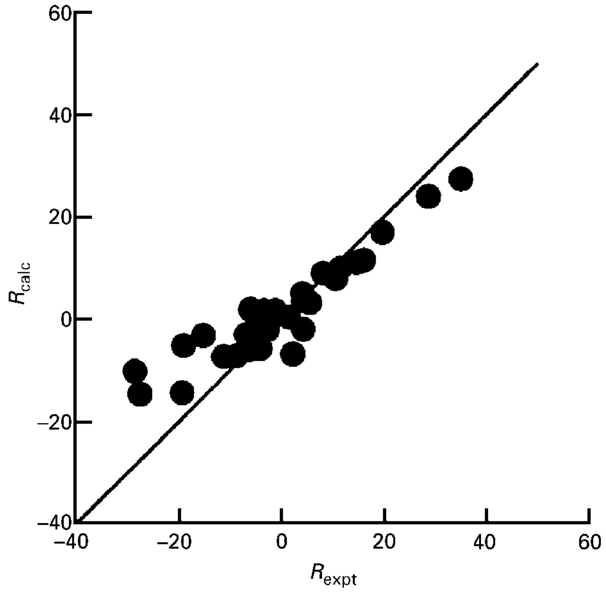



**Figure 2** Comparison of calculated and experimental rotational strengths for (1*R*, 4*R*)-(+)-camphor. *R* values are in $10^{-44}$ $esu^2$ $cm^2$. The straight line has a slope +1.

We turn now to the VCD spectrum. Each VCD band corresponds to an absorption band. The assignment of the experimental absorption spectrum can thus be transferred directly to the experimental VCD spectrum. Comparison of predicted and experimental VCD intensities, fundamental by fundamental, can then be carried out. As seen in **Figure 1**, the agreement is qualitatively excellent. Quantitative comparison of calculated and experimental rotational strengths, the latter obtained by Lorentzian analysis of the experimental VCD spectrum, is shown in **Figure 2**.

## Discussion

At the present time, calculation of absorption and VCD spectra within the harmonic approximation using DFT is computationally straightforward at the 6-31G* basis set level using GAUSSIAN 98 and standard parallel computers for organic molecules containing ≤ 200 atoms. For smaller molecules, larger basis sets can be used, yielding results of higher accuracy. For molecules with > 200 atoms one awaits further developments in computational algorithms and machine speed. The accuracy of DFT/6-31G* calculations using hybrid density functionals, illustrated in **Figure 1**, is already impressive. Further improvements in functionals, to be expected in the near future, will bring predicted spectra into even closer agreement with experiment.

There are two major deficiencies in the theoretical treatment of VCD described above: the neglect of anharmonicity and of solvent effects. Anharmonicity is relatively unimportant in the mid-IR spectral region, but becomes much more important at higher frequencies, e.g. in the C–H stretching region. Its inclusion for large molecules constitutes one of the remaining major theoretical challenges for the theory of VCD. Solvent effects are relatively unimportant in simple non-polar solvents such as $CCl_4$. However, they become much more important in solvents interacting much more strongly with solute molecules. Inclusion of solvent effects in such solvents, especially water, constitutes another major challenge for the theory of VCD.

We are here restricted to the theory of VCD. Its applications, past and future, are beyond the scope of our discussion. It is worth emphasizing, nevertheless, that the development of an accurate, computationally efficient methodology for the prediction of VCD spectra is of enormous consequence for the practical application of VCD spectroscopy to problems of stereochemistry in chiral molecules. In particular, the determination of absolute configuration in organic molecules containing ≤ 200 atoms is now practicable using VCD spectroscopy.

## List of symbols

$A$ = absorbance; $c$ = concentration (mol $L^{-1}$); $D$ = dipole strength; $-e$ = electronic charge; $f$ = normalized line shape function; $H_{el}$ = adiabatic electronic Hamiltonian; $k$ = force constant; $l$ = pathlength (cm); $M^{\lambda}_{\alpha\beta}$ = atomic axial tensor; $p_i$ = electronic momentum; $P^{\lambda}_{\alpha\beta}$ = atomic polar tensor; $P_{\lambda}$ = nuclear momentum; $Q$ = normal coordinate; $r_i$ = electronic position; $R$ = rotational strength; $R_{\lambda}$ = nuclear position; $W$ = energy; $Z_{\lambda}e$ = nuclear charge; $\alpha_g$ = fraction of molecules in state $g$; $\Delta A$ = circular dichroism; $\varepsilon$ = molar extinction coefficient; $\mu$ = dipole moment operator; $\nu$ = frequency; $\Psi$ = wavefunction.

*See also:* **Vibrational CD, Applications; Vibrational CD Spectrometers.**

## Further reading

Cheeseman JR, Frisch MJ, Devlin FJ and Stephens PJ (1996) *Ab initio* calculation of atomic axial tensors and vibrational rotational strengths using density functional theory. *Chemical Physics Letters* **252:** 211–220.

Devlin FJ, Stephens PJ, Cheeseman JR and Frisch MJ (1997) *Ab initio* prediction of vibrational absorption and circular dichroism spectra of chiral natural products using density functional theory: camphor and fenchone. *Journal of Physical Chemistry* **101:** 6322–6333.

Devlin FJ, Stephens PJ, Cheeseman JR and Frisch MJ (1997) *Ab initio* prediction of vibrational absorption and circular dichroism spectra of chiral natural products using density functional theory: α-pinene. *Journal of Physical Chemistry* **101:** 9912–9924.

Hehre WJ, Schleyer PR, Radom L and Pople JA (1986) *Ab initio Molecular Orbital Theory*. New York: Wiley.

Laird BB, Ross RB and Ziegler T (eds) (1996) *Chemical Applications of Density Functional Theory*, ACS Symposium Series 629. ACS.

Schellman JA (1975) Circular dichroism and optical rotation. *Chemical Reviews* **75:** 323–331.

Stephens PJ (1985) Theory of vibrational circular dichroism. *Journal of Physical Chemistry* **89**: 748–752.

Stephens PJ (1987) Gauge dependence of vibrational magnetic dipole transition moments and rotational strengths. *Journal of Physical Chemistry* **91**: 1712–1715.

Stephens PJ and Lowe MA (1985) Vibrational circular dichroism. *Annual Reviews of Physical Chemistry* **36:** 213–241.

Stephens PJ, Cheeseman JR, Frisch MJ, Ashvar CS and Devlin FJ (1996) *Ab initio* calculation of atomic axial tensors and vibrational rotational strengths using density functional theory. *Molecular Physics* **89:** 579–594.

Yamaguchi Y, Osamura Y, Goddard JD and Schaefer HF (1994) *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*. OUP.

# Vibrational, Rotational and Raman Spectroscopy, Historical Perspective

**AS Gilbert**, Beckenham, Kent, UK

Copyright © 1999 Academic Press

| VIBRATIONAL, ROTATIONAL & RAMAN SPECTROSCOPIES |
| --- |
| Historical Overview |

After a slow start in the nineteenth century, infrared (IR) spectroscopy saw a rapid increase in use and for a while, from 1945 onwards, was the most widespread method for determining the chemical structures of molecules. In the 1960s it became increasingly supplemented and overshadowed by other techniques, but recent developments in instrumentation have vastly improved its sensitivity and speed and enabled it to be applied to previously intractable problems. While research applications are many, there is overwhelming analytical and industrial usage.

The development of Raman spectroscopy has generally lagged behind that of IR owing to greater technical difficulties. The advent of the laser was the most important event in its history and has enabled many special and esoteric Raman experiments to be conceived. It is no longer confined mainly to the laboratory and is now used extensively in industry.

## Beginnings: IR and Raman spectroscopy before the Second World War

The astronomer William Herschel, better known as the discoverer of the planet Uranus, first detected IR radiation in 1800. Using a glass prism to refract the rays of the sun, he observed a rise in temperature in a thermometer positioned beyond the red limit of the visible spectrum (**Figure 1**).

Over the next one hundred years the essential nature of IR radiation was slowly established, electrical methods of measurement were developed and a number of materials such as rock salt were found to be largely transparent to it and thus useful as refracting elements. Interest was mainly directed to the physics of the subject, with little attention given to any possibilities for application to chemistry.

Near the end of the nineteenth century, however, a number of workers observed that many specific classes of organic compounds absorbed in characteristic regions of the IR spectrum. A monumental compilation of such bands, mostly from his own measurements, was published by W.W. Coblentz in 1905. One of his spectra is shown in **Figure 2** and can be compared with a modern-day version of the same compound in **Figure 3**. At the time, the precise mechanism for absorption was unclear, but it was soon realized that it was derived from what could be considered to be intramolecular vibrations. By the 1930s, a fairly complete theory encompassing the relation of dipole moment change to band intensity, selection rules, molecular symmetry and anharmonicity was available for both vibrational and rotational motions.

In 1928, C.V. Raman announced the discovery of the effect that now bears his name. In fact he had already observed it a few years before as a weak 'residual fluorescence' from highly purified organic liquids. It is generally considered, however, that the effect had been predicted by A. Smekal. Raman was awarded the Nobel Prize, the second Indian to be so honoured. Although the Raman effect was found to be very weak, the relation of selection rules to symmetry differed, so that results could complement those from IR absorbance.

### Instrumentation

Detection of IR radiation could be done by using bolometers or thermocouples, but point-by-point plotting of galvanometer readings made measurement of spectra somewhat tedious. IR-sensitive photographic film was available later and was occasionally used. Technical obstacles were many, however; for example the DC output from detectors could not be amplified and most galvanometers had long response times. Material for dispersion of IR radiation was scarce, rock salt of the quality and size suitable for prisms being difficult to obtain. Machines were single-beam and baselines were a major problem owing to the normal variations in laboratory heat background. Most instruments were built in-house, though commercial models were available, the first being introduced by Adam Hilger Ltd of the UK in 1913. It is illustrated in **Figure 4**.

The KBr disk method was not to be invented until 1952, so solids were usually sampled by mulling or reflection from whole crystals. The majority of samples studied were either liquids or gases.

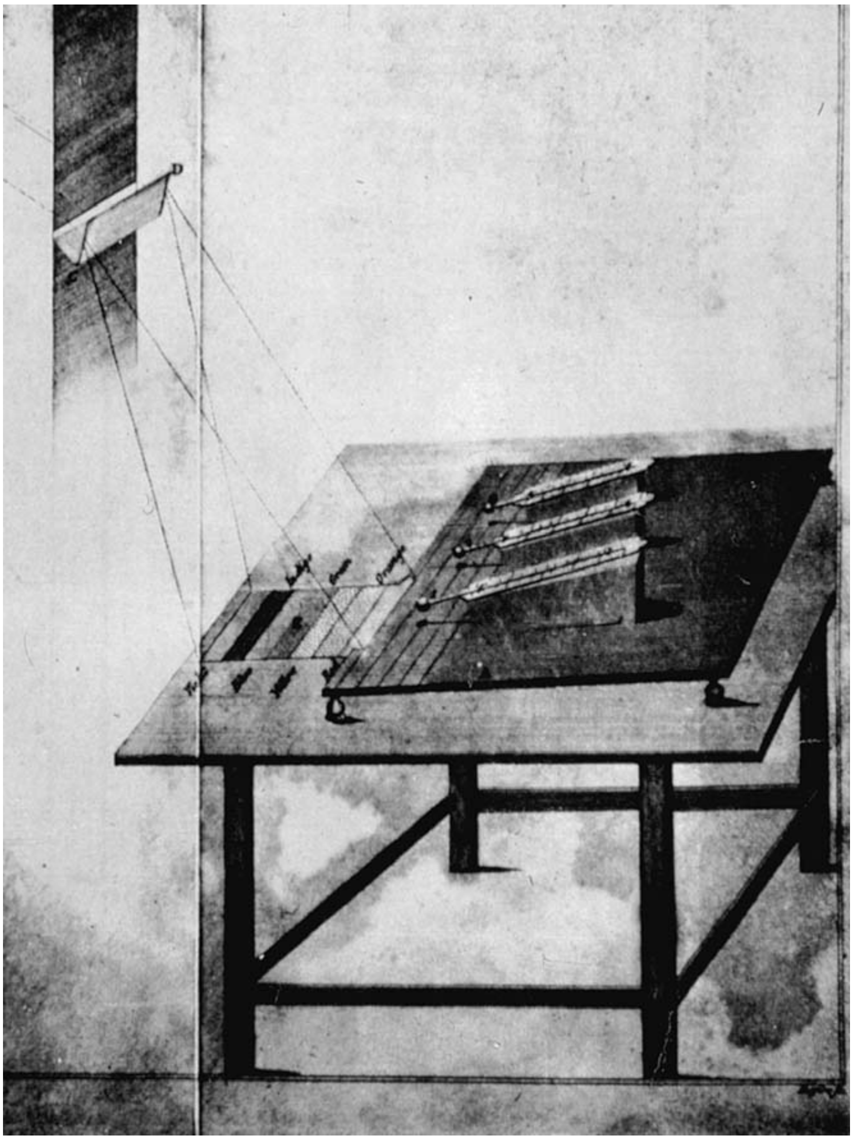


**Figure 1** Picture of William Herschel's experimental setup. Using blackened-bulb mercury thermometers, he observed a difference in temperature between one positioned at the red end of the visible and one positioned beyond. Originally from *Philosophical Translations*, Pt II **80**: 289 (1800). Photograph courtesy of Professor N. Sheppard, FRS.

By contrast, Raman spectra were usually easier to acquire so long as the sample was not coloured or turbid. Working in the visible region allowed simple spectrographs with silica/glass optics to be used and suitable photographic films were readily available, though long exposures were generally required. Discharge lamps, usually mercury vapour, provided a rather weak source of excitation radiation with a broad, diffuse background.

## Studies between the wars

The benefits of theory soon allowed vibrational spectroscopy to move away from simple characterization to deductions about the shape and symmetry of simple molecules. The observation that gaseous $CO_2$ yielded three fundamental modes of vibration, of which two were seen in the IR only and the other in the Raman only, immediately determined that it must be linear and symmetric. The tetrahedral structure of methane was confirmed by the lack of any observable pure rotational Raman spectrum. Such results gave valuable support to the developing theories of chemical bonding.

Knowledge of the actual frequency values enabled the calculation of the vibrational partition functions. Thermodynamic quantities of interest such as the heat capacity in the gas phase at different temperatures

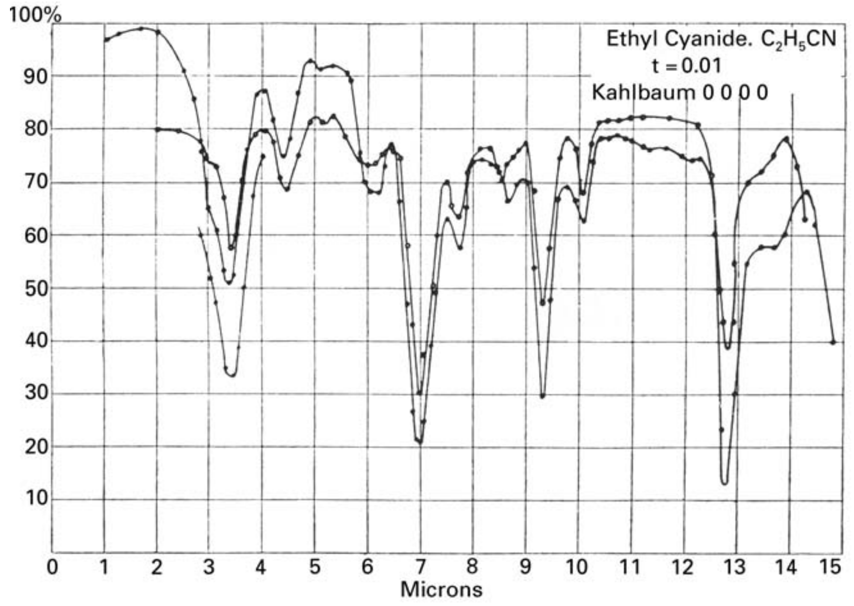



**Figure 2** Point-by-point IR spectrum of ethyl cyanide. From Coblentz WW (1905) *Investigations of Infrared Spectra*. Washington DC: Carnegie Institution of Washington. Photograph courtesy of Professor N. Sheppard, FRS.

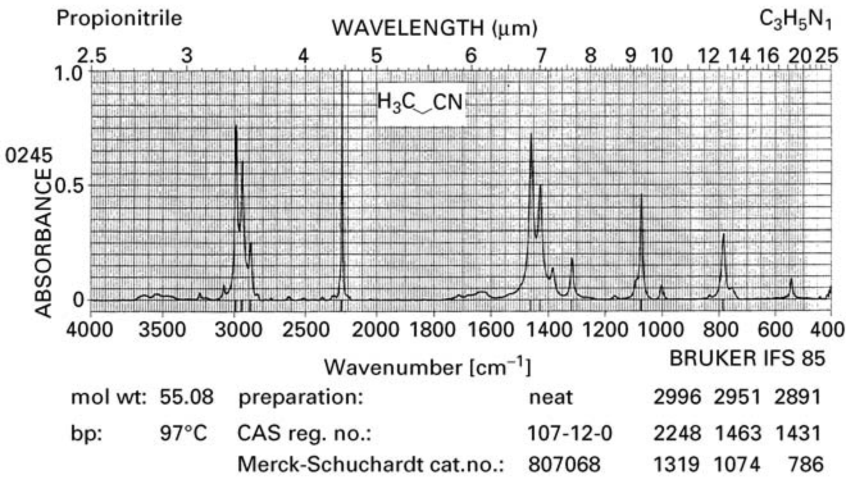



**Figure 3** An FT-IR spectrum of ethyl cyanide. Reproduced with permission from Pachler KGR, Matlok F and Gremlich H-U (1998) *Merck FT-IR Atlas*. New York: VCH.

could then be estimated with great precision. In addition, the force constants of very simple molecules could be determined.

IR spectroscopy gave a dramatic illustration of the existence of hydrogen bonding, a suspected new type of molecular attraction. For instance, at high temperatures in the gas phase, formic acid yielded the expected single fundamental from O–H stretching at about 3570 $cm^{-1}$. But at lower temperatures and in the liquid, this band disappeared to be replaced by a new one at near 3080 $cm^{-1}$.

While most attention was naturally given to the IR region below 4000 $cm^{-1}$, some measurements were made in the near IR (NIR), i.e. above 4000 $cm^{-1}$, often for no other reason than ease of experimentation. Overtones and combination bands were of course of interest in their own right but could also sometimes be used to make deductions about lower-frequency fundamental modes.

## First flowering: IR spectroscopy to the 1970s

Developments in electronics during the 1930s made measurements of IR spectra somewhat easier.

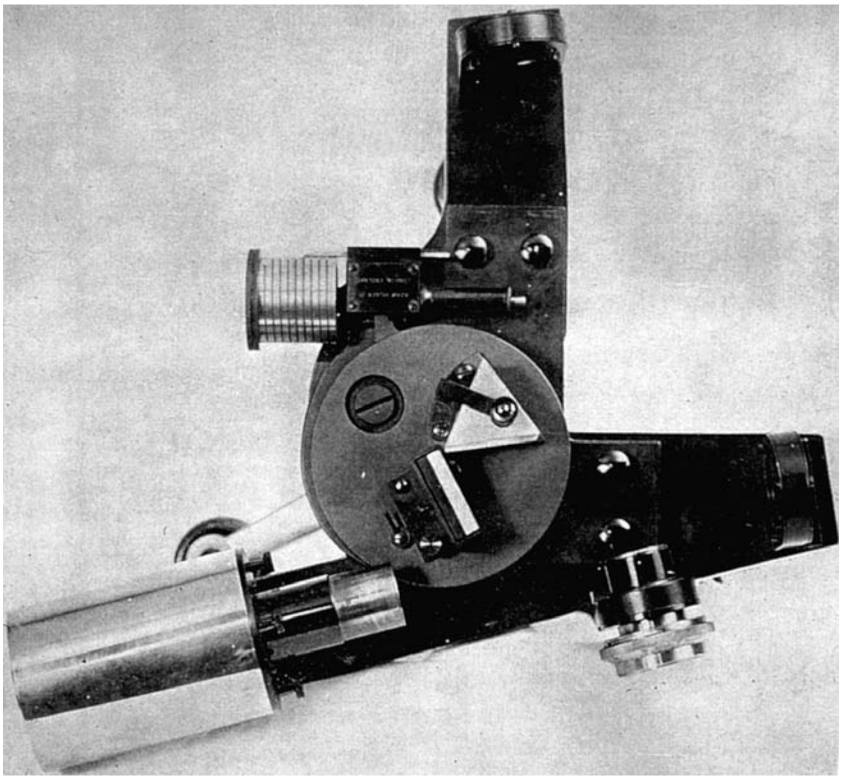


**Figure 4** Plan view of a Hilger IR spectrometer made in 1918. A constant-deviation Wadsworth optical arrangement with a 60° rock salt prism was employed. The detector was a thermopile and the source a Nernst filament. From *Hilger Journal*, August 1955. Photograph courtesy Professor N. Sheppard, FRS, and reproduced with permission from Hilger Analytical.

Successful application to a number of tasks connected with the Second World War, such as monitoring the composition of Axis petroleum samples to determine origin, led to a considerable expansion in activities.

Considerable effort was put into organic group frequency correlations, and compilations soon became available for general use. These were utilized extensively for detailed structural characterization of organic compounds. The boon to organic chemists can be appreciated by considering the methods that had been available to them before. Elemental analysis could only give overall composition and ultraviolet–visible (UV-visible) spectroscopy little more than some idea of the degree or type of unsaturation. A limited number of chemical group types were detectable by chemical spot tests.

Moreover, IR spectroscopy was also applied to other work such as quantitative analysis and the study of physical interactions between molecules. The general rise of scientific activities in both universities and industry naturally led to a rise in demand for IR facilities and the 1940s saw the first commercial automatically recording spectrometers. Spectra were now produced routinely as continuous plots. Single-beam instruments were rapidly superseded by double-beam ones and gratings came into general use.

### Dispersive double-beam spectrometers

In double-beam instruments compensation was, with one or two exceptions, by the optical null method whereby a servomotor drove a comb into the reference beam to minimize the signal difference between it and the sample beam as it was attenuated to varying degrees by absorbance. Movement of the comb was mechanically geared to a pen, which thus recorded the spectrum as a trace on a moving chart.

Sample and reference beams were alternately presented to the detector by means of mirrors and a chopper rotating at a frequency of several hertz, thus allowing selective amplification of the signal difference, which was fed to the servomotor. Thermal detectors continued to be used for many years but with developments, such as the pneumatic Golay cell, for greater sensitivity.

Good-quality large alkali halide crystals could now be grown on an industrial scale, so that fabrication of prisms and other optical components did not present any difficulties. In particular, use of synthetic KBr allowed spectra to be measured down to 400 $cm^{-1}$; pre-war natural NaCl (rock salt) prism spectrometers had a lower limit of about 650 $cm^{-1}$ owing to the higher frequency absorption edge of the salt.

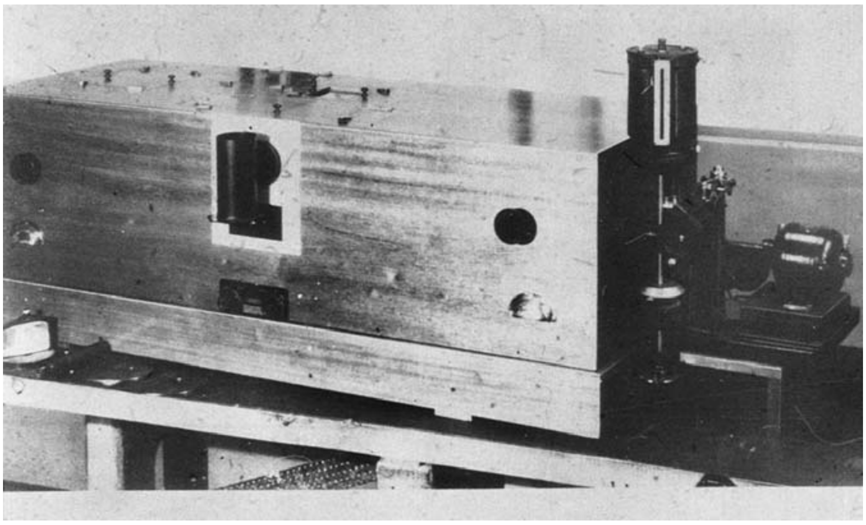


**Figure 5** An early automatically recording IR spectrometer, the Hilger model D209 introduced in 1940. The spot from a mirror galvanometer was focused onto photographic film fixed to a rotating drum. Later during the Second World War, this machine was developed into the first commercial double-beam (ratio recording) IR spectrometer. Photograph courtesy of Professor N. Sheppard, FRS, and reproduced by permission of Hilger Analytical.

Rotation of the prism or grating was controlled by a cam to allow the spectral trace to be recorded in either constant wavelength or wavenumber. This cam was coupled to the slits to vary their width in order to keep the radiation reaching the detector roughly constant. Resolution therefore varied over the spectrum. Positional accuracy and repeatability were poor owing to backlash and wear in the mechanical linkages; most specifications quoted values around ±0.5 $cm^{-1}$.

The inherent disadvantage of the optical null mechanism was the attenuation of the reference beam, so that accuracy and response time became worse as sample transmittance (%$T$) decreased. Linearity of response depended on how well the comb had been machined. Theoretically it was impossible for the equipment to measure 0 %$T$. Scanning had to be slow to avoid excessive pen lag. Reliable quantitative analysis was difficult though possible with heed.

The amounts of sample required for good spectra of a solid was generally of the order of a milligram. By utilizing beam condensers and the best instruments, adequate spectra could be obtained from small samples of the order of a hundred nanograms.

## Commercial instruments

Several companies built spectrometers; based in the USA were Perkin-Elmer, Beckman, Baird and Cary; in the UK, Grubb-Parsons, Hilger (**Figure 5**) and Unicam (later Pye-Unicam). Instruments were also built in France, Germany (both West and East), Japan and the USSR. The commonest optical layout used was the Littrow arrangement which had the virtue of compactness. Low-cost instruments were introduced in the 1950s, some only with prisms though all were soon sold with gratings instead. Their relative cheapness and ease of use were major factors in the rapid expansion in the practice of IR spectroscopy. Research-grade models usually had a grating monochromator with a fore-prism. Two or more gratings with auto changeover were required to cover the whole range.

During the latter part of the period, prisms disappeared and more advanced features were built into dispersive spectrometers such as computer interfaces, principally for data acquisition only. Ratio recording became standard in place of the optical null. However, the Perkin-Elmer model 983, introduced in 1982, demonstrated the benefits that computerization could bring to instrument control. Grating rotation was carried out directly by a stepper motor under the command of a microprocessor. With no mechanical cam, position repeatability (±0.005 $cm^{-1}$) was, in consequence, almost as good as that of FT-IR instruments.

## Applications

A pre-eminent use already described was structure determination of organic compounds. The KBr disk technique allowed complete and unmarred spectra of solids to 400 $cm^{-1}$. Extensive collections of spectra (both literature and commercial) of organics, inorganics, pharmaceuticals and industrial chemicals now appeared.

A result of great topical interest in the late 1940s was the recognition that penicillin contained a fused four-membered lactam ring on the basis of several bands, including one at an unusually high value, near 1780 $cm^{-1}$, from the amidic carbonyl stretching mode.

Because of the distinctive fingerprint nature of the spectra of individual substances, a very common application was, and continues to be, compound verification in forensic identification and quality assurance to meet pharmacopoeial and other standards set by regulatory authorities.

The sensitivity of IR bands to changes in physical state led to numerous studies concerned with the thermodynamics of hydrogen bonding in solvents and the influence of solvent effects (donor/acceptor capacities). Phase transitions (polymorphism) in solids could easily be detected by changes in band intensity and position.

Inorganic and metal–organic compounds also received attention in structural studies. For instance, IR spectroscopy of metal carbonyl complexes could easily distinguish between terminal and bridging ligands.

Access to computing (mainframe) facilities from the 1950s onwards provided the means to analyse the fundamental mode frequencies of a polyatomic molecular structure on the basis of harmonic oscillation and obtain force constants. These could be transferred to other structures and the procedure reversed so as to estimate frequencies for cases where a spectrum was difficult to interpret. While fraught with difficulties, mainly because there were usually many more force constants to estimate than frequencies, such work did at least highlight the fact that many vibrational modes are strongly coupled to each other and that few bands can normally be assigned to specific bond motions.

Data handling was never more than very primitive. Although the largest commercial collections of spectra offered searching facilities, for example by matching of most prominent bands, these were not very successful in practice even for pure compounds.

The rapid spread of mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy from the 1960s caused IR spectroscopy to be sidelined in many areas of analysis. NMR was much better at elucidating many fine details of molecular structure, while detection limits were lower using MS.

## Renaissance: The (Fourier) transformation of IR spectroscopy

IR interferometers seem first to have been tried out for examining the very low signal levels of IR emission from astronomical sources. They were shortly also employed for far-IR spectroscopy where low source output made measurement extremely difficult using conventional spectrometers below 200 $cm^{-1}$.

There was sufficient interest for several far-IR instruments to be made available commercially from the late 1950s, but one drawback was the difficulty of numerically converting the interferograms to recognizable spectra. With only mainframe computers available, instrument-based digital computing was not feasible. RIIC/Beckman marketed a machine that stored data, after analogue-to-digital conversion, on magnetic core memory. The data were then reconverted back to analogue to be fed into an electronic wave analyser to generate the spectrum.

Ironically, these interferometers could not be used to advantage in the mid-IR region (4000–400 $cm^{-1}$) in most circumstances. This was because, with scan speeds in minutes, signal levels were so large that the noise was well below the threshold of even the smallest bit of the best (16-bit) ADCs obtainable. However, by the early 1960s, Block Engineering of the USA had developed a fast scanning interferometer. Providentially, lasers (for accurate path difference referencing), minicomputers and the Cooley–Tukey fast Fourier transform (FFT) algorithm soon all appeared. A Block subsidiary, Digilab, exploited these devices to produce the first commercial mid-IR Fourier transform (FT-IR) spectrometer, the model FTS-14, in 1969.

They were followed during the 1970s by other companies, most of whom had had no previous experience in the IR field but made computers and data loggers, such as Nicolet and Bruker. The machines offered were principally research grade, though Digilab made an early and substantial entry into the specialist market for quality control of semiconductors.

Most manufacturers of dispersive instruments opted not to enter the new arena, presumably because of the costs involved in investment in the new technology, and in consequence left the business altogether as demand for dispersive instruments dried up. A major exception was Perkin-Elmer who, although starting late, had the resources to participate in a big way. Although offering research-grade instruments as well, they particularly targeted the analytical market, in which they were soon to be a major supplier.

At least twenty manufacturers worldwide have offered FT-IR spectrometers in recent years, several of them for niche markets such as dedicated gas analysis and space science.

### Advantages of interferometry in the mid IR

Two factors were judged to be of major importance: The Fellgett (multiplex) and Jacquinot (throughput)

advantages. The Jacquinot advantage, though large, and foremost at higher wavenumbers, was found, however, to be largely cancelled out by the much decreased performance (at the high modulation frequencies imposed by fast scanning interferometers) of the early pyroelectric detectors used, both factors having a roughly similar wavenumber dependency. Traditional thermal effect detectors such as thermocouples were far too slow.

As other factors (e.g. beam splitter versus grating efficiency) were of lesser importance, this meant that overall advantage in signal to noise (S/N) ratio of mid-IR FT instruments over dispersives was basically determined by multiplexing only. Strict comparisons were not possible, though, because FT instruments ran at constant resolution unlike dispersives where resolution varied somewhat across the spectrum.

Thus for spectra run at moderate resolution and over limited spectral range, FT machines with pyroelectric detectors were not strikingly better than dispersives. This was because the multiplex advantage (for FTIR) is no better than $\frac{1}{2}\sqrt{M}$, where $M$ is the number of resolution elements.

However, for low signal levels a quantum detection, MCT (mercury cadmium telluride), was soon developed, that had much greater sensitivity at high modulation rates and thus did not trade off the Jacquinot advantage. This allowed FT spectrometers to tackle situations that dispersives could not even attempt and to go some way in competing with and complementing MS in trace analysis. Various advances in component design (e.g. sources) saw a considerable improvement in general performance during the 1970s.

The Connes advantage claimed for interferometers was that laser referencing would yield much greater wavelength accuracy and reproducibility. The latter was crucial for both signal averaging and spectral subtraction but, as the Perkin-Elmer company had shown, this could almost be matched by dispersives.

Photometric accuracy and reproducibility were actually worse than with the best ratio-recording grating instruments, owing mainly to sample reflection and emission back into the interferometer.

Fashion probably played a significant role in the cessation of dispersive spectrometer manufacture as the high S/N provided by FT machines was not required when sample was abundant and sample preparation took much longer than measurement.

## Applications

Interferometers made far-IR spectroscopy possible, and enabled observation of lattice vibrations of crystals, low-frequency skeletal motions of organics and the stretching modes of heavy atoms. In the mid-IR region interest was not so much in new vibrations as in old ones in new and difficult situations, now accessible because of the high S/N available. Multitudes of hitherto indifferent samples suffering from low transmission were now directly amenable without resort to complex pre-preparation. Because data were obtained and stored in digital form, spectra could readily be manipulated and spectral subtraction to remove solvent bands from solution spectra, assess purity and reveal impurities became a popular activity.

High S/N, and in addition fast scanning, enabled events as rapid as fractions of seconds to be monitored. Thus FT-IR spectroscopy was applied with considerable success as a detection method for gas chromatography (GC) and manufacturers were obliged to offer suitable devices. However, optimizing the interface took a surprisingly long time; waiting on GC column technology was partly to blame. Some 20 years elapsed from the early 1970s before minimum identifiable limits, starting in micrograms, reached the mid-picogram level with the advent of cryogenic trapping methodology.

While many biological/biochemical samples had been studied by dispersive instruments, the difficulties of dealing with aqueous solutions was a severe limitation. Now, however, this general field became a major area of research attention for application of FT-IR in the 1980s. Of interest was the sensitivity of vibrational modes to subtle changes in molecular environment of the type crucial to biological mechanism and structure studies. Particular topics were how cell membrane conformations were affected by interactions with other chemicals or general physicochemical effects, and analyses of protein secondary structure.

## Instrumentation

The GC–IR interface was the first of many specialized accessories that began to appear in increasing numbers during the 1980s. They included interfaces to various forms of chromatography (the combination being known as a hyphenated method), photoacoustic and diffuse reflectance cells and microscopes. Spectrometer design was influenced so as to avoid the inconvenience of changeover. Interferometers were therefore built with multiple beam ports to allow two or more accessories to be permanently connected. With such developments and the continuing miniaturization of computing hardware, the large floor-standing FT spectrometers of the 1970s were superseded by the modular benchtop machines of the 1980s.

The most significant development of the interferometer in recent years has been the step scanning

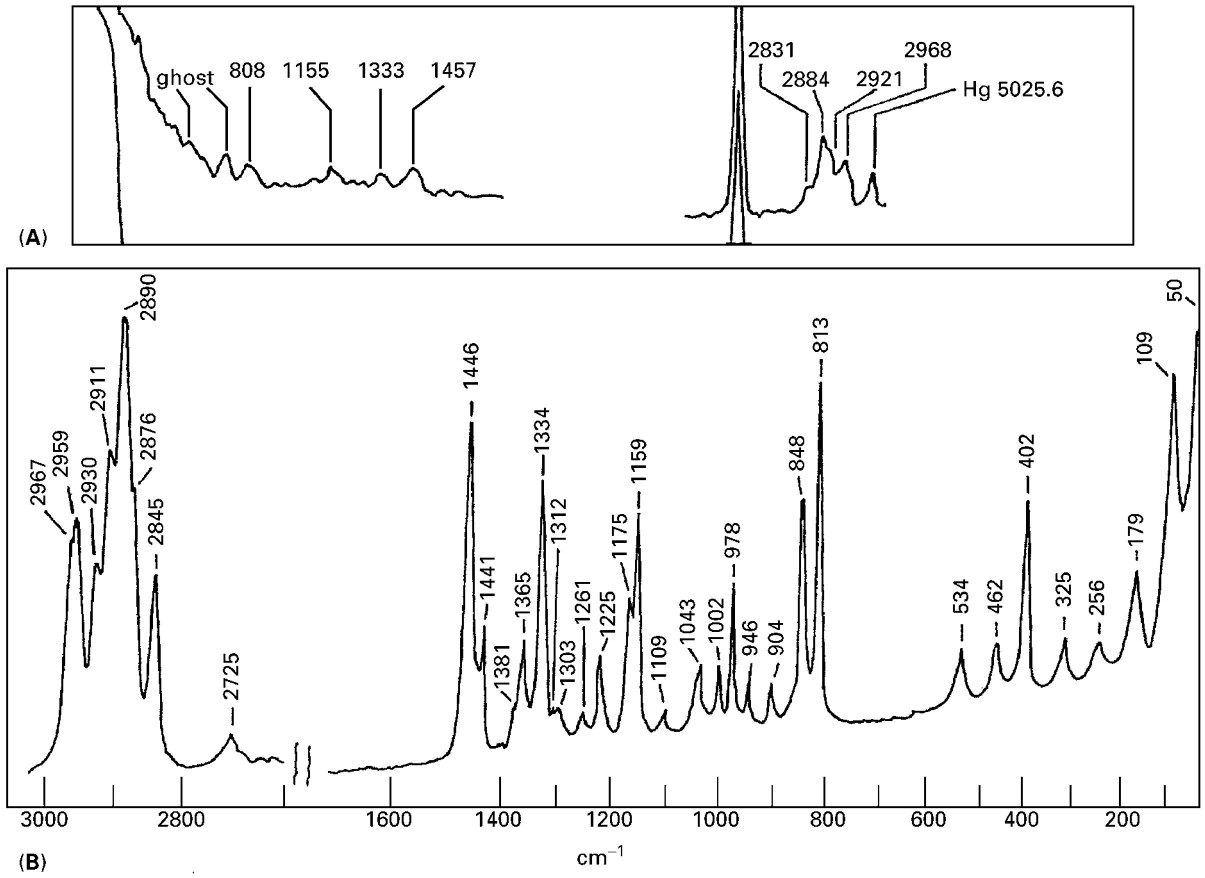



**Figure 6** Comparison of the Raman spectra of isotactic polypropylene. (A) Mercury arc, 435.8 nm; densitometer trace of a photographic negative. (B) Laser, 632.8 nm; photoelectric recording. Reproduced by permission from Tobin MC (1971) *Laser Raman Spectroscopy.* New York: Wiley.

modification. This has allowed, among other things, the monitoring of fast processes on the timescale of single data point acquisition so that, for example, vibrational spectra of some excited states can be obtained.

Finally, the interferometer was not to be all-conquering, however, as simple filter spectrometers were found to be useful for many dedicated monitoring purposes. Tuneable IR lasers have also been applied to gas monitoring, but their general usefulness has been limited by the range of wavelengths that can be output.

### The coming of the black box

The necessary provision of computers in FT-IR soon led to much more than simple data manipulation. After much unfulfilled promise and wasted effort, data transfer systems became fairly standardized and reliable by the late 1980s. With the universal adoption of the JCAMP format, spectra in digital form became truly portable and could become part of the business of laboratory information management systems (LIMS). More interestingly, they could be directly matched into large digital spectral databases for automatic identification or structural analysis by expert systems.

In parallel, the 1980s also saw considerable development in the application of software for data analysis and instrument control. By this time, software was accounting for more than half the development cost of an FT-IR system.

## Acronyms galore: The impact of the laser on Raman spectroscopy

Unlike the case for IR spectroscopy, there was little development in the immediate post-war years, so that for a long time practice was mainly confined to a few academic institutions.

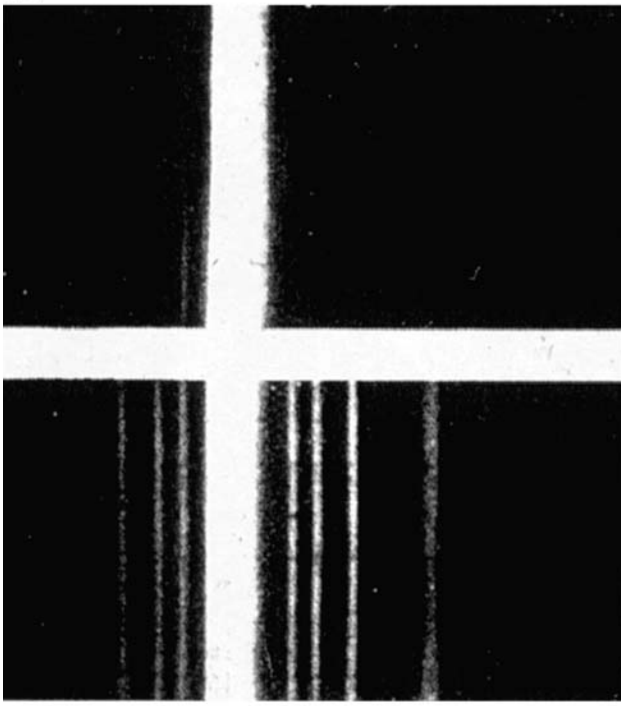
An improved source, the mercury arc, a helical discharge lamp surrounding a cylindrical sample tube, yielded much-increased light energy, though probably equal in importance was the advent of photoelectric recording. The latter facility enabled the production of scanning instruments. Two companies, Cary and Hilger+Watts, offered models, the former fielding an image slicing device that

attempted to overcome the optical incompatibility of the extended area of the mercury arc source and the narrow spectrometer slits.

Apart from the weakness of the spectra, the range of materials that could be examined was restricted by problems thrown up by the sample itself. Coloured samples absorbed source and scattered signal and possessed the propensity to fluoresce, which could easily drown out the Raman signal.

The advantages offered by the laser to Raman spectroscopy were recognized almost as soon as it had been demonstrated in 1960. It was to revolutionize the technique by providing dramatic increases in sensitivity (**Figure 6**) and the opportunity for unusual experiments by virtue of nonlinear effects that it could induce in many materials.

(A)



(B)

**Figure 7** Mercury arc-excited Raman spectrum of carbon tetrachloride with photographic recording. (A) The spectrum of the mercury arc itself for reference. (B) The four Stokes lines (right side of the exciting line) and the weaker antiStokes (left side of the exciting line) lines, of which only three can be seen. From part of Plate 1, Raman CV and Krishnan KS (1929) The production of new radiations by light scattering. *Proceedings of the Royal Society (London)* **A122**: 23–35. Reproduced from a photograph courtesy of Professor N. Sheppard, FRS, and with permission of the Royal Society.

The continuous gas lasers were found to be most useful. The principal advantage of the laser in general over the mercury arc lay in the small point-source area that allowed larger flux throughput for a given spectrometer étendue.

Where formerly grams of material were needed, now milligram quantities, as solid powders, liquid or solutions in capillary tubes, were routinely amenable to examination.

The beckoning opportunities soon led several manufacturers (e.g. Spex, Jarrel-Ash, Coderg and Perkin-Elmer) to bring out new equipment, while older designs (Cary) were modified to accommodate the new type of source. Instruments were of course expensive compared to dispersive IR spectrometers as the optical requirements were far more stringent. Two coupled monochromators were necessary to cut down stray light from the exciting line. The Czerny–Turner layout was most common, the monochromators usually, but not always, being arranged for additive dispersion. **Figure 7** shows a very early photographic/mercury arc Raman spectrum of $CCl_4$ for comparison with what could be achieved with an automatically recording spectrometer of the late 1960s (**Figure 8**).

Photoelectric detection for Raman spectroscopy initially suffered from poor sensitivity in the red, but this was rectified during the 1980s when many cheap photomultiplier tubes became available. Rather more expensive were multichannel detectors, originally used for fast time-resolved experiments in the 1970s. Their take-up was slow but by the late 1980s very sensitive charge-coupled devices (CCDs) developed for astronomy were being incorporated in spectrographs for relatively mundane analytical work.

### Complementing infrared

Compared to IR spectroscopy, the Raman technique was found to possess some advantages, one particular being the comparative ease in dealing with aqueous solutions. In the biochemical sphere this meant, for example, that ionization behaviour and pH change could be studied; even before the arrival of the laser, amino acids had been demonstrated to exist as zwitterions.

For a long time a major disadvantage was the common occurrence of interference from fluorescence, especially often encountered in biological material. A redeeming feature of Raman spectroscopy, however, was the many special experiments that could be performed.

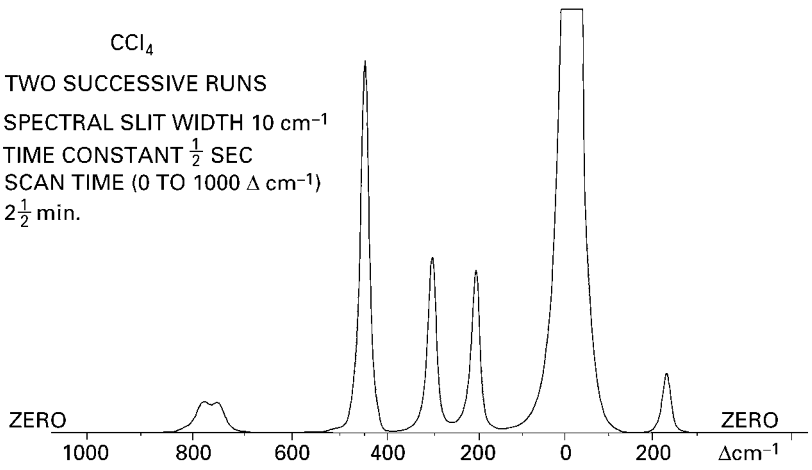



**Figure 8** Laser-excited Raman spectrum of carbon tetrachloride with photoelectric recording. Reproduced with the permission of Professor N. Sheppard, FRS.

#### Nonlinear effects and other esoterica

The high power densities available from lasers were found to be capable of inducing a number of strange (nonlinear) Raman effects, though many were of limited applicability to problems in chemistry. Typical examples included SIRS (stimulated inverse Raman scattering), RIKES (Raman induced Kerr effect spectroscopy) and the hyper Raman effect.

Probably the most useful was found to be CARS (coherent antiStokes Raman scattering), in which the signal could be detected largely free of background such as fluorescence. It was particularly beneficial in examining combustion processes in flames. Hopes that it might provide the complete answer to the fluorescence problem were to be largely unfulfilled owing mainly to its technical complexity.

Other interesting and useful discoveries did not require high-power excitation. Resonance Raman scattering (RRS) arises in certain samples when the laser excitation wavelength falls within an electronic absorbance band. RRS bands can be several orders of magnitude stronger than normal. One application was to employ small-molecule ligands exhibiting RRS as probes in the active sites of enzymes, thus avoiding major interference from the complex spectrum of the protein.

In 1974 news came that the Raman spectrum of pyridine was considerably enhanced when absorbed on a roughened silver substrate. This effect was named SERS (surface enhanced Raman scattering) and has since been observed from many other compounds and other metals. Electrochemistry was an early application. Inevitably, experimenters were to combine the effect with RRS to create SERRS, which in some circumstances can equal fluorescence for sensitivity of detection.

#### Instrumental developments

FT-Raman spectroscopy was introduced in 1986 and it is now available as a bolt-on to many FT-IR machines. Interestingly, interferometers might have been used earlier for Raman spectroscopy if the laser had not been invented, as their large circular aperture could have coped advantageously with the extended source area of the mercury arc. As it was, the multiplexing capability was needed to boost sensitivity so as to satisfactorily observe the weak spectra produced by a near-IR laser. The rationale was that fluorescence was largely eliminated. Thus, high-quality spectra of dyes, for instance, could be obtained that had formerly been impossible. Unfortunately, as overtones and combinations of $H_2O$ vibrations possess significant absorbance in the near IR, spectra from aqueous solutions were affected.

An earlier development was the reintroduction of the spectrograph made possible by the availability of multichannel detectors. CCDs, with sensitivity equal to the traditional photomultiplier and up to 1024 channels, thus provided a considerable multiplexing advantage. They allowed the construction of small, rugged instruments able to acquire good-quality spectra very rapidly. Coupled to fibreoptic sampling devices, such spectrographs have recently found considerable use for process monitoring in industry.

A very useful accessory has been the microscope. Here there is a significant advantage over IR spectroscopy as spatial resolution is higher owing to the shorter wavelength of the source radiation.

## Out of the orphanage: The rise of near-IR spectroscopy

The NIR region was for long neglected, a curiosity for users of many research-grade UV-visible and mid-IR spectrometers that had been provided with extended range capabilities. This was not unexpected, as absorption bands in the region originate from either uncommon electronic transitions in inorganic compounds or broad and heavily overlapped overtones and combinations of vibrational fundamentals. The latter are mostly derived from X–H stretching modes in organic compounds. In consequence, spectra, particularly of mixtures, are not easy to interpret.

Experimentally, however, the spectra are easy to observe, thick samples being tractable in either transmission or reflection without preparation. As the spectra seldom possess many narrow features that could be unduly affected by instrumental or other factors, robust methods for quantitation were possible. With the arrival of cheap instrumental computing from the 1980s onwards and the development of multivariate analysis methodology, NIR spectroscopy has undergone considerable expansion in use. It is now widely applied to automated, rapid and precise quantitative analyses in agriculture, industrial process control and noninvasive medical examinations. A typical and early example was determination of the protein content of grain and flour.

Most work is now done with dispersive and simple filter analysers. Though S/N is not usually a problem, FT-NIR and filter/CCD machines are becoming more popular.

## Full circle: IR and Raman spectroscopy into the new millennium

Originally IR spectra were measured point by point; they then became continuous and now are again digital in nature. Spectrometers were single-beam, then double-beam, and now are almost all single-beam once more. With the introduction of spectrographs coupled to multichannel detectors, Raman spectroscopy is also in a sense back where it used to be, when photographic film effectively provided a multiplicity of channels. NIR CCDs are presently available that cover some of the range and the future possibilities of usable array detection right down into the mid IR could eventually spell a general return to dispersive techniques. Further into the future, fully tuneable IR lasers would remove even the need for a dispersing element for IR spectroscopy. If the rationale for interferometers, hitherto their advantage in throughput, is lost, then their disadvantages such as poor photometric accuracy and mechanical complexity could mean retention only in special circumstances.

Hand in hand with increasing sophistication of spectrometers and software, much of which runs transparently to the user, has gone an overall deskilling of operatives. This is inevitable (and usually desirable) given the workload and demands on modern analytical laboratories and industrial processes. There are obvious dangers, however, as the theory behind the methodology is beyond many scientific workers, as are the ramifications of many spectrometer function operations on the input data. The days are long gone when spectroscopists built and maintained their own machines, synthesized the chemicals they studied and contributed to advance of theory.

*See also:* **Chromatography–IR, Applications; Fourier Transformation and Sampling Theory; FT-Raman Spectroscopy, Applications; Hydrogen Bonding and other Physicochemical Interactions Studied By IR and Raman Spectroscopy; IR Spectral Group Frequencies of Organic Compounds; Industrial Applications of IR and Raman Spectroscopy; IR Spectrometers; IR and Raman Spectroscopy of Inorganic, Coordination and Organometallic Compounds; Near-IR Spectrometers; Nonlinear Raman Spectroscopy, Applications; Raman Spectrometers; Raman and IR Microspectroscopy.**

## Further reading

Bellamy LJ (1975) *The Infra-red Spectra of Complex Molecules*. London: Chapman and Hall.

Brugel W (1961) *An Introduction to Infrared Spectroscopy*. London: Methuen.

Ferraro JR (1996) A history of Raman spectroscopy. *Spectroscopy* **11**(3): 18–25.

Griffiths PR, Sloane HJ and Hannah RW (1977) Interferometers vs monochromators: separating the optical and digital advantages. *Applied Spectroscopy* **31**(6): 485–495.

Herzberg G (1945) *Infra-red and Raman Spectra of Polyatomic Molecules*. New York: Van Nostrand.

Hibben JH (1939) *The Raman Effect and Its Chemical Applications*. New York: Reinhold.

Johnston SF (1991) *FTIR: A Constantly Evolving Technology*. Chichester: Ellis and Horwood.

Jones RN (1985) Analytical applications of vibrational spectroscopy – a historical review. In Durig JR (ed) *Chemical, Biological and Industrial Applications of Infrared Spectroscopy*. Chichester: Wiley.

Long DA (1988) Early history of the Raman effect. *International Reviews in Physical Chemistry* **7**(**4**): 317–349.

White RG (1964) *Handbook of Industrial Infrared Analysis*. New York: Plenum.

# X

# Xenon NMR Spectroscopy

**Jukka Jokisaari**, University of Oulu, Finland

Copyright © 1999 Academic Press

MAGNETIC RESONANCE
Applications

Nuclear magnetic resonance experiments were originally planned for measuring nuclear magnetic moments. This explains why such exotic nuclei as $^{129}$Xe (spin-$\frac{1}{2}$) and $^{131}$Xe (spin-$\frac{3}{2}$) were the subjects of NMR investigations already as early as 1951, only about 5 years after the first successful NMR experiments at Harvard and Stanford. In the 1960s and 1970s, researchers were mostly interested in performing shielding and relaxation experiments of the two xenon isotopes (as well as of the other noble gas nuclei: $^3$He, $^{21}$Ne and $^{83}$Kr) in gas, liquid and solid states. The expansion of xenon NMR spectroscopy started in the early 1980s when it was proposed that the physical properties of zeolites and clathrates, and in general, of porous solids, could be characterized by adsorbing xenon to the sample and recording the $^{129}$Xe NMR spectrum. Since then, xenon NMR has been applied to derive information on (besides porous solids) various isotropic liquids and liquid mixtures, liquid crystals, proteins and membranes, myoglobin and haemoglobin, and polymers. The finding that $^{129}$Xe can be spin-polarized by optical pumping, leading to a sensitivity increase up to a factor of $10^5$, has widened the field of application to low surface area solids, human blood and perfused tissue, and imaging of organs of small animals, and recently, even of humans.

## Properties of the nuclei

Xenon possesses nine stable isotopes, but only two of them, $^{129}$Xe and $^{131}$Xe, have a nonzero spin necessary for magnetic resonance. NMR properties of the isotopes are collected in **Table 1**. Xenon being a monatomic gas, the $^{129}$Xe and $^{131}$Xe NMR spectra recorded in isotropic solutions, for example, consist of a single resonance line. In anisotropic environments with nonzero static electric field gradient (EFG), the $^{131}$Xe NMR spectrum is a triplet. Despite this simplicity, the spectra provide much information on the environment into which the xenon is introduced. This is due to the fact that the xenon shielding is extremely sensitive to the changes taking place in its local environment. Furthermore, as xenon is very inert it does not disturb the system into which it is taken to. In the literature, the change of the shielding of atomic xenon, arising from bulk and local effects, is often called the chemical shift. This is somewhat misleading. However, in this article chemical shift ($\delta = \sigma_0 - \sigma_{Xe}$, $\sigma_0$ and $\sigma_{Xe}$ being the shielding of the reference gas, usually free xenon gas, and of xenon introduced into the environment under study, respectively) and shielding ($\sigma_{Xe} - \sigma_0 = -\delta$) are used in parallel.

The 32-fold sensitivity of $^{129}$Xe compared to $^{13}$C makes it an easy nucleus for NMR detection; the

**Table 1** Properties of the NMR active xenon isotopes $^{129}$Xe and $^{131}$Xe

| *Isotope* | *Spin* | *Natural abundance* (%) | *Gyromagnetic ratio* ($10^7$ rad T$^{-1}$ s$^{-1}$) | *Quadrupole moment* ($10^{-28}$ m$^2$) | *NMR frequency*$^a$ (MHz) | *NMR sensitivity*$^b$ |
|---|---|---|---|---|---|---|
| $^{129}$Xe | $\frac{1}{2}$ | 26.44 | −7.441 | – | 110.632 | 31.82 |
| $^{131}$Xe | $\frac{3}{2}$ | 21.18 | 2.206 | −0.12 | 32.795 | 3.318 |

$^a$ At the magnetic field $B_0$ = 9.39 T.

$^b$ Absolute sensitivity (product of the relative sensitivity and natural abundance) with respect to that of the $^{13}$C isotope.

spectrum may be obtained on a single pulse from samples with ~1 atm or higher pressure of gas. This is a great advantage because often the spin–lattice relaxation time, $T_1$ is long, up to hundreds of seconds in solutions and up to over 100 min in the gas phase. In many applications, relaxation agents can be utilized to shorten $T_1$. The $^{131}$Xe isotope possesses an electric quadrupole moment and thus its spin–lattice relaxation is dominated by the quadrupole interaction, the $T_1$ values are at the millisecond level allowing a relatively fast pulse repetition in accumulation.

A drastic improvement in the NMR sensitivity of $^{129}$Xe can be achieved by the optical pumping method. In a conventional NMR experiment, the nuclear magnetization arises from the population difference due to Boltzmann distribution between energy states. For example, for $^{129}$Xe at the magnetic field of 9.4 T and at thermal equilibrium at $T = 300$ K, the relative population difference between the $m = -\frac{1}{2}$ and $m = +\frac{1}{2}$ states is $9 \times 10^{-6}$ = 9 ppm. When applying optical pumping with a high-powered laser, this figure can be increased up to >0.3, i.e. the magnetization increases by a factor of $\sim 3 \times 10^5$. This increase is performed by placing xenon and nitrogen gas together with alkali-metal vapour (usually rubidium) in a glass cell and illuminating by circularly polarized laser light with a wavelength of 749.7 pm. During binary collisions, spin polarization is transferred from alkali-metal atoms to xenon atoms.

## Xenon in gases

The shielding of pure xenon gas is often expressed as a virial expansion

$$\sigma(\rho, T) = \sigma_0 + \sigma_1(T)\rho + \sigma_2(T)\rho^2 + \sigma_3(T)\rho^3 \qquad [1]$$

where $\sigma_0$ is the shielding constant of the atom *in vacuo*, $\rho$ is density (given in amagat, the density of xenon under standard conditions at 298 K, $\sim 2.5 \times 10^{19}$ atoms cm$^{-3}$), and for the virial coefficients the following values have been reported (at 298 K): $\sigma_1 = -0.548 \pm 0.004$ ppm amagat$^{-1}$, $\sigma_2 = (-0.17 \pm 0.02) \times 10^{-3}$ ppm amagat$^{-2}$, and $\sigma_3 = (0.16 \pm 0.01) \times 10^{-5}$ ppm amagat$^{-3}$. The coefficients $\sigma_1$, $\sigma_2$ and $\sigma_3$ arise from two-, three- and four-body interactions, respectively. At low densities the shielding constant depends linearly on density, whereas at high pressures many-body collisions become important as well and cause deviation from linearity. The second virial coefficient arises from the Xe–Xe pair interactions (with the potential $V(r)$, $r$ is the interatomic separation) and can be presented in the form.

$$\sigma_1(T) = 4\pi \int_0^\infty \sigma(r) \exp[-V(r)/kT]\, r^2\, \mathrm{d}r \qquad [2]$$

When xenon is mixed with another gas, G, the collisions Xe–G also contribute to the $^{129}$Xe shielding. If only binary collisions are considered to be important, the shielding of xenon is given by

$$\sigma_{\mathrm{exp}} = \sigma_0 + \sigma_1(\mathrm{Xe-Xe})\rho_{\mathrm{Xe}} + \sigma_1(\mathrm{Xe-G})\rho_{\mathrm{G}} \qquad [3]$$

where $\rho_{\mathrm{Xe}}$ and $\rho_{\mathrm{G}}$ are the densities of Xe and G, respectively. Once $\sigma_1$ (Xe–Xe) is determined in pure xenon gas, the $\sigma_1$ (Xe–G) term can be solved from Equation [3]. Mixtures of xenon with other noble gases as well as with some small molecules, such as CO, $N_2$, $O_2$, $CO_2$, $CH_nF_{4-n}$, $CH_4$, $SiF_4$, $SF_6$, etc., have been studied. The very recent coupled Hartree–Fock calculations with gauge-including atomic orbitals on the shielding surfaces of the Xe–$CO_2$, Xe–$N_2$ and Xe–CO systems predict second virial coefficients in fair agreement with the corresponding experimental ones. The calculations revealed the shielding surfaces to be highly anisotropic.

In the early days of $^{129}$Xe NMR, the spin–lattice relaxation time was assessed to be very long, and therefore a paramagnetic substance ($Fe_2O_3$) was used to shorten it in the first NMR experiments. The estimate of long $T_1$ was based on the dipolar interaction between two nuclei when they collide with each other. This model leads to the relaxation rate, $R_1 = 1/T_1$, which is inversely proportional to the mean collision time. Another interaction, more effective than the dipolar one, is the spin–rotation interaction, in which the nuclear spin couples with the angular momentum of a transient diatomic molecule formed during the collision. This model accounts for the experimental finding of $R_1$ being linearly dependent on the gas density; $R_1 = (5.0 \pm 0.5) \times 10^5 \rho$, where $R_1$ is given in s$^{-1}$ and $\rho$ in amagats. The $T_1$ value determined in hyperpolarized $^{129}$Xe has been found to be dependent upon the magnetic field strength: 155 min (gas pressure 790 torr) and 185 min (896 torr) at 2.0 T, and 66 min (790 torr) and 88 min (896 torr) at 7.05 T, experimental error being ~5% and temperature 20 °C. The slight variation of $T_1$ at a constant magnetic field is assumed to arise from differences in the cell wall structure, whereas the field dependence was interpreted as a consequence of the less effective wall interaction at

the higher magnetic field where the nuclear presession frequency is also higher.

The spin–lattice relaxation of the $^{131}Xe$ isotope is predominantly due to the interaction of the nuclear electric quadrupole moment with the electric field gradient (EFG) induced during binary collisions. Experiments have shown that also in this case the relaxation rate is linearly dependent upon the density: $R_1 = 0.039\,\rho$, where $R_1$ is in $s^{-1}$ and $\rho$ in amagats. Theory has given a similar relation with a proportionality coefficient of 0.046.

## Xenon as a probe in liquid and solid environments

### Isotropic liquids

For xenon dissolved in an isotropic liquid, the solvent effect on the shielding, $\sigma_m$, can be represented as

$$\sigma_{\rm m} = \sigma_{\rm exp} - \sigma_0 - \sigma_{\rm b} = \sigma_{\rm a} + \sigma_{\rm w} + \sigma_{\rm E} \qquad [4]$$

where $\sigma_{exp}$ is the experimental shielding constant, $\sigma_0$ is the shielding in the free atom (obtained by extrapolation of $\sigma_{Xe}$ to zero pressure), $\sigma_b$ arises from bulk susceptibility, $\sigma_a$ from the magnetic anisotropy of the nearest neighbouring solvent molecules, $\sigma_w$ from the van der Waals interactions, and $\sigma_E$ is the shielding contribution caused by the permanent electric dipole of the solvent. Solvent-induced change of the $^{129}Xe$ shielding is about 250 ppm, as can be seen from **Table 2**. On the other hand, the $^{129}Xe$ gas-to-solution shifts, i.e. the change of the shielding compared to the shielding of free xenon, are over 330 ppm.

Various models have been developed for explaining the solvent-induced changes in the $^{129}Xe$ shielding. For example, it has been proposed, based on the reaction field theory of Onsager, that the medium shift is proportional to the function $f(n) = [(n^2 - 1)/(2n^2+1)]^2$ (this is called the van der Waals continuum model), where $n$ is the refractive index of the solvent. Part of the experimental data indeed follows this relation, but most does not. An alternative model is provided with the pair interaction structureless approximation (PISA). The xenon–solvent dispersion energy, $E_{dis}$, calculated on this approximation is found to correlate better with the $^{129}Xe$ medium shift than $f(n)$.

One possible approach to gain insight into the solvent effects is to perform group contribution analysis. $^{129}Xe$ gas-to-solution shifts have been determined for pure *n*-alkanes, *n*-alkyl alcohols, *n*-alkyl carboxylic acids, di-*n*-alkyl ketones and cycloalkanes, and in solutions of lauric acid in *n*-heptane. It was found that the medium shifts corrected for solvent density are linearly dependent on the number of carbon atoms, except for the shortest members of the series of linear solvents (see **Figure 1**). Not only the structure of the environment but also temperature affects the $^{129}Xe$ shielding significantly. For example, in $CD_3CN$ the shielding increases with increasing temperature (i.e. with decreasing density) at the rate of 0.30 ppm $K^{-1}$. This is 33 Hz $K^{-1}$ at the magnetic field of 9.4 T. The position of the xenon resonance can be determined often with accuracy to better than 0.5 Hz, and consequently, $^{129}Xe$ shielding provides a good basis for a thermometer; accuracy may even be 0.02 K. A modified continuum

**Table 2** Solvent effect on the $^{129}Xe$ shielding. Values are referenced to zero-pressure xenon gas.

| *Solvent* | $\sigma_m$ (ppm) | *Solvent* | $\sigma_m$ (ppm) |
|---|---|---|---|
| Hexafluorobenzene | −85 | Water | −196 |
| Methanol | −148 | Chlorobenzene | −202 |
| Methyl chloride | −153 | Bromobenzene | −219 |
| Tetramethylsilane | −158 | Carbon tetrachloride | −222 |
| Ethanol | −165 | Methyl iodide | −239 |
| Fluorobenzene | −176 | Iodobenzene | −248 |
| Toluene | −190 | Methylene iodide | −335 |

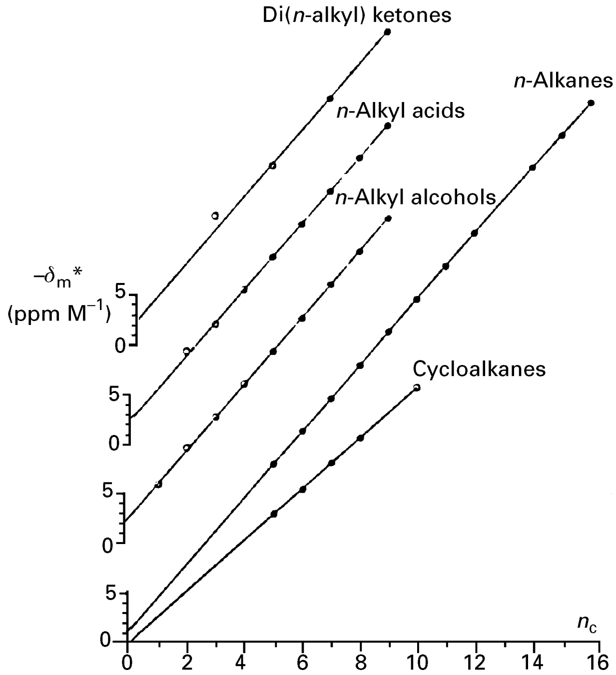



**Figure 1** Molar medium effect on the $^{129}Xe$ gas-to-solution shifts, $-\sigma^*_m$ as a function of the number of carbon atoms, $n_c$. $-\sigma^*_m = -\sigma_m/\rho$, where $\rho$ is density. Adapted with permission of the American Chemical Society from Luhmer M and Bartik K (1997) *Journal of Physical Chemistry A* **101**: 5278–5283.

model of van der Waals shifts has been presented to include also the effect of temperature. However, this model predicts temperature shifts in reasonable agreement with experiments only for the *n*-alkanes.

Although xenon NMR has been applied fairly widely to study physical properties of various materials, surprisingly little attention has been paid to its relaxation. The situation is, however, changing with the application of hyperpolarized $^{129}Xe$; the longitudinal hyperpolarized magnetization decays with the spin–lattice relaxation time, $T_1$. The relaxation mechanisms of the $^{129}Xe$ isotope are exclusively due to interparticle (xenon–solute molecule) interactions. In pure xenon gas, the relaxation mechanism has been proposed to arise from spin–rotation (SR) coupling during atomic collisions or during the transient existence of diatomic molecules. The SR interaction may also partly explain the relaxation of $^{129}Xe$ in benzene; the xenon atom is located on the $C_6$ symmetry axis of benzene with a binding energy of 10.4 KJ mol$^{-1}$. The dominating interaction in protonated solvents, however, is the $^{129}Xe$–$^1H$ dipolar interaction; in benzene its contribution is over 50% and in cyclohexane over 90% of the total relaxation rate, $R_1 = 1/T_1$. The $T_1$ ranges from ~70 s to ~1000 s for xenon in typical isotropic solvents. In blood cells and plasma, the $T_1$ is 4.5 s and 9 s, respectively, whereas in blood foam it is 21 s (oxygenated) and 40 s (deoxygenated).

The relaxation of the quadrupolar $^{131}Xe$ nucleus is predominantly due to the interaction between the nuclear electric quadrupole moment and the fluctuating EFG at the nuclear site. The origin of the EFG contributing in a solution is, however, still partly an open question. Various models, both electrostatic and electronic, have been developed. The electrostatic models assume the EFG to be due to solvent molecules represented by point charges, point dipoles or quadrupoles, or a dielectric continuum. In the electronic approach, EFG is considered to be a consequence of the deformation of the spherical electron distribution of $^{131}Xe$. The deformation arises from the collisions between xenon and solvent molecules. It is obvious (evidence is provided, for example, by $^{131}Xe$ NMR experiments in liquid–crystal solutions, and by first principles calculations) that neither of these approaches alone is sufficient. In typical isotropic solvents, the $^{131}Xe$ $T_1$ ranges from ~4 ms to ~40 ms.

## Liquid crystals

Thermotropic liquid crystals (LC) are anisotropic liquids that possess a mesophase (a phase with crystal and liquid properties) within a certain temperature range. In a spectrometer magnet, LC molecules tend to orient to a common direction which defines the director of the liquid crystal. The director may orient either along the external magnetic field or perpendicular to it, depending upon the sign of the anisotropy of the diamagnetic susceptibility. When xenon is dissolved in a liquid crystal and its $^{129}Xe$ NMR spectrum is recorded at variable temperatures, a series of spectra, as shown in **Figure 2**, may be obtained. This kind of experiment provides information on phase transition, isobaric thermal expansion coefficient, liquid–crystal orientational order parameters and anisotropy of the $^{129}Xe$ shielding tensor ($\Delta\sigma_d$). The latter property arises from the fact that in a mesophase, the originally spherical electron cloud of xenon is deformed leading to an axially symmetric shielding tensor with nonzero $\Delta\sigma_d$. The above-mentioned quantities can be derived from Equation [5] by least-squares fitting:

$$\begin{aligned}\sigma_{\exp}(T) - \sigma_0 = [1 - \alpha(T - T_0)]\{\sigma_d[1 + 2c\tau_1{}^2(T)] \\ + (2/3)P_2(\cos\phi)\Delta\sigma_d[S(T) \\ + 2c\sigma_1(T)\tau_1(T)]\} \qquad [5]\end{aligned}$$

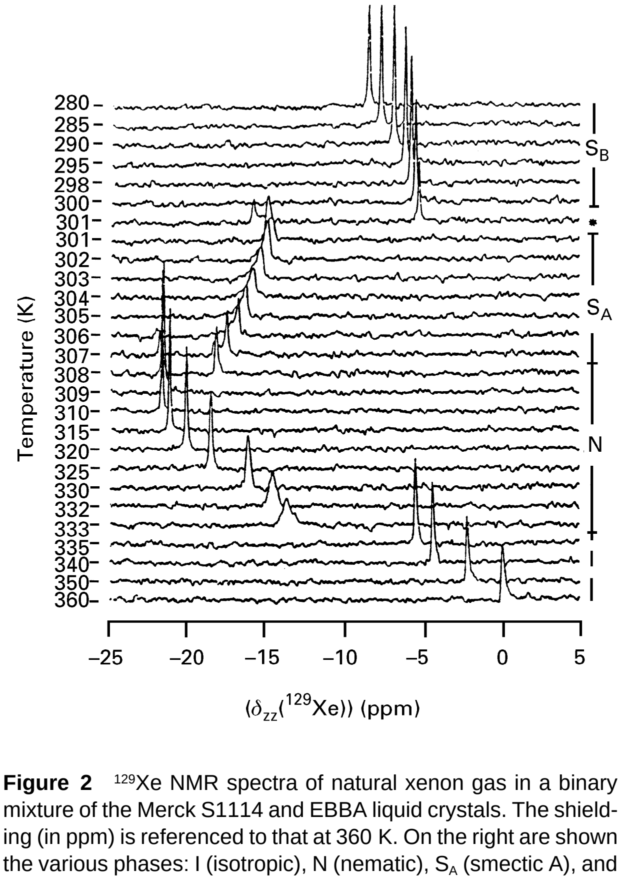



**Figure 2** $^{129}Xe$ NMR spectra of natural xenon gas in a binary mixture of the Merck S1114 and EBBA liquid crystals. The shielding (in ppm) is referenced to that at 360 K. On the right are shown the various phases: I (isotropic), N (nematic), $S_A$ (smectic A), and $S_B$ (smectic B). Adapted with permission of Gordon and Breach Publishers from Jokisaari J, Diehl P and Muenster O (1990) *Molecular Crystals and Liquid Crystals* **188**: 189–196.

where $\sigma_{exp}(T) - \sigma_0$ is the shielding difference for xenon in liquid–crystalline and gaseous phases, $\alpha$ is the isobaric thermal expansion coefficient, $T_0$ is the reference temperature (for example, the isotropic–nematic or nematic–smectic A phase transition temperature), $\sigma_d$ and $\Delta\sigma_d$ are the shielding constant and shielding anisotropy of xenon, $S(T)$, $\sigma_1(T)$ and $\tau_1 T$ are the conventional order parameter, translational–orientational order parameter and translational order parameter (the last two are present only in smectic phases), respectively, $P_2$ is the second Legendre polynomial, and $\phi$ is the angle between the external magnetic field and the liquid-crystal director, and the coefficient $c$ describes how much the positional distribution function deviates from a uniform distribution.

As mentioned above, the $^{131}$Xe nucleus possesses an electric quadrupole moment. In a liquid–crystalline solution the quadrupole moment interacts with the EFG at the nuclear site, and consequently, instead of a single resonance peak detectable in isotropic phases, a triplet with theoretical relative intensities of 3:4:3 is observed. (Generally, the multiplet consists of $2I$ resonance lines, where $I$ is the spin of the nucleus.) An example is given in **Figure 3**. The quadrupole splitting, i.e. the separation of the resonance peaks in a spectrum, can be used for determining external EFGs, i.e. EFGs arising from the electric multipoles of LC molecules, EFGs arising from the deformation of the electron cloud of xenon when it collides with LC molecules and LC orientational order parameters.

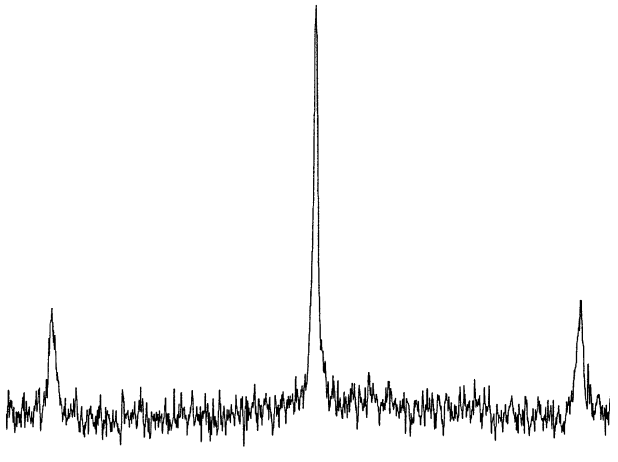


**Figure 3** The $^{131}$Xe triplet of xenon in the thermotropic Merck ZLI1167 liquid crystal. The frequency separation of the two outmost peaks is ~56 kHz. The intensity ratios are distorted because of experimental instabilities. Run parameters: $^{131}$Xe resonance frequency 49.218 MHz ($B_0$ = 14.1 T), acquisition time ~7 min, $T$ = 325 K. (Unpublished data from this laboratory.)

## Polymers

The physical and mechanical properties of polymeric systems are connected with their solid state morphology. NMR spectroscopy of the nuclear spins attached to a polymeric system is a very applicable means to gain insight into the microstructure as well as into the dynamics of the system. An alternative way is to make use of a probe, such as a xenon atom, which diffuses over the environment and gives information on the microscopic heterogeneity.

Since the xenon shielding is sensitive to the density of the surrounding medium, one may expect that it is affected by the glass transition of an amorphous polymer. Indeed this is the case, but, however, a more distinct change can be detected in the $^{129}$Xe line width, as shown for poly(ethyl methacrylate) in **Figure 4**. $^{129}$Xe NMR has proven to be particularly useful in studies of polymer blends whose components possess almost identical glass transition temperatures. Namely, for a phase-separated two-component blend, the $^{129}$Xe spectrum consists of two resonance signals, while the homogeneous morphology of a miscible blend yields a single resonance. The application of thermal analysis techniques is restricted by the fact that the different glass transitions can only be detected if they differ at least by 20 K.

When xenon is adsorbed, for example, into a solid EPDM rubber (a terpolymer composed of ethylene, propylene and ethylidene norbornene) at least four distinct $^{129}$Xe resonance signals can be observed indicating the presence of physically distinct domains; the intensity ratios may be used for the determination of the size of the domains, whereas the shielding differences reveal the variation of the destiny in the domains. When the rubber is cross-linked, the spectrum is clearly different from the one before

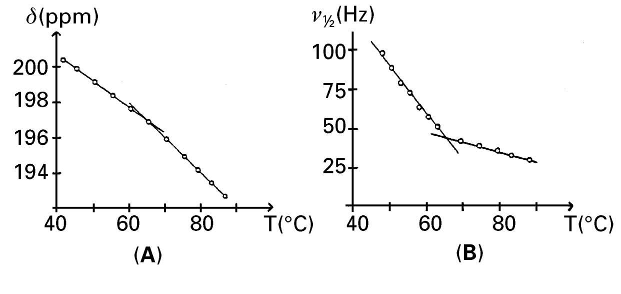



**Figure 4** (A) $^{129}$Xe chemical shift, and (B) $^{129}$Xe line width at 24.79 MHz as a function of temperature for xenon adsorbed in poly(ethyl methacrylate), the glass temperature, $T_g$ is 65 °C. Adapted (redrawn) with permission of the American Chemical Society from Stengle TR and Williamson KL (1987) *Macromolecules* **20**: 1430–1431.

cross-linking. As **Figure 5** shows, the cross-linking leads to the disappearance of the signal corresponding to the highest shielding of $^{129}$Xe, i.e. the largest amorphous voids. This is consistent with the fact that cross-linking produces a more condensed polymer matrix.

One possibility for investigating microheterogeneity in polymers with the xenon probe is to apply the cross-polarization (CP) technique, in which polarization is transferred from polymer protons to $^{129}$Xe. The necessary condition for $^{1}$H–$^{129}$Xe CP is for the xenon atom to be trapped long enough near a proton for the dipolar coupling between the nuclei to be effective. **Figure 6** displays the normal single-pulse $^{129}$Xe NMR spectrum, together with the $^{1}$H–$^{129}$Xe CP spectrum of xenon in a polymer blend of a copolymer (a mixture of 2/3 polyethylene, PE, and 1/3 polypropylene, PP) dispersed in a polymer matrix. The latter spectrum consists of a single resonance peak arising from xenon in the PP matrix where the translational mobility of the xenon atom is slow enough in order not to interrupt the dipolar coupling between the nuclei. Because of the CP between the PP protons and xenon adsorbed in the PP matrix, it is possible to obtain a correlation spectrum between the two spins, making it possible to identify the protons involved in the polarization transfer.

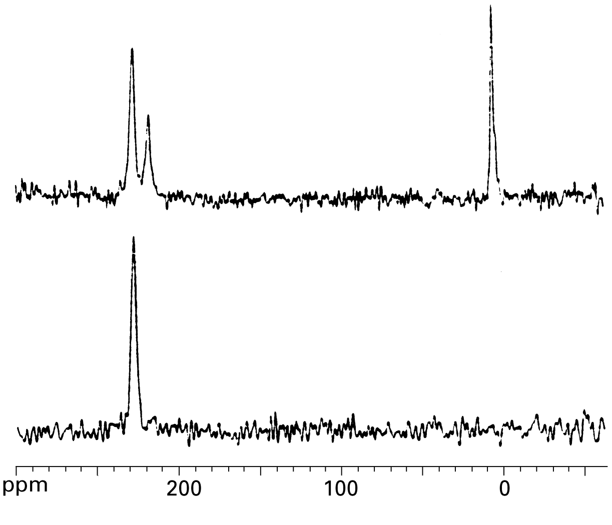



**Figure 6** (Top) The conventional single-pulse $^{129}$Xe NMR spectrum and (bottom) the $^{1}$H–$^{129}$Xe CP spectrum of xenon in a polymer blend. The mixing time in the CP experiment was 3 ms. The signal at 0 ppm arises from free xenon gas, the signal at 216 ppm from xenon in copolymer and the signal at 226 ppm from xenon in PP. Note: the scale is the chemical shift scale, which is opposite to the shielding scale. Adapted with permission of Elsevier Science Ltd from Mansfeld M and Veeman WS (1994) *Chemical Physics Letters* **222**: 422–424.

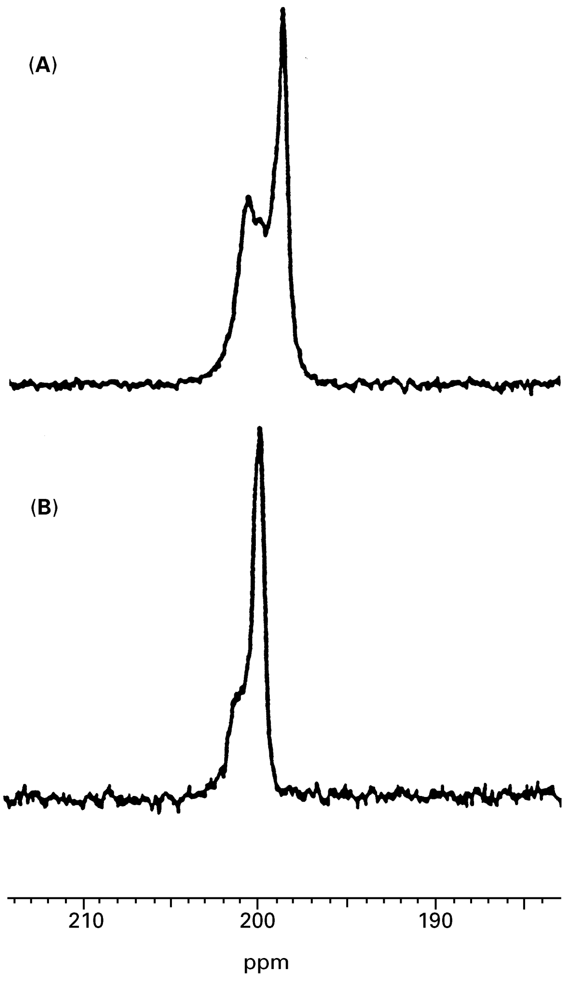



**Figure 5** $^{129}$Xe NMR spectra of xenon adsorbed in solid EPDM: (A) before, and (B) after crosslinking. Note: the scale is the chemical shift scale, which is opposite to the shielding scale. Adapted with permission of Springer-Verlag from Kennedy GJ (1990) *Polymer Bulletin* **23**: 605–606.

The efficiency of the polarization transfer depends upon the internuclear distance according to $r^{-6}$, and consequently, it is restricted to the nearest-neighbour protons of xenon. Thus CP experiments yield information only on the spatial proximity of distinguishable domains in a polymer. A much wider range of distance can be covered by two-dimensional (2D) exchange spectroscopy (EXSY). Its application is most useful in cases where the exchange rate of xenon between domains is slow compared to the chemical shift difference, and separate resonance signals from xenon in each domain can be observed. **Figure 7** displays results for $^{129}$Xe 2D EXSY experiments on a model blend system of poly vinyl chloride (PVC) and poly vinyl methyl ether (PVME). The system consists of thin alternating layers (thickness 2–6 μm) of the two polymers. The EXSY experiments were performed with two mixing times, 0.8 s and 8 s. It is seen that during the shorter mixing time, xenon samples all the local environments in the

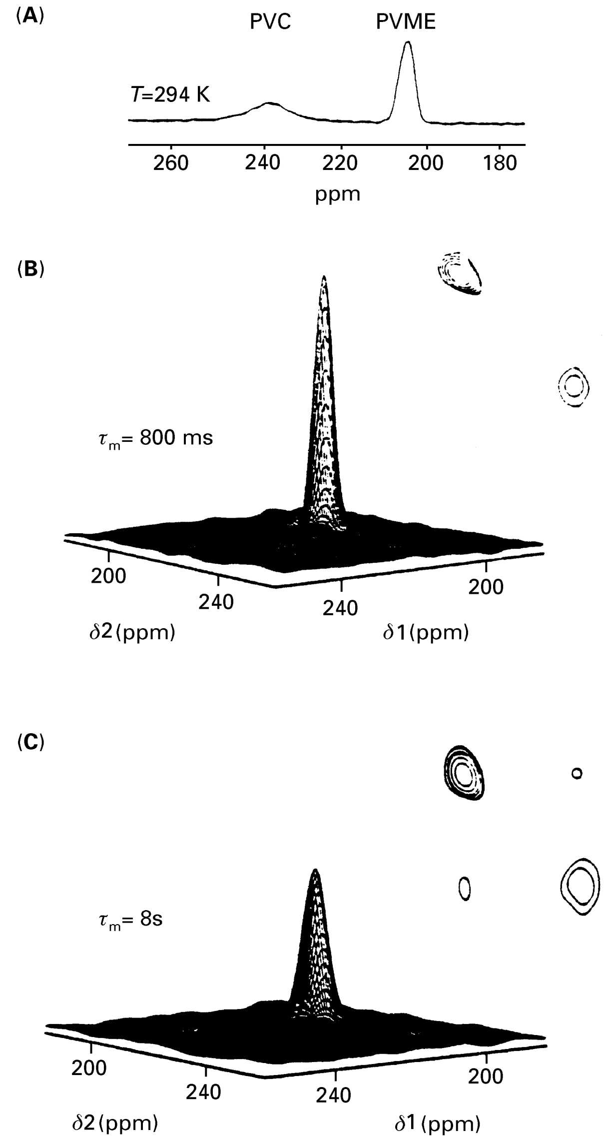



**Figure 7** $^{129}$Xe spectrum (A), and $^{129}$Xe 2D EXSY spectra (B and C) of xenon in the PVC/PVME model blend. $\tau_m$ is the mixing time. The insets show contour plots of the same data. Adapted with permission of Elsevier Science Ltd from Tomaselli M, Meier BH, Robyr P, Suter UW and Ernst RR (1993) *Chemical Physics Letters* **205**: 145–152.

PVC domains (this is indicated by the round shape of the diagonal peak), whereas no exchange between the PVC and PVME is taking place (there is no cross-peak in the spectrum). In contrast, during the mixing time of 8 s, exchange between the two phases is also seen.

## Zeolites, molecular sieves and clathrates

Most attention has been drawn to the application of $^{129}$Xe NMR to studies of zeolites, molecular sieves and clathrates. The main goal is to derive information on the size and shape of the pores of unknown structure through the $^{129}$Xe shielding measurements. In principle this information is available but not very straightforwardly since the xenon shielding is affected not only by the two factors but also by xenon–xenon collisions, and the presence of strong absorption sites (SAS), paramagnetic species and adsorbed molecules, etc. The experimental shielding of xenon in zeolites and molecular sieves is usually represented in the form

$$\sigma_{\rm exp} = \sigma_0 + \sigma_{\rm S} + \sigma_{\rm Xe} + \sigma_{\rm SAS} + \sigma_{\rm E} + \sigma_{\rm M} \qquad [6]$$

where $\sigma_0$ is the shielding of the reference (usually a zero pressure gas), $\sigma_S$ arises from the interaction of xenon with the pore walls, $\sigma_{Xe} = \sigma_{XeXe}\rho_{Xe}$ is the shielding contribution of the xenon–xenon collisions, $\sigma_{SAS}$ stems from Xe interaction with strong adsorption sites, $\sigma_E$ in turn takes into account electric field effects (due to charge–compensating cations) and $\sigma_M$ is the contribution of paramagnetic species.

**Figure** 8 shows the $^{129}$Xe chemical shift as a function of the number of xenon atoms in different

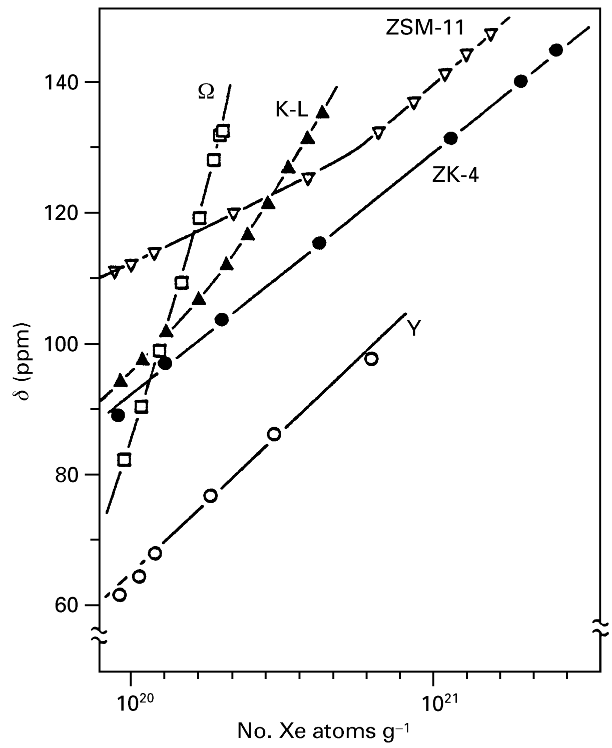



**Figure 8** Dependence of the $^{129}$Xe chemical shift (negative shielding) upon the number of xenon atoms adsorbed per gram of zeolite. Fraissard J (1996) In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance Spectroscopy*, Vol. 5, pp. 3058–3064. Chichester: Wiley © John Wiley & Sons Limited. Reproduced with permission.

zeolites. The shielding values (= $-\delta$), obtained by extrapolation to zero xenon number, range from ~ –110 ppm to ~ –60 ppm (in fact, in mordenite, which is not shown in the Figure, it goes down to –250 ppm), indicating the sensitivity of xenon shielding to zeolite structure.

No comprehensive theory exists for interpreting the experimental Xe shielding results in porous materials, although very significant progress in this direction has taken place recently; in particular, simulation calculations have improved our knowledge. In order to develop the method to a level giving as diversified information as possible of the shape and size of void space (and possibly cation distribution), it is important to investigate systems with varying properties and well-defined structure. The most simple systems are zeolites without charge compensating cations, and paramagnetic species when the two last shielding contributions in Equation [6] can be neglected. Very illustrative examples are the siliceous zeolite Si-ZSM-12 and molecular sieve $AlPO_4$-11. In the first approximation, both possess 1D channels with elliptical cross section. A closer inspection, however, reveals that the structure is more complex, consisting of series of cells, which have to be taken into account when interpreting experimental shielding data. The static $^{129}Xe$ spectrum in these two systems is a CSA (chemical shift anisotropy) powder pattern whose shape changes continuously from axially symmetric to asymmetric upon the xenon loading, as shown in **Figure 9**. Powder-like spectra have been observed for xenon only in a few zeolites. This may partly be due to misinterpretation of observed, slightly asymmetric line shapes; the xenon shielding is significantly affected by magnetic field inhomogeneity as well as by temperature gradients and fluctuations, and thus the CSA contribution, when small, has been masked by these instabilities during spectral recordings. The line shapes for xenon in Si-ZSM-12 and $AlPO_4$-11 have been explained by a dynamic model. In this model the shielding tensor is a dynamic, population-weighted average of tensors corresponding to three sites at which xenon has 0, 1 or 2 neighbouring xenons, and xenon samples rapidly fill the space available to it. When the elements of the shielding tensor are presented as a function of xenon loading, a linear dependence is observed at low loadings and thus the extrapolation to zero loading is straightforward and gives the shielding tensor elements corresponding to xenon with no neighbouring xenon atom.

The averaging of the xenon shielding tensor, and consequently, the derivation of structure information from the $^{129}Xe$ shielding data necessitates the knowledge of xenon dynamics. (One should emphasize here also that the dynamics of the zeolite framework affects significantly the averaging process. This is a consequence of the decrease of the effective potential barrier to intercellular jumps of xenon.) The observed line shape and shielding are averages over the shielding tensors of xenon sampling intracage volume and exchanging between cages, i.e. the mobility of xenon. For the first time, a CSA powder pattern was observed for xenon trapped in a clathrate where xenon is truly immobile. Xenon dynamics is available through spin–lattice relaxation time and diffusion measurements and 2D EXSY experiments.

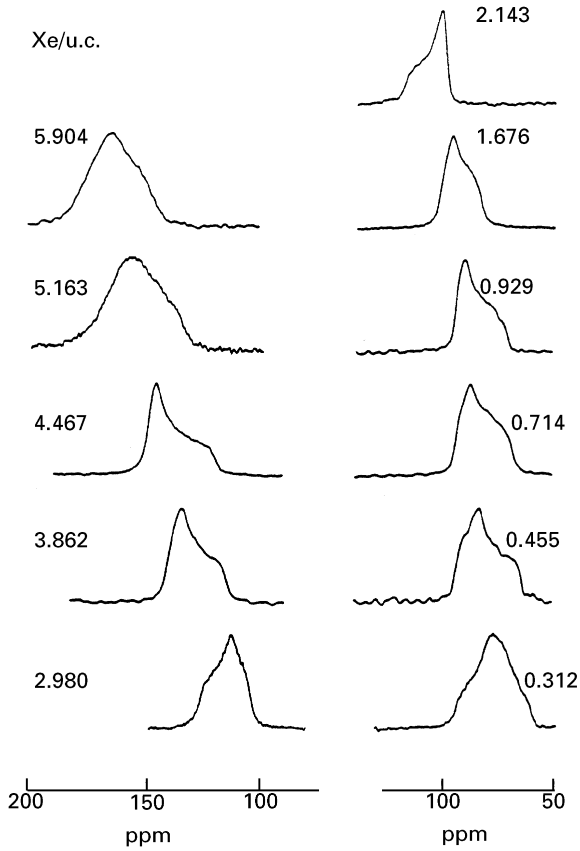



**Figure 9** Static $^{129}Xe$ NMR spectra of xenon adsorbed in Si-ZSM-12 zeolite for different loading levels (xenon atoms per unit cell, Xe/u.c.). All spectra were recorded at 295 K. Adapted with permission of Springer-Verlag from Moudrakovski IL, Ratcliffe CI and Ripmeester JA (1996) *Applied Magnetic Resonance* **10**: 559–574.

2D EXSY has been applied, for example, to study xenon dynamics in the NaA zeolite whose structure is composed of large α-cages (inner diameter approximately 11.4 Å) and smaller β-cages (6.6 Å). At elevated temperatures and pressures the Xe atoms are distributed among the α-cages. This is seen in the $^{129}Xe$ 1D NMR spectrum which displays several

distinct resonance signals (see **Figure 10**); xenon shielding decreases with increasing number of atoms in a cage. **Figure 10** also shows the results of the 2D EXSY experiments performed with three different mixing times. The emergence of cross-peaks arises from intercage motion of xenon during the mixing time. Performing the experiment at variable temperatures allows for the derivation of the rates of intercage motion as well as the adsorption and activation energies of the xenon atoms.

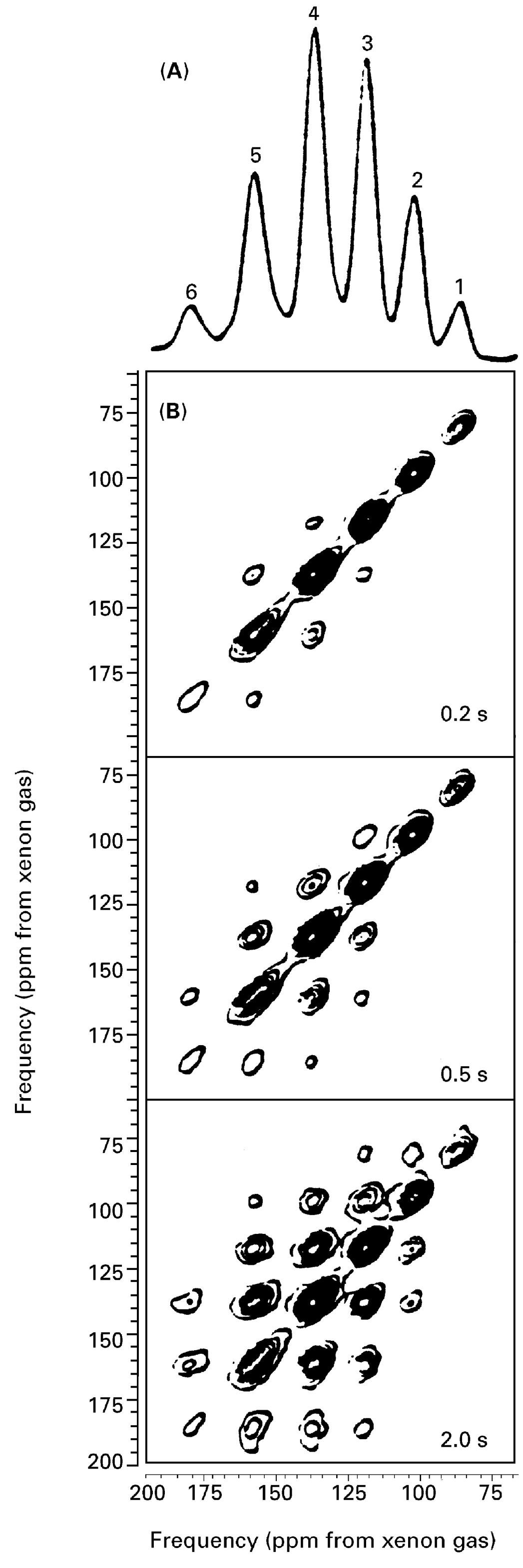



**Figure 10** (A)$^{129}$Xe 1D NMR spectrum (on top of each peak is shown the number of Xe atoms in the cage) and (B) $^{129}$Xe 2D EXSY spectra of xenon adsorbed in NaA zeolite at 523 K and 30 atm (1 atm = 101 325 Pa). The mixing times are: 0.2, 0.5 and 2.0 s as shown. Adapted with permission of Elsevier Science Ltd from Larsen RG, Shore J, Schmidt-Rohr K *et al* (1993) *Chemical Physics Letters* **214**: 220–226.

## Applications of laser-polarized xenon

The $^{129}$Xe magnetization can be enhanced compared to the thermal equilibrium magnetization by a factor of ~$10^5$ by utilization of optical pumping and spin-exchange between xenon and alkali-metal atoms. The resulting state is a nonequilibrium state, and the magnetization is often called hyperpolarized magnetization. One has to take into account the following facts when performing NMR experiments: (i) the hyperpolarized spins decay toward thermal equilibrium with the spin–lattice relaxation time $T_1$, and thus the hyperpolarization cannot be recovered by waiting for thermal equilibrium, (ii) when a radiofrequency $\theta$ pulse is applied to the spin systems, the longitudinal hyperpolarized magnetization decreases by cos $\theta$. On the other hand, when applying repeated pulses the pulse repetition time can be short because there is no need to wait for the spin system to return to thermal equilibrium. In most applications the $^{129}$Xe NMR spectrum of hyperpolarized xenon can be obtained on a single pulse. It is also possible to apply continuously flowing hyperpolarized $^{129}$Xe.

Hyperpolarized $^{129}$Xe can be utilized in two ways: firstly, by observing directly the $^{129}$Xe NMR spectrum of xenon introduced to the environment under study, and secondly, by transferring hyperpolarization from xenon to other nuclei and recording their spectra. The polarization transfer takes place through cross-relaxation between spins without the need for irradiation of the spins. This technique has been denoted SPINOE (spin polarization induced nuclear Overhauser effect). The SPINOE may be either positive or negative depending upon whether right or left circularly polarized light is applied.

When hyperpolarized $^{129}$Xe gas is dissolved, for example, into benzene, the $^1$H NMR signal is enhanced because of SPINOE. This finding opens up new views for deriving information, for example, on Xe–protein interactions as well as on blood and other biological systems. The very high sensitivity of hyperpolarized $^{129}$Xe makes it possible to investigate also low surface area (1–10 m$^2$ g$^{-1}$) nonporous solids

which otherwise, due to the low inherent sensitivity of NMR, are difficult if not impossible to investigate. On the other hand, magnetization transfer from hyperpolarized $^{129}Xe$ to other nuclei, e.g. $^{1}H$ and $^{13}C$ on a surface, allows for the direct observation of the magnetic resonances of these isotopes. Experiments have shown that enhancement of proton magnetization even by a factor of 20 may be achieved at low temperatures. The enhancement is more pronounced at low temperatures because of decreasing mobility of xenon atoms. An example of the $^{1}H$ magnetization enhancement is given in **Figure 11**.

Clathrates were the first systems investigated by $^{129}Xe$ NMR of natural xenon gas. The xenon atom is of the same size and shape as methane and it also forms a clathrate hydrate with water. The xenon shielding is much more sensitive (by a factor of about 30) than the $^{13}C$ shielding of methane to the changes taking place in the structure of their environment. The use of hyperpolarized $^{129}Xe$ allows the investigation of large and small cages of clathrate hydrate; two distinct resonance signals are seen in the spectrum as shown in **Figure 12**. Much interest has been drawn to the application of hyperpolarized $^{129}Xe$ to materials, and in particular, to medical imaging (xenon gas is a safe general anaesthetic) and to spectroscopy in blood systems and in tissue, specifically the heart, lungs, brain and other organs. Recently, the technique was used to obtain human lung images. In this case, ~0.5 L of hyperpolarized $^{129}Xe$ is needed, necessitating higher power (100–120 W) lasers.

## Xenon compounds

Xenon is known to covalently bond to fluorine, oxygen, nitrogen, carbon and to itself. In such circumstances, $^{129}Xe$ exhibits a large range of chemical

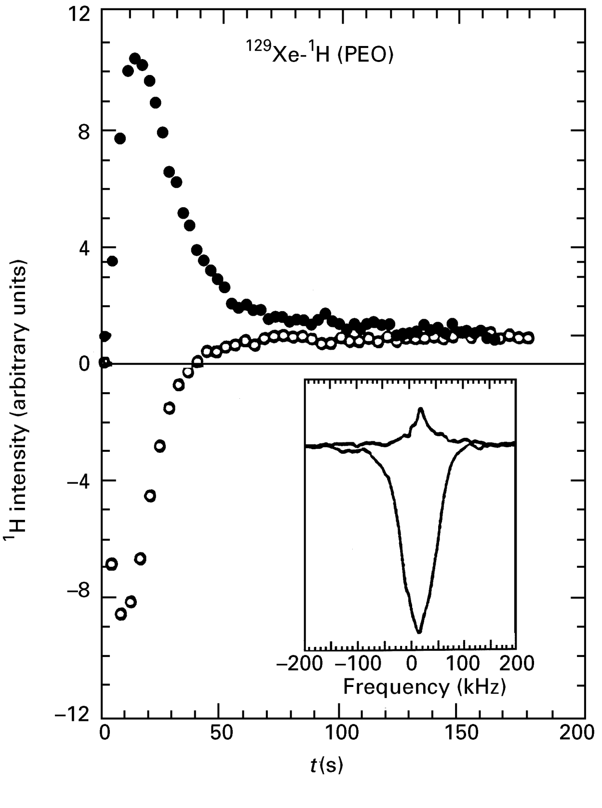



**Figure 11** Evolution of the $^{1}H$ spin magnetization of PEO (polyethylene oxide)-coated Aerosil1130 as a function of time after exposure of the surface to hyperpolarized $^{129}Xe$. The initial Xe pressure is 160 torr and the sample temperature is 130 K for positive SPINOE (●) and 125 K for negative SPINOE (○). The inset displays two single-shot $^{1}H$ spectra from a negative SPINOE run, one taken at the Boltzmann equilibrium for the unpolarized sample (positive peak) and the other taken at the time $t_0$, when the negative SPINOE enhancement has reached its maximum absolute value. Adapted with permission from Rõõm T, Appelt S, Seydoux R, Hahn EL, and Pines A (1997) *Physical Review B* **55**: 11604–11610.

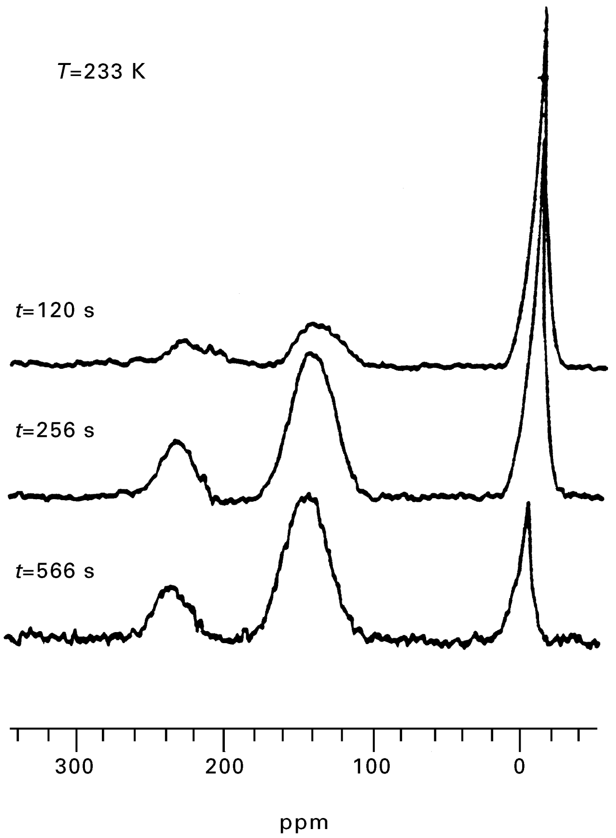



**Figure 12** The $^{129}Xe$ NMR spectra of the formation of a xenon clathrate hydrate at 233 K and time *t* after admission of the xenon to the powdered ice sample. The signal at 160 ppm is attributed to xenon in the large tetrakaidecahedral cages and the one at 240 ppm to xenon in the smaller dodecahedral cages. Adapted with permission of the American Chemical Society from Pietrass T, Gaede HC, Bifone A, Pines A and Ripmeester JA (1995) *Journal of the American Chemical Society*, **117**: 7520–7525.

shifts, about 7500 ppm. **Figure 13** displays chemical shifts of some selected xenon compounds with different oxidation states. Also the $^{129}Xe$ shielding anisotropy may be large. For example, the theoretical estimate of the shielding anisotropy for $XeF_2$ is 5125–7185 ppm, for $XeF_4$ it is 3940 ppm and for $XeOF_4$ it is 365–1500 ppm. Spin–lattice relaxation time measurements have given an anisotropy of ~4700 ppm for $XeF_2$. In this particular case, the relaxation is dominated by the CSA and SR interactions and the $T_1$ values range from 150 to 430 ms depending upon magnetic field strength (CSA interaction depends upon the square of the magnetic flux density) and temperature (SR interaction depends linearly upon temperature). In general, the $^{129}Xe$ $T_1$ of various species in solution varies between 285 and 780 ms. Thus the relaxation is much faster than that of atomic xenon, and renders possible fast repetition rates in data accumulation. The one-bond spin–spin couplings of xenon with $^{19}F$ are relatively large and include a sizable relativistic contribution. The change of the absolute value of $^1J(^{129}Xe–^{19}F)$ can be used as a diagnostic tool to confirm the formal oxidation number of xenon as the coupling decreases in the order: Xe(II)> Xe(IV)>Xe(VI). Xenon couplings to other nuclei are smaller. No absolute sign determination has been made for any of the couplings to xenon. **Table 3**, shows absolute spin–spin coupling values of some selected xenon compounds.

## List of symbols

$c$ = distribution coefficient; $E_{dis}$ = dispersion energy; $I$ = nuclear spin; $P_2$ = second Legendre polynomial;

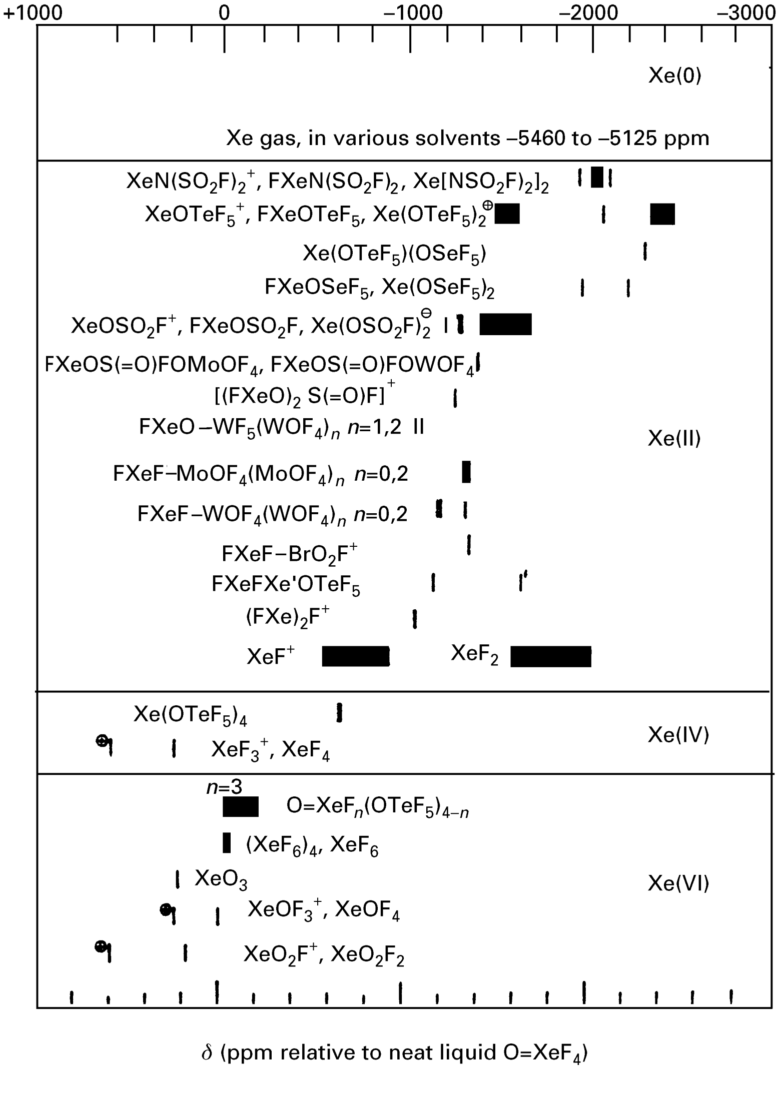



**Figure 13** $^{129}Xe$ NMR chemical shifts of a few selected xenon compounds. Adapted with permission from Jameson C (1987) The noble gases. In: Mason J (ed) *Multinuclear NMR*, Chapter 8, pp. 463–477. New York: Plenum Press.

**Table 3** Absolute values of xenon spin–spin couplings for selected molecules.

| *Molecule* | *Coupling* | *Value* (Hz)[a] |
|---|---|---|
| Xe(II) | | |
| $XeF_2$ | $^1J(^{129}Xe,^{19}F)$ | 5644 |
| $CH_3C{\equiv}CN{-}XeF^+$ | $^1J(^{129}Xe,^{19}F)$ | 6020 |
| $C_6F_5C{\equiv}N{-}XeF^+$ | $^1J(^{129}Xe,^{19}F)$ | 6610 |
| Xe(IV) | | |
| $XeF_5^-$ | $^1J(^{129}Xe,^{19}F)$ | 1056–1082 |
| *cis*-$F_2Xe(OTeF_5)_2$ | $^1J(^{129}Xe,^{19}F)$ | 3714 |
| *trans*-$F_2Xe(OTeF_5)_2$ | $^1J(^{129}Xe,^{19}F)$ | 3503 |
| $XeF_4$ | $^1J(^{129}Xe,^{19}F)$ | 3801–3900 |
| Xe(VI) | | |
| $OXeF(OTeF_5)_3$ | $^1J(^{129}Xe,^{19}F)$ | 1056–1082 |
| $OXeF_3(OTeF_5)$ | $^1J(^{129}Xe,^{19}F)$ | 1127–1148 |
| $XeOF_4$ | $^1J(^{129}Xe,^{19}F)$ | 1115–1131 |
| $(XeF_6)_4$ | $^1J(^{129}Xe,^{19}F)$ | 330–331.7 |
| $XeOF_4$ | $^1J(^{129}Xe(II),^{17}O)$ | 692–704 |
| $CH_3{\equiv}N{-}XeF^+$ | $^1J(^{129}Xe(II),^{14}N)$ | 313 |
| $C_6F_5Xe^+$ | $^1J(^{129}Xe(II),^{13}C)$ | 119 |
| $Xe(OSeF_5)_2$ | $^2J(^{129}Xe(II),^{77}Se)$ | 130 |
| $Xe(OTeF_5)_2$ | $^2J(^{129}Xe(II),^{125}Te)$ | 470 |
| $HC{\equiv}NXeF^+$ | $^3J(^{129}Xe(II),^1H)$ | 24.7–26.8 |

[a] In some cases, the range of experimental values is given.

$r$ = interatomic separation; $R_1$ = longitudinal relaxation rate; $T_0$ = reference temperature; $T_1$ = spin–lattice relaxation time; $V(r)$ = potential function for Xe–Xe pair interactions; $\alpha$ = isobaric thermal expansion coefficient; $\delta$ = chemical shift; $\Delta\sigma_d$ = shielding anisotropy of xenon; $\rho$ = density; $\sigma_0$ = shielding constant of reference gas; $\sigma_a$ = magnetic anisotropy shielding constant; $\sigma_b$ = bulk susceptibility shielding constant; $\sigma_{exp}$ = experimental shielding constant; $\sigma_E$ = permanent electric dipole shielding constant; $\sigma_M$ = paramagnetic species shielding constant; $\sigma_S$ = surface induced shielding constant; $\sigma_{SAS}$ = strong absorption sites shielding constant; $\sigma_W$ = van der Waals interaction shielding constant; $\sigma_{Xe}$ = shielding constant of xenon gas; $\tau$ = time; $\tau_M$ = mixing time; $\phi$ = angle between external magnetic field and liquid–crystal director.

*See also:* **Chemical Exchange Effects in NMR; Diffusion Studied Using NMR Spectroscopy; Gas Phase Applications of NMR Spectroscopy; Liquid Crystals and Liquid Crystal Solutions Studied By NMR; MRI Applications, Clinical; NMR in Anisotropic Systems, Theory; NMR Microscopy; NMR Relaxation Rates; Nuclear Overhauser Effect; Polymer Applications of IR and Raman Spectroscopy.**

## Further reading

Albert MS, Cates GD, Driehuys B *et al* (1994) Biological magnetic resonance imaging using laser-polarized $^{129}$Xe: *Nature (London)* **370**: 199–201.

Barrie JP and Klinowski J (1992) $^{129}$Xe NMR as a probe for the study of microporous solids: A critical review. *Progress in NMR Spectroscopy* **24**: 91–108.

Dybowski C and Bansal N (1991) NMR spectroscopy of xenon in confined spaces: clathrates, intercalates, and zeolites. *Annual Review on Physical Chemistry* **42**: 433–464.

Fraissard J and Ito T (1988) $^{129}$Xe N.M.R. study of adsorbed xenon: A new method for studying zeolites and metal–zeolites. *Zeolites* **8**: 350–361.

Jameson C (1987) The noble gases. In: Mason J (ed) *Multinuclear NMR*; pp 463–477. New York: Plenum Press.

Jokisaari J (1994) NMR of noble gases dissolved in isotropic and anisotropic liquids. *Progress in NMR Spectroscopy* **26**: 1–26.

Miller JB (ed) (1995) Special issue on magnetic resonance studies of noble gases. *Applied Magnetic Resonance* **8**: 337–595.

Raftery D and Chmelka BF (1994) Xenon NMR spectroscopy. *NMR Basic Principles and Progress*, Vol **30**: pp. 112–158. Berlin: Springer-Verlag.

Ratcliffe CI (1998) Xenon NMR. *Annual Reports on NMR Spectroscopy* **36**: 124–208.

Schrobilgen GJ (1996) Noble gas elements. In: Grant DM and Harris RK (eds) *Encyclopedia of Nuclear Magnetic Resonance*, Vol 5; pp 3251–3262. Chichester: Wiley.

Walker TG and Happer W (1997) Spin-exchange optical pumping of noble gas nuclei. *Review of Modern Physics* **69**: 629–642.

# X-Ray Absorption Spectrometers

**Grant Bunker**, Illinois Institute of Technology, Chicago, IL, USA

Copyright © 1999 Academic Press

HIGH ENERGY SPECTROSCOPY

**Methods & Instrumentation**

## Introduction

The task of an X-ray absorption spectrometer is the precise and accurate measurement of the linear X-ray absorption coefficient of a substance. A principal use of such spectrometers is the measurement of X-ray absorption fine structure (XAFS) spectra of solids, liquids, and molecular gases. XAFS consists of modulations in the X-ray absorption coefficient in the vicinity of an X-ray absorption edge, which may extend more than one KeV beyond the edge. Applications and theory of X-ray Absorption Spectroscopy are covered elsewhere in this volume. This article is directed primarily to instrumental requirements for X-ray absorption spectroscopy over the energy range from several KeV X-ray photon energy to approximately 100 KeV, with emphasis on synchrotron radiation based instruments.

## Absorption, fluorescence, and fluorescence excitation spectra

In the simplest case, the X-ray absorption coefficient ($\mu$) of a homogeneous sample of thickness $x$ is given by

$$\frac{I}{I_0} = \exp[-\mu(\varepsilon)x]$$

where $I_0$ and $I$ respectively are the incident and transmitted X-ray intensities at X-ray photon energy $\varepsilon$. An analogous quantity in optical absorption spectroscopy is the product of the molar extinction coefficient and the concentration.

X-ray *absorption* spectrometers are distinct from X-ray *fluorescence* spectrometers (often referred to as 'X-ray spectrometers'), which measure the intensity of fluorescence radiation emitted by atoms in a specimen following their excitation by high energy photons, charged particle beams, or other interactions. X-ray fluorescence spectrometers are primarily used for measuring the elemental composition of samples, or measuring shifts in energy of fluorescence, to obtain chemical information about a sample.

A principal use of X-ray absorption spectrometers is in the measurement of X-ray absorption fine structure (XAFS) spectra, which provide quantitative information on the local structural and chemical environment within the region several ångstroms around selected atomic species in a sample. Absorption and fluorescence spectroscopies are complementary, and both can be used for spatially resolved spectroscopic mapping of samples. Fluorescence spectrometers also have considerable applicability for determining elemental composition in astrophysical research, planetary sciences, and numerous other areas.

There are many situations in which the X-ray absorption spectrum is most easily measured (indirectly) by monitoring the fluorescence produced following absorption of X-rays. One measures variations in the fluorescence intensity of a particular atomic species as the energy of incident photons is varied over an absorption edge of a selected element. This 'fluorescence excitation spectrum', distinct from the fluorescence spectrum, provides an indirect measurement of the X-ray absorption coefficient, albeit one that is subject to several well-known instrumental effects. If care is taken in sample preparation, systematic errors can be minimized. Fluorescence detection is the method of choice for dilute systems, because it provides an improved signal-to-noise ratio.

## X-ray absorption spectra

Over the range from several KeV to 100 KeV, X-ray photons propagating through a sample are absorbed, scattered without loss of energy (elastic scattering), or scattered with loss of energy (inelastic scattering). X-rays require energies in excess of one MeV for positron–electron pair production to occur.

Cross sections vary significantly as a function of energy, as shown in **Figure 1A** and **B**. Absorption cross sections between the absorption edges decrease approximately as $1/\varepsilon^3$.

## Absolute and relative measurements

Although it is essential to use samples of appropriate thickness and concentration in XAFS experiments, in practice it is seldom necessary to precisely determine

the *absolute* absorption coefficient. The quantities of interest (e.g. inter-atomic bond lengths) are intrinsic to the material, while the absolute absorption coefficient depends on the thickness of the specimen, concentration of the element of interest, and other irrelevant factors. The standard methods of XAFS analysis treat the data in such a way that the structural parameters ultimately determined are invariant with respect to change of multiplicative scale factors, and additive background, provided it is a sufficiently smooth function of energy. To the extent that the X-ray scattering cross-sections vary slowly with energy, or are negligible compared to the photo electric cross-section, they do not affect the structure determination. It should be noted, however, that both the elastic and inelastic scattering cross-sections do have small contributions that vary near absorption edges and may need to be accounted for in some circumstances.

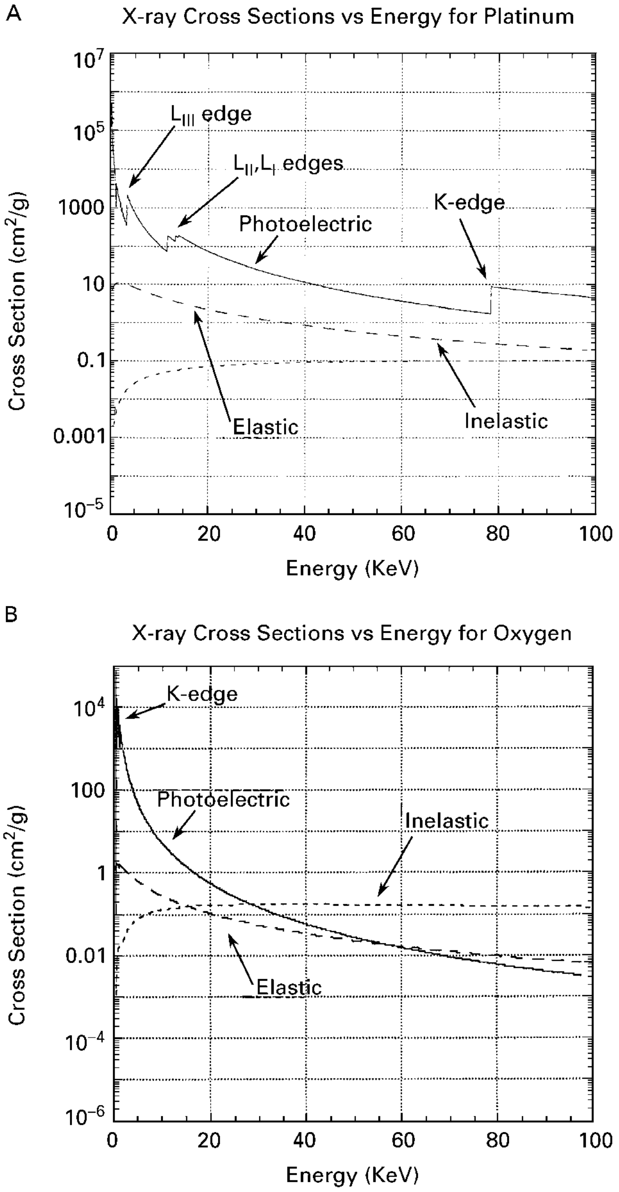



**Figure 1** Absorption and scattering cross-sections for (A) platinum, and (B) oxygen.

For some purposes the absolute cross-section is needed; for example, to determine the areal concentration of a particular atomic species in the sample, or to quantify the elemental composition of samples. The simplest approach is to measure the absorption coefficient with and without the sample inserted into the beam, as is done in a single beam optical spectrophotometer. Measurement precision can in principle be improved by modulation, i.e. rapidly performing differential measurements by rapidly inserting and removing the sample, for example by using a rotating disk with apertures that alternately contain a sample or a blank. This approach may be impractical for some samples, particularly if they require a special environment such as ultra high vacuum, high pressure, etc. In such cases it is possible to deflect the X-ray beam with a glancing incidence X-ray mirror, or by introducing an X-ray beam splitter.

Measurements of highest accuracy discriminate between the scattered, refracted, and transmitted beams. This can be made possible through the use of crystal analysers following the sample. If the X-ray beam is polarized and statistically non-isotropic it is important to account for orientation-dependent effects. Magic-angle spinning of the sample can be an effective means of averaging over anisotropies.

## Requirements for X-ray absorption spectroscopy

X-ray absorption fine structure experiments require energy resolution on the order of several electron volts or less, to resolve modulations in the spectra. Spectra are intrinsically broadened by the 'core-hole lifetime broadening', which is an effect stemming from the relatively rapid ($<10^{-15}$ s) filling of the core-hole state (initial state vacancy) produced after the X-ray absorption event. Heisenberg's time–energy uncertainty relation $\Delta E\ \Delta t \geq h/2\pi$, implies that the rapid decay of the core-hole state broadens the energy spectrum with an energy level width $\Delta E$ inversely proportional to the lifetime $\Delta t$. The level width increases rapidly with atomic number.

Several stringent criteria must be met before an apparatus can be regarded as a suitable X-ray absorption spectrometer for XAFS. First, the device must have an appropriate energy resolution (several

eV or less); tunability (i.e. smooth, reliable scanning over an energy range of more than ~1000 eV range above selected absorption edges; and high flux generally $>10^{10}$ photons $s^{-1}$). The harmonic content, i.e. contributions from high energy photons at multiples of the selected energy, should be limited to less than 0.1% of the fundamental. Beam intensity variation over a scan should be kept within ~20% depending on the linearity of detectors. **Figure 2** shows the layout of a typical instrument for transmission XAFS experiments.

## Sources

The collimated beams, smooth energy spectrum, and high intensity of synchrotron radiation sources offer compelling advantages for XAFS experiments, compared to conventional fixed and rotating anode X-ray generators, although the latter are useful in some situations.

Historically, synchrotron radiation sources were associated with particle accelerators constructed for high energy physics experiments – so-called first generation sources. Electrons or positrons (anti-electrons) are accelerated in a closed orbit through an evacuated pipe at speeds exceedingly close to the speed of light:

$$\frac{v}{c} \sim 1 - \frac{1}{2\gamma^2}$$

where $\gamma = E/mc^2$ and $E$ is the particle energy. These particles are used because particles of low mass radiate energy much more efficiently than those of high mass. Through the use of magnets the path of the particles is bent into a closed path of several hundred to one thousand metres circumference, through which the particles circulate at frequencies of hundreds of kilohertz to megahertz. Energy lost from the particle beam by synchrotron radiation is replenished through the use of radio-frequency cavities which apply a force to the particles along the direction of motion as they pass by. The particles travel in discrete 'bunches', and therefore the radiation produced has a pulsed structure that can be used to advantage for time-resolved experiments. The electrons circulate for many hours; scattering of the electrons from residual gas atoms in the ultra high vacuum environment in the ring, intra-bunch electron–electron interaction, quantum perturbations from spontaneous emission of photons, and other effects cause a slow loss of particles from the beam, which therefore must be periodically refilled. It is technically feasible to periodically replenish the beam and preserve nearly constant current in the ring.

All accelerating charged particles radiate energy, and if it were not for relativistic effects, the radiation produced by a synchrotron would be in the radio frequency spectrum. The relativistic motion causes the familiar dipole radiation lobes of an accelerating charge (seen in an inertial frame co-moving with the charge) to tilt forward along the direction of motion, in the observer's reference frame. The radiation pattern from a bend magnet is a horizontal fan of several milliradians angular width in the horizontal direction (in the orbital plane), and it is very well collimated in the vertical direction to an opening angle of order $1/\gamma$, where $\gamma = E/mc^2$, $E$ is the particle beam energy, $m$ is the rest mass of the electron, and $c$ is the speed of light. For all bend magnets the broad spectrum (integrated over vertical opening angle) is described by a universal function

$$g_1(x) \equiv x \int_x^\infty K_{5/3}(t)\,\mathrm{d}t$$

where $K_{5/3}$ is the modified Bessel function of order $\frac{5}{3}$. The spectrum is parameterized by the 'critical energy'

$$\varepsilon_c = 3\frac{hc}{4\pi\rho}\gamma^3$$

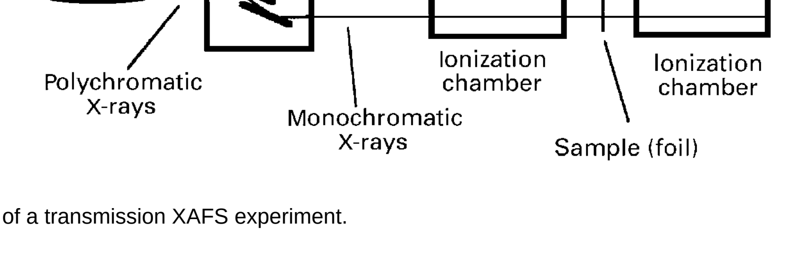



**Figure 2** Schematic of a transmission XAFS experiment.

where $\varepsilon$ is the X-ray energy, $h$ is Planck's constant, and $\rho$ is the bend radius [$\rho_\beta$ (metres) ~3.336 $E$(GeV)/$B$(Tesla)]. The root mean square angular width in the vertical direction can be approximated by $(0.57/\gamma)(\varepsilon_c/\varepsilon)^{0.43}$. **Figure 3** shows the universal spectral curve for bend magnets and planar wigglers (see below).

Second generation sources are constructed specifically to produce bend magnet radiation for experimental use. Third generation sources are optimized for the use of 'insertion devices' – magnetic structures inserted into the particle beam path that modify the trajectory so as to produce synchrotron radiation of the desired energy spectrum, spatial characteristics, and polarization. They are designed for low emittance, which is the product of the spread in momentum and the spread in position of the particle beam (the phase space volume). According to Liouville's theorem the emittance is conserved as the beam propagates through the dipole, quadrupole, and sextupole magnets used to guide and focus the particles.

Wigglers are arrays of alternating magnetic poles that apply an approximately sinusoidal magnetic field to the particles and cause their trajectory to oscillate. The spectrum and angular radiation pattern is similar to that of an array of bend magnets of alternating curvature, but with the advantages of higher flux owing to multiple magnetic poles, and the critical energy determined by experimental requirements rather than geometrical constraints.

Undulators are similar to wigglers in that they produce an alternating magnetic field that causes the particle trajectory to oscillate in an approximately sinusoidal manner. The tangent direction of the particle trajectory is kept within the intrinsic width of the synchrotron radiation cone, which allows the X-rays emitted at each successive magnetic pole to interfere. This interference concentrates the energy into narrow energy bands and a narrow angular divergence in both horizontal and vertical directions. The amplitude of oscillation of the particles is characterized by the 'deflection parameter' $K = \gamma\delta_w$, where $\delta_w = \lambda_0/2\pi\rho_0$, $\lambda_0$ is the undulator period, and $\rho_0$ is the bend radius corresponding to the peak magnetic field. For small oscillations, there is a single peak in the spectrum at 2 $\gamma^2$ times the frequency of oscillation of the particle $\Omega_w$, as measured in the inertial frame of the average particle velocity. Undulators in use for X-ray spectroscopy generally have sufficiently high fields that they have characteristics intermediate between an ideal undulator and wigglers. In this case the particle motion becomes relativistic even in its co-moving reference frame, and harmonics are generated. The X-ray frequency of the fundamental observed at an angle $\theta_0$ is given approximately by

$$\frac{2\,\gamma^2\,\Omega_w}{1 + (K^2/2) + \gamma^2\,\theta_0{}^2}$$

This expression shows that the positions of peaks in the spectrum can be controlled by adjusting the deflection parameter $K$, by controlling the magnetic field presented to the particle beam. The energy width of undulator peaks is typically of the order of 100 eV, decreasing with the number of poles. The fluxes from even-order harmonics are of significantly lower amplitude, particularly on the undulator axis. **Figure 4** shows the spectrum from an APS type A undulator.

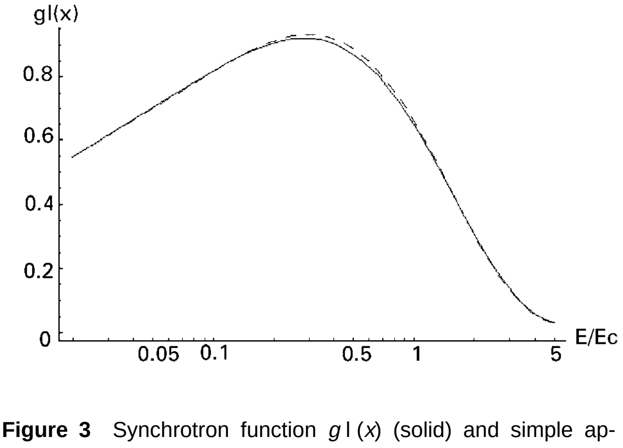



**Figure 3** Synchrotron function $gI(x)$ (solid) and simple approximation (dashes): $f(x) = 1.8\,x^{0.3}\exp(-x)$, where $x = \varepsilon/\varepsilon_c$. A more accurate approximation (not shown) is $gI(x) = ax^b\exp(-cx)$ with $a = 1.71857$, $b = 0.281526$, $c = 0.968375$. The spectral photon flux (photons/sec/0.1% bandwidth ($\Delta\varepsilon/\varepsilon$)/mA beam current/mrad) integrated over the full vertical opening angle is $1.256 \times 10^7\,\gamma g I\,[x]$, with $\gamma = E/mc^2$.

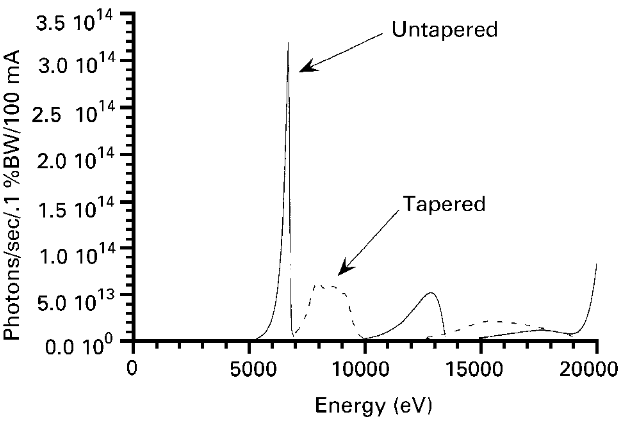



**Figure 4** Integrated spectral flux for Advanced Photon Source (APS) undulator. The position of the peaks is adjustable by varying the undulator magnetic gap.

## Optics

### Monochromators

The desired energy bandwidth of approximately 1 eV is selected by allowing the beam to impinge at a selected angle θ onto a cooled single crystal of silicon, germanium, diamond, or other substance. The crystals reside on a goniometer inside a vacuum chamber or inert atmosphere to minimize ozone production and absorption, and the angle θ can be remotely scanned by a computer. X-rays that meet the Bragg diffraction condition $n\lambda = 2\, d_{hkl} \sin(\theta)$ are diffracted through an angle 2θ; the rest are absorbed. In this equation, λ is the X-ray wavelength, which is related to the photon energy $\varepsilon = hc/\lambda$; $n$ is the harmonic number, and the spacing between diffracting atomic planes in the crystal for 'reflection' $hkl$ is $d_{hkl} = a_0/(h^2+k^2+l^2)^{1/2}$, where $a_0$ is the lattice constant (0.5431 nm for Si). The crystals used are sufficiently perfect that they are well described by dynamical diffraction theory instead of kinematic theory. Some of the lower index allowed reflections are 111, 220, 311, 400, 331, 422, 333, 511, 440, and 531. Higher index crystals are used to obtain better energy resolution but at the expense of lower integrated reflectivity.

Normally, a parallel second crystal is placed after the first crystal to deflect the beam in a direction parallel to the incident beam direction, so that the X-ray beam angle is maintained constant as the energy is scanned. The first and second crystal faces can be formed from the same piece of silicon by cutting a channel in it, making a so-called channel-cut configuration. Alternatively, separate crystals can be mounted independently of each other. In such a double crystal Bragg monochromator, the beam is displaced by a distance $2H\cos(\theta)$ from the height of the incident beam, where $H$ is the perpendicular separation of the crystals; consequently the beam height varies with energy, typically by less than a millimetre. The beam motion can be tracked by moving the sample and detectors under computer control; alternatively, in appropriately constructed monochromators, $H$ can be adjusted to preserve a fixed beam height. A translation of the second crystal parallel to the first to keep the diffracted beam centred on the second crystal also may be beneficial. A fine adjustment of the relative orientation of first and second crystal is essential; this is accomplished with piezoelectric transducers or highly gear-reduced motors.

The 'rocking curve' – the reflectivity of the crystal versus θ for monochromatic X-rays – is approximately rectangular in shape, with a typical (energy dependent) width of 3–10 arc-seconds (25–50 microradians). The rocking curves have small, but long-range, tails that can degrade the energy resolution, and may distort X-ray absorption spectra in the near-edge region where there may be rapid changes in absorption with energy. The contribution of these tails can be reduced by using a second pair of crystals following the first pair, but at the cost of considerably greater instrumental complexity.

Focusing in the sagittal (usually horizontal) direction can be accomplished by bending the second crystal to an appropriate radius $R$, given by $2 \sin(\theta)/R = (1/u+1/v)$, where $u$ is the source to optic distance, and $v$ is the optic to focal point distance. Substantial vertical (meridional) focusing cannot be achieved by bending the second crystal because doing so would cause the incidence angle of the beam on parts of the second crystal to fall outside of the rocking curve.

The high power density produced by undulator beams presents challenges for monochromator designers. The heat deposited in the first crystal creates a 'thermal bump' in the first crystal because the local heating creates thermal expansion in the silicon that degrades the rocking curves, and hence the resolution and throughput. The relevant parameter is the ratio of the thermal conductivity to the thermal expansion coefficient. Several approaches have been devised to deal with this problem. One is to cool the silicon with liquid nitrogen to a temperature around 100 K, which is beneficial because the thermal expansion coefficient is greatly reduced, and the thermal conductivity is increased. Another approach is to cut the crystal so the diffracting planes are at an inclined angle relative to the crystal face, so that the beam is spread over a larger area of the crystal. A third approach is to use diamond crystals instead of silicon, because the thermal conductivity of diamonds greatly exceeds that of silicon. A fourth approach is to use an X-ray mirror or synthetic multilayer as a pre-filter to reduce the power load on the first crystal.

### Mirrors

X-ray mirrors are used for rejecting harmonics, focusing, power filtering, displacing the beam, and improving collimation of the beam incident on the first crystal in order to improve energy resolution.

At small angles of incidence (on the order of milliradians), X-rays are totally externally reflected from the surfaces of materials. This effect is the X-ray analogue of ordinary total *internal* reflection that is observed at visible wavelengths when looking from a dense medium into a less dense medium. For most materials, the index of refraction at X-ray energies is a complex number: $\tilde{n} = 1 - \delta - i\beta$, where $\delta = ne^2\lambda^2/$

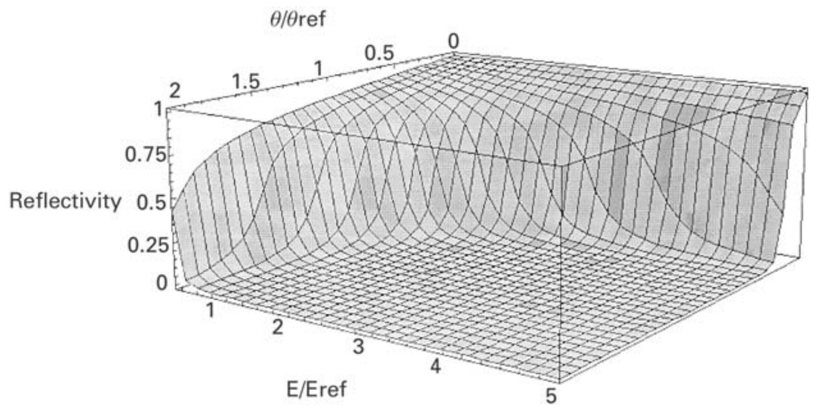



**Figure 5** Mirror reflectivity as a function of angle and energy. $\theta_{ref}$ is the critical angle at (arbitrary) energy $E_{ref}$.

$2\pi mc^2$ and $\beta = \mu\lambda / 4\pi$. The real and imaginary parts describe dispersion and absorption, and are connected through a Kramers–Kronig transform. Here $n$ is the number of dispersive electrons per volume of the material, $e$ and $m$ are the charge and mass of an electron, $\lambda$ is the wavelength, and $\mu$ is the X-ray absorption coefficient. For elemental materials this reduces to $\delta = N(Z/A)\rho e^2\lambda^2/2\pi mc^2$ where $N$ is Avogadro's number, $Z$ is the atomic number, $A$ is the atomic mass, and $\rho$ is the mass density.

Total external reflection occurs at angles $\theta < \theta_c$, where the 'critical angle' $\theta_c=(2\delta)^{1/2}$. $\theta_c$ is approximately inversely proportional to the X-ray energy, as shown in **Figure 5**. This allows the experimenter to eliminate harmonics by selecting an angle so that the fundamental is reflected from the mirror, but the harmonics are not.

The reflectivity from a mirror can be expressed as a function of the reduced angle $\phi = \theta/\theta_c$ as

$$R(\phi) = \frac{(\phi - A)^2 + B^2}{(\phi + A)^2 + B^2}$$

where

$$2A^2 = \left[(\phi^2 - 1)^2 + \left(\frac{\beta}{\delta}\right)^2\right]^{1/2} + (\phi^2 - 1)$$

and

$$2B^2 = \left[(\phi^2 - 1)^2 + \left(\frac{\beta}{\delta}\right)^2\right]^{1/2} - (\phi^2 - 1)$$

For a given angle, materials or coatings of high atomic number have larger critical energies, which allows them to reflect X-rays at higher energies. The disadvantage of using high $Z$ coatings are lower reflectivity and a less sharp cutoff of reflectivity against energy. This is a consequence of absorption in the material, the effect of which is shown in **Figure 6**. The product of energy and $\theta_c$ is an intrinsic property of the material coating: representative measured values (in KeV mrad, ±2%) are Si (31), Ni (59), Pd (62), Rh (67), Pt (82), Au (80).

X-ray mirrors are fabricated from polished glass (float glass, ultra low thermal expansion titanium silicate), silicon, silicon carbide, appropriate ceramics, or metal substrates. Typically they are tens to hundreds of centimeters in length in order to accept the vertical divergence of the beam. Surface roughness degrades the reflectivity, and accordingly the best mirrors are highly polished to as little as ~0.2 nm root mean square roughness. With present technology, ångstrom-level roughness and microradian rms slope errors are achievable over mirror lengths of more than 1 m. Meridional focusing in the vertical direction can be accomplished by bending a short mirror to an appropriate radius, given by $2/(R\theta) = (1/u+1/v)$, where $u$ is the source to optic distance, and $v$ is the optic to focal point

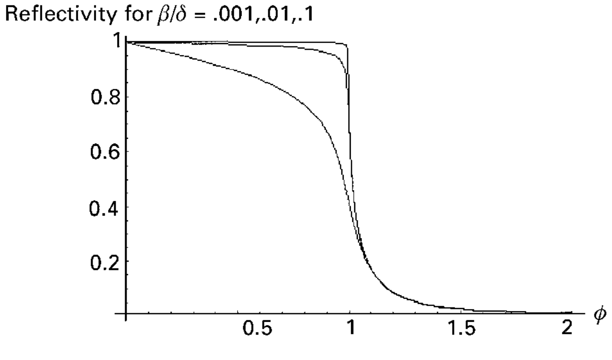



**Figure 6** Effect of X-ray absorption from mirror coatings on reflectivity. See text for explanation.

distance. Longer mirrors can be used provided the local radius of curvature at each point on the mirror satisfies this focusing equation.

## Detectors

Ionization chambers are the most commonly used detectors for X-ray absorption spectroscopy. They have the virtue of being partially transparent so that the incident beam intensity can be monitored. Typically they consist of a pair of parallel conducting plates several cm in each dimension and separation approximately 1 cm, inside a gas-tight housing in which is placed an appropriate gas (e.g. helium, nitrogen, argon, krypton). The gas can be flowed through the chamber, or sealed inside under positive, negative or ambient pressure. A constant potential of hundreds to thousands of volts is applied between the plates using a high voltage power supply or battery, and the small current flowing between the plates through the gas between them is measured. The X-ray beam is allowed to pass between the plates, which ionizes the gas molecules, rendering the 'fill-gas' partially conductive. The current (microamperes to nanoamperes) is proportional to the photon flux, given approximately as $N\ (E/E_{cc})\ (1-\exp(-\mu_{fg}(E)l))$, where $N$ is the number of photons per second of energy $E$, $E_{cc}$ is the mean energy required to produce a charge carrier in the gas (typically around 32 eV), $\mu_{fg}$ is the absorption coefficient of the fill gas (or mixture), and $l$ is the active length of the plates. The current is amplified with a transconductance (current to voltage) amplifier, and read by a computer with an analogue to digital converter; or it is converted to pulses with a voltage to frequency (V/F) converter and counted by a scaler.

Absorption within an ionization chamber is controllable by selection of fill-gas composition and pressure. For reliable absorption measurements, ionization chambers must be carefully constructed and operated at sufficiently high bias voltage that they are linear, i.e. the output current is linearly related to the absorbed photon flux. Under these conditions the chamber is said to be in the 'plateau region' of the flux versus bias voltage curve.

PIN diodes (positive-intrinsic-negative) are semiconductor devices that act essentially as solid state ion chambers. X-rays absorbed by the diodes create electron hole pairs that act as charge carriers. The electric field acting to separate them in the intrinsic (undoped) region is produced by the adjacent positively and negatively doped regions. The charge collected is amplified in the same manner as in an ionization chamber. A bias voltage can also be applied to alter the operational characteristics. PIN diodes are capable of excellent linearity but the X-ray thickness is not as easily experimentally controllable as it is for ionization chambers.

Ionization chambers with X-ray transparent plates made of aluminized plastic or thin metallic mesh are used for detection of fluorescence radiation. Detection of the fluorescence from dilute species requires a means of rejecting elastically scattered background from the sample. X-ray filters and slits used with ionization chambers are a standard method of rejecting background. Alternatively, arrays of solid state detector elements with appropriate electronics can provide useful energy resolution and adequate count rates for many purposes. Eliminating the scattered radiation with synthetic multilayer or crystal analysers is a promising approach for third generation synchrotron radiation sources.

## List of symbols

$E$ = particle beam energy; $I(I_0)$ = transmitted (incident) X-ray intensities; $K$ = deflection parameter; $u$ = source to optic distance; $v$ = optic to focal point distance; $\gamma = E/mc^2$; $\varepsilon$ = X-ray photon energy; $\varepsilon_c$ = critical energy; $\theta$ = angle of incidence on an optical surface; $\lambda_0$ = undulator period; $\lambda$ = photon wavelength; $\rho(\rho_0)$ = bend radius (bend radius at peak magnetic field).

*See also:* **Light Sources and Optics; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Creagh DC and Hubbel JH (1987) Problems associated with the measurement of X-ray attenuation coefficients. I. Silicon. Report on the International Union of Crystallography X-ray Attenuation Project. *Acta Crystallographica* **A43**: 102–112.

Heald SM (1988) EXAFS with synchrotron radiation. In: Koningsberger D and Prins R (eds) *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS and XANES*. New York: John Wiley.

Koningsberger DC (1988) Laboratory EXAFS facilities. In: Koningsberger D and Prins R (eds) *X-ray Absorption: Principles, Applications, Techniques of EXAFS. SEXAFS and XANES*. New York: Wiley.

Knoll G (1989) *Radiation Detection and Measurement* 2nd edn. New York: Wiley.

Krinsky S (1983) Characteristics of synchrotron radiation and its sources. In: Koch EE (ed) *Handbook on Synchrotron Radiation*. Amsterdam: North-Holland.

Matsushita S and Hashizume H (1983) X-ray monochromators. In: Koch, EE (ed) *Handbook on Synchrotron Radiation*. Amsterdam: North-Holland.

# X-Ray Emission Spectroscopy, Applications

**George N Dolenko**, Lermontova 35A/16, 664033 Irkutsk, Russia
**Oleg Kh Poleshchuk**, Tomsk Pedagogical University, Tomsk, Russia
**Jolanta N Latośińska**, Adam Mickiewicz University, Poznań, Poland

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Applications**

## MO structure investigation

Any variations in the composition, structure, stereochemistry or coordination character of a molecule change its chemical properties and MO (molecular orbital) structure. MO changes are clearly observed by the fine structure of XFS (X-ray fluorescence spectroscopy). This makes it possible to relate some features of the chemical behaviour of compounds to their electronic structure and opens a way to various 'chemical property–electronic structure parameter' correlations which are frequently of help for explaining and predicting the chemical properties of compounds.

The transitions from valence atomic levels to vacancies in inner shells form X-ray valence emission lines, reflecting the structure of the valence levels or zones. Electron transitions between different inner levels form inner X-ray emission lines. The study of the fine structure of different valence emission lines of all the atoms in a compound allows detailed investigation of the structure of valence levels or zones. Research into the shifts of inner X-ray emission lines allows one to investigate effective charges on the corresponding atoms.

For example, consider the X-ray emission spectra of sulfur. The initial state of a sulfur atom for X-ray emission is that with a vacancy in the K or $L_{2,3}$ level. This vacancy is rapidly filled (within $10^{-16}-10^{-14}$ s) as a result of transitions obeying the dipole selection rules, i.e. $2p \rightarrow 1s$ ($K\alpha$ lines), $3p \rightarrow 1s$ ($K\beta$ lines) or $3s \rightarrow 2p$, $3d \rightarrow 2p$ ($L_{2,3}$ lines) transitions. The energy released in this case is emitted from the atom as either an Auger electron or an X-ray quantum (**Figure 1**).

Whereas $SK\alpha$ are inner lines, $SK\beta$ and $SL_{2,3}$ are valence lines. The energies of atomic $np \rightarrow 1s$ transitions can be represented by the equations

$$h\nu = E_{\text{fin}} - E_{\text{init}} = E^{-\text{np}} - E^{-1\text{s}} \quad [1]$$

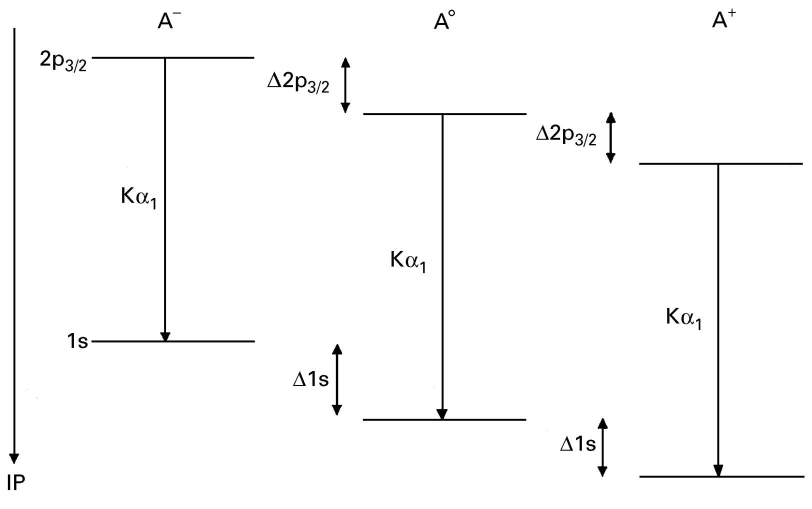



**Figure 1** Scheme showing the change of one-electron energies of 1s, 2p levels and the energy of the A $K\alpha$ line in the ions $A^+$ and $A^-$ with respect to the neutral atom $A^0$.

$$\Delta h\nu = \Delta E^{-\text{np}} - \Delta E^{-1\text{s}} = \Delta E^{-\text{np}} \approx \Delta(E^0 - \varepsilon_{\text{np}}) = -\Delta\varepsilon_{\text{np}} \quad [2]$$

where $h\nu$ is the energy of the emission quantum, $E_{\text{fin}}$ is the final energy of the system, $E_{\text{init}}$ is the initial energy of the system, $E^{-\text{nl}}$ is the energy of the system with an nl electron removed and $\varepsilon_{\text{np}}$ is the one-electron energy of the np level.

Thus, in a one-electron approximation the distance between individual maxima in a spectral series is equal to the difference in one-electron energies of the corresponding atomic levels. In molecules the 3s, 3p and 3d electrons of the sulfur atom are involved in chemical bonding to form an MO system. In this case, the SKβ spectrum (S3p → S1s interatomic transitions), for example, corresponds to $\text{MO}_i \rightarrow \text{S 1s}$ transitions, and the distances between spectral maxima correspond to energy differences of the appropriate occupied molecular levels:

$$\Delta E_{ij}(\text{SK}\beta) = \varepsilon(\text{MO}_j) - \varepsilon(\text{MO}_i) \quad [3]$$

The intensity of X-ray emission lines is determined by the relation (for the SK series as an example):

$$I(\text{Snp} \rightarrow \text{S1s}) \sim N_{\text{Snp}} E_1^3 \left|\langle \Phi_{\text{Snp}} | r | \Phi_{\text{S1s}} \rangle\right|^2 \quad [4]$$

where $N_{\text{Snp}}$ is the np level population of the sulfur atom, $E_1$ is the energy of Snp → S1s transitions, $\Phi_{\text{Snp}}$ is the wavefunction of sulfur np orbitals and $\Phi_{\text{S1s}}$ is the wavefunction of the sulfur 1s orbital.

For molecules this expression is transformed to the equation

$$I(\text{MO}_i \rightarrow \text{S1s}) \sim N_i E_1^3 \left|\langle \Psi^i | r | \Phi_{\text{S1s}} \rangle\right|^2 \quad [5]$$

where $\Psi^i$ is the wavefunction of *i*th MO.

In the generally accepted MO LCAO model the wavefunction of any *i*th MO can be represented as an AO sum:

$$\Psi^i = \sum_j c_j^i \Phi_j \quad [6]$$

where $\Phi_j$ is the wavefunction of *j*th AO and $c^i_j$ are the coefficients.

Then, with account taken of the dipole selection rules, Equation [5] is transformed into

$$I(\text{MO}_i \rightarrow \text{S1s}) \sim N_i E_1^3 \left| \left\langle \sum_j c_j^i \Phi_j | r | \Phi_{\text{S1s}} \right\rangle \right|^2 = N_i E_1^3 (c^i)^2 \left|\langle \Phi_{\text{S3p}} | r | \Phi_{\text{S1s}} \rangle\right|^2 \quad [7]$$

where $c^i$ is the coefficient of $\Psi^i$ at the wavefunction of S3p AO ($\Phi_{\text{S3p}}$). Thus, the ratio of partial intensities in the SKβ spectrum of a molecule is equal to that of the squares of the coefficients of the S3p wavefunction in the expansion of the corresponding MOs in terms of AOs:

$$I_i : I_j : I_k = (c^i)^2 : (c^j)^2 : (c^k)^2 \quad [8]$$

The above is also true for other X-ray emission series. Important features of X-ray emission spectra are the comparative ease of interpretation and the possibility of investigating the electronic structure from the 'viewpoint' of any atom of the molecule investigated.

#### Example: electronic structure of the sulfate ion

Information concerning the electronic structure of molecules provided by XFS can be well illustrated with the sulfate ion as an example. The wavefunction of any *i*th valent MO of the sulfate ion can be described by the equation

$$\Psi^i = c^i{}_1\Phi_{\text{S3s}} + c^i{}_2\Phi_{\text{S3p}} + c^i{}_3\Phi_{\text{S3d}} + c^i{}_4\Phi_{\text{O2s}} + c^i{}_5\Phi_{\text{O2p}} \quad [9]$$

All possible X-ray fluorescence spectra of the sulfate ion are presented in **Figure 2** whereas **Figure 3** shows an MO diagram constructed from a full set of these spectra. 'Adjustment' to the scale of the ionization potentials [IP] of valence electrons is effected by subtracting the X-ray transition energies from the IPs of the corresponding inner levels (S1s, S2p, O1s) determined by the use of data of X-ray photoelectron spectroscopy:

$$E_i(\text{SK}\beta) = \text{IP(S1s)} - \text{IP}(\text{MO}_i) \quad [10]$$

$$E_j(\text{SL}_{2,3}) = \text{IP(S2p)} - \text{IP}(\text{MO}_j) \quad [11]$$

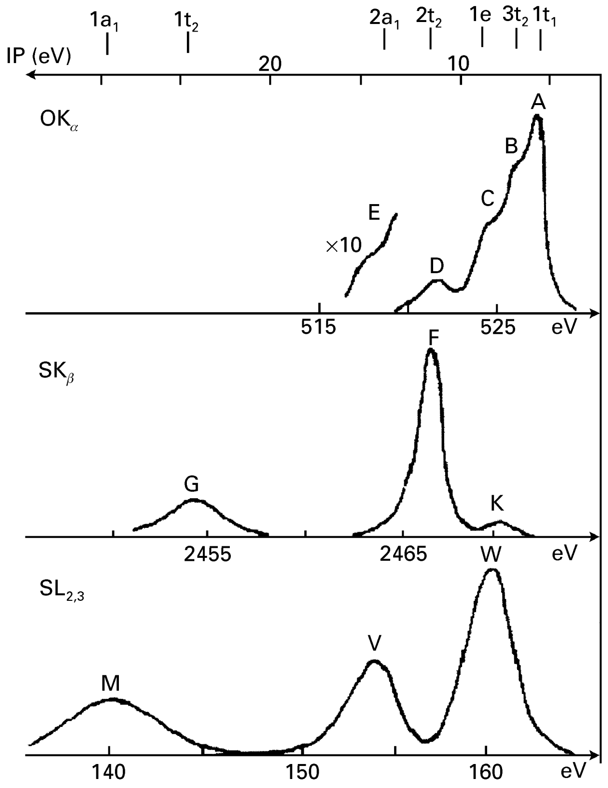



**Figure 2** Full set of X-ray fluorescence spectra of the sulfate ion on an energy scale corresponding to the ionization potentials of valence electrons.

$$E_k(\mathrm{OK}\alpha) = \mathrm{IP(O1s)} - \mathrm{IP}(\mathrm{MO}_k) \qquad [12]$$

From these, the following equations are derived

$$\begin{aligned}\mathrm{IP}(\mathrm{MO}_i) &= \mathrm{IP(S1s)} - E_i(\mathrm{SK\beta})\\ &= \mathrm{IP(S2p)} + E(\mathrm{SK}\alpha) - E_i(\mathrm{SK\beta})\end{aligned} \qquad [13]$$

$$\mathrm{IP}(\mathrm{MO}_j) = \mathrm{IP(S2p)} - E_j(\mathrm{SL}_{2,3}) \qquad [14]$$

$$\mathrm{IP}(\mathrm{MO}_k) = \mathrm{IP(O1s)} - E_k(\mathrm{OK}\alpha) \qquad [15]$$

All MOs with $c_5 \neq 0$ are displayed in the O Kα spectrum (O2p → O1s transitions), those with $c_2 \neq 0$ in the SKβ spectrum (S3p → S1s transitions) and those with $c_1 \neq 0$ and $c_3 \neq 0$ in the $SL_{2,3}$ spectrum (S3s → S2p and S3d → S2p transitions). From the spectra shown in **Figure 3** it follows that the highest occupied MO $1t_1$ (maximum A in the OKα spectrum) consists of only O2p electrons; next, the $3t_2$ and 1e levels (maxima B, C and W) significantly correspond to the $\pi$ bond S3d – O2p; the $2t_2$ level then follows (maxima D and F), which corresponds to the strong S3p – O2p $\sigma$ bond. The $2a_1$ level (maxima E and V), consisting of the S3s AO and, possibly, the O2s with a small admixture of O2p AO, lies even deeper. Deep $1t_2$ and $1a_1$ MOs consisting mainly of the O2s AO are seen as low intensity long-wavelength maxima G and M respectively.

Consequently, much information on the MO structure of a chemical species under investigation can be derived from X-ray emission spectra.

## Determination of the effective atomic charges

The concept of the effective charge on atoms in molecules is known to be fundamental in the field of theoretical chemistry. It assumes that the entire electron distribution of an investigated atom can be considered as a point charge that coincides with the coordinates of the nucleus. This simple and obvious form of the electron density distribution in the examined species is rather approximate: in real molecules the outer electron shells of atoms substantially lose their 'individuality' because of the delocalization of valence electrons of atoms. The approximation procedure (the replacement of the real distribution of the outer electron density of an atom by the point charge on its nucleus) is not simple in general and depends largely on the actual definition of effective charge. A number of calculations and experimental methods used for the determination of the latter are known. The effective charges on atoms ($q_A$, where A is an element) do not belong to the class of directly observed physical characteristics, and therefore the so-called 'experimental determination' of $q_A$ values usually means the result of the interpretation of various experimental data in terms of the corresponding model with $q_A$ as a parameter.

Shifts of the energy of inner nl levels of the A atom (ΔAnl), defined by X-ray photoelectron spectroscopy, are sensitive to $q_A$ and the effective charges of all other atoms, and in terms of the so-called potential model can be written as

$$\Delta\mathrm{Anl} = k(\mathrm{Anl})q_\mathrm{A} + \sum_{i\neq \mathrm{A}} \frac{q_i}{r_{i\mathrm{A}}} + \Delta E_\mathrm{r} \qquad [16]$$

where K(Anl) is the coefficient that is characteristic of the A nl level, $\Sigma_{i\neq A}\,(q_i/r_{iA})$ is the Madelung potential and $E_r$ is the relaxation energy. Shifts of inner $AK\alpha_1$ line ($2p_{3/2}$ → 1s electron transitions are more intense than for $2p_{1/2}$ → 1s) are determined by the difference between the shifts of one-electron energies

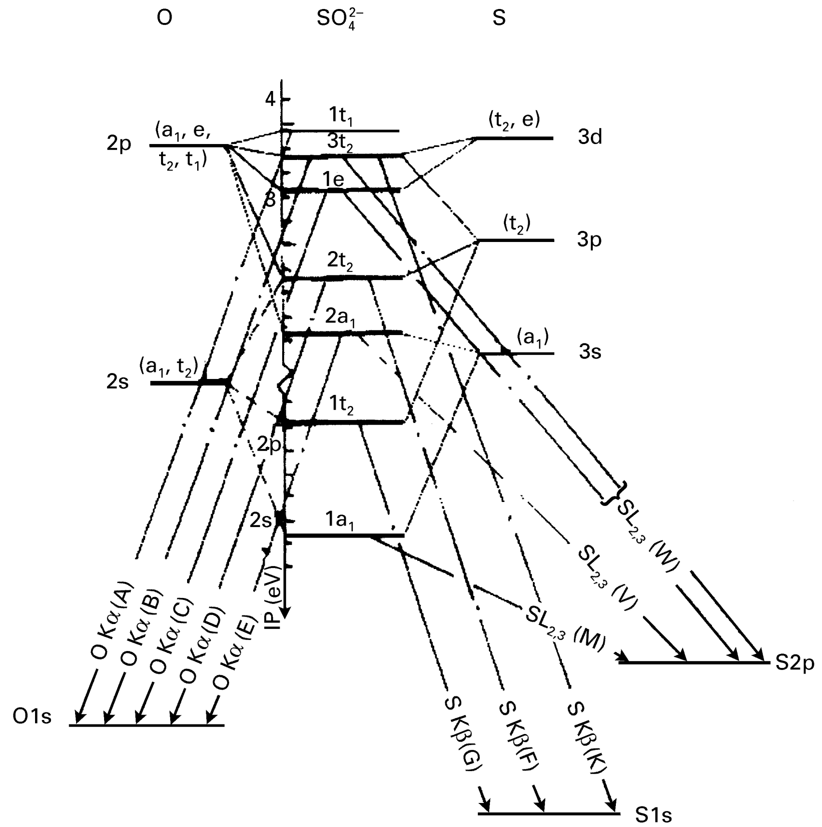



**Figure 3** Scheme of the sulfate ion MOs obtained from the full set of X-ray fluorescence spectra.

of A1s and $A2p_{3/2}$ levels:

$$\Delta AK\alpha_{3/2} = (E^{-2p} - E^{-1s}) - (E_0^{-2p} - E_0^{-1s})$$
$$= \Delta E^{-2p} - \Delta E^{-1s} \approx \Delta A1s - \Delta A2p_{3/2} \qquad [17]$$

where $E_0^{-nl}$ is the total energy of the neutral A atom ($A^0$) containing a vacancy in the nl level.

The difference in the energy changes for different Anl levels is rather small and therefore the $\Delta AK\alpha_1$ values are determined by subtraction of two large and very similar quantities, ΔA1s and $\Delta A2p_{3/2}$ (the energy shift of the A1s level always prevails over that of $A2p_{3/2}$). As a result, the $\Delta AK\alpha_1$ values are normally about ten times smaller than X-ray photoelectron Anl shifts. However, $AK\alpha_1$ shifts have an obvious advantage. The corresponding electron transitions are localized in the same potential well as that created by the Coulomb field of other atoms of the system investigated. Hence, the potential model for AKα shifts, analogous to Equation [16], can be written as

$$\Delta AK\alpha_1 = k_A q_A \qquad [18]$$

or, in a more universal form

$$\Delta AK\alpha_1 = f(q_A) \qquad [19]$$

Many authors have investigated the dependencies of Equation [19] for different inner X-ray emission lines and atoms by different methods. **Table 1** gives the $AK\alpha_1$ shifts for some phosphorus-, sulfur- and chlorine-containing compounds and $q_A$ values obtained by the correlation of Mulliken charges calculated by the SCF *ab initio* method using a 4-31G** basis set for a sufficiently large series of A-containing molecules with the experimental $AK\alpha_1$ shifts.

## Participation of the 3d atomic orbitals in L-emission spectra

From the X-ray L-emission spectra of S, Cl, P, within the framework of a MO method, one can estimate the 3d-population. The basic problem is the need to calculate the matrix elements of the transitions 2p → 3s and 2p → 3d. For Kβ spectra of elements in period 3 in the dipole approximation, transitions are permitted from levels involving 3pAO, from which

**Table 1** Experimental $AK\alpha_1$ shift and $q_A$ values

| *A* | *Class of compounds* | $\Delta AK\alpha_1$ (eV) *relative to* $P_{red}$, $S_8$, $Cl_2$ | $q_A(e)$ *in 4-31G** charge scale* |
|---|---|---|---|
| P | $R_3P$ | –0.01 – 0.25 | –0.03 – 0.25 |
| | $R_4P^+$ | 0.16 – 0.38 | 0.32 – 0.76 |
| | $P(OR)_3$ | 0.48 – 0.52 | 0.97 – 1.05 |
| | $R_3PO$ | 0.27 – 0.57 | 0.54 – 1.15 |
| | $R_3PS$ | 0.17 – 0.24 | 0.34 – 0.48 |
| | $(RO)_3PO$ | 0.69 – 0.76 | 1.39 – 1.53 |
| | $PO_4^{3-}$ | 0.76 – 0.80 | 1.54 – 1.61 |
| S | =S | –0.14 – –0.02 | –0.22 – 0.09 |
| | RSH | –0.08 – 0.00 | –0.11 – 0.05 |
| | $R_2S$ | –0.09 – 0.14 | –0.18 – –0.32 |
| | $R_3S^+$ | 0.00 – 0.12 | 0.05 – 0.28 |
| | RNSNR′ | 0.18 – 0.25 | 0.40 – 0.54 |
| | $R_2SO$ | 0.36 – 0.42 | 0.75 – 0.87 |
| | $R_2SO_2$ | 0.78 – 0.85 | 1.57 – 1.61 |
| | $SO_4^{2-}$ | 1.00 – 1.20 | 2.00 – 2.40 |
| Cl | RCl | –0.19 – 0.01 | –0.28 – 0.02 |

the MO are constructed. The relative intensity of separate emission lines is proportional to the squares of the factors $c_i^2$. In L-emission spectra of atoms in period 3 the MO coefficients owing to dipole rules of selection ($\Delta l = \pm 1$), will be simultaneously displayed MO, which are constructed with participation of 3s AO and 3d AO. If the symmetry of a molecule is such that the MO of the systems can be constructed with simultaneous participation of 3s- and 3d AOs of a period 3 atom the intensity of the emission line will depend on the contributions of both the 3d and the 3s AOs to the appropriate MO. Thus estimations of $c_i^2$ values with 3d AO and 3s AO from X-ray spectra need a knowledge of matrix elements of transitions $|\langle 2p | r | 3d \rangle|$, $|\langle 2d | r | 3s \rangle|$. The determination of such matrix elements becomes complicated by the problem of choosing a 'good' 3d wavefunction. It is known that the atomic 3d functions are too diffuse and their electronic density, appropriate to them, is located far from the nuclei of atoms and to be unsuitable for participation in chemical bonds. In the case of X-ray transitions the important behaviour of 3d wavefunctions is not in the area of valence electrons but in the field of core electrons of atoms, in particular in the area of 2p AO. The 2p AO wavefunctions are located near the nucleus of an atom and have atomic character. It is possible to consider that 3d AO in this area also has atomic character. On this basis the estimation of matrix elements $|\langle 2p | r | 3d \rangle|$ and $|\langle 2p | r | 3s \rangle|$ is carried out to account for the intensity of X-ray atomic transitions.

The analysis of spectra of molecules and ions shows that the short-wave maximum W in L-spectra of S and Cl (**Figure 2**) is basically connected with an MO, in which there is a significant contribution from the 3d AO, while the contributions of the 3s AO to these MOs are insignificant. The maxima V and M are connected with an MO in which the 3s AO participates. Hence, from L-spectra it is possible to obtain experimental values of the relative intensity of various lines and to determine the contribution of AO and MO.

Estimations for the ion $SO_4^{2-}$ and the molecule $SF_6$ give $I_{3d}/I_{3s}$ values that correspond to theoretical results. The consideration of results from sulfur- and chlorine-containing compounds indicate that the participation of 3d AO in various MO becomes appreciable. The study of shifts to Kα lines over a range of molecules shows that the experimental relation I 3d/I 3s increases as the Kα shift grows, physically this is connected to the growth of a positive charge on an atom of sulfur. **Table 2** gives the relative experimental 3d occupations for sulfur in some compounds.

## Application of SKα spectra to characteristic compounds

It is known that the energy of valence electrons of heteroatoms in periods 2 and 3 varies linearly with changes in the energy of their core electrons. The concept of an energy level of a hypothetical electron lone pair of a sulfur atom (hnS), whose energy depends only on the charge on the sulfur atom (qS), has been suggested. The position of this level in the SKβ spectrum was related to the values of the SKα shift, which are proportional to the net charge on the sulfur atom. The following relationship was

**Table 2** Relative 3d sulfur population from X-ray spectral data

| *Molecule* | $I_{3d}/I_{3s}$ | $\Delta E\,K\alpha_{1,2}$ (eV) | $q_s$ | $|\langle 2p | r | 3d \rangle| / |\langle 2p | r | 3s \rangle|$ |
|---|---|---|---|---|
| $(CH_3)_2SO$ | 0.2 | 0.25 | 1.0 | 0.1 |
| $Cl_2SO$ | 0.3 | 0.27 | 1.2 | 0.3 |
| $(C_6H_5)_2SO$ | 0.4 | 0.31 | 1.2 | 0.3 |
| $SO_2$ | 0.7 | 0.45 | 1.5 | 0.6 |
| $(CH_3)_2SO_2$ | 0.8 | 0.75 | 1.9 | 0.9 |
| $SF_4$ | 1.0 | 0.75 | 1.9 | 0.9 |
| $SO_3^{2-}$ | 1.0 | 0.65 | 1.6 | 0.7 |
| $(C_6H_5)_2SO_2$ | 1.2 | 0.82 | 2.1 | 1.1 |
| $SF_6$ | 1.0 | 1.45 | 2.6 | 1.4 |
| $SO_4^{2-}$ | 1.8 | 1.10 | 2.3 | 1.2 |

obtained by comparing the short-wave maximum in the SK spectra of saturated sulfides with the corresponding values of Δ(SKα): $hn_S(K\beta)(eV) = E(n_S \rightarrow 1S_S) = 0.0056 \times 10^3 \Delta SK\alpha\ (eV) + 2468.37$, with $r = 0.973$, $s = 0.06$, $n = 26$.

Using the equation it is possible, knowing ΔSKα values to predict the location of a hnS level in the SKβ spectra of any saturated sulfide compound. In fact, this level can be treated as a reference level in the analysis of the changes in the spectral structure caused only by orbital interactions devoid of the effect of charge changes on the sulfur atom. As an example one can consider the application for complex compounds with dimethyl sulfide.

It follows from **Figure 4** that the intensity of the short-wave maximum A, which in the SKβ spectra of sulfides corresponds to the transition from the nS level to the vacancy K of the sulfur atom, significantly decreases and is considerably shifted towards longer wavelength with respect to the hnS level (Kβ). This shift, ($\Delta n_S = E_A(SK\beta) - hn_S(SK\beta)$) (**Table 3**), characterizes quantitatively the bonding nature of the highest occupied molecular orbital. The observed shift towards longer wavelength indicates that the nS level interacts with the vacant levels of the acceptor, being mainly of the d type. The differences in shapes of the SKβ spectra of the complexes studied (**Figure 4**) can be explained by the presence of partly populated valence d orbitals, apart from the vacant ones, in Ti, in contrast to Sn and Sb.

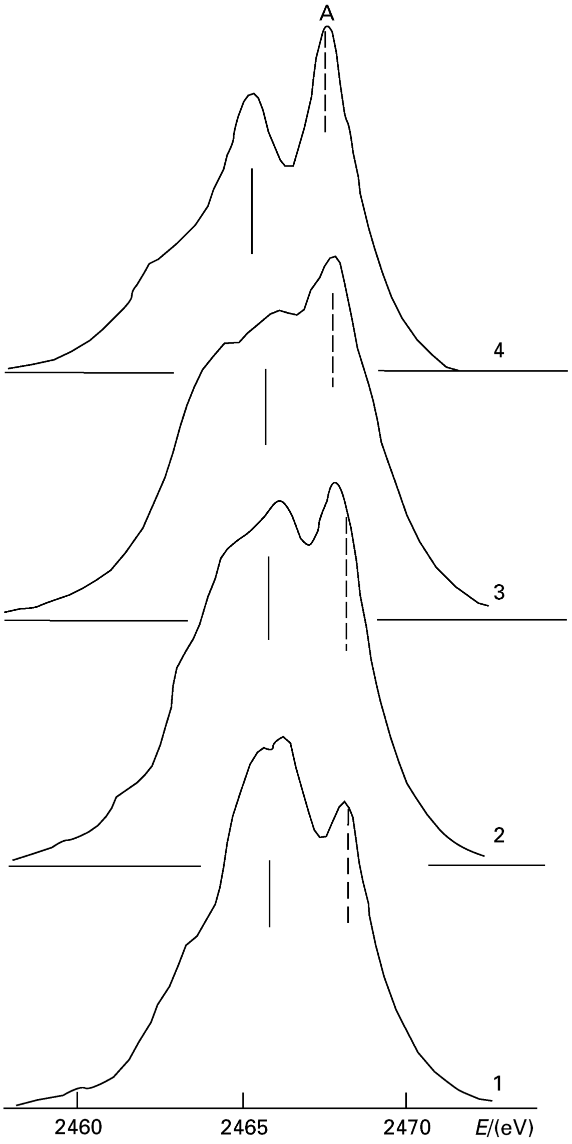



**Figure 4** SK$\beta$ spectra of some $Me_2S$ complexes. (—) centre of gravity; (---) $hn_S(K\beta)$. 1 = $TiCl_4 \cdot 2SMe_2$; 2 = $SbCl_5 \cdot SMe_2$; 3 = $SnCl_4 \cdot 2SMe_2$; 4 = $SMe_2$.

## Application of SK β and SKα spectra for rodano-group

It is known that the NCS-group in compounds can be coordinated with a metal in three ways: M–NCS (a), M–SCN (b) and M–NCS–M (c). Inorganic thiocyanates with coordination of type (a) have a characteristically large (negative) total electronic density on the sulfur atom, lower intensity of long-wave maxima and lower energy of a short-wave maximum (**Figure 5**). In organic isothiocyanates the ΔSKα values vary in an interval −7400–5800 eV with the intense short-wave maximum A in the SKβ spectra ($E_A$) in the range − 2467.1 – 2467.8 eV. In organic thiocyanates these parameters become ΔSKα = 4.7 – 13.3 eV $10^2$, $E_A = 2468.1 - 2469.1$ eV (**Table 4**).

**Table 3** Parameters determined from X-ray spectra of the sulfur atoms for some of the complexes studied

| *Compound* | *Δ(SKα)[a] $10^{-3}$*(eV) | *$q_s$ (e) in 4–31G** charge scale* | $hn_s$(Kβ) (eV) | $E_A$(Kβ) (eV) | $\Delta n_s$ (eV) |
|---|---|---|---|---|---|
| $(CH_3)_2S$ | −63(6)[b] | −0.07(2) | 8.02(4) | 8.1(1) | −0.1(1) |
| $SnCl_4 \cdot 2S(CH_3)_2$ | 2(14) | 0.05(2) | 8.38(8) | 8.2(1) | 0.2(1) |
| $SbCl_5 \cdot S(CH_3)_2$ | 17(7) | 0.08(2) | 8.47(5) | 8.3(1) | 0.2(1) |
| $TiCl_4 \cdot 2S(CH_3)_2$ | −3(10) | 0.04(2) | 8.35(6) | 8.11(5) | 0.24(8) |

[a] Relative to $S_8$.
[b] The mean-square errors in the last significant digit, taken for 95% confidence interval by Student's criterion, are given in parentheses.

**Table 4** X-ray spectral character of some thiocyanates and isothiocyanates

| *No.* | *Compound* | Δ(SKα)* $10^{-2}$ (eV) | $q_s$ *(e) in 4–31G** charge scale* | $E_A$(Kβ) (eV) | $hn_s$ (Kβ) (eV) *relative to 2467.0 eV* | $\Delta n_s$ (eV) | *Coordination type* |
|---|---|---|---|---|---|---|---|
| (I) | KNCS | –4.5(12) | –0.04(3) | 0.2 | 1.1 | –0.9 | a |
| (II) | $NH_4NCS$ | –9.2(6) | –0.13(2) | 0.3 | 0.9 | –0.6 | a |
| (III) | $Ba(NCS)_2$ | –3.6(10) | –0.02(2) | 0.2 | 1.2 | –1.0 | a |
| (IV) | $Sn(NCS)_2$ | –7.2(8) | –0.09(2) | 0.3 | 1.0 | –0.7 | a |
| (V) | CuSCN | 4.1(8) | 0.13(2) | 2.0 | 1.6 | 0.4 | b |
| (VI) | $CH_3SCN$ | 4.7(8) | 0.14(2) | 1.8 | 1.7 | 0.1 | b |
| (VII) | $C_6H_5CH_2SCN$ | 5.4(11) | 0.16(3) | 1.6 | 1.7 | –0.1 | b |
| (VIII) | $C_6F_5SCN$ | 9.4(8) | 0.23(2) | 1.9 | 1.9 | 0.0 | b |
| (IX) | $C_8H_4OPhSCN$ | 13.3(9) | 0.31(3) | 2.1 | 2.1 | 0.0 | b |
| (X) | $C_6H_5SCN$ | 5.0(8) | 0.15(2) | 1.7 | 1.7 | 0.0 | b |
| (XI) | $CH_3NCS$ | –4.7(8) | –0.04(2) | 0.5 | 1.1 | –0.6 | a |
| (XII) | $P(NCS)_3$ | 0(2) | 0.05(2) | 0.2 | 1.4 | –1.2 | a |
| (XIII) | $C_6F_5(NCS)_2$ | –7.4(6) | –0.09(2) | 0.1 | 1.0 | –0.9 | a |
| (XIV) | $C_8H_4OPhNCS$ | 5.8(9) | 0.16(2) | 0.8 | 1.7 | –0.9 | a |
| (XV) | $C_6H_5NCS$ | –5.3(7) | –0.05(2) | 0.3 | 1.1 | –0.8 | a |

The mean-square errors in the last significant digit, taken for 95% confidence interval by Student's criterion, are given in parentheses.
[a] Relative to $S_8$.

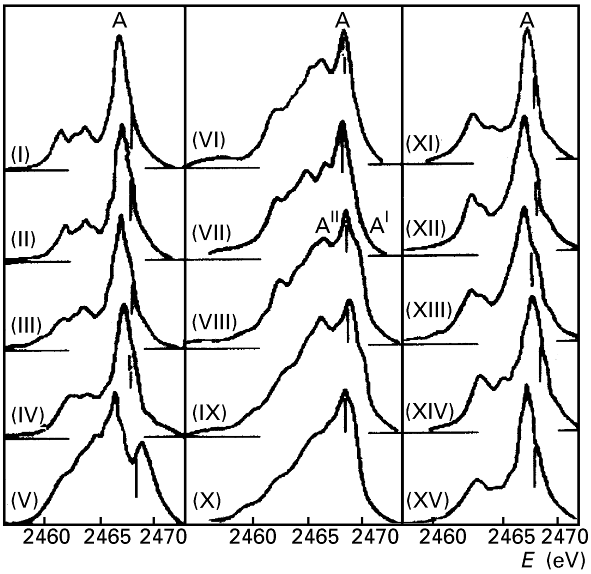



**Figure 5** SKβ spectra of some inorganic and organic thiocyanates and isothiocyanates. The compound numbers are defined in **Table 4**. The energy levels of the hypothetical lone electron pair are marked by the vertical lines.

However, the ΔnS values of thiocyanates and isothiocyanates are divided almost unequivocally: ΔnS ≥ 0 for thiocyanates, ΔnS < 0 for isothiocyanates (in thiocyanates $3p_{\pi}$– electronic density of atom of sulfur enters into a strongly bonding MO).

From the presence in the SKβ spectra of thiocyanates (VI), (VII) of a single short-wave maxima A that is coincident with the level hnS, it follows that the level nS does not practically couple with the $\pi_{C\equiv N}$ orbitals. The existence of the advanced short-wave structure in SKβ spectra thiocyanates (VIII)–(X), beyond the hnS level, must relate to $n_S$–$\pi_{Ar}$ interactions. Thus, the spectra of thiocyanates (VIII)–(X) indicates two orthogonal conformations of the molecules, one in which $n_S$–$\pi_{Ar}$ interactions are absent giving rise to a peak coincident with the hnS level, and another in which the $n_S$ and $\pi_{Ar}$ couple to give two maxima, A′ and A″, corresponding to levels $n_S \pm \pi_{Ar}$.

## Applications of Kα shifts for electronic density redistribution

It was of interest to use the data obtained for the investigation of the electron density redistribution on complexation between a donor molecule $PCl_3$ or of $SPCl_3$ (where one can define the effective charges on all atoms by their Kα shifts) and an acceptor molecule $AlBr_3$ containing no interfering atoms. The data presented in **Table 5** show that, in spite of a positive $q_P$ growth on complexation, the total ligand electron density does not decrease (in the range of accuracy achieved) due to a strong dampening effect of the Cl atoms. This leads to a sufficient growth of

**Table 5** Change of atomic charges and bond ionicities of free ligands on their complexation

| Compound | ΔA Kα[a](eV) | | | *Ionicities of bonds[b] (%)* | | *Change of effective atomic charge ($\delta q_A$)[c] of ligands on their complexation in 4-31G** charge scale* | | | |
|---|---|---|---|---|---|---|---|---|---|
| | A = P | A = S | A = Cl | P–S | P–Cl | A = P | A = S | A = Cl | Σ δqi |
| $PCl_3$ | 0.373(7) | – | –0.09(1) | – | 44(5) | | | | |
| $AlBr_3PCl_3$ | 0.599(8) | – | –0.156(15) | – | 72(7) | 0.5(1) | – | –0.10(7) | 0.2(2) |
| $SPCl_3$ | 0.430(8) | –0.10(2) | –0.11(2) | 50(6) | 51(6) | | | | |
| $AlBr_3SPCl_3$ | 0.616(8) | –0.080(9) | –0.171(14) | 68(7) | 74(8) | 0.4(1) | 0.04(6) | –0.1(1) | 0.1(2) |

[a] The mean-square errors in the last significant digit, taken for 95% confidence interval by Student's criterion, are given in parentheses.
[b] The ionicity of the AB bond is equal to ($|q_A – q_B|/2$)100%.
[c] The positive sign of $\delta q_A$ corresponds to a decrease of the A atom electron density on complex formation.

the ionicity of all bonds of the ligands and acceptor. From **Table 5** it also follows that the positive charge on the central acceptor atom grows sufficiently on complexation while the electron density on acceptor geminal of atoms does not decrease.

## List of symbols

$c^i$ = coefficient of $\Psi^i$; $E_0^{-nl}$ = total energy of neutral atom A with a vacancy at the nl level; $E_{init}$ = initial energy of system; $E_{fin}$ =final energy of system; $E_r$ = relaxation energy; $h\nu$ = energy of emission quantum; $I$ = intensity of X-ray emission lines; $N_{Snp}$ = np level population of sulfur atom; $q_A$ = effective charge on atom A; $\varepsilon_{np}$ = one-electron energy of np level; $\Phi_{S1s}$ = wavefunction of the sulfur 1s orbital; $\Psi^i$ = wavefunction of the *i*th MO.

*See also:* **X-Ray Emission Spectroscopy, Methods; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading

Dolenko GN (1993) X-ray determination of effective charges on sulphur, phosphorus and chlorine atoms. *Journal of Molecular Structure* **291**: 23–57.

Dolenko GN, Litvin AL, Elin VP and Poleshchuk OKh, (1991) X-ray investigation of electron density redistribution on complexation. *Journal of Molecular Structure* **251**: 11–27.

Dolenko GN, Latajka Z and Ratajczak H (1995) X-ray spectral determination of the effective charges on P, S, and Cl atoms in chemical compounds with a nonempirical charge scale. *Heteroatom Chemistry* **6**: 553–557.

Dolenko GN, Poleshchuk OKh and Koput J (1998) Antimonium pentachloride electron density redistribution on complexation. *Heteroatom Chemistry* **9**: 543–548.

Mazalov LN and Yumatov VD (1984) *Electronnoe stroenie ekstragentov*. Novosobirsk: Nauka, 199 p.

Nogaj B, Poleshchuk OKh, Kasprzak J, Koput J, ElinVP, Dolenko GN (1997)Changes in electron density distribution resulting from formation of antimony pentachloride complexes studied by X-ray fluorescence spectroscopy. *Journal of Molecular Structure* **406**: 145–151.

Poleshchuk OKh, Nogaj B, Dolenko GN and Elin VP (1993) Electron density redistribution on complexation in non-transition element complexes. *Journal of Molecular Structure* **297**: 295–312.

Poleshchuk OKh, Nogaj B, Kasprzak J, Koput J, Dolenko GN, Elin VP, Ivanovskii AL (1994) Investigation of the electronic structure of $SnCl_4L_2$, $TiCl_4L_2$ and $SbCl_5L$ complexes by X-ray fluorescence spectroscopy. *Journal of Molecular Structure* **324**: 215–222.

# X-Ray Emission Spectroscopy, Methods

**George N Dolenko**, Lermontova 325A/16, 664033 Irkutsk, Russia
**Oleg Kh Poleshchuk**, Tomsk Pedagogical University, Tomsk, Russia
**Jolanta N Latośińska**, Adam Mickiewicz University, Poznań, Poland

Copyright © 1999 Academic Press

HIGH ENERGY SPECTROSCOPY
Methods & Instrumentation

## General characteristics

Recently, intensive development of the theory of the electronic structure of chemical compounds has revealed a great need for physical experimental methods to which modern methods of quantum chemistry can be applied. Primary information provided by quantum chemical methods of electronic structure studies concerns the energy spectrum of molecules as well as the structures of wavefunctions of molecular orbitals (MOs). There is a need for physical methods which would allow direct measurement of MO energies as well as the determination of the degree of participation of different atomic electrons in chemical bonding. X-ray emission spectroscopy (XES), which provides both integral information about electronic structure (effective charges on atoms, atomic nl populations) and differential information (relative energies of occupied MOs and characteristics of their AO components), can be referred to these methods.

X-rays, discovered by Röentgen in 1895, are a form of electromagnetic radiation that occupy the spectral area between UV and γ radiation in the range of wavelength $\lambda = 10^{-3}$–$10^{2}$ nm, corresponding to energies $h\nu = 10$–$10^{6}$ eV ($\nu = c/\lambda$). X-ray spectroscopy is divided into X-ray emission and X-ray absorption spectroscopy and is subclassified into short wavelength ($\lambda \leq 0.2$ nm), long wavelength ($0.2 \leq \lambda \leq 2$ nm) and ultralong wavelength ($\lambda > 2$ nm) regions.

X-ray emission spectroscopy is used for the study of electronic structure and for the qualitative and quantitative analysis of substances. With the help of X-ray emission spectroscopy one may investigate all elements of the periodic table (with $Z > 2$) in compounds of any phase.

X-rays are divided into continuous and characteristic. Continuous X-rays occur as a result of the stopping of very fast charged particles (e.g. electrons) in the target substance and have a 'white' spectrum. Characteristic X-ray emission radiation is emitted by target atoms after their collisions with hot electrons (primary excitation) or with X-ray photons (secondary excitation, fluorescence radiation) and produces a line spectrum. These collisions may also remove an electron of any inner shell of a target atom. The resulting vacancy is filled by an electron transition from another inner or outer electron shell. As a consequence of this electron transition, energy is released which may be in the form of an X-ray quantum (**Figure 1**).

## Characteristic X-ray emission spectra

Characteristic X-ray emission spectra consist of spectral series (K, L, M, N...), whose lines have a common initial state with the vacancy in the inner level. Labels of basic X-ray transitions are shown in **Figure 2**. All electron levels with the principal quantum number *n* equal to 1, 2, 3, 4, etc. are named as K, L, M, N etc. levels and denoted with corresponding Greek letters and digit indexes. The electron transitions which satisfy the dipole selection rules

$$\Delta l = \pm 1; \Delta j = 0, 1\ (j = \pm \tfrac{1}{2}); \Delta n \neq 0 \qquad [1]$$

are most intense. The dependence of X-ray emission line energy on atomic number $Z$ is defined by Moseley's law:

$$h\nu \sim (Z - \sigma)^2 \qquad [2]$$

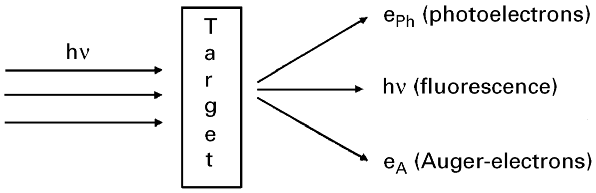



**Figure 1** Scheme of radiation interaction with a substance.

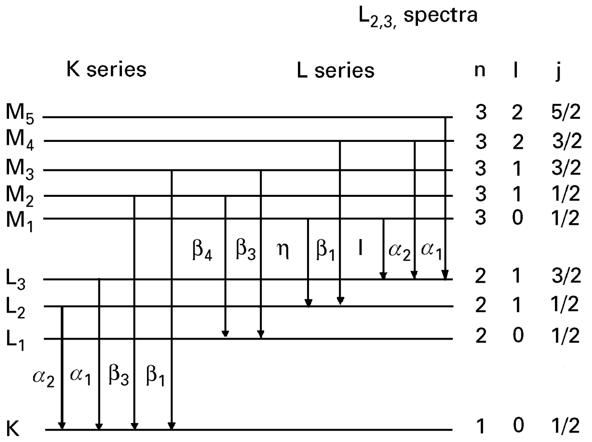



**Figure 2** Scheme of the most important X-ray emission transitions; *n*, *l* and *j* are correspondingly the principal, orbital and total quantum numbers of K, $L_1$, $L_2$, $L_3$ levels, etc.

where $Z$ is the atomic number and $\sigma$ is the shielding constant, which varies from series to series. Therefore, any X-ray emission spectral line is the fingerprint of an element.

With X-ray emission excitation by electron bombardment (primary emission) all emission lines of the *i*th series appear when the X-ray tube voltage $U$ exceeds the ionization potential of *i*th level ($V_i$). At higher $U$ the intensity of all lines of the *i*th series, $I_i$, increases because the electrons penetrate deeper into the target substance and, therefore, the number of excited atoms in the target increases. In the $V_i < U < 3V_i$ region, the intensity obeys the rule $I_i \sim (U - V_i)^2$. With a further increase in $U$ X-ray emission begins to be absorbed by the target atoms, therefore the increase in $I_i$ is reduced. At $U \geq 11V_i$, $I_i$ decreases because now most of the excited atoms are so deep in the target that their emitted radiation is absorbed by the target substance.

X-ray emission spectra are usually excited by X-ray photons because most chemical compounds are decomposed by electron bombardment. With X-ray emission spectra excitation by photons [secondary emission or fluorescence (XFS)] the fluorescent line intensity depends on the exciting photon energy $hv$ $I_i = 0$ if $h\nu <$ å$V_i$. All lines of the *i*th series appear if $h\nu =$ å$V_i$; however, $I_i$ decreases little with further increase in $hv$. Therefore, to excite X-ray fluoresence one must use a target that contains a substance with intense characteristic X-ray lines whose energy just exceeds $eV_i$. Using the continous radiation of an X-ray tube with a target consisting mostly of heavy elements it is possible to excite X-ray fluorescence.

The intensity of a characteristic X-ray spectrum (both primary and fluorescent) depends on the probability $p_r$ of a radiation transition in the atom having the vacancy in the *i*th level. The value of $p_r$ is determined by the total probability of photon emission when this vacancy is filled by outer electrons. However, with a probability $p_A$ the vacancy may be filled by outer electrons without radiation as the result of the Auger-effect (see **Figure 1**). For the K series of medium and heavy elements $p_r > p_A$, for the light elements $p_r < p_A$. For all others series of any elements $p_r << p_A$. The ratio $f = p_r/(p_r + p_A)$ is called the yield of characteristic radiation.

However, X-ray characteristic lines appear because of single atom ionization; in X-ray emission spectra weaker lines are found to occur as a result of binary (or multiple) atom ionization when two (or more) vacancies are formed simultaneously in different electron shells. If, for example, one vacancy is formed in the K shell of atoms and filled by electrons belonging to the $L_{2,3}$ shell, atoms emit an $\hat{E}\alpha_{1,2}$ doublet. If another vacancy is formed simultaneously which too is filled by electrons from the $L_{2,3}$ shell, then the final state will have a binary ionization $L_{2,3}L_{2,3}$, and would correspond to the emission of radiation with energy exceeding that of the $\hat{E}\alpha_{1,2}$ doublet. As a result, in an X-ray emission spectrum a short wavelength $\hat{E}\alpha_{3,4}$ doublet, called a satellite of the main $\hat{E}\alpha_{1,2}$ doublet, would appear. Because of such processes of multiple ionization X-ray emission spectra may have a large number of satellites of the main lines. Usually, the satellite intensity is some orders of magnitude less than that of the main lines. However, if target atoms are bombarded by heavy ions with great energy, the probability of multiple atom ionization becomes higher than that of single ionization. Therefore, in this case the intensity of the main emission lines is essentially less than that of the satellites.

## The continuous X-ray spectrum

The continuous X-ray radiation occurs because of electron deceleration in the target substance. Electron energy losses by radiation have quantum character, the emitted photon having an energy $hv$, which can not exceed the electron kinetic energy $\varepsilon$, i.e. $hv \leq \varepsilon$. The energy $hv_i = \varepsilon$ is called the quantum boundary of a continuous spectrum. The corresponding wavelength $\lambda_i$ depends on the X-ray tube voltage charge $U$ as follows:

$$\lambda_i(\text{nm}) = hc/(\text{eU}) = 1239.8/U(\text{V}) \qquad [3]$$

At $\lambda < \lambda_i$ the intensity of continuous radiation is absent. With an increase in $\lambda$ from $\lambda_i$ to $3\lambda_i/2$ the

continuous radiation intensity increases and with further increases in $\lambda$ it decreases.

## X-ray sources

The most widespread source of X-rays is the X-ray tube. In an X-ray tube, electrons emitted from the cathode are accelerated by an electrical field and bombard the metal target (anode). Target atoms, excited by electron impact, and electrons losing their kinetic energy when decelerating in the anode substance, emit X-rays. The primary radiation of the X-ray tube consists of two parts – characteristic (line) and continuous radiation. As a result of the primary radiation impinging on a substance its atoms emit characteristic fluorescence (secondary) radiation.

Other sources of X-rays are radioactive isotopes which can directly emit X-rays or electrons or $\alpha$ particles. In the last two cases charged particles can bombard the target substance which then emits X-rays. The intensity of X-ray isotope sources is some orders of magnitude less than that of X-ray tubes; however the dimensions, weight and cost of X-ray isotope sources are less than that of X-ray tubes.

X-rays are also generated as synchrotron radiation. It can be selected by a crystal analyser and may be used as an X-ray source. The intensity of X-rays selected from synchrotron radiation is some orders higher than that from an X-ray tube.

The characteristic radiation of the X-ray tube is spread in space isotropically, whereas its continuous radiation has maximal intensity in directions in a plane perpendicular to the trajectory of electrons bombarding the target. The X-ray component of synchrotron radiation is polarized and spread only in the plane of the synchrotron ring.

## Obtaining X-ray emission spectra

A schematic of an X-ray fluorescence instrument is presented in **Figure 3**. The X-ray tube is used as the source of primary radiation $h\nu_1$. The vacancies in inner shells of atoms of the substance investigated are formed as a result of primary radiation action. These vacancies are filled by other inner or outer electrons. This is accompanied by X-ray fluorescent photons $h\nu_2$ being emitted. This fluorescence radiation is spread out into the spectrum by means of a crystal analyser (or, for the ultrasoft X-ray region, by means of diffraction gratings) in accordance with Bragg's law

$$2d \sin \phi = n\lambda \qquad [4]$$

where $n$ is the order of the spectrum, $\lambda$ is the wavelength, $d$ is the grating constant of the crystal analyser and $\phi$ is the angle of incidence of the collimated X-ray fluorescent beam on the specific set of parallel planes in the crystal from which the beam is diffracted (see **Table 1**).

The X-ray fluorescence spectrum is then registered on a photographic film or by Geiger, proportional or scintillation counters, semiconductor detectors, etc.

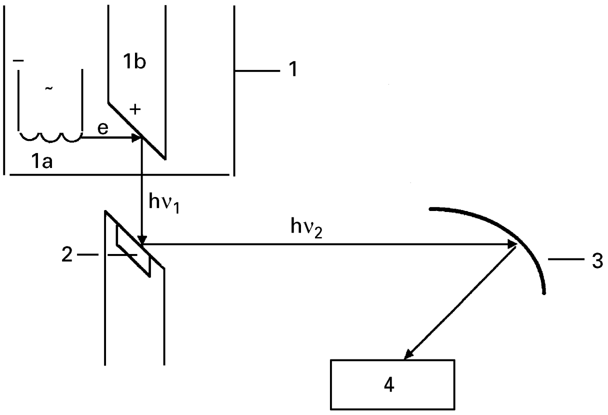



**Figure 3** An X-ray fluorescence spectrometer. 1, X-ray tube; 1a, electron source; 1b, target; 2, substance investigated (secondary anode); 3, crystal analyser; 4, registration device; $h\nu_1$, primary radiation; $h\nu_2$, secondary radiation; and $h\nu_3$, registered radiation.

**Table 1** Parameters of typical crystal analysers

| *Crystal* | *Reflecting plane* | *2d* (nm) | *Maximal solving ability (λ/Δ λ)* | *Relative coefficient* |
|---|---|---|---|---|
| KAP | 001 | 2.7714 | 1400 | 8–18 |
| Mica | 002 | 1.9884 | 2 000 | 2–3 |
| ADP | 101 | 1.0659 | 10 000 | 1–10 |
| EDDT | 020 | 0.8808 | – | – |
| PET | 002 | 0.8726 | 8 000 | 10–20 |
| Quartz | $10\bar{1}0$ | 0.8512 | 20 000 | 1–10 |
| Quartz | $10\bar{1}\bar{1}$ | 0.671 53 | 10 000 | 2–14 |
| Plumbago | 002 | 0.6696 | 100 | 50–200 |
| Ge | 111 | 0.653 27 | 6 000 | – |
| Si | 111 | 0.6271 | 10 000 | 2–10 |
| Calcite | 211 | 0.6069 | 15 000 | 2–30 |
| Quartz | $10\bar{2}0$ | 0.4912 | 30 000 | 0.4–3.3 |
| LiF | 200 | 0.4028 | 2 000 | 10 |
| Ge | 220 | 0.400 | 13 000 | 17–23 |
| Si | 220 | 0.383 99 | 29 000 | 1–6 |
| Calcite | 422 | 0.3034 | 64 000 | 0.4–0.9 |
| LiF | 220 | 0.2848 | 1 300 | 10–20 |
| Quartz | $20\bar{2}\bar{3}$ | 0.2806 | 90 000 | 0.3–0.9 |
| Quartz | $22\bar{4}\bar{3}$ | 0.2024 | 144 000 | 0.2–0.45 |
| Calcite | 633 | 0.202 | 122 000 | 0.3–0.6 |

The wavelengths and energies of the characteristic emission lines have been accurately measured and tabulated in handbooks, monographs and reference works for all chemical elements with $Z > 2$.

## X-ray fluorescence analysis

X-ray fluorescence analysis (XFA) is based on the X-ray emission line's intensity dependence on the concentration of the appropriate element. XFA is widely used for the quantitative analysis of various materials, especially in black and colour metallurgy and geology. XFA is distinguished by rapidity and a high degree of automation. The detection limits depend on the element, matrix composition and spectrometer used and lie in the region $10^{-3}$–$10^{-10}$%. Defining any element (with $Z > 4$) is possible by means of XFA in both a solid and a liquid phase.

However, the fluorescence line intensity $I_A$ of an investigated element A depends not only on its concentration $C_A$ in the sample, but also on the concentration of other elements, $C_i$, because other elements promote both absorption and excitation of fluorescence of the element A (matrix effect). Moreover, the measured value $I_A$ essentially depends on the sample surface, phase distribution, grain sizes, etc. Numerous methods have been developed to account for such effects. Most notable are the empirical methods of external and internal standards, using the background of scattered primary radiation and the method of dilution.

In the external standard method the unknown concentration $C_A$ of the element A is determined by comparing the intensity $I_A$ in the sample investigated with analogous $I_{st}$ values of standards for which defined element concentrations $C_{st}$ are known:

$$C_A = C_{st} I_A / I_{st} \quad [5]$$

This method allows one to take into account corrections connected with the equipment used; however, the composition of the standard should be close to that of the investigated sample to precisely match the matrix effect.

In the internal standard method some amount $\Delta\tilde{N}_A$ of a defined element A is added to the sample investigated. This leads to an increase in the fluorescence intensity of $\Delta I_A$. In this case:

$$C_A = I_A \Delta C_A / \Delta I_A \quad [6]$$

This method is especially effective for the analysis of complex samples but needs the special requirements of sample preparation.

The use of the background of scattered primary radiation is based on the fact that the ratio $I_A{:}I_b$ ($I_b$ is the background intensity) mainly depends on $C_A$ and only weakly depends on the concentration of other elements, $C_i$.

In the method of dilution a great amount of a weak absorber or a small amount of strong absorber is added to the sample investigated. These additions should reduce the matrix effect. This method is effective for water solution analysis and for the analysis of complex samples when the internal standard method is inapplicable.

There are also models in which the measured intensity $I_A$ is corrected on the basis of the intensities $I_i$ and concentrations $C_i$ of other elements. For example, $C_A$ may be represented as:

$$C_A = a_{A0} + a_{A1} I_A + a_{a2} {I_A}^2 + I_A \sum_{m,i} m_{Ai} I_i + \sum_{m,i} b_{Ai} I_i + \sum_{m,i} d_{Ai} I_i^2 \ (i \neq A) \quad [7]$$

$a$ and $b$ are values determined by the least-squares method with the help of $I_A$ and $I_i$ values measured in several standards with known concentrations $C_A$ of element A. Such models are widely used for the serial analysis of many samples via computers.

## X-ray microanalysis

X-ray microanalysis is a local analysis, fulfilled by means of microanalyser electron probe, for sample sites of ~ 1–3 μm². The electron probe is formed by electrostatic and magnetic fields to obtain a parallel electron beam with a diameter of ~ 1 μm. The analysis is via primary X-ray sample emission which is spread out into a spectrum by means of X-ray spectrometer. In this method corrections for the atomic number of the element, the absorption of its radiation in the sample, its fluorescence, and the characteristic spectra of other elements contained in the sample must be accounted for. Microanalysis is used for the investigation of two- and three-component systems such as mutual diffusion, crystallization processes, local variations of alloy structure, etc.

## List of symbols

$d$ = grating constant; $I_A$ = fluorescence line intensity of an investigated element A; $I_{st}$ = fluorescence intensity of a standard; $n$ = order of spectrum;

$p_A$ = probability of a vacancy being filled without emission; $p_r$ = probability of photon emission when vacancy is filled; $U$ = X-ray tube voltage; $V_i$ = ionization potential of the $i$th level; $Z$ = atomic number; $\varepsilon$ = electron kinetic energy; $\lambda$ = wavelength; $\delta$ = shielding constant; $\phi$ = angle of incidence.

*See also:* **X-Ray Absorption Spectrometers; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications.**

## Further reading


Bearden JA (1967) X-ray wavelengths. *Review of Modern Physics* **39**: 78–124.

Ehrhardt H (1981) *Röntgenfluoreszenzanalyse*. Leipzig: VEB Deutscher Verlag für Grundstoffindustrie, 250p.

Mazalov LN, Yumatov VD, Murakhtanov VV, Gelmukhanov FK, Dolenko GN, Gluskin ES and Kondratenko AV (1977) *Rentgenovskie Spectry Molekul*. Novosobirsk: Nauka, 331p.

Meisel A, Leonhardt G and Szargan R (1977) *Röntgenspektren und Chemische Bindung*. Leipzig: Geest & Portig, 320p.

Nemoshkalenko VV and Aleshin VG (1979) *Teoreticheskie osnovy rentrenovskoj emissionnoj spektroskopii*. Kiev: Naukova Dumka, 384p.

Siegbahn K, Nordling C, Fahlman A, Nordberg R, Hamrin K, Yedman J, Johansson G, Bergmark T, Karlsson SE, Lindgren I and Lindberg B (1967) *ESCA. Atomic, Molecular and Solid State Structure by Means of Electron Spectroscopy*. Uppsala: Nova acta Regiae Societatus Scientiarum Upsaliensis, 493p.


# X-Ray Fluorescence Spectrometers

**Utz Kramar**, University of Karlsruhe, Germany

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**

**Methods & Instrumentation**

X-ray fluorescence (XRF) spectrometers are widely used for the determination of elements with atomic numbers from 4 (beryllium) to 92 (uranium) at concentrations from 0.1 μg g$^{-1}$ to high percentage levels. These elements can be analysed using characteristic $K_\alpha$-lines ($KL_{III}$) from 11.4 nm/0.1885 keV (Be $K_\alpha$) to 0.0126 nm/98.4 keV (U $K_\alpha$). Nevertheless, elements of higher atomic number (e.g. Cd $L_{III}M_V$ 0.3956 nm/3.133 keV; U $L_{III}M_V$ 0.09106 nm/13.61 keV) are often determined using their L-lines. The characteristic X-ray lines of these elements can be determined either with sequential or with simultaneous wavelength-dispersive spectrometers by Bragg diffraction, using the wave phenomena of X-rays or in energy-dispersive systems using their energy characteristic. Coherent and incoherent scattering of primary X-rays in the sample may cause increased background effects, and matrix-dependent absorption of the characteristic secondary X-rays may also cause severe matrix effects. Since XRF methods are routinely used, methods for correcting the matrix effects such as fundamental parameters have been developed and instruments with drastically improved peak to background ratios such as polarized X-ray fluorescence (PXRF) and total reflection X-ray fluorescence (TXRF) spectrometers have been designed during the 1990s. The principles of XRF spectroscopy and descriptions of the different kinds of instrumentation are given in an increasing number of monographs, books and reviews, with extensive data compilations.

## Principles

If a target is irradiated with photons, or charged particles (electrons or ions) with energies exceeding the binding energy of the bound inner electrons, an electron from inner orbitals of the target atoms can be ejected.

$$E_\infty = 13.56 \cdot \frac{[Z_{\text{eff}}]^2}{n} \text{ (eV)}$$

where $Z_{\text{eff}}$ = effective atomic number and $n$ = principal quantum number. If the total energy of the photon is transferred to the electron this interaction is called the photoeffect.

The resulting atom is unstable and regains its ground state by transferring an electron from a high-energy outer orbital to the vacancy in the inner electron shell. The energy difference between the initial and final energy state of the transferred electron is

released as a photon of the energy

$$E_{\text{photon}} = E_{\text{initial}} - E_{\text{final}}$$

In XRF spectroscopy photons are used to excite the characteristic elemental X-rays from the sample. Alternatively the incident photons can be scattered coherently (Rayleigh scattering) at an electron of the inner shell or incoherently at an electron of the outer shell (Compton scattering). In the second case the wavelength/energy of the scattered photon depends on the initial wavelength/energy of the original photon and the scattering angle $\Theta$

$$E' = E/[1 + (1 - \cos\Theta)(E/E_e)]$$

where $E'$ = photon energy after Compton scattering, $E$ = initial energy of the photon and $E_e$ = energy equivalent of the electron mass. The XRF spectrum of a sample is a mixture of the different characteristic X-rays emitted from the atoms in the sample and of coherent and incoherent scattered components of the primary radiation source. The task of XRF spectrometers is to separate the different spectral components, to determine their intensities and based on this to calculate the elemental concentrations. Typically, XRF spectrometers consist of a photon source for the excitation of the secondary X-rays, a sample support, an X-ray detection unit and a data evaluation unit.

## X-ray sources

In most XRF spectrometers an X-ray tube is used as the photon source. Alternatively radioactive isotopes can be applied for the excitation of the characteristic X-rays from the sample. X-ray tubes can be used in both wavelength-dispersive and energy-dispersive systems. Due to their lower beam intensities, application of radionuclides is restricted to energy-dispersive systems.

### X-ray tubes

In X-ray tubes, the X-rays are produced by the bombardment of matter with accelerated electrons. The X-ray tubes are built as a vacuum-sealed metal glass cylinder. The electrons are emitted from a heated tungsten filament which serves as the cathode and are accelerated by a high voltage applied between the filament and a metal anode. Two effects can occur if the accelerated electrons interact with the atoms of the anode material. (1) An electron enters the electric field of an atomic orbital and is slowed down, and the loss of kinetic energy during slowing down is emitted as electromagnetic radiation, called bremsstrahlung. (2) An inner-shell electron of an atom is ejected if the kinetic energy exceeds the binding energy, and from this the characteristic X-rays of the anode material are emitted. Thus the spectrum emitted from the anode consists of the continuous bremsstrahlung and the characteristic X-ray lines of the anode material.

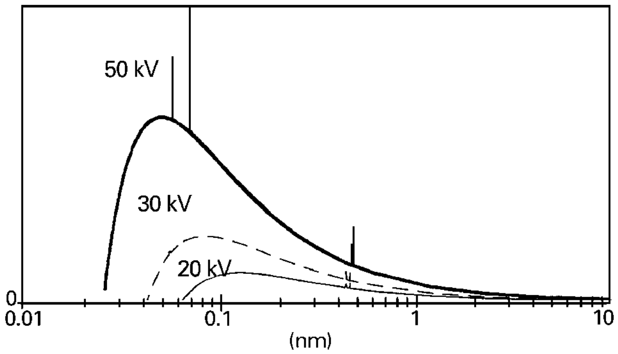



**Figure 1** Primary spectrum of an X-ray tube with a rhodium anode.

The maximum X-ray energy primarily emitted from the tube is determined by the applied acceleration voltage.

$$E_0 = e \cdot U_o$$

The radiation intensity is a function of the tube current and the applied high voltage (**Figure 1**).

$$I(\lambda)\mathrm{d}\lambda = kiZ\left(\frac{\lambda}{\lambda_0} - 1\right)\frac{1}{\lambda^2}\mathrm{d}\lambda$$

Excitation of the characteristic X-rays from the sample can be optimized by selecting the appropriate anode material, voltage and tube current (**Table 1**).

**Table 1** Anode materials and application ranges of X-ray tubes used for X-ray fluorescence analysis

| *Anode material* | *Element range* | | *Operating voltage* (kV) |
|---|---|---|---|
| | *K-spectra* | *L-spectra* | |
| Cr | $^{8}$O–$^{22}$Ti | $^{42}$Mo–$^{55}$Cs | 55 |
| W | $^{23}$V–$^{27}$Co | $^{56}$Ba–$^{72}$Hf | 55 |
| Au | $^{28}$Ni–$^{30}$Zn; $^{40}$Zr–$^{92}$U | $^{73}$Ta–$^{75}$Re | 65/100 |
| Mo | $^{31}$Ga–$^{39}$Y | $^{76}$Os–$^{92}$U | 100 |
| Rh | $^{4}$Be–$^{56}$Ba | $^{42}$Mo–$^{92}$U | 60 |

In wavelength-dispersive XRF (WDXRF), TXRF and PXRF methods, efficiencies of primary-to-detected-secondary radiation are very low. Therefore tubes for these methods are generally operated at high power [~3 kW per kV up to 100 kV]. Because most of the tube power is converted to heat, these tubes have to be water cooled. Depending on the kind of application, different tube designs such as end-window (e.g. **Figure 9**), side-window (**Figure 2A**) and line focus are employed. When extremely high primary intensities are necessary rotating anodes are used (**Figure 2B**). In energy-dispersive systems radiation efficiency is several orders of magnitude higher than in WDXRF, therefore compact air-cooled low-power tubes (3–100 W) are used in most cases.

## Radionuclide sources

In mobile EDXRF systems, where small dimensions are essential, sealed radionuclides are often used as primary radiation sources instead of X-ray tubes. The radionuclides have to be selected with respect to their decay scheme, half-life and radiotoxicity (**Table 2**).

### $\beta^-$ Decay

From isotopes decaying in $\beta^-$-mode, high-energy electrons ranging from several keV to MeV are emitted from the nucleus. The electrons interact with the matter and bremsstrahlung is produced. Most $\beta^-$-sources are used for secondary target excitation.

### Electron capture (EC)

The nucleus captures an electron of the innermost electron shell and a proton of the nucleus changes into a neutron. Due to this process a vacancy in the innermost electron shell occurs. By filling this vacancy with an electron from the outer shells, the characteristic X-rays of the daughter atom are emitted.

### $\gamma$ Decay

In most of the radioactive decays the daughter isotope is formed in an excited state. These nuclides are transferred into the ground state by emitting electromagnetic radiation, the $\gamma$-radiation.

### Selection criteria

The choice of the radionuclide source depends on the type of sample, the elements to be determined and the detection technique. Generally a high specific activity of the radionuclide and emission energies suitable for the application are required. The most convenient isotopes for energy-dispersive XRF are those that decay exclusively by EC without emitting $\gamma$-radiation.

Radionuclides emitting high-energy photons (>150 keV), or decaying by $\beta^+$ or high-energy $\beta^-$, or those with short half-lives are not recommended for use in XRF systems. A selection of suitable radionuclides with suitable energies and high specific activity is given in **Table 2**.

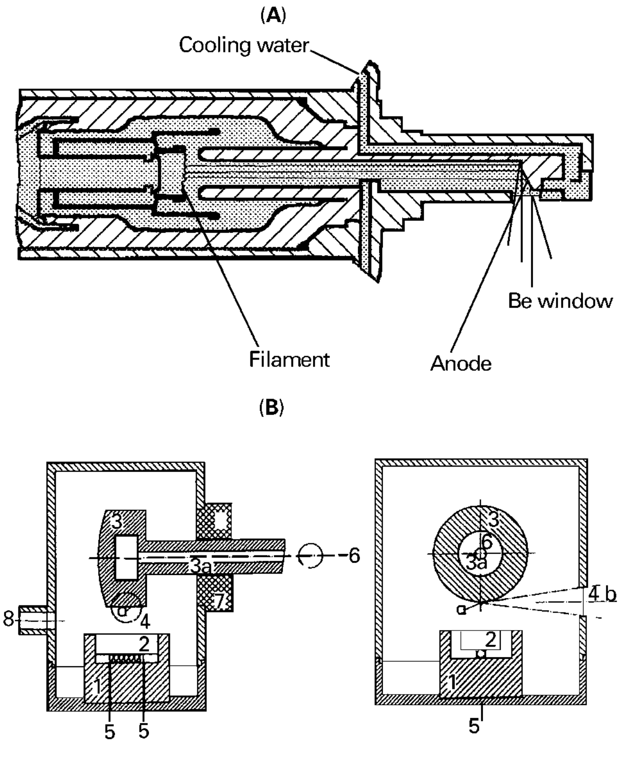



**Figure 2** (A) Schematic sketch of a side window tube. (B) Rotating anode tube in two sectional views: (1) cathode unit; (2) filament; (3) cylindrical anode with (3a) rotary shaft; (4) window; (5) electrical connections; (6) cooling water connection; (7) sealing gasket; (8) vacuum flange. Reproduced with permission from Klockenkämper R (1997). *Total Reflection X-ray Fluorescence Analysis.* New York: Wiley.

## X-ray dispersion and detection units

XRF spectrometers have two major objectives: (a) to determine the spectral distribution of the X-rays emitted from the sample; (b) the measurement of the intensity of the selected spectral component. In wavelength-dispersive XRF, the spectrum is dispersed into different wavelengths by Bragg diffraction at different crystals. Intensities are measured by electronic detectors. In energy-dispersive spectrometers both energy dispersion and intensity measurement are performed by electronic detectors; thus for

**Table 2** Radionuclides used as primary sources in energy-dispersive X-ray fluorescence

| *Isotope* | *Decay mode* | *Half-life* (years) | *Energy* (keV) | *Elemental range* *K X-rays* | *L X-rays* | *Recomm. activity* (MBq) | *Working life* (years) |
|---|---|---|---|---|---|---|---|
| $^{241}$Am | α | 433 | 59.5 Np L X-rays (12–22) | Zr–Ce | W–U | 370–1 110 | 15 |
| $^{109}$Cd | EC | 1.26 | 88 Ag K X-rays (22–26) | Ti–Nb | Tb–U | 111–740 | 5 |
| $^{57}$Co | EC | 0.74 | 122; 136 Fe K X-rays (6–7) | Ba–U | | 37–370 | 5 |
| $^{244}$Cm | α | 17.8 | 43; 99; 152 Pu L X-rays (12–23) | Ti–Se | Ce–Pb | 370–3 700 | 10 |
| $^{55}$Fe | EC | 2.69 | Mn K X-rays (5.9–6.5) | Si–V | Nb–Sn | 185–1 850 | 5 |
| $^{153}$Gd | EC | 0.66 | 97; 103 Eu K X-rays (41–48) | | | | |
| $^{125}$I | EC | 0.17 | 35 Te K X-rays (27–32) | | | 370–3 700 | 1 |
| $^{3}$H | $\beta^-$ | 12.43 | $E_{max}$: 18.6 (beta) | | | | |
| $^{238}$Pu | α | 433 | 43 U L X-rays (11–22) U K X-rays (94–115) | Ti–Se | Ce–Pb | 370–1 110 | 10 |

detectors used in energy-dispersive spectrometers extremely good energy resolutions are required.

## Dispersing crystals

At the X-ray diffraction crystals the X-ray spectrum is dispersed into different wavelengths according to Bragg's law

$$n \cdot \lambda = 2d \sin \Theta$$

The wavelengths of characteristic lines used in X-ray spectrometry range from ~0.03 nm (Ba K) to ~10 nm (Be K). This range cannot be covered by use of a single diffraction crystal. The detectable wavelength and high-order reflections are limited by the relation between $d$ and $\lambda$.

$$n\lambda \leq 2d$$

For example with LIF 200, the crystal with the broadest application range, the elements with atomic numbers <19 (K) are not detectable. At shorter wavelength, the dispersion between neighbouring lines decreases and they cannot be resolved from each other. The diffraction crystals have to be selected with respect to their reflectivity and to be suitable for the lines to be detected. Commonly used X-ray diffraction crystals and their application range are compiled in **Table 3**.

**Table 3** Diffraction crystals commonly used in wavelength-dispersive X-ray fluorescence

| *Analyser crystal* | *Material* | *2d* *(nm)* | *Detectable elements* *K* | *L* | *Efficiency* |
|---|---|---|---|---|---|
| | Topaz | 0.27 | V–Ta | Ce–U | Average |
| LIF(220) | Lithium fluoride | 0.29 | V–Ta | Ce–U | High |
| LIF(200) | Lithium fluoride | 0.4 | K–La | Cs–U | High |
| Ge | Germanium | 0.65 | P–Zr | | Average |
| PET | Pentaerythrite | 0.87 | Al–Ti | Kr–Xe | High |
| AdP | Ammonium dihydrogen phosphate | 1.06 | Mg | | Low |
| TAP/TIAP | Thallium biphthalate | 2.58 | F–Na | | High |
| OVO-55 | W/Si multilayer | 5.5 | N–Si | Ca–Br | High |
| OVO-160 | W/C multilayer | 16 | Be–O | | ? |
| OVO-C | V/C multilayer | 12 | C | | ? |
| OVO-B | Mo/B4C multilayer | 20 | Be–B | | ? |
| PbSD | Lead stearate decanoate | 10 | B–F | | Average |

## X-ray detectors

In X-ray detectors the energy transported by the radiation is converted into forms that can be recognized visually or electronically. Generally the photons are absorbed by the detector material and energy transfer takes place by ionization. The number of ionizations $N$ per photon is proportional to the energy $E$ of the absorbed photon and depends on the average energy $e$ necessary to produce an

electron–ion pair in the detector material.

$$N = E/e$$

The energy resolution of a detector is determined by the statistical scatter of the number of ionizations per photon.

$$\sigma = \sqrt{(FN)}$$

with $F$ = the Fano factor. Due to this relation it is possible to separate the radiation absorbed in the detector into the different energies.

Critical parameters for selection of an appropriate detector are: efficiency, energy resolution and deadtime, i.e. the pulse processing time of the detector. A measure for the energy resolution is the full width of half maximum (FWHM).

$$\text{FWHM} = 2.36 \times \sqrt{(eFE)}$$

The detector efficiency depends on the radiation energy to be determined and the density and type of detector material. In XRF spectrometers three different type of X-ray detectors are used: gas-filled detectors, scintillation detectors and semiconductor detectors.

## Gas proportional (flow-) counters

Gas-filled detectors are used for the determination of low-energy X-rays. An isolated thin wire is mounted in a metal cylinder filled with a suitable counting gas (e.g. 10% Ar, 90% $CH_4$). The wire is held at a high positive potential (1.2–2 kV) and the metal cylinder is grounded (**Figure 3**). A thin window (Be or Mylar) allows the radiation to enter the detector. The gas is ionized by the photons and the electrons produced are accelerated towards the anode and the ions towards the cathode. The energy necessary to produce an ion pair is 20–30 eV, depending on the type of counting gas. Additional ionizations are produced by collisions with other gas atoms. The electrons are collected at the anode, producing a short-term voltage drop. Applying an appropriate high voltage to the detector, the voltage drop is proportional to the radiation energy (**Figure 4**). High radiation doses degrade the counting gas. In addition, the thin detector windows, necessary for the detection of low-energy photons, can lead to a permanent loss of counting gas by diffusion. To avoid these problems the detectors are operated as

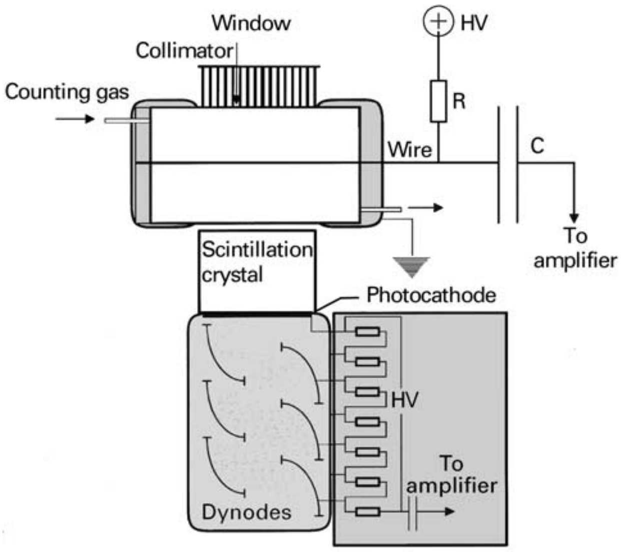



**Figure 3** Schematic of gas proportional counter (flow counter) and scintillation detector operated as tandem detectors in wavelength-dispersive X-ray fluorescence spectrometers. R = resistor; C = capacitor.

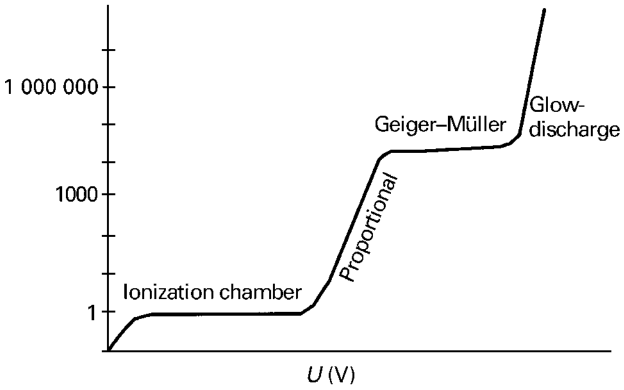



**Figure 4** Gas-amplification characteristics of gas-filled detectors. Most flow counters for WDXRF are operated in the change-over region from proportional to Geiger–Müller.

flow counters, in which the counting gas is permanently exchanged.

## Scintillation detectors

Scintillation detectors are used for the determination of the high-energy part of the X-ray spectrum. In scintillation detectors the material of the detector is excited to luminescence (emission of visible or near-visible light photons) by the absorbed photons or particles. The number of photons produced is proportional to the energy of the absorbed primary photon. The light pulses are collected by a photocathode. Electrons, emitted from the photocathode,

are accelerated by the applied high voltage and amplified at the dynodes of the attached photomultiplier (**Figure 3**). At the detector output an electric pulse proportional to the absorbed energy is produced. The average energy necessary to produce one electron at the photocathode is approximately 300 eV. For X-ray detectors, in most cases NaI or CsI crystals activated with thallium are used. These crystals offer a good transparency, high photon efficiency and can be produced in large sizes.

### Semiconductor detectors

Semiconductor detectors can be considered analogous to ionization chambers with the gas being replaced by the semiconductor material. The group 4 elements silicon and germanium are most widely used as semiconductor detectors when excellent energy resolution is required. Semiconductors have a small energy gap between their valence and conducting bands. At energies above absolute zero, thermal excitation will move some electrons from the valence band to the conductor band and holes or positive charges are left in the valence band. The quasi-free electrons in the conductor band move towards the anode if an electric field is applied. If ionizing radiation interacts with a semiconductor, electrons are transferred from the valence to the conducting band. The energy necessary to produce an ion pair is low (3–5 eV). Thus a large number of electrons, proportional to the absorbed radiation energy, can be collected at the anode by applying a high voltage of 1–2 kV to the detector, and statistical scatter is low. The current of the radiation-induced free electrons is superimposed by the temperature-dependent free charge density (leakage current) and electron recombination by acceptor-type impurities. Today silicon cannot be produced at purity levels that allows electron recombination by acceptor-type impurities to be neglected. In Si(Li) detectors acceptor-type impurities are compensated for by drifting the silicon with lithium as donator.

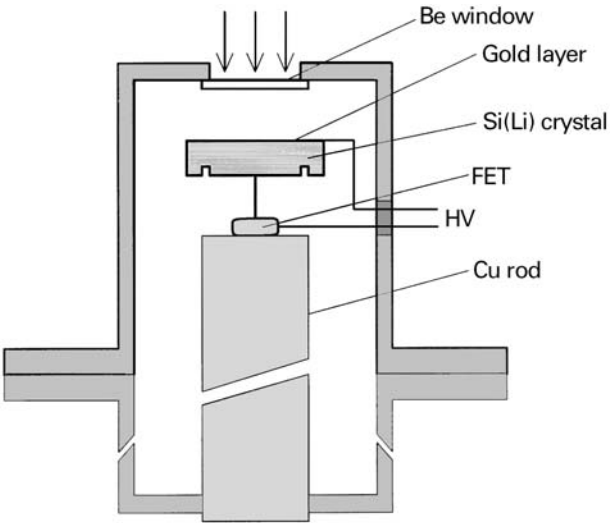



**Figure 5** Schematic of a planar semiconductor detector. HV = high voltage; FET = field-effect transistor.

In order to reduce the thermal charge carrier generation (noise) to acceptable levels Si(Li) detectors and Ge detectors are operated at liquid nitrogen ($LN_2$) temperatures. LN-cooled detectors are mounted in a vacuum chamber, which is inserted in or attached to a $LN_2$ Dewar flask (**Figure 5**). The detector is in thermal contact with the $LN_2$ and the radiation enters the detector housing via a thin Be window. In mobile systems, and for applications requiring long-term unattended operation, detectors cooled by Joule–Thomson coolers, He-cycle refrigerators or Peltier elements are used, but all of them can show degraded energy resolution, with the Peltier exhibiting the most (~20%). Extremely compact room temperature detectors using semiconductors with higher energy gaps (CdZnTe, $HgI_2$, GaAs) have been developed in the last decade. Energy resolution of these detectors is poorer than for Si(Li) detectors but better than with proportional or scintillation detectors. The characteristic properties of the different detector types are compared in **Table 4**.

**Table 4** Properties of radiation detectors used in XRF systems

| *Detector type* | *Material/filler gas* | *Ionization energy* (eV) | *Band gap* (eV) | *FWHM at 5.9 keV* (eV) *Theor.* | *Typical* | *Dead-time* (µs) | *Energy range* (keV) | *Remarks* |
|---|---|---|---|---|---|---|---|---|
| Proportional counter | Ar/methane | 26.4 | | 840 | 840 | 1 | <1–10 | |
| Scintillation | NaI(Tl) | 300 | | 3 150 | 3 150 | 0.2 | 3–100 | |
| Semiconductor | Si(Li) | 3.61 | 1.12 | 120 | 150 | 2–4 | 1–60 | $LN_2$ |
| | Ge | 2.98 | 0.74 | 108 | 145 | | 1–200 | $LN_2$ |
| | CdZnTe | ~5 | ~1.5 | 135 | 285 | | >2 | |
| | $HgI_2$ | 6.5 | 2.13 | 160 | 270 | | >2 | |

## Wavelength-dispersive spectrometers

In wavelength-dispersive spectrometers the characteristic radiation emitted from the sample is dispersed into different wavelengths by Bragg diffraction and the resulting radiation is registered by electronic detectors. Since the wavelength selection is performed by diffraction of the radiation to different angles, only minor energy resolution of the radiation detectors is required. Nevertheless according to Bragg's law higher order diffraction radiations of the wavelength $\lambda/2$, $\lambda/3$, etc. are registered at the same angle as those of wavelength $\lambda$. These high-order signals cannot be separated by the diffraction crystals, but they are significantly different in their energies and can be separated using the energy-resolution capabilities of the detectors. **Figure 6** gives an example of elimination of second-order radiation by energy discrimination in WDXRF.

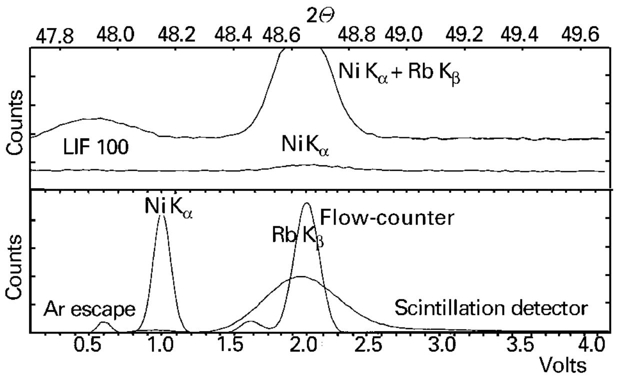



**Figure 6** Wavelength and energy spectrum of Ni K$\alpha$ in the presence of high Rb concentrations with open window (3 V) and window width set to 1.3 V to remove second-order radiation of Rb $K_\beta$. Spectrum of reference rock MA-N, measured with WDXRF SRS 303 AS.

Thus, a wavelength-dispersive spectrometer consists of the excitation unit (X-ray tube with high-voltage supply), the sample chamber, the diffraction unit and the detector unit with amplifier, lower level and window discriminator, and registration unit (**Figure 7**). Generally the sample, the diffraction unit and the low-energy detector are mounted together in a vacuum chamber to avoid absorption of low-energy X-rays along the radiation path. For some applications, primary beam filters are used to optimize the tube spectrum. Wavelength-dispersive spectrometers are either designed as fixed (simultaneous) spectrometers or as sequential (scanning) spectrometers. Generally wavelength-dispersive spectrometers have a mass of > 400 kg and are operated in well established laboratories, equipped with cooling water facilities and high-current electric power supply. Nevertheless, mobile systems operated with low power tubes are also possible.

### Sequential spectrometers

In most sequential XRF spectrometers a flat crystal is mounted on the central axis of a rotating goniometer and a proportional counter and a scintillation counter are mounted at the moving goniometer arm. The Bragg angle can be simply varied, rotating the crystal mount by $\Theta$ and the detectors by $2\Theta$. Primary and secondary collimators between sample and crystal and crystal and detectors are used to limit the beam divergence. Most common types of collimator consist of a pack of parallel plates of highly absorbing material (e.g. tungsten). Radiation (nearly) parallel to these plates can pass, whereas diverging radiation is absorbed by the plates. Narrow spacing

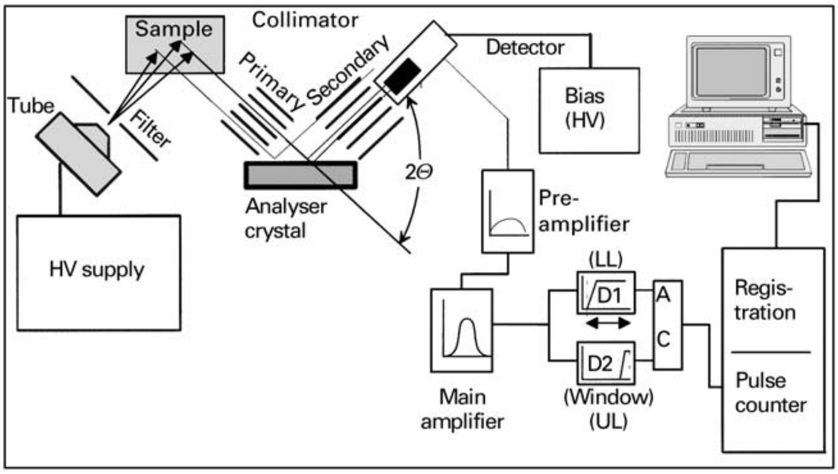



**Figure 7** Schematic drawing of a typical sequential X-ray fluorescence setup. HV = high voltage; AC = anti-coincidence; D = discriminator; LL = lower level; UL = upper level.

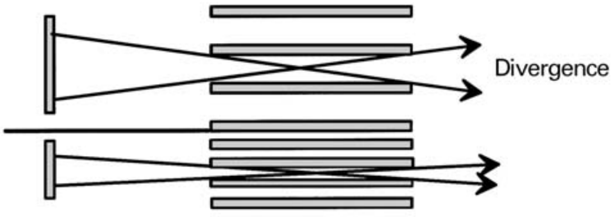



**Figure 8** Schematic of coarse and fine collimator assignment.

of the plates results in optimum resolution but intensity is reduced by an amount equivalent to the space angle (**Figure 8**). Therefore most sequential spectrometers are equipped with coarse and fine spaced primary collimators.

Curved focusing scanning crystals are rarely used in wavelength-dispersive instruments, but are common in microprobes.

### Simultaneous spectrometers

Simultaneous spectrometers are designed with fixed crystal arrangements. In modern simultaneous spectrometers up to 30 crystal–detector combinations are fixed at suitable angle positions for the peak and background measurements of the application. Generally, focusing crystal shapes are used. For fixed diffraction geometries the best focusing is achieved with logarithmically curved crystals (**Figure 9**).

## Energy-dispersive spectrometers

In energy-dispersive systems the separation of the different components in the X-ray spectrum is performed exclusively by the detector and the subsequent electronic components. Thus good energy resolution, low electronic noise, low temperature drift and excellent linearity of the electronic components are required. Generally, an energy-dispersive spectrometer consists of a radiation source (X-ray tubes or radionuclides), sample support, pre- and main amplifier, pile-up rejector, analogue-to-digital converter and multichannel analyser.

As radiation sources, X-ray tubes or radionuclides are applied. The polychromatic radiation of the tube is scattered by the Compton or Rayleigh effect towards the detector, which results in a high bremsstrahlung background and line interferences caused by the characteristic lines of the anode material and the resulting Compton peak.

Primary beam filters are used routinely in energy-dispersive XRF (EDXRF) to optimize the tube spectrum for the specific application (**Figure 10**). Filters with a sharp absorption edge slightly above the characteristic tube lines are used for filtering out the high-energy bremsstrahlung and a quasi-monochromatic excitation spectrum is achieved (e.g. Rh tube and Pd filter). The characteristic lines can be removed by heavy-element filters without an absorption edge in the high-energy region (e.g. Rh tube and Cu filter) (**Figure 11**). The use of primary beam filters causes some intensity loss but this is overcompensated for by the background reduction.

Secondary targets are used to provide optimal excitation conditions for specific elements, but the drastic loss of intensity in the beam restricts their application to tubes with higher output powers or

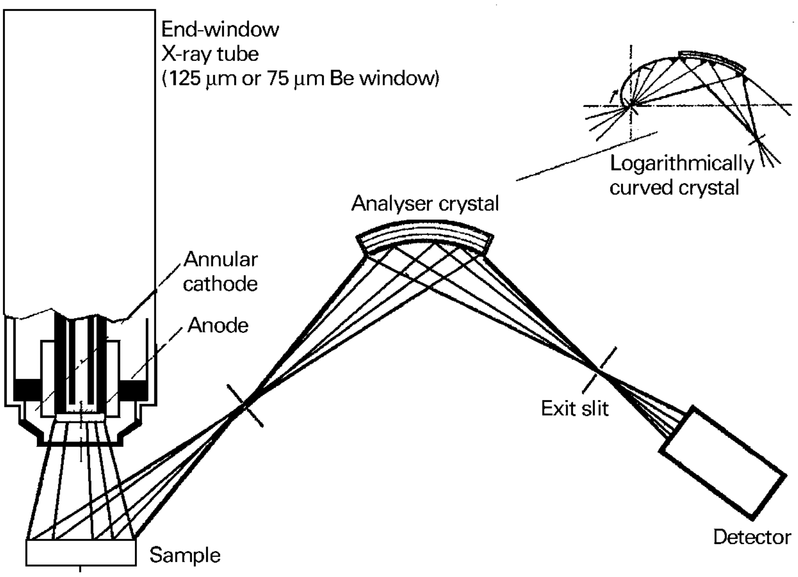



**Figure 9** Schematic of a simultaneous WDXRF with end-window tube and logarithmically curved focusing analyser crystal. Redrawn from a sketch in a brochure from Bruker AXS AG.

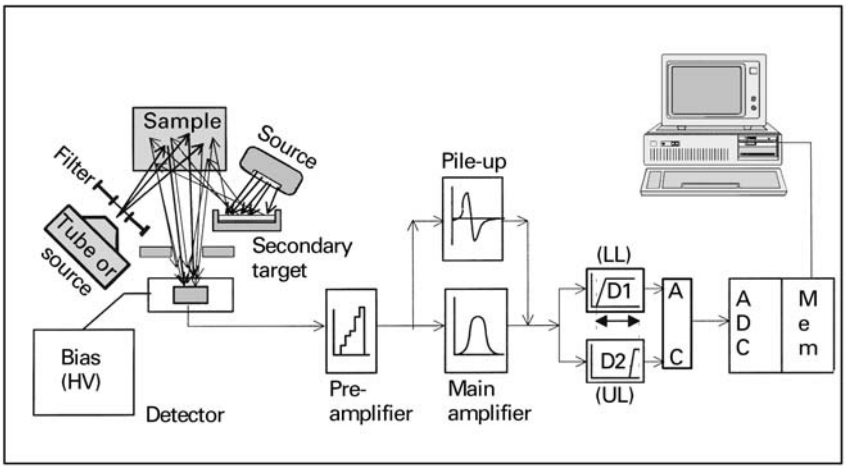



**Figure 10** Schematic drawing of a typical EDXRF setup with excitation by tube or radionuclide with optional primary filter or secondary target excitation. ADC = analog to digital converter.

radionuclides of higher activities. Most radionuclides provide monoenergetic excitation, but the excitation energies are not tunable. For radionuclide sources a licence for handling and transportation is required.

In EDXRF systems, semiconductor detectors are used exclusively, apart from some mobile systems for low-energy application equipped with high-resolution proportional counters. Most systems are equipped with cooled Si(Li) or Ge detectors. The energy-proportional electric charge pulse is integrated by the preamplifier. Today state of the art preamplifiers are of the pulsed optical feedback type with the first active element, a field-effect transistor (FET), mounted in the detector cryostat in order to reduce electronic noise. The preamplifier output is converted to an energy-proportional voltage pulse of some microseconds duration, depending on the shaping time of the amplifier. Events arriving at the detector and amplifier during their signal processing period cannot be processed correctly. A pile-up rejector eliminates these pulses by inspecting the slopes of the pulses. The amplifier output signal is digitized by a subsequent analogue-to-digital converter (Wilkinson type) and collected in the memory of a multichannel analyser. Each channel is equivalent to a small energy interval and the channel content holds the counts of this energy interval, accumulated during the measuring time. All spectral components are registered simultaneously.

EDXRF systems are available as stationary systems, to be used for fast simultaneous multielement determinations in analytical laboratories, as compact equipment suitable for mobile field laboratories and as hand-held probes to be used directly on-site for screening analyses. The dimensions of the equipment range from several hundred kg for stationary systems down to 1.3 kg for the smallest radionuclide-excited hand-held probe.

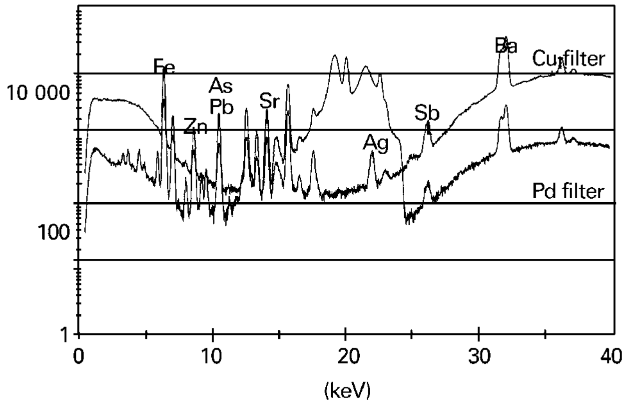



**Figure 11** Energy-dispersive spectra of reference soil sample GXR-2, measured with Spectrace 5000 using a Rh anode tube and Pd and Cu primary beam filters to optimize excitation conditions.

Special instruments have been developed for trace element analysis in the picogram region in microsamples using total reflection of the X-ray beam (TXRF) and the sub-μg $g^{-1}$ range in bulk samples using polarized X-rays.

## Total reflection XRF (TXRF)

At radiation incident angles of less than the critical angle $\alpha_{crit}$, the primary beam is totally reflected at the surface of a specimen.

$$\alpha_{crit} \approx \frac{1.65}{E}\sqrt{\frac{Z}{A}\rho}$$

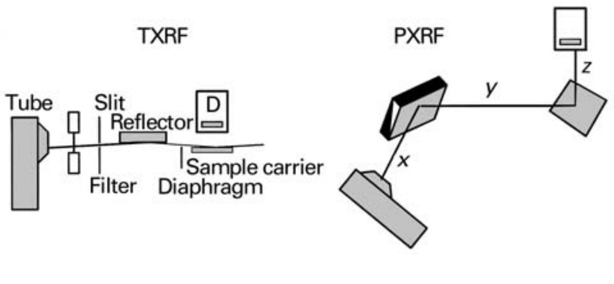



**Figure 12** Schematic of TXRF and *x*–*y*–*z* geometry of PXRF.

where $\rho$ = density, $A$ = atomic mass and $Z$ = atomic number. At angles $<\alpha_{crit}$ the X-rays can penetrate into the substrate only for a few nm. At these angles interaction of the X-rays with the total reflecting specimen is at a minimum. This effect is used in two types of TXRF instruments. In total reflection devices for trace element analysis, the sample is prepared as a thin film on a well polished quartz block. In conventional thin film methods the characteristic X-rays of elements in the sample are excited, but excitation and Compton scattering occurs within the material of the sample support as well. In TXRF the geometry is arranged to provide a total reflection of the primary X-ray beam (**Figure 12**). The low total reflecting incident angle provides an excellent interaction between the primary beam (absorption path ~1 mm per 1 µm sample thickness) and the sample, whereas absorption of the secondary radiation in the sample can be neglected in most cases. Generally, it is not necessary to use matrix correction methods as it is with other XRF methods. Once calibrated, samples of all the matrices can be analysed without recalibration and matrix corrections. Detection limits at the µg $g^{-1}$ level can be obtained from a few micrograms of solid sample and at the ng $g^{-1}$ level from a few microlitres of a liquid sample. For surface and thin layer analysis, instruments with angle variable sample positioning devices are used. With these instruments the angle-dependent intensity profile can be recorded as the basis for surface and thin layer analysis. For the determination of thin layer and surface contaminations, e.g. for wafer analysis, the measurement is performed at the critical angle with the beam grazing the surface. With these devices impurities of some $10^{11}$ – $10^{13}$ atoms $cm^{-2}$ can be determined.

## Polarized X-ray fluorescence (PXRF)

In PXRF the primary X-ray beam is polarized by Barkla scattering or Bragg reflection. Under orthogonal conditions the scattered radiation is highly polarized. The polarized radiation is used to excite the characteristic X-rays of the elements in the sample. The secondary radiation is measured in orthogonal geometry to the polarization plane. Background, produced by Compton and Rayleigh scattering in the sample, is effectively reduced. Detection limits in bulk samples are improved by approximately a factor of 3 compared to normal EDXRF. Thus sensitivity and accuracy are equivalent (K–Mo, Hf–U) or better (Ag–Nd) than those obtained by WDXRF. Therefore, PXRF combines the capability of nondestructive multielement analysis with low detection limits and a minimum of sample preparation.

## Sample preparation for XRF

With XRF solid as well as liquid samples can be analysed. Liquid samples have to fill an adequate sample container, with an X-ray transparent window (e.g. Mylar). For the determination of light elements the sample chamber has to be flushed with He instead of applying a vacuum. Solid samples can be analysed as pressed powder pellets, bulk powder samples, fused pellets or as pills (diluted with a flux, e.g. Spectromelt). For quantitative analyses a flat and homogenous sample surface is needed. Powder samples have to be pulverized to a fineness, where absorption in single grains can be neglected. Solid and liquid samples can be prepared either by thin film techniques, or as samples of infinite thickness. In thin film techniques the amount of sample in the irradiated spot has to be known. This can be established by adding an internal standard. A specimen can be regarded to be of infinite thickness if

$$D_\infty \geq 3 \cdot d_{1/2_{\text{prim}}} \quad \text{with} \quad d_{1/2_{\text{prim}}} = \frac{\ln 2}{\mu(E_{\text{prim}})}$$

where $d_{1/2}$ = half absorption thickness and $\mu$ = mass attenuation coefficient.

## Matrix correction methods

Thin specimens show a strictly linear dependency of the intensity of a characteristic line on the amount of the analyte. In thick specimens, primary radiation is absorbed along its path in the sample and secondary radiation on its way out. Additionally, secondary excitation by intense characteristic lines with energies above the absorption edge of the analyte can occur. These intensity variations (up to a factor of ~20) are strongly matrix dependent and have to be corrected. Different methods can be used.

Normalization of the Compton peak can be applied, if no absorption edge occurs between the Compton peak and the line of the analyte.

$$c_i = (a_i I_i + b_i) \frac{I_{C_{\mathrm{ref}}}}{I_{C_{\mathrm{sample}}}}$$

Empirical, semiempirical and theoretical matrix correction models are used for the determination of major components, where line intensities are influenced by absorption edges and enhancement. Routinely applied is the fundamental parameters model based on theoretical or experimentally determined $\alpha$-coefficients:

$$c_i = [D + E \cdot I(\lambda_i)] \times [1 + \Sigma(c_j \cdot \alpha_{ij})]$$

or empirical corrections such as

$$c_i = \left(b_{0i} + b_{1i}I_i + b_{2i}I_i^2 + I_i\Sigma m'_{ij} \cdot I_j + \Sigma m''_{ij} \cdot I_j\right)$$

or under consideration of elemental concentrations

$$c_i = \left[b_{0i} + b_{1i}(I_{\mathrm{F}})_i + b_{2i}(I_{\mathrm{F}})_i^2\right] \times \left(1 + \Sigma k'_{ij} \cdot c_j + \frac{\Sigma k''_{ij} \cdot c_j}{1 + c_j}\right)$$

where $I_{\mathrm{F}}$ = fluorescence intensity.

## Recent trends

All modern X-ray fluorescence systems are computer controlled and equipped with automatic sample changers. Different matrix correction models are included in the data evaluation software, and further developments on expert systems will reduce manual calibration work. With WDXRF, the detectability of light elements (down to Be) will be optimized, e.g. by using X-ray tubes and detectors with ultrathin windows and widely spaced (focusing) analyser crystals. In EDXRF, trends are for miniaturization, development and optimization of high-resolution room temperature detectors and extension of the application range towards the determination of light elements.

## List of symbols

$A$ = atomic mass; $b_i$ = calibration coefficient for element $i$; $c_i$ = concentration of element $i$; $C_j$ = concentration of element $j$; $d$ = d-spacing of a crystal; $d_{1/2}$ = half absorption thickness; $D_\infty$ = saturation thickness; $e$ = charge of an electron; $E$ = initial energy of the photon; $E_0$ = maximum energy of accelerated electron; $E_\infty$ = energy necessary to remove an electron from the orbital $n$; $E'$ = photon energy after Compton scattering; $E_{\mathrm{e}}$ = energy equivalent of the electron mass or energy necessary to produce an electron–ion pair; $E_{\mathrm{final}}$ = final energy of the transferred electron; $E_{\mathrm{inital}}$ = initial energy of the transferred electron; $E_{\mathrm{photon}}$ = energy of the emitted photon; $F$ = Fano factor; $i$ = tube current; $I$ = radiation intensity; $I_{\mathrm{Cref}}$ = intensity of reference Compton peak; $I_{\mathrm{Csample}}$ = intensity of sample Compton peak; $I_{\mathrm{F}}$ = intensity of fluorescence radiation; $I_i$ = radiation intensity element; $k$ = constant; $k'_{ij}$ = coefficient for absorption correction; $k''_{ij}$ = coefficient for enhancement effect; $m'_{ij}$ = interelement coefficient for absorption and enhancement; $m''_{ij}$ = coefficient for line interference; $n$ = principal quantum number or order of diffraction; $N$ = number of electrons or number of ionizations per photon; $U_0$ = acceleration voltage; $Z$ = atomic number; $Z_{\mathrm{eff}}$ = effective atomic number; $\alpha_{\mathrm{crit}}$ = critical angle of total reflection; $\alpha_{ij}$ = correction coefficient for interelement matrix effects; $\Theta$ = angle between primary and secondary radiation; $\mu$ = mass attenuation coefficient; $\lambda$ = wavelength of radiation; $\rho$ = density; $\sigma$ = standard deviation.

*See also:* **Quantitative Analysis; Scanning Probe Microscopy, Theory; X-Ray Fluorescence Spectroscopy, Applications; X-Ray Spectroscopy, Theory.**

## Further reading

Argawal BK (1979) *X-ray Spectroscopy*. Berlin: Springer.

Bertin EP (1978) *Introduction to X-ray Spectrometric Analysis*. New York: Plenum Press.

Buhrke VE, Jenkins R and Smith DK (1998) *Preparation of Specimens for X-ray Fluorescence and X-ray Diffraction Analysis*. New York: Wiley.

Hahn-Weinheimer P, Hirner A and Weber-Diefenbach K (1995) *Röntgenfluoreszenzanalytische Methoden. Grundlagen und praktische Anwendung in den Geo-, Material- and Umweltwissenschaften*. Wiesbaden: Vieweg, Braunschweig.

Heckel J (1995) Using BARKLA polarized X-ray radiation in energy dispersive X-ray fluorescence spectrometry. *Journal of Trace Microprobe Techniques* 13: 97–108.

Jenkins R, Gould RW and Gedcke D (1981) *Quantitative X-ray Spectrometry*. New York: Marcel Dekker.

Klockenkämper R (1997) *Total Reflection X-ray Fluorescence Analysis*. New York: Wiley.

Lachance GR and Claisse F (1995) *Quantitative X-ray Fluorescence Analysis*. New York: Wiley.

Whiston C (1987) *X-ray Methods. Analytical Chemistry by Open Learning*. Chichester: Wiley.

Williams KL (1987) *Introduction to X-ray Spectrometry*. London: Allen & Unwin.

# X-Ray Fluorescence Spectroscopy, Applications

**Christina Streli, P Wobrauschek** and
**P Kregsamer,** Atominstitut of the Austrian
Universities, Wien, Austria

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**

**Applications**

## Introduction

X-ray fluorescence spectrometry (XRF) has been applied during the 1970s to 1990s as a versatile tool to many analytical problems. The analysis of major, minor and trace elements in various kinds of samples can be performed qualitatively as well as quantitatively. The working principle is based on the excitation of the sample atoms by high-energy X-rays, followed by the emission of characteristic photons with a certain energy, well correlated to the atomic number *Z* of each element (Moseley's law). The determination of the energy (or wavelength) of the emitted photon allows qualitative analysis and the determination of the number of emitted characteristic photons allows quantitative analysis. The fundamental physical principle of X-ray fluorescence is described in the article about the theory of X-ray fluorescence spectroscopy. One of the features of XRF besides the accurate, rapid, multielement capacity, is that the analysis can be performed nondestructively. In fact, one has to consider that XRF is a surface-sensitive method, because of the energy of the excited and emitted radiation which is in the range 1–115 keV (Na to U K-radiation). The penetration depth of the primary radiation is some μm or so for low-*Z* elements and some 100 μm or so for heavy elements, depending also on the matrix or type of sample (solid, liquid and powder samples are common). In any case the surface of the object of investigation has to be representative of the entire volume, and thus requires a homogenous sample. Therefore, because of the sample preparation, the 'nondestructiveness' is lost. To perform XRF a spectrometer is required that consists of an excitation source, sample and a detection system, which can be either wavelength-dispersive or energy-dispersive.

The excitation is mostly performed by X-rays produced in and emitted by an X-ray tube. The spectral distribution of the emitted radiation is partly the bremsstrahlung, with a maximum energy corresponding to the applied voltage, and is partly superimposed by the characteristic lines of the respective anode material. The intensity of the emitted radiation depends on the atomic number of the target and the applied voltage. The intensity of the measured fluorescence signal depends on the intensity and energy of exciting photons hitting the sample atoms. Low-power X-ray tubes operating in the few W range and standard X-ray tubes dissipating up to 3 kW, as well as X-ray tubes with a rotating anode, up to 18 kW, are in use. Photons from radioisotopes are used for special applications, offering an excitation source independent of any power supply. The brightest excitation source is synchrotron radiation. It is emitted when bunches of electrons or positrons with energies in the GeV range are travelling along curved sections in a storage ring. An intensive continuous spectrum with a strong natural collimation in the forward direction is emitted. The radiation is continuous from the eV region to some hundreds of keV and linear polarized in the orbital plane.

Due to the different working principles of wavelength-dispersive (WD) and energy-dispersive (ED) XRF the applications differ strongly and it is necessary to work out the special techniques and their advantages and disadvantages.

## Instrumentation and methodology

### Wavelength-dispersive spectrometers

Two types of WD instruments are in use, the sequential spectrometer and the simultaneous spectrometer. The sequential spectrometer scans the radiation emitted by the sample by changing the angle sequentially. For different wavelength regions different analyser crystals must be used to fulfil Bragg's equation. A set of 6 to 8 crystals with various lattice spacings *d* is properly mounted and can be changed automatically. The simultaneous spectrometers consist of various combinations of analyser crystals and detectors, arranged around the sample at fixed angle settings. So each 'channel' is optimized to detect an individual wavelength corresponding to an element. Mostly these channels use focusing optics at the detector to increase the signal. These spectrometers are called simultaneous multielement spectrometers, but the number of simultaneously detected elements depends on the number of channels of the spectrometer. LiF, topaz and other natural crystals are used for the

medium-$Z$ elements. The use of synthetic multilayer structures (consisting of alternating layers of a high-$Z$ and a low-$Z$ material with a bilayer thickness of ~1–10 nm) as dispersive elements and measurement in an evacuated environment allows the efficient determination of low-energy characteristic radiation down to even Be–K lines ($\lambda = 11$ nm) if a flow counter with an ultrathin entrance window is used. The big advantage of the WD spectrometers is their excellent wavelength-to-energy resolution, especially in the low-energy region and the high count-rate range ($10^6$ cps) in operation.

### Energy-dispersive spectrometers

An ED detector consists of a semiconductor crystal (Si, Ge) prepared as p-i-n diode, mounted under vacuum, generally operated at 77 K and cooled with liquid nitrogen (LN) and the necessary connected electronics. Therefore a large, heavy dewar (7–30 L of LN) is required for an ED detector. LN consumption is ~1 L a day. The crystal environment has to be a vacuum, and as an entrance window in front of the crystal generally Be is used. Be is chosen as it is available with 8–25 μm thickness, vacuum and light-tight and its absorption of low-energy photons is tolerable. If very low energy photons ($E < 1$ keV) are to be measured, an ultrathin (< 1 μm) entrance window is required. Generally, the energy resolution is much larger with ED detectors than with WD systems, so ED is more susceptible to line overlaps. Especially in the low-energy regions this leads to difficulties in the interpretation of the measured spectrum and mathematical procedures for spectrum deconvolution are required. To overcome the problem of LN cooling, new Peltier cooled detectors are available offering smaller size and lighter weight, but worse resolution. As common practice the value for the energy resolution is given at 5.9 keV and is in the range 130–180 eV for LN-cooled detectors, and 175–200 eV for Peltier cooled detectors.

ED spectrometers measure all photons coming from the sample simultaneously. On one hand this is an advantage, because many elements can be detected within a short time, but on the other hand it is a disadvantage, because the maximum count-rate is limited to 50–80 kcps. The processing of the measured signal from each photon requires a certain length of time and during that interval the system is not ready to process the signal from the next arriving photon. The result is a deadtime, which has to be corrected. The reason for the saturation is not the number of fluorescence photons from the sample mainly, but the exciting radiation being scattered by the sample and the sample carrier. Therefore special techniques have been developed to reduce the scattered radiation.

### Use of monochromatic radiation

The scattering can be drastically reduced, if monoenergetic radiation is used for excitation. Then only monoenergetic photons are arriving at the sample and can be scattered and so contribute to spectral background. Various methods of producing monoenergetic radiation are in use.

**Filtered radiation** The easiest way to reduce the number of photons not used for sample excitation is the insertion of a filter. It mainly absorbs the low-energy bremsstrahlung, but also a filter with $Z$ as $Z_{anode}$–1 can be used to effectively reduce the characteristic $K_\beta$ radiation from the anode.

**Secondary target excitation** The radiation from an X-ray tube is used to excite a suitable target and the fluorescence radiation from the secondary target is used to excite the sample. This method of achieving quasi-monochromatic radiation suffers from a tremendous loss of photon flux, but this can be compensated for by using higher current from high-power tubes.

**Radioisotope excitation (RIXRF)** A few radioisotope sources (with acceptable half-life) emit radiation in the X-ray region, such as Am-241 (59.5 keV), Cd-109 (Ag–K lines, 22 and 25 keV) and Fe-55 (Mn lines, 5.9 keV), and can be used for excitation. The advantage of radioisotope sources is their independence of a generator and power supply, which makes their use interesting for portable instruments, in field, on-line and even extraterrestrial applications. The disadvantage is the low photon flux in comparison to tube excitation.

**Crystal monochromators and multilayer structures** Crystal monochromators in the beam path of an X-ray tube allow the selection of either the characteristic line from the anode or the selection of an energy from the continuous spectrum. Mono-chromators with a high reflectivity as well as a large energy band width are usually preferred, because the product of these two parameters determines the photon flux on the sample. In comparison to crystal monochromators, the multilayer offers high reflectivity as well as larger band width ($dE/E = 10^{-2}$). Higher photon fluxes could be obtained. They are used with either X-ray tubes or synchrotron radiation.

### XRF using a linear polarized beam

If linear polarized radiation is scattered, in the ideal case no scattering radiation is emitted in the

direction of the polarization vector. This effect can be used to reduce the scattering from the sample itself. Barkla polarizers or Bragg polarizers can be used. The Barkla polarizer scatters the entire spectrum of the exciting radiation; the Bragg polarizer acts additionally as a monochromator. Both polarizers use the scattering of the unpolarized radiation through an angle of 90°. Using polarized radiation the spectral background is drastically reduced in comparison to nonpolarized radiation, but again losses in intensity occur.

#### Total reflection X-ray fluorescence analysis (TXRF)

TXRF is an EDXRF technique, utilizing the total external reflection of X-rays on the smooth plane surface of a reflector material, e.g. polished quartz. If a low divergent beam impinges on the reflector surface at an angle smaller than the critical angle for total reflection, most of the beam is reflected from the surface; only a small part penetrates into the reflector, causing scattered radiation. This leads to a reduced spectral background. The fluorescence signal is enhanced because the primary and also the reflected beam excite the sample, which is deposited on the reflector. Due to the small incidence angle the detector can be brought very close to the sample, so the detection efficiency is high. All these features lead to detection limits in the range of pg. Generally the samples have to be in liquid form. A droplet of 2–100 µL is pipetted onto the sample reflector and the liquid matrix is evaporated. Also, thin films or thin layers, as well as atoms implanted in a reflecting material such as Si-wafers, can be measured and thickness and depths determined.

#### Microfluorescence analysis (MXRF)

Microfluorescence analysis indicates the analysed area to be very small, leading to spatially resolved information of the sample composition. There are several methods of obtaining a beam with a small diameter, from a simple pinhole to highly sophisticated X-ray optical elements with focusing characteristics. One very effective method is the use of capillaries using the principle of total reflection of X-rays on the inner walls of a glass capillary. Small diameters down to 1 µm can be obtained with satisfactory intensity.

#### Synchrotron radiation induced XRF (SRXRF)

Synchrotron radiation, as described above, offers several advantages for use as an excitation source for XRF, especially the higher intensity, orders of magnitude greater than that offered by an X-ray tube. In combination with a monochromator the exciting radiation can be tuned to the energy with the optimum value of the photoelectric cross section for the investigated sample. It is also possible to tune the energy below the absorption edge of a main element in the sample to excite an element at trace levels with $Z < Z_{\text{main element}}$. This method is called *selective excitation* and offers several advantages in trace element analysis. To perform microanalysis, focusing elements are inserted to produce high-intensity microbeams. Also TXRF can be done using SR as exciting radiation.

#### Trace element analysis

To detect elements at trace levels – µg g$^{-1}$ (ppm), ng g$^{-1}$(ppb) or even pg g$^{-1}$ (ppt) – with XRF, generally special techniques as well as special sample preparation methods have to be used. The relevant quantity for trace element analysis is the limit of detection (DL), which is given by the formula

$$\mathrm{DL} = \frac{3 \cdot \sqrt{I_{\mathrm{B}}/t}}{S}$$

$I_B$ is the background intensity, $t$ the measuring time and $S$ the sensitivity (cps ppm$^{-1}$ or cps ng$^{-1}$). Either increasing the sensitivity or reducing the background leads to a reduction of detection limits. Therefore, special techniques of XRF, mainly EDXRF techniques, are applied, like TXRF, MXRF or SRXRF, as methods for trace element analysis. Detection limits range from ng to fg or µg g$^{-1}$ to pg g$^{-1}$.

## Applications

The applicability of XRF is almost unlimited with respect to type of sample and concentration range, but depends very much on the chosen technique. It can be used for on-line analysis in production processes or in-field measurements of geological samples with portable instruments, or quality control, such as impurity determination on Si-wafer surfaces at ultratrace levels, and environmental investigations. Also sample preparation is an important factor influencing the applicability (see **Figure 1**). For fine art object investigation or precious objects from museums, absolute nondestructiveness is required, contrary to environmental sampling where homogenizing and pressing to a pellet is no problem.

To describe the various fields of application in detail they are divided into categories.

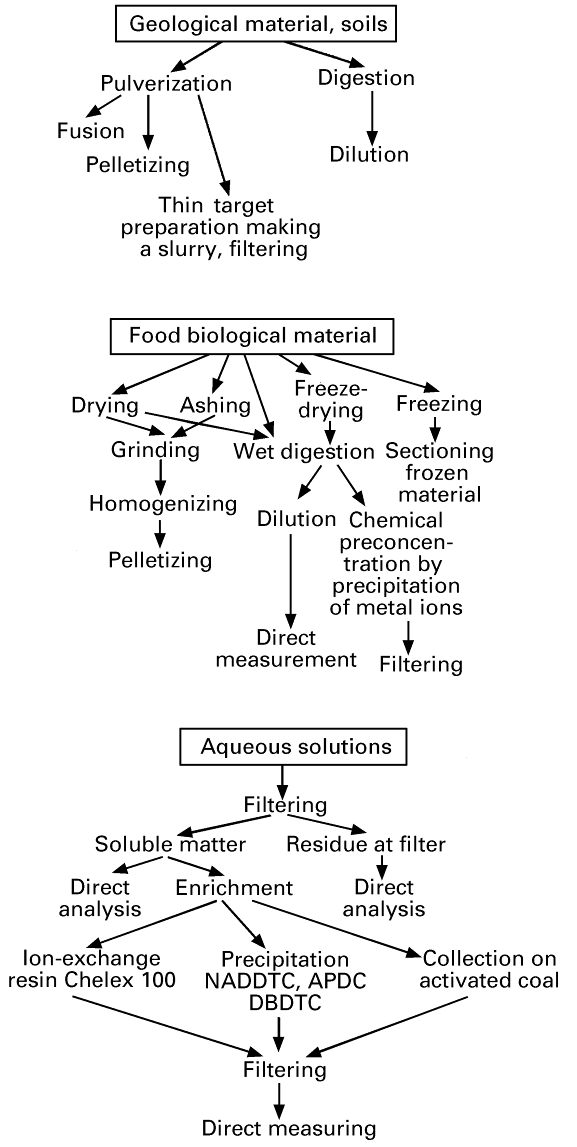



**Figure 1** Overview of sample preparation techniques typical for XRF.

## Metals and alloys

Process control in today's highly automated production facilities is strongly dependent upon fast, precise and accurate chemical analysis, and XRF has been found to be widely applicable in the metal industries. XRF of metallic samples includes several solvable problems, especially in the areas of sample preparation and modelling calculations to convert intensities into concentration data. In general, metallic samples do not need complicated sample preparation, but the analytical information is derived from a volume close to the surface which must be polished. XRF is applied to various kinds of alloys, such as Na–Mg alloys, and Al, Ti, ferrous, Ni, Cu, Zr, W and Au alloys, bronzes and brass. Typically, simultaneous WD spectrometers with an automatic sample changer are used.

## Mining and ore processing

In mining and ore processing XRF is used for quality and process control. The spectrometers used differ very much depending on their application. Laboratory spectrometers for quality control may be WD or ED systems whith tube excitation. The on-stream spectrometers are located at in-plant locations, can be WD and ED systems, with radioisotope excitation or X-ray tube excitation and equipped with flow cells. In-stream instruments can be installed in slurry streams, mainly equipped with radioisotope excitation and scintillation counters for single-element determination. Field instruments must be portable and battery operated. The big advantage of XRF over other analytical tools for that application is its simplicity and speed. The usefulness of X-ray instruments to selective mining is well established, the information being used for ore–waste sorting.

## Cement analysis

Finishing cements and raw mixes typically contain Ca, Si, Al and O at high concentrations, plus Fe, K, Mg and Na at low concentrations. One of the major problems in accurate cement analysis is homogeneity, particularly in the case of raw mixes, where the source of raw material may be variable. Most of the elements to be determined are of low atomic number, hence the penetration of their characteristic lines will be of the order of a few micrometres only. Careful grinding and pelletizing will suffice, but the fusion bead technique with $Li_3BO_4$ is strongly recommended. Simultaneous multichannel WD systems are ideally suited for this kind of application.

## Lubricating oil analysis

Raw or used oils are usually analysed for additive-element content including Ba, Zn, Mn, Ca, P and Cl, plus naturally occurring elements including S, N, Ni and Na. In blended stocks the concentration of these elements would typically lie in the range 0.01–2.5%. In a standard case the analysis will be performed in the liquid phase and under a He atmosphere. Large matrix effects are likely because of the variable concentration levels of relatively heavy elements in a very low average atomic number matrix.

## Geology

XRF offers a rapid, accurate, low-cost method of analysis. Fully automated XRF spectrometers, sequential as well as simultaneous WD instruments,

and ED instruments are in use. The analysis of geochemical samples often involves the analysis of samples having concentrations ranging from 0.0001 to 80%. Elements from Na to U are routinely determined. Detection limits depend very much on sample preparation and are in the range 20–1000 μg $g^{-1}$ for low-*Z* elements, 5–10 μg $g^{-1}$ for medium-*Z* elements and 1–20 μg $g^{-1}$ for high-*Z* elements.

One spectacular application of EDXRF was the Mars Pathfinder which was designed to inspect the rocks on the surface of Mars. Radioisotope excitation was used and one of the new electrically cooled ED detectors mounted on a small vehicle. Surface analysis of rocks could be performed quickly and the data could be transferred to Earth.

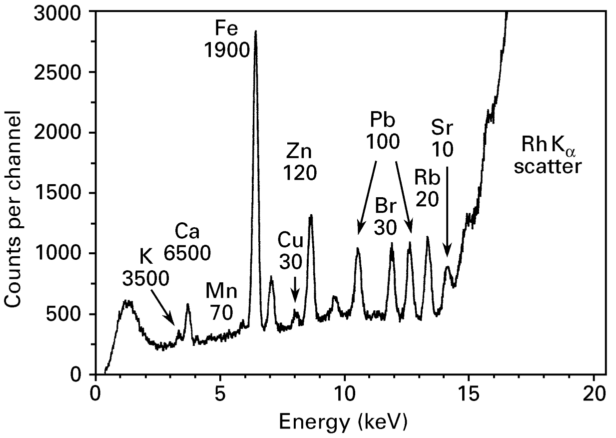



**Figure 2** Spectrum of lichen sample. Pressed pellet, measured with a standard EDXRF spectrometer, Rh tube 30 kV, 0.2 mA, 500 s, Pd filter, Si(Li) detector; concentration values of respective elements given in μg $g^{-1}$.

### Fine art and archaeological objects

In principle, with EDXRF, nondestructive analysis can be performed, which opens up the wide field of art and museum objects. Coins, bronzes, paintings, pottery, ceramics and ancient glasses can all be analysed. Paintings can be analysed pixel by pixel. With a spectrometer mounted on an *x–y* stage, a selected area can be analysed and the whole painting can be scanned. In particular, MXRF techniques are very well adapted to analysing specific points on figures or vessels, as well as lines on ancient documents.

### Environmental analysis

This is probably the most versatile and important application as all kinds of biological material such as plants, roots, needles, food-stuff, algae and lichens as biomonitors can be analysed with XRF. Lichens offer the advantage of a natural sampler collecting aerosols without exchange with the substrate. See **Figure 2** as example of a sample lichen measured with an EDXRF spectrometer. Soil, river sediments, sewage sludge, dust, coal fly ash, car exhaust and fog condensate or the aerosols, sampled in impactor stages, are well suited for XRF. All kinds of liquids can be analysed: river water, sea water, snow, ice. However, special techniques for sample preparation should be applied especially when trace element analysis is required.

### Microfluorescence analysis

MXRF allows spatial resolution of analysis. To perform it with sufficient intensity in the laboratory with X-ray tubes, special optics should be applied such as capillaries. Using capillaries, X-ray tube excitation with SR allows much higher intensities, and so lower detection limits can be obtained. Applications range from microelectronics and plating thickness, maps of bone cross-sections, superconductor films, human hair, pig heart muscles, metal alloys, leaves, chinaware, environmental particles, tree rings and glass fragments.

### Medicine

*In vivo* measurements as well as *in situ* measurements and analysis of malignant cells, as well as tissue sampling have been performed. *In vivo* measurements started with the determination of iodine in the thyroid and range from Cd in liver and kidney and Pt in kidneys and tumours, to Hg in the wrists and skulls of dentists, Pb in various near-surface bones, Cu in the eye and Fe in the skin. For the *in vivo* measurements, the use of polarized radiation offers big advantages. Hg and Pb can be analysed by their K-radiation thus giving information from even deeper tissues. Also, the analysis of biopsy samples, whole blood and blood serum (Se) can be performed using EDXRF. Trace element content in malignant and benign tissues was investigated, as well as lung tissue from different factory workers, showing different elements at higher concentrations corresponding to their profession. Various body fluids were analysed, and correlation between trace element content and diseases found. EDXRF, TXRF and SRXRF were used, depending on the task.

### TXRF applications

TXRF extends EDXRF to a method for trace and ultratrace element analysis. A special feature is the small sample amount required, which is in the range μg–ng of a solid material and less than 100–10 μL of a liquid. Therefore, TXRF is a microanalysis method

and samples can seldom be analysed as-received. Pretreatment is generally required to prepare the samples as solutions, suspensions, fine powders or thin sections. For a determination of ultratraces, the matrix of the sample should be separated and removed. All preparation techniques can be applied that have been tested with other methods of atomic spectrometry, e.g. AAS, inductively coupled plasma mass spectrometry. The quantification is generally very simple, because the sample forms a thin layer, so the thin-film approximation is valid. One element of known concentration has to be added as internal standard. Quantification can be performed after the determination of the sensitivity factors of all elements relative to the internal standard. Also, surface and thin-layer analysis, as well as analysis of atoms below a reflecting surface, can be performed by varying the angle of incidence in the region of total reflection. This angle-dependent intensity profile allows a qualitative differentiation between contaminations on the surface, in the layers, in implantations or in so-called residues after evaporation of liquid samples. Quantitative determinations can be made by applying an algorithm deduced from theory in combination with an external standard.

It is obvious that the sample preparation technique used influences the detection limits. **Table 1** shows this influence on various samples from different fields of application. **Table 2** gives an overview of applications of TXRF already analysed. **Figure 3** shows a spectrum of a water standard reference sample (NIST 1643c) obtained with a TXRF vacuum chamber, constructed at Atominstitut, Vienna. Generally, an excellent field of application of TXRF in trace element analysis can be seen in liquid samples. All kinds of liquids, ranging from different kinds of water to acids and oils, as well as body fluids, can be analysed. Environmental samples, like airborne particles, plant material or medical and biological samples such as tissue can be analysed directly on a reflector.

The main industrial application of TXRF is the surface quality control of Si-wafer material. Wafers offer the required flatness and are polished, so that they can be directly analysed by TXRF. Several commercial instruments have been developed as wafer analysers and some 100 instruments are utilized in the semiconductor industry. TXRF is capable of checking the contaminations brought in by different steps during the production process. The required sensitivity is now $10^9$ atoms $cm^{-2}$ for transition

**Table 1** Influence of sample preparation on detection limits in TXRF (after Klockenkämper R (1997))

| *Sample* | *Drying* | *Freeze-drying* | *Chemical matrix separation* | *Open digestion* | *Ashing* | *Suspension* | *Solution* | *Pressure digestion* | *Freeze cutting* |
|---|---|---|---|---|---|---|---|---|---|
| Rain, river water | 0.1–3 ng $mL^{-1}$ | 20–100 pg $mL^{-1}$ | 3–20 pg $mL^{-1}$ | 1–3 ng $mL^{-1}$ | | | | | |
| Blood, serum | | | Digest: 2–30 ng $mL^{-1}$ | 20–80 ng $mL^{-1}$ | 40–220 ng $mL^{-1}$ | | | | |
| Air dust, ash, aerosols | | | | 5–200 µg $g^{-1}$ | | 10–100 µg $g^{-1}$ | 0.1–3 µg $g^{-1}$ | | |
| Air dust on filter | | | | | 0.6–20 ng $cm^{-2}$ | | | 0.2–6 ng $cm^{-2}$ | |
| Suspended matter | | | | 3–25 µg $g^{-1}$ | | 10–100 µg $g^{-1}$ | | | |
| Sediment | | | | | | 10–100 µg $g^{-1}$ | | 15–300 µg $g^{-1}$ | |
| Powdered biomaterial | | | | | | 1–10 µg $g^{-1}$ | | 0.2–2 µg $g^{-1}$ | |
| Fine roots | | | | 1–10 µg $g^{-1}$ | Digest: 0.1–1 µg $g^{-1}$ | | | | |
| High-purity acids | 5–50 pg $mL^{-1}$ | | | | | | | | |
| Tissue, foodstuff, biomaterial | | | | | | | | | 0.5–5 µg $g^{-1}$ |
| Mineral oil | | | | | | | 1–15 µg $g^{-1}$ | | |
| Mussel, fish | | | | | | | | 0.1–1 µg $g^{-1}$ | |
| High purity water | | 1 pg $mL^{-1}$ | | | | | | | |

**Table 2** Applications of TXRF

| | |
|---|---|
| *Environment* | |
| Water | Rain, river, sea, drinking water, waste water. |
| Air | Aerosols, airborne particles, dust, fly ash. |
| Soil | Sediments, sewage sludge. |
| Plant material | Algae, hay, leaves, lichen, moss, needles, roots, wood. |
| Foodstuffs | Fish, flour, fruits, crab, mussel, mushrooms, nuts, vegetables, wine, tea. |
| Various | Coal, peat. |
| *Medicine/biology/pharmacology* | |
| Body fluids | Blood, serum, urine, amniotic fluid. |
| Tissue | Hair, kidney, liver, lung, nails, stomach, colon. |
| Various | Enzymes, polysaccharides, glucose, proteins, cosmetics, biofilms. |
| *Industrial/technical applications* | |
| Surface analysis | Si-wafer surfaces, GaAs-wafer surfaces. |
| Implanted ions | Depth and profile variations. |
| Thin films | Single layers, multilayers. |
| Oil | Crude oil, fuel oil, grease. |
| Chemicals | Acids, bases, salts, solvents. |
| Fusion/fission research | Transmutational elements in Al Cu, iodine in water. |
| *Mineralogy* | |
| Ores, rocks, minerals, rare earth elements. | |
| *Fine arts/archaeological/forensic* | |
| Pigments, paintings, varnish. | |
| Bronzes, pottery, jewellery. | |
| Textile fibres, glass, cognac, dollar bills, gunshot residue, drugs, tapes, sperm, fingerprints. | |

elements like Cr, Fe, Co, Ni, Cu and Zn. TXRF has an 'up-time' of 90% and is nondestructive. Surface mapping can be performed and differentiation between film type or particle type is possible. The detection limits can be improved by more than two orders of magnitude, if the impurities of the entire surface of the wafer are collected and preconcentrated prior to TXRF analysis. The native oxide layer is dissolved by HF vapour and the impurities remaining on the surface are collected by scanning the wafer with a drop of a suitable liquid. This method has the advantage of higher sensitivity, but nondestructiveness is lost.

It is also possible to measure the thickness of near-surface layers in the range 1–500 nm on reflecting substrates with TXRF. Single layers as well as multilayer samples can be analysed. Also, atoms implanted in the reflecting surface can be detected. The implantation depth as well as the depth profile can be determined.

TXRF can also be applied to fine art and museum objects. The sampling technique – a dry cotton bud can be used to rub off a small amount of paint – can only be applied during restoration, because the varnish has to be removed. For analysis the bud is dipped onto a sample carrier by a single tip. An amount of less than 100 ng is transmitted and can be analysed.

## Application of SRXRF

The rapid development of the SR X-ray sources since about 1975 is starting to have an impact on X-ray analysis. Due to the features of SR, especially the small source size and therefore the high brilliance, the use of microprobes is obvious. There are several approaches to producing a microbeam; the simplest is to use a pinhole collimator, but more sophisticated systems use focusing optics. Monochromatic as well as continuous radiation is used. Because SR is linearly polarized in the orbital plane, scattering from the sample is reduced, leading to low detection limits. SRXRF is a trace element analytical method as well as an MXRF method. SRXRF is performed at several SR facilities, most prominently NSLS Brookhaven, HASYLAB Hamburg, SSRL Stanford, SLS Daresbury, Photon Factory Tsukuba, DCI Lure and ESRF Grenoble. The available spot sizes are in the region of 10 μm and the detection limits in the low pg to fg range. Interesting applications were found in the fields of geology (mineral inclusions can be analysed), as well as in biology and medicine (distribution of trace elements in bone, tooth, brain, hair or algae strands, tree rings and aerosol particles), giving interesting information. Also, application in archaeology is found, letters in different ancient papers being analysed to allow the differentiation of the ink used to help identify the workshop. Even extraterrestrial minerals and rocks have been analysed.

TXRF can also be done using SR as the excitation source (SR-TXRF) offering the advantage of higher photon flux and improved detection limits. Experiments are performed at SSRL, HASYLAB, Photon Factory and ESRF. The main application is the surface quality control of Si wafers. With SR-TXRF, detection limits on wafer surfaces of $10^7$ atoms cm$^{-2}$ have been obtained. At SSRL there is a beamline dedicated for routine wafer analysis. SR is the ideal source for the excitation of low-*Z* elements. It offers high intensity also in the low-energy region, in comparison to standard X-ray tubes. They do not emit enough intensity in the low-energy region; therefore the analysis of low-*Z* elements always lacks intensity

**Table 3** Overview of various applications (from Török S and Van Grieken R (1994) X-ray spectrometry. *Analytical Chemistry* **66:** 186R–206R; Török S, Labar J, Injuk J and Van Grieken R (1996) *Analytical Chemistry* **68**: 467R–485R)

| *Field* | *Method* | *Field* | *Method* |
|---|---|---|---|
| *Archaeology* | | Se in soil | SR-TXRF |
| General | MXRF | Soil, marine sediments | WDXRF |
| Paintings | TXRF | Impurities in ice | TXRF |
| Obsidian | EDXRF | V, Ni, in oil, asphaltene | EDXRF |
| Medals | WDXRF | Bitumen solutions | EDXRF |
| Pottery | WDXRF | *Geology* | |
| Pigments | TXRF | Rocks | WDXRF |
| *Biomedical* | MXRF | Mineral grains | SRXRF |
| General | SR-TXRF | Soils, sediments | MXRF |
| Single-cell analysis | EDXRF, TXRF | Fossilized bone | WDXRF |
| Biopsy samples | EDXRF, RIXRF | Crysolite | EDXRF |
| Blood | EDXRF | Geological samples | WDXRF |
| Skin *in vivo* | RIXRF | Phosphate in rocks | EDXRF |
| Bone | EDXRF | Oxides, silicates, carbonates | WDXRF |
| Bone *in vivo* | RIXRF | Liquid petroleum products | EDXRF |
| Leaves | TXRF, WDXRF | Microlayer of Fe–Mn nodules | SRXRF |
| Hair | EDXRF | Au in micas | SRXRF |
| Vegetables | TXRF | *Materials science* | |
| Plants | WDXRF | Thin-film characterization | XRF, EDXRF |
| Lichens | WDXRF | Impurities on Si-wafers | TXRF, SRXRF |
| Moss | RIXRF | Multilayers | EDXRF |
| Cd in kidney | RIXRF | Superconductors | WDXRF |
| Mussel shells | WDXRF | Zirconium oxide | WDXRF |
| Cu, Se, Zn in kidney | SR-TXRF | Alumina | WDXRF |
| Marine bivalve shells | SRCXRF | Cu corrosions | EDXRF |
| Pt in tumour tissues | EDXRF | Ultrapure reagents | TXRF |
| Cu in human serum | EDXRF | Glasses | WDXRF, EDXRF |
| Pb in bones, serum, blood | RIXRF | Ferroalloys | EDXRF |
| Fe *in vivo* | RIXRF | High-purity Cu | EDXRF |
| Hg *in vivo* | RIXRF | GaAs-wafers | TXRF |
| Amniotic fluids | TXRF | Electrolytic solutions | RIXRF |
| Plankton in polluted lakes | EDXRF | $Al_2O_3$ thin films | WDXRF |
| Meadow moths | SRXRF | Plastic materials | EDXRF, WDXRF |
| *Environmental* | | Ceramic materials | WDXRF |
| Aerosols | WDXRF, EDXRF, | Hf in Zr matrix | EDXRF |
| | TXRF, SRXRF | Ga in polyurethane foam | WDXRF |
| Fly ash | SRXRF | P in PbO films | EDXRF |
| Rain water | TXRF | Cu, Sr, Bi film on MgO | RIXRF |
| River water | TXRF | Molybdate crystals | EDXRF |
| Sea water | TXRF | Ferrous alloy | WDXRF |
| Sediments, suspensions | EDXRF, TXRF | Ta in Ti–Ta alloys | WDXRF |
| Waste | EDXRF | Pb in houseware | RIXRF |
| Coal | EDXRF, WDXRF | Textile fibres | TXRF |
| Dust | EDXRF | W analysis | TXRF |
| Pb in dust | WDXRF | HTSC films | SRXRF |

HTSC = High-temperature semiconductor.

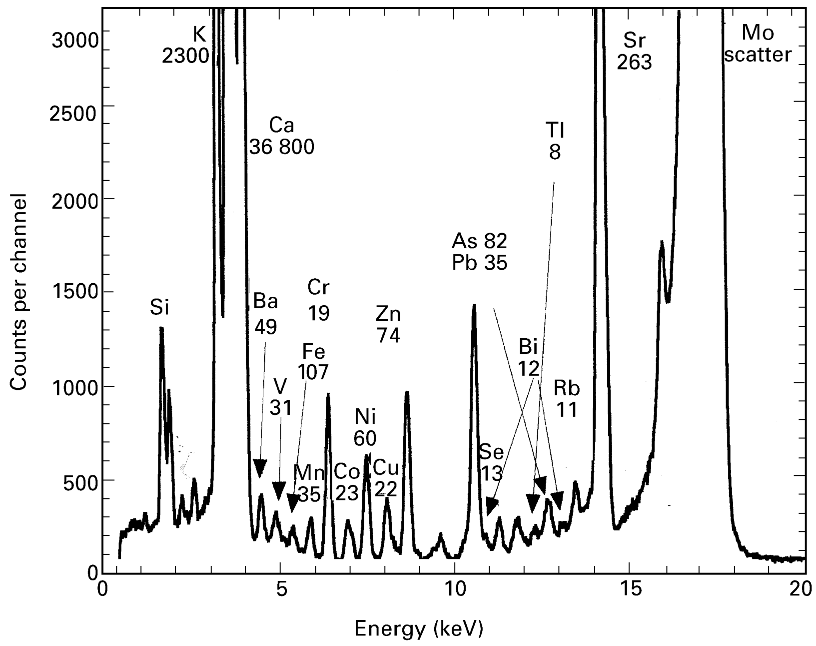



**Figure 3** Spectrum of water sample, NIST 1643c standard reference material, 10 μL, dried on a quartz reflector, measured with the TXRF vacuum chamber of Atominstitut, Mo monochromatic excitation, 40 kV, 50 mA, 1000 s; concentration values of respective elements given in μg $L^{-1}$.

of the fluorescence lines. TXRF is also applied to the determination of low-Z elements using a special detector. Detection limits of 60 fg for Mg have been achieved using SR and have to be compared to 7 pg with windowless Si-anode tube (prototype) excitation.

SR-TXRF generally is a very fast growing field and the problem of reduced access to SR sources for routine analytical applications is becoming less severe due to the large number of dedicated facilities.

**Table 3** gives an overview of applications and techniques published from 1994 to 1998.

## Conclusions

Applications range from on-line analysis and in-field inspections to ultratrace analysis of semiconductor surfaces. There is almost no sample that cannot be analysed by XRF as long as elemental analysis is required. The achievable detection limits depend on the method used and range from μg $g^{-1}$ (ppm) to pg $g^{-1}$ (ppt). Nondestructive analysis can be performed but sometimes sophisticated sample preparation techniques are required. The elemental range (Be to U) depends on the excitation source as well as the detection system. Generally XRF can be seen as a work-horse for elemental analysis and is easy to use.

## List of symbols

*d* = lattice spacing; DL = detection limit; $I_B$ = background intensity; *S* = sensitivity; *t* = measuring time; *Z* = atomic number; λ = wavelength.

*See also:* **Environmental and Agricultural Applications of Atomic Spectroscopy; Environmental Applications of Electronic Spectroscopy; Geology and Mineralogy, Applications of Atomic Spectroscopy; Inorganic Compounds and Minerals Studied Using X-Ray Diffraction; IR and Raman Spectroscopy Studies of Works of Art; IR Spectroscopy Sample Preparation Methods; MRI of Oil/Water in Rocks; X-Ray Fluorescence Spectrometers.**

## Further reading

Bertin EP (1978) *Introduction to X-ray Spectrometric Analyis*. New York: Plenum Press.

Carpenter DA (ed) (1997) Special Issue on Micro X-ray Fluorescence Analysis. *X-ray Spectrometry* **26**(6):

Ellis A, Potts Ph, Holmes M, Oliver GL, Streli C and Wobrauschek P (1996) Atomic spectrometry update: X-ray fluorescence spectrometry. *Journal of Analytical Atomic Spectrometry* **11**: 409R–442R.

Ellis A, Potts Ph, Holmes M, Oliver GL, Streli C and Wobrauschek P (1997) Atomic spectrometry update: X-ray

fluorescence spectrometry. *Journal of Analytical Atomic Spectrometry* **12**: 461R–490R.
Herglotz HK and Birks LS (eds) (1978) *X-ray Spectrometry*. New York: Marcel Dekker.
Holynska B (1993) Sampling and sample preparation in EDXRS. *X-ray Spectrometry* **22**: 192–198.
Iida A and Gohshi Y (1991) Trace element analysis by X-ray fluorescence. In: Ebashi S, Koch M and Rubenstein R (eds) *Handbook on Synchrotron Radiation*, Vol. 4, pp 307–349. Amsterdam: North Holland, Elsevier.
Jenkins R, Gould RW and Gedcke D (1981) *Quantitative X-ray Spectrometry*. New York: Marcel Dekker.
Klockenkämper R (1997) *Total-Reflection X-ray Fluorescence Analysis*. New York: Wiley.
Sparks CJ (1982) X-ray fluorescence microprobe for chemical analysis. In: Winick H and Doniach S (eds) *Synchrotron Radiation Research*, pp 459–509. New York: Plenum Press.
Török S and Van Grieken R (1994) X-ray spectrometry. *Analytical Chemistry* **66**: 186R–206R.
Török S, Labar J, Injuk J and Van Grieken R (1996) X-ray spectrometry. *Analytical Chemistry* **68**: 467R–485R.
Van Grieken R and Markowicz A (eds) (1993) *Handbook of X-ray Spectrometry*. New York: Marcel Dekker.
Wielopolski L and Ryon RW (eds) (1995) Workshop at the Denver X-ray conference on *in vivo* XRF measurement of heavy metals. *Advances in X-ray Analysis* **38**: 641.

# X-Ray Spectroscopy, Theory

**Prasad A Naik**, Centre for Advanced Technology, Indore, India

Copyright © 1999 Academic Press

| HIGH ENERGY SPECTROSCOPY |
|---|
| Theory |

X-ray is the region of the electromagnetic spectrum lying between gamma rays and extreme ultraviolet (XUV / EUV) corresponding to a wavelength range of about 0.1 to 100 Å. The radiation on the lower end of the XUV region, up to about 300 Å, is also sometimes referred to as X-ray. On the lower wavelength side, radiation of shorter wavelengths is termed X-ray if it is nonnuclear in origin. The wavelength of the radiation is related to the photon energy by the standard relation $E$ (keV) = 12.4/λ(Å). In terms of energy, the X-ray region is roughly between 125 eV and 125 keV. Being electromagnetic radiation, X-rays can be reflected, refracted, scattered, absorbed, polarized etc. They also show interference and diffraction effects.

There are several sources of X-rays such as a Coolidge tube, vacuum sparks, hot-dense fusion plasmas, synchrotron, pinch devices, muonic atoms, beam-foil interaction, stellar X-ray emitters, solar flares, etc. The X-rays originating from all these sources can be broadly categorized into main types: (1) atomic inner shell transitions, (2) emission by free electrons, (3) X-rays from few electron systems. The basic spectroscopic aspects of the various types of X-rays are discussed in this article.

## X-rays from inner shell transitions in atoms

X-rays are produced when an electron in an outer shell of an atom jumps to an inner shell to fill an electron vacancy. The difference in energy is emitted as an X-ray photon. The vacancy giving rise to such a transition can be produced by an energetic photon, bombardment of charged particles ($e^-$, p, α ..), or by nuclear processes such as internal conversion, K-capture, etc. If a charged particle collision or a nuclear process produces the vacancy, the resulting X-ray emission is called primary. If the vacancy is produced by an X-ray photon, the subsequent emission is called secondary or fluorescence radiation.

In all these cases the singly ionized atom lowers its energy by emission of a photon of definite wavelength which is characteristic of the emitting atom. Hence, these X-rays are also called characteristic X-rays.

### Characteristic X-rays

The most energetic X-ray emission comes when a vacancy in a K shell ($n = 1$) is filled by an outer electron. Removal of a 1s electron from a neutral atom raises it to the highest energy state represented by $1s1s^{-1}$ or $1\ ^2S_{1/2}$ or $K_I$. Removal of a 2s electron

($2s2s^{-1}$) gives rise to $L_1$ state ($2\ ^2S_{1/2}$). Removal of a 2p electron from a filled 2p shell ($2p^5p^{-1}$) gives rise to $L_{II}$ ($2^2p_{1/2}$) and $L_{III}$($2\ ^2P_{3/2}$) states, respectively. Similarly, removal of a 3s,3p,3d electron from filled shells gives rise to ($3s3s^{-1}$, $3p^53p^{-1}$, $3d^93d^{-1}$). $M_I$($3\ ^2S_{1/2}$), $M_{II}$($3\ ^2P_{1/2}$), $M_{III}$($3\ ^2P_{3/2}$), $M_{IV}$($3\ ^2D_{3/2}$), $M_V$($3\ ^2D_{5/2}$) states.

The selection rules applicable to optical dipole transitions also apply to X-ray transitions. The rules are $\Delta L = \pm 1$, $\Delta j = 0, \pm 1$. Intensity rules are also the same as those applicable to optical transitions. The transitions obeying selection rules are called normal transitions. Not all transitions allowed by selection rules are observed. On the contrary, some transitions, which are not allowed by selection rules, are sometimes observed. These are called forbidden transitions. The observed lines were initially given names as per their observed line intensities. The Siegbahn notation used to name various observed lines is given in **Table 1**. As this nomenclature is intensity-based (e.g. $K_{\alpha 1}$ more intense than $K_{\alpha 2}$) and was adopted long before the origin of the lines was explained spectroscopically, this notation is somewhat confusing.

Moseley's law gives the frequency of line transition as $\nu(\text{cm}^{-1}) = KR(Z-\delta)^2$, where $R$ is the Rydberg constant, and $Z$ is the atomic number. For example, for $K_\alpha$: $\delta = 1$ and $K = 3/4$, which gives $\nu(\text{cm}^{-1}) = (3/4)R(Z-1)^2$. Similarly, for $L_{\beta 1}$, $\nu(\text{cm}^{-1}) = (1/2^2-1/3^2)R(Z-7.4)^2$.

The energy levels formed by single electron removal are similar to those of a hydrogen-like atom, obtained by replacing $Z$ by ($Z-\sigma$) and ($Z-s$) in the Sommerfeld formula, where $\sigma$ and $s$ are screening constants. The energy of a level is given by the modified Sommerfeld formula as

$$T = R\,(Z-\sigma)^2/n^2 + R\,\alpha^2\,(Z-s)^4$$
$$(n/k - 3/4)/n^4 + \mathrm{O}(\alpha^4, \alpha^6, \ldots)$$

**Table 1** Siegbahn notation for various inner shell X-ray transitions

| *Final level* | | | | | | *Initial level* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Shell structure | n | l | i | Optical notation | X-ray notation | K series | L series | | | M series | | | | |
| $1s\ 1s^{-1}$ | 1 | 0 | 1/2 | $^2S_{1/2}$ | $K_I$ | | | | | | | | | |
| | | | | | | $K_I$ | | | | | | | | |
| $2s\ 2s^{-1}$ | 2 | 0 | 1/2 | $^2S_{1/2}$ | $L_I$ | - | | | | | | | | |
| $2p^5\ 2p^{-1}$ | 2 | 1 | 1/2 | $^2P_{1/2}$ | $L_{II}$ | $\alpha_2$ | | | | | | | | |
| $2p^5\ 2p^{-1}$ | 2 | 1 | 3/2 | $^2P_{3/2}$ | $L_{III}$ | $\alpha_1$ | | | | | | | | |
| | | | | | | | $L_I$ | $L_{II}$ | $L_{III}$ | | | | | |
| $3s\ 3s^{-1}$ | 3 | 0 | 1/2 | $^2S_{1/2}$ | $M_I$ | - | - | $\eta$ | l | | | | | |
| $3p^5\ 3p^{-1}$ | 3 | 1 | 1/2 | $^2P_{1/2}$ | $M_{II}$ | $\beta_3$ | $\beta_4$ | - | - | | | | | |
| $3p^5\ 3p^{-1}$ | 3 | 1 | 3/2 | $^2P_{3/2}$ | $M_{III}$ | $\beta_1$ | $\beta_3$ | - | - | | | | | |
| $3d^9\ 3d^{-1}$ | 3 | 2 | 3/2 | $^2D_{3/2}$ | $M_{IV}$ | [$\beta_5$] | [$\beta_{10}$] | $\beta_1$ | $\alpha_2$ | | | | | |
| $3d^9\ 3d^{-1}$ | 3 | 2 | 5/2 | $^2D_{5/2}$ | $M_V$ | [$\beta_5$] | [$\beta_9$] | - | $\alpha_1$ | | | | | |
| | | | | | | | | | | $M_I$ | $M_{II}$ | $M_{III}$ | $M_{IV}$ | $M_V$ |
| $4s\ 4s^{-1}$ | 4 | 0 | 1/2 | $^2S_{1/2}$ | $N_I$ | - | - | $\gamma_5$ | $\beta_6$ | - | √ | √ | - | - |
| $4p^5\ 4p^{-1}$ | 4 | 1 | 1/2 | $^2P_{1/2}$ | $N_{II}$ | $\beta_2(\gamma_2)$ | $\gamma_2$ | - | - | √ | - | - | $\xi_2$ | - |
| $4p^5\ 4p^{-1}$ | 4 | 1 | 3/2 | $^2P_{3/2}$ | $N_{III}$ | $\beta_2(\gamma_1)$ | $\gamma_3$ | - | - | √ | - | - | √ | $\xi_1$ |
| $4d^9\ 4d^{-1}$ | 4 | 2 | 3/2 | $^2D_{3/2}$ | $N_{IV}$ | [$\beta_4$] | | $\gamma_1$ | $\beta_{15}$ | - | √ | $\gamma_2$ | - | - |
| $4d^9\ 4d^{-1}$ | 4 | 2 | 5/2 | $^2D_{5/2}$ | $N_V$ | [$\beta_4$] | - | - | $\beta_2$ | - | - | $\gamma_1$ | - | - |
| $4f^{13}\ 4f^{-1}$ | 4 | 3 | 5/2 | $^2F_{5/2}$ | $N_{VI}$ | - | - | - | - | - | - | - | $\beta_1$ | $\alpha_2$ |
| $4f^{13}\ 4f^{-1}$ | 4 | 3 | 7/2 | $^2F_{7/2}$ | $N_{VII}$ | - | - | - | - | - | - | - | - | $\alpha_1$ |
| | | | | | | | | | | | | | | |
| $5s\ 5s^{-1}$ | 5 | 0 | 1/2 | $^2S_{1/2}$ | $O_I$ | - | | $\gamma_8$ | $\beta_7$ | - | √ | √ | - | - |
| $5p^5\ 5p^{-1}$ | 5 | 1 | 1/2 | $^2P_{1/2}$ | $O_{II}$ | $\delta_2$ | $\gamma_4$ | - | - | √ | - | - | √ | - |
| $5p^5\ 5p^{-1}$ | 5 | 1 | 3/2 | $^2P_{3/2}$ | $O_{III}$ | $\delta_1$ | $\gamma_4$ | - | - | √ | - | - | √ | √ |
| $5d^9\ 5d^{-1}$ | 5 | 2 | 3/2 | $^2D_{3/2}$ | $O_{IV}$ | - | - | $\gamma_6$ | $\beta_5$ | - | √ | √ | - | - |
| $5d^9\ 5d^{-1}$ | 5 | 2 | 5/2 | $^2D_{5/2}$ | $O_V$ | - | - | | $\beta_5$ | - | - | $\varepsilon$ | - | - |

- : Forbidden by selection rules

[ ] : Forbidden but observed.

√: Allowed by selection rules but no name given

where α is the fine structure constant, $n$ is the total (principal) quantum number, and $k$ is the Sommerfeld original azimuthal quantum number; $k$ = 1,2 and 3 for s,p and d electrons, respectively.

A pair of terms having the same $n$, $s$, L but different $j$ (in the $^{2s+1}L_j$ Russel–Saunders notation) is called a spin-relativity doublet (earlier called a regular doublet). $L_{II}$–$L_{III}$, $M_{II}$–$M_{III}$,$M_{IV}$–$M_V$ are examples of such doublets. The screening constant σ (same for a doublet) depends on s, p, d sub levels but $s$ is the same for a spin-relativity doublet, almost independent of $Z$ ($s$ values: $L_I$:2.0, $L_{II}$–$L_{III}$: 3.5, $M_I$:6.8, $M_{II}$–$M_{III}$:8.5, $M_{IV}$–$M_V$:13).

A pair of terms having the same $n$, $s$ and $j$ but different L values is called as screening doublet (or irregular doublet). $L_I$–$L_{II}$, $M_I$–$M_{II}$, $M_{III}$–$M_{IV}$ are examples of such screening doublets.

From the first term of the Sommerfeld formula, one obtains the screening doublet law, which states that the difference between the square roots of the term values of a given doublet is constant, i.e. independent of $Z$. This term also gives the irregular doublet law which states that the difference between term values of an irregular (screening) doublet is a linear function of $Z$.

The second term of the Sommerfeld formula gives the separation in energy for a spin-relativity doublet as proportional to the fourth power of the screened atomic number, i.e. $(Z–s)^4$. This is referred to as the regular doublet law. For example, for the $L_{II}$–$L_{III}$ doublet (same σ), $\Delta\nu(\text{cm}^{-1}) = (R\alpha^2/16)(Z–3.5)^4$.

### Satellite lines

These are weaker lines appearing on the shorter wavelength side of the normal (characteristic) lines. They were initially referred to as nondiagram lines because unlike the normal lines, these lines did not fit conveniently in the energy level diagrams of that time. Later, it was realized that the energies can also be predicted using energy level diagrams of multiply charged ions (**Figure 1**).

The satellite lines are due to additional electron vacancy in a doubly ionized atom. Due to the absence of a second electron, the energy levels shift to the higher energy side (relative to those of a singly ionized atom) due to reduced Coulomb screening. As a result, single electron transitions in such atoms (doubly ionized) are at a slightly higher energy compared to those in singly ionized atoms. As the probability of creation of two electron vacancies in an atom is much smaller than that of a single vacancy, the intensity of satellite lines is much less than that of normal lines.

Satellite lines are denoted as α′, α″, α‴ ... where the higher number of primes implies lower intensity. If K–M (i.e. $K_I$–$M_{II}$,$M_{III}$) denotes a $K_\beta$ transition, then KL–LM denotes the satellite line $K_\beta'''$ which is due to an additional vacancy in the L shell (**Figure 1**).

**Hypersatellite** This is a special case of satellite lines wherein X-rays originate from atoms with two holes in the same inner shell. A K-hypersatellite line appears when an atom has initially two vacancies in its

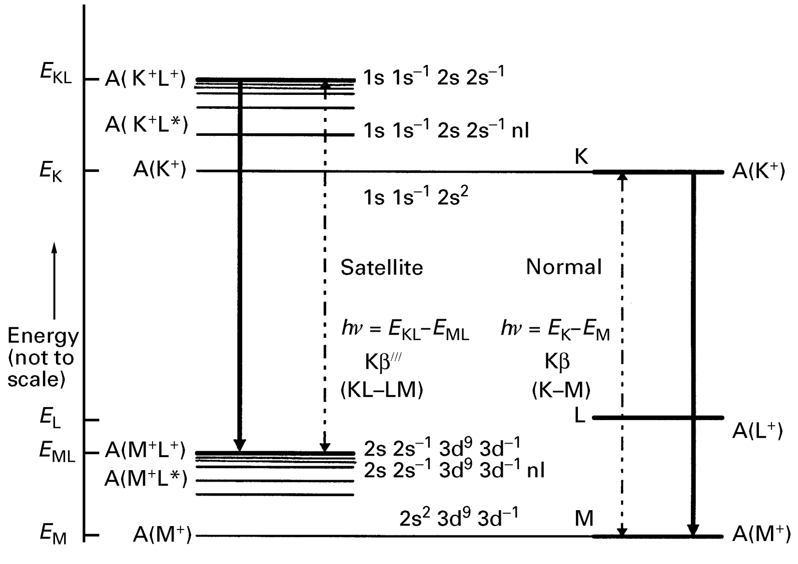



**Figure 1** Energy diagrams for a $K_\beta$ transition and Kβ‴ satellite transition. A() denotes the state of an atom. For example, A(M⁺L*) denotes an atomic level where an M shell electron has been removed and an L shell electron is in an outer bound (excited) state. The energy difference $E_{KL}$–$E_K$ is more than $E_{ML}$–$E_M$ as more energy is required to remove an L shell electron in the presence of a K shell vacancy than in the presence of a M shell vacancy. Solid downward arrows show vacancy transitions, and dashed lines with double arrows show radiative transitions.

K shell. They are denoted by superscript H. For example, $K^H_{\alpha1,2}$ is a hypersatellite of the $K_{\alpha1,2}$ line (**Figure 2**). Due to the strong reduction of screening, there is a large energy difference between a normal line and its hypersatellite. Hypersatellites are more easily observed in heavy ion collision spectra or radioactive decay by K-electron capture.

**Plasma satellites** These are low intensity structures, which can appear on either side of a parent line. They are equally spaced and correspond to an energy difference of $(h/2\pi)\omega_p$, where $\omega_p$ is the plasma frequency given by $\omega_p^2 = (N_e e^2/\varepsilon_0 m)$ where $N_e$ is the electron density in the conduction band of the solid target material.

If an X-ray photon loses its energy in exciting a plasmon, one obtains a low energy plasmon satellite. On the other hand, if plasmons already exist in the solid at the time of X-ray emission, they can lead to high energy satellites.

### Auger and allied processes

These processes compete with the radiative process in de-excitation of the atom. Thus they influence the energy-width of X-ray lines. Moreover, they transfer a vacancy from one level to another and thereby affect the intensity of X-ray lines. Many of these processes also lead to double ionization which (as discussed earlier) gives rise to satellite lines. Some of these processes are briefly described below.

**Auger ionization** In this case, an inner shell vacancy is filled by an electron from an outer shell and the excess energy instead of being emitted as a photon, is consumed in ejecting a second electron (called an Auger electron) from the atom. This is a radiationless process leading to double ionization. A transition of an electron from an L shell to a K shell vacancy, accompanied by the ejection of an L shell electron, is represented as a KLL transition. The energy of the ejected electron in KLL transition is $E_{Auger} = E_K - E_{LL}$. Here $E_K$ is the energy of the atom with one electron missing in the K shell and $E_{LL}$ is the energy of the level formed by ejection of an L shell electron from a singly ionized atom with an L shell electron vacancy (**Figure 3A**).

**Radiative Auger process** In this process, the Auger electron, instead of carrying away all the energy, receives only a part of it and the rest is emitted as an X-ray photon. The energy of the photon is given by $h\nu = E_{Auger} - E_{kin}$ where $E_{Auger}$ is the energy of the electron in the Auger transition involving the same three levels (e.g. KLL), and $E_{kin}$ is the kinetic energy of the ejected electron. The energy spectrum of the emitted X-rays is continuous, with a maximum energy $h\nu = E_{Auger}$, corresponding to zero kinetic energy of the ejected electron (**Figure 3B**).

**Semi-Auger process** If, in the above radiative Auger process, the electron, instead of being ejected out of the atom, is transferred to some outer bound state, then the energy of the emitted X-ray photon would be discrete having the value $(E_K - E_L) - E_b$. Here, $E_b$ is the difference in energy between the outer bound state and the energy level of the electron in an atom with a L shell vacancy (**Figure 3C**). This process is referred to as a semi-Auger process.

**Coster–Kronig process** This is a special case of Auger ionization in which the vacancy transition is

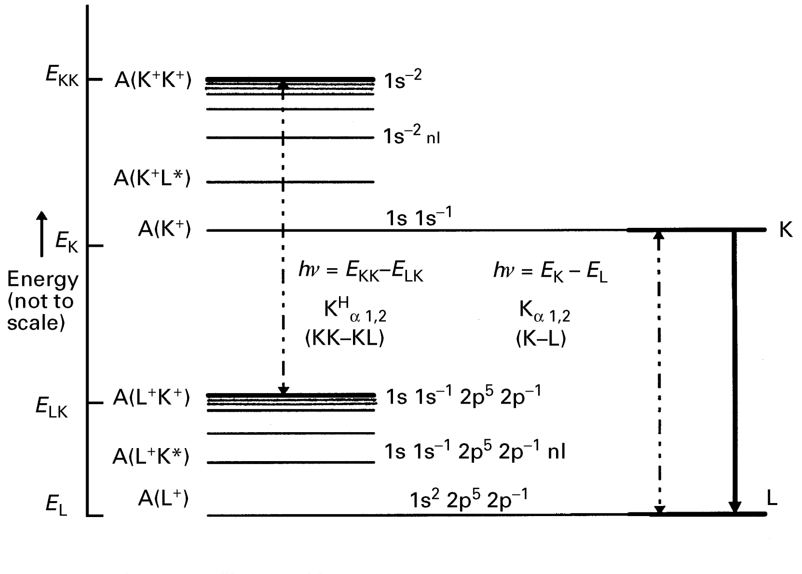



**Figure 2** Energy diagram for $K^H_{\alpha1,2}$ hypersatellite transition.

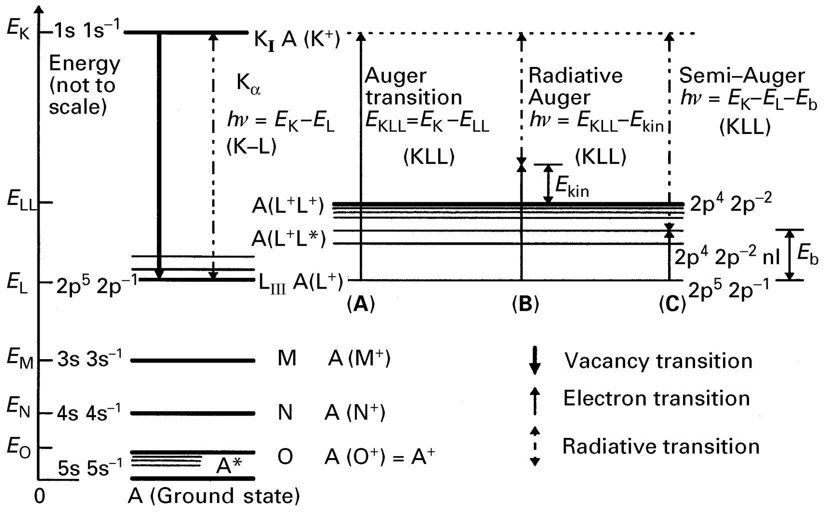



**Figure 3** Energy diagrams for: (A) Auger transition, (B) radiative Auger transition, and (C) semi-Auger transition.

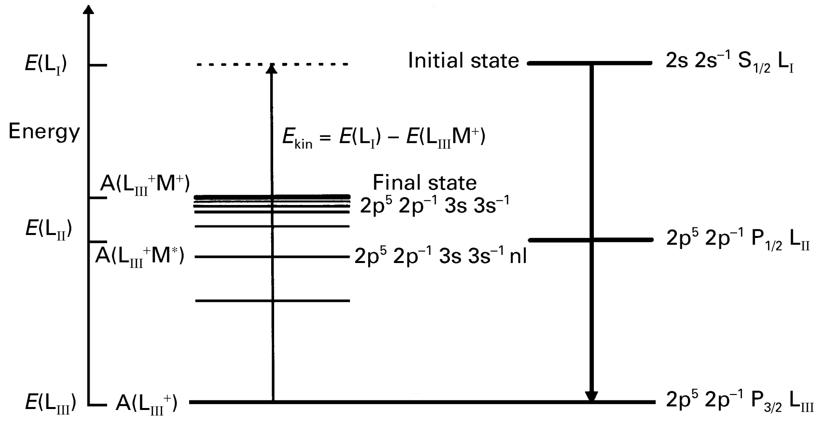



**Figure 4** Energy diagram for LLM Coster–Kronig transition.

between two levels with the same principal quantum number (i.e. within the same shell; $\Delta n = 0$) and an electron from some outer shell (different $n$) is ejected (e.g. LLM transition: see **Figure 4**).

**Super Coster–Kronig process** This is an Auger transition wherein not only the vacancy transition is within the same shell ($\Delta n = 0$ as in a CK transition), but the electron is ejected from the same shell (e.g. MMM transition: see **Figure 5**).

**Autoionization** This process is similar to Auger ionization. In this case a vacancy is created by promoting an inner shell electron to an outer bound level, instead of being ejected from the atom. When this vacancy is filled by an outer shell electron, if the difference in energy exceeds the ionization potential of any electron, then that electron is ejected leading to a singly ionized atom (**Figure 6**).

The differences between auto ionization and Auger ionization are as follows: a) Auto ionization results from an electron vacancy produced by the excitation of a core electron to an outer bound level whereas Auger ionization results from an electron vacancy created by ejecting a core electron from an atom. b) Autoionization takes place in a neutral atom leading to a singly ionized atom whereas Auger ionization takes place in a singly ionized atom leading to a doubly ionized atom.

There is a quite high probability (at least for K shell ionization) of the order of 20%, that the first ionization producing the initial vacancy is simultaneously accompanied by the excitation or ionization of a second electron. If the second electron is excited to some bound state, the process is called shake-up. If the second electron is ejected, the process is called shake-off.

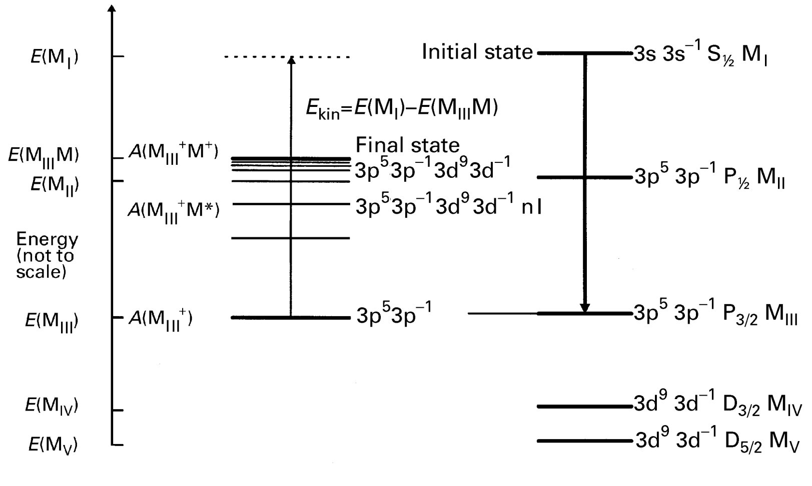



**Figure 5** Energy diagram for MMM super Coster–Kronig transition.

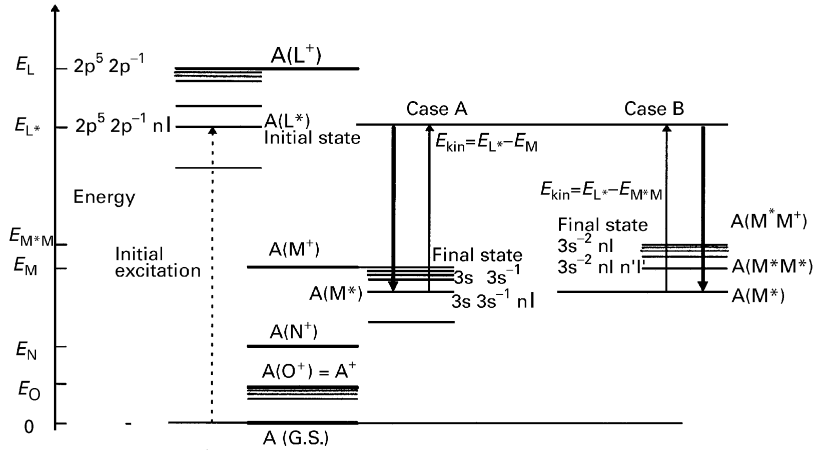



**Figure 6** Energy diagram for auto-ionization. An L shell electron is excited to an outer nl shell. Subsequently, an M shell electron fills up the L shell vacancy. In case A, the excited electron in the nl shell leaves the atom, and in case B, another M shell electron leaves the atom while the nl shell electron remains a spectator electron.

The Auger processes compete with radiative decay. The probability of radiative decay is proportional to $Z^4$, whereas the probability of Auger ionization is constant, independent of Z. The fluorescence yield ($\omega$) is defined as the ratio of the number of vacancies filled by X-ray emission to the total number of vacancies (filled by all processes). For K shell, $\omega_K$ = number of K X-ray photons/(number of K X-ray photon + number of Auger electrons) = $Z^4/(Z^4 + \alpha)$, where $\alpha = 1.12\times10^6$. Hence, for low Z atoms, the fluorescence yield is low due to strong competition from Auger processes (for $Z < 33$, $\omega_K < 1/2$).

Because satellite lines are due to doubly ionized atoms, and since Auger processes dominate for low Z atoms, satellite lines are more prominent in the spectra of low Z atoms compared to high Z atoms.

## X-ray emission by free electrons

From electromagnetic theory, any accelerating charge will radiate, with the intensity of emission being proportional to the square of the acceleration. Electrons, due to their smaller mass, undergo higher acceleration for a given force and hence emit more intense radiation. The spectrum of the emitted radiation called Bremsstrahlung, (braking radiation) depends on the nature and magnitude of the acceleration. We discuss here three main sources of X-ray based on electron acceleration (deceleration), namely (i) X-ray tubes (linear acceleration, monoenergetic electrons), ii) hot plasma sources (linear acceleration, Maxwellian distribution of energy) and iii) synchrotron sources (transverse acceleration). The X-ray spectra from these differ from each other.

### X-ray tube

In this case, thermionically generated electrons are accelerated by a d.c. (or pulsed) potential ($V$) to several keV energy and are abruptly slowed down by impinging them on a solid target. Due to this deceleration, Bremsstrahlung radiation is emitted. The maximum emission is in the direction perpendicular to the acceleration. Accordingly, the target surface in an X-ray tube is kept at an angle to the electron beam.

The total intensity of the X-ray emitted is given by $I \propto ZV^2$.

The efficiency of X-ray conversion is given by $\varepsilon$ = X-ray power/electron beam power. Experimentally, the value of $\varepsilon$ is about $1.1\times10^{-9}Z(V+10.3Z)$ or $\sim1.1\times10^{-9}\ ZV$ for large accelerating voltage. This means that the X-ray conversion efficiency is better for higher $Z$ targets and at higher accelerating voltages. However, even for a high-$Z$ material like tungsten ($Z = 74$) at a voltage of 50 kV, the efficiency is only about 0.4%. The rest of the $e^-$ beam energy is spent mostly in heating the target. It is therefore necessary to have an X-ray tube anode with a high melting point in either a rotating or cooled (or both) configuration to prevent it from melting.

The X-ray energy distribution for a thin target is given by

$$\begin{aligned} I_\nu &= I_o \quad \text{for } \nu \le \nu_{\max} \\ &= 0 \quad \text{for } \nu > \nu_{\max} \end{aligned}$$

where $h\nu_{\max}$ is the maximum energy of the X-ray photon, equal to the electron beam energy. This corresponds to the case where the electron loses its full energy in a single collision. It is also referred to as the Duane–Hunt limit.

The above spectral distribution is true for a thin target where the electron undergoes a single collision. However, in practice, the target is thick and the electron is scattered (decelerated) many times. As a result, the energy spectrum becomes

$$\begin{aligned} I_\nu &= KZ\,(\nu_{\max} - \nu) \quad \nu \le \nu_{\max} \\ &= 0 \quad \nu > \nu_{\max} \end{aligned}$$

The spectral distribution, in terms of X-ray wavelength, is given by (**Figure 7**).

$$\begin{aligned} I_\lambda &= c\, I_\nu/\lambda^2 \sim Kc^2 Z(\lambda - \lambda_{\min})/\lambda^3\ \lambda_{\min} \quad \text{for } \lambda \ge \lambda_{\min} \\ &= 0 \quad \text{for } \lambda < \lambda_{\min} \end{aligned}$$

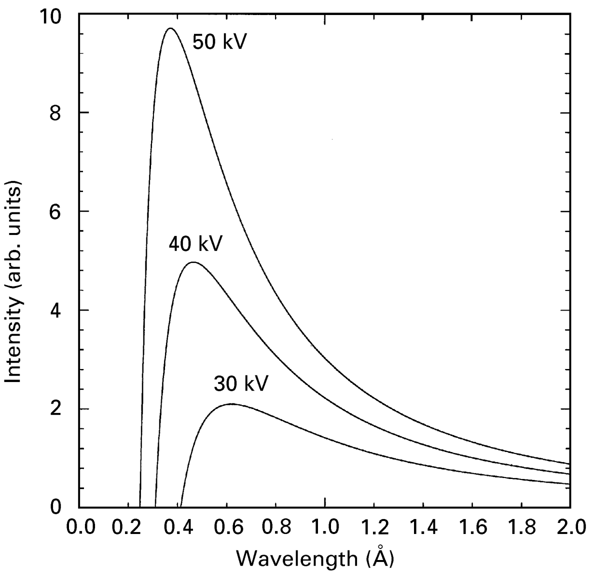



**Figure 7** Bremsstrahlung emission spectrum from an X-ray tube.

Here $\lambda_{\min}$(Å) = 124 000/V. The peak of the X-ray emission is at $\lambda \cong (3/2)\lambda_{\min}$. For example, for an accelerating voltage of 50 kV, $\lambda_{\min} = 0.25$ Å and the spectral distribution has a peak at $\lambda = 0.37$ Å.

The Bremsstrahlung radiation from a thick target is only partially polarized. However, at wavelengths near $\lambda_{\min}$, the radiation is strongly polarized in the plane containing the X-ray and the electron beam directions.

It may be noted that if the energy of the electron exceeds the binding energy of the inner shells of the target material, the impinging electron can also knock out core electrons from the target atoms, creating inner shell vacancies, leading to emission of characteristic X-rays of the target material. These X-ray lines are superimposed on the continuous Bremsstrahlung spectrum.

### Hot plasma sources

These include sources such as laser produced plasmas, tokamak plasmas, pinch plasmas, solar flares, stellar X-ray emitters, etc. In such plasmas, the electron temperature (corresponding to a Maxwellian velocity distribution) can be a few hundreds of eV to several keV. On collision with plasma ions these energetic electrons undergo acceleration/deceleration and thereby emit Bremsstrahlung radiation. Electron–electron collisions do not emit any net radiation as the two colliding electrons undergo exactly equal and opposite accelerations. The radiation emitted by the two electrons is therefore equal in magnitude and opposite in phase. Hence, there is no net radiation emitted.

For a Maxwellian velocity distribution of the electrons in the plasma, the spectral distribution of the emitted Bremsstrahlung radiation is given by (**Figure 8**)

$$I_\lambda = C\, N_e\, \Sigma_i\, N_i\, z^2\, T_e^{-1/2}\, g_f \exp(-hc/\lambda\, k\, T_e)/\lambda^2$$

where $N_e$ is the electron density, $N_i$ is the ion density of charge $i$, and $z$ is the average ion charge. The factor $g_f$ is of the order of unity and it represents a departure of quantum mechanical calculations from the classical results. This factor is called the Gaunt factor. The peak of this spectrum is at $\lambda_p$(Å) = $hc/2k\ T_e \sim 6.2/T_e$(keV).

The spectral distribution of the X-ray Bremsstrahlung emitted from a hot plasma is shown in **Figure 8**, which shows a strong temperature dependence of the spectrum for $\lambda < \lambda_p$. This fact is often used for estimation of the electron temperature of the plasma.

It may be noted that, unlike the Bremsstrahlung emission from X-ray tubes, there is no short wavelength limit here as the electrons have all possible energies.

**Recombination radiation** In addition to Bremsstrahlung radiation, a hot plasma also emits recombination radiation. This radiation is emitted when a free electron is captured in a bound state of an ion. If $E$ is the kinetic energy of a free electron and $\chi_n$ is the ionization potential of the energy level in which the electron is captured, the radiation is emitted with a photon energy of $h\nu = E + \chi_n$ (**Figure 9**). As the free

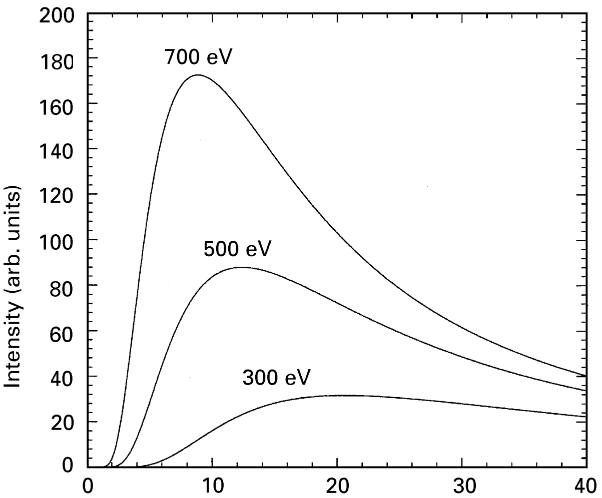



**Figure 8** Bremsstrahlung emission spectrum from hot plasma.

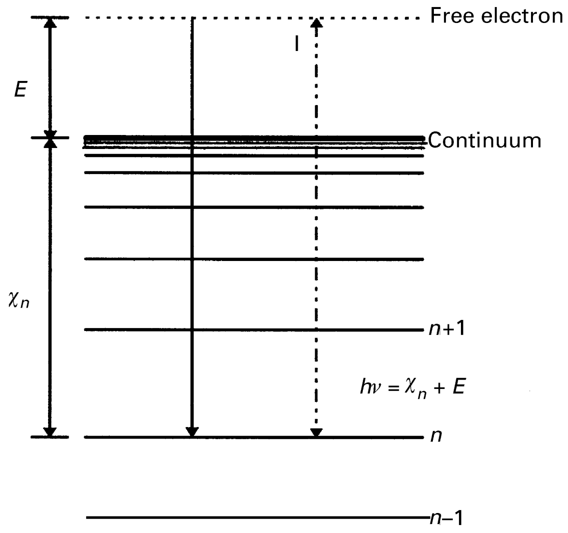



**Figure 9** Energy diagram for recombination radiation.

electron has a continuous energy distribution, the emitted radiation spectrum is also continuous for $h\nu \geq \chi_n$. Further, since the recombination can occur in different energy levels of the ion, the overall spectrum is quasi-continuous showing discontinuities at energies equal to the ionization potential energies of various levels. The overall shape of the spectrum is similar to that of plasma Bremsstrahlung radiation shown in **Figure 8**.

Interestingly, whereas in an X-ray tube, the radiation is on the longer wavelength side of the Duane–Hunt limit, here the spectrum is on the shorter wavelength side of the ionization potential.

**Cyclotron radiation** If the hot plasma happens to be magnetized (like the tokamak/pinch plasma, X-ray binary stars, etc.), then the electrons also emit cyclotron radiation (also called magnetic Bremsstrahlung) due to their gyration around the magnetic lines of force. The acceleration is due to the Lorentz force acting on the moving electrons.

The radiation spectrum is a discrete line spectrum at an electron cyclotron frequency ($\omega_c$) and its harmonics:

$\omega = n\omega_c$, where $\omega_c = eB/m \cong 1.76\times10^{11}B$ (T) rad s$^{-1}$.

Although the spectrum is independent of electron temperature, the total power radiated is proportional to both the electron temperature ($T_e$) and electron density ($N_e$) and is given by $P_c$ (W/m$^3$) $\cong 4.4\times10^{-28}\, N_e$ (m$^{-3}$)$B$(T)$^2$ $T_e$ (eV).

**Synchrotron radiation sources** In a synchrotron source, electrons move with relativistic speed ($v$~$c$). Although this speed is almost constant, when the electron trajectory is bent using bending magnets, the

electrons undergo transverse acceleration and radiate energy. This radiation is called synchrotron radiation.

The power radiated is given by

$$S = (e^2c/6\pi\varepsilon_0)\,\gamma^4/R^2$$

where $\gamma = E/m_0c^2$ and $R$ is the radius of the electron trajectory. For an electron beam of energy $E$(GeV) and a bending magnet field of $B$(T), the radius of the electron trajectory is given by $R = 3.33\ E/B$ m.

Since the motion of the electron is relativistic, the radiation (emitted perpendicular to the direction of acceleration) is highly concentrated in the direction of velocity within a narrow cone of nominal angular width $1/\gamma$.

Since the radiation is emitted in the form of a searchlight-like cone, an observer receives this radiation for a very short time period $\Delta t \sim 4R/3c\gamma^3$. As a result, the emitted radiation has a large bandwidth given by $\Delta\omega = 2\pi/\Delta t = 3\pi c\gamma^3/2R$. Thus the observed radiation is a continuous spectrum over a large frequency range.

Critical frequency is defined as the frequency which divides the synchrotron radiation power spectrum into two equal parts. This frequency is given by $\omega_c = 3c\gamma^3/2R$, and the corresponding critical energy is given by

$$E_c\ (keV) = 3hc\gamma^3/4\pi R = 0.665\ B\ (T)\ E^2\ (\text{GeV})$$

Also, the critical wavelength is given by

$$\lambda_c = 2\pi c/\omega_c = 4\pi R/3\gamma^2 = 5.6\ R\ (\text{m})/E^3\ (\text{GeV})$$
$$= 18.6/B\ (\text{T})\ E^2\ (\text{GeV})$$

The synchrotron radiation from a bending magnet is linearly polarized when observed in the plane of the $e^-$ orbit. Out of this plane, it is elliptically polarized with opposite helicity on either side of the plane.

**Periodic magnetic structure** In order to enhance the X-ray emission from the synchrotron, periodic magnetic structures are sometimes inserted in the linear sections of the synchrotron. This makes the electron notation sinusoidal in a horizontal plane. An important parameter characterizing the electron motion is the deflection parameter ($K$) given by $K = 93.4\ \lambda_u$ (m) $B$ (T), where $\lambda_u$ is the period of the magnetic structure.

*Undulator* For low magnetic field ($K<1$) the angular excursion of the electron is within the nominal $1/\gamma$ radiation cone. In this case, the electron beam breaks up into equally spaced bunches and the radiation from these bunches adds coherently (in phase). Such a magnetic structure is called an undulator. The emitted coherent radiation is at harmonics of $\lambda_L = (1+0.5\ K^2)\lambda_u/2\gamma^2$. The relative bandwidth of the $n$th harmonic is given by $\Delta\lambda/\lambda = 1/(nN)$, where $N$ is the number of magnetic periods. Along the axis, only odd harmonics are observed. However, if a helical magnetic structure is used, the radiation will be circularly polarized and there will be no harmonics present in the on-axis radiation. The emission cone is further narrowed down to $\theta \sim 1/(\gamma\sqrt{N})$. The total radiation flux is $N^2$ times the flux due to a single bending magnet.

If such a structure is placed in an optical resonator or used as an amplifier for radiation of wavelength $\lambda = \lambda_L$, then it is referred to as a free electron laser.

*Wiggler* If the electron excursion angle exceeds $1/\gamma$ (i.e. when $K >> 1$) then the magnetic structure is called a wiggler. Here, the radiation from different sections of the electron trajectory (where the direction of electron motion makes an angle less than $1/\gamma$ with the axis) adds up incoherently. The total radiation flux is $2N$ times the flux due to a single bending magnet. The emission cone is several times larger than that of the synchrotron radiation.

The total power emitted by an undulator or wiggler is given by

$$P(\text{kW}) = 0.633\ E^2\ (\text{GeV})\ B^2\ (\text{T})\ L\ (\text{m})\ I\ (\text{A})$$

where $L$ is the total length ($=\lambda_u N$) of the magnetic structure, and $I$ is the beam current.

For a planar magnetic field in the vertical direction, the on-axis radiation is polarized in the horizontal plane, as in the case of a bending magnet. However, unlike the bending magnet case, for a periodic magnetic structure, the off-axis radiation is also plane-polarized in the horizontal plane. This is because the vertical component of polarization emitted in one half is cancelled by the radiation in the next half (out of phase) as both have the same direction of polarization. However, the horizontal polarization in the two halves being opposite in direction, the net radiation adds up.

## X-rays from few electron systems

X-rays from a few electron systems such as hydrogen-like, helium-like, lithium-like ions are observed in hot

plasmas (laser produced / tokamak / Z pinch), beam-foil experiments, heavy ion collisions, solar flares, etc. Muonic atoms also emit X-rays.

### Ionic X-rays

Transitions in the highly charged ions are in the X-ray region and are interesting because they are relatively simple to interpret. They are also one of the few cases in atomic physics wherein high order multipole transitions are observed.

**Hydrogen-like ions** These are ions having a single electron left. The energy levels of these ions are exactly the same as those of the hydrogen atom except that they are increased by a factor $Z^2$, where $Z$ is the atomic number of the ion.

$$E\,(n, l\,) = Z^2\; E_{\mathrm{H}}(n, l\,)$$

The hydrogen-like series is composed of transitions of the type $(\mathrm{np})^2\ \mathrm{P}_{3/2,1/2}$–$(1\mathrm{s})\ {}^2\mathrm{S}_{1/2}$, where $n\geq 2$. Each line of the series is a doublet. The limit of this series is the highest energy X-ray line that can be emitted by a given element.

Unlike the hydrogen Lyman-α doublet ($\Delta v = 0.36$ cm$^{-1}$) which cannot be easily resolved due to Doppler broadening even in moderate temperature (~$10^4$K) plasmas, the fine structure in high $Z$ elements can be easily resolved even in high temperature (~$10^6$K) plasmas. For example, in H-like calcium, the wavelengths of the $\mathrm{Ly}_\alpha$ doublet are 3.018 Å and 3.24 Å, which can be easily resolved with a standard crystal spectrometer.

**Helium-like ions** These are ions with only K shell electrons left. Here, except for the ground state which is a singlet (${}^1\mathrm{S}_0$), all the excited levels ($n\geq 2$) have both singlet and triplet states (**Figure 10**). The helium-like series is composed of transitions of the type (np 1s) ${}^1\mathrm{P}_1$–$(1\mathrm{s}^2){}^1\mathrm{S}_0$. The end point of the series is the ionization potential of the 1s electron in its $(1\mathrm{s}^2){}^1\mathrm{S}_0$ ground state.

**Types of transitions** The terminology for X-ray transitions in ions is the same as that of optical transitions in an atom. The three main types of transitions observed in an ionic line spectrum are: 1) resonance transitions, 2) intercombination transitions, and 3) satellite transitions.

*Resonance transitions* These are transitions from an excited state to the ground state (**Figure 10**). The H-like series and He-like series discussed earlier are resonance transitions. The oscillator strengths of these transitions are higher than those of other types of transition. For the same reason, these lines can be strongly reabsorbed if they are emitted by a hot-dense plasma. They follow the normal selection rules applicable to optical transitions.

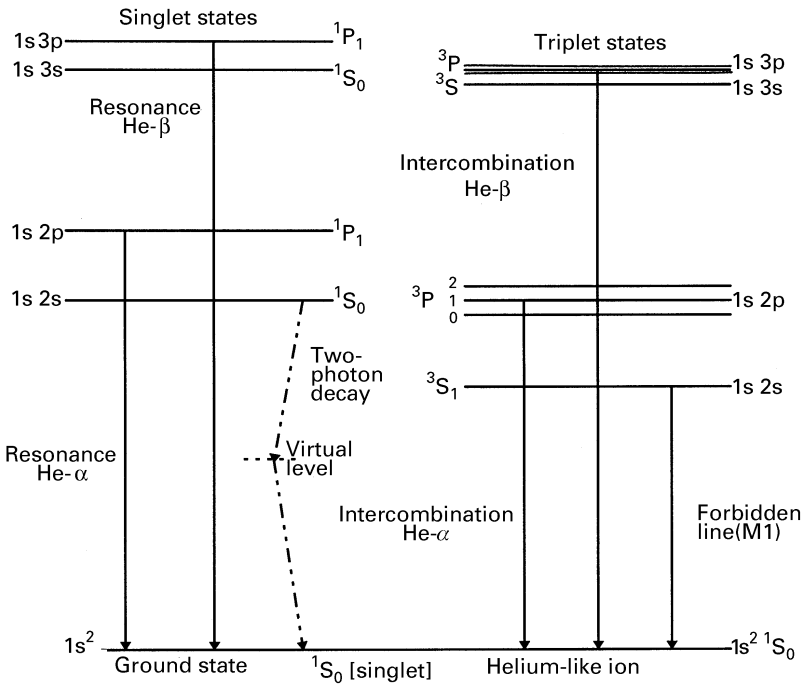



**Figure 10** Energy level diagram for helium-like ions showing resonance, intercombination, and other transitions.

*Intercombination transitions* These transitions are similar to resonance transitions except that they are between states of different multiplicity (e.g. triplet to singlet). Corresponding to a resonance transition (1s2p) $^1P_1$–(1s$^2$) $^1S_0$ (which is a singlet–singlet transition), the intercombination transition (triplet–singlet) is (1s$^2$p)$^3$P–(1s$^2$) $^1S_0$ (**Figure 10**). These lines appear on the lower energy side of the resonance lines. Though $^3P_1$ – $^1S_0$ is a spin-forbidden transition, in high *Z* ions, the $^3P_1$ state decays by dipole transition through mixing with $^1P_1$ states.

*Satellite lines* These are weaker lines arising from doubly excited ions of higher ionization states. The satellite lines of H-like ions are due to He-like ions, those of He-like ions are due to Li-like ions and so on. For example, transitions in Li-like ions of the type 1s nln'l' → 1s$^2$ n'l' will appear as a satellite to 1s nl–1s$^2$ resonance transitions in He-like ions (**Figure 11**). The n'l' electron is a spectator electron. Due to the presence of this electron, Coulomb shielding decreases, which results in the transition occurring at a slightly lower energy than that of the resonance transition. Since two electrons are involved in such transitions (one active, one spectator), these lines are also referred to as dielectronic satellites.

The largest separation from a parent line occurs when the spectator electron is in the lowest *n* level (i.e. $n = 2$), when the Coulomb shielding effect is maximum. These satellites are referred to as 2s or 2p satellites. For example, the 1s 2s 2p – 1s$^2$ 2s transition in Li-like ions will be a 2s satellite for the 1s 2p – 1s$^2$ (He-$\alpha$) transition in He-like ions (**Figure 11**) Satellites of H-like ions in the Gabriel notation are denoted by capital letters A, B,...J,... and those of He-like ions are denoted by lower case letters a, b, c,......u, v.

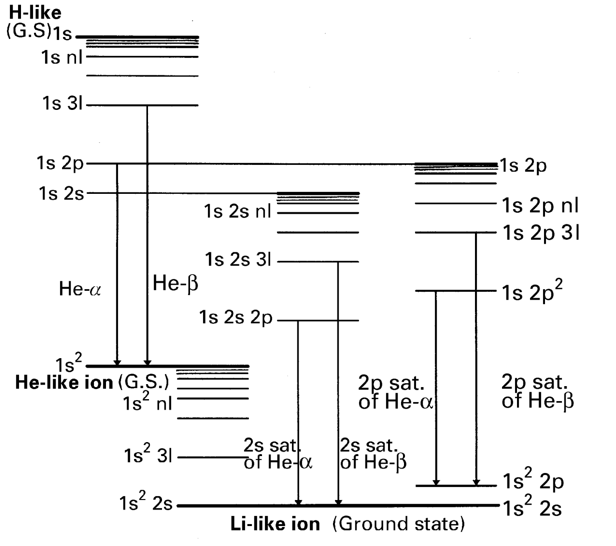



**Figure 11** Energy diagram for 2s and 2p satellite transitions in lithium-like ions corresponding to He-α and He-β transitions in helium-like ions.

### Muonic X-rays

When a negatively charged particle ($\mu^-$, $\pi^-$, $K^-$ meson) replaces an electron in an atom, a mesonic atom is formed. For example, when a $\mu^-$ meson is brought to rest in a target, muonic atoms of the target element are formed by replacement of a valence electron by $\mu^-$. The energy levels in a muonic atom are analogous to the electronic energy levels of an H-like ion except that the muon mass is higher ($m_\mu \sim 207\ m_e$). The mesonic atom energy levels are related to those of the hydrogen atom by $E(n,l) = Z^2\ (M^*/m^*)\ E_H(n,l)$, where $M^*$ and $m^*$ are the reduced masses of $\mu^-$ and the electron, respectively, and $E_H(n,l)$ denotes energy levels in hydrogen. The newly formed muonic atom is thus in a highly excited state and lowers its energy by ejecting electrons by successive Auger processes until the principal quantum number falls to less than 5 (in heavy atoms). At this point, the radiative transition probability becomes more prominent. The energy level differences become of the order of several keV. As a result, X-rays are emitted until the muonic atom reaches its ground state.

The average radius of the lowest (Bohr) orbit of the muonic atom is given by $r \sim (m^*/M^*)\ (a_H/Z)$ which is considerably smaller than that of a normal atom. For example, for silver, *r* is $5 \times 10^{-15}$m, which is of the order of the nuclear size. As a result, the muon spends considerable time inside the nucleus. Consequently, the 1s level is strongly affected by the nucleus. This is corroborated by the fact that whereas the Balmer series ($n \to 2$) transition energies are found to be exactly $Z^2(M^*/m^*)$ times those of the H-atom, the Lyman series ($n \to 1$) transition energies are lower than expected.

## List of symbols

$E$ = photon energy; $g_f$ = Gaunt factor; $I$ = intensity; $I$ = beam current; $k$ = Sommerfeld quantum number; $K$ = deflection parameter; $L$ = length of magnetic structure; $M^*, m^*$ = reduced mass of muon and electron; $n$ = total quantum number; $N$ = number of magnetic periods; $N_e$ = electron density; $N_i$ = ion density; $P$ = total power; $r$ = radius of Bohr orbit; $R$ = radius of electron trajectory; $R$ = Rydberg constant; $S$ = power; $T$ = term value; $T_e$ = electron temperature; $V$ = voltage; $W$ = work function; $z$ = ion charge; $Z$ = atomic number; $\Delta\omega$ = bandwidth; $\alpha$ = fine structure constant; $\chi_n$ = ionization potential; $\varepsilon$ = efficiency; $\gamma$ = relativistic factor = $E/m_0c^2$; $\lambda$ = wavelength; $\nu$ = frequency; $\theta$ = cone angle; $\sigma$, $s$ = screening

constants; ω = fluorescence yield; $\omega_c$ = electron cyclotron frequency; $\omega_p$ = plasma frequency.

*See also:* **Photoelectron Spectrometers; X-Ray Absorption Spectrometers; X-Ray Emission Spectroscopy, Applications; X-Ray Emission Spectroscopy, Methods; X-Ray Fluorescence Spectrometers; X-Ray Fluorescence Spectroscopy, Applications; Zero Kinetic Energy Photoelectron Spectroscopy, Theory.**

## Further reading

Agarwal B.K (1991) *X-ray Spectroscopy: an Introduction*, Berlin: Springer-Verlag.

Azaroff LV (1974) *X-ray Spectroscopy*, New York: McGraw-Hill.

Bertin EP (1975) *Principles and Practice of X-ray Spectrometric Analysis*, New York: Plenum Press.

Bonnelle C and Mande C (1982) *Advances in X-ray Spectroscopy*, New York: Pergamon Press.

Craseman B (1985) *Atomic Inner Shell Physics*. New York: Plenum Press.

Herglotz HK and Birks, LS (1978) *X-ray Spectrometry*. New York: Dekker.

Janev RK, Presnyakov LP and Shevelko VP (1985) *Physics of Highly Charged Ions*. Berlin: Springer-Verlag.

Jenkins R (1976) *An Introduction to X-ray Spectrometry*, London: Heyden.

Kauffman RL and Richard P (1976) X-ray region. In: Williams D (ed) *Methods of Experimental Physics Vol. 13 Part A (Spectroscopy)*. London: Academic Press.

Michette AG and Buckley CJ (eds) (1993) *X-ray Science and Technology*. London: IOP Publishing Ltd.

Thompson, M. Baker, MD Christie, A and Tyson JF (1985) *Auger Electron Spectroscopy (Chemical Analysis, Vol. 74)* New York: Wiley Interscience.

Williams KL (1987) *Introduction to X-ray Spectrometry* London: Allen and Unwin.

White HE (1986) *Introduction to Atomic Spectroscopy*. Singapore: McGraw-Hill Book Co.

# Z

# Zeeman and Stark Methods in Spectroscopy, Applications

**Ichita Endo** and **Masataka Linuma**, Hiroshima University, Japan

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
| --- |
| **Applications** |

## Introduction

An atomic system is influenced by an external electric and magnetic field. Due to an interaction of the magnetic field with the magnetic moment of an atom, an electronic energy level in the atomic system is shifted in accordance with the formula of the Zeeman effect, while an external static electric field would polarize the atom, resulting in an energy shift referred to as the Stark effect. As the amount of energy shift depends on the magnetic quantum number of the level, the Zeeman and Stark effects resolve the otherwise degenerate energy levels into sublevels. From the pattern of level splitting we can assign the quantum numbers of the observed electronic level. The absolute value of separation between the split levels tells us about the magnetic moment and the polarizability for Zeeman and Stark spectroscopy, respectively.

A straightforward application of Zeeman spectroscopy is a magnetic-field determination using atoms with a known magnetic moment, and one of Stark spectroscopy is a measurement of an electric field using atomic levels with predetermined polarizability. Such measurements are useful when field-measuring probes based on other principles are either unusable or difficult to apply as in the astrophysical environment or in a plasma.

Indirect but important usage of Zeeman and Stark effects is found in fundamental physics researches: for example measurements of violation of symmetry in physical laws under time and space inversion known as the T-violation and the parity violation, respectively. Such measurements would eventually give us clues of new physics beyond the standard model of unified electromagnetic and weak interactions.

We present here some selected topics in the application of Zeeman and Stark spectroscopy of atoms in the gas phase with special emphasis on parity-violation experiments.

## Fundamental physics research

A steady state of an isolated atom is described by a quantum-mechanical state specified by the energy and the total angular momentum in accordance with the translation invariance in the time coordinate and the rotational symmetry of the space coordinates, respectively. If physical laws were completely invariant under the parity operation, i.e. space inversion, the wavefunction $\psi$ of the atom remains exactly the same except for its sign; $P\psi$, where $P = \pm 1$. The parity quantum number $P$ introduced in this way would also be a good quantum number in the atomic system if the forces acting on atomic electrons were due only to the classical electromagnetic interaction which is invariant under space inversion. Recent findings in fundamental physics, however, predict that parity is not conserved in the atomic system, though the amount of violation is extremely small.

When an electric field is applied to an atom, the space symmetry is destroyed so that an even-parity state is slightly contaminated by an odd-parity state and *vice-versa*. This makes the otherwise-forbidden E1 transition between the same parity levels to be observable; a Stark-induced transition. The interference term between the Stark-induced E1 transition

amplitude and that due to the intrinsic parity violation changes sign when the direction of the electric field is reversed. Therefore, the parity nonconservation effect can be measured by comparing a small amount of change in the transition rates as we reverse the electric field. There have been several experiments based on this principle to acquire quantitative information of parity violation, the most precise experiments being laser spectroscopy of an atomic beam of Cs under an external electric and magnetic field.

Various attempts are being made to achieve higher accuracy in the parity nonconservation (PNC) measurements. One of the possibilities is to use a heavier atom and to observe a transition to a level with a relatively large amount of parity mixing. Rare earth atoms are deemed as good candidates, because they have many close-lying level pairs of opposite parity. However, there has been no experimental support for them to have a sizable enhancement in parity mixing.

Some examples of Zeeman and Stark spectroscopy of rare earth atoms are shown below. They were obtained in a series of studies aiming at finding the atomic states suitable for the PNC experiment.

In **Figure 1** Stark spectra of samarium atoms are shown. They were obtained by detecting the fluorescence from the level excited with the laser beam. The observed transition is from the ground state ($J = 0$) to the 1.9404 eV ($J = 1$) excited level, in which $J$ is the electronic total angular momentum.

The strength of the electric field denoted by $E$ is 0.0, 17.2 and 26.1 kV cm$^{-1}$ for the upper, middle and lower part of the figure, respectively. The peaks labelled by open and solid circles correspond, respectively, to the transition from $|m| = 0 \rightarrow |m| = 0$ and from $|m| = 0 \rightarrow |m| = 1$, in which $m$ is the magnetic quantum number. We see that each peak for $^{152}$Sm and $^{154}$Sm is split into two peaks of which the separation increases as the electric field is strengthened. In this case only the Stark effect on the upper level ($^7G_1$) is responsible for the splitting because the lower level has $J = 0$. The energy separation of the split is expressed by

$$\Delta W = \frac{3}{2}\alpha_2 J_u E^2 \qquad [1]$$

where $\alpha_2$ is the tensor polarizability and $J_u$ is the electronic total angular momentum of the upper level. The $E^2$-dependence of splitting of the $^{154}$Sm peak for the same transition is shown in **Figure 2**. It is clearly seen that the energy interval of the splitting is proportional to $E^2$. The tensor polarizability is determined from the slope of the straight line: $\alpha_2 = -554.6 \pm 1.3$ kHz (kV)$^{-2}$ cm$^2$ for the data shown in **Figure 2**.

The Zeeman spectra for the transition between the 0.0363 eV level ($J = 1$) and the 1.9301 eV level ($J = 2$) of Sm are shown in **Figure 3**. The peaks labelled by squares and circles correspond to $^{154}$Sm and $^{152}$Sm. The applied magnetic field is 0, 115×10$^{-4}$, 224×10$^{-4}$ and 352×10$^{-4}$ T for the spectra shown in **Figures 3A**, **B**, **C** and **D** respectively. The open and solid symbols are for the $\sigma$ and $\pi$ components of the transition, respectively. The $\sigma$ component is the transition associated with a change in magnetic quantum number, $\Delta m = \pm 1$, caused by a photon with its polarization perpendicular to the direction of the magnetic field. The $\pi$ component is defined as the transition with $\Delta m = 0$. The energy interval, represented by the frequency, $f$, of Zeeman splitting, is given by the formula,

$$f = \mu_B (g_u m_u - g_l m_l)B + C \qquad [2]$$

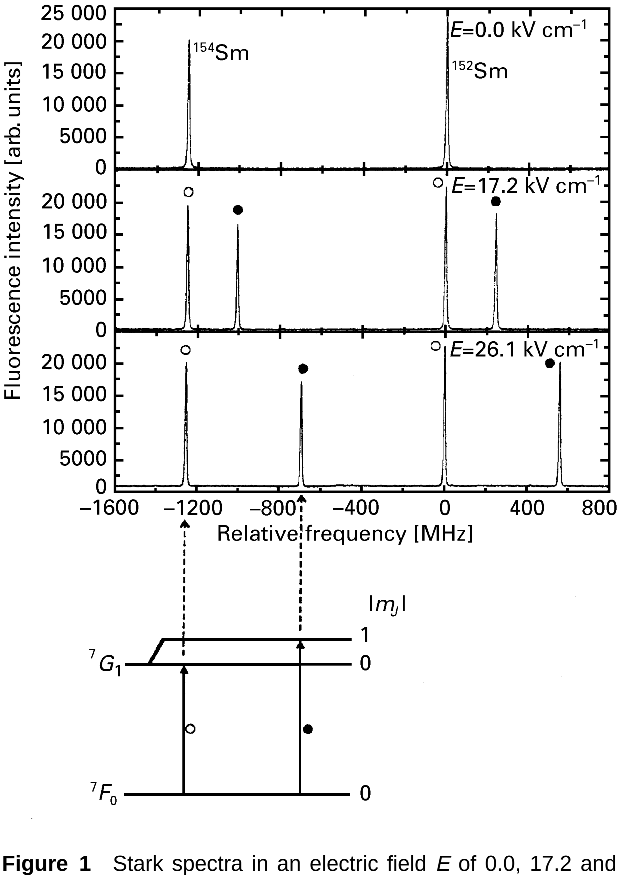



**Figure 1** Stark spectra in an electric field $E$ of 0.0, 17.2 and 26.1 kV cm$^{-1}$ for $^{152}$Sm and $^{154}$Sm for the transition from the ground state with $J = 0$ to the 1.9404 eV level with $J = 1$. The energy levels responsible for these spectra are schematically shown in the lower part together with their electronic configuration.

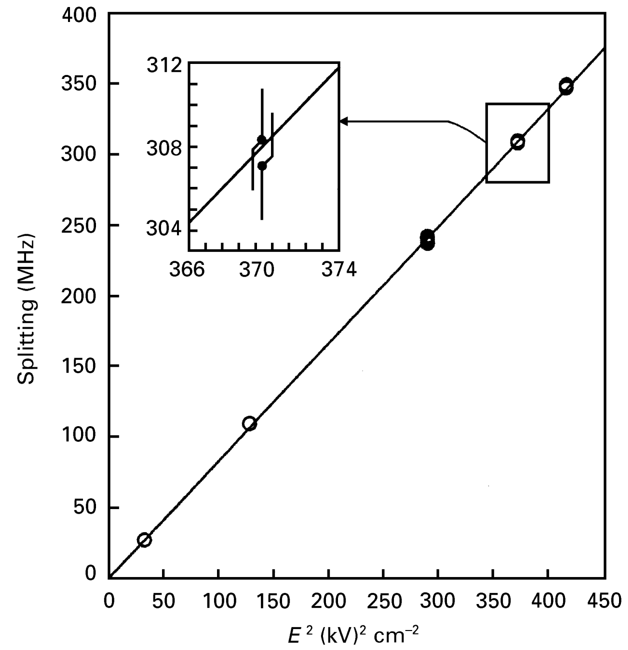



**Figure 2** Stark splitting as a function of squared electric field $E^2$.

where $\mu_B$ is the Bohr magneton, $m_u$ and $g_u$ are the magnetic quantum number and the $g$-factor for the upper level, respectively, while $m_l$ and $g_l$ are those for the lower level. A constant $C$ gives the original frequency without an external magnetic field. In the case of **Figure 3** the $g$-factor for the lower level is known to be zero so that the constants $g_u$, $B$ and $C$ are determined by least-squares fitting of Equation [2] to the relative frequencies among the Zeeman peaks.

## Field measurements

If the $g$-factors (polarizabilities) are known in advance, it is possible to measure a static magnetic (electric) field by means of the Zeeman (Stark) effect. This is useful particularly in such situations as in hot plasma and in astronomical objects where the standard field-measuring probes, e.g. a nuclear magnetic resonance probe and a Hall probe, are unusable.

In plasma diagnostics, for example, Stark spectroscopy is used for determining the local electric field. Since the Stark splitting is large for the Rydberg levels, the excitation to the level with high principal quantum number $n$ is used. A small amount of probe atoms mixed in the plasma are excited to metastable

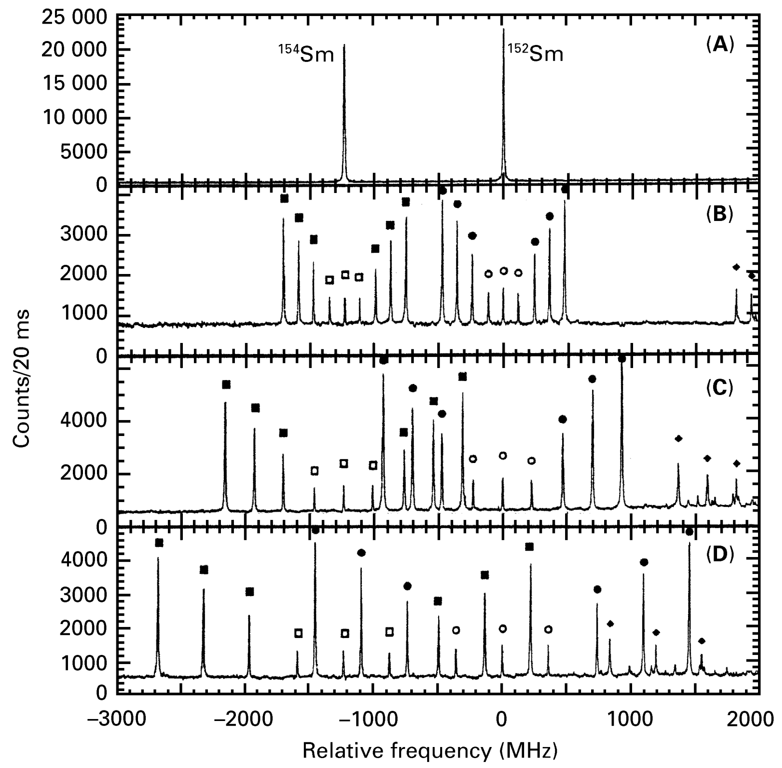



**Figure 3** Zeeman spectra of samarium atoms. The applied magnetic field is 0, $115 \times 10^{-4}$, $224 \times 10^{-4}$ and $352 \times 10^{-4}$ T for (A), (B), (C) and (D) respectively. The peaks represented by squares and circles correspond, respectively, to the transition of $^{154}$Sm and $^{152}$Sm. The open and solid symbols represent the $\sigma$ and $\pi$ components of the transition, respectively.

states by an electric discharge. They are further pumped up to the Rydberg levels by a laser beam, followed by decay to intermediate levels due to collisional transitions. The fluorescence from the intermediate levels bears information on the electric field at the particular position in the plasma at which the laser is focused with its frequency being scanned. The electric field can also be evaluated from the ratio of the intensity of the forbidden transition induced by the Stark mixing to that of the allowed transition.

Another example is related to astronomical researches: the spectra of starlight usually reveal some emission and absorption lines. The position and the depth of the absorption lines tell us about the atomic species and their relative abundance in the cold outer gas of the star. If a strong field exists in a star, Zeeman and Stark splitting are identifiable in the spectra. From the pattern of absorption or emission lines, it is sometimes possible to determine the strength of magnetic and electric fields in astronomical objects.

## List of symbols

$E$ = electric-field strength; $f$ = frequency of Zeeman splitting; $J$ = electronic total angular momentum; $m$ = magnetic quantum number; $P$ = parity quantum number; $W$ = energy separation of Stark splitting; $\alpha_2$ = tensor polarizability; $\mu_B$ = Bohr magneton; $\psi$ = atomic wavefunction;

*See also:* **Atomic Fluorescence, Methods and Instrumentation; Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Zeeman and Stark Methods in Spectroscopy, Instrumentation.**

## Further reading

Dalgarno A and Layzer D (1987) *Spectroscopy of Astrophysical Plasmas*. Cambridge: Cambridge University Press.

Demstroder W (1998) *Laser Spectroscopy*, 2nd edn. Berlin: Springer-Verlag.

Greenberg KE and Hebner GA (1993) Electric-field measurements in 13.56 MHz helium discharges. *Applied Physics Letters* **63**: 3282–3284.

Hanle W and Kleinpoppen H (1978) *Progress in Atomic Spectroscopy*. New York: Plenum Press.

Khriplovich IB (1991) *Parity Nonconservation in Atomic Phenomena*. Philadelphia: Gordon and Breach Science Publishers.

Kobayashi T, Endo I, Fukumi A *et al* (1997) Measurement of hyperfine structure constants, *g* values and tensor polarizability of excited states of Sm I. *Zeitschrift für Physik D* **39**: 209–216.

Shimoda K (1976) *High-Resolution Laser Spectroscopy*. Berlin: Springer-Verlag.

Svanberg S (1992) *Atomic and Molecular Spectroscopy*, 2nd edn. Berlin: Springer-Verlag.

# Zeeman and Stark Methods in Spectroscopy, Instrumentation

**Ichita Endo** and **Masataka Linuma**, Hiroshima University, Japan

Copyright © 1999 Academic Press

| ELECTRONIC SPECTROSCOPY |
| --- |
| Methods & Instrumentation |

## Introduction

The energy difference between the Zeeman and Stark sublevels is usually far smaller than the line width of the optical transition in atoms at normal temperature due to Doppler broadening. Doppler-free techniques are necessary for obtaining the values of *g*-factors and polarizabilities in optical spectroscopy.

Variations of coherent spectroscopy, such as level-crossing, quantum beat, and pulsed-field spectroscopy, are examples of Doppler-free techniques. They make use of the interference effect in the transition amplitudes of simultaneous excitation from a level to two closely separated higher levels.

Another technique widely used is atomic beam spectroscopy. In a gas jet ejected from a small orifice, the transverse motion of atoms is much reduced. The line width of the transition induced by a laser beam perpendicularly crossing the atomic beam can be narrow enough to resolve the Zeeman and Stark splitting.

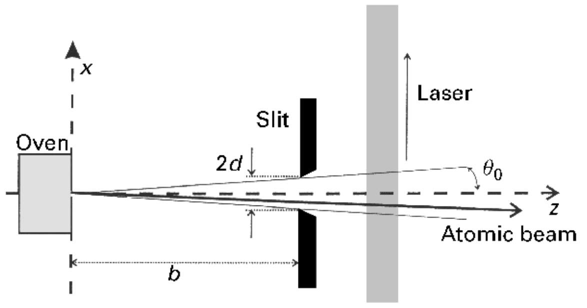



**Figure 1** Schematic illustration of an atomic-beam technique to reduce Doppler broadening. Atomic vapour effuses from a small orifice of an oven. The angular divergence of atoms in the beam is limited to $\theta_0 = \tan^{-1} b/d$ by a slit whose aperture is $2d$ placed at a distance $b$ from the orifice.

## Atomic beam spectroscopy

The Doppler effect broadens absorption or emission lines from atoms in the gas phase at thermal equilibrium. Assume that an atom at rest is excited by a photon with wave vector $\boldsymbol{k}$ and de-excited to emit light with angular frequency $\omega_0$. If the atom is moving at a velocity $\boldsymbol{v}$ the angular frequency of the emitted light is shifted to a value $\omega_0'$ according to the formula,

$$\omega_0' = \omega_0 - \boldsymbol{k} \cdot \boldsymbol{v} \quad [1]$$

At thermal equilibrium, the velocities of atoms in the gas phase obey a Maxwellian distribution. This results in the broadened intensity profile around $\omega_0$ as approximately represented by a Gaussian form,

$$I(\omega) = I_0 \exp\left[-\left(\frac{c(\omega - \omega_0)}{\omega_0 v_{\mathrm{th}}}\right)^2\right] \quad [2]$$

where $m/c$ is the velocity of light and $v_{\mathrm{th}} = (2k_{\mathrm{B}}T/m_{\mathrm{a}})^{1/2}$ is the most-probable velocity of atoms with mass $m_{\mathrm{a}}$, temperature $T$ and Boltzmann constant $k_{\mathrm{B}}$. The Doppler width defined by the full width at half maximum of the Gaussian profile is

$$\Delta\omega_{\mathrm{D}} = 2\sqrt{\ln 2}\,\frac{\omega_0 v_{\mathrm{th}}}{c} = \left(\frac{\omega_0}{c}\right)\sqrt{\frac{k_{\mathrm{B}}T \ln 2}{m_{\mathrm{a}}}} \quad [3]$$

Let us consider a case where the atoms are effusing into a vacuum chamber, as shown in **Figure 1**, from an orifice of an oven filled with vapour at a temperature $T$. Let the atomic beam travel along the $z$-axis, while the laser beam is parallel to the $x$-axis. One can reduce the Doppler broadening by limiting the beam divergence with a slit with a small aperture $2d$ in the $x$ direction at a distance $b$ from the orifice. This makes the beam divergence in the $x$-$z$ plane smaller than $\theta_0 = \arctan d/b$ and the Doppler broadening is reduced to $\Delta\omega_{\mathrm{D}}' = \Delta\omega_{\mathrm{D}} \sin\theta_0$.

An example of laser spectrometers for Zeeman and Stark spectroscopy using a collimated atomic beam is shown schematically in **Figure 2**. It consists of a continuous-wave (CW) tunable dye-laser system, a frequency calibration system, a vacuum chamber with a fluorescence detector, and a data-acquisition system. The interaction point of the atomic beam with the laser is inside the vacuum chamber.

A magnified view around the interaction point is illustrated in **Figure 3**. The oven made of molybdenum

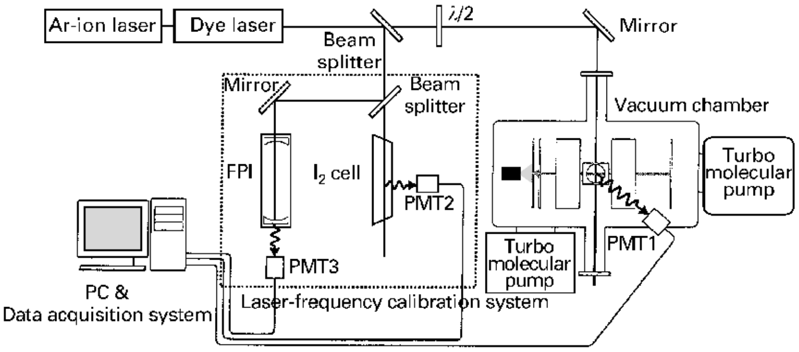



**Figure 2** Typical setup for a laser spectrometer based on the atomic-beam method. The apparatus is composed of a continuous-wave (CW) tunable laser system, a laser-frequency calibration system, a vacuum chamber, and a data acquisition system. The fluorescence light from the excited $I_2$ molecules in a cell and the excited atoms in the vacuum chamber, and the transmitted light from a Fabry-Perot interferometer (FPI) are detected simultaneously with three photomultiplier tubes (PMTs). The signals are transformed to digital pulses event-by-event and introduced to the inputs of a multi-channel scaler (MCS).

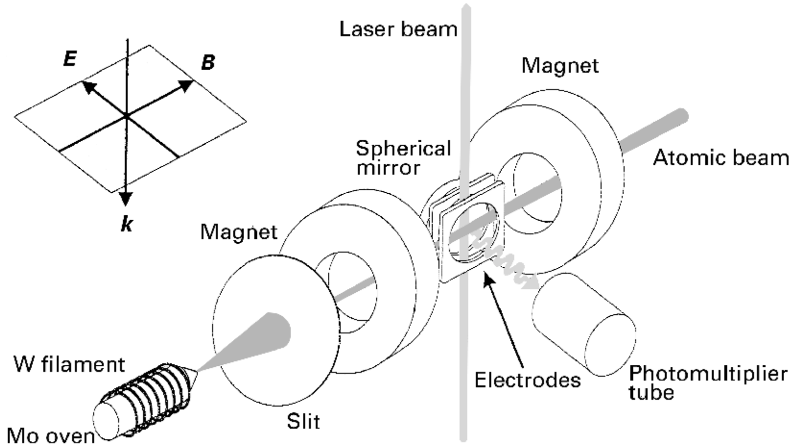



**Figure 3** Magnified view around the interaction point of the atomic beam with the laser beam. The oven made of molybdenum is heated by a tungsten filament wound around it to eject a gas jet from the orifice with a diameter of 0.8 mm. The atomic beam is collimated with the slit to a diameter of about 4 mm at the interaction point, where the two electrodes for applying the electric field and a pair of Helmholtz coils to produce the magnetic field are installed. The photomultiplier tube (PMT) for detecting the fluorescence light from the atoms and a spherical mirror to collect light are shown.

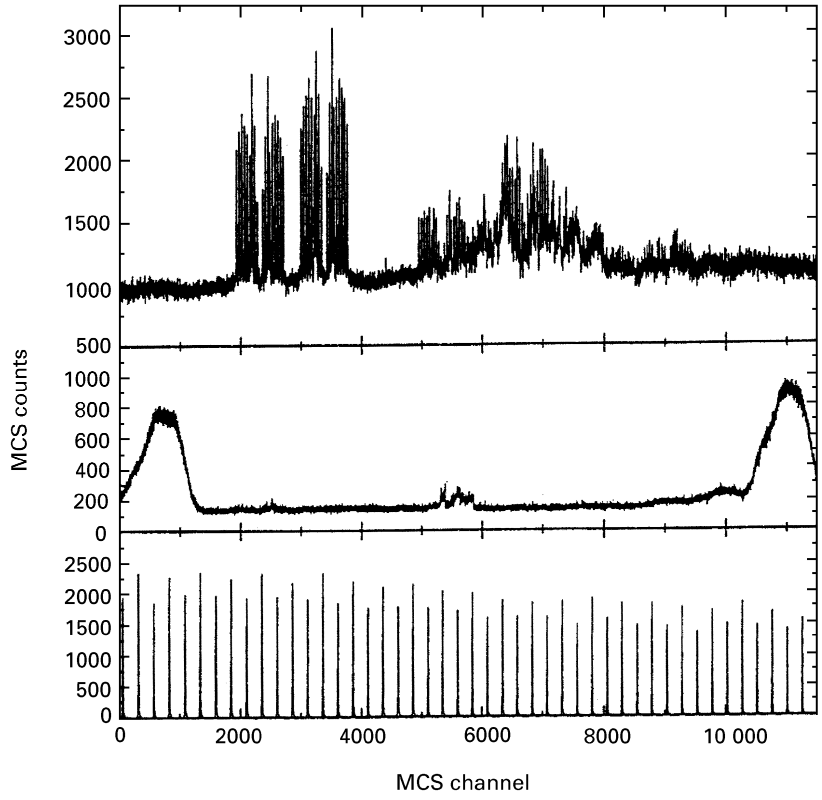



**Figure 4** Typical set of raw data for Zeeman spectrum of samarium atoms with the natural isotopic abundance under the magnetic field of $167.38 \times 10^{-4}$ T. The number of counts of detected photons per 20 ms is plotted against the MCS channels corresponding to the elapsed time from the starting point of the frequency sweep of the laser. The top part corresponds to the Zeeman spectrum in the transition from the level of $E = 0.184\,68$ eV($J = 3$) to the one of $E = 2.076\,5$ eV($J = 3$). In the middle and the bottom, the spectrum of $^{127}I_2$ and the spectrum of the transmitted light from the FPI are shown, respectively.

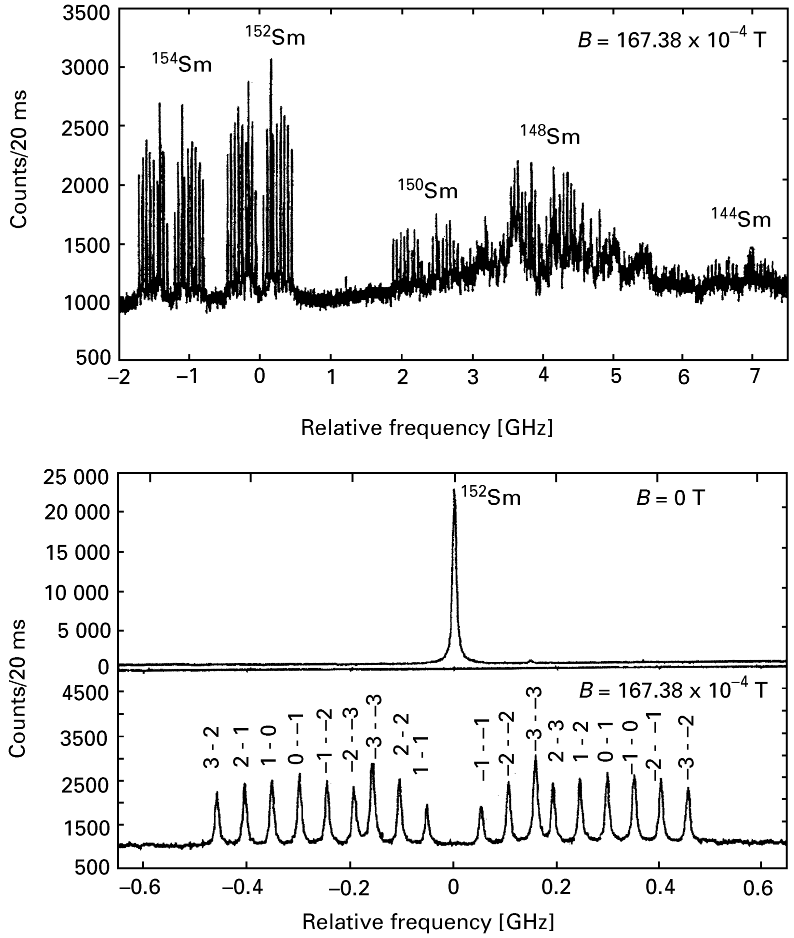



**Figure 5** Zeeman spectra after calibration on the horizontal axis and peak assignments. The top part of the spectrum is the same as the one shown in **Figure 4**. Magnified spectra of just the area of the peaks for the $^{152}$Sm atoms are shown in the middle and the lower parts, in which the magnetic field is 0 T and 167.38 × 10$^{-4}$ T respectively. The label above each peak is to indicate the relevant change in the magnetic quantum number, *m*, to the one with *m*′, associated with the optical transition.

is attached to an end plate of the vacuum chamber in which the pressure is kept to about 10$^{-6}$ torr. The oven is heated by a tungsten filament wound around it to eject a gas jet from the orifice with a diameter of 0.8 mm. The temperature can be increased to about 1000 K and is monitored with a Pt-Rh thermocouple. The atomic beam is collimated by the slit and led to the interaction point, where the two electrodes made of BK7 glass plates coated with ITO($InSnO_2$) on one side are installed to apply the electric field *E*. The magnetic field *B* in parallel with the atomic beam is applied by a set of Helmholtz coils. The fluorescence from the atoms is detected with a photomultiplier tube (PMT) which is cooled to reduce thermal noise. In order to increase the collection efficiency of the emitted photons, a spherical mirror is installed on the opposite side of the PMT.

Linearly polarized light from a laser is introduced to the inside of the vacuum chamber as shown in **Figure 2**. The CW dye-laser is capable of continuously changing its frequency with time, sweeping over a certain frequency range. The laser polarization is adjusted with a half-wave plate when necessary. The signal from the PMT is fed to one of the inputs of a multi-channel scaler (MCS) where the number of counts in each time interval, corresponding to a small frequency segment, is recorded.

The spectra from molecular iodine, $^{127}I_2$, together with the transmitted light through a Fabry-Perot interferometer (FPI), are recorded synchronously with the fluorescence from the excited atoms to give frequency marks separated by the free-spectral range (FSR) of the FPI.

In **Figure 4** a set of raw data obtained in Zeeman spectroscopy of Sm at *B* = 167.38 × 10$^{-4}$ T is shown as an example. The uppermost part corresponds to the Zeeman spectra for samarium atoms with the natural isotopic abundance ($^{144}$Sm: 3.1%, $^{147}$Sm: 15.0%, $^{148}$Sm: 11.3%, $^{149}$Sm: 13.8%, $^{150}$Sm: 7.4%, $^{152}$Sm: 26.7%, $^{154}$Sm: 22.7%) in the transition from

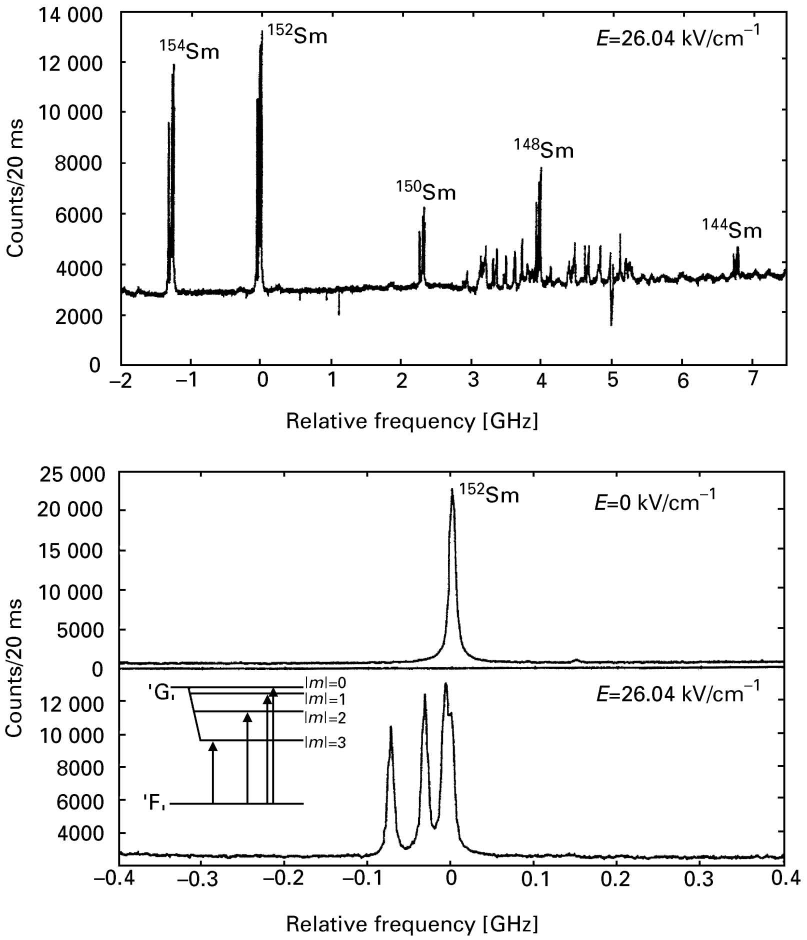



**Figure 6** Stark spectra obtained by the analogous method to the one used for **Figure 5**. The transitions are the same as in **Figure 5**. The electric field of 26.04 kV $cm^{-1}$ is applied. The middle and lower graphs correspond to the spectra for the $^{152}Sm$ atoms under the electric field of 0 kV $cm^{-1}$ and 26.04 kV $cm^{-1}$ respectively.

the level of $E = 0.184\,68$ eV ($J = 3$) to the one of $E = 2.076\,5$ eV ($J = 3$) where $J$ is the electronic total angular momentum. The spectrum of $^{127}I_2$ and that of the transmitted light from the FPI is shown in the middle and the lowest parts, respectively. After calibration on the horizontal axis and assignment of each peak, we obtain the spectra shown in **Figure 5** for which the Zeeman splitting is completely resolved. Combining the spectra measured with different magnetic field strengths, we can determine the $g$-factor for either the upper or lower level if one of them has been known in advance.

The Stark spectrum for the same transition at $E = 26.04$ kV $cm^{-1}$ is shown in **Figure 6**. Here again, the splitting is clearly seen thanks to the Doppler-free technique applied here.

## Coherent spectroscopy

Although the line widths in the optical transitions observed in atoms in the gas phase at room temperature are larger than the spacing of the Zeeman and Stark splitting, the intervals of the sublevels themselves are scarcely altered by the thermal motion. In the coherent techniques, the relative energies between the sublevels are determined from the observation of the interference of amplitudes of coherent optical excitation followed by de-excitation.

Let us consider the case of two close-lying levels $|1\rangle$ and $|2\rangle$. It is possible that the atoms are simultaneously excited to these two levels from a common lower level $|i\rangle$ by a short-pulse laser with the pulse width of $\Delta t < \hbar/|E_2 - E_1|$, where $\hbar$ is the Planck constant divided by $2\pi$. Assume, for simplicity, that the populations of levels 1 and 2 decay into another common lower level $|f\rangle$ with the same decay constant $\gamma$. The total intensity, $I(t)$, of fluorescence emitted from either level will vary with time according to the form

$$I(t) = C\,e^{-\gamma t}\,(A + B\cos\omega_{21}t) \qquad [4]$$

where $A$, $B$ and $C$ are constants depending on both the relevant atomic wave-functions and the experimental arrangement, and $\omega_{21}$ is given by $\omega_{21} = |E_2 - E_1|/\hbar$. This behaves as an exponential decay $\exp(-\gamma t)$ suffering from a sinusoidal modulation with the angular frequency $\omega_{21}$, which is called a quantum beat.

In the case of the Zeeman (Stark) splitting, the Zeeman (or Stark) spectrum is reproduced in a Fourier transform of the time dependence in the fluorescence similar to Equation [4]. It is essential that the time response of the detection system is fast enough to observe oscillations with the characteristic period $2\pi/\omega_{21}$.

For other measuring techniques based on coherent excitation, see the 'Further reading' section for details.

## List of symbols

$b$ = distance of slit from orifice; $B$ = magnetic field; $2d$ = slit aperture; $E$ = electric field; $I$ = intensity; $k$ = photon wave vector; $k_B$ = Boltzmann constant; $m$ = magnetic quantum number; $m_a$ = atomic mass; $T$ = temperature; $\boldsymbol{v}$ = atomic velocity; $\gamma$ = decay constant; $\hbar$ = Planck constant divided by $2\pi$; $\omega$ = angular frequency of emitted light.

*See also:* **Atomic Fluorescence, Methods and Instrumentation; Laser Applications in Electronic Spectroscopy; Laser Spectroscopy Theory; Zeeman and Stark Methods in Spectroscopy, Applications**.

## Further reading

Demtroder W (1998) *Laser Spectroscopy*, 2nd edn. Berlin: Springer-Verlag.

Fukumi A, Endo I, Horiguchi T *et al.* (1997) Stark and Zeeman spectroscopies of $4f^66s6p^7G_{1-6}$ levels in Sm I under external electric and magnetic fields. *Zeitschrift für Physik D* **42**: 243–249.

Hanle W and Kleinpoppen H (1978) *Progress in Atomic Spectroscopy*. New York: Plenum Press.

Shimoda K (1976) *High-Resolution Laser Spectroscopy*. Berlin: Springer-Verlag.

Svanberg S (1992) *Atomic and Molecular Spectroscopy*, 2nd edn. Berlin: Springer-Verlag.

# Zero Kinetic Energy Photoelectron Spectroscopy, Applications

**K Müller-Dethlefs** and **Mark Ford**, University of York, UK

Copyright © 1999 Academic Press

**HIGH ENERGY SPECTROSCOPY**
**Applications**

## Introduction

NO has been studied extensively by photoelectron spectroscopy. A study using vacuum ultraviolet photoelectron spectroscopy by Turner and co-workers can be compared with a ZEKE study through the $A^2\Sigma^+$ state, using a $1 + 1'$ photon experiment. As can be seen in **Figure 1** the resolution is improved by approximately three orders of magnitude; resolving the rotational structure of the NO cation. Benzene and paradifluorobenzene have been studied using time-of-flight photoelectron spectroscopy. These two systems have also been studied using ZEKE spectro-scopy. With benzene, rotational resolution has again been obtained, as shown in the comparison in **Figure 2**. The two techniques are compared for paradifluorobenzene in **Figure 3**. Both of the above systems exhibit a breakdown in the Born–Oppenheimer approximation, and were useful indicators of the Herzberg–Teller, and Jahn–Teller effects.

ZEKE spectroscopy has been applied to a wide variety of molecular ions, clusters, van der Waals molecules, free radicals, reactive intermediates, and even to elusive transition states of chemical reactions. Examples of such typical applications of high-resolution ZEKE spectroscopy to molecules and clusters are given here. Compared to conventional photoelectron spectroscopy, ZEKE spectroscopy offers greatly increased spectral resolution, allowing the rotational structure of large molecular cations such as the benzene cation and the intermolecular vibrations of molecular clusters like phenol-water to be obtained.

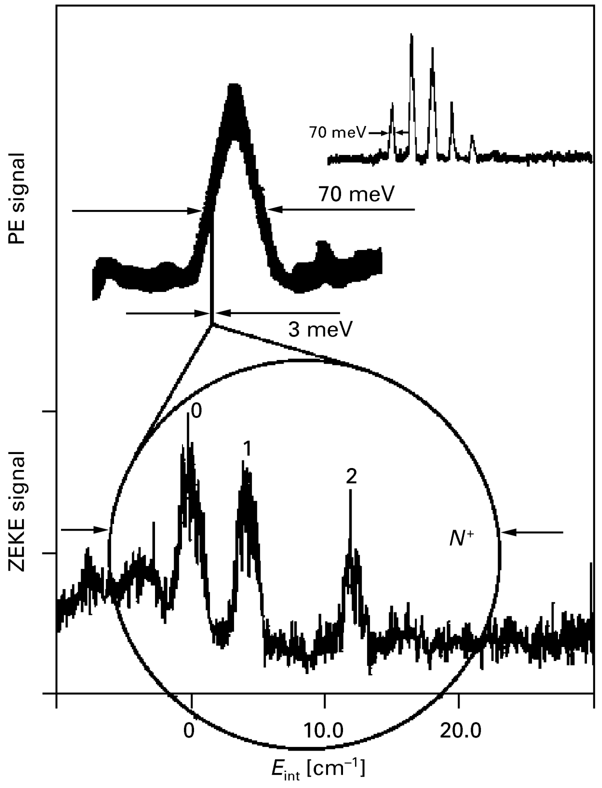



**Figure 1** A comparison between conventional VUV PES and ZEKE spectroscopy on NO; with the latter technique rotational resolution is attained.

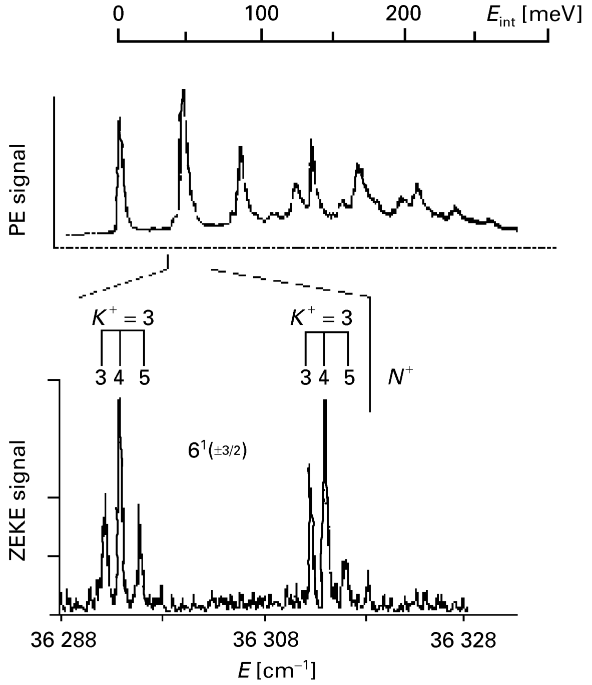



**Figure 2** The rotationally resolved ZEKE spectrum of benzene compared with time-of-flight PES; again the resolution is improved by several orders of magnitude.

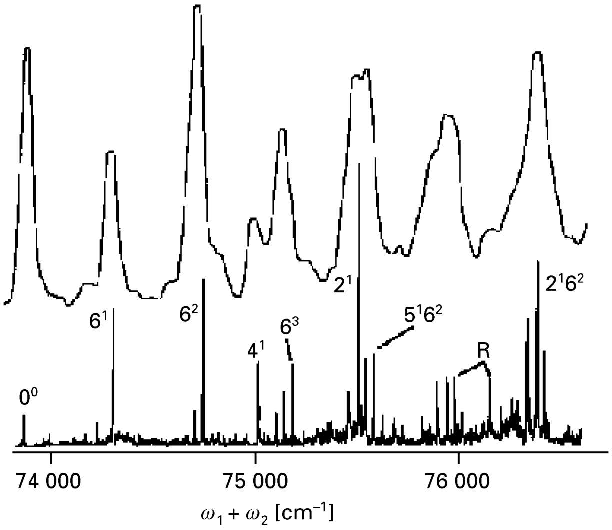



**Figure 3** In *para*-difluorobenzene, the vibrational structure of the cation was not fully resolved until the introduction of ZEKE spectroscopy.

## Smaller molecules

### Iodine ($I_2$)

Iodine has been studied extensively by ZEKE spectroscopy, including 2 + 1′ and 1 + 2′ schemes carried out by Cockett and co-workers. In the first case, a number of centrosymmetric Rydberg excited states acted as resonant intermediate states, and in the second case, the valence $B^3\Pi_0$ state was the intermediate. These studies demonstrate how well ZEKE spectroscopy can give a detailed vibrationally resolved spectrum and how autoionization is unavoidable in the photoelectron spectroscopy of small molecules.

The 2 + 1′ ZEKE spectra of $I_2$ exhibited non-Franck–Condon behaviour, having intense off-diagonal peaks in $v^+$, $v$, due to vibrational autoionization. **Figure 4** gives the spectra resulting from ionization through the band origin of the $[^2\Pi_{3/2}]_{core}\,5d;2_g$ state at about 62 600 cm$^{-1}$, and also through the first three vibrationally excited levels. This state was ionized into the lower spin-orbit state of the ion. Conversely **Figure 5** gives the spectra recorded through the first three vibrational levels of the $[^2\Pi_{1/2}]_{core}\,5d;2_g$ Rydberg state at about 68 000 cm$^{-1}$, which was ionized into the upper spin-orbit state of the ion. For the spectrum given in **Figure 4A**, which was through the origin, the $\Delta v = 0$ transition was most intense; the total transition energy to this level is 75 066 ± 2 cm$^{-1}$, which is the adiabatic ionization energy (to the ion in

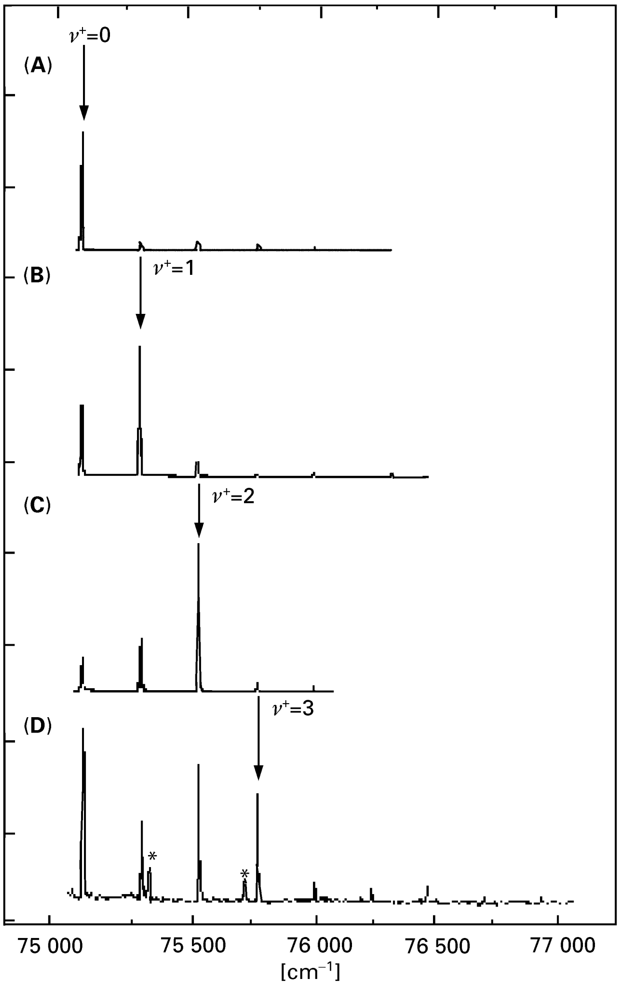



**Figure 4** ZEKE spectrum of $I_2$ recorded through the $[^2\Pi_{3/2}]_{core}$ 5d;$2_g$ state; the arrows indicate diagonal transitions, and the asterisks accidental resonances with A ← X transitions.

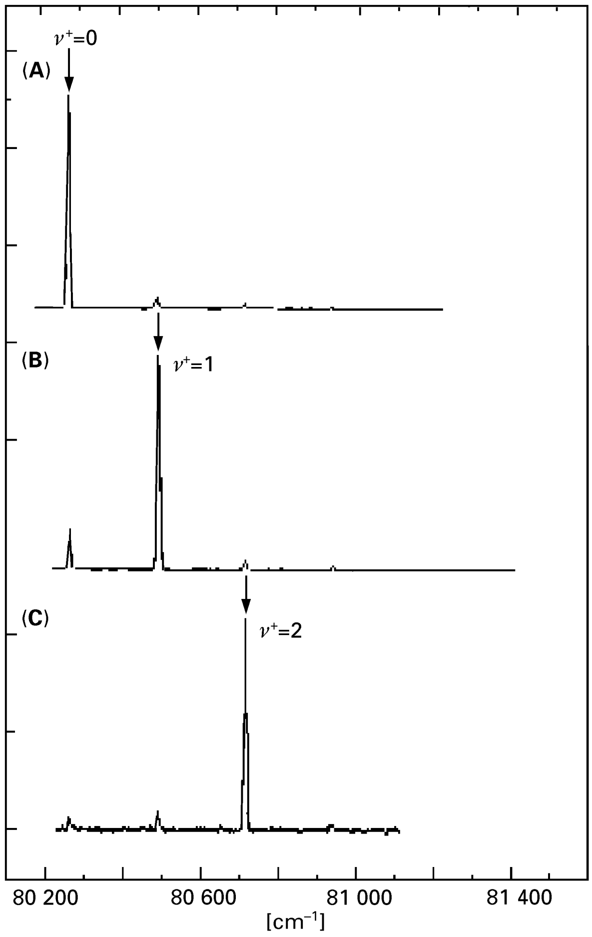



**Figure 5** ZEKE spectrum of $I_2$ recorded through the $[^2\Pi_{3/2}]_{core}$ 5d;$2_g$ state; the arrows indicate diagonal transitions.

its ground vibrational state). There is a weak vibrational progression from the band origin up to $v^+ = 4$. The dominance of the origin peak indicates a minimal change in geometry on ionization. This minimal change in geometry is expected after ionization from the (intermediate) Rydberg state; however, the progression does extend to higher $v^+$ than is expected merely on the basis of Franck–Condon factors. For the $v = 1$ intermediate state (**Figure 4B**) the non-Franck–Condon behaviour is even more pronounced: the stretching progression can be followed up to $v^+ = 7$; the $\Delta v = 0$ transition remains dominant; however, the peak is of about the same intensity as the $v^+ = 0$ peak. For the intermediate states $v = 2$ (**Figure 4C**) and $v = 3$ (**Figure 4D**) the progressions become longer. Although the spectra show a $\Delta v = 0$ propensity, a Franck–Condon envelope does not fit the intensity distribution.

The pattern for the vibrational propensity was found to be quite different in the upper spin-orbit state of the ion. In the spectrum recorded through the origin (**Figure 5A**) the most intense peak corresponds to the $\Delta v = 0$ transition (to the $v^+ = 0$ level of the ion). The corrected total transition energy to the origin is 80 266 ± 2 cm$^{-1}$; this again corresponds to the adiabatic ionization energy. This result, combined with the ionization energy for the lower $^2\Pi_{3/2}$ spin-orbit state, gives an improved spin-orbit splitting constant for $I_2^+$ in its ground electronic state of 5197 ± 4 cm$^{-1}$. For the upper spin-orbit component, the propensity for the $\Delta v = 0$ transition remains strong even as the vibrational level of the intermediate is increased. This corresponds to classic Franck–Condon behaviour.

The non-Franck–Condon intensities observed in the lower ($^2\Pi_{3/2}$) spin-orbit state spectrum were attributed to autoionization involving another Rydberg series which converges to a higher

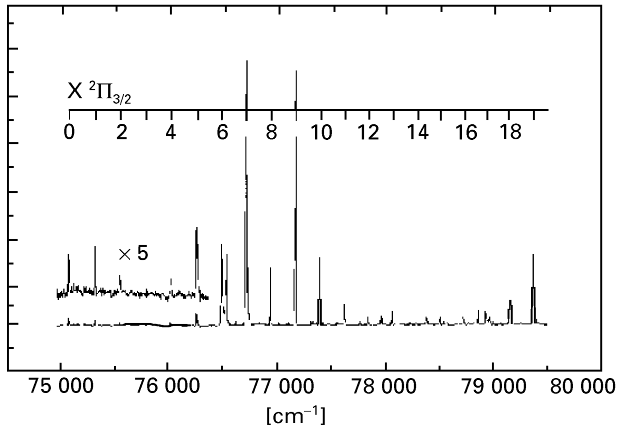



**Figure 6** The long Franck–Condon forbidden progression exhibited in the ZEKE spectrum of the lower spin-orbit state recorded through $\nu$ = 15 of the $B^3\Pi_u 0^+_u$ state.

vibrational state of the upper spin-orbit state in the ion. An interaction between the two Rydberg series, at the ionization threshold of the lower series, gave rise to autoionization from the higher series. No interaction occurs with the upper ($^2\Pi_{1/2}$) spin-orbit state, as there are no nearby Rydberg series.

The 1 + 2′ photon study, via the valence $B^3\Pi_0$ state, extended the previous work. Long Franck–Condon progressions, arising from the valence character of the intermediate state, are evident in the ZEKE spectra of both spin-orbit components. In the lower spin-orbit component, the vibrational progression extends to at least $\nu^+ = 62$, and in the upper state as high as $\nu^+ = 34$. The spectrum in the range 75 000 to 80 000 cm $^{-1}$ of the lower spin-orbit state, which was recorded via $\nu$ = 15, is shown in **Figure 6**.

The vibrational progressions can be adequately simulated through the calculation of Franck–Condon factors; however the observed spin-orbit branching ratio, along with the intensity distribution, reflects a considerable contribution from both spin-orbit and field-induced resonant autoionization processes. Also, accidental resonances at the two-photon level with ion-pair states further perturb the distribution of peak intensities.

### HCl

From the ZEKE spectra of hydrogen halides, the rotational-state distribution of the product ion provides a direct measure of the angular momentum of the outgoing electron; this is a sensitive probe of ionization dynamics. Autoionization occurs very readily in these molecules, *via* rotational, vibrational and electronic pathways, and is often evident in the spectra recorded. A further motivation in much of the work on the hydrogen halides has been to investigate the artefacts of autoionization, particularly the role of rotational and spin-orbit autoionization processes. In the cation there is a $\pi$-vacancy in the ground electronic configuration. Hence, spin-orbit coupling is evident in the spectra.

HCl has been studied by both single-photon and two-colour multiphoton experiments. The studies focused on the vibrational ground state of the ion and observed a tendency to large changes in angular momentum on ionization, $|\Delta J| \leq 7/2$, indicating a preference for an outgoing d-partial wave. A preference for negative values of angular momentum transfer for both spin-orbit components was observed, which has been evident in many ZEKE spectra which exhibit rotational resolution; this is attributed to rotational autoionization. In the single-photon experiment by White and co-workers, anomalous branch intensities in the ZEKE spectrum were interpreted in terms of field- or dipole-induced mixing of Rydberg states converging on higher-ion rotational levels. Intensity anomalies were observed in the spin-orbit and rotational branching ratios of two-colour ZEKE spectra of de Lange and co-workers recorded *via* the $F\,^1\Delta_2$, $D^1\Pi_1$ and $f^3\Delta_2$ Rydberg states. The branching ratios were dependent on three experimental parameters:

(i) The delay time employed between excitation and ionization;
(ii) The magnitude of the bias electric field;
(iii) The magnitude of the applied pulsed electric field.

The results were rationalized on the basis of the increasing number of autoionization decay channels that become available to the high-*n* Rydberg states as each ionization threshold is reached. An analysis of the decay-dependence of the ZEKE spectra *via* the $F\,^1\Delta_2$ state provided evidence for a non-exponential decay of the high-*n* Rydberg states.

Studies on the other hydrogen halides have provided similar results; from HF, however, it was concluded that the s-channel dominates the photoionization process as opposed to the d-channel in HCl, HBr and HI.

### Ammonia ($NH_3$)

The $NH_3$ molecule, when studied by Habenicht and co-workers, was the first polyatomic molecule studied by ZEKE spectroscopy for which full rotational resolution in the cation was achieved. Ionization was achieved by a 2 + 1′ process, excited through the two-photon transition $\tilde{B} \leftarrow \tilde{X}$. The excitation

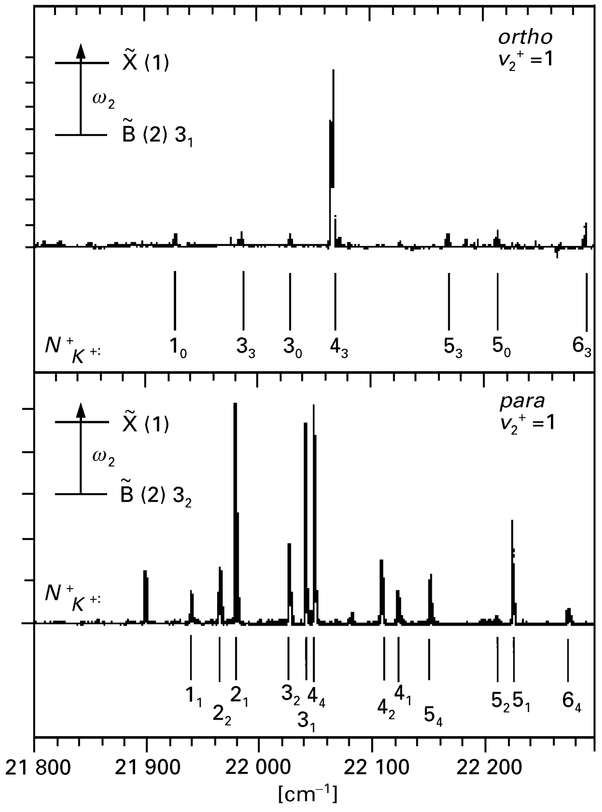



**Figure 7** The symmetry-based selection rules applied to ZEKE spectroscopy are borne out by the different spectra obtained using the *ortho*- and *para*-nuclear spin states of ammonia.

spectrum for the $2^2_0$ transition obtained by REMPI shows clearly resolved rotational structure with a Coriolis interaction giving *l*-type doubling. This allows for rotational-state selectivity in the intermediate state. The two rotational states corresponding to the *ortho*- ($J'_{K'} = 3_1$) and para- ($J'_{K'} = 3_2$) nuclear spin states of $NH_3$ were chosen as intermediate states for the ZEKE spectra. These two ZEKE spectra, recorded through the B̃-state, are given in **Figure 7**, taking the $2^2_1$ vibronic transition. There is a clear difference between the spectra obtained for *ortho*- (top spectrum) and *para*-$NH_3$ (bottom spectrum). The ZEKE spectrum of *ortho*-$NH_3$ shows one strong transition into the ion rotational state with $N^+_{K+} = 4_3$, and other transitions with $K^+ = 0$ and 3, which are considerably weaker. On first sight this ZEKE spectrum appears similar to an atomic photoionization spectrum. On the other hand the ZEKE spectrum of *para*-$NH_3$ is much fuller giving the strongest lines observed for $K^+ = 1$ and $N^+_{K+} = 4_4$; also there are weaker transitions into $K^+ = 2$. These spectra are in very good agreement with the symmetry selection rules that apply to ZEKE transitions.

## Larger molecules

### Benzene ($C_6H_6$)

The neutral benzene molecule has a hexagonal, planar structure with $D_{6h}$ symmetry; in the electronic ground state the electronic configuration is $a_{2u}^2\ e_{1g}^4$. When benzene is ionized, one of the $e_{1g}$ electrons is removed, leaving one $e_{1g}$ electron unpaired. Thus, the cation has a doubly degenerate $^2E_{1g}$ electronic ground state. The Jahn–Teller theorem predicts that for any nonlinear polyatomic molecule in a degenerate electronic state, there exists a distortion of nuclear geometry along at least one non-totally symmetric normal coordinate that results in a splitting of the potential-energy function such that the potential minimum is no longer at the symmetrical position.

The structural distortions of the benzene cation have been discussed at length; quantum-chemical *ab initio* calculations predict three equivalent $D_{2h}$ in-plane distortions corresponding to elongation, or compression along three of the twofold-symmetry axes of benzene. These give structures more stable than the hexagonal structure, with an experimentally determined stabilization energy of 266 cm$^{-1}$. This is approximately half the zero-point energy of the lowest-frequency Jahn–Teller active normal vibration.

For weak Jahn–Teller coupling, the stabilization energy for the distorted symmetry is much smaller than the zero-point energy of the Jahn–Teller active mode. Under collision-free conditions, the three equivalent $D_{2h}$ structures of the cation would dynamically interconvert rapidly, and the ground state of the cation would still be described in the $D_{6h}$ symmetry group. For strong Jahn–Teller coupling the cation would spend much time in one of the three structures, and would therefore be described in $D_{2h}$. The knowledge of the structure and the symmetry of the isolated benzene cation is desirable not only for testing quantum-mechanical model calculations: it also has a fundamental importance for organic chemistry. Through group-theoretical considerations, rotationally resolved ZEKE spectroscopy gives a clear and unambiguous determination of the symmetry of the cation. Thus, if the molecule were statically distorted to lower symmetry, transitions would appear in the rotationally resolved ZEKE spectra, which are forbidden in the $D_{6h}$ structure. Thus, the observed rotational transitions are a sensitive and clear indication of the symmetry of the cation.

If one quantum of a Jahn–Teller active normal vibration in the benzene cation (these are the modes $v_{6-9}$ with $e_{2g}$ symmetry) is excited, the linear dynamic Jahn–Teller coupling leads to a splitting into two vibronic states with $j = \pm 1/2$ ($E_{1g}$ symmetry) and

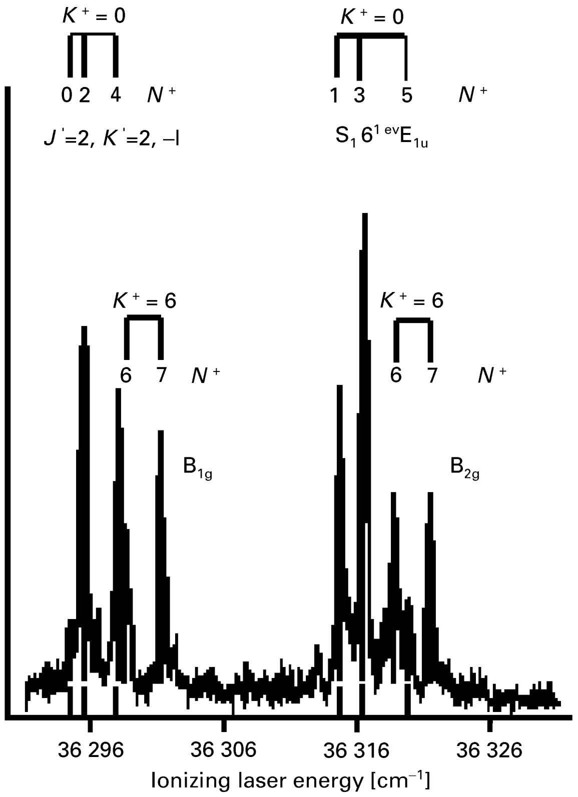



**Figure 8** In the lower-energy band, recorded through the $S_1\ ^16^1$ $E_{1u}$, $2_2$-l state, only even *K* are observed, whereas in the higher-energy band odd *K* are observed, indicating rigorously the symmetry of the vibronic state associated with each band.

$j = \pm 3/2$ ($B_{1g} \oplus B_{2g}$ symmetry) vibronic angular momentum. The $j = \pm 3/2$ states are further split by quadratic dynamic Jahn–Teller coupling, but the $j = \pm 1/2$ state remains doubly degenerate.

Detailed vibronic structure is seen in the low-resolution scan of the ZEKE spectrum of benzene, recorded via the $6^1$ vibrational level in the $S_1$ state of the neutral. Bands with fundamental frequencies characteristic of the $\nu_6$, $\nu_{16}$, $\nu_4$ and $\nu_1$ vibrational modes are seen. However, no harmonic progressions can be observed, with the higher-energy portion of the spectrum exhibiting a highly irregular and dense system of vibronically active states. The active mode of lowest energy is $\nu_6$, which is along the coordinate predicted for the Jahn–Teller distortion by *ab initio* methods. A key pair of bands, which appear at about 350 cm$^{-1}$, corresponding to the ion internal energy, are the $B_{1g}$ and $B_{2g}$ vibronic components of the $\nu_6$ fundamental, shifted to lower energy by linear Jahn–Teller coupling. However, the relative ordering of $B_{1g}$ and $B_{2g}$ is unclear in this spectrum. The conservation of symmetry of the nuclear spin wavefunction restricts the possible transitions to these two vibronic states; thus, depending on the vibronic symmetry, only certain rotational progressions can be observed.

A high-resolution ZEKE scan of the $B_{1g}$ and $B_{2g}$ components of the $6^0$ band recorded by exciting through the $6^1$, $J'_{K'} = 2_2$, -$l'$, $S_1$ state is shown in **Figure 8**. In this spectrum, only the $K^+ = 0$ projections of even $N^+$ are seen in the lower-energy vibronic component, whereas only those from odd $N^+$ are seen in the higher-energy vibronic component. This effect can be attributed to nuclear spin statistics, and indicates unambiguously that the lower-energy vibronic component has $B_{1g}$ symmetry, and that the higher-energy vibronic component has $B_{2g}$ symmetry. Thus the rotational structure in the ZEKE spectrum has established that the $B_{2g}$ level in the quadratically split $6^1$ ($j = 3/2$) levels of the benzene cation, lies above the $B_{1g}$ level.

The cation is apparently distorted to a small extent by quadratic Jahn–Teller coupling in $\nu_6$. It has been concluded from the rotational intensities in the ZEKE spectrum that the wells in the pseudorotation coordinate correspond to local $B_{1g}$ electronic configurations (the elongated structure) whereas the saddle points are locally $B_{2g}$ (compressed). From the coupling parameters that fit the vibronic structure in the coarse ZEKE spectrum, an energy difference between the stationary states of only 8 cm$^{-1}$ is established. This is much less than the $\nu_6$ zero-point energy of 413 cm$^{-1}$, indicating that the benzene cation is fluctuational, and therefore must be viewed in $D_{6h}$ symmetry rather than in terms of the three $D_{2h}$ structures with locally non-degenerate electronic configurations.

## Toluene

The toluene molecule and its torsional states are classified according to its irreducible representations in the molecular symmetry (MS) group $G_{12}$, which is isomorphic to the point group $D_{3h}$. The problem of an unhindered, rigid methyl rotor attached to a rigid frame reduces to a one-dimensional Schrödinger equation with eigenfunctions, $\Psi_{tor}(\alpha) = (2\pi)^{-1/2}e^{\pm im\alpha}$ and eigenvalues, $E_M = Bm^2$, where $B$ is an effective constant for the methyl-group rotation relative to the rigid molecular framework and $m = 0, \pm 1, \pm 2, \ldots$ is the rotational quantum number. In toluene, the torsional potential has 6-fold symmetry. A small torsional perturbation of the form $V(\alpha) = (V_6/2)(1-\cos 6\alpha)$ leaves the $m = \pm 1$, $\pm 2$, $\pm 4$ and $\pm 5$ states doubly degenerate but lifts the degeneracy of the $m = \pm 3$ and $m = \pm 6$ levels.

$\alpha = 0$ is defined as an eclipsed geometry, and $\alpha = \pi/6$ a staggered geometry. The rotor states are termed $0a_1'$, $1e''$, $2e'$, $3a_2''$, $3a_1''$, $4e'$, etc. Band intensities in the $S_1$-$S_0$ and cation-$S_1$ spectra reveal which $|m| = 3$ state lies lower in energy and thus the absolute phase of the torsional potential. For

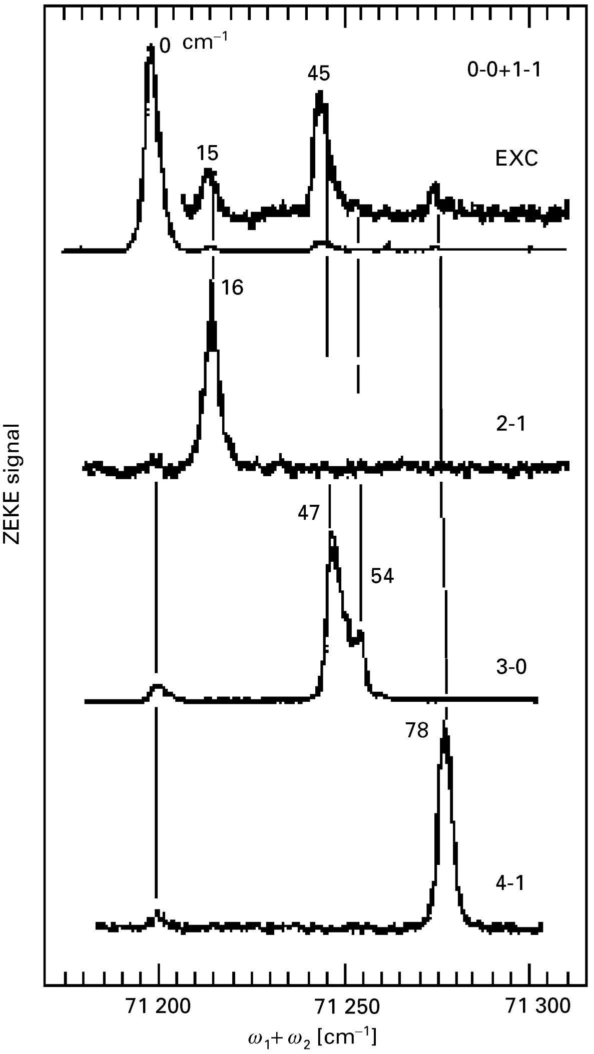



**Figure 9** ZEKE spectra recorded through various torsional levels of the $S_1$ state in toluene. The label EXC indicates the torsional transition to the intermediate: $S_0 \leftarrow S_1$.

negative $V_6$ (potential minimum at $\alpha = \pi/6$), $3a_2''$ lies below $3a_1''$, whereas for positive $V_6$ (potential minimum at $\alpha = 0$), $3a_1''$ lies below $3a_2''$.

Weisshaar and co-workers recorded ZEKE spectra of toluene through different intermediate resonances of $S_1$ and they are presented in **Figure 9**. From the assignment it can be shown that there is a positive $V_6$ with $3a_1''$ lower in energy than the $3a_2''$ torsional state. The torsional states of the toluene cation (0, 15, 54 and 75 cm$^{-1}$) are not very different from the torsional states in the $S_1$ state (1, 15, 55 and 77 cm$^{-1}$), leading to the conclusion that the torsional barriers in the cation and in $S_1$ are quite similar.

For both $0a_1'$–$0a_1'$ and $3a_1''$–$0a_1'$ excitations of $S_1$, the ZEKE band at 46 cm$^{-1}$ is stronger than the 54 cm$^{-1}$ band. Using the assumption that torsion–electronic coupling allows all *a*–*a* torsional transitions and forbids *a*–*e* transitions and the fact that the 47 cm$^{-1}$ band is three times more intense than the 54 cm$^{-1}$ band for ionization through the $3a_1''$ state of $S_1$, the 47 cm$^{-1}$ band is assigned to the $3a_1''$–$3a_1''$ transition and the 54 cm$^{-1}$ band to $3a_2''$–$3a_1''$. These assignments for the ZEKE spectra constitute a major step in understanding large-amplitude motions and the role of torsional–electronic couplings.

## Weakly bound molecules

### Phenol–methanol

The study of the PhOH-MeOH complex, by Müller-Dethlefs and co-workers, was carried out with a 1 + 1′ REMPI spectrum. This was difficult to interpret owing to the comparatively dense vibrational structure. Various vibrational levels in the $S_1$ state were used as intermediate states on the way to ionization. The ZEKE spectrum obtained by exciting *via* the $S_1$ vibrationless level, given in **Figure 10**, is striking with progressions of about ten quanta in a low-frequency vibrational mode of 34 cm$^{-1}$, denoted $h_1$, appearing in combination with components of an anharmonic progression of the intermolecular stretch of 278 cm$^{-1}$. The pattern suggests a rather substantial change of geometry upon ionization. The adiabatic ionization energy was derived as 63 207 ± 4 cm$^{-1}$ which is a red shift from the $S_0$ state of 5421 ± 8 cm$^{-1}$, indicating a large increase in bond strength. The latter point is also exemplified by the large increase in the energy of the intermolecular stretch compared with 176 cm$^{-1}$ in the $S_1$ state and 162 cm$^{-1}$ in the $S_0$ state. Additionally, between the latter components, another set of progressions of the intermolecular mode of 34 cm$^{-1}$ were seen, this time in combination with a third intermolecular mode of 52 cm$^{-1}$.

The ZEKE spectrum obtained *via* the $S_1$ state, with one quantum of the lowest-frequency intermolecular mode excited, showed the same vibrations but with a substantially changed Franck–Condon envelope which allowed the identification of a fourth intermolecular mode, denoted $h_4$, at 153 cm$^{-1}$. A slightly different envelope for the 34 cm$^{-1}$ vibration was also obtained when exciting through the $S_1$ state with one quantum of the intermolecular stretch excited. The other two intermolecular modes of the Ph-MeOH cationic complex were identified from the ZEKE spectrum obtained *via* a combination band; their values being 76 cm$^{-1}$ and 158 cm$^{-1}$.

### $I_2$-Ar

Iodine–argon was one of the first complexes to be studied in the early jet spectroscopy experiments

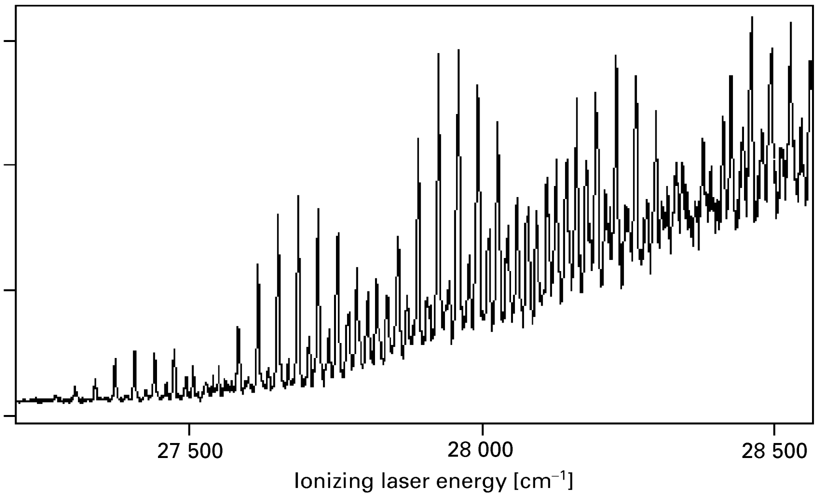



**Figure 10** The striking vibrational progression seen in the ZEKE spectrum of phenol–methanol.

conducted by Levy and co-workers in the 1970s. Most of the work carried out on this complex since then has been concerned with the $\tilde{B}^3\Pi_0^+ \leftarrow \tilde{X}^1\Sigma^+_g$ system studied using laser-induced fluorescence (LIF) spectroscopy. This state is not well-suited as a resonant intermediate in ZEKE spectroscopy, as it lies only 15 800 cm$^{-1}$ above the electronic ground state. Since this study a number of grade ns and nd Rydberg excited states, based upon both spin-orbit states of the $\tilde{X}^2\Pi_{\Omega,g}I_2^+$-Ar core, have been characterized using 2 + 1 mass-resolved REMPI spectroscopy, by Cockett and co-workers. These Rydberg states lie between 53 000 and 69 000 cm$^{-1}$ and are better suited as resonant intermediate states for ionization into the two spin-orbit components of the $\tilde{X}^2\Pi_{\Omega,g}$ ionic ground state.

Substantial differences in the binding energy were observed for the 6s, 5d and 6d Rydberg states of the $I_2$-Ar complex, which correlate with the degree to which the positive charge of the core is shielded from the argon atom by the Rydberg electron. More-penetrating Rydberg orbitals reduce the charge-induced dipole forces between iodine and argon. The observed binding energy increases are seen in the REMPI spectra as progressive increases in spectral red shifts and intermolecular van der Waals stretching frequencies.

An issue which had been the subject of considerable debate was whether $I_2$-Ar adopts a T-shaped or linear geometry. Initial speculation suggested that it should be linear by analogy with the known linear geometry of the ClF-Ar complex. However the first direct experimental evidence for the geometry of $I_2$-Ar emerged from a partially rotationally resolved B-X fluorescence excitation spectrum, which showed that $I_2$-Ar adopts a T-shaped geometry. It was also suggested that a linear isomer was responsible for an observed fluorescence excitation continuum underlying the discrete B-X transitions.

The mass-resolved 2 + 1 REMPI spectrum recorded by monitoring the $I_2^{\bullet+}$-Ar mass channel is shown in **Figure 11**. The spectrum is composed of partially overlapping vibrational progressions arising from the $[^2\Pi_{3/2}]_c$ 5d; $2_g$ (**Figure 11A**; recorded with circularly polarized light) and $[^2\Pi_{3/2}]_c$ 5d; $0^+_g$ (**Figure 11B**; recorded with linearly polarized light) Rydberg states of $I_2$-Ar. The vibrational structure for both states essentially arises from simultaneous excitation of both the I-I stretch ($\nu_1$) and the $I_2$-Ar van der Waals stretch ($\nu_3$). For the $[^2\Pi_{3/2}]_c$ 5d; $0^+_g$ state, the progression terminates abruptly at $\nu_3\nu_1$ which suggests that the complex dissociates at this point. The $\nu_3\nu_1$ band appears at an internal vibrational energy of 758 cm$^{-1}$, but the spectral red shift of the band origin dictates a lower limit to the zero-point dissociation energy for this state of 563 ± 5 cm$^{-1}$. Thus, it would appear that the coupling between the two vibrational modes is sufficiently weak to enable the complex to accommodate an excess 195 cm$^{-1}$ of internal energy before it dissociates. This is consistent with the geometry of the complex being T-shaped.

The $[^2\Pi_{3/2}]_c$ 5d; $2_g$ Rydberg state progression shown in **Figure 11A** has an additional complexity of structure when compared to the progression observed for the $0^+_g$ state. For both $\nu_1 = 0$ and $\nu_1 = 1$, each vibrational band appears significantly broader than observed for the $0^+_g$ progression, and on most of the bands, an apparent splitting can be resolved. However, for $\nu_1 = 2$ the doublet structure has disappeared and the peaks have adopted the narrower profile seen in the $0^+_g$ progression. In fact, the

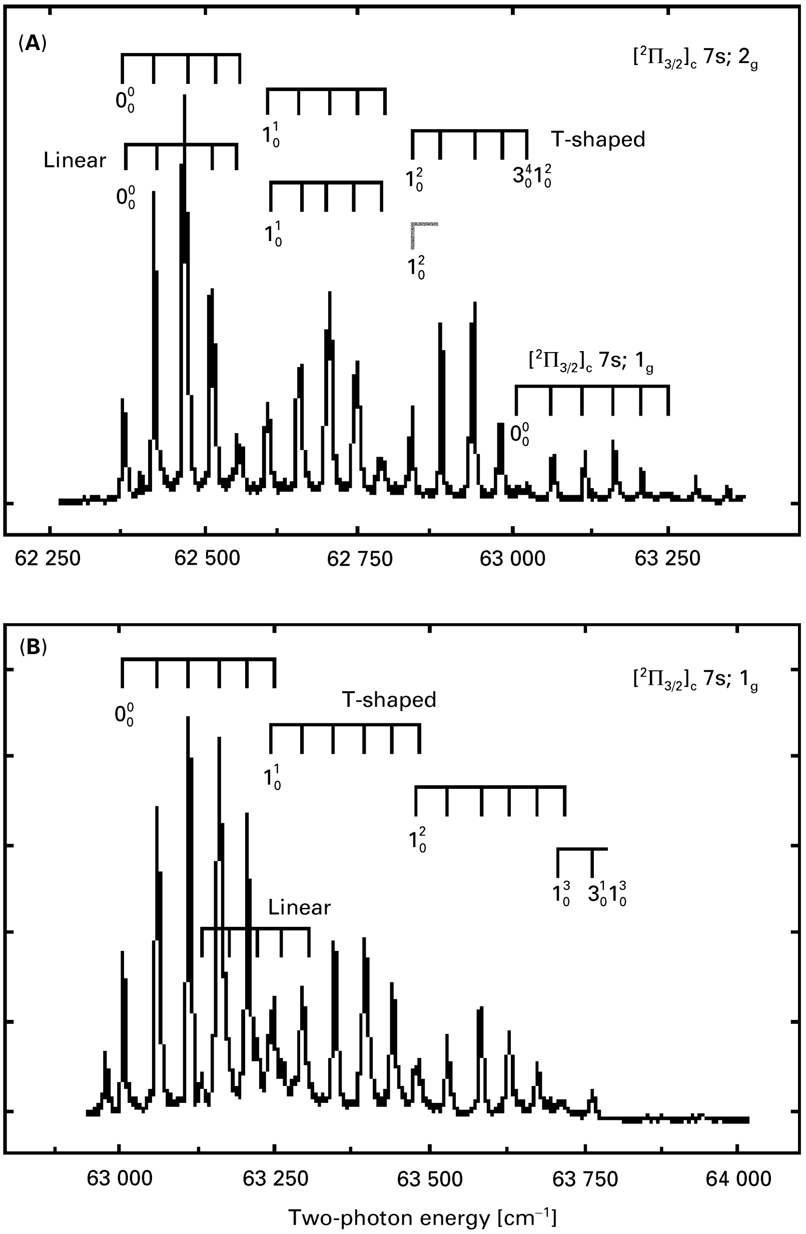



**Figure 11** REMPI spectra of $I_2$-Ar showing the assignments attributed to each of the two structural isomers; (A) was recorded through the $[^2\Pi_{3/2}]_{core}$ 5d; $2_g$ state with circularly polarized light, and (B) was recorded through the $[^2\Pi_{3/2}]_{core}$ 5d; $0^+_g$ state with linearly polarized light.

doublet structure arises, not from any splitting of the peaks, but from two partially overlapping vibrational progressions with near-identical band origins. It appears from the spectrum that one of the progressions terminates at $\nu_1 = 1$ while the other continues at least as far as $\nu_1 = 2$.

Assuming that the point at which the progression terminates represents the point at which the complex dissociates, the conclusion drawn was that for the shorter progression, the onset of dissociation occurs at 463 cm$^{-1}$ internal vibrational energy (compared with a calculated value for $D_0$ of about 503 cm$^{-1}$), while for the longer progression, dissociation occurs at a lower limit of 655 cm$^{-1}$. On this basis, a provisional assignment was made of the shorter progression to the linear isomer, for which the van der Waals stretch might be expected to couple more efficiently with the $I_2$ stretch, and the longer progression to the T-shaped isomer.

Although the REMPI spectrum certainly provided a great deal of circumstantial evidence for the existence of two isomers, the assignment was to be greatly strengthened by extending the study to the ionic state by using ZEKE spectroscopy to probe

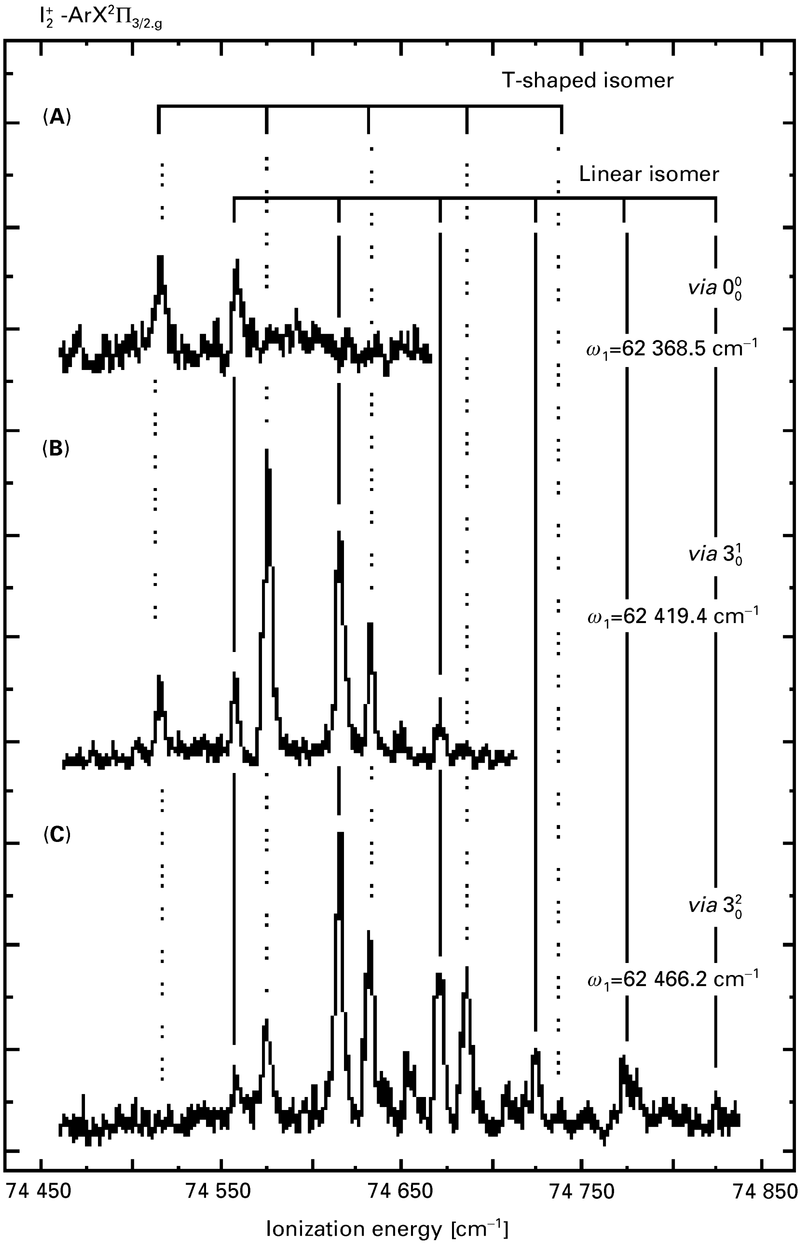



**Figure 12** The ZEKE spectra of $I_2$-Ar recorded through the $[^2\Pi_{3/2}]_{core}$ 5d; $2_g$ state. (A) is through the overlapping ($0^0_0$) vibrational levels, (B) the ($3^1_0$) levels and (C) the ($3^2_0$) levels.

the $[^2\Pi_{3/2}]_c$ 5d; $2_g$ Rydberg state in a two-colour 2 + 1′ ionization scheme, and recording isomer-specific ZEKE spectra. The ZEKE spectra of the ground electronic state of the ion recorded *via* several intramolecular vibrational levels in the $[^2\Pi_{3/2}]_c$ 5d; $2_g$ Rydberg state show, in each case, two well-separated vibrational progressions (**Figure 12**). As the level of vibrational excitation is increased in the intermediate Rydberg state, so the vibrational activity in the resulting ZEKE spectra increases.

The general propensity for diagonal transitions was consistent with the fairly small changes in geometry that occur on exciting from the Rydberg state to the ion. The experimental observations in this case were consistent with an assignment of the two overlapping Rydberg state progressions to two geometrical isomers. The measured difference in the ionization energies of 43 $cm^{-1}$ for the two isomers provided an indication of their relative stabilities in the ion, with the linear isomer being the more weakly bound.

## List of symbols

$B$ = constant for methyl-group rotation; $j$ = vibronic angular momentum; $J$ = total angular momentum with a projection of $K$ in the molecular frame; $m$ = rotational quantum number; $n$ = principal quantum number; $N$ = total angular momentum excluding spin; $\alpha$ = torsional angle; $\nu$ = vibrational quantum number; $\omega$ = corresponding vibrational constant.

*See also:* **Photoelectron Spectrometers; Photoelectron Spectroscopy; Zero Kinetic Energy Photoelectron Spectroscopy, Theory.**

## Further reading


Cockett M, Müller-Dethlefs K and Wright TG (1994) Recent applications and developments in ZEKE spectroscopy. *Royal Society of Chemistry Annual Reports Section C* **94**: Chapter 9, pp 327–373.

Habenicht W, Reiser G and Müller-Dethlefs K (1991) High resolution zero kinetic energy electron spectrum of ammonia. *Journal of Chemical Physics* **95**: 4809–4820.

Haines S, Dessent C and Müller-Dethlefs K Mass analysed threshold ionization of phenol-CO, intermolecular binding energies of a hydrogen bonded complex. *Journal of Chemical Physics* (submitted).

Lindner R, Müller-Dethlefs K, Wedum E, Haber K and Grant ER (1996) On the shape of $C_6H_6^+$. *Science* **721**: 1698–1702.

Müller-Dethlefs K (1995) High resolution spectroscopy with photoelectrons: ZEKE spectroscopy of molecular systems. In: Powis I, Baer T and Ng C-Y (eds) *High Resolution Laser Photoionization and Photoelectron Studies*, Chapter 2, pp 22–78. Wiley.

Müller-Dethlefs K and Cockett M (1998) *Nonlinear Spectroscopy for Molecular Structure Determination*, Chapter 7, pp 167–201. Oxford: Blackwell Science.

Müller-Dethlefs K and Schlag EW (1998) Chemical applications of zero kinetic energy (ZEKE) photoelectron spectroscopy. *Angewandte Chemie, International Edition English* **37**: 1347–1374.

Müller-Dethlefs K, Dopfer O and Wright TG (1994) ZEKE spectroscopy of complexes and clusters. *Chemical Reviews* **94**: 1845–1871.

Wang K and McKoy V (1995) High resolution photoelectron spectroscopy of molecules. *Annual Review of Physical Chemistry* **46**: 275–304.


# Zero Kinetic Energy Photoelectron Spectroscopy, Theory

**K Müller-Dethlefs** and **Mark Ford**, University of York, UK

Copyright © 1999 Academic Press

HIGH ENERGY SPECTROSCOPY

Theory

## Introduction

Understanding of chemistry and the chemical bond is greatly influenced by molecular orbital theory. The power of the molecular orbital approach in providing an understanding of the structure and reactivity of molecules lies in its description of chemical bonds. The legitimacy of a molecular orbital description is attested to by the results of molecular electronic spectroscopy. In the single-electron molecular orbital picture, photoionization involves the excitation of an electron from a bound orbital into the ionization continuum. The energy from the photon is partitioned between the kinetic energy of the electron and the ionization energy as in Equations [1] and [2]. Thus by selecting the electron kinetic energy a spectrum can be recorded.

$$h\nu = E_{\mathrm{ionization}} + E_{\mathrm{kinetic}} \quad [1]$$

$$E_{\mathrm{ionization}} = E_{\mathrm{internal}}(\mathrm{ion}) - E_{\mathrm{internal}}(\mathrm{neutral}) \quad [2]$$

To a first approximation the ionization energy is equal to the energy of the orbital from which the electron originates in the ground-state molecule. This is known as Koopmans' theorem, and allows each signal in the spectrum to be assigned to an orbital. Koopmans' theorem provides evidence that molecular orbitals are conceptually valid and has assumed a very important place in the development of our understanding of the electronic structure of molecules.

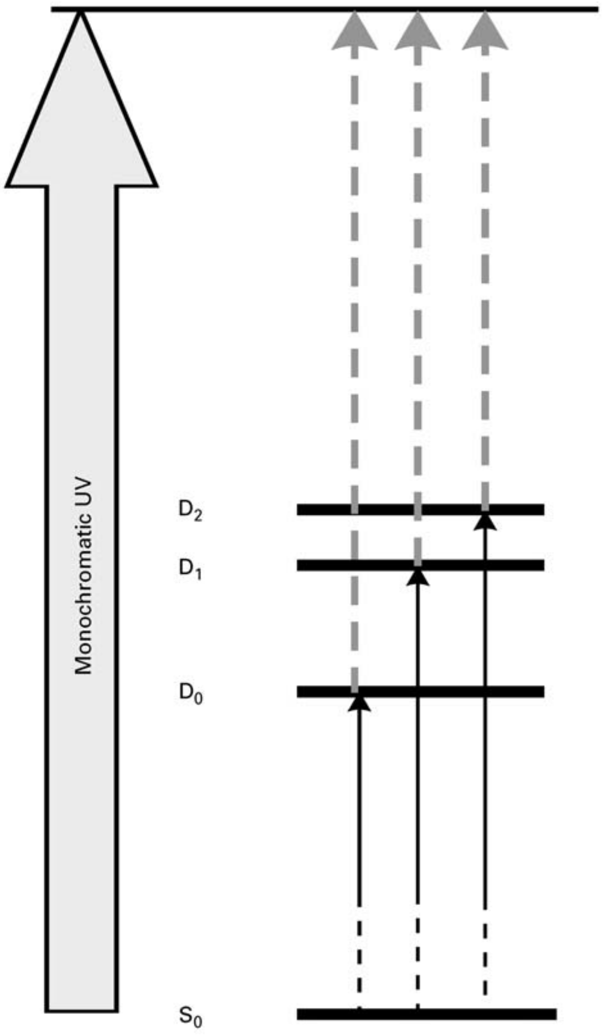



**Figure 1** An ionization scheme for conventional PES.

Koopmans theorem is limited by its neglect of molecular orbital reorganization, and transitions between states of correct symmetry species that are allowed; hence it can be quite misleading. Thus a better picture is to say that, at low resolution, the signals observed correspond to different electronic states in the cation. At higher resolution, fine structure in the photoelectron spectrum corresponds to the vibronic levels of the cation; a broad band, with significant vibronic structure, indicates a large change in structure between the neutral molecule and the cation. This is due to the improved overlap with vibrational overtones, arising as a result of the Franck–Condon principle. An ionization scheme for conventional PES is given in **Figure 1**.

The setup for a conventional photoelectron spectrometer is illustrated in **Figure 2**. In modern spectrometers the sample is introduced in a molecular beam, with a very low temperature. The major restriction on the resolution of PES, however, arises

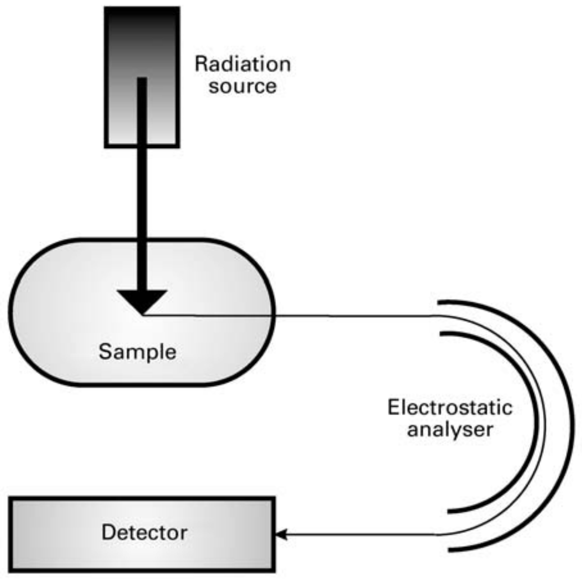



**Figure 2** The setup for conventional photoelectron spectroscopy.

through the practical resolution of the geometrical or time-of-flight photoelectron analysers used to analyse the kinetic energy; in order to have a reasonable signal strength this is about 10 meV (80 $cm^{-1}$).

The energy of the radiation source used in photoelectron spectroscopy determines the configuration of the ion. If a high-energy source is used as in X-ray photoelectron spectroscopy (XPS) the electron is emitted from a core orbital, whereas ultraviolet photoelectron spectroscopy (UVPS) will only give signals from the valence orbitals of molecules with relatively low ionization energies.

## ZEKE photoelectron spectroscopy

The method of ZEKE spectroscopy was invented by Müller-Dethlefs, in 1984, to bypass the resolution limitation. The technique uses a delayed electric field pulse to extract only the electrons ionized without kinetic energy, as shown in **Figure 3**. A signal is only seen when the radiation used is precisely that required to excite the electrons to the ionization threshold; thus, in theory, the resolution is limited only by the bandwidth of the radiation used in ionization. Dye lasers are used as the source of radiation in the ZEKE experiment, as these can have a very low bandwidth (of the order of 0.05 $cm^{-1}$). This allows the prospect of resolving rotational states of larger molecular cations. A graphic illustration of the greater resolving power of ZEKE spectroscopy can be seen in **Figure 4** in which the conventional photoelectron spectrum of $I_2$ is compared with a ZEKE spectrum.

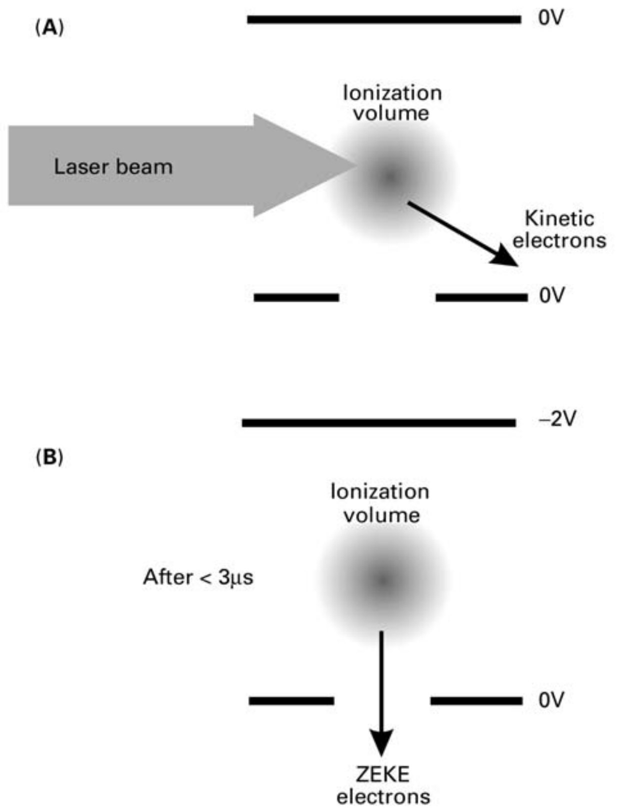



**Figure 3** (A) A pulsed laser beam ionizes the sample, and kinetic electrons are scattered randomly. (B) A delayed electric-field pulse is used to extract ZEKE electrons from the ionization volume.

The photon energy from a dye laser, even after frequency-doubling, is normally insufficient to ionize a molecule; hence the experiment involves a multiphoton process, normally via a resonant intermediate state. An advantage from this resonant intermediate state selection is, for instance, that different vibrational levels of the intermediate state can be selected, thus allowing the Franck–Condon factors for ionization to change. In order to find the intermediate states a resonance-enhanced multiphoton ionization (REMPI) spectrum must first be recorded. This is usually done by incorporating a time-of-flight mass spectrometry apparatus into the ZEKE photoelectron spectroscopy apparatus. For the extraction of the heavier ions a high-voltage pulse (~1 kV) is used and the ion signal is detected by multichannel plates. The REMPI signal can be mass-selected, to ensure that it represents an excited state of the appropriate species; this property is very important when studying van der Waals clusters in molecular beams. The ionization scheme for a two-colour $1 + 1'$ REMPI spectrum is shown in **Figure 5**.

With the exception of electron photodetachment from anions, it was soon realized that the ZEKE signal detected in a typical experiment does not actually arise from direct photoionization. With the application of an electric field to extract the ZEKE electrons there is a lowering of the ionization energy by the Stark effect, resulting in the ionization of high-$n$ Rydberg states ($n > 100$) which are said to be within a 'magic region' lying about 5 to 10 cm$^{-1}$ below each ion threshold, as indicated in **Figure 6**. As it turned out, the pulsed-field ionization of the Rydberg states sometimes leading to the name ZEKE-PFI, is in fact preferable to the 'true' ZEKE technique, since the neutral Rydberg states are less susceptible to stray electric fields in the apparatus than the 'true' ZEKE electrons.

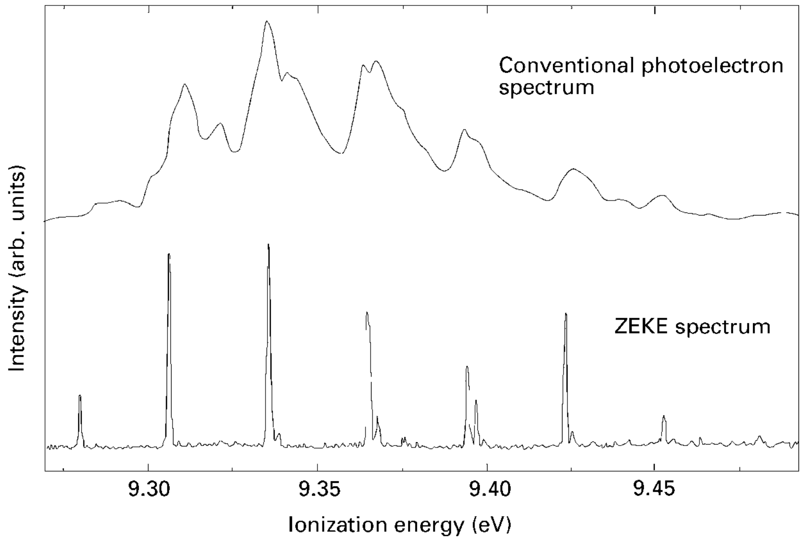



**Figure 4** The improvement in resolution gained by using ZEKE spectroscopy.

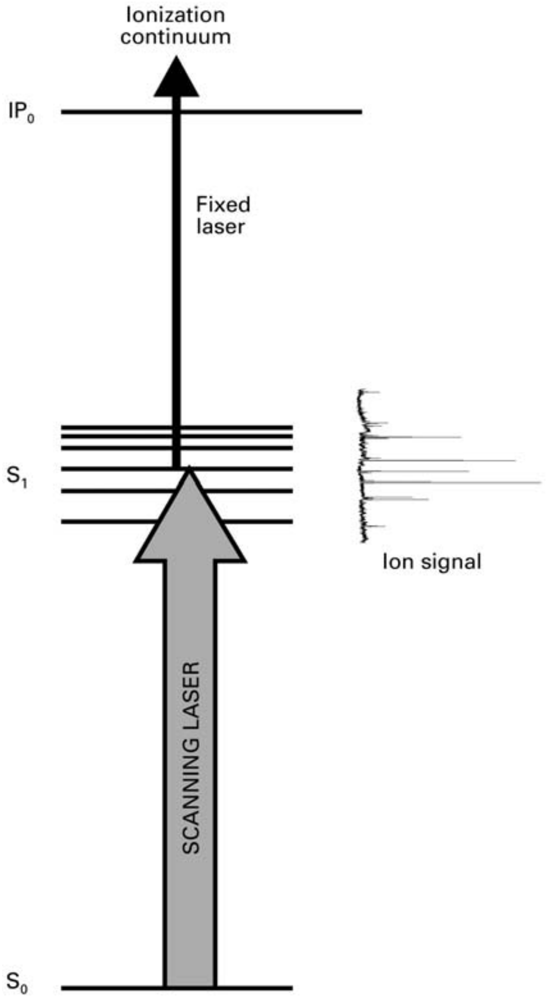



**Figure 5** An ionization scheme for a 1 + 1′ REMPI experiment.

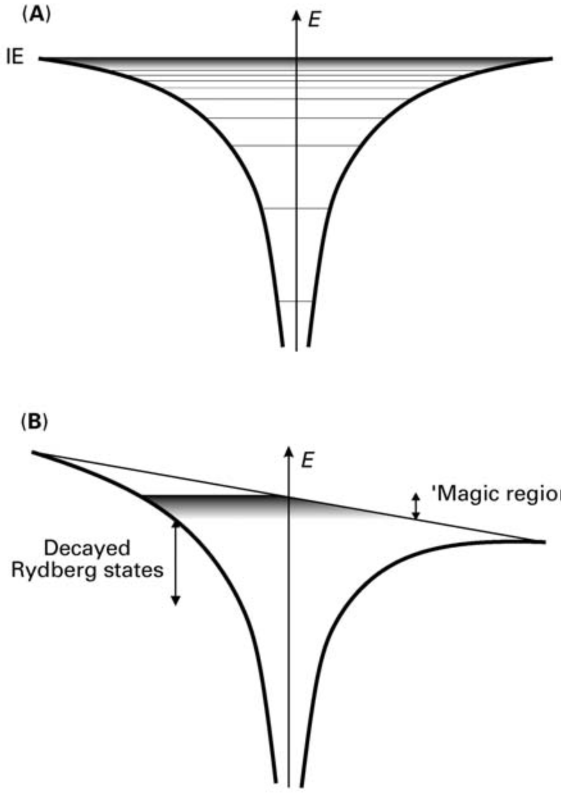



**Figure 6** (A) the Rydberg states converge on the ionization energy. (B) the long-lived states in the 'magic' region are field-ionized by the extraction pulse.

The ionization scheme for ZEKE is given in **Figure 7**. In the case of ZEKE photodetachment, the only mechanism by which one can obtain a signal is the detection of free electrons with zero kinetic energy.

A great deal of research effort has been expended in trying to understand the nature of the Rydberg states in the 'magic region'. The origin of the exceptionally long lifetime of the ZEKE Rydberg states has been a matter of some considerable debate and although the discussion is continuously evolving, a degree of consensus has been reached concerning the principal contributory effects. The extended lifetime is attributed to a combination of the effects of small homogeneous fields (which typically originate from electronic equipment in the laboratory) and inhomogeneous electric fields (associated with regions of localized charge in the spectrometer, e.g. ions). The highly diffuse nature of high-lying Rydberg states renders them susceptible to $l$ and $m_l$ mixing through external perturbations. Stray DC electric fields may cause substantial $l$ mixing through the Stark effect, while inhomogeneous fields inducing $m_l$ randomization arise from the presence of ions in low to medium concentrations. This slows down the rate of intramolecular relaxation considerably due to the conservation of angular momentum, as depicted in **Figure 8**. In the 'magic region' the Rydberg electron acquires a non-penetrating character and no longer interacts with the positive-ion core. Conversely, in the lower-lying Rydberg states the Rydberg electron still undergoes regular collisions with the core, which leads to intramolecular relaxation processes such as predissociation into neutral fragments and autoionization. This decay of the lower Rydberg states during the typical delay times used in ZEKE experiments is the underlying reason why peak widths observed in ZEKE spectroscopy are limited, even when high field strengths are used.

If laser-limited resolution is required, it is necessary to design sophisticated field-ionization schemes. The current resolution benchmark is the rotationally resolved ZEKE spectrum of benzene. From this it was demonstrated that the benzene cation is planar and adequately described in the $D_{6h}$ molecular point group, despite being subject to Jahn–Teller

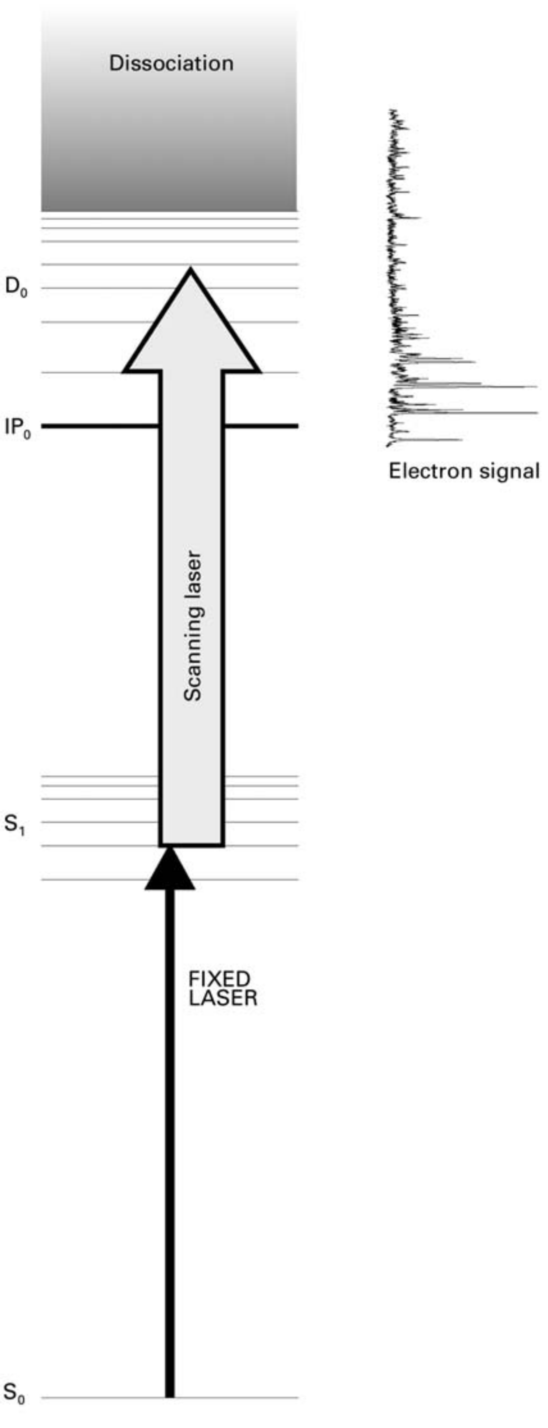



Figure 7 An ionization scheme for ZEKE spectroscopy.

distortion. ZEKE spectroscopy has also been successfully applied to studies of the vibrational structure of large organic molecules, molecular and metal clusters, and hydrogen-bonded systems. One of the more routine strengths of the technique is that ionization energies can be determined with an accuracy

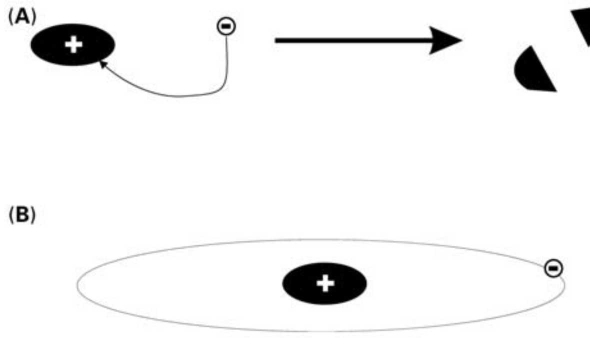



**Figure 8** (A) Collisions between the Rydberg electron and the core in lower Rydberg states cause decay by intramolecular processes such as predissociation. (B) The nonpenetrating character of higher Rydberg states results in a longer lifetime.

comparable to that of Rydberg extrapolations but with less experimental effort.

Variation of the slope of the pulse allows the spectral resolution to be adjusted according to the laser bandwidth limitations and to the demands of the system under study (whether one requires vibrational or rotational resolution). **Figure 9** shows schematically the effect of pulse-slope risetime on the time-of-flight (TOF) of the corresponding electrons produced by PFI. A fast pulse generates all the signal within a narrow TOF distribution, whereas a slow pulse spreads the different 'slices' of Rydberg states into a broader TOF distribution. Thus, for a particular TOF gate a smaller spectral 'Rydberg slice' will be collected when the photon energy of the light source is scanned for a slow rather than for a fast risetime pulse.

There are many ways to improve the resolution by varying the pulse sequence. The use of an extraction pulse causes the $2l + 1$ degeneracy of a given state to be lifted, due to the Stark effect. Those with a negative $m_l$ are raised in energy, and those with a positive $m_l$ are lowered in energy; this has the effect of broadening the signal about the field-free ionization energy. However, it has been shown that under field ionization the Stark states shifted to the blue (with negative $m_l$ values) have a longer lifetime than those shifted to the red. Hence the signal seen arises principally from states which have been red shifted. This can be accounted for using a multi-step staircase-like extraction pulse, to determine exactly the ionization energy under field-free conditions. It has also been observed that if a double extraction pulse is used, the second being of opposite polarity to the first, the Stark states which were blue shifted, and therefore were not ionized by the first pulse, kept the same orientation, and were red shifted with respect to the second pulse. The consequence of this fractional Stark-state selection is that a narrower slice of Stark

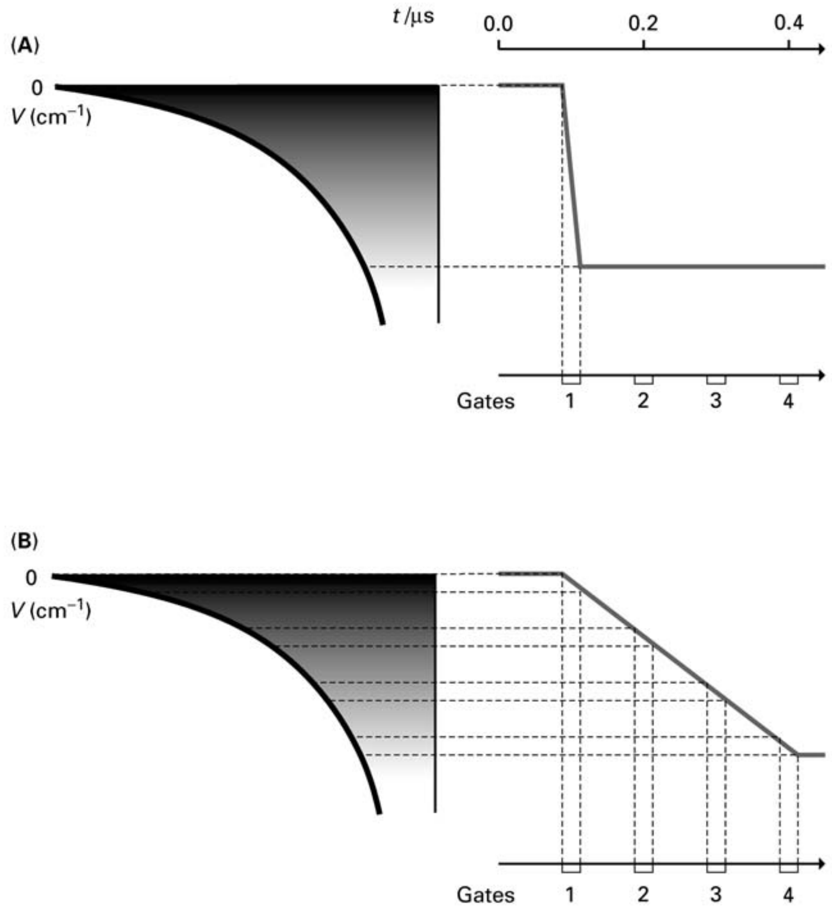



**Figure 9** The effect of slope risetime on the resolution: in (A) the fast pulse gives a low resolution as a large slice of the Rydberg manifold is ionized; in (B) the slower pulse enables the detection of much smaller slices of the manifold, giving a higher resolution at the cost of signal strength.

states is selected, and a higher-resolution signal is obtained.

A typical experimental setup for a ZEKE experiment is shown schematically in **Figure 10**. It consists of a laser system and a vacuum apparatus which includes the molecular beam source, the extraction plates and a μ-metal-shielded flight tube with electron/ion detectors (dual multichannel plates) at each end. In a typical two-colour experiment, both dye lasers (often frequency-doubled) are pumped simultaneously by an excimer laser or a Nd:YAG laser. The first dye laser excites a specific vibronic or rovibronic level of the intermediate state and the second laser ionizes the molecules or promotes them into long-lived Rydberg states ($n > 150$) converging to (ro)vibronic levels of the electronic ground state or an electronically excited state of the cation. After a delay time of a few μs, an extraction pulse is applied by either a simple electric pulsing device or by an arbitrary-function generator. The electrons are detected at the multichannel plates and their time of flight is recorded with boxcar integrators or a transient digitizer by setting narrow time gates (10–30 ns).

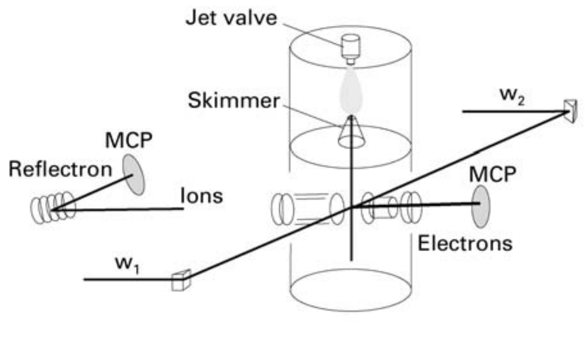



**Figure 10** A schematic representation of the ZEKE apparatus.

## Intensities in ZEKE spectra

The selection rules for ZEKE transitions are governed by the usual principle that the transition-moment integral must be totally symmetric to give

a non-zero intensity. For ZEKE transitions these can be affected by coupling to other Rydberg states. For rotational transitions one has to consider the possibility of transfer of angular momentum to the Rydberg electron; hence the selection rule for the quantum number representing the total angular momentum excluding spin is no longer $\Delta N = \pm 1, 0$. One model to account for the intensities of the rotational transition is the spectator model. In the spectator model it assumed that the Rydberg electron wave function is atomic hydrogen-like, with angular momentum $l_0$ and has no interaction with the core; hence the selection rules for $N''$, the total angular momentum of the Rydberg state, and $N^+$, the total angular momentum of the ion, are bound by the triangle condition $|N'' - N^+| \leq l_0 \leq |N'' + N^+|$. The intensities of ionizing transitions in the spectator model are dependent only on the transitions into the Rydberg state; a more complicated model is the compound model, where it is not assumed that the Rydberg state is initially fully decoupled from the core; hence further transfer of angular momentum can occur before the extraction pulse.

A further factor which affects the intensity of transitions is the role of autoionization from the Rydberg states. When a given Rydberg state is near the ionization threshold of another, lower-energy Rydberg series, the state has a shortened lifetime with respect to autoionization; as a consequence the intensity of higher-energy Rydberg series is usually depleted. In rotationally resolved spectra this is observed as a propensity for negative changes in angular momentum.

## Mass-selected ZEKE spectroscopy

An extension of ZEKE spectroscopy is mass-analysed threshold ionization (MATI), ‘photoelectron spectroscopy without photoelectrons’. This is effectively the same experiment; for every ZEKE electron produced, there must be a cation, and in MATI detection a signal is recorded from these ions. It is much harder to separate the ions produced from pulsed-field ionization of the ZEKE Rydberg states from the ever-present directly produced ions. Ions are much heavier than electrons and hence move more slowly, so a higher-voltage extraction pulse is required for the separation and the subsequent extraction and selection of the cations. The obvious advantage of this combination of ZEKE with mass spectrometry is the ability to select the cations on the basis of their mass.

The MATI signal also allows the study of fragmentation processes. It is interesting that at levels of ion internal energy at which a complex dissociates, the ZEKE spectrum can still be observed. When only looking at the ZEKE signal, it is not obvious that such a fragmentation has occurred; however, by looking at the MATI signal, fragmentation can be observed, as the spectrum switches from the parent cation mass channel to a fragment mass channel. This gives a direct measure of the dissociation energy of the cation. Predissociation can also be observed directly by this technique, and the dissociation products observed, which is useful for obtaining a more complete rovibronic structure of molecules.

Another development in photoionization spectroscopy is the technique called photoinduced Rydberg ionization (PIRI). In this, the neutral molecule absorbs radiation to produce a high-$n$ Rydberg state, as well as prompt ions. The prompt ions are separated using a delayed electric pulse, and the remaining Rydberg states, rather than being field-ionized, as in MATI, are photoexcited to form core-excited Rydberg states, which autoionize, and can be separated from the remaining Rydberg states. This technique effectively gives the absorption spectrum of the cation; also the problems arising from Stark shifts are no longer relevant as the molecules are not field-ionized.

## List of symbols

$E$ = energy; $h$ = Planck’s constant; $l$ = orbital angular momentum of an electron with projection $m$ on the laboratory $Z$-axis; $n$ = principle quantum number; $N$ = angular momentum; $\nu$ = frequency.

*See also:* **Photoelectron Spectrometers; Photoelectron Spectroscopy; Zero Kinetic Energy Photoelectron Spectroscopy, Applications.**

## Further reading

Dietrich H-J, Müller-Dethlefs K and Baranov LY (1996) Fractional Stark state selective electric field ionization of very high-*n* Rydberg states in molecules. *Physical Review Letters* **76**: 3530–3533.

Fischer I, Lindner R and Müller-Dethlefs K (1994) State to state photoionization dynamics probed by zero kinetic energy photoelectron spectroscopy. *Journal of the Chemical Society, Faraday Transactions* **90**: 2425–2442.

Haines SR, Geppert WD, Chapman DM *et al* (1998) Evidence for a strong intermolecular bond in the phenol $N_2$ cation. *Journal of Chemical Physics* **109**: 9244–9251.

Müller-Dethlefs K and Schlag EW (1991) High resolution zero kinetic energy (ZEKE) photo electron spectroscopy of molecular systems. *Annual Review of Physical Chemistry* **42**: 109–136.

Müller-Dethlefs K (1995) High resolution spectroscopy with photoelectrons: ZEKE spectroscopy of molecular systems. In: Powis I, Baer T and Ng C-Y (eds) *High*

*Resolution Laser Photoionization and Photoelectron Studies*, Chapter 2, pp 22–78. Chichester: Wiley.

Müller-Dethlefs K (1995) Applications of ZEKE spectroscopy. *Journal of Electron Spectroscopy and related Phenomena* 75: 35–46.

Müller-Dethlefs K, Schlag EW, Grant ER, Wang K and McKoy BV (1995) ZEKE spectroscopy: High resolution spectroscopy with photoelectrons. *Advances in Chemical Physics* 90: 1–104.

Müller-Dethlefs K (1991) Zero kinetic energy electron spectroscopy of molecules – rotational symmetry selection rules and intensities. *Journal of Chemical Physics* **95**: 4821–4839.

Reiser G and Müller-Dethlefs K (1992) Rotationally resolved zero kinetic energy photoelectron spectroscopy of nitric oxide. *Journal of Physical Chemistry* **96**: 9–12.

Wright TG, Reiser G and Müller-Dethlefs K (1994) Good vibrations. *Chemistry in Britain* 30: 128–132.

# Zinc NMR, Applications

*See* **Heteronuclear NMR Applications (Sc–Zn).**

# Zirconium NMR, Applications

*See* **Heteronuclear NMR Applications (Y–Cd).**

# 1. Periodic Table of Elements

| IA | IIA | | IVA | VA | VIA | VIIA | | VIII | | IB | IIB | IIIB | IVB | V | VIB | VIIB | O |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H 1.008 | | | | | | | | | | | | | | | | | -le 003 |
| Li ;.941 | 4 Be 9.012 | | | | | | | | | | | .81 | 6 C 12.01 | 7 N 14 | 8 O 16.00 | 9 F 19.00 | Ne 0.18 |
| Na 22.99 | 12 Mg 24.31 | | | | | | | | | | | .98 | 14 Si 28.09 | 15 P 30 | 16 S 32.06 | 17 Cl 35.45 | Ar 9.95 |
| K 39.10 | 20 Ca 40.08 | 2' 4 | 22 Ti 47.90 | 23 V 50.9 | 4 Cr 52.00 | 25 Mn 54.94 | 5 | 27 Co 58.93 | 28 Ni 58.: | 29 Cu 63.55 | 30 Zn 65.37 | ia .72 | 32 Ge 72.60 | 33 A 74 | 34 Se 78.96 | 35 Br 79.90 | Kr 3.80 |
| Rb 85.47 | 38 Sr 87.62 | 3! | 40 Zr 91.22 | 41 Nb 92.9 | 2 Mo 95.94 | 43 Tc (99) | 1 | 45 Rh 102.9 | 46 Pd 106. | 47 Ag 107.9 | 48 Cd 112.4 | 4.8 | 50 Sn 118.7 | 51 S 12 | 52 Te 127.6 | 53 I 126.9 | Xe 31.3 |
| Cs 132.9 | 56 Ba 137.3 | 5 | 72 Hf 178.5 | 73 Ta 180.: | 4 W 183.9 | 75 Re 186.2 | s 2 | 77 Ir 192.2 | 78 Pt 195 | 79 Au 197.0 | 80 Hg 200.6 | 4.4 | 82 Pb 207.2 | 83 B 20 | 84 Po (210) | 85 At (210) | Rn !22) |
| Fr (223) | 88 Ra (226) | 8 ( | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lanthes | 58 Ce 140. | 9 Pr 140.9 | 60 Nd 144.2 | 6 ) | 62 Sm 150.4 | 63 Eu 152. | 64 Gd 157.3 | 65 Tb 158.9 | .5 | 67 Ho 164.9 | 68 E 16 | 69 Tm 168.9 | 70 Yb 173.0 | .u 5.0 |
| Actis | 90 Th 232. | 1 Pa (231) | 92 U 238.0 | ) | 94 Pu (242) | 95 Am (24 | 96 Cm (248) | 97 Bk (247) | 1) | 99 Es (254) | 100 F (25 | 101 Md (256) | 102 No (254) | .r 257) |

# 2. Tables of SI and related units

## SI base units

| Quantity | Name | Symbol |
|---|---|---|
| length | metre | m |
| time | second | s |
| electric current | ampere | A |
| thermodynamic temperature | kelvin | K |
| amount of substance | mole | mol |
| luminous intensity | candela | cd |

| Name | Symbol | Value in SI units |
|---|---|---|
| minute | min | 1 min = 60 s |
| hour | h | 1 h = 60 min = 3600 s |
| day | d | 1 d = 24 h = 86 400 s |
| degree<br>second | °<br>" | 1° = (π/180) rad<br>1" = (1/60)' = (π/648 000) rad |
| litre | L | 1 L = 1 dm$^3$ = 10$^{-3}$ m$^3$ |
| tonne | t | 1 t = 10$^3$ kg |

## cgs units with special names

| | | |
|---|---|---|
| erg | erg | 1 erg = 10$^{-7}$ J |
| dyne | dyn | 1 dyn = 10$^{-5}$ N |
| poise | P | 1 P = 1 dyn s cm$^{-2}$ = 0.1 Pa s |
| stokes | St | 1 St = 1 cm$^2$ s$^{-1}$ = 10$^{-4}$ m$^2$ s$^{-1}$ |
| gauss | Gs, G | 1 Gs corresponds to 10$^{-4}$ T |
| oersted<br>stilb | Oe<br>sb | 1 Oe corresponds to (1000/4 π) A m$^{-1}$<br>1 sb = 1 cd cm$^{-2}$ = 10$^4$ cd m$^{-2}$ |
| phot | ph | 1 ph = 10$^4$ lx |

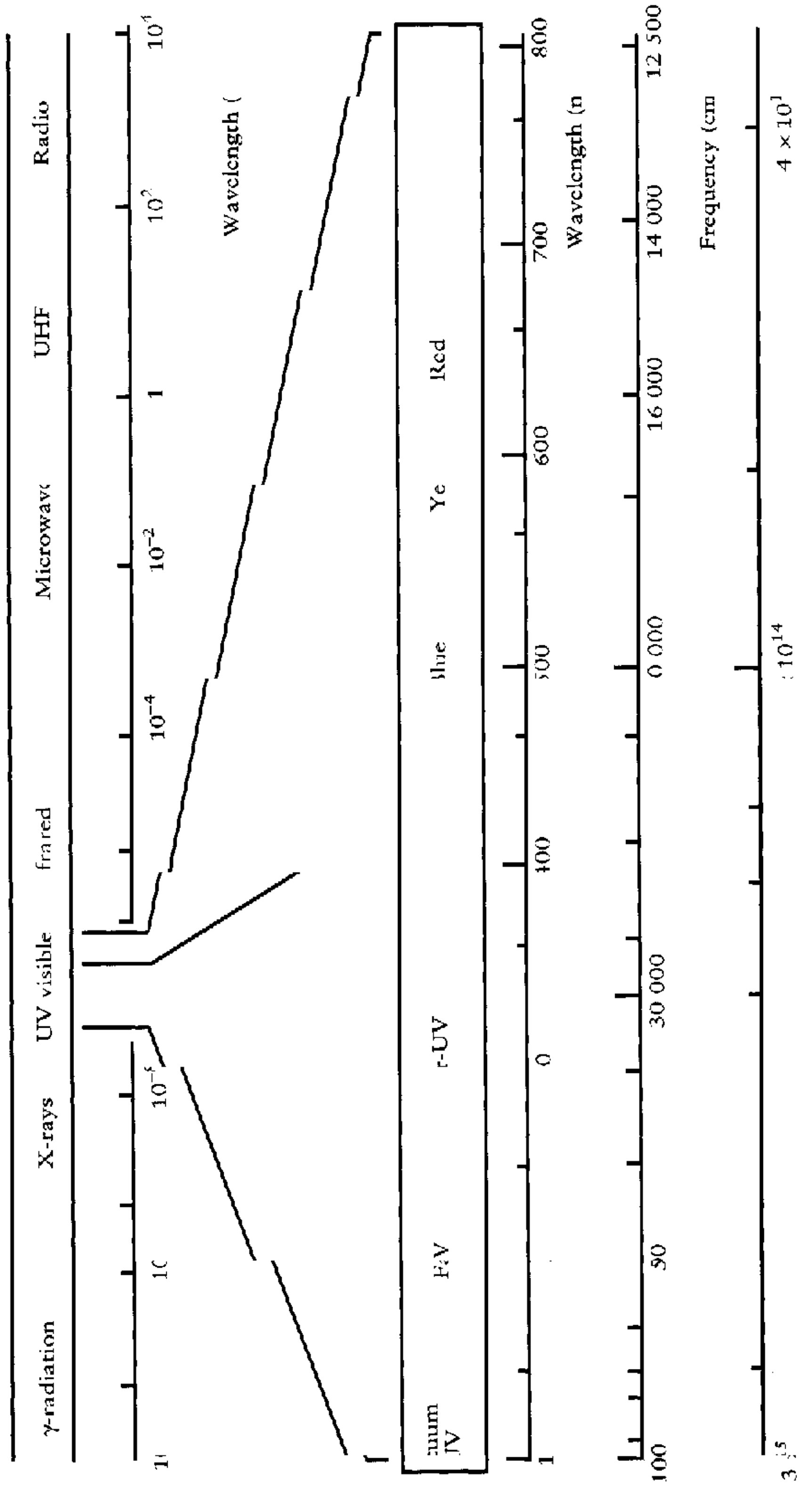
## SI prefixes

| Factor | Prefix | Symbol | Origin |
|---|---|---|---|
| 10$^{15}$ | peta | P | 'πεντε', Greek 'five' (10$^3$ to fifth power is 10$^{15}$) |
| 10$^{12}$ | tera | T | 'τερασ', Greek 'monster' |
| 10$^9$ | giga | G | 'γιγασ', Greek 'giant' |
| 10$^6$ | mega | M | 'μεγασ', Greek 'large' |
| 10$^3$ | kilo | k | 'χιλιοι', Greek 'thousand' |
| 10$^2$ | hecto | h | 'εκατον', Greek 'hundred' |
| 10$^{-1}$ | deci | d | 'decima pars', Latin 'tenth' |
| 10$^{-2}$ | centi | c | 'pars centesima', Latin 'hundredth' |
| 10$^{-3}$ | milli | m | 'pars millesima', Latin 'thousandth' |
| 10$^{-6}$ | micro | μ | 'μικροσ', Greek 'little' |
| 10$^{-9}$ | nano | n | 'nanus', Latin 'dwarf' |
| 10$^{-15}$ | femto | f | 'femten', Danish, Norwegian 'fifteen' |
| 10$^{-18}$ | atto | a | 'atten', Danish, Norwegian, 'eighteen' |

## SI derived units

| Quantity | SI unit | | | |
|---|---|---|---|---|
| | Name | Symbol | Expression in terms of other units | Expression in terms of SI base units |
| Frequency | hertz | Hz | | |
| **Force** | newton | N | | m kg s$^{-2}$ |
| **Pressure, stress** | pascal | Pa | N m$^{-2}$ | m$^{-1}$ kg s$^{-2}$ |
| **Energy, work, quantity of heat** | joule | J | N m | m$^{2}$ kg s$^{-2}$ |
| **Power, radiant flux** | watt | W | J s$^{-1}$ | m$^{2}$ kg s$^{-3}$ |
| **Quantity of electricity, electric charge** | coulomb | C | [illegible] | [illegible] |
| **Capacitance** | farad | F | C V$^{-1}$ | m$^{-2}$ kg$^{-1}$ s$^{4}$ A$^{2}$ |
| **Electric resistance** | ohm | Ω | V A$^{-1}$ | m$^{2}$ kg s$^{-3}$ A$^{-2}$ |
| **Conductance** | siemens | S | A V$^{-1}$ | m$^{2}$ kg$^{-1}$ s$^{3}$ A$^{2}$ |
| **Magnetic flux** | weber | Wb | V s | m$^{2}$ kg s$^{-2}$ A$^{-1}$ |
| **Magn. flux density, magn. induct., magn. field** | tesla | T | Wb m$^{-2}$ | kg s$^{-2}$ A$^{-1}$ |
| **Luminous flux** | lumen | lm | | cd sr |
| **Illuminance** | lux | lx | lm m$^{-2}$ | m$^{-2}$ cd sr |
| **Dynamic viscosity** | pascal second | Pa s | | m$^{-1}$ kg s$^{-1}$ |
| **Moment of force** | metre newton | N m | | m$^{2}$ kg s$^{-2}$ |
| **Surface tension** | newton per metre | N m$^{-1}$ | | kg s$^{-2}$ |
| **Heat capacity, entropy** | joule per kelvin | J K$^{-1}$ | | m$^{2}$ kg s$^{-2}$ K$^{-1}$ |
| **Specific heat capacity, specific entropy** | joule per kilogram kelvin | J kg$^{-1}$ K$^{-1}$ | | m$^{2}$ s$^{-2}$ K$^{-1}$ |
| **Specific energy** | joule per kilogram | J kg$^{-1}$ | | m$^{2}$ s$^{-2}$ |
| **Thermal conductivity** | watt per metre kelvin | W m$^{-1}$ K$^{-1}$ | | m kg s$^{-3}$ K$^{-1}$ |
| **Electric field strength** | volt per metre | V m$^{-1}$ | | m kg s$^{-3}$ A$^{-1}$ |
| **Electric charge density** | coulomb per cubic metre | C m$^{-3}$ | | m$^{-3}$ s A |
| **Electric displacement, electric flux density** | coulomb per square metre | C m$^{-2}$ | | m$^{-2}$ s A |
| **Permeability** | henry per metre | H m$^{-1}$ | | m kg s$^{-2}$ A$^{-2}$ |
| **Molar energy** | joule per mole | J mol$^{-1}$ | | m$^{2}$ kg s$^{-2}$ mol$^{-1}$ |
| **Molar entropy, molar heat capacity** | joule per mole kelvin | J mol$^{-1}$ K$^{-1}$ | | m$^{2}$ kg s$^{-2}$ K$^{-1}$ mol$^{-1}$ |

# 3. Wavelength scale




Tange of electromag radiation. The lopart is an enlargeuof the UV-visible m.

# 4. Colour, wave length, frequency, wave number and energy of light

| Colour | $\lambda$/nm | $\nu$/Hz | $\tilde{\nu}$/cm$^{-1}$ | E/eV | E/kJ mol$^{-1}$ |
|---|---|---|---|---|---|
| Infrared | 1000 | $3.00 \cdot 10^{14}$ | $1.00 \cdot 10^{4}$ | 1.24 | 120 |
| Red | 700 | $4.28 \cdot 10^{14}$ | $1.43 \cdot 10^{4}$ | 1.77 | 171 |
| Yellow | 580 | $5.17 \cdot 10^{14}$ | $1.72 \cdot 10^{4}$ | 2.14 | 206 |
| Green | 530 | $5.66 \cdot 10^{14}$ | $1.89 \cdot 10^{4}$ | 2.34 | 226 |
| Blue | 470 | $6.38 \cdot 10^{14}$ | $2.13 \cdot 10^{4}$ | 2.64 | 254 |
| Violet | 420 | $7.14 \cdot 10^{14}$ | $2.38 \cdot 10^{4}$ | 2.95 | 285 |
| Near Ultraviolet | 300 | $1.00 \cdot 10^{15}$ | $3.33 \cdot 10^{4}$ | 4.15 | 400 |

# 5. Magnetic susceptibilities at 25°C

| | $\kappa$ | $\chi$ |
|---|---|---|
| Water | $-9.0 \cdot 10^{-5}$ | $-9.0 \cdot 10^{-9}$ |
| Cyclohexane | $-7.9 \cdot 10^{-6}$ | $-1.02 \cdot 10^{-8}$ |
| Carbon tetrachloride | $-8.9 \cdot 10^{-6}$ | $-5.4 \cdot 10^{-9}$ |
| NaCl (s) | $-1.39 \cdot 10^{-5}$ | $-6.4 \cdot 10^{-9}$ |
| Cu (s) | $-9.6 \cdot 10^{-5}$ | $-1.07 \cdot 10^{-9}$ |
| S (s) | $-1.29 \cdot 10^{-5}$ | $-6.2 \cdot 10^{-9}$ |
| $CuSO_4 \cdot 5H_2O$ (s) | $1.76 \cdot 10^{-4}$ | $7.7 \cdot 10^{-8}$ |
| $MnSO_4 \cdot 4H_2O$ (s) | $2.64 \cdot 10^{-3}$ | $8.12 \cdot 10^{-7}$ |
| $NiSO_4 \cdot 7H_2O$ (s) | $4.16 \cdot 10^{-4}$ | $2.01 \cdot 10^{-7}$ |
| $FeSO_4(NH_4)_2SO_4 \cdot 6H_2O$ (s) | $7.55 \cdot 10^{-4}$ | $4.06 \cdot 10^{-7}$ |
| Al (s) | $2.2 \cdot 10^{-5}$ | $8.2 \cdot 10^{-9}$ |
| Na (s) | $7.3 \cdot 10^{-6}$ | $7.5 \cdot 10^{-9}$ |
| K (s) | $5.6 \cdot 10^{-6}$ | $6.5 \cdot 10^{-9}$ |

The CGS values of the susceptibility (per gram) are obtained by forming $1000\,\chi / 4\pi$, and the value per gram mole from $1000\,\chi\,M_r / 4\pi$.

| Z | Element | K | L | | M | | | N | | | | O | | | | P | | | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1s | 2s | 2p | 3s | 3p | 3d | 4s | 4p | 4d | 4f | 5s | 5p | 5d | 5f | 6s | 6p | 6d | 7s |
| | | 1st period | | | | | | | | | | | | | | | | | |
| 1 | H | 1 | | | | | | | | | | | | | | | | | |
| [illegible] | [illegible] | [illegible] | | | | | | | | | | | | | | | | | |
| 3 | Li | 2 | 1 | | | | | | | | | | | | | | | | |
| 4 | Be | 2 | 2 | | | | | | | | | | | | | | | | |
| 5 | B | 2 | 2 | 1 | | | | | | | | | | | | | | | |
| 6 | C | 2 | 2 | 2 | | | | | | | | | | | | | | | |
| 7 | N | 2 | 2 | 3 | | | | | | | | | | | | | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | | | | | | | | | | | | | | |
| 10 | Ne | 2 | 2 | 6 | | | | | | | | | | | | | | | |
| | | 3rd period | | | | | | | | | | | | | | | | | |
| 11 | Na | 2 | 2 | 6 | 1 | | | | | | | | | | | | | | |
| 12 | Mg | 2 | 2 | 6 | 2 | | | | | | | | | | | | | | |
| 13 | Al | 2 | 2 | 6 | 2 | 1 | | | | | | | | | | | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | | | | | | | | | | | | |
| 16 | S | 2 | 2 | 6 | 2 | 4 | | | | | | | | | | | | | |
| 17 | Cl | 2 | 2 | 6 | 2 | 5 | | | | | | | | | | | | | |
| 18 | Ar | 2 | 2 | 6 | 2 | 6 | | | | | | | | | | | | | |
| | | 4th period | | | | | | | | | | | | | | | | | |
| 19 | K | 2 | 2 | 6 | 2 | 6 | | 1 | | | | | | | | | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | | | | | | | | | | |
| 22 | Ti | 2 | 2 | 6 | 2 | 6 | 2 | 2 | | | | | | | | | | | |
| 23 | V | 2 | 2 | 6 | 2 | 6 | 3 | 2 | | | | | | | | | | | |
| 24 | Cr | 2 | 2 | 6 | 2 | 6 | 5 | 1 | | | | | | | | | | | |
| 25 | Mn | 2 | 2 | 6 | 2 | 6 | 5 | 2 | | | | | | | | | | | |
| 26 | Fe | 2 | 2 | 6 | 2 | 6 | 6 | 2 | | | | | | | | | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | | | | | | | | | | |
| 29 | Cu | 2 | 2 | 6 | 2 | 6 | 10 | 1 | | | | | | | | | | | |
| 30 | Zn | 2 | 2 | 6 | 2 | 6 | 10 | 2 | | | | | | | | | | | |
| 31 | Ga | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 1 | | | | | | | | | | |
| 32 | Ge | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 2 | | | | | | | | | | |
| 33 | As | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 3 | | | | | | | | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | | | | | | | | | |
| 36 | Kr | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | | | | | | | | | | |
| | | 5th period | | | | | | | | | | | | | | | | | |
| 37 | Rb | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | | | 1 | | | | | | | |
| 38 | Sr | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | | | 2 | | | | | | | |
| 39 | Y | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 1 | | 2 | | | | | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | [illegible] | | | | | | | |
| 42 | Mo | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 5 | | 1 | | | | | | | |
| 43 | Tc | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 5 | | 2 | | | | | | | |
| 44 | Ru | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 7 | | 1 | | | | | | | |
| 45 | Rh | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 8 | | 1 | | | | | | | |
| 46 | Pd | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | | | | | | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | [illegible] | | | | | | | |
| 49 | In | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 1 | | | | | | |
| 50 | Sn | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 2 | | | | | | |

[illegible]

| Z | Element | K | L | | M | | | N | | | | O | | | | P | | | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1s | 2s | 2p | 3s | 3p | 3d | 4s | 4p | 4d | 4f | 5s | 5p | 5d | 5f | 6s | 6p | 6d | 7s |
| 51 | Sb | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 3 | | | | | | |
| 52 | Te | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 4 | | | | | | |
| 54 | Xe | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 6 | | | | | | |
| | | 6th period | | | | | | | | | | | | | | | | | |
| 55 | Cs | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 6 | | | 1 | | | |
| 56 | Ba | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 6 | | | 2 | | | |
| 57 | La | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | | 2 | 6 | 1 | | 2 | | | |
| 58 | Ce | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 1 | 2 | 6 | 1 | | 2 | | | |
| 60 | Nd | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 4 | 2 | 6 | | | 2 | | | |
| 61 | Pm | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 5 | 2 | 6 | | | 2 | | | |
| 62 | Sm | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 6 | 2 | 6 | | | 2 | | | |
| 63 | Eu | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 7 | 2 | 6 | | | 2 | | | |
| 64 | Gd | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 7 | 2 | 6 | 1 | | 2 | | | |
| 65 | Tb | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 8 | 3 | 6 | 1 | | 2 | | | |
| 67 | Ho | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 10 | 2 | 6 | 1 | | 2 | | | |
| 68 | Er | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 11 | 2 | 6 | 1 | | 2 | | | |
| 69 | Tm | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 13 | 2 | 6 | | | 2 | | | |
| 70 | Yb | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | | | 2 | | | |
| 71 | Lu | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 1 | | 2 | | | |
| 72 | Hf | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 2 | | 2 | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | [illegible] | | | |
| 75 | Re | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 3 | 6 | 5 | | 2 | | | |
| 76 | Os | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 6 | | 2 | | | |
| 77 | Ir | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 7 | | 2 | | | |
| 78 | Pt | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 9 | | 1 | | | |
| 79 | Au | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 1 | | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | [illegible] | [illegible] | | |
| 82 | Pb | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 2 | | |
| 83 | Bi | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 3 | | |
| 84 | Po | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 4 | | |
| 85 | At | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 5 | | |
| 86 | Rn | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 6 | | |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | [illegible] | [illegible] | | [illegible] |
| 88 | Ra | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 6 | | 2 |
| 89 | Ac | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 6 | 1 | 2 |
| 90 | Th | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | | 2 | 6 | 2 | 2 |
| 91 | Pa | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 2 | 2 | 6 | 1 | 2 |
| 92 | U | 3 | 3 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 3 | 2 | 6 | 1 | 2 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 95 | Am | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 7 | 2 | 6 | | 2 |
| 96 | Cm | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 7 | 2 | 6 | 1 | 2 |
| 97 | Bk | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 8 | 2 | 6 | 1 | 2 |
| 98 | Cf | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 9 | 2 | 6 | 1 | 2 |
| 99 | Es | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 10 | 2 | 6 | 1 | 2 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 102 | No | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 13 | 2 | 6 | 1 | 2 |
| 103 | Lr | 2 | 2 | 6 | 2 | 6 | 10 | 2 | 6 | 10 | 14 | 2 | 6 | 10 | 14 | 2 | 6 | 1 | 2 |

| Solv | Abbreviation | Formula | Molecular weight | F.p. | B.p. (°C) | Den | Dipole Mom.** | Sustib. ×1( | Dielectr. Const. | Rct. In | Chem Shift δ 1H (ppm) | 13C (ppm) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Acecld | | $C_2H_4C$ | $60.0_5$ | 1 | 117.9 | 1.04 | 1.2 | 0.5! (3 | 6.1 | 13 | 2.1 | 21.1 177.3 |
| Ace | | $C_3H_6C$ | 58.1 | –9 | 56.3 | 0.79 | 2.88 | 0.4( | 20.7 | 17 | 2.2 | 30.2 205.1 |
| Acetrile | | $C_2H_3N$ | $41.0_5$ | –4 | 81.6 | 0.77 | 3.92 | 0.5: | 37.5 | 13 | 2.0 | 0.3 117.2 |
| Bere* | | $C_6H_6$ | 78.1 | | 80.1 | 0.87 | 0 | 0.6! (2 | 2.3 | 11 | 7.4 | 128.7 |
| t-Bulcohol | OH | $C_4H_{10}$ | 74.1 | 2 | 82.2 | 0.78 | 1.66 | 0.5: | 1.8 | 18 | 1.3 | 31.6 68.7 |
| Cardisulfide | | $CS_2$ | 76.1 | –11 | 46.2 | 1.27 | 0 | 0.5 | 2.6 | 13 | | 192.8 |
| Cartetrachloride* | | $CCl_4$ | 153.8 | –2 | 76.7 | 1.58 | 0 | 0.6 | 2.2 | 14 | | 96.7 |
| Chlorm* | | $CHCl$ | 119.4 | –6 | 61.1 | 1.48 | 1.01 | 0.7 | 4.8 | 13 | 7.3 | 77.7 |
| Cycxane | | $C_6H_{12}$ | 84.2 | | 80.7 | 0.77 | 0 | 0.6 | 2.0 | 15 | 1.4 | 27.8 |
| Cycntane | | $C_5H_{10}$ | 70.1 | –9 | 49.3 | 0.74 | | 0.6 | 2.0 | 15 | 1.5 | 26.5 |
| Deca | | $C_{10}H_1$ | 138.2 | –4 | 191.7 | 0.87 | | 0.6 | | 18 | 0.9 to 1.8 | 24 to 44 |
| Dibomethane | | $CH_2B$ | 173.8 | –5 | 97.0 | 1.54 | 1.43 | 0.9 | 7.5 | 2. | 5.0 | 21.6 |
| o-Dorobenzene* | CB | $C_6H_4C$ | 147.0 | –1 | 180.5 | 1.30 | 2.50 | 0.7 | 9.9 | 15 | 7.0 to 7.4 | 8 to 133 |
| Diether | | $C_4H_{10}$ | 74.1 | –1 | 34.5 | 0.71 | 1.15 | | 4.3 | 16 | 1.2 3.5 | 17.1 67.4 |
| 1,2-loroethane | | $C_2H_4C$ | 99.0 | –3 | 83.5 | 1.24 | 1.75 | | 10.4 | 11 | 3.7 | 51.7 |
| 1,2-loroethylene Z | | $C_2H_2C$ | 96.9 | –8 | 60.6 | 1.28 | 1.90 | 0.6 (1 | 9.2 | 10 | 6.4 | 119.3 |
| 1,2-loroethylene E | | $C_2H_2C$ | 96.9 | –4 | 47.7 | 1.25 | 0 | 0.6 (1 | 2.1 | 12 | 6.3 | 121.1 |
| 1,1-loroethylene | | $C_2H_2C$ | 96.9 | –1 | 31.6 | 1.21 | 1.34 | 0.6 (1 | 4.6 | 17 | 5.5 | 115.5 128.9 |
| Dicbmethane | | $CH_2C$ | 84.9 | –9 | 39.8 | 1.31 | 1.60 | 0.7 | 8.9 | 11 | 5.3 | 54.2 |
| Dimoxymethane | | $C_3H_8C$ | 76.1 | –1( | 42.3 | 0.86 | | 0.6 | 2.6 | 13 | 3.3 4.4 | 54.8 97.9 |
| Dimylacetamide | IA | $C_4H_9N$ | 87.1 | –2 | 166.1 | 0.93 | 3.8 | | 37.8 | 16 | 2.1 3 | 1/38 169.6 |
| Dimylcarbonate* | | $C_3H_6C$ | 90.1 | | 90.5 | 1.06 | | | | 17 | 3.65 | 54.8 156.9 |
| Dimylether | | $C_2H_6C$ | 46.1 | –1 | –24 | | | | | | | 59.4 |
| N,N Dimylformamide | IF | $C_3H_7N$ | 73.1 | –6 | 153.0 | 0.94 | 3.86 | | 36.7 | 12 | 2.9 8.0 | 31–36 161.7 |
| Dimylsulfoxide | ISO | $C_2H_6S$ | 78.1 | – | 189.0 | 1.09 | 3.96 | | 46.7 | 13 | 2.6 | 39.5 |
| Dio* | | $C_4H_8C$ | 88.1 | · | 101.3 | 1.02 | 0.45 | 0.6 (3 | 2.2 | 13 | 3.7 | 67.8 |

| Solve | Abbtation | Formula | )lecular ight | F.p. (°C) | B.p. (°C) | Density | le l.** | Susceptib. ×10⁶ | electr. nst. | Refr Inde | δ ¹H (ppm) | δ ¹³C (ppm) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | Cheal Shift | |
| Ethyl ate | | $C_4H_8O_2$ | 8.1 | −83.9 | 77.1 | 0.895 | | 0.554 | 5.0 | 1.36 | 1.2/2.0<br>4.1 | 14.3<br>60.1<br>70.4 |
| Ethylcarbonate | EC | $C_3H_4O_3$ | 8.1 | 36.4 | !38 | 1.321 | | | 3.6 | 1.42 | 4.4 | 65.0<br>55.8 |
| Freon | | $CF_2Cl_2$ | 0.9 | −158 | -29.8 | 1.18 | | 0.642 (−30°) | | | | |
| Freon | | $CHClF_2$ | 6.5 | −146 | -40.8 | 1.491 | | | | | | |
| Formde | F | $CH_3ON$ | 5.0 | 2.6 | !10.5 | 1.133 | | 0.551 | 9 | 1.44 | 7.2<br>8.1 | 65.1 |
| Hexacoacetone | | $C_3Cl_6O$ | 4.8 | −30 | !03 | 1.744 | | | | 1.51 | | 23.7<br>26.4 |
| Hexafodichloro-prop | | $C_3F_6Cl_2$ | 0.9 | −136 | 35 | 1.589 | | | | | | |
| Hexanyl-phosamide* | HMI | $C_6H_{18}N_3PO$ | 9.2 | 7.2 | !33 | 1.027 | | | 0.0 | 1.45 | 2.4<br>2.6 | 36.6 |
| Metha | MeC | $CH_4O$ | 2.0 | −97.7 | 64.7 | 0.787 | | 0.530 | 2.7 | 1.32 | 3.5 | 49.3 |
| Methyoride | | $CH_3Cl$ | 0.5 | −97.7 | -24.1 | 0.916 | | | 2.6 | 1.33 | | 25.1 |
| Morphe | | $C_4H_9NO$ | 7.1 | −3.1 | 28.9 | 1.005 | | 0.631 | 7.4 | 1.45 | 2.6<br>3.9 | 46.8<br>68.9 |
| Nitrorane | | $CH_3NO_2$ | 1.0 | −28.5 | 01.2 | 1.131 | | 0.391 (25°) | 5.9 | 1.37 | 4.3 | 57.3 |
| Pyridi | | $C_5H_5N$ | 9.1 | −41.6 | 15.3 | 0.978 | | 0.611 | 2.4 | 1.50 | 7.1 to 8.8 | 24 to 150 |
| Quinc | | $C_9H_7N$ | 9.2 | −14.9 | !37.1 | 1.098 | | 0.729 | 9.0 | 1.62 | 7 to 8.8 | 21 to 151 |
| Sulfuxide | | $SO_2$ | 4.1 | −72.7 | -10.0 | 1.434 | | | 7.6 | | | |
| 2,2-Tehloro-etha | | $C_2H_2Cl_4$ | 7.8 | −43.8 | 46.2 | 1.578 | | 0.856 | 3.2 | 1.48 | 6.0 | 74.0 |
| Tetracoethane | | $C_2Cl_4$ | 5.8 | −22.3 | 21.2 | 1.631 | | 0.802 (15°) | 2.3 | 1.50 | | 20.4 |
| Tetralofuran | THF | $C_4H_8O$ | 2.1 | −66 | 66.0 | 0.889 | | | 7.6 | 1.40 | 1.8<br>3.7 | 26.7<br>68.6 |
| Tetranylurea | TMU | $C_5H_{12}N_2O$ | 6.2 | −1.2 | 76.5 | 0.969 | | 0.634 (15°) | 3.1 | 1.44 | 2.8 | 38.5<br>65.6 |
| Toluel | | $C_7H_8$ | 2.1 | −94.9 | 10.6 | 0.867 | | 0.618 | 2.4 | 1.49 | 2.3<br>7.2 | 21.3<br>125 to 138 |
| Trichl:thylene* | | $C_2CHI_3$ | 1.4 | −73 | 87.2 | 1.476 | | 0.734 | 3.4 | 1.48 | 6.4 | 16.7<br>24.0 |
| Water | | $H_2O$ | 8.02 | 0 | 00 | 0.997 | | 0.719 | 3.5 | 1.33 | 4.8 | |

B.p.: bg point F.p.: freg point Refce of the shifts: ¹H ¹³C (TMS) * Carcinogens cncer suspect ager

** Dipnoment in Debye t (1D = $10^{-18}$e.s.u.≙ 3.3356 x $10^{-30}$ C

## 8. Important acronyms in organic chemistry

| | |
|---|---|
| **A** | adenine |
| **AA** | anisylacetone |
| **AAO** | acetaldehyde oxime |
| **Ac** | acetate |
| **ACAC (acac)** | acetylacetone |
| **ACTH** | adrenocorticotropic hormone (corticotropin) |
| **ADMA** | alkyldimethylamine |
| **ADP** | adenosine 5′-diphosphate |
| **Ala** | alanine |
| **AMP** | adenosine 5′-monophosphate |
| **AN** | acetonitrile |
| **APS** | adenosine 5′-phosphosulfate |
| **Ar** | aryl |
| **Arg** | arginine |
| **Asp** | aspartic acid |
| **ATA** | anthranilamide |
| **ATP** | adenosine 5′-triphosphate |
| **B** | nucleoside base (adenine, cytosine, guanine, thymine, or uracil) |
| [illegible] | benzo[a]pyrene |
| **BBP** | benzyl butyl phthalate |
| **BHC** | benzene hexachloride |
| **Bn** | benzyl (also Bz, BZL, or Bnz) |
| **BN** | benzonitrile |
| **BON** | β-oxynaphthoic acid |
| **BSA** | bovine serum albumin |
| **BTA** | benzoyltrifluoroacetone |
| *n* **Bu** | *n*-butyl |
| *t* **Bu** | *tert*-butyl |
| **Bz** | benzoyl |
| **CD** | cyclodextrin |
| **CDP** | cytidine 5′-disphosphate |
| **CE** | cyanoethyl |
| **CMP** | cytidine monophosphate |
| **CoA** | coenzyme A |
| **12-Crown-4** | 1,4,7,10-tetraoxacyclododecane |
| **CTA** | citraconic anhydride |
| **Cys** | cysteine |
| **DAA** | diacetone acrylamide |
| **DAP** | dodecylammonium propionate |
| **DCEE** | dichloroethyl ether |
| **DDD** | 2,2′-dihydroxy-6,6′-dinaphthyl disulfide |
| **DDH** | 1,3-dibromo-5,5-dimethylhydantoin |
| **DDM** | diphenyldiazomethane |
| **DEA** | *N,N*-diethylaniline |
| **DEC** | diethylaminoethyl chloride hydrochloride |
| **Diglyme** | diethylene glycol dimethyl ether |
| **Diox** | dioxane |
| **DMA** | dimethylacetamide |
| **DMAA** | *N,N*-dimethylacetoacetamide |
| **DME** | 1,2-dimethoxyethane (glyme) |
| **DML** | dimyristoyl-lecithin |
| **DMS** | dimethylsiloxane |
| **DMSO** | dimethyl sulfoxide |
| **DMT** | dimethyl terephthalate |
| **DNA** | deoxyribonucleic acid |
| **DOCA** | deoxycorticosterone acetate |
| **DPG** | 2,3 diphosphoglycerate |
| **DPI** | dipalmitoyl-lecithin |
| **dpm** | dipivaloylmethanato |
| **DPPH** | diphenylpicrylhydracyl |
| | (usually as the sodium salt) |
| **DST** | disuccinimidyl tartrate |
| **DTBN** | di-t-butyl nitroxide |
| **EAA** | ethyl acetoacetate |
| **EAK** | ethyl amyl ketone |
| **EBA** | *N*-ethyl-*N*-benzylaniline |
| **EDC** | ethylene dichloride |
| **EDTA** | ethylenediaminetetraacetic acid |
| **EGS** | ethylene glycol bis(succinimidyl succinate) |
| **en** | ethylenediamine |
| **Et** | ethyl |
| **FA** | furfuryl alcohol |
| **FAD** | flavin adenine dinucleotide |
| **Fl** | flavin |
| **FMA** | fluoroscein mercuric acetate |
| **G** | guanine |
| **gln** | glutamine |
| **Glu** | glutamic acid |
| **Gly** | glycine |
| **Glyme (glyme)** | 1,2-dimethoxyethane |
| **HAB** | 4,4′- bis(heptyl)azoxybenzene |
| **hex** | hexane (or hexyl) |
| **HFA** | hexafluoroacetone |

| | |
|---|---|
| **His** | histidine |
| **HMPA** | hexamethylphosphoramide |
| **HMPT** | hexamethylphosphorous triamide |
| **HOAB** | *p*-n-heptyloxyazoxybenzene |
| **HOAc** | acetic acid |
| **IH** | immobilized histamine |
| **IHP** | inositolhexaphosphate |
| **Ile** | isoleucine |
| **IMP** | inosine 5′-monophosphate |
| **IPN** | isophthalonitrile |
| **LAP** | leucine aminopeptidase |
| **LDH** | lactic dehydrogenase |
| **Leu** | leucine |
| **Lys** | lysine |
| **M** | metal |
| **MAA** | menthoxyacetic acid |
| **MBBA** | *N*-(*p*-methoxybenzylidene)-*p*-butylaniline |
| **MCA** | monochloroacetic acid |
| **Me** | methyl |
| **MeLeu** | *N*-methylleucine |
| **MMH** | methylmercuric hydroxide |
| **MOM-** | methoxymethyl- |
| **MSA** | methanesulfonic acid |
| **NAD** | nicotinamide adenine dinucleotide |
| **NADH(P)** | nicotinamide adenine dinucleotide |
| **NAI** | *N*-acetylimidazole |
| **NCA** | *N*-chloroacetamide |
| **NM** | nitromethane |
| **NMA** | *N*-methylolacrylamide |
| **NMF** | *N*-methylformamide |
| **NU** | nucleophile |
| **OCBA** | *o*-chlorobenzoic acid |
| **OCT** | *o*-chlorotoluene |
| **ODCB** | *o*-dichlorobenzene |
| **P** | polymer substituent |
| **PAS** | *p*-aminosalicylic acid |
| **PBA** | pyrene butyric acid |
| **PBLG** | poly(L-benzyl μ-glutamate) |
| **PC** | propylene carbonate |
| **PCB** | polychlorinated biphenyl |
| **PCP** | pentachlorophenol |
| **PEG** | polyethylene glycol |

| | |
|---|---|
| **Ph** | phenyl |
| **Phe** | phenylalanine |
| **PMA** | poly(methacrylic acid) |
| **PMMA** | poly(methyl methacrylate) |
| **POC** | cyclopentyloxycarbonyl |
| **PPA** | polyphosphoric acid |
| **Pr** | propyl |
| ***i* Pr** | isopropyl |
| **Pro** | proline |
| **PS** | phosphatidylserine |
| **PVA** | polyvinyl alcohol |
| **PVC** | polyvinyl chloride |
| **PVF** | poly(vinyl flouride) |
| **PVP** | poly(vinyl pyrrolidone) |
| **Pyr (or Py)** | pyridine |
| **SDS** | sodium dodecyl sulfate |
| **Ser** | serine |
| **SLS** | sodium lauryl sulfate |
| **TAB** | trimethylammonium bromide |
| **TBE** | tetrabromoethane |
| **TCNQ** | tetracyanoquinodimethane |
| **TEA** | triethylamine |
| **TFA** | trifluoroacetic acid |
| **THE** | tetrahydrocortisone |
| **THF** | tetrahydrofuran |
| **TMB** | *N*,*N*,*N*′,*N*′-tetramethylbenzidine |
| **TMM** | trimethylenemethane |
| **TMS** | tetramethylsilane |
| **TMU** | tetramethylurea |
| **TNM** | tetranitromethane |
| **TP** | thymolphthalein |
| **TPC** | thymolphthalein complexone |
| **TPE** | tetraphenylethylene |
| **Tr** | trityl |
| **Triglyme** | triethylene glycol dimethyl ether |
| **Ts** | tosyl (or *p*-toluenesulfonyl-) |
| **Tyr** | tyrosine |
| **U** | uridine |
| **UTP** | uridine triphosphate |
| **Val** | valine |
| **VTC** | vinyltrichlorosilane |
| **ZDBC** | zinc dibutyldithiocarbamate |

# 9. Equilibrium constants at 25°C/concentration units for solutions

## Equilibrium constants at 25°C

### Dissociation constants of weak acids

| | | |
|---|---|---|
| $H_3BO_3$ | boric acid | $6.0 \times 10^{-10}$ |
| HCN | hydrogen cyanide (hydrocyanic acid) | $4.0 \times 10^{-10}$ |
| $HC_2H_3O_2$ | acetic acid | $1.8 \times 10^{-5}$ |
| [illegible] | | [illegible] |
| $H_2C_2O_4$ | oxalic acid | $5.4 \times 10^{-2}$ ($K_1$)<br>$5.0 \times 10^{-5}$ ($K_2$) |
| $H_2C_6H_6O_6$ | ascorbic acid | $5.0 \times 10^{-5}$ ($K_1$)<br>$1.5 \times 10^{-12}$ ($K_2$) |
| [illegible] | [illegible] | |
| $H_3PO_3$ | phosphorous acid | $1.6 \times 10^{-2}$ ($K_1$)<br>$7 \times 10^{-7}$ ($K_2$) |
| $H_3PO_4$ | phosphoric acid | $7.6 \times 10^{-3}$ ($K_1$)<br>$6.3 \times 10^{-8}$ ($K_2$)<br>$4.4 \times 10^{-13}$ ($K_3$) |
| $H_2S$ | hydrogen sulfide (hydrosulfuric acid) | $1.1 \times 10^{-7}$ ($K_1$)<br>$1.0 \times 10^{-14}$ ($K_2$) |
| $H_2O + SO_2$ | sulfurous acid | $6.3 \times 10^{-8}$ ($K_2$) |
| $HSO_4^-$ | hydrogen sulfate ion (bisulfate ion) | $1.2 \times 10^{-2}$ |
| HF | hydrofluoric acid | $6.7 \times 10^{-4}$ |
| HOCl | hypochlorous acid | $3.2 \times 10^{-8}$ |
| $HClO_2$ | chlorous acid | $1.1 \times 10^{-2}$ |

## Equilibrium constants at 25°C

| Formula | Name | $K_b$ |
|---|---|---|
| $NH_3$ | ammonia | $1.8 \times 10^{-5}$ |
| $NH_2OH$ | hydroxylamine | $9.1 \times 10^{-9}$ |
| $CH_3NH_2$ | methylamine | $4.4 \times 10^{-4}$ |
| $C_{10}H_{14}N_2$ | nicotine | $7.4 \times 10^{-7}$ ($K_1$)<br>$1.4 \times 10^{-11}$ ($K_2$) |
| $PH_3$ | phosphine | $1 \times 10^{-14}$ |

| Name | Symbol | Definition |
|---|---|---|
| Weight percent | % | (Grams of solute per grams of solution) × 100 |
| Mole fraction[a] | $X_A$ | Moles of A per total number of moles |
| Molarity | $M$ | Moles of solute per litre of solution[b] |
| Normality | $N$ | Equivalents of solute per litre of solution[b] |
| Molality | $m$ | Moles of solute per kg of solvent |
| Weight formality[c] | $f$ | Formula weight of solute per kg of solvent |

[a] The symbol $Y_A$ is often used for the mole fraction of A in a gas phase that is in equilibrium with a liquid solution.

[b] Note that 1 litre = 1 $dm^3$. The SI concentration unit of mol $m^{-3}$ is not very attractive for most work in chemistry, and non-SI concentrations based on the litre are widely used.

[c] [illegible]

| | |
|---|---|
| **AM1** | Austin Method 1,<br>a semiempirical procedure |
| **AO** | Atomic Orbital |
| **CI** | Configuration Interaction<br>a semiempirical procedure |
| **DZ** | Double Zeta,<br>a basis set consisting of two STOs for each atomic orbital |
| **EH** | Extended Hückel theory |
| **EPR** | Electron Paramagnetic Resonance (see ESR) |
| | (see photoelectron spectroscopy) |
| **ESR** | Electron Spin Resonance |
| **FEMO** | Free-Electron Molecular Orbitals |
| **GTO** | Gaussian Type atomic Orbital |
| **HF** | Hartree-Fock method for self-consistent fields |
| **HOMO** | Highest Occupied Molecular Orbital |
| **INDO** | Intermediate Neglect of Differential Overlap,<br>a semiempirical procedure |
| **IR** | Infrared Radiation,<br>as in vibrational spectroscopy |
| | [illegible] |
| **LCAO** | Linear Combination of Atomic Orbitals |
| **LUMO** | Lowest Unoccupied Molecular Orbital (see HOMO) |
| **MINDO** | Modified Intermediate Neglect to Differential Overlap,<br>a semiempirical procedure |
| | Modified Neglect of Differential Overlap |
| **MO** | Molecular Orbital |
| **NMR** | Nuclear Magnetic Resonance |
| **PES** | Photoelectron Spectroscopy |
| **PPP** | Pariser-Parr-Pople method,<br>a semiempirical procedure |
| **SCF** | Self-Consistent Field |
| **STO** | Slater Type Orbital |
| **STO-*NG*** | A Slater type orbital expressed as the sum of *N* guassian orbitals |
| **TFD** | Thomas-Fermi-Dirac method,<br>a statistical treatment of electron density |
| **UV** | Ultraviolet Radiation |
| **VB** | Valence Bond theory |
| **VSEPR** | Valence-Shell Electron Pair Repulsion,<br>a theory of molecular geometry |
| | Wentzel-Kramers-Brillouin method |
| **ZDO** | Zero Differential Overlap,<br>a semiempirical approximation |

Selected half-reaction potentials are listed in order of increasingly positive (anodic) assignments, providing a convenient guide to redox couples in a given range of standard potentials. The synopsis consists of two tables, one each for acid and basic solutions, which have been advisedly restricted to selected half-reactions.

**Standard potentials in acid solutions**

| | |
|---|---|
| $(3/2)N_2 + H^+ + e \rightarrow HN_3$ | −3.10 |
| $Li^+ + e^- \rightarrow Li$ | −3.045 |
| $K^+ + e^- \rightarrow K$ | −2.925 |
| $Rb^+ + e^- \rightarrow Rb$ | −2.925 |
| $Cs^+ + e^- \rightarrow Cs$ | −2.923 |
| $Ra^{2+} + 2e^- \rightarrow Ra$ | −2.916 |
| $Sr^{2+} + 2e^- \rightarrow Sr$ | −2.89 |
| $Ca^{2+} + 2e^- \rightarrow Ca$ | −2.84 |
| $Na^+ + e^- \rightarrow Na$ | −2.714 |
| $No^{2+} + e^- \rightarrow No$ | −2.5 |
| $Fm^{2+} + 2e^- \rightarrow Fm$ | −2.37 |
| $La^{3+} + 3e^- \rightarrow La$ | −2.37 |
| $Y^{3+} + 3e^- \rightarrow Y$ | −2.37 |
| $Ce^{3+} + 3e^- \rightarrow Ce$ | −2.34 |
| $Nd^{3+} + 3e^- \rightarrow Nd$ | −2.32 |
| $Gd^{3+} + 3e^- \rightarrow Gd$ | −2.29 |
| $Mg^{2+} + 2e^- \rightarrow Mg$ | −2.356 |
| $Lu^{3+} + 3e^- \rightarrow Lu$ | −2.30 |
| $1/2H_2 + e^- \rightarrow H^-$ | −2.25 |
| $Cf^{2+} + 2e^- \rightarrow Cf$ | −2.2 |
| $Am^{3+} + 3e^- \rightarrow Am$ | −2.07 |
| $AlF_6^{3-} + 3e^- \rightarrow Al + 6F^-$ | −2.067 |
| $Cm^{3+} + 3e^- \rightarrow Cm$ | −2.06 |
| $Sc^{3+} + 3e^- \rightarrow Sc$ | −2.03 |
| $Bk^{3+} + 3e^- \rightarrow Bk$ | −2.01 |
| $Es^{3+} + 3e^- \rightarrow Es$ | −2.0 |
| $Be^{2+} + 2e^- \rightarrow Be$ | −1.97 |
| $Fm^{3+} + 3e^- \rightarrow Fm$ | −1.96 |
| $Th^{4+} + 4e^- \rightarrow Th$ | −1.83 |
| $Np^{3+} + 3e^- \rightarrow Np$ | −1.79 |
| $Zr^{4+} + 4e^- \rightarrow Zr$ | −1.70 |
| $Al^{3+} + 3e^- \rightarrow Al$ | −1.67 |
| $U^{3+} + 3e^- \rightarrow U$ | −1.66 |
| $Ti^{2+} + 2e^- \rightarrow Ti$ | −1.63 |
| $Hf^{4+} + 4e^- \rightarrow Hf$ | −1.56 |
| $SiF_6^{2-} + 4e^- \rightarrow Si + 6F^-$ | −1.2 |
| $TiF_6^{2-} + 4e^- \rightarrow Ti + 6F^-$ | −1.191 |

*Continued*

| | |
|---|---|
| $Mn^{2+} + 2e^- \rightarrow Mn$ | −1.18 |
| $V^{2+} + 2e^- \rightarrow V$ | −1.13 |
| $Nb^{3+} + 3e^- \rightarrow Nb$ | −1.1 |
| $H_3BO_3 + 3H^+ + 3e^- \rightarrow B + 3H_2O$ | −0.890 |
| $SiO_2(vit) + 4H^+ + 4e^- \rightarrow Si + 2H_2O$ | −0.888 |
| $Ta_2O_5 + 10H^+ + 10e^- \rightarrow 2Ta + 5H_2O$ | −0.81 |
| $Zn^{2+} + 2e \rightarrow Zn$ | −0.7626 |
| $TlI + e^- \rightarrow Tl + I^-$ | −0.74 |
| $Te + 2H^+ + 2e^- \rightarrow H_2Te$ | −0.740 |
| $TlBr + e^- \rightarrow Tl + Br^-$ | −0.658 |
| $TlCl + e^- \rightarrow Tl + Cl^-$ | −0.5568 |
| $Ga^{3+} + 3e^- \rightarrow Ga$ | −0.529 |
| $U^{4+} + e^- \rightarrow U^{3+}$ | −0.52 |
| $Sb + 3H^+ + 3e^- \rightarrow SbH_3(g)$ | −0.510 |
| $H_3PO_2 + H^+ + e^- \rightarrow P(w) + 2H_2O$ | −0.508 |
| $Fe^{2+} + 2e \rightarrow Fe$ | −0.44 |
| $Cr^{3+} + e^- \rightarrow Cr^{2+}$ | −0.424 |
| $Cd^{2+} + 2e \rightarrow Cd$ | −0.4025 |
| $Ti^{3+} + e^- \rightarrow Ti^{2+}$ | −0.37 |
| $PbI_2 + 2e^- \rightarrow Pb + 2I^-$ | −0.365 |
| $Eu^{3+} + e^- \rightarrow Eu^{2+}$ | −0.35 |
| $In^{3+} + 3e^- \rightarrow In$ | −0.3382 |
| $Tl^+ + e^- \rightarrow Tl$ | −0.3363 |
| $PbBr_2 + 2e^- \rightarrow Pb + 2Br^-$ | −0.280 |
| $Co^{2+} + 2e^- \rightarrow Co$ | −0.277 |
| $PbCl_2 + 2e^- \rightarrow Pb + 2Cl^-$ | −0.268 |
| $Ni^{2+} + 2e^- \rightarrow Ni$ | −0.257 |
| $V^{3+} + e^- \rightarrow V^{2+}$ | −0.255 |
| $2SO_4^{2-} + 4H^+ + 4e^- \rightarrow S_2O_6^{2-} + 2H_2O$ | −0.253 |
| $SnF_6^{2-} + 4e^- \rightarrow Sn + 6F^-$ | −0.25 |
| $As + 3H^+ + 3e^- \rightarrow AsH_3$ | −0.225 |
| $Mo^{3+} + 3e^- \rightarrow Mo$ | −0.2 |
| $CuI + e^- \rightarrow Cu + I^-$ | −0.182 |
| $CO_2 + 2H^+ + 2e \rightarrow HCOOH(aq)$ | −0.16 |
| $AgI + e^- \rightarrow Ag + I^-$ | −0.1522 |
| $Sn^{2+} + 2e^- \rightarrow Sn$ | −0.136 |
| $Pb^{2+} + 2e^- \rightarrow Pb$ | −0.1251 |

| Couple | $E°$ (V) |
|---|---|
| $WO_3(c) + 6H^+ + 6e^- \rightarrow W + 3H_2O$ | −0.090 |
| $P(w) + 3H^+ + 3e^- \rightarrow PH_3$ | −0.063 |
| $O_2 + H^+ + e \rightarrow HO_2$ | −0.046 |
| $Hg_2I_2 + 2e^- \rightarrow 2Hg + 2I^-$ | −0.0405 |
| $GeO_2 + 4H^+ + 4e^- \rightarrow Ge(hex) + 2H_2O$ | −0.009 |
| $2H^+ + 2e \rightarrow H_2$ | 0.000 |
| $CuBr + e^- \rightarrow Cu + Br^-$ | 0.033 |
| $HCOOH(aq) + 2H^+ + 2e^- \rightarrow HCHO(aq) + H_2O$ | 0.056 |
| $AgBr + e^- \rightarrow Ag + Br^-$ | 0.0711 |
| $CuCl + e^- \rightarrow Cu + Cl^-$ | 0.121 |
| $C + 4H^+ + 4e^- \rightarrow CH$ | 0.132 |
| $Hg_2Br_2 + 2e^- \rightarrow 2Hg + 2Br^-$ | 0.13920 |
| $S + 2H^+ + 2e \rightarrow H_2S$ | 0.144 |
| $Np^{4+} + e^- \rightarrow Np^{3+}$ | 0.15 |
| $Sb_4O_6 + 12H^+ + 12e^- \rightarrow 4Sb + 6H_2O$ | 0.1504 |
| $SO_4^{2-} + 2H^+ + 2e^- \rightarrow H_2SO_3 + H_2O$ | 0.158 |
| $Cu^{2+} + e^- \rightarrow Cu^+$ | 0.159 |
| $UO_2^{2+} + e^- \rightarrow UO_2^+$ | 0.16 |
| $BiOCl + 2H^+ + 3e^- \rightarrow Bi + H_2O + Cl^-$ | 0.1697 |
| $ReO_2 + 4H^+ + 4e^- \rightarrow Re + 2H_2O$ | 0.22 |
| $AgCl + e^- \rightarrow Ag + Cl^-$ | 0.2223 |
| $HCHO(aq) + 2H^+ + 2e^- \rightarrow CH_3OH(aq)$ | 0.232 |
| $(CH_3)_2SO_2 + 2H^+ + 2e^- \rightarrow (CH_3)_2SO + 2H_2O$ | 0.238 |
| $HAsO_2(aq) + 3H + 3e^- \rightarrow As + 2H_2O$ | 0.248 |
| $HCNO + H^+ + e^- \rightarrow 1/2C_2N_2 + H_2O$ | 0.330 |
| $VO^{2+} + 2H^+ + e^- \rightarrow V^{3+} + H_2O$ | 0.337 |
| $ReO_4^- + 8H^+ + 7e^- \rightarrow Re + 4H_2O$ | 0.34 |
| $Cu^{2+} + 2e^- \rightarrow Cu$ | 0.340 |
| $AgIO_3 + e^- \rightarrow Ag + IO_3^-$ | 0.354 |
| $C_2N_2 + 2H^+ + 2e^- \rightarrow 2HCN(aq)$ | 0.373 |
| $UO_2^+ + 4H^+ + e^- \rightarrow U^{4+} + 2H_2O$ | 0.38 |
| $H_2N_2O_2 + 6H^+ + 4e^- \rightarrow 2NH_3OH^+$ | 0.387 |
| $2H_2SO_3 + 2H^+ + 4e^- \rightarrow S_2O_3^{2-} + 3H_2O$ | 0.400 |
| $Ag_2CrO_4 + 2e^- \rightarrow 2Ag + CrO_4^{2-}$ | 0.4491 |
| $PdBr_4^{2-} + 2e^- \rightarrow Pd + 4Br^-$ | 0.49 |
| $RhCl_5 + 3e^- \rightarrow Rh + 6Cl^-$ | 0.5 |
| $H_2SO_3 + 4H^+ + 4e^- \rightarrow S + 3H_2O$ | 0.500 |
| $2H_2SO_3 + 4H^+ + 6e^- \rightarrow S_4O_6^{2-} + 6H_2O$ | 0.507 |
| $ReO_4^- + 4H^+ + 3e^- \rightarrow ReO_2 + 2H_2O$ | 0.51 |
| $TeO_2(c) + 4H^+ + 4e^- \rightarrow Te + 2H_2O$ | 0.53 |
| $I_2 + 2e^- \rightarrow 2I^-$ | 0.5355 |
| $I_3^- + 2e^- \rightarrow 3I^-$ | 0.536 |
| $AgBrO_3 + e^- \rightarrow Ag + BrO_3^-$ | 0.546 |
| $Cu^{2+} + Cl^- + e^- \rightarrow CuCl$ | 0.559 |
| $TeOOH^+ + 3H^+ + 4e^- \rightarrow Te + 2H_2O$ | 0.559 |
| $MnO_4^- + e^- \rightarrow MnO_4^{2-}$ | 0.56 |
| $S_2O_6^{2-} + 4H^+ + 2e^- \rightarrow 2H_2SO_3$ | 0.569 |
| $CH_3OH(aq) + 2H^+ + 2e^- \rightarrow CH_4 + H_2O$ | 0.59 |
| $Sb_2O_5 + 6H^+ + 4e^- \rightarrow 2SbO^+ + 3H_2O$ | 0.605 |
| $Au(SCN)_4^- + 3e^- \rightarrow Au + 4SCN^-$ | 0.636 |
| $AgC_2H_3O_2 + e^- \rightarrow Ag + C_2H_3O_2$ | 0.643 |
| $Cu^{2+} + Br^- + e^- \rightarrow CuBr$ | 0.654 |
| $Ag_2SO_4 + 2e^- \rightarrow 2Ag + SO_4^{2-}$ | 0.654 |
| $NpO_2^+ + 4H^+ + e^- \rightarrow Np^{4+} + 2H_2O$ | 0.66 |
| $O_2 + 2H^+ + 2e^- \rightarrow H_2O_2$ | 0.695 |
| $PtBr_4^{2-} + 2e^- \rightarrow Pt + 4Br^-$ | 0.698 |
| $2NO + 2H^+ + 2e^- \rightarrow H_2N_2O_2$ | 0.71 |
| $H_2SeO_3 + 4H^+ + 4e^- \rightarrow Se + 3H_2O$ | 0.739 |
| $PtCl_4^{2-} + 2e^- \rightarrow Pt + 4Cl^-$ | 0.758 |
| $Rh^{3+} + 3e^- \rightarrow Rh$ | 0.76 |
| $Fe^{3+} + e^- \rightarrow Fe^{2+}$ | 0.771 |
| $Hg_2^{2+} + 2e^- \rightarrow 2Hg$ | 0.7960 |
| $Ag^+ + e^- \rightarrow Ag$ | 0.7991 |
| $2NO_3^- + 4H^+ + 2e^- \rightarrow N_2O_4 + 2H_2O$ | 0.803 |
| $IrBr_6^{2-} + e^- \rightarrow IrBr_6^{3-}$ | 0.805 |
| $OsO_4(c) + 8H^+ + 8e^- \rightarrow Os + 4H_2O$ | 0.84 |
| $AuBr_4^- + 3e^- \rightarrow Au + 4Br^-$ | 0.854 |
| $2HNO_2 + 4H^+ + 4e^- \rightarrow H_2N_2O_2$ | 0.86 |
| $IrCl_6^{3-} + 3e^- \rightarrow Ir + 6Cl^-$ | 0.86 |
| $Cu^{2+} + I^- + e^- \rightarrow CuI$ | 0.861 |
| $2Hg^{2+} + 2e^- \rightarrow Hg$ | 0.9110 |
| $Pd^{2+} + 2e^- \rightarrow Pd$ | 0.915 |
| $NO_3^- + 3H^+ + 2e^- \rightarrow HNO_2 + H_2O$ | 0.94 |
| $NO_3^- + 4H^+ + 3e^- \rightarrow NO + 2H_2O$ | 0.957 |
| $AuBr_2^- + e^- \rightarrow Au + 2Br^-$ | 0.960 |
| $HNO_2 + H^+ + e^- \rightarrow NO + H_2O$ | 0.996 |
| $AuCl_4^- + 3e^- \rightarrow Au + 4Cl^-$ | 1.002 |
| $Pu^{4+} + e^- \rightarrow Pu^{3+}$ | 1.01 |
| $PuO_2^{2+} + e^- \rightarrow PuO_2^+$ | 1.02 |
| $PuO_2^{2+} + 4H^+ + 2e^- \rightarrow Pu^{4+} + 2H_2O$ | 1.03 |
| $PuO_2^+ + 4H^+ + e^- \rightarrow Pu^{4+} + 2H_2O$ | 1.04 |
| $Sb_2O_5 + 2H^+ + 2e^- \rightarrow Sb_2O_4 + H_2O$ | 1.055 |

| Couple | $E^\circ$ (V) |
|---|---|
| $Br_2(l) + 2e^- \rightarrow Br^-$ | 1.065 |
| $ICl_2^- + e^- \rightarrow 2Cl^- + 1/2I_2$ | 1.07 |
| $N_2O_4 + 2H^+ + 2e^- \rightarrow 2HNO_3$ | 1.07 |
| $Cu^{2+} + 2CN^- + e^- \rightarrow Cu(CN)_2^-$ | 1.12 |
| $SeO_4^{2-} + 4H^+ + 2e^- \rightarrow H_2SeO_3 + H_2O$ | 1.151 |
| $ClO_3^- + 3H^- + 2e^- \rightarrow HClO_2 + H_2O$ | 1.181 |
| $ClO_2 + H^+ + e^- \rightarrow HClO_2$ | 1.188 |
| $S_2Cl_2 + 2e^- \rightarrow 2S + 2Cl^-$ | 1.19 |
| [illegible] | [illegible] |
| $O_2 + 4H^- + 4e^- \rightarrow 2H_2O$ | 1.229 |
| $MnO_2 + 4H^+ + 2e^- \rightarrow Mn^{2+} + 2H_2O$ | 1.23 |
| $N_pO_2^{2+} + e^- \rightarrow NpO_2^+$ | 1.24 |
| $N_2H_5^+ + 3H^+ + 2e^- \rightarrow 2NH_4^+$ | 1.275 |
| $2HNO_2 + 4H^+ + 4e^- \rightarrow N_2O + 3H_2O$ | 1.297 |
| $NH_3OH^+ + 2H^+ + 2e^- \rightarrow NH_4^+ + H_2O$ | 1.35 |
| $Cl_2 + 2e^- \rightarrow 2Cl^-$ | 1.3583 |
| $Cr_2O_7^{2-} + 14H^+ + 6e^- \rightarrow 2Cr^{3+} + 7H_2O$ | 1.36 |
| $2NH_3OH^+ + H^+ + 2e^- \rightarrow N_2H_5^+ + 2H_2O$ | 1.41 |
| $PbO_2(\alpha) + 4H^+ + 2e^- \rightarrow Pb^{2+} + 2H_2O$ | 1.468 |
| $BrO_3^- + 6H^- + 5e \rightarrow 1/2Br_2 + 3H_2O$ | 1.478 |
| $Mn^{3+} + e^- \rightarrow Mn^{2+}$ | 1.5 |
| $MnO_4^- + 8H^+ + 5e^- \rightarrow Mn^{2-} + 4H_2O$ | 1.51 |
| $Au^{3+} + 3e^- \rightarrow Au$ | 1.52 |
| $NiO_2 + 4H^+ + 2e^- \rightarrow Ni^{2+} + 2H_2O$ | 1.593 |
| $H_5IO_6 + H^+ + 2e^- \rightarrow IO_3^- + 3H_2O$ | 1.603 |
| $HBrO + H^+ + e^- \rightarrow 1/2Br_2 + H_2O$ | 1.604 |
| $HClO + H^+ + e^- \rightarrow 1/2Cl_2 + H_2O$ | 1.630 |
| $Bk^{4+} + e \rightarrow Bk^{3+}$ | 1.67 |
| $HClO_2 + 2H^+ + 2e^- \rightarrow HClO + H_2O$ | 1.674 |
| $PbO_2(\alpha) + SO_4^{2-} + 4H^+ + 2e^- \rightarrow PbSO_4 + 2H_2O$ | 1.698 |
| $MnO_4^- + 4H^+ + 3e^- \rightarrow MnO_2 + 2H_2O$ | 1.70 |
| $Ce^{4+} + e \rightarrow Ce^{3+}$ | 1.72 |
| $AmO_2^+ + 4H^+ + 2e^- \rightarrow Am^{3+} + 2H_2O$ | 1.72 |
| $Au^+ + e^- \rightarrow Au$ | 1.83 |
| $Co^{3+} + e^- \rightarrow Co^{3+}$ | 1.92 |
| $HN_3 + 3H^+ + 2e^- \rightarrow NH_4^+ + N_2$ | 1.96 |
| $S_2O_8^{2-} + 2e^- \rightarrow 2SO_4^{2-}$ | 1.96 |
| $Ag^{2+} + e^- \rightarrow Ag^+$ | 1.980 |
| $F_2O + 2H^+ + 4e^- \rightarrow 2F^- + H_2O$ | 2.153 |
| $OH + H^- + e^- \rightarrow H_2O$ | 2.38 |

| Couple | $E^\circ$ (V) |
|---|---|
| $O(g) + 2H^+ + 2e^- \rightarrow H_2O$ | 2.430 |
| $Am^{4+} + e^- \rightarrow Am^{3+}$ | 2.62 |
| $F_2 + 2e^- \rightarrow 2F^-$ | 2.87 |
| $F_2 + 2H^- + 2e^- \rightarrow 2HF(aq)$ | 3.053 |
| $Ca(OH)_2 + 2e^- \rightarrow Ca + 2OH^-$ | −3.026 |
| $Ba(OH)_2 + 2e^- \rightarrow Ba + 2OH^-$ | −2.99 |
| $Sr(OH)_2 + 2e^- \rightarrow Sr + 2OH^-$ | −2.88 |
| $Y(OH)_3 + 3e^- \rightarrow Y + 3OH^-$ | −2.85 |
| $Ho(OH)_3 + 3e^- \rightarrow Ho + 3OH^-$ | −2.85 |
| $Tm(OH)_3 + 3e^- \rightarrow Tm + 3OH^-$ | −2.83 |
| $Lu(OH)_3 + 3e^- \rightarrow Lu + 3OH^-$ | −2.83 |
| $Gd(OH)_3 + 3e^- \rightarrow Gd + 3OH^-$ | −2.82 |
| $Tb(OH)_3 + 3e^- \rightarrow Tb + 3OH^-$ | −2.82 |
| $La(OH)_3 + 3e^- \rightarrow La + 3OH^-$ | −2.80 |
| $Dy(OH)_3 + 3e^- \rightarrow Dy + 3OH^-$ | −2.80 |
| $Pr(OH)_3 + 3e^- \rightarrow Pr + 3OH^-$ | −2.79 |
| $Ce(OH)_3 + 3e^- \rightarrow Ce + 3OH^-$ | −2.78 |
| $Nd(OH)_3 + 3e^- \rightarrow Nd + 3OH^-$ | −2.78 |
| $Pm(OH)_3 + 3e^- \rightarrow Pm + 3OH^-$ | −2.76 |
| $Mg(OH)_2 + 2e^- \rightarrow Mg + 2OH^-$ | −2.687 |
| $Sc(OH)_3 + 3e^- \rightarrow Sc + 3OH^-$ | −2.60 |
| $ThO_2 + 2H_2O + 4e^- \rightarrow Th + 4OH^-$ | −2.56 |
| $Am(OH)_3 + 3e^- \rightarrow Am + 3OH^-$ | −2.53 |
| $Cm(OH)_3 + 3e^- \rightarrow Cm + 3OH^-$ | −2.5 |
| $Al(OH)_4^- + 3e^- \rightarrow Al + 4OH^-$ | −2.310 |
| $Al(OH)_3(g) + 3e^- \rightarrow Al + 3OH^-$ | −2.300 |
| $Np(OH)_3 + 3e^- \rightarrow Np + 3OH^-$ | −2.2 |
| $TiO + H_2O + 2e^- \rightarrow Ti + 2OH^-$ | −2.13 |
| $U(OH)_3 + 3e^- \rightarrow U + 3OH^-$ | −2.10 |
| $Ti_2O_3 + H_2O + 2e^- \rightarrow 2TiO + 2OH^-$ | −1.95 |
| $B(OH)_4^- + 3e^- \rightarrow B + 4OH^-$ | −1.811 |
| $SiO_3^{2-} + 3H_2O + 4e^- \rightarrow Si + 6OH^-$ | −1.7 |
| $HPO_3^{2-} + 2H_2O + 2e^- \rightarrow H_2PO_2^- + 3OH^-$ | −1.57 |
| $Mn(OH)_2 + 2e^- \rightarrow Mn + 2OH^-$ | −1.56 |
| $PuO_2 + 2H_2O + e^- \rightarrow Pu(OH)_3 + OH^-$ | −1.4 |
| $2TiO_2 + H_2O + 2e^- \rightarrow Ti_2O_3 + 2OH^-$ | −1.38 |
| $Zn(CN)_4^{2-} + 2e^- \rightarrow Zn + 4CN^-$ | −1.34 |
| $Cr(OH)_3 + 3e^- \rightarrow Cr + 3OH^-$ | −1.33 |
| $Cr(OH)_4^- + 3e^- \rightarrow Cr + 4OH^-$ | −1.33 |
| $CdS + 2e^- \rightarrow Cd + S^{2-}$ | −1.255 |
| $Zn(OH)_2 + 2e^- \rightarrow Zn + 2OH^-$ | −1.248 |

| Couple | $E°$ (V) |
|---|---|
| $H_2GaO_3^- + H_2O + 3e^- \rightarrow Ga + 4OH^-$ | −1.22 |
| $Te + 2e^- \rightarrow Te^{2-}$ | −1.14 |
| $PO_4^{3-} + 2H_2O + 2e^- \rightarrow HPO_3^{2-} + 3OH^-$ | −1.12 |
| $WO_4^{2-} + 4H_2O + 6e^- \rightarrow W + 8OH^-$ | −1.074 |
| $Zn(NH_3)_4^{2+} + 2e^- \rightarrow Zn + 4NH_3$ | −1.04 |
| $HGeO_3^- + 2H_2O + 4e^- \rightarrow Ge + 5OH^-$ | −1.03 |
| $MnO_2 + 2H_2O + 4e^- \rightarrow Mn + 4OH^-$ | −0.980 |
| $CNO^- + H_2O + 2e^- \rightarrow CN^- + 2OH^-$ | −0.97 |
| $Cd(CN)_4^{2-} + 2e^- \rightarrow Cd + 4CN^-$ | −0.943 |
| $PbS + 2e^- \rightarrow Pb + S^{2-}$ | −0.923 |
| $MoO_4^{2-} + 4H_2O + 6e^- \rightarrow Mo + 8OH^-$ | −0.913 |
| $Sn(OH)_6^- + 2e^- \rightarrow HSnO_2^- + H_2O + 3OH^-$ | −0.91 |
| $P + 3H_2O + 3e^- \rightarrow PH_3 + 3OH^-$ | −0.89 |
| $2H_2O + 2e^- \rightarrow H_2 + 2OH^-$ | −0.828 |
| $VO + H_2O + 2e^- \rightarrow V + 2OH^-$ | −0.820 |
| $HFeO_2^- + H_2O + 2e^- \rightarrow Fe + 3OH^-$ | −0.8 |
| $MoO_4^{2-} + 2H_2O + 2e^- \rightarrow MoO_2 + 4OH^-$ | −0.780 |
| $CdCO_3 + 2e^- \rightarrow Cd + CO_3^{2-}$ | −0.734 |
| $Co(OH)_2 + 2e^- \rightarrow Co + 2OH^-$ | −0.733 |
| $CrO_4^{2-} + 4H_2O + 3e^- \rightarrow Cr(OH)_4^- + 4OH^-$ | −0.72 |
| $Ag_2S + 2e^- \rightarrow Ag + S^{2-}$ | −0.691 |
| $FeO_2^- + H_2O + e^- \rightarrow HFeO_2^- + OH^-$ | −0.69 |
| $AsO_2 + 2H_2O + 3e^- \rightarrow As + 4OH^-$ | −0.68 |
| $Se + 2e^- \rightarrow Se^{2-}$ | −0.67 |
| $Sb(OH)_4^- + 3e^- \rightarrow Sb + 4OH^-$ | −0.639 |
| $Cd(NH_5)_4^{2+} + 2e^- \rightarrow Cd + 4NH_3$ | −0.622 |
| $ReO_4^- + 2H_2O + 7e^- \rightarrow Re + 8OH^-$ | −0.604 |
| $ReO_4^- + 2H_2O + 3e^- \rightarrow ReO_2 + 4OH^-$ | −0.594 |
| $2SO_3^{2-} + 3H_2O + 4e^- \rightarrow S_2O_3^{2-} + 6OH^-$ | −0.58 |
| $Cu_2S + 2e^- \rightarrow 2Cu + S^{2-}$ | −0.542 |
| $HPbO_2^- + H_2O + 2e^- \rightarrow Pb + 3OH^-$ | −0.502 |
| $Ni(NH_3)_6^{2+} + 2e^- \rightarrow Ni + 6NH_5$ | −0.476 |
| $Sb(OH)_6^- + 2e^- \rightarrow Sb(OH)_4^- + 2OH^-$ | −0.465 |
| $Bi_2O_3 + 3H_2O + 6e^- \rightarrow Bi + 6OH^-$ | −0.452 |
| $NiCO_3 + 2e^- \rightarrow Ni + CO_3^{2-}$ | −0.45 |
| $Cu(CN)_2^- + e^- \rightarrow Cu + 2CN^-$ | −0.44 |
| $TeO_3^{2-} + 3H_2O + 4e^- \rightarrow Te + 6OH^-$ | −0.42 |
| $Cu_2O + H_2O + 2e^- \rightarrow 2Cu + 2OH^-$ | −0.365 |
| $SeO_3^{2-} + 3H_2O + 4e^- \rightarrow Se + 6OH^-$ | −0.36 |
| $O_2 + e^- \rightarrow O_2^-$ | −0.33 |
| $Ag(CN)_2^- + e^- \rightarrow Ag + 2CN^-$ | −0.31 |
| $Cu(SCN) + e^- \rightarrow Cu + SCN^-$ | −0.310 |
| $CuO + H_2O + 2e^- \rightarrow Cu + 2OH^-$ | −0.29 |
| $Mn_2O_3 + 2H_2O + 2e^- \rightarrow 2Mn(OH)_2^- + 2OH^-$ | −0.25 |
| $2CuO + H_2O + 2e^- \rightarrow Cu_2O + 2OH^-$ | −0.22 |
| $O_2 + H_2O + 2e^- \rightarrow HO_2^- + OH^-$ | −0.0649 |
| $Tl(OH)_3 + 2e^- \rightarrow TlOH + 2OH^-$ | −0.05 |
| $MnO_2 + H_2O + 2e^- \rightarrow Mn(OH)_2 + 2OH^-$ | −0.05 |
| $AgCN + e^- \rightarrow Ag + CN^-$ | −0.017 |
| $NO_3^- + H_2O + 2e^- \rightarrow NO_2^- + 2OH^-$ | 0.01 |
| $Co(NH_3)_6^{3+} + e^- \rightarrow Co(NH_3)_6^{2+}$ | 0.058 |
| $TeO_4^{2-} + H_2O + e^- \rightarrow TeO_3^{2-} + 2OH^-$ | 0.07 |
| $HgO(red) + H_2O + 2e^- \rightarrow Hg + 2OH^-$ | 0.0977 |
| $N_2H_4 + 2H_2O + 2e^- \rightarrow 2NH_3(aq) + 2OH^-$ | 0.1 |
| $VO_4^{3-} + 4H_2O + 5e^- \rightarrow V + 8OH^-$ | 0.120 |
| $HO_2^- + H_2O + e^- \rightarrow OH + 2OH^-$ | 0.184 |
| $O_2^- + H_2O + e^- \rightarrow HO_2^- + OH^-$ | 0.20 |
| $PbO_2(\beta) + H_2O + e^- \rightarrow HPbO_2 + OH^-$ | 0.208 |
| $PbO_2(\beta) + H_2O + 2e^- \rightarrow PbO(red) + 2OH^-$ | 0.247 |
| $IO_3^- + 3H_2O + 6e^- \rightarrow I^- + 6OH^-$ | 0.257 |
| $PuO_2(OH)_2 + e^- \rightarrow PuO_2OH + OH^-$ | 0.3 |
| $PbO_3^{2-} + 2H_2O + 2e^- \rightarrow HPbO_2^- + 3OH^-$ | 0.330 |
| $Ag_2O + H_2O + 2e^- \rightarrow 2Ag + 2OH^-$ | 0.342 |
| $Ag(NH_3)_2^+ + e^- \rightarrow Ag + NH_3$ | 0.373 |
| $ClO_3^- + H_2O + 2e^- \rightarrow ClO_2^- + 2OH^-$ | 0.374 |
| $NH_2OH + H_2O + 2e^- \rightarrow NH_3 + 2OH^-$ | 0.42 |
| $Ag(SO_3)_2^{3-} + e^- \rightarrow Ag + 2SO_3^{2-}$ | 0.43 |
| $Ag_2CO_3 + 2e^- \rightarrow 2Ag + CO_3^{2-}$ | 0.47 |
| $IO^- + H_2O + 2e^- \rightarrow I^- + 2OH^-$ | 0.472 |
| $NiO_2 + 2H_2O + 2e^- \rightarrow Ni(OH)_2 + 2OH^-$ | 0.490 |
| $BrO_3^- + 3H_2O + 6e^- \rightarrow Br^- + 6OH^-$ | 0.584 |
| $RuO_4^- + e^- \rightarrow RuO_4^{2-}$ | 0.593 |
| $MnO_4^{2-} + 2H_2O + 2e^- \rightarrow MnO_2 + 4OH^-$ | 0.62 |
| $2AgO + H_2O + 2e^- \rightarrow Ag_2O + 2OH^-$ | 0.640 |
| $H_3IO_6^{2-} + 2e^- \rightarrow IO_3^- + 3OH^-$ | 0.656 |
| $ClO_2^- + H_2O + 2e^- \rightarrow ClO^- + 2OH^-$ | 0.681 |
| $Ag_2O_3 + H_2O + 2e^- \rightarrow 2AgO + 2OH^-$ | 0.739 |
| $BrO^- + H_2O + 2e^- \rightarrow Br^- + 2OH^-$ | 0.766 |
| $HO_2^- + H_2O + 2e^- \rightarrow 3OH^-$ | 0.867 |
| $ClO^- + H_2O + 2e^- \rightarrow Cl^- + 2OH^-$ | 0.890 |
| $O_3 + H_2O + 2e^- \rightarrow O_2 + 2OH^-$ | 1.246 |
| $OH + e^- \rightarrow OH^-$ | 1.985 |

# 12. Typical UV absorptions of unconjugated chromophores

Values are typically those found in nonpolar solvents

| Chromophore | Transition | $\lambda_{max}$ (nm) | log $E_{max}$ (approx) |
|---|---|---|---|
| $C_2H_6$ | $\sigma\rightarrow\sigma^*$ | 135 | |
| ROH | $n\rightarrow\sigma^*$ | 180–185 | 2.70 |
| RSH | $n\rightarrow\sigma^*$ | 190–200; 225–230 | 3.18; 2.18 |
| RCl | $n\rightarrow\sigma^*$ | 170–175 | 2.48 |
| RBr | $n\rightarrow\sigma^*$ | 200–210 | 2.60 |
| $H_2O$ | $n\rightarrow\sigma^*$ | 180–185 | 3.48 |
| $R_2S$ | $n\rightarrow\sigma^*$ | 210–215; 235–240 | 3.10; 2.00 |
| RSSR | $n\rightarrow\sigma^*$ | 250 | 2.60 |
| Amines | $n\rightarrow\sigma^*$ | 190–200 | 3.40–3.60 |
| $C_2H_4$ | $\pi\rightarrow\pi^*$ | 163; 174 | 4.18; 3.74 |
| R–C≡CH | $\pi\rightarrow\pi^*$ | 185; 223 | 3.34; 2.08 |
| R–C≡C–R | $\pi\rightarrow\pi^*$ | 178; 196–223 | 4.00; 3.30; 2.20 |
| C=C=C | $\pi\rightarrow\pi^*$ | 170–185; 230 | 3.70–4.00; 2.78 |
| C=C=O | $n\rightarrow\pi^*$ | 380 | 1.30 |
| [illegible] | $n\rightarrow\pi^*$ | [illegible] | [illegible] |
| $RCO_2H, RCO_2R'$ | $n\rightarrow\pi^*$ | 195–210 | 1.60–2.00 |
| $RCONH_2$ | $n\rightarrow\pi^*$ | 175 | 3.85 |
| RCN | | <170 | |
| RCONHCOR | $n\rightarrow\pi^*$ | 230–240 | 1.90–2.00 |
| $RNO_2$ | $n\rightarrow\pi^*$ | 270–280 | 1.30–1.48 |
| | $\pi\rightarrow\pi^*$ | 200–210 | 4.18 |
| RONO | $n\rightarrow\pi^*$ | 350 | 2.18 |
| | $\pi\rightarrow\pi^*$ | 220 | 3.00 |
| RNO (monomer) | $n\rightarrow\pi^*$ | 600–650; 300 | 1.30; 2.00 |
| | $\pi\rightarrow\pi^*$ | <200 | |
| RCHO | $n\rightarrow\pi^*$ | 290 | 1.18 |
| | $\pi\rightarrow\pi^*$ | 185–195 | |
| $R_2CO$ | $n\rightarrow\pi^*$ | 270–290 | 1.00–1.30 |
| RCOCOR | $n\rightarrow\pi^*$ | 420–460; 280–285 | 1.00; 1.30 |
| RSOR | $n\rightarrow\pi^*$ | 210–230 | 3.18–3.40 |
| $RSO_2R$ | | <190 | |

## 13. Typical UV absorption maxima of substituted benzenes

| Bands arising from the same transition type and grouped together | | | | |
|---|---|---|---|---|
| **Substituent** | **Solvent** | **$\lambda_{max}$ nm (log E in brackets)** | | |
| –OH | Water | 210.5 (3.78) | 270 (3.16) | |
| $–O^-$ | Water | 235 (3.97) | 287 (3.42) | |
| –SH | Hexane | 236 (4.00) | 269 (2.85) | |
| $–NH_2$ | Water | 230 (3.93) | 280 (3.16) | |
| $–NH_3^+$ | Water | 203 (3.88) | 254 (2.20) | |
| $–NO_2$ | Hexane | 252 (4.00) | 280 (3.00) | 330 (2.10) |
| –CHO | Ethanol | 244 (4.18) | 280 (3.18) | 328 (1.30) |
| $–CO_2H$ | Water | 230 (4.00) | 270 (2.90) | |
| $–CO_2^-$ | Water | 224 (3.94) | 268 (2.75) | |
| –CN | Water | 224 (4.11) | 271 (3.00) | |
| –F | Ethanol | 204 (3.80) | 254 (3.00) | |
| –Cl | Water | 209.5 (3.87) | 263.5 (2.28) | |
| –I | Water | 207 (3.85) | 257 (2.85) | |
| $–CH_3$ | Water | 207 (3.85) | 261 (2.35) | |
| $–CH=CH_2$ | Ethanol | 244 (4.08) | 282 (2.65) | |
| $–C≡C–C_6H_5$ | Hexane | 236 (4.10) | 278 (2.81) | |
| $–C_6H_5$ | Ethanol | 246 (4.30) | | |

# 14. Typical UV absorption maxima of aromatic and heteroaromatic compounds

| Bands for different compounds arising from the same transition type are grouped together. For bands with fine structure, only $\lambda_{max}$ of the band centre is given. | | | | | | |
|---|---|---|---|---|---|---|
| Benzene | Cyclohexane | | | [illegible] | [illegible] | [illegible] |
| Napthalene | Ethanol | 167 (4.48) | 190 (4.00) | 220 (5.12) | 286 (3.97) | 312 (2.46) |
| Anthracene | Cyclohexane | 186 (4.51) | 221 (4.16) | 256 (5.26) | 375 (3.95) | (buried) |
| Napthacene | Benzene | 187 (4.20); 211 (4.64) | 230 (3.23) | 272 (5.26) | 474 (4.10) | |
| Azulene | Cyclohexane | 700 (2.48) | 193 (4.26) | 236 (4.34) | 269 (4.67) | 357 (3.60) |
| Phenanthrene | Cyclohexane | | [illegible] | [illegible] | [illegible] | [illegible] |
| Isoquinoline | Cyclohexane | | | 218 (4.80) | 265 (3.62) | 313 (3.26) |
| Acridine | Ethanol | | | 250 (5.30) | 358 (4.00) | |
| Pyridine | Hexane | 195 (3.88) | | 251 (3.30) | 270 (2.65) | |
| Pyrimidine | Cyclohexane | | | 243 (3.31) | 298 (2.48) | |
| Pyrazine | Cyclohexane | | | 260 (3.80) | 327 (2.00) | |
| Purine | Water | <220 (3.48) | | 263 (3.90) | | |
| Pyrrole | Hexane | 210 (3.71) | | | 240 (2.48) | |
| Furan | Hexane | 205 (3.81) | | | | |
| Thiophene | Hexane | 231 (3.85) | | | | |
| Imidazole | Ethanol | 207 (3.70) | | | | |
| Isoxazole | Ethanol | 211 (3.60) | | | | |
| Thiazole | Ethanol | 240 (3.60) | | | | |

# 15. Common isotopes for Mössbauer spectroscopy

| Isotope | Natural abundance (%) | Spin states | | $E_\gamma$ (keV) | Source isotope | Half-life | Line width (mm $s^{-1}$) | Sign of $\Delta R/R$ |
|---|---|---|---|---|---|---|---|---|
| | | (gd) | (ex) | | | | | |
| $^{57}Fe$ | 2.2 | 1/2 | 3/2 | 14.4 | $^{57}Co$ | 270 d | 0.2 | –ve |
| $^{121}Sb$ | 2.1 | 5/2 | 7/2 | 37.2 | $^{121m}Sb$ | 77 y | 1.8 | –ve |
| $^{127}I$ | 100 | 5/2 | 7/2 | 57.6 | $^{127m}Te$ | 105 d | 2.0 | –ve |
| $^{129}I$ | 0.6 | 7/2 | 5/2 | 27.7 | $^{129m}Te$ | 33 d | 0.8 | +ve |
| $^{99}Ru$ | 12.8 | 3/2 | 5/2 | 89.4 | $^{99}Rh$ | 16 d | 0.3 | +ve |
| $^{193}Ir$ | 61.5 | 1/2 | 3/2 | 73.0 | $^{193m}Os$ | 30 h | 1.4 | +ve |

| Isot | Spin | Nat. Abun danc | Sensly rel.* | Sensly bs.** | NMR-frequendHz) at a field (T) 2.3488 | 4.6975 | 5.87 | 7.0463 | 50 | 11.7440 | 14.0926 | 17.6 | 18.7900 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1/2 | 99.9 | 1.00 | | 100.00( | 200.000 | 250.( | 300.000 | 400 | 500.000 | 300.000 | 750.( | 800.000 |
| 2 | 1 | $1.5^{-2}$ | $9.65\times10^{-3}$ | $\times10^{-6}$ | 15.35 | 30.701 | 38.: | 46.051 | ·02 | 76.753 | 92.102 | 115.' | 122.804 |
| 3 | 1/2 | 0 | 1.21 | | 106.66: | 213.327 | 266.( | 319.990 | 464 | 533.317 | 539.980 | 799.! | 853.328 |
| 3 | 1/2 | $1.3^{-4}$ | 0.44 | $\times10^{-7}$ | 76.17! | 152.355 | 190.: | 228.533 | 310 | 380.888 | 457.066 | 571.: | 609.420 |
| 6 | 1 | 7.4 | $8.50\times10^{-3}$ | $\times10^{-4}$ | 14.71( | 29.431 | 36.7 | 44.146 | 62 | 73.578 | 88.292 | 110.: | 117.724 |
| 7 | 3/2 | 92.5 | 0.29 | | 38.86: | 77.727 | 97.1 | 116.590 | 154 | 194.317 | 233.180 | 291.: | 310.908 |
| 9 | 3/2 | 100 | $1.39\times10^{-2}$ | $\times10^{-2}$ | 14.05: | 28.106 | 35.1 | 42.160 | 13 | 70.267 | 84.320 | 105.: | 116.426 |
| 10 | 3 | 19.5 | $1.99\times10^{-2}$ | $\times10^{-3}$ | 10.74( | 21.493 | 26.8 | 32.239 | 86 | 53.732 | 64.478 | 80.5 | 85.972 |
| 11 | 3/2 | 80.4 | 0.17 | ( | 32.08: | 64.167 | 80.2 | 96.251 | 135 | 160.419 | 192.502 | 240.( | 256.670 |
| 13 | 1/2 | 1.1 | $1.59\times10^{-2}$ | $\times10^{\ 4}$ | 25.14: | 50.288 | 62.8 | 75.432 | 177 | 125.721 | 150.864 | 188.5 | 201.154 |
| 14 | 1 | 99.6 | $1.01\times10^{\ 3}$ | $\times10^{\ 3}$ | 7.22: | 14.447 | 18.( | 21.671 | 94 | 36.118 | 43.342 | 54.1 | 57.788 |
| 15 | 1/2 | 0.3 | $1.04\times10^{\ 3}$ | $\times10^{\ 6}$ | 10.13: | 20.265 | 25.: | 30.398 | 31 | 50.664 | 60.796 | 75.9 | 81.062 |
| 17 | 5/2 | $3.7^{\ 2}$ | $2.91\times10^{-2}$ | $\times10^{-5}$ | 13.55: | 27.113 | 33.8 | 40.670 | 27 | 67.784 | 81.340 | 101.( | 108.454 |
| 19 | 1/2 | 100 | 0.83 | ( | 94.07 | 188.154 | 235.1 | 282.231 | 308 | 470.385 | 564.462 | 705.5 | 752.616 |
| 21 | 3/2 | 0.2 | $2.50\times10^{-3}$ | $\times10^{-6}$ | 7.89: | 15.788 | 19.7 | 23.683 | 77 | 39.472 | 47.366 | 59.2 | 63.154 |
| 23 | 3/2 | 100 | $9.25\times10^{-2}$ | $\times10^{-2}$ | 26.451 | 52.902 | 66.1 | 79.353 | 105 | 132.256 | 158.706 | 198.: | 211.610 |
| 25 | 5/2 | 10.1 | $2.67\times10^{-3}$ | $\times10^{-4}$ | 6.119 | 12.238 | 15.2 | 18.358 | 77 | 30.597 | 36.716 | 45.8 | 48.954 |
| 27 | 5/2 | 100 | 0.21 | ( | 26.05: | 52.114 | 65.1 | 78.172 | 129 | 130.287 | 56.344 | 195.4 | 208.458 |
| 29 | 1/2 | 4.7 | $7.84\times10^{-3}$ | $\times10^{-4}$ | 19.86: | 39.730 | 49.6 | 59.595 | 60 | 99.325 | 19.190 | 148.9 | 159.280 |
| 31 | 1/2 | 100 | $6.63\times10^{-2}$ | $\times10^{-2}$ | 40.481 | 80.961 | 101.2 | 121.442 | 123 | 202.404 | 242.884 | 303.6 | 323.846 |
| 33 | 3/2 | 0.7 | $2.26\times10^{-3}$ | $\times10^{-5}$ | 7.67( | 15.339 | 19.1 | 23.009 | 78 | 38.348 | 46.018 | 57.5 | 61.356 |
| 35 | 3/2 | 75.5 | $4.70\times10^{-3}$ | $\times10^{-3}$ | 9.79: | 19.596 | 24.4 | 29.395 | 93 | 48.991 | 58.790 | 73.4 | 78.386 |
| 37 | 3/2 | 24.4 | $2.71\times10^{-3}$ | $\times10^{-4}$ | 8.15( | 16.311 | 20.3 | 24.467 | 23 | 40.779 | 48.934 | 61.1 | 65.246 |
| 39 | 3/2 | 93.1 | $5.08\times10^{-4}$ | $\times10^{\ 4}$ | 4.667 | 9.333 | 11.6 | 13.999 | 66 | 23.333 | 27.998 | 34.9 | 37.332 |
| 41 | 3/2 | 6.8 | $8.40\times10^{\ 5}$ | $\times10^{\ 6}$ | 2.561 | 5.122 | 6.4 | 7.684 | 45 | 12.806 | 15.368 | 19.2 | 20.490 |
| 43 | 7/2 | 0.1 | $6.40\times10^{\ 3}$ | $\times10^{-6}$ | 6.72: | 13.456 | 16.8 | 20.184 | 13 | 33.641 | 40.368 | 50.4 | 53.826 |
| 45 | 7/2 | 100 | 0.30 | ( | 24.29( | 48.588 | 60.7 | 72.882 | 76 | 121.470 | 45.764 | 182.2 | 194.352 |
| 47 | 5/2 | 7.2 | $2.09\times10^{-3}$ | $\times10^{-4}$ | 5.637 | 11.273 | 14.0 | 16.910 | 47 | 28.184 | 33.820 | 42.2 | 45.094 |
| 49 | 7/2 | 5.5 | $3.76\times10^{-3}$ | $\times10^{-4}$ | 5.63: | 11.276 | 14.0 | 16.914 | 52 | 28.191 | 33.828 | 42.2 | 45.104 |
| 50 | 6 | 0.2 | $5.55\times10^{-2}$ | $\times10^{-4}$ | 9.97( | 19.940 | 24.9 | 29.911 | 31 | 49.852 | 59.822 | 74.7 | 79.762 |
| 51 | 7/2 | 99.7 | 0.38 | ( | 26.28: | 52.576 | 65.7 | 78.864 | 162 | 131.440 | 57.728 | 197.1 | 210.304 |
| 53 | 3/2 | 9.5 | $9.03\times10^{-4}$ | $\times10^{-3}$ | 5.652 | 11.304 | 14.1 | 16.956 | 08 | 28.260 | 33.912 | 42.3 | 45.216 |

* at consfield for equal nur of nuclei. ** pro of relative sensitivnd natural abund:.

| Isœ | Spin | Na Ab da% | Sensy rel.* | abs.** | NMR-frequen(Hz) at a field (T) 2.348 | 4.6975 | 5.87 | 7.0463 | 50 | 11.7440 | 14.0926 | 17.6 | 18.7900 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 55 n | 5/2 | 10 | 0.18 | | 24.66 | 49.328 | 61. | 73.993 | 57 | 123.322 | 147.986 | 184. | 197.314 |
| 57 e | 1/2 | | $3.37\times10^{-5}$ | $\times10^{-7}$ | 3.23 | 6.462 | 8. | 9.693 | 25 | 16.156 | 19.386 | 24. | 25.850 |
| 59 o | 7/2 | 10 | 0.28 | | 23.61 | 47.228 | 59. | 70.842 | 57 | 118.071 | 141.684 | 177. | 188.914 |
| 61 li | 3/2 | | $3.57\times10^{-3}$ | $\times10^{-5}$ | 8.93 | 17.872 | 22. | 26.808 | 44 | 44.681 | 53.616 | 67. | 71.488 |
| 63 u | 3/2 | 6 | $9.31\times10^{-2}$ | $\times10^{-2}$ | 26.50 | 53.010 | 66. | 79.515 | 20 | 132.525 | 159.030 | 198. | 212.040 |
| 65 u | 3/2 | 3 | 0.11 | $\times10^{-2}$ | 28.39 | 56.788 | 70. | 85.183 | 77 | 141.972 | 170.366 | 212. | 227.154 |
| 67 n | 5/2 | | $2.85\times10^{-3}$ | $\times10^{-2}$ | 6.25 | 12.508 | 15. | 18.762 | 60 | 31.271 | 37.524 | 46. | 50.320 |
| 69 a | 3/2 | 6 | $6.91\times10^{-2}$ | $\times10^{-2}$ | 24.00 | 48.006 | 60. | 72.009 | 12 | 120.016 | 144.018 | 180. | 192.024 |
| 71 a | 3/2 | 3 | 0.14 | $\times10^{-2}$ | 30.49 | 60.990 | 76. | 91.485 | 80 | 152.476 | 182.970 | 228. | 243.960 |
| 73 e | 9/2 | | $1.4\times10^{-3}$ | $\times10^{-4}$ | 3.48 | 6.976 | 8. | 10.465 | 53 | 17.442 | 20.930 | 26. | 27.906 |
| 75 s | 3/2 | 10 | $2.51\times10^{-2}$ | $\times10^{-2}$ | 17.12 | 34.253 | 42. | 51.380 | 07 | 85.634 | 102.760 | 128. | 137.014 |
| 77 e | 1/2 | | $6.93\times10^{-3}$ | $\times10^{-4}$ | 19.06 | 38.135 | 47. | 57.203 | 70 | 95.338 | 114.406 | 143. | 152.540 |
| 79 r | 3/2 | 5 | $7.86\times10^{-2}$ | $\times10^{-2}$ | 25.05 | 50.107 | 62. | 75.160 | 14 | 125.267 | 150.320 | 187. | 200.428 |
| 81 r | 3/2 | 4 | $9.85\times10^{-2}$ | $\times10^{-2}$ | 27.00 | 54.012 | 67. | 81.018 | 25 | 135.031 | 162.036 | 202. | 216.050 |
| 83 r | 9/2 | 1 | $1.88\times10^{-3}$ | $\times10^{-4}$ | 3.84 | 7.695 | 9. | 11.543 | 91 | 19.238 | 23.086 | 28. | 30.782 |
| 85 b | 5/2 | 7 | $1.05\times10^{-2}$ | $\times10^{-3}$ | 9.65 | 19.310 | 24. | 28.965 | 20 | 48.276 | 57.930 | 72. | 77.240 |
| 87 b | 3/2 | 2 | 0.17 | $\times10^{-2}$ | 32.72 | 65.442 | 81. | 98.163 | 85 | 163.606 | 196.326 | 245. | 261.770 |
| 87 r | 9/2 | | $2.69\times10^{-3}$ | $\times10^{-4}$ | 4.33 | 8.667 | 10. | 13.001 | 35 | 21.669 | 26.002 | 32. | 34.670 |
| 89 r | 1/2 | 10 | $1.18\times10^{-4}$ | $\times10^{-4}$ | 4.89 | 9.798 | 12. | 14.697 | 96 | 24.496 | 29.394 | 36. | 39.192 |
| 91 r | 5/2 | 1 | $9.48\times10^{-3}$ | $\times10^{-3}$ | 9.33 | 18.660 | 23. | 27.991 | 21 | 46.651 | 55.982 | 69. | 74.642 |
| 93 b | 9/2 | 10 | 0.48 | | 24.44 | 48.885 | 61. | 73.328 | 71 | 122.214 | 146.656 | 183. | 195.542 |
| 95 o | 5/2 | 1 | $3.23\times10^{-3}$ | $\times10^{-4}$ | 6.51 | 13.029 | 16. | 19.544 | 59 | 32.574 | 39.088 | 48. | 52.118 |
| 97 o | 5/2 | | $3.43\times10^{-3}$ | $\times10^{-4}$ | 6.65 | 13.304 | 16. | 19.957 | 09 | 33.261 | 39.914 | 49. | 53.218 |
| 99 u | 3/2 | 1 | $1.95\times10^{-4}$ | $\times10^{-5}$ | 3.38 | 6.779 | 8. | 10.169 | 59 | 16.949 | 20.338 | 25. | 27.118 |
| 101 u | 5/2 | 1 | $1.41\times10^{-3}$ | $\times10^{-4}$ | 4.94 | 9.882 | 12. | 14.824 | 65 | 24.707 | 29.648 | 37. | 39.530 |
| 103 h | 1/2 | 10 | $3.11\times10^{-5}$ | $\times10^{-5}$ | 3.14 | 6.295 | 7. | 9.442 | 90 | 15.737 | 18.884 | 23. | 25.180 |
| 105 d | 5/2 | 2 | $1.12\times10^{-3}$ | $\times10^{-4}$ | 4.57 | 9.152 | 11. | 13.728 | 05 | 22.881 | 27.456 | 34. | 36.610 |
| 107 g | 1/2 | 5 | $6.62\times10^{-5}$ | $\times10^{-5}$ | 4.04 | 8.093 | 10. | 12.139 | 86 | 20.233 | 24.278 | 30. | 32.372 |
| 109 g | 1/2 | 4 | $1.01\times10^{-4}$ | $\times10^{-5}$ | 4.65 | 9.304 | 11. | 13.956 | 08 | 23.260 | 27.912 | 34. | 37.216 |
| 111 d | 1/2 | 1 | $9.54\times10^{-3}$ | $\times10^{-3}$ | 21.20 | 42.410 | 53. | 63.616 | 21 | 106.027 | 127.232 | 159. | 169.642 |

* at cont field for equal ner of nuclei. ** pct of relative sens and natural abur:e.

| Isoi | | Spin | Nat Abu dan | Sensiy rel.* | bs.** | NMR-frequencHz) at a field (T) 2.3488 | 4.6975 | 5.87 | 7.0463 | S0 | 11.7440 | 14.0926 | 17.61 | 18.7900 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 113 | d | 1/2 | 12 | $1.09\times10^{-2}$ | $\times10^{-3}$ | 22.18 | 44.365 | 55.4 | 66.548 | 81 | 110.914 | 33.096 | 166.3 | 177.462 |
| 113 | | 9/2 | 4 | 0.34 | $\times10^{-2}$ | 21.86 | 43.733 | 54.6 | 65.600 | 66 | 109.333 | 31.200 | 163.9 | 174.932 |
| 115 | b | 9/2 | 95 | 0.34 | | 21.91 | 43.828 | 54.7 | 65.742 | 56 | 109.570 | 31.484 | 164.3 | 175.312 |
| 115 | n | 1/2 | 0 | $3.5\times10^{-?}$ | $\times10^{-4}$ | 32.69 | 65.399 | 81.7 | 98.099 | 139 | 163.498 | 96.198 | 245.2 | 261.598 |
| 117 | n | 1/2 | 7 | $4.52\times10^{-?}$ | $\times10^{-3}$ | 35.62 | 71.250 | 89.0 | 106.875 | 131 | 178.126 | 13.750 | 267.1 | 285.002 |
| 119 | n | 1/2 | 8 | $5.18\times10^{-2}$ | $\times10^{-3}$ | 37.27 | 74.544 | 93.1 | 111.817 | 189 | 186.362 | 23.634 | 279.5 | 298.178 |
| 121 | b | 5/2 | 57 | 0.16 | $\times10^{-2}$ | 23.93 | 47.860 | 59.8 | 71.791 | 21 | 119.652 | 43.582 | 179.4 | 191.442 |
| 123 | b | 7/2 | 42 | $4.57\times10^{-2}$ | $\times10^{-2}$ | 12.95 | 25.918 | 32.3 | 38.878 | 37 | 64.796 | 77.756 | 97.1 | 103.674 |
| 123 | e | 1/2 | C | $1.80\times10^{-2}$ | $\times10^{-4}$ | 26.20 | 52.415 | 65.5 | 78.623 | 131 | 131.039 | 57.246 | 196.5 | 209.662 |
| 125 | e | 1/2 | 6 | $3.15\times10^{-2}$ | $\times10^{-3}$ | 31.59 | 63.193 | 78.9 | 94.790 | 187 | 157.984 | 89.580 | 236.9 | 252.774 |
| 127 | | 5/2 | 10C | $9.34\times10^{-2}$ | $\times10^{-2}$ | 20.00 | 40.014 | 50.0 | 60.021 | 29 | 100.036 | 20.042 | 150.0 | 160.058 |
| 129 | e | 1/2 | 26 | $2.12\times10^{-2}$ | $\times10^{-3}$ | 27.66 | 55.321 | 69.1 | 82.981 | 142 | 138.302 | 65.962 | 207.4 | 221.284 |
| 131 | e | 3/2 | 21 | $2.76\times10^{-3}$ | $\times10^{-4}$ | 8.19 | 16.399 | 20.4 | 24.598 | 98 | 40.998 | 49.196 | 61.4 | 65.596 |
| 133 | s | 7/2 | 10C | $4.74\times10^{-2}$ | $\times10^{-2}$ | 13.11 | 26.234 | 32.7 | 39.351 | 58 | 65.585 | 78.702 | 98.3 | 104.916 |
| 135 | a | 3/2 | 6 | $4.90\times10^{-3}$ | $\times10^{-4}$ | 9.93 | 19.868 | 24.8 | 29.802 | 36 | 49.670 | 59.604 | 74.5 | 79.472 |
| 137 | a | 3/2 | 11 | $6.86\times10^{-3}$ | $\times10^{-4}$ | 11.11 | 22.226 | 27.7 | 33.339 | 52 | 55.566 | 66.678 | 83.3 | 88.904 |
| 138 | a | 5 | 0 | $9.19\times10^{-2}$ | $\times10^{-5}$ | 13.19 | 26.386 | 32.9 | 39.579 | 72 | 65.965 | 79.158 | 98.9 | 105.544 |
| 139 | a | 7/2 | 99 | $5.92\times10^{-2}$ | $\times10^{-2}$ | 14.12 | 28.252 | 35.3 | 42.378 | 04 | 70.631 | 84.756 | 105.9 | 112.808 |
| 141 | r | 5/2 | 10 | 0.29 | ) | 29.29 | 58.582 | 73.2 | 87.872 | 163 | 146.454 | 75.744 | 219.6 | 234.326 |
| 143 | d | 7/2 | 12 | $3.38\times10^{-3}$ | $\times10^{-4}$ | 5.43 | 10.875 | 13.5 | 16.313 | 50 | 27.188 | 32.626 | 40.7 | 43.500 |
| 145 | d | 7/2 | 8 | $7.86\times10^{-4}$ | $\times10^{-5}$ | 3.34 | 6.690 | 8.3 | 10.036 | 81 | 16.727 | 20.072 | 25.0 | 26.762 |
| 147 | m | 7/2 | 1 | $1.48\times10^{-3}$ | $\times10^{-4}$ | 4.12 | 8.256 | 10.3 | 12.384 | 12 | 20.640 | 24.768 | 30.9 | 33.024 |
| 149 | m | 7/2 | 1 | $7.47\times10^{-4}$ | $\times10^{-4}$ | 3.28 | 6.578 | 8.2 | 9.868 | 56 | 16.446 | 19.736 | 24.6 | 26.312 |
| 151 | u | 5/2 | 4 | 0.18 | $\times10^{-2}$ | 24.80 | 49.601 | 62.0 | 74.401 | 02 | 124.002 | 148.802 | 186.0 | 198.404 |
| 153 | u | 5/2 | 5 | $1.52\times10^{-2}$ | $\times10^{-3}$ | 10.95 | 21.903 | 27.3 | 32.854 | 05 | 54.757 | 65.708 | 82.1 | 87.610 |
| 155 | d | 3/2 | 1 | $2.79\times10^{-4}$ | $\times10^{-5}$ | 3.81 | 7.639 | 9.5 | 11.458 | 78 | 19.097 | 22.916 | 28.6 | 30.556 |
| 157 | d | 3/2 | 1 | $5.44\times10^{-4}$ | $\times10^{-5}$ | 4.77 | 9.548 | 11.9 | 14.323 | 97 | 23.871 | 28.646 | 35.8 | 38.194 |
| 159 | b | 3/2 | 10 | $5.83\times10^{-2}$ | $\times10^{-2}$ | 22.67 | 45.357 | 56.6 | 68.035 | 13 | 113.391 | 136.070 | 170.0 | 181.426 |
| 161 | y | 5/2 | 1 | $4.17\times10^{-4}$ | $\times10^{-5}$ | 3.29 | 6.588 | 8.2 | 9.883 | 77 | 16.471 | 19.766 | 24.7 | 26.354 |
| 163 | y | 5/2 | 2 | $1.12\times10^{-3}$ | $\times10^{-4}$ | 4.58 | 9.166 | 11.4 | 13.750 | 33 | 22.917 | 27.500 | 34.3 | 36.666 |
| 165 | o | 7/2 | 10 | 0.18 | 3 | 20.51 | 41.026 | 51.2 | 61.538 | 51 | 102.564 | 123.076 | 153.8 | 164.102 |
| 167 | r | 7/2 | 2 | $5.07\times10^{-4}$ | $\times10^{-4}$ | 2.89 | 5.780 | 7.2 | 8.671 | 60 | 14.451 | 17.342 | 21.6 | 23.120 |

* at cont field for equal ner of nuclei. ** pct of relative sensi and natural abun:e.

| otope | | Spin | at. bun-ance% | rel | abs.** | 88 | 4.6975 | 3719 | 7.0463 | 9.3950 | 11.) | 14.0926 | 5157 | 18.7900 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | sitivity | | NMR-freocy (MHz) at a fiel) of | | | | | | | | |
| 1( | Tm | 1/2 | 00 | 5.66× | $5.66\times10^{-4}$ | :71 | 16.543 | ).679 | 24.814 | 33.086 | 41. | 49.628 | 037 | 66.172 |
| 1: | Yb | 1/2 | 14.31 | 5.46× | $7.81\times10^{-4}$ | ;13 | 35.226 | ..032 | 52.839 | 70.452 | 88. | 105.678 | 096 | 140.904 |
| 1: | Yb | 5/2 | 16.13 | 1.33× | $2.14\times10^{-4}$ | ;52 | 9.704 | :.130 | 14.556 | 19.409 | 24. | 29.112 | 390 | 38.818 |
| 1' | Lu | 7/2 | 97.41 | 3.12× | $3.03\times10^{-2}$ | ı07 | 22.815 | ;.518 | 34.222 | 45.629 | 57. | 68.444 | 554 | 91.258 |
| 1' | Lu | 7 | 2.59 | 3.72× | $9.63\times10^{-4}$ | )28 | 15.858 | ).822 | 23.786 | 31.715 | 39. | 47.572 | 466 | 63.430 |
| 1' | Hf | 7/2 | 18.5 | 6.38× | $1.18\times10^{-4}$ | ;20 | 6.240 | '.801 | 9.361 | 12.481 | 15. | 18.722 | 403 | 24.962 |
| 1' | Hf | 9/2 | 13.75 | 2.16× | $2.97\times10^{-5}$ | 369 | 3.739 | ı.674 | 5.609 | 7.479 | 9. | 11.218 | 022 | 14.958 |
| 1: | Ta | 7/2 | 99.98 | 3.60× | $3.60\times10^{-2}$ | )70 | 23.940 | ).925 | 5.910 | 47.880 | 59. | 71.820 | 775 | 95.760 |
| 1: | W | 1/2 | 14.4 | 7.20× | $1.03\times10^{-5}$ | ı61 | 8.322 | ).402 | 12.483 | 16.644 | 20. | 24.966 | 206 | 33.288 |
| 1: | Re | 5/2 | 37.07 | 0.13 | $4.93\times10^{-2}$ | ;13 | 45.027 | ;.284 | 67.541 | 90.055 | 112. | 135.082 | 852 | 180.110 |
| 1: | Re | 5/2 | 62.93 | 0.13 | $8.62\times10^{-2}$ | '44 | 45.488 | ;.861 | 68.233 | 90.977 | 113. | 136.466 | 583 | 181.954 |
| 1: | Os | 1/2 | 1.64 | 1.22× | $2.00\times10^{-7}$ | 303 | 4.606 | ;.758 | 6.909 | 9.212 | 11. | 13.818 | 274 | 18.424 |
| 1: | Os | 3/2 | 16.1 | 2.34× | $3.76\times10^{-4}$ | '58 | 15.517 | ).397 | 23.276 | 31.035 | 38 | 46.552 | 191 | 62.070 |
| 1' | Ir | 3/2 | 37.3 | 2.53× | $9.43\times10^{-6}$ | '18 | 3.437 | ı.296 | 5.156 | 6.875 | 8 | 10.312 | 888 | 13.750 |
| 1 | Ir | 3/2 | 62.7 | 3.27× | $2.05\times10^{-5}$ | 371 | 3.743 | ı.678 | 5.614 | 7.486 | 9 | 11.228 | 034 | 14.972 |
| 1 | Pt | 1/2 | 33.8 | 9.94× | $3.36\times10^{-3}$ | ı99 | 42.998 | 3.747 | 64.497 | 85.996 | 107 | 128.994 | 241 | 171.992 |
| 1 | Au | 3/2 | 00 | 2.51× | $2.51\times10^{-5}$ | '12 | 3.425 | ı.281 | 5.138 | 6.850 | 8 | 10.276 | 843 | 13.700 |
| 1 | Hg | 1/2 | 16.84 | 5.67× | $9.54\times10^{-4}$ | 327 | 35.654 | ı.568 | 53.481 | 71.309 | 89 | 106.962 | 704 | 142.618 |
| 2 | Hg | 3/2 | 13.22 | 1.44× | $1.90\times10^{-4}$ | ;99 | 13.199 | ;.499 | 19.799 | 26.399 | 32 | 39.598 | 497 | 52.798 |
| 2 | Tl | 1/2 | 29.5 | 0.18 | $5.51\times10^{-2}$ | ı49 | 114.298 | 2.873 | 171.448 | 228.597 | 285 | 342.896 | 619 | 457.194 |
| 2 | Tl | 1/2 | 70.5 | 0.19 | 0.13 | '08 | 115.416 | ı.270 | 173.124 | 230.832 | 288 | 346.248 | 810 | 461.664 |
| 2 | Pb | 1/2 | 22.6 | 9.16× | $2.07\times10^{-3}$ | 321 | 41.843 | 2.304 | 62.765 | 83.687 | 104 | 125.530 | 912 | 167.374 |
| 2 | Bi | 9/2 | 00 | 0.13 | 0.13 | )69 | 32.139 | ).174 | 48.208 | 64.278 | 80 | 96.416 | 522 | 128.556 |
| 2 | U | 7/2 | 0.72 | 1.21× | $8.71\times10^{-7}$ | 790 | 3.580 | ı.475 | 5.371 | 7.161 | 8 | 10.742 | 425 | 14.322 |

* ;nstant field for equmber of nuclei. roduct of relative itivity and natural ıdance.

**Some representative $^{19}F$ NMR chemical shifts referenced to $CFCl_3$**

| | δ / ppm | | δ / ppm | | δ / ppm |
|---|---|---|---|---|---|
| MeF | −271.9 | $CFBr_3$ | 7.4 | $FCH{=}CH_2$ | −114 |
| $CF_2H_2$ | −1436 | $CFH_2Ph$ | −207 | $F_2C{=}CF_2$ | −135 |
| $CF_3R$ | −60 to −70 | $CF_2Cl_2$ | −8 | $C_6F_6$ | −163 |
| $AsF_5$ | −66 | $[AsF_6]^-$ | −69.5 | $[BeF_4]^-$ | −163 |
| $BF_3$ | −131 | $ClF_3$ | 116; −4 | $ClF_5$ | 247; 412 |
| $IF_7$ | 170 | $MoF_6$ | −278 | $ReF_7$ | 345 |
| $[SiF_6]^{2-}$ | −127 | $TeF_6$ | −57 | $WF_6$ | 166 |
| $XeF_2$ | 258 | $XeF_4$ | 438 | $XeF_6$ | 550 |

**Some representative $^{31}P$ NMR chemical shifts referenced to 85% $H_3PO_4$**

| | δ / ppm | | δ / ppm | | δ / ppm | | δ / ppm |
|---|---|---|---|---|---|---|---|
| $PMe_3$ | −62 | $PMeF_2$ | 245 | $Me_3PO$ | 36.2 | $Me_3PS$ | 59.1 |
| $PEt_3$ | −20 | $PMeH_2$ | −163.5 | $Et_3PO$ | 48.3 | $Et_3PS$ | 54.5 |
| $PPr^i_3$ | −33 | $PMeCl_2$ | 192 | $[Me_4P]^+$ | 24.4 | $[Et_4P]^+$ | 40.1 |
| $PBu^n_3$ | −32.3 | $PMeBr_2$ | 184 | $[PO_4]^{3-}$ | 6.0 | $[PS_4]^{3-}$ | 87 |
| $PBu^i_3$ | −45.3 | $PMe_2H$ | −99 | $PCl_5$ | −80 | $[PCl_4]^-$ | 86 |
| $PBu^s_3$ | 7.9 | $PMe_2Cl$ | 96.5 | $MePF_4$ | −29.9 | $[PCl_6]^-$ | −295 |
| $PBu^t_3$ | 63 | $PMe_2Br$ | 90.5 | $Me_3PF_2$ | −158 | $Me_2PF_3$ | 8.0 |

| Nucleus | Spin | Chemical shift range δ (ppm) | Standard |
|---|---|---|---|
| $^{1}H$ | 1/2 | 12 to −1 | $SiMe_4$ |
| $^{6}Li$ | 1 | 5 to −10 | 1M LiCl in $H_2O$ |
| $^{7}Li$ | 3/2 | 5 to −10 | 1M LiCl in $H_2O$ |
| $^{13}C$ | 1/2 | 240 to −10 | $SiMe_4$ |
| $^{15}N$ | 1/2 | 1200 to −500 | $MeNO_2$ |
| $^{17}O$ | 5/2 | 1400 to −100 | $H_2O$ |
| $^{19}F$ | 1/2 | 100 to −300 | $CFCl_3$ |
| $^{23}Na$ | 3/2 | 10 to −60 | 1M NaCl in $H_2O$ |
| $^{29}Si$ | 1/2 | 100 to −400 | $SiMe_4$ |
| $^{31}P$ | 1/2 | 230 to −200 | $H_3PO_4$ |
| $^{43}Ca$ | 7/2 | 40 to −40 | $CaCl_2$ |
| $^{51}V$ | 7/2 | 0 to −2000 | $VOCl_3$ |
| $^{67}Zn$ | 5/2 | 100 to −2700 | $ZnClO_4$ |
| $^{93}Nb$ | 9/2 | 0 to −2000 | $NbCl_6^-$ |
| $^{99}Ru$ | 3/2 | 3000 to −3000 | $RuO_4/CCl_4$ |

| Nucleus | Spin | Chemical shift range δ (ppm) | Standard |
|---|---|---|---|
| $^{119}Sn$ | 1/2 | 5000 to –3000 | $SnMe_4$ |
| $^{121}Sb$ | 5/2 | 1000 to –2700 | $Et_4NSbCl_6$ |
| $^{129}Xe$ | 1/2 | 2000 to –6000 | $XeOF_4$ |
| $^{133}Cs$ | 7/2 | 300 to –300 | CsBr |
| [illegible]Hg | 1/2 | [illegible] | $HgMe_2$* |

* highly poisonous (see Chem. & Engin. News 6/1997).

# 19. Abbreviations and acronyms used in magnetic resonance

| | |
|---|---|
| [illegible] | [illegible] |
| **ADA** | Alternated Delay Acquisition |
| **ADC** | Analog-to-Digital Converter |
| **ADLF** | Adiabatic Demagnetization in the Laboratory Frame |
| **ADRF** | Adiabatic Demagnetization in the Rotating |
| **AEE** | Average Excitation Energy approximation |
| **AJCP** | Adiabatic J Cross Polarization |
| **APT** | Attached Proton Test |
| **AQ** | Acquire |
| **ARP** | Adiabatic Rapid Passage |
| **ASTM** | American Society for Testing and Materials |
| | |
| **BB** | Broadband, as in decoupling |
| **BIRD** | Bilinear Rotation Decoupling |
| **BLEW** | A windowless multiple-pulse decoupling sequence |
| **BR-24** | Burum & Rhim (pulse sequence) |
| **BURP** | Band-selective Uniform Response Pure-phase |
| **BWR** | Bloch/Wangsness/Redfield (theory) |
| | |
| [illegible] | [illegible] |
| **CCPPA** | Coupled Cluster Polarization Propagator Approximation |
| **CH-COSY** | Carbon-Hydrogen Correlation Spectroscopy |
| **CHESS** | Chemical Shift Selective Imaging Sequence |
| **CHF** | Coupled Hartree-Fock molecular orbital |
| **CIDEP** | Chemically Induced Dynamic Electron Polarization |
| **CIDNP** | Chemically Induced Dynamic Nuclear Polarization |
| **COCONOESY (CONOESY)** | Combined COSY/NOESY |
| | Coupling |
| **COSY** | Correlated Spectroscopy |
| **COSY-45** | COSY with 45° mixing pulse |
| **COSYLR** | COSY for Long-Range couplings |
| **CP** | Cross Polarization |
| **CPD** | Composite-Pulse Decoupling |
| **CPMAS** | Cross Polarization Magic-Angle Spinning |
| **CPMG** | Carr-Purcell-Meiboom Gill Sequence |
| | Spectroscopy |
| **CRAZED** | Correlated Spectroscopy Revamped by Asymmetric Z-gradient Echo Detection |
| **CSA** | Chemical Shift Anisotropy |
| **CSCM** | Chemical Shift Correlation Map |
| **CSI** | Chemical Shift Imaging |
| **CW** | Continuous Wave |
| **CYCLOPS** | Cyclically Ordered Phase Sequence |
| **CYCLPOT** | Cyclic Polarization Transfer |
| | |
| **DAC** | Digital-to-Analog Converter |
| **DANTE** | Delay Alternating with Nutation for Tailored |
| **DAS** | Dynamic Angle Spinning |
| **DCNMR** | NMR in Presence of an Electric Direct Current |
| **DD** | Dipole-Dipole |
| **DECSY** | Double-quantum Echo Correlated Spectroscopy |
| **DEPT** | Distortionless Enhancement by Polarization Transfer |
| **DEPTH** | Spin-echo sequence for spatial localization |
| **DFT** | Discrete Fourier Transformation |
| **DIGGER** | Discrete Isolation from Gradient-Governed Elimination of Resonances |
| | of Scalar Interactions |
| **DISCO** | Differences and Sums within COSY |
| **DLB** | Differential Line Broadening |
| **DNMR** | Dynamic NMR |
| **DNP** | Dynamic Nuclear Polarization |
| **DOPT** | Dipolar Order Polarization Transfer |
| **DOUBTFUL** | Double Quantum Transition for Finding Unresolved Lines |

Abbreviations and acronyms used in magnetic resonance (continued)

| | |
|---|---|
| **DQ** | Double Quantum |
| **DQC** | Double Quantum Coherence |
| **DQF** | Double Quantum Filter |
| **DQF-COSY** | Double Quantum Filtered COSY |
| **DRESS** | Depth Resolved Spectroscopy |
| **DSA** | Data-Shift Acquisition |
| | |
| **ECOSY** | Exclusive Correlation Spectroscopy |
| **EFG** | [illegible] |
| **ELD** | Energy Level Diagram |
| **ENDOR** | Electron-Nucleus Double Resonance |
| **ENMR** | Electrophoretic NMR |
| **EOM** | Equations of Motion |
| **ESR** | Electron Spin Resonance |
| **EXORCYCLE** | 4-step phase cycle for spin echoes |
| **EXSY** | Exchange Spectroscopy |
| | |
| **FC** | Fermi Contact or Field Cycling |
| **FID** | Free Induction Decay |
| **FIRFT** | Fast Inversion-Recovery Fourier Transform |
| **FLASH** | Fast Low-angle Shot Imaging |
| **FLOPSY** | Flip-Flop Spectroscopy |
| **FOCSY** | Foldover-Corrected Spectroscopy |
| **FPT** | Finite Perturbation Theory |
| **FT** | Fourier Transform |
| **FUCOUP** | Fully Coupled Spectroscopy |
| **FWHM** | Full (line) Width at Half Maximum |
| [illegible] | Rectangular Pulse |
| **GE** | Gradient Echo |
| **GES** | Gradient-Echo Spectroscopy |
| **GIAO** | Gauge Included Atomic Orbitals |
| **GRASS**<br>**GRASP** | Gradient-Recalled Acquisition in the Steady<br>Gradient-Accelerated Spectroscopy |
| **GRECCO** | Gradient Enhanced Carbon Coupling |
| **GROPE** | Generalized compensation for Resonance Offset and Pulse Length Errors |
| **GS** | Gradient Spectroscopy |
| **H, X-COSY** | $^{1}$H, X-nucleus chemical-shift correlation spectroscopy |
| **HECTOR** | Heteronuclear Correlation Spectroscopy |
| **HEHAHA** | Heteronuclear Hartmann Hahn |
| **HMBC** | Heteronuclear Multiple-Bond Correlation |
| **HMQ** | Heteronuclear Multiquantum |
| **HNCO** | triple resonance experiments involving $^{15}$N, $^{13}$C and $^{1}$H |
| **HOESY** | Heteronuclear Overhauser Effect Spectroscopy |
| **HOHAHA** | Homonuclear Hartmann-Hahn Spectroscopy |
| **HRPA** | Higher Random Phase Approximation |
| **HSQC** | Heteronuclear Single Quantum Coherence |
| **IDESS** | Improved Depth Selective single surface coil Spectroscopy |
| **IGLO**<br>**INADEQUATE** | Individual Gauge for different Localized<br>[illegible]<br>Quantum Transfer Experiment |
| **INDO** | Intermediate Neglect of Differential Overlap |
| **INDOR** | Internuclear Double Resonance |
| **INDO/S** | Intermediate Neglect of Differential Overlap Calculations for Spectroscopy |
| **INVERSE** | H, X corelation via $^{1}$H detection |
| **IR** | Inversion-Recovery |
| **ISIS** | Image-Selected in vivo Spectroscopy |
| **IST** | Irreducible Spherical Tensor |
| | |
| **JR** | Jump-and-Return sequence ($90_y-\tau-90_{-y}$) |
| | |
| **LAS** | Laboratory Axes System |
| **LIS** | Lanthanide Induced Shift |
| **LORG** | Local Origin |
| **LP** | Linear Prediction |
| **LSR** | Lanthanide Shift Reagent |
| | |
| **MAGROFI** | Magnetization Grid Rotating-Frame Imaging |
| **MARF** | Magic Angle in the Rotating Frame |
| **MASS** | Magic-Angle Sample Spinning |
| **MEDUSA** | Technique for the Determination of Dynamic Structures |
| **MEM** | Maximum Entropy Method |
| **MINDO** | Modified INDO |
| **MP** | Multi Pulse |
| **MQ** | Multiple-Quantum |

| | |
|---|---|
| **MQC** | Multiple-Quantum Coherence |
| **MQF** | Multiple-Quantum Filter |
| **MQS** | Multi Quantum Spectroscopy |
| **MREV** | Mansfield-Rhim-Elleman-Vaughan sequence |
| **MRI** | Magnetic Resonance Imaging |
| **MRS** | Magnetic Resonance Spectroscopy |
| **MRSI** | Magnetic Resonance Spectroscopic Imaging |
| **MSPGSE** | Multiple-Stepped PGSE |
| | |
| **NOE** | Nuclear Overhauser Effect |
| **NOESY** | Nuclear Overhauser Effect Spectroscopy |
| **NQCC** | Nuclear Quadrupole Coupling Constant |
| **NQR** | Nuclear Quadrupole Resonance |
| | |
| **OSIRIS** | modification of ISIS |
| | |
| **PAR** | Phase-alternated Rotation of magnetization |
| **PAS** | Principal Axes System |
| **P.COSY** | Purged COSY |
| **PENDANT** | Polarization Enhancement During Attached Nucleus Testing |
| **PFG** | Pulsed Field Gradient |
| **PGSE** | Pulsed Gradient Spin Echo |
| **PMFG** | Pulsed Magnetic Field Gradient |
| **ppm** | Parts per million |
| **PRE** | Proton Relaxation Enhancement |
| **PRESS** | Point-Resolved Spectroscopy |
| **PRFT** | Partially Relaxed Fourier Transform |
| **PSD** | Phase-sensitive Detection |
| | |
| **QF** | Quadrupole moment/Field gradient (interaction or relaxation mechanism) |
| **QPD** | Quadrature Phase Detection |
| | |
| **RARE** | Rapid Acquisition Relaxation Enhanced |
| **RE-BURP** | Refocused Band selective Uniform Response Pure phase |
| **RECSY** | Multistep Relayed Coherence Spectroscopy |
| **REDOR** | Rotational Echo Double Resonance |
| **RELAY** | Relayed Correlation Spectroscopy |
| **REX** | orbital theory |
| **RF** | Radio Frequency |
| **RIDE** | Ring Down Elimination |
| **RODI** | Rotating-frame relaxation Dispersion Imaging |
| **ROESY** | Rotating Frame Overhauser Effect Spectroscopy |
| **RPA** | Random Phase Approximation |
| | |
| **SA** | Shielding Anisotropy |
| **SC** | Scalar Coupling |
| **SCPT** | Self Consistent Perturbation Theory |
| **SD** | Spin Dipolar |
| **SE** | Spin Echo |
| **SECSY** | Spin-Echo Correlated Spectroscopy |
| **SEDOR** | Spin-Echo Double Resonance |
| **SEDUCE** | Selective Decoupling Using Crafted Excitation |
| [illegible] | Spin-Echo Fourier Transform Spectroscopy |
| **SELINCOR** | Selective Inverse Correlation |
| **SELINQUATE** | Selective INADEQUATE |
| **SELRESOLV** | Selective Resolution of C,H Coupling |
| **SEMUT** | Subspectral Editing Using a Multiple-Quantum Trap |
| **SFORD** | Single Frequency Off-Resonance Decoupling |
| **SGSE** | Steady-Gradient Spin-Echo |
| **SKEWSY** | Skewed Exchange Spectroscopy |
| **SL** | Spin-Lock pulse |
| **SLITDRESS** | Slice interleaved Depth Resolved Surface coil-Spectroscopy |
| **SNR or S/N** | Signal-to-noise Ratio |
| **SOPPA** | Second Order Polarization Propagator Approach |
| **SPACE** | Spatial and Chemical-Shift Encoded Excitation |
| **SPT** | Selective Population Transfer |
| **SQC** | Single-Quantum Coherence |
| **SQF** | Single-Quantum Filter |
| **SR** | Saturation-Recovery |
| **SSFP** | Steady-State Free Precession |
| **STE** | Stimulated Echo |
| **STEAM** | Stimulated Echo Acquisition Mode for imaging |
| | |
| **TANGO** | Testing for Adjacent Nuclei with a Gyration Operator |
| **TART** | Tip Angle Reduced $T_1$ Imaging |
| **TE** | Time delay between excitation and Echo maximum |

## Abbreviations and acronyms used in magnetic resonance (continued)

| | |
|---|---|
| **TMR** | Topical Magnetic Resonance |
| **TOCSY** | Total Correlation Spectroscopy |
| **TOE** | Truncated NOE |
| **TORO** | TOCSY-ROESY Relay |
| **TPPI** | Time-proportional Phase Incrementation |
| **TQ** | Triple Quantum |
| **TQF** | Triple-Quantum Filter |
| **TR** | Time for Repetition of excitation |
| **TRCF** | Tilted Rotating Coordinate Frame |
| **UE** | Unpaired Electron (relaxation mechanism) |
| | |
| **VAS** | Variable Angle Spinning |
| **VOSING** | Volume-selective Spectral editing |
| **VOSY** | Volume-Selective Spectroscopy |
| **WAHUHA** | Waugh-Huber-Haeberlen Sequence |
| **WALTZ** | CPD Sequence Containing the Elements 1-2-3 |
| **WATERGATE** | Water suppression pulse sequence |
| **WEFT** | Water Eliminated Fourier Transform |
| | |
| **XCORFE** | H, X Correlation using a Fixed Evolution time are coupled to X |
| **X-HALF-FILTER** | Selection of $^1$H-$^1$H correlation when one H is coupled to X |
| | |
| **Z-COSY** | COSY with z-Filter |
| **ZECSY** | pulse sandwich for elimination of signal Zero Quantum Echo-Correlation Spectroscopy |
| **ZQ** | Zero-Quantum |
| **ZQC** | Zero-Quantum Coherence |
| **ZQF** | Zero-Quantum Filter |
| **ZZ-Spectros-** | Selection of coherences involving ZZ or |
| **$\beta$-COSY** | COSY with small flip angle mixing pulse |
| **$\Psi$-COSY** | psuedo-COSY using incremented frequency-selective excitation |

## 20. Symbols used in magnetic resonance

| | |
|---|---|
| $B$ | magnetic-flux density |
| $B_{loc}$ | local magnetic-flux density contribution by secular interaction |
| $B_0$<br>$\eta$ | stationary main magnetic-flux density in $z$<br>asymmetry parameter (or anisotropy constant) or nuclear Overhauser enhancement factor |
| $\delta$ | chemical shift (ppm) |
| $G$ | gradient of the main magnetic-flux density |
| $\gamma_n$ | gyromagnetic ratio of nucleus n |
| $J$ | (indirect) spin-spin coupling constant (in Hz) |
| $M$ | magnetization in the laboratory frame |
| $M'_x, M'_y, M'_z$ | magnetization components in the rotating frame along the $x'$, $y'$ and $z'$ axes, respectively |
| $M_0$ | equilibrium (Curie) magnetization |
| $\mu_0$ | magnetic dipole moment |
| $\mu_B$ | Bohr magneton |
| $\mu_N$ | nuclear magneton |
| $\omega_L$ | Larmor (angular) frequency |
| $\omega_m$ | modulation angular frequency (in rad $s^{-1}$) |
| $\omega_r$ | sample rotation frequency (rad $s^{-1}$) |
| $\omega_1$ | Larmor (angular) frequency in $B_1$ |
| $Q$ | nuclear quadrupole moment |
| $p$ | density operator |
| $R_1$ | spin-lattice relaxation rate |
| $R_2$ | transverse relaxation rate |
| $\sigma$ | chemical shift shielding tensor |
| $T_1$ | longitudinal (spin-lattice) relaxation time for $M_z$ |
| $T_2$ | transverse (spin-spin) relaxation time for $M_{xy}$ |
| $T_2^*$<br>$T_{2e}$ | time constant of the FID in the presence of $B_0$ inhomogeneities<br>multiple-pulse irradiation |
| $T_{2\rho}$ | transverse relaxation time during a spin-lock pulse |
| $t_1$, $t_2$ | time domains in multi-dimensional experiments |
| $T_{1\rho}$ | $T_1$ of spin-locked magn. in rotation frame |
| $T_E$, $t_e$ | echo time |
| $t_n$ | time domain of the n-th dimension |
| $F_n$ | frequency domain of the n-th dimension |
| $t_m$ | mixing time |
| $T_R$, $t_r$ | repetition time |
| $\tau_{col}$ | mean time between molecular collision in the liquid state |

| | |
|---|---|
| $\tau_i$ | angular momentum correlation time |
| $\tau, \tau_1, \tau_2, ...$ | intervals in pulse sequences |
| $W_0, W_1, W_2$ | transition probabilities for zero-, single-, and double-quantum transitions |
| $\Omega$ | frequency offset |
| $\chi$ | magnetic susceptibility |
| $Y_{l,m}$ | spherical harmonics |

# 21. EPR/ENDOR frequency table

| Z | Isotope* | Natural Abundance | Spin | ENDOR Frequency | $\mu$ | $g_n\beta_n$ | Electric Quadrupole Moment Q $\lvert e \rvert \times 10^{-24}$ cm² |
|---|---|---|---|---|---|---|---|
| 0 | $n^{1*}$ | ... | 1/2 | 10.20769 | −3.82608422 | $1.040674 \times 10^{-3}$ | ... |
| 1 | $H^1$ | 99.985 | 1/2 | 14.90218 | 5.5856912 | $1.519278 \times 10^{-3}$ | ... |
| | $H^2$ | 0.0148 | 1 | 2.287575 | 0.8574376 | $2.332185 \times 10^{-4}$ | 0.002875 |
| | $H^{3*}$ | ... | 1/2 | 15.89525 | 5.957920 | $1.62052 \times 10^{-3}$ | ... |
| 3 | $Li^6$ | [illegible] | [illegible] | 2.193107 | 0.8220514 | $2.233937 \times 10^{-4}$ | −0.000644 |
| | $Li^7$ | 92.5 | 3/2 | 5.791950 | 2.170961 | $5.904900 \times 10^{-4}$ | −0.040 |
| 4 | $Be^9$ | 100. | 3/2 | 2.094 | −0.7850 | $2.135 \times 10^{-4}$ | 0.053 |
| 5 | $B^{10}$ | 19.8 | 3 | 1.60133 | 0.600216 | $1.63256 \times 10^{-4}$ | 0.08608 |
| | $B^{11}$ | 80.2 | 3/2 | 4.782043 | 1.792424 | $4.875299 \times 10^{-4}$ | 0.040 |
| 6 | $C^{13}$ | 1.11 | 1/2 | 3.74705[illegible] | 1.40482[illegible] | $3.80104 \times 10^{-4}$ | [illegible] |
| | $N^{15}$ | 0.366 | 1/2 | 1.511052 | −0.5663784 | $1.540519 \times 10^{-4}$ | ... |
| 8 | $O^{17}$ | 0.038 | 5/2 | 2.02099 | −0.757516 | $2.06040 \times 10^{-4}$ | −0.026 |
| 9 | $F^{19}$ | 100. | 1/2 | 14.02721 | 5.257732 | $1.430075 \times 10^{-3}$ | ... |
| 10 | $Ne^{21}$ | 0.27 | 3/2 | 1.17708 | −0.441197 | $1.20003 \times 10^{-4}$ | 0.1029 |
| 11 | $Na^{22*}$ | | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 12 | $Mg^{25}$ | 10.00 | 5/2 | 0.91291 | −0.34218 | $9.3071 \times 10^{-5}$ | 0.22[v] |
| 13 | $Al^{27}$ | 100. | 5/2 | 3.886094 | 1.456601 | $3.961878 \times 10^{-4}$ | 0.150 |
| 14 | $Si^{29}$ | 4.67 | 1/2 | 2.9630 | −1.1106 | $3.0208 \times 10^{-4}$ | ... |
| 15 | $P^{31}$ | 100. | 1/2 | 6.03804 | 2.26320 | $6.15579 \times 10^{-4}$ | ... |
| 17 | $Cl^{35}$ | 75.77 | 3/2 | 1.461795 | 0.5479157 | $1.490302 \times 10^{-4}$ | −0.08249 |
| | $Cl^{36*}$ | ... | 2 | 1.71477 | 0.642735 | $1.74820 \times 10^{-4}$ | −0.0180 |
| | $Cl^{37}$ | 24.23 | 3/2 | 1.216790 | 0.4560820 | $1.240519 \times 10^{-4}$ | −0.06493 |
| 18 | $Ar^{39*}$ | ... | 7/2 | 0.99 | −0.37 | $1.0 \times 10^{-4}$ | ? |
| 19 | $K^{39}$ | 93.26 | 3/2 | 0.6963030 | 0.2609909 | $7.098816 \times 10^{-5}$ | 0.054 |
| | $K^{40*}$ | 0.0117 | 4 | 0.86582 | −0.32453 | $8.8270 \times 10^{-5}$ | −0.067 |
| 20 | $Ca^{41*}$ | ... | 7/2 | 1.215641 | −0.4556514 | $1.239348 \times 10^{-4}$ | ? |
| | $Ca^{43}$ | 0.135 | 7/2 | 1.00424 | −0.376414 | $1.02983 \times 10^{-4}$ | <0.23[v] |
| 21 | $Sc^{45}$ | 100. | 7/2 | 3.62586 | 1.35906 | $3.69657 \times 10^{-4}$ | −0.22[u] |
| 22 | $Ti^{47}$ | 7.4 | 5/2 | 0.84144 | −0.31539 | $8.5784 \times 10^{-5}$ | 0.29[u] |
| | $Ti^{49}$ | 5.4 | 7/2 | 0.841667 | −0.315477 | $8.58081 \times 10^{-5}$ | 0.24[u] |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 24 | $Cr^{53}$ | 9.50 | 3/2 | 0.8396 | −0.3147 | $8.560 \times 10^{-5}$ | −0.0285/+0.022 |

| Z | Isotope[a] | Natural Abundance percent | Spin | ENDOR Frequency in MHz, for 3.5 kG Field | $g_n$ | $\frac{g_n\beta_n}{g_e\beta_e}$ | Quadrupole Moment Q in multiples of $\lvert e\rvert \times 10^{-24}$ cm² |
|---|---|---|---|---|---|---|---|
| 25 | $Mn^{53*}$ | ... | 7/2 | 3.828 | 1.435 | $3.903 \times 10^{-4}$ | ? |
| | $Mn^{55}$ | 100. | 5/2 | 3.6868 | 1.3819 | $3.7587 \times 10^{-4}$ | $0.33^u$ |
| 27 | $Co^{59}$ | 100. | 7/2 | 3.516 | 1.318 | $3.585 \times 10^{-4}$ | 0.42 |
| | $Co^{60*}$ | ... | 5 | 2.025 | 0.7589 | $2.064 \times 10^{-4}$ | $0.44^u$ |
| 28 | $Ni^{61}$ | 1.13 | 3/2 | 1.3340 | −0.50001 | $1.3600 \times 10^{-4}$ | 0.162 |
| 29 | $Cu^{63}$ | 69.2 | 3/2 | 3.959 | 1.484 | $4.036 \times 10^{-4}$ | −0.222 |
| | $Cu^{65}$ | 30.8 | 3/2 | 4.237 | 1.588 | $4.319 \times 10^{-4}$ | −0.195 |
| 31 | $Ga^{69}$ | 60.1 | 3/2 | 3.58673 | 1.34439 | $3.65667 \times 10^{-4}$ | 0.168 |
| | $Ga^{71}$ | 39.9 | 3/2 | 4.55729 | 1.70818 | $4.64616 \times 10^{-4}$ | 0.106 |
| 32 | $Ge^{73}$ | 7.8 | 9/2 | 0.5214100 | −0.1954371 | $5.315787 \times 10^{-5}$ | $-0.19^J$ |
| 33 | $As^{75}$ | 100. | 3/2 | 2.56026 | 0.959647 | $2.61019 \times 10^{-4}$ | $0.29^u$ |
| 34 | $Se^{77}$ | 7.6 | 1/2 | 2.8528 | 1.0693 | $2.9084 \times 10^{-4}$ | ... |
| 35 | $Br^{79}$ | 50.69 | 3/2 | 3.746469 | 1.404266 | $3.819530 \times 10^{-4}$ | $0.293^u$ |
| | $Br^{81}$ | 49.31 | 3/2 | 4.038446 | 1.513706 | $4.117201 \times 10^{-4}$ | 0.27 |
| 36 | $Kr^{83}$ | 11.5 | 9/2 | 0.575481 | −0.215704 | $5.86704 \times 10^{-5}$ | $0.260^u$ |
| | $Kr^{85*}$ | ... | 9/2 | 0.5957 | 0.2233 | $6.074 \times 10^{-5}$ | $0.45^u$ |
| 37 | $Rb^{85}$ | 72.17 | 5/2 | 1.44402 | 0.541253 | $1.47218 \times 10^{-4}$ | 0.273 |
| | $Rb^{87}$ | 27.83 | 3/2 | 4.89369 | 1.83427 | $4.98912 \times 10^{-4}$ | 0.130 |
| 39 | $Y^{89}$ | 100. | 1/2 | 0.7332410 | −0.2748361 | $7.475398 \times 10^{-5}$ | ... |
| 40 | $Zr^{91}$ | 11.2 | 5/2 | 1.39118 | −0.521448 | $1.41831 \times 10^{-4}$ | ? |
| 41 | $Nb^{93}$ | 100. | 9/2 | 3.6583 | 1.3712 | $3.7296 \times 10^{-4}$ | $-0.28^u$ |
| 42 | $Mo^{95}$ | 15.9 | 5/2 | 0.9754 | −0.3656 | $9.944 \times 10^{-5}$ | $-0.019^u$ |
| [illegible] | $Mo^{97}$ | [illegible] | 5/2 | [illegible] | [illegible] | [illegible] | [illegible] |
| 44 | $Ru^{99}$ | 12.7 | 5/2 | 0.664 | −0.249 | $6.77 \times 10^{-5}$ | $0.076^u$ |
| | $Ru^{101}$ | 17.0 | 5/2 | 0.744 | −0.279 | $7.59 \times 10^{-5}$ | $0.44^u$ |
| 45 | $Rh^{102*}$ | ... | (6) | 1.83 | 0.685 | $1.86 \times 10^{-4}$ | ? |
| | $Rh^{103}$ | 100. | 1/2 | 0.4717 | −0.1768 | $4.809 \times 10^{-5}$ | ... |
| 46 | $Pd^{105}$ | 22.2 | 5/2 | 0.683 | −0.256 | $6.96 \times 10^{-5}$ | $0.66^u$ |
| | $Ag^{109}$ | 48.17 | 1/2 | 0.698309 | −0.261743 | $7.11927 \times 10^{-5}$ | ... |
| 48 | $Cd^{111}$ | 12.8 | 1/2 | 3.17597 | −1.19043 | $3.23791 \times 10^{-4}$ | ... |
| | $Cd^{113*}$ | 12.2 | 1/2 | 3.3226 | −1.2454 | $3.3874 \times 10^{-4}$ | ... |
| 49 | $In^{113}$ | 4.3 | 9/2 | 3.27791 | 1.22864 | $3.34184 \times 10^{-4}$ | 0.846 |
| | $In^{115*}$ | 95.7 | 9/2 | 3.28498 | 1.23129 | $3.34904 \times 10^{-4}$ | 0.861 |
| | $Sn^{117}$ | 7.75 | 1/2 | 5.34139 | −2.00208 | $5.44555 \times 10^{-4}$ | ... |
| | $Sn^{119}$ | 8.6 | 1/2 | 5.58812 | −2.09456 | $5.69709 \times 10^{-4}$ | ... |
| 51 | $Sb^{121}$ | 57.3 | 5/2 | 3.5897 | 1.3455 | $3.6597 \times 10^{-4}$ | $-0.33^u$ |
| | $Sb^{123}$ | 42.7 | 7/2 | 1.9443 | 0.72876 | $1.9822 \times 10^{-4}$ | −0.68 |
| | $Sb^{125*}$ | ... | 7/2 | 2.005 | 0.7514 | $2.044 \times 10^{-4}$ | ? |
| 52 | $Te^{123}$ | 0.89 | 1/2 | 3.9314 | −1.4736 | $4.0081 \times 10^{-4}$ | ... |
| 53 | $I^{127}$ | 100. | 5/2 | 3.00221 | 1.12530 | $3.06076 \times 10^{-4}$ | $-0.789^v$ |
| | $I^{129*}$ | ... | 7/2 | 1.9979 | 0.74886 | $2.0369 \times 10^{-4}$ | $-0.553^v$ |

| Z | Isotope[a] | Natural Abundance percent | Spin | ENDOR Frequency in MHz, for 3.5 kG Field | $g_n$ | $\frac{g_n\beta_n}{g_e\beta_e}$ | Electric Quadrupole Moment Q in multiples of $\lvert e \rvert \times 10^{-24}$ cm$^2$ |
|---|---|---|---|---|---|---|---|
| 54 | $Xe^{129}$ | 26.4 | 1/2 | 4.15115 | −1.55595 | $4.23210 \times 10^{-4}$ | ... |
| | $Xe^{131}$ | 21.2 | 3/2 | 1.23055 | 0.461240 | $1.25455 \times 10^{-4}$ | −0.120[u] |
| | $Cs^{134*}$ | ... | 4 | 1.9967 | 0.74842 | $2.0357 \times 10^{-4}$ | 0.389 |
| | $Cs^{135*}$ | ... | 7/2 | 2.0828 | 0.78069 | $2.1234 \times 10^{-4}$ | 0.051 |
| | $Cs^{137*}$ | ... | 7/2 | 2.1658 | 0.81180 | $2.2081 \times 10^{-4}$ | 0.052 |
| 56 | $Ba^{133*}$ | ... | 1/2 | 4.11 | −1.54 | $4.19 \times 10^{-4}$ | ... |
| | $Ba^{135}$ | 6.59 | 3/2 | 1.4909 | 0.55884 | $1.5200 \times 10^{-4}$ | 0.20[u] |
| | $Ba^{137}$ | 11.2 | 3/2 | 1.6679 | 0.62515 | $1.7004 \times 10^{-4}$ | 0.34[u] |
| | $La^{138*}$ | 0.089 | 5 | 1.9817 | 0.74278 | $2.0203 \times 10^{-4}$ | 0.51 |
| | $La^{139}$ | 99.911 | 7/2 | 2.1215 | 0.79520 | $2.1629 \times 10^{-4}$ | 0.20 |
| 59 | $Pr^{141}$ | 100. | 5/2 | 4.3 | 1.6 | $4.4 \times 10^{-4}$ | −0.041[J] |
| 60 | $Nd^{143}$ | 12.2 | 7/2 | 0.8207 | −0.3076 | $8.367 \times 10^{-5}$ | −0.56 |
| | $Nd^{145}$ | 8.3 | 7/2 | 0.507 | −0.190 | $5.17 \times 10^{-5}$ | −0.29 |
| 62 | $Sm^{147*}$ | 15.1 | 7/2 | 0.6195 | −0.2322 | $6.316 \times 10^{-5}$ | −0.18[L] |
| | $Sm^{149}$ | 13.9 | 7/2 | 0.5109 | 0.1915 | $5.209 \times 10^{-5}$ | 0.056[u] |
| | $Sm^{151*}$ | ... | 5/2 | 0.379 | 0.142 | $3.86 \times 10^{-5}$ | 0.52[v] |
| 63 | $Eu^{151}$ | 47.9 | 5/2 | 3.706 | 1.389 | $3.778 \times 10^{-4}$ | 1.53[u] |
| | $Eu^{152*}$ | ... | 3 | 1.7265 | 0.64713 | $1.7602 \times 10^{-4}$ | 3.16[u] |
| | $Eu^{153}$ | 52.1 | 5/2 | 1.637 | 0.6134 | $1.668 \times 10^{-4}$ | 3.92[u] |
| | Eu | ... | 5/2 | 2.00 | 0.772 | $2.10 \times 10^{-4}$ | ? |
| 64 | $Gd^{155}$ | 14.8 | 3/2 | 0.4597 | −0.1723 | $4.686 \times 10^{-5}$ | 1.30[u] |
| | $Gd^{157}$ | 15.7 | 3/2 | 0.6011 | −0.2253 | $6.128 \times 10^{-5}$ | 1.34 |
| 65 | $Tb^{157*}$ | ... | 3/2 | 3.5 | 1.3 | $3.5 \times 10^{-4}$ | ? |
| | $Tb^{158*}$ | ... | 3 | 1.563 | 0.5860 | $1.594 \times 10^{-4}$ | 2.7 |
| | $Tb^{159}$ | 100. | 3/2 | 3.580 | 1.342 | $3.650 \times 10^{-4}$ | 1.34 |
| | $Dy^{163}$ | 2.49 | 5/2 | 0.710 | 0.266 | $7.24 \times 10^{-5}$ | 2.51 |
| 67 | $Ho^{165}$ | 100. | 7/2 | 3.180 | 1.192 | $3.242 \times 10^{-4}$ | 2.73[u] |
| 68 | $Er^{167}$ | 22.9 | 7/2 | 0.4317 | −0.1618 | $4.401 \times 10^{-5}$ | 2.827[u] |
| 69 | $Tm^{169}$ | 100. | 1/2 | 1.24 | −0.466 | $1.27 \times 10^{-4}$ | ... |
| | $Tm^{171*}$ | ... | 1/2 | 1.229 | 0.4606 | $1.253 \times 10^{-4}$ | ... |
| | Yb | 16.2 | 5/2 | 0.72554 | −0.27195 | $7.3969 \times 10^{-5}$ | 2.8 |
| 71 | $Lu^{173*}$ | ... | 7/2 | 1.78 | 0.669 | $1.82 \times 10^{-4}$ | ? |
| | $Lu^{174*}$ | ... | (1) | 5.18 | 1.94 | $5.28 \times 10^{-4}$ | ? |
| | $Lu^{175}$ | 97.39 | 7/2 | 1.7059 | 0.63943 | $1.7392 \times 10^{-4}$ | 5.68[u] |
| | $Lu^{176*}$ | 2.61 | 7 | 1.21 | 0.452 | $1.23 \times 10^{-4}$ | 8.0[u] |
| 72 | $Hf^{177}$ | 18.6 | 7/2 | 0.6048 | 0.2267 | $6.166 \times 10^{-5}$ | 4.5 |
| 73 | $Ta^{181}$ | 99.9877 | 7/2 | 1.8070 | 0.67729 | $1.8422 \times 10^{-4}$ | 3.44[u] |
| 74 | $W^{183}$ | 14.3 | 1/2 | 0.6284800 | 0.2355694 | $6.407365 \times 10^{-5}$ | ... |
| 75 | $Re^{185}$ | 37.40 | 5/2 | 3.4011 | 1.2748 | $3.4674 \times 10^{-4}$ | 2.33[u] |
| | $Re^{187*}$ | 62.60 | 5/2 | 3.4357 | 1.2878 | $3.5027 \times 10^{-4}$ | 2.22[u] |
| 76 | $Os^{187}$ | 1.6 | 1/2 | 0.3498 | 0.1311 | $3.566 \times 10^{-5}$ | ... |
| 77 | $Ir^{191}$ | 37.3 | 3/2 | 0.259 | 0.097 | $2.64 \times 10^{-5}$ | 0.78[u] |
| | $Ir^{193}$ | 62.7 | 3/2 | 0.285 | 0.107 | $2.91 \times 10^{-5}$ | 0.70[u] |

| Z | Isotope[a] | Natural Abundance percent | Spin | ENDOR Frequency in MHz, for 3.5 kG Field | $g_n$ | $\frac{g_n\beta_n}{g_e\beta_e}$ | Electric Quadrupole Moment Q in multiples of $\lvert e\rvert \times 10^{-24}$ cm$^2$ |
|---|---|---|---|---|---|---|---|
| 78 | $Pt^{195}$ | 33.8 | 1/2 | 3.2523.. | 1.2190.. | $8.3156\times10^{-4}$ | [illegible] |
| 80 | $Hg^{199}$ | 16.8 | 1/2 | 2.699321 | 1.011770 | $2.751961\times10^{-4}$ | ... |
| | $Hg^{201}$ | 13.2 | 3/2 | 0.996423 | −0.373483 | $1.01585\times10^{-4}$ | 0.42 |
| 81 | $Tl^{203}$ | 29.5 | 1/2 | 8.656103 | 3.244514 | $8.824909\times10^{-4}$ | ... |
| | $Tl^{204*}$ | ... | 2 | 0.1211 | 0.0454 | $1.235\times10^{-5}$ | ? |
| | $Tl^{205}$ | 70.5 | 1/2 | 8.7385 | 3.2754 | $8.9089\times10^{-4}$ | ... |
| 83 | $Bi^{207*}$ | ... | 9/2 | 2.59 | 0.970 | $2.64\times10^{-4}$ | −0.50[v] |
| | $Bi^{209}$ | 100. | 9/2 | 2.50 | 0.938 | $2.55\times10^{-4}$ | −0.46[v] |
| 84 | $Po^{209*}$ | ... | 1/2 | 4.0 | 1.5 | $4.1\times10^{-4}$ | ... |
| 89 | $Ac^{227*}$ | ... | 3/2 | 1.9 | 0.73 | $2.0\times10^{-4}$ | 1.7[v] |
| 90 | $Th^{229*}$ | ... | 5/2 | 0.43 | 0.16 | $4.4\times10^{-5}$ | 4.4[v] |
| 92 | $U^{233*}$ | ... | 5/2 | 0.69 | 0.26 | $7.1 \times 10^{-5}$ | 3.5/7.9[v] |
| | $U^{235*}$ | 0.720 | 7/2 | 0.27 | −0.10 | $2.7 \times 10^{-5}$ | 4.3[v] |
| 93 | $Np^{237*}$ | ... | 5/2 | 3.2 | 1.2 | $3.3 \times 10^{-4}$ | 4.5[v] |
| 94 | $Pu^{239*}$ | ... | 1/2 | 1.08 | 0.406 | $1.10 \times 10^{-4}$ | ... |
| | $Pu^{241*}$ | ... | 5/2 | 0.755 | −0.283 | $7.70 \times 10^{-5}$ | 5.6[v] |
| | $Am^{243*}$ | ... | 5/2 | 1.72 | 0.644 | $1.75 \times 10^{-4}$ | 4.9[v] |
| 96 | $Cm^{243*}$ | ... | 5/2 | 0.43 | 0.16 | $4.4 \times 10^{-5}$ | ? |
| | $Cm^{245*}$ | ... | 7/2 | 0.3 | 0.1 | $3 \times 10^{-5}$ | ? |
| | $Cm^{247*}$ | ... | 9/2 | 0.21 | 0.08 | $2.2 \times 10^{-5}$ | ? |

[a] Stable isotopes and those with half-lives >1 yr.
* Means radioactive
[v] Means unclear whether the correction was made

Table by courtesy of Prof. J. A. Weil
and R. S. Rao, Dept. of Chemistry
Saskatoon, Canada S 7N OWO

## 22. Some useful conversion factors in EPR

0.07144775 and divided by the $g$ factor, and 2) a hyperfine coupling constant $A$ expressed in the units cm$^{-1}$ is multiplied by $2.9979246 \times 10^4$ to convert it to MHz.

| | | | | |
|---|---|---|---|---|
| $\nu_e$ (GHz) | $= 13.9962\ g\ B$ (T)<br>$= 1.39962\ g\ B$ (kG) | | $A$ (MHz) | $= 2.997\,9246 \times 10^4 A$ (cm$^{-1}$)<br>$= 13.9962\ g\ A$ (mT)<br>$= 1.399\,62\ g\ A$ (G) |
| | | | $A$ (cm$^{-1}$) | $= 0.333\,564\,10 \times 10^{-4}\ A$ (MHz) |
| $\nu_p$(MHz) | $= 42.576\,38\ B$ (T) | ($H_2O$) | 1 mT | = 10 G |
| $B$ (T) | $= 0.023\,487\,20\ \nu_p$ (MHz) | ($H_2O$) | | |
| $g$ | $= 0.071\,447\,75 \frac{\nu_e\text{(GHz)}}{B\text{(T)}}$ | | | |

**Atomic weights of the isotopes of all elements based on $^{12}$C = 12.00000000**

| Isotope | Exact Mass | Abundance (%) | Isotope | Exact Mass | Abundance (%) |
|---|---|---|---|---|---|
| $^{1}$H | 1.00782504 | 100.0000 | $^{46}$Ca | 45.9536890 | .0041 |
| D | 2.01410179 | 100.0000 | $^{45}$Sc | 44.9559136 | 100.0000 |
| $^{3}$He | 3.01602929 | .0001 | $^{46}$Ti | 45.9526330 | 10.8401 |
| $^{4}$He | 4.00260325 | 100.0000 | $^{47}$Ti | 46.9517650 | 9.8916 |
| $^{6}$Li | 6.01512320 | 8.1081 | $^{48}$Ti | 47.9479470 | 100.0000 |
| $^{7}$Li | 7.01600450 | 100.0000 | $^{49}$Ti | 48.9478710 | 7.4526 |
| $^{10}$B | 10.0129380 | 24.8439 | $^{50}$V | 49.9471610 | .2510 |
| $^{11}$B | 11.0093053 | 100.0000 | $^{51}$V | 50.9439630 | 100.0000 |
| $^{12}$C | 12.0000000 | 100.0000 | $^{50}$Cr | 49.9460460 | 5.1915 |
| $^{13}$C | 13.0033548 | 1.1122 | $^{52}$Cr | 51.9405100 | 100.0000 |
| $^{14}$N | 14.0030740 | 100.0000 | $^{53}$Cr | 52.9406510 | 11.3379 |
| $^{16}$O | 15.9949146 | 100.0000 | $^{55}$Mn | 54.9380460 | 100.0000 |
| $^{17}$O | 16.9991306 | .0381 | $^{54}$Fe | 53.9396120 | 6.3236 |
| $^{18}$O | 17.9991594 | .2005 | $^{56}$Fe | 55.9349390 | 100.0000 |
| $^{19}$F | 18.9984033 | 100.0000 | $^{57}$Fe | 56.9353960 | 2.3986 |
| $^{20}$Ne | 19.9924391 | 100.0000 | $^{58}$Fe | 57.9332780 | .3053 |
| $^{22}$Ne | 21.9913837 | 10.1867 | $^{60}$Ni | 59.9307890 | 38.2305 |
| $^{23}$Na | 22.9897697 | 100.0000 | $^{61}$Ni | 60.9310590 | 1.6552 |
| $^{24}$Mg | 23.9850450 | 100.0000 | $^{62}$Ni | 61.9283460 | 5.2585 |
| $^{25}$Mg | 24.9858392 | 12.6598 | $^{64}$Ni | 63.9279680 | 1.3329 |
| $^{26}$Mg | 25.9825954 | 13.9380 | $^{59}$Co | 58.9331980 | 100.0000 |
| $^{28}$Si | 27.9769284 | 100.0000 | $^{65}$Cu | 64.9277921 | 44.5710 |
| $^{29}$Si | 28.9764964 | 5.0634 | $^{64}$Zn | 63.9291450 | 100.0000 |
| $^{30}$Si | 29.9737717 | 3.3612 | $^{66}$Zn | 65.9260350 | 57.4074 |
| $^{31}$P | 30.9737634 | 100.0000 | $^{67}$Zn | 66.9271290 | 8.4362 |
| $^{32}$S | 31.9720718 | 100.0000 | $^{68}$Zn | 67.9248460 | 38.6831 |
| $^{34}$S | 33.9678677 | 4.4306 | $^{69}$Ga | 68.9255810 | 100.0000 |
| $^{36}$S | 35.9670790 | .0220 | $^{71}$Ga | 70.9247010 | 66.3890 |
| $^{35}$Cl | 34.9688527 | 100.0000 | $^{70}$Ge | 69.9242500 | 56.1644 |
| $^{37}$Cl | 36.9659026 | 31.9780 | $^{72}$Ge | 71.9220800 | 75.0685 |
| $^{39}$K | 38.9637079 | 100.0000 | $^{73}$Ge | 72.9234640 | 21.3698 |
| $^{41}$K | 40.9618254 | 7.2170 | $^{76}$Ge | 75.9214030 | 21.3698 |
| $^{36}$Ar | 35.9675456 | .3380 | $^{75}$As | 74.9215960 | 100.0000 |
| $^{38}$Ar | 37.9627322 | .0630 | $^{79}$Br | 78.9183360 | 100.0000 |
| $^{40}$Ar | 39.9623831 | 100.0000 | $^{81}$Br | 80.9162900 | 97.2776 |
| $^{40}$Ca | 39.9625907 | 100.0000 | $^{74}$Se | 73.9224770 | 1.8145 |
| $^{43}$Ca | 42.9587704 | .1393 | $^{77}$Se | 76.9199080 | 15.3226 |
| $^{44}$Ca | 43.9554848 | 2.1518 | $^{78}$Se | 77.9173040 | 47.3790 |

| | | |
|---|---|---|
| $^{80}Se$ | 79.9165210 | 100.0000 |
| $^{82}Se$ | 81.9167090 | 18.9516 |
| $^{78}Kr$ | 77.9203970 | .6140 |
| $^{80}Kr$ | 79.9163750 | 3.9474 |
| $^{82}Kr$ | 81.9134830 | 20.3509 |
| $^{84}Kr$ | 83.9115060 | 100.0000 |
| $^{86}Kr$ | 85.9106140 | 30.3509 |
| $^{85}Rb$ | 84.9117996 | 100.0000 |
| $^{87}Rb$ | 86.9091840 | 38.5710 |
| $^{84}Sr$ | 83.9134280 | .6781 |
| $^{87}Sr$ | 86.9088900 | 8.4766 |
| $^{88}Sr$ | 87.9056250 | 100.0000 |
| $^{89}Y$ | 88.9058560 | 100.0000 |
| $^{90}Zr$ | 89.9047080 | 100.0000 |
| $^{91}Zr$ | 90.9056440 | 21.9048 |
| $^{94}Zr$ | 93.9063190 | 33.6832 |
| $^{96}Zr$ | 95.9082720 | 5.4033 |
| $^{93}Nb$ | 92.9063780 | 100.0000 |
| $^{92}Mo$ | 91.9068090 | 61.5002 |
| $^{94}Mo$ | 93.9050860 | 38.3340 |
| $^{96}Mo$ | 95.9046760 | 69.1256 |
| $^{97}Mo$ | 96.9060180 | 39.5773 |
| $^{98}Mo$ | 97.9054050 | 100.0000 |
| $^{100}Mo$ | 99.9074730 | 39.9088 |
| $^{96}Ru$ | 95.9075960 | 17.4684 |
| $^{99}Ru$ | 98.9059370 | 40.1899 |
| $^{100}Ru$ | 99.9042180 | 39.8734 |
| $^{101}Ru$ | 100.905581 | 53.7975 |
| $^{102}Ru$ | 101.904348 | 100.0000 |
| $^{104}Ru$ | 103.905422 | 59.1772 |
| $^{102}Pd$ | 101.905609 | 3.7322 |
| $^{104}Pd$ | 103.904026 | 40.7611 |
| $^{105}Pd$ | 104.905075 | 81.7051 |
| $^{106}Pd$ | 105.903475 | 100.0000 |
| $^{108}Pd$ | 107.903894 | 96.8167 |
| $^{107}Ag$ | 106.905095 | 100.0000 |
| $^{109}Ag$ | 108.904754 | 92.9050 |
| $^{106}Cd$ | 105.906461 | 4.3508 |
| $^{108}Cd$ | 107.904186 | 3.0978 |
| $^{110}Cd$ | 109.903007 | 43.4737 |
| $^{112}Cd$ | 111.902761 | 83.9888 |
| $^{113}Cd$ | 112.904401 | 42.5339 |

| | | |
|---|---|---|
| $^{114}Cd$ | 113.903361 | 100.0000 |
| $^{116}Cd$ | 115.904758 | 26.0703 |
| $^{113}In$ | 112.904056 | 4.4932 |
| $^{115}In$ | 114.903875 | 100.0000 |
| $^{112}Sn$ | 111.904823 | 3.0864 |
| $^{115}Sn$ | 114.903344 | 1.2346 |
| $^{116}Sn$ | 115.901744 | 45.3704 |
| $^{117}Sn$ | 116.902954 | 23.7654 |
| $^{118}Sn$ | 117.901607 | 75.0000 |
| $^{119}Sn$ | 118.903310 | 26.5432 |
| $^{122}Sn$ | 121.903440 | 14.1975 |
| $^{124}Sn$ | 123.905271 | 17.2840 |
| $^{121}Sb$ | 120.903824 | 100.0000 |
| $^{123}Sb$ | 122.904222 | 74.5201 |
| $^{127}I$ | 126.904477 | 100.0000 |
| $^{122}Te$ | 121.903055 | 7.6923 |
| $^{123}Te$ | 122.904278 | 2.6864 |
| $^{124}Te$ | 123.902825 | 14.2485 |
| $^{125}Te$ | 124.904435 | 21.1243 |
| $^{126}Te$ | 125.903310 | 56.0651 |
| $^{130}Te$ | 129.906229 | 100.0000 |
| $^{124}Xe$ | 123.906120 | .3717 |
| $^{126}Xe$ | 125.904281 | .3346 |
| $^{128}Xe$ | 127.903531 | 7.1004 |
| $^{129}Xe$ | 128.904780 | 98.1413 |
| $^{131}Xe$ | 130.905080 | 78.8104 |
| $^{132}Xe$ | 131.904148 | 100.0000 |
| $^{134}Xe$ | 133.905395 | 38.6617 |
| $^{136}Xe$ | 135.907219 | 33.0855 |
| $^{133}Cs$ | 132.905433 | 100.0000 |
| $^{132}Ba$ | 131.905042 | .1409 |
| $^{134}Ba$ | 133.904490 | 3.3710 |
| $^{135}Ba$ | 134.905668 | 9.1939 |
| $^{136}Ba$ | 135.904556 | 10.9540 |
| $^{137}Ba$ | 136.905816 | 15.6625 |
| $^{138}La$ | 137.907114 | .0901 |
| $^{139}La$ | 138.906355 | 100.0000 |
| $^{136}Ce$ | 135.907140 | .2147 |
| $^{138}Ce$ | 137.905996 | .2825 |
| $^{140}Ce$ | 139.905442 | 100.0000 |
| $^{141}Pr$ | 140.907657 | 100.0000 |
| $^{142}Nd$ | 141.907731 | 100.0000 |

| Isotope | Exact Mass | Abundance (%) | Isotope | Exact Mass | Abundance (%) |
|---|---|---|---|---|---|
| $^{143}$Nd | 142.909823 | 44.8949 | $^{176}$Hf | 175.941420 | 14.7727 |
| $^{144}$Nd | 143.910096 | 87.7258 | $^{177}$Hf | 176.943233 | 52.8409 |
| $^{145}$Nd | 144.912582 | 30.5934 | $^{178}$Hf | 177.943710 | 76.9886 |
| $^{146}$Nd | 145.913126 | 63.3616 | $^{179}$Hf | 178.945827 | 39.0341 |
| $^{150}$Nd | 149.920900 | 20.7888 | $^{180}$Ta | 179.947489 | .0120 |
| $^{144}$Sm | 143.912003 | 11.6540 | $^{181}$Ta | 180.948014 | 100.0000 |
| $^{147}$Sm | 146.914907 | 56.7670 | $^{180}$W | 179.946727 | .4239 |
| $^{148}$Sm | 147.914832 | 42.4810 | $^{182}$W | 181.948225 | 85.7515 |
| $^{149}$Sm | 148.917193 | 52.2560 | $^{183}$W | 182.950245 | 46.6254 |
| $^{152}$Sm | 151.919741 | 100.0000 | $^{186}$W | 185.954377 | 93.2507 |
| $^{154}$Sm | 153.922218 | 84.9620 | $^{185}$Re | 184.952977 | 59.7444 |
| $^{151}$Eu | 150.919860 | 91.5709 | $^{187}$Re | 186.955765 | 100.0000 |
| $^{153}$Eu | 152.921243 | 100.0000 | $^{184}$Os | 183.952514 | .0488 |
| $^{152}$Gd | 151.919803 | .8052 | $^{186}$Os | 185.953852 | 3.8537 |
| $^{155}$Gd | 154.922629 | 59.5813 | $^{188}$Os | 187.955850 | 32.4390 |
| $^{156}$Gd | 155.922130 | 82.4074 | $^{189}$Os | 188.958156 | 39.2683 |
| $^{157}$Gd | 156.923967 | 63.0032 | $^{190}$Os | 189.958455 | 64.3902 |
| $^{158}$Gd | 157.924111 | 100.0000 | $^{192}$Os | 191.961487 | 100.0000 |
| $^{160}$Gd | 159.927061 | 88.0032 | $^{191}$Ir | 190.960603 | 59.4896 |
| $^{156}$Dy | 155.924287 | .2128 | $^{190}$Pt | 189.959937 | .0296 |
| $^{158}$Dy | 157.924412 | .3546 | $^{192}$Pt | 191.961049 | 2.3373 |
| $^{160}$Dy | 159.925203 | 8.2979 | $^{194}$Pt | 193.962679 | 97.3373 |
| $^{161}$Dy | 160.926939 | 67.0212 | $^{195}$Pt | 194.964785 | 100.0000 |
| $^{162}$Dy | 161.926805 | 90.4255 | $^{196}$Pt | 195.964947 | 74.8521 |
| $^{164}$Dy | 163.929183 | 100.0000 | $^{197}$Au | 196.966560 | 100.0000 |
| $^{165}$Ho | 164.930332 | 100.0000 | $^{196}$Hg | 195.965812 | .5059 |
| $^{162}$Er | 161.928787 | .4167 | $^{198}$Hg | 197.966760 | 34.0641 |
| $^{164}$Er | 163.929211 | 4.7917 | $^{199}$Hg | 198.968269 | 57.3356 |
| $^{166}$Er | 165.930305 | 100.0000 | $^{200}$Hg | 199.968316 | 77.9089 |
| $^{168}$Er | 167.932383 | 79.7619 | $^{202}$Hg | 201.970632 | 100.0000 |
| $^{170}$Er | 169.935476 | 44.3452 | $^{204}$Hg | 203.973481 | 22.9342 |
| $^{169}$Tm | 168.934225 | 100.0000 | $^{203}$Tl | 202.972336 | 41.8922 |
| $^{168}$Yb | 167.933908 | .4088 | $^{205}$Tl | 204.974410 | 100.0000 |
| $^{170}$Yb | 169.934774 | 9.5912 | $^{204}$Pb | 203.973037 | 2.6718 |
| $^{172}$Yb | 171.936393 | 68.8679 | $^{207}$Pb | 206.975885 | 42.1756 |
| $^{173}$Yb | 172.938222 | 50.6918 | $^{208}$Pb | 207.976641 | 100.0000 |
| $^{174}$Yb | 173.938873 | 100.0000 | $^{209}$Bi | 208.980388 | 100.0000 |
| $^{176}$Yb | 175.942576 | 39.9371 | $^{232}$Th | 232.038054 | 100.0000 |
| $^{175}$Lu | 174.940785 | 100.0000 | $^{234}$U | 234.040947 | .0055 |
| $^{174}$Hf | 173.940065 | .4545 | $^{238}$U | 238.050785 | 100.0000 |

Table by courtesy of Prof. J. Seibl (ETH Zürich, Switzerland)

# 24. Conversion table of transmittance and absorbance units

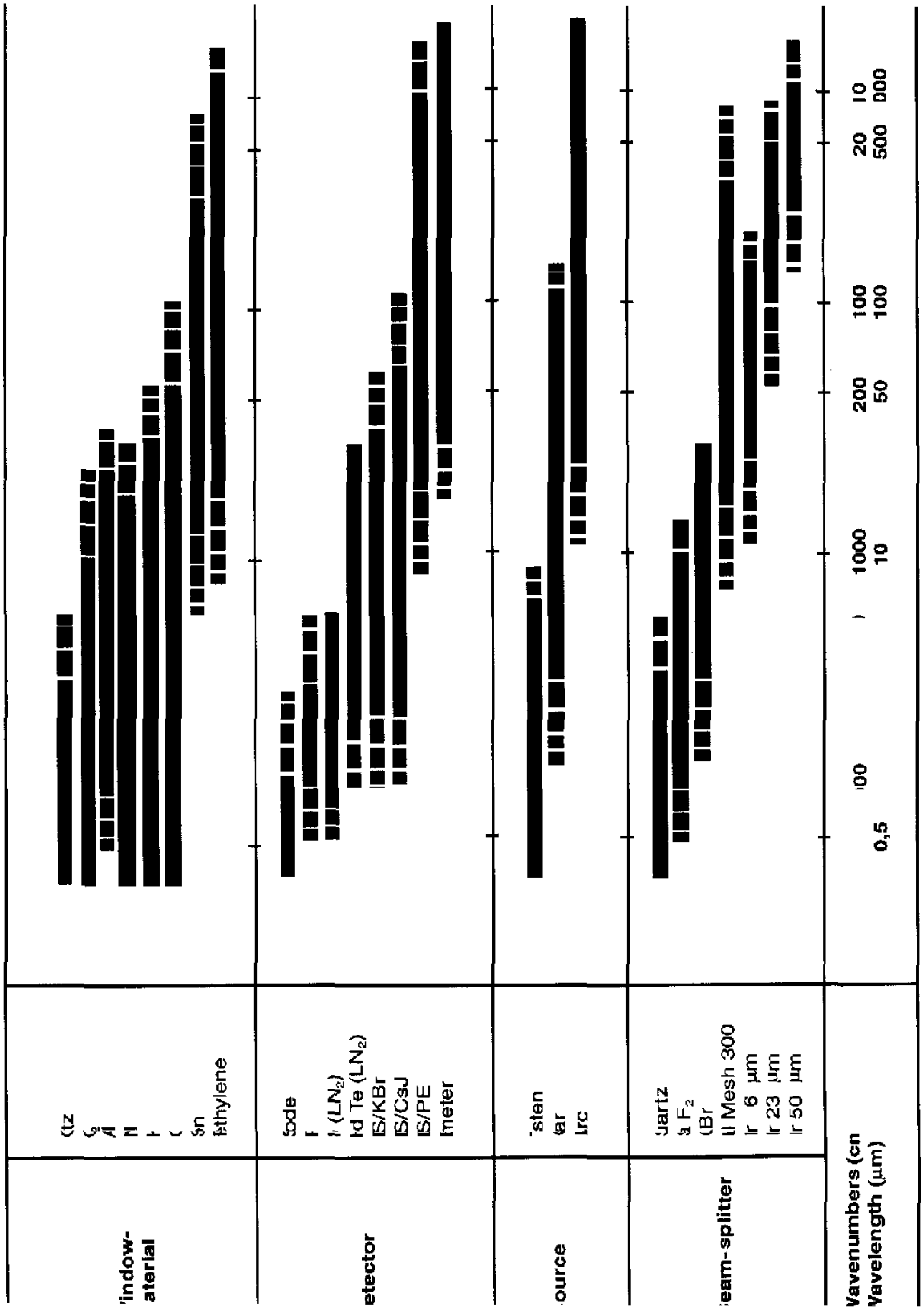
| Transmittance [%] | Absorbance | Transmittance [%] | Absorbance |
|---|---|---|---|
| 1.0 | 2.000 | 51.0 | .292 |
| 2.0 | 1.699 | 52.0 | .284 |
| 4.0 | 1.398 | 54.0 | [illegible] |
| 5.0 | 1.301 | 55.0 | .260 |
| 6.0 | 1.222 | 56.0 | .265 |
| 7.0 | 1.155 | 57.0 | .244 |
| 8.0 | 1.097 | 58.0 | .237 |
| 9.0 | 1.046 | 59.0 | .229 |
| 11.0 | .959 | 61.0 | .215 |
| 12.0 | .921 | 62.0 | .208 |
| 13.0 | .886 | 63.0 | .201 |
| 14.0 | .854 | 64.0 | .194 |
| 15.0 | .824 | 65.0 | .187 |
| 17.0 | .770 | 67.0 | [illegible] |
| 18.0 | .745 | 68.0 | .167 |
| 19.0 | .721 | 69.0 | .161 |
| 20.0 | .699 | 70.0 | .155 |
| 21.0 | .678 | 71.0 | .149 |
| 22.0 | .658 | 72.0 | .143 |
| 24.0 | .620 | 74.0 | .131 |
| 25.0 | .602 | 75.0 | .125 |
| 26.0 | .585 | 76.0 | .119 |
| 27.0 | .569 | 77.0 | .114 |
| 28.0 | .553 | 78.0 | .108 |
| 30.0 | .523 | 80.0 | [illegible] |
| 31.0 | .509 | 81.0 | .092 |
| 32.0 | .495 | 82.0 | .086 |
| 33.0 | .481 | 83.0 | .081 |
| 34.0 | .469 | 84.0 | .076 |
| 35.0 | .456 | 85.0 | .071 |
| 37.0 | .432 | 87.0 | .060 |
| 38.0 | .420 | 88.0 | .056 |
| 39.0 | .409 | 89.0 | .051 |
| 40.0 | .398 | 90.0 | .046 |
| 41.0 | .387 | 91.0 | .041 |
| 43.0 | .367 | 93.0 | [illegible] |
| 44.0 | .357 | 94.0 | .027 |
| 45.0 | .347 | 95.0 | .022 |
| 46.0 | .337 | 96.0 | .018 |
| 47.0 | .328 | 97.0 | .013 |
| 48.0 | .319 | 98.0 | .009 |
| 50.0 | .301 | 100.0 | .000 |

| Wavenumber [cm⁻¹] | Wavelength [Micron] | Wavelength [nm] | Frequency [GHz] | Electron Volt [eV] |
|---|---|---|---|---|
| 2.0 | 5 000.00 | 5 000 000 | 60 | .00 025 |
| 4.0 | 2 500.00 | 2 500 000 | 120 | .00 050 |
| 8.0 | 1 250.00 | 1 250 000 | 240 | .00 099 |
| 10.0 | 1 000.00 | 1 000 000 | 300 | .00 124 |
| 12.0 | 833.33 | 833 333 | 360 | .00 149 |
| 14.0 | 714.29 | 714 286 | 420 | .00 174 |
| 16.0 | 625.00 | 625 000 | 480 | .00 198 |
| 20.0 | 500.00 | 500 000 | 600 | .00 248 |
| 22.0 | 454.55 | 454 545 | 660 | .00 273 |
| 24.0 | 416.57 | 416 667 | 719 | .00 298 |
| 26.0 | 384.62 | 384 615 | 779 | .00 322 |
| 28.0 | 357.14 | 357 143 | 839 | .00 347 |
| 30.0 | 333.33 | 333 333 | 899 | .00 372 |
| 34.0 | 294.12 | 294 118 | 1 019 | .00 422 |
| 36.0 | 277.78 | 277 778 | 1 079 | .00 446 |
| 38.0 | 263.16 | 263 158 | 1 139 | .00 471 |
| 40.0 | 250.00 | 250 000 | 1 199 | .00 496 |
| 50.0 | 200.00 | 200 000 | 1 499 | .00 620 |
| 70.0 | 142.86 | 142 857 | 2 099 | .00 868 |
| 80.0 | 125.00 | 125 000 | 2 398 | .00 992 |
| 90.0 | 111.11 | 111 111 | 2 698 | .01 116 |
| 100.0 | 100.00 | 100 000 | 2 998 | .01 240 |
| 110.0 | 90.91 | 90 909 | 3 298 | .01 364 |
| 130.0 | 76.92 | 76 923 | 3 897 | .01 612 |
| 140.0 | 71.43 | 71 429 | 4 197 | .01 736 |
| 150.0 | 66.67 | 66 667 | 4 497 | .01 860 |
| 160.0 | 62.50 | 62 500 | 4 797 | .01 984 |
| 170.0 | 58.82 | 58 824 | 5 096 | .02 108 |
| 190.0 | 52.63 | 52 632 | 5 696 | .02 356 |
| 200.0 | 50.00 | 50 000 | 5 996 | .02 480 |
| 220.0 | 45.45 | 45 455 | 6 595 | .02 728 |
| 240.0 | 41.67 | 41 667 | 7 195 | .02 976 |
| 260.0 | 38.46 | 38 462 | 7 795 | .03 224 |
| 300.0 | 33.33 | 33 333 | 8 994 | .03 720 |
| 320.0 | 31.25 | 31 250 | 9 593 | .03 967 |
| 340.0 | 29.41 | 29 412 | 10 193 | .04 215 |
| 360.0 | 27.78 | 27 778 | 10 792 | .04 463 |
| 380.0 | 26.32 | 26 316 | 11 392 | .04 711 |
| 500.0 | 20.00 | 20 000 | 14 990 | .06 199 |
| 600.0 | 16.67 | 16 667 | 17 987 | .07 439 |

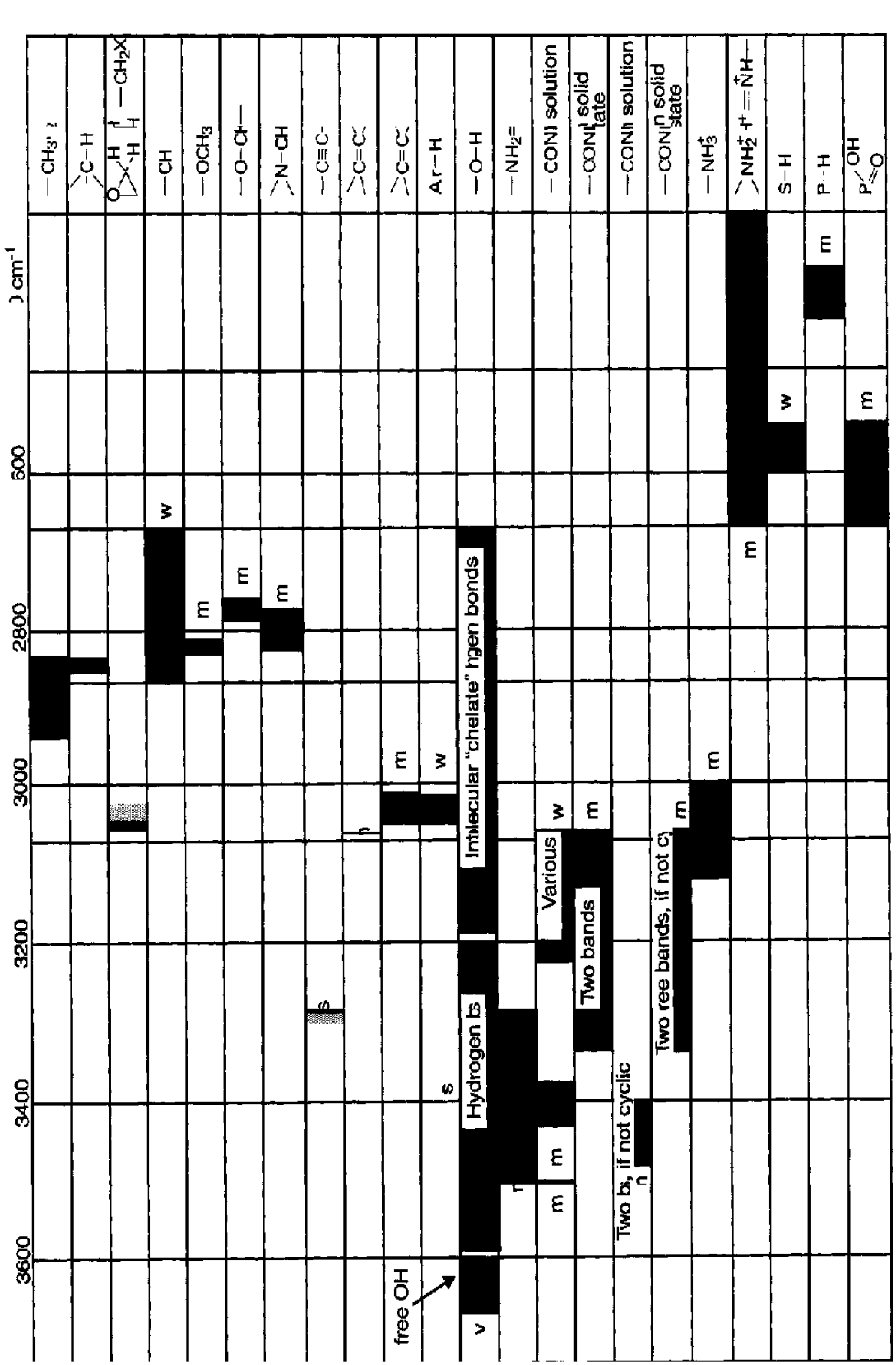
Conversion table of energy and wavelength units (continued)

| Wavenumber [cm$^{-1}$] | Wavelength [Micron] | Wavelength [nm] | Frequency [GHz] | Electron Volt [eV] |
|---|---|---|---|---|
| 700.0 | 14.29 | 14 286 | 20 985 | .08 679 |
| 800.0 | 12.50 | 12 500 | 23 983 | .09 919 |
| 1 000.0 | 10.00 | 10 000 | 29 979 | .12 398 |
| 1 100.0 | 9.09 | 9 091 | 32 977 | .13 638 |
| 1 200.0 | 8.33 | 8 333 | 35 975 | .14 878 |
| 1 300.0 | 7.69 | 7 692 | 38 973 | .16 118 |
| 1 400.0 | 7.14 | 7 143 | 41 971 | .17 358 |
| 1 500.0 | 6.67 | 6 667 | 44 968 | .18 598 |
| 1 700.0 | 5.88 | 5 882 | 50 964 | .21 077 |
| 1 800.0 | 5.56 | 5 556 | 53 962 | .22 317 |
| 1 900.0 | 5.26 | 5 263 | 56 960 | .23 557 |
| 2 000.0 | 5.00 | 5 000 | 59 958 | .24 797 |
| 2 200.0 | 4.55 | 4 545 | 65 954 | .27 276 |
| 2 600.0 | 3.85 | 3 846 | 77 945 | .32 236 |
| 2 800.0 | 3.57 | 3 571 | 83 941 | .34 716 |
| 3 000.0 | 3.33 | 3 333 | 89 937 | .37 195 |
| 3 200.0 | 3.13 | 3 125 | 95 933 | .39 675 |
| 3 400.0 | 2.94 | 2 941 | 101 929 | .42 155 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 4 000.0 | 2.50 | 2 500 | 119 916 | .49 594 |
| 5 000.0 | 2.00 | 2 000 | 149 895 | .61 992 |
| 6 000.0 | 1.67 | 1 667 | 179 874 | .74 390 |
| 7 000.0 | 1.43 | 1 429 | 209 853 | .86 789 |
| 8 000.0 | 1.25 | 1 250 | 239 832 | .99 187 |
| 10 000.0 | 1.00 | 1 000 | 299 790 | 1.23 984 |
| 11 000.0 | .91 | 909 | 329 769 | 1.36 382 |
| 12 000.0 | .83 | 833 | 359 748 | 1.48 781 |
| 13 000.0 | .77 | 769 | 389 727 | 1.61 179 |
| 14 000.0 | .71 | 714 | 419 706 | 1.73 578 |
| 16 000.0 | .62 | 625 | 479 664 | 1.98 374 |
| 17 000.0 | .59 | 588 | 509 643 | 2.10 773 |
| 18 000.0 | .56 | 556 | 539 622 | 2.23 171 |
| 19 000.0 | .53 | 526 | 569 601 | 2.35 570 |
| 20 000.0 | .50 | 500 | 599 580 | 2.47 968 |
| 24 000.0 | .42 | 417 | [illegible] | [illegible] |
| 26 000.0 | .38 | 385 | 779 454 | 3.22 358 |
| 28 000.0 | .36 | 357 | 839 412 | 3.47 155 |
| 30 000.0 | .33 | 333 | 899 370 | 3.71 952 |
| 32 000.0 | .31 | 312 | 959 328 | 3.96 749 |
| 34 000.0 | .29 | 294 | 1 019 286 | 4.21 546 |
| 38 000.0 | .26 | 263 | 1 139 202 | 4.71 139 |
| 40 000.0 | .25 | 250 | 1 199 160 | 4.95 936 |

# 26. Optical components used in FT-IR-Spectroscopy

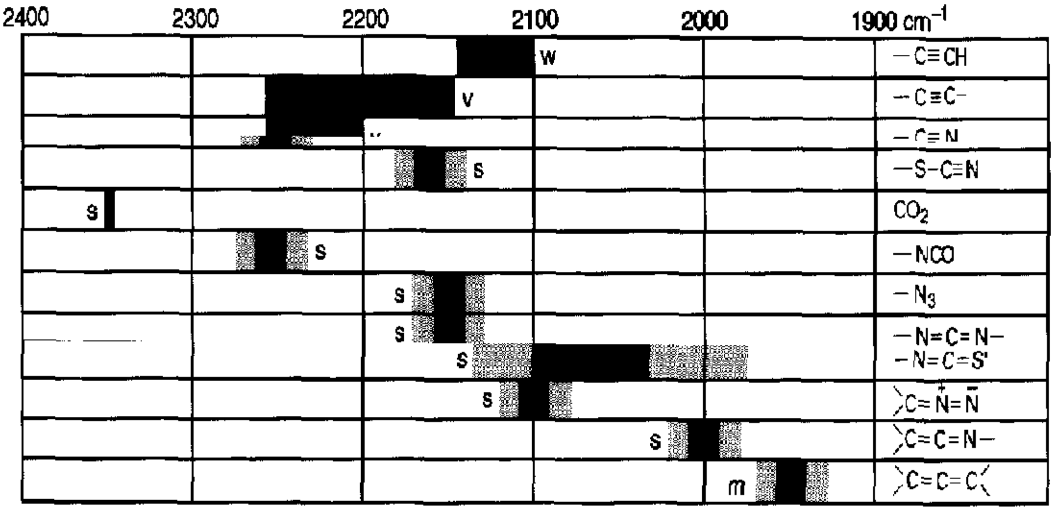


27. Infrared and Raman tables

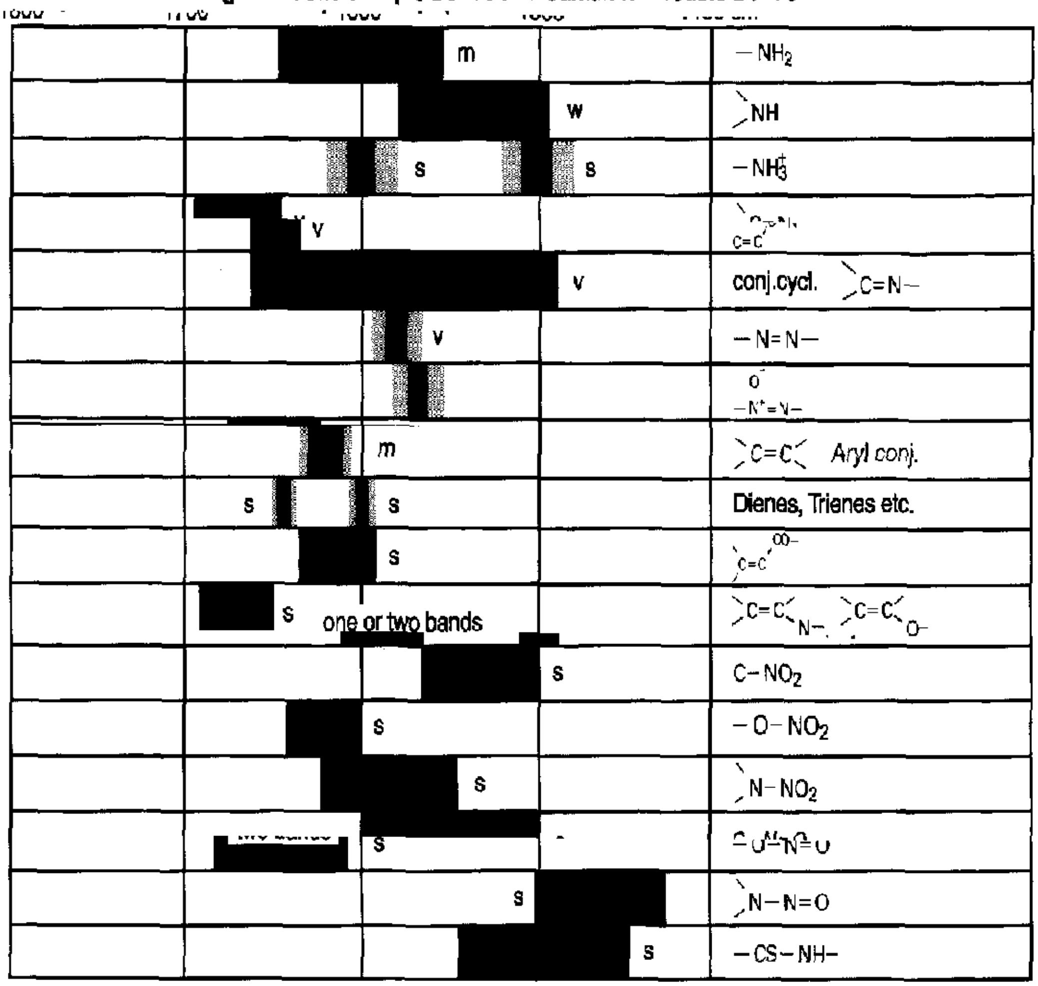



**Positions of Stretg Vibrations of Hgen** (in the hatchenges the boundarie not well defined)
Band intensity: s = ıg, m = medium, weak, v = varying.

## Infrared and Raman tables (continued)

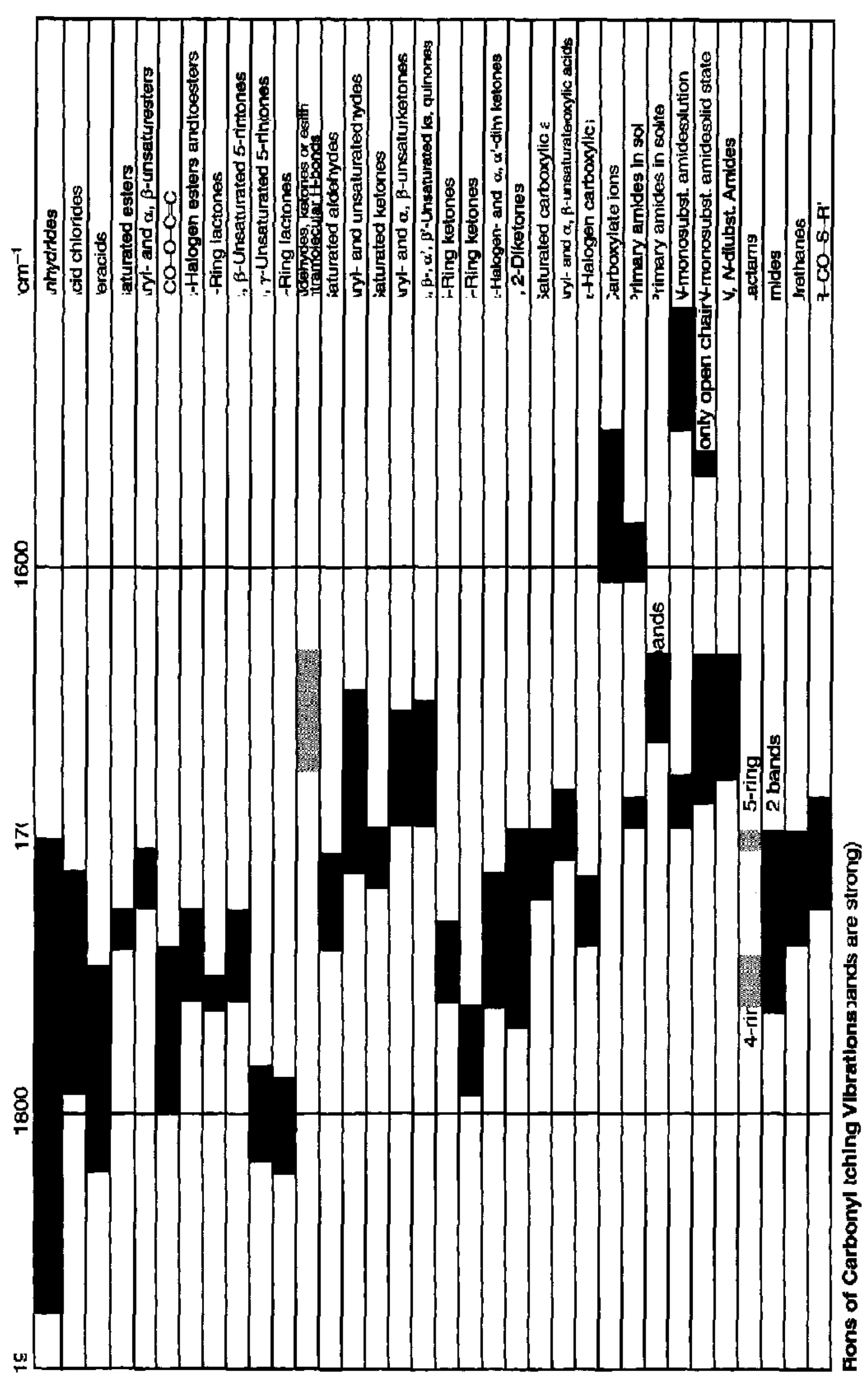



**Positions of Stretching Vibrations of Triple Bonds and Cumulated Double Bonds**

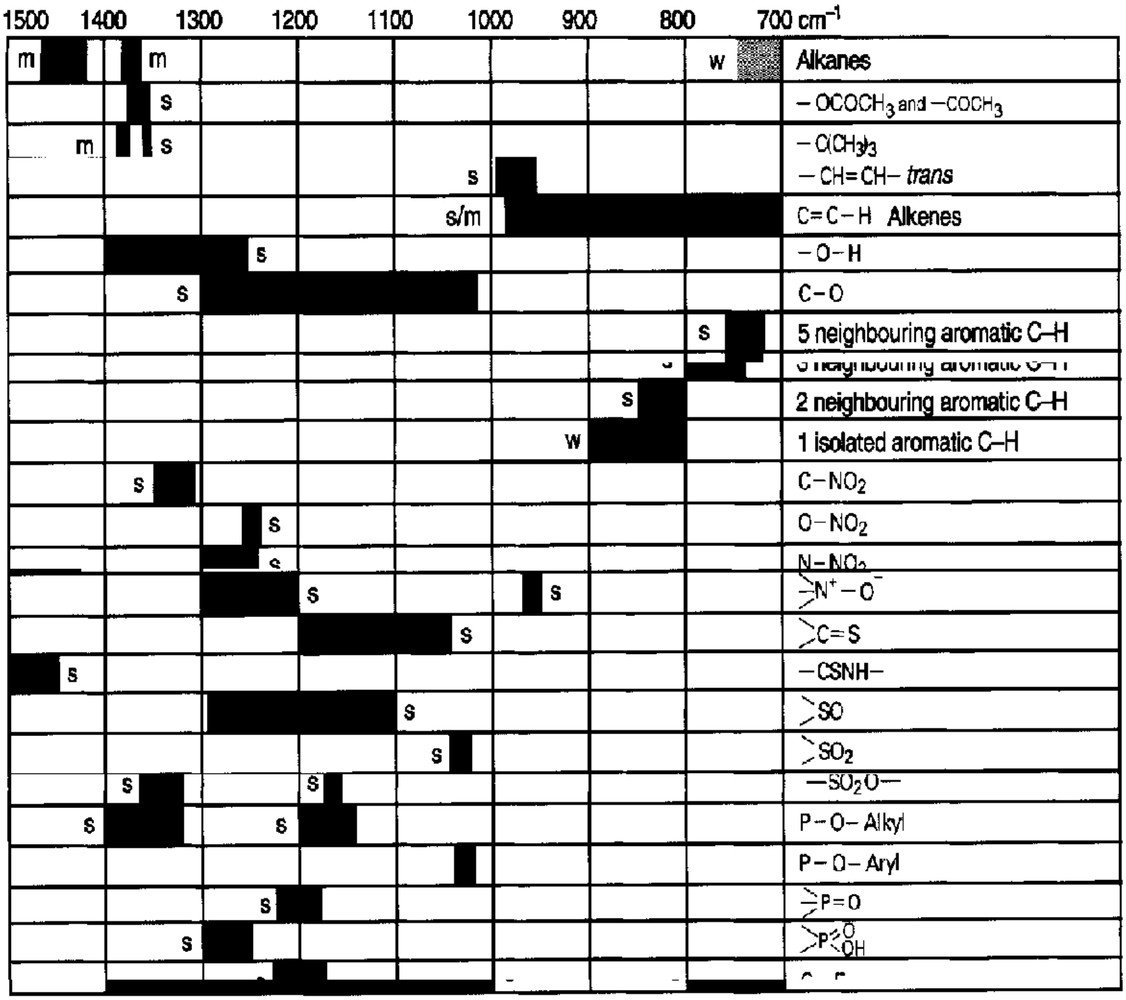



**Positions of the Double Bond Stretching Vibrations and N-H Bending Vibrations**
(s = strong, m = medium, w = weak, v = varying)

## Infrared and Raman tables (continued)

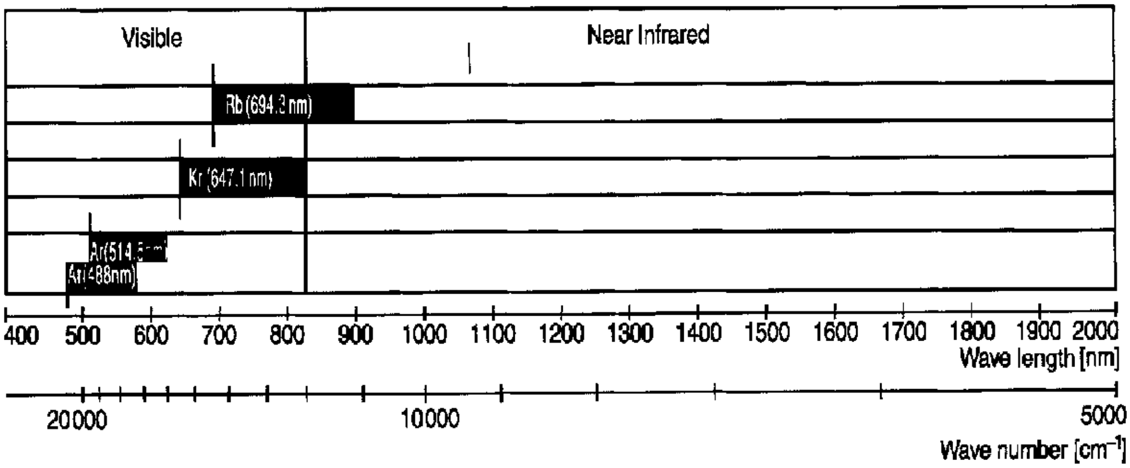


Infrared and Raman tables (continued)




**Characteristic Absorptions in the Fingerprint Region** (s = strong, m = medium, w = weak)

## Stokes shifts (0–3500 cm$^{-1}$) of various laser sources



## 28. Selected force constants and bond orders (according to Siebert) of organic and inorganic compounds

| Bond A-B | Force Const. f (N cm$^{-1}$) | Bond Order | Compound |
|---|---|---|---|
| H–H | 5.14 | 0.77 | $H_2$ |
| B–B | 3.58 | 1.2 | $B_2$ |
| C–C | 16.5 | 3.2 | HCCH |
| N–N | 22.42 | 3.2 | $N_2$ |
| O–O | 11.41 | 1.4 | $O_2$ |
| F–F | 4.45 | 0.58 | $F_2$ |
| Si–Si | 4.65 | 2.0 | $Si_2$ |
| Si–Si | ~1.7 | ~0.9 | $Si_2H_6$ |
| P–P | 5.56 | 2.1 | $P_2$ |
| P–P | 2.07 | 0.95 | $P_4$ |
| S–S | 4.96 | 1.7 | $S_2$ |
| [illegible] | [illegible] | [illegible] | [illegible] |
| Ni–Ni | 0.11 | 0.2 | Ni solid |
| As–As | 3.91 | 1.8 | $As_2$ |
| Se–Se | 3.61 | 1.6 | $^{80}Se_2$ |
| Br–Br | 2.36 | 1.1 | $Br_2$ |
| Rh–Rh | 0.08 | 0.2 | Rh[illegible] |
| Sb–Sb | 2.61 | 1.9 | $Sb_2$ |
| Te–Te | 2.37 | 1.7 | $Te_2$ |
| I–I | 1.70 | 1.2 | $I_2$ |
| Hg–Hg | 1.69 | 1.5 | $Hg_2^{2-}$ |
| Pb–Pb | [illegible] | [illegible] | $Pb_2$ |
| H–B | 2.75 | 0.68 | $BH_3$ |
| H–C | 5.50 | 1.0 | $CH_4$ |
| H–N | 7.05 | 1.1 | $NH_3$ |
| H–O | 8.45 | 1.1 | $H_2O$ |
| H–O | 7.40 | 1.0 | $HO^-$ |
| H–Al | 1.76 | 0.60 | $AlH_4^-$ |
| H–Si | 2.98 | 0.84 | $SiH_4$ |
| H–P | 3.11 | 0.82 | $PH_3$ |
| H–S | 4.29 | 1.0 | $H_2S$ |
| H–Cl | 4.81 | 1.0 | HCl |
| H–As | 2.85 | 0.81 | $AsH_3$ |
| H–Se | 3.51 | 0.93 | $H_2Se$ |
| H–Br | 3.84 | 0.98 | HBr |
| H–Sn | 2.03 | 0.76 | $SnH_4$ |
| H–Sb | 2.09 | 0.77 | $SbH_3$ |
| C–H | 5.50 | 1.0 | $CH_4$ |
| C–B | 3.82 | 1.1 | $B(CH_3)_3$ |

| Bond A-B | Force Const. f (N cm$^{-1}$) | Bond Order | Compound |
|---|---|---|---|
| C–C | 16.5 | 3.2 | HCCH |
| C–C | [illegible] | 1.7 | $C_6H_6$ |
| C–C | 4.4 | 1.1 | $H_3CCH_3$ |
| C–N | 18.07 | 3.0 | HCN |
| C–N | 11.84 | 2.1 | $CN_2^{2-}$ |
| C–N | 6.54 | 1.3 | $NNCH_2$ |
| C–O | 16.01 | 2.4 | $CO_2$ |
| C–O | 12.76 | 2.0 | $OCH_2$ |
| C–O | 7.86 | 1.3 | $CO_3^{2-}$ |
| C–O | 5.1 | 0.96 | $O(CH_3)_2$ |
| C–F | 6.98 | 1.1 | $CF_4$ |
| C–S | 7.67 | 2.0 | $CS_2$ |
| C–S | 3.3 | 1.0 | $S(CH_3)_2$ |
| C–Cl | 3.12 | 0.93 | $CCl_4$ |
| C–Ni | 2.91 | 1.2 | $Ni_4CO$ |
| C–Ni | 1.43 | 0.68 | NiCO |
| C–Br | 2.42 | [illegible] | $CBr_4$ |
| C–Rh | 2.4 | 1.2 | $(Rh(CN)_6)^{3-}$ |
| C–Ag | 2.0 | 0.99 | $(Ag(CN)_2)^-$ |
| C–I | 1.69 | 0.79 | $CI_4$ |
| N–H | 7.05 | 1.1 | $NH_3$ |
| N–C | 18.07 | 3.0 | HCN |
| N–N | 22.42 | 3.2 | $N_2$ |
| N–N | 16.01 | 2.4 | N–NNH |
| N–N | 13.15 | 2.0 | $N–N–N^-$ |
| N–O | 25.07 | 3.1 | $N–O^+$ |
| N–O | 15.49 | [illegible] | $NO^+$ |
| N–O | 15.18 | 2.0 | ONCl |
| N–O | 11.78 | 1.7 | NNO |
| N–F | 4.16 | 0.66 | $NF_3$ |
| N–Si | 3.8 | 1.1 | $((CH_3)_3Si)_2NH$ |
| [illegible] | [illegible] | 2.5 | [illegible] |
| N–S | 3.1 | 0.87 | $H_3N–SO_3$ |
| O–Li | 1.58 | 0.66 | LiO |
| O–Be | 7.51 | 1.8 | BeO |
| O–B | 13.66 | 2.5 | BO |
| O–O | 16.39 | 2.0 | [illegible] |
| O–O | 11.41 | 1.4 | $O_2$ |

## Selected force constants and bond orders (according to Siebert) of organic and inorganic compounds (continued)

| Bond A-B | Force Const. f (N cm$^{-1}$) | Bond Order | Compound |
|---|---|---|---|
| O–O | 6.18 | 0.89 | $O_2^-$ |
| O–Na | ~3.2 | ~1.1 | Na–OH |
| O–Mg | 3.5 | 1.1 | MgO |
| O–Al | 5.66 | 1.5 | AlO |
| O–Al | 3.8 | 1.1 | $Al(OH)_4^-$ |
| O–Si | 9.25 | 2.1 | SiO |
| O–P | 9.41 | 2.0 | PO |
| O–P | 6.16 | 1.4 | $PO_4^{3-}$ |
| O–S | 10.01 | 2.0 | $SO_2$ |
| O–Cl | 4.26 | 1.0 | $ClO_2^-$ |
| [illegible] | [illegible] | [illegible] | [illegible] |
| O–Ti | 7.19 | 2.4 | TiO |
| O–Cr | 5.82 | 1.9 | CrO |

| Bond A-B | Force Const. f (N cm$^{-1}$) | Bond Order | Compound |
|---|---|---|---|
| O–V | 7.36 | 2.3 | VO |
| O–Fe | 5.67 | 1.7 | FeO |
| O–Cu | 2.97 | 0.93 | CuO |
| O–Ge | 7.53 | 1.8 | $^{74}GeO$ |
| O–Se | 6.45 | 1.5 | SeO |
| O–Mo | 3.05 | 1.2 | $Ba_2CaMoO_6$ (solid) |
| O–Ag | 2.00 | 0.79 | AgO |
| O–Sn | 5.53 | 1.7 | SnO |
| O–Te | 5.31 | 1.6 | TeO |
| O–Ba | 3.79 | 1.8 | BaO |
| [illegible] | [illegible] | [illegible] | [illegible] |
| O–Nd | 3.5 | 1.6 | $NdAc_3 \cdot H_2O$ (polymer) |

# 29. Fundamental physical constants

| Physical constant | Symbol | Numerical value (SI units) | Physical constant | Symbol | Numerical value (SI units) |
|---|---|---|---|---|---|
| Proton rest mass<br>Electron rest mass | $M_p$<br>$m_e$ | $1.6726231 \times 10^{-27}$ kg<br>$9.109389 \times 10^{-31}$ kg | Proton Lande factor<br>Proton-electron ratios[a] | $\gamma_h$<br>$\mu_e/\mu_p$<br>$\gamma_e/\gamma_p$ | 5.58569<br>658.2106881<br>658.2275841 |
| Elementary charge | e | $1.6021773 \times 10^{-19}$ As | Electron Lande factor | $g_e$ | 2.0023193043 |
| Charge-to-mass ratio for electron | $e/m_e$ | $1.758819 \times 10^{11}$ As kg$^{-1}$ | Electron magnetic moment<br>gyromagnetic ratio | $\mu_e$ | $-9.2847701(31) \times 10^{-24}$ J/T |
| Bohr radius | $a_0$ | $5.291772 \times 10^{-11}$ m | Neutron mass | $m_n$ | $1.6749286(10) \times 10^{-27}$ kg |
| Electron radius | $r_e$ | $2.817940 \times 10^{-15}$ m | Fine structure constant | $1/\alpha = 2h/\mu_0 ce^2$ | 137.0359895(61) |
| Planck constant<br>Boltzmann constant | h<br>$\hbar = h/2\pi$<br>$k_B$ | $6.62607 \times 10^{-34}$ Js<br>[illegible]<br>$1.38065 \times 10^{-23}$ J K$^{-1}$ | Velocity of light in a vacuum<br>Permeability of vacuum | c<br>$\mu_0 = 4\pi \times 10^{-7}$ | $2.99792458 \times 10^8$ m s$^{-1}$<br>[illegible] |
| Gas constant | R | 8.3145 J mol$^{-1}$ K$^{-1}$ | Permittivity of a vacuum | $E_0 = 1/\mu_0 c^2$ | $8.854187817 \times 10^{-12}$ c$^2$/Jm |
| Avogadro number | $N_A$ | $6.0221367 \times 10^{23}$ | Gravitational constant | G | $6.6725 \times 10^{-11}$ m$^3$ kg$^{-1}$s$^{-2}$ |
| | | | | | |
| Faraday constant | $F = N_A \cdot e$ | $9.64853 \times 10^4$ C mol$^{-1}$ | Compton wavelength of the electron ($h/m_e c$) | $\lambda_c$ | $2.426310 \times 10^{-12}$ m |
| Bohr magneton | $\beta_e$ | $9.27401 \times 10^{-24}$ J T$^{-1}$ | | | |
| Nuclear magneton<br>gyric ratio | $\beta_N$ | $5.05078 \times 10^{-27}$ J T$^{-1}$ | | | |
| $\pi = 3.14159$ | | e = 2.71828 | 1/e = 0.36788 | | ln(10) = 2.30259 |

[a] The magnetic moment $\mu_p$ is for a bare proton and the gyromagnetic ratio $\gamma_p$ is for protons in a spherical water sample which has [illegible]

# 30. List of suppliers

## Magnetic Resonance

| | |
|---|---|
| Bruker Analytik GmbH<br>Silberstreifen<br>Germany | NMR, EPR |
| Bruker Medizintechnik GmbH<br>Rudolf-Plank-Strasse 23<br>D-76275<br>Germany | MRI |
| Varian Inc<br>Palo Alto<br>CA 94304<br>USA | NMR |
| Jeol Ltd<br>1-2 Musashino 3-chome<br>Akishima<br>Tokyo 196 | NMR, EPR |
| Oxford Instruments Ltd<br>Head Office<br>Old Station Way<br>Eynsham<br>Witney<br>Oxon OX8 1TL | Magnets for NMR and MRI |
| Oxford Magnet Technology Ltd<br>Wharf Road<br>Eynsham<br>Witney<br>Oxon OX8 1BP<br>UK<br>47697 Westinghouse Drive<br>Fremont<br>CA 94539<br>USA | Magnets for MRI<br><br><br><br><br><br>systems |
| MR Resources Inc<br>PO Box 880<br>Gardner<br>MA 01440<br>USA | Reconditioned NMR spectrometers and |
| Fonar Corporation<br>110 Marcus Drive<br>Melville<br>NY 11747 | MRI |
| Picker International, Inc<br>World Headquarters<br>595 Miner Road<br>Cleveland<br>Ohio 44143<br>USA | MRI |
| Doty Scientific, Inc<br>700 Clemson Road<br>South Carolina 29229<br>USA | NMR Accessories |
| Magnex Scientific Ltd<br>21 Blacklands Way<br>Abingdon<br>Oxon OX14 1DY | MRI and NMR magnets |
| Resonance Instruments Ltd<br>Unit 13<br>Thorney Leys Business Park<br>Witney OX8 7GE<br>UK | NMR Spectrometers |
| Tecmag, Inc<br>Houston<br>TX 77081<br>USA | NMR Spectrometers |
| Goss Scientific Ltd<br>100 Vicarage Lane<br>Great Baddow<br>Essex CM2 8JP | NMR accessories, deuterated solvents, sample tubes |
| Martek Biosciences Corp<br>6480 Dobbin Road<br>Columbia<br>MD 21045<br>USA | Labelled compounds for NMR |

## Mass Spectrometry

| | |
|---|---|
| Micromass Ltd<br>Floats Road<br>Wythenshawe<br>Manchester M23 9LF | Sector mass spectrometers<br>Time-of-flight mass spectrometers<br>Quadrupole time-of-flight<br>Triple quadrupole mass spectrometers<br>Quadrupole mass spectrometers |
| AMD Intectra GmbH<br>Königsberger Strasse 1<br>D-27243 Harpstedt | Sector mass spectrometers |
| Bruker Daltonics<br>15 Fortune Drive<br>Billerica<br>MA 01821<br>USA | Fourier transform ion cyclotron resonance mass spectrometers<br>Ion trap mass spectrometers<br>Time-of-flight mass spectrometers |

| | |
|---|---|
| Hitachi Instruments, Inc<br>Separation Systems<br>3100 North First Street<br>San Jose<br>CA 95134<br>USA | Ion trap mass spectrometers |
| Suite 122<br>Irvine<br>CA 92618<br>USA | spectrometers |
| Jeol Ltd<br>1-2 Musashino 3-chome<br>Tokyo 196<br>Japan | Sector mass spectrometers |
| Kratos Analytical Shimadzu<br>Wharfside<br>Trafford Wharf Road<br>Manchester M17 1GP | Sector mass spectrometers<br>Time-of-flight mass spectrometers |
| PE Sciex<br>Applied Biosystems Division<br>850 Lincoln Center Drive<br>Foster City<br>CA 94404-1128<br>USA | Time-of-flight mass spectrometers<br>Triple quadrupole mass spectrometers |
| Framingham<br>MA 01701<br>USA | |
| ThermoQuest Finnigan Corporation<br>355 River Oaks Parkway<br>CA 95134<br>USA | Sector mass spectrometers<br>Ion trap mass spectrometers<br>Time-of-flight mass<br>Fourier transform ion cyclotron resonance mass spectrometers<br>Triple quadrupole mass spectrometers<br>Inductively coupled plasme |

## Spatially Resolved Spectroscopic Analysis

| | |
|---|---|
| Bruker AXS GmbH<br>Germany | X-ray diffraction equipment |
| Digital Instruments GmbH<br>Janderstrasse 9<br>D-68100 Manneheim<br>Germany | Scanning probe microscopes |
| ThermoMicroscopes<br>Sunnyvale<br>CA 94089<br>USA | Scanning probe |

| | |
|---|---|
| Perkin-Elmer Ltd<br>Post Office Lane<br>Beaconsfield<br>Bucks HP9 1QA<br>UK | Fluorescence, fluorescence accessories, UV-Vis-NIR Absorption, UV-Vis-NIR Accessories |
| 4-21 Sennin-cho 2-chome<br>Hachioji<br>Tokyo 193<br>Japan | accessories, UV-Vis-NIR Absorption, UV-Vis-NIR Accessories, CD Spectrometers |
| Jasco (UK)<br>18 Oak Industrial Park<br>Essex<br>CM6 1XN<br>UK | |
| Edinburgh Instruments Ltd<br>Riccarton, Currie<br>Edinburgh EH14 4AP | Fluorescence, fluorescence accessories |
| Instruments SA UK Ltd<br>2-4 Wigton Gardens<br>Stanmore<br>Middlesex<br>HA7 1BG<br>UK | Fluorescence, fluorescence accessories, UV-Vis-NIR Absorption, UV-Vis-NIR Accessories, CD Spectrometers, Ellipsometers |
| Eugene, OR 97402-0469<br>4849 Pitchford Ave.<br>Eugene, OR 97402-9165<br>USA | |
| Molecular Probes Europe BV<br>For customers in Europe, the Middle East, contact:<br>PoortGebouw<br>Rijnsburgerweg 10<br>2333 AA Leiden<br>The Netherlands | Fluorescence accessories |
| Hellma GmbH & Co.<br>Germany<br><br>Hellma (UK)<br>Cumberland House<br>24-28 Baxter Avenue<br>Southend on Sea<br>Essex SS2 6HZ | Fluorescence accessories, |
| Specac, Inc.<br>500 Technology Ct. Southeast<br>Smyrna, GA 30082<br>USA | Fluorescence accessories |
| Oxford Cryosystems<br>3 Blenheim Office Park<br>Long Hanborough<br>Oxford OX8 8LN<br>UK | Fluorescence accessories, UV-Vis-NIR Accessories |

| | |
|---|---|
| 3120 Hansen Way<br>Palo Alto<br>CA 94304<br>USA | UV-Vis-NIR Accessories |
| Hitachi Instruments, Inc<br>Separation Systems<br>CA 95134<br>USA | UV-Vis-NIR Absorption,<br>UV-Vis-NIR Accessories |
| Aviv Associates Inc.<br>750 Vassar Avenue<br>Lakewood NJ 08701-6907<br>USA<br>Unit 3<br>Pinnacle Business Centre<br>Gordon Street<br>Stockport<br>SK4 1RS<br>UK<br>1 Mole Business Park<br>Leatherhead<br>Surrey<br>KT22 7AU<br>UK | CD Spectrometers |
| Newport Corporation<br>USA | UV-Vis-NIR Accessories |
| Hammamatsu Photonics UK Ltd<br>Lough Point<br>2 Gladbeck Way<br>Windmill Hill<br>Middlesex EN2 7JA<br>UK | UV-Vis-NIR Accessories |
| Spectroscopy Central<br>Maple House<br>Green Lane<br>Warrington | UV-Vis-NIR Accessories |
| Osram<br>Hellabrunner Straße 1<br>81543 München<br>Germany | UV-Vis-NIR Accessories |

## Vibrational, Rotational Spectroscopy

| | |
|---|---|
| ASI SensIR Technologies<br>15 Great Pasture Road<br>Danbury<br>CT 06810<br>USA | Standard equipment for reflection spectroscopy |
| Perkin-Elmer Ltd<br>Bucks HP9 1QA<br>UK | Infra-red spectrometers, |

| | |
|---|---|
| 450 Saint-Jean-Baptiste Ave<br>Quebec<br>Canada G2E 5S5 | |
| Midac Corporation<br>17911 Fitch Avenue<br>Irvine | Infra-red spectrometers,<br>Raman and IR accessories |
| Analect Instruments<br>2771 North Garey Ave<br>Pomona<br>CA 91714<br>USA | Infra-red spectrometers |
| 5225 Verona Road<br>Madison<br>WI 53711-4495<br>USA | Raman and IR accessories |
| Bruker Analytische Messtechnik GmbH<br>D-76189<br>Germany | Infra-red spectrometers,<br>Raman and IR accessories |
| Bio-Rad (Digilab division)<br>237 Putnam Ave<br>Cambridge<br>MA 02139 | Infra-red spectrometers,<br>Raman and IR accessories |
| Innova Air Tech Instruments A/S<br>Energivej 30<br>2750 Ballerup<br>Denmark | Infra-red spectrometers |
| Block Engineering<br>Marlborough<br>MA 01752-1178<br>USA | Infra-red spectrometers |
| Optical Solutions Inc<br>9477 Greenback Lane<br>Suite 521<br>CA 95630<br>USA | Infra-red spectrometers |
| Jasco Inc<br>8649 Commerce Drive<br>Easton<br>MD 21601 | Infra-red spectrometers,<br>Raman and IR accessories |
| Instruments SA (Yobin Yvon – Spex)<br>16018 rue du Canal<br>91165 Longjumeau Cedex<br>France<br>and<br>NJ 08820<br>USA | Raman spectrometers |

| | |
|---|---|
| Renishaw (Raman Group)<br>Old Town<br>Wotton-under-Edge<br>Glos GL12 7DH<br>UK | Raman spectrometers,<br>Raman and IR accessories |
| Kaiser Optical Systems Ltd<br>371 Parkland Plaza<br>Ann Arbor<br>MI 48106<br>USA | Raman spectrometers, |
| Graseby Specac Ltd<br>97 Cray Avenue<br>UK | Raman and IR accessories |
| Coherent Inc<br>5100 Patrick Henry Drive<br>Santa Clara<br>CA 95054<br>USA<br>(TSL)<br>Thunbergsvägen 5A<br>Box 533<br>S-75121 Uppsala<br>Sweden | Raman and IR accessories |
| Specac Inc<br>Smyrna, GA 30082<br>USA | Raman and IR accessories |
| Spectroscopy Central<br>Maple House<br>Green Lane<br>Warrington | Raman and IR accessories |
| European Crystal Laborotories<br>Unit 3<br>Pinnacle Business Centre<br>Gordon Street<br>Stockport<br>SK4 1BS | Raman and IR accessories |
| Analytical Spectral Devices<br>Boulder, CO<br>USA | Near IR spectrometers |
| Bran & Luebbe<br>Buffalo Grove, IL<br>USA<br>Baltimore, MD<br>USA | Near IR spectrometers |
| Bühler Inc<br>Minneapolis, MN<br>USA | Near IR spectrometers |
| USA | |

| | |
|---|---|
| CVI Spectral Instrument<br>Division<br>Putnam, CT<br>USA | Near IR spectrometers |
| Focus Engineering<br>Budapest<br>Foss NIR Systems<br>Höganäs<br>Sweden | Near IR spectrometers<br><br>Near IR spectrometers |
| Infrared Engineering<br>Concord, MA<br>USA<br>Villa Park, CA<br>USA | Near IR spectrometers |
| Leco<br>Saint Joseph, MI<br>USA | Near IR spectrometers |
| Rockville, MD<br>USA | |
| Mattson Instruments<br>Madison, WI<br>USA | Near IR spectrometers |
| Midac<br>USA | Near IR spectrometers |
| Ocean Optics<br>Dunedin, FL<br>USA | Near IR spectrometers |
| OLIS: On-line Instruments | Near IR spectrometers |
| Oriel Instruments<br>Stratford, CA<br>USA | Near IR spectrometers |
| Perten Instruments<br>Reno, NV<br>Rosemount Analytical<br>Orrville, OH<br>USA | Near IR spectrometers<br><br>Near IR spectrometers |
| OUP Guided Wave<br>El Dorado Hills, CA<br>USA<br>South Norwalk, CT<br>USA | Near IR spectrometers |
| Zeiss Jena<br>Jena<br>Germany | Near IR spectrometers |
| Zeltex<br>USA | Near IR spectrometers |

| | |
|---|---|
| Wright Instruments Ltd<br>Unit 10<br>26 Queensway<br>Enfield<br>Middlesex EN3 4SA<br>UK [illegible]<br>1250 W.Middlefield Road<br>PO Box 7013<br>Mountain View<br>CA 94039-7013<br>USA | CCD detectors<br>[illegible] |
| Coherent Inc<br>5100 Patrick Henry Drive<br>Santa Clara<br>CA 95054<br>USA | Lasers |
| Isotec Inc<br>3858 Benner Road<br>[illegible]<br>USA | Deuterated solvents for NMR; labelled |
| Optispec<br>Rigistrasse 5<br>CH-8173 Neerach<br>Switzerland | Special equipment for SBSR-ATR and ME-ATR Spectroscopy |
| [illegible] Optical Drive<br>Azusa<br>CA 91702-3251<br>USA | |
| Hamamatsu Corp<br>360-T Foothill Road<br>[illegible]<br>NJ 08807-0910<br>USA | Light sources |
| Omega Optical, Inc<br>3 Grove Street<br>PO Box 573<br>Brattleboro<br>USA | Dielectric filters |
| Princeton Instruments Inc<br>3660-T Quakerbridge Road<br>Trenton<br>NJ 08619<br>USA<br>400 York Avenue<br>Duryea<br>PA 18642<br>USA | Detectors |
| Heraeus Amersil, Inc<br>3473 Satellite Blvd<br>[illegible]<br>USA | Glass production |

| | |
|---|---|
| 18019N 25th Avenue<br>Phoenix<br>AZ 85023<br>USA | |
| Fiberguide Industries Inc<br>1-T Bay St<br>NJ 07980<br>USA | Fibers |
| Ceramoptec GmbH<br>Siemenstrasse 8-12<br>D-53121 Bonn 1<br>Germany<br>620 Magnolia Street<br>PO Box 2167<br>Orangeburg<br>SC 29116<br>USA | Fibers |
| Kaiser Optical Systems Inc<br>371 Parkland Plaza<br>Ann/Arbor<br>MI 48106-0983<br>USA | Imaging spectrographs |
| Breault Research Organiza-tion Inc<br>Tucson<br>AZ 85715<br>USA | Illumination analysis |
| MicroQuartz Sciences<br>4420-T S 32nd St<br>Phoenix<br>USA | Custom fibreoptic probes |
| Rare Earth Medical Inc<br>126B Mid-Tech Drive<br>West Yarmouth<br>MA 02673<br>USA<br>151-T Osigian Blvd<br>Dept 50<br>Warner Robins<br>GA 31088<br>USA | Custom fibreoptic probes |
| [illegible]<br>Vancouver<br>British Columbia<br>Canada V5Z 4H5 | [illegible] |
| Labsphere<br>PO Box 70A<br>North Sutton<br>USA | Reflectance and scattering standards |

| | |
|---|---|
| Ltd<br>Fiber Optics Division<br>5-11-3, Shimbashi<br>Minato-ku<br>Tokyo 105<br>Japan | |
| Via del Sole<br>CH-6598<br>Tenero | |
| Sapphire Engineering Inc<br>53 C Portside Dr<br>Pocasset | Sapphire products |
| Saphirwerk Industrieprodukte AG<br>Erlenstrasse 36<br>CH-2555 Brügg, Biel<br>Switzerland | Sapphire products |
| France | |
| Sandoz SA<br>Clos de la Gare<br>CH-1482, Cugy<br>Switzerland | Sapphire products |
| CH-3608, Thun<br>Switzerland | |
| Wilmad<br>PO Box 688<br>1002 Harding Highway<br>Buena<br>USA | NMR and EPR glassware |
| AMT<br>2570 E Cerritos Avenue<br>Anaheim<br>CA 92806<br>USA | |

## High Energy Spectroscopy

| | |
|---|---|
| ARL<br>Eclubens, CH-1024<br>Switzerland<br>Austin<br>TX 78759<br>USA | X-ray fluorescence systems |
| Bruker AXS GmbH<br>D-76181 Karlsruhe<br>Germany<br>Bucks HP7 9NA<br>UK | X-ray fluorescence and diffraction systems |
| Oberschleissheim<br>D-85764<br>Germany | |
| BAIRD<br>Bedford<br>MA 01730-1468 | X-ray fluorescence systems |
| CIANFLONE<br>Pittsburg<br>PA 15275-1002<br>USA | X-ray fluorescence systems |
| EDAX<br>Prairie View | X-ray fluorescence systems |
| EG & G Ortec | X-ray fluorescence systems |
| Meteorex<br>Espoo, FIN 02201<br>Finland | X-ray fluorescence systems |
| N V Philips<br>The Netherlands | X-ray fluorescence systems |
| Niton<br>Bedford<br>MA 01730-0368<br>USA | X-ray fluorescence systems |
| Oxford Instruments<br>UK | X-ray fluorescence systems |
| Kratos Analytical<br>Wharfside<br>Trafford Wharf Road<br>Manchester M17 1GP<br>UK | X-ray fluorescence systems<br>accessories |
| Siemens | X-ray diffractometers, X-ray accessories |
| Oxford Cryosystems<br>3 Blenheim Office Park<br>Lower Road<br>Oxford OX8 8LN<br>UK | X-ray accessories |

## Atomic Spectroscopy

| | |
|---|---|
| Cetac Technologies, Inc.<br>Omaha, NE 68107<br>USA | |
| Edinburgh Instruments Ltd<br>Riccarton, Currie<br>Edinburgh EH14 4AP<br>UK | |
| Arlington Heights, IL 60004<br>USA | |

| | |
|---|---|
| Perkin-Elmer Ltd<br>Post Office Lane<br>Beaconsfield<br>Bucks HP9 1QA<br>UK | |
| GmbH<br>Boschstr. 10<br>D-47533 Kieve<br>Germany | |
| Thermo Jarrell Ash Corp<br>27 Forge Parkway, Franklin,<br>USA | |
| ThermoQuest Corporation<br>San Jose<br>CA 95134<br>USA | |

## Software

| | |
|---|---|
| ACD (Advanced Chemistry Development)<br>133 Richmond Street West<br>Toronto<br>Ontario<br>Canada M5H 2L3 | Chemical structure and spectroscopic manipulation |
| Acorn NMR Inc<br>46560 Fremont Blvd #418<br>Fremont<br>CA 94538<br>USA | NMR data processing |
| Galactic Industries Corporation<br>New Hampshire 03079<br>USA | |
| Leeman Labs Inc<br>6 Wentworth Drive<br>Hudson<br>NH 03051<br>Shimadzu Corporation<br>1, Nishinokyo-Kuwabaracho<br>Nakagyo-kur<br>Kyoto 604-8511<br>Japan | |
| CIC Photonics Inc<br>[illegible]<br>USA | |
| Bio-Rad Laboratories, Sadtler Division<br>3316 Spring Garden Street<br>Philadelphia, PA 19104-2596<br>USA<br>380 Main Street<br>Dunedin, FL 34698<br>USA | |